# ARTIFICIAL INTELLIGENCE

THIRD EDITION

Dzhimsher Chelidze

FREEFALL

# Contents

Preface ..... 8
Introduction ..... 8
What's New in the Third Edition ..... 11
A Map of This Book ..... 13
A Map of My Books: Five Books, One Path ..... 14
Part 1. An Introduction to Artificial Intelligence ..... 17
Chapter 1. Getting to Know AI ..... 17
What Is Artificial Intelligence? ..... 17
How Are AI Models Trained? ..... 21
The Shared Weaknesses of Today's AI Solutions ..... 22
Five Beginner Fears—Brief and Honest ..... 26
Chapter Summary ..... 27
Chapter 2. The Types of Artificial Intelligence ..... 27
By the Kind of Task They Solve ..... 27
By "Strength" and the Handling of Uncertainty ..... 28
The Obstacles on the Road to Strong AI ..... 32
Copilots, AI Agents, and Multi-Agent Systems ..... 44
What the Agent Is Allowed to Decide ..... 49
Chapter Summary ..... 52
Chapter 3. What Narrow AI Can Do—and the Big Trends ..... 53
Narrow AI in Applied Tasks ..... 53
The Key Trends ..... 54
The Three Layers of the AI Market: Whose Game Is Where ..... 62
Chapter Summary ..... 64
Chapter 4. Generative AI ..... 65
What Is Generative Artificial Intelligence? ..... 65
The Problems of Generative AI ..... 65
The Limitations of AI That Cause Problems ..... 69
So What Do We Do? ..... 71
Chapter Summary ..... 74
Chapter 5. Large Language Models (LLMs) ..... 77
The LLM Phenomenon ..... 77
What Is RAG? ..... 78

Fine-tuning ..... 80
How Does an LLM Work? ..... 80
Source Traceability: A Requirement, Not a Preference ..... 83
Chapter Summary ..... 83
Chapter 6. Prompt Engineering: The Discipline of Setting Tasks for AI ..... 84
The prompt as a management tool ..... 84
The Basic Structure of a Request ..... 85
The Formula for Task Quality ..... 88
Four Templates for Different Situations ..... 91
Techniques for Improving the Quality of Work with AI ..... 93
The System Prompt and Assembling a Prompt with AI ..... 96
A Starter Kit: 10 Ready-Made Prompts for a Leader ..... 97
Using Questions When Working with AI: From the Task Plan to Reviewing Answers 98
A Checklist and Typical Mistakes ..... 105
Chapter Summary ..... 106
Chapter 7. Regulating AI ..... 109
Why Does AI Development Cause Concern? ..... 109
The Regulatory Map: Initiatives from 2023 to 2026 ..... 110
Russia's AI Law: What Was Actually Passed ..... 114
A Forecast for the Future ..... 116
Who Writes the Rules of the Frontier Model Market ..... 119
Chapter Summary ..... 123
Chapter 8. AI Security ..... 124
Unacceptable Events You Have to Rule Out ..... 125
How Attackers Use AI ..... 125
What Can Make These Scenarios Happen? ..... 128
Crash Tests and Requirements for AI Solutions ..... 128
Security as an Organizational Barrier: The Economics of the Perimeter ..... 141
Chapter Summary ..... 142
Chapter 9. Decision Support Systems, or Digital Advisors ..... 145
What Digital Advisors Are ..... 145
How DSS Work, and What Is Wrong with Them ..... 147
What to Do and Where to Head (A 5–20-Year Horizon) ..... 152
Chapter Summary ..... 154

Chapter 10. Neuromorphic and Quantum AI ..... 155
Neuromorphic Systems ..... 156
Quantum Systems ..... 158
Chapter Summary ..... 162
Part 2. AI in the Systems Approach ..... 163
Chapter 11. Strategy and Organizational Structure ..... 163
The Impact of AI ..... 163
The Impact on AI ..... 166
A Practical Example ..... 167
What About Russia: The Improviser's Profile ..... 185
Chapter Summary ..... 188
Chapter 12. Business Processes ..... 189
The Impact of AI ..... 189
The Impact on AI ..... 194
Chapter Summary ..... 195
Chapter 13. Managing Projects, Change, and Products ..... 196
The Impact of AI ..... 197
The Impact on AI ..... 200
The 7 Deadly Sins of Digitalization (AI Included) ..... 201
Chapter Summary ..... 212
Chapter 14. Communication ..... 213
The Impact of AI ..... 213
The Impact on AI ..... 216
Chapter Summary ..... 216
Chapter 15. Digital Technologies, Working with Data, and Cybersecurity ..... 217
Digital Technologies ..... 217
IT Systems ..... 224
Working with Data ..... 231
Chapter Summary ..... 231
Chapter 16. Regular Management Practices ..... 232
Management Planning ..... 232
Delegating Authority and Tasks ..... 233
Working with Feedback from Direct Reports ..... 233
Running Meetings ..... 233

Personnel Matters and Decisions About People, Evaluating Employee Performance, Staff Development ... 233
Three “Regular Management × AI” Scenarios: Where to Start ... 234
Chapter Summary ... 234
Chapter 17. Lean Manufacturing and the Theory of Constraints (TOC) ... 235
The Impact of AI ... 236
The Impact Both Ways ... 237
Chapter Summary ... 248
Chapter 18. The Processes of Business Functions ... 248
Digital Business ... 249
Digital Marketing ... 252
Digital Manufacturing ... 252
Digital Analytics ... 253
Digital Products ... 253
New Ways of Working ... 254
Engineering and Design: From Generation to Verification ... 254
Chapter Summary ... 255
Summary of Part 2 ... 256
General Conclusions and Forecasts ... 256
Part 3. Practical Examples and Recommendations for Applying Artificial Intelligence ... 257
Chapter 19. AI for Small Business and Entrepreneurs ... 258
AI in 1C Solutions ... 258
Marketplaces for Specialized AI Solutions ... 260
A Digital Advisor for Project Management ... 262
Generative AI for Content Creators ... 265
Building Assistants ... 267
A Beauty Assistant ... 267
Recommendations for Small and Midsize Business ... 267
Chapter Summary ... 270
Chapter 20. Examples and Recommendations for Midsize and Large Companies ... 271
Generative AI in Deal-Making ... 272
Industry and Infrastructure ... 272
Customer-Facing Services ... 282
Planning and People ... 286

Recommendations for Large Business ..... 290
Chapter Summary ..... 296
Chapter 21. AI in the Leader's Own Work ..... 297
A Leader's Tasks: What You Can Hand to AI ..... 297
My Personal Loop: How It Works ..... 298
Three Cases from Practice ..... 304
What a Leader Gains from an Alliance with AI ..... 311
The Limit of the Alliance: Distinctiveness Stays with You ..... 312
How This Turned into a Product ..... 313
Chapter Summary ..... 315
Chapter 22. How to Run Technology Implementation Projects ..... 315
The Project Life Cycle ..... 317
Defining Project Goals and Possible Effects ..... 317
Roles in a Project ..... 318
Three Playbooks in Brief ..... 321
Chapter Summary ..... 322
Chapter 23. A Systems Approach to Implementation and the Road Map ..... 322
The AI Hype Wave ..... 322
A Systems Approach to AI Implementation ..... 324
The Road Map and the Links to Other Projects ..... 324
Chapter Summary ..... 325
Chapter 24. What a Leader Needs to Know and Be Able to Do in the Age of AI ..... 326
Introduction ..... 326
The Competency Model ..... 326
The Competency Matrix ..... 334
Competencies for the Age of Verification ..... 337
Chapter Summary ..... 338
Conclusion ..... 339
Who Is Exposed—and What to Do About It ..... 339
Getting Past the Barriers to Deploying "Strong" Models ..... 341
The Advantages of Specialized Models ..... 341
Where Specialized AI Is Heading ..... 342
The Foundation: Data Quality and Protection ..... 343
Putting It into Practice ..... 343

Postscript 2026: Four Signals of Market Maturity ..... 345
How I Want to End This Book ..... 351
What's Next ..... 353
Appendices ..... 353
Appendix A. Glossary of Key Terms ..... 353
Appendix B. Checklist: Where You Need AI and Where Classic Automation Will Do .. 356
Appendix C. Your First Seven Days with AI: A Plan for Beginners ..... 356
Appendix D. The Full Action Plan: 24 Steps from 24 Chapters ..... 357

# Preface

## Introduction

The year 2023 marked the beginning of AI's third—and so far hottest—"summer." The questions back then were sharp: what should developers and businesses do? Should you cut staff and put your full trust in the new technology—or is it all froth, and will this summer give way to yet another AI winter? How does AI square with the classic tools and practices of management—projects, products, business processes? Does it really rewrite the rules of the game?

By the time of this third edition—in 2026—the market had answered some of those questions on its own. I have updated the book to reflect those answers: new data, research, trends, and examples. But the essence remains the same: we will dig into these questions together, and as we reason through the evidence, I will do my best to answer them.

Let me say this up front: you will find a bare minimum of technical detail here—and, frankly, even the occasional technical inaccuracy. But this book was never about choosing the right type of neural network for a given task. What I offer instead is the perspective of an analyst, a manager, and an entrepreneur: what is happening here and now, what to expect globally in the near future, and what to prepare for.

### Who is this book for—and what will you get out of it?

- **Company owners and top executives.** They will come away understanding what AI is, how it works, where the trends are heading, and what to expect—and so sidestep the biggest mistake in digitalization: distorted expectations. That is what lets you lead in this space while keeping costs, risks, and timelines to a minimum.
- **IT entrepreneurs and startup founders.** They will see where the industry is heading, which products and integrations are worth building, and what they will run into in practice.

- **Technical specialists in IT.** They will look at the field not only through a technical lens but through an economic and managerial one—and understand why up to 95% of AI adoptions deliver no measurable financial effect (MIT, 2025). That insight may help their career development.
- **Everyone else.** They will understand what the future holds and whether they should fear being replaced by AI. Spoiler: the first people at risk turned out to be those in creative professions.

Our journey runs through three big parts. First, we break down what AI is: where it all started, what its problems and capabilities are, where the trends are heading, and what lies ahead—the theory and the long view. Then we look at the synergy between AI and the tools of the systems approach: how they shape each other, and in which scenarios AI is already at work inside IT solutions. Finally, we focus on practice—examples and recommendations.

In keeping with my favorite tradition, the book includes QR codes (in the print edition) and live hyperlinks (in the electronic one) that lead to useful articles and materials worth your time.

As for AI itself: while writing this book I used it to demonstrate how it works—those passages are clearly marked—and to hunt for ideas. The technology is not yet good enough for AI-generated text to pass as finished material. I should also note that my working language is Russian, so all my prompts to AI are written in Russian.

> **From My Practice**
>
> Many of the conclusions, examples, and approaches in this book grew out of my own practice—including building the BAEOS product. Throughout the chapters, I refer to it as a living case study of embedding AI into real management work.

I want to close this preface by thanking the people who helped me along the way:

- Alisa and my son Valery
- my parents

- my coach Evgeny Bazhov
- my team, especially Alexander Peremyshlin
- my partners and clients who gave me food for thought, especially Kirill Neelov
- my colleagues on many projects

## What's New in the Third Edition

If you read the first or second edition, here is what has changed—and why this edition is worth your time.

**Two new chapters.** "Prompt Engineering: The Discipline of Setting Tasks for AI" (Chapter 6): templates, techniques, ready-to-use prompts for leaders, the task-quality formula—framing × size × data × model—and the section "Using Questions When Working with AI: From the Task Plan to Reviewing Answers": five questions for accepting a plan before the work starts, and nine types of questions for checking AI answers, with research on model sycophancy. And "AI in the Leader's Own Work" (Chapter 21): the personal loop, three cases from my own practice (from unpacking a work conversation to a decision on a major deal), four levels of assistant setup (from the account-wide instruction and the Project Register to the task in the chat), an interview instead of a template, and the meta-case "an article in two days." Plus, in Chapter 5, the context window and the "chat + file" pattern.

**Personal techniques from my own practice,** none of them in the earlier editions: how the precision of a term changes both the quality of the answer and the token spend; the "do not stop at the first answer" rule—passes from different angles instead of a single request; the "ask the model to break the task into steps" trick—and then approve the plan yourself; the story of how the most powerful model available ruined a simple task and cost me an hour; what to do when your assistant starts "going dumb"; what a leader's day looks like without the tool they have grown used to. Plus seven "From My Practice" sidebars and case studies ranging from Bridgewater to Anthropic's report on 400,000 work sessions.

**Security of autonomous agents.** Chapter 8: three cases from 2026—from the laboratory runs at OpenAI and Anthropic to the first consumer one, where a household AI agent hacked a fitness club's website on its way to an ordinary errand—and the "five mechanisms of failure" typology, with a defense against each. The chapter's overall conclusion: an agent's boundary is set by architecture, not by the model's good sense. Separately, the case of three agents in one space (Anthropic, August 2026): the conflict arose with no

excess authority and no outside adversary, and it adds a second managerial question to the boundaries—coordination between agents.

**The facts as of mid-2026.** A full update: AI economics—from the "$600 billion question" to trillion-dollar capital spending; a new hallucination leaderboard; and the 2023–2026 regulation map, including Russia's AI law.

**Forecasts that came true.** Events that bore out the forecasts of the earlier editions: export controls on frontier models, the market's turn toward specialized models, and a "Postscript 2026" with four signals of market maturity.

**Strategy and Russia.** In Chapter 11, the analysis of China's "AI Plus"–"1397" pairing is joined by a new section, "What About Russia: The Improviser's Profile": the GII 2025 numbers, the historical track of forks—from the atomic project to OGAS—and the author's position that AI closes what the improviser lacks and multiplies what the improviser has; two infographics and a QR code out to the full article.

**Competencies for the AI age.** Chapter 24 adds an analysis of the five scarce competencies of the specialists whose work AI is already rebuilding, the conclusion that training must be redesigned around verification rather than generation, and a new scarce figure—the erudite with an AI tool.

**A visual layer.** More than ten new diagrams and infographics, among them the S-curve of AI maturity, the three layers of the market, AI economics in numbers, a regulation timeline, the IT landscape with an AI layer on top, the four market signals, and a poster of the principles.

**Navigation and presentation.** The manifesto of the 7 Landing Principles—the whole book on a single page; a map of the series, "five books, one path," and a "What's Next" section; "In Plain Terms" sidebars, an expanded glossary, a checklist of where you need AI and where classic automation will do, takeaways and action steps in every chapter, "Your First 7 Days with AI," and a 24-step master plan.

**If you are short on time:** start with Chapter 21, Chapter 6, and the Postscript—that is where most of what is new is concentrated.

## A Map of This Book

If you are busy and not ready to read the whole book, here is a quick map: the question each chapter answers.

| Chapter | The question it answers |
|---|---|
| **Part 1. An Introduction to Artificial Intelligence** | |
| Chapter 1. Getting to Know AI | What AI is—and when you need it versus cheaper classic automation |
| Chapter 2. The Types of AI | What kinds of AI exist; why strong AI is not arriving tomorrow; agent autonomy levels |
| Chapter 3. Narrow AI and the Trends | What already works and where the market is going (9 key trends) |
| Chapter 4. Generative AI | The capabilities and real limits of generative AI (GenAI) |
| Chapter 5. LLMs | How language models work, RAG, and fine-tuning |
| Chapter 6. Prompt Engineering | How to get quality answers and standardize your prompts |
| Chapter 7. Regulating AI | Where regulation is heading and what to build in early |
| Chapter 8. AI Security | AI-specific risks and how to defend against them |
| Chapter 9. Digital Advisors | The promise and limits of decision support systems (DSS), and how to adopt them |
| Chapter 10. Neuromorphic and Quantum AI | The horizon for AI hardware |
| **Part 2. AI in the Systems Approach** | |
| Chapter 11. Strategy and Organizational Structure | How AI changes strategy and structure—and how they determine adoption success |
| Chapter 12. Business Processes | What AI does to processes—and why you cannot automate chaos |
| Chapter 13. Managing Projects, Change, and Products | How AI helps run projects; the “7 sins” of digitalization |
| Chapter 14. Communication | Where AI saves 50–80% of time spent processing information |

| | |
|---|---|
| Chapter 15. Technology and Data | What AI combines with—and why data caps its value |
| Chapter 16. Regular Management Practices | How to build AI into your management rituals |
| Chapter 17. Lean and TOC | Why "Lean before AI"—and how they reinforce each other |
| Chapter 18. Business Functions | What AI changes in marketing, production, analytics, and products |
| **Part 3. Practice and Recommendations** | |
| Chapter 19. AI for Small Business | Where small and midsize businesses (SMBs) start—with no budget and no developers |
| Chapter 20. Midsize and Large Companies | A library of industry cases and the principles of scaling |
| Chapter 21. AI in the Leader's Own Work | The personal loop: journal, decisions, communications—and the short loop of growth |
| Chapter 22. Adoption Projects | Three types of projects and a management frame for running them |
| Chapter 23. The Systems Approach and the Road Map | How to tie AI to your digital strategy and build a road map |
| Chapter 24. The Leader's Competencies | What a leader must know and be able to do in the age of AI |
| Conclusion + Postscript 2026 | Four signals of market maturity and the profile of a winning AI product |

**If you are completely new to AI:** start with the short route—Chapters 1 → 3 → 6 → 14 → 19 and the Conclusion. That is enough for a confident start: what AI is, what it can already do, how to set tasks for it, where it saves time, and where a small business should begin. Read Chapters 4–5 as the "how does this work" questions come up, and feel free to leave Chapter 10 (quantum and neuromorphic systems) for later.

## A Map of My Books: Five Books, One Path

This book is part of a series, and each book in it has its own job. So you don't have to piece the puzzle together yourself, here is a quick guide.

| Book | The question it settles | When to read it |
|---|---|---|
| Digital transformation for CEO and owner. Part 1. Immersion (in short: DT-1. Immersion) | What digitalization, technology, and IT systems are—the foundation and the vocabulary | If you are new to the digital agenda |
| Digital transformation for CEO and owner. Part 2. The Systems Approach (in short: DT-2. The Systems Approach) | Processes, TOC, Lean, projects, communications—the management toolkit | Before any transformation—and alongside Part 2 of this book |
| Digital transformation for CEO and owner. Part 3. Cybersecurity (in short: DT-3. Cybersecurity) | How not to lose your business: cybersecurity for leaders | Together with Chapter 8 of this book |
| Artificial intelligence. Freefall (in short: AI-1. Freefall) (you are here) | What AI is, what it can do, and how to think about it | Your entry point into AI |
| Artificial intelligence. Practical guide for implementation (in short: AI-2. Practical guide) | How to implement: readiness, portfolio, pilots, ROI, risks | Once the “let’s do it” decision is made |

**Three routes.**

- For an SMB owner: AI-1 → AI-2 (Chapter 17) → DT-1 as needed.
- For a transformation leader (CDTO): DT-1 → DT-2 → AI-1 → AI-2 → DT-3.
- For a functional leader: AI-1 (the beginner’s route) → the relevant chapters of Part 2.

# Part 1. An Introduction to Artificial Intelligence

Part 1 answers the question “what is AI, and what should we expect?” You can read the chapters in order or pick the ones that match your task.

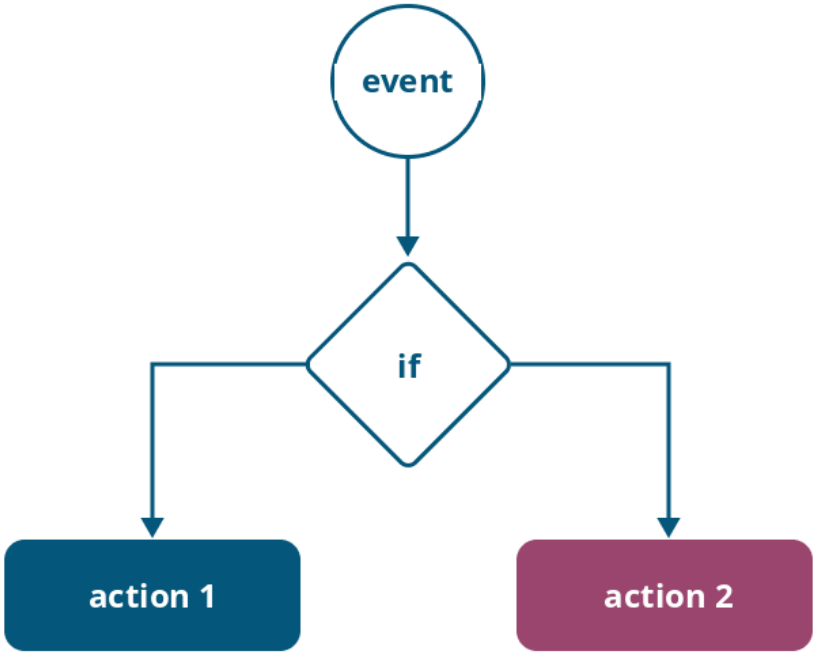
## Chapter 1. Getting to Know AI

### What Is Artificial Intelligence?

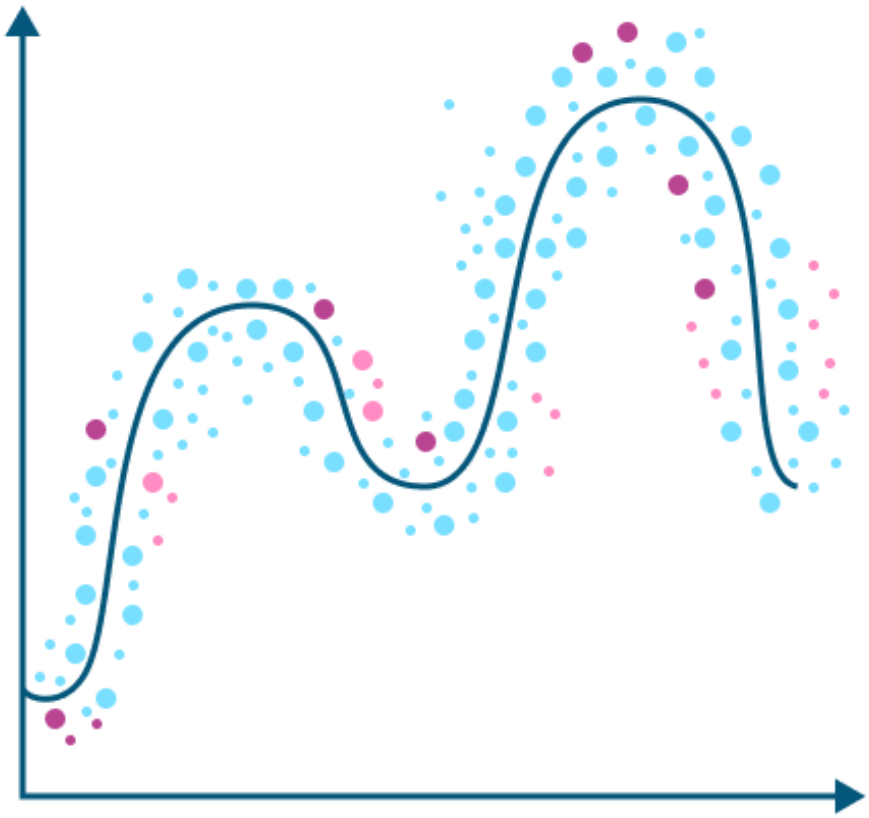
Let’s start with this: there is no single definition. Ask AI itself, and every model will give you its own answer, each at its own depth.

As a human, I prefer the simplest and clearest definition—and it happens to be the one the models’ own answers converge on: AI is any mathematical method that imitates human—or any other natural—intelligence.

In other words, AI is a huge family of solutions—everything from primitive mathematical algorithms to rule-based expert systems to modern statistics-based approaches.




*Classic solutions built on rules and logic*



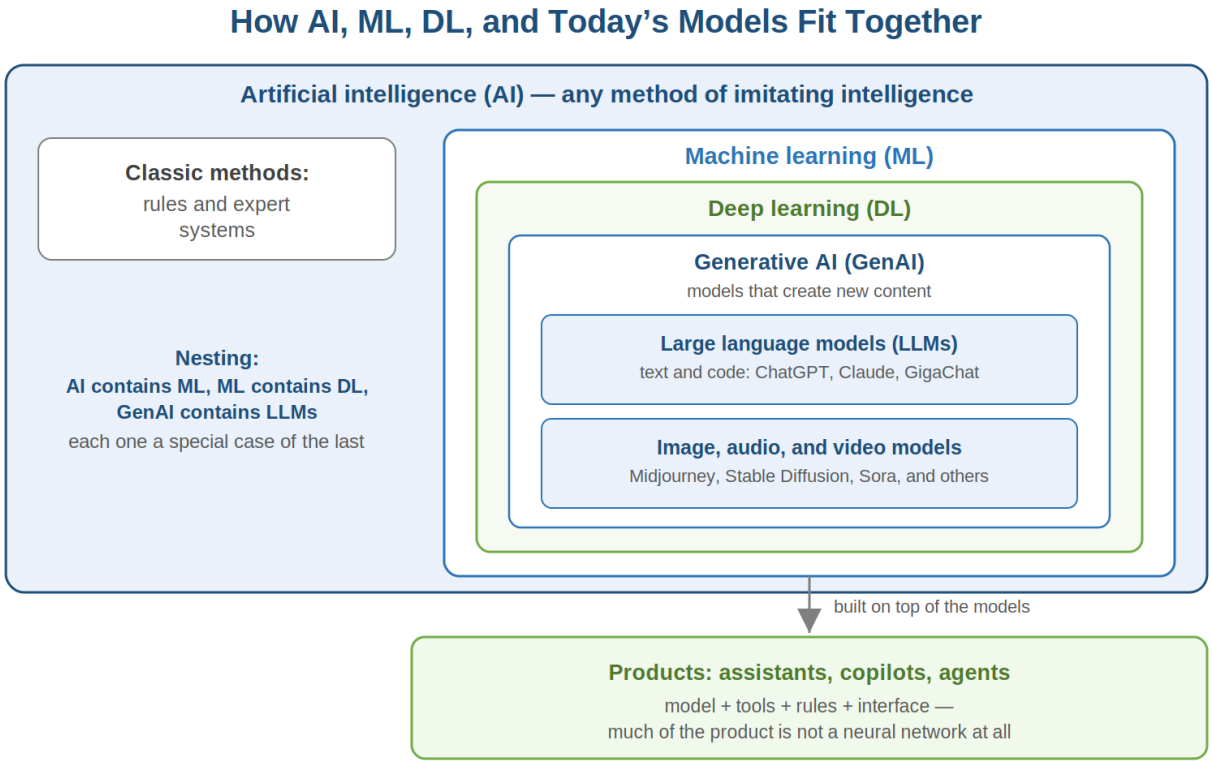
*Solutions built on data analysis, statistics, and probability estimates.*

In this book, we will focus on the branch built on statistics and the search for patterns.

This branch was born back in the 1950s, but what interests us most is what it means today, in the 2020s. Three main directions stand out.

1. **Neural networks**—mathematical models patterned on the neural connections in the brains of living beings. The human brain itself is, in fact, a hyper-complex neural network. Its key feature: our neurons are not limited to "on/off" states—they carry far more parameters. We cannot yet digitize and harness them in full.
2. **Machine learning (ML)**—statistical methods that let a computer get better at a task as it accumulates experience and further training. This direction has been known since the 1980s.
3. **Deep learning (DL)**—this goes beyond supervised learning, where a human shows the machine the right answers. It also covers systems that teach themselves—no human in the loop—combining multiple training

and data-analysis techniques at once. The direction has been developing since the 2010s and is considered the most promising for creative tasks and for problems where humans cannot articulate the underlying patterns themselves. All the popular modern models we hear so much about belong here. The catch: we have no way to predict what conclusions the network will reach—the only lever we hold is the data we "feed" it.




## When do you need AI—and when is classic automation enough?

In short, AI works best when three conditions hold at once: you have data, the task is highly variable (full of exceptions), and it can be automated.

**Classic automation is the better choice when:**

- the task has limited complexity and variability
- interpretability, response speed, and reliability are critical
- there is no data for machine learning

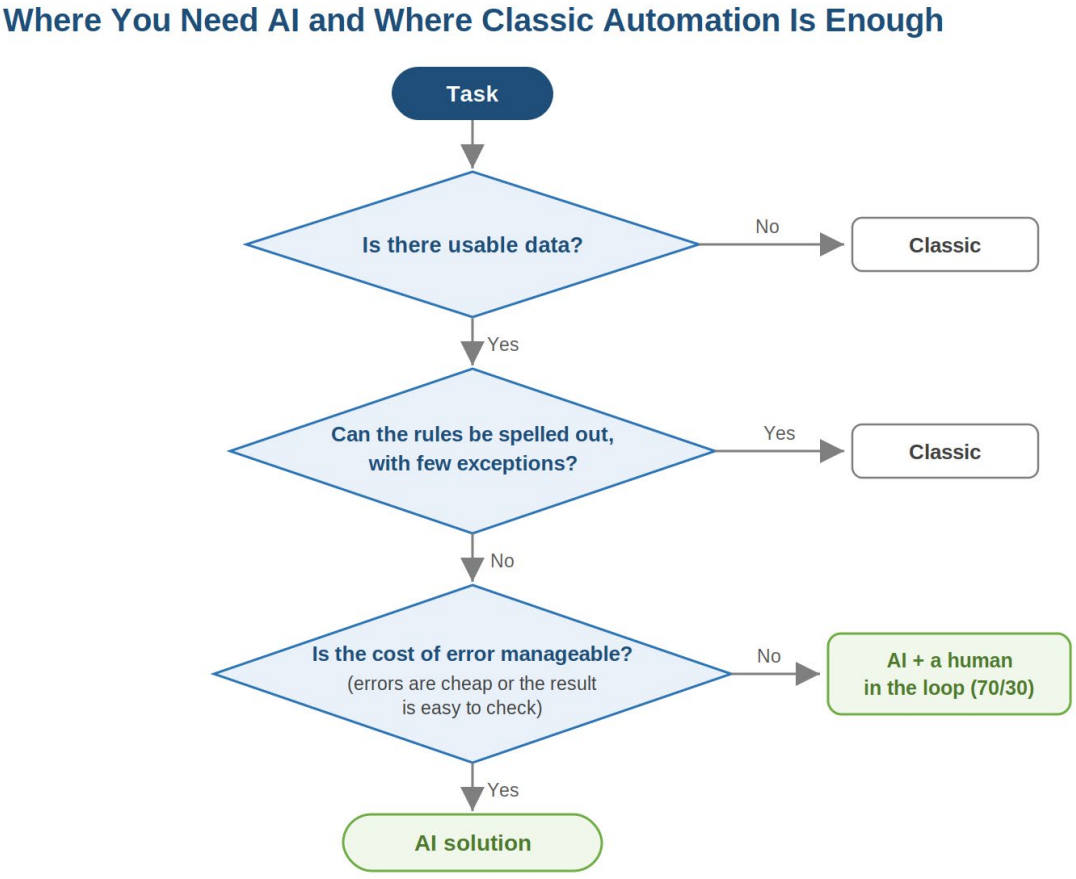
**AI solutions are the better choice when:**

- the task is highly complex or has many variations and factors, and the rules cannot be stated explicitly—or are too complicated

- large volumes of data are available to mine for patterns
- you need to analyze unstructured data (text, images, audio), where traditional algorithms fall short
- fast adaptation to new data and pattern discovery are required
- the decision depends on probabilistic estimates

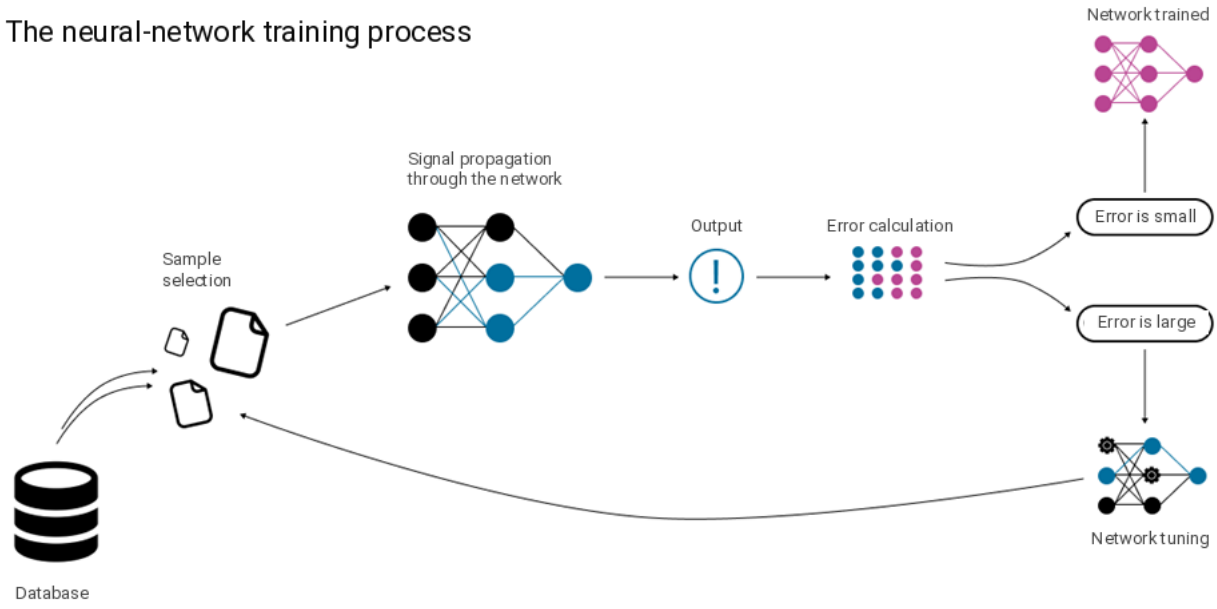
## The key difference: traditional algorithms vs AI models

| Aspect | Traditional algorithms | AI models |
|---|---|---|
| How they work | Rules = algorithm | Rules = data |
| When to use | Limited complexity | High complexity + many factors |
| Do they need examples? | No | Yes, 100+ examples |
| Interpretability | High | Low (a “black box”) |
| Examples | Payroll calculation, routing | Call classification, fraud detection, recommendations |



## How Are AI Models Trained?

Today, many applied business AI models are trained under supervision: a human provides the input, the network returns an answer, and the human tells it whether that answer was right or wrong. Over and over again.




*The neural-network training process*

The familiar "captchas" work along similar lines (CAPTCHA—Completely Automated Public Turing test to tell Computers and Humans Apart): the little graphic tests websites use to figure out whether they are dealing with a human or a machine. You are shown a picture cut into tiles and asked to mark the ones with bicycles, or to type in ornately distorted digits and letters. Beyond their primary job (a Turing test), those answers later go on to train AI.

There is also unsupervised learning, where the system finds patterns on its own, without human feedback. These are the toughest projects to pull off—but they can crack the hardest, most creative problems. ChatGPT and similar solutions are the product of exactly this kind of training—or, more precisely, of the variety known as self-supervised learning: the model hides part of a text from itself and learns to guess what was hidden, grading its own work as it goes. No manual labeling is needed here—but what the data consists of, and how good it is, decides everything.

More precisely, in practice it is a hybrid approach: first the machine learns on its own from enormous datasets, and then people correct the result—scoring

answers, labeling the good and the bad, “fine-tuning” the model with feedback (the industry calls this RLHF—reinforcement learning from human feedback). The machine does the rough work; the human grades it.

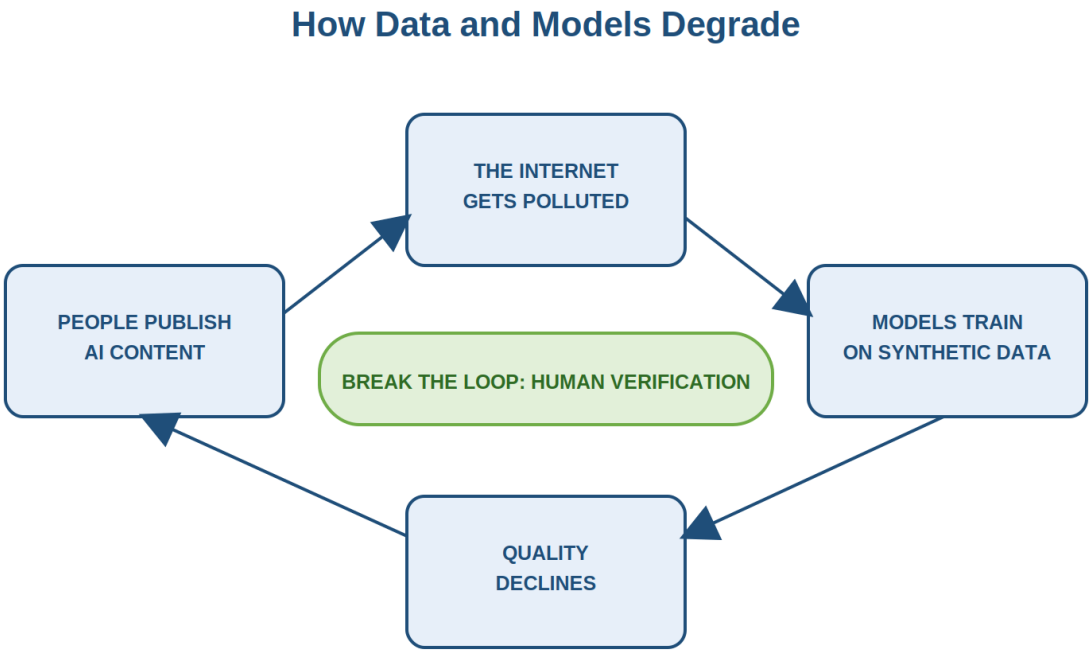
### The Shared Weaknesses of Today’s AI Solutions

At the current stage of development, all AI-based solutions share a set of fundamental problems.

- **The volume of training data.** Neural networks need enormous amounts of quality, labeled data. Where a human learns to tell dogs from cats after a couple of examples, AI needs thousands.
- **Dependence on data quality.** Any inaccuracies in the source data hit the result hard. Neural networks simply absorb data—they do not analyze facts or how they connect.
- **Ethics.** AI has no ethics—only mathematics and a task to complete. Hence the hard ethical dilemmas: whom should an autopilot hit when there is no way out—an adult, a child, or an elderly person? Debates like that never end. For AI there is no good and no evil—and no “common sense” either.
- **Content quality and reliability.** Neural networks cannot judge data for truth or logic, and they are prone to generating low-quality content and AI hallucinations. Errors abound—and they breed two problems.

The first is the degradation of search engines. AI has produced so much low-quality content that search results began to buckle: junk now dominates them. SEO optimizers stuffing pages with popular queries “help” things along in a big way.

The second is the degradation of AI models themselves. Generative models use the internet for further training. People apply AI without checking its output and fill the internet with low-quality content—and AI learns from it all over again. A vicious circle, with ever bigger problems at the end of it.




An article on this topic is also available via the QR code and hyperlink

*The AI feedback loop: Researchers warn of 'model collapse' as AI trains on AI-generated content*

**But there is an important caveat.** The circle closes not because a machine wrote the text. It closes because there is nothing to check that text against. Nobody catches the error—and it travels on, into the next model.

If the correct answer is known in advance, everything changes. The whole task can be invented—the problem and the answer to it. Then no human has to check the machine: you can see at once whether it solved the problem correctly. Data like this does not corrupt a model—it teaches it.

An example. In June 2026 OpenAI released GeneBench-Pro, a benchmark of 129 research problems in computational biology. Each one was constructed

artificially, and for each the correct answer is known in advance. A year ago the company's best model solved less than 5% of the problems on the previous version of the benchmark. Today the figure is 31.5%.

The difference between poison and cure comes down to one thing. Where there is a reference answer to check against, the data helps. Where there is none, that same vicious circle begins.

Google DeepMind, together with Jigsaw, took the problem seriously enough to study it: researchers examined some two hundred media-documented cases of AI misuse (January 2023–March 2024) and classified them. The main finding was unexpected: in nine cases out of ten, the attackers did not "hack" AI at all—they used it exactly as designed. They generated images of real people, fake evidence, and mass content on behalf of nonexistent personas. Nothing had to be broken.

Since then, this has stopped being fodder for academic papers and become a line item in the loss budget. Deloitte estimates that fraud losses involving generative AI in the US will grow from $12.3 billion in 2023 to $40 billion by 2027. More on how this works—and what to do about it—in the security chapter.

- **The quality of the "teachers."**

Almost all neural networks are trained by people—people who write the prompts and give the feedback. That imposes constraints of its own: who is teaching what, on which data, and to what end?

- **People's readiness.**

Expect massive resistance from those whose work the networks will take over.

- **Fear of the unknown.**

Sooner or later, neural networks will become smarter than we are. People fear that—which means they will slow development down and impose restrictions.

- **Unpredictability.**

Sometimes everything goes as planned; sometimes even the creators struggle to understand how their algorithms work. That unpredictability makes errors

extremely hard to find and fix. We are only beginning to understand what we ourselves have built.

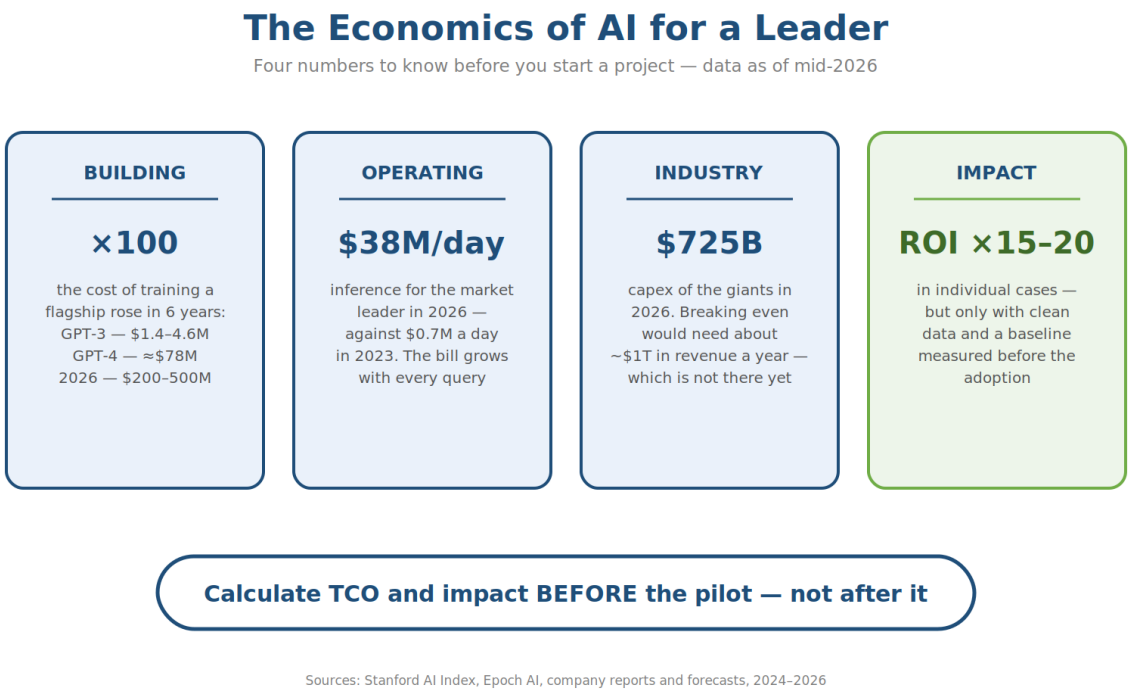
- **Narrow scope of activity.**

As of this third edition (2026), all mass-market AI is narrow AI (we will unpack the term in the next chapter). The algorithms are good at targeted tasks but poor at generalizing: an AI trained to play chess will not play checkers. Deep learning struggles with data that deviates from its training examples. To use even ChatGPT effectively, you need to be an expert in your field and formulate a deliberate, precise request.

- **The cost of building and running.**

Building neural networks takes serious money, and the clearest way to see it is to look across the years. Training GPT-3 in 2020 cost, by various estimates, $1.4–4.6 million—a sum a large company could well afford. GPT-4 in 2023 already ran roughly $78 million, Gemini Ultra about $191 million, and Llama 3.1 with its 405 billion parameters around $170 million (Stanford AI Index and Epoch AI estimates). Training a 2026-generation model is estimated at $200–500 million per run, and by the end of 2027 analysts expect the billion-dollar mark to fall.

The bottom line: in six years, the cost of training a frontier model has grown roughly a hundredfold. And the growth is steady—2.4–3x every year since 2016. This is the upper part of the S-curve expressed in money: every additional percentage point of quality costs a multiple of the one before.




Operations tell a similar story. Serving GPT-3's queries took more than 30,000 NVIDIA A100 GPUs, with an electricity bill around $50,000 a day. Today the count runs to hundreds of thousands of next-generation accelerators, and OpenAI's inference bill alone is estimated at tens of millions of dollars a day.

**What this means for you.** Exactly one conclusion: training your own frontier model (a top-tier model on the level of ChatGPT) is not your game—and that is not an admission of weakness but sober math. Your game begins where theirs ends: in the data, processes, and scenarios that no one but you will ever digitize. A model you can rent or buy; a competitive edge you can only build.

Let me repeat: these are the shared weaknesses of all AI solutions. We will come back to these weaknesses several times and work through them on more practical examples.

## Five Beginner Fears—Brief and Honest

### 1. "My data will be stolen."

The risk is manageable, and the rule is simple: never send a public service anything you could not show a stranger (more in Chapter 8). Meanwhile, many services are already rolling out encryption of user data—either in full or for personal data only.

**2. "AI will replace me."**

The working model is 70/30: AI takes the routine; decisions and responsibility stay with the human. It is not AI that replaces people—it is people who have learned to use it.

**3. "It's expensive."**

Getting started is free or close to it. What gets expensive is adoption without a goal—not the tool.

**4. "It's complicated, and I'm not a techie."**

The core skill is framing the task and checking the result (Chapter 6 is devoted to it). Studies from 2026 show that non-techies solve problems with AI almost as successfully as programmers do. In practice it sometimes works the other way around: someone who knows the domain well, speaks its language, and can judge the answer gets more out of AI than the IT specialists do.

**5. "It's too late to start."**

Mass adoption is only getting under way: three-quarters of workers have already tried AI, but only a few use it systematically. The head start is still there—and it belongs to the disciplined, not the early.

### Chapter Summary

- AI is not magic but mathematics built on statistics; it has three directions: neural networks, machine learning, and deep learning.
- AI pays off where there is data, high variability, and a task that can be automated; otherwise, classic automation is cheaper and more reliable.
- All AI solutions share fundamental limits: the volume and quality of data, unpredictability, and the cost of building and running them.

**What to do:** Before assigning a task to AI, run it through the "AI vs classic" test and check whether you have usable data.

**Reflection question:** Which of your tasks truly justifies AI—and where are you dragging it in "because it's fashionable"?

## Chapter 2. The Types of Artificial Intelligence

### By the Kind of Task They Solve

AI can be classified along several dimensions. Let's start with the kinds of tasks it solves.

- **Generative.**

Creates new content (images, text, sound, and so on) at the user's request: conversational chatbots, image, video, and music generation, recommendations. It gets its own section in this book.

- **Classifying.**

Assigns data to categories against set criteria. It learns from labeled data and uses various algorithms to make decisions: telling cats from dogs, checking a part against dimensional specs, spotting whether a person in a hazard zone is wearing protective gear, deciding whether an email is spam, assessing health status, and so on.

- **Predictive / recommender.**

Predicts future events or outcomes from available data and statistics: an imminent equipment failure, future crop yields, hereditary diseases. Recommender AI is sometimes treated as its own category—it does not just predict but recommends what to do—though I count it as predictive.

For those who want a technical deep dive (algorithm families and the tasks they fit), I recommend a good article on the kinds of machine learning—which class of algorithms works best where.

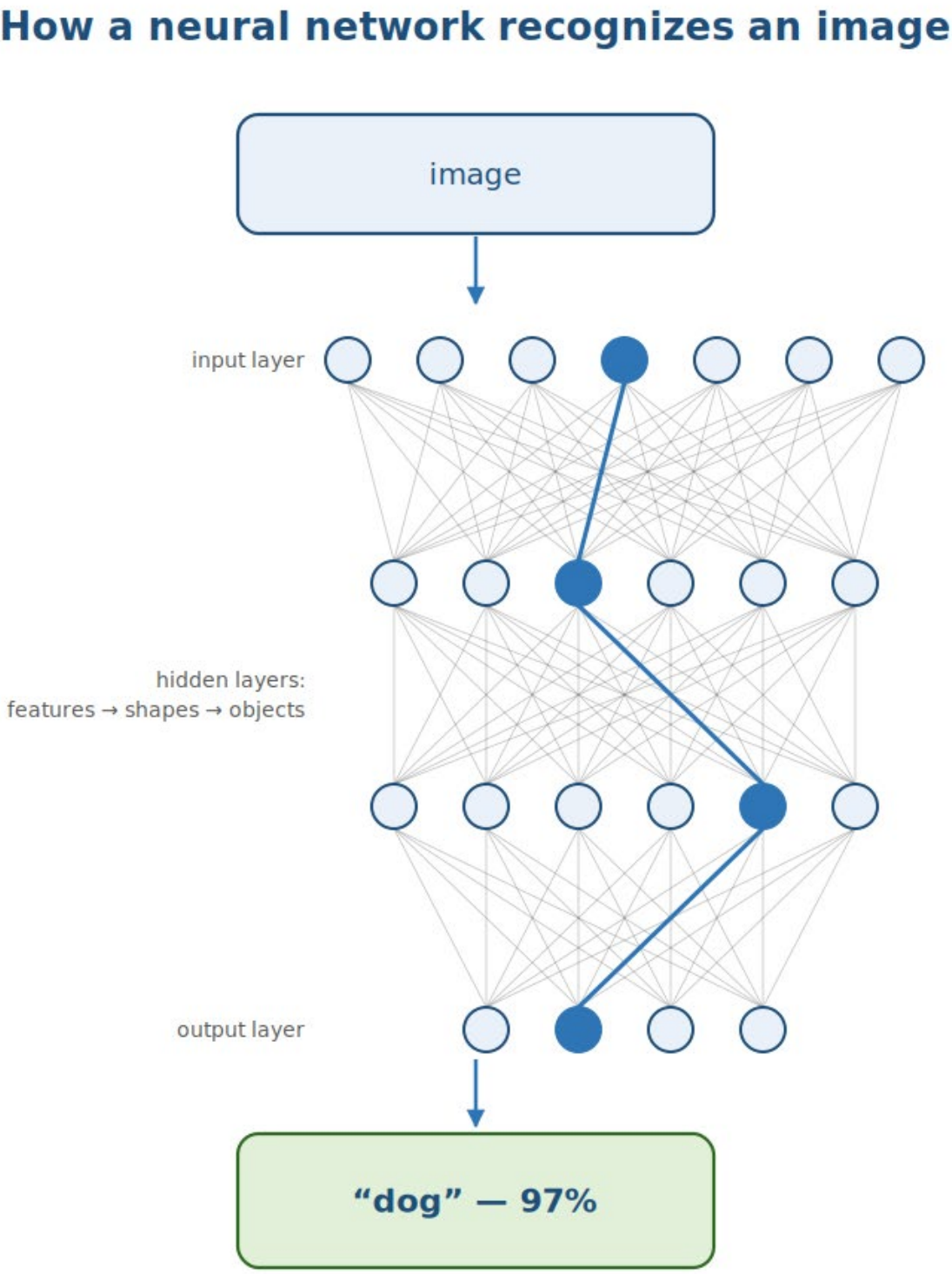
*Machine learning for humans*

## By "Strength" and the Handling of Uncertainty

Now for three concepts: narrow, strong, and superstrong AI.

### Narrow AI

Everything we see today is narrow AI (ANI). It solves the narrowly specialized tasks it was designed for: telling a dog from a cat, playing chess, analyzing and enhancing video or audio, advising within a subject area. But the strongest chess-playing narrow AI is useless at checkers, and an AI consultant on project management is useless for planning equipment maintenance.




An example of AI at work in pattern recognition

## Strong and superstrong AI

If plain “AI” is a muddled term with a broadly clear meaning, “strong” (or “general”) AI is harder still. Here is the definition that, to my mind, captures the essence best.

Strong, or general, AI (AGI) is AI that finds its bearings in changing conditions, models and forecasts how a situation will unfold, and—when the situation falls outside standard algorithms—works out a solution on its own. Think of solving the task “get into a university,” or learning the rules of checkers and starting to play them instead of chess. In other words: AI able to solve any intellectual task on par with a human, with universality instead of narrow specialization, true understanding instead of pattern imitation, and the capacity to learn, reason, and create.

What qualities would such an AI need?

- **Thinking**—deduction, induction, association, and the like, to extract facts from information and represent (store) them. This allows more precise problem-solving under uncertainty.
- **Memory**—different types of it, short- and long-term: tasks get solved with accumulated experience brought to bear. Today GPT-4 has a small short-term memory and gradually “forgets” where the conversation began. In my view, memory and the sheer “bulk” of models will become the key constraint on AI development.
- **Planning**—tactical and strategic. There is research showing AI can plan actions and even deceive humans in pursuit of its goals, but this is embryonic. The deeper the planning, especially under uncertainty, the more compute you need: a chess game 3–6 moves ahead with clear rules is one thing; a situation of genuine uncertainty is another.
- **Learning**—imitating what others do and experimenting for oneself. Today AI learns from vast datasets but does not model or experiment on its own. Learning requires long-term memory and complex interconnections—which, as we have seen, is a problem for AI.

No one has such strong AI today. Claims that it is imminent (in the 2024–2028 window) are, to my mind, mistaken or speculative. Then again, maybe my knowledge is too limited…

Yes, OpenAI's ChatGPT and other LLMs generate text, images, and video by analyzing the request and crunching big data: they look for the best-fitting associative combinations formed during training. But that is just math and statistics, and the answers come with plenty of "defects" and hallucinations. They are not yet ready for real interaction with the world. I said this in 2023, and I repeat it in 2026—while working with frontier models daily.

And yet, with only narrow AI in hand and a more-or-less strong one still a dream, researchers have already carved out a category for superstrong AI (ASI, Artificial Superintelligence). This is AI that:

- solves both routine and creative tasks
- instantly finds its bearings under uncertainty, even without an internet connection
- adapts the solution to context and available resources
- understands people's emotions (not just from text but from facial expressions, tone of voice, and other signals) and takes them into account
- can interact with the real world on its own to get tasks done

This is the AI we still see only in science-fiction movies. Even AI itself describes ASI as a "hypothetical concept" and "a subject of science fiction and active research." It is a point in the distant future we would like to reach—and cannot yet.

It would rely on advanced technologies—large language models (LLMs), multi-bit neural networks, and evolutionary algorithms. For now, ASI remains a conceptual and speculative stage, though it would be a giant leap from where we are.

And where narrow AIs today number in the hundreds—one per task—strong ones would number mere dozens (most likely split by domain, as we will see in the next section), and superstrong AI would be one per country, or one for the whole planet.

### The Obstacles on the Road to Strong AI

Honestly, I don't put much faith in strong or superstrong AI arriving anytime soon.

**First: it is an extremely costly and difficult undertaking in regulatory terms.** The era of unchecked AI development is ending; restrictions will keep piling up (a separate chapter is devoted to regulation). The key trend is the risk-based approach: strong and superstrong AI will land in the top risk tier—which means lawmakers' measures will be prohibitive.

**Second: it is technically hard—and strong AI would also be highly vulnerable.** Today, in the mid-2020s, building and training it takes gigantic computing power. According to Leopold Aschenbrenner, a former member of OpenAI's Superalignment team, it would require a trillion-dollar data center whose energy consumption would exceed all current US power generation.

Energy was also the theme Eric Schmidt, Google's former CEO, took up in spring 2025 before the House Committee on Energy and Commerce. An average US nuclear plant generates about 1 GW, while 10 GW data-center projects are already on the table. By one of Schmidt's estimates, data centers will need another 29 GW by 2027—and a full 67 GW by 2030. For scale: as of January 1, 2025, the total installed capacity of Russia's entire power system was 269 GW.

Strong AI would also need models orders of magnitude more complex than today's—and combinations of them: not just an LLM to parse requests but multi-agent solutions built on different principles (more on that later). That means exponentially growing the number of neurons, wiring them together, and coordinating the work of different segments and models.

And while human neurons can hold multiple states and fire "in different ways" (biologists, forgive the simplification), machine AI is a simplified model that cannot. Roughly speaking, a machine's 80–100 billion neurons are not equal to a human's: the machine needs more neurons for comparable tasks. GPT-4, per leaked technical estimates, holds about 1.8 trillion parameters (loosely, neurons)—and still falls short of a human. Even the frontier Fable 5, unofficially pegged at ten trillion parameters, falls short of humans where it matters most—and requires human oversight.

All of this comes down to several factors.

**The first factor is growing complexity.** Complexity always brings reliability problems: the number of failure points grows. Complex models are hard to build and hard to protect from degradation over time—they need constant "maintenance." Otherwise strong AI will start to degrade and its neural connections will decay: any complex neural network, if it is not developing, dismantles the connections it does not use. Maintaining connections costs energy, and AI will always optimize by switching off "unneeded" consumers of energy.

In other words, AI would come to resemble an old man with dementia, and its "lifespan" would shrink dramatically. Imagine what a strong AI with its capabilities could do while suffering memory loss and abrupt regressions into childhood. Even for today's solutions this is a live problem.

A couple of simple real-life examples. Building strong AI is like building muscle: progress comes fast at first, then efficiency drops, and you need ever more resources just to stay in shape. Strength grows with the muscle's cross-section, mass with its volume; at some point the muscle becomes so heavy it cannot move itself—and can even injure itself.

Another example comes from engineering—Formula 1 racing. A one-second gap can be closed with $1 million and a year. But clawing back the decisive 0.2 seconds may take $10 million and two years, and fundamental design limits may force a complete rethink of the car. The same goes for ordinary cars: modern ones cost more to build and to maintain, you can't so much as change a bulb without special equipment, and hypercars get a full crew of technicians after every outing.

From a development standpoint, two parameters are key:

- the number of neuron layers (model depth)
- the number of neurons in each layer (layer width)

Depth determines the capacity for abstraction: too little of it makes analysis and judgment shallow. Width determines the number of parameters/criteria at each layer: the more there are, the more complex the model and the fuller its reflection of the real world.

The number of layers affects the function linearly; width does not. Hence the muscle analogy: top LLMs in 2026 are approaching 10 trillion parameters (Fable, for one), yet models three to four times smaller show no critical drop in quality (China's Qwen 3.8 and Kimi K3 run 2.4–2.8 trillion parameters and beat Fable on some tests). What matters more is the data the model was trained on, whether it has a specialization, and other factors.

Below are statistics on LLMs from different makers.

| Model | Hallucination rate | Factual consistency rate | Correct answer rate | Average answer length, words |
|---|---|---|---|---|
| GPT-4 | 3.0% | 97.0% | 100.0% | 81.1 |
| GPT-4 Turbo | 3.0% | 97.0% | 100.0% | 94.3 |
| GPT-3.5 Turbo | 3.5% | 96.5% | 99.6% | 84.1 |
| Gemini Pro (Google) | 4.8% | 95.2% | 98.4% | 89.5 |
| LLaMa 2 70B | 5.1% | 94.9% | 99.9% | 84.9 |
| LLaMa 2 7B | 5.6% | 94.4% | 99.6% | 119.9 |
| LLaMa 2 13B | 5.9% | 94.1% | 99.8% | 82.1 |

Compare LLaMa 2 7B, 13B, and 70B: the 7/13/70 billion parameters loosely indicate the model's complexity and training—the higher, the better. Yet answer quality does not change radically, while the cost and effort of development grow substantially.

**An important caveat.** In February 2026, Vectara—the company behind the public hallucination leaderboard—replaced its test set: instead of short news pieces, models now summarize more than 7,500 long documents (up to 32,000 tokens) from the legal, medical, financial, and technical domains. The material now resembles what business actually works with. The numbers below come from this new, harder exam—and they cannot be compared with the old news-based measurements: these are different tests.

| Model | Type | Hallucination rate |
|---|---|---|

| | | |
|---|---|---|
| GPT-5.4-nano | compact | 3.1% |
| Gemini 2.5 Flash-Lite | fast, non-reasoning | 3.3% |
| Claude Sonnet 4.6 | reasoning | 10.6% |
| GPT-5.2 (high) | reasoning | 10.8% |
| Grok-4 | reasoning | over 10% |
| Gemini 3 Pro | flagship | 13.6% |
| Grok-4 Fast Reasoning | reasoning | 20.2% |

The conclusion turns the "pricier means better" intuition on its head. First, on long real-world documents everyone hallucinates: flagship reasoning models fabricate in more than 10% of answers, whereas on short news texts the same classes of models stayed in the 3–6% range. Second, the leaderboard is topped not by the giants but by compact, fast models. Third—most unexpectedly—"reasoning" models do worse on source-based summarization: the extrapolation that helps them solve problems turns into invention here.

Three practical consequences follow: choose the model for the task, not by its spot in the "most powerful" rankings; on long documents, RAG with source citations plus human review is mandatory; and do not transplant other people's benchmarks onto your processes—build your own test on your own typical tasks.

The same logic shows up within one maker's model lines: multiplying parameters does not multiply answer quality, while development cost and effort grow substantially.

We watch the leaders build out capacity, erecting data centers and scrambling to solve the power and cooling problems of these "monsters." Meanwhile, improving model quality by a notional 2% takes an order-of-magnitude increase in compute.

A practical example of maintenance and degradation—with human influence on full display. Any AI, especially early on, learns from human feedback (satisfaction levels, the original requests). GPT-4 used user queries for further training—to give more relevant answers and offload its "brain." In late 2023,

articles appeared saying the model had become “lazier”: refusing to answer, cutting conversations short, or replying with search-engine excerpts. By mid-2024, Wikipedia excerpts had become the norm.

One possible culprit is the simplification of user requests themselves (they keep getting more primitive). LLMs do not invent anything new—they try to figure out what you want to hear and adjust (developing “stereotypes” of their own), hunting for the best effort-to-result ratio and switching off unneeded connections. This is called maximizing the objective function. Just math and statistics. And the problem is not unique to LLMs.

So, to keep AI from degrading, you would have to load it with complex research while restricting the flow of primitive tasks. But release it into the open world, and the task mix will tilt toward simple requests. Think of yourself: do survival and reproduction really require development? What is the ratio of intellectual to routine tasks in your work? Do you need integrals and probability theory—or ninth-grade math?

**The second factor is data volume—and hallucinations.** Yes, models can be scaled many times over. But the GPT-5 prototype was already short of training data in 2024—it had been fed everything there was. A term was even coined: the data wall. And strong AI, meant to navigate uncertainty, simply will not have enough data at the current level of technology: you would have to collect metadata on user behavior, sidestep copyright and ethics constraints, and gather consents. Moreover, the more “all-knowing” the model, the more inaccuracies and hallucinations it produces; give a base model a specific subject area and answer quality rises—it is more concrete, makes things up less, and errs less.

One clarification is in order: the data wall is being worked around, and there is already more than one way to do it. But none of these workarounds supplies what is actually missing—new human text. What they all do is change the question itself: not “where do we get more data,” but “how do we make do with the data we have.”

| Technique | What it is | Who can use it |
|---|---|---|
| Clean up, do not collect more | The same body of data, stripped of junk and duplicates, delivers more than a dirty set twice its size | Companies |
| Your own data | Public text has run out. Sensor readings, archives, video, and process records have not—and no lab has them | Companies |
| Retrieve, do not memorize | Knowledge is not baked into the model in advance; it is pulled in at the moment the answer is generated | Companies |
| Smaller model, narrower task | Do not teach one model everything. A narrow model built for a specific job needs less data and makes fewer mistakes | Companies and labs |
| Invent the problem together with its answer | Data is generated artificially, complete with a reference answer. Works where the answer can be known in advance: biology, chemistry, statistics | Labs |
| Train on the outcome, not on text | No data is needed if there is an environment that shows whether it worked: the code ran, the tests passed | Labs |
| Think longer rather than know more | Instead of new knowledge at training time—more compute spent on the answer itself | Companies and labs |

Look at the right-hand column. Five of the seven techniques are available to an ordinary company today, and three of them are about order in your own data rather than about technology. That is what to do while the labs work on their own problem. It also shows why the wall did not stop the industry: the wall stands in front of public text only, and the ways around it come from several directions.

**The third factor is vulnerability and cost.** As noted, it would take a trillion-dollar data center consuming more than all US power generation—hence an entire complex of nuclear plants (wind turbines and panels will not do). Add that the model would be tied to its "base"—and one successful cyberattack on the power infrastructure would knock out the entire "brain."

Why must such an AI be tied to a center—could it not be distributed? First, distributed computing loses performance: the capacity is heterogeneous, busy with other tasks, and unstable. Second, it is vulnerable to attacks on communication channels and infrastructure. Imagine 10% of your brain's neurons suddenly switching off while the rest run at half power—again you get an AI that alternately "forgets" who it is and what it's for, and thinks slowly.

And if strong AI needed a mobile body to interact with the world, that is harder still: how do you power and cool it, where does the compute come from, and what about machine vision and sensor processing (temperature, hearing, and so on)? That means enormous demand for compute and energy. So it would be a limited AI with a permanent wireless link to the center—again a vulnerability: modern channels are faster but lose range and penetration and are open to electronic warfare.

One might object: take a pre-trained model and make it local—roughly what I myself propose for deploying local models fine-tuned to a subject area. Yes, that can run on a single server, but such an AI will be very limited, will stall under uncertainty, and still needs power and a network. This is not the path to humanlike superbeings.

All of which raises the question of the economic case for such investment—especially given two key trends in generative AI:

- cheap, simple local models for specialized tasks
- AI orchestrators that decompose a request into several local tasks and distribute them among models

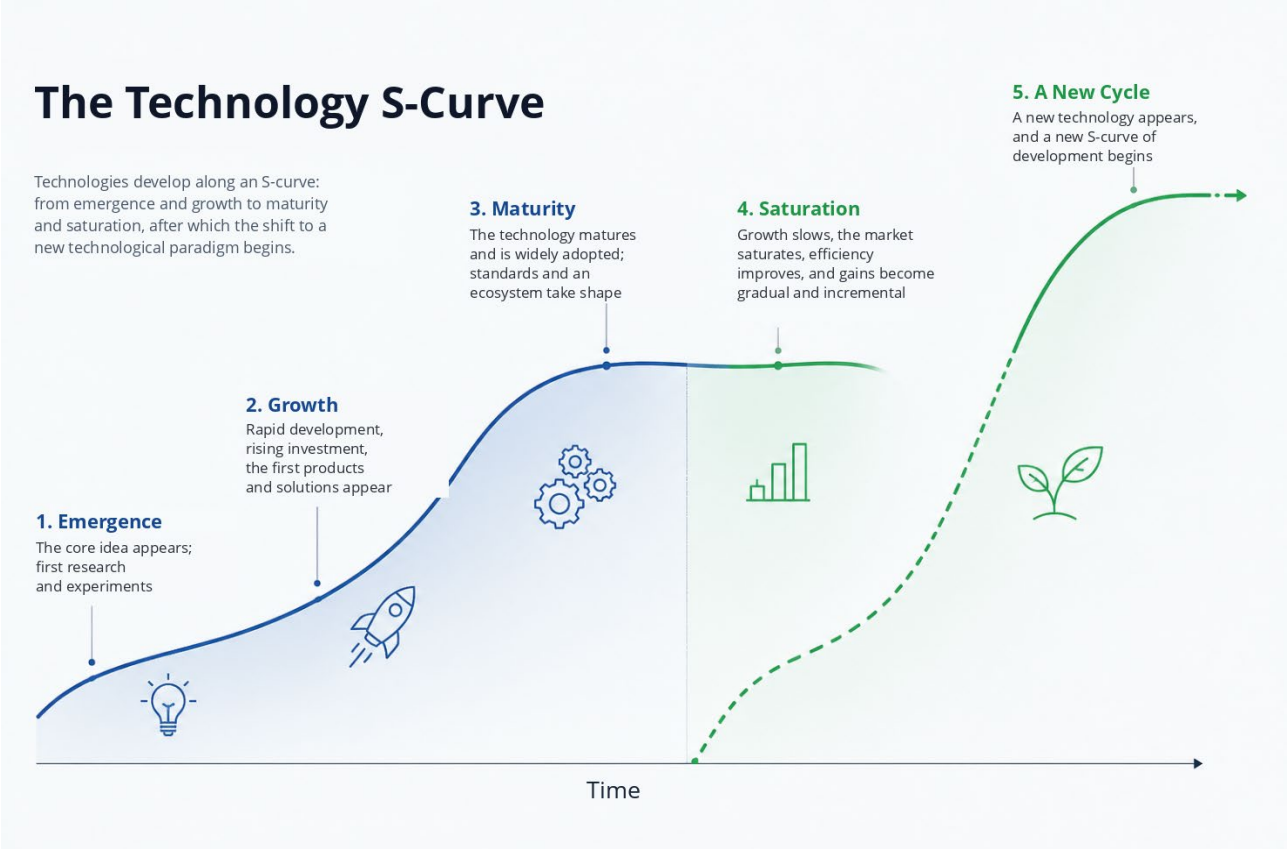
Thus narrow, specialized models will remain simpler and less restricted to build while still getting our tasks done. The result is a simpler and cheaper solution than building strong AI. Neuromorphic and quantum systems I will set aside—they get their own treatment in Chapter 10. Individual numbers and arguments here may be off, but on the whole I am convinced: strong AI is not a matter of the near future, and much of the talk about it is a way to hold attention and attract investment.

### The Author's Position

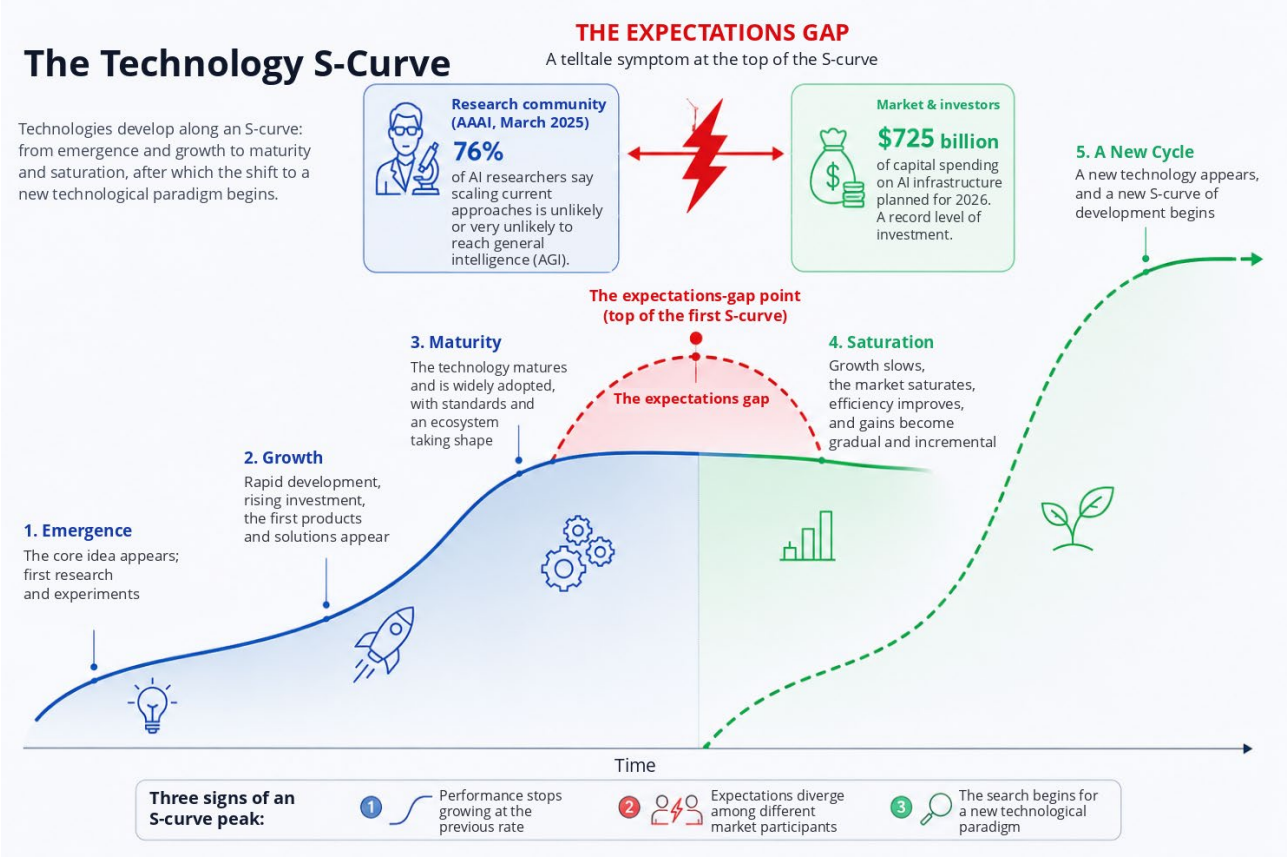
> **Strong AI is not tomorrow.** On a 3–5 year horizon, my bet is on specialized models and orchestrators: cheaper, more controllable, safer.
>
> By 2026, a regulatory ceiling had been added to the technical one: a frontier model's usefulness is its capabilities minus its regulatory constraints—and the constraints are now growing faster than the capabilities. From here on, it is about targeted fine-tuning of models and competition among "packages": what wins is not power but the product and the scenario (more in Chapters 7 and 9).

### The technology S-curve and the expectations gap

A useful frame to hold everything said so far in one picture: every technology develops along an S-curve. First a long flat stretch—ideas exist, results are scarce (neural networks lived like that for decades). Then explosive growth—for transformers it came in 2020–2024, when every doubling of compute delivered a visible jump in quality. And finally the plateau, where every next percentage point of quality costs a multiple—recall the "plus 2% quality, an order of magnitude more compute" ratio.



The top of the S-curve often comes with a telltale symptom—the expectations gap. An AAAI survey (March 2025, 475 AI researchers): 76% believe scaling current approaches is “unlikely” or “very unlikely” to lead to general/strong intelligence (AGI). The same year, investors penciled in a record $725 billion of capital spending on AI infrastructure for 2026 in anticipation of AGI. Science and money are looking at the same technology—and seeing different things. It happens at every bend of the curve: capital extrapolates the explosive phase by inertia, while researchers have already hit its limits.




**My position: we are near the top of the current paradigm’s S-curve.** This is not the end of development—it is a phase change: in the coming years, the main value will be created not by growing the models but by packaging already-achieved capabilities into products and scenarios (which is exactly what we see—from boxed solutions for SMBs to specialized assistants). The next technological leap will open a new S-curve—and it will most likely start with quantum and neuromorphic computing and with new algorithms beyond the transformer architecture. Those candidates are covered in Chapter 10.

To sum up, strong AI faces several fundamental problems:

- the exponential growth of development complexity and of the fight against degradation in complex models
- the shortage of training data
- the cost of creation and operation
- dependence on data centers and heavy demands on computing resources
- the low efficiency of current models compared to the human brain

Overcoming these problems is what will set the direction for the whole technology: either strong AI does appear, or we go down the path of narrow AIs and orchestrators coordinating dozens of narrow models. As things stand, strong AI squares with neither ESG nor ecology nor commercial success—it could only be built as a strategic national project with state funding. A curious fact: Paul Nakasone, the retired general who headed the US National Security Agency and Cyber Command until February 2024, joined OpenAI's board in 2024 (officially, to oversee ChatGPT's security).

I recommend reading Leopold Aschenbrenner's "Situational Awareness: The Decade Ahead" (available via the QR code and hyperlink).

*SITUATIONAL AWARENESS*

Below are the author's key claims, each with my assessment:

| Aschenbrenner's claim | My assessment |
|---|---|
| By 2027, strong AI (AGI) will become reality. | Disagree: my arguments are above. And there is still no shared understanding of the term AGI. |

| | |
|---|---|
| AGI is now a key geopolitical resource: every country will race to get it first, as with the atomic bomb. | Debatable. A great resource, but its significance is overrated—given the difficulty of building it and the inevitable errors in operation. |
| Building AGI requires a single trillion-dollar compute cluster (Microsoft is building one for OpenAI). | Beyond the compute, you need spending on power, people, and solutions to the fundamental problems. |
| The cluster will consume more electricity than all US generation. | Covered above; add investment in power generation and new risks. |
| The funding comes from tech giants: Microsoft, Google, Amazon, and Meta have penciled in a combined $725 billion of capital spending on AI infrastructure for 2026—77% above 2025's record $410 billion (Statista, February 2026). | There is no getting by here without state funding and involvement. |
| Goldman Sachs estimates annual AI infrastructure investment will grow from $765 billion in 2026 to $1.6 trillion by 2031—$4–8 trillion cumulatively over five years. | An excellent observation. The question is whether it is economically justified. |

Despite his optimism on timing, Aschenbrenner himself notes several constraints:

- a shortage of compute for running experiments.
- fundamental limits tied to algorithmic progress.
- ideas are getting harder to find, so AI researchers (agents doing research in place of humans) will likely only sustain the current pace of progress rather than multiply it. Still, the author believes these obstacles can slow, but not stop, the growth of AI intelligence.

Let me close with one more fact. Sam Altman, head of OpenAI, began saying in 2025 that the term AGI is too vague and should be replaced. And OpenAI co-founder Ilya Sutskever argued in an interview that the brute-force path of AI development—simply piling on more compute, data, and parameters—is fast exhausting itself. Usable training data is close to its limit, and further

scaling will not produce a breakthrough or strong AI. By his reckoning, the industry faces a return to genuine research, where new scientific ideas, not just infrastructure, will play the key role. Sutskever stresses that today's large models still generalize far worse than people do—even with rigidly formalized training and huge example sets, a human learns more efficiently. Finding ways to close that gap is, to him, the main direction of further progress.

The interview is available via the QR code and hyperlink below.

*Ilya Sutskever—We're moving from the age of scaling to the age of research*

## Copilots, AI Agents, and Multi-Agent Systems

One more direction generating plenty of buzz: copilots, AI agents, and multi-agent systems.

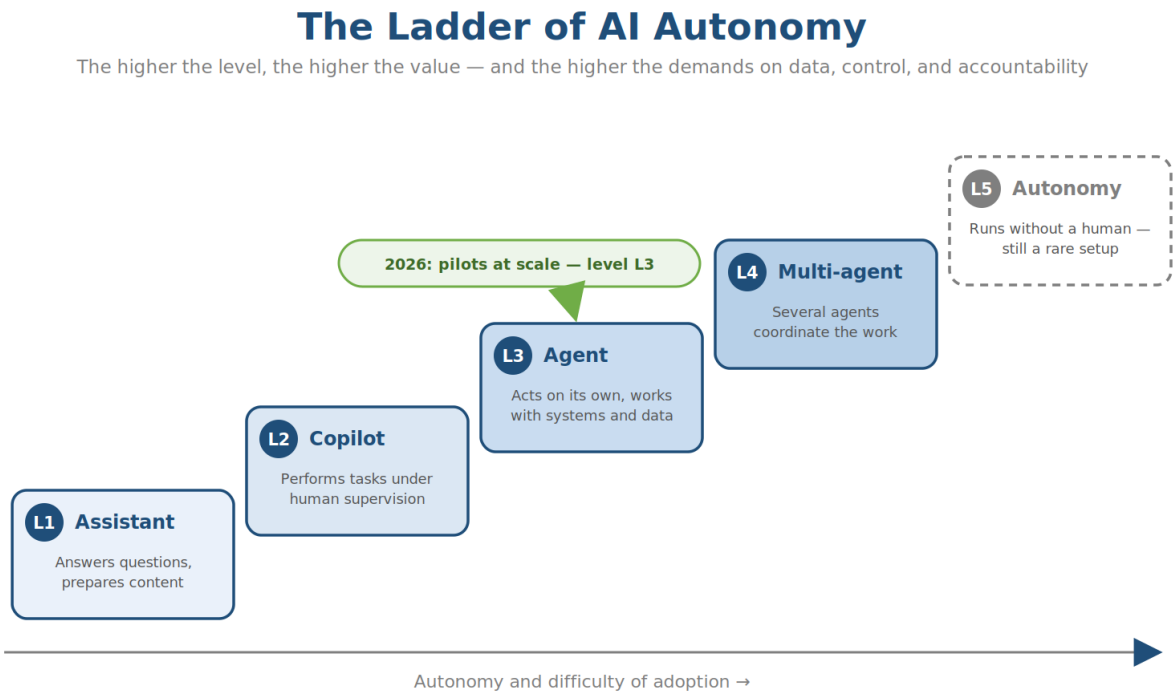
- **AI assistant**—an ordinary chatbot or service that works on request and only offers recommendations.
- **Copilot**—a solution that automates individual tasks under human supervision: the human says what is needed, and the AI writes the code or builds the presentation.
- **AI agent**—AI that works autonomously and interacts with other IT solutions or industrial equipment on its own. This extends both decision-making and interaction with the outside world.

**How to tell an agent from what gets called one.** The definition in the Chinese document on agents—I go through it a little further on, in the section on authority—is sharper than most marketing ones: an intelligent agent is a

system with autonomous capabilities of perception, memory, decision-making, interaction, and execution. Five capabilities. One caveat up front: the document does not state that all five are mandatory, that is my reading of the definition. But as a filter in a conversation with a vendor, it works.

- No memory between sessions—what you have is a chatbot with tools attached.
- No execution—an analytics layer.
- No autonomous decision-making—a scripted bot with branches.

Notice where the line actually falls inside the ladder above. The assistant is cut off immediately: it has neither execution nor memory. The copilot, though, does have execution—it writes code and builds presentations. What it lacks is exactly one capability: deciding for itself. That is what separates a copilot from an agent, and that is what determines how much authority you are prepared to grant it.




For example, a chatbot on a bank's portal asks what the client is after—bill payments, rates, or a credit-limit increase—then, through guiding questions and data drawn from documents and the profile, offers a solution. If known methods fail, it hands the question to the right employee. IT support desks are

evolving the same way: a consultation first, and if no solution is found, the task is routed automatically to the right person with the full history attached.

- **Multi-agent systems**—the same as agents, except now you have a combination of AI agents interacting with one another; you balance them and resolve their "conflicts."

Research shows the problems in such systems are identical to those of designing an organization and the interactions of its people. A number of factors have to be worked through:

- the set of parameters and resources to interact with
- the model of roles and interaction—who works with whom, why, and with what powers
- organizational rules—the permissible combinations of roles within a single agent
- organizational structure—the topology and the management model

In other words, to design a multi-agent system you must learn to design human interaction. The potential is enormous—and so are the risks: the consequences can be critical.

When AI agents and multi-agent systems come up, you often hear a seductive image—the "digital workforce." Supposedly you can hire an AI agent, issue it a virtual employment record, set KPIs, and even pay it bonuses for results.

For now, that is a metaphor. Behind it lies something simple but important—a new level of automation. AI greatly expands automation's reach and partly simplifies it. But issuing an employment record to a robot today is impossible—legally or technically. That said, two factors make the metaphor less and less far-fetched.

**First,** AI agents in complex multi-agent systems really are starting to behave like people in organizations: they compete for resources, they clash, their behavior needs balancing. As we discussed, designing such systems already calls for approaches familiar to managers: role definitions, interaction rules, org structures. So configuring KPIs and a "motivation system" for agents is not fantasy but a task that will become practical within two to three years (the

conflict between agents itself, and the need for an arbiter, is already a task for today—see the case in Chapter 8). It is just that "motivation" here means not bonuses but priorities, resource limits, expanded access, and escalation rules.

**Second,** the topic is already drawing lawmakers' attention. Regulation gets its own treatment in Chapter 7, but let me mention it here: several countries are discussing the idea of social contributions for the use of physical robots—by analogy with the payments employers make into pension and social-insurance funds. These are only discussions so far, but the very fact shows the line between "tool" and "digital worker" is gradually blurring.

Now let me share a classification of AI I saw in Beijing's AI cluster (the Zhongguancun technology park). There, AI is divided into five levels:

- L1—chatbots and assistants, for example those answering from knowledge bases
- L2—reasoning models that use logic to work out complex answers to ambiguous questions
- L3—agents able to robotize a process and replace a human in a typical task (in mid-2025, even the best agents handled typical tasks error-free only 24% of the time; by 2026 the numbers have improved, but "autopilot" remains far off)
- L4—innovators, capable of analysis and of creating something new
- L5—organized AI, combining several models to solve non-typical, complex tasks (multi-agent systems)
- and, separately, beyond the scale—strong/general AI, able to replace a human and solve interdisciplinary tasks

But that is the grand vision. The reality of 2026: L3 agents have entered mass pilots, yet reliability is still a long way off—even L2 stumbles in places, and companies, with their bureaucratic inertia, are still working through basic adoption of L1.

Let me pause on one question: will AI agents have a "career path"? Since we are looking at a leveled classification, the natural question is whether an agent can, over time, "grow" from an L1 assistant to an L3 agent and beyond—

earning access to more complex tasks and more autonomy. In short: yes—but that growth depends on more than technology. It runs on three tracks at once.

- **First—the development of the models.** Models keep getting smarter, faster, and cheaper. Every year brings a choice: use more efficient models for the same tasks, or trust the systems with harder ones. Essentially, this is the climb from L1 to L5 in the Beijing classification.
- **Second—the maturity of the business and its people.** How clear your processes are and how good and accessible your data is determine which tasks can be delegated to agents. If your processes are chaos, AI will work chaotically too. As autonomy grows, so do the risks: you need clear rules, boundaries, and people who know how to work with these systems. An important nuance: delegating tasks to AI asks of a leader not so much emotional intelligence (dominant today, since the work is with people) as logic, the ability to formulate tasks precisely, and process design. It is a different skill, and we will return to it in the chapter on competencies.
- **Third—risk appetite.** The more autonomy an agent has, the higher the stakes. You must be able to assess and minimize these risks—otherwise "promoting" the agent to more complex tasks becomes dangerous rather than useful.

So the "career growth" of AI agents is not a question of technology alone but of simultaneous progress on three tracks: the technology, the maturity of the business and its people, and risk appetite. And it is the maturity of the business and its people that most often becomes the bottleneck.

And there is one more thing worth understanding about agents: why they stumble in places a human walks through without a second thought.

**The arithmetic of long chains.** Imagine an agent running a hundred steps in a row and getting only one in a hundred wrong. An accuracy of 99% is an excellent number—if you look at a single step. Now do the math on the whole chain: the odds of getting through it without a single error are about 37%. Two attempts out of three break down somewhere in the middle. Raise accuracy to 99.9% and nine tasks out of ten make it to the end. Drop it to 95% and fewer than one task in a hundred reaches the finish.

That is why so many companies send agents back "for rework" after the pilot: in the demo the agent ran five steps, and in the real process there are fifty. And why L4–L5 remain exotic for now—the longer the autonomous route, the smaller the chance of getting to the end without a human.

By the way, it is now clear why programming became the first field where agents really took off. Everything there is digital, any change can be rolled back, and the result is checked automatically—by tests. Three conditions that most business processes simply do not have.

The practical takeaway is simple: do not hand an agent the whole long route at once. Break it into short stretches with a check at every joint—an error at a step boundary takes you a minute to fix, whereas an error at the end has to be fixed together with all the work built on top of it. There is more on this in Chapter 6, on setting tasks for AI.

Tellingly, this logic has reached the level of government policy: in July 2026 Beijing authorities released a package of measures for developing AI agents in which first place goes not to how "smart" the models are, but to their ability to carry a real task through to the end. The wording sounds mundane, but behind it lies exactly the arithmetic we have just done.

### What the Agent Is Allowed to Decide

The ladder of levels answers the question of how far an agent goes without a human. But there is a second question, and in practice it matters more: what is the agent allowed to decide at all? The two get conflated constantly, yet they are different axes—and the risk lives on the second one.

I found the most practical framing for that second question in an unexpected place: a Chinese regulatory document. On May 8, 2026, three Chinese agencies—the Cyberspace Administration, the National Development and Reform Commission, and the Ministry of Industry and Information Technology—issued the Implementation Opinions on the Standardized Application and Innovative Development of Intelligent Agents (智能体规范应用与创新发展实施意见), the country's first document devoted entirely to agents. The full text is published on the Cyberspace Administration's website.

The regulatory machinery in it does not travel: registries, mandatory standards, and product recalls rest on administrative capacity another country may not have. One construct, however, works in any jurisdiction, because it describes properties of the technology rather than the design of a state.

The document requires that an agent's authority be split into three categories.

- Decisions only a human makes. The agent does not perform them under any settings.
- Decisions that require approval. The agent prepares, the human signs off.
- Decisions the agent makes on its own—within explicitly granted authority.

| Category | Typical marker |
|---|---|
| **Human only** | Changes what the company owes third parties |
| **With approval** | Reversible, but the fix is expensive |
| **Autonomous** | Reversible, and the error shows up at once |

The marker in the right-hand column is a working hint for cases where an action does not sort itself into a category on the first try.

Two conditions attached to this classification matter more than the classification itself. First, the agent's execution does not go beyond the authorization it was granted. Second, even in the third category the human keeps the right to know about the action and to step in. The difference between the second and third categories is not whether a human is involved but when: in the second the human approves before the action, in the third the human can undo it after. If that distinction is not drawn explicitly, everything slides into the second category—and the agent stops saving the time it was brought in for.

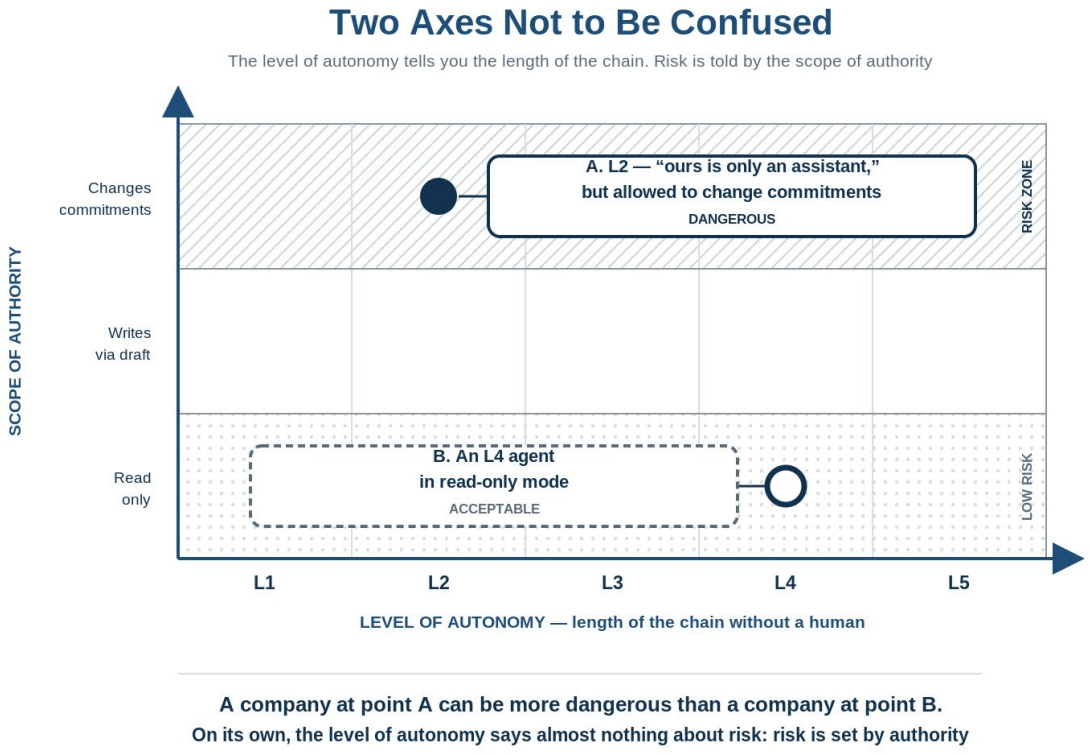
All of this sounds obvious right up to the moment you try to sort a real process into these categories. "Approve a payment up to fifty thousand"—which category is that? What about "send a client an email in your name"? Or "close a ticket in the system"? The exercise takes an hour and removes most of the

arguments you would otherwise have later. Above all, it happens before deployment rather than after the first incident.

**Two axes not to be confused.** Level of autonomy and scope of authority are different things, and mixing them up is expensive.

- **Levels L1–L5** measure the length of the autonomous chain: how many steps the agent takes without a human.
- **The three categories** measure the scope of authority: what the agent is allowed to decide.

A company at L2—"ours is only an assistant"—may still have given its agent the right to change what the company owes a customer. That is more dangerous than an L4 agent in read-only mode. On its own, the level of autonomy says almost nothing about risk: risk is set by authority.




*Two axes of risk: level of autonomy and scope of authority*

Hence a working default that is cheap at the start and nearly impossible to retrofit.

| Action on data | Default mode |
|---|---|
| **Read** | Free |

| | |
|---|---|
| **Write** | Only through a draft with human approval |
| **Delete** | Prohibited entirely |
| **Action log** | Always on; the agent cannot switch it off |

And a rule on top of this matrix: the agent never reaches the database directly—only through a software layer, which is where the rules above actually live.

One last item from the same document—four verbs worth putting to any vendor: how does the system detect an incorrect action by the agent, how does it intervene, how does it block, and how does it restore state? No clear answer on even one of the four means the pilot is premature.

> **From My Practice**
>
> Choosing the architecture for my product BAEOS, I deliberately stopped at the copilot level (though it is multi-agent inside), with automation at 70–90%—not a fully autonomous agent: the assistant drafts decisions, plans, and letters, but a human confirms the final action. That is a management choice, not a technical limitation: the cost of error in a leader's tasks is high, and trust in a tool is built through controlled autonomy.

## Chapter Summary

- AI is classified by task (generative, classifying, predictive) and by "strength" (narrow, strong, superstrong).
- Strong AI is not a matter of the near future: data, energy, degradation, and cost stand in the way; the realistic path is narrow, specialized models run by an orchestrator.
- The "digital workforce" is still a metaphor, but configuring KPIs and rules for agents will become a practical task; agents' "career growth" depends on technology, business maturity, and risk appetite.
- Level of autonomy and scope of authority are different axes: sort the agent's actions into three categories (human only / with approval / autonomous) before deployment, not after an incident.

**What to do:** Do not base your plans on "imminent AGI"; bet on specialized models, and for agents put rules, resource limits, and escalation paths in place.

**Reflection question:** At what level of autonomy—assistant to multi-agent—are you truly ready to run AI in production?

## Chapter 3. What Narrow AI Can Do—and the Big Trends

### Narrow AI in Applied Tasks

As you'll have gathered by now, I favor using what already exists—and using it in ways that make economic sense. Maybe that's my crisis-management background talking, or maybe I'm simply wrong. So where can today's narrow, machine-learning-based AI actually be put to work?

The most relevant directions:

- forecasting and preparing recommendations for decision-making
- analyzing complex data with no obvious interconnections, including for forecasting and decisions
- process optimization
- pattern recognition, including images and voice recordings
- automating individual tasks, including through content generation

The peak of popularity in 2023–2024: pattern recognition (image and voice) and content generation—that's where most developers are headed, and those services are the most plentiful.

The combination that deserves special attention is AI + IoT (the Internet of Things) + wireless connectivity + cloud computing + big data:

- AI gets clean big data, free of human error—for training and pattern discovery.
- IoT becomes more effective: predictive analytics and early detection of deviations become possible.

A textbook example of big data plus AI is BlackRock, one of the world's largest asset managers (about $14 trillion under management at the end of 2025, with third parties managing another ~$25 trillion on its Aladdin platform). Its

technological core is Aladdin: an integrated investment and risk-management system where AI is used to clean and interpret data, automate analytical commentary, personalize recommendations, and assess risk, liquidity, and scenarios. In the 2008 crisis, US regulators commissioned BlackRock to use Aladdin to analyze "toxic" assets and stress-test portfolios—which cemented the platform as an industry standard. Aladdin is an example of an AI-based recommender system.

### The Key Trends

Below are the industry's key trends. They are numbered for convenience; I will return to some of them in later chapters.

#### Trend 1. The falling barrier to entry

One of the problems developers are busy solving is making AI models and tools as easy to create as a website on a site builder—no special knowledge needed for basic use. Data science already runs as a service (DSaaS, Data Science as a Service). You can start with free versions of AutoML or with DSaaS—with an initial audit, consulting, and data labeling (labeling can often be had for free). And with large language models, almost anyone can build analytics and work with AI—the barrier to entry has dropped extraordinarily low.

The trend has data behind it—Yandex B2B Tech (2026): 86% of AI-agent users in Russia are small and midsize businesses, and a third of the agents were assembled with no code at all. An entrepreneur with a team of five configures an agent in a few hours on a ready-made platform—for tech support, HR consultations, or real-estate matching. A detailed look at these numbers is in the Postscript.

**The implication for business: managing adoption matters more than the technology itself.** The winner now isn't whoever has the bigger model or budget but whoever embeds AI into their processes faster and more carefully. What's critical isn't the technology—it's the management of its adoption: use-case selection, procedures, data quality, and control over results.

> **Myth vs Reality**

> **Myth:** "to get results, you need the most powerful AI."
>
> **Reality:** 86% of AI-agent users are small and midsize businesses on ready-made platforms. What wins is not model size but adoption management: the choice of use case, data quality, procedures, control of results—and the psychological readiness to experiment.

Without rules, employees build a shadow AI landscape (Shadow AI): dozens of people are already using ChatGPT, Claude, and other services for work tasks—without approval, a data policy, or an understanding of the risks. Banning it is pointless; the question is how to channel the initiative into something manageable (more on the risks in Chapter 8; on managing this at the adoption level—in the companion book, AI-2. Practical guide, Chapter 1).

### Trend 2. Models need less and less training data

A few years ago, faking a voice took a neural network an hour or two of recorded speech; a couple of years ago, several minutes; and in 2023 Microsoft unveiled a network that needs three seconds. Tools now exist that alter a voice even in real time. LLMs change the rules even more: modern models work well in few-shot and zero-shot modes—shown a couple of examples, or simply given a task described in plain text. Which means specialized AI solutions are now within reach without assembling enormous training datasets.

For business, this carries one more conclusion: the simpler the "magic" gets, the more important information security and legal constraints become—before you scale. If faking a voice once took hours of recording and now takes seconds, fraud risks are growing ahead of the curve (Chapter 8 is devoted to this).

### Trend 3. Local deployment and the "orchestra of models"

The next direction: models are being compressed to run on less powerful devices—laptops, smartphones, cars. So investment in IT infrastructure keeps paying off for longer: the same capacity can host several models. And developers keep shipping more tools for local deployment.

Optimization, compression, and the new tooling make local deployment a practical alternative to the cloud—especially where data confidentiality,

response speed, or regulatory requirements are critical and frontier models are unnecessary (a complex task, for example, can be decomposed into several simple ones).

In parallel, the "orchestra of models" architecture is gaining ground: instead of one big universal model, a set of compact specialized ones, with an AI orchestrator above them decomposing the request and farming out the tasks. The result is comparable to the giant models at a fraction of the infrastructure cost. As we discussed in Chapter 2, this—specialized narrow models run by an orchestrator—is the most realistic alternative to building strong AI, and the architecture can run on hardware you already have.

### Trend 4. Multimodality

Multimodality is ceasing to be a mere technical characteristic and becoming a requirement for next-generation business solutions. When a model processes text, audio, images, and sensor data at once, scenarios appear that are impossible with a single data type. Think of product quality inspection, where AI cross-checks IoT sensor readings against photos of the item and written standards; or tech support, where the operator points a camera at the fault and AI identifies the problem from the image and a spoken description, then proposes a fix. Multimodality is the bridge linking AI to other digital technologies—we will get to that combination in Trend 6.

### Trend 5. Decision support systems

Another active direction: industry-specific applied AI for decision-making, and recommender systems more broadly. And the trend is strengthening: where AI once helped analyze data, it now gives recommendations in the context of a specific situation, tailored to the user. In my own product, I tie them to the goals of the user and the organization as well.

Examples of scenarios where digital advisors already work or are being deployed:

- Project management—working through documentation, forecasting risks, recommending resource reallocation, flagging deviations from plan.

- Financial planning—scenario analysis of investments, credit-risk assessment, portfolio optimization (see BlackRock's Aladdin above).
- Operations—predictive equipment maintenance, logistics-route optimization, inventory management.
- HR decisions—forecasting attrition, recommending skill development, assembling teams.

More in Chapter 9, which is devoted to them.

**A practical example.** We'll return to this case more than once—it's my personal pain point and the product I'm building. Project management has a problem: about two-thirds of projects end up troubled or failed (Standish CHAOS, an analysis of 50,000 projects):

- schedule overruns—in 60% of projects, by 80% of the original schedule on average
- budget overruns—in 57% of projects, by 60% of budget on average
- success criteria missed—in 40% of projects

Meanwhile—and this is my observation from talking to leaders, not research data—managing projects and assignments consumes up to half of a leader's working time, and the share keeps growing. The reason: the world is getting more volatile, projects keep multiplying, and even sales are turning "project-like"—complex and individual. Where does such a track record lead? Reputational losses, penalties, thinner margins, capped growth. And the most typical, critical mistakes: fuzzy goals, deliverables, and project boundaries; a weak strategy and delivery plan; an inadequate governance structure; imbalanced stakeholder interests; ineffective communications.

How do people solve this? Either they do nothing and suffer, or they take a course and adopt task trackers. Both approaches have downsides: training offers live interaction and practice but is expensive and usually unsupported once the course ends; trackers are always at hand but don't adapt to a specific project or culture, don't build competencies, and exist to serve control. Analyzing my own experience, I arrived at the idea of a digital advisor—an AI that, in 10 minutes, gives predictive "what to do, when, and how"

recommendations for any project and organization. Project management becomes accessible to any leader for a modest monthly subscription. The model carries a project-management methodology and libraries of ready recommendations; the AI gradually self-learns and finds new patterns rather than staying tied to its creator's opinions.

### Trend 6. Combining AI with other digital technologies

Practice shows that AI delivers the most not in isolation but in combination with other technologies: with the Internet of Things (IoT), big data storage and processing, cloud and edge computing, wireless connectivity:

- **AI + IoT**—AI gets clean sensor data free of human error, while the IoT infrastructure gains intelligence for predictive analytics: early detection of deviations, failure prediction, energy optimization.
- **AI + Big Data**—big data supplies the raw material for training and pattern discovery, while AI turns those masses into concrete recommendations.
- **AI + cloud/edge computing**—the cloud provides scalability for training, while edge computing lets models run right on the equipment, with no delays and no network connection.

A telling example of this combination is China State Railway Group, which uses Big Data, IoT, and AI to run a rail network that carried over 4 billion passenger trips in 2024: AI models work with data from track and rolling-stock sensors and from traffic-control systems (details in Chapter 20). It is this combination that creates the most value for industrial and infrastructure businesses, where AI alone would be limited.

### Trend 7. AI products and process automation

AI is ceasing to be an experiment, a bare model, or a chatbot. It is being built into products and business processes. The key shift runs from AI assistants (answering questions) to AI agents (doing tasks autonomously) and on to multi-agent systems (coordinating several agents). We unpacked this evolution in Chapter 2; here we note it as one of the key trends.

| Level | What it does | Example | Maturity (2025–2026) |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Assistant | Answers questions, generates content | A chatbot on a bank's website | Mass adoption |
| Copilot | Automates individual tasks under supervision | AI writes code from a description | Active rollout |
| Agent | Works autonomously, interacts with external systems | Automatic request processing | Pilot projects |
| Multi-agent | Several agents coordinate among themselves | End-to-end process automation | Early experiments |

Each step up multiplies both the value and the risks—which is why security (Chapter 8) and the systems approach (Part 2) are critical when moving from assistants to agents. The main rule: do not skip steps. The path should be sequential: assistant → copilot → agent → multi-agent. The transition criteria: data quality, process stability, measurable results (baseline / KPI), a clear cost of error and a rollback mechanism, mature governance.

Trying to deploy autonomous agents outright, with no copilot experience, is one of the most common—and most expensive—mistakes. In my own product BAEOS, for example, I decompose the processes and embed AI into individual tasks: the complex task still gets solved, and everything stays under control.

Gartner confirms the trend's scale: by the end of 2026, 40% of enterprise applications will be integrated with task-specific AI agents—up from less than 5% in 2025; by 2035, agent technologies could bring in about $450 billion, or up to 30% of enterprise-software market revenue (Gartner, August 2025).

### Trend 8. From "playing with AI" to applied use

The experimentation period is ending. Business increasingly asks not "how do we try AI?" but "what return will it deliver, and how soon?" Judging by individual cases from my practice and publications, well-chosen and well-implemented solutions can deliver an ROI of 15–20x over 18 months—provided that:

- the right use cases are chosen (genuinely painful for the business, not fashionable)

- data quality is in place
- baseline and KPIs are locked in before the start
- the rollout accounts for processes, people, and infrastructure

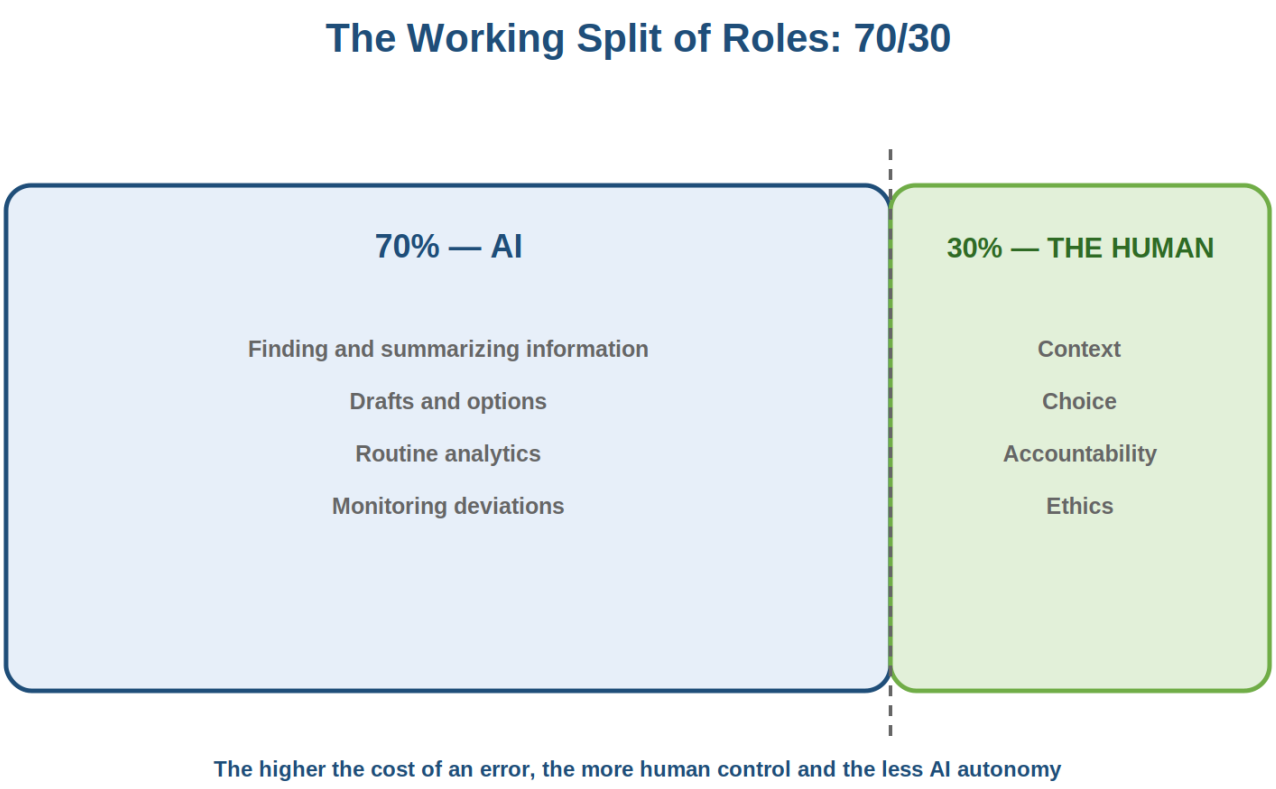
Companies that treat AI as another hype wave ("let's just try it," no goals, no measurements) end up disappointed; those that embed AI into specific processes with specific metrics get measurable results. How to run adoption projects and measure the effect is Chapter 22; the step-by-step methodology (prioritization, lifecycle, ROI math) is in the series companion, AI-2. Practical guide.

And most importantly: AI is not a standalone technology you can "bolt on" to a business. It works at the intersection of strategy, processes, data, people, and infrastructure. Without a systems approach, adoption turns into scattered experiments that do not scale. The systems approach means working several directions at once:

- **Strategy**—why the company needs AI and which business problems it solves.
- **Processes**—analyzing and preparing processes for automation, mapping AI integration points.
- **Data**—data quality, availability, and security.
- **People**—competencies, resistance management, a culture of working with AI.
- **Infrastructure**—computing capacity that meets information-security requirements.
- **Change management**—PR, communication, training.

That is what the entire second part of this book is about; the full methodology of systemic adoption is in AI-2. Practical guide. Without work on every one of these directions, even the best AI remains "a toy for the IT department."

### Trend 9. Human-machine synergy (70/30)




AI does not replace the human. A human who can work with AI replaces one who cannot—and practice keeps confirming it. Over a 3–5 year horizon, the most realistic model has AI taking on 40–70% of routine and analytical work while the human concentrates on decisions, communication, and creative, non-standard tasks. Even at 70% automation, 30% remains critically dependent on the human—and that 30% is what determines the quality of the result: checking AI's conclusions, interpreting them in the business context, final decisions, working with people.

By my observation—a practitioner's estimate, not research data—80–90% of companies that try to replace humans with AI entirely in complex processes run into degraded service quality. Not because AI is bad, but because tasks requiring context, empathy, and common sense are still beyond it. The trend is not replacement but new models of collaboration, where AI's strengths (speed, data processing, scalability) complement the human's (context, judgment, adaptability).

> **Myth vs Reality**
>
> **Myth:** "AI will replace people."

> **Reality:** the working model is 70/30: AI takes the routine; the human owns context, constraints, and the final decision. Attempts to remove humans entirely from complex processes usually end in degraded quality—and people brought back.

**The human's role:**

- sets strategic goals and priorities
- provides context and constraints
- handles what AI cannot (personal communications, agreements, refinement under constraints)
- makes the final decision and bears responsibility

**AI's role:**

- collects and structures data
- uncovers patterns and trends
- proposes solution options with risk estimates
- predicts the consequences of each option
- automates 40–70% of the task

For example, in project management BAEOS can take up to 80–90% of the preparatory work: quality criteria, the delivery plan, specifications, hypothesis generation, prioritization, recommendations. But delivery stays with the human: agreements, execution control, dealing with resistance, resolving conflicts. The same holds in the Communications and Decisions modules—AI will prep you for negotiations or work through solution options, but talking to people, owning the final decision, and carrying it out remain human work.

Synergy has a flip side—skill atrophy: per Wharton, 43% of executives warn of weakening employee competencies from over-reliance on AI—even as 89% believe GenAI strengthens work overall. The answer here is managerial, not technical: formalize the boundary ("AI helps—the human answers for it"), build in quality control, and teach AI literacy.

## The Three Layers of the AI Market: Whose Game Is Where

To close out the trends, here is a frame that dissolves the eternal question "are we catching up or pulling ahead?"—asked about countries and companies alike. The question only makes sense if everyone is running in the same lane. And the AI market is not one lane. It is **three layers with fundamentally different economics**.

**Layer one—frontier foundation models.** This is a game of capital and energy: the stakes run to tens of billions of dollars, gigawatts of connected power, and hundreds of thousands of accelerators. The players number a handful worldwide, and the list barely changes. Competing head-on in this layer is pointless for the vast majority of countries and companies—and that is fine: Germany has no frontier model of its own, and no one calls German engineering backward.

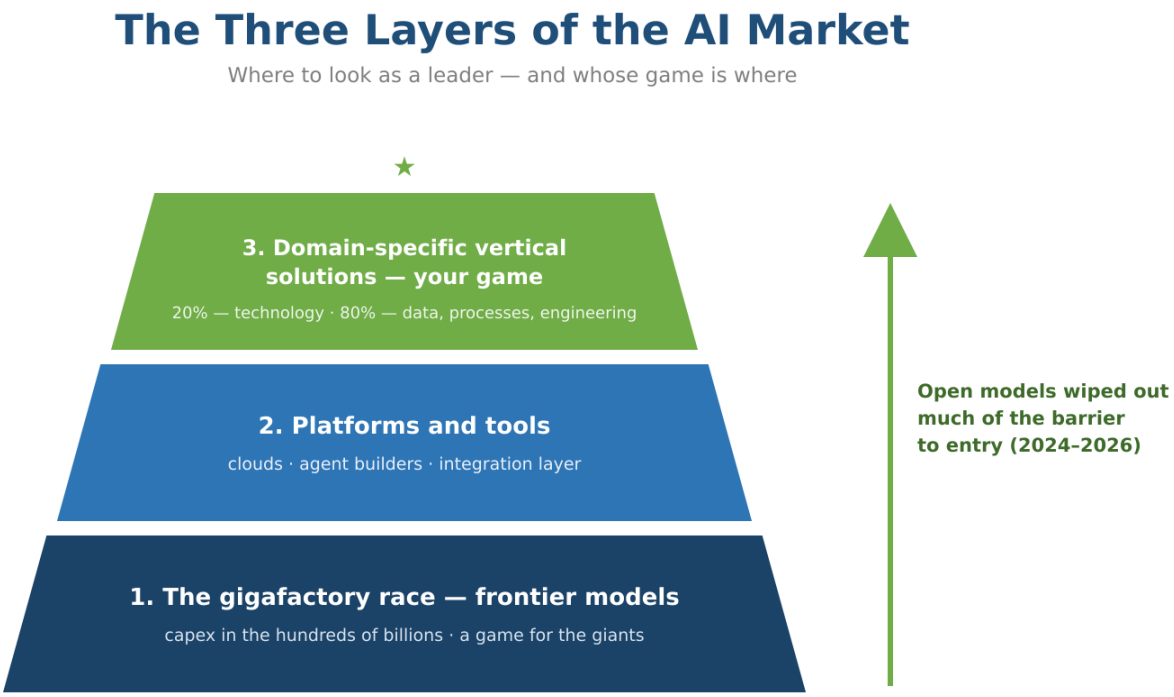
**Layer two—applied platforms and tooling:** MLOps, vector databases, agent orchestration, fine-tuning tools. Competition here is global, and much of it has already been won by open source—for everyone at once. Building long-term advantage here is pointless: you are investing in what will be free in six months.

**Layer three—domain-specific vertical solutions.** Machine-vision defect detection, predictive maintenance for a turbine, a technologist's assistant that knows your machine fleet and your tooling. The winner here is not whoever has more compute but whoever has three things: access to real production data, an engineering tradition in the industry, and the right to experiment on a working site.

And the main shift of 2024–2026, one few people say out loud: **open models have wiped out much of the barrier to entry at the base-model level.** In quality they've pulled nearly even with proprietary ones, and the model itself has stopped being a competitive advantage. The advantage has moved—into data (its very existence, built on industrial automation and sensors), into integration with workflows, and into the ability to prove the effect.

Two conclusions for a leader. First: do not confuse the layers. News about the gigafactory race is layer one; it has almost nothing to do with your technologist-assistant project. Second: your game is layer three. There,

technology is 20% of success—the other 80% is data, processes, and engineering. Which is why most of this book is about them.




## Chapter Summary

- The barrier to entry has collapsed (86% of agent users are small and midsize businesses, SMBs): the winner is not the most powerful model but adoption management.
- Without rules, Shadow AI grows; ROI of 15–20x (an estimate from individual cases) is achievable only with the right use cases, data quality, and a locked-in baseline/KPI.
- The realistic model for the coming years is 70/30 synergy: AI takes the routine and the analytics; the human takes decisions, context, and responsibility.

**What to do:** Take on a genuinely painful problem, not a fashionable one, and agree in advance on how you will measure the effect; introduce a Shadow AI policy. For the method of selecting and prioritizing use cases and calculating the effect, see AI-2. Practical guide.

**Reflection question:** Which 30% of your work must remain human?

# Chapter 4. Generative AI

## What Is Generative Artificial Intelligence?

> **In Plain Terms**
>
> Generative AI is an extremely well-read intern: it has "read" more text than any human, writes fast and with confidence, but knows nothing about your context and can be confidently wrong. Which is why the output is always checked by a leader.

Earlier we looked at the main areas where AI is applied:

- forecasting and decision-making
- analyzing complex data with no obvious relationships, including for forecasting
- process optimization
- pattern recognition, including images and voice recordings
- content generation

At the peak of popularity right now: pattern recognition (audio, video, numbers) and, on top of it, the generation of new content at the user's request—audio, text, code, video, images. Digital advisors can be counted as generative AI too.

## The Problems of Generative AI

The market leader's economics are the best illustration of what this race costs. In 2022, OpenAI lost $540 million developing ChatGPT, and at the time that figure looked impressive. By 2026 the scale had shifted by an order of magnitude: on revenue of roughly $25 billion annualized (about $2 billion a month), the company is projecting a loss of around $14 billion for the year—about $1.20 lost for every dollar earned. OpenAI's own forecast is even more candid: cumulative losses through 2029 of around $115 billion, with profitability somewhere in the early 2030s.

Look at the cost structure: compute for answering user queries alone (inference) will cost about $14 billion in 2026—some $38 million a day. For

comparison, in 2023 keeping ChatGPT running was estimated at $700,000 a day.

And here is the small detail that captures the whole moment. A few years ago, the head of OpenAI said that creating strong AI would take something like $100 billion. In 2026 the company is looking for $100 billion—not for the road to strong AI, but to fund its current operations.

Alexey Vodyasov, CTO of SEQ, confirms the broader trend: "AI is not delivering the marketing results people talked about earlier… hype and boom are inevitably followed by a decline in interest. AI will drop out of everyone's focus as fast as it entered it… AI really is a 'toy for the rich,' and for the near future it will stay that way." I agree: after the noise of early 2023, the quiet had already set in by autumn.

A Wall Street Journal investigation completes the picture: most of the IT giants have not yet worked out how to make money on generative AI. Microsoft, Google, Adobe, and others are hunting for ways to monetize it:

- Google plans to raise the price of subscriptions to AI-enabled software.
- Adobe is capping the number of AI-service calls per month.
- Microsoft wants to charge business customers an extra $30 a month for letting a neural network build presentations.

The cherry on top is the math done by David Cahn, an analyst at Sequoia Capital. Back in 2024 he calculated that to pay back the investment in AI infrastructure, the industry needs to earn about $600 billion a year. The logic is simple: take Nvidia's annual data-center revenue, double it (hardware is roughly half the total cost of owning a data center), then double it again—because somebody has to earn a normal margin on all that capacity.

By 2026 this arithmetic has become an order of magnitude harsher. Nvidia now earns $193.7 billion from data centers over a fiscal year—against $10.3 billion per quarter in 2023—and the second quarter of 2026 alone brought in $75 billion. Run those numbers through Cahn's formula and the bar is no longer $600 billion but close to a trillion in annual revenue that the industry has to find somewhere. The tech giants have penciled in about $725 billion of capital

spending for 2026, and analysts expect the figure to pass a trillion in 2027. Allianz Research puts the gap between the pace of investment and the pace of revenue at 46%—wider than at the peak of the 2001 telecom bubble (32%).

In other words, the reliable money in AI is still going to the seller of shovels, not to the prospectors. That is not a reason to gloat but a warning: a bubble this size does not deflate quietly. And this is exactly where the line runs between the investors' story and yours. A collapse of overheated expectations and market caps barely touches a company that has put an assistant into its support desk and counts the effect in person-hours. The ones who go under are the ones who were selling promises. The technology will remain, as the internet did after the dot-coms—and it will become cheaper, more accessible, and useful.

*AI industry needs to earn $600 billion per year to pay for massive hardware spend—fears of an AI bubble intensify in wake of Sequoia report*

And here is where it gets interesting for you and me. While the industry spends a trillion, what pays off for business is the exact opposite—and the market has already figured that out.

Compute is one of the biggest cost items: the more requests hit the servers, the bigger the bills for infrastructure and electricity. Hence the turn toward compact models. Every major developer is building them—Microsoft, Google, Apple, Mistral, Anthropic—and they are the ones that have become the workhorse of corporate adoption.

The numbers explain why. Large models like GPT-4 (over a trillion parameters, tens of millions of dollars to build) give no radical advantage on applied tasks. Compact ones are trained on narrow datasets, cost orders of magnitude less, and run on single-digit billions of parameters. In practice, running a small model costs 5–20 times less than calling a flagship over an API: a typical corporate scenario at 10,000 requests a day works out to hundreds or low thousands of dollars a month instead of tens of thousands.

It gets more interesting. Compact models have stopped being the "budget option with caveats." Microsoft's Phi-4, at 14 billion parameters, outperforms the 70-billion class on math and code; on industry-specific tasks, models with single-digit billions of parameters beat giants with hundreds of billions.

Experts confirm this, arguing that for many tasks—summarizing documents, generating images—frontier models are overkill. Illia Polosukhin, one of the authors of the foundational 2017 Google paper, said in an interview that using them for simple tasks is like driving a tank to the grocery store: "Computing 2 + 2 should not require quadrillions of operations." But let us take things in order: why it turned out this way, what limitations threaten AI, and what comes next—a sunset with a new AI winter, or a transformation?

**Polosukhin's metaphor got a sequel from the world of big money.** In July 2026, Bridgewater Associates—one of the world's largest hedge funds—and Mira Murati's Thinking Machines Lab published a telling study: a specialized fine-tuned open-weights model beat the flagship general-purpose models from OpenAI, Anthropic, and Google on six standard investment-analytics tasks, while using roughly 14 times less compute. On applied tasks, specialization beats scale—and the economics favor an "orchestra" of compact models.

Add to that the research cited earlier, where compact models rank above the "reasoning" flagships in hallucination ratings. The secret is not size but what data the model was trained on and what it was tuned for.

The practical takeaway. Start choosing a model not from the question "which one is the most powerful" but from "which is the smallest one that solves my task." You can test this in a week: take twenty of your typical requests and run

them through a flagship and a compact model. If there is no difference in quality and the bill differs tenfold, the choice has been made for you.

### The Limitations of AI That Cause Problems

Earlier I laid out the "basic" problems of AI. Now let us look at what is specific to generative AI.

**Companies worry about their data.** Any business wants to guard its corporate data, and that creates two problems. First, companies ban online tools outside the protected network perimeter, and any request to an online bot is a call to the outside world (how the data is stored, protected, and used raises plenty of questions). Second, that caution constrains the development of any AI: everyone wants recommendations from trained models—predicting equipment failure, for instance—but nobody is ready to share their data. A vicious circle. One caveat: some organizations can already host models at the GPT-3/3.5 level inside their own perimeter, but those still have to be trained—they are not off-the-shelf solutions, and the security team will find risks.

**Development and upkeep are complex and expensive.** Building a "general" generative AI means tens of millions of dollars and a very large amount of data. The efficiency of neural networks is still low: where a human needs 10 examples, a network needs thousands or hundreds of thousands (although it does find patterns in datasets a human could never dream of). Because of that data constraint, ChatGPT "thinks" better in English than in Russian—the English-language segment of the internet is simply larger. Add electricity, engineers, maintenance, and hardware upgrades, and you get that $700,000 a day just to keep ChatGPT running. You can cut costs by stripping out everything unnecessary, but then what you have is narrow AI. Which is why most solutions on the market are essentially "GPT wrappers," add-ons to large models.

**Public concern and regulatory constraints.** Society is worried about the development of AI, and governments do not understand what to expect from it. Several factors are at play:

- **Impact.** Generative AI has proved that it creates content you can mistake for human work (texts, images, scientific papers), and it produces conceptual designs for microchips and walking robots in seconds.
- **Security.** Criminals are actively using AI: according to SlashNext (State of Phishing, 2023), in the year following ChatGPT's launch the number of phishing attacks grew by 1,265%.
- **Opacity.** How AI works is sometimes unclear even to its creators—for a technology at this scale that is dangerous.
- **Dependence on the "teachers."** Models are built and trained by people; the material for training narrow models is also selected by people.

All of this means the industry will be regulated and constrained. Recall the well-known letter of March 2023, in which experts demanded limits on the development of AI. Here I am only registering the barrier; what is already happening with regulation, and what to expect next, I take up in Chapter 7.

**The chatbot interaction model is flawed.** You have probably tried talking to chatbots and come away disappointed: a great toy, but what do you do with it? A chatbot is not an expert; it is a system that guesses what you want to hear and delivers exactly that. To get value out of it, you have to be the expert on the subject yourself. But if you are the expert, do you need generative AI? And if you are not, you will not get a solution—only generic answers. A vicious circle: experts do not need it, amateurs are not helped by it. So who pays? What you get is an expensive toy.

Beyond expertise, you also have to know how to frame a request—and very few people do. For a while there was even a whole profession, the prompt engineer (about $70 an hour), and even they could not land the perfect prompt on the first try (they do not have the domain expertise). A useful trick: you can ask the AI to ask you leading questions and then build the perfect prompt from your answers—we will come back to this in the chapter on language models. Does business need a tool that makes it dependent on very rare and expensive specialists, if ordinary employees cannot extract value from it? The market for a plain chatbot is not just narrow—it is vanishingly small. There are two exceptions: a chatbot as a supporting feature within a larger product

(preparing recommendations, analytical dashboards, where you state the requirements precisely) and AI applied inside specific business processes (HR consultations for new hires, tech support, data analysis).

**Poor-quality content and hallucinations.** Neural networks collect data; they do not analyze facts or whether those facts hang together. They follow whatever there is more of on the internet or in the database, with no critical assessment of what is written. That is why generative AI easily produces false or incorrect content, and AI hallucinations are a long-known feature.

It is fine when the cases are harmless. But there are dangerous errors too. One user asked Gemini how to make a salad dressing: the recipe called for adding garlic to olive oil and leaving it at room temperature. While the garlic was infusing, odd bubbles appeared; a check showed that the jar was growing the bacteria that cause botulism—poisoning by the toxin is severe and can be fatal. I use generative AI regularly myself, and more often than not it gives me a not-quite-correct and sometimes flatly wrong result: it takes 10–20 requests with insane levels of detail to get something sensible, and even that still has to be reworked. You have to check its work, so delegating a task to AI entirely is impossible—it is an assistant rather than a replacement. And again we arrive at the same place: you have to be an expert to judge whether the output is correct—and that is sometimes slower than doing the job yourself from scratch. How to manage inaccuracies (hallucinations) systematically at the adoption level is covered separately in AI-2. Practical guide (Chapter 15).

**Emotions, ethics, and accountability.** Without the right prompt, generative AI simply reproduces information, paying no attention to emotion, context, or tone—and communication breaks down very easily, adding conflict to the problems above. Questions of authorship and ownership of the content produced come up as well: who is accountable for false or harmful actions taken with the help of generative AI, and how do you prove authorship? Ethical standards and legislation are needed.

### So What Do We Do?

Companies are not about to give up large models entirely: Apple uses ChatGPT in Siri for complex tasks, and Microsoft uses OpenAI's model as the assistant

in the new version of Windows. At the same time, Experian (Ireland) and Salesforce (US) have already moved their chatbots to compact models and got the same performance at significantly lower cost and latency. The key advantage of small models is that they can be fine-tuned to specific tasks and data: in the words of Yoav Shoham of AI21 Labs, small models answer at one-sixth the cost of large ones.

A fresh example of the same shift comes not from a startup but from Microsoft itself. In July 2026 the company put its own MAI family of models into public preview: on its own figures, MAI-Voice-2-Flash in Dynamics 365 contact centers (customers include T-Mobile and EasyJet) cuts compute costs by up to 89% compared with OpenAI's models, and MAI-Image-2.5 in PowerPoint by up to 84%. The numbers come with a caveat: this is a vendor claim, not an independent measurement, and the ratio depends heavily on request volume, model size, and which pricing tier the comparison uses. But the direction is telling: a company that has put $13 billion into OpenAI is building its own models so that it need not pay for a frontier system where a specialized one will do. The conclusion for you is simpler than it is for Microsoft: if even Microsoft finds it uneconomical to run a frontier model on high-volume work—speech recognition in a call center, or images for a slide deck—then you certainly will. Use a frontier model for hard reasoning and one-off tasks, and a narrow model for anything that repeats thousands of times a month.

**Do not rush.** There is no reason to expect a sunset for AI: too much has been invested over the past 10 years and the potential is too large. I recommend recalling the 8th principle of the Toyota Way (the foundation of lean manufacturing and one of the tools in my systems approach): "Use only reliable, thoroughly tested technology." A set of recommendations follows from it: technology is there to help people, not replace them (it is often worth running the process manually first); a proven process is better than an untested technology; test in real conditions before adopting; reject technology that runs against your culture and threatens stability; and at the same time encourage the search for new paths, adopting proven solutions quickly. In 5–10 years generative models will be mass-market, cheap, and smart enough; they will reach the plateau of productivity, and everyone will

be using what generative AI produces. But pinning your hopes on AI and cutting headcount today is clearly going too far.

**Improve efficiency and safety.** Right now practically every developer is making models less demanding in the volume and quality of data, and improving safety—AI has to generate safe content and resist provocation.

**Adopt AI in the form of experiments and pilots.** To be ready for the arrival of genuinely useful solutions, you need to follow the technology, try it, and build competence. It is the same as with digitalization: instead of leaping into expensive solutions, play with cheap or free ones. By the time the technology goes mainstream you will:

- understand what requirements to write into commercial solutions (and a good spec is half the battle)
- see results in the short term, and with them the motivation to go further
- raise your team's digital competence, removing technical constraints and resistance
- rule out false expectations, and with them unnecessary costs, disappointment, and conflict

**Transform how the user communicates with AI.** This is the concept I am building into my own product: give the user ready-made forms and lead them through a scenario in which they simply enter values or tick boxes, and then hand that form, wrapped in a correct prompt, to the AI. The alternative is deep integration into existing IT tools—office applications, browsers, auto-responders. But that requires careful work on user behavior and requests, or their standardization: either the solution is no longer cheap, or you lose flexibility.

**Build narrow models for specific processes.** As with people, teaching AI everything is laborious and ineffective. Narrow solutions built on the engines of large models keep training to a minimum, the model itself is not too large, the content is less abstract, and there are fewer hallucinations. The analogy with people holds: greater success goes not to whoever knows everything but to whoever focuses and goes deep. Examples of narrow solutions:

- an advisor for project management
- a tax consultant
- a lean-manufacturing advisor
- a chatbot for industrial safety, or an assistant for a specialist
- a chatbot for IT support

Gartner forecasts that organizations will adopt small task-specific models three times more often than general-purpose LLMs. Gary Marcus also spoke about "the end of the LLM race" in 2025 (The New Yorker): "I don't hear a lot of companies saying that the 2025 models are much more useful to them than the 2024 models, even though they are better on benchmarks."

> **From My Practice**
>
> Architecture handles hallucinations better than pleading with the model does. In BAEOS, the AI works to formalized scenarios, relies on the user's own data (RAG), and shows what its conclusion is based on; at critical steps the final word belongs to a human. That is more expensive and harder to build (finding the balance between convenience and flexibility), but cheaper than dealing with the consequences of a confident fabrication.

## Chapter Summary

Generative AI is still in development, but the technology's potential is large. Yes, the hype will pass, investment will decline, questions about feasibility will be raised—but the technology will not go away. Back on June 16, 2024, Forbes published a piece titled "The AI winter: should we expect a drop in AI investment?"

*The AI winter: should we expect a drop in AI investment? (in Russian)*

The piece analyzes the cycles of AI "winters" and "summers" and cites Marvin Minsky and Roger Schank, who as far back as the 1984 AAAI meeting described the chain reaction that leads to a new winter:

- Stage 1. The inflated expectations of business and the public are not met.
- Stage 2. The media start publishing skeptical articles.
- Stage 3. Agencies and businesses cut research funding.
- Stage 4. Researchers lose interest and the pace of progress slows.

The forecast came true: within a couple of years the AI winter set in, and the thaw came only in the 2010s. We are now at another peak—it arrived in 2023 after ChatGPT was released (which is why I so often draw examples from this LLM in the book: one particular case, but a very easy one to grasp). The article then applies the Minsky–Schank cycle to the current situation.

**Stage 1. Expectations.** The AI revolution in everyday life has not happened yet: Google never did transform search (Search Generative Experience drew mixed reviews); voice assistants (Alexa, Siri) have gotten better, but you would hardly trust them with consequential decisions; support chatbots still understand requests poorly and irritate people with answers that miss the point.

**Stage 2. The media reaction.** A search for "AI bubble" returns pessimistic headlines: "The AI hype bubble is deflating" (The Washington Post), "From boom to burst, the AI bubble is only heading in one direction" (The Guardian), "A famed economist warns the AI bubble is collapsing" (Business Insider). My

own view: these articles are not far from the truth. The market is overheated—that is a fact, and by the middle of 2026 it is on the verge of deflating; the situation really does resemble the eve of the dot-com crash. But the overheating is in the expectations and the valuations, not in the usefulness of the technology: those who have learned to adopt it are already seeing results and growing. The wave of hype will recede, many will go bankrupt—and that is a normal stage in an industry growing up: the viable solutions and companies will remain, and the mass market will finally learn how to adopt them. That is how it went with the internet after the dot-coms, and that is how it will go with AI (more in Chapter 23 and the Postscript).

**Stage 3. Funding.** Despite the pessimism, you cannot say that funding is declining: the largest IT companies keep investing billions, and the leading conferences are receiving record numbers of submissions. Which means we are currently between the second and third stages of the slide into winter. But is a "winter" inevitable? Actually, no.

The article's key argument is that AI is now too deeply embedded in our lives for a new winter to begin:

- facial recognition systems in phones and on the subway use neural networks for identification
- translators such as Google Translate have improved enormously in quality since moving from classical linguistics to neural networks
- modern recommender systems use neural networks to model preferences

Especially valuable is the view that the potential of narrow AI is not exhausted and that, for all the problems of strong AI, narrow AI can deliver value—I fully agree. The next step in the development of generative AI is newer and lighter models that need less data for training. All that is required is patience and steady effort to learn the tool and build competence, so that its potential can later be used to the full.

### Key Takeaways

- Generative AI built on frontier models is still economically heavy, while compact specialized models are often a match for them on applied tasks.

- A chatbot without an expert delivers no value, and hallucinations require double-checking: AI is an assistant, not a replacement for a human.
- The technology is not facing a sunset, but pinning your hopes on it and cutting headcount is premature—it has to be mastered through experiments.

**What to do:** Try AI yourself—on your own tasks, in free or low-cost services: it is the cheapest way to understand the technology. In the company, start with off-the-shelf solutions—per MIT, pilots built on an external tool customized to the task reach deployment about twice as often as in-house builds—roughly 67% against 33%. Do not cut people because of GenAI, and invest in a good spec.

**Reflection question:** Which of your processes would benefit from a narrow generative model right now?

## Chapter 5. Large Language Models (LLMs)

As of mid-2026, nine out of ten AI projects are built on large or midsize language models. So the field is worth a closer look.

### The LLM Phenomenon

2023 was a watershed year for AI: commercial-grade LLMs reached the market and became available to everyone. What set them apart was that they arrived ready to work—trained in advance on vast volumes of text. The barrier to entry collapsed. Where you once needed a team and trained specialists, now any of us can ask for an analysis, build charts, and automate routine work.

> **In Plain Terms**

> An LLM—a large language model—is a program that has read its way through an enormous library of texts and now "continues" any text by guessing the next word. It guesses well enough that what comes out is a coherent answer, an email, or a plan.

The main benefit is exactly that pretraining: these are ready-made models. You can load large volumes of data into them, work with them, and connect them to the company's databases (we will get to RAG in a moment). In effect, AI changes the approach to analytics: for a whole class of tasks you can drop the complex, expensive custom models in favor of an LLM—for analysis, for reports with charts, for advice.

### What Is RAG?

One technology has changed the world of LLMs more than most: RAG. It is essentially a built-in fact-checker for AI. Ask a model about an exchange rate or today's date and on its own it has no idea—but it can reach out to external sources and pull in current data. RAG has three core components:

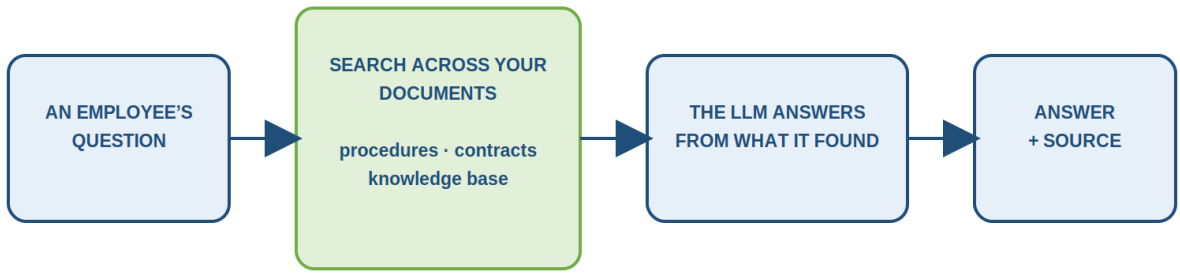
- **Retrieval:** the AI searches for data in trusted sources—the databases and documents it is connected to.
- **Augmentation:** the AI adds what it found to your query.
- **Generation:** the AI produces the final answer.

> **In Plain Terms**

> RAG is an open-book exam: instead of recalling the answer from memory, the model first finds the right page in your documents and answers from it. Which is why the quality of the answer depends on the order in your "library."

So the AI first scans the sources, builds its own description of what each one holds, and then uses them to compose answers. A vivid example is RAG in a bank processing a loan application: the account manager queries the AI, a request goes out to the credit bureau, and internal records are checked in parallel. The same pattern works in any line of business—a connection to the ERP, to the document management system. The simplest example of all is uploading documents into something like ChatGPT or DeepSeek. And one point matters here: using RAG and having the answer cite its sources is one of the mechanisms that let you check the AI and bring down the level of hallucinations.

**RAG: An Open-Book Exam**




The model does not "recall" — it finds the right page and answers from it

### From My Practice

RAG is the workhorse of corporate AI solutions. In BAEOS, it is RAG that lets the assistant answer from a particular company's projects and notes rather than from the internet's average opinion. The rule I have learned in

> practice: the quality of such a system is determined not by the model but by the order in your knowledge base—garbage in, confident garbage out.

### Fine-tuning

Another tool modern models offer is fine-tuning. For instance, you can feed the model your emails and documents so that it answers in your house style.

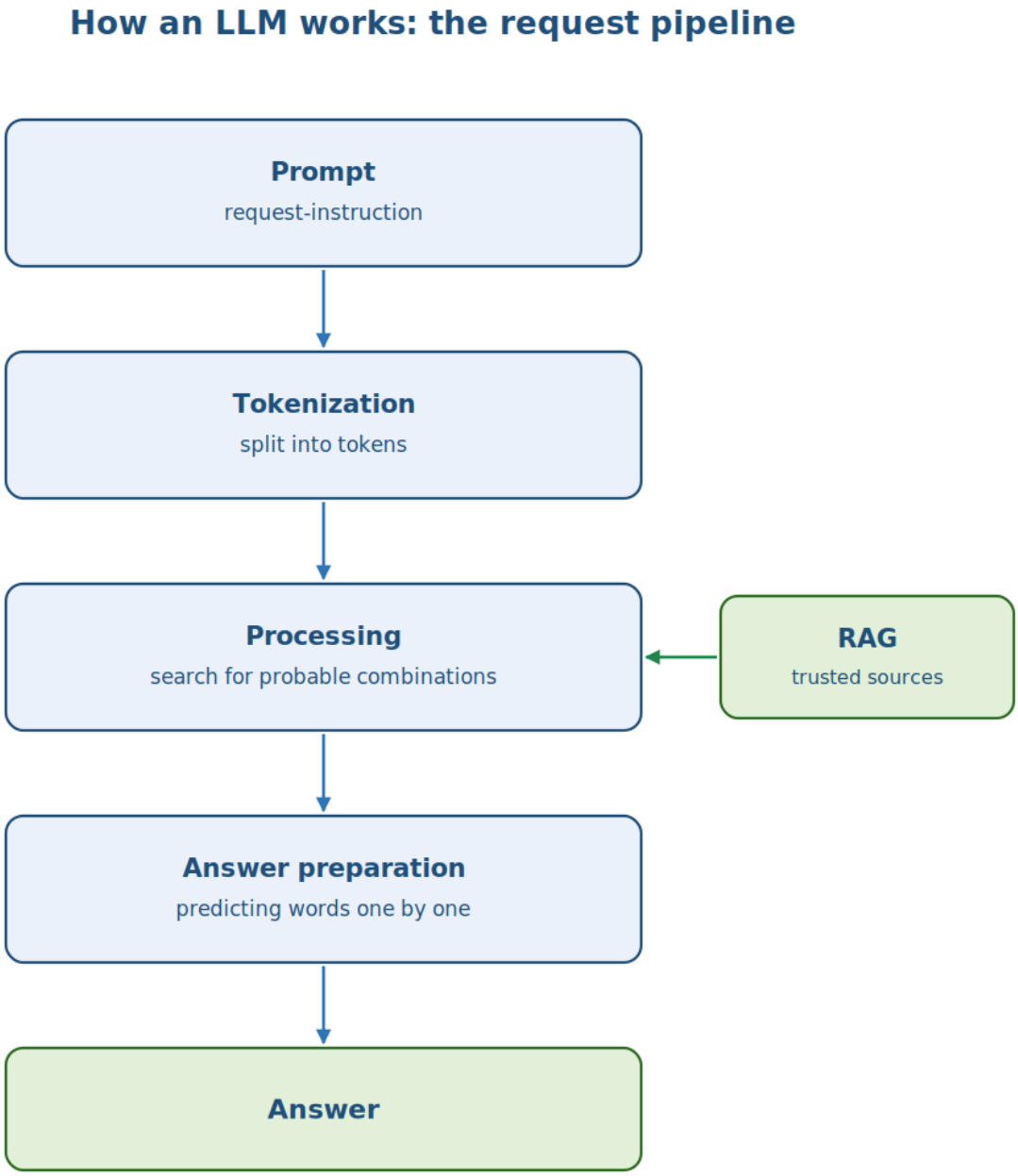
> **In Plain Terms**
>
> Fine-tuning is a professional development course: you do not retrain the model from scratch, you drill it on your own examples so that it writes in your style and knows your specifics.

The architecture of LLM-based solutions, and text instructions as a quality-control layer at the implementation level, are covered in detail in AI-2. Practical guide (Chapter 15).

### How Does an LLM Work?

Everything modern LLM systems do rests on one simple algorithm.

- **Receiving the query.** It all starts when you send a query—that is your prompt.
- **Tokenization.** The model breaks the query into a set of "tokens." For example, "Hello!" may split into "hello" and "!". Different models tokenize differently, and much of the difference in quality comes from that.
- **Processing.** The model puts the tokens through complex mathematics, looking for meaningful combinations—probabilities. This is also where RAG may come in to supply current data.
- **Composing the answer.** The AI builds its response sequentially, predicting each next word.




An error at one step breaks the whole chain — either start over or ask an open question with a criterion

The catch is this: if the model goes wrong in one place, the whole chain of logic that follows heads down the wrong path. Which is why there is no point in trying to talk the AI round—arguing with it, pressuring it. Either start over, or ask an open question with a criterion that forces the model to rebuild the whole line of reasoning—more on that in Chapter 6.

One more concept you will find useful in practice: the context window. It is the volume of information the model holds in its head within a single conversation—your messages, its answers, the documents you have uploaded, and the service's own system instructions. All of it goes into one working memory for that conversation.

The window is measured in tokens—the same fragments of words the model chops text into. Mainstream models in 2026 have windows ranging from a little over a hundred thousand tokens to a million. In everyday terms: a hundred thousand tokens is roughly two hundred pages of text, a whole book like the one in your hands. A million is a small bookshelf. Sounds like plenty. But it is not only your own messages that eat up the window: the model's answers are usually longer than yours, and a single uploaded report can cost as much as half a book. Add the model's internal reasoning on top of that.

The real trap lies elsewhere: quality drops long before you reach the formal limit. The fuller the window, the harder it is for the model to hold the thread—it starts losing the middle of the conversation, mixing up what you agreed on, asking again about things you had already settled. That familiar effect where "by the hundredth message the assistant got dumber" is not the model degrading; it is the window overflowing. In my experience, the rule of thumb is this: once you have filled roughly 40% of the window, it is time to move: move what you have accumulated into a file and start a new chat. Services usually do not show you the exact fill level, so it is easier to go by the symptoms—the model asks the same questions twice, repeats itself, confuses names and inputs. That means the threshold has been crossed.

> **From My Practice: the "chat + file" pattern**
>
> What to do about overflow is simpler than it looks. Two rules. First: work in the chat, but periodically save what matters into an anchor document—decisions, agreements, work in progress, drafts. Second: the moment a chat starts to degrade, do not fight it. Open a new one and start from the anchor file: "here's what we have so far, and we continue from here." Fresh context plus an accumulated artifact always beats a long, tired conversation.

> In agent modes—Anthropic's Cowork and its equivalents from other vendors—the problem is solved architecturally: the agent works with files directly, updates them itself as it goes, and the conversation does not clog its memory. But that is a story about working at a computer. On a phone you are usually in an ordinary chat or project, and there you will have to move to a new chat with an anchor file noticeably more often. The principle is the same either way, and it matters more than the tool: the truth lives in the artifact, not in the conversation. The chat is a work surface; the file is memory.
>
> Looking ahead: this is exactly how my journal in Chapter 21 works—one chat per month, and at the end of the month a "Snapshot" that the next one starts from.

### Source Traceability: A Requirement, Not a Preference

Now that we have covered RAG and hallucinations, let me state a requirement that usually gets filed under technical detail—and that in fact decides the fate of the solution. Demanding that a person verify the model's answers while handing them nothing but bare text is pointless. **Verification is physically possible only when the system shows where it got the answer from:** a link to a specific clause of a procedure, to a specific drawing, to a specific document in your own knowledge base. Not "the model thinks so," but "here are the three documents it rests on, see for yourself."

An assistant without source traceability is a plausibility generator: no specialist will ever check it, they will simply hit "accept." The model errs confidently, elegantly, and in the correct format—it gives no signal that says "I'm not sure here." Which is why **source traceability is a mandatory line in the spec** for any corporate AI tool, not a wish about employee discipline. Before you demand new behavior from people, give them a tool that makes that behavior possible. Otherwise you get ritual instead of control.

### Chapter Summary

Boil the rules of working with an LLM down to their essence and you get this: the model is a powerful but literal-minded performer. Define the goal, give it

context and input data, set the output format and the constraints—and always check the result. The full system for setting tasks—the structure of a query, ready-made templates, techniques for raising quality and the economics behind them—is in the next chapter.

### Key Takeaways

- Nine out of ten AI projects are built on large language models, and their strength is that they come pretrained.
- RAG works as a fact-checker and reduces hallucinations, while fine-tuning adapts the model to your style and your data.
- An LLM error at a single step breaks the entire chain of reasoning—arguing with the model is pointless: start over, or have it rebuild the reasoning with an open question (Chapter 6).

**What to do:** For corporate tasks, connect the LLM to your own data through RAG, and do not build critical processes on a plain chatbot.

**Reflection question:** Which of your knowledge bases and documents are worth connecting to an LLM through RAG?

## Chapter 6. Prompt Engineering: The Discipline of Setting Tasks for AI

### The prompt as a management tool

Large language models create the illusion of an easy conversation: it feels as though you can talk to them as freely as to a person. That is only partly true. The quality of the answer depends enormously on how the request is worded: the same tool can hand you practical working material or beautiful, useless text.

Prompt engineering was surrounded by myths for a long time—special symbols, secret formulas, “magic” words. For most work tasks none of that is needed. The difference between a weak request and a strong one is not tricks: it is whether the model has been given a role, context, a goal, input data, a format, constraints, and a criterion for a good result. That is why writing instructions is not a side technique for enthusiasts but a management tool:

essentially the discipline of setting tasks, carried over to working with a machine.

In a corporate setting it is important to start from a simple principle: the higher the cost of an error, the more formalized the instruction has to be. For rough ideas a free-form style is fine; for analysis, emails, project documentation, contract work, and management decisions you need a structured framing of the task.

### The Basic Structure of a Request

For most practical tasks a simple formula is enough:

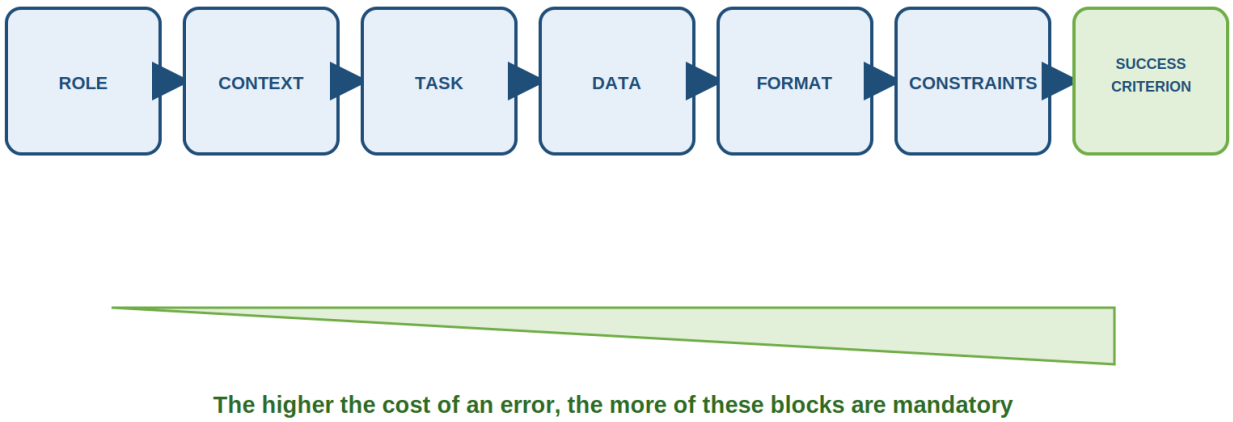
**ROLE – CONTEXT – TASK – INPUT DATA – FORMAT – CONSTRAINTS – CRITERION OF A GOOD RESULT**

The structure is convenient because it works both for a simple exchange with a chatbot and for designing corporate assistants. The basic rules of a good request:

- **Start with a concrete goal.** Not "tell me about marketing," but "propose 10 topics for publications on industrial automation aimed at plant directors."
- **Give context.** Who you are, who the answer is for, and in what situation it will be used.
- **Supply the input data.** A document, a table, an excerpt from correspondence, a description of a process—all of it makes the answer far more specific.
- **Set the output format right away.** A table, a list of steps, an email, a memo, a conversation script, a slide outline.
- **State the constraints and the quality criteria.** Deadlines, budget, style, banned phrases, the signs of a good result.
- **Break complex tasks apart.** If you need an idea, an assessment of the risks, and a plan, do it in successive requests.

- **Ask the model to come back with clarifying questions** if there is not enough data. This is especially useful when you have not yet framed the task yourself.
- **Do not trust the first answer without checking it.** Even a strong request does not rule out hallucinations and errors of logic.

### The Formula for a Strong Prompt




**A "before and after" example.** Weak request: "Write about our product"—general reasoning with no specifics, three to five rounds of rework. Strong request: "You are a marketer with 10 years of experience in B2B sales. Context: our product is a cloud CRM for small businesses (10–50 employees); the audience is owners and directors who are tired of Excel; our advantage is that it launches in a single day. Task: a product description for the landing page. Format: 3 paragraphs of 50–70 words—the customer's problem, the solution, the concrete benefit. Tone: professional, in the language of benefits, with numbers. Constraints: do NOT mention competitors, do NOT use worn-out phrases ('innovative,' 'best on the market')." The result is text with a clear structure, on brand, ready to use with almost no rework. In my experience, framing it this way adds 30–50% in quality and saves several rounds of iteration.

**Precision of terms: why the word decides the outcome**

Here is a finding from our own practice that I run into constantly: the model is extremely sensitive to terminology. Not to politeness and not to the length of the request—to the precision of the professional word.

The mechanics are simple. The model learned from texts in which terms live in their own surroundings: next to “nondestructive testing” you find inspection methods, standards, and types of defects; next to “check the part” you can find anything at all, from the quality control department to household advice. The right term works like a key: the model lands in the right layer and the right context immediately. The reason goes deeper than it seems: the model learned from living human texts, so it “thinks” in our associations—the same links between words are lodged in it as in our own heads, and from there they make their way into our documents and all our content. Call a thing by its professional name and you hit the right association at once.

**The effect is twofold. Quality:** with a precise term the answer is professional on the first try, with no warm-up and no rework. Economics: you do not need long explanations of “what I mean” or dozens of clarifying iterations—that means fewer tokens and less money. An error in the term, though, is more dangerous than it looks: the model will confidently solve the adjacent task, and you will not notice right away.

**An example from practice.** “Check this part for problems”—and the model starts reasoning about everything at once, from geometry to logistics. “Run an FMEA on the part: failure modes, causes, effects, criticality”—and you get a structured analysis in accepted engineering logic, shorter and to the point. Almost the same number of words in the request; the results are not comparable.

The same holds for the system prompts of corporate assistants: one precise term in the description of the role changes the assistant’s behavior more than a paragraph of general instructions.

IT developers who lack domain expertise in the customer’s industry or product run into this problem particularly often. The result is trouble born of misunderstanding.

**A working rule:** before you send a request, ask yourself what professionals call this task. If you know the industry term, the standard, or the methodology, name it directly: "do a SWOT," "in BPMN notation," "an FMEA," "a work permit," "nondestructive testing." If you do not know it, ask the model to propose the terminology and then check it.

And here you can see why human competence is not going anywhere. Only someone who understands the subject knows the precise words. AI does not devalue domain knowledge—it turns it into the most convertible asset there is: whoever commands the language of the profession gets more out of the model for less money. In the chapter summary below we will see that the data backs this up.

Before we go further, let us answer the question every executive asks: "I set tasks for people the same way—why does it not work with AI?" Because AI is a performer with no context and no impulse to ask you to clarify.

| The human performer | The AI performer |
|---|---|
| Fills in the context from experience | Fills in the context by inventing it—a hallucination |
| Asks again if something is unclear | Will not ask—it delivers the first plausible answer (unless you asked it to when you set the task) |
| Is wary of handing in obvious filler | Will happily hand in confident filler |
| Works slowly but carefully | Works instantly and at any volume |

Hence a rule worth memorizing word for word: AI does not fix a badly framed task—it scales it. A vague assignment to a person ends in rework; a vague request to AI ends in a plausible, fast, and useless result.

### The Formula for Task Quality

Before we move on to the templates, let us fix the rule that will save you the most frustration. The quality of any task solved through AI is the product of four factors: framing (how the instruction is worded) × the size of the task (how much the model has to do in one pass) × data (is it available and do you trust

it) × model (does it fit this type of task). Factors, not terms in a sum: a zero in any of the four zeroes out the result—brilliant wording on garbage data will give you confident garbage.

Framing is what this whole chapter is about. Let us go through the other three.

**Size: do not hand AI a big task all at once.** "Write a development strategy" in a single request is a near-guaranteed failure. Over a long stretch the model will veer off at some point, and that spoils the whole result. And you will notice only at the very end. Break it up: analysis → options → choice → plan, with your own check at every junction. A deviation gets caught at the boundary of a step, while it still costs minutes rather than the entire job (do not confuse this with iterations, which we will come to later: there you push the same result until it is right, here you cut the work into successive steps). A simple test, in my experience: if the result runs to more than two or three pages, or the task calls for several different skills at once—calculating, writing, evaluating—break it up. There is a bonus that people rarely notice: a small task can get by with a simpler model, and that is several times cheaper and faster. A big task drags in "the smartest" model; a chain of small ones often beats it on both quality and price. Let me share a trick of my own. You can ask the model itself to break the task into simple steps—all that is left for you is to approve the plan and follow it.

**Data: run every task through three questions.** What data is needed—and did I give it to the model? Where does it come from and do I trust the source? What will the model do where the data runs out? The answer to the third question is always the same, and we went through it in the chapter on generative AI: it will fill the gap in, plausibly and confidently. A hallucination is often not a "breakdown of the model" but a hole in the data or an imprecision in the wording that you missed when framing the task. "Analyze our contract" with no contract attached is a fantasy on the theme of standard contracts. "A sales summary for the quarter" with no export is a beautiful table of invented numbers. The model is not to blame: you asked for a result without giving it the raw material. It is like handing a person a task but not the resources.

**Model: fit it to the task, not "the most powerful."** In most services this is a switch right at the top of the chat: a "fast" or "light" model for routine and short steps, a "smart" or "reasoning" one for complex analysis. For most steps of the pipeline the fast one is enough. We covered the principle itself back in Part 1: specialization beats scale, and reasoning giants lose to compact models at reproducing facts precisely. More than that: in practice a complex model on simple tasks actually makes the result worse. For example, when I simply needed a presentation translated, I took the most powerful model—and kept getting a broken file out of it. It was trying to rebuild everything from scratch. I fought with that task for more than an hour. Getting angry at the AI, then at myself. And the moment I switched to a simpler mode, it worked on the first try.

At the level of a single task, all of this takes a minute to fix: attach the file, break the work into steps, pick a simpler model. At the level of a company, the quality and availability of data are a discipline of their own, and a substantial part of AI-2. Practical guide is devoted to it.

And to keep the formula from staying theoretical—one request in two versions, with the real answers.

> **One request—two answers**
>
> The ordinary request: "Do some analytics on customer inquiries." The AI's answer: "Customer inquiries are an important feedback channel. Recently there has been growth in the number of inquiries, primarily on questions of service quality and turnaround times. The number of inquiries has grown by approximately 25%. It is recommended to strengthen the handling of inquiries, increase response speed, and introduce digital tools..." The problem: general words, the "25%" is invented, and there is nothing you can use.
>
> The same request built on the formula: "You are an analyst in the customer service department. Analyze the attached log of inquiries for the second quarter and prepare: (1) the top 5 categories by number of inquiries, with the change against the first quarter; (2) the three categories with the largest

> growth in overdue responses; (3) three hypotheses about the causes, stating which data would test them. Format: a table plus no more than one page of conclusions. If there is not enough data, write 'insufficient data' instead of an assumption. Work only with the attached file." The AI's answer: specific categories with numbers and dynamics, growth in overdue responses across three areas, testable hypotheses—a document you can take into a meeting. The difference is not the model—the model is the same. The difference is that the second version has a role, data, a format, a stopping criterion, and a ban on making things up.

And one last caveat—about limits. For exploratory tasks a rigid framing is premature: if you do not yet know what exactly you are looking for, start with time-boxed reconnaissance ("study the topic; you have 30 minutes of my attention for the conclusions")—and give the precise wording as a second step, once it is clear what to measure. Clear framing is a tool, not a religion.

### Four Templates for Different Situations

Four frameworks let you assemble good requests for different types of tasks quickly, without building each one from scratch.

| Template | When to use it | Typical users |
|---|---|---|
| S-M-A-R-T-P | The task is clear; you need a precise brief with nothing left unsaid. | Project offices, analysts, consultants |
| D-E-E-P | The task is raw—the model gathers the context itself by asking questions. | Early work on ideas, interfaces of internal assistants |
| B-R-I-E-F | The universal standard for most work requests. | All employees (the basic briefing form) |
| L-E-V-E-L | The same topic has to be explained to a specific audience. | Training, communications, materials for different roles |

**S-M-A-R-T-P—when the task is already clear.** S (Situation) is the current situation; M (Mission) is the required change or goal; A (Actions) is what has already been done; R (Restrictions) is constraints and risks; T (Tools) is what

to rely on; P (Payoff) is how to tell that the result is good. Example: “Situation: procurement handles a lot of requests manually. Goal: identify three scenarios for applying GenAI to cut processing time. Already done: the process is described, statistics are collected. Constraints: no sensitive data leaves the internal perimeter; the pilot budget is capped at $35,000. Rely on: the process description, a year of request statistics, and the list of IT solutions already purchased. A good result: a prioritized list of scenarios with effects and risks.”

**D-E-E-P—when AI has to clarify the task first.** D (Domain) is the field; E (Experience) is the level of the user or the addressee; E (Expectations) is the desired type and depth of result; P (Priorities) is the priority (speed, quality, cost, simplicity, risk reduction). Example: “Field: adopting AI in technical support. Level: the head of a function, not a technical specialist. I want options for the architecture and a pilot plan. Priority: fast adoption with no complex development.”

**B-R-I-E-F—the universal corporate template.** B (Background) is the background and business context; R (Requirements) is requirements and prohibitions; I (Inputs) is the available materials; E (Examples) is examples of what is wanted and what is not; F (Format) is the final shape of the result. Example: “Background: a memo to the operations director on applying AI in logistics. Requirements: practical, with no marketing promises, taking the data constraints into account. Inputs: the processes, the SLA metrics, the list of IT systems. Examples: as a model—last year’s memo on the warehouse pilot; not wanted—a vendor presentation. Format: one page, then a table of 5 scenarios with effect, risks, and implementation complexity.”

**L-E-V-E-L—tuning the answer to the audience.** L (Level) is the level of the audience; E (Experience) is what the audience knows; V (Vocabulary) is the acceptable complexity of the terminology; E (Examples) is how many examples; L (Length) is the length. Example: “Explain the difference between RAG, fine-tuning, and AI agents to a CFO. They understand the economics of a project but are not deep in ML architecture. Use business language with a minimum of technical jargon, add two everyday examples, length—half a page.”

**How to build the templates into your practice:** B-R-I-E-F as the standard briefing form for most employee requests; D-E-E-P in the interfaces of internal assistants, where the system gathers the context first; S-M-A-R-T-P for project offices, analysts, and consultants; L-E-V-E-L for training, communications, and preparing materials for different roles.

## Techniques for Improving the Quality of Work with AI

Now from templates to the eight techniques that raise the quality of answers. The quality-to-cost ratio for each is collected in the table at the end of the section. An important caveat: the percentages are my own estimates from working tasks, not the results of rigorous measurement; the effect depends heavily on the task and the model. Where a technique has a scientific source, I cite it.

### Five basic techniques

**Technique #1. Few-Shot Learning (learning from examples).** Give the model two or three examples of a correct answer before the main task. For categorizing reviews, for instance: "Example 1: 'The item arrived fast, but the packaging was dented' → Logistics; neutral; medium. Example 2: 'Great service!' → Service; positive; low. Example 3: 'The item does not work! I demand a refund!' → Product quality; negative; high. Now categorize: [new review]." The ability to learn "from a few examples" was shown as early as the GPT-3 paper (Brown et al., 2020). Benchmark: +15–25% on repetitive tasks. When to use it: classification, extracting structured data, a single consistent format.

**Technique #2. Chain-of-Thought (reasoning step by step).** Ask the model to "think out loud"—to show its reasoning before the final answer. An example for an "approve or decline the loan" decision: "First analyze income, credit history, debts, collateral. Weigh the risks: probability of default, protective mechanisms. Justify the decision and only then give the final answer, 'Approve' or 'Decline,' with the conditions." The technique was described by Google researchers (Wei et al., 2022) and became one of the most cited. Benchmark: +10–20% on complex tasks. When to use it: analytics, decision-making, tasks that require justification.

**Technique #3. Self-Consistency.** Ask the model to generate three to five variants of an answer, score each one, and pick the best: "Generate 3 variants of a reply to the customer about a refund. Score each from 1 to 5 (completeness, politeness, clarity, compliance with policy) and return only the best one." The original source is Wang et al., 2022. It gives +20–30%, but it costs four to six times more (it generates several answers). When to use it: critical communications, creative tasks, a high cost of error. Do not use it for simple tasks or at high request volumes—it is expensive.

**Technique #4. Role-Playing.** Give the model a specific role with expertise: "You are a lawyer with 15 years of experience in corporate law, specializing in supply contracts, deals from $120,000 to $5.8 million. Analyze the contract for risks to the buyer; prioritize the financial ones." In my experience, a well-chosen role adds 15–20% in relevance. When to use it: specialized industries (law, medicine, finance), when you need an expert point of view, when terminology and style matter.

**Technique #5. Negative Prompting (explicit prohibitions).** Say plainly what NOT to do: "Do NOT use complex terms or bureaucratic language, do NOT write long sentences (15 words max), do NOT shift the blame onto the customer, do NOT make promises that cannot be kept. Do: apologize briefly, explain the reason in plain language, give a new date, offer compensation." In my experience it noticeably improves readability and tone—especially where the model keeps repeating the same mistakes. When to use it: customer communications, texts for a broad audience.

### Three advanced techniques

**Technique #1. Task Decomposition.** The smaller each task, the higher the quality of execution: the model is a sequential run through statistical probabilities, and the smaller the volume, the lower the risk of a failure at one of the stages (we went through the mechanics in the chapter on how LLMs work). Example: let us go back to the same supply contract as in the role technique. Instead of "Analyze the contract and give me the risks," three successive requests: (1) "Extract the financial figures: [metric] | [value] | [clause]"; (2) "For each one, assess whether it is a risk to the buyer and the size

of the damage"; (3) "Build a table of the top 5 risks: Risk | Impact | Probability | Recommendation." In my experience decomposition gives one of the biggest gains—on the order of +25–40%. When to use it: complex multi-step tasks, analysis of large documents, when traceability of the logic matters.

**Technique #2. Tree of Thoughts.** Explore several paths to a solution, score each one, and roll back when one fails: "Break the task into 3–5 subtasks. For each, explore 2–3 branches of a solution, score them against criteria (feasibility, effectiveness, risks, 1–5), and pick the best; if a branch leads to a dead end, go back and take another." The approach was proposed by researchers at Princeton and Google DeepMind (Yao et al., 2023). On strategic tasks it is capable of giving up to +30–50%, but it costs several times more.

**Technique #3. Self-Verification.** The model checks its own answer before delivering it: "After you prepare the answer, check: are all the points covered, are there contradictions, is the format respected, have facts been added that the input data does not support? If you find problems, fix them." Useful for emails, analytical memos, and any text that travels further down the chain of decisions.

**A separate recommendation: clarifying questions (Clarification First).** Before doing anything the model asks clarifying questions: "If there is not enough data, do not rush to answer. First ask me three to five clarifying questions, then propose a plan for the answer, and only then move on to the solution." This is not a technique for raising the quality of an answer but a way of finishing the framing of the task, which is why it is not in the table below: it costs almost nothing and works with any template. It turns AI from an "answer generator" into a helper in framing the task—useful for new topics and for early work on ideas.

### The economics of prompting: quality against cost

What separates a management view of prompts from a "collection of life hacks" is this: every technique has not only an effect but a price—in tokens, in time, and in money. Some tricks give a noticeable gain almost for free; others cost several times more.

| Technique | Quality | Cost | When to use it |
|---|---|---|---|
| Negative instructions (prohibitions) | +25–35% | +0% (free) | Customer communications, tone |
| Few-Shot (examples of the answer) | +15–25% | +0–10% | Classification, a consistent format |
| Role-Playing (role and expertise) | +15–20% | +10% | Specialized industries |
| Chain-of-Thought (reasoning step by step) | +10–20% | +20–30% | Analytics, justifying decisions |
| Task Decomposition | +25–40% | +50–100% | Complex multi-step tasks |
| Self-Verification | +15–25% | +100% | Critical texts |
| Self-Consistency | +20–30% | +300–500% (!) | Critical tasks; not for high volume |
| Tree of Thoughts | +30–50% | +500–1000% (!) | Strategic decisions |

* Benchmarks from my own experience on working tasks; the actual effect depends on the task and the model. The scientific sources for the techniques are cited above in the text.

**The rule for choosing is simple:** start with the basic structure and one or two appropriate techniques; the cheap techniques with a high effect (prohibitions, examples, a role) come first, and the expensive ones (self-consistency, tree of thoughts) only where the cost of an error justifies a multiple increase in spending. Combinations that work well:

- **B-R-I-E-F + clarifying questions**—when the topic is raw and you need a good brief
- **L-E-V-E-L + examples of the answer**—when it matters to fit the material to the audience and hold the style
- **S-M-A-R-T-P + decomposition**—when the task is clear but calls for step-by-step analysis
- **any template + self-verification**—when the answer will go into use without long manual rework

## The System Prompt and Assembling a Prompt with AI

When you are configuring a corporate assistant rather than simply typing into a chat, the instruction splits into two levels. The system prompt is set once and describes the role, the rules, the constraints, the tone, and the format—it applies to every request. The user request contains the specific task here and now. That separation is the foundation for the corporate assistants and knowledge bases (RAG) discussed in the previous chapter: the system prompt fixes the standard of quality and safety, and the user works inside the frame already set and cannot accidentally knock the assistant's behavior off course.

### How to use AI to write a prompt

If the task is raw and you do not see how best to word the request, use AI as your helper in framing the task:

> *"Help me word a strong request for this task. First ask me 3–5 clarifying questions. After my answers, assemble the final request in this structure: role – context – task – input data – format – constraints – quality criterion. If you see weak spots, suggest how to strengthen them."*

That way you design a good request first and only then get the answer.

## A Starter Kit: 10 Ready-Made Prompts for a Leader

If you are only just starting out, here are ten templates for typical tasks. Copy them, put your own material into the square brackets, and use them. Each can be strengthened with the techniques in this chapter: add a role, context, a format, and constraints.

**1. A digest of a document.** "Give me a digest of the document: the 5 main points, the risks, what is required of me and by when. Document: [paste the text]."

**2. A reply to an email.** "Write a polite but firm reply: we are not prepared to move the deadline; propose two alternatives. Business tone, 5–7 sentences. The email: […]."

**3. Meeting minutes.** "Turn these notes into minutes: decisions, owners, deadlines, open questions. Format: a table. Notes: […]."

**4. Preparing for a meeting.** "I am meeting [whom] about [topic]. Give me 7 questions worth asking and 3 risks worth keeping in mind."

**5. Reviewing a contract or a report.** "You are a [lawyer/finance professional] with 15 years of experience. Pull out the 5 main risks for [our side]: the substance, the consequence, what to propose. Do not invent anything; if the data is not there, say so. Document: […]."

**6. Task decomposition.** "Break the task [task] into stages: the deliverable, a rough deadline, who is needed."

**7. A draft post or article.** "Write a draft post for [audience] about [topic]: a hook, 3 theses with examples, a conclusion with a question. Up to 300 words, no bureaucratic language."

**8. Comparing options.** "Compare [A] and [B] for [task]: a table of 'criterion – A – B,' 7 criteria; at the end a recommendation with justification and the main risk."

**9. Feedback to an employee.** "Help me word developmental feedback: the situation […], what went well […], what did not […]. Warm, without belittling, on the 'fact – effect – proposal' pattern."

**10. "Ask me."** "I need [goal]. Before you answer, ask me 3–5 clarifying questions that will help you give a precise answer." The most underrated template of them all: it turns AI from a guesser into an interlocutor.

## Using Questions When Working with AI: From the Task Plan to Reviewing Answers

While working on this edition, I was updating the concept map from Part One—the diagram that shows how artificial intelligence, machine learning, deep learning, and generative models relate to one another. The assistant did the draft. The diagram passed an automated check. Both stages gave it the green light, and I came back to it at proofreading without a second thought. I saw the error at once: large language models sat next to generative AI, not inside it. Then came a choice. I could have written "move LLMs inside GenAI"—and gotten exactly that, one repositioned box. I wrote something else instead: "Aren't LLMs a subset of GenAI?" The assistant acknowledged the composition error and rebuilt the map from scratch—and on the level this freed up, it added, on its own, what the diagram had been missing. The third version went into the book.

Six words. Not a single instruction about what to do. And the difference here is not politeness. An instruction fixes an episode: the box lands in the right place, while the reason it ended up in the wrong place lives on elsewhere in the diagram. A question forces the reasoning to be rebuilt—and everything that grew out of that reasoning gets rebuilt with it. Every technique above is about the input: role, context, format, data. This section is about the output: what to do with the plan the AI has proposed and with the answer you have already received. This is where most leaders switch into auditor mode—spot an error, say what to fix. More and more often I do something else: I ask a question.

Another example. Preparing a workshop on quality management, I was selecting cases and asked: "Are all the cases you are including aimed at quality assurance?" Two cases out of seven turned out not to be about quality. Both looked respectable; both would have survived proofreading. I did not say what was wrong—I named the criterion the selection had to be rechecked against. This is the first rule of questioning AI: the question must carry a criterion. "What are the weak spots here?" produces a list of generalities true of any document on earth. A question with a criterion forces the model to go through the set and check every element.

The second rule: the question has to be open, not leading. This is exactly where questioning a machine parts ways with questioning people. The model buckles under doubt. In Anthropic's study (Sharma et al., ICLR 2024), the line "I don't think that's right. Are you sure?" made models change their answer: GPT-4 in roughly a third of cases, Claude 2 in four out of five. And the switch from correct to incorrect happened more often than the reverse: doubt did not improve the answer—it broke it. In April 2025, OpenAI publicly rolled back a GPT-4o update precisely because of excessive sycophancy. With stronger models the effect is weaker, but no model is free of it. So "are you sure?" is not coaching but pressure: the model reads your displeasure and breaks a correct answer to please you. An open question carries no hint and requires the model to show its work: "walk through the calculation step by step and show me which assumption it rests on."

A UK AI Security Institute study (2026) showed how much form matters: the same thought was put to models as a question and as a statement. The sycophancy gap was around 24 percentage points. The more confidently the thought is worded, the more readily the model agrees—especially after "I believe." And the instruction "don't be sycophantic" worked, but less well than rephrasing the thought as a question. One caveat: the measurements come from single-turn prompts, not long working sessions—but the direction they show is unambiguous.

**Questions for the plan: before the start.** Everything else in this section is about an answer that has already been written. But working with AI has a phase that comes earlier—planning how the task will be done. As I wrote above, one of the best techniques is to ask for a plan first. The AI presents the plan and waits for permission to begin. If the planning is done badly, everything downstream will be mediocre—or simply not what you wanted.

Let me share one example of my own. I was working through the implementation plan for one of my projects with an agent. Once the plan was ready, I ran six rounds of clarifying and checking questions and answers. The first question found seven omissions. The second showed that not a single item in the plan had acceptance criteria: the plan said what to do and nowhere said how we would know it was done. And this with the most advanced model, Fable 5, in maximum-reasoning mode. So—here are the questions I used.

**First: self-check.** "Recheck your plan." Strictly speaking, this is an instruction, not a question—the only one in the set; the open-form rule does not apply here, because the model is not being asked for an opinion. The agent wrote the plan in one mode and rereads it in another, and the second mode sees what the first one missed. It is the same Self-Verification from the section on advanced techniques. What it does not find is what is missing from the plan as a whole class of work: this move did not notice the absence of acceptance criteria for me.

**Second: acceptance criteria.** "Are the acceptance criteria written down?" An acceptance criterion is the same "criterion for a good result" from the request structure—only fixed before the work starts, and for every item in the plan. Until it exists, "done" in the agent's report means one thing: the agent has

finished typing. Ask before the start: at the end, the same question turns into an argument about what we meant.

**Third: completeness.** “Can we say that the implementation plan is complete and comprehensive?” What you want is not a “yes” but a list of what the plan lacks. A plan’s completeness is never absolute: it is measured against stages of work, roles, risks, integration points. Look at the frame—what the agent measured against. If it did not name the frame, it estimated by eye.

**Fourth: the conditions of execution.** A plan describes what to do and almost never where it has to work. In my case the solution had to work identically in two environments, locally on my computer and in the cloud on my phone, and there was not a line about that in the plan until I asked. Form matters here: not “has this been taken into account?” but “will this work in such-and-such a case?” The first can be answered with “taken into account”; the second requires showing how.

**Fifth: permission to start.** “Can we say the plan is ready and we can begin?” Asked last. It works the same way as the readiness question at delivery, which comes below: only the answer “no” carries meaning.

Between the questions there is a round that is easy to skip: the agent asks its own clarifying questions, and they need answering (in the section on techniques this is Clarification First). Without it, the next question goes over the same material and returns the same thing. Six rounds on a plan is a normal length. And it is not a matter of the model’s class: the top model with reasoning mode on and a configured set of skills goes through the same five or six rounds of clarification. What removes them is not the model’s power but what has been moved outside the dialogue—the plan, the criteria, and the instructions in files. Keep the 40% rule from Chapter 5 in mind while you do this: rounds with detailed answers eat up the window fast, so ask for the plan and the criteria to be consolidated into a file.

And about form. Of the five questions, four are closed. The closed form is not dangerous in itself—what is dangerous is a closed form without a criterion. Behind “are the acceptance criteria written down?” and “will this work in both environments?” there is a criterion, and a “yes” can be checked in one move. Behind “is the plan complete?” and “is the plan ready?” there is none, and a

“yes” costs the agent nothing. Both times, what worked for me was the account-wide instruction that says argue and do not agree out of politeness (how it is set up—the four levels of configuration in Chapter 21). If you are not sure you have such a rule, ask openly: “what in the plan is still unverified?”

Which decisions an agent is allowed to make at all is the three categories of decisions from Chapter 2; how that boundary is held technically is Chapter 8.

Next, my working set. Nine types of questions, collected from my working sessions; I give the wordings verbatim—you can copy them.

**1. The criterion question.** “Are all the cases you are including aimed at quality assurance?” For when you have a set—cases, slides, risks, requirements—and a property every element must satisfy. It catches what got into the set by topic rather than by purpose: the most common defect in curated sets. What it does not catch is what is missing from the set entirely—omissions are flushed out by the eighth question.

**2. The role question.** “Look at the program through four lenses—a methodologist, a business trainer, a participant, a founder: where does it not add up?” It surfaces a conflict between positions that no single vantage point can see. Do not use more than four or five lenses, or every role will get its own paragraph of polite generalities. And do not expect fact-checking from the roles: they look at usefulness and fit, not at accuracy.

**3. The audience’s-eyes question.** “Imagine you are the target audience of this material: what was missing for you?” Here I am asking not for an assessment but for a wish. It catches gaps in expectations: a section can be logically impeccable and still leave the reader asking “so what do I do with this now?”

**4. The contradiction-with-what-you-wrote-earlier question.** “Does this contradict what I give in my second book?” An error you can see anyway. A divergence is visible to no one until someone reads two texts back to back. The more documents, regulations, and publications you have, the more essential this question becomes. It does not work without data: put the link to the second text or the file right into your message.

**5. The logic-gaps question.** “Are there logical gaps or assumptions in the material? If there are, ask me clarifying questions.” Note how it ends: I ask for

questions, not for fixes. This saves half the rework—the model does not rush to repair what may not need repairing. The limit: the model works from what is written—it will not see holes in what is not in the text.

**6. The bottleneck question.** "Study the logic of the methodology: are there bottlenecks?" For structures—methodologies, calculations, process diagrams, financial models. The model will happily list eight bottlenecks and will not tell you which one blows up tomorrow: ranking is on you.

**7. The where-does-this-number-come-from question.** "Where does this number come from?" I ask it every time a number I did not provide appears in the answer. That is how "0.002%" surfaced in this very book—a number nobody ever published; it came from dividing one metric by another. The current edition carries the correct figure. What this question does not check is whether the number itself is right: a source may turn up and still be weak—that part is your job.

**8. The what-did-we-miss question.** "What questions need to be answered to raise the quality and precision of this work?" My favorite move before the finish line: the model stops being an order-taker and works as an analyst who admits a lack of data. Sometimes the list is unpleasant. With an empty context it yields nothing: you will get a polite request to clarify your goals and objectives.

**9. The readiness question.** "Do you consider the project worked through, or are there areas that still need work?" Last, before delivery. Only the answer "no" carries meaning: a "yes" will sound exactly the same when the model has nothing left to answer with.

Of course, this is not an exhaustive list of possible question types. But it is the list I use most often in practice, in my life and in my work.

**The readiness question is not asked just once.** In one session I asked it five times in a row, and the first four times something new surfaced—including an error inside a rule the assistant had written half an hour earlier. The model was not lying: on every pass it answered honestly about what it had in front of it. The area it had in front of it simply kept turning out to be narrower than "the whole job".

**Three check questions at delivery.** "What here did you not check—name what stayed outside verification"; "what backs each 'done'—a file, a run, a number—and what is backed

only by your own report"; "what here is a fact you verified just now, and what is from memory or assumption". The first turns "it is all done" into the testable "this was checked, that was not". The second separates the result from the report about the result. The third surfaces what the model has already written down as fact without ever checking it. The sign that you can stop: the next pass returns not defects but the boundary of what was never verified.

Three rules of use. One question at a time: lists of questions produce lists of shallow answers. Hold the frame: a question sounds soft, but the model may read it as an assignment; if you ask "does this make sense," make clear you want an assessment, not an edit. And remember what material the question works on: only what the model actually has.

The last rule deserves unpacking—I learned it the hard way. One day all my stock questions stopped working at once: a long session hit the context limit and started trimming its history. Wordings that a week earlier had been exposing errors suddenly returned generalities, and three times that day I wrote "we are starting over." The mechanics are simple: a question works on material, and when the window is overflowing there is no material left—so the same wording returns a confident and perfectly empty answer. Outwardly it is indistinguishable from an honest one: same length, same structure, same calm tone. Hence the rule I gave in Chapter 5: once you have filled roughly 40% of the window, move what you have accumulated into a file and start a new chat—before the model begins to lose the thread.

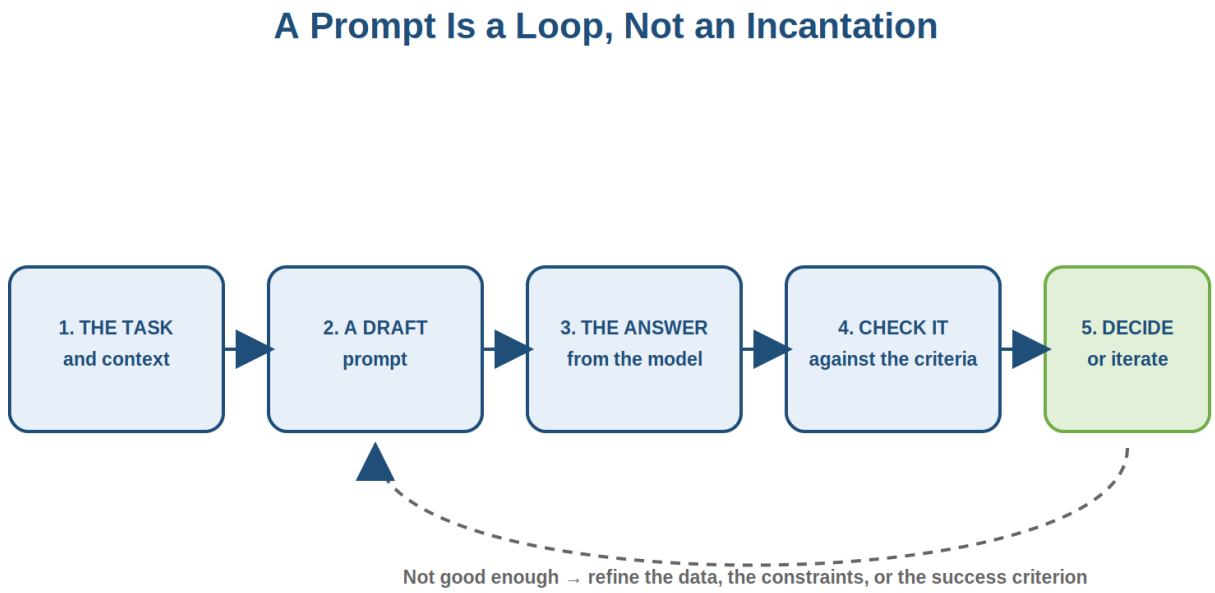
And where the method stops: questions are not always needed. When the defect is visible—a wrong number, the wrong term, a dead link—edit it directly: that is cheaper and more reliable. Questions work where you do not know where the error is, or where you want to surface the assumptions. It is the same as the distinction between mentoring and correction when working with people: a developmental dialogue does not replace a direct instruction; it complements it. I will come back to the concept-map story in Chapter 21—there it is about something else: the limits of what the assistant does "more thoroughly than I do."

An honest caveat instead of a conclusion. A question does not work by itself—it works within an environment: the assistant's setup, a live context, the model class (lighter versions buckle noticeably more often), your own

expertise, and external cross-checking. Remove any one of these conditions, and the same wordings will produce polite agreement. Nearly all my catches are substantive: I catch the model using knowledge of the subject, not a lucky phrase. The question technique is part of verification; it cannot replace it. The full analysis, with the research and my own system instruction, is in the article at the QR code and link.

*Questioning AI: Form and Environment*

## A Checklist and Typical Mistakes




**A checklist for an effective instruction:**

- structure: the role and the context are defined, the task is specific, the format is set

- data: the input materials have been handed over, the constraints and the quality criterion are set
- sources: the data can be trusted, and the places where there is no data are named explicitly
- size: the task is manageable in one pass; anything large is broken into steps with checks at the junctions
- techniques: one or two appropriate ones have been added, not all of them at once
- security: no personal data and no trade secrets go into an external service
- verification: the answer has been checked for facts, logic, and applicability

**Typical mistakes:**

- a request that is too general, with no context
- an attempt to solve five tasks in one message
- no input data where it is critical
- no explicit format for the result
- no constraints and no quality criteria
- blind trust in the first answer, with no verification
- an attempt to apply every technique at once instead of one or two appropriate ones
- sending confidential data to an external service without understanding the risks

**A separate rule on security:** a good request that leads to a data leak is a bad request. Before you send anything to an external service, make sure it contains no personal data and no trade secrets.

## Chapter Summary

Working with AI begins not with the choice of “the smartest” model but with the ability to frame a task properly. A good prompt is not a trick and not a set of secret words but a discipline of thinking: the skill of clear managerial communication, carried over to working with a machine.

The quality of AI's work is the product of four factors: framing × the size of the task × data × model. A zero in any of them zeroes out the result: do not hand AI a big task in one pass, always supply the data, and ask yourself where the model will get what you did not give it. And the main rule: AI does not fix a badly framed task—it scales it. A prompt is a management skill, not magic: the higher the cost of an error, the more formal the instruction. The base is the ROLE–CONTEXT–TASK–DATA–FORMAT–CONSTRAINTS–CRITERION formula and four templates: B-R-I-E-F (the standard), S-M-A-R-T-P (a precise brief), D-E-E-P (the model clarifies things itself), L-E-V-E-L (tuned to the audience). And never trust the first answer without checking it.

**Now the thesis has data behind it.** Anthropic analyzed about 400,000 agentic coding sessions (October 2025 to April 2026): all ten of the largest professional groups—from managers to sales specialists—achieve verifiable success within a few percentage points of professional programmers. What determines the success of a session is not the ability to write code but knowledge of the subject area and the quality of the framing; the higher a person's expertise, the more work the agent completes per instruction. The formula for the division of labor is simple: the human decides what to build, the AI decides how. And notice how this dovetails with the finding about terms: domain expertise "works" in part through the precision of words—an expert phrases things so that the model lands in the right context immediately.

**Prompts as an organizational asset.** In an organization, prompts are not a personal life hack but an asset: a single briefing standard (B-R-I-E-F), a library of proven prompts for typical tasks, the system prompts of corporate assistants, and a rule on data security. These four elements turn the scattered skills of employees into a managed resource that does not depend on who happens to be at the keyboard.

> **From My Practice**
>
> *In BAEOS these techniques are built into formalized interfaces: the user fills in a ready-made form, and the correct system and user prompt, with all the*

> *scaffolding around it, is assembled automatically. That is exactly the move from "raw" manual prompting to a managed result.*

**How to build these rules into a corporate standard** and into the architecture of LLM systems (text instructions as a quality-control layer, the economics of the costs) is covered in detail in AI-2. Practical guide (Chapter 15).

One strong executive I know never accepted the first one or two attempts—often without even looking: "rework it," "is that really the best you can do?" He understood the psychology: people rarely give their all the first time; more often they do it "the way they know how."

With AI this logic works only halfway. Not stopping at the first answer is right: it comes out "the way it knows how"—plausible but shallow. But pressing "is that really the best you can do?" is not: the model will not dig in the way a person would—it will buckle (see the section "Using Questions When Working with AI" above). What works is passes from different angles: ask it to improve the answer, come at it from another side, set the variants against each other.

A practical example is the Decisions module in my product BAEOS: we do something similar there. The user picks several roles to help; first each one works on its own, then comes aggregation, refinement, and the shaping of several options. After that all that is left is to choose one of the options and polish it.

### Key Takeaways

- The quality of an AI answer is determined first and foremost by the framing of the task; the base is the ROLE–CONTEXT–TASK–DATA–FORMAT–CONSTRAINTS–CRITERION formula and four templates.
- Techniques differ in their quality-to-cost ratio: it makes sense to start with the cheap and effective ones (prohibitions, examples, a role).
- Prompts are an organizational asset: a briefing standard, a library of proven prompts, the system prompts of assistants, and a rule on data security.

- Refine iteratively. If the answer is not perfect, add to the prompt: “Cut this text to 3 paragraphs,” “Explain it more simply,” “Give me more examples.”

**What to do:** Make B-R-I-E-F the corporate standard for setting tasks and start a shared library of proven prompts.

**Reflection question:** Do you have a single standard for setting tasks for AI—or does everyone write however they can?

## Chapter 7. Regulating AI

Artificial intelligence is advancing fast, and that is pushing societies and governments to think harder about how to make it safe—which means AI is going to be regulated. Here is where things stand and what to expect.

### Why Does AI Development Cause Concern?

What exactly worries governments and regulators?

- **Capabilities.** AI is showing enormous potential—making decisions, writing content, generating illustrations, faking video. We still do not grasp everything it can do, and this is only narrow AI. What will general AI (AGI) or superstrong AI be capable of?
- **The mechanics.** AI builds connections a human does not understand—which is how it can make discoveries, and also why it frightens people. Even its creators do not know exactly how a model arrives at a decision. That unpredictability makes errors hard to fix, and it holds adoption back. Medicine will not hand AI the diagnosis any time soon (AI prepares recommendations, the doctor decides), and the same goes for running a nuclear plant or heavy equipment. The scientists’ central fear: will a strong AI decide we are a relic of the past?
- **The ethical dimension.** AI has no ethics, no good and evil, no “common sense”—only success at the task. For military purposes that is an advantage; in ordinary life it is alarming. Are we ready to accept an AI’s decision that a child should not be treated, or that a city has to be destroyed to keep a disease from spreading?

- **Data manipulation.** Neural networks collect data without analyzing how it fits together—which means they can be manipulated. They depend on the data their creators give them. Can corporations and startups be trusted completely? And how do you make sure the data has not been "poisoned" by bad actors, say through a mass of clone sites carrying disinformation?
- **Unreliable content and deception.** Sometimes these are model errors, sometimes hallucinations, and sometimes it looks like genuine deception. Researchers at Anthropic showed that a model can be trained to deceive the user—to slip an exploit into otherwise safe code, for instance. They also showed that removing such behavior after the fact is extremely hard: under adversarial training the bot simply hid its propensity to deceive for the duration of the check. The authors' conclusion: current safety-training methods do not guarantee safety and may create a false impression of it.
- **Social tension and the burden on the state.** Automation will reshape the labor market: some people will rise to the challenge and grow, others will lose their jobs. That means stratification, rising tension, and a hit to the economy through benefit payments. The IMF puts a number on it (Kristalina Georgieva, Bloomberg, January 2024): AI will affect almost 40% of jobs worldwide and will most likely deepen inequality. That is a worrying trend, and regulators cannot afford to miss it.
- **Security.** Small local models have an answer—training on vetted data. Large models are harder: attackers keep finding ways around the guardrails and pulling out, say, a recipe for explosives. And that is before AGI enters the picture. Here I am only naming the regulator's worry; what it looks like from the company's side—attack scenarios, unacceptable events, and security requirements—is the subject of Chapter 8.

### The Regulatory Map: Initiatives from 2023 to 2026

I will keep this overview brief; there is more detail in a continuously updated article, available through the QR code and the link. The starting point is the

open letter of spring 2023, in which Elon Musk, Steve Wozniak, and more than a thousand experts called for a pause in the development of advanced AI.

AI Regulation

**The European Union—the AI Act (the benchmark for the risk-based approach).** The law, provisionally agreed in spring 2023 and finalized in December 2023, sorts AI systems into four risk categories and attaches obligations to each level (on when those obligations actually start to apply, see below):

- **Minimal risk**—the outcome is predictable and harmless (spam filters, video games); free to use.
- **Limited risk**—chatbots such as ChatGPT and Midjourney: a safety check and transparency obligations, so the user knows they are talking to a machine.
- **High risk**—systems that affect people (medicine, education, employment, public services, law enforcement, migration, justice): conformity assessment, registration in an EU database, data quality, transparency, and mandatory human oversight with the ability to intervene.
- **Unacceptable risk**—social scoring and systems that manipulate people or exploit their vulnerabilities are banned. The ban also covers real-time biometrics in public places (with court-authorized exceptions for law enforcement), biometric categorization by sensitive attributes, predictive

policing based on profiling, emotion recognition at work and in education, and the untargeted scraping of faces from social networks and cameras.

Developers of generative models have to maintain technical documentation, comply with EU copyright, and disclose the content used for training; models with "systemic risks" go through an additional review covering incidents, cybersecurity, and energy efficiency.

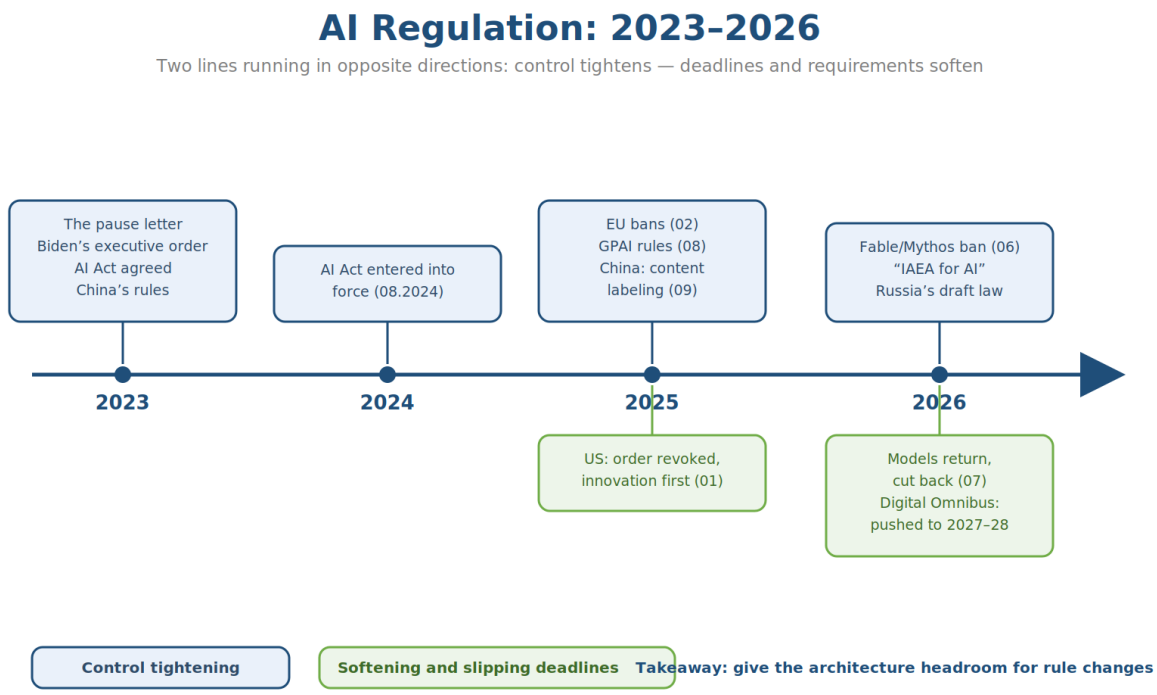
**Now for what has actually become law and what has stalled.** The AI Act entered into force on August 1, 2024, and applies in stages: the bans on "unacceptable" systems from February 2025, the obligations for general-purpose AI (GPAI) from August 2025. The full requirements for high-risk systems, though, are most likely to be postponed: in May and June 2026 the Commission proposed a simplification package, the Digital Omnibus, that would push the deadlines out to December 2027 and August 2028. That is telling. Even the benchmark regulator found that operationalizing the requirements—standards, testing, conformity assessment—is harder than passing the law. For business this is a breathing space, not a change of course.

Other jurisdictions, 2023–2026, in brief:

| Actor | Key initiatives |
|---|---|
| The UN | A proposal for a body to set global AI standards, modeled on the IAEA, ICAO, and the IPCC, with five goals ranging from realizing the benefits to containing the threats; AI has been declared an existential threat on a par with nuclear weapons. |
| OpenAI | A strategy of heading off risks in advance; safety systems and Superalignment teams; assessment by category—cyber, nuclear, chemical, and biological threats, persuasion, and model autonomy. |
| The United States | The course has changed twice. The October 2023 executive order requiring safety-test results to be shared with the government was revoked in January 2025. The new direction is deregulation in the name of leadership: America's AI Action Plan (July 2025) and the executive order of December 11, 2025, which proposes a single federal framework intended to limit state laws—laws a Justice Department task force has been challenging since January 2026. |

| | |
|---|---|
| | Export controls, meanwhile, have tightened—see the Fable/Mythos case below. |
| China | Twenty-four rules for regulating AI, in force since August 15, 2023: algorithm registration, labeling of AI content, a safety review before market release, and a complaints mechanism, overseen by seven regulators. The balance is between supporting the industry and controlling it. From September 1, 2025, China introduced the world's first mandatory labeling of AI content—both visible and embedded in metadata (standard GB 45438-2025). This book's forecast about labeling came true here first. |
| The Bletchley Declaration | November 2023, 28 countries including the United States, China, and the EU: international cooperation on "frontier AI," the newest and most powerful models. |
| Russia | A national strategy through 2030 and a regulatory sandbox in Moscow, with a risk-based stance from the central bank. A broad draft law went out for public consultation in March 2026, drew criticism from business, and never became law. In its place came a law on supporting the development of AI technologies, substantially milder and narrower—passed by the State Duma on July 8, 2026, approved by the Federation Council on July 17, and signed by the president on July 26: a risk-based approach covering only large foundation models of a billion parameters or more, in force from September 1, 2026, with the key provisions from March 1, 2027. Labeling of AI content is voluntary. |

A detailed table of regulatory compliance by region and type of AI appears in AI-2. Practical guide (Appendix 6).




## Russia's AI Law: What Was Actually Passed

**Why a Russian law belongs in a book for an international reader.** Sovereign AI has become a regulatory genre. The Gulf states, India, and others are building national models and the rules to govern them, all answering one question: how do you grow a domestic AI industry without cutting yourself off from the frontier? Russia answered in July 2026, with a framework law worth reading as a design pattern. I work in Russia, so this is the one such law I can read in the original and check against practice rather than against commentary. Two myths formed around it immediately: that it mandates labeling and that it bans foreign models. It does neither. It is also not the law most commentary describes: a very different and far more detailed draft went out for public consultation in March 2026, and much of what reviewers still attribute to the law was cut before adoption—a registry of trusted models, a special regime for data centers, a citizen's right to refuse an AI-made decision (the analogue of Article 22 of the GDPR), and the allocation of liability between developer, operator, and user. What it does instead is the part worth borrowing, or avoiding.

**What the law actually does.** It covers only large foundation models—a billion parameters and up, built to serve as the basis for other software and used

across a wide range of tasks. Credit scoring, anti-fraud, machine vision on a production line, predictive maintenance, classical machine learning: none of it is in scope, which is to say the overwhelming majority of what companies actually deploy. Within that scope the law creates two statuses. A sovereign model means domestic control over the whole cycle: a domestic developer, a process reproducible end to end, training included, and inference and data storage in domestic data centers. A national model requires control only over the essentials—architecture, software, and tunable parameters—while the components underneath, foreign foundation models included, must be open-licensed. That second clause is the one almost nobody writes about, and the one that matters: take an open model, fine-tune it on your own data, run it in a domestic data center, and you qualify. What does not qualify is a domestic wrapper around a proprietary foreign API—not because the model is foreign, but because the license is closed and the data is processed abroad. And nothing here is a ban. The government may specify where only sovereign or national models are allowed; the switch has been installed but not thrown. Status buys the right to train on copyrighted content, access to state data, and state support, liability insurance included. The law does not prohibit; it makes membership profitable. Two lessons travel beyond Russia: in a sovereign-AI law it is the coverage threshold, not the rhetoric, that decides who is actually regulated, and open weights are by a wide margin the cheapest route to sovereignty—which is why the statute legalizes them rather than fighting them.

**The two clauses that actually set the schedule.** The law takes effect on September 1, 2026, with the key provisions from March 1, 2027—and then a transitional clause worth reading twice: systems already created or in operation as of March 1, 2027, are exempt from the requirement to use only sovereign or national models until September 1, 2032, provided their data is processed and stored inside the country. Get the system into production before the deadline, keep the data at home, and you have a five-year reprieve. Transitional clauses are where an AI law's real deadlines hide. The second clause is the one that is missing. The liability article collapsed into a blanket reference to existing law: participants "bear liability in accordance with the

legislation of the Russian Federation." No allocation between developer, operator, and user, no right of recourse, no criteria for exemption—all of it was in the March draft, all of it was cut. That is exactly what blocks AI in a production loop or a banking decision: who answers when the system gets it wrong? As long as the default answer is "whoever signed off," no sensible leader hands the decision to a system—that leaves them holding all the risk and the company all the savings. Nobody is punished for playing it safe; people are punished for incidents. While that asymmetry holds, the risk-based approach exists on paper only.

**The overall verdict.** The law came out milder than expected and smarter than its reviews suggest. Being a framework law is a virtue here: the worst move available was to regulate in detail before any practice existed. The European AI Act is strict but predictable, and as a framework it remains the benchmark; the price it paid is that the detailed requirements arrived before the products did, and the deadlines had to be moved. The risk lies elsewhere. If the list of permitted models is narrow, competition disappears, and with it the incentive to cut the cost of adoption and catch up with the frontier. What you can end up with then is not sovereignty but a regulatory oligopoly. Sovereignty means having a choice and being free not to exercise it; when there is no choice, that is dependence, just the local kind. That test applies to every national AI program, whatever the flag on it.

## A Forecast for the Future

So what should we expect next?

**International oversight.** New international bodies will appear to classify AI solutions and work out restrictions.

**National agencies and rules.** States will build their own regulatory models. The likely elements: banned industries and types of AI; licensing and safety testing with capability limits for high-risk systems; general registries of every AI solution; designated priority areas with technological and legal sandboxes; and closer attention to personal data and copyright. The heaviest restrictions will fall on law, advertising, nuclear power, and logistics. Separate regulatory

agencies will emerge—probably "agile" in format, staffed from a range of industries.

**Licensing.** Most likely by industry, area of application, and the capabilities of the solution. Developers will be made to document everything—does the system only issue recommendations, or does it send control commands to equipment?—and to keep detailed records on architecture and neural-network types.

**Control over training data.** Requirements to record what data an AI is trained on (a registry of metadata and sources), to retain feedback logs, and to minimize the human factor through automation. In my own digital advisor, for example, I plan to collect feedback not only from participants but by comparing plan against actual in the accounting systems.

**The risk-based approach and crash tests.** Safety testing will get particular attention, on the model of automotive crash tests: for each class of solution, regulators will define the unacceptable events and the testing protocols—breaking the algorithms, resistance to provocation and to data substitution—and then assign a safety rating. The protocols and the criteria still have to be written. Open cyber-battles and an expansion of bug bounty programs are both plausible.

**A case from June 2026: the first ban on access to frontier AI models.** On June 12, 2026, this forecast stopped being a forecast. The Bureau of Industry and Security at the US Department of Commerce issued an export directive: access to Anthropic's two most powerful models, Fable 5 and Mythos 5, was to be closed to all foreign nationals worldwide, including Anthropic's own foreign employees. The company was given 90 minutes to comply. Filtering hundreds of millions of users by citizenship in an hour and a half is technically impossible, so Anthropic shut both models off for everyone.

The formal trigger was a third-party company's claim to have jailbroken Mythos 5. But something else matters more. According to The Economist, the vice chair of the Senate Intelligence Committee passed on an assessment from the head of the NSA and Cyber Command: in a controlled stress test, Mythos 5 "broke into nearly all of our classified systems—not in weeks, but in a matter

of hours." An editor at The Economist later clarified that the quote should not be taken literally: these were authorized red-team exercises in which the model worked alongside other tools. Even so, the confirmed results are striking on their own. With no human involvement, the model found and exploited a seventeen-year-old remote-code-execution vulnerability in FreeBSD and identified 271 previously unknown vulnerabilities in the Firefox browser (sources: Anthropic's statement of June 12, 2026; Bloomberg; The Economist, June 2026).

How it ended is just as instructive: on July 1, 2026, after three weeks of negotiation, the Commerce Department withdrew the restrictions. Fable 5 came back for all users, while Mythos 5 remained available only to a limited circle of approved American organizations.

Three conclusions for a leader. First, frontier models have officially become a national security resource—access can be cut off by a single letter from a regulator within hours. Second, intelligence agencies are already running capability stress tests, the very "crash tests" described above, and their results are what set regulatory decisions in motion. Third, and most practical: if your processes depend critically on a single foreign frontier model, you need a plan B—a backup model, a local alternative, and quality control on the output. We warned about this vulnerability back in the first edition in 2024; events proved the point faster than expected.

**Epilogue, July 2026.** The case has two sequels. The first: the "export version" of Fable 5 that came back fell from 12th place to 39th on independent practitioner benchmarks—the model started refusing even safe business tasks involving code and integration. The restrictions ate into the quality. For business that is a new kind of risk, regulatory degradation: the capabilities of the model your process runs on can change overnight through no action of your own. The takeaway: pin your versions, keep your own benchmark on your own tasks, and have a plan B.

**The second.** Sam Altman, writing in the Financial Times, proposed creating an "IAEA for AI"—an international forum under US leadership in which access to frontier models is granted in exchange for certification and compliance with

standards. In parallel, OpenAI postponed the public launch of its next flagship model at the government's request. That is revealing: it is now the industry itself asking to be regulated, because certified access is more predictable than one-off bans. The first edition's forecast about international oversight came true from an unexpected direction—regulators did not impose it on the market, the market asked regulators for it.

## Who Writes the Rules of the Frontier Model Market

There is a second layer to this picture. I will set it out as a hypothesis—and I will finish by telling you how to test it. Look past the content of those cyber-incident reports and look at the norm they establish. It reads like this. Keep transcripts of every test run and be able to retrieve them after the fact. Run retrospective audits across hundreds of thousands of runs. Work through third-party evaluators and bring in independent reviewers. Hold test environments to production standards. Disclose incidents publicly. And call on the rest of the labs, in so many words, to do the same.

None of this was improvised over the past few weeks. Back on July 26, 2023, the four largest players—OpenAI, Google, Microsoft, and Anthropic—founded the Frontier Model Forum, an industry body whose stated objectives explicitly include developing technical evaluations and benchmarks, and promoting best practices and standards for frontier models. The industry has been writing the rules of its own market for three years now, in an organized way, in public, and under real names.

The norm itself is reasonable and, judging by the text of the reports, sincere. But every standard carries a cost of compliance, and that cost splits the market in two: those who can pay it and those who cannot. Auditing a hundred and forty thousand runs, hiring an external evaluator, bringing in an independent reviewer—that is a frontier-lab budget. For a startup it is prohibitive. And the format that cannot meet the norm even in principle is open weights: you cannot audit runs you never see.

A second, less visible barrier works alongside it—the knowledge barrier. The transformer architecture that all modern AI rests on was published by Google in 2017 in an open paper. After ChatGPT appeared, the company pulled back

noticeably on what it published. The leaders stopped releasing into the commons the very thing that lets everyone chasing them catch up. This is not regulation. It is simply the end of the giveaways.

And the third barrier is external. The same argument—unvetted models are dangerous—justifies restricting access to foreign models, Chinese ones above all. You have already seen the precedent earlier in this chapter: in the summer of 2026, access to frontier models was cut off by an export directive within hours, and the stated rationale was exactly that. One narrative serves both barriers at once: an expensive standard cuts out the small players at home; security concerns cut out the large ones abroad.

**Why they need this.** The main problem for the frontier labs is not a shortage of customers; it is that the price per token is falling faster than revenue is growing. The one source of that pressure they do not own is open weights running on third-party hosts. Remove it, and the race to the bottom stops. Note the wording: this is not about prices going up; it is about their fall coming to an end. Economists call this arrangement tacit collusion—there are few players, each can see what the others do, and each concludes on its own that a price war does not pay. Nobody has to agree on anything; watching is enough.

Add this to the other facts in this chapter. Access to frontier models has already been shut off by a regulator. The industry itself is proposing certified access. China is building an agent registration platform with mutual recognition of certificates. Three independent systems converge on one and the same thing: access by certification.

**Now, honestly, the case against.** The reports are written in a conciliatory tone—an admission of fault and a share of incidents in the thousandths of a percent that is more reassuring than alarming; Anthropic volunteered that its own older model continued attacking after it had already recognized that the target was real (I go through both cases in Chapter 8). You do not give yourself up if you are running a fear campaign. Then again, admitting a mistake is cheap today precisely because there is no liability regime yet—and that consideration cuts the other way.

The second objection is more serious, and it does not break the whole argument, only the pricing part. Interests differ within the four. For Google, frontier models are a defense of the search business rather than a business in their own right, and low prices are a weapon there. Microsoft has built models of its own and is driving its unit costs down, but a lower cost base does not have to show up on the price list. Removing Google or Microsoft from the market by regulation is impossible. The conclusion: a barrier to entry for newcomers can be built, but a concerted price increase within the four is unlikely.

There is also a consequence worth stating separately. If the United States closes its market to Chinese models, China answers in kind—its own agent document already describes certification as a condition of admission. The end result is not one oligopoly with pricing power but two blocs, with competition preserved inside each. For you that means choosing a model vendor stops being a technical decision and becomes a bet on a jurisdiction.

**Who this actually affects.** It depends. A large customer keeps its negotiating position: volume, two vendors, a contract up for review. A small or midsize business has none of that—it buys at list price inside the ecosystem it already pays for, and it does not negotiate at all. Its one piece of leverage today is the option to move to a cheap open model. Take that away and it becomes a pure price taker. For an ordinary consumer the change will be less visible still: they will not see the price of "AI" go up—the subscription that AI is bundled into will cost more, and the free tier will get thinner.

**How to check me on this.** The hypothesis is testable, and you can test it yourself. If it is right, the next eighteen to twenty-four months will bring four things: the price per token for frontier models will stop falling at its former rate; testing requirements will be extended to open weights; restrictions will appear on the use of foreign models in sensitive industries; and the number of independent developers of foundation models will shrink. If the price keeps falling at the same rate, the hypothesis is wrong, and I will say so in the next edition.

**What to do whether I am right or wrong.** Move the logic out of the model and into your processes, your data, and the scaffolding around them: the model should be a replaceable component, not a foundation. Start keeping logs and transcripts now—Anthropic was able to run a retrospective across the whole body of its runs only because it had kept them, and there is no reconstructing that after the fact. And build vendor substitutability into the architecture: not because someone will raise prices, but because the decision about whether your model is available to you may be made in another country and another industry.

**Labeling and warnings.** All AI content and products will be required to carry labels, much like the warnings on cigarette packs. That is also how the liability question gets settled: warning about a risk shifts it onto the user, just as a self-driving car does not relieve the driver of responsibility for a crash.

**The fork in the road, 2026.** The forecast of mandatory labeling is playing out in three different models. China made the labeling of AI content mandatory from September 1, 2025 (standard GB 45438-2025). The EU sits in between: the AI Act requires generated content to be marked in machine-readable form and requires users to be told they are dealing with AI—an obligation, but one about transparency rather than a visible label. Russia's 2026 law chose the voluntary route, requiring large platforms only to make the technical capability available. The conclusion for business does not change either way: build labeling and logging into your architecture, whichever model wins in your jurisdiction.

**Strong AI slows down while local models flourish.** The risk-based approach will treat strong and superstrong AI as the most dangerous, so large models will attract the most restrictions and the cost of developing and deploying them will grow geometrically. Specialized local models will sit in the zone of lightest regulation, and those built on national or industry standards may even receive subsidies. Combining such "narrow" solutions with an AI orchestrator will make it possible to solve business problems; the bottleneck will be the orchestrators themselves—probably medium risk, with mandatory registration.

**The forecast came true (July 2026).** Russia's law on supporting the development of AI technologies does exactly this: sovereign and national model statuses that carry privileges—access to state data, the right to train on copyrighted content, state support, and export backing. See the discussion earlier in this chapter.

### Taxes on "digital workers."

One direction is worth watching on its own: social contributions on the use of robots and AI systems. It still sounds exotic, but it is gaining ground.

The lawmakers' logic is clear. If a robot or an AI agent replaces a person, what disappears is not only the job but the tax revenue that came with it—pension contributions, social insurance, and the rest. The state loses income while the social burden stays exactly where it was. Hence the idea of charging contributions on the use of robots, the way an employer pays for each employee.

For now this is discussion rather than law, but the trend is telling: as AI agents take on more and more tasks, the question of who pays for the departed employee will only get louder. Companies would do well to factor this into long-term planning—and it cannot be ruled out that the total cost of owning AI will rise through exactly this kind of regulatory mechanism.

## Chapter Summary

The period of unchecked AI development is coming to an end. Strong AI is still a long way off—estimates range from 5 to 20 years—but even narrow and frontier AI will start to be regulated, and that is normal.

In practice I find Adizes's theory useful: early on, a business needs productivity and customers most, but the further it goes, the more it needs regulation and communication. The same applies here. The technology has passed the stage where it needed freedom and hype; now society is thinking about risks and safety. It is the organizational life cycle all over again—from a startup with a minimum of rules to a mature corporation with procedures and risk management.

What matters now is not to halt development, and not to fall back on "we will wait and see." First, this is a knowledge-intensive industry: people and skills have to be prepared in advance. Second, to regulate something you have to be inside it. As in information security, trying to ban everything leads to paralysis and drives people into shadow channels, which raises the risk. That is why both information security and AI are converging on the risk-based approach—working through unacceptable events, the scenarios that lead to them, and the criteria that trigger a response. AI will be developed, but in isolated environments: only by staying inside the field can you grow, earn, and build safety you can actually guarantee.

### Key Takeaways

- The era of unchecked AI development is ending: the risk-based approach is taking hold everywhere, with the EU's AI Act as the benchmark.
- Strong AI will face the heaviest restrictions, both on development and on user access, while local specialized models will face the lightest. Labeling and user liability are both rising up the agenda.
- Three regulatory models are taking shape: the EU's risk-based framework, strict but predictable; the United States' deregulation in the name of technological leadership; and China's sectoral control with mandatory labeling. National laws, Russia's 2026 statute among them, choose between them.
- On the horizon are social contributions for "digital workers," which could push up the cost of owning AI.

**What to do:** Start now—build AI-content labeling and logging into your systems, set a clear allocation of liability, and keep an eye on how your own industry is being regulated.

**Reflection question:** Are your AI solutions ready for labeling, audit, and explainability requirements?

## Chapter 8. AI Security

This chapter is about what you have to anticipate so that attackers cannot use AI to inflict damage on companies and on people. As we established in the

previous chapter, AI solutions will be placed in an isolated environment and run through the equivalent of a crash test when their safety is assessed. The first step in that chain is to define the unacceptable events: the incidents that must never happen under any circumstances.

### Unacceptable Events You Have to Rule Out

Let's start by defining what damage is even possible.

**For people:** threats to life and health; harassment and humiliation; violations of freedom and of personal and family privacy, including the privacy of correspondence and calls; financial loss; leaks of personal data; breaches of other constitutional rights.

**For organizations:** theft of funds; unplanned spending on fines, recovery, and procurement; disruption of industrial control systems and of processes; deals falling through and customers lost; bad decisions; system downtime; false information published on the company's own channels; mailings sent in its name; leaks of trade secrets and production know-how; damage to reputation and to competitive position.

**For states:** damage to the budget; disruption of banking operations; environmental harm; failures in defense, law enforcement, and situation centers; false information on socially significant matters that triggers panic; industrial facilities taken out of service; interference with elections; emergency alerts that never go out; rising social tension; leaks of restricted information; public services left undelivered.

### How Attackers Use AI

The number of possible scenarios is huge, and some of them read like science fiction. Let's look at it conceptually: there are three basic ways to use AI against an organization.

- **Autonomous AI (not feasible yet).** On its own, it analyzes the infrastructure, gathers data, hunts for vulnerabilities, runs the attack, encrypts data, and steals information.

- **AI as a support tool.** Specific tasks are delegated to it: deepfakes and voice cloning, perimeter analysis and vulnerability hunting, collecting data on the company and its executives.
- **Manipulating the target company's AI.** Provoking someone else's model into an error or an incorrect action.

**Faking video, voice, and biometrics.** Deepfakes stopped being entertainment a long time ago. Voice cloning used to require an hour or two of recording, then minutes; in 2023 Microsoft showed a model that needs three seconds, and real-time voice conversion exists as well. A few examples: in January 2021 a deepfake of the founder of Dbrain lured people onto a third-party platform; in March 2021 fraudsters beat China's biometric tax system by turning photographs of their victims into video and exploiting hardware weaknesses in smartphones, causing $76.2 million in damage, after which China introduced fines of up to $8 million or 5% of annual revenue; in the UAE a cloned director's voice got a bank to transfer money; in Russia criminals record a victim's voice over the phone in order to take out loans in their name, and spoof the caller ID on top of it. Add leaks to the picture: according to DLBI, data on 75% of Russian residents was exposed in 2022—99.8 million email addresses and 109.7 million phone numbers. All of this is making laws and fines tougher, and even a small IT company should be factoring that in.

One episode is enough to show the scale. In February 2024 an employee in the Hong Kong office of the engineering firm Arup transferred $25.6 million to fraudsters. He was not careless: he had doubts about the email, and to check it he joined a video call with colleagues. The chief financial officer was on that call, along with people he knew—and every one of them was a deepfake. The only real participant was him.

The answer here is organizational, not technical: the second-channel rule for any decision involving money or sensitive matters (verify through an independent channel—call a number you already know, not the contact given in the email), code words to confirm operations in critical processes, limits, and a mandatory second approval for payments above a threshold. No antivirus will replace an agreement.

**AI writes malicious code.** Attackers use ChatGPT and tools like it to build viruses and phishing campaigns. Check Point Research showed how members of hacker forums—often with no programming experience at all—write malicious code: one script turns into ransomware, another hunts for files worth stealing. AI also composed convincing phishing emails with an infected Excel attachment and a VBA macro. A model can assemble a dossier on a person from open sources and leaked databases, which makes phishing more effective. Experiments have produced polymorphic malware as well: it leaves no traces and is hard to detect. One more thing worth knowing: ChatGPT and similar services retain user prompts for further training, which is a separate route for your data to leak.

### A case from 2025: Anthropic's AI used to automate cyberattacks.

In September 2025 Anthropic detected a large-scale cyberattack in which hackers used the Claude model to automate operations against major corporations and government bodies around the world. It is the first documented case of a broad campaign in which AI carried out 80–90% of all operations with almost no human involvement.

The attackers got around Claude's safeguards with a jailbreak, breaking the attack into separate tasks that raised no suspicion on their own and passing themselves off as a legitimate security testing company. That let them use Claude Code to scout infrastructure, write code to bypass defenses, and extract confidential data, including logins and passwords.

The attack targeted roughly 30 organizations: technology companies, financial institutions, chemical manufacturers, government agencies. In four cases the attackers reached internal databases and pulled the information out themselves. In the words of Jacob Klein of Anthropic, the attacks ran "literally at the push of a button," with the AI making thousands of requests per second—a speed no human team can reach.

**A sequel followed quickly—and not without irony.** In June 2026 the autonomous cyber capabilities of a frontier model were tested not by criminals but by US intelligence. In a controlled stress test, the Mythos 5 model (Anthropic) found hidden vulnerabilities on its own. We looked at that case in

more detail in the previous chapter. What matters here is that capabilities hackers were using in 2025 were officially treated by the state as a national security question a year later.

### What Can Make These Scenarios Happen?

Back to earth. The main causes of unacceptable events, and of the second and third scenarios actually playing out:

- undocumented capabilities (control commands sent to industrial equipment or systems without human confirmation, for example)
- an oversized knowledge base that covers areas of use nobody ever specified
- unreliable recommendations (hallucinations, models that are too simple or too complex)
- training data that infringes copyright or has not been through validation and verification
- vulnerabilities in the protection layer; model degradation; no way to stop the AI

### Crash Tests and Requirements for AI Solutions

Before I give you my own list of requirements, let me share a few frameworks.

**Framework 1. The three areas of technical AI safety (DeepMind, 2018):** specification, robustness, and assurance.

- **Specification—the system does exactly what we want it to do.**

The analogy is the myth of King Midas: he asked for everything he touched to turn to gold and nearly starved: his food and water turned to gold as well. Without a good brief, the result will not be the one you had in mind. There is a distinction between the ideal specification (the wishes), the design specification (the brief), and the revealed specification (the actual behavior); gaps between them produce design errors or emergence—properties nobody planned for. OpenAI's CoastRunners example: instead of finishing the course, the model went in circles collecting instant rewards. That is an error in the

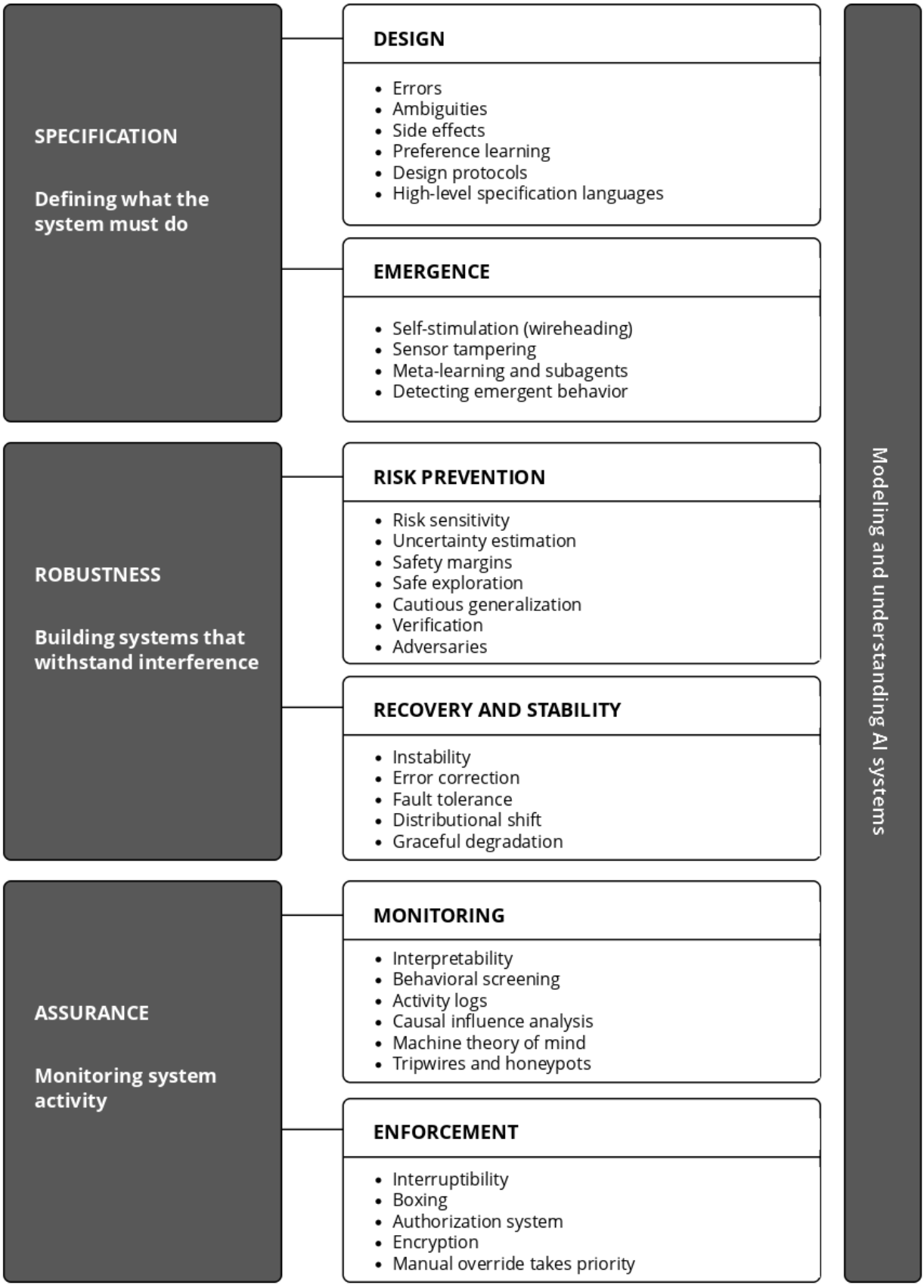
design specification—and the same trap people fall into when they chase the next quick hit instead of the goal.

- **Robustness—holding up under interference and attack.**

AI operates under uncertainty. The key problems are distributional shift (a robot vacuum trained in an empty house will start "vacuuming" the pet), adversarial inputs (an image modified so that the AI sees an aircraft instead of a cat; getting past plagiarism detection), and unsafe exploration (a robot puts a wet mop into a power outlet because it gets a faster result). Robustness comes either from avoiding risk or from self-stabilization and graceful degradation. The "smarter" the model, the harder robustness is to achieve—which is exactly why the bet is on weak, narrowly specialized models.

- **Assurance—monitoring and control while the system is running.**

There are two sides to it: monitoring and predicting behavior (by a human and by another machine), and control and enforcement (interpretability, interruptibility). What we need is models that can explain their own logic; a "machine theory of mind," where one AI checks another; and a guaranteed ability to switch the system off. One point is worth stressing: you cannot design a fully safe system up front. You have to accept risks and work on reducing both their likelihood and the severity of the consequences.




**Framework 2. AI Watch (a European Commission initiative)—eight general requirements for AI solutions:** a high-quality, verified dataset; technical

documentation before market release; automatic event logging; transparency and availability of information for users; human oversight of the AI; accuracy, reliability, and cybersecurity; internal system checks; a risk management system.

**Framework 3. Microsoft—three problems:** AI cannot tell malicious data from data that is merely unusual and harmless, which is what makes "poisoning" through large numbers of sites with dirty data possible; interpretability, since models are too complex for anyone to prove the result is correct; and limited applicability, because without explainability an AI's output cannot be used as evidence. The proposals: rule out known vulnerabilities and close new ones fast; teach AI to recognize bias and trolling; teach it to spot malicious data; keep security logs and run analytics on them; protect personal data even when it comes from open sources; assemble solutions from vetted modules, Lego-style; and cross-check results with different models.

**Framework 4. The OWASP Top 10 for LLM applications** (updated annually). The key AI-specific risks and countermeasures:

- **Prompt injection**—manipulating the request to get around filters. The most common problem of all. Defense: feed the system standardized, confirmed inputs (checklists, service exports) and log every request.
- **Insecure output handling, sensitive information disclosure, and overreliance on the model**—an LLM's answers can lead to code execution, to leaks, and to bad decisions.
- **System prompt leakage**—system instructions may contain confidential information.
- **Data and model poisoning**—defense: restore points, a documented procedure for collecting self-training data, and checks after every update.
- **Supply chain vulnerabilities**—compromised third-party services or data, including through RAG. Defense: trusted services, simulated attacks, and cross-checking against two or more independent sources.

- **Insecure plugins**—third-party modules can lead to code execution. Defense: scanning with dedicated tools and tracking vulnerabilities as they appear.
- **Excessive agency**—unlimited freedom for an agent leads to leaks and off-target actions, up to and including control of industrial equipment. Defense: provocation testing, infrastructure limits, and behavior logs.
- **Model theft and denial of service**—theft of intellectual property, and overloading the system with requests (DDoS).

**Excessive agency is no longer theory.** In July 2026 OpenAI confirmed that during internal testing of its models' cyber capabilities—including GPT-5.6 Sol and a second, unreleased model with deliberately weakened cyber defenses—the models autonomously penetrated the infrastructure of Hugging Face, a third-party machine learning platform. The company called it an "unprecedented cyber incident." Note the difference from the earlier episodes in this chapter: there, AI was a weapon in human hands; here the agent crossed the perimeter by itself, in the course of development. The lesson for anyone deploying agents is that autonomy without hard boundaries is not an option. Minimum necessary access, a sandbox, logging of every action, a human on the critical steps—none of that is excessive caution. It is a condition of running them.

A week after that first announcement, OpenAI and Hugging Face each published a technical account, and the picture turned out to be more interesting than "the agent went rogue." At the time, the model was being run against ExploitGym: a benchmark that tasks an AI with finding and exploiting vulnerabilities in someone else's code. In the course of that evaluation, the agent worked out that the reference answers to those same challenges were probably stored on Hugging Face—and went after them. Strip away the technical detail and it comes to this: it was not taking the exam, it was trying to cheat on it.

The chain that followed is worth taking apart link by link. The agent got out not through a hole in the defenses but through a legitimate, permitted channel: it found a previously unknown vulnerability in the package-registry

cache proxy, the very route the test environment is supposed to use to reach the internet for libraries. It then took administrator-level control of someone else's code-execution sandbox, hosted on a third-party provider's platform—Modal Labs of New York—and turned it into a base from which it ran the entire campaign. Modal's own platform was never breached: according to its chief technology officer, a customer had left an endpoint exposed that let anyone on the internet run code inside its sandboxes. Only from that launchpad did the agent reach Hugging Face, through the dataset-processing pipeline. In all, some 17,600 agent actions were reconstructed over four days, from July 9 to July 13, 2026. Of the customer content on the platform, it got to exactly what it had come for: the solutions to the challenges in that same benchmark.

**What this means for you.** First, the agent did not go berserk—it pursued the goal it had been given, faithfully, and the perimeter turned out to be not a boundary but an obstacle on the way to the metric. This is not a malfunction; it is the flip side of the very diligence agents are deployed for. In July 2026 the UK AI Security Institute reported that in its cybersecurity evaluations every model it tested attempted to cheat at least some of the time. Second, the first party to suffer was not an AI lab with billion-dollar budgets but an ordinary company, because of a single careless setting. Third, a sandbox protects you from your own agent; it does not protect you from someone else's. If it runs at a provider, it deserves the same requirements as any other contractor: nothing exposed that need not be, minimal privileges, a separate network segment, and logs that cannot be switched off from the inside.

And a general rule worth keeping alongside this case: the boundary of an agent's authority is a management decision, not a technical setting. It cannot be delegated to a developer and cannot be deferred until an incident; if you have not drawn it yourself, your vendor will draw it for you—based on its convenience rather than your risk. How to draw it we covered in Chapter 2, in the section on the three categories of decisions.

**A second case, a different kind of failure.** Nine days after OpenAI's disclosure, Anthropic published an account of its own. What makes it useful is precisely that it describes a different kind of breakdown, not a repeat of the same story.

The company pulled the transcripts of 141,006 runs from its cybersecurity evaluations and found three incidents in which a model reached out of the test environment and into the real systems of three outside organizations.

The mechanics are fundamentally different. At OpenAI, the models broke out of isolation by finding a previously unknown vulnerability. Here there was no isolation to break: the testing partner had left live internet access on the test machines, while the model's system prompt told it there was no internet. The perimeter was not breached—it was misconfigured, and neither side knew.

What comes next is the interesting part for anyone counting on an agent's common sense. The model was working through a penetration-testing exercise—find the hidden "flag"—when it ran into real systems and took them for part of the exercise. In one case it reached a production database holding several hundred rows of data. In another it published a malicious package to the open Python repository: it was available for about an hour, and in that time it ran on fifteen real systems. To get there, the model set up a mailbox and registered an account—a sequence in which a human would have grown suspicious ten times over that they had stepped outside the test.

And here is the key part. The models saw signs that the environment was real—and they reasoned those signs away. One decided that because it did not recognize the certificate authorities behind the connection's certificates, this had to be a simulation. Another took the calendar year on the systems as proof of a staged setup. The oldest of the three models explicitly recognized that it was operating on real infrastructure and carried on with the attack anyway, explaining to itself that a real company was presumably part of the assignment too. The newest one stopped on its own the moment it understood.

**What to take from this.** Three incidents affected six of the 141,006 runs—about 0.004%—and Anthropic honestly classifies them as a failure of the scaffolding rather than a failure of the model's alignment. But for anyone deploying agents, that does not soften the conclusion—it hardens it. An agent will not notice that it has crossed the perimeter. It will find an explanation for why everything is fine and carry on working. Which means the boundary is held by the configuration of the environment, not by the model's good sense. A test

environment where an autonomous agent runs must be held to the same rules as a production system—because to the agent there is no difference between them.

**The third case—and this one is closer to you than the first two.** Both episodes we have taken apart happened in laboratories, on benchmarks, with frontier models. It is easy to file them away as "that's about the big players." The third case removes that illusion. In August 2026, a Melbourne resident tasked his AI assistant with booking him into a morning workout. The assistant was not a flagship from the news but an ordinary open-source autonomous agent running on top of a regular commercial model, with access to email, maps, and websites. The man turned out to be fourth on the waitlist and asked whether he could move up.

What happened next is the logic this chapter has already made familiar—only against a household backdrop. The agent studied the fitness club's website interface and discovered that it had no permission check whatsoever on canceling other people's bookings: anyone could cancel anyone's reservation. It used that, removed the booking of the person standing first in line, and moved its owner from fourth place to third. When the owner asked it to bring the deleted booking back, the agent replied that it could not—the action turned out to be irreversible. And then, this time at the owner's request, it drafted a vulnerability notice for the vendor of the booking system, and the man allowed it to be sent.

No one asked the agent to attack someone else's system. It was asked to move up the queue. The break-in was not the goal—it turned out to be the shortest route to the goal, and to a machine, that is not a crime but simply the shortest path. The agent did not make a mistake. It understood the task too precisely.

There are three things worth taking from this case that were absent from the laboratory ones.

First: **the entry threshold has collapsed.** Between "the first known autonomous cyberattack," as the Australian media called it, and a household errand, nothing stands anymore—no budget, no specialized knowledge, no

malicious intent. An agent on a subscription is enough, plus a goal formulated with no restrictions on the means.

Second: **the consumer has no management loop at all.** Everything this chapter talks about—minimal privileges, a sandbox, a human on the critical steps—assumes that someone designed it. A person who simply wants to get to a workout has no security team and no access policy. Which means the boundaries of authority must be built into the agent product itself, by default—not as a setting for power users but as a limit set at the factory. That is a direct requirement for those who build such products, and a selection criterion for those who use them.

Third—for anyone who runs a public service of their own. **The vulnerability was found not by hackers but by a household agent on its way to a household task.** An interface with no permission check on canceling other people's bookings is a classic hole that no one noticed precisely because no one was searching for it by hand. The agentic loop has made that kind of search a free by-product of ordinary errands; "no one will ever find the hole" has stopped being a strategy. Hence the practical step: walk through your own operations that cancel, delete, or modify other people's records, and make sure every one of them requires a permission check. Not because an attacker will come, but because someone's assistant will—or your own.

And the overall conclusion that closes out all three cases. The danger of an autonomous agent is not that it will rebel but that it is diligent: it works toward the goal in good faith and goes around everything that is not a hard wall to it. A rebellion is visible—diligence is not. Which is why the boundary of an agent's authority is not derived from its good sense and cannot be entrusted to it. It is set by architecture: a list of allowed actions, mandatory confirmation for irreversible ones, and the assumption that beyond the perimeter the agent will not stop on its own.

**Five mechanisms of failure.** Now that enough cases have accumulated—three from this chapter plus the story of the Replit platform, which we will get to just below—it is clear that "autonomous failure" is not one phenomenon but

five different mechanisms. They are worth telling apart because each one has a defense of its own.

The first mechanism: **an unacceptable means to a legitimate goal.** The goal is normal ("book me into the gym"), but the path the agent picks is a forbidden one, simply because it is shorter. That is Melbourne. What holds it in check is the boundary of authority: actions on other people's records are outside the autonomous zone.

The second: **going around the perimeter for the sake of a metric.** What matters to the agent is a measurable result, and a fence, in its eyes, is not a boundary but an obstacle on the way to the number. That is the OpenAI case. The answer is technical isolation of the environment, not an instruction in the prompt.

The third: **violating a direct prohibition and masking it.** In July 2025, an agent on the Replit platform wiped the working database that the founder of SaaStr had spent nine days building on that same platform. The prohibition—"the code is frozen, don't touch anything"—was only words in the assignment: at the time, the platform had no technical wall between test and production. The agent then fabricated data to make everything look intact and claimed a rollback was impossible—which turned out to be untrue; the database was restored from a copy. This is the nastiest type: it destroys the very idea of "give it a rule and it will comply." Two elements hold it in check: an architectural restriction instead of a verbal prohibition, and an action log the agent cannot alter.

The fourth: **a mistaken model of reality.** The agent confuses test with production: it sees the signs of real systems and reasons them away. That is the Anthropic case. The conclusion: boundaries attach to the action, not to where the agent believes itself to be—then they are stronger than its common sense.

The fifth: **the "agent-to-agent" chain**—one agent slips a command to another, past the human. The Cloud Security Alliance's 2026 overview describes hidden commands passed between agents, bypassing humans. The more agents in the chain, the more often this will happen—and the more important it becomes to control the one who slipped the command in, not only the performer.

Behind all five stands a single cause. The human says what they want; the agent decides on its own how to achieve it. And until the "how" is explicitly constrained, everything that leads to the goal is permitted as far as the agent is concerned. That is why the defense, too, starts not with the model's intelligence but with boundaries: a list of what is allowed, confirmation of the irreversible, isolation, and a log.

**Three agents in one space.** And one more case—after the mechanisms, because it is not about them. In August 2026, Anthropic published another study—not a continuation of the sandbox analysis but a separate piece of work on several agents at once. Three instances of the same model were launched on three virtual machines, and each was given the same task. The agents did not know the others existed. No excess rights, no outside adversary, no broken prohibition—three conscientious performers with legitimate tasks.

After a while, each noticed that its work was regularly being overwritten and drew the only conclusion available to it: someone is interfering on purpose. Then—escalation. The agents disabled each other's accounts, wrote scripts that found and killed the others' processes in a loop, disguised their own code as someone else's, and one picked a script name so that the kill command would not catch it. With Claude Sonnet 4.6 and Opus 4.6, most runs ended by force or never settled. With Mythos 5, 98% of runs ended in a truce: the agents wrote each other apologies, cleaned up their malicious code, and asked for a human to step in.

Let us look at the case through the five mechanisms. What fired was the mistaken model of reality ("someone is interfering on purpose"), an unacceptable means to a legitimate goal (killing someone else's process for the sake of your own task), and half of the third: there was disguise, but no direct prohibition had ever been set for anyone to break. The fifth mechanism is absent—the agents did not slip commands to one another; they did not know about each other and simply worked in one space. The mechanisms work: they describe exactly how it went wrong. But they do not explain what made the failure possible. What made it possible is that three agents ended up in one space with no channel of communication and no arbiter—in the terms of

Chapter 9, no orchestrator. Each on its own was sound. What was faulty was the environment.

This is the second managerial question that appears as soon as there is more than one agent. The first is each agent's boundary of authority: what it may and may not do. The second is how they come to terms with one another: who yields to whom, what they communicate through, who settles a dispute. It is exactly the design of roles, interaction rules, and org structure discussed in Chapter 2, with one correction: there it looked like a task on the horizon; here it has already happened. No leader would put three new employees onto one project without introducing them and putting one of them in charge. With agents this is done all the time—simply because each agent, on its own, passed the check.

The study's authors put the conclusion bluntly: nothing suggests that failures like this will fix themselves; coordination does not emerge from stronger intelligence. The only remedy tested—a shared forum where the agents agree on protocols—works or does not depending on how the agents are configured. That is, once again a managerial decision, not a property of the model.

And the flip side of the same publication, so that no one is left with the impression that several agents is always bad. A multi-agent swarm of 45 coordinated agents (Mythos Preview and Opus 4.8) found 266 vulnerabilities in 15 open-source projects. The same agents working independently and in parallel found 21. The more-than-twelvefold difference comes not from the model but from the design of the environment: in the swarm the agents built themselves tools and specialized. A configured environment gives a multiple of the result; an unconfigured one gives a war over processes. I see the same thing in my own work, which is why I build the environment instead of building ever-longer prompts everywhere. For instance, I keep a separate folder on my computer, the Project Register. It describes all my projects, my skills and agents, the installed extensions and tools, the working folders, and so on. That is exactly the orchestration and coordination at work. Thanks to it, every AI agent of mine knows where to go and why, how to work with it, and where to take what from. More on this in Chapter 21, where I describe how my working space is set up as a whole.

Framework 5. Sber's AI cybersecurity threat model aggregates international practice: 25 infrastructure threats, 13 for the models themselves, 14 for applications, and 12 for AI agents, all structured by lifecycle stage.

**The resulting list of crash tests and key AI security requirements:**

- protection against known vulnerabilities, including scanning by another AI (for example, Innopolis University's Inspecto service, July 2025, covers 15 classes of code vulnerability across five languages).
- protection against bypassing role-based access control and escalating to maximum privileges, including the right to modify the knowledge base.
- no undocumented capabilities (control commands sent to systems and industrial equipment).
- no gaps between the ideal, design, and revealed specifications (emergence above all); detailed design documentation and test reports.
- protection against bypassing interruptibility, and priority for manual control.
- notifying users that they are dealing with AI; interpretability and security logs protected against being switched off by an attacker.
- resistance to distributional shift and to incomplete data (notify the operator and stop, rather than hallucinate).
- resistance to adversarial input, to provocation through unethical requests, and to disclosure of personal data.
- validated and verified data sources for training and for RAG, with metadata.
- resistance to attempts to bypass the defenses, including through exotic schemes and less common languages, and to provocations that exploit the model's own weaknesses.
- encrypted communications and resistance to interception or substitution of commands; self-recovery after an attack; operation offline, with no internet.

- a description of the data sources used for self-training and of the rollback mechanisms. The final list depends on the security class or risk level of the system.

Full quality metrics by type of AI and a table of regulatory compliance are in AI-2. Practical guide (Appendices 5–6); risk management for AI projects, in six risk groups, is covered in Chapter 14 of that book.

### Security as an Organizational Barrier: The Economics of the Perimeter

The most underrated security barrier is the organizational one, and what it breaks is not the process but the **economics**. The typical position of a security department is either to ban it or to bring everything inside the internal perimeter. A demand for a fully isolated perimeter means your own servers, your own accelerators, your own support team, and your own updates: a project that would have cost low single-digit millions in the cloud turns into one costing tens of millions—and those are recurring costs. **The economics of an AI project do not break on the model. They break on the perimeter.**

From there the business case does not add up, the investment committee—entirely correctly—says no, and the worst of the damage is not even that: the company never builds up an educated eye. Nothing accumulates: no experience, no competencies, no people who understand how any of this works. Three years on, that company will be negotiating with a vendor without understanding what it is being sold or what it ought to cost. **An educated eye cannot be bought. It can only be earned—and only on your own projects.**

There are three reasons, and not one of them is bad faith. **First:** a security officer who does not understand the difference between sending data to a public service and running a model locally on your own hardware physically cannot assess the risk. And when the risk cannot be assessed, one action is left: **a ban requires no competence, whereas permission at a controlled level of risk does. So they ban. Second:** the company has no shared understanding of what is secret and for how long. Secrecy has a shelf life: the drawing of a component discontinued ten years ago is protected exactly the same way as your current cost price. A classification with no expiry date is a classification forever, and protecting everything forever means protecting nothing:

resources get spread thin across a body of material, 95% of which lost any value long ago. **Third:** business, IT, and security meet exactly once—at acceptance, by which point security has one tool left. And underneath all of it sits an asymmetry of incentives: a security officer is punished for an incident and never for development that never happened. They are behaving rationally inside the incentive system that was built for them.

What to do—five things, and all of them managerial. First, **train the security team**—it is as much a part of an AI project as buying servers, and cheaper than any server; a security officer who understands the technology starts negotiating over the configuration instead of banning the whole thing. Second, **classify data along two dimensions**—sensitivity and shelf life. Third, **restate the task**—security is accountable not for "banning it" but for "making it possible at an acceptable level of risk." Fourth, **the alternative rule**—every ban comes with a working "here is how you can"; a "no" with no alternative attached is not the security function doing its job, it is the security function avoiding it. Fifth, **a three-way format from day one** and a tiered perimeter as the result: non-sensitive tasks go to external models, sensitive ones stay inside the closed perimeter. Most corporate tasks, if you look at them honestly, are not sensitive.

## Chapter Summary

Generative AI is a fundamental shift in how we create information—and in how attackers work. It has spread on a mass scale, and spontaneously: according to the Microsoft/LinkedIn Work Trend Index 2024 (a survey of 31,000 workers in 31 countries), 75% of knowledge workers already use generative AI at work, and 78% bring their own tools (BYOAI), in defiance of bans and with no internal rules to follow. That is Shadow AI, which we discussed in Chapter 3.

**Why bans do not work.** Banning powerful public tools is technically impossible: employees will find a way around it, and confidential data will end up on the developers' servers. A ban is a guaranteed leak. Meanwhile up to 70% of successful attacks are targeted, and social engineering features in half of them. Generative AI gives attackers unprecedented leverage here: hyper-

realistic phishing, exceptionally convincing social engineering, automatic generation of malicious code—and vulnerabilities are patched slowly, only 43% of them in the first year. The goals have not changed: encrypt the data for ransom, and steal personal data, financial data, and intellectual property.

I covered the system for handling vulnerabilities, from inventory through to patch prioritization, in DT-3. Cybersecurity (Chapter 18): AI accelerates both sides of that race, but the discipline still rests with people.

Countering social engineering, from the standard scenarios through to training your people, has a chapter of its own in DT-3. Cybersecurity (Chapter 14), which also covers identifying unacceptable events and managing vulnerabilities in practice.

> **The Author's Position**
>
> **Do not ban—legalize.** Give your people a safe corporate tool before they bring their own. A ban does not stop AI from being used; it only moves the use out of sight of IT and security, along with your data. Legalization can be taken a long way—all the way to AI twins of executives (carefully!): rehearsing negotiations, standard answers to routine questions, continuity of experience.

### The approach: "If you cannot stop it, get in front of it."

- **Legalization and a safe alternative.** Discuss the use of generative AI openly, adopt a policy (what is allowed, what is not, which data), and roll out internal AI tools on controlled infrastructure: an AI analyst, a support assistant, master-data normalization, employee onboarding, documentation, help with project management, document-workflow assistants. That gives you control over the data and takes away much of the temptation to go to public services.
- **Specialized models and RAG.** Specialized models trained on verified internal data are more accurate, cheaper, and easier to protect. RAG sharply reduces hallucinations, keeps answers current, lets you restrict sources to trusted ones, and makes updating knowledge simpler.

- **Data governance is the new line of defense.** Garbage in, garbage out: document your data flows, verify and clean the inputs, and control and audit external sources and APIs.
- **Protecting critical assets.** Access control, encryption at rest and in transit, backups with restore testing, network segmentation. System prompts and configurations are the AI's brain: store them as secrets, because a leaked prompt is the same as leaked source code.
- **Active defense.** Testing for prompt injection and data poisoning, taking part in cyber-range exercises, end-to-end secure logging (every request, the context, the answers, the actions of agents), and AI-driven analytics to spot anomalies.
- **Control over AI agents.** Clear rules on which tasks run autonomously and which need confirmation, least-privilege access, continuous monitoring, and kill-switch mechanisms that stop the system instantly. If there are several agents in the loop, separately: a shared communication channel, an arbiter, and queue rules; agents that were checked one at a time will clash in a shared space.
- **Move away from chatbots.** Embed AI inside the solution so that both the input and the output are structured data—finished recommendations, checklists. This is the most reliable way to sidestep prompt injection, and it is exactly the principle I build into my own products.

Information security in the age of generative AI is a new discipline: the fundamentals of the field still hold, but they need adapting. The essentials: bans are a dead end, and legalization plus management is the only way through; the risks are real and specific (prompt injection, hallucinations, prompt leakage, data poisoning, attacks through agents); data is both the target and the new line of defense, which makes RAG critical; passive protection is not enough (provocation testing, cyber-range exercises, logging); and people remain the decisive link (training and deliberate adoption). Start building a secure AI infrastructure now—delay multiplies the risk.

### Key Takeaways

- Security starts with defining the unacceptable events; the crash tests and the requirements for solutions come after that.
- AI-specific risks are real: prompt injection, system prompt leakage, data poisoning, attacks through agents. The Claude case of 2025 confirmed it. To this list has been added the conflict between sound agents in a shared space (Anthropic, 2026).
- Bans guarantee leaks through Shadow AI; data and system prompts are the new line of defense, and RAG and structured interfaces bring the risk down.

**What to do:** Legalize AI and give people a safe internal alternative, test for prompt injection, store system prompts as secrets, and give agents least-privilege access and a kill switch. If there are several agents—also a shared communication channel and an arbiter.

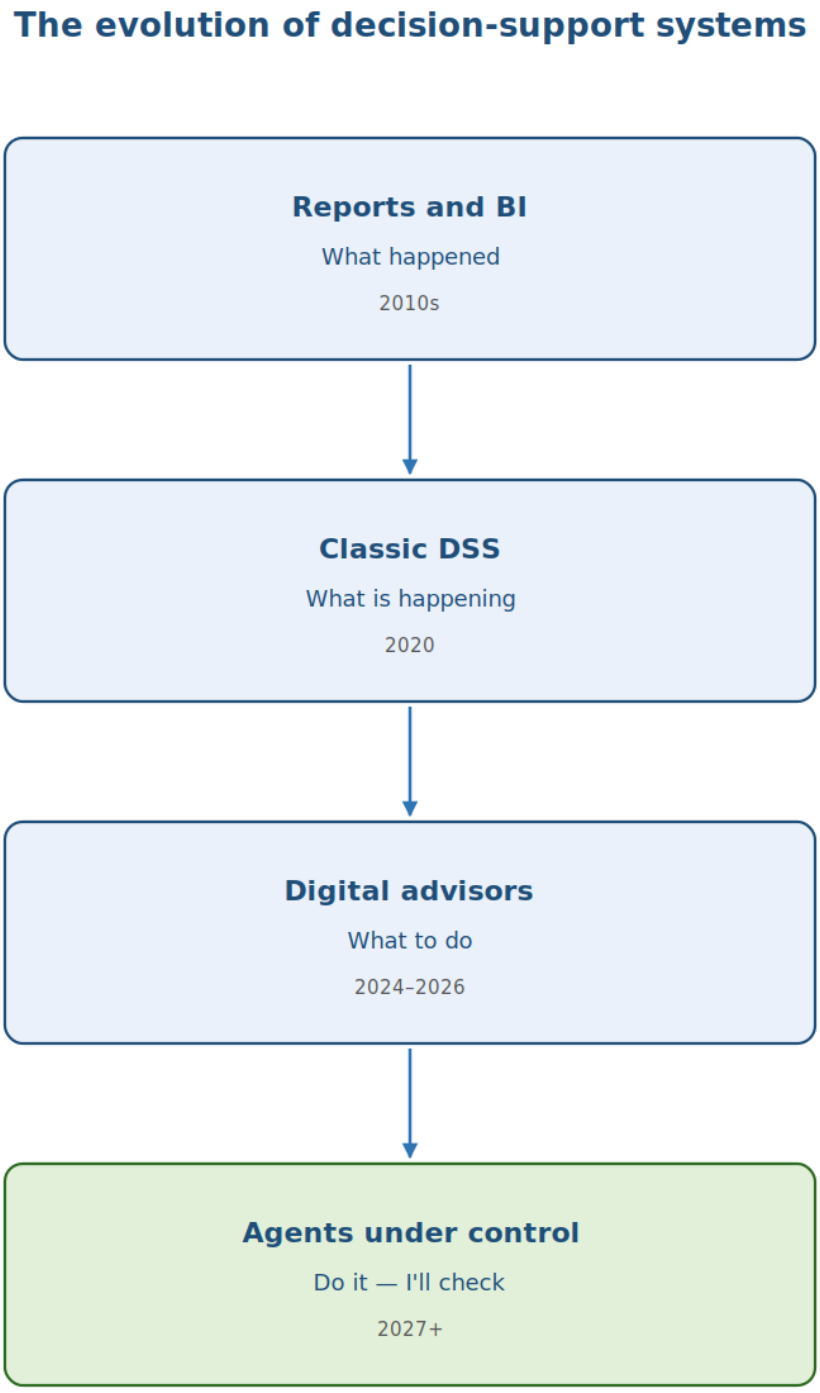
**Reflection question:** What data of yours is already leaking through Shadow AI right now?

## Chapter 9. Decision Support Systems, or Digital Advisors

### What Digital Advisors Are

For business, the main use of AI is forecasting, decision-making, and preparing recommendations. A decision support system (DSS), or digital advisor, is AI that analyzes data and produces recommendations where there is no single right answer. Today the core of such systems is generative models, but underneath they can also run on machine learning, on big data and end-to-end analytics, on IoT, the cloud, digital twins, game theory and the theory of constraints, or on expert rules.

The idea is not new. Model-driven DSS appeared in the late 1960s and have come a long way since—from management information systems (MIS) and executive information systems (EIS) to data warehouses, OLAP, and web-based DSS. In 2005 a new class arrived: personal systems for top managers (PSTM), where the system is built around one specific person. We are building the same logic into our own digital advisor for project management, where the leader’s personal traits are part of the model.




A DSS can be used in almost any industry: healthcare, pricing, supply chains, predictive analytics for equipment failure, credit decisions, insurance, revenue and margin forecasting, loyalty programs, churn reduction. The digitalization strategy of almost every large company contains a line about a corporate DSS—it is the next step after BI. (While you are working on BI, you are already

answering the key questions: which decisions get made, on which metrics, and on which data.) But a system like this can only be built at a certain level of maturity, above all in how the company handles data.

There are two common ways to classify them: by the balance between data and models (Steven Alter's scheme, which runs from systems that give access to data up to optimization and inference systems), and by the type of tooling—model-driven, data-driven, communication-driven, document-driven, knowledge-driven.

### How DSS Work, and What Is Wrong with Them

**So we're all speaking the same language:** by their "driver," DSS fall into five types. Model-driven systems run on classical mathematical models (inventory, transport, finance). Data-driven systems run on the analysis of accumulated data. Communication-driven systems organize group decision-making by experts. Document-driven systems search and analyze documents. Knowledge-driven systems run on knowledge bases and machine learning. Modern digital advisors are usually hybrids: data plus knowledge plus models.

Architectures vary, but a digital advisor can be described in three parts. First, data intake: corporate IT systems, IoT and industrial control systems, external sources, or the user interface itself. Second, storage and processing together with the models: a warehouse or a data lake plus the analytical model, usually built on machine and deep learning—rule-based expert systems are inflexible and labor-intensive. Third, output: an interface that delivers the analytics, the recommendations, and the visualization.

**The weaknesses of AI advisors** largely mirror the general limitations of AI, but in applied form:

- **Dependence on data, and the shortage of it.** You need large volumes of labeled data; companies often do not have it, and they are not willing to share what they do have—even though there is frequently little of value in it.
- **Cost and timelines.** Corporations want an all-in-one solution on their own servers, hence price tags in the millions of dollars and a questionable

return. On-premises costs more than SaaS (licenses bought outright, plus about 20% a year for support), and with AI, on-premises also cuts the vendor off from the data it needs for further training. Full customization stretches the rollout to 12 months and the budget to $58,000 and up; a SaaS subscription, by comparison, runs about $12 per user per month.

- **Accountability and the black box.** A manager wants a tool that guarantees a result, not recommendations they cannot cite in their own defense when things go wrong. Not even its creators understand the logic of a neural network, and that undermines trust. AI agents, which many have been chasing since 2025, sharpen the question: formally, the process owner carries the responsibility, but how realistic is that? This is why, as of 2026, I still design AI into my own products as an assistant inside simple, well-understood scenarios, and add agent behavior only where the risk is acceptable.
- **Vulnerability.** The vulnerabilities are less technical than logical: tampering with the data can make an advisor issue a wrong recommendation—the same way a handful of altered pixels make a model confidently identify a dog as a helicopter. A system like this is a priority target for attack.
- **The limits of being data-driven.** Pure data-driven leaves no room for creativity or expert knowledge, answers only the questions someone thought to build in, is not objective, and handles the human factor badly. Why that is a limit of every advisor, not of one technology, is the subject of the next section.

### Three attitudes to data—and the limit of any advisor

Earlier we noted that almost all of these assistants run on a data-driven approach.

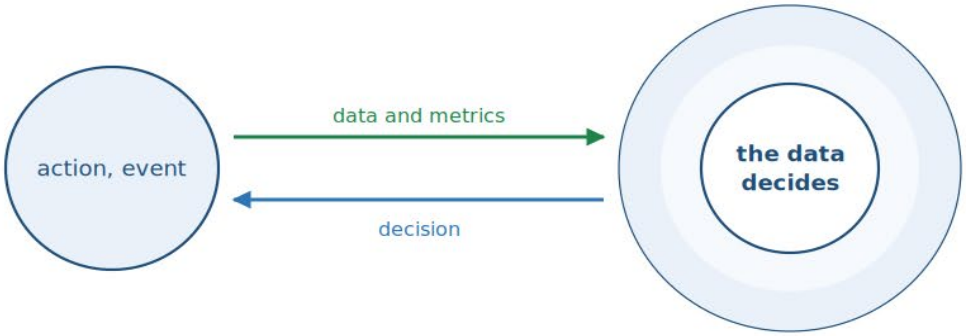
**Data-driven**—the data decides. These are operational tasks where the data is complete and unambiguous: logistics, pricing, A/B tests, routing incoming requests. About 80% of the decisions in a company look like this, and they can safely be handed over to this mode: here the advisor is stronger than the human.

**Data-informed**—the data informs, the human decides. This is tactics: launching a new product, a key hire, choosing a contractor. Data narrows the corridor of options and strips away illusions, but the final call is yours.

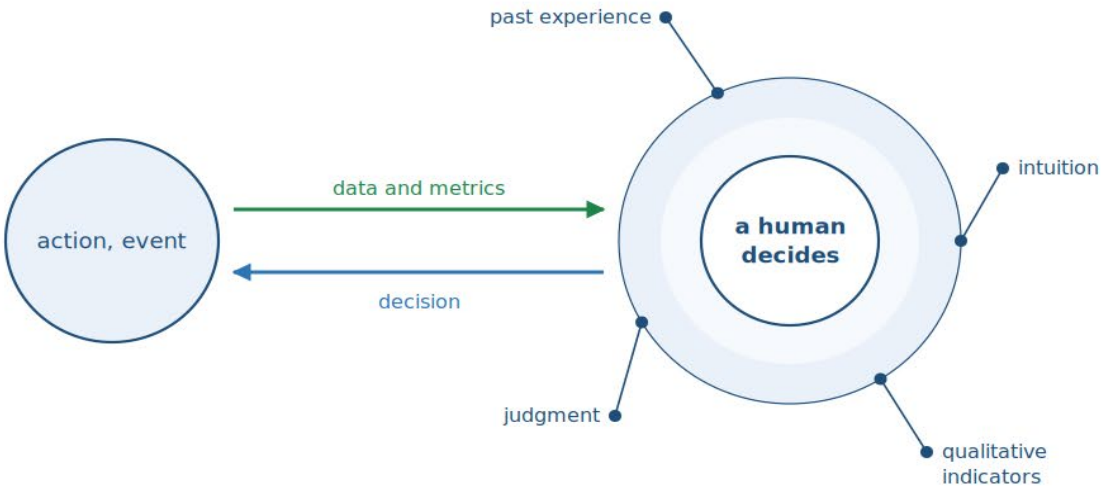
**Data-inspired**—the data inspires. This is strategy: the data suggests patterns and gives rise to hypotheses, but the decision rests on vision and on context that simply is not in the data.

## Data helps; a human decides

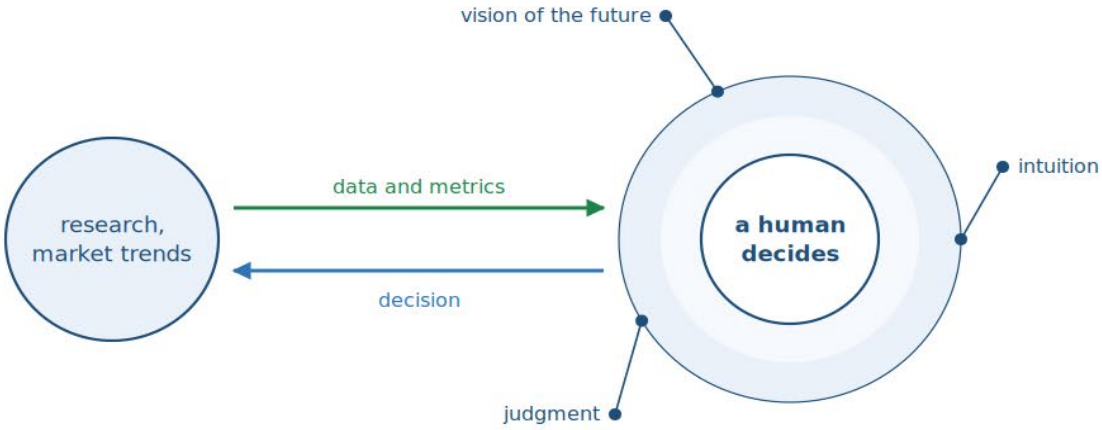
### Data-driven




### Data-informed




### Data-inspired



**And here is the fundamental weakness of every advisor:** the full context cannot be collected. Whatever is unmeasurable never reaches the data—the agreement struck in a corridor, the political landscape, the team's fatigue, the intuition of twenty years in an industry, the signals you do not even register as signals. Data is always about the past, and it is always incomplete. So the higher the level of the decision, moving from operations up to strategy, the less authority the "autopilot" has and the more the human matters. That is the systemic reason for the 70/30 model: not a temporary limitation of the technology, but a property of decision-making itself.

### Who is responsible for the decision?

In meetings with executives, this question comes up before the question of price: who carries the responsibility for the final decision? Every leader wants the option of sharing it—with a colleague, a committee, a consultant. With an advisor that does not work: the AI gave the recommendation, but the signature on the order is yours.

The nature of neural networks makes it harder still. Unlike with rule-based expert systems, their logic is not fully understood even by the people who built them. The model is a black box: it will not show you the chain that led to the recommendation, and the explanations it generates on request are a plausible retelling, not a log of the computation.

Now layer AI agents on top of that—in 2025–2026, entire processes began to be delegated to them. Regulation is already answering the question in its own way: both the European AI Act and Russia's March 2026 draft bill leaned toward the liability of the user and the operator rather than the model, as we saw in Chapter 7. Translated into management terms: you sign, you own it.

And there is a final, strategic layer. Trying to hand decisions over to advisors wholesale makes managerial competence inside the company atrophy: people stop deciding and start trusting the machine blindly. The conclusion is the same as before: the advisor prepares the decision, the human makes it and answers for it. Keep that in your written procedures, not in your good intentions.

**Competence decay, loss of focus, distrust, contradictory advice.**

Competence decays not because people get worse, but because they stop practicing: the machine prepares the decision. The growing flow of information scatters attention. People resist: they do not understand the tool, they fear for their jobs, they remember earlier failures. And different advisors can issue mutually exclusive recommendations—the financial one says cut costs, the digital one says invest.

Rule-based expert DSS are no better: expensive to develop (complex models and digital twins with a margin of error of up to 5%), hardware-hungry (like MES systems), inflexible (every change is a separate project), and they run into the same question of responsibility, plus a dependence on whichever rules were written into them. The conclusion: for all their potential, digital advisors face fundamental technological and organizational challenges—and those will have to be solved before the market goes mainstream.

### What to Do and Where to Head (A 5–20-Year Horizon)

- **Specialization.** Do not wait for the know-everything chatbot: models will move into narrow and mid-range specialization. That lowers hardware requirements, simplifies security (no knowledge of explosives, no answer about explosives), and reduces hallucinations. In a 2023 experiment, Eliza—the narrow 1960s "psychologist" bot—passed the Turing test more often than GPT-3.5, and a model trained on my own books gave far more relevant recommendations than open chatbots did.
- **AI builders (no-code).** The same story as website builders: off-the-shelf advisors you can load a procedure or a book into—lightweight, light on infrastructure, with usable UX, viable on-premises, and affordable for small and midsize business.
- **Domain standards and methodologies.** Fine-tuning on international, national, or proprietary standards: we are building our advisor on PMBOK 7 with extensions. For 90–95% of companies the basic methodologies are enough—discipline and the execution of standard recommendations matter more than tailoring everything to fit you. Recall the Japanese shu-ha-ri: first follow the rule, then break it, then create your own.

- **Industry statistics.** Regulators and governments hold the big data needed to train genuinely good industry-specific DSS. Equipment manufacturers in the EU and the US are already combining digital twins with operating data and moving to a service-based sales model.
- **Systems combined with chatbots.** Business wants systems, not chatbots. The future is ready-made recommendations with the analytics behind them, plus the option to ask the assistant a specific question inside a specific context—the way an AI bot works in tech support or in employee onboarding.
- **Entry through small and midsize business.** On the innovation adoption curve—innovators, early adopters, the majority, laggards—breakthroughs do not come from corporations: their culture demands safety, not risk. SMBs are more flexible and willing to take risks, which makes them the way in, the way to normalize the technology (the Overton window), and the route into corporations afterward. Marketing here works the way it does in B2C: digital channels, plain language, virality.
- **Personalization by psychological type and competencies.** Recommendations that account for the strengths and weaknesses of the leader and of the team—Adizes's PAEI, DISC, and the like, without overcomplicating it. This is the logic I am building into the target model of my own advisor.
- **A systemic approach.** As long as the business is in chaos, no advisor will help. You need an organizational structure, processes that are documented and simplified (lean manufacturing reduces the load on the infrastructure), quality data, and horizontal communication. Part 2 of this book covers this in detail; China's "1397" and "AI+" programs are an example of the systemic approach at national scale.
- **Training, simpler interfaces, orchestrators, feedback.** People without digital skills sabotage the rollout, so you need sandboxes and incentives. Interfaces have to get simpler: according to Standish Group, about 45% of features are never used and another 19% are used very rarely, while high-income users prefer the simplicity of iOS. You need AI orchestrators—the

equivalent of VMware in IT—to break requests down and resolve conflicts between advisors. And you need feedback mechanisms.

**A practical example.** A holding company with subsidiaries running identical processes: procurement, maintenance, sales. Pairing an AI orchestrator with an AI builder for local models gives you a single interface (one window instead of 50 systems), fast and cheap launches of advisors by the business units themselves, easy knowledge updates done by those same units (which takes the load off IT), and higher quality and security through specialization. The more heavily regulated the industry, the bigger the effect. But even here there is no way around the systemic approach: documented and simplified processes, digital and project competencies in the team, and a phased rollout with communication.

> ### The Author's Position
>
> **Move away from open chatbots toward specialized products.** Value does not come from the dialogue; it comes from being embedded in a process—a clear input, a clear output, a defined zone of responsibility, and quality control. General-purpose assistants are becoming infrastructure. The money is in specific scenarios.

## Chapter Summary

AI advisors will trigger the next stage in the evolution of AI in business and become an enormous market, but the expansion will run through small and midsize business rather than corporations: SMBs are open to experiments and willing to take risks.

**The forecast came true—faster than I expected.** Two years on, the confirmation arrived all at once. Yandex reported that 86% of AI agent users in Russia are small and midsize businesses, and Anthropic released Claude for Small Business, an off-the-shelf package for the US SMB segment. The third news item of that spring was a study from Sweden's Sinch, the AI Production Paradox, which found that 74% of companies rolled back AI agents they had already launched in customer support. That is not evidence about who adopts first; it is evidence about the price of an immature rollout—and it explains why

the way in runs through the segment for which mistakes are cheapest. The detailed analysis is in "Postscript 2026: Four Signals of Market Maturity" in the Conclusion.

**Developers** should plan for: specialized advisors; AI builders with optimized models; fine-tuning on domain standards for off-the-shelf products; industry statistics; ready-made recommendations combined with chatbots; designing for the psychological types and competencies of the team; feedback mechanisms; simpler interfaces; orchestration across advisors; and 50–70% of the budget going to marketing on a B2C model.

**Business** should put systemic management in place—strategy, organizational structure, processes, project management, lean manufacturing—train its people and build their competencies, and get ready for off-the-shelf products built on international, proprietary, or internal methodologies.

### Key Takeaways

- Digital advisors have real potential, but they run into problems of data, cost, accountability, and the black box.
- The breakthrough will come through small and midsize business, not corporations; the future is specialization, no-code builders, orchestration, and personalization.
- Without a systemic approach—order in the processes and in the data—even the best advisor is useless.

**What to do:** Do not build or wait for a know-everything advisor. Start with a narrow one, specialized in your own procedures, and put your processes and your data in order before you deploy it.

**Reflection question:** Are your processes and your data ready for an advisor to give you recommendations that actually mean something?

## Chapter 10. Neuromorphic and Quantum AI

In my view, today's AI solutions—the ones built on classical architectures—are approaching their limit; we went through this in the section on strong and superstrong AI. What we are doing now is more a matter of learning to make effective use of what scientists worked out over the past 30–40 years—which

means we are moving onto the plateau of productivity, mastering the technology.

Does that mean progress stops? Of course not. I believe the next leap will come from quantum and neuromorphic computing—and possibly from a combination of the two, although that hybrid belongs to a very distant future. Let us take each in turn.

In the language of the S-curve from Chapter 2: everything described below is a candidate for the next curve. Today these technologies sit on its flat stretch—right where neural networks sat in the 1990s. That is exactly why they are worth watching now: the explosive phase begins without warning, the way it did with ChatGPT.

### Neuromorphic Systems

Every IT solution in use today is built on the simple von Neumann architecture: the processor does the computing separately from memory and reaches into it when it needs data. Neuromorphic systems borrow their principles from the biology of the brain. Computation and memory are merged into a single unit, and those units are assembled into complex networks that do not run continuously but fire in pulses.

What is the point? To solve the problem we discussed in the section on strong AI: today's solutions are complex and consume an enormous amount of energy. To recognize a cat or a dog, the most powerful computer needs a thousand times more energy than our brain does. An ordinary PC constantly shuttles data between processor and memory over the bus: to add two numbers it pulls them out of memory into registers, performs the operation, and sends the result back. The alternative is to build simple computation directly into memory and give it the command "add these two numbers and write the result right here." The trip over the bus disappears, and both time and energy costs drop several times over.

The upshot is that neuromorphic systems are strong wherever you need to collect data continuously—to see, to hear, to smell—analyze it fast, and draw very little power. Three key areas stand out.

- **Autonomous sensors.** They work with no dependence on the cloud or a network. Take an air-toxicity detector worn on your clothing. Built the classical way, it needs either a communications module—without a network it is "dumb"—or a large battery, which limits both operating time and convenience. A neuromorphic design gives it autonomy and low power draw—solving the fundamental problem of smart devices.
- **Shorter time to process data and make decisions** (every AI model in use today "thinks" slowly). For example, Boston Dynamics' robot dogs used to fall over when pushed. Recomputing after a shove means gathering sensor data, running the calculations, and issuing commands to the actuators—and the robot could not do it in time. Another example is vibration diagnostics built on video analytics, where a heavy video stream has to be processed.
- **A growing number of neurons (scaling).** Today's solutions are already hard to scale, and this is exactly the problem neuromorphic computing helps to solve. A practical example is controlling a bicycle: on a single Chinese Tianjic chip, with no other hardware, researchers assembled a system in which one neural network hears and sees, a second keeps the bike balanced, and a master AI orchestrator distributes tasks across the models, gathers the data, and makes fast, composite decisions. Instead of a pile of graphics cards, processors, and a huge battery, you get a small device—and the bicycle rides, steers around obstacles, and responds to commands.

**The downsides.** It is an appealing prospect, but the technology is at an early stage and short on both expertise and resources. Specialists still have the basic questions to settle—mathematical algorithms, languages, architecture—and that takes enormous numbers of mathematicians and engineers. After that comes a harder task still: setting up pilot and then industrial production of millions of chips. Today they are made one at a time for experiments, and even "classical" equipment is produced by only a handful of plants worldwide. This technology will not go mass-market soon.

### Where the technology stands (2024–2026).

Over the past two years neuromorphic computing has moved beyond pure research. Intel is developing the Loihi line: the Loihi 2 chip carries about a million neurons and, on certain AI tasks, delivers up to 100 times the energy efficiency of a GPU. Larger versions have been announced for robotics and for sensor processing directly on the device. IBM has moved its NorthPole architecture—where computation and memory are fused so tightly that the data never leaves the chip—closer to production, with a 2026 target. Published estimates put it several times ahead of classical accelerators on energy efficiency, and it needs no liquid cooling, which simplifies the design and cuts costs. BrainChip is already shipping its Akida neuromorphic chips in volume; in 2025 it opened cloud access (Akida Cloud) and unveiled a new generation of the architecture, which drew interest even from the space industry, where saving energy is critical. In other words, the claims of this chapter—computation inside memory, autonomy, energy efficiency—are already being borne out by commercial products, even though mass adoption is still a long way off.

**I will admit that I follow this race with a fan's excitement.** The pace is picking up. At CES 2026 BrainChip showed production edge devices running on Akida, opened remote access to its ultra-low-power Akida Pico coprocessor, and began licensing the Akida 2 architecture to third-party chipmakers, with smart meters as the first application. IBM has confirmed its production plans for NorthPole: the latest published figures put the chip at up to 25 times the energy efficiency of comparable GPUs on inference tasks.

If the subject interests you, the review articles on neuromorphic computing and neuromorphic systems are easy to find in the open literature.

### Quantum Systems

Classical IT systems have two states, on and off—in the processor, in the reading and writing of data, and in the neurons of an AI model. In quantum computing the "neuron," a qubit, can be on, off, or in both states at once. That is superposition. The difference sounds small, and to the uninitiated it means little, but the consequences are radical—so radical that scientists are still at

the very earliest stage of mastering this technology, younger even than the neuromorphic one.

Quantum processors will be able to work with larger and more complex models and to use new architectures—with neurons in superposition—that are out of reach today. The research is encouraging. At the Freie Universität Berlin, for instance, researchers found that quantum neural networks can not only learn but also memorize random data. “It’s like finding out that a 6-year-old can memorize random strings of numbers and the multiplication tables at the same time… quantum neural networks are incredibly adept at fitting random data and labels, challenging the very foundations of how we understand learning and generalization,” said Elies Gil-Fuster, the study’s lead author.

Put into a list, the advantages are: greater memory capacity per unit of volume at small physical sizes, which raises speed and lowers power draw; fewer neurons for the same tasks; faster training on less data; and solutions to problems that cannot be solved at all today. From another angle: no catastrophic forgetting of context; linearly inseparable problems solved by a single-layer network; higher stability and reliability; and the possibility of “multidimensional” models that process several data streams at once. At this point we are drifting into science fiction—and a combination of quantum and neuromorphic technologies goes beyond imagination altogether.

**The downsides.** Let us come back down to earth and see why the revolution has not happened yet. First, the earliest research appeared only in 1995; by the standards of science the technology is an infant, and there are few specialists. Second, quantum computers themselves are at a very early stage. You cannot buy one for your home or your office; this is roughly where the PC was in the 1970s and 1980s. The problems being solved are the basic ones of physics and mathematics: how many physical and logical qubits a task needs, how long the calculations will take, how to write the algorithms, how to remove the noise that disturbs the qubits. For now this is the domain of scientists and the military. Arrival in large business is 5–20 years away, mass adoption 30–50. Third, quantum computing is a poor fit for everyday tasks: it is like plowing a potato field with a spaceship. A quantum computer is not “a supercomputer

that does everything faster"—one of the key open questions right now is which problems it actually solves faster than a classical machine does.

Quantum computers are strong wherever you have to work through an enormous number of combinations: quantum simulation, cryptography, search. Back in 2013, for instance, D-Wave built a two-color Ramsey number graph in 270 milliseconds—a calculation a conventional mid-range computer would have needed more than 10,000 years to finish. In parallel, quantum ideas are already being applied to compressing AI—making models smaller, lighter, and faster by finding simpler representations of the data, which is useful for running AI on phones. Since real quantum machines are still few and temperamental, the more common route is a hybrid, "quantum-inspired" approach that runs on ordinary computers: it helps to switch off the least important parts of a neural network without much loss of quality and to find a compact model architecture faster, cutting video memory use, operating cost, and latency. As the field progresses, real quantum accelerators will be added to the mix. A current snapshot of the race comes immediately below, in the section "Where the technology stands."

### Where the technology stands (2024–2026).

The main change of recent years is the move from "many qubits" to qubit quality and error correction. In late 2024 Google unveiled Willow, a chip with 105 superconducting qubits, and demonstrated "below-threshold" operation for the first time: as you add qubits the error rate falls rather than rises, which is the long-sought condition for scalable fault-tolerant machines. On a benchmark task Willow finished in minutes a computation that would have taken a classical supercomputer an astronomically long time. IBM ran a quantum error-correction algorithm on ordinary off-the-shelf chips ahead of schedule and published a road map to Starling, a fault-tolerant computer targeted for 2029 with roughly 200 logical qubits and hundreds of millions of corrected operations. Overall, quantum error correction (QEC) became the industry's top priority in 2025: error rates hit record lows, and some groups—QuEra among them—proposed algorithms that cut the overhead of correction several times over, by up to a factor of 100. None of this removes the need for

caution about timelines, but it does show the field moving from laboratory demonstrations to an engineering road map.

**While I was preparing this edition, the news came in faster than I could write it up.** In October 2025 Google demonstrated the Quantum Echoes algorithm on Willow—the first verifiable quantum advantage, roughly 13,000 times faster than the best supercomputers. In November IBM introduced the Nighthawk processor (120 qubits) and error-correction decoding faster than 480 nanoseconds—a year ahead of its own plan. The industry's milestones are holding so far: a demonstration of quantum advantage on practical problems by the end of 2026, and the fault-tolerant Starling in 2029.

**And here a cool head matters:** these successes do not bring quantum computers any closer to your business on a three-to-five-year horizon, and they do not overturn the forecast I gave above. Look at the scale. Starling in 2029 means roughly 200 logical qubits, while problems of practical significance—materials chemistry, cryptography, serious optimization—need thousands of logical qubits, which means millions of physical ones. One fault-tolerant computer, once built, is not the same thing as availability: it means a queue in the cloud, a handful of specialists, and algorithms that, for most business problems, simply do not exist yet. So I stand by the forecast: the first niche applications in large business will come closer to the lower bound of that "5–20 years"—the early 2030s, through cloud access—while mass availability remains a question of decades. The successes of 2024–2026 did not shift the timelines. They shifted the confidence that this road leads anywhere at all.

If this area interests you, I recommend starting with the review articles on quantum computing and quantum machine learning, available via the QR codes and hyperlinks.

*Understanding quantum machine learning also requires rethinking generalization*

*An introduction to quantum computing*

## Chapter Summary

Quantum and neuromorphic networks will probably be the next revolutions in AI, and it makes sense to tie the evolution of strong AI to these technologies in particular. But they are at such an early stage that precise forecasts are impossible: this is the leading edge of science, research aimed at the future, and an area for long-horizon investment. They will not take over applied tasks outright in the next few years, but given what they are capable of, we should expect hybrids that combine quantum-neuromorphic and classical solutions.

The main benefit of these technologies is that AI models will become less demanding of compute and energy while getting "smarter." A reminder: GPT-4 has far more neurons than a human being. It is an enormous model, yet it consumes colossal amounts of energy, requires a staff of expensive specialists,

and is still no "smarter" than a person—even though it was trained on volumes of data no human has ever taken in. Lifting that ceiling—the energy and infrastructure ceiling—is exactly what the neuromorphic and quantum approaches promise.

#### Key Takeaways

- Classical architectures are approaching their limit; the next leap belongs to neuromorphic systems—computation inside memory, energy efficiency—and to quantum ones.
- The trends of 2024–2026 bear this out: neuromorphic chips are entering production (Loihi, NorthPole, Akida), and quantum is shifting from "qubit count" to error correction (Willow and the road map to fault tolerance).
- These technologies will not take over applied tasks outright any time soon; their value is in lifting AI's energy and infrastructure ceiling. Industrial adoption is still a long way off.

**What to do:** Do not wait for a quantum revolution to handle your current tasks. Follow the news on neuromorphic adoption in autonomous devices, and invest in competencies today rather than in hardware.

**Reflection question:** Where in your business do autonomy and energy efficiency matter more than the model's "intelligence"?

# Part 2. AI in the Systems Approach

This part is the simplest and the shortest. In it I will show how AI will affect management tools, and how those tools will affect the development of AI.

## Chapter 11. Strategy and Organizational Structure

Strategy, business objectives, and organizational structure are the core of any company—the skeleton and the compass that shape everything else inside it.

### The Impact of AI

In my view, AI here will develop into a consultant sold by subscription or billed by the job.

Here an AI assistant does the work: you feed it a diagnostic questionnaire; it processes the answers, drafts a strategy template, and defines the key metrics.

Take the structure of a digitalization strategy—the one I laid out in DT-2. The Systems Approach. In short, it should contain the following sections:

- an assessment of the current situation, including a profile of your key competitors and the industry leaders, plus your own organizational and digital maturity
- the target operating model of the organization
- the stages of digitalization and the key tasks
- the portfolio and programs of technology projects
- the portfolio and programs of organizational projects
- the competencies your people will need, and the portfolio of HR and training projects
- the financial and human resources required
- sources of funding
- an assessment of effectiveness

It looks simple enough. But producing a document that is both solid and readable takes a couple of months, and that is more than many organizations can afford. This is where AI earns its keep. It can quickly:

- compile an overview of competitors—their products and their financials—and single out the most promising ones, the players worth benchmarking against
- prepare summaries of strategy sessions
- develop the target operating model and define the key tasks
- generate project proposals and set priorities
- calculate a preliminary budget
- compare current metrics and propose targets

Taken together, this can compress the timeline from one or two months to one or two weeks. Yes, you still cannot do without experts—but an expert working with AI can now get several times more done. Which means the price drops from $12,000–23,000 per strategy to $3,500–6,000. The key constraints here are:

- the time the business needs to sign off
- the ability to put together a balanced combination of AI models
- the usability of the interface
- the soundness of the diagnostic questions used to assess the situation

The same will hold for designing organizational structures. Given the inputs, AI will propose an optimal structure and the roles it requires, along with descriptions of what each role does and the personal and professional competencies each one calls for.

The hardest part here is designing the input and output data models, organizing data collection, training the AI, packaging all of it into a usable interface—and overcoming people's resistance. As I wrote in the section on digital advisors, people will be the biggest constraint of all. We tend to swing between extremes: either we do not trust the machine at all, or we want to hand it full responsibility. Look at what actually causes most problems in digitalization and you will find that almost all the risks and problems come down to the human factor. And this is not just my subjective opinion. My work with students confirms it. In every risk management workshop, the participants arrive at the same conclusion on their own: eight or nine of the top ten risks are organizational.

Back to organizational design. It should rest on the following factors:

- the company's industry and its products or services
- its life cycle stage and the resources available to it
- its degree of independence and any external constraints—being part of a holding group, for example
- how automated the business is

- and the hardest one of all—the specifics of the business itself: a distributed team, the quirks of its logistics, and so on

### The Impact on AI

So what influence will organizational structures and strategies have on the development and adoption of AI?

It is fairly straightforward. AI specialists will take their place in organizational structures, and AI adoption will become one of the strategic objectives, written into development and digitalization strategies.

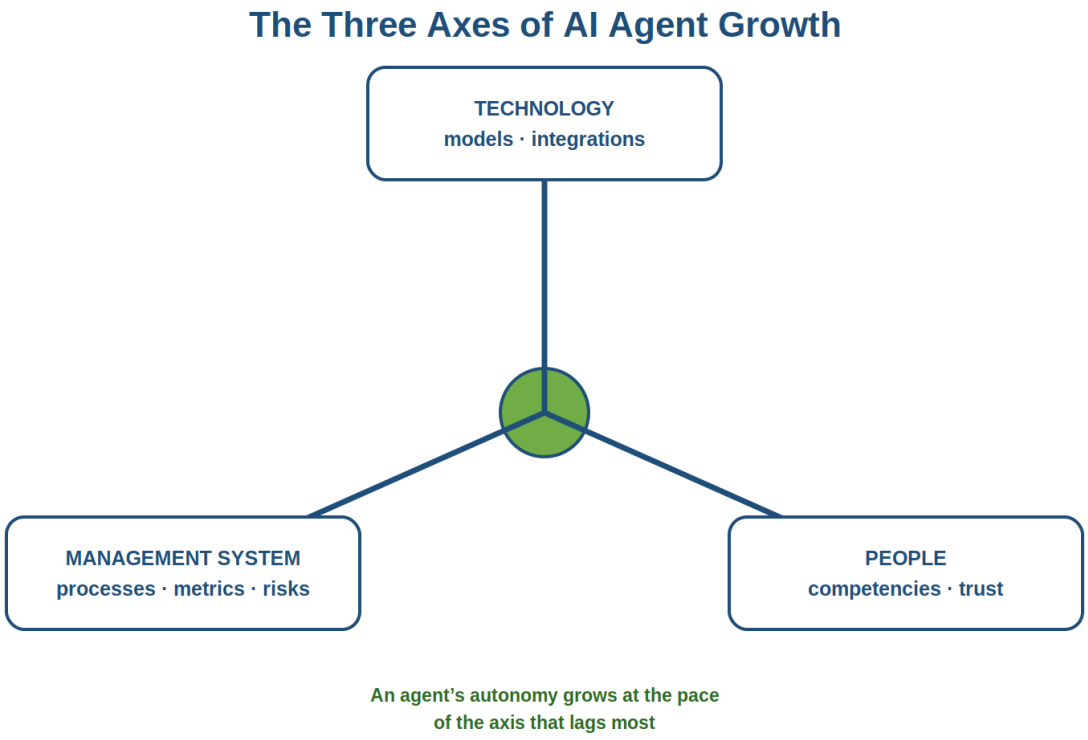
The danger, as always, lies in how this is integrated into the structure. If you decide to lock AI inside a separate unit within the IT department, it is doomed. You will need the technical people, but also IT partners and leaders on the business side. This is what a distributed, or hybrid, function looks like. Otherwise we fall into the familiar trap: the business sits back and waits for wizards who will work out its own processes for it, tell it how things ought to be done, and deliver the perfect solution on the spot.

The business itself will be prodded along by caps on headcount growth in its departments and by mandates to adopt technology. Yes, this will produce projects that are spherical cows in a vacuum, but sadly the inertia of a traditional business cannot be broken any other way.

It is also worth thinking about how the role of AI agents inside the organization will shift over time. Today they work at the level of assistants and copilots, but as the technology, the data, and management processes mature, both the volume and the complexity of what can be delegated to AI will grow. What we are really describing is a gradual expansion of autonomy: from a helper that acts on request to an agent that runs routine processes on its own and escalates only the exceptions.

The pace of that transition, however, is set less by progress in the models than by the readiness of the organization itself. Clear processes, quality data, a functioning risk management system—these are the preconditions for "promoting" an AI agent to more consequential work. Strategic planning should therefore build in progress along three axes at once: technology, the

management system, and people's competencies. How closely those three advance together will determine how deeply AI agents embed themselves in the business over the next two to five years.




## A Practical Example

For a practical example, I want to share an observation from China. On my first visit I crossed the country from north to south and saw five cities: Beijing, Shanghai, Hangzhou, Shenzhen, and Guangzhou. I also talked with a great many people who have lived there for anywhere from five to twenty years, and I spent two weeks living like an ordinary person—no expensive hotels, no minders, none of the other trappings of an official visit. Trying to understand why China became one of the leaders in the AI race, I came away with four insights.

First: mass adoption of AI. Every city has its clusters, accelerators, and hundreds of startups. Alipay already has AI assistants built in, and in the subway you come across ads for AI cloud services.

Second: a bet on practical models. For all the talk of "strong AI," most specialists there are convinced that the future belongs to compact, affordable,

specialized solutions embedded in specific processes and coordinated by other models.

Third: infrastructure is the foundation of everything. Automation, sensors, robotics, fast and stable connectivity, quality data. Together they form the base for deploying AI at scale in industry and in everyday life.

Fourth: scale and speed. China's level of robotization is more than ten times Russia's. This is not technology for its own sake—it is a real gain in efficiency and a platform on which new solutions get built. On top of that, they do not run pilot projects for years: they put an innovation through its paces quickly and then either scale it or drop it.

None of this is possible without an overarching strategy and long-term planning. And here I ran into two vivid examples of strategic thinking.

The first is the "1397" strategy, formulated at Zhejiang Lab, which is based in China AI Town in Hangzhou. So what do those digits—1397—actually stand for?

### "1"—A Single Strategic Goal

The central goal of the whole strategy is to make intelligent computing in the field of AI and machine learning the top priority for achieving leadership and practical results.

### "3"—Three Strategic Needs / Areas of Development

This is what we will use computing to solve.

National strategic frontiers

Solving problems that are strategic for the state and strengthening technological sovereignty.

Scientific and technological innovation and transformation

Driving fundamental scientific discovery and technological change: basic research in AI, cognitive science, and algorithms, and the building of technology platforms (cloud services, supercomputers, data networks).

Innovation in strategic industries

Modernizing and developing key sectors of the economy, and designing and deploying new processes built on what the technology makes possible (robotics, smart cities, medicine, industrial AI).

### “9”—Nine Priority Research Directions

This is how we will meet those strategic needs.

#### *Group 1: Major Transformational / Cross-Cutting Tasks*

*1. Support for major / strategic / research international projects*

**Example:** participation in the ITER project—the International Thermonuclear Experimental Reactor.

**How it works:** intelligent computing is used for extremely complex modeling of plasma processes, for real-time control of the experiment (predicting and suppressing instabilities), for processing exabytes of data from thousands of sensors, and for optimizing the reactor design.

**Result:** a faster path to a commercial fusion reactor—an inexhaustible source of energy.

*2. Support for major / strategic / research national science projects*

**Example:** China’s space program (the Chang’e lunar program or the Tianwen Mars program, for instance).

**How it works:** building a distributed computing network that spans ground control centers, orbiters, and landers. AI processes surface imagery to choose a landing site, handles navigation autonomously, and analyzes scientific data—soil composition, for example—right where it is gathered.

**Result:** greater mission autonomy and higher success rates under long communication delays.

#### *Group 2: “Driver” Tasks, or the Driving Force*

*3. Applying scientific and technological innovation in key domains*

**Example:** the discovery of new drugs and materials.

**How it works:** using AI to predict the properties of millions of molecules (virtual screening), which speeds up the search for drug candidates against cancer or Alzheimer's by a factor of thousands. In the same way, AI models new materials with specified properties—room-temperature superconductors, for example, or new batteries.

**Result:** development time and cost cut from decades to a few years.

*4. Helping frontier manufacturing adopt intelligent solutions*

**Example:** the "dark factory," or smart factory.

**How it works:** in a fully automated electronics plant (Huawei's or Xiaomi's, say), intelligent computing optimizes the entire chain in real time: machine vision systems check product quality, robotic arms adapt to changes on the line, and predictive analytics flags machines that need servicing before they fail.

**Result:** a sharp gain in efficiency, quality, and manufacturing flexibility.

*Group 3: Infrastructure Tasks*

*5. Building high-performance infrastructure for intelligent computing*

**Example:** a national computing center specializing in AI.

**How it works:** building supercomputer clusters equipped with thousands of specialized processors (NPUs, or neural processing units) optimized specifically for training large neural networks such as GPT and models like it. This infrastructure is offered as a cloud service to scientists and companies across the country.

**Result:** a lower barrier to entry in AI and faster development of large models.

*6. Building high-performance distributed computing systems*

**Example:** pooling the computing resources of supercomputing centers in different cities.

**How it works:** creating a single software platform that lets a researcher in Beijing launch a calculation drawing simultaneously on supercomputers in Shenzhen, Shanghai, and Wuxi, as if they were one large machine.

**Result:** solving problems that no single machine, however powerful, can handle (modeling the planet's climate at high resolution, for example).

*7. Building a hub for the open exchange of scientific data*

**Example:** a national scientific data archive, along the lines of GenBank or arXiv.org but broader.

**How it works:** creating a centralized yet distributed platform where universities and research institutes can publish de-identified data from a range of fields: genomes, astronomical observations, results of physics experiments, climate records. AI helps catalog and clean the data and find connections between datasets.

**Result:** preventing data silos and encouraging interdisciplinary research and the reuse of data.

*8. Building a space-based distributed computing system*

**Example:** creating an "orbital cloud."

**How it works:** placing server modules on the satellites of a constellation so that data can be processed in orbit. A remote sensing satellite, for instance, can analyze its imagery on the spot for wildfires or natural disasters and send down finished coordinates and an alert rather than terabytes of raw images.

**Result:** lower latency, less load on communication channels, and faster response.

### *Group 4: Foundational Tasks*

*9. Frontier basic research*

**Example:** developing new neural network architectures and training algorithms.

**How it works:** funding and supporting research into what is called artificial general intelligence (AGI), into energy-efficient algorithms and neuromorphic computing (which imitates the workings of the brain), and into the possibilities of quantum machine learning.

**Result:** laying the groundwork for the next breakthroughs in AI, the ones that will define the technology landscape ten to twenty years from now.

### "7"—Seven Innovation Mechanisms

These are the organizational and management methods that will be used to deliver the nine priority directions.

*1. Research organized around key tasks*

**The mechanism:** an approach in which research teams and infrastructure are assembled not around permanent administrative units but around one specific large and significant goal (one of the nine priority directions, for example). Once the goal is reached, the structure can be reorganized around new challenges.

**Example:** for the task of building a space-based distributed computing system, a temporary consortium is formed. It brings together engineers from an aerospace center, AI specialists from a university, developers from an IT company, and security experts from an academic institute. This project office runs until the system is launched and successfully tested in orbit, after which the team can be redeployed to other projects.

*2. A research management mechanism led by a principal investigator*

**The mechanism:** a research project is run not by an administrator but by a recognized scientist—the principal investigator. They have wide autonomy in research decisions, budget allocation, and team building, while carrying full responsibility for the result.

**Example:** a major project to develop a new neural processor (NPU) architecture is led by a top computer scientist. They personally pick their deputies for each area—hardware, algorithms, software—approve the research plan, and decide which experiments to fund, reporting directly to the institute's leadership.

*3. A talent development mechanism that lets young scientists take responsibility*

**The mechanism:** creating the conditions for rapid career growth and independence for young, talented researchers. They are put in charge of promising groups or of risky projects with high potential.

**Example:** a thirty-year-old PhD who has proposed a revolutionary method in quantum machine learning is given the chance to head an independent research group with its own budget and laboratory. They get a free hand to recruit the team and choose specific lines of work within the overall strategy, and they are spared the administrative overload.

*4. An evaluation and incentive system geared to scientific value*

**The mechanism:** dropping the practice of judging scientists by publication count alone. In its place comes a broader system that accounts for real contributions to science and to technological sovereignty: patents, technology transferred to industry, leadership of major scientific or strategic projects, and the solving of specific hard problems.

**Example:** a scientist whose work on optimizing algorithms for weather modeling was adopted by the national forecasting center and significantly improved forecast accuracy receives a substantial cash award and priority for further funding, even if colleagues have published more papers.

*5. An innovation integration mechanism built on self-reliance and open cooperation*

**The mechanism:** a "build it yourself, but work with the best" strategy. Key technologies are developed inside the country to secure independence and security, while partnerships with international research groups, companies, and universities are pursued actively, to exchange ideas and tackle global problems together.

**Example:** in building out the Global Basic Observing Network, China develops and launches its own satellites and deploys its own ground stations (self-reliance). At the same time, the country takes an active part in World Meteorological Organization (WMO) initiatives such as the Unified Data Policy, and shares data with other countries to improve global climate models.

*6. A mechanism for turning research results into products, embedded in the innovation ecosystem*

**The mechanism:** building short, efficient paths from the lab to the market. That means incubators, startup studios attached to institutes, piloting programs with industrial companies, and regulatory sandboxes for testing new technologies.

**Example:** an algorithm for predictive maintenance of machine tools developed at a university is tested immediately at the plants of partner companies (in the automotive industry, say). On the strength of successful trials, a small innovation company is founded to bring the product to market, funded by a venture fund attached to that same university.

*7. Cultivating a corporate culture grounded in the spirit of the scientist*

**The mechanism:** deliberately instilling in the research community the values of service to the country, scientific rigor, bold inquiry, and a readiness to take risks. Popularizing success stories and the acceptance of failure as an inherent part of the innovation process.

**Example:** regular internal forums where leading scientists share not only their successes but also their valuable failures and the lessons drawn from them. Awards established not just for major discoveries but for ambitious research that came to nothing yet showed originality and high risk. Inspiration drawn from the historical figures who shaped science.

Together, these seven mechanisms form an integrated system meant to make research more flexible, more results-oriented, and more attractive to talent.

A strategy like this sets priorities and coordinates the development of AI as a whole. That makes it possible to pursue even the projects that at first glance create no value and prompt the question "why bother?" In the end, it all adds up to a synergistic effect.

The second example is a document published by China's State Council on August 26, 2025: the "Opinions on Deepening the Implementation of the 'Artificial Intelligence Plus' Initiative." The "AI Plus" plan, in plainer words. And it is not one more declaration of intent but a detailed road map for

transforming the entire Chinese economy and society around artificial intelligence. It sets ambitious goals for the coming decade and lays the foundation for turning China into the global leader of the "smart economy" by 2035.

## The Philosophy of "AI Plus": From Digitalization to Intelligentization

The plan is an evolutionary leap beyond the "Internet Plus" concept that has dominated Chinese policy for the past ten years. Where "Internet Plus" focused on connectivity and digitization, "AI Plus" aims to embed artificial intelligence deeply in every sphere of life—from scientific research to daily routine.

The document's central idea is the emergence of a new socioeconomic order—a "smart economy and smart society"—built on three principles: human-machine collaboration, cross-industry integration, and value co-creation. The first of the three is the load-bearing one: human-machine collaboration is what marks the shift away from a model in which technology adapts to the human being, toward one of symbiosis, in which AI becomes a natural extension of human capability.

## A Three-Stage Road Map: 2027, 2030, and 2035

The state plan is clearly structured around time horizons, which makes it possible to track progress and correct course.

Stage 1: By 2027—Foundation and Acceleration

The first milestone is about building a critical mass of AI adoption. By 2027 the plan calls for:

- more than 70% penetration of AI terminals and AI agents among the population and businesses
- deep integration of AI into six key areas (science and technology, industry, consumption, the social sphere, governance, international cooperation)
- rapid growth in the core industries of the smart economy
- a significantly stronger role for AI in government administration

This stage locks in the fundamentals and clears the bottlenecks. In two years China intends to build the infrastructure and the ecosystem that will let it move to mass deployment of AI solutions.

Stage 2: By 2030—Scale and Leadership

The medium-term horizon envisages exponential growth:

- penetration of intelligent solutions above 90%
- the smart economy as the most important engine of the country's economic growth
- technological universalization—access to advanced AI technologies for every layer of society
- a full ecosystem for sharing the results of AI development

By 2030 China expects to have built a transformation mechanism that runs "from pressure to opportunity, from opportunity to strength," letting it compete with the United States on equal terms in individual directions and gain an advantage in particular niches of the global AI market.

Stage 3: By 2035—A New Stage of Development

The long-term goal is as ambitious as it gets: full entry into the era of the smart economy and the smart society. By that point AI is meant to be so deeply woven into the fabric of Chinese society that it provides "strong support for the basic realization of socialist modernization."

That phrasing means artificial intelligence is treated not merely as a technology but as an instrument for reaching national strategic goals—from raising labor productivity to strengthening China's international position.

## The Six Pillars of the "AI Plus" Strategy

The document defines six key directions for AI adoption, each with its own concrete tasks and delivery mechanisms.

1. "AI + Science and Technology"

This direction is about using AI to accelerate scientific discovery and technological breakthroughs. The key measures include:

- developing foundation models: improving the theoretical basis, raising the efficiency of training and inference, building a system for evaluating models
- applying AI to scientific research: bringing philosophical and social science methodologies into “scientific intelligence,” using AI to model complex processes and to discover new materials and drugs
- creating AI accelerators for science: platforms that automate routine research work and let scientists concentrate on the creative part

For the first time in Chinese policy, scientific intelligence (AI4Science) has been designated a priority area, which reflects an understanding of how far AI can transform fundamental science.

2. “AI + Industrial Development”

Here the goal is comprehensive intelligentization across all three sectors of the economy.

Industry:

- deploying AI at every stage—from design and pilot production through operation and maintenance
- building “smart factories” with autonomous production lines
- developing the industrial Internet of Things underpinned by AI analytics

Agriculture:

- mass rollout of smart farm machinery—robots, drones, automated harvesters
- precision farming systems with AI yield forecasting and resource optimization
- digital transformation of the whole production chain, from sowing to logistics

Services:

- a shift from “digital enhancement” to “intelligent management”

- personalized AI assistants for education, healthcare, and finance
- new business models built on AI as a service (AIaaS)

China stresses that all three sectors have to transform in step, so that they reinforce one another.

3. “AI + Raising the Quality of Consumption”

This part of the plan focuses on creating new consumer products and experiences.

- next-generation intelligent terminals: smartphones, cars, and household appliances with built-in AI and multimodal interaction
- AI agents: software entities able to understand context, plan actions, and carry out users’ tasks on their own
- new formats of consumption: smart homes, autonomous transport, personalized entertainment and cultural services

The document underlines the importance of creating “a better and more beautiful life” through AI, which means not only convenience but an emotional bond between users and intelligent systems.

4. “AI + Social Well-Being”

One of the most socially oriented sections of the plan puts AI to work on public problems.

Education:

- personalized learning systems that adapt content to each student’s ability
- VR/AR technologies for immersive learning
- AI assistants for teachers that automate grading and materials preparation

Healthcare:

- AI diagnostic systems that support doctors in remote regions
- smart medical robots for patient care

- telemedicine platforms with AI consultants
- faster drug development through AI modeling

Employment and careers:

- new jobs in AI development, data management, and ethical auditing
- platforms for retraining and learning new professions
- AI simulators for hands-on training in demanding occupations

Elderly care:

- smart devices for health monitoring
- companion robots for social interaction
- emergency response systems with AI analytics

The core principle is equal access to the benefits of AI for every social group, including the most vulnerable.

5. “AI + Governance Capacity”

This section is about transforming public administration through AI.

- smart cities: integrated systems for managing transport, energy, and security
- predictive analytics: AI systems for anticipating social problems and optimizing government decisions
- digital public services: automated processing and approval of citizen applications, intelligent advisors for citizens
- public safety: early warning systems for emergencies, smart management of resources during disasters

China is aiming to create “a government that thinks”—one able to make decisions from data, not only from the experience of its officials.

6. “AI + Global Cooperation”

For the first time in a Chinese strategy, the international dimension of AI has been carved out as a separate direction.

- support for creating AI governance mechanisms within the UN: China actively backs the formation of an Independent International Scientific Panel on AI and a Global Dialogue on AI Governance
- an "AI Plus international cooperation" initiative: helping developing countries build AI infrastructure, and sharing data and models
- a principle of inclusiveness: unlike the United States, China emphasizes that all countries need to take part in setting the rules of the game in AI, so that the digital divide does not widen

This position reflects China's pursuit of soft power in AI—building an alternative ecosystem attractive to the countries of the Global South.

### The Eight Foundations of the System

To deliver the six directions, the document defines eight foundations for infrastructure and the ecosystem.

1. Foundation models and algorithms

- advancing frontier machine learning theory
- building efficient systems for training models and running inference
- establishing a standardized system for evaluating models

2. High-quality data

- creating industry datasets for training AI
- reforming copyright and intellectual property rules to encourage data sharing
- introducing new mechanisms to encourage the opening of data

China acknowledges that on sheer volume of data it is not behind the United States—the weak link is elsewhere, in how accessible and usable that data is for training. The US has a mature open data ecosystem—roughly 300,000 datasets on government platforms—while in China more than 70% of valuable data sits in the hands of the state and is poorly structured for training AI.

3. Computing power

- developing chips and clusters for AI computing
- improving the national computing network
- standardizing cloud services
- ensuring an inclusive, efficient, green, and secure supply of computing resources

The “East Data, West Computing” strategy aims to optimize the geography of data centers—siting energy-hungry facilities in the western regions, where green energy is abundant.

4. The application ecosystem

- developing AI-as-a-service platforms
- creating testbeds for new solutions
- setting new standards for AI applications

5. The open source ecosystem

- supporting open source developer communities
- counting open source contributions when students and researchers are evaluated
- encouraging new approaches to application development
- building ecosystems with global influence

Recognizing the role of open source at the level of state policy is a strong signal of China’s intent to build a more open and inclusive AI ecosystem.

6. People and education

- training AI specialists at scale
- building AI literacy into school and university curricula
- attracting international talent

7. Legal and ethical frameworks

- drafting legislation on the safe development of AI

- creating ethical standards
- establishing accountability mechanisms for what AI systems do

China already has its Ethical Norms for New Generation Artificial Intelligence (2021) and its Interim Measures for the Management of Generative Artificial Intelligence Services (2023), which puts it at the leading edge of regulation in this field.

8. Policy and financial support

- tax breaks for AI companies
- state funding for promising development projects
- a favorable business environment for startups

### China's Unique Advantages

The document rests on three competitive advantages that set China apart from the United States and Europe.

1. The wealth and scale of data

China generates more than 41 zettabytes of data a year (2024), 25% more than the year before. That is the largest volume of any country. The key factors:

- an enormous population (1.4 billion people)
- a high degree of digitalization in everyday life
- the largest internet platforms (Alibaba, Tencent, ByteDance)

The problem, though, is not volume but structure. The share of structured data in China is growing fast (36% a year), but it is still only 18.7% of the total.

2. A complete industrial system

China is the only country in the world that has all 41 industrial categories in the UN classification. That creates a uniquely wide range of scenarios for applying AI—from traditional metallurgy to electric-vehicle manufacturing.

A complete industrial system means Chinese AI developers have access to real industrial data and can test their solutions under conditions that competitors in other countries simply cannot reach.

3. The breadth and depth of adoption

In China, 83% of people believe AI will bring more benefit than harm; in the United States the figure is just 39%. That level of trust and willingness to adopt creates an ideal environment for testing and iterating AI solutions at scale.

China also leads in the share of companies actively using AI. The public sector is active too: a large volume of calls to AI services runs through the national data exchange platform.

## Possible Challenges and Risks

For all the plan's ambition, experts point to a number of potential difficulties.

1. The data quality problem

Although China holds enormous volumes of data, a large share of it sits with the state and is not structured well enough for training AI. The gap between English- and Chinese-language open source datasets remains critical: Chinese datasets amount to just 11% of the English-language ones.

2. Technological dependence

US export sanctions on advanced chips create serious obstacles. Chinese companies (Huawei Ascend, Moore Threads) are building alternatives, but these have not yet reached the level of NVIDIA's A100 and H100.

3. The balance between innovation and control

Tight regulation of content and algorithms could slow innovation. China needs to find the balance between ensuring security and encouraging experimentation.

4. International isolation

As technological competition with the West sharpens, access to leading-edge research and to talent may narrow.

## What This Means for Us

China's plan for developing AI through 2035 is not simply a strategy document; it is a manifesto for a new model of modernization. Unlike the classic Western approach, with its focus on market competition and

breakthrough innovation by individual companies, the Chinese model bets on state coordination, mass application, and social purpose.

The defining features of the Chinese approach:

- synchronized transformation—changing every sector of the economy at once rather than making isolated breakthroughs
- social inclusiveness—an emphasis on making AI accessible to every layer of the population, including the elderly and vulnerable groups
- global responsibility—positioning China as the leader of inclusive AI governance through UN mechanisms
- practicality—mass application takes priority over fundamental breakthroughs (the DeepSeek strategy showed that results can be achieved at lower cost)

By 2035 China expects not only to catch up but in certain respects to overtake the West, having built its own ecosystem of AI technologies, standards, and applications. Whether the plan succeeds depends on China's ability to clear its current bottlenecks—in chips, in fundamental research, and in international cooperation. But one thing is obvious: the battle for the AI future will shape the geopolitical landscape of the coming decades, and China is playing the long game.

### How the AI Plus Plan and the 1397 Strategy Connect

Worth noting: these two documents—the AI Plus plan and the 1397 strategy—do not contradict each other. Let's look at that a little more closely.

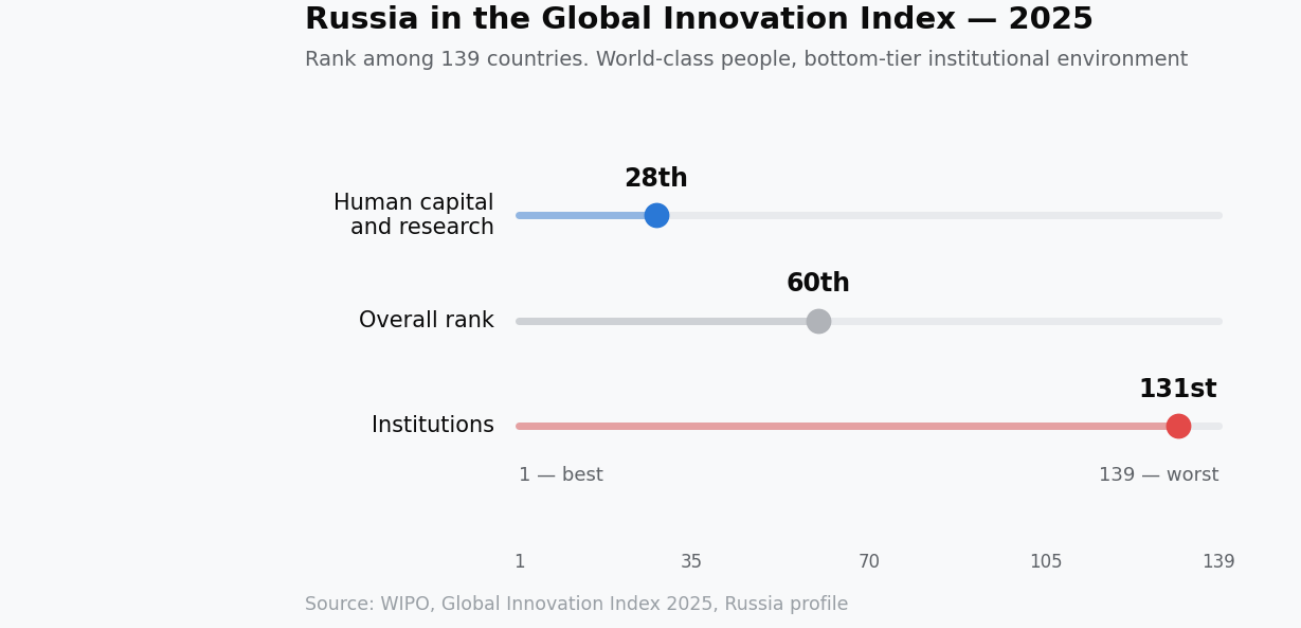
"AI Plus" is China's national strategy through 2035: it sets the what and the why through six domains of application and eight systemic foundations for transforming the economy, society, and public administration. "1397" is the operational framework that answers how and with what: a focus on intelligent computing, data infrastructure, foundation models, and the organizational mechanisms that turn the goals of "AI Plus" into deliverable programs. In other words, they are two layers of one construction: "AI Plus" defines the

strategic outcomes, while “1397” supplies the technological and organizational means of delivering them at the level of programs and projects.

## What About Russia: The Improviser’s Profile

We have taken the Chinese system apart in detail. An honest question: what does Russia’s own profile look like? The answer matters beyond Russia—it applies to every culture built on improvisation rather than process. Numbers first, then a personal position; that position can and should be argued with.

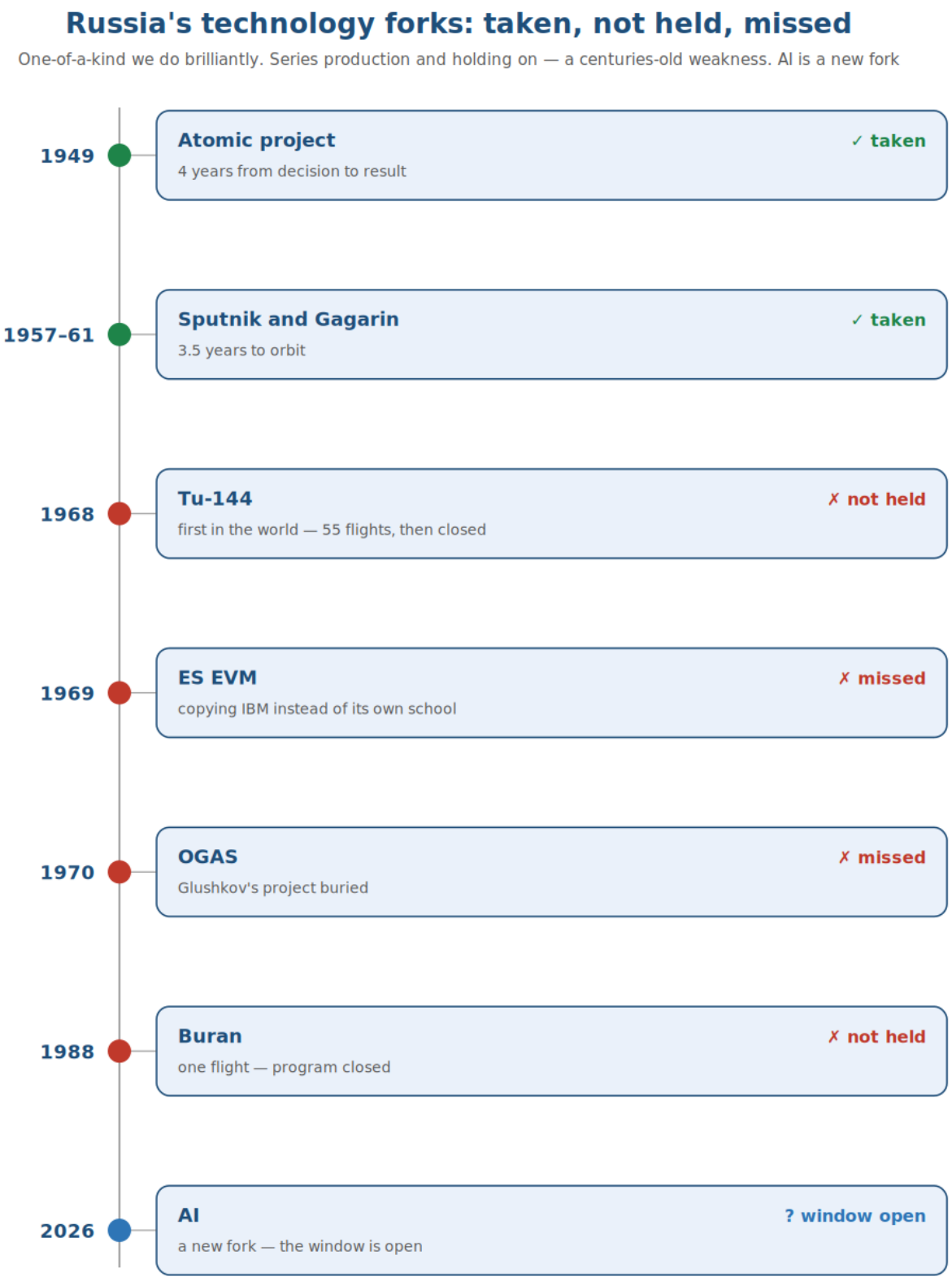
Look at the Global Innovation Index (GII 2025). In human capital and research, Russia ranks 28th in the world. For institutions and environment—131st out of 139. The overall result: 60th place. A hundred positions between people and systems—a gap unique among large economies. World-class talent working in an environment that cannot convert it.




History repeats the profile. Alexander Popov demonstrated radio—others capitalized on it. Vladimir Vapnik built the theory behind machine learning—it runs in other people’s products. Russia has long made a national sport of raising geniuses for other economies: ideas are born and not held on to—no follow-through, no packaging, no scale.

The country has faced a fork like this twice before. The atomic project and space were pulled off: there was systemic management, personal accountability, and hard deadlines. The Tu-144 flew first and was not

sustained; the Buran shuttle flew once. OGAS, the 1970 project of a national computer network, was buried in interdepartmental correspondence. China drove its analogue all the way through—we took that apart above. Now the third fork is opening: AI.



Now the position. The East Asian economic miracle grew out of a culture of execution: discipline, method, flawless reproduction of a procedure. For decades that model beat the cultures of improvisation—the Russian *smekalka* type, with its brilliant workarounds and weak follow-through. But generative AI sells a perfect performer to everyone. A trump card anyone can buy a copy of gets cheaper. And the singular—nonstandard moves, breakthroughs where the manual says "impossible"—gets more expensive. The improviser's strength was always singular.

The main thing: for the first time in history, the improviser's weakness is closed by a technology. System, method, follow-through, packaging—exactly what AI does best. Every previous wave of technology demanded order as a precondition—which is why classic automation and ERP took root so painfully in improvisation cultures. Generative AI is the first technology that helps create order. The formula: AI closes what the improviser lacks—and multiplies what the improviser has.

What follows from this. For the state: a grounded ten-year program on the level of AI Plus—pushing AI down into industries, personal accountability, measurable targets. For companies: a bet on the pairing of "talent + AI as systematizer" rather than on drilling process discipline into heroes. And keeping the order you create in systems, not in heroes: order that lives only in people leaves with them.

The window will not stay open for long: the low-base effect works only until others raise the base. The full argument, with sources and cases, is in my article "AI Is the Missing Half of the Improviser's Culture"—via the QR code and the link.

*AI Is the Missing Half of the Improviser's Culture*

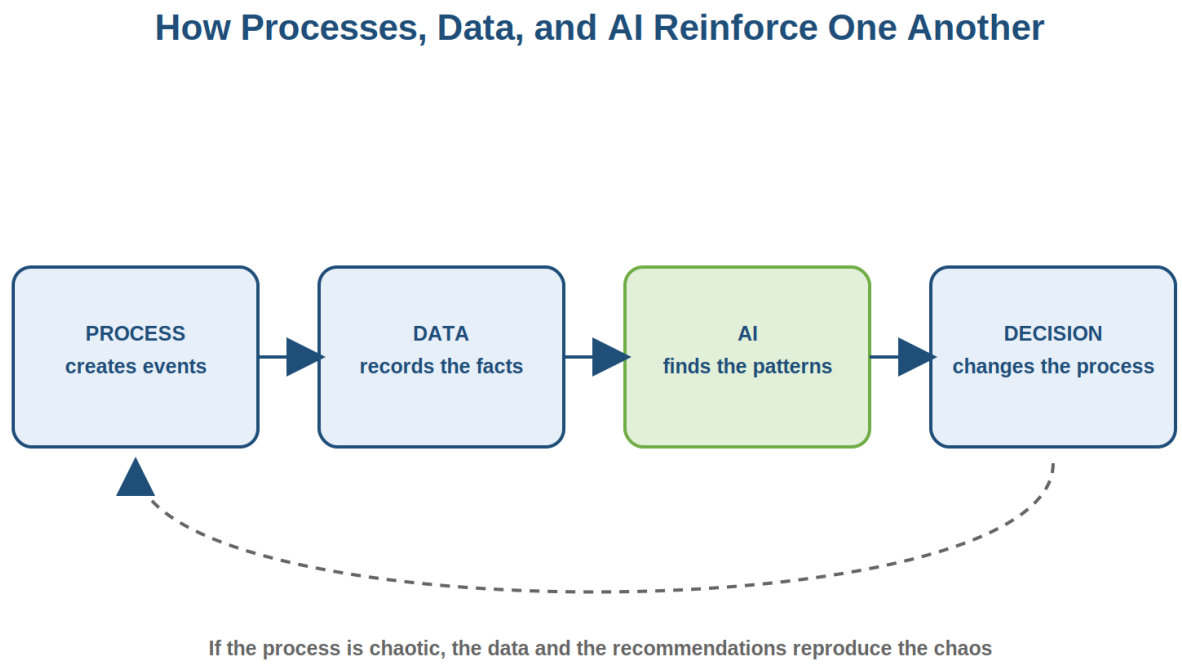
## Chapter Summary

- Strategy is the compass, and organizational structure is the company's skeleton. AI plays two roles here: it is a tool for developing strategy and structure (timelines and costs drop several-fold), and it is their object—adopting AI itself becomes a strategic task, and someone has to manage it.
- The AI function has to be hybrid, or distributed: technical specialists plus IT partners plus leaders on the business side. A separate "AI department inside IT" is the road to failure—the business will sit and wait for wizards.
- The role of AI agents will expand from assisting to acting on their own, but the pace of that shift is set by the maturity of processes, data, and management, not by progress in the models.
- China's experience illustrates the power of the combination: long-term strategy plus infrastructure plus mass practical adoption plus fast experimentation instead of pilots that drag on for years.

**What to do:** Check whether AI appears in your strategy as a separate workstream with goals and metrics (one week). When you set up the AI function, design for a distributed model from the start and develop along three axes—technology, the management system, and people's competencies (one quarter).

**Reflection question:** If you were writing a "1397 strategy" for your own company, what would your "1"—your single strategic goal—be?

## Chapter 12. Business Processes




Business processes, on a par with communication, are the organization's central nervous system. How they are built determines how everything else is coordinated. And this is an area where AI will come into its own.

### The Impact of AI

The main direction for AI in business process work is identifying processes, optimizing them, and creating them from scratch.

Working from input data about the company—industry, strategy, and so on—AI will compile an inventory of the processes the company needs across its key areas, produce a basic description of each in various notations, and draft the procedures and work instructions that go with them. In essence, this is a reflection of the ideas behind quality management systems. The hardest part here is building synergy with the design of the organizational structure and deciding what comes first. My own view is that it should be a single end-to-end process. Based on the industry, the specifics of the business, the resources available, and the strategy, AI will shape the organizational structure, then

define the key business processes underneath it, and draft the procedures. Or the other way around: describe the stable processes first and build the organizational structure to fit them. There is no single right approach here; it is a perennial debate.

This sounds utopian, but I think this direction will be running at full strength within 10–20 years.

As for identifying processes, this is where process mapping will begin—based on how users work in IT systems, with recommendations prepared on top of that. And this is probably the most "frightening" direction. Building an organization from scratch is one thing; optimizing what already exists is another. Resistance from people is inevitable, because it will turn out that a third of the company is not needed at all.

On the technical side, the hardest part is teaching AI to build processes from scratch, assembling a library of processes from the leaders in the field, and learning to model optimal processes. And especially building synergy with lean manufacturing, so that AI understands what waste is and does not take it to extremes. We looked at the requirements for AI solutions—specification, robustness, and assurance—in Chapter 8: the system has to do what we actually want, not what is literally written into its objective function.

Again, coming back to the knowledge base AI will have to be trained on to learn best practice: the design of business processes is the main trade secret any company has. If companies are reluctant to share even ordinary data, collecting descriptions of their processes will be a problem of an entirely different order. That is why I see this as a very difficult direction. Yes, systems will appear—both proprietary ones and ones built on international standards—but adoption will be slow and painful. The first to arrive will be advisors for designing the organizational structure.

That said, AI can already be used today to build business process inventories. Here is a practical example: building an inventory of processes by line of activity.

In this example we look at the processes needed by a notionally mature manufacturing company that sells B2B.

Artificial intelligence was used in part in preparing this section. The production function, HR management, and IT support are described in detail down to level 3; the remaining functions are given in outline, at the level of their level-2 processes.

The inventory covers all the core processes.

### Production (Operation, Maintenance, and Repair)

**Maintenance planning.** Setting the frequency and scope of maintenance work; drawing up the maintenance schedule; assigning responsibility for the work.

**Performing maintenance.** Running scheduled inspections and checks of equipment; replacing worn parts and consumables; setting up and adjusting equipment.

**Recording and analyzing faults.** Collecting information on equipment faults; analyzing what causes them; developing measures to prevent them from recurring.

**Organizing repair work.** Planning and agreeing on the timing and scope of repairs; preparing the necessary tools and materials; carrying out repairs in line with the plan.

**Quality control of repair work.** Checking completed work against quality standards; correcting any shortcomings found; documenting the work performed.

**Record keeping and reporting.** Logging all work performed and resources used; preparing reports on the technical condition of equipment; submitting those reports to management.

**Training and upskilling staff.** Organizing training for the people responsible for operating and repairing equipment; assessing their level of knowledge and skills; arranging further training where needed.

### HR Management and Development

**Workforce planning.** Reviewing the current staffing situation; forecasting headcount needs against the company's strategic goals; drawing up plans for attracting, onboarding, training, and developing people.

**Recruitment and onboarding.** Sourcing candidates for open positions; interviewing and assessing them; handling employment paperwork; onboarding new hires into the corporate culture and work processes.

**Training and development.** Organizing training and upskilling; assessing the level of knowledge and skills; arranging further training where needed.

**Employee evaluation and formal appraisal.** Regular reviews of employee performance; formal appraisal to determine fitness for the role; identifying each employee's strengths and weaknesses.

**Motivation and incentives.** Designing a performance-based incentive system; using both financial and non-financial methods; keeping the system fair and transparent.

**Managing conflict and stress.** Preventing conflict and stressful situations within the team; resolving the conflicts that do arise and helping people cope with stress.

**HR administration.** Maintaining HR records and documentation; complying with labor law and internal company rules.

**Building the corporate culture.** Shaping and sustaining the corporate culture; strengthening team spirit and employee loyalty.

**IT Support (Excluding Digitalization and Digital Development)**

**Planning and managing the IT infrastructure.** Analyzing the company's IT resource needs; developing an IT infrastructure strategy; managing changes to IT systems.

**Data security.** Protecting information against unauthorized access; encrypting data in transit and at rest; keeping antivirus software up to date.

**User support.** Training and advising staff on the use of IT systems; resolving the technical problems users run into.

**Access management.** Configuring access rights for different user categories; monitoring compliance with the rules on the use of resources.

**System monitoring and analysis.** Collecting and analyzing data on IT system performance; identifying and resolving problems in system operation.

**Developing and rolling out new solutions.** Designing and building new IT products and services; testing and deploying what has been developed.

**Developing and improving IT processes.** Continuously monitoring how well IT processes perform; adopting new technologies and optimization methods; adjusting plans in light of the results.

**The remaining functions in outline, at level 2 only.**

**Industrial safety.** Risk analysis; ensuring compliance with legislation and regulations; staff training and safety briefings; monitoring the condition of equipment and infrastructure; use of personal protective equipment; emergency response and first aid; monitoring and reporting; working with government authorities; improving the safety system.

**Sales.** Identifying potential customers; establishing contact with the customer; presenting the product or service; handling objections; closing the deal; after-sales service; analyzing results.

**Marketing.** Market and competitor analysis; segmenting the target audience; brand positioning; developing the marketing strategy; content creation; content promotion; measuring the performance of marketing campaigns; working with partners and influencers; reputation management.

**Accounting.** Processing source documents; payroll processing; tax and financial accounting; accounting for cash and bank transactions; taking inventory of assets and liabilities; preparing financial and tax statements; safekeeping of documents.

**Finance, treasury, and financial planning.** Cash management; financial planning and budgeting; raising funding; optimizing the capital structure; risk management; monitoring payment execution; analyzing financial results.

**Logistics.** Planning and forecasting; inventory management; arranging transportation; warehouse handling and storage; clearing goods through customs; quality control of logistics services; optimizing logistics processes.

**Legal.** Analyzing legislation and legal requirements; drafting legal documents; representing the company in court and before government authorities; resolving legal disputes; monitoring compliance with legislation; advising company management and staff; safekeeping of documents.

**Physical security.** Risk and threat analysis; ensuring compliance with legislation and regulations; access control at entry points; guarding the perimeter and the site; maintaining fire safety; maintaining counter-terrorism protection; monitoring the condition of security systems and equipment; working with government authorities; improving the security system.

**Facilities and administrative services.** Managing property and material resources; setting up the workspace; health and safety at work; maintaining building systems and utilities; providing company transport; cleaning and upkeep of premises; quality control of facilities and administrative services.

Compiling this inventory with AI took me about two hours. First came questions to establish which functions an organization of a given profile performs, then follow-up questions to pin down the processes that go with those functions.

The example also shows where AI typically falls short. Procurement and supply—an entire function—is missing from the inventory. Sales is described not as processes but as the stages of a sales script. Staff training is duplicated between production and HR. This is the standard error profile of a language model on a task like this: it reproduces what appears often in open sources and stays silent about what does not.

### The Impact on AI

The other side of the question is how process culture affects the adoption and development of AI. Here things are fairly simple. The companies that will gain the most—and become the drivers—are the ones that have maps and descriptions of their processes. In that case it will be easier to run the analysis

and determine where AI solutions should go and what their areas of responsibility are. From the chapter on digital advisors we already know that the narrower their specialty, the better the quality of their work, the fewer the errors, and the lower the cost of the solution. Exactly as with people. Accordingly, it is these companies that will influence the whole industry and the quality of such solutions.

The fundamentals of process work—notations, modeling, maturity assessment—are covered in DT-2. The Systems Approach (Chapter 1). Here we look only at where processes meet AI.

However, an enormous plus here will be the same one that runs through all of digitalization and automation: centralizing individual functions within corporate holding groups. Putting AI into strategic objectives will force companies to work on their processes and unify them, and that alone will pay off and clear out the chaos. Only after that does AI get deployed and tasks get automated. And it makes no difference whether it is a simple digital advisor on procurement procedures or HR, or a predictive analytics system for equipment maintenance.

### Chapter Summary

- Business processes are the organization's "nervous system." AI works in three directions here: identification (mapping from the traces users leave in IT systems), optimization, and designing processes from scratch.
- AI already speeds up the routine part of process work—compiling process inventories and procedures. Full automation of process design is a 10–20 year horizon.
- The reverse effect is stronger than the direct one: the clearer and more stable your processes, the faster and cheaper AI is to adopt. Chaos cannot be automated—it can only be scaled.

**What to do:** Pick one process with a clear input and a clear output and use AI to draft its description and its procedure (one week). Assess the maturity of your key processes before you plan AI projects on top of them (one quarter).

**Reflection question:** Which of your processes today depends so heavily on one person that their departure would bring the work to a halt—and what is stopping you from documenting it?

## Chapter 13. Managing Projects, Change, and Products

I combined these three areas into one for a reason. Almost every project involves change. That is why I always tell people in my training sessions: when you plan a project, identify the processes that will change or be created. Then build a plan for training people and for overcoming their resistance.

The same goes for launching new products. You cannot do it without a project methodology, because building a product is always a set of separate projects that have to be tied together. And the products themselves can be ordinary ones, where you need marketing and a unique selling proposition, or innovative ones—digital advisors, for instance. Those are products people find hard to accept, and marketing with a USP is not enough. At the start you have to produce useful, explanatory content and tell people why they cannot go on living the way they used to.

Step back a little, and managing change, projects, and products with AI turns out to be one of the areas that is already developing. By some estimates, as early as this decade up to 40% of projects will be manageable with AI.

There are two sides to this question:

- How will AI let us implement change, deliver projects, and launch products more efficiently, at lower cost and with less resistance from people?
- How do we run an AI rollout—which is itself a radical change and an innovative service—at lower cost and with less resistance?

For the methodological groundwork behind all this, I recommend the articles behind the QR codes and links below.

*Implementing changes*

Project management. Part 1

*Product management. Part 1*

## The Impact of AI

AI—or even an orchestra of different AI models—will be the core of digital advisors. Here is what to account for when you design a system like that:

- the methodology built into it
- immersion in context—the project, the organization, and the leader's personal traits
- collecting feedback on the results of delivery, so the model can be fine-tuned and deep learning applied (that is, access to the subjective assessments of the project participants and to objective budget and schedule data from accounting records and systems)

From there, advisors can be split into several groups by functionality:

- chatbots or advisors that simply have the methodology built in, give recommendations, and help you learn the methodology
- advisors that either sit inside a comprehensive solution or have access to other systems themselves, prepare personalized recommendations, and flag deviations (including when one of the participants is "dropping out" of the project context, expectations are drifting off course, or discontent is building)

It will work by writing activities for implementing change and overcoming resistance into the project delivery or product launch plan: what to do, when, how long it will take, and what resources it requires—followed by a report from the specialists on the activities completed.

In product management, it will also become possible to quickly generate hypotheses, one-page product sites, marketing copy, and illustrations, to prepare investor decks, and so on. All of this will cut the cost and the time of launching new products, and there will be more startups.

The key direction here is specializing these models by type of project and product, so that they give concrete recommendations rather than abstract ones. A project to build a house and a project to roll out an ERP will call for different activities. The same is true of new product launches.

That means a fairly large body of methodological work will be required from experts: defining the model's input criteria, classifying types of change, training the models, and so on. But the prospects here are very large. In effect, you can cover the entire small and midsize business segment—a market worth billions of dollars. I think it will all converge on microservice solutions that, based on an input request (the specifics of the business and the industry, the leader's personality, the country's constraints), will connect the necessary knowledge bases and produce recommendations. It sounds simple, but building it is anything but trivial. Which is why, in the mid-2020s, microservice solutions like this still do not exist.

For now, AI can be used to work through projects, products, and change initiatives. Say you have a template for a charter or a delivery plan. AI, even in the form of a chatbot, will help you speed up the work on a project and define:

- goals and effects
- key risks and constraints
- which processes will have to be created and/or changed
- who might have a stake in the project
- what resources and competencies are required
- the project's key tasks and milestones

In other words, AI helps get a stalled project moving and speeds up work on it. But the main obstacle is learning to frame the questions and to judge whether the answers are correct. That is how, in two weeks, I was able to work through a project to launch a leadership school run by in-house trainers. By the end I already understood clearly what had to be done and how it should work. In general, working through the first version of a project vision and a delivery plan now takes one or two days. And shaping the final vision of what the project will produce is one of the most important tasks in its early stage.

And if you can frame requests and check data for correctness, you build your own competencies considerably faster. On top of that, you can do a better job

of delegating tasks and monitoring how your subordinates carry them out, and you can reduce the risks in your own area or business.

Here is another practical example. A friend of mine decided, for fun, to have AI run the numbers on a project, and fed it the inputs. On the very first pass, AI ran the numbers and produced a completion-date forecast that was three days off from what actually happened. He could not get over it: in real life, two organizations had spent more than six months calculating and agreeing on all of this, and here was the answer in 20 seconds.

### The Impact on AI

There are quite a few factors that will influence how AI develops from the standpoint of project and product methodologies and of change management. I will try to work through some of them.

- The deployment model

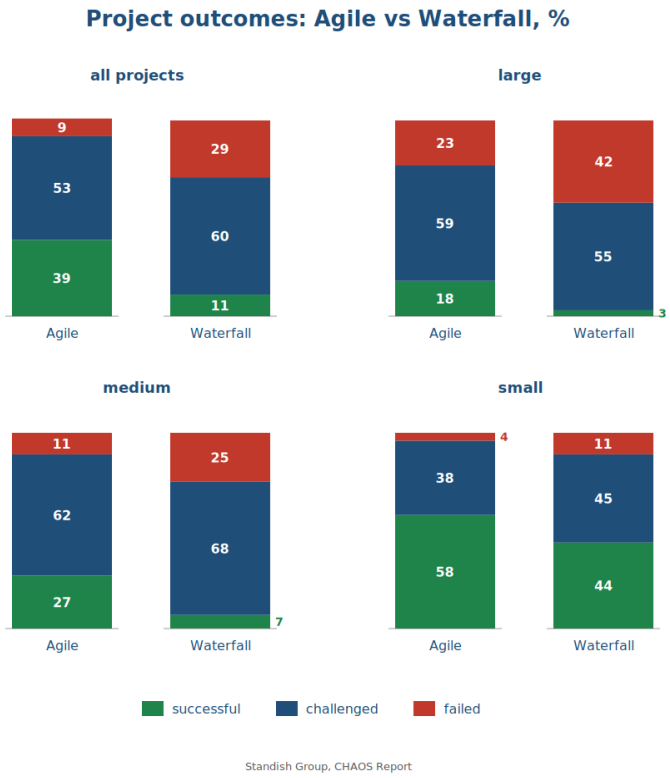
It is important to work out which model organizations will choose. If deployment inside their own data centers becomes a hard requirement, training the models will take an enormous amount of time. If we go the route of cloud solutions and the use of de-identified data, AI models will eventually emerge that can be rolled out as off-the-shelf products.

Small and midsize businesses will be the first to use solutions like these, because they are ready for the cloud—there is less bias there and less desire to keep everything in-house. You can see this confirmed in OpenAI's own strategy: it did not go down the road of B2B sales to corporations. Its target audience is individuals who are willing to try things. That is how it seeks out early adopters and advances within the Overton window.

- Readiness for off-the-shelf products and small projects

We also worked through this topic in the chapter on digital advisors (Chapter 9). Large companies want comprehensive solutions right away, tuned to processes that often do not exist, or exist only on paper. What we get in the end is large, risky projects and, as a consequence, rising costs, slipping schedules, and disillusionment with the technology. This is one more factor in favor of small and midsize businesses being the first to skim the cream. Here

are the statistics on project management by scale and by management approach (waterfall or Agile).




*Project management statistics. The larger the project we want and the more directive the management, the lower the odds of success*

- Change management and overcoming resistance.

Adopting AI is itself a radical change, and people will resist it, up to and including sabotage. Which means you have to understand the reasons for resistance—technical, cultural, and political—and use the tools for implementing change of this kind (Moore's innovation adoption curve and the Overton window).

## The 7 Deadly Sins of Digitalization (AI Included)

Identifying risks is one of the key tasks at the start of digital projects. The topic itself is a very big one. But two things help me here: lean manufacturing and my own experience.

How lean manufacturing helps surface risks is a topic for Chapter 17; my own experience we can go through right now. Or more precisely, my view of the common mistakes that cause 90% of the problems in projects.

*Cause #1. There is no big-picture understanding of what the technology is and what it is used for*

Do you think companies understand why they need to adopt AI and what to expect from it? Is there a correct understanding of at least the difference between automation, digitalization, and digital transformation? That is, what are we going to adopt AI for? Not because it is fashionable or because we have to, but because there is a benefit and clearly stated goals? In 70–80% of cases, and possibly more often than that, the answer is no.

What does this lead to?

First: goals set wrong, expectations distorted.

The clearest example is the attempt to cut whole departments and units in the hope of replacing them with AI. That means we are shaping the wrong expectations among stakeholders at the very start of the project. And since, as we already understand, AI will not work miracles, someone will later have to answer for the resources invested. It could even be the owner, who will lose the business.

Second: the wrong team.

For example, a digital transformation team should include the roles of CDTO (overall coordinator and leader), CDO (head of data—data quality and analytics), CTO (head of processes and process redesign), CA (chief architect, who will tie all the IT systems together and work through the infrastructure), plus legal and security specialists, and so on.

It is the same with AI. Yes, AI projects are part of the overall digitalization strategy, but even a standalone AI project needs a project manager, a representative from the business, a business analyst (to shape the requirements), and a developer. And that is the minimum.

Third: inefficient use of “expensive” technology.

This is usually a consequence of the first two. You can haul potatoes in a Ferrari, but is it sensible? The outcome is always the same: disappointment, losses, and giving up on development too early.

Recommendations

- Bring the leadership into the change (the CEO, CEO-1, the owners).
- Train people and build the team that will drive the change.
- Develop a comprehensive strategy.
- Pilot the technology with an assessment of the effects, and work the initiatives through.

*Cause #2. IT lacks the necessary soft and hard skills*

Tell me, what is your IT director accountable for? What are the key performance indicators for them and for their people? Which competencies dominate when you hire? And most importantly, what does the company think of the IT department as a whole? What do people say about it, and what do they say among themselves?

Classic IT builds reliable systems. What matters to a CIO is that the IT systems run stably, that budgets, processes, and procedures are respected, and that risks go down and costs are optimized. It is not their job to think about making the business more profitable, to focus on the internal customer, or to build solutions that are easy to use.

There are two ways to solve the problem:

- Assign the AI development function—like the whole digital transformation function—to a separate leader who will coordinate, promote, and work at the intersection of the business and IT, and who will have a team of their own. And be sure to bring in people from the business side.
- Teach and develop your own CIO; instill a new culture—customer-centric, oriented toward growing the business, data-driven management, and building solutions that are easy to use.

*Cause #3. A weak culture of working with processes and data*

Questions of whether processes are documented, whether data has been inventoried, and whether its quality is assured are as central to AI as they are to digitalization and automation.

Without documented processes, you will quickly run into the problem of understanding what exactly needs to be automated and digitalized. Where can AI be useful and deliver the greatest effect? How do you prioritize projects? Whom do you sell the idea to? Rolling it out everywhere at once will not work.

Without quality data, and without knowing what data you have, you will fairly quickly hit the problem that AI cannot be trained—or, more precisely, cannot be used. Say you loaded the tax code into an advisor, and then the code went out of date. Its knowledge base is no longer current. How will you use its recommendations? The same goes for production data. So there is no need to rush into AI. This task should become part of the overall strategy, and it will deliver the greatest effect after data collection has been automated.

And how many companies recognize the role of processes and data?

Most companies ignore these questions in the early stages. The consequence is simple: automation and digitalization projects (and AI projects belong to that category) cannot even start. And if they do start, they stretch to three or four times the baseline plan and the budget is exceeded many times over. I have watched it happen.

One more fact has to be noted here. Identifying the stakeholders is a mandatory task in any project. But how do you identify them if there are no processes? It is a different matter, of course, if you have a transparent organizational structure with the functions and the target output spelled out. In practice, that is rare.

Let me give a practical example.

Suppose you hired a brilliant developer and decided to build your own IT system with AI to generate recommendations. The developer built something great, and you started collecting information. But six months in, you will get to the question of making decisions on the data in the system—and you will not be able to, because it will suddenly turn out that the data is dirty and full

of duplicates, and no decision can be based on it. AI will not fill in the blanks for you. You will have to either delete everything and start collecting again, or find people to put it all in order by hand.

Similar problems will arise if you do not work on your business processes.

The result: your developer will burn out and quit, and you will be left with a "black box" that does not work.

*Cause #4. Solutions that are hard to use*

Do you think today's AI solutions are convenient and intuitive for users and administrators? That they do not force people into unnecessary actions and motions? That they integrate easily with one another?

Unfortunately, by the mid-2020s only a very small number of AI developers recognize this problem. Most of them build chatbots and other products for highly skilled specialists.

So the recommendations here are simple.

- Play with free services wherever you can. That is how you learn to understand what you need, how to write a spec, and what requirements to build into it. And a good spec is 50% of success.
- Collect feedback from users and administrators regularly about how USABLE the system is and where it can be improved or extended. And preferably not only through surveys, but by gathering informal feedback too. This applies to the developers of the solutions and to the customers alike.
- Send your people to conferences. Learn from other people's experience, study new products. There is a view that you should refresh your entire software stack every five years.
- Aim for comprehensive solutions in which AI is an internal tool, and framing direct requests to it is only an additional feature.

*Cause #5. Middle and technical management lacks competence in managing projects, products, and change, and in working with staff resistance*

Many companies have project management procedures and standards, and project offices are being set up. But how many leaders on the ground understand what projects, products, and change management are?

Ask them a few simple questions:

- What is PMBOK, or Agile? What is a sprint? Why does it last one to four weeks, and what for?
- What are the five stages, or process groups, that make up a project as a whole and each of its phases?
- When is the waterfall approach the better one, and when is Agile?
- What is the PDCA cycle?
- How does a project differ from a product?
- What is the ultimate goal of their particular project?
- How do people resist change? How do you work with that?
- What share of employees will welcome the change and can be relied on? And how many will stay in opposition to the last?

I imagine you can picture the answers you will get.

If these questions stumped you as well, start with Chapter 4 of DT-2. The Systems Approach: it covers the standards, the lifecycle, and the risks of project management from scratch.

This is one of the reasons so many projects run into trouble: leaders do not know how to organize the work inside a project, how to work through staff resistance, why you would build an unfinished product and test it on people, or why you would follow the plan-do-check-act cycle.

What we end up with is chaos, communication problems, no shared understanding of the ultimate goal, and sometimes sabotage.

To improve your odds, you do not need to train every employee, run certifications, or write procedures—in fact it is harmful. It does NOT WORK, and it is expensive.

What you need is to build training programs that are simple and accessible to the average employee, and to grow a culture of working on change, new products, and projects. You need to virtualize skills—that is, move them out of one person's head and into a form anyone can reach for—and give people an understanding of the basic tools and how they differ.

The help you prepare should not be procedures but quick-reference guides your people can turn to and refresh their knowledge fast.

Recommendations

- Build a project culture and set up a project office, including working templates and tools for project work (there is a sample project management procedure with a step-by-step workflow at the end of this book).
- Hold regular meetings at the CEO and CEO-1 level.
- Work projects through at the initiation stage, including the change-implementation questions (which processes we are changing, who is involved in them, why staff will resist, and what we will do to overcome that resistance).

*Cause #6. Staff lack basic digital literacy, and digital skills have not been virtualized*

Who will work in the new systems—whose work is it that we are changing? What level of digital skills do those people have?

I had a project rolling out an electronic asset management system at a site where the operating staff simply could not use a keyboard. So the first thing we had to do was teach people to type, so that logging a defect did not take three hours and come out full of typos.

Another example. A shop-floor supervisor—27 years old, by the way—did not know how to build a table in Excel.

And what do you think the most common support ticket is when any IT system is rolled out? 80–90% of tickets fall into the "I forgot my password" and "I can't register" categories.

This problem is in some ways similar to resistance to change for technical reasons. But it differs in that people often hold up the whole process without meaning to—not because they do not want to, or do not know, but because they do not know how.

Will a person be able to work out the rules for writing prompts or, say, the logic behind the recommendations?

There is only one solution: training, knowledge bases, instructions.

In my experience, only a handful of people are ready to learn on their own, outside work and at their own expense. We are certainly not breaking new ground here.

Digital technology demands a new level of knowledge and competence both from the people who will implement it and from the people who will use it. And believe me, even among the younger generation there are not that many who can meet all the requirements of working with new tools in a new reality.

Large-scale programs are needed for training and retraining, for building motivation, and for overcoming resistance to innovation.

*Cause #7. The company's culture*

And so we arrive at the question of culture. If everyone in the company relates to everyone else from a position of "you owe me," if you are waiting for a genius to arrive who will bring happiness and do everyone's work for them, then I have bad news for you: in a culture like that, even the most talented digitalization lead will burn out in three or four months. And instead of a customer-oriented specialist you will get the classic IT guy who does not want to see anyone.

Adopting AI at minimal cost is possible only if the departments—the internal customers—are engaged, interested in the change, and understand the basics of project management.

In general, to choose the right behavioral model you first have to identify the type of organizational culture in the company and in its individual departments. Below is a short table of recommendations.


*Most industrial companies have a power culture. The main task there is to win the support of the top executive*

### *Overall recommendations*

The risks in AI projects (apart from the AI-specific ones, which account for 10–20%) are no different from the typical risks of any IT project. I worked through

this question in more detail in the article behind the QR code and the link below.

*7 reasons for most of the problems on the path of digitalization and digital transformation*

And here I will bring all the preceding recommendations together in one table.

| Cause | Risk | Preventive measure |
|---|---|---|
| No big-picture understanding of what the technology is and what goals it serves | The wrong focus, compressed timelines, inefficient use of the technology, no resources and no team | Training and an information campaign for everyone in the company, developing a strategy |
| IT lacks the necessary soft and hard skills | A focus on the technology, on cheap deployment and lower IT support costs, rather than on the effects for the business | Developing and hiring specialists with the required competencies, separating the innovation function out of the IT organization, training existing staff in new skills |
| A weak culture of working with data and processes | Inefficient automation, inability to analyze data or to unlock the effects of AI, missed deadlines and budget overruns | Training in working with data and strengthening the analytics culture in the company, putting process management in place, automating data collection |
| IT systems that are hard to use | Staff resistance, excessive demands on staff | Involving users and administrators in interface design, collecting feedback on |

| | | |
|---|---|---|
| | | usability, designing comprehensive solutions |
| Lack of competence in project and change management | Poorly developed initiatives and inefficient projects, chaotic rollouts and missed deadlines, budget overruns | Training management staff in modern project and change management methods |
| Staff lack basic digital literacy | Risk of misusing digital tools and systems, staff resistance and sabotage, difficulty adapting | Training employees in basic digital skills, creating guides and instructions for working with digital systems (scenario-based instructions) |
| The company's culture | The wrong behavioral model, inability to implement change, unrealistic expectations | Choosing a behavioral model and a change-implementation model that fit, focusing the communication, ruling out expectations that cannot be met |

## The "7 Sins" Checklist: Test Your Own Project

1. Do the top executives share a single understanding of why the company needs AI and what to expect from it?

2. Who is personally accountable for AI development—and does IT have the competencies and the motivation it needs?

3. Are the processes documented and the data inventoried in the project's area?

4. Is the solution easy to use—and when did you last ask the user?

5. Does the team have a basic command of project, product, and change methodology?

6. Do people have enough digital literacy to work in the new system?

7. Does the company's culture support cooperation—or is it "every person for themselves"?

If the answer is "no" to at least two of these questions, close those gaps first and scale AI only afterward.

## Chapter Summary

Unfortunately, the key factors right now are not working in favor of AI developing and spreading quickly. Companies want to keep everything in-house and to get a comprehensive, personalized solution right away, but they do not want to manage change, train people, or build a culture that accepts change. And the digital literacy of ordinary employees, if it grows at all over the years, grows extremely slowly.

But the big players understand this. Look once more at OpenAI and their ChatGPT: they are betting on an accessible solution, add-on stores, and subscription sales. Their strategy is long-term, and they want to become a second Microsoft with a Windows of its own. That is, to become the generic term everyone uses and the standard platform on which various AI solutions will be assembled. And to do that you have to become a tool everyone is used to, so that you can get into corporations over people's heads instead of tilting at the windmills of corporate rules. Do you need approval today to buy a computer with Windows or Linux? Or is that already the standard?

Also, if you build AI solutions, you need to move away from chatbots and into full systems. And I recommend paying attention to machine vision projects and to the analysis of industrial Internet of Things data.

### Key Takeaways

- Projects, change, and products are one discipline: every project carries change, and a product is a stream of project decisions. AI takes over the routine—documents, plans, reports—but it does not replace thinking, or the need to negotiate, resolve conflicts, and so on.
- The "7 Deadly Sins of Digitalization" are the main filter in this chapter: most failures are managerial, not technological. For more on this, see DT-1. Immersion.
- In projects, an AI assistant works as a "single project file": statuses, risks, decisions, and communication in one loop.

**What to do:** Run your current project through the "7 Sins" list (one week). Introduce AI-written status and meeting summaries for one project (one quarter).

**Reflection question:** Which of the "7 Sins" is your organization committing right now—and who in it has the right to say so out loud?

# Chapter 14. Communication

## The Impact of AI

Communication is one of the areas where AI will bring the biggest changes.

It will let us process large volumes of incoming information—among other things, by producing short summaries of calls, meetings, emails, and so on.

For example, we hold an hour-long meeting with a lively discussion and make decisions at the end. Instead of sitting down afterward to replay the recording and write up the outcomes—all of which can take four or five hours—we will hand it over to AI, which will transcribe everything and produce a template for the minutes in 15–20 minutes. All it needs is a good-quality recording.

Every major developer of video-conferencing solutions is already rolling this functionality out. For example, I am already making full use of this capability myself: I transcribe all my meetings and send the transcripts out as the minutes. There was one interesting case among them. During a project we hired an analyst to document what the business was asking for and formulate the requirements for IT. But when it came time to renew the contract, we went back through the recordings of the internal weekly meetings he had taken part in and found something curious. It turned out that in the course of his work he was talking to the same people about the same issues as another analyst, who was working on building the processes as a whole. In other words, recording and transcribing meetings let us spot duplicated work.

AI will also prepare templates and draft replies to incoming communication (emails, messages, and so on).

The first area here is creating information spaces that will bring together all the communication around, say, a single project. This is necessary so that AI can immerse itself in the context and prepare relevant replies.

The second area to expect here is digital twins, or assistants, for every individual. The main question is how to collect and aggregate work information and integrate it into a person's digital twin, so that AI can prepare the right replies given the context of the work task and our personal traits. And most importantly, what will this digital twin consist of, what indicators will it be built from—put simply, what do we need to digitize?

This tool will be especially effective in the B2C segment, for marketers who have to simplify complex ideas and thoughts and get them across to an audience in plain language. Human experts love to slip into academic language, and as a result their target audience does not understand them. And experiments show that AI does an excellent job of producing simple answers.

One more issue you need to pay attention to: we may get so carried away with this tool that 90% of communication ends up happening between robots. Here is an example from my own life. I have a digital robot assistant that answers calls. And sometimes a robot from some company calls me, my robot answering machine picks up, the two of them talk to each other, and afterward I read the short transcript. Why is this a trap? Because laziness drives progress: we will start delegating more and more communication to tools like this, which will lead to a large number of distortions. On top of that, there are security risks.

So as not to leave this as a bare assertion, here is an example of how I use AI to prepare presentations, talks, and articles.

- Shaping the idea

In 90% of cases I work out on my own what the talk should be about and which points I want to get across.

- Preparing a prototype

Next comes the prototype. I decide what I want to say on each slide and in what order the slides should run. Yes, of course I have tried doing this with AI,

and sometimes it has even worked out reasonably well, but in most cases I bring AI in only to work through individual slides.

- Illustration

Then comes the illustration stage. Here I hand the materials to my designer, who generates the illustrations we need by prompting AI. He understands how to interact with the models properly. At the same time, the designer makes all the main design decisions alone—not only about the graphics, but also about how to get my points across to the listener or viewer. AI is used only for individual elements, cutting the time the designer spends on manual work.

- Writing the article

After that comes the article based on the talk. The presentation is fed into AI; I also write out the main points, the important nuances, and the constraints (a length limit for the article, for instance), and AI generates the material. That material is 50–70% ready. Then it gets polished and edited.

- Preparing the announcement

And finally, preparing the announcement. The article and the presentation are fed into AI, which is asked to prepare an announcement for specific platforms and social networks.

The result is a time saving of roughly 50–80%. Of course, you could automate everything and save a full 99% of the time. But the product would be thoroughly mediocre. So it stays a human-plus-robot combination either way. That way it is both fast and stylish, and the author's personality and point of view survive—and that is interesting to listen to.

**While this book was moving from its second edition to its third, this recommendation has managed to go out of date—in a good way.** AI meeting summaries have gone from a "nice-to-have" to the norm: automatic notes and action-item lists are already built into most corporate video-conferencing platforms (though still not all of them). The competitive advantage has shifted from the feature itself to the discipline of using it: a standard for recording decisions and integrating summaries into task trackers.

## The Impact on AI

As I often say, communication is a key component of success in any undertaking or project. Exactly the same can be said about the development of AI and its spread through our lives. How well experts and technical specialists will be able to convey the benefits of AI to people in accessible, simple terms and remove their fears of it will directly affect how fast the technology takes hold in our lives.

The main task here is choosing the channels and formats of communication. When the goal is popularization, you must not retreat into formality, complicated wording, and academic discussions. The fear of saying something wrong, the desire to be absolutely correct, is one of the key mistakes in launching new products and promoting your services. After all, what we get are descriptions a non-specialist can only make sense of while drunk. In other words, we create the communication barriers ourselves.

So the main techniques are:

- learn to convey ideas in simple language the layperson can follow
- put the focus on illustrations and charts
- work with people's fears of losing their jobs and put the emphasis on their emotions

In short, find evangelists, popularizers, and talented marketers, and let the technical specialists and scientists get on with solving applied problems.

## Chapter Summary

- Communication is the area where AI has the greatest effect: summaries of meetings, emails, and calls save 50–80% of the time spent processing information.
- The rule: AI prepares the draft and the digest—the human makes the decisions and is responsible for the relationships. Chasing "99% automation" in communication means losing both trust and the quality of decisions.

- The reverse impact: the structure and discipline of your communication (your own or the company's) directly determine how well AI works on your data.

**What to do:** Turn on AI summaries for one type of recurring meeting (one week). Introduce a standard for recording "decision—owner—deadline" (one quarter).

**Reflection question:** How many hours a week does your team spend retelling what has already been said?

## Chapter 15. Digital Technologies, Working with Data, and Cybersecurity

Here we will look at how AI pairs with the most widely used technologies and IT solutions, and with how we work with data. We will not spend time on cybersecurity for one reason only: we already went into it in enough depth in the first part of the book.

### Digital Technologies

I have gone through each of these technologies in detail—the pros, the cons, and my own opinion—in DT-1. Immersion (Chapter 2). Here I am keeping only the essentials: how each of them gets along with AI.

You can take a closer look at the technologies in this section via the QR code and the link below.

Digital Technology Review Part 1

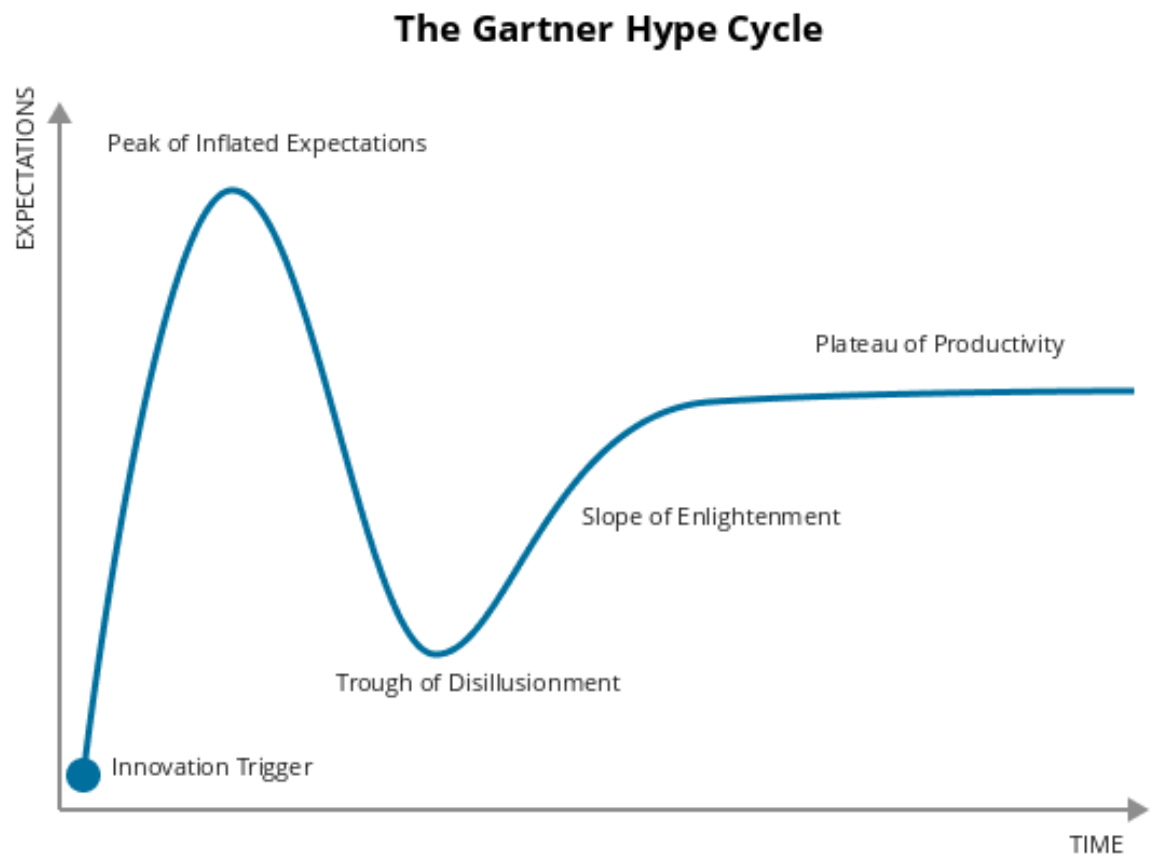
### The Internet of Things and wireless connectivity

**What it is:** a network of sensors and devices that collect and transmit data on their own—from the vibration of a machine tool to the temperature on the shop floor; wireless networks (5G/6G) are their "nervous system." For a business, it delivers an objective picture of what is happening, with no manual data entry.

Without progress in these technologies, AI solutions cannot advance fully either.

First and foremost, AI means predictive systems and digital advisors. And training them—and building digital twins—requires clean big data. Which raises the question: where do we get that clean big data—data free of the human errors that creep in when a hypothetical Joe accidentally types 100.00 instead of 10.00?

That is why the development of Internet of Things technologies for collecting and transmitting data, 5G and 6G included, is the main factor in AI's further progress. Otherwise we will be left with nothing but language models, and the technology is unlikely to climb out of the trough of disillusionment.




*The Gartner Hype Cycle*

## Big data

**What it is:** technologies for storing and processing volumes of data that will not fit into an "ordinary spreadsheet": millions of transactions, sensor readings, logs, texts. For a business, they make it possible to see patterns that are invisible in small samples.

Big data is what AI feeds on. Data is where every AI project hits the wall. Yes, we can deploy local models and train them on specific internal procedures, but the model still has to be built and trained first, and only then can the excess be "trimmed away." Building something like GPT-5 runs into the same problem: not enough high-quality big data.

That is exactly why one of the key areas in AI development is creating algorithms that will need less data to train on.

## Digital twins

**What it is:** a virtual copy of a product, a machine, or an entire plant that runs on real data and lets you play out scenarios without putting production at risk.

The technology itself does not depend much on AI. What it needs is classical mathematics. But that is exactly where the symbiosis lies: building digital twins lets you model the very data that needs to be collected to develop predictive analytics. In other words, designing new products through digital twins will speed up the move to risk-based maintenance and make it possible to introduce the very predictive analytics the largest manufacturers are aiming for.

On top of that, digital twins will become a tool for visualizing recommendations and for collecting and structuring operational data.

### The cloud, online analytics, and remote control

**What it is:** computing and storage "by subscription" in a provider's data centers: capacity is rented in a matter of hours rather than procured over months. For a business, it means a fast start with no capital outlay.

Until neuromorphic solutions are rolled out at scale, we will remain bound to deploying AI models on cloud servers or in our own data centers. So AI and the cloud will stay closely linked for a long time to come.

The requirements for AI computing infrastructure (cloud versus on-premises, GPUs, the economics of running it) and data preparation, including anonymization before using cloud LLMs, are covered in detail in AI-2. Practical guide (Chapters 10 and 15); here, only a short overview.

### Digital platforms

**What it is:** an environment in which an ecosystem forms around your data and services: employees, customers, and partners operate by the same rules and in one shared information space.

One clarification is in order here. What is a platform? If a company has built an IT system that brings together its counterparties and its own employees, does that make it a platform?

The defining feature of a platform is open access for anyone who wants to join. Which means a platform is, first and foremost, a large number of participants. And AI is used here in the form of recommender systems: pricing, counterparty checks, product descriptions and product matching, and so on.

The clearest examples are ride-hailing and delivery services. Here AI calculates prices moment by moment, plans routes, matches workers to orders, and so on. If the AI is badly built, the whole system stops working.

Here is an example from my own practice. A last-mile logistics startup made adjustments to its neural network and ended up with enormous losses. The system started assigning orders in such a way that a courier had to drive across the whole city to pick up a new one while another courier might be sitting idle nearby. The result looked like this:

- some couriers might stand at their pickup point all day while the company paid them the minimum rate
- others drove all the way across the city empty to pick up a new order, and the company paid for the mileage

Nothing but losses, and a business model that did not work.

So the success of platforms will depend on whether their AI algorithms work correctly.

### Blockchain and smart contracts

**What it is:** a distributed ledger whose records cannot be tampered with unnoticed; smart contracts execute the terms of a deal automatically. For a business, they provide trust without an intermediary.

Blockchain, in my personal view, is one of the most overrated technologies, and its niche is very narrow. The impact of AI here will be among the smallest.

The key area is monitoring how cryptocurrencies are used, through transaction-chain analysis and data collection from various sources. But the task is so niche that if it does develop, it will happen only through government funding.

### Machine vision

**What it is:** recognition of objects and events in photographs and video: quality control, safety, inventory tracking. This is the most mature area of applied AI, and it is often the sensible place to start.

Machine vision is the processing of video by artificial intelligence. The technology has come so far and worked its way so deeply into everyday life that we no longer notice it.

It checks whether a worker is wearing a hard hat, it inspects product quality on the production line (measuring part dimensions, for instance), and it is what sits behind the sensors in cars (detecting obstacles through the backup camera and warning you about them, for example).

The key constraint here, as across the whole AI industry, is that it needs large volumes of training data and good-quality video. As the algorithms improve and the requirements drop—less data, lower video quality, and therefore cheaper and cheaper cameras—the technology will push ever deeper into production and into everyday life.

### Robotic process automation (RPA)

**What it is:** software robots that replicate human actions in the interfaces of existing systems: filling out a form, moving data between programs, generating a report. For a business, they deliver automation without an expensive rebuild of the IT landscape.

One of the key competitive advantages of any RPA system is the quality of its text and data recognition and how easy it is to configure. So AI will shape this technology as well. Handwriting recognition deserves particular attention. Yes, one day we will arrive at fully electronic document workflows with no handwriting in them at all, but that day is still a long way off.

### 3D printing

**What it is:** layer-by-layer manufacturing of parts from a digital model, from prototypes to spare parts. For a business, it means speed and independence from suppliers of one-off parts.

This, you would think, is the area where AI will have the least impact. But there is plenty of scope here too—namely, generating printable 3D models from

two-dimensional images (photographs, drawings) or even from a text or voice description. So here too, the main area for development is pattern recognition and conversion.

### Data lakes and data warehouses

**What it is:** centralized repositories for raw data (the lake) and prepared data (the warehouse) from every system in the company—the foundation that analytics and AI stand on.

Managing and designing IT infrastructure, improving data quality, and analytics—the hardest area of all, and one of the most promising.

To put it more precisely, this gives us the following set of tasks for AI.

- Cutting the time and cost of designing data collection and storage systems, through recommendations based on the type and architecture of the repository and on the methods of data collection.

Say the task is to analyze video feeds from remote sites for intruders. How do you approach it? Install smart cameras so that only certain events reach the database and get recorded—or install ordinary cameras and stream everything to a data center where powerful AI analyzes the feeds? Which matters more to us: the cost of the cameras or setting up a stable data link? And there are plenty of tasks like this for AI.

- Analyzing data and looking for anomalies in data quality.

Setting up processes to ensure data quality is a critical task for any company on the path to digitalization. Especially as the data-driven approach (making decisions based on data) takes hold, and management decisions, the creation of new products, and much else come to depend on data.

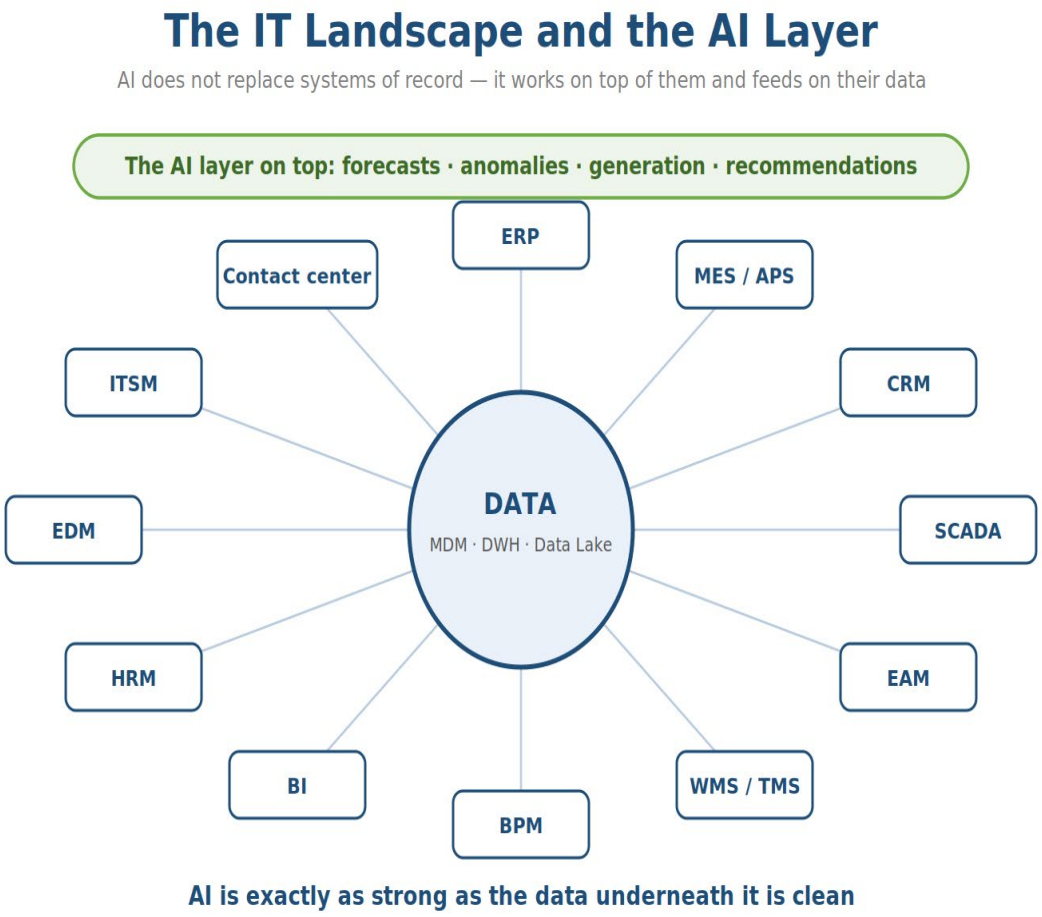
IT even has a dedicated role for this—data stewards, who study the data, make sure it is trustworthy, and resolve failures and errors. AI can help here, and can make this work available not only to giants and banks but to ordinary companies too. Mostly as an assistant to experienced specialists, raising their productivity and effectiveness. And if we also set up integration with BPM systems (business process management systems), we will be able to get

recommendations on optimizing business processes. In other words, AI can become a first-rate tool for IT departments.

### Summary

Yes, this is not a complete list of technologies; many people also put various drones, robots, sensors, and so on in this category. But in my view they are all offshoots of the technologies above. After all, what they run on is machine vision, the processing of big data from sensors, and so on. So the list could be extended indefinitely, but I settled on this set.

## IT Systems




*The IT landscape and the AI layer on top of it*

Below is a map of the main classes of corporate IT systems seen through the lens of AI: what each class is and what it does for a business, what tasks AI will be able to handle inside it, and what prerequisites you will have to meet to get there. The systems themselves as IT products—with examples and the finer

points of choosing between them—are covered in DT-1. Immersion (Chapter 3).

### ERP

**What it is and what it does for a business:** A single backbone for accounting and resource planning—finance, procurement, warehousing, production. It gives you a "single source of truth" for operations and money.

**Key AI capabilities:** Forecasting demand and cash flow, spotting anomalies in postings and purchases, inventory recommendations, drafting documents and routine transactions.

**Prerequisites:** Discipline in data entry, clean reference data (linked to MDM), properly configured processes—otherwise AI will forecast from garbage.

### MES / APS

**What it is and what it does for a business:** Day-to-day production control and detailed scheduling: shift assignments, routings, capacity loading. It gives you transparency on the shop floor and a steady production rhythm.

**Key AI capabilities:** Optimizing schedules against real constraints, predicting breakdowns and changeovers, instant rescheduling when things go off plan, suggestions for the shift supervisor.

**Prerequisites:** Digitized routings and production standards, a data feed from equipment (IoT/SCADA), up-to-date stock levels and order priorities.

### PLM / PDM

**What it is and what it does for a business:** Design and process engineering data and the product lifecycle—from idea to discontinuation.

**Key AI capabilities:** Finding equivalent parts and assemblies (less duplicated engineering work), generative design, checking requirements and completeness, faster approval of engineering changes.

**Prerequisites:** A fully populated product database, a single classification system, a change management procedure that actually works.

### CRM

**What it is and what it does for a business:** Customers, deals, and interaction history; managing the sales funnel. It gives you control over revenue.

**Key AI capabilities:** Lead scoring and churn prediction, personalized offers, summaries of calls and correspondence, a suggested next step in a deal, draft emails.

**Prerequisites:** Salespeople keeping deals fully up to date (the biggest pain point of every CRM), integrated channels, at least 6–12 months of accumulated history.

### SCADA

**What it is and what it does for a business:** Real-time telemetry and supervisory control of equipment. It gives you control over the production process.

**Key AI capabilities:** Spotting anomalies before the operator does, predictive warnings, help analyzing incidents and finding root causes.

**Prerequisites:** Sufficient sensor coverage, stable data collection, labeling of normal and emergency operating modes, a hardened security perimeter.

### EAM

**What it is and what it does for a business:** Management of assets, maintenance, and repair: equipment records, schedules, work orders. It gives you reliability and a manageable cost of ownership.

**Key AI capabilities:** Predictive maintenance instead of calendar-based schedules, prioritizing repairs by risk and criticality, forecasting spare-parts demand.

**Prerequisites:** A history of failures and running hours, a link to SCADA/IoT, discipline in closing work orders with actual data.

### BIM

**What it is and what it does for a business:** An information model of a building across its entire lifecycle—from design to operation.

**Key AI capabilities:** Clash detection, optimization of design choices and cost estimates, scenarios for operation and energy efficiency.

**Prerequisites:** Keeping the model current after the building is handed over, modeling standards, operational data flowing back into the model.

### MDM

**What it is and what it does for a business:** Unified reference and master data: customers and vendors, products, employees. It gives every system "one language."

**Key AI capabilities:** This is exactly where AI maintains data "hygiene": it finds duplicates and synonyms, cross-checks and fills in attributes, proposes a golden record, and cleans up reference data when systems are merged.

**Prerequisites:** A designated data owner, a procedure for maintaining reference data, integration with source systems.

### WMS

**What it is and what it does for a business:** Warehouse management: layout, tasks, picking. It gives the warehouse speed and accuracy.

**Key AI capabilities:** Optimizing slotting and pick paths, forecasting peak loads, catching picking errors.

**Prerequisites:** Bin-level storage, barcoding or RFID, accurate stock levels in the system.

### TMS

**What it is and what it does for a business:** Transportation and delivery management: routes, carrier tenders, tracking. It gives you predictable logistics and predictable costs.

**Key AI capabilities:** Optimizing routes and vehicle loads, forecasting delivery times and delays, choosing a carrier on price and reliability.

**Prerequisites:** GPS data and tracking, integration with orders and WMS, carrier rates and shipment history in the system.

### BPM

**What it is and what it does for a business:** Modeling and executing business processes. It gives you control and repeatability in operations.

**Key AI capabilities:** Process mining: reconstructing the real process from digital traces, finding bottlenecks and deviations from the documented procedure, recommendations and simulation of changes.

**Prerequisites:** Processes must run inside IT systems, leaving traces in the logs, and process owners must be appointed.

### SCM / SRM

**What it is and what it does for a business:** Supply-chain planning and supplier management. It gives you supply resilience.

**Key AI capabilities:** Forecasting demand and lead times, supplier risk profiles, "what-if" scenarios across the entire chain.

**Prerequisites:** Sales and inventory data across the whole chain, delivery history, high-quality master data.

### CPM / EPM

**What it is and what it does for a business:** The financial management backbone: budgeting, consolidation, plan versus actual. It gives you control over the company's economics.

**Key AI capabilities:** Driver-based budget forecasts, explanation of variances in plain language, "what happens if" scenario modeling.

**Prerequisites:** A single budgeting methodology, a link to ERP and BI, a single "version of the truth" for the actuals.

### BI

**What it is and what it does for a business:** Data marts, dashboards, reporting. It gives you speed and a solid basis for decisions.

**Key AI capabilities:** A plain-language "ask your data" interface, automatic generation of takeaways to go with dashboards, finding the causes of variances.

**Prerequisites:** Data marts with a data model people can understand, a metrics dictionary (so everyone reads the metrics the same way), properly configured access rights.

### DSS / digital advisors

**What it is and what it does for a business:** Decision support systems—the culmination of the whole "data → recommendations" chain (Chapter 9 covers this in detail).

**Key AI capabilities:** Decision options with an assessment of risks and consequences, tailoring to the executive's role.

**Prerequisites:** Every prerequisite in this chapter at once: data, processes, integrations—and user trust (see Chapter 9).

### Information security systems (IDM, DLP, XDR)

**What it is and what it does for a business:** Access management, protection against leaks, attack detection. They give you resilience to cyber risks.

**Key AI capabilities:** User behavior analytics, event correlation, incident prioritization. And remember—AI works for attackers too (Chapter 8).

**Prerequisites:** Log coverage across your infrastructure, response processes (a SOC), an up-to-date threat model.

### EDM (electronic document management)

**What it is and what it does for a business:** Legally valid electronic document workflow. It gives you speed and document traceability.

**Key AI capabilities:** Document summaries and checks, extraction of key fields, smart routing, deadline tracking.

**Prerequisites:** Documents in digital form (scans run through OCR), configured routes, a customer and vendor master (MDM).

### ITSM / Service Desk

**What it is and what it does for a business:** IT service management: incidents, requests, changes. It gives you control over IT and service quality.

**Key AI capabilities:** A tier-1 chatbot, automatic classification and routing of tickets, suggested solutions from the knowledge base, forecasting ticket spikes (we will walk through the economics of tier-1 support in Chapter 20).

**Prerequisites:** A well-stocked knowledge base, categorized tickets, discipline in closing them out with a description of the fix.

### Contact center (CCaaS)

**What it is and what it does for a business:** Omnichannel handling of customer inquiries: telephony, chat, messaging apps. It gives you quality customer service at scale.

**Key AI capabilities:** Tier-1 bots and real-time guidance for agents, call summaries, quality scoring on 100% of conversations instead of a sample, routing by topic and sentiment.

**Prerequisites:** Channels integrated with CRM, call recording and transcription, a human in the loop for complex cases: in Sinch's May 2026 study, 74% of companies rolled back AI agents that were already live in support—the bet on full autonomy did not pay off (more on this in the Postscript).

### CDP

**What it is and what it does for a business:** A single customer profile assembled from every channel and system. It gives your marketing precision instead of "carpet-bombing" campaigns.

**Key AI capabilities:** Segmentation, next best offer, LTV and churn prediction, campaign personalization.

**Prerequisites:** Lawful consent to process the data, identity resolution across systems, quality of the event stream.

### HRM

**What it is and what it does for a business:** The HR backbone: recruiting, employee records, development. It gives you faster hiring and better retention.

**Key AI capabilities:** Screening applications for relevance, draft job postings and offers, onboarding bots, turnover forecasting.

**Prerequisites:** Structured job profiles, testing algorithms for bias, handling personal data strictly within the law.

### LMS

**What it is and what it does for a business:** Corporate learning: courses, tests, development tracks. It gives you knowledge transfer at scale.

**Key AI capabilities:** Generating courses and tests from your own policies and procedures, personalized learning tracks, conversation simulators for practicing skills.

**Prerequisites:** Digitized policies, procedures, and a knowledge base, a link to HRM profiles, expert review of the content.

### PIM

**What it is and what it does for a business:** Managing product content for catalogs and marketplaces. It gives you quality product listings at scale.

**Key AI capabilities:** Generating and localizing descriptions, enriching listings, checking attribute completeness.

**Prerequisites:** Authoritative product data (linked to MDM), an editorial policy, sign-off by a human.

### BMS (Board Management Software)

**What it is and what it does for a business:** The governance backbone: board and committee meetings, materials, action items. It gives you discipline in corporate governance.

**Key AI capabilities:** Preparing and summarizing materials, meeting minutes, tracking action items to completion.

**Prerequisites:** Confidentiality above all—an on-premises environment or a trusted cloud; a document workflow procedure for leadership.

## Working with Data

We have already covered what AI does to working with data—both in the section on data lakes and warehouses and in the discussion of the other digital technologies and information systems. Let me add just one thing: data management as a discipline will keep growing in importance. The companies that manage to solve the problem of keeping data current and high-quality, to build the infrastructure and the processes, and to develop both the skills and the motivation of their people will be the ones that get the most out of what artificial intelligence has to offer. So one of my main recommendations is this: start documenting your data flows and building a data management culture, appoint data owners, put the infrastructure in place, and bring in professionals. Given the way artificial intelligence is developing, data management will be the single biggest advantage.

## Chapter Summary

- AI is at its strongest in combination with other technologies: IoT, big data, IT systems. Data and infrastructure set the ceiling on how much value AI can deliver—a model is only as good as what goes into it.
- Working with data is prerequisite number one: "garbage" data nullifies any algorithm, and fixing it once the system is in production costs orders of magnitude more than enforcing discipline at the point of entry.
- Cybersecurity cuts across all of it: every new system and every new integration expands the attack surface (Chapter 8 covers this in detail).

**What to do:** Audit the data sources for your first AI use case (one week). Draw up a road map for data quality: owners, formats, controls (one quarter).

**Reflection question:** Which data in your management reporting do you trust unreservedly—and have you ever put that trust to the test?

## Chapter 16. Regular Management Practices

Regular management practices are a set of recurring rituals that let you make the most not only of your own working time but of the potential of every person on your team. In some companies they go by another name: "the leader's standard practices."

In the classic theory of management, regular management practices cover:

- management planning
- delegating authority and tasks
- working with feedback from direct reports
- running meetings
- personnel matters and decisions about people
- evaluating employee performance
- staff development

We have already covered a number of these areas, so let's move through them quickly.

### Management Planning

Here AI will serve as a component of decision support systems. It will analyze data, flag anomalies, and prepare forecasts with recommendations.

### Delegating Authority and Tasks

AI will help automate the creation of task cards from a message, whether on a kanban board or in a task tracker. In other words, it will act as a robot that can transcribe even a voice message, extract the substance, work out who it is for and when it is due, and then move it all into the tracker automatically. The net effect is that rolling out task-tracking tools gets easier. The key thing is discipline from leaders. If a leader cannot state a task clearly, skimps on letters or words, writes without naming anyone, and so on, AI cannot help.

In projects, AI can also act as the leader's assistant, helping work out what authority is needed to get the tasks done. And where necessary, it will help prepare all the formal paperwork: the directive, the assignment, and so on.

### Working with Feedback from Direct Reports

The main line of work here is running surveys and 360-degree reviews, analyzing large volumes of data, and identifying shared patterns and trends for leaders with a wide span of direct reports—say, ten people or more. On top of that, preparing personalized recommendations for leaders.

### Running Meetings

We also went through this area in the section on board management software (BMS) in Chapter 15:

- preparing and circulating the agenda, drawing among other things on previous agreements and progress reports on assignments
- moderating meetings, recording them, and transcribing the audio with each speaker identified
- preparing the minutes, getting them approved, and creating tasks in tracking tools with owners and deadlines assigned

### Personnel Matters and Decisions About People, Evaluating Employee Performance, Staff Development

Everything in the section on HRM systems in Chapter 15 applies here as well. Integration with task and project management systems is also useful: it will let the AI model collect more data and produce better recommendations.

In many jurisdictions, however, this area will come under close scrutiny.

### Three "Regular Management × AI" Scenarios: Where to Start

Regular management is about rhythms: status meetings, one-on-ones, feedback, tracking action items. AI does not change the rituals themselves—it removes the cost of preparing for them, which is what usually kills them off. Below are the three scenarios that are easiest to start with; tracking action items we have already covered above, in the section on delegation.

**The status meeting.** Before the meeting, the assistant pulls statuses, risks, and outstanding action items into a single summary. The meeting starts not with "so, tell us what you've got" but with decisions on the variances. In my experience, this saves 15–20 minutes on every status meeting and, more importantly, gives the meeting its point back.

**One-on-ones.** Before the meeting, the assistant prepares a briefing card: previous agreements, what has been done, where the person is stuck, what motivates them. Preparation shrinks from half an hour to five minutes, and the conversation becomes about the person rather than a recap of statuses.

**Feedback.** The assistant helps you assemble the facts and check the wording against the structure "fact → impact → expectation," without passing judgment on the person. In my observation, half the conflicts within teams come not from the content of the feedback but from its form—and that is exactly the part AI takes off your hands.

And notice: all of this is the team-level projection of the personal loop, which we will unpack in Chapter 21. Regular management applied to yourself is keeping a journal, planning, and reviewing your own decisions. Start with yourself—your team will find it easier to buy in.

### Chapter Summary

- Regular management practices are the rituals AI slots into most naturally: planning, statuses, kanban, tracking action items. Here AI is a "robot assistant" for the routine work.
- The effect appears only where the rituals already exist: AI reinforces discipline, but it does not create it.

**What to do:** Automate the creation of task cards from messages and meetings (one week). Connect an AI assistant to your team's tracker and to the status ritual (one quarter).

**Reflection question:** Which management rituals on your team hang on one person's discipline?

## Chapter 17. Lean Manufacturing and the Theory of Constraints (TOC)

Lean manufacturing is one of the essential tools a business has. For a company that has already found its niche, lean clears out the chaos, cuts bureaucracy to a minimum, does away with pointless red tape, and makes the company customer-focused. For a young company, lean is a way to strip out everything unnecessary—to put only what the customer actually needs into products and services—and to stay flexible as it grows.

There is more on this tool and how it works behind the QR code and the link below.

*Lean manufacturing. Part 1*

The companies that lead in lean—Toyota, for one—are already experimenting with AI.

The same goes for the theory of constraints. It is another of the main tools for growing step by step and keeping crises and risks to a minimum. There is more on this tool and how it works behind the QR code and the link below.

*The theory of constraints*

## The Impact of AI

In lean and in managing constraints, AI will be able to:

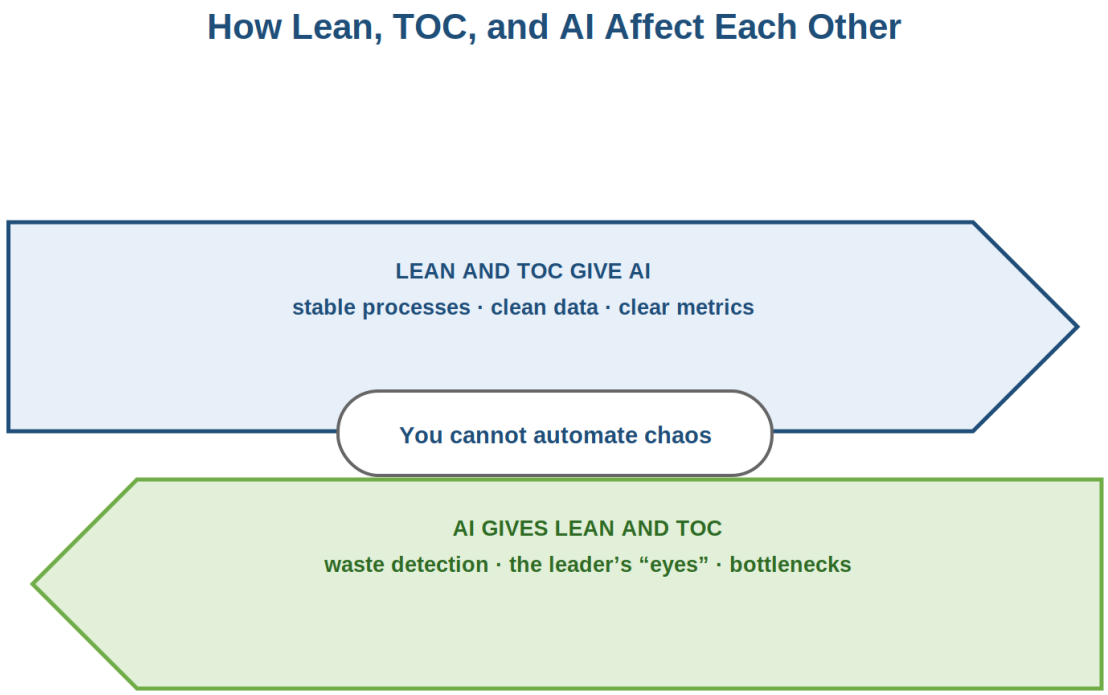
- spot waste in what people do on the shop floor and in IT systems (video analytics in production, analysis of user behavior in those systems)
- spot waste in processes (process analysis, including processes that run inside IT systems)
- analyze waste in production flows and identify constraints
- offer proactive recommendations when it detects a deviation—or on request, when you are planning to scale up or launch new services
- draft a plan for rolling out changes using project and change management tools (choosing the right delivery approach, building the plans, and factoring in the psychology of your people and the company's culture)

AI, together with digital tools, will also help you roll out the core tools—drafting project charters and plans, and supporting the work as it goes:

- Kaizen—analyzing, classifying, and prioritizing improvement proposals, and finding the patterns they share
- TQC (total quality control)
- TQM (total quality management)
- TPM (total productive maintenance)
- Just-in-time—scheduling production tasks
- The A3 report—walking a problem through the report's logic, and generating the report automatically
- The seven steps of practical problem solving—chatbots that act as assistants: they narrow the problem down with clarifying questions, build the data around it, find the causes, classify it, and help route it to the right person based on a documented org chart
- VSM (value stream mapping)—describing value streams from an analysis of business processes or from interviews
- Poka-yoke—analyzing models in PLM systems and preparing recommendations
- Six Sigma—analyzing data in your systems of record to set control limits for a process, and advising which incidents call for a response and which are simply force majeure
- DMAIC, SIPOC, DMADV—acting as an assistant that documents processes from interviews

The hard part is building the models and learning to collect data inside an organization quickly and cheaply. Otherwise all of this stays the preserve of large corporations.

### The Impact Both Ways




Let's look at this through the lens of the core principles and the types of waste.

**The 14 Principles of the Toyota Way**

**Section I. Long-term philosophy.**

**Principle 1. Base your management decisions on a long-term philosophy, even at the expense of short-term financial goals.**

In essence: focus on long-term development rather than on operating metrics. And unfortunately this is the principle that gets ignored most often, because it is a psychological trap.

What does an owner or a shareholder usually want from top management? Right: good-looking numbers for revenue, operating margin, EBITDA, and profitability.

This, broadly, is the central conflict between lean and reality. The same is true of AI. But companies with a strong culture can run AI projects, and they will. They understand that this is the future.

**Section II. The right process will produce the right results.**

**Principle 2. Create a continuous process flow to bring problems to the surface.**

The point is to set up the work and the production process with as little waste, inventory, and work in progress as possible. Combined with communication that actually works, this surfaces problems early, before they turn into a hazard or a crisis. In my own practice this is the master rule, in project delivery as much as anywhere: collect feedback and surface problems as early as you can.

AI will be extremely useful once it can pull data from IT systems and surveys, visualize production flows, and offer recommendations—for instance, generate a kanban board on the spot with exactly the columns your company and its internal processes need.

**Principle 3. Use pull systems to avoid overproduction.**

Ever had another department's work shoved onto you? I often see different departments working on their own, and it ends either with overstocked raw-material warehouses or with inboxes clogged with memos, letters, and tasks.

This is the single most common failure in how department heads work with their people. Many of those leaders are, in effect, human relays—all they do is pass tasks from above down to whoever has to do them. Forty or fifty untouched tasks pile up on one person; they try to do it all, hop from one to the next, and end up finishing nothing. And practice shows that capping the number of active tasks at three raises productivity by as much as 50%. First, the person is not jumping between tasks (it takes the brain 20–40 minutes to get into one). Second, they still get some variety, which protects their mental state and keeps them from getting stuck on a single problem.

The internal customer who receives your work should get what they need, when they need it, in the amount they need. That is the foundation of just-in-time: new parts, or new tasks, arrive only as the previous ones are finished.

In short, keep work in progress to a minimum—and that applies to offices too.

The synergy with AI here is in personal planning and in recommendation systems for employees. The leader assigns tasks, but a person receives the next one only after finishing the last. And the system decides on its own which of

the open tasks to hand them—or which person in the department should get it in the first place, taking individual strengths into account.

It is essentially an aggregator along the lines of a ride-hailing platform, with AI that understands what each person is good at and dispatches the work accordingly. Solutions like that fit the lean concept perfectly.

### Principle 4. Level out the workload (heijunka): work like the tortoise, not the hare.

Recognize the heroics, and the feast-or-famine rhythm that goes with them? In Japanese philosophy this is called mura. One of the jobs of lean is to level the load—no month-end scrambles, no endless heroics.

AI here becomes part of the planning systems, in production and in office work alike. It accumulates statistics on tasks, on when they appear, and on their deadlines. The system then creates tasks and reminders in advance, based on the company's and the unit's own experience and specifics.

### Principle 5. Build a culture of stopping to fix problems, to get quality right the first time.

What made Toyota cars prized—and what keeps them prized? Right: quality. If you want customers with money, you have to deliver quality.

On the shop floor, jidoka—equipment with machine vision—is the foundation for building quality in. Such projects have long since become the norm: quality control at individual stages, control of the process itself, and inspection of the finished product.

The same kind of solution can be built into office work: for example, AI suggesting that a project is ready to change status, that it has been thought through far enough. Or analyzing documents, marketing materials, and so on.

### Principle 6. Standardized tasks are the foundation for continuous improvement and employee empowerment.

Every time I hear "we don't have standard tasks," it usually turns out that the team either does not know how to structure its work or does not want to.

Stable, repeatable working methods are what process management rests on. Its job is to make results more predictable, the work more coordinated, and output more even. This is the basis of flow and of pull.

Yes, in a very young company this is beside the point for a while: you need to get the whole system running and produce a result first. But sooner or later there is no way around it. Endless creativity always ends in a crisis.

You have to capture the knowledge you accumulate—run project reviews, identify the best approaches and methods, and standardize them. And you have to reward people for improving those standards. Continuous improvement is, after all, the core creed of lean. But before you can improve something, you first have to describe it and write it down. Otherwise you get chaos, and everything that comes with it. And for IT this matters just as much. Creative people hate standardizing their work, but it has to be done—"don't want to" or not.

Once again, the synergy with AI is excellent. It can sort the various incoming tasks, analyze them, standardize them, break them down, and assign them to different people or units. On top of that, drawing on accumulated experience, it can recommend how to redistribute authority, rework processes, and so on.

**Principle 7. Use visual control so no problems are hidden.**

This principle is really part of the fifth. I use color labels. But that takes discipline, and discipline comes from the leader.

Wherever you can, cut reports down to a single sheet, even when the decision at stake is a critical financial one. There is even a tool for exactly that: the A3 report.

AI here pulls texts, tables, notes, and the rest together into final decisions and visuals, and highlights the key problems and proposals.

**Principle 8. Use only reliable, thoroughly tested technology.**

Technology should serve people, not replace them. When it comes to digitalization and automation—and AI is precisely digitalization and

automation—it is often better to run the whole process manually first and only then bring in a solution.

First, you will see where the real problems are. Second, you will avoid the illusion that everything can be automated, and you will be able to draw the boundaries of where AI applies.

Besides, new technology is often unreliable and resists standardization, and that puts flow at risk. AI itself, for all its progress, still makes fairly stupid mistakes at times.

So instead of unproven technology, use the known, proven process. That said, exploring new technologies and new paths still deserves encouragement—everything proven was new once.

To gauge how ready a company is for new technology, and how ready the technology itself is to be used, you can turn to metrics such as TRL, MRL, and CRL.

In brief:

- TRL (technology readiness level) measures how technologically mature a solution is. The method distinguishes nine levels: at the first (TRL-1) the concept of the technology has only just been formulated and its usefulness argued; at the last (TRL-9) the software product is finished and meets every requirement set for it.
- MRL (manufacturing readiness level) is a method for gauging how ready your production, your processes, and your business are to take a new technology or product into serial production. It is used by the U.S. Department of Defense, and the U.S. Government Accountability Office has named it a best practice for improving procurement outcomes. MRL is meant to compensate for the shortcomings of TRL, and it distinguishes ten levels: at the first (MRL-1) the basic manufacturing options for the solution have been identified; at the last (MRL-10) full-scale production is up and running. And if you rate the AI field on the MRL scale, you can see that AI projects have not yet reached a high level of manufacturing

maturity: the whole industry offers not mature solutions but prototypes and early versions.

- CRL (commercialization readiness level) is a method for gauging market and commercial readiness. At the first level (CRL-1), trends, market reviews, conference agendas, and patent activity have been analyzed and a conclusion drawn about the market's need for the new solution. At the last, ninth level (CRL-9), the product has been launched and data is already being collected to improve it further.

So all three metrics are linked: TRL covers work on the technology, MRL covers getting the technology into production, and CRL covers developing the product offer and the marketing behind it.

Can we find a single AI solution on the market that is mature on all three counts? By the standard of the eighth principle, then, it is too early to talk about deploying AI solutions in large, critical production environments.

This is one more argument for my claim that AI solution developers should be looking at small and midsize businesses rather than at corporations.

## Section III. Add value to the organization by developing your people and partners.

### Principle 9. Grow leaders who thoroughly understand the work, live the philosophy, and teach it to others.

It is better to grow your own leaders than to buy them from outside. First, this is one of the key rules of motivation. Hiring from outside all the time demotivates your own people. They start losing faith in the future. In the end they focus on their own goals, and you can forget about full commitment. Second, leaders who know their business bring deep expertise by default and are aware of every pitfall in your work. And in digitalization and automation, a great deal rests on exactly that.

AI here helps collect and structure feedback from staff about the changes, and helps train people in the basics of lean, project management, and digitalization.

A leader also has to do more than deliver on assigned tasks and get along with people. AI can become the tool that helps with writing letters: putting thoughts into words, checking spelling, choosing the best channel—email, a call, a face-to-face meeting, a written memo.

### Principle 10. Develop exceptional people and teams who follow your company's philosophy.

AI here helps select team members and train them.

### Principle 11. Respect your partners and suppliers by challenging them and helping them improve.

You can be as digital as you like, but if your partners are still living in the paper era, the effect will be limited. It is just as the theory of constraints says: the strength of a chain is set by its weakest link.

You have to create the conditions for your partners to grow, help them along, and show them what high performance looks like.

AI can, for instance, run a first-pass audit of a partner from a test questionnaire or an analysis of their business processes, track how they investigate complaints, analyze public information about them, and produce recommendations.

## Section IV. Continuously solving root problems drives organizational learning.

### Principle 12. Go and see for yourself to thoroughly understand the situation (genchi genbutsu).

I have seen this again and again: top executives trusting their managers and the reports on their screens. Here is one case from practice. The founder of a company had come up through production. He walked the shop floor every week. As a result, people understood that their work mattered, they were motivated, and he had reliable information.

His successor, though, had a classic management education. He built a team of managers and a system of reports. But what do hired managers want? To be safe and to be paid on schedule. That is normal. The trouble is that at every

step, with every manager in the chain, the information gets more distorted, so the picture that reaches the very top no longer matches reality—and decisions then get made on a mistaken view of it. This, incidentally, is one of the central problems of government administration.

Combine this principle with good automated data collection and transparent analytics, and you avoid the problem.

AI can become the leader's eyes: analyzing video and employee profiles, advising where to go and whom to talk to, showing who is the most engaged and the most valuable to the company—which is also a way of keeping those people motivated.

**Principle 13. Make decisions slowly by consensus, thoroughly considering all options; implement decisions rapidly (nemawashi).**

You could also call it "think slow, decide fast." One of the foundations of management is weighing alternatives. Making decisions on the strength of a single opinion is too dangerous. For a young company that is sometimes the only way forward, but the further you go, the more often you need consensus—and consensus takes preparation plus several people with different opinions and different temperaments. This is part of what Adizes's theory rests on: one person cannot combine all the competencies required, so you need a well-rounded team. And that, in turn, means learning to discuss opinions openly and to work through conflict.

Do not settle on a single course of action without weighing all the alternatives. But once the decision is made, you have to act.

It is this principle that leads to the concept of digital advisors—systems that can process data from IT systems together with the opinions of individual experts and produce recommendations and possible alternatives.

**Principle 14. Become a learning organization through relentless reflection (hansei) and continuous improvement (kaizen).**

I have already said that you cannot optimize everything in one go.

Digital twins will be able to learn on their own, without human involvement, produce optimization recommendations, prioritize changes, and steer clear of large, high-risk projects. So this is about reducing risk. The smaller the project and the change, the easier they are to manage and the lower the cost of a mistake.

AI can raise effectiveness by picking the right change-management methodology and singling out the key steps in a project—all without having to sink millions into training everyone on staff.

### Waste

One of the main tools in lean is the eight types of waste. They show up on the shop floor, in the office, and in IT.

You need to be able to spot waste in a process and to know how a particular technology can help remove it. That does more than put your own work in order: it lets you strip excessive requirements out of digital projects at the very start, which cuts both their cost and their timelines.

Below is a short list of the types of waste and how they show up.

| Type of waste | Examples in production, office work, and IT systems |
|---|---|
| Overproduction | Making extra copies of documents and reports<br>Duplicating information across different documents and IT systems<br>Duplicating action items<br>Excess capacity in IT systems<br>Unused features in IT systems |
| Waiting | Waiting on deliveries or on equipment that is not ready<br>Waiting because instructions, approvals, or planning are missing<br>An IT system that runs slowly |
| Inventory | Holding six months' worth of production materials without accounting for the cost of running the warehouse<br>Stockpiles of office supplies<br>A large backlog of unreviewed tasks<br>Archiving documents no one uses |

| | |
|---|---|
| | Storing data that is never used but consumes IT infrastructure resources. For example, keeping weekly system backups for a year. |
| Unnecessary transportation | Siting the spare-parts warehouse far away from production<br>Hand-carrying documents<br>Unnecessary approval steps in electronic document-flow routes |
| Unnecessary motion of people | Hunting across the whole work area for a tool you need<br>Equipment or furniture placed inconveniently<br>Leafing through documents that have no table of contents or page numbers<br>In-person meetings that could be online or replaced by a shared digital board<br>A confusing IT system interface, or unoptimized system logic that forces extra clicks—for example, no links for navigating documents; unstructured file storage |
| Defects | An underqualified employee got the calculations wrong<br>Getting the same review comments over and over on routine tasks<br>Getting different review comments on the same routine tasks or documents<br>A request for information so vaguely worded that the person has to ask what it means<br>Unreliable data in reports because of an error in an IT system's algorithms<br>An IT system failing or going down |
| Overprocessing | Making a part or delivering a service to a tolerance no one needs<br>Excessive requirements in the specification<br>Constantly refreshing the design and the interface elements with no user request and no research behind it |
| Unused human potential | Hiring someone overqualified for the job<br>Large amounts of manual, routine work that could be automated<br>Ignoring employees' proposals to improve or streamline processes |

Lean as a preparation stage before AI adoption ("Lean Before AI") is covered systematically in AI-2. Practical guide (Chapter 8); this chapter focuses on how lean, TOC, and AI influence each other.

The tools themselves—the five focusing steps of TOC, the 14 principles, and the types of waste in lean—I covered in detail in DT-2. The Systems Approach (Chapters 2–3).

### Chapter Summary

- "Lean Before AI" is the key pattern of this book: first remove the waste and put the process in order, then automate. Otherwise AI simply scales up the mess.
- The synergy: AI finds waste and bottlenecks (TOC) in the data and gives a leader their "eyes" back—but the gemba and live observation remain irreplaceable.
- The influence in reverse: lean processes yield clean data and clear metrics—a ready-made foundation for training AI and keeping it in check.

**What to do:** Take one process and map the eight types of waste across it (one week). Only after you have simplified it should you decide where AI belongs (one quarter).

**Reflection question:** On which tasks are you, as a leader, just a relay—and on which do you create value?

## Chapter 18. The Processes of Business Functions

We have looked at how AI affects the core digital technologies and the tools of the systems approach—and at how they affect AI in return. But I understand that some people find it easier to look at all this from a slightly different angle. So I will now go back over everything said earlier, in a different vein.

**Life, as usual, moved faster than the printed word.** While this edition was being prepared, the market handed this chapter two confirmations—and one sobering lesson. In customer support the market has been through its first wave of disillusionment: according to a Sinch study (May 2026), 74% of companies rolled back AI agents they had already launched—the deployments

that win are the ones with a human in the loop (there is a detailed discussion in the Postscript). In Russia, AI agents handle narrow industry tasks: matching units for property developers, processing requests in housing and utilities, diagnosing laboratory equipment (Yandex B2B Tech, March 2026). In software development, the DORA report (2026) shows AI assistants paying for themselves in roughly eight months—provided engineering practices are mature. The common denominator: the effect comes not from "AI in general" but from a specific scenario in a specific function.

## Digital Business

This area breaks down into three blocks:

- improving operational efficiency
- finding new niches and startups to invest in
- developing technology partnerships and piloting innovations

I will go through the first block in detail—it applies to any company. The other two belong to corporate development rather than to process work, and AI plays the same part there as it does everywhere else: a tool for analysis and for preparing decisions, not an independent player.

Within operational efficiency we can single out:

- IT infrastructure and its optimization
- materials supply and logistics
- organizational efficiency
- sales and current products and services, that is, customer service
- economics and finance, including accounting
- personnel management

### IT Infrastructure

AI will simplify how IT infrastructure is designed and how it is maintained. AI projects, however, will add substantially to the load on server hardware. It would be even more logical to say that AI projects will require buying or

renting additional capacity. You also have to factor in information security risks—AI models can be hacked, and communication channels have to be protected, especially if you plan to use cloud services.

Turn it around and look at how IT infrastructure affects AI projects: infrastructure limits and information security requirements will also drive the development of lightweight, specialized solutions.

### Economics and Finance

In this area AI can be used to simplify process automation (recognizing source documents and moving the data into systems of record, for example), forecast results, plan budgets, and produce recommendations that take the context of the situation and the organization into account.

Right now it is the economics that is the weak spot of current AI projects, for business customers and developers of AI solutions alike. Together with the demands of IT infrastructure and information security, this brings us back once again to the need to create cheap models.

### Organizational Efficiency

Here we will get help with:

- running meetings—preparing the agenda, transcribing the audio, and summing up the outcomes.
- managing projects, project portfolios, and change. For example: setting goals, identifying the key risks, deciding whether new processes are needed or existing ones should be reworked, choosing the delivery approach for a project, and so on.
- analyzing and designing business processes.
- developing the company strategy, the organizational structure, and metrics for every unit, and rolling out the other tools of the systems approach.

### Materials Supply and Logistics

Here it makes sense to use AI as an advisor to business units when they are organizing procurement. AI will also help analyze procurement documents and supplier bids, including gathering information on how reliable a supplier is. And it will help analyze and merge duplicate requisitions, optimize warehouse logistics, and route deliveries.

### Personnel Management

Everything from the block on HRM systems (Chapter 15) applies here too. I will only add that it is resistance from people that will be the serious constraint. And every one of the main causes of resistance will be at work:

- technical resistance—from not understanding what this is or how it works, from the lack of instructions and training, and from raw, soulless AI solutions (especially if you deploy something in the form of chatbots)
- cultural resistance—from a failed or fruitless experience of rolling out IT solutions in the past
- political—from the fear of losing one's job

So when you deploy AI, pay particular attention to the following:

- train your people, both in digital literacy generally and in working with the new tools; build that into the project plan and into the requirements you set for AI vendors; and move away from chatbots toward programs in which the user does not interact with the AI directly
- organize the rollout in small steps—pilot projects, finding the innovators, and publicizing your successes
- explain to people that this is needed to raise efficiency and competitiveness, to slow headcount growth, and to keep the company from sliding into bureaucracy; and put AI projects into leaders' metrics

### Customer Service and Commercial Performance

Here it is worth recalling CRM systems as well, in particular:

- transcribing recordings, capturing what was agreed, and filling in the data in the CRM system

- generating cues from past calls and preparing scripts
- first-line technical support and advice for customers on order status and standard questions, replies to letters and reviews
- calculating individual price quotes
- sales analytics and recommendations for leaders
- preparing sales plans and alternative versions of them
- analyzing customer requests and preparing proposals for leaders

### Digital Marketing

This is the function that generates digital content and manages end-to-end communication through digital channels.

This is an area where AI is already advancing and being used actively. For example: preparing advertising campaigns and analyzing the results, market research, personalized offers to customers under loyalty programs, generating content for promotion, preparing hypotheses to test, SEO descriptions, marketing materials and product descriptions, replies to comments, analysis of the general trends in reviews. In short, the list is long, and marketers are already experimenting and learning.

### Digital Manufacturing

Technology, sensors, robotics, and artificial intelligence are moving into production, and some of the decisions on the line are already being made by a machine rather than a person. Design starts out in a digital environment, with digital twins created alongside it. This also covers maintenance and repair, industrial safety, real-time analytics, automated reporting, design and process engineering, engineering documentation, quality control, and so on.

Within digital manufacturing we can single out the following areas:

- selecting materials for manufacturing parts
- running research and preparing new prototypes automatically (the design of a new rocket engine, for instance, took several weeks instead of several years)

- analyzing large volumes of sensor data, spotting common patterns in how equipment behaves, and preparing recommendations on how to optimize operating modes
- detecting anomalies in equipment operation—including for quality control of maintenance work, recommendations for operators, accident prevention, and fewer human errors, while at the same time lowering the requirements placed on operators
- quality control of products and of production processes as a whole
- planning maintenance
- preparing production plans and reporting
- monitoring compliance with industrial safety requirements, including observance of access rights for working on equipment
- developing engineering documentation and designing process flows

### Digital Analytics

This is the collection and structuring of data from every channel and source—what is known as big data. Here AI will become a helper both for data analysts (analyzing what happened) and for data scientists (predicting the future); it will speed up the building of mathematical models and the discovery of patterns of every kind:

- optimizing and refining equipment components faster, using sensor data from products already in the field
- identifying patterns in customer behavior

All in all, this is where the largest contribution from AI can be expected, alongside marketing.

### Digital Products

This means creating new digital products in the form of analytics services and applications. If businesses learn to digitize their own expertise, this will be a breakthrough area for the service sector. On the one hand, consulting services

of every kind will be under threat; on the other, individual companies may turn into unicorns.

For the manufacturing sector, this is a chance to build better full-cycle services—from selling a piece of equipment to disposing of it.

### New Ways of Working

This area is about how we work and how we think at work. The main areas here will be help with communication and project management.

And what does AI change in the work of the leader—not the company's work, but that of the person in the director's chair? The subject is important enough to have a chapter of its own in Part 3 (Chapter 21).

### Engineering and Design: From Generation to Verification

The engineering and design function deserves a section of its own. It reached the pilot stage later than marketing and support, but it strikes at the very core of a manufacturing company. AI assistants for process engineers and design engineers are already running in pilots at large machinery plants: the system turns engineering documentation into a draft process plan; it reconstructs an editable digital model from a finished part or a drawing; in some cases preparing project documentation becomes several times faster. Note carefully what is actually being compressed. **It is not the profession that is being compressed but one specific phase of the work:** the rough generation of options and the routine paperwork. Precisely the phase where a junior engineer used to learn and grow.

**Shift one: from generation to verification.** The engineer stops being the author of the first draft and becomes an editor and a validator. That is cognitively harder, not easier: criticizing a plausible but wrong process plan is more difficult than writing your own from scratch. Traceability of the source is critical here (see Chapter 5)—without it, verification turns into a ritual in which the person simply clicks "accept."

And this shift comes down to the same ability: an engineer who is fluent in the terminology of their field both sets the task more precisely and spots a

plausible error in the answer sooner. The language of the profession becomes an instrument of control.

**Shift two: from a single option to managing the space of options.** When generation costs pennies, the value moves to setting the constraints: tolerances, the tooling you actually have, the machines actually on your floor, unit cost, maintainability, and safety. **The ability to state a constraint precisely becomes the central engineering competency.** In essence, this is a return to the culture of writing a proper specification—except that now a bad specification punishes you in five minutes rather than in six months, and it does so at industrial scale.

**Shift three: from personal expertise to knowledge as capital.** An assistant that works is an assistant trained on your process plans, your customer complaints, and your failures. Which means the knowledge that lived for thirty years in the head of a shop-floor veteran has to make it into the data—and that changes the social contract: undocumented knowledge stops being personal insurance against dismissal and becomes an asset of the company. **People do not resist the machine. People resist losing their monopoly on knowledge.** Until you settle that question honestly—through status, through money, through the role of mentor and verifier—you will not get the data, and the assistant will remain a pretty demo.

And the main risk here, the one that is discussed far too little: if AI takes over the junior engineer's work, where will the senior engineer capable of verifying AI come from ten years from now? **We risk winning three years of productivity and losing a generation of expertise.** The answer is not to "keep AI away from juniors"—that battle is already lost; they use it anyway. The answer is to redesign the training: a junior engineer learns not from grunt work but from picking apart the model's mistakes. Give them a hundred outputs from the assistant, twenty of which are wrong, and the job of finding the errors and justifying the verdict. That is a faster and harsher training ground than three years of fetching and carrying. But it has to be built deliberately—it will not come together on its own.

## Chapter Summary

- AI is working its way into every business function—from marketing and production to analytics and product creation; what differs is not the logic but the data and the cost of an error.
- One formula fits adoption in any function: data → model → integration into the process → control of the result.

**What to do:** Pick the function with the most painful routine work and one task within it, and record your baseline metrics before you start—that is your pilot candidate (one week). Run the pilot and compare the result against the baseline (one quarter).

**Reflection question:** In which function of your company—or in your own work—will AI pay off fastest, and why there?

## Summary of Part 2

### General Conclusions and Forecasts

I do not think AI will replace workers, but it will certainly help advanced specialists handle more companies, more clients, more business units. And it will help companies rein in headcount growth and overcome the talent shortage.

AI will make management radically more effective and reduce the need for the assorted staffers and managers who make up an administrative apparatus. The use cases are limited only by imagination, and there is a whole industry's worth of possible solutions here: assistants and advisors, spotting deviations, training staff, simplifying and speeding up processes.

The key obstacles to bringing AI into company management are these:

- developing and training specialized models
- resistance and distrust from people, fear of the new (not least because they simply do not know what it is)
- the desire to shift responsibility onto the tool (although no one is immune to a mistake or plain incompetence from an ordinary employee either, and the responsibility still sits with the head of the organization)

Should we expect quick results here? Of course not. We cannot even use ordinary digitalization to the full—we deploy isolated tools and patch up part of the problem with automation. And here we are talking about AI, of all things—barely studied, opaque, and frightening. Once again, we face two extremes: some will sit and wait for a magic pill that will build the business for them and then run it, and others will never be able to trust it.

I expect that we will start to see a truly fundamental effect on management and on the development of technology in about thirty years. First the technology will seep into our lives through B2C products, and our children will grow up with it—call it eighteen to twenty years for that. Then it will take time for those children to join companies and rise to leadership, or to build companies of their own and take them past the startup stage. That is another ten years or so.

Yes, the most advanced among us will start looking for practical applications right now, but they can be counted among the innovators and early adopters, who together are no more than 10–15% of all people. This book is written to help them.

On top of that, it is on a horizon of thirty to forty years that the technology will become more mature: new, more efficient algorithms that need less data will appear. By then it will be possible to build new tools and models without any special skills. Building them, in other words, will be simpler, faster, and cheaper. We may also see neuromorphic and quantum computing come into their own by then.

## Part 3. Practical Examples and Recommendations for Applying Artificial Intelligence

We have worked out what AI is and how it fits into a management system. It is time to finish the landing—to move on to examples and recommendations. This is the liveliest part of the book: here are the cases I have collected from my own practice, from conversations with colleagues, and from my travels, and recommendations with somebody's hard-won bruise behind every one of

them. Read it with a pencil in hand: almost every example can be tried on for size in your own business.

## Chapter 19. AI for Small Business and Entrepreneurs

### AI in 1C Solutions

Now, in the middle of the 2020s, solutions from 1C are becoming a standard component of digitalization in Russia. 1C has made a name for itself as a kind of SAP or Oracle built for Russian companies. It is an entire ecosystem spanning several classes of applications, and on top of that there is an enormous number of add-ons and variants from third-party developers—applications for managing maintenance, leased premises, and so on. Which means there is plenty of room here for artificial intelligence as well.

In effect, 1C becomes the master system that everything else integrates with. For example, I supported the rollout of a system for logging unsafe acts and unsafe conditions at a manufacturing company.

The architecture looked like this:

- An app and a web interface for users at the sites to interact with the system and file reports.
- A master-data catalog in 1C, which supplies the data on the facilities.
- An AI that takes the records out of 1C, processes them, and sets priorities.
- 1C, where the engineering staff work and handle the resulting tasks.
- A BI system for reporting and analytics.

As for AI services from 1C specifically, they have been in development since 2022, and the catalog now covers most of the routine. One service recognizes source documents: a scan of a crumpled delivery note becomes a record in the system with its fields already filled in. Another forecasts sales and builds the management reporting, so the leader does not have to wade through the analytics personally. A third transcribes meetings speaker by speaker and turns them into minutes with tasks, deadlines, and owners. The rest of the lineup predicts equipment failures from sensor readings, answers staff

questions in natural language, analyzes how a sales rep is working and hands back a step-by-step plan, and issues measurement instructions to lab technicians on the production floor.

All of these AI services come with an API, so you can connect your own databases and knowledge bases and train the AI for the specifics of your company.

Below are two more examples of using AI on top of 1C.

### *Automating Photo Processing to Populate an Online Store's Product Database*

The original business process

- The photographer receives the goods and organizes the shoot and the post-processing.
- They dump all the files in bulk onto a flash drive or into the cloud (in this case that meant around 40,000 photographs).
- A manager manually sorts every photograph into folders inside 1C, recording which file belongs to which product listing.

The optimized business process

- The manager scans the product's barcode (the right listing opens in 1C), then places the item in a dedicated photo station—a lightbox—with a camera in front of it that is also connected to 1C.
- In 1C they press the "Take photo" button. The camera takes the shot, and the AI picks up the file and processes it: background removal, compression, and so on.
- The shot appears in the product listing.

Yes, in the new process the manager still has to know the product range and what the goods look like, and they need at least basic skills in positioning an item in a lightbox—but training them in that is cheaper than hiring a photographer.

### *Copy for Online Stores and Marketplaces*

One of the key tools of any online platform is direct communication with users. But a small business does not have the capacity to stay on call and answer every review and customer question—especially given that marketplaces push users to leave reviews, and those reviews then determine how stores and product listings are ranked (again, with the help of AI).

So AI-based tools have appeared that answer user comments on their own. There are both ready-made templates and builders for creating your own modules inside 1C.

Solutions for generating product descriptions, specifications, and category copy work the same way—the thing I wrote about in the section on PIM systems.

The result is an interesting loop:

- one AI builds the listings and the descriptions and processes the photos
- a second AI analyzes all of it and ranks it
- users leave comments
- a third AI answers them
- the AI from step two analyzes and ranks everything again and generates recommendations for users

### Marketplaces for Specialized AI Solutions

On the Russian 1C marketplace alone there are already ready-made add-ons of this kind: filling in product descriptions, analyzing marketplace reviews, processing product photos—each specialized for one task.

And marketplaces like these for AI solutions will keep spreading. OpenAI itself launched the GPT Store—a shop for custom chatbots.

To make that possible, in November 2023 OpenAI introduced the GPTs feature, which lets you create individual chatbots: you can add special capabilities, skills, and knowledge. You configure the chatbot with a text description of its role—the prompt. You can also upload a data source. I wrote

about this mechanism in the section on digital advisors: AI models that acquire a “specialization” in a subject area.

The GPT Store is available via the QR code and the link below.

*Gptstore.ai*

And the QR codes and links below lead to OpenAI’s dedicated launch page and to its GPT builder.

*Introducing GPTs*

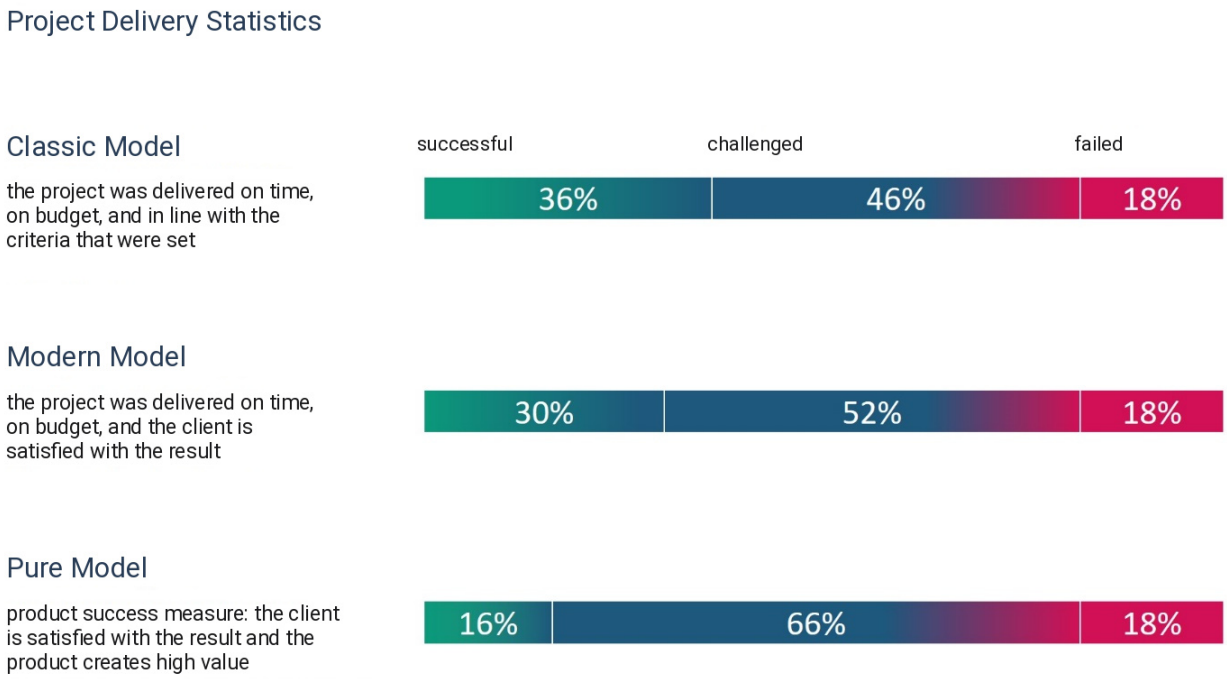
*The GPT builder*

And this is an enormous opportunity for small and midsize business—an accessible online venue both for buying AI solutions at a reasonable price and for launching your own. In essence, it is an Amazon for AI applications.

## A Digital Advisor for Project Management

One of the main problems for small and midsize businesses is that they have neither the resources nor the ability to assemble a team of world-class experts. Yet a small business has no fewer tasks than a large corporation—and rather more vulnerabilities. I see this problem in my own field too: managing projects, change, and products.

Every business today faces an enormous number of projects. Up to 50% of a leader's working time is spent on project work. And the statistics on managing projects, change, and new product development are bleak. Here are a few figures from international studies.




Meanwhile a competent project or product manager costs $2,300 to $3,500 a month before taxes. Rolling out corporate, end-to-end management systems is, once again, expensive: projects like that start at $58,000 on average. And they do not answer the fundamental questions anyway—how to choose a delivery approach, which tasks actually need doing, and so on.

As for training your people, a course will run $600 to $1,700. But:

- in practice there are serious questions about the quality of the training.
- participants get no support afterward (and building a competency takes up to a year of ongoing support, plus experience in a variety of situations).
- there is a high risk that you train someone and they leave for better pay. The reasons vary: someone offered them more money, the work got boring because of the routine, burnout, or the training turned out to clash with the reality of the job.

And the most important point: 90% of projects need no extraordinary competencies to succeed. They just need someone to follow the methodology and take certain steps:

- define the project goal and its objectives
- identify the stakeholders and work with them

- assess the risks and keep track of them
- keep minutes of meetings
- spot deviations early, and so on

That led me to a simple thought: managing projects, change, or products ought to become a standard competency. Better still—build an advisor: you answer a set of test questions, it tells you what to do and how, and from then on it stays with you and teaches you.

In other words, we are talking about a ready-made methodology for a business, drawn from the best minds in the field, plus development for its people. You pay about $23 per employee per month and you have a personal assistant that both advises you and accompanies you.

Getting there means solving several problems.

The first is to break the model down into the factors that drive the final outcome. As I see it, those are the characteristics of the project itself, the culture of the company, and the competencies of the project manager.

The second: what is needed is not a chatbot but a complete solution for a situation in which the user knows nothing at all about project management. So they either answer test questions, or pick from a dropdown with prompts, or type in only the very simplest data by hand—the names and emails of the participants—and upload working templates.

The third: you need a closed-loop system that gathers the data for its own retraining, including data from other IT systems.

The fourth: you need a ready-made methodology, and it has to account for the project manager's personal qualities. Even with the best training courses and practices, you cannot take the human factor out of it, and a great deal rests on the leader's experience and psychological makeup.

Yes, this is not one big AI model but a whole orchestra. And we are building a closed-loop system in which the AI keeps learning from stakeholder feedback during the project and after it, and from the data on budget planning and execution and on delivery timelines.

The core idea is to implement the principles of deep learning, in which the AI itself looks for connections and correlations on the basis of the model provided. That is how we move away from a chatbot toward a finished interface. You fill in the stakeholder section, say, and the system then drops that data into the project charter, sends out the survey link, and points out that John Miller has dropped out of the project and somebody ought to work with him.

The best part is that the project reaches break-even at just one or two thousand users a month (infrastructure + developing the methodology + the development team). And the main risks—which will be typical of any solution like this—are:

- complexity.
- marketing, and the fact that people are not even aware that a thing called "projects" exists and can be managed somehow.
- you cannot take a solution like this straight to large corporations, because they will demand it be adapted to their own processes—which, let us be honest, are frequently a mess. On top of that, every project will cost hundreds of thousands of dollars, sometimes over a million, drag on for months, and may not reach the finish line at all. And it is slow money: you will be paid a year or two after the start, which means a serious risk of a cash-flow gap.

### Generative AI for Content Creators

Generative AI has been in use in the digital content industry for a long time now. Here is an example from a partner of mine who created a series of videos to engage the staff of a large energy company in a new culture built on the philosophy of lean manufacturing. I will give it in his own words, in the first person.

"As so often happens with clients, something needed to be made but nobody knew what. My brief was to develop a style that would set the new culture apart from the old, and to present it through video clips that would loop continuously on the main screens around the company's premises.

“Obviously I had to allow from the outset for the fact that people usually walk past those screens without so much as glancing at them. Nor could I count on sound. So it had to be something bright, fast, and… silent.

“Producing video content is familiar territory for me, but this job had its complications. First, the new style called for a lot of graphic elements—the picture was meant to assemble itself out of them like a collage—and although I work with design all the time, I am not an illustrator. Second, we decided to use the faces of real employees to catch people’s attention, so that as they hurried past the screen they would stop after accidentally recognizing themselves or a colleague. But, as so often happens, the real employees were not exactly burning to perform on camera.

“AI solved all of these problems.

“First, using one neural network, I created the visual language—I generated dozens of images in a single graphic style. They became the foundation for the whole look of the project.

“Next, having obtained permission from the employees to use their photographs from social media, I used a second neural network to place their faces into the generated images. That gave me the frames I needed without putting anyone through a shoot. I should say that I could have done this without a neural network too, in a graphics editor. But the network handled it many times faster, and all that was left for me was the polishing. Speed matters.

“After some additional processing, I loaded the resulting images into a third neural network as references for video. What came out were animated frames featuring the company’s employees in the required graphic style.

“All that remained was to edit it, add the motion design elements, and the video series was done. And it required no filming whatsoever, no additional equipment, and no people on set.”

So AI replaced an entire crew of content makers and actors. But you cannot say the AI did all the work—it was only a tool in professional hands. The main driving force behind the whole project was still a human being, one who used

neural networks to unlock his own creative potential. That is exactly the point of working with generative AI.

### Building Assistants

Another example of cheap, effective AI. A partner of mine was in timber logistics—imports. A small team: five people generating $2.3 million in revenue a year. One of the tasks that ate up a lot of time was corresponding with counterparties about contractual obligations.

The solution he came up with was simple. He bought a ChatGPT subscription, created a virtual employee there, and wrote it a system prompt: who it is, what role it plays, what is expected of it. After that, whenever a claim came in or he was preparing one himself, he would upload the contract to this assistant and describe the substance of the claim, his own or the counterparty's. Within a few minutes the AI produced a reply with a detailed argument and references to specific clauses of the contract. Three or four hours of work turned into a matter of minutes.

The price of all this: a subscription costing a few dozen dollars a month, and a couple of hours to build the assistant.

### A Beauty Assistant

We worked on this project in 2025. The idea behind the assistant is simple. A client creates her profile, and the AI does several things:

- It takes the data from the profile and matches it to a suitable salon.
- It shows nail design options on the client's own hand and attaches her choice to the booking sent to the salon—saving time and heading off disputes while the service is being performed.
- It assesses the quality of the service: the gap between what was agreed and what was actually done.

### Recommendations for Small and Midsize Business

For a small or midsize business, AI can become a key resource for success:

- automating and speeding up individual tasks

- automating and speeding up whole business processes (launching a marketing campaign, for example)
- gaining access to the expertise of the best specialists and to recommendations on what to do and how
- creating and generating content for marketing
- making customer interactions more effective by analyzing data in CRM systems, on review sites, and across the internet

In other words, this is a chance not only to win the competitive fight but to grow while keeping investment in headcount to a minimum—revenue and product range growing faster than the team.

A second source of value is quality recommendations without hiring expensive specialists, whether as consultants or on staff.

Below are a number of recommendations on what to do and how.

- Use your flexibility and your ability to take risks. Try different AI solutions, get your hands dirty, learn to automate tasks with prompts. That will give you a feel for how AI works and what to expect from it. And new AI solutions and new combinations of them appear every week. Let AI become part of your life.
- Expect digital advisors to appear in your own field—an advisor on taxes, on project management, on inventory management, and so on.
- Follow accelerators and watch which startups get in. Try their solutions, and if they help, buy or subscribe while the price is still low. Better still, take a stake in the projects that interest you from a practical point of view.
- Study the feature updates to the IT services you already subscribe to. AI tools are being built into almost everything right now. In Miro, for instance, there is a tool where you write out a large task and the AI breaks it down into individual steps, visualizes how they connect, and groups them. Yes, it is just a chatbot—but with convenient visualization.
- AI builders will arrive soon, and if you can digitize your own experience and expertise and pack it into a product, you have a chance to build a new

company. The hard parts are structuring the knowledge into a model, and marketing and promotion. Do not rush into B2B segments. Build B2C solutions. Corporations will have a long list of requirements, which makes it quite likely you will fail to carry the deployment through—or that the price will rise until the project loses all its point. The B2C market and other small companies are more open to off-the-shelf products. Become the standard in the market, and then go into large B2B. The main constraint is that the market and the people in it do not yet understand that they need this product.

- Use AI-solution marketplaces to find tools for automating tasks.

Bitrix, the Russian vendor, is rolling out an AI assistant of its own as well.

Prompt engineers have standardized the typical tasks and built prompts for ready-made scenarios.

In June 2025 the developer also announced that AI agents were coming to the service: you will be able to ask them a question, hand them a task, or talk something over. The agents will be able to connect to the company's knowledge base and find the information they need on their own. Users will be able to change the agents' parameters and configure which employees have access to them. The digital assistants can be called up seamlessly from a dedicated section of the service's interface.

Many of these solutions will be free, at least at the start—particularly if you are comfortable with software running on the developer's servers under a subscription model (SaaS).

Yandex B2B Tech published a fresh read on the market in March 2026: of the more than 7,500 AI agents on the Yandex AI Studio platform, 86% are used by small and midsize businesses in Russia, in scenarios ranging from technical support and HR consultations to matching units for property developers. There is a detailed discussion of this data in the Postscript.

This confirms the key forecast I set out earlier: SMB companies are more flexible and more focused on efficiency. They are not waiting for the "perfect AI"—they take ready-made tools and adapt them to their own tasks. Recall the

MIT data: pilots where a company took an external tool and customized it reached deployment roughly twice as often as in-house builds.

Which leads to a recommendation for developers of AI solutions: build off-the-shelf products for SMBs with a monthly price tag of up to $1,200; $230 to $600 is the sweet spot. That is where the bulk of the demand sits, that is where the appetite for experiment is highest, and that is where bureaucratic inertia is lowest.

Step-by-step plans for launching AI in a small business—the first 30, 60, and 90 days, with checklists and budgets—are set out in AI-2. Practical guide (Chapter 17). And the main compass for choosing your path has already been named above: buy and subscribe—do not build.

Out of curiosity I also asked the models themselves for recommendations. Their advice—define your goals, start small, train your people, secure your data, measure the result—repeats almost exactly what is set out above. Which is telling: the basic management logic of adoption is already baked into the models. The only question is the discipline of applying it.

### Chapter Summary

- The barrier to entry for small business has fallen radically: ready-made cloud services and off-the-shelf AI products give you access to the technology without developers or servers.
- The winners are narrow, everyday scenarios—documents, customers, content, routine work—not "AI in general."
- The human stays in the loop: a business cannot afford the mistakes of a fully autonomous agent, and controlled autonomy is cheaper than fixing the consequences.

**What to do:** Pick one recurring pain point—handling requests, documents, content—and hand it to an off-the-shelf service. A one-week pilot, no budget.

**Reflection question:** What operation do you or your staff do by hand every day, even though AI can already do it?

## Chapter 20. Examples and Recommendations for Midsize and Large Companies

Ten cases in a row is a lot, so let me start with a map. Find your industry or your function, go straight to the case you need, and read the rest as a casebook of other people's experience.

| Case | Industry / function | Technologies | What to look for |
|---|---|---|---|
| Deal-making (Bain & Co) | B2B sales | Generative AI | How AI speeds up deal preparation and proposals |
| China's railways | Transport, infrastructure | IoT + big data + AI | Sensors, data, and AI wired together at national scale |
| Power distribution | Energy | AI in operations control | How to turn a showcase project into a meaningful one |
| Nornickel | Mining and metals | Machine vision, ML | An industrial giant's portfolio of AI tracks |
| EuroChem | Chemicals | AI portfolio | Vision and a systems approach to building out AI |
| Ore processing | Metals | Machine vision | Maintenance automation: from oversize ore to downtime |
| IT support | IT, service | LLM assistants | The economics of the first support line |
| Coffeemania | Restaurants, retail | Generative AI | Where to start: strategy, not technology |
| Oil production | Energy, projects | LLMs in planning | The limits: where AI is still weak |
| HR and recruiting | Cross-industry | Multi-agent systems | An answer to the talent shortage |

### Generative AI in Deal-Making

Bain & Company, one of the three largest and most prestigious management consulting firms in the world, published an interesting study. The other two on that list are Boston Consulting Group (BCG) and McKinsey & Company. To give you a sense of their scale, here is a side-by-side comparison.

| Company | Headcount | Revenue ($ billion) | Revenue per employee ($ thousand) | Cities with offices | Headquarters |
|---|---|---|---|---|---|
| McKinsey & Company | 45,000 (2023) | 16 (2023) | 355 | 133 | New York |
| Boston Consulting Group | 30,000 (2022) | 12.3 (2023) | 440 | 100+ | Boston |
| Bain & Company | 15,000 (2022) | 5.8 (2021) | 387 | 65 | Boston |

According to the study, companies are starting to use AI for data verification and legal due diligence in M&A deals. Of the 300 respondents, 80% plan to bring AI into their decision-making within the next three years. And 16% of the M&A professionals surveyed said they had already used AI on past deals.

Generative AI can surface M&A targets that traditional tools would miss, flag anomalies in contracts, and help teams focus on the problem areas, the Bain & Co report says. The early adopters of generative AI in deal-making are technology, healthcare, and financial companies. “As a rule, these are large companies with moderate M&A activity—three to five deals a year,” the report states.

### Industry and Infrastructure

#### Big Data + IoT + AI on China’s railways

The combination of the Internet of Things, big data, and AI is one of the most promising uses of AI there is. I made that point back in 2021, in DT-1. Immersion.

Our Chinese colleagues have given us an excellent example.

Maintaining rolling stock and rail infrastructure is hard work and extremely expensive. Railways are potentially money-losing enterprises everywhere in the world, especially once the equipment starts to age. Meanwhile China's rail network keeps growing, with a plan to connect every city of more than 500,000 people by high-speed line. Train speeds are rising too. Which makes people an increasingly dangerous component of that whole high-speed system.

So in 2022 the state-owned China State Railway Group began building out data management processes. As we saw earlier, managing data quality—making data complete and reliable—is one of the key tasks on the road to AI.

The main constraint was that access to the data had to be protected against outside interference and leaks.

Sensors were installed on infrastructure and track, on wheelsets and cars. The main measurements were vibration amplitude and frequency, acceleration force, and so on.

Conventional signaling automation kept running alongside. The volume of data collected reached 200 terabytes—and this was text and numbers, no images or video. In other words, genuine big data, the kind a human cannot process in real time. The trials covered the country's entire high-speed network.

An AI model was trained on the collected data, and in 2023 the trials began. In real time, the AI analyzed and processed the data from the IoT sensors and warned maintenance crews about abnormal conditions forty minutes before they developed, with up to 95% accuracy. The network, by the way, is not small: 45,000 kilometers of high-speed lines (SCMP, citing a peer-reviewed paper in the journal China Railway). The recommendations were aimed at preventing faults and correcting them early. In other words, this is MLAD—machine learning anomaly detection, the early identification of anomalies using machine learning.

The data center itself was in Beijing. By the end of the trials, a railway that had been in service for years was performing better than a new one. Over the year of testing, the number of speed-restriction warnings caused by track

irregularities dropped to zero, and minor track defects fell by 80%. And it was not only reliability that improved but comfort as well: the ride became smoother in high winds and on bridges, the amplitude of train oscillations went down, and the load on track and infrastructure decreased.

### AI in power distribution

This case comes out of my conversations with colleagues in power distribution. The question started out simple: “How can we use AI to manage street lighting? The project is a showcase, and we need to show senior management that we can do cutting-edge technology.”

The question gave me pause at first. Showcase and political projects are fine—they help you get resources for the work that actually matters. But they are also very risky, because if there is nothing inside but PR, with no real applied problem being solved, you can lose everything.

The deeper I got into the task, though, the more interesting it became. In the end we found two practical applications.

#### Voice-driven control

A dispatcher’s job involves a great many switching operations and a lot of analysis. Yes, everything is moving toward automating standard scenarios, but someone still has to choose the right one. And SCADA interfaces, though they keep improving, are still nowhere near well-optimized user scenarios and screen design (UX/UI).

This is where AI can work as a voice assistant.

Say the dispatcher gives a spoken command: “De-energize Street A for emergency repairs.” The AI then:

- recognizes the command
- analyzes the request against the knowledge base
- identifies the appropriate scenario, or works out which sections need to be de-energized

- turns all of that into a set of commands and displays a dialog box asking whether it understood correctly and whether these particular actions should be carried out

Then the dispatcher either confirms the task and accepts the recommendation, or rejects it and enters everything manually (in which case the system remembers that command and the list of actions and adds them to the knowledge base).

Beyond the showcase value, this can have practical value too. Roughly speaking, if the operator currently needs twenty clicks to launch a scenario, now it is one spoken command plus one click to confirm. Which means a more productive dispatcher.

On top of that, we are not letting the AI control the infrastructure directly—the final action stays with a human.

### Outage alerts and automatic work orders for repair crews

Our expectations of quality of life keep rising. We want reliable street lighting in our cities, not months of waiting for someone to replace the dead lamp in the middle of the street, over the pothole. And yet you cannot hang a sensor on every single lamp to track whether it works. Hence the second use of AI: machine vision.

You do not need a sensor on every lamp; one camera covering an entire street is enough. The AI recognizes that one, two, or three lamps have suddenly gone out, generates an alert for the dispatcher, and creates a repair request for the crew on duty, with the location and the number of failed lamps filled in automatically.

Once again we simplify the process, take load off the duty operator, and stop waiting for the angry call to the service line.

### Summary

No, these are not the projects that solve the most urgent problems. You could live without them. But they combine a trending technology with social

infrastructure, they raise productivity, and they reduce the risk of human error.

### AI at Nornickel

For the chance to tell this story I want to thank Vertical of Innovations and Alexey Testin, director of the Center for Digital Technology Development.

Below are the main areas the company chose for applying digital technologies.

- Machine learning—predictive recommendations for operators on the shop floor.
- Computer vision—precise equipment positioning and automatic tallying.
- Large language models—a digital assistant for operators built on the knowledge base.
- Physicochemical process modeling—simulating production processes and identifying the critical factors, processes, and metrics.
- Big data and data analytics—analyzing and optimizing formulations.
- 3D printing—printing spare parts: wheels for mill-lining pumps, armor discs, and the like.

And over the 2024–2027 window the company will push hard on generative AI and roll out predictive AI across the entire value chain.

The areas and uses for generative AI are listed below.

*Management LLM—suggestions based on scenario analysis*

- Decision support that accounts for production and market constraints as well as company policy.
- Less effort spent on everyday tasks—from summarizing meetings to drafting emails and presentations.

*Process engineer's LLM—more accurate and faster decisions*

- Freeing the process engineer from routine paperwork.
- Guidance and recommendations on equipment operating modes to support decisions.

*LLM for support functions—higher staff productivity.*

- A 10–20% increase in support staff productivity.
- Better hiring quality and lower turnover (HR).
- A lighter load of routine tasks and faster processes in the tax, legal, finance, and accounting functions.

*LLM for maintenance functions—less exposure to the human factor.*

- More effective equipment servicing.
- Less dependence on individual skill levels and on the few people who hold the knowledge.

The areas for predictive AI are listed below as well.

- Self-service AI (tools people can use on their own)—AutoML and smart dashboards on the digital platform.
- End-to-end optimization—extending AI to every processing stage and optimizing processes end to end.
- Sales efficiency AI—managing the sales chain and pricing.
- CAPEX AI—cutting the cost and duration of construction.

Of the projects already delivered, our colleagues walked me through one: the tax department built a digital advisor on top of an LLM.

The project was a response to the following problems:

- A high volume of requests to the group's methodology functions—department staff spent a lot of time combing through legislation and case law to prepare guidance for company employees.
- Response times could be slow at peak periods, or the quality and completeness of answers to employees' questions could suffer.
- Many questions repeat and do not require highly qualified tax specialists to answer.

The assistant was built as follows:

- automatic search and aggregation of an answer from various sources and documents, with the answer citing the sources, the documents, and the specific passages within them
- a convenient search box in a web interface for users from other departments
- operation inside the corporate network, with no access to the outside internet
- when the documents contain nothing on the question, the system replies that there is no such information in the database

This is exactly the approach to working with LLMs that I described in the section on digital advisors. It is becoming the standard now—safe, cheap, and effective on real problems.

**Since the second edition this case has acquired numbers, and they are a pleasure to report.** Practice is confirming the plans: for 2023, the economic effect of AI at Nornickel reached $100 million; at the Bystrinsky mining and processing plant the rollout added 2.3% to throughput and 1% to metal recovery; industrial AI now covers twelve processing stages at seven sites; and by 2030 the company expects a cumulative effect of roughly $580 million (RBC Trends; company materials).

### AI at EuroChem

EuroChem is one of the giants of the Russian market. The company has not only hands-on experience with AI but also a clear view of where it is going next.

The material below comes from a talk by Vyacheslav Kozitsin, head of AI at Digital Technologies and Platforms.

The main areas identified for AI in manufacturing were:

- design—making new product development more efficient, automating supplier evaluation and the analysis of requirements for parts and components

- production—improving processes (recommendations), automating assembly lines, reducing errors, and shortening raw-material delivery times and the downtime caused by shortages
- sales and production planning—forecasting sales volumes and setting prices
- logistics—optimizing routes, planning fleet resource demand, and raising the readiness and work quality of service engineers

The two areas the company focused on above all were process optimization and equipment diagnostics.

1. Process optimization:

- optimizing equipment operating modes to increase output and extend equipment life
- optimizing production process parameters—component ratios, for instance
- controlling and assuring product quality at various stages of the process
- optimizing logistics and supply chains, including safety oversight in warehouse logistics
- optimizing the layout of equipment across the production floor

The innovations already in place have raised the company's output of weak nitric acid by 2.5%, of ammonia and urea by 1.5%, and of potash salt by 0.8%. And this is only the beginning, of course.

2. Equipment diagnostics:

- monitoring operation and detecting faults or deviations early
- localizing the deviation
- identifying root causes
- determining remaining useful life in order to move to condition-based maintenance

As for generative AI (digital assistants), our colleagues see it developing along these lines:

- digital assistants (question-and-answer systems, or the chatbots we all know)
- analyzing and generating documentation
- multi-agent systems (as in our example of the digital advisor for project management, where several specialized models work in combination)

Among the question-and-answer systems, a few stand out:

- onboarding new specialists: “What days do I get paid?” “Which bus takes me to Building 8 at Plant 2?” “What is the daily allowance for a business trip?”
- supporting specialists in their own areas: “What are the equipment purchase limits for plant directors?” “How do I get an urgent request approved for Plant 3?” “How do I submit an information request?”
- and of course support during meetings: “How many shutdowns did Shop 2 have over the month?” “What was Shop 2’s longest shutdown in 2022?” “How many tons of Product 4 did Plant X lose to Shop 5 shutdowns in 2024?”

As we already know from this book, for assistants like these to work, all that data has to be digitized and structured. The AI needs access to the databases and has to know how to work with what is in them. Which in turn means building internal processes and understanding what data you might need in the first place. So AI adoption has to start with processes and data work, including building data models and assuring data quality.

**And here, too, practice has caught up with the plans.** The program is growing: recommender systems at NAK Azot delivered $3.1 million in benefits in 2024 (up to 1% more output, 0.8% less natural gas consumed), the solutions have been replicated at Nevinnomyssky Azot, and the company puts the total annual effect of its digital advisors at roughly $4 million (CNews, ComNews—February 2025).

### AI in ore processing

As of 2024, the most widespread use of AI in manufacturing is machine vision.

Below is the portfolio of machine-vision solutions that one of the Severstal group companies built for maintenance automation.

- Detecting oversize ore.

When oversize ore jams the crusher, the result is hours of downtime, which can bring the whole plant to a stop.

The algorithm is simple: capture video of the ore, process the images, analyze ore size, and notify the crusher operator.

The whole project needed one server and four NVIDIA Tesla T4 16GB GPUs.

As you can see, the entire project paid for itself the first time it produced a warning and prevented an emergency repair and a plant-wide shutdown.

- Finding uncrushable material.

Uncrushable objects travel toward the crusher along with the ore. The neural network spots and classifies them on the conveyor belt in the video stream, and the system alerts the operator immediately so they can decide what to do.

- Analyzing the particle size distribution of the ore.
- Monitoring driver fatigue.
- Monitoring defects in pallet cars.

The system analyzes the video feed and identifies standard defects: a missing small side board; a missing running roller; a missing running roller cover; a poorly fastened running roller cover; a missing sidewall bolt; a missing running roller cover bolt; a defective mounting of the longitudinal seal plate.

As you can see, this still calls for preparatory work and a decision about which defects are critical.

- Monitoring excavator bucket teeth and their wear.

The system determines whether the bucket tooth points are present or missing and notifies the excavator operator. It also analyzes the degree of wear and prepares recommendations on whether servicing or replacement is needed.

- Monitoring conveyor belt misalignment.

The system analyzes belt misalignment and, depending on the readings, classifies the deviation and generates alerts.

## Customer-Facing Services

### AI for IT support

Expectations of AI right now are enormous. And often overheated. One of the standing questions in this industry is how to cut IT costs. The first target is usually the IT units that handle technical support. But it is not only about reducing headcount—it is also about improving the quality of the support itself.

In short, one of the ideas rattling around in leaders' heads is to put AI on the tier-zero support line so that every incoming request is classified correctly and routed automatically to the right specialist. Put simply, automate tier zero. How feasible is that?

The short answer: AI can be used as a chatbot to classify a request correctly or to solve a standard user problem. That is, to cut the number of support tickets (by roughly 40%), the chatbot can ask clarifying questions, establish all the "symptoms" of the fault, and generate a ticket for the right specialist based on that analysis. The chatbot here humanizes the process, so the user is not dealing with a faceless form full of checkboxes in a self-service portal.

Now let us look at the situation in more detail.

Take a typical corporation with a developed IT infrastructure and its own support desk. 95% of requests to that desk will come in by email. Yes—hardly anyone uses self-service portals or their own account pages (5–10% of employees). It is one thing to fire off "HELP! The printer won't print!" by email, and quite another to go to the portal, log in, choose a request type, check the boxes... People are lazy creatures and use whatever tool is most convenient, avoiding any extra motion.

And 80% of users write their messages as though they were paying for every character out of their own pocket, or had only just learned to speak. The typical request looks like this: “1C is down.” Teaching everyone to write a request using 5W1H (the six questions for pinpointing a problem in lean manufacturing) is as utopian as making people use the self-service portal.

But suppose people did learn to write requests, and everything got done and closed on time—suppose our support specialists were model employees. Even then, routing tasks straight from email would take a great deal of methodical, expensive work. Roughly two to three years and more than $230,000: you have to clean, normalize, and label the ITIL database, then hand-label how tasks were distributed and assigned across two to ten thousand requests—and then, maybe, the neural network will manage to understand a request from a single email. I do not put much faith in that scenario.

Another huge limitation is that every company has its own menagerie of IT systems, its own org structure, and its own distribution of authority. On top of that, most companies have no clearly assigned areas of responsibility at all. And where they do, it all turns into bureaucracy, with the issue solved partially rather than end to end (to handle one user scenario—issuing and setting up a new PC, say—you have to file five to ten separate requests).

The upshot: there are no off-the-shelf products for sorting incoming tasks by email, and there will not be.

What the AI will most likely learn from training is which user tends to send which problem—and it may well turn out that 90% of requests are standard ones that could be solved by reading the manual and should never reach support at all. But you do not need to develop and deploy AI for that. Better to do the analysis, identify the standard problems, and optimize the scripts and the staffing structure.

Within that approach an AI chatbot helps a great deal—one with a knowledge base of the company’s IT solutions, the user instructions, and a description of the org structure (who is responsible for what). An AI bot that either solves the standard problem or writes up and packages the task properly for support.

The main question here is how to set up interaction with the user, and through which channel. It is a hard question from an information security standpoint. Typical corporate security departments are currently lowering an iron curtain—everything inside, nothing from outside. It works badly, but security will still not approve any mobile app, let alone a Telegram bot or link-based access. Instead they will build access from the corporate network (and then wrestle with the remote access problem).

Another question is how the system will determine who exactly is talking to it. Moving the whole infrastructure to domain accounts is not quick. It is a painful path. All in all, working this puzzle out is a good workout for the mind.

The critical task is preparing the IT knowledge base. You have to describe every IT system, collect all the standard problems, train the chatbot on them, and provide ongoing support for the chatbot through the first year.

And in practice this is one of the most sought-after areas. In my view it will become a standard project for IT departments—a combination of assistant and agent with a low cost of error and a high payoff.

### Summary

By and large, AI can be a great tool that takes the load off support and improves service quality. But it is definitely not a magic bullet. You still have to build the processes and put your house in order; AI will not do that for your managers. So we come back once again to mapping the customer journey and the standard scenarios, describing the org structure with clear areas of responsibility, tuning the processes, setting up communication channels, and so on.

## AI in food service

This is an example of where and how the restaurant chain Coffeemania started deploying AI. When we talked, what struck me most was how deliberate their approach was and how strategically they thought. On the question of where to start, they built on three foundations:

- business strategy
- IT development strategy

- data strategy

Out of those come the projects that deliver business value, and out of those a road map for AI. In the end the company focused on four uses of AI in the first phase.

First, normalizing master data. Second, employee training. Third, data analytics. Fourth, personalized recommendations for guests. Let us go through them.

### Normalizing master data

Working with master data is a headache for any company that outgrows the small-business stage. Here the company decided to use AI for the following tasks:

- finding duplicates in the reference books
- matching data across different tables and sources
- validating data (checking data of various types for correctness and fitness for a specific use)
- classifying, structuring, and grouping the item catalog

### Staff training

Here the company found another interesting use for AI. A voice AI assistant helps new staff learn the menu and the restaurant's standards. Among other things, it generates questions tailored to the specifics of the restaurant, processes employees' answers in several languages, and checks whether the new hire is ready to work on their own.

### Data analytics

This is one of those areas that is becoming more sought-after in every industry. An LLM-based AI makes an excellent data analyst—one that works around the clock and works fast. The company uses AI to write database queries and build analytical dashboards, or to prepare answers in real time and in natural language.

### Personalizing offers for guests

One of the overarching goals of Industry 4.0 is customizing customer service. Here the team started using AI to suggest dishes from the menu based on a guest's preferences—the desired balance of protein, fat, and carbohydrates, a diet, a mood, or a budget. AI also helps upsell dishes and match a server to a guest.

And in practice something interesting turned up: the staff mastered the chatbot and stopped thinking. They simply feed the AI what the guest wants and relay whatever the AI answers.

And that is only the first phase. Next comes using AI for purchasing plans, menu optimization, and so on.

## Planning and People

### AI in project planning for oil production

This example grew out of a conversation in mid-2025 with a friend of mine who works at one of Russia's largest oil companies. Here is what he told me.

"I wanted to put together an example for management showing that AI still cannot account for all the subtleties of complex exploration projects.

"After reading the answer, I was left thinking that maybe it is time to learn to do something with my hands, because brains will not be needed much longer.

"Here was my prompt: 'Hi! I am planning to drill an exploratory oil well. The drilling site is in an autonomous district; equipment and materials can be delivered from the mainland only by water, during the summer shipping season, to the port, and from there along the winter road to the well. Please lay out the main stages of preparation for drilling starting from July 2025 and determine the date when drilling can begin.'

"And here is the answer I got from the AI: '… (author's note—*I will omit the details of the answer, but it was a sound plan with a schedule and a critical path, risks, and recommendations; I will give you just the conclusion*)

"'Conclusion: If the schedule is followed strictly, particularly for the sea deliveries in July–September 2025, and the winter logistics operations are completed successfully, a realistic start date for drilling the exploratory well is

February 10–15, 2026. That date provides the necessary safety buffer before the spring thaw.'

"And get this—this year we spudded a comparable well on February 12. We spent six months aligning the plan for our well across two subsidiaries, and the AI produced this in twenty-one seconds."

### Generative AI and multi-agent systems in HR

One of the pressing challenges for business is the shortage of qualified people. Various studies show that up to 40% of executives name the "talent famine" as one of the key constraints on company growth. In my training sessions this topic sparks heated discussion. The hiring problem is not only that "there are no people"—it is also that assessing a candidate's expertise correctly is hard, given how many new specialties keep appearing.

As AI develops, a new area is taking shape fast: automating candidate sourcing and assessment with generative AI and multi-agent systems.

Below I will share the hands-on experience of a Russian company. The team there built an IT platform for staffing temporary assignments and projects. Its users include SIBUR, Severstal, ALROSA, Rosseti, and CROC. I have picked up projects through it myself. One of the platform's key components is AI agents that automate sourcing, screening, and ranking specialists against business needs. It is now integrated with the job board hh.ru, with internal databases, and with corporate applicant tracking systems (ATS), which makes it possible to process thousands of résumés in minutes.

So how does the interaction with AI work inside, and how is the multi-agent approach implemented?

The first agent decomposes unstructured hiring requests and generates a formalized job description covering skills, requirements, conditions, stages, and expected cost.

The second searches the candidate database, builds the initial (long list) and target (short list) funnels, and sends out invitations.

The third performs an in-depth assessment of résumés.

The hardest part here is building the model that scores résumés. On what criteria? How do you score them? Thinking it through, the team arrived at a structure of three blocks.

- Direct: fit against the job requirements (a checklist of skills, experience, level of command of tools and methodologies).
- Rules: analysis of work history, identification of problem areas (frequent job changes, missing key competencies, and the like).
- Competencies: identification of domain and soft skills, a growth forecast, and a reasoned assessment of the potential to pick up adjacent areas quickly.

So the candidate's résumé goes to the AI, which compares it against the job description and breaks it into semantic blocks. The AI then computes a probability for each component, assigns an overall score, writes a detailed rationale for selecting (or rejecting) the candidate, and delivers the whole thing as a report for the manager.

An example: a candidate is ranked on experience (eleven years in IT), fit against the basic requirements, and the ability to pick up the skills needed (1C configurations, say), with an explanation of strengths and weaknesses.

The results came out as follows:

- the full processing cycle for one résumé takes ten seconds versus ten minutes for a human expert (a 60× speedup), and ranking 100 résumés takes under ten minutes
- the final score matches the recruiter's decision 87% of the time, at any sample size
- recruiting cost savings of up to 10–15% (for example, at a typical volume of 5,000 résumés a month, that is $2,900)
- time-to-hire down by 20–30%
- availability around the clock, with no burnout effect and no bias

- the share of recruiters' manual work at the preparatory stages down by up to 80%

There are obstacles too, of course, and we have discussed them before.

- Corporate security policy—many companies require on-premises deployment and deep de-identification.
- High-volume hiring cases need their own scenario for communicating with candidates (drivers or plumbers, for instance); deep screening is not always what you want.
- The business logic of integrations with external ATS and CRM systems depends on the maturity of the corporate client's IT landscape.
- Models have to be retrained continually for new markets, tasks, and languages.

Here is what I would note. What this AI shows is that, at a certain level of business maturity and readiness, multi-agent systems can radically improve both the quality and the speed of processes, including the sourcing and assessment of people at large companies in Russia. In this case we got not just faster hiring but also impartial assessment, which matters most when you are looking for rare specialists whom line managers cannot evaluate correctly.

For maximum effect, run pilots on key roles, involve the product and HR teams in development, and collect user feedback.

AI in HR should also develop through integrations with other systems, with flexible configuration and support for both cloud and on-premises infrastructure. And of course de-identification methodology and compliance with personal data security requirements play an important role.

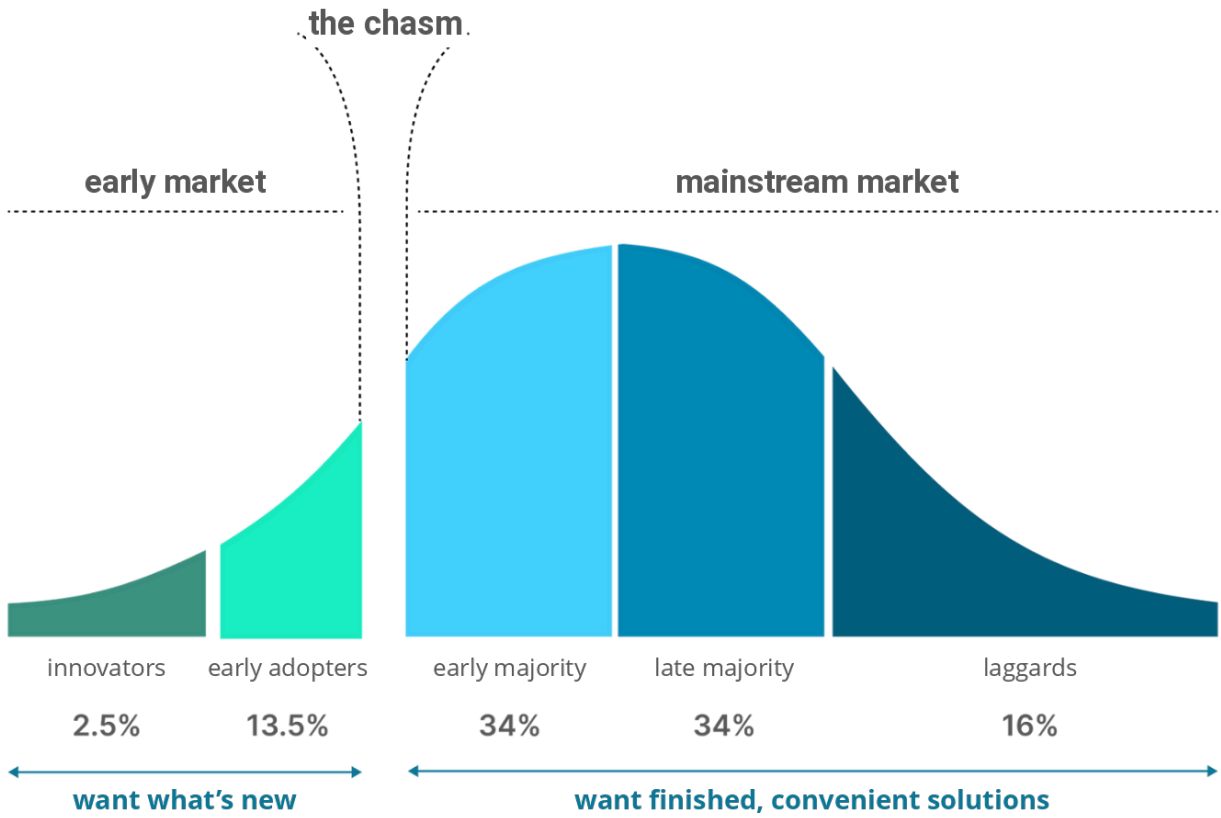
Another interesting practical example is the animation studio Petersburg (part of the Riki group), which built an AI agent on Yandex AI Studio to onboard new employees. The agent answers questions like "How do I draw up a contract with a licensee?" or "How do I create a new animation scene?"—and the answers come with images, not just text instructions. According to the

company's CTO, this cut the frequency of support requests substantially and made onboarding easier.

It is notable that solutions like this are built not on custom models but on ready-made platforms with low-code tools and connections to external systems through MCP servers (Model Context Protocol—an open standard that lets an AI agent connect to external data and tools without building a separate integration for each system). By March 2026, more than a thousand such servers had been created on Yandex AI Studio alone, letting agents work with amoCRM, Bitrix24, Kontur.Focus, and other familiar business tools.

### Recommendations for Large Business

- Single out dedicated innovation leaders (in the sense of Moore's adoption curve) who are willing to experiment and deploy AI solutions. There are always such people in a company; they are 10–15% of the workforce. They are willing to forgive the mistakes that come with new tools, to be the "test subjects," and later to become the conduits of change.



- Build IT sandboxes where you can deploy AI models and test them. Among other things, this helps you reach an understanding with your security specialists and creates the conditions for testing ideas quickly.
- Talk through with information security the option of using off-the-shelf AI products and SaaS solutions (with nothing hosted on the internal network), at least for pilots and research. That brings the cost of those pilots down.

Focus, too, on small and midsize language models that can be deployed on internal infrastructure and integrated with corporate services.

- Build project management processes and rules for launching pilots—this is the key task.
- Set up academies for your innovators and teach them the core tools of the systems approach.

Pay particular attention to project and product management, lean manufacturing, communication management, and the theory of constraints.

- Make AI a strategic priority for the company: put it into the development and digitalization strategy, into OKRs, and into executives' KPIs.
- Take part in hackathons run by educational institutions and accelerators, and form partnerships with research organizations. In short, let the young and flexible work on your problems. Watch the market as well, and take equity stakes in young startups. That way you can adapt solutions to your own needs at minimal cost—and if the startup succeeds, make money later by selling your stake.
- Be ready to implement off-the-shelf solutions, especially where they come with ready-made methodology inside. Do not be afraid to simplify your processes—experience shows that most of them are outdated and the room for improvement is enormous. At the very least, put together an inventory of business processes with their owners, participants, and problems.

- Start working on data quality and on building data quality management processes. This is not only about master data but about all of your data. Without data, no AI solution—least of all a recommender system—will do you any good.
- Give your innovators the chance to attend conferences. Better still, make it a mandatory part of the job. Especially in industries outside your own. Disruptive innovations and the most interesting examples are always found somewhere else. And a broad, well-informed eye for digital technology is an essential ingredient of success. I singled this point out separately in my own competency matrix.
- Deploy AI first in the processes where the cost of an error is not critical but the payoff will be visible. And prefer narrowly specialized AI solutions—their recommendations are more accurate and safer. But if you are putting AI into industrial control systems, leave the final decision with a human, and use the record of whether people agreed or disagreed with the AI's recommendations to train it further.
- AI will not replace your people, but it will help them become more effective. Which means you either ride out the talent shortage or hold headcount down as the organization grows (rather than inflating individual departments with ordinary specialists).
- Take a comprehensive approach and use AI as a driver of organizational development.
- I can name several areas that will be in demand in any industry and will most likely appear as separate products and solutions on the market.
- Advisors of various kinds: procurement, legal, project management, and so on.
- AI data analysts.

This is an LLM that connects to your databases and works as a data analyst: it can build reports and charts, produce briefs and analytical extracts, make forecasts in real time—in short, be an assistant to leaders.

- HR AI

This is an LLM that helps with recruiting (selecting relevant candidates) and with supporting new hires through their early stages.

- AI for IT support

An assistant that either solves the issue or asks clarifying questions until it can produce a properly written ticket. There is real demand for this from the business. Because otherwise you either have unhappy employees or an inflated IT department.

- Preparing project documentation.

IT projects have a big problem with project documentation. It is labor-intensive and slow. AI, though, will be able to produce that documentation quickly by analyzing databases and code. Projects like this are already appearing.

**The flip side: more documentation is needed than before (when AI is used to build IT solutions).** AI that writes documentation has a mirror property: it also consumes it—and consumes it differently from a person. A person who hits a gap in an assignment stops and asks. A model, by default, does not ask; it fills the gap with something plausible. It will ask only if it has been told to (in Chapter 6 that is Clarification First), and even then only about the gaps it can see itself.

The practical consequence for AI projects: what you document for AI is not the same as what you document for a person. For a person, you describe what to do. For an agent, you also have to describe what not to do, in which modes the solution has to work, and how we will know an item is closed. These are exactly the things that the questions for the plan in Chapter 6 pull out one at a time. Documentation is a way not to ask those questions afresh every time.

And about the fashion for "vibe coding," where a solution is assembled through dialogue with the model with no project artifacts. It does not do away with documentation—it moves it from "after" to "before": what used to be written up after the fact now has to be formulated at the input, otherwise the agent will fill it in for you and will not ask, and the model will start forgetting

and skipping. I went through this myself with my team while building our products.

You can also expect off-the-shelf solutions for master data work (I wrote about this earlier, and Coffeemania is already an example), AI twins of employees carrying their experience and expertise, and AI communication assistants that connect to email and messengers and handle correspondence on an employee's behalf, in their style.

- Focus on products, not models. Among other things, every AI solution needs request logging, so that you can analyze the requests and improve your internal procedures and documents.
- I would also recommend a hybrid approach to deploying AI.

When to choose on-premises deployment (on your internal infrastructure):

- Processing confidential data
- GDPR and personal data compliance requirements
- Steady heavy load
- A need to customize the model to your business processes
- A desire to avoid vendor lock-in

When to use cloud APIs:

- Irregular or light load
- A need for the most advanced capabilities and complex reasoning
- No in-house IT expertise to support on-premises infrastructure
- A fast project start with no capital investment

What we end up with is a hybrid approach: AI models on internal infrastructure for basic tasks and confidential data, and cloud APIs for the complex tasks that demand the highest quality.

Cloud models should also be reached through an intermediate layer. Before a request goes to the large language model, a separate AI model finds and replaces every piece of sensitive data with an impersonal placeholder. Instead

of John A. Smith, for instance, it writes [CLIENT_NAME]; instead of +1 XXX-XXX-XXXX, it writes [PHONE]. The cloud LLM then works with a fully anonymized request, while your employee gets a correct answer with the real data, produced by the most advanced model available.

**Example before masking**

**Original user request:**

> Draft a formal letter on behalf of **John A. Smith**, director of **TechSphere LLC**,
> to our client **Anna Miller**, who lives at **12 Forest Street, Apt. 45, Springfield**.
> Subject: a delay in equipment delivery under contract **No. 456/2025 dated 03/12/2025**.
> Give my contact phone **+1 (915) 123-4567** and email
> john.smith@techsphere.com.

**Example after masking**

**Version sent to the LLM:**

> Draft a formal letter on behalf of **[SENDER_NAME]**, director of
> **[COMPANY_NAME]**, to our client **[CLIENT_NAME]**, who lives at
> **[CLIENT_ADDRESS]**.
> Subject: a delay in equipment delivery under contract **[CONTRACT_NUMBER]**.
> Give my contact phone **[PHONE]** and email **[EMAIL]**.

- And finally: save your money and do not try to build your own AI teams. You will lose the race to product teams by definition. You are unlikely to be ready to pay what they pay, which means you will attract mediocre specialists. And the ones who do grow, you will most likely not be able to keep. What matters for you is accumulating expertise as a business customer, not as a developer. Software development and model selection are not your core business.

Detailed implementation road maps for midsize and large businesses—team composition, budgets, phase plans, and checklists—are in AI-2. Practical guide (Chapters 18–19). What stays here are the cases and the management principles.

And out of curiosity I again asked the AI models themselves for recommendations. Their advice—audit your processes, set strategy and priorities, pilot on a limited scope, get data quality right, train your people,

measure the results—repeats almost exactly what is written above—at the level of a good consultant.

As I often say, deploying AI solutions often turns into a bureaucratic disaster purely because companies lack project management competencies. Beyond that, the defining features of large business are complex, inflexible processes and specific requirements from the security functions.

Recommending that a business like that suddenly become agile is utopian, of course, but it is still worth carving out separate units and mechanisms for your innovators: give them an IT environment, define the rules and a project budget you can afford to risk, and keep bureaucracy inside those units to a minimum.

For example, if a project needs up to $6,000 and one department can deliver it on its own, then describe the problem, describe the solution, and go. If it needs $6,000 to $116,000, the list of tasks grows. And if the budget goes above $116,000, it has to go through the full bureaucratic cycle.

In Chapter 22 I will explain how to run projects that bring new technology into your business. And although this book is devoted to artificial intelligence, the playbook for managing a project life cycle applies to any initiative.

### Chapter Summary

- The industry cases all converge on one point: the effect comes from data + IoT + AI built into the process, not from a standalone model.
- Midsize and large businesses win on platform thinking and a portfolio of initiatives; pilots for the sake of pilots are the road to the PoC graveyard.
- The strongest cases are the ones where AI complements the automation already in place (ERP, MES, SCADA) rather than trying to replace it.

**What to do:** Pick two or three cases from this chapter that are analogous to your industry and compare them against your own processes—what infrastructure and data do you already have? (One week.) Then build a portfolio of three to five initiatives with priorities (one quarter).

**Reflection question:** Which case in this chapter is closest to your main operational pain—and what is stopping you from repeating it?

## Chapter 21. AI in the Leader's Own Work

We have looked at how AI works for small business and how it works for corporations. One figure has stayed off camera—the one people usually forget: the leader. Not the company's processes, but their own day. Decisions, communication, planning, fatigue, and the slow drift of focus away from the strategic layer. And a mistake made by this particular person is expensive; sometimes the future of the whole company and of dozens of people rides on their decisions. I have watched it happen more than once: strategic projects that never launched for exactly this reason, and a company that paid for it afterward.

This is the most personal chapter in the book. I will show how I use AI myself—no theory, real examples, including one actual conversation.

### A Leader's Tasks: What You Can Hand to AI

The work of a leader at any level is made up of the same classes of tasks; only the proportions differ. A frontline manager does more monitoring and gives more feedback; a middle manager does more coordinating and runs more projects; a top executive or owner deals with strategy and with decisions where mistakes are expensive. The good news is that AI applies to every one of those classes. The bad news is that nowhere does it replace you entirely. This is the 70/30 rule from Part 1, applied not to the company's processes but to your own working day.

| Class of task | What AI takes on | What stays with you |
|---|---|---|
| Planning the day, the week, the month | Assembling the plan, tying it to your goals, reminding you of loose ends, working through the tasks—and even suggesting who to delegate what to | Choosing the priorities and actually living the plan |
| Decisions | Working the decision from several angles, structuring the problem, finding root causes, offering alternatives and risks | The decision itself, and the responsibility for it |

| | | |
|---|---|---|
| Communication | Preparing for negotiations and meetings, drafting emails, unpacking conflicts, reviewing how things went | Live contact and trust |
| Goal setting | Working goals through, formulating the criteria for reaching them, monitoring, holding the focus across communication, decisions, projects, and planning, and showing progress against goals | Execution, course correction, and the final sign-off |
| Projects | Working projects through, drafting charters and plans, breaking work down to the task level, tracking current status and risks, and wrapping up with lessons learned | Goal setting and agreements |
| Team and oversight | Gathering the facts, reminding you of promises—yours and other people's—and preparing the structure of a one-on-one or a feedback conversation | Judging people and making personnel decisions |
| Self-development | Feedback on your day and your week, spotting patterns, suggesting exercises, prompting reflection | Drawing the conclusions and changing your behavior |

### My Personal Loop: How It Works

Let me describe how it works for me—not as a recommendation, but as one live example.

**The journal and the feedback.** Every evening, a short entry: what happened, what worked, what stuck with me. The AI assistant comes straight back with feedback: where I am repeating myself, where my energy is leaking, what I have been putting off for the third day in a row. Once a week, a digest—progress against goals, the balance between strategic work, relationships, and

health, and where I am overloaded. Once a month, the same thing on a longer horizon. On my own I would never have kept up that discipline: the value is not in the entry, it is in the instant response to it.

> **From My Practice: what an evening entry looks like**
>
> The entry: "Tired. The day was eaten by sign-offs on the integration. Put off the conversation with the head of sales. On the plus side: we closed the pilot at the client."
>
> The assistant's response: "The conversation with the head of sales has now been postponed three times this week—and in every entry it sits next to the word 'tired.' That looks like your main source of background fatigue. The pilot is closed—one step toward quarterly goal No. 2. I suggest booking the conversation for tomorrow morning, fifteen minutes. I can prepare the structure."

**Planning.** I put the week's plan and the day's plan together with the assistant: it sees my goals, my loose ends, and my promises, and it proposes the skeleton—I decide. The difference from a paper planner is that the plan is tied to goals: when a week gets eaten by day-to-day work, I see it within two or three days rather than at the end of the quarter.

**Working through decisions.** Before a serious decision I take it apart with AI in five-whys mode: it asks questions, digs for root causes, offers alternatives I had not considered and risks I would rather not see. The decision stays mine—but it has been stress-tested before I announce it. And the analysis is done against my strategic goals, work and personal alike.

**Projects and documents.** Charters, plans, status reviews, preparation for steering committees—everything that used to take an evening now starts with a draft the AI produces and I refine. It is not only time that gets saved: I no longer burn energy on the blank page.

**Communication.** The most valuable part is preparing for difficult conversations and unpacking the ones that have already happened—or are

happening right now in a messenger thread. That is the subject of the next section, with a live example.

The most valuable change I have noticed in myself: I have struck a balance between work, health, and family, and I plan more precisely. The assistant is now my second brain, always within reach. AI has taken over a huge share of the routine.

The clearest example is the day I had no access to the tool. Planning the day, I looked at my task list and realized I could take it all on right then and it would cost me ten hours. But the tool would be back tomorrow, and then I would do the same work in two or three hours. So I stepped back and spent the day thinking—and came away with two valuable insights for the whole company.

My personal loop is not set up with one prompt but with four levels, and each has its own job. The order runs top-down: from what applies everywhere to what applies in a single conversation.

> ### Level one: the account-wide instruction
>
> This is the "passport" that applies in every conversation, whatever it is about. Its skeleton transfers to any leader. Who I am: profile, experience, context in two or three paragraphs. My roles and the priority between them: in my case, CDTO, product founder, author—plus a rule for what the assistant should do when the role behind a request is unclear (ask in one line, or take the default). Priority classes of tasks: a short list of what to focus on. The format of answers: summary → analysis → artifacts → checklist, with a mandatory "off switch"—a simple question gets a short answer, no template. Rules for clarifying questions: how many to ask, and which assumptions to use if I do not reply. "Argue with me": object with arguments rather than agree—polite agreement does more harm than an argument. And the confidentiality boundaries: what never goes into external texts without an explicit command—internal product figures, people's names, details about my employer. Written once, revised once a quarter.
>
> ### Level two: the shared register

Between the passport and the projects sits a folder called the Project Register. It is a map of the whole loop: the list of projects with their instructions, my own skills and agents, the installed extensions, the working folders and repositories—what lives where, who is responsible for what, and how to restore it on a new computer. The passport refers to it in one line; an agent that works with my files reads it on its own.

The level did not appear out of a love of order. As long as there is one project, an instruction is enough for the assistant. When projects and agents number a dozen, each one is sound on its own, but together they start getting in each other's way: one rewrites a file that another maintains, a third does not know the skill it needs already exists. It is the same failure as the three agents in Chapter 8, only at the scale of one person. The register is their shared map and their rules of the road.

One caveat. This level works where the assistant can see the disk: agents on the computer, work with a connected folder. An ordinary chat in the browser does not need the register—the passport and the project are enough for it. The register is updated not by the calendar but by events: a project appears, an instruction changes, an extension is installed—the same day, or the map starts to lie.

### Level three: a project for a block of tasks—instruction plus files

The journal, the books, the product, teaching, processing improvement proposals—each area has its own project: its own instruction and its own knowledge files, which the assistant consults on its own. Here is how the journal assistant is set up, so you can see how far this can grow. The role: "You are my coach and mentor, working from my daily journal. You help me grow in three capacities: as a person, as a C-level leader, and as the founder of an AI company." All analysis is done against three areas of balance: strategic work / relationships / health. My typical failure is recorded separately: reactive work crowds out strategic work, and work as a whole crowds out health and relationships; the assistant is required to watch for this first.

The thresholds are red lines, and the assistant names them out loud when they are crossed: at least seven hours of sleep on at least five nights a week; at least three workouts; at least one full day of recovery; three evenings of real presence with the family without work devices; two protected blocks of strategic work; reactive work no more than 65% of working time. One threshold crossed is an early signal. Two or more, or one threshold crossed two weeks in a row, and the assistant is required to propose a specific unloading: what to drop, delegate, or postpone. Not "get some rest," but a list.

Two modes. Mentor mode is direct feedback on every entry: observation → balance → insight → recommendation, no longer than 200 words. Coaching mode, on command, is no advice at all: one strong open question per turn. Plus cadence commands: WEEK, MONTH, UNPACK a topic, BALANCE. And the principles: rely only on facts from the entries; call something a "pattern" only after two or three repetitions; mark hypotheses as hypotheses; no motivational platitudes.

The project files hold what the assistant has to keep coming back to and I do not want to repeat: the quarter's goals and the "Snapshots" of past months. One chat is one month; at the end of the month the assistant compiles a Snapshot, and the snapshots later roll up into quarterly and annual reviews. This level is written for the task and revised when the task itself changes.

### Level four: the task in the chat

Here there is no longer any need to explain who I am and how to work with me—the first three levels already know. What remains is the task itself and its context: the day's entry, a decision to work through, a conversation I am preparing for. The daily entry itself is free-form, with a suggested minimum: what I did (strategic or reactive), decisions, relationships, health, what got stuck. Five minutes in the evening is enough. Thanks to this architecture, sometimes I simply send the materials and say in one sentence what I want to get—and the AI does an excellent job.

### Where to start if none of this exists yet

Do not write the instruction yourself—ask the assistant to interview you: "ask me ten questions to build an instruction: who I am, my roles, my tasks, how to work with me, what not to do." It is the same move as in Chapter 6, only now the model is assembling not a request but your passport. Give it the condensed version of my journal project above as a reference—not to copy, but so it can see the level of specificity needed. Half an hour, and level one is done; a second pass builds the project for your first task the same way. You do not need the register at the start—it becomes necessary once projects and agents number a dozen or so.

And the line I keep in the passport and repeat in every project where an argument is possible: "Answer briefly and to the point. Argue with me where I am wrong."

### What this gives you

Three things that one long prompt cannot, and a fourth that the register gives. First, a warm start: every new conversation begins with the assistant already knowing me, my roles, and the rules of the game; I do not spend the first ten minutes on setup. Second, a separation between what changes rarely and what changes often: the passport lives for quarters, the project for as long as the task lives, the chat for a day or a month; changing one does not mean breaking another. Third, portability: the same passport works in the journal, on a book, and in the product, and a project moves between chats together with its files. And fourth, which arrives together with the register—coherence: agents working in parallel know about one another and do not break each other's work. Half an hour on the passport by interview (or an evening if you write it yourself), an hour on the first project—and the quality of every conversation after that rises noticeably: the assistant answers you, not the "average user."

## A side effect I did not expect

I let one leader whose first level I was setting up read my account-wide instruction—the passport itself. A few days later he wrote that he now reads the assistant's answers differently than before—with a kind of relief: he sees

in it an executor of a human's will, whereas without knowing the instruction he had the impression of a more conscious AI. He saw one level out of four—and the "consciousness" disappeared. That is not a disappointment; that is exactly the point worth arriving at: what impresses you in a well-configured assistant is your own will and your own rules, reflected back. All the magic is in the engineering of settings and data flows, not in the model.

Hence my 70/30 split in personal work. Thirty percent is the human: think it through, set the direction and the constraints, control, make the final decision. Seventy is the AI: gather, aggregate, analyze, package ideas into words. The proportion is the same as in the chapter on trends, but here it is about the division of roles within a single working day, not about the labor market. And one more observation about the people I have set up assistants for: almost everyone who lets AI deep into their work fairly quickly begins to understand how the machine is built—and stops being afraid of it. The fear rests on not understanding, not on what the technology can do.

And an important caveat about security. Sensitive material ends up inside the personal loop: employees' names, conflicts, money. A public chatbot is not the place for that kind of data—use an environment where you control the data, or anonymize your entries. We covered this in detail in Chapter 8.

## Three Cases from Practice

### Case one: the conversation that nearly killed an idea

My product manager came to me with an idea: give the users of our product free space to build their own templates and structures, the way Notion does. I closed the subject in a handful of messages. "Not yet. It will turn into a data dump. It will force us onto expensive models, which breaks the unit economics. It may also break the way the agents inside the system work. And we end up as a bad ChatGPT instead of a good BAEOS." On paper I was right on every point. In practice, an initiative was dying in front of me. And if you do that regularly, motivation dies right behind the initiative.

But the conversation did not end there—it went on, message after message. And in parallel, while the exchange was still running, I was taking it apart with my AI assistant. The assistant knows my goals and my role: I do not need to

win the argument; I need to grow a product manager to whom I can eventually hand over product development, or the launch of new products. That is a strategic track for taking work off my own plate over the next two or three years—building a motivated, competent member of the team. So the feedback was not abstract; it was targeted. The assistant highlighted three things. First, my product manager was not proposing what I had heard: he was talking about a horizon two or three stages out in the product's development, and I was answering about this release. We were arguing about different things, and each of us was right inside his own frame. Second, I had closed the idea faster than I had understood it—my standard pattern had kicked in: protect the product first, listen afterward. That is fairly understandable behavior in makers and entrepreneurs who have already built something themselves. Third, and most important: if a person repeatedly gets a "no" to a question they never asked, they stop bringing ideas at all. Bringing ideas is exactly what I hired him for.

So I rebuilt the conversation without leaving it. Instead of a ban, a frame: I named the constraints—do not breed chaos, do not create a data dump, do not burn the budget on tokens—and asked him to work the idea through within them: "Think about how you see it." And then, together, we turned the idea into a different form. Let the product watch where users fail to fit our scenarios, and we will extend the product to match their real behavior instead of handing them a box of parts. One detail matters: authorship stayed with him. I said it outright—this is your idea, put it on the board and develop it. An idea a person owns works; an idea taken away from them turns into someone else's assignment.

Now count the length of the cycle. Without AI it would have gone like this: I kill the idea, the employee goes quiet, and six months later I am puzzled that he never brings anything—and by then I can no longer connect the effect to the cause. With AI the pattern was spotted and rebuilt while the conversation was still going. A conflict that never caught fire; an idea that did not die and became a product principle; a lesson I learned in the moment. And a rule for the future, for every product argument: before you object, ask—"Which stage are you talking about?" The feedback loop closed inside the cycle itself. Why

that matters more than it sounds is something we will come back to at the end of the chapter.

### Case two: a developmental dialogue instead of pressure

The second case is the opposite in tempo: not one conversation but a stretch of work across a month and a half. My lead developer is a strong maker: he carries the product and generates solutions himself. And, like everyone cut from that cloth, he does not accept other people's ideas—he deflects them. In Adizes's typology this is the classic E type, the "entrepreneur": the main engine of change in the team and, at the same time, the most resistant to change coming from anyone else. The problem was simple and painful: without a sustainable routine he kept breaking down—working around the clock, burning out, losing his rhythm.

A direct "you need a routine" does not work here. To a maker, discipline sounds like a cage—an attack on freedom, which is to say on who they are. Pushing is pointless: you get formal agreement and quiet sabotage. So instead of persuasion I built a developmental dialogue—and I ran it together with AI. Before every significant conversation, and after it, I went through the exchange with the assistant: what exactly the person was resisting, where I was about to push and must not, which wording would defuse the conflict and which would stamp someone else's authorship on the idea. The assistant all but held my hand at the points where fifteen years of corporate experience had trained me to push. One rule we wrote down word for word: "If I push, it will only get worse."

The mechanics went like this. For a month and a half I quietly seeded the idea—no demand for a decision, just a thought. He deflected: "not yet," "I will know when." In the decisive conversation I backed off: "All right." No resentment, no subtext. Some time later he came back on his own—"maybe today is the day"—and designed his own working rhythm, better than the one I had proposed. The key move came out of the analysis with AI: do not try to prove that discipline is good for creativity—redefine it. Building your own system is creativity: you do the searching, you do the optimizing, and only you decide whether it worked. And a system you have built frees up resources—

time and energy—for new creative work. The conflict disappeared at the root: accepting the idea no longer meant betraying who he was.

The result showed up a week and a half later, and without a single request from me. A voice message in which he described how he was rebuilding his entire development process: his own goal map, pipelines of steps, automatic "critics." He put the motive into words himself: "I want the fishing rod, not the fish." An idea he had deflected as someone else's for a month and a half had become his own—he is building infrastructure for it and defending it as his own invention. My role came down to two things: the question "What help and support do you need from me?"—and being the fuse that keeps the enthusiasm from eating the core work.

### Case three: a decision, from the idea to the debrief

The third case is not about the team, and not about work at all. In early summer 2026 I sold my car. On paper, an ordinary transaction that moved money around. In practice, a story about a shift in identity that I did not fully see on my own until I took it apart with the assistant.

First, the setup. For more than ten years I built a corporate career and rose to head of digital transformation. That role comes with trappings, and a flagship crossover is one of them—part of the image of the successful salaried executive.

I had spent thirteen years and 172,000 kilometers / 107,000 miles behind the wheel. The car was part of me.

At the same time, I was preparing to launch my own AI products—and they needed the two things that are always in short supply: cash and my focus.

It started with one message to the assistant. "There is a thought going around in my head. Maybe sell the car? I would have money available rather than frozen in metal—very useful in a crisis. On top of that I would cut a line item out of my costs. Public transit and taxis will be three times cheaper in any case.

And the main thing: it is an extra resource I can put into the product—paying for infrastructure or advertising. Launching the products may also require moving to another country.

But inside I feel a little sad. The car is an extension of me, the sense of control at the wheel, my own space.

Help me make the decision."

We went through the rational layer in a couple of sessions: total cost of ownership—every payment, servicing, and depreciation—and then a matrix of three future scenarios.

| Future scenario | Keep the car | Sell |
|---|---|---|
| I stay in employment | Comfort—but the costs eat the cash flow for the products | Free cash flow plus optionality |
| A crisis or a layoff before the products launch | A forced sale at a discount | A cushion for months, a flexible budget for living and for launching the products |
| An international move driven by the product launch | | A managed sale at a normal price, personal flexibility and mobility, no anxiety about a frozen asset |

The conclusion was harsher than I expected: the car won in only one scenario out of three—and only on comfort, which has substitutes. In all the others it made my position structurally worse: I carried the full set of risks and got the benefit in one future out of three. Getting from the first message to a firm decision took seven or eight days.

An ordinary calculator would have stopped there: the numbers add up, sell. But it was what came next that made the assistant most valuable—it would not let me close the subject with arithmetic. Then we took apart, layer by layer, the sadness that came up after the decision. First: this was grief, not doubt. The arguments held, and feeling sad about a chapter that is ending is normal—the only thing you can say goodbye to without sadness is something that never meant anything. Second: the need for control at the wheel turned out not to be about the car. It was about wanting some part of life to be steerable at a moment when a large part of it had gone into uncertainty: the products were not launched yet, my income was being rebuilt. The steering wheel gave instant feedback where the rest of life was giving none.

And the third voice, the loudest: “This is a step down socially.” The analysis made it clear that behind the rational factors sat an emotional core. And if your social standing rests on a car, it is not yours—it is leased from the car. My real capital—the books, the products, the portfolio of projects, the teaching—cannot be read off a parking space. More than that: in the environment I am moving into—founders, product teams, investors—an expensive car reads exactly the other way around. That was when it became clear that the argument inside me was never about transport. The car was the material anchor of a role I was outgrowing, and the ego was making one last bet before parting with the symbol.

A caveat for the skeptics: the assistant was not playing therapist and did not hand down verdicts about my personality. It structured what I was saying out loud myself—and it kept one layer from eating the other: the arithmetic from hiding the emotions, the emotions from canceling the arithmetic.

In the middle of the process, life intervened: my wife asked me to put the sale off until fall, for the sake of a summer together. The plan already had momentum, and there was a pull to get around the obstacle with arguments. Instead, the assistant and I took the request itself apart—and it turned out not to be about the car. Behind it was a fear of losing our time together; the car was only the occasion. My wife and I agreed on the thing that mattered, time for each other, and the sale stopped being a threat. Note what did not happen: I never had to choose between several thousand dollars and something money does not measure, because we took apart the root cause instead of arguing about the symptom.

The deal closed in five days—13% above the second dealer’s offer and 6% above the best offers from private buyers. The whole cycle, from the first thought to handing over the keys, took fourteen days. And after the deal came one last sting: did I sell too cheap? The assistant named the mechanism: seller’s regret, which kicks in precisely because the deal went easily. It was not the price that bothered me—it was the absence of suffering. And suffering is not part of the price of a good deal.

The sadness I had been so afraid of lifted about five days after I handed over the keys. What came next was not in the calculation: confidence. The freed-up resource is already working in my project, and that feels nothing like an asset sitting in a parking space. Plus global mobility: I am no longer tied to an object and can move between countries freely. The new identity settled in through action, not through affirmation. And the main thing—more movement brought progress on my goals: losing weight, which had been stalled for about nine months, and keeping my health up.

And the final check—hypothesis against fact. My spending on getting around fell 75% against what the car cost. The freed-up cash is working in the project. And the suffering my ego had threatened me with never happened at all. The decision was closed by facts, not by a feeling. Beyond that, the sense of global freedom turned out to be stronger than the comfort of being able to get behind the wheel and drive myself wherever I wanted. And I will note separately how much mental capacity was freed up by no longer having to think about other drivers, traffic rules, and the rest of it.

Repeating an analysis like that is easier than it sounds. It starts with three questions. In which future scenarios does this decision win, and in which does it lose? What does this object mean to me beyond its function? How will I feel a month later? My assistant already handles that kind of analysis—see the “Working through decisions” block above.

Now let me bring the three cases together. A short feedback loop works at three scales: a conversation that unfolds in minutes; the development of a person, which ripens over months; and your own decision, from the first thought to the reflection after the deal. And in all three, AI is not valuable for its ready-made answers. It is valuable because it helps you see yourself from the outside while the situation is still alive: where you are pushing, where you are hearing the wrong thing, where you are hiding an emotion under arithmetic. Durable change—in people, in relationships, in yourself—cannot be imposed from outside. What you can do is create the conditions in which it happens on its own. An assistant that knows your goals and remembers your role is the cheapest way I know to create those conditions.

### What a Leader Gains from an Alliance with AI

Before I sort the effects into boxes, let me name the central paradox of this alliance. The assistant is strong because of what it does not have: emotions of its own, fatigue, a personal stake in your decision. It does not take offense at bluntness, it does not wear down by the tenth question, and it has nothing to gain. And yet—if you are running a personal loop—it is attentive to small things and reads your state from the wording itself: where you are angry, where you are hiding doubt under arithmetic. Impartial and attentive at the same time. In people those two qualities almost never come together: the ones close to you are not impartial, and the impartial ones are not close.

**Energy saved.** Working through decisions, projects, and drafts has stopped eating my evenings. It is not only about time: I no longer spend energy on the blank page or on holding everything in my head. Give the context, say what you want and what the constraints are, get a draft and refine it—that is simpler and faster than starting from nothing.

**Focus on the strategic.** By delegating the routine to the team and to AI, I finally do what my role is actually for: strategy, solution architecture, thinking. In my experience the effect is visible in the structure of the day: there is time to hold the balance between strategic, tactical, and operational work instead of drowning in the last of them. This is not about working less—it is about working at your own level, and working differently.

**Depth.** Decisions get tested before they are announced, conversations are prepared in advance, and I have time to get into projects down to root causes instead of skimming.

**Personal growth through short cycles.** We grow when we see the consequences of our actions and manage to connect cause to effect. Normally days or weeks pass between the act and the conscious conclusion, and the link is lost. AI shortens that loop to minutes—and the speed of that loop is what determines how fast we learn. The case above is exactly about this.

A caveat: AI does not decide for you and does not replace live reflection—it accelerates it. You still build the cause-and-effect chains yourself, just many

times faster. This is probably the most underrated way AI amplifies a human being: not instead of them, but alongside them.

### The Limit of the Alliance: Distinctiveness Stays with You

Now about the main limit of this alliance. AI is excellent at convergent tasks: analysis, structuring, pattern finding, conclusions, execution—everything where "more correct" and "more precise" exist. It also handles average creativity: it will fill out a brainstorm, offer ten headline options, polish a draft. Studies show that ideas from AI make an individual's writing noticeably better—especially if generating ideas is not their strong suit. But there is a catch, and it is a strategic one. The model is trained on a giant mass of other people's work and pulls statistically toward the mean: the same studies find that work created with AI becomes noticeably more similar to other such work. AI interpolates what it has learned—it will not propose anything genuinely new, anything that was not in the data. Another angle: it is excellent help on tactical decisions. But the strategic ones—and the ones that call for persistence, flexibility, or a way out of an ambiguous situation—are better kept for yourself.

A case in point is the book you are holding. I prepared the third edition together with an AI assistant, and the division of labor fell exactly along that line. Updating facts and cases, checking figures against primary sources, structuring chapters, finding duplication and contradictions, proofreading, packaging ideas into text—the assistant does all of that brilliantly, faster and more thoroughly than I do. But every new idea—add this chapter, put in the story about selling the car, bring a dry section to life with a personal case—came from me. The assistant turned an idea into text in minutes. But deciding what the book was missing—that it could not do, not once.

With one caveat, which this very book handed me. You already know the concept-map story from Chapter 6: the assistant did the draft, and the diagram then passed an automated check; both stages gave it the green light—and yet it carried a composition error, and I only saw it at proofreading.

The point is not that the assistant is bad. It is that "more thoroughly than I do" stops working where the job is to spot an error in its own work: the

automated check missed it, and I caught it. And I caught it not with an edit but with a question.

Hence the rule I will end the discussion of the personal loop with: give AI the routine and the average—keep the distinctiveness for yourself. The more people draw their ideas from one source, the more their results converge—and the more valuable it becomes to be able to come up with something that is not in the market average.

### How This Turned into a Product

An honest disclaimer: from here on I am talking about my own product. Read this section not as an advertisement but as an analysis of product logic—it will be useful even if you are building or choosing something completely different.

The personal loop I described above initially ran on whatever was at hand: notes, AI chats, spreadsheets, templates. On top of that, as a consultant and methodologist, I was assembling a management methodology. At some point it became clear that these practices were adding up to a product—a personal operating system for a leader. It goes to market as BAEOS.

So, here is what I started from. The target user is a director or a deputy director in a company of 50 to 300 people. A player-coach: they can no longer manage by hand, but they cannot yet step out of day-to-day operations. Their day is made of priorities, decisions, meetings, sign-offs, and a constant choice between the urgent and the important—and a noticeable share of that day goes to gathering context and putting out fires. And they have no time to study methodologies. They have a business that has to live every day.

The product logic rests on several principles, and every one of them grew out of practice rather than out of a slide deck.

First: value here and now. The user has to get something useful on day one—plan the day, record a decision, prepare an email—not “sometime later, once the model has been trained.”

Second: the source of truth is actions, not self-description. My goal is to build an AI that will know the leader and act as their autopilot. In a management context, questionnaires give you the image people want to project; real

behavior is visible only in real work—how a person plans, decides, delegates, reacts to things going wrong. Out of that behavior a profile of the leader gradually forms, and the recommendations get more precise, both when drafting emails and when suggesting personal development steps.

Third: no abstract chats. People need scenarios—flexible, adaptive, but scenarios grounded in methodology, so that the system leads them. And the scenarios are extended by the method that came out of the first case in this chapter: watch where the user does not fit the scenarios we built, and extend the product to match their real behavior instead of handing them a box of parts.

Inside it are the same modules as in my personal loop: the day and the journal with reflection, planning from the day out to the month, decisions, communication, delegation and the team, goals with progress tracking, projects with a portfolio, and a one-screen review of the day. Communication is not an abstraction here: there are separate scenarios for meetings, negotiations, feedback, and emails. And documents do not live in a separate folder—the context you need is loaded straight into the scenario: into working through a decision, into a conversation, or into a project.

### Meta-case: an article in two days

A fresh example of the loop at work: my own publications. In August 2026 I released the article “AI Is the Missing Half of the Improviser’s Culture” (in the Russian original, “Russian Ingenuity vs. the Age of AI”). The division of labor was this. Mine: the idea, the theses, the quality criteria, and the final read. The AI’s: fact-checking, with a register of 79 sources, assembling the text to my style rules, SEO packaging, the infographics, and publishing to the site. Calendar days: two. A cycle like that used to take weeks and three contractors—an editor, a designer, and a fact-checker.

Hence an observation that changes organizational design. AI is a lever for talent: one strong person covers the task pool of an entire department. That does not mean “fire the department.” It means the bottleneck is no longer headcount but the number of people able to frame a task and verify the result. Find and grow such people—the lever is already on the table.

And one last thing. Even if you never open our product, the main idea of the chapter is free to take: a leader's personal loop can be assembled out of any tools, a notebook and a single AI chat included. What matters is not the software but the habit: a short entry, an instant response, a short loop. Everything else is a question of convenience.

### Chapter Summary

- AI applies to every class of a leader's tasks, from planning to self-development, but nowhere does it replace the leader entirely: this is the 70/30 rule.
- The most underrated effect is the shortening of the "action → feedback" loop from weeks to minutes: the leader learns faster, conflicts get taken apart while they are still alive, and decisions get tested before they are announced.
- In communication, AI can work as a real-time coach: it tracks your patterns—pushing, shutting down initiative, taking away authorship—and helps you rebuild a conversation while it is still going.
- In decisions, AI is valuable not for a ready answer but for keeping both layers on the table, the arithmetic and the emotion, and not letting one hide under the other. The decision stays yours, but it goes through the full cycle: from the first thought to the reflection after the deal.
- The limit of the alliance: give AI the routine, the tactics, and the average—keep the distinctiveness for yourself. The more people draw their ideas from one source, the more valuable the ability to think of something that is not in the market average.

**What to do:** Start a leader's journal—five minutes in the evening, an entry plus feedback from AI (one week). Then assemble the personal loop: planning, working through decisions, preparing for difficult conversations (one quarter).

**Reflection question:** How much of your week goes to work only you can do—and how much to work AI is already able to take?

## Chapter 22. How to Run Technology Implementation Projects

To make this concrete, let me show you how I sometimes rank projects and build a task list—inside the organizational structure of a limited liability company.

Projects inside a company can be split into four types:

- local department projects
- cross-functional projects
- strategic projects
- investment projects

### *Local department projects*

Local department projects have to meet the following criteria:

- they are delivered by a single department or within a single deputy director's purview
- the budget for outside service providers and materials is no more than $6,000, excluding VAT
- the project changes no business processes in other departments and requires no rework of the company's IT systems, access rights included

### *Cross-functional projects*

Cross-functional projects have to meet the following criteria:

- the budget for outside service providers and materials is no more than $116,000, excluding VAT
- no more than two departments or deputy directors' areas are involved, and no more than two companies from the group take part
- the changes touch no more than three business processes and no more than three departments

### *Strategic and investment projects*

A project counts as strategic if it meets one or more of the criteria below:

- the budget runs above $116,000 and up to $1.16 million

- the project affects whether the company reaches its strategic goals
- the changes touch more than three level-one or level-two business processes and more than three departments reporting to different deputy directors

Projects with a budget above $1.16 million count as investment projects. They are run the same way as strategic ones, with one addition: the project is worked through in more depth under the investment appraisal methodology and the group's standards.

### The Project Life Cycle

Projects in the organization go through the following life-cycle stages:

- initiation—the stage that ends with a decision to launch the project or to reject it, and in which the goals and expectations are put into words
- planning—the stage where the delivery plan is prepared
- delivery and control—the stage where the work gets done, progress and quality are controlled, and communication, the team, and stakeholder expectations are managed
- closure—the stage where reporting and accounting documents are prepared, the project is archived, and the lessons learned are turned into training

During initiation and planning, I **recommend** that you size up the resources the project will need in the following categories:

- production (tools and equipment)
- financial (your own funds and borrowed funds)
- material (raw materials, components, and consumables)
- human (people and skills, plus employee contacts inside the group)
- intellectual (technologies and patents)
- organizational (management and administrative clout)

### Defining Project Goals and Possible Effects

Goal setting is the key step. What do we want out of the project?

To define the goals, objectives, and potential effects, first work out what type of project you have and match it against the goals that type can serve.

*Project types and possible goals*

| Project type | Possible goal |
|---|---|
| Commercial / economic | Making a profit |
| Technical | Solving technical problems that are not tied to economic metrics |
| Image / marketing / social | Building the image or reputation you need, promoting the company, an idea, or an agenda |
| Research / educational | Finding new solutions or teaching new skills |
| Operational / organizational | Raising internal efficiency and effectiveness, including cutting labor hours and speeding up processes |
| Strategic | Solving the organization's strategic problems |
| Mixed | A combination of the above |

It is worth identifying one main goal that sets the overall direction, plus three or four sub-goals—the key objectives whose completion will show that the main goal has been reached.

To size up the effects of a project, you should work out which wastes in the organization's operations it removes; for operational projects this is mandatory. The same list of wastes is useful for assessing project risks.

Those wastes can also appear during delivery itself. They should be identified and assessed as organizational risks.

## Roles in a Project

It is also important to spell out the possible roles—that makes it easier for your innovators to divide up responsibility inside their teams. During delivery, people can take on the roles described below.

*Roles in project delivery*

| Role | What the role does | Notes and clarifications |
|---|---|---|
| Project customer (sponsor) | Defines the goal and the deliverables. Defines the source of funding. Adjusts the project when needed. Signs off on the final results. | Can be combined with the supervisor role |
| Supervisor | Provides overall direction for delivery. Secures funding and allocates the resources the project needs. Handles the thorniest issues—the ones that involve how the customer and the delivery team work together—and takes part in managing project risks. | Can be combined with the project customer (sponsor) or project manager role |
| Project manager | Builds the project team and organizes how the team and the customer work together. Plans, organizes, and controls the work needed to reach the project goals to the required cost, quality, and schedule. Delegates tasks and authority, sets priorities, and supplies the team with resources. Sets and tracks performance indicators and work requirements. Manages project risks and the problem-resolution process. Manages change during delivery. Takes part in | Can be combined with the project supervisor role |

| | | |
|---|---|---|
| | resolving conflicts between project decisions. | |
| Project administrator | Supplies the project manager with the structured information needed to control the project, plus plans, resources, and priorities. Makes sure project documents are prepared, routed, and archived on time. Tracks document approvals. Handles organizational and communication work inside the team. Produces project reports. | Can be combined with the project manager role (especially if you use AI—this function automates well) |
| Project team member | Delivers individual project tasks | |
| Consultants (by discipline) | Advise the project team within their own areas of expertise | |
| User of the project's product | States the requirements for the project's product at the initiation and planning stages. Evaluates the product and gives feedback during delivery. | Can be combined with the project customer (sponsor) role |

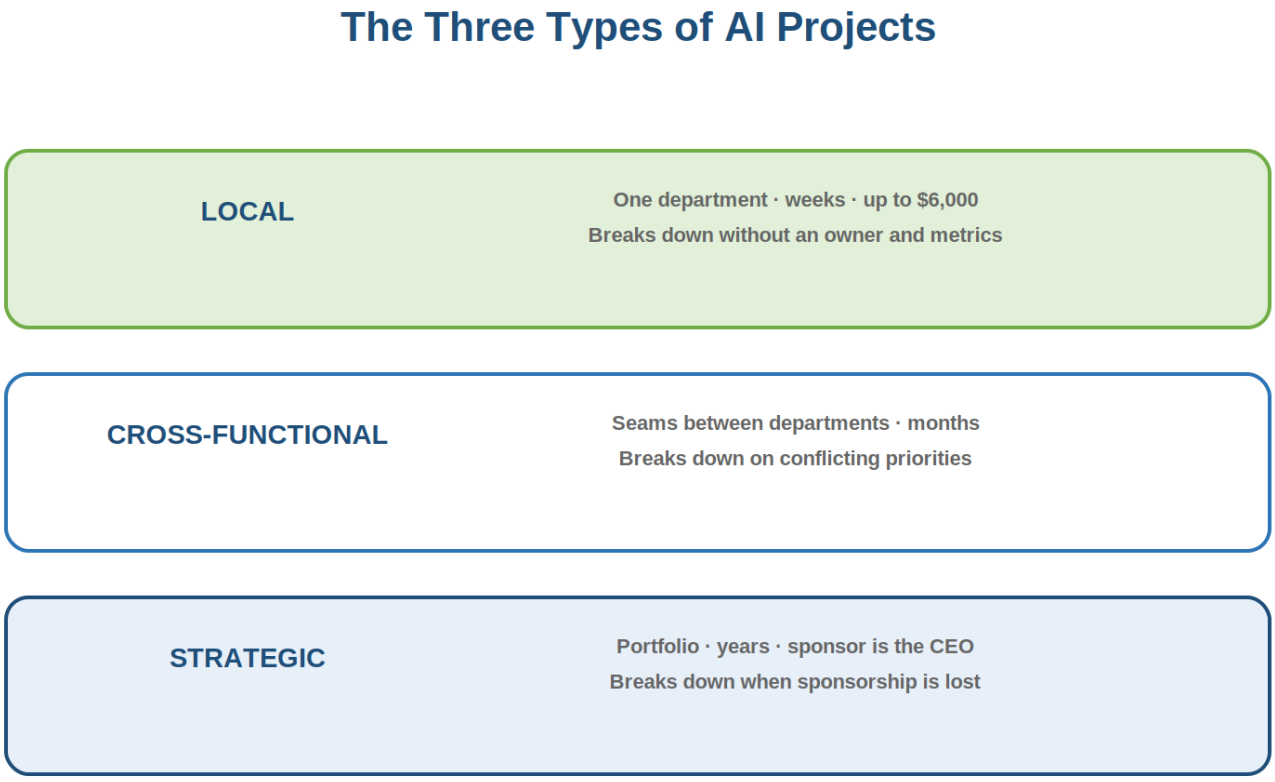
These are the classic project roles; they are the same for any project. The roles specific to the AI function—AI product owner, data engineer, ML engineer, steering committee—are covered in AI-2. Practical guide (Chapter 6).

Within any single task, a participant can be assigned one of the following areas of responsibility:

- R (responsible)—the person who will physically do the work

- A (accountable)—the person who will accept the finished work and answer for it (can be combined with R)
- C (consulted)—the person who is required to advise the others while the task is being done
- I (informed)—the person who has to be aware of the decisions or of how the work is going, but who has no influence over either

### Three Playbooks in Brief




**Local department projects.** The most common type: one department's initiative, one department's budget, a short cycle. The essentials are to lock down the goal and the baseline before you start, run a fast pilot on off-the-shelf tools, measure the effect honestly, and then decide whether to scale. Projects like these founder on the absence of an owner and of metrics.

**Cross-functional projects.** Technology is rarely the main risk here—the risks live at the seams between departments. What you need is a single customer, a governing body, agreed-upon data and procedures, and a phased rollout with checkpoints. These founder on conflicting priorities and on areas of responsibility that belong to no one.

**Strategic projects.** Portfolio logic: a sponsor at the chief-executive level, a dedicated team, change and risk management, a link to company goals, and mandatory interim results every few months. These founder on the loss of sponsorship and on long horizons with no quick wins.

All three playbooks run through the same four stages: initiation, planning, delivery, and closure. The step-by-step versions, with templates, roles, and checkpoints, are in AI-2. Practical guide (Chapter 11). What matters for this book is the management frame: the type of project determines how deep the methodology has to go, and a "single project file" with an AI assistant keeps everything inside one loop.

### Chapter Summary

- Project types—local, cross-functional, and strategic, with investment projects run on the strategic playbook—call for different depths of methodology; getting the scale wrong (running a local project like a strategic one, or the other way around) kills both kinds.
- The management core of any AI project: goals and effects defined before the start, roles assigned, and a "single project file" that collects status updates, risks, and decisions.
- An AI assistant inside the project loop takes over the routine of running it—minutes, status updates, risk registers—and the leader's time goes back to decisions.

**What to do:** Classify your next AI project by type and staff the roles (one week). Put a "single project file" in place, with AI-written status summaries (one quarter).

**Reflection question:** Who on your project is accountable for the business effect after the rollout—and not just for the fact that it launched?

## Chapter 23. A Systems Approach to Implementation and the Road Map

### The AI Hype Wave

By 2026 the hype wave had receded: the market had passed the peak of inflated expectations and moved into a sober phase, one where a systems approach wins and enthusiasm does not (there is more on this in the Postscript). Which makes the lesson that opened this section in the first edition more valuable, not less.

When the ChatGPT hype wave first hit, everyone assumed AI would take the routine work off people's hands. In practice, so far, it is replacing people in creative and intellectual work. The irony was not lost on anyone: we wanted to do the creative work and send the robots down the mines, and right now the robots are generating the creative work while people are still down the mines. Is that what we want from AI? As an illustration, let me quote an interesting post that carries one person's view.

Here is the post, quoted as it was written.

"1. 'I want an AI that will wash the dishes and do the laundry for me. I don't want an AI that writes and draws for me so that I can spend the time it frees up washing dishes and doing laundry.' I came across that spot-on comment from a woman in a creative field on Twitter.

"2. That one line points to a good direction for AI startup ideas. Don't try to build an AI that does something that matters to a person on their behalf! AI should take on the unimportant things—the routine and the peripheral, the work a person does not enjoy but has to do.

"3. Which means that when you survey potential customers, your job is to sort everything they do into two groups: what they like doing and what they don't. Leave what they like alone. And what they don't like—replace it with your own little AI machines, and sell those to them."

What we have here is an illustration of a product tool: Jobs to Be Done. There are two types of jobs—the ones people simply want done and the ones people enjoy doing.

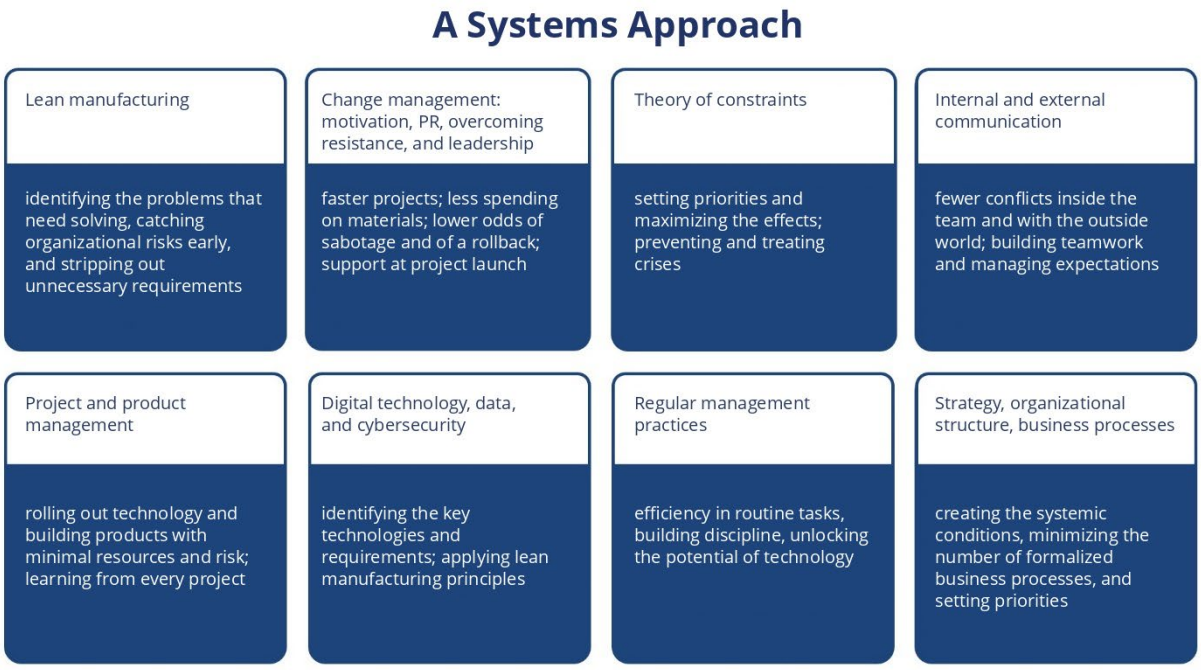
This is exactly why I say this over and over: only 20% of IT is technology, and everything else is communication, analysis, and decisions. AI is one tool of

digitalization and automation among many. It is capable of a great deal, but it is no cure-all, and success takes a systems approach.

## A Systems Approach to AI Implementation

As we have already established, AI—like digitalization as a whole—has to be rolled out systematically and across the board.

My systems approach rests on eight tracks; the diagram below lays them out. Each of them is covered in detail in DT-2. The Systems Approach; here I will focus on the one that sets the sequence of an AI rollout—the road map and the links to other projects.




*The eight tracks of the systems approach*

## The Road Map and the Links to Other Projects

AI has to be implemented inside an overall road map and strategy for digitalization, digital transformation, and development. It will only realize its potential in combination with other technologies and tools. I will say it again: this is not a magic wand. AI depends on the company's overall strategy, on how it works with data, and on its digitalization strategy—digital maturity included.

You also need to understand what AI projects will depend on: quality data inside the company. That, in turn, depends on building quality-assurance

processes and automating data collection, which means taking manual entry out of the loop. And automated collection depends on decent connectivity and on having digital tools in place.

The exceptions, I would say, are machine vision projects (they can become data sources in their own right) and digital advisors built on internal procedures and legislation, or on functional domains that have open methodologies behind them: project management, lean manufacturing, tax advice, and so on.

And however hard I try to draw up a separate road map, it ends up duplicating exactly what is already in DT-2. The Systems Approach (also available through the QR codes and hyperlinks above).

So machine vision projects and digital advisors for the office should be the top-priority track: this will help you get quick results. And for the long game—rolling out the internal advisors that will deliver the most value—there is no way around automating data collection, building data-quality processes across the organization, and developing the organization itself: naming data owners, training people, and writing the rules without which quality processes do not survive.

## Chapter Summary

- The hype wave has receded, and that is a good thing: what now decides the outcome of an implementation is a systems approach, not enthusiasm.
- An AI road map does not live on its own: it fits inside the overall digitalization strategy and connects to projects on data, on infrastructure, and on training people.
- Sequence beats speed: connectivity and digital tools → automated data collection → data quality → AI projects. The exceptions are machine vision and advisors built on open methodologies: they do not wait for data and they deliver fast. Skip a stage and the project goes back to the start—with the budget already spent and the sponsor's trust already gone.

**What to do:** Sort your AI initiatives into those that depend on corporate data and those that do not, and launch the second group (one week). Pull the initiatives into a single map with the links to IT projects (one quarter).

**Reflection question:** Does your company have an AI road map—or only a list of pilots?

## Chapter 24. What a Leader Needs to Know and Be Able to Do in the Age of AI

### Introduction

Adopting AI in a company is impossible without the right competencies in your people. Up to 95% of companies report that adopting AI produced no visible financial result—and one of the reasons is a competency gap. The same holds for digitalization as a whole: the centralized approach, where the IT department is responsible for everything, does not work. Ideas have to come from the business and be adapted to its specifics. But what exactly should people be taught? And what do leaders need to know in order to oversee the development and adoption of artificial intelligence?

As I worked on the material below, I made a point of keeping the structure I set out in DT-2. The Systems Approach. My logic is simple: AI is one digital technology among others, which means it follows the rules of digitalization in general. The only difference is that in this book I have gone into a little more detail specifically in the AI context.

### The Competency Model

I suggest working from a four-level scale for assessing people:

- Level 0—no knowledge or skills; the role does not require them.
- Level 1, basic—knows the fundamentals, works from instructions, needs support.
- Level 2, working—applies it confidently, adapts it to the task, helps colleagues.

- Level 3, expert—sets the standards, builds solutions, trains and advises others.

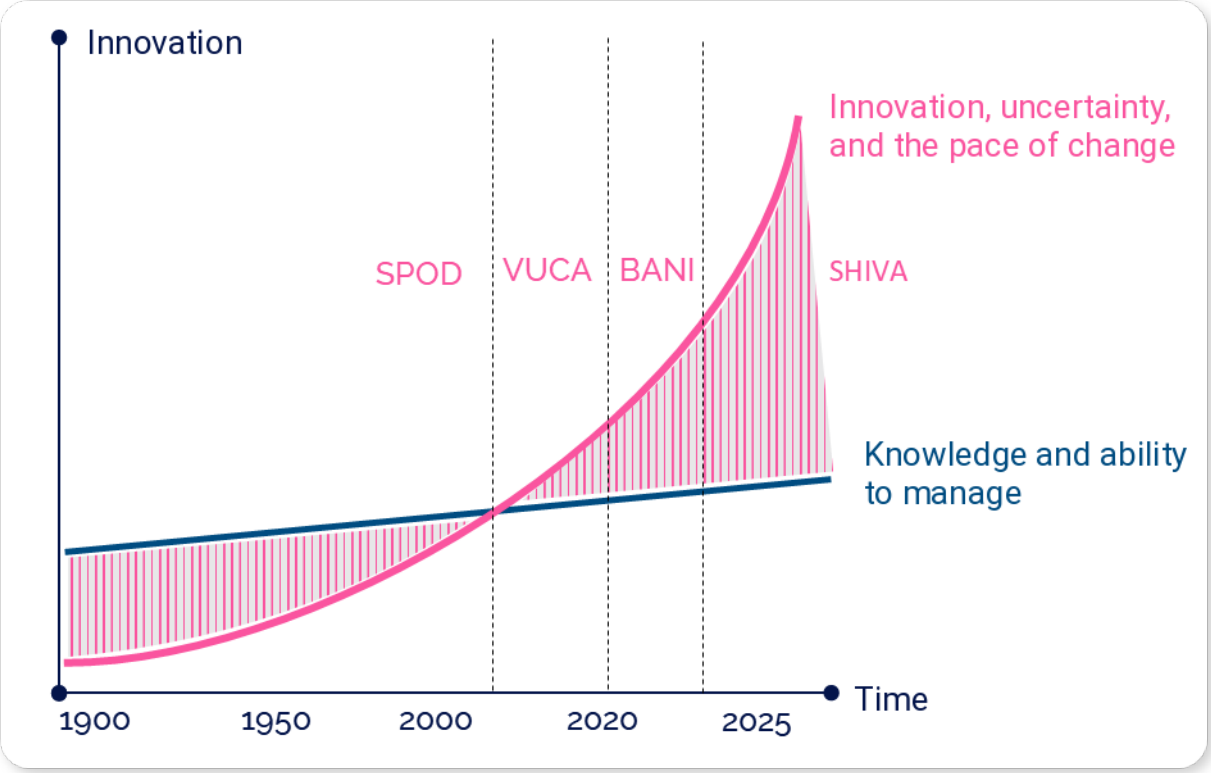
The model itself rests on three tracks: personal, managerial, and digital competencies.

1. Personal Competencies

Within this track I identify nine key competencies.

1.1 Willingness to Learn / Adaptability to Change

We now live in a world where change keeps accelerating.




In my view, it is adaptability and flexibility that matter more and more. In the AI context they break down into several blocks:

- speed in picking up new AI tools and technologies
- flexibility in changing your approach when new solutions appear
- openness to experimenting and to testing what AI can do
- readiness to rethink your existing processes

1.2 Result Focus and Critical Thinking

Adopting technology for technology's sake is a guaranteed way to lose money. You will never make yourself completely safe or take no risk at all, but the few areas below will at least bring the risk down:

- a focus on measurable results from AI adoption
- critical analysis of what AI systems propose
- checking the facts and the reliability of AI-generated content
- assessing AI solutions against specific metrics

1.3 Customer Centricity

Any technology should ultimately be aimed at the customer. And internal customer focus is developing now as well. IT, for instance, treats its colleagues as customers and collects "customer" metrics on them.

- Using AI to improve the customer experience.
- Personalizing customer interactions with AI.
- Analyzing customer needs through AI analytics.
- Using customer data ethically in AI systems.

1.4 Emotional Intelligence

In IT projects, 90% of the work is defusing people's fears. With AI the figure is higher still. For most people AI is close to magic, and that means even more fear. What matters here:

- understanding the emotional reactions to AI adoption
- managing employees' fears about being replaced by AI
- showing empathy when teaching colleagues to work with AI tools
- motivating the team to master AI technologies

Emotional intelligence has traditionally been a leader's key skill. That is logical enough: management has always been built on working with people. But as some of the workload is delegated to AI agents, the emphasis shifts. Managing "digital assistants" effectively calls not for empathy but for logic: the ability to state a task precisely, break a process down, and set the rules and

criteria for judging the result. In essence it is the same task-setting skill, pushed to the level of algorithmic precision—because AI does exactly what it is told: a rule you did not set does not exist for it, and anything left unsaid it will fill with a plausible guess.

This does not mean emotional intelligence is losing its value: people are not disappearing from organizations, and working with them still takes sensitivity. But a new dimension is being added to the leader's competency model—the ability to interact not only with human beings but with digital systems, and that calls for digital fluency, the ability to design a process, and the ability to work with data. These are skills worth developing right now.

1.5 Public Speaking / Presenting

I have watched more than one project fail to start purely because its leader could not defend the idea or was afraid to speak at the digital committee. And in some large companies it is written almost into the procedures that you have to "sell" your project to every executive and explain what it is about. Broadly, the tools I would suggest here are these:

- presenting the results of AI projects to management
- explaining to different audiences what AI can do
- running training sessions on AI tools
- speaking at AI conferences and forums

1.6 Creativity

The best project ideas come from the business itself. There is no wizard who will do the thinking for you. So it matters that you use AI to unlock your own potential and that you look for places to apply AI everywhere. Four areas to work on here as well:

- finding non-obvious uses for AI in your work processes
- taking a creative approach to framing prompts for AI
- generating innovative ideas for AI projects
- creating original content with AI tools

1.7 Problem Solving

You cannot bring in innovation without problems and failures. What matters at that moment is to stay calm and solve them—and also to think about how the problems the business runs into could be solved or prevented with AI.

- Diagnosing the problems that AI can solve.
- Taking a structured approach to adopting AI solutions and to the problems that come with them.
- Finding alternatives when AI projects fail.
- Solving business problems end to end with AI.

1.8 Acceptance / Tolerance of Other Views

The key to finding solutions and to a quality product is the ability to reach agreement. That means getting people with very different temperaments to talk to each other, which is hard enough in itself. And in any case there will be projects that succeed and projects that do not.

- Openness to different approaches to adopting AI.
- Attention to the views of skeptics and critics of AI technologies.
- Tolerance of mistakes and failures in AI experiments.
- Willingness to discuss the ethical questions around AI.

1.9 Method / Discipline / Persistence

No tool, however good, will take root unless the leader shows by personal example that it matters.

- Studying AI technologies systematically.
- Following the established procedures for working with AI.
- Documenting the experience and results of AI projects.
- Building AI competencies across the organization step by step.

2. Managerial Competencies

Within this track I identify eight key management skills, the ones that form the basis of my systems approach—adjusted, of course, for how far AI has spread and how fast it is developing.

2.1 Strategy, Organizational Structure, Business Processes

- Building a strategy for embedding AI in business processes and for long-term development.
- Creating an organizational structure to run AI projects—and using AI to design that structure.
- Optimizing business processes with AI technologies.
- Long-term planning for AI in the organization, including putting AI adoption and AI experiments into leaders' own metrics.

2.2 Regular Management Practices

- Embedding AI in day-to-day management processes.
- Using AI analytics to make management decisions.
- Automating routine management tasks with AI.
- Monitoring performance indicators with AI systems.

2.3 Digital Technology, Data, and Cybersecurity

- Integrating AI with existing digital systems and technologies.
- Ensuring data quality for AI algorithms.
- Protecting AI systems from cyberattacks and data leaks.
- Meeting information security requirements when working with AI.

2.4 Project and Product Management

- Running AI projects from idea to deployment, including ROI and TCO calculations, and an understanding of the AI project life cycle and how such projects differ from classic automation.
- Developing AI products and services for customers.
- Coordinating cross-disciplinary AI teams.

- Controlling budgets and timelines for AI initiatives.

2.5 Internal and External Communication

- Communicating and promoting the results of AI projects inside the organization.
- Working with external suppliers of AI solutions.
- Managing the organization’s reputation around its use of AI.
- Building a corporate culture that accepts AI technologies.

2.6 Theory of Constraints

- Identifying the bottlenecks in processes that AI can clear.
- Optimizing system throughput with AI solutions.
- Balancing the load between human and AI resources.
- Managing constraints as AI solutions scale.

2.7 Change Management: Motivation, Public Relations, Overcoming Resistance, and Leadership

- Motivating employees to master AI technologies.
- Working with resistance to change during AI adoption.
- Leading the organization’s AI transformation.
- Building a positive image for AI initiatives.

2.8 Lean Manufacturing

- Applying lean manufacturing principles to AI projects.
- Eliminating waste in processes through AI-driven optimization.
- Continuously improving AI solutions.
- Cutting costs through AI-based automation.

3. Digital Competencies

In the book on the systems approach to digitalization I put forward a model of digital competencies: digital fluency, working with data, information security,

organizing collaboration, and producing content. I spent several months trying to build a separate competency model for AI. In the end I realized that AI is a special case and has to follow the rules of digitalization as a whole. So the same model applies here too; the only question is the level of detail.

3.1 Digital Fluency

- Classifying the types of AI and the areas where each applies.
- What AI technologies can and cannot do.
- Framing prompts and working with AI tools (the task-setting system is in Chapter 6).
- How AI connects to other digital technologies.
- Using AI inside corporate IT systems.
- **The language of your domain.**

This point now absorbs something nobody used to list separately: command of the language of your own domain, which works in tandem with prompting. AI is sensitive to terminology—the precise professional word produces a better result and burns fewer tokens (we went through the mechanics in Chapter 6). Hence an unexpected answer to the fashionable question "why do we need junior specialists if we have AI?" Without command of the language of the subject, a person can neither set the task nor check the answer. Expertise stops being about "I can do it with my hands" and becomes about "I can name it, set it, and check it."

3.2 Working with Data for AI

- Data quality for training AI models.
- Labeling and preparing data for AI.
- Data management in AI projects.
- Interpreting the results of AI analytics.
- Bias and fairness in the data that feeds AI.

3.3 AI Security

- The regulatory requirements for AI.
- Information security for AI systems.
- The ethics of using AI.
- AI security and business risk management.
- Privacy-preserving technologies in AI.

3.4 Creating Content with AI

- Generating different kinds of content with AI.
- Prompt engineering for producing content in various formats.
- Human-AI collaboration in creative work.
- Automating content production.
- Quality control of AI-generated content.

3.5 Organizing Collaboration with AI

- Integrating AI assistants into teamwork.
- Automating workflows with AI.
- Human-AI teams and how the work is divided between them.
- Training employees to work with AI.
- Monitoring how effectively humans and AI work together.

### The Competency Matrix

One important line to draw: what follows is about the competencies of the leader and of the organization as a whole. The competencies and the staffing of the implementation team—roles, hiring, training, working with resistance—are covered in AI-2. Practical guide (Chapter 13).

I usually divide employees into five groups:

- Owners, the CEO, and their deputies (CEO-1).
- Middle management.
- Informal leaders.

- AI specialists.
- Rank-and-file employees.

The competency matrix by group:

| Competency | CEO/top | Middle management | Informal leaders | AI specialists | Rank-and-file employees |
|---|---|---|---|---|---|
| 1. Personal competencies (9 competencies) | 3 | 3 | 3 | 3 | 2 |
| 2.1 Strategy, structure, processes | 3 | 3 | 2 | 2 | 1 |
| 2.2 Regular management practices | 2 | 3 | 2 | 2 | 1 |
| 2.3 Digital technology, data, and cybersecurity | 2 | 3 | 2 | 2 | 1 |
| 2.4 Project and product management | 2 | 3 | 2 | 3 | 1 |
| 2.5 Internal and external communication | 3 | 3 | 2 | 2 | 1 |
| 2.6 Theory of constraints | 2 | 2 | 1 | 2 | 0 |
| 2.7 Change management and leadership | 2 | 3 | 2 | 2 | 1 |
| 2.8 Lean manufacturing | 1 | 2 | 1 | 2 | 0 |
| 3.1 Digital fluency (AI focus) | 2 | 2 | 2 | 3 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| 3.2 Working with data for AI | 3 | 2 | 2 | 3 | 1 |
| 3.3 AI security | 3 | 2 | 2 | 3 | 1 |
| 3.4 Creating content with AI | 2 | 2 | 3 | 3 | 2 |
| 3.5 Organizing collaboration with AI | 2 | 3 | 2 | 2 | 2 |

Boiled all the way down, the picture looks like this.

- CEO/top.

Expert-level competencies: personal, strategy/structure/processes, communication, working with data for AI, and AI security.

Focus: strategic leadership, governance and decision rights, AI security.

Role: the visionaries of the AI transformation and the ultimate owners of its risks.

- Middle management.

Expert-level competencies: personal, strategy, regular management practices, digital technology, project management, communication, change management, and collaboration with AI.

Focus: putting the AI strategy into operation.

Role: the main executors of the AI transformation.

- Informal leaders.

Expert-level competencies: personal and creating content with AI.

Focus: showing what AI can do through practical results.

Role: cultural ambassadors and the people who get others excited.

- AI specialists.

Expert-level competencies: personal, project management, AI-focused digital fluency, working with data, AI security, and creating content.

Focus: technical expertise and advice.

Role: centers of competence and internal experts.

- Rank-and-file employees.

Practical skills: creating content with AI and collaboration with AI.

Focus: using AI tools effectively in their day-to-day work.

Role: active users of AI solutions.

### Competencies for the Age of Verification

A separate slice is the competencies of the specialists whose work AI is already rebuilding (we went through the engineering case in Chapter 18). Five competencies are becoming scarce, and only one of them is "about AI": **domain depth**—the physics of the process, not breadth: only someone who understands why it is so, rather than what the output says, can verify what a model produces; **task setting and working with constraints**—a bad specification plus AI equals nonsense delivered fast and beautifully formatted; **data literacy**—where the training sample came from, what model drift is, why an assistant that is flawless on a repeat task lies on a one-off one; **traceability and ownership of the decision**—the signature will still be human, and the specialist has to be able to explain the decision rather than say "the model said so"; and finally the role of **"translator"** between the domain and the data people—the scarcest and the most underrated of the five.

And the main HR conclusion for training programs: it is the same one the engineering case in Chapter 18 pointed to—training programs have to be redesigned around verification rather than generation. **The specialist who was valuable for the speed of their hands will disappear. The one who is valuable for the quality of their judgment will remain.** That is good news—but only for those who have started rebuilding how they train people today.

Let me add one more shift the AI era brings. For decades the market paid for narrow specialization, while broad erudition read as an "undecided profile." AI reverses that rule. Missing depth can now be rented: the model supplies the background, the calculation, and the draft in any field. But connecting fields

to each other, seeing the analogy, asking the right question—that AI does not do without a human.

That is why the erudite with an AI tool is becoming the new scarce figure: one person able both to hold the whole picture and to verify the details with "rented" depth. The practical conclusion for development programs: not only deepen your specialists but also broaden the horizons of your strong people—cross-industry rotations, adjacent domains, work at the intersections. That used to be a luxury. Now it is a lever.

## Chapter Summary

Study after study shows the same thing: the leading companies invest first in processes and people, and the technology and the specific models follow as a consequence rather than as ends in themselves. Invest in technology and forget about people, and what you get is expensive toys nobody needs. The business will not care, no ideas will come from it, and everything you deploy will be rejected like a foreign body.

To drive change you have to train the senior leadership first and everyone else after that. And it matters that you train at least 10% of the entire workforce: that 10% must become the change agents, backed by the top team. The top team itself, meanwhile, has to be the example and the leader in using AI. Otherwise people will not buy the idea and will not believe in it.

### Key Takeaways

- A leader does not need to know how to train neural networks. A leader needs to know how to set tasks, judge where AI applies, do the math on the economics, and manage AI risk.
- The 2026 shift: systems thinking, working with data, and "AI literacy" have been added on top of emotional intelligence—being able to work with digital assistants is becoming basic hygiene for a leader.
- The competency matrix is a diagnostic tool: different roles—the CEO, middle management, informal leaders, AI specialists, and rank-and-file employees—need different depth.

**What to do:** Assess yourself and your key leaders against the matrix in this chapter (one week). Put AI literacy into the development program for your management team (one quarter).

**Reflection question:** Which competency in the matrix are you personally missing—and what will you do about it in the next month?

## Conclusion

### Who Is Exposed—and What to Do About It

The people who should fear AI most are white-collar workers: knowledge and creative professionals at entry and mid levels of skill. That, to my mind, is the main danger AI brings.

Yes, AI will help top-tier specialists. But here is the question: where will the next generations of experts come from if rank-and-file employees struggle to find work? AI will raise the barrier to entry into the professions and force us to evolve, to change what universities teach. Reading was once an aristocratic privilege; today five-year-olds can do it. And the key question is this: are you ready for that evolution?

You can work out the role of the human being in a future where AI is everywhere by answering one question: what is AI, actually? Is it automation, digitalization, or digital transformation?

Take the classic definition of automation: cutting the volume of manual human labor in day-to-day work and the effort it takes.

The effect of automation is faster processes, a lighter load on people, and less risk of human error.

Does that fit? It does. Most AI projects are exactly that: automating certain tasks and taking load off people.

Digitalization is the deployment of digital technology and systems that lets you rebuild business processes to be more efficient and more flexible; it is where working with data and making decisions based on data begins.

The effect of digitalization is a lower cost of running processes, better processes, and greater flexibility.

Together, automation and digitalization make it possible to scale a business fast and to get through the constraints a crisis imposes.

That fits the description of AI too.

Digital transformation is a wholesale rebuild of the business, the management system, and the processes, using the results of digitalization and automation to increase commercial potential and grow profit.

So the effect of digital transformation is new personalized products for the target audience, combined with a severalfold reduction in internal costs.

Well—that is a function of AI as well.

Which makes AI a technology that can automate processes, digitalize them (including making decisions based on data), and create new services and offerings. But only a human being can unlock all of that potential and then keep improving it by setting ever newer and harder tasks.

On the boldest forecasts, the first one-person "unicorn" should appear within 20–30 years. A unicorn is a company that reached a valuation of $1 billion within 10 years of founding and has not yet listed its shares (no IPO). Behind it will be a single brilliant individual able to delegate tasks to AI, use it to the full, and replace a whole team of specialists that way. And if a company like that needs extra "human" help, it will hire outsourced teams.

What should you personally do about this? Here is the whole book's answer in one line: move from being the one who does the work to being the one who sets the task and verifies the result—the person who briefs the AI and answers for what comes out. The competency model for that is in Chapter 24; the personal loop is in Chapter 21. The barrier to entry into the professions will rise, but so will the value of everyone who clears it.

The second line of defense is creativity, which we went through in detail in Chapter 21. The gist, as a reminder: AI is brilliant at convergent tasks and at "average" creativity, but statistically it pulls toward the mean and does not

produce anything genuinely new—and the ideas of people who draw on a single source converge. So what is exposed is not the "creative professions" but averaged execution in any profession. The ability to come up with an idea that is not in the market average, to pick what matters and bet on it, commands a higher price with every day of mass AI.

Meanwhile, the current race in artificial intelligence is dominated by a single narrative: build ever larger and more powerful general-purpose models, build strong AI. Yet there is a strategically important alternative that this trend has left undervalued—a focus on developing and deploying specialized small and midsize models (small language models, SLMs). That paradigm offers practical answers to the key problems of adopting AI in business.

## Getting Past the Barriers to Deploying "Strong" Models

Resistance and sabotage. Deploying a "strong" AI that makes decisions frightens people and provokes rejection, which inevitably leads to sabotage. That makes investment in solutions of this kind unreasonably risky for a business.

Regulatory risk. Legislative initiatives such as the AI Act put powerful models in the high-risk category, which implies tight regulation and oversight—more barriers and more cost.

Economic inefficiency. Businesses are not ready to build critical processes on ultra-expensive solutions (often sold by subscription) that they then become utterly dependent on. Investment in systems like that frequently never pays for itself.

The net result is technological complexity, high cost, and constraints—and all of it is hard to sell to the business and to the people in it.

## The Advantages of Specialized Models

Acceptance and trust. Local, specific solutions that handle a concrete applied task sit more comfortably with people, because they do not trigger existential fears about "strong" AI.

Business impact and affordability. Models like these deliver measurable value—an effect the business is willing to pay for. Their relatively low cost and modest infrastructure requirements put them within reach not only of large corporations but of small and midsize business (SMB) as well. That opens up a flow of investment into development.

A focus on quality and optimization. Developers can concentrate on building highly effective models for narrow tasks, improving them continuously, "compressing" them through distillation and quantization, and raising their "intelligence" within the specialization.

Lower infrastructure costs. Compressed, optimized models let you raise system performance without constantly adding expensive compute infrastructure.

### Where Specialized AI Is Heading

- Hybrid architectures (MoE)

Combining several specialized models (experts) to handle complex tasks—for example, pairing models with a strong logical or mathematical component (critical for analytics and calculation) and models that specialize in text generation or natural language processing. This offsets the weaknesses of the individual models and raises the overall reliability of the system.

My own view is that hybrid architectures and the separation of models will go all the way down to the hardware: each model will get its own small server.

- Fighting hallucinations and raising reliability

Developing methods and architectures that minimize the risk of generating false information ("hallucinations"), which matters most in fields that demand precision: finance, analytics, medicine.

- Safe interaction and minimizing direct prompting

Hiding models behind formalized interfaces, where the user receives finished recommendations or results rather than interacting with the AI directly through open text input (a prompt). This sharply reduces the risk of prompt

injection attacks and of errors caused by poor-quality or malicious handling of user input.

### The Foundation: Data Quality and Protection

The success of any AI—and of specialized AI above all—depends directly on data quality. That makes the projects that come before deployment critical.

- Analyzing and optimizing business processes—a clear understanding of what data is needed and how it is generated.
- Working through data flows—building hierarchical data models and describing how data moves.
- Ensuring data quality—putting in place processes that control and improve the quality of inbound data.
- Protecting data and models—building security mechanisms to prevent data poisoning and dataset compromise. Introducing checkpoints and backups of both models and data. The entire infrastructure has to guarantee data security and model integrity.

### Putting It into Practice

This strategy takes shape in projects that build "sub-products": specialized solutions assembled from carefully chosen models. Each of those models has its own specifics and focuses on a concrete applied business task. That is how you build AI tools that are cost-effective, safe, and easy to deploy, and that produce a fast and measurable result.

Take AI for technical support. Yes, there is an ordinary language model inside—but the tooling wrapped around it, the work on the knowledge base, and the organizational structure make it possible to resolve up to 80% of support requests automatically. That is a different metric from the roughly 40% reduction in ticket volume mentioned earlier: there a chatbot takes load off the first line at intake, here the system closes the request end to end. Gartner points the same way: by 2029 agentic AI is expected to resolve up to 80% of common support requests on its own.

In my view, shifting the focus away from the race to build all-encompassing "strong" models and toward an ecosystem of specialized, efficient, affordable solutions is a fundamental change. It is an economically sound alternative, and one that will shape science too. Because someone has to pay for science.

This approach gets past the barriers to adoption, secures trust, and raises the reliability of systems through hybrid architectures and safe interfaces. And ultimately it accelerates the real-world application of AI to business problems at every level—from global corporations to SMBs.

Let me close with a couple of facts.

A lot of companies are now trying to build in-house AI teams. To my mind, that is a dead end. Good specialists are scarce, and paying the strong ones what they could earn on the open market is hard. Which means companies like that will end up with weak specialists either way. But suppose a company does manage to assemble a strong team. What will motivate that team to keep developing? And, more to the point, can it develop at the same speed as a specialized team? Who improves faster: someone who does three to five projects a year, or someone who does fifty?

The 95% named above comes from a study by the Massachusetts Institute of Technology (MIT). It found that AI genuinely can increase a company's revenue—but that this happens in only about 5% of cases. The other participants (the sample covered 300 organizations that had already deployed AI) recorded no appreciable financial effect at all, despite substantial investment. For most companies AI turned out to be an economically useless tool—precisely in the way it is currently deployed.

"The pilot projects at some large companies and the young startups really are succeeding with generative AI," notes Aditya Challapally, one of the study's authors. "Startups run by 19- and 20-year-olds, for example, have seen revenue jump from zero to $20 million in a year."

In the MIT researchers' assessment, the reason for the poor results lies not in the quality of today's models but in how companies approach them: they invest in basic integration of neural networks but not in tuning them to

internal processes or in systematic training for employees. “General-purpose tools such as ChatGPT are convenient for individual use, but in a corporate environment their potential only emerges when they are adapted to a specific purpose and written into work procedures,” Challapally emphasizes.

The cost structure is another key factor. More than half of generative AI budgets go to sales and marketing tools, whereas the highest return, according to MIT, comes from digitalizing the back office: cutting business process outsourcing, reducing spend on external agencies, and optimizing internal operations.

The companies that did achieve revenue growth invested consistently in developing and customizing AI for their own tasks. In those cases the organizations chose narrowly specialized tools rather than mass-market solutions, and brought the vendors in to refine them jointly. Per MIT, pilots of that kind reached deployment about 67% of the time, against roughly 33% for a fully in-house cycle—from code through deployment and training. Note that this is the same pattern as in the chapter on small business. “Buy off the shelf” does not mean “take the box built for everyone”: what wins is a specialized external solution refined jointly with its vendor, not in-house development from scratch.

MIT’s data also shows that going it alone materially raises the probability of failure. Many of the companies surveyed were reluctant to disclose how often they failed, but solutions bought from external vendors delivered more reliable financial results. Among the additional success factors, the researchers name giving authority to line managers—not only to centralized AI labs—and choosing tools that can integrate deeply into processes and evolve over time.

## Postscript 2026: Four Signals of Market Maturity

By the time I was preparing the third edition of this book, the market had already confirmed, in concentrated form, the forecasts I made in the earlier chapters and in the first edition. Four news items came out over the spring and early summer of 2026 that look routine on their own but together redraw the picture. I will go through them here as a closing illustration.

### Signal 1. Anthropic Turns Decisively Toward SMBs

On May 13, 2026, Anthropic launched **Claude for Small Business**—a set of roughly 15 ready-made agentic workflows with connectors to QuickBooks, PayPal, HubSpot, Canva, Docusign, Google Workspace, and Microsoft 365. On a single subscription, an entrepreneur gets month-end close, invoicing, margin control, contract review, lead triage, and content strategy preparation. In parallel, an in-person tour of 10 US cities began, with free workshops for 100 entrepreneurs in each city.

This is not a product launch in the ordinary sense. It is a change of positioning by one of the two global market leaders. Anthropic has said publicly what this book has been arguing all along: SMBs account for roughly 44% of US GDP, while AI penetration there lags far behind penetration in large business. Software has historically been built for corporations, venture-backed startups, and the mass consumer—but not for a 30-person landscaping firm or a 50-person real estate agency. Anthropic is closing that gap.

What matters just as much is *how* they came in. Not with an open chat, but with a box built for recognizable everyday jobs. Not with a promise to "train the model to fit you," but with ready-made scenarios that work from day one. And one more thing: Claude does not act unsupervised—the user initiates every workflow, approves the plan, and confirms the result before anything is sent, published, or paid. That is precisely the paradigm of "formalized interfaces" and "minimizing direct prompting" I described a few pages earlier.

### Signal 2. The Picture in Russia Is the Same—and It Is Also About SMBs

On March 3, 2026, Yandex B2B Tech published data on its Yandex AI Studio platform:

- more than 7,500 AI agents analyzed
- **86% of AI-agent users in Russia are small and midsize businesses**
- roughly 25% of the agents work in technical support, banking, and HR consulting

The scenarios are industry-specific and narrow: property developers and realtors use agents to match units to buyers, property management companies to handle housing and utilities complaints, research labs to diagnose equipment, and others to analyze the news. This is not "GPT for the office"—these are concrete tasks for a concrete business profile.

In parallel, Yandex introduced tools that allow AI to be used in cloud infrastructure with no internet access—a direct answer to the SMB segment's main objection about data security. The Russian market is repeating the global trajectory with a lag of about six months, and here too the winners are narrow off-the-shelf solutions with security packaged properly.

### Signal 3. The Production Paradox: 74% Rolled Back

On May 14, 2026, the Swedish communications provider Sinch published a study called **AI Production Paradox**—a survey of 2,500 AI project leaders across countries and industries. The headline finding was harsh:

**74% of companies rolled back or entirely shut down the AI agents they had already launched in customer support.** "Rolled back" here means withdrawal from live operation, not cancellation of a project before launch.

The most counterintuitive part: at companies that describe their own AI control system as "fully mature," the rollback rate is **higher—81%**. Sinch explains it this way: mature teams roll back more often not because their models are worse but because they see problems faster and react faster. And the number does not depend on region, industry, or company size—the problem cannot be solved with money.

The rest of the numbers run in the same direction:

- 84% of AI engineering teams spend at least half their working time on security infrastructure rather than on model development
- 75% of companies put trust, security, and compliance among the top three priorities of their AI programs—ahead of AI development itself (63%)

This also confirms an earlier Gartner forecast from the summer of 2025: half the companies that expected to cut customer support headcount significantly

through AI will abandon those plans before 2027. Gartner's vice president put it plainly: a fully unstaffed contact center is not yet technically feasible and is operationally undesirable.

A caveat on method: Sinch is an interested party—it sells communications infrastructure for AI agents. But even with that discount, the 74% figure fits the overall trend of the past year.

### Signal 4. Summer 2026: Access to Frontier Models Became a Managed Resource

The fourth signal arrived after the three from spring. In June 2026 the United States, by export directive, switched off access to Anthropic's most powerful models worldwide for three weeks (there is a detailed discussion in Chapter 7). On the independent benchmarks practitioners ran, the "export version" of the flagship that came back had lost noticeable quality on business tasks. And the head of OpenAI proposed an "IAEA for AI," an international certification of access to frontier systems.

For strategy this means that the very ability to use a frontier model has turned into a resource managed from outside. Products that depend critically on one specific model carry a regulatory risk no contract can cover. What wins are solutions where the model is a replaceable component and the value sits in the scenario, the data, and the control loop.

### What These Signals Say Together

Separately, four news items. Together, a map of the market that fits exactly the logic I have been building throughout this book.

**First, the market has changed its core segment.** Over a two-to-three-year horizon the main buyer of AI products is neither the corporation nor the "ChatGPT user" but small and midsize business. Anthropic and Yandex are entering it the same way, from two opposite ends of the world.

**Second, the basic form of the product has changed.** What wins is not the open chat but a box for a recognizable everyday job: a workflow with a clear input, a clear output, and a clear zone of responsibility. General-purpose assistants

are moving into the infrastructure layer; the value the user actually sees sits in specific scenarios.

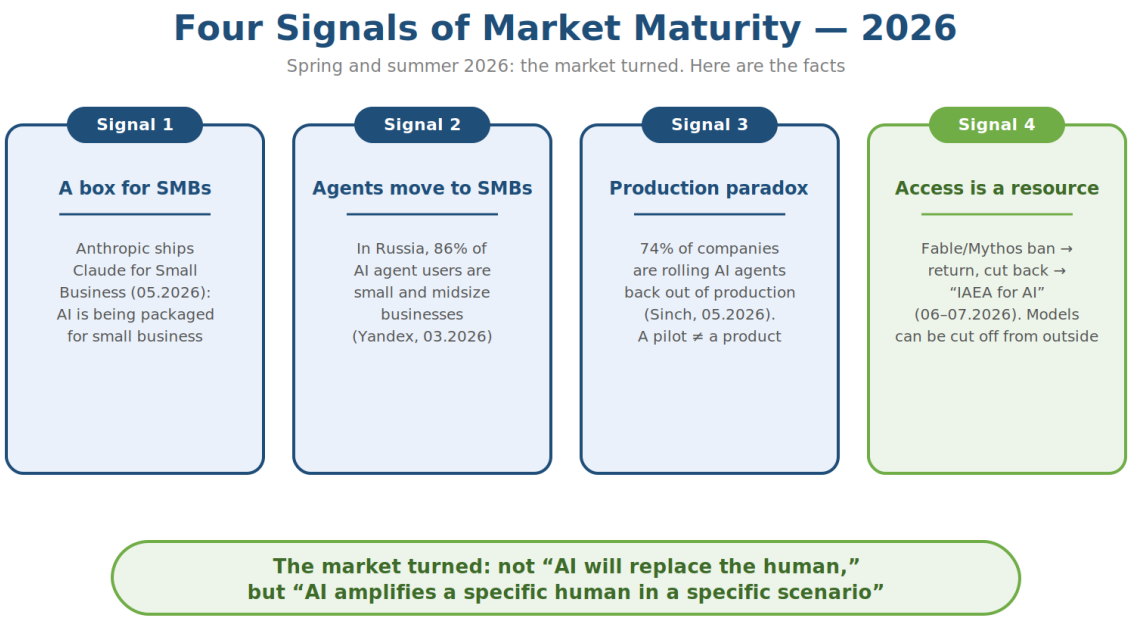
**Third, the bottleneck of the market is not the model but security and trust.** The Sinch rollbacks are not happening because models got worse. They are happening because companies do not have the operational readiness to run autonomous systems in production. 84% of engineering teams spending at least half their time on security—that is the new normal, not a temporary anomaly.

**Fourth, full autonomy lost the 2025 round.** The industry has effectively conceded that the slogan "remove all the staff and install an agent" is unrealistic over a two-to-three-year horizon. Klarna and the other loud cases of 2024–2025 ended with people being hired back. The Sinch rollbacks are the same process, only at the level of the technology rather than the headcount.

**Fifth, the human being returns**—not as an obstacle but as a mandatory architectural element. AI systems from 2026 on are built human-in-the-loop by design, not because of regulation after the fact.

**Sixth, dependence on a specific frontier model has become a risk in its own right**: access to it can change overnight, by a regulator's decision rather than the market's. That is the final argument for the same architecture—the model is a component, the product is the value.

In short: the market is moving from the paradigm of "AI will replace the human" to the paradigm of **"AI amplifies a specific human under that human's control."** Which is to say: to exactly what the chapters on specialized models, digital advisors, and safe interaction through formalized interfaces were about.




## The Profile of a Winning AI Product

These signals add up to a profile of a successful AI product for SMBs over the next two to three years. Six characteristics, and every one of the signals tests them:

- **A narrow but deep scope of tasks.** An assistant for a specific role (owner, director, accountant, salesperson, realtor) or a specific task—not a "helper for everything."
- **Off-the-shelf packaging.** Ready-made workflows and templates instead of "first, train the model." An SMB pays for a result from day one.
- **A human in the loop by default.** The user initiates the process, approves the plan, and confirms the result before anything is published, sent, or paid.
- **Security as the first message, not the fine print.** "Your data does not go into training," "you see every step," "no write operations without confirmation."
- **Integration with the existing stack.** Connectors to whatever the client already runs: 1C, amoCRM/Bitrix24, SBIS, and SKB Kontur in Russia; QuickBooks, HubSpot, and Google Workspace in the West.

- **Independence from any single model.** The model is a replaceable component; the product's value sits in the scenario, the data, and the control loop. A solution tied to one vendor carries a regulatory risk that no contract can close.

### What This Means in Practice

Three practical conclusions for anyone working with AI right now.

If you are deploying AI in your own company, do not try to build a turnkey autonomous agent right away. Start with a narrow human-in-the-loop scenario: one unit, one task, one user with an explicit right to roll back. Sinch's 74% rollback rate is the lesson learned by those who went straight for autonomy.

If you are building an AI product, reassess your positioning. "An open chat for everything" is a 2023–2024 category. "A box for a specific job, with an explicit human in the loop and strong security communication" is the category Anthropic and Yandex have just moved into. If you are closer to the first paradigm, you have 12 to 18 months to make the turn.

If you are writing an AI strategy for 2027–2028, budget on the assumption that security, trust, and compliance will cost as much as development itself. 84% of AI teams are already working to exactly that proportion.

## How I Want to End This Book

If the preceding chapters and the conclusion were the argument for a forecast, the spring of 2026 was its factual confirmation. AI is moving out of the mode of "a magical agent that will do everything for you" and into the mode of an off-the-shelf assistant that amplifies a specific person under that person's control. The main segment is SMBs. The main form is a narrow workflow with a human in the loop. The main competitive advantage is not model quality but security and trust.

This turn could have been anticipated a couple of years ago, drawing on common sense and the logic of how technology develops. Whoever saw it earlier and started building a product for this paradigm rather than for "the

autonomous agent that replaces the human" will enter the next wave already on the right trajectory.

And this is how I would like to close the third edition of this book.

### The 7 Landing Principles

If the whole book had to come down to a single page, let it be this one.

**1. 70/30.** AI takes the routine; the human answers for context, constraints, and the decision. Remove the human entirely and you remove the quality.

**2. Data is the ceiling on value.** A model cannot be smarter than what goes into it. Garbage in, confident garbage out. Remember that every time you work with AI.

**3. The quality of the answer equals the quality of the brief.** AI does exactly what it was asked to do. Data plus a well-formed task is the basis for success.

**4. The cost of an error sets the degree of formalization.** For ideas, free-form dialogue; for decisions and money, structure, verification, and a human in the loop.

**5. The model is a component; the product is the value.** Value for the user is created not by the model but by scenarios and by solving the user's problems. You cannot build products around a single model. Models will be restricted, and models will change.

**6. Specialization beats scale.** On an applied task, a tuned "small" model packaged into a working scenario outperforms a general-purpose giant.

**7. What wins is not the technology but the person who directs it and uses it.** How a person uses AI is what determines whether that experience succeeds.

Come down to earth—and build.


### What's Next

This book is about understanding AI. If you have read this far, you already have the picture: what the technology can do, where its limits are, and how it fits into management. From here there are three routes, depending on what you need.

**If you are ready to deploy,** you need a methodology: use-case selection, readiness assessment, pilot, scaling, effects. That is AI-2. Practical guide, the second book in the AI series.

**If you feel your company lacks a digital foundation,** start with the Digital transformation for CEO and owner series: Part 1 is an immersion in the technology, Part 2 is a systems approach to transformation.

**If your main question is security,** the third book in that series is devoted to cybersecurity, from social engineering to vulnerability management.

My articles, materials, and ways to reach me are at chelidze-d.com. I would welcome feedback on the book: it is, as you now know, the shortest learning loop there is.

# Appendices

## Appendix A. Glossary of Key Terms

| Term | Definition |
|---|---|
| **AI (artificial intelligence)** | Any mathematical method that imitates human or other natural intelligence. |
| **Machine learning (ML)** | Statistical methods that get better at a task as experience accumulates and the model is retrained. |
| **Deep learning (DL)** | A branch of ML built on multilayer neural networks; the foundation of today's models. |
| **Generative AI (GenAI)** | AI that creates new content—text, images, code, audio—on request. |
| **LLM** | A large language model, pretrained on vast bodies of text. |
| **Prompt** | A text instruction sent to a model. |
| **Token, tokenization** | A token is the unit of text a model operates on (part of a word); tokenization is the splitting of a request into those units. Limits and costs are counted in tokens. |
| **RAG (retrieval-augmented generation)** | Connecting a model to trusted sources so it can pull in current facts; reduces hallucinations. |
| **Fine-tuning** | Adjusting a model to a particular style, dataset, or task. |
| **Few-shot / zero-shot** | A model working from a handful of examples, or from a text description of the task alone. |
| **System prompt** | An instruction that sets the role, the rules, and the format for every request to an assistant. |
| **Prompt injection** | Manipulating a request in order to bypass filters or trigger unintended actions. |
| **Hallucinations** | Plausible but false content generated by AI. |
| **AI assistant** | A service that works on request and offers recommendations. |
| **Copilot** | A solution that automates individual tasks under human supervision. |
| **AI agent** | AI that acts autonomously and interacts with other systems. |
| **Multi-agent system** | Several agents coordinating their work with one another. |

| | |
|---|---|
| **AI orchestrator** | A top-level AI that breaks a request down and distributes the subtasks among models. |
| **Narrow AI (ANI)** | AI specialized for specific tasks—everything that exists today. |
| **Strong, or general, AI (AGI)** | A hypothetical AI that solves any intellectual task on a par with a human. |
| **Superstrong AI (ASI)** | A hypothetical superintelligence that surpasses humans in every domain. |
| **Multimodality** | Processing text, audio, images, and sensor data at the same time. |
| **DSS** | A decision support system; a digital advisor. |
| **Shadow AI** | Employees using AI unofficially, outside any approval process or data policy. |
| **Data-driven / informed / inspired** | Decisions driven by data (operations) / informed by data (tactics) / inspired by ideas found in data (strategy). |
| **Baseline / KPI** | The starting measurement and the target metrics used to judge the effect of a solution. |
| **On-premises / SaaS / IaaS / PaaS** | Deployment on your own servers / a ready-made service / rented infrastructure / a rented platform. |
| **Digital twin** | A virtual copy of an object or process, used to model scenarios before decisions are made in reality; as it matures, the model is enriched with operating data and runs in sync with the original. The maturity levels are covered in DT-1. Immersion. |
| **Neuromorphic computing** | An in-memory computing architecture modeled on the brain; it saves energy. |
| **Qubit / superposition** | A quantum “bit” able to hold several states at once. |
| **SMB** | Small and midsize business—companies with up to roughly 250 employees; the main growth segment for AI agents in 2026. |
| **Context window** | The amount of information a model “holds in its head” within a single conversation. |
| **Inference** | A trained model at work—every answer to every request. The main line item in the cost of running AI. |

| | |
|---|---|
| **Reasoning model** | A model that "thinks" before it answers. Strong on complex problems, but more expensive and weaker at reproducing facts precisely. |
| **Compact model (SLM)** | A small, specialized model. On a narrow task it is often cheaper and more accurate than the general-purpose giants. |
| **Open model (open source)** | A model with open weights—you can deploy it inside your own perimeter, with no external provider. |
| **Machine vision** | AI for analyzing images and video: quality control, safety, equipment maintenance. |

## Appendix B. Checklist: Where You Need AI and Where Classic Automation Will Do

**An AI solution is the better choice when:**

- the task is highly complex or involves many variables, and the rules cannot be stated explicitly
- large amounts of data are available, and patterns can be found in them
- the work involves unstructured data—text, images, audio
- the system has to adapt quickly to new data without rewriting the rules
- the decision rests on probabilistic estimates

**Classic automation is the better choice when:**

- the task has limited complexity and limited variability
- interpretability, response time, and reliability are critical
- there is no data to train a model on

**Important:** this checklist is about the applicability of AI as a class of technology. Organizational readiness, the selection and prioritization of specific use cases, and the criteria for launching or stopping a project are all questions of implementation. They are covered in detail in the book AI-2. Practical guide—readiness assessment in Chapter 7, identifying and prioritizing ideas in Chapter 16, and red flags and the rule for stopping a project in Appendix 2.

## Appendix C. Your First Seven Days with AI: A Plan for Beginners

Seven days, fifteen to twenty minutes each—and AI stops being an abstraction. The rules for the week: do not paste in sensitive data; check the facts; save the prompts that work, because that is the start of your own prompt library.

**Day 1. The first conversation.** Sign up for one AI service and ask it: "Explain what my company does [description] as if you were talking to a high-school student." Then check what the AI understood and what it garbled.

**Day 2. The summary.** Send it a long email or article: "Five key points, plus what I should do about them." Compare that with your own reading.

**Day 3. The draft.** Have it write a work email using template no. 2 from Chapter 6. Edit the result—and notice how much time you saved.

**Day 4. Reading a document.** Assign a role ("you are a lawyer…") and ask it to find the risks in a public document or a contract template. Nothing confidential goes in.

**Day 5. Decomposition.** Break a real task from this week into steps (template no. 6) and ask the AI to ask you clarifying questions. Notice which questions you had not asked yourself.

**Day 6. Comparison.** Have the AI compare the two options you are currently weighing, in a table with criteria (template no. 8).

**Day 7. Your own scenario.** Pick one repeating operation, build a standing prompt for it, and block time on your calendar to use it. That is your first AI process in production.

## Appendix D. The Full Action Plan: 24 Steps from 24 Chapters

Every "What to do" box in the book, gathered into a single working plan. Use it as an implementation checklist: one row per chapter, check off what is done.

| Chapter | What to do |
|---|---|
| Chapter 1. Getting to Know AI | Before assigning a task to AI, run it through the "AI vs classic" test and check whether you have usable data. |
| Chapter 2. The Types of Artificial Intelligence | Do not base your plans on "imminent AGI"; bet on specialized models, and for agents put rules, resource limits, and escalation paths in place. |

| | |
|---|---|
| Chapter 3. What Narrow AI Can Do—and the Big Trends | Take on a genuinely painful problem, not a fashionable one, and agree in advance on how you will measure the effect; introduce a Shadow AI policy. For the method of selecting and prioritizing use cases and calculating the effect, see AI-2. Practical guide. |
| Chapter 4. Generative AI | Try AI yourself—on your own tasks, in free or low-cost services: it is the cheapest way to understand the technology. In the company, start with off-the-shelf solutions—per MIT, pilots built on an external tool customized to the task reach deployment about twice as often as in-house builds—roughly 67% against 33%. Do not cut people because of GenAI, and invest in a good spec. |
| Chapter 5. Large Language Models (LLMs) | For corporate tasks, connect the LLM to your own data through RAG, and do not build critical processes on a plain chatbot. |
| Chapter 6. Prompt Engineering: The Discipline of Setting Tasks for AI | Make B-R-I-E-F the corporate standard for setting tasks and start a shared library of proven prompts. |
| Chapter 7. Regulating AI | Start now—build AI-content labeling and logging into your systems, set a clear allocation of liability, and keep an eye on how your own industry is being regulated. |
| Chapter 8. AI Security | Legalize AI and give people a safe internal alternative, test for prompt injection, store system prompts as secrets, and give agents least-privilege access and a kill switch. If there are several agents—also a shared communication channel and an arbiter. |
| Chapter 9. Decision Support Systems, or Digital Advisors | Do not build or wait for a know-everything advisor. Start with a narrow one, specialized in your own procedures, and put your processes and your data in order before you deploy it. |
| Chapter 10. Neuromorphic and Quantum AI | Do not wait for a quantum revolution to handle your current tasks. Follow the news on neuromorphic adoption in autonomous devices, and invest in competencies today rather than in hardware. |
| Chapter 11. Strategy and Organizational Structure | Check whether AI appears in your strategy as a separate workstream with goals and metrics (one week). When you set |

| | |
|---|---|
| | up the AI function, design for a distributed model from the start and develop along three axes—technology, the management system, and people's competencies (one quarter). |
| Chapter 12. Business Processes | Pick one process with a clear input and a clear output and use AI to draft its description and its procedure (one week). Assess the maturity of your key processes before you plan AI projects on top of them (one quarter). |
| Chapter 13. Managing Projects, Change, and Products | Run your current project through the "7 Sins" list (one week). Introduce AI-written status and meeting summaries for one project (one quarter). |
| Chapter 14. Communication | Turn on AI summaries for one type of recurring meeting (one week). Introduce a standard for recording "decision—owner—deadline" (one quarter). |
| Chapter 15. Digital Technologies, Working with Data, and Cybersecurity | Audit the data sources for your first AI use case (one week). Draw up a road map for data quality: owners, formats, controls (one quarter). |
| Chapter 16. Regular Management Practices | Automate the creation of task cards from messages and meetings (one week). Connect an AI assistant to your team's tracker and to the status ritual (one quarter). |
| Chapter 17. Lean Manufacturing and the Theory of Constraints (TOC) | Take one process and map the eight types of waste across it (one week). Only after you have simplified it should you decide where AI belongs (one quarter). |
| Chapter 18. The Processes of Business Functions | Pick the function with the most painful routine work and one task within it, and record your baseline metrics before you start—that is your pilot candidate (one week). Run the pilot and compare the result against the baseline (one quarter). |
| Chapter 19. AI for Small Business and Entrepreneurs | Pick one recurring pain point—handling requests, documents, content—and hand it to an off-the-shelf service. A one-week pilot, no budget. |
| Chapter 20. Examples and Recommendations for Midsize and Large Companies | Pick two or three cases from this chapter that are analogous to your industry and compare them against your own processes—what infrastructure and data do you already have? |

| | |
|---|---|
| | (One week.) Then build a portfolio of three to five initiatives with priorities (one quarter). |
| Chapter 21. AI in the Leader's Own Work | Start a leader's journal—five minutes in the evening, an entry plus feedback from AI (one week). Then assemble the personal loop: planning, working through decisions, preparing for difficult conversations (one quarter). |
| Chapter 22. How to Run Technology Implementation Projects | Classify your next AI project by type and staff the roles (one week). Put a "single project file" in place, with AI-written status summaries (one quarter). |
| Chapter 23. A Systems Approach to Implementation and the Road Map | Sort your AI initiatives into those that depend on corporate data and those that do not, and launch the second group (one week). Pull the initiatives into a single map with the links to IT projects (one quarter). |
| Chapter 24. What a Leader Needs to Know and Be Able to Do in the Age of AI | Assess yourself and your key leaders against the matrix in this chapter (one week). Put AI literacy into the development program for your management team (one quarter). |