

# AI as a working model.

A working model that brings AI into the daily flow of an organisation, without damaging its culture or its craft. For organisations between 15 and 500 people.

---

AUTHOR	PUBLISHER	DATE	VERSION
Dusty Baars	Mono by Dusty	May 2026	v1.0

## CONTENTS

00 **Summary** Four core points of the method

I **The problem** Why AI often stalls between 15 and 500 people

II **The working model** Four lines, role map, zoom-in

III **Building blocks and rhythm** Skill and agent, two speeds, three phases

IV **Five cases** Performance marketing, services, SaaS, recruitment, finance

V **What stays human and governance** Including the EU AI Act up to 2 August 2026

VI **Where this is heading** Four shifts in the field in 2026 and 2027

· **Closing** Bio, an invitation to explore

## What this document argues, in four points.

In organisations of fifteen to five hundred people, AI often stalls in scattered tools and good intentions. The distance to the daily workflow stays large, and the gain it could deliver barely becomes tangible. This document describes a working model that closes that distance, in a way that fits the pace and the scale of such organisations.

The model is built from four connected choices. None of them is revolutionary on its own. The combination is what makes AI embed rather than hover.

01

### Start with roles, not with tools

Before a tool is chosen, we map per role which tasks lend themselves to automation, augmentation, agent execution or deliberate human work. That puts the focus on the workflow of the people doing the work, instead of on what vendors happen to offer.

02

### Work in building blocks, not in projects

A building block is a contained skill or agent with its own outcome definition. Done means the output is usable in the daily flow, not that a number of hours has been spent. A building block stands on its own, delivers value on its own, and can be made to work in five to seven working days.

03

### Two speeds run in parallel

Building happens in days, adoption in weeks. While building block one is being adopted, building block two is already being built. Waiting for a linear sequence turns a six-week effort into a three-month one for no reason.

04

### Deliberately protect what is human

Strategy, the client relationship, concept choices, final approvals and ultimate accountability stay in human hands. AI speeds up production, but it does not automatically raise the quality of the substantive work. By making that distinction explicit, an organisation holds on to what makes it itself.

In the rest of this document I work out those four choices. First I describe why AI so often stalls in mid-sized organisations. Then comes the working model itself, with the role map and the zoom-in as its core instruments. After that I explain the building blocks and the rhythm, followed by five cases from a range of sectors. Finally I turn to what stays human, and what introducing this model asks of an organisation.

## PART I

# Why AI so often stalls in mid-sized organisations.

In conversations with founders, COOs and department heads I hear almost the same picture every time. The licences have been bought, part of the team is trying ChatGPT or Copilot, there is an internal channel where people share prompts. And yet little changes in the way the work gets done day to day. The gain most vendors hold out remains a feeling, not a measurable result.

The cause rarely lies in the tools and almost always in the distance between tool and workflow. An AI tool that sits separately next to the existing process is a new browser tab. People open it when they think of it, use it for one-off questions, and close it again. The work itself, with its fixed handovers, formats, deadlines and quality requirements, stays unchanged. The tool delivers a few minutes of gain per person per day. For the organisation as a whole, that gain shows up nowhere in the numbers.

The distance rarely sits between people and technology, and almost always between tool adoption and workflow change. Tool adoption is the phase in which people learn what they can do with a tool. Workflow change is the phase in which the process itself has been redesigned, with the AI component as a fixed part of it. Most organisations in the scale of fifteen to five hundred people have gotten stuck somewhere halfway through phase one.

## Why this scale in particular

An organisation of fifteen to five hundred people sits in an awkward middle position. Too small to employ its own AI team that can build full time, too large to pull the whole company along on the loose experiments of a few enthusiastic employees. Large companies have a central capability that can invest in building, testing and rolling out. Small companies have few processes for AI to land against and solve a lot ad hoc. In between lies the hardest territory: enough complexity to gain from AI, and not enough capacity to organise it themselves.

The external support available in this territory rarely fits the scale. Strategy consultancy delivers thick reports that have captured the problem on slide fifteen and leave the implementation to the organisation itself. Implementation partners build from their own stack and treat the workflow as a given to fit into, rather than as the design question it is. Both approaches leave an organisation with the same gap it had before the start.

### THE OBSERVATION

The question at this scale is not which tool to choose. The question is how the work in the department itself is redesigned, with AI as a fixed part of the flow instead of an optional tab beside it.

## Symptoms I come across in almost every organisation

When tool adoption has started but the workflow has not, a number of recognisable patterns appear. I name them here because recognising them is often the starting point of a level-headed conversation about what is needed.

The first pattern is that AI use becomes person-dependent. A few employees use it intensively and gain a lot on their own work, while colleagues in the same role barely touch it. The work the company delivers becomes inconsistent rather than consistently better.

The second pattern is that productivity gains cannot be traced. People are convinced they work faster, but the lead times in the operation stay the same. The

gain leaks away into extra rounds, into interim output that nobody uses, or into time lost somewhere else in the process.

The third pattern is that AI use lands mostly in production and rarely in the chain before or after it. The copywriter has AI make first drafts, but the briefing he receives and the review that follows have not been adjusted. The building block stands alone, and the flow around it no longer fits.

The fourth pattern is that AI lands mostly on individual tasks and rarely on decision-making or coordination. That produces many small optimisations without the broader coordination improving. The department does not get better as a whole, only visibly faster in a few places.

Anyone who sees these patterns in their own organisation need not conclude that adoption has failed. It usually means that the transition from tool adoption to workflow change has not yet been made explicit. The rest of this document is about that.

## PART II

### **The working model.**

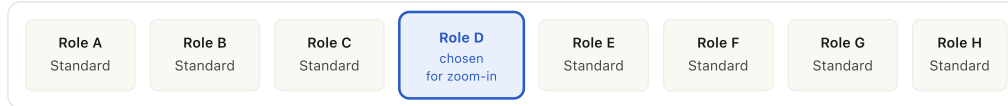
The working model consists of three instruments used alongside each other. The four lines give direction to where AI lands in an organisation. The role map makes visible where within roles the opportunities lie. The zoom-in gives the approach per role concrete shape. Each instrument is usable on its own, and together they form the basis for a workflow change that is achievable and measurable.

## THE WORKING MODEL IN ONE IMAGE

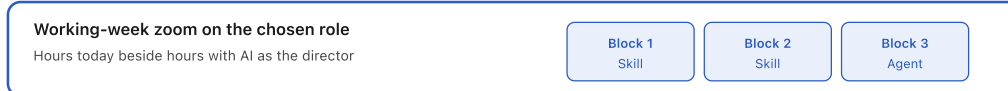
### INSTRUMENT 1 · FOUR LINES



### INSTRUMENT 2 · ROLE MAP



### INSTRUMENT 3 · ZOOM-IN



The three instruments of the working model run from broad to deep. The four lines give direction, the role map shows where within roles there is gain to be had, and the zoom-in makes concrete, for one chosen role, which building blocks land first.

## Instrument 1, the four lines

In an organisation of fifteen to five hundred people, AI lands on four connected lines at once. The lines run in parallel and reinforce each other when they are addressed at the same time.

### Production

The work that currently takes many person-hours to make. Copy, content, variants of something existing, report drafts, briefings, contracts and standard documents. The most visible time gain sits here, because the output is directly comparable to the existing situation.

### Analysis

Gathering, cross-referencing and summarising data. Noticing patterns before clients or colleagues do. Reports whose lead time goes from hours to minutes. Here the work shifts from data gathering to interpretation.

### Sales and onboarding

Client and proposition work that now takes a lot of manual effort. Account briefings, proposition adjustments, first client research, onboarding flows that feel personal without being assembled by hand. Here the organisation gains consistency and lead time at once.

### Optimisation

Continuous monitoring and hypothesis forming. Agents that flag and propose, so specialists can spend their time on execution and strategic choices. Here the way the work is steered changes: no longer reactive to client questions, but ahead of signals that become visible earlier.

In practice an organisation rarely chooses to address four lines at once. I usually advise one line to start, sometimes two, with a clear priority for the line that pinches most in the current operation. That is usually production or analysis. Sales and onboarding often follow in the second round, optimisation last, because it benefits most from the structure the other lines build up.

## Instrument 2, the role map

The role map is the second instrument and in practice the most discussed. It is a simple overview of the roles within a department or organisation, with the core tasks of each role classified. Every task gets a classification that indicates how AI can land on that task.

● Automate · rule-based, no human needed

● Augment · AI helps, human decides

● Agent · AI does, human reviews

● Human · stays human work

The map has three advantages that make it valuable in practice. First, it makes the conversation between leadership and team concrete. Instead of a general discussion about AI possibilities, we talk about specific tasks people do today. Second, it separates the roles that can structurally gain a lot from the roles that have barely any overlap. Finally, it produces a working document that can keep being refined after a first pass, rather than a report that closes and then disappears into a drawer.

A role map is rarely filled in completely in one session. The first version comes from what is publicly visible, possibly supplemented with a few short conversations. The second version follows after interviews with the role holders themselves, because reality always turns out to be a little different from what the org chart suggests. From the third version on, the map becomes a living document that is updated regularly during the engagement.

### **Instrument 3, the zoom-in**

The zoom-in is the instrument that links the role map to concrete building blocks. For a chosen role, an average working week is mapped, with the current time spent per core task. That same week is then recalculated in a scenario where AI lands on the relevant tasks. The difference that becomes visible is the basis for the conversation about which building blocks return the most.

The zoom-in has a specific function within the working model. It forces us not to talk about the whole organisation at once, while making it concrete enough to take decisions. A department head sees which hours go where, and which hours might free up. A team member recognises their own week and can point out where it pinches. A management team sees how one role relates to the broader capacity of the organisation.

#### THE CHOICE IN THE MODEL

Not every role is a logical place to start. I choose the role where the capacity pressure is greatest, where the work most visibly consists of sub-tasks, and where building a first building block in five to seven working days is realistic. Whatever turns out to work in the first role almost always turns out to have value in adjacent roles too.

In the cases later in this document, a chosen role is worked out for each example. They show how the zoom-in works in concrete terms: which hours sat where, which building block lands on it first, and what the outcome definition of that building block is. In most cases the pattern turns out to be recognisable, even though the sectors differ strongly from one another.

## Building blocks and rhythm.

A building block is the smallest form in which AI can land in a workflow. It is contained, it has its own input and output, and it has an outcome definition that describes in concrete terms when the building block is done. Working in building blocks rather than in projects or time packages creates a rhythm that fits how AI actually gets built and how organisations actually adapt.

### Two forms, with a different division of roles

Within the working model I use two forms of building block, with a distinction that turns out to matter in practice.

#### FORM 1 · SKILL

##### A contained capability on request

Trigger	The person calls the skill at the moment they need it.
Input	A briefing, a dataset, a document or a question the person supplies.
Output	A contained deliverable that fits directly into the workflow, for example a set of variants, a first draft or a structured analysis.
Human role	Initiates, supplies input, assesses output, decides on use.
Examples	Ad Copy Generator, Contract Reviewer, ICP Matcher, Brief Builder, KYC Triage.

## A continuous process that acts on its own

Trigger	Time- or event-based, not called by a person for each action.
Input	Data from running systems, monitored signals or incoming requests.
Output	A digest, a routed request, a structured alert or a prepared action awaiting approval.
Human role	Sets the boundaries and the quality, reviews periodically, steps in on exceptions.
Examples	Performance Digest, Customer Health Monitor, Compliance Triage, Inbox Router.

The distinction seems small, but it decides a lot. A skill carries the rhythm of the person. An agent carries its own rhythm and asks the organisation to adapt to it. For most organisations it is wise to start with skills, because they can be placed into the existing workflow without the flow itself having to change. Agents usually come in a second round, once the organisation has learned to work with AI output and is ready to hand parts of the work over to an agent.

## Outcome definition over hours

A building block is defined in advance on outcome, not on time spent. For every skill or agent I set out beforehand what done means, expressed in concrete observable terms. For an Ad Copy Generator that is, for example, that eighty percent of the generated copy is usable without rewriting. For a Performance Digest, that it is ready on time every morning and that specialists actually base their work on that digest. For a Contract Reviewer, that a standard NDA is reviewed in fifteen minutes with its risk detection preserved.

Choosing outcome definition has three consequences I deliberately seek out. The organisation pays for a result, not for the time it takes to reach that result. The building block comes under pressure to be usable rather than to look complete. And the conversation between builder and organisation becomes more concrete, because there is a shared definition of when a building block does its job.

## HOW MONO WORKS

I do not sell time packages of three, six or twelve months. For every building block, what done means is agreed in advance. A practical time indication is perfectly possible alongside that, but the agreement is about the outcome, not about the hours.

### Two speeds run in parallel

The rhythm of an engagement is set by two speeds that run alongside each other and have a different character. One speed is tech-driven, the other people-driven.

#### Building runs in days

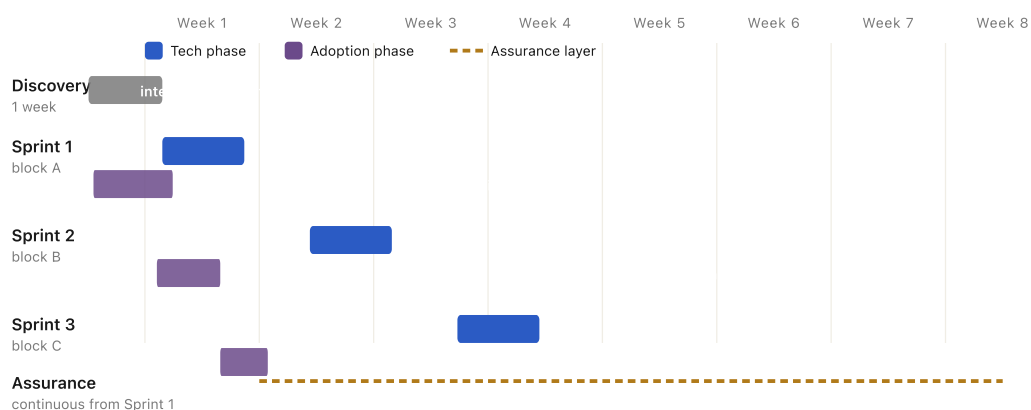
A working skill or agent can be made operational in five to seven working days. On day one or two there is a proof of concept. At the end of the first week a refined version follows. In the second week the building block is production-ready. After that, the ball is no longer with the build.

#### Adoption takes weeks

Getting a team to actually work with a new building block usually takes two to three weeks. Getting used to it, edge cases that surface, feedback from daily use and small adjustments all take time. That is part of human change and cannot be sped up by building better.

The difference between those two speeds sets the cadence of the whole engagement. While building block one is in adoption with the specialists, building block two can already be built. Waiting for a linear sequence turns a six-week engagement into a three-month one for no reason. By letting the speeds run in parallel on purpose, the pace keeps matching the absorption capacity of the team instead of the build capacity of the partner.

## TWO SPEEDS RUN IN PARALLEL, NOT IN SEQUENCE



Three building blocks end to end in six to eight weeks. The tech phase of the next building block starts during the adoption of the previous one. Assurance runs continuously from Sprint 1, rather than as a separate phase at the end.

## The three phases of an engagement

Within the working model, every engagement has three phases that follow each other. Discovery is short and focused. Sprint is the phase in which building blocks are delivered, at pace. Assurance runs alongside from the first building block rather than as a separate phase at the end.

### Discovery

One working week. Interviews with leadership and a sample of role holders, a walkthrough of the workflow, an audit of the tool stack and a data quality check. By the end of the week there is a confirmation of the first three building blocks, or an adjusted scope when the facts call for it. The result is a clear conclusion on which Sprint 1 can start, without an extensive interim report.

### Sprint

Per building block, five to seven working days of tech, two to three weeks of adoption, with overlap. Sprint 2 starts as soon as the adoption of Sprint 1 is on track. The phase is complete once the outcome definitions of the building blocks have been met. Six to eight weeks end to end for three building blocks is a normal lead time.

### **Assurance**

Continuous from Sprint 1. A prompts library, governance agreements on who sees what and where which data ends up, review steps on output, documentation and training material. By the end of the run, this belongs to the way the organisation works on its own, independent of my involvement.

### **Handover**

At the end of Sprint and Assurance, there is no package left behind that only I can maintain. The skills and agents are documented, the prompts library sits in a central place, and there is an internal owner who takes on the management. From that point on, the organisation can formulate building blocks four and five itself, or do that together with me.

## **PART IV**

### **Five cases, from five sectors.**

The following five cases show the working model applied to a range of sectors. Each case highlights the chosen role, describes a building block that lands first, and gives a mini zoom-in on the effect such a building block has. The cases are assembled from Mono projects, sector conversations and public information. They illustrate the pattern, and are not meant as a promise of an identical outcome in every organisation.

## CASE 01 · PERFORMANCE MARKETING

# A performance agency that no longer solves growth and capacity pressure by hand alone.

### SECTOR

Performance marketing

### CHOSEN ROLE

SEA Specialist

### BUILDING BLOCK

Ad Copy Generator

### FORM

SKILL

A performance agency lives on hours delivered against the responsibility it carries for client results. Every hour that production costs is an hour not spent on strategy, client contact or an additional account. In this case the SEA role is chosen as the entry point, because the capacity pressure is greatest here and the platforms themselves are already automating heavily within the discipline.

The building block that lands first is an Ad Copy Generator. The input consists of a product briefing, USPs, tone of voice and an audience description. The output consists of thirty responsive search ad variants, ready for implementation in Google Ads, plus ten display headlines. The SEA Specialist reviews the variants, picks the top five and makes the final adjustments where needed. Done means that eighty percent of the generated copy is usable without rewriting.

#### NOW, PER CAMPAIGN

## 2 hours

Brainstorm, writing, alternating between variants, quality check and final review by the specialist.

#### WITH THE BUILDING BLOCK

## 15 minutes

Supply the briefing, assess thirty variants, pick the top five, make the final adjustments. The review shifts from production to direction.

With three building blocks on the SEA side, used by six specialists, the agency-wide gain sits around **30 to 48 hours per week** of capacity that now goes to production and could later go to direction, strategy and

additional accounts. On an annual basis that comes down to effectively one FTE of capacity, without needing a new hire for it.

## CASE 02 · PROFESSIONAL SERVICES

# A law firm that lifts away standard work without losing its risk detection.

### SECTOR

Law, legal advisory

### SIZE

35 people

### CHOSEN ROLE

Senior Associate

### BUILDING BLOCK

Contract Review Agent

### FORM

AGENT

A mid-sized law firm carries two different workflows at once. On one side the bespoke work that differs with every matter, where the craft of the lawyers comes through. On the other side a considerable amount of standard work, with NDAs, confidentiality clauses, simple supplier contracts and repetitive provisions. The standard work often lands with the Senior Associate, who performs the detailed checking that a legally sound review requires. Forty-five minutes per NDA is not unusual.

The building block that lands first here is a Contract Review Agent. The agent receives the incoming contract, compares it against the firm's annotated clause library, flags deviations by risk category, proposes suggestions and delivers a briefing for the Senior Associate who does the final review. Done means that a standard NDA is reviewed in ten minutes, with the risk detection preserved as the firm defines it.

#### NOW, PER NDA

**45 minutes**

Reading through, comparing with the house standard, flagging deviations, drafting change proposals.

#### WITH THE BUILDING BLOCK

**10 minutes**

Going through the agent's briefing, assessing the flagged clauses, final review and sign-off on the change proposals.

For the firm, the time gain matters less than what is done with it. The freed-up room goes to the work where the margin return is highest, and where the firm's substantive distinctiveness comes into its own. For

partners there is often a second effect here, because repetitive standard work is a form of wear that measurably decreases once it is automated.

## CASE 03 · SAAS SCALE-UP

# A B2B SaaS that sees churn coming two weeks earlier than before.

### SECTOR

B2B SaaS

### SIZE

120 people

### CHOSEN ROLE

Customer Success  
Manager

### BUILDING BLOCK

Customer Health Digest

### FORM

AGENT

In a SaaS scale-up of around a hundred people, the Customer Success department almost always has the same challenge. The account portfolio per CSM is larger than they can hold in their head day to day, and the signals that point to a coming churn are scattered across several systems. Product usage in the database, ticket volume in the support tool, NPS scores in the feedback tool, deal status in the CRM and the account team's channel in chat. In practice, churn usually only becomes visible once a client has already decided to cancel.

The building block that lands first here is a Customer Health Digest. The agent cross-references the signals from the relevant systems daily, calculates a health score per account that takes sector and account size into account, and delivers a morning digest per CSM with the three accounts that need the most attention and the three accounts where the score is improving. Done means that churn risk is structurally visible two weeks earlier, and the average response time to a flagged signal drops from twenty-four to four hours.

#### NOW, AVERAGE SIGNAL TIME

## 24 hours

Signals scattered across systems, noticed at the moment the CSM happens to be in that tool or a colleague points it out.

#### WITH THE BUILDING BLOCK

## 4 hours

Morning digest at 08:00 with three prioritised accounts. The CSM starts the day with a list, instead of an empty screen.

What the organisation gains here is not only time, but a head start in time. A client conversation that begins two weeks earlier has fundamentally different outcomes than the same conversation two weeks later. For a SaaS organisation of a hundred people, a few percent less churn translates directly into monthly recurring revenue, and into a higher lifetime value per account.

## A recruitment agency that halves sourcing time without losing quality.

SECTOR	SIZE	CHOSEN ROLE
Recruitment, staffing	50 recruiters	Recruiter
BUILDING BLOCK	FORM	
ICP Matcher	<b>SKILL</b>	

In a recruitment agency, a considerable part of the daily work consists of matching incoming vacancies to available candidates. A recruiter reads a job description, forms a picture of the ideal candidate profile, searches LinkedIn, the agency's own database and adjacent sources, and draws up a first list of potential candidates. A large part of the lead time of an assignment rests on this sourcing phase.

The building block that lands first here is an ICP Matcher. The skill reads the job description, extracts the hard and soft criteria, searches the connected sources and delivers a first long list of fifteen to twenty candidates, each with a rationale for why the match is relevant. The recruiter assesses the list, cuts where needed and builds the short list for the client from this long list. Done means that sourcing time per assignment has been halved, while the interview rate per submitted candidate stays the same or rises.

NOW, SOURCING PER ASSIGNMENT

**6 hours**

Reading the vacancy, forming the profile, searching across several sources, assembling the long list, making the first cut.

WITH THE BUILDING BLOCK

**3 hours**

Supplying the vacancy, assessing the long list, making the qualitative cut based on the rationale provided.

The gain sits in the lead time per assignment and in the volume the agency can handle without extra hires. For an agency of fifty recruiters, halving the sourcing time means substantially more assignments can run in parallel,

while the quality of the submitted candidates stays at the same level. The substantive conversation with client and candidate stays human work, and gains room.

## A wealth management firm that brings down KYC lead time and steers specialists toward the real exceptions.

SECTOR	SIZE	CHOSEN ROLE
Wealth management, financial services	80 people	Compliance Officer
BUILDING BLOCK	FORM	
KYC Triage Agent	AGENT	

In a wealth management firm, KYC lead time structurally demands a lot of capacity. A Compliance Officer assesses every new client file on identification, source of funds, sanctions list check, PEP status and the appropriate risk classification. A considerable share of the files follows a standard pattern that, with the current staffing, still gets the same attention as the genuinely complex cases. The average lead time is therefore carried by the simple cases, and the complex cases do not always get the attention they deserve.

The building block that lands first here is a KYC Triage Agent. The agent receives the incoming file, runs the standardised checks, compares against the firm's risk matrix and classifies the file as standard, complex or blocked. A standard file is given a structured summary, ready for final sign-off by the Compliance Officer. A complex or blocked file is routed to the right specialist, with the concerning elements up front in the briefing. Done means that eighty percent of incoming files get a standard sign-off within fifteen minutes, and the remaining twenty percent land with the specialist with sharper context.

NOW, AVERAGE LEAD TIME

**2 days**

Every file goes through the same steps, regardless of complexity. The queue sets the pace, not the substantive work.

WITH THE BUILDING BLOCK

**15 min · 2 days**

Standard files done within fifteen minutes, complex files keep their original attention and lead time,

but land directly with the right specialist.

For a firm in this sector, governance matters as much as efficiency. The agent delivers an audit trail with every file, showing the checks that were run and the signals that were weighed. The final signature stays with the Compliance Officer, and with blocked files always with a human. The gain sits in the combination of faster turnaround for standard work and sharper context for specialists on the cases that genuinely matter.

#### WHAT THE CASES SHOW

The cases differ in sector, but show the same pattern every time. A chosen role, a contained building block on top of it with an outcome definition formulated in observable terms, and as a result a shift from production to direction within the workflow. The working model is sector-agnostic, and that is exactly why it fits the scale of fifteen to five hundred people this document focuses on.

#### PART V

## **What stays human, and what this asks of an organisation.**

The working model only works when two questions are taken as seriously as the build question itself. The first question is about what deliberately stays human work, even though AI could technically take it over. The second question is about what introducing this model asks of an organisation in terms of governance, ownership and steering.

## What stays in human hands

AI speeds up production, but it does not automatically raise the quality of the concept or the character of the organisation. The following parts therefore stay deliberately human work, on the conviction that this is exactly where distinctiveness comes from.

### STRATEGY

The choice of which direction to take for which client or market, depending on phase, sector and context.

### CLIENT RELATIONSHIP

Building trust, the personal contact and the conversations in which an unexpected insight emerges.

### CONCEPT AND CREATION

The substantive direction of a campaign, proposition or product, and the voice in which it is told.

### APPROVALS

The final decision on spend, copy, publication and execution. That signature stays with people.

### TEAM AND CULTURE

Hiring, coaching, development and the energy that keeps the organisation the kind of place where people like to work.

### ULTIMATE ACCOUNTABILITY

The choices that have to be made when something goes wrong. That accountability belongs to people, not to an agent.

Making these parts explicit serves a purpose that goes beyond protecting them. It also brings clarity for the rest. When team and organisation know where the human stays indispensable, room opens up to look with more confidence at the places where AI can land. The conversation then becomes about design, not about fear.

## What this asks of an organisation

A working model rarely works just because it is right on paper. Introducing it asks four things of the organisation itself, separate from the builder on the other side of the table.

The first is an internal owner. A building block that rests on me is a dependency for the organisation. A building block for which an internal owner has been appointed becomes part of the company itself. That owner does not need to be the sharpest

prompt engineer, but does need a mandate to take decisions about scope, quality and further development.

The second is an agreement on governance. Who sees which data, where that data ends up, who has access to an agent's logs and who signs off on exceptions. For most organisations at this scale that is not a complicated document, but it is a topic that belongs on the table in advance. Privacy, traceability and liability become concrete through it, rather than abstract.

The third is a review rhythm. A building block in production should be held up to the light periodically. Does it still work as agreed, does it still deliver the outcome it was built for, is it drifting from its original quality. A short monthly review with the internal owner and the specialist who works with it daily is usually enough.

The fourth is a willingness to adjust the workflow itself. A building block that gives a specialist back two hours a week only delivers value when those two hours go somewhere else. That seems obvious, but in practice it is the step most often underestimated. The freed-up time calls for a conversation about what happens with it, before it leaks away into extra rounds or new meetings.

## **Connecting to the EU AI Act, with 2 August 2026 as the key date**

The EU AI Act has been entering into force in phases since August 2024, and on 2 August 2026 it passes a milestone with direct consequences for organisations at the scale of this document. On that date the transparency obligations of Article 50 become fully enforceable, all member states must have a working AI sandbox in place, and providers of high-risk AI systems face a mandatory conformity assessment before a system reaches the market. Most organisations in the audience of this whitepaper are deployers, users that is, and not providers of AI systems, but deployers too acquire their own set of obligations on that date.

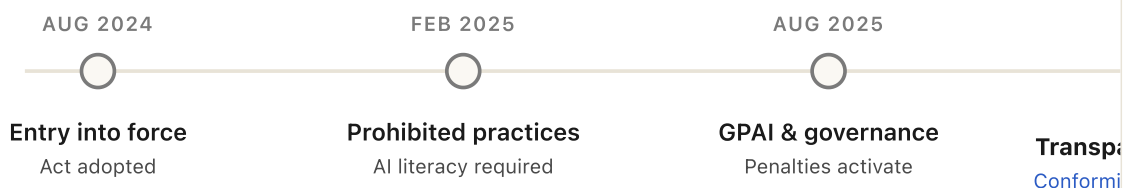
The deployer obligations that take effect touch four domains at once. First, keeping system logs of high-risk AI systems, so that traceability remains possible afterwards. Second, transparency toward the people the system affects, particularly in HR, credit and public services. Third, technical and organisational measures to use the system the way the provider intended. Fourth, for organisations operating as a public entity or a public-service provider, a Fundamental Rights Impact Assessment before putting a high-risk system into use.

**Recent development.** In early 2026 the EU institutions reached a provisional political agreement on the Digital Omnibus on AI, which postpones part of the high-risk obligations under Annex III to December 2027 and embedded AI under Annex I to August 2028. This agreement still has to be formally adopted. The transparency obligations, the sandbox requirements and the provider obligations for conformity assessment stay unchanged on 2 August 2026. For organisations in the audience of this document, the practical conclusion is that there is more time to set up the governance side of AI use, and no extra time to put off starting on it.

For the cases in Part IV specifically, the Act is directly relevant in two places. Recruitment applications fall under Annex III category 4 (employment, workers management and access to self-employment). KYC and creditworthiness applications fall under category 5 (access to essential private services). Both are classified as high-risk, with the corresponding deployer obligations. For the other three cases the requirements stay limited to the transparency layer of Article 50, with the note that copywriting output presenting itself as human work must be marked as such from 2 August 2026.

The working model connects to these requirements without needing a separate compliance layer on top. Because every building block has an outcome definition, an internal owner, a review rhythm and an audit trail in the logs, the documentation the Act requires is largely already present. For organisations still unsure where to start, this offers a second reason to move in this direction. The setup that delivers value for productivity is the same setup that makes them ready for the regulation.

#### EU AI ACT TIMELINE, FOR ORGANISATIONS AT THIS SCALE



The solid blue node is the next date that matters for organisations at this scale. The dashed line marks the horizon by which the governance side needs to be in place.

#### THE CONNECTION

A working model in which AI lands deliberately is usually also a working model in which AI lands responsibly. Outcome definitions, ownership and review rhythms are not only instruments for productivity, but also for governance. Taking the workflow question seriously solves a considerable part of the governance question at the same time.

#### PART VI

## Where this is heading.

Designing a working model for today only makes sense if it still holds tomorrow. In this part I describe four shifts that visibly shape the field in 2026 and 2027, and how the working model is positioned for them. The observations come from conversations with founders and operators, recently published research and Mono's own building practice.

### **Shift 1, from scattered tools to connected agents**

Until 2024, AI in companies was largely a collection of separate tabs. A copywriting tool here, a meeting summarizer there, a datasheet AI in the browser. From 2025 on, that collection begins to take on structure. Open standards such as the Model Context Protocol, which Anthropic donated to the Linux Foundation in late 2025, make it possible to connect AI systems to data sources and internal tools in a standardised way. That changes what a building block is. It is no longer an isolated skill in a browser, but a component that plugs into the organisation's existing data and applications and behaves there as a fixed part of them.

The practical consequences are considerable. An Ad Copy Generator no longer has to be fed product information by hand every time. A Customer Health Digest no longer needs a separate ETL pipe to the CRM. The integration layer, for years the most expensive and most fragile piece of every AI implementation, shifts to a lighter and reusable level. Organisations that design their building blocks with a standard like MCP in mind from the start avoid an expensive rebuild twelve months later.

## Shift 2, from software seats to completed work

The second shift runs deeper and touches the commercial model of a whole generation of SaaS companies. Until now, software mostly sold access to users. An organisation bought a hundred seats, and the value lay in what those hundred people then did with the tool. Vertical AI agents turn that logic around. What such an agent delivers is the completed work itself, instead of the ability to do that work. A KYC Triage Agent delivers finished files, a Performance Digest Agent delivers prepared morning briefings. The invoice is thereby tied to an outcome the buyer can check, instead of to a licence that may or may not be used.

For organisations in the audience of this document, that shift has two sides. On the purchasing side it means that a growing part of what used to be bought as software now presents itself as an operational service with a fixed outcome. That changes how budgets are structured and how vendors are chosen. On the other side, new opportunities open up for the organisation itself to place internal building blocks exactly where generic SaaS no longer fits. The choice to build your own agents around your own workflows, instead of accepting what the market offers, becomes realistic for the first time at the scale of fifty to five hundred people.

### RESEARCH 2026

Gartner expects that by the end of 2026 around **40 percent** of all enterprise applications will have task-specific agents embedded, compared to less than **5 percent** in 2025. For mid-sized organisations, the distinction between what they buy now and what they will get built in is therefore less important than the question of who designed the workflow around it.

### **Shift 3, mid-market runs ahead and hits the governance wall**

A surprising result from recent research is that mid-sized organisations adopt agentic AI faster than large enterprises in 2026. The reason is obvious. There are fewer approval layers, the distance between decision and execution is shorter, and turnkey platforms from Salesforce, Microsoft and specialised newcomers make building more accessible without an in-house platform team. At the same time, the same group also shows the highest abandonment figures. Partial deployments strand more often, and the gain that was visible at pilot level returns less often in the structural numbers.

The cause is the same in all the research: the governance gap. Sixty-eight percent of organisations point to weak governance as the primary reason AI agents do not scale beyond the pilot phase. At large enterprises, that gap is usually closed by internal compliance and data teams. At mid-sized organisations that capacity is not present, and the gap becomes a wall as soon as the first real incident occurs or the first audit comes knocking.

The combination of head start and pitfall makes 2026 and 2027 a window for mid-sized organisations with a high return, and a high price for those whose foundation is not in order. Anyone who puts the build speed to work now without building the assurance alongside it will, within eighteen months, have to redo the entire setup. Anyone who sets up both at once takes a structural lead over both larger and smaller players in the same market.

### **Shift 4, from advisor to architect who builds alongside**

The fourth shift touches directly on the way external parties take part in this kind of work. Until now, the norm at this scale was that strategy consultancy sat at the front and implementation partners at the back. Between them lay a translation problem that became visible in almost every engagement. The reports from the front did not fit the practice at the back, and what the back built often diverged from what the front had conceived.

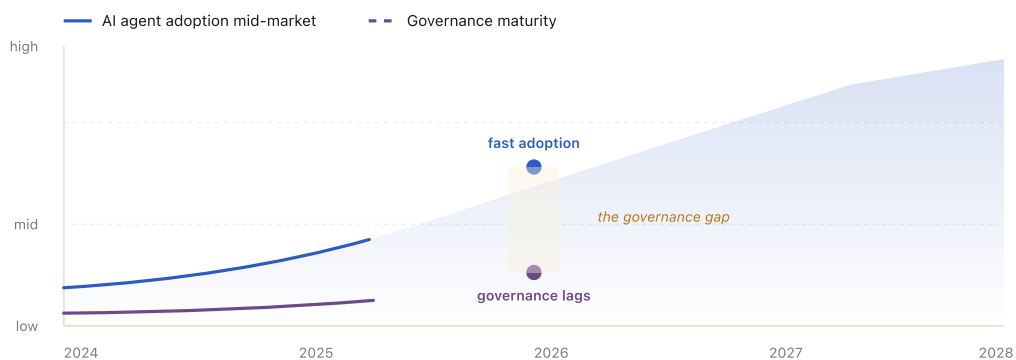
In a world where building takes days and adoption takes weeks, that distinction falls away. What remains is a role that designs and builds in one continuous movement, with the organisation in the lead. For the Mono practice that is not an announcement, but the way of working described in this document. For the wider market it is a shift that is slowly becoming visible and will reach the mainstream in 2027.

## How the working model is positioned for this

Four shifts, four ways in which the working model from this document is well positioned. Standardisation through MCP and comparable protocols fits the building-block way of thinking seamlessly, because a building block is by definition a contained component with clear input and output. The shift from seats to completed work coincides with the outcome definition by which every building block is delivered within the working model. The governance wall that mid-market organisations run into largely disappears through the combination of internal owner, review rhythm and audit trail the working model produces. And the shift from advisor to architect who builds alongside is exactly the role Mono fills in projects.

That a working model happens to fit four future shifts is not luck. It is the result of designing from the workflow rather than from the tool, and from the outcome rather than from the hours. Anyone who chooses a working model based on those principles now is choosing, at the same time, a working model that still holds two years from now.

### MID-MARKET IN 2026: HEAD START AND PITFALL



Adoption of AI agents in mid-sized organisations runs ahead of the governance setup. The blue curve rises faster than the purple one, and the zone between the two is where most pilots strand in 2026. The working model in this document moves both curves up at once.

## **An invitation to explore.**

Anyone who has read this document to here probably recognises a number of patterns from their own organisation. The gain from AI feels present, and at the same time keeps lagging in the numbers. A department is trying, but the work barely changes. The question of where to start is bigger than the question of what to use.

Mono by Dusty works with organisations at this scale along the working model described above. An engagement begins with a short Discovery of one working week, followed by building blocks delivered at a pace that matches the team's absorption capacity. There are no time packages, and there is no report-first-implementation-later construction. For every building block, what done means is agreed in advance, and that is what gets steered toward.

Anyone considering a first exploration does not need to know in advance exactly which role or which building block returns the most. The working model begins with that very question, and delivers a concrete answer within a working week. For most organisations that is a comfortable way to take the first step, without immediately needing a multi-year engagement for it.

## About the author

Dusty Baars works under the name Mono by Dusty from Amsterdam. With a background of twelve years in enterprise UX at, among others, Shell, Roche, Philips and ING, he has spent the past two years focused on building AI products and setting up AI workflows in mid-sized organisations. Memortium for the funeral sector, Open Brain as persistent memory for AI systems, and Mono Dash for agent orchestration are now in production.

The role Dusty fills in projects is that of an architect who builds alongside, not that of an external advisor. An engagement begins with a limited scope that is made to work, and expands where it delivers value in practice. The documentation, the build and the ownership stay with the organisation itself.

---

Mono by Dusty · Amsterdam · [contact@dustybaars.com](mailto:contact@dustybaars.com)

For a first exploration or to get acquainted, a short email is the quickest starting point.