MEMORANDUM · MAIN DOCUMENT 3 OF THE SITUATIONAL-PICTURE SERIES

AI and the Situational Picture – From Report to Situation Service

How Artificial Intelligence Changes the Creation and Provision of Situational Information

Version 2.1 · As of 2026-08-18 · Audience: decision-makers in an enterprise organization – no AI, military or agile background required

About this document: Curated, conceived and editorially owned by **Robert Seebauer**; researched and produced with the support of **Claude** (Anthropic – model **Claude Fable 5**, via Claude Code). This is **version 2.1**: the deepening of v2.0 after the curator's reading and appraisal of *Prediction Machines*, *Noise* and *Co-Intelligence* (**Part A2 as the "judgment block"**, chapter **C.6**, **decision hygiene** as the seventh AI principle, an extended competency field 7, a noise audit in the pilot, roadmap consequences for Phase D) – supplemented by the **legal frame of the AI Act** as of August 2026 (legal review box in D.2: deployer obligations under Art. 4 and Art. 50, the high-risk trap of person-level data, legal guardrails for Phase D) and by the transparency notice per the series standard. Sister documents: "Team of Teams & Staff Training" (people & rituals) and "Enterprise Architecture & Solution Situational Picture" (content & structure) – together with this memorandum (tooling & provision) the triad of the series. Earlier versions: release archive `denkschriften/<date>` .

Transparency notice: This memorandum was produced with the support of an AI system (Claude, Anthropic – model Claude Fable 5, via Claude Code): drafting, research preparation, typesetting and chart code. All content was substantively reviewed, edited and approved by Robert Seebauer, who holds editorial responsibility.

Executive Summary

Starting question: The series has shown *how* a situational picture is run (rituals, staff training) and *what* belongs in it (capabilities, debt, confidence). What remains open is the question of tooling: **To what extent can Artificial Intelligence help with creating a situational picture and – above all – with the provision of situational information?**

THE SINGLE MOST IMPORTANT SENTENCE IN THIS DOCUMENT

AI makes situational information cheap while judgment remains scarce – therefore AI belongs on the staff, not on the commander's hill, and its greatest lever is **provision**: the right information, at the right time, in the language of the recipient.

Findings in five sentences:

1. **The economics reverse.** AI drastically lowers the cost of information processing and prediction; in return, the value of the complementary good rises – human judgment (Agrawal/Gans/Goldfarb). The bottleneck of the situational picture migrates from *creation* to *interpretation and delivery*.
2. **The military supplies the mature model.** For decades, the intelligence cycle has treated provision ("dissemination") as *a phase of its own with a doctrine of its own* – and the AI practice of NATO and armed forces shows a clear line: AI condenses situation and analysis, decision and responsibility stay with the human.

3. **The greatest civilian lever is the situation service.** For the first time, AI makes a *standing, queryable* situation service for the portfolio affordable – the shift from the periodic report to a service you can query at any time and that reports in on its own at decision points. The Security Operations Center has been proving for years that companies can run such a thing.
4. **No provenance, no truth – and no hygiene, no judgment** (deepened in v2.0). An AI-generated situational picture without a statement of sources and confidence is the perfect "Information Masquerade" (Milchman). And judgments suffer not only from bias but from **scatter (noise)** – AI can dampen both, but amplifies noise when it is trusted blindly (Kahneman/Sibony/Sunstein). The empirical record confirms: ~90% AI adoption, ~30% distrust, and AI *amplifies* what is there (DORA 2025). **Seven principles** draw the guardrails from this – newly among them, decision hygiene.
5. **Consequence for us:** Every decision can be decomposed into six components (Agrawal/Gans/Goldfarb) – chapter C.6 assigns, per component, where our software helps today, where AI can help and where the human remains. In addition, competency field 7 and curriculum module 7, a 90-day pilot with delta briefing and **noise audit**, and the software roadmap's **Phase D** behind the deterministic core.

What "AI" means here (taxonomy, applies to the whole document)

Level	What	In our context today
1 · Rule-based automation	Deterministic pipelines, thresholds, traffic lights	The entire report pipeline
2 · Statistics & classical ML	Monte Carlo forecasts, anomaly detection, fault prediction	Simulate module; *Fault Detection*
3 · Generative AI / LLMs (focus)	Summarizing, translating (languages **and** readings), Q&A, drafts	Memoranda production, manuals in 5 languages
4 · Agentic AI	Autonomous collection, multi-step assignments, tool use	Outlook only (D.2, guardrails)

Anyone deciding on "AI in the situational picture" should always be able to say which level they are talking about – opportunities, risks and guardrails differ considerably by level.

Part A – The sources: AI in our own corpus

The four AI titles of the 22-volume Apress corpus (full texts in the curator's working archive), read through the situational-picture lens – plus Milchman as the hinge to the EA memorandum.

A.1 *Architecting Enterprise AI Applications* – drawing the strategic boundary

The strategic-conceptual work of the corpus on the question of **where AI should decide and where it should not**: human-vs.-AI division of labor, meta-systems, overfitting and proxy traps. Core stance: deploy AI deliberately *within limits* where context and responsibility matter.

For us: the corpus' principal witness for the role model of this memorandum – AI as a staff member with a defined mission, not as a decision-maker.

A.2 *Mastering Machine Learning Architecture and Solutions* – the lifecycle

ML systems as architecture building blocks with a **lifecycle** of their own: data pipelines, model selection, operations, drift – MLOps as "DevOps for models".

For us: Whoever builds AI into the situational picture takes on operational responsibility (risk catalogue D.2: model drift); AI building blocks are not a one-time installation.

A.3 *AI and Microservices* – the applied view

AI-supported development and API/testing work in microservice landscapes.

For us: Evidence that level-3 AI already has tool status in everyday engineering – the question is not *whether* but *with which guardrails* (congruent with the DORA findings in A2.6).

A.4 *Fault Detection in Microservice Architectures* – analysis in pure form

Data-driven **fault prediction** from structural metrics (coupling/cohesion → fault probability, evaluated with precision/recall).

For us: the concrete pattern for turning existing metrics into *leading indicators* – level-2 AI for the analyze phase of the cycle (C.1), long before an LLM enters the picture.

A.5 The hinge: Milchman, *Unit Oriented Enterprise Architecture*

The only work in the corpus that literally speaks of **"situational awareness"** – subtitle: *Constructing Large Sociotechnical Systems in the Age of AI*. In the AI age, his diagnosis of the **"Information Masquerade"** (information flows worse than everyone claims) becomes a double warning: AI can *fight* the masquerade (provenance, confidence, availability) or *perfect* it (fluent text without substance). Which side wins is decided by the principles in C.4.

Part A2 – The judgment block: economics, scatter, the jagged frontier – and automation research

*Seven sources outside the corpus, each carrying one flank. The first three – in v1.x only sketched from general knowledge – the curator has since read in full and appraised; together with Duke (A2.7) they form the **judgment block** of this memorandum: why judgment is the scarce good (Prediction Machines), what ails it (Noise), how to bring AI in without*

*gambling judgment away (Co-Intelligence), and how to train it (Duke). Every appraisal ends with **the curator's position** – the point at which reading turns into a stance.*

A2.1 Agrawal/Gans/Goldfarb – *Prediction Machines* (2018, updated ed. 2022): the anatomy of a decision

The thesis of the Rotman economists: economically speaking, AI is a **price collapse for prediction** – understood in the broad sense as filling in missing information. When a good becomes cheap, the value of its **complements** rises: data and above all **judgment** (assessing what a prediction *means* and how much each outcome is worth). That carries the single most important sentence of this memorandum – but the book's real working tool is the **decomposition of every task into six components** (the "AI Canvas"):

Component	Question	Who/what delivers it
Prediction	What will we know once we know it? Which missing information gets filled in?	the machine's domain – this is where the price falls
Judgment	How do we value the possible outcomes? What are a hit and a miss each worth to us?	human; the complement that rises in value
Action	What do we do – which action follows from prediction plus judgment?	human or automated – depending on the dial (C.2)
Outcome	By what do we measure whether the action succeeded?	the feedback to all other components
Input data	Which data does the prediction need in ongoing operation?	precondition – this is where quality is decided
Training/feedback	From which data does the system learn, and how do outcomes flow back?	the learning loop – without it every prediction goes stale

The power of the decomposition is that it **makes the question "Where does AI help?" answerable**: not wholesale, but per component. And it carries an economic warning: whoever confuses prediction with judgment – that is, leaves it to the machine what outcomes are *worth* – has not automated but abdicated.

> THE CURATOR'S POSITION
> Full agreement. The decomposition is coherent and useful – it is the reason this memorandum gets a chapter of its own in v2.0 (C.6), which checks per component exactly where our software helps today, where AI can help and where the human remains.

A2.2 Kahneman/Sibony/Sunstein – *Noise* (2021): judgments scatter – and AI can do both

The book draws a distinction that decision-makers usually do not: **bias** is the *systematic* deviation of judgments (everyone is off in the same direction) – **noise** is their *unwanted scatter* (same situation, same facts, different judgments – depending on the person, the form of the day, the sequence, the weather). The authors show that noise in organizations is almost everywhere large, almost never measured and almost always underestimated – from insurance premiums to court sentences to personnel decisions. The two error sources are independent of each other; **whoever fights only bias leaves half the error on the table.** Their answer is **decision hygiene** – rules that reduce noise without having to see it in the individual case:

- **Independent judgments first, discussion afterwards** – whoever knows the opinion of the most senior person or of the group no longer judges independently.
- **Aggregation** – the average of several independent judgments is more reliable than the single one, especially where expertise is involved.
- **Structured assessment** – break complex judgments into partial judgments, assess each on its own, hold back the overall judgment until the end.
- **Relative instead of absolute scales** – ranking and comparison are more consistent than absolute grades.
- **Noise audits** – present the same case to several judges separately and measure the scatter; only then does an organization know how noisy it is.
- **Algorithms and rules as consistency anchors** – a simple, transparent model often beats the scatter of human judgments because it does not tire and has no form of the day.

The last point makes *Noise* ambivalent for this memorandum: AI can **dampen** noise (always the same rule, the same attention) – but a *large language model* is not a simple model; it is itself **non-deterministic** and has **error sources of its own** (hallucination, false sources, sensitivity to how the question is phrased). Whoever trusts it blindly imports its scatter into their own judgment – and, because of its air of objectivity, no longer even notices.

> THE CURATOR'S POSITION
>
> Full agreement – bias and noise are different, and one must always look at **both**. For AI in the situational picture the danger prevails: I see AI rather as a **noise amplifier**, because it is trusted blindly, without critical questioning – while it hallucinates or builds on false sources. The decision-hygiene rules matter and are adopted: as the seventh principle (C.4), in competency field 7 (C.5), as a noise audit in the pilot (D.3).

A2.3 Mollick – *Co-Intelligence* (2024): the jagged frontier and the human in the loop

The Wharton professor distills everyday work with language models into four rules: **always invite AI to the table** (try what it can do in your own work), **be the human in the loop**, **treat AI like a person – but tell it what kind of person it is** (role, context, assignment), and **assume this is the worst AI you will ever use** (the models evolve faster than organizations learn). His most important finding is the **"jagged frontier"**: the capability boundary of the models runs jagged – surprisingly strong at tasks that look hard, surprisingly weak at tasks that look easy – and this boundary is **invisible from the outside**. One finds it only by use, and one finds it anew at every point.

> THE CURATOR'S POSITION
>
> The concept is coherent and matches my own observations – precisely *because* AI is not deterministic. What follows is not hesitation but a double consequence: **involve AI heavily in order to recognize its limits at all** – but **always with the human in the loop**, because today we do not yet really see the limits and the bright spots, and the models evolve too fast to gamble. Trying it out is a duty; trust without verification is recklessness.

A2.4 Bainbridge – "Ironies of Automation" (1983): the rebuttal

The classic of automation research (doi.org), in four pages: automation takes over the *easy* parts of a task and leaves the human the *hardest* ones – monitoring and exception handling – while at the same time depriving them of the daily practice they would need for exactly that. The better the automation, the **more important and at the same time worse prepared** the human.

For us: perhaps the most important single warning in this memorandum – principle 5 "The human keeps practicing" (C.4) and the AI-free katas in the curriculum (D.4) are the direct answer.

A2.5 Parasuraman/Sheridan/Wickens (2000): the dial

The standard model of human-automation research distinguishes **four functional classes** – information acquisition, information analysis, decision selection, action implementation – and, per class, **levels of automation** from "the human does everything" to "the machine decides autonomously". The point: the level of automation is **to be chosen separately per functional class**.

For us: the scientific foundation of the automation dial in C.2 – "staff member, not commander" becomes a tunable setting instead of a ban.

A2.6 DORA 2025 – *State of AI-assisted Software Development*: the empirical findings

The annual DORA study (Google Cloud, ~5,000 respondents worldwide) delivers three findings for decision-makers: (1) **~90% of software professionals use AI** at work – the adoption question is settled. (2) The **"amplifier effect"**: AI amplifies existing strengths and weaknesses; high-performing organizations get better, weak ones see their problems more clearly – AI *repairs* nothing. (3) Despite broad usage, **~30% report little or no trust** in AI-generated output; the greatest returns come not from the tools but from platform quality and clear workflows.

For us: (2) is the empirical confirmation of the ToT single sentence ("ritual before tool") transferred to AI; (3) makes provenance and confidence a question of acceptance, not of style.

A2.7 Duke – *Thinking in Bets* (2018): the hygiene of judgment (since v1.1)

The former poker professional addresses the act of evaluating under uncertainty itself – three contributions carry directly: **resulting** (separating decision quality from outcome quality) is the literacy without which probabilistic situational information does not survive politically – a 15% scenario that materializes does not refute a P85 forecast. **Truthseeking groups** (a truthseeking charter: sharing information, organized skepticism, accountability) supply the *social* antidote to automation bias – the group cross-check before an AI suggestion is adopted. **Ulysses contracts** (advance commitments that bind the later self) give the decision-point wake-up call (M4) its foundation.

For us: Duke trains exactly the judgment that, per *Prediction Machines*, is the scarce good – and her truthseeking groups are the social form of the decision hygiene from *Noise*: judge independently, then aggregate, then discuss.
(Full appraisal: ToT memorandum, Part A2.5.)

THE JUDGMENT BLOCK IN ONE SENTENCE

Judgment is the scarce good (Prediction Machines), it scatters more than we think (Noise), AI finds its own limits only with the human in the loop (Co-Intelligence) – and it can be trained through bets, truthseeking and the separation of decision and outcome (Duke).

Part B – Reference practice: military intelligence and its AI

The method of the series: first understand the mature outside discipline, then transfer. For situational information, the most mature system is military intelligence – it has been answering our guiding question institutionally for decades.

B.1 The intelligence cycle in plain language

US doctrine (Joint Publication 2-0, *Joint Intelligence*) organizes intelligence work as a cycle of **planning/direction** → **collection** → **processing** → **analysis** → **dissemination/integration**, with constant evaluation and feedback. Three properties make it valuable for us:

1. **It starts with the decision-maker, not with the data**
   At the beginning stands the leader's prioritized information need (**CCIR/PIR** – Commander's Critical / Priority Intelligence Requirements): *Which information would change my decision?* That is exactly Hubbard's EVI question in doctrinal form – what gets collected is what can change decisions, not what is measurable.

2. **Dissemination is a phase of its own with a doctrine of its own**
   The best analysis is worthless if it does not reach the decision-maker in time, comprehensibly and in the right format – which is why the doctrine explicitly regulates *push* (pre-planned delivery to defined recipients) and *pull* (access to holdings on demand), recipient fit and timeliness. **Exactly this phase is what our guiding question addresses with "above all the provision".**

3. **It is a cycle**
   Every decision creates new information needs; evaluation runs continuously alongside. A situational picture is never "finished".

B.2 What the doctrine teaches about provision

From dissemination doctrine and staff practice (cf. the Team-of-Teams memorandum, Part B), five provision lessons can be distilled that hold independently of AI – AI merely makes them affordable:

- **BLUF – Bottom Line Up Front:** core message first, derivation afterwards; the situation briefing is a practiced form, not a piece of prose.
- **Recipient-appropriate:** the same facts are prepared differently per level (commander ↔ specialist staff) – *without* the facts changing.
- **Push by plan, pull on demand:** critical items are actively delivered (at decision points), everything else is available on call.
- **Timeliness before perfection:** a timely 80% picture beats a perfect one from yesterday (McChrystal's O&I lived on this principle).
- **Provenance and reliability are part of the report:** intelligence has always rated its sources explicitly – the civilian counterpart is our confidence flags (EA principle 4).

B.3 AI in the alliance: what NATO and research report – and what stays with the human

- **NATO ACT, *Decision-Making in the Age of Big Data and AI*** (2021): Multimodal sensor data plus analytics are to enable a **consolidated, multidimensional situational picture in near real time**; AI relieves analysts of sifting and fusion work and accelerates decision cycles. Just as clearly, the report names the preconditions: data quality, interoperability, user trust – and human responsibility.
- **RAND (RRA4316-1)** analyzes AI along four fundamental military competitions – mass vs. quality, **hiding vs. finding**, **centralized vs. decentralized command**, cyber offense vs. defense. Two are central for us: AI is primarily a *finding technology* (it lifts signals out of noise – that is situational-picture work), and it can strengthen **both** command

models. That is a warning: whoever *can* see everything centrally with AI will be tempted to *decide* everything centrally again – McChrystal's coupling (situational picture **plus** decentralized authority) must be actively defended against this recentralization reflex.

- **Atlantic Council** (report on NATO's AI integration): AI as an enabler of consolidated multi-domain situational pictures in C4ISR – the same thrust from the alliance perspective.
- **NATO itself draws the line:** its *Principles of Responsible Use* (2021) demand, among other things, accountability (humans remain answerable), traceability and governability of AI deployments. Translated into our image: **AI serves on the staff, it does not lead** – it is humans who steer it and who own the decision. The European legislator lays down the same principle: the **AI Act** (Regulation (EU) 2024/1689) requires the *deployer* of an AI system to ensure AI literacy (Art. 4) and transparency (Art. 50) and leaves responsibility and judgment with the human – the consequences for the situation service are set out in the legal review box in D.2.

B.4 The civilian proof of existence: the Security Operations Center

Companies have long been running a standing situation service – where the fire is: the **SOC** maintains a cyber situational picture around the clock. Sensors deliver events, correlation systems (SIEM) condense them, **AI-supported triage** prioritizes, alerts arrive with context, and a human analyst decides on the response. Two lessons from twenty years of SOC practice transfer directly: **alert fatigue** is the death of any situation service (hence prioritization by decision relevance – EVI), and **automation ends before the response** – even in the SOC, the machine rarely triggers action on its own. For the portfolio, where millions are decided, nothing comparable exists – not for lack of need, but because a humanly staffed situation service would be unaffordable. **AI changes that math.**

Part C – Transfer: the AI-supported situation service for the enterprise

C.1 The cycle for the enterprise situational picture – and the four roles of AI

The intelligence cycle maps directly onto the situational-picture work of the solution/portfolio level (and onto our software pipeline):

DIRECT What must the decision-maker know? (PIR/CCIR, civilian: define information needs by EVI)

↓

COLLECT Jira, Git, CI/CD, monitoring, registers `[get_data]`

↓

PROCESS pool, canonicalize, flag quality `[transform]`

↓

ANALYZE trends, anomalies, forecasts, scenarios `[build_reports, simulate]`

↓

PROVIDE readings, delta, Q&A, alerts, situation briefing ← **focus** `[report + situation service]`

↓

DECIDE & FEED BACK sync/O&I, AAR, new questions → back to DIRECT `[ritual + human]`

Across the cycle, AI's contribution can be captured in **four roles** – all four are staff roles, not command roles:

Role	Cycle steps	Examples	Maturity
1 · Sensor & collector	Collect, process	Extraction, classification ("which capability does this epic affect?"), duplicates, justifying quality flags	productive today
2 · Staff aide (editor)	Process, provide	Summaries, translation between readings, delta briefings, situation-briefing drafts, multilingual output	productive today, with mandatory review
3 · Analyst & sparring partner	Analyze	Anomaly/trend detection, forecasts (MC/ML), scenario drafts for wargaming, premortem questions, attacking assumptions (red team)	partly productive, partly worth piloting
4 · Situation service	Provide	Queryable situational picture, decision-point wake-up calls, audience-appropriate delivery	worth piloting – the new territory

C.2 The automation dial: how much AI per step?

Following Parasuraman/Sheridan/Wickens (A2.5), the level of automation is set **per functional class** – for the situational picture we recommend four dial settings:

1. **Information acquisition and processing: dial high**
   This is where the cost reduction happens; human drudgery in data collection is not a mark of quality but waste.

2. **Analysis: dial medium to high**
   With a confidence statement on every result and spot-check verification as staff routine.

3 **Decision selection: dial low**

AI may draft, compare and recommend options – the choosing is done by humans (which is also the MDMP logic: the staff develops options, the commander decides).

4 **Execution: only with human release**

In the situational-picture context this means in practice: no automatic escalations or reprioritizations without an agreed rule and a named responsible person.

This turns "AI is a staff member, not a commander" from a metaphor into a tunable setting – differentiated enough for real architecture decisions, simple enough for the board slide.

C.3 The situation service: seven provision patterns (the centerpiece)

The emphasis of the guiding question lies on provision – here are the seven patterns with which AI turns the report into a situation service:

#	Pattern	What it delivers	Main risk	Guardrail
M1	**Audience-specific readings**	Executive BLUF and engineering detail from the **same** numbers; extends EA principle 3 to "one picture, n readings"	Personalization fragments the single truth	One data basis, a shared core of metrics, provenance per statement
M2	**Delta briefing**	"What has changed since the last sync – and what is accelerating?" (Hohpe's "first derivative" as a product)	Noise gets sold as change	The deterministic diff supplies the facts, AI only the wording
M3	**Queryable situational picture**	Q&A in everyday language over the report data; drill-down without a tooling hurdle – more people can *read* the situational picture	Hallucination: eloquent answers without a basis	**Citation requirement** – no answer without a source reference to the data path; "I don't know" is permitted
M4	**Decision-point wake-up call**	Leading indicator + threshold + agreed response; AI monitors and reports **with context** instead of bare alerts	Alert fatigue (SOC lesson)	Thresholds and responses are agreed in the ritual, not invented by the AI – a **Ulysses contract** (Duke): The board binds in advance its later self in the heat of the moment
M5	**Audience-appropriate delivery**	Role, channel, timing, **language**; including the **situational-picture subscription**: periodic role-specific digests as a plannable push form	Delivery illusion: "delivered" ≠ "understood"	Backbrief spot checks in the ritual
M6	**Situation-briefing draft**	AI drafts the BLUF briefing for the sync; the human reviews, shortens, presents	The ritual degenerates into a reading-out of AI text	The presenter signs responsible; draft ≠ briefing
M7	**Institutional memory**	AAR archive, decision/assumption log searchable and queryable – "Have we seen this before? What did we decide back then – and why?"	Outdated decisions as current truth	Every answer carries the date and validity status of its source

Round-the-clock availability vs. ritual. Does the 24/7 situation service devalue the portfolio sync? No – it upgrades it: if the information reached everyone beforehand, the sync turns from a reporting meeting into a **decision meeting** (McChrystal's O&I was, precisely, *daily*). The only danger is cancelling the ritual "because it's all in the chat anyway" – then the stage dies on which leadership demonstrates culture (ToT success factor 2: rituals before tools).

One picture, n readings. How much personalization can the single truth bear? Readings may differ in selection, depth and language – **never in the numbers and never in the risk picture**. The shared core (traffic lights, top risks, confidence) cannot be personalized; whoever personalizes it builds watermelon reports with AI acceleration.

The third tension – radical transparency vs. protection of employees – is so fundamental that it is answered not as a trade-off but as a principle: see P6.

C.4 The seven AI situational-picture principles

Designed to connect to the five situational-picture principles of the EA memorandum – extended by a sixth (the aggregation rule, since v1.0) and, in v2.0, by a seventh (decision hygiene, from the judgment block):

1. **AI is a staff member, not a commander.**
   It collects, drafts, translates and contradicts – it does not decide and is accountable for nothing. Responsibility cannot be delegated to something that can bear no accountability. (ToT B.1; NATO principles; Prediction Machines.)

2. **Provenance before eloquence.**
   Every AI statement in the situational picture carries source and confidence – otherwise it is the perfect Information Masquerade. (Milchman; EA principle 4; DORA trust gap; Art. 50(4) AI Act – human review and editorial responsibility are at the same time the condition under which AI-assisted texts require no labelling.)

3. **The LLM writes, it does not calculate.**
   Numbers originate in the deterministic core (pipeline, simulation); the AI phrases, translates and explains them. A freely invented number is a system defect, not a cosmetic one. (Our own architecture; Hubbard calibration.)

4. **To provide is to prioritize.**
   What gets delivered is what can change decisions (EVI) – otherwise AI merely automates the KPI graveyard and produces alert fatigue. (Hubbard; SOC lesson; Goodhart warning.)

5. **The human keeps practicing.**
   Whoever leaves every draft to the AI loses the competence to spot its errors – training, katas and AI-free exercises are the protection against deskilling. (Bainbridge; Marquet "competence"; Martin "practice".)

6. **The situational picture knows teams, not individuals.**
   Aggregation at team/ART level is a hard rule, not a stylistic device: no drill-down to individuals, no individual measurement – all the more so once AI makes such analysis cheap. At the same time the most effective compliance simplification (co-determination, data protection, the AI Act: whoever evaluates no individuals stays outside the high-risk area of Annex III, point 4). (Standing line of the series; legal review box D.2.)

7. **Judge independently first, then ask the AI – and always check for both errors.** (new in v2.0)
   Judgments suffer from bias **and** scatter; AI dampens both only if it does not become a substitute for one's own judgment. Hence decision hygiene applies: an independent assessment *before* the AI suggestion, aggregation of several judgments, structured partial assessment, relative instead of absolute scales – and noise audits, to know how noisy one's own organization is. An AI suggestion that replaces the scatter of the humans without stating its own has not reduced noise but hidden it. (Kahneman/Sibony/Sunstein; Duke truthseeking; the curator's position, A2.2.)

C.5 Competency field 7: AI-assisted staff work

The Team-of-Teams memorandum defines six competency fields for portfolio staffs. AI adds a seventh – **not** "prompt tricks" but staff craft (and at the same time exactly what Art. 4 of the AI Act has required of every deployer since February 2025: fostering the AI literacy of one's own people):

- **Being able to task:** formulating precise assignments for AI (context, sources, format, limits) – the civilian form of drafting an order.
- **Being able to verify:** running spot checks, checking source references, holding numbers against the deterministic core; knowing *when* full review is required (decision papers) and when a spot check suffices (routine texts).
- **Reading and writing confidence:** understanding calibrated uncertainty language and demanding it from AI (connecting to the calibration training from curriculum module 3); evaluating forecasts without "resulting" (Duke): A 15% scenario that materializes does not refute a P85 forecast.
- **Recognizing automation bias:** knowing one's own tendency to follow fluent suggestions under time pressure – and mastering counter-rituals (four-eyes principle, anonymous advance votes *before* the AI suggestion, the truthseeking cross-check in the group per Duke).
- **Mapping the jagged frontier:** learning through guided practice where AI is strong in one's own context and where it is unreliable (Mollick) – this knowledge goes stale with every model change and must be maintained. *The curator's position (v2.0):* therefore **involve AI heavily**, in order to see the limits at all – but never without a human in the loop; the models evolve too fast to gamble.
- **Practicing decision hygiene** (new in v2.0): writing down one's own assessment *before* asking the AI; for important judgments, aggregating several independent assessments; breaking evaluations down into partial judgments; knowing that bias and noise are two different errors – and that a fluent AI suggestion can mask both.
- **Mastering the decomposition** (new in v2.0): being able to break a pending decision down into prediction, judgment, action, outcome, input data and feedback (Prediction Machines) – and to say, per component, what software, AI and the human take on (C.6).

C.6 Where software, AI and the human belong, per component of a decision (new in v2.0)

The decomposition from A2.1 makes the core question of this memorandum precisely answerable. For a typical portfolio decision – *"Do we pull initiative X forward even though Y then slips?"* – the result, per component, is:

Component	What our software delivers today	Where AI can help (level 3, with guardrails)	What remains with the human
Input data	Extraction, transformation, canonicalization, confidence flags (pipeline; roadmap phases A–C)	Classification, duplicates, gap finding, justification of quality flags (role 1)	Deciding *which* data count – the information need (EVI, Direct phase)
Prediction	Monte Carlo forecasts with confidence intervals (Simulate), trend/anomaly statistics – **deterministic, reproducible**	Formulating scenario variants, *explaining* forecast results, questioning assumptions (role 3) – **never generating the number itself** (principle 3)	Checking calibration; knowing that a forecast is a distribution, not a promise (anti-resulting rule)
Judgment	Metric set (flow + outcome + quality + people) as the basis for assessment; ROAM/decision log as structure	Listing assessment criteria completely, juxtaposing trade-offs, supplying counter-arguments (red-team assistant D5)	**The judgment itself**: what a hit and a miss are worth – principles 1 and 7; judging independently before the AI speaks
Action	No automatism – deliberately	Formulating the decision paper and the delta briefing (role 2, D1/D2)	Deciding and taking responsibility; execution only with release (dial 4)

Outcome	Report diff between two snapshots, target attainment per metric (D2)	Explaining deviations, preparing AAR questions (roles 2/3)	AAR with the anti-resulting rule: decision quality ≠ outcome quality
Training/feedback	Historical reports, decision/assumption log, AAR archive as the data basis (M7)	Recognizing patterns across initiatives, answering "Have we seen this before?" (M7, with citation requirement)	Learning as an organization: tracking calibration, checking assumptions with an expiry date

Three lessons from the table: **(1)** The software is strongest where determinism counts – data, prediction, outcome measurement; exactly there AI may *explain* but not *replace*. **(2)** AI is most effective in the transitions – raw material for judgment, language for action, patterns for feedback – and everywhere with a provenance obligation. **(3)** Judgment is the only row in which the "human" column is in bold: it is the component whose value rises with every cheaper prediction – and the component in which noise does the most harm. Whoever takes the decomposition seriously therefore also knows what the roadmap will **not** build: a recommendation button that automates the judgment row.

Part D – Implementation

D.1 Fields of application by maturity (decision-maker matrix)

Field of application (role)	Maturity	Benefit	Main guardrail
Data extraction & classification (role 1)	**productive**	Collection effort drops drastically	Quality flags instead of blind trust
Summaries, readings, translation (role 2)	**productive with mandatory review**	Rosetta-Stone work becomes affordable	Principles 2 + 3; mandatory review before dispatch
Anomaly/trend/forecast (role 3, level 2)	**productive** (statistics)	Leading indicators instead of the rear-view mirror	Confidence intervals instead of point values
Scenario/premortem assistance (role 3, level 3)	**worth piloting**	Red-team capacity for every staff	Human moderates; AI supplies raw material
Queryable situational picture, wake-up calls, delivery (role 4)	**worth piloting**	From report to situation service	Citation requirement; EVI prioritization; P6
Agentic collection/action (level 4)	**outlook**	In prospect: standing research assignments	Only with execution approvals (C.2, dial 4)

D.2 Risks & governance

Risk	Mechanism	Countermeasure
Hallucination	The LLM invents plausible details where data is missing	Citation requirement (M3), generation only from supplied data, rewarding "I don't know", principle 3
Automation bias	AI suggestions are adopted uncritically under time pressure	Confidence statements, spot-check routine, four-eyes review for decision papers, truthseeking cross-check (Duke)
Noise amplification (new in v2.0)	Non-deterministic models scatter themselves (phrasing sensitivity, hallucination, false sources); blind trust imports this scatter as supposed objectivity – bias and noise become invisible at the same time	Principle 7: independent judgment *before* the AI, aggregation, structured partial assessment; **noise audit** (put the same question several times/to several judges, measure the scatter – the AI's too); citation requirement against false sources
Deskilling	Automation withdraws the very practice that exceptions require (Bainbridge)	AI-free katas (D.4), a "do the math yourself once" rotation, principle 5
False certainty	Perfectly worded text suggests certainty in the Clausewitzian fog	Confidence belongs to the number *and* to the wording; standardized uncertainty language
Goodhart amplification	Teams optimize for AI-readable signals once AI reads and writes reports	A metric set instead of a single metric; qualitative cross-checks in the ritual
Confidentiality & data leakage	The situational picture bundles sensitive company information; external AI services as an outflow channel	Explicit deployment decision (on-prem / EU cloud / zero retention); data classification **before** connecting AI

Prompt injection	Manipulated content in source data steers the answers of the queryable situational picture	Treat inputs as data, separation of privileges, answer provenance
Model drift & operations	AI building blocks have a lifecycle of their own (MLOps)	Name operational responsibility; monitoring of the AI components themselves
Regulatory misclassification (new in v2.1)	A situational picture that admits AI evaluations at the level of individuals slides into the high-risk area of the AI Act (Annex III, point 4: performance/conduct of employees) – with conformity assessment, registration and oversight; missing transparency notices violate Art. 50	Principle 6 as an **architecture rule in the deterministic core** (aggregation boundary, no drill-downs to individuals – not merely in the AI layer); AI notices from day one (Art. 50); legal review box below

WARNING BOX – THE MASQUERADE TRAP

An AI-generated situational picture without proof of provenance is the perfect Information Masquerade: fluent, convincing, unverifiable. It feels like an information gain and is a loss of control. Whoever adopts only a single rule from this document, let it be the **citation requirement**.

Legal review box (have this legally reviewed at introduction; information as of 2026-08-18): AI Act (Regulation (EU) 2024/1689) – staggered entry into application: prohibitions (including emotion recognition in the workplace) and the **AI-literacy obligation (Art. 4)** since 2 February 2025, obligations for general-purpose AI models since 2 August 2025, **transparency obligations (Art. 50) since 2 August 2026** (a grace period until 2 December 2026 applies only to the machine-readable marking by *providers* under paragraph 2 for systems already on the market); the high-risk obligations were postponed by the *Digital Omnibus* (2026) – Annex III to 2 December 2027, Annex I to 2 August 2028. Four points affect the AI-supported situation service directly:

(1) We are the deployers. Whoever uses an AI system professionally is a *deployer* (Art. 3(4)) – with the obligation to foster the AI literacy of one's own people (Art. 4; competency field 7 and module 7 are exactly that) and with the transparency obligations of Art. 50.

(2) Transparency. A queryable situational picture must make it recognizable that an AI is answering (Art. 50(1)). AI-generated texts that inform the public must be labelled – *unless* they have undergone substantive human review and a person holds editorial responsibility (Art. 50(4)): that is exactly what the citation requirement, the four-eyes principle and this series' transparency notice deliver. The Commission guidelines (July 2026) set the bar high: a cursory sign-off is not enough, and substantive AI editing *after* the review forfeits the exemption – sign-off is the last step.

(3) The high-risk trap. Systems relating to the performance/conduct of employees (Annex III, point 4) carry the full burden of obligations (risk management, conformity assessment, registration, human oversight). Principle 6 ("teams, not individuals") keeps the situational picture away from it – as an architecture rule in the core, not as a request to the AI – and is thus the most effective compliance simplification of all.

(4) Open source does not exempt. The exemption for open-source AI (Art. 2(12)) expressly does not apply to Art. 50 nor to high-risk systems – our software, too, builds the transparency rules in (memo items below).

On top of that, nationally: **§ 87 Abs. 1 Nr. 6 BetrVG** (German Works Constitution Act: co-determination for technical monitoring of performance/conduct) and the **GDPR** – whoever evaluates no personal data has half the review behind them here as well. Involve the works council early, not after the fact. *(Sources: Regulation (EU) 2024/1689; Commission guidelines on Art. 50 and the Code of Practice on transparency of AI-generated content, July 2026 – see bibliography.)*

D.3 The 90-day pilot

1. **Weeks 1–2 – Direct:** survey the decision-makers' information needs by EVI ("Which information would change which decision?"); make the data-classification and deployment decision for AI services; inform the works council (P6 as a commitment).
2. **Weeks 3–6 – first pattern productive:** introduce the **delta briefing (M2)** – deterministic report diff plus AI wording, sent to all participants before each sync. Lowest risk, immediately tangible, a clean demonstration of principle 3.
3. **Weeks 7–12 – new territory, controlled:** the **queryable situational picture (M3)** as a supervised pilot for a defined user group, with citation requirement and logging; in parallel, competency-field-7 training for the staff. *New in v2.0 – the **noise audit** (weeks 9–10):* A real prioritization question is (a) judged independently by five staff members, (b) put to the AI five times in slightly varied wording, (c) only then discussed in the group. What gets measured is the scatter in all three rounds. The result is the most honest number of the pilot: How noisy is our staff – and how noisy is our AI?
4. **Wrap-up – a retro on the situation service itself:** Do users trust the answers – and *why* (not)? Which questions could the situational picture not answer (= gaps in the data)? What did the noise audit show – and which hygiene rule becomes routine from now on? Decision on expansion (M4, M5, M7).

D.4 Curriculum module 7: AI-assisted staff work

Extension of the 12-month curriculum of the Team-of-Teams memorandum (there Part D.2) by a seventh module – alongside the job, exercise-heavy:

- **Content:** the four roles and the automation dial; tasking/verifying/confidence (C.5); risk catalogue and legal situation (D.2); the jagged frontier in one's own context; *new in v2.0:* the decomposition of a decision into six components (A2.1, C.6) and decision hygiene (A2.2, principle 7).
- **Exercises:** (1) produce the same situation briefing once with and once without AI, and compare; (2) audit a prepared AI briefing with built-in errors ("find the hallucination"); (3) moderate a premortem with AI raw material; (4) a monthly **AI-free kata** – a situation analysis in 30 minutes by hand, so the form sits when the tool fails (the Bainbridge answer); *new in v2.0:* (5) a **decomposition kata** – break a pending decision down into the six components in 20 minutes and assign software/AI/human per component; (6) a **mini noise audit** – the same assessment question first silently and individually, then with AI, then in the group; make the scatter visible and discuss it.
- **Artifact:** a one-page **AI usage policy for staff work** (which level for what, review duties, citation requirement, P6, *new:* the hygiene rule "own judgment first") – the governance document most organizations do not yet have.

D.5 The purchasable market (vendor-neutral, by category)

- **Enterprise LLM platforms** with corporate data protection (private instances, EU options, zero-retention contracts) – the basis for roles 2 and 4.
- **BI/analytics copilots** (natural-language querying of existing dashboards) – the fastest route to M3-like functionality, but check the citation requirement and the data path.
- **AIOps/anomaly tools** (operational data, level 2) – mature; deliver leading indicators for the technical health picture (connecting to the SLO/DORA dimension of the EA roadmap).
- **Forecasting tools** (Monte Carlo/probabilistic) – already present as our own Simulate module; external ones only needed if third-party data sources are added.
- **AI competence trainings** – buy them tailored to staff work (verification, bias, law), not as a prompting course; for method training (red teaming, calibration, superforecasting) see the program table of the Team-of-Teams memorandum D.1.

Connectivity to our software (memo items)

The roadmap of the `portfolio` module (`docs/portfolio_roadmap_solution-lagebild.md`) is extended by **Phase D – AI assistance**, strictly **behind** the deterministic core:

#	Feature	Pattern	Prerequisite
D1	LLM executive summary (numbers from `summary.py`, LLM does the wording)	M1/M6	–
D2	**Delta briefing** (deterministic report diff + narration) ⟵ *first pilot*	M2	two report snapshots
D3	Queryable situational picture (Q&A over report data with citation requirement)	M3	provenance model
D4	Anomaly notices (statistics in the core, AI explains)	M4	–
D5	Red-team assistant (premortem questions from the assumption/decision log)	Role 3	Roadmap B4
D6	Multilingual report output	M5	–
D7	Noise-audit support (n-fold varied query, scatter statement) (new in v2.0)	Principle 7	D3

Architecture principle: The AI layer is optional and replaceable (own model/own key, local or API); the core remains able to run offline and true to the stdlib. **No core feature may require AI – AI refines, it does not carry.**

Consequences of the judgment block for Phase D (new in v2.0; adopted into the roadmap): **(a)** Every D feature is assigned to a decision component from C.6 – D1/D2/D6 to *action* (wording), D3/D7 to *prediction/feedback* (explaining), D4 to *outcome*, D5 to *judgment* (raw material, never the judgment itself). **(b)** The queryable situational picture (D3) gets a **hygiene provision**: it can be configured to ask for the questioner's own assessment *before* answering and to place both side by side. **(c)** New: **D7 · noise-audit support** – put the same question n times in varied wording, log and report the scatter of the answers; the AI measures its own noise level. **(d)** Confirmed: no recommendation button for the judgment row.

Legal guardrails for Phase D (new in v2.1; from the legal review box D.2, adopted into the roadmap): **(a)** The queryable situational picture (D3) indicates that an AI is answering (Art. 50(1) AI Act); AI-worded report texts (D1/D2/D6) carry a visible AI notice until a human has reviewed and approved them – the sign-off is at the same time the exemption from the labelling obligation (Art. 50(4)). **(b)** The aggregation boundary (principle 6) is enforced in the deterministic core: no D feature receives personal data – so the situational picture stays outside Annex III, point 4. **(c)** The open-source licence does not exempt from Art. 50 (Art. 2(12)) – the notices are part of the product, not of the configuration. **(d)** Deployer evidence: model, version, system instructions and data classification per deployment are logged – the basis for being able to demonstrate compliance with the AI-literacy and transparency obligations.

> **Reflexive evidence:** This series practices what it describes. Every memorandum is produced with AI support and carries the transparency notice (versioning ground rules 5 and 6 of the series: AI support named, content reviewed and editorially owned, the curator's sign-off as the last step before publication – at the same time the conditions of the exemption in Art. 50(4) AI Act); the software's manuals appear AI-translated in five languages; the forecasts are computed by a deterministic Monte Carlo module whose results AI may explain but never generate. We do not write about AI staff work – we run it, with exactly the guardrails this document demands.

Connectivity to the series

- **ToT memorandum:** The O&I ritual remains the decision forum – the situation service supplies it; empowered execution is defended against the recentralization reflex (B.3/RAND); competency field 7 and module 7 dock onto the curriculum and the competence catalogue.
- **EA memorandum:** The seven AI principles continue the five situational-picture principles; the EA dimensions (capability, NFR, ROAM, dependencies) are the data basis the situation service makes queryable; with M2/M4 the control loop gets its pacemakers.
- **Architecture Thought Leaders companion memorandum:** Hohpe's "first derivative" becomes a product as the delta briefing; Fowler's pattern vocabulary underpins the shared language of the AI usage policy; Martin's culture of practice carries principle 5; Hohmann's rituals remain the human stage on which AI raw material is negotiated.
- **Situational-awareness reading list** (in the curator's working archive): takes up the new sources on AI & situation service and on judgment (Duke).

Bibliography

Own corpus (full texts in the curator's working archive): *Architecting Enterprise AI Applications* · *Mastering Machine Learning Architecture and Solutions* · *AI and Microservices* · *Fault Detection in Microservice Architectures* (all Apress) · Milchman: *Unit Oriented Enterprise Architecture – Constructing Large Sociotechnical Systems in the Age of AI*. Apress, 2024.

Newly archived (freely available, PDFs in the curator's working archive):

- NATO ACT: *NATO Decision-Making in the Age of Big Data and Artificial Intelligence* (2021) – act.nato.int
- RAND: *How Artificial Intelligence Could Reshape Four Essential Competitions in Future Warfare* (RRA4316-1) – rand.org
- Bainbridge, Lisanne: "Ironies of Automation". *Automatica* 19(6), 1983 – doi.org.

Online (cited):

- DORA: *State of AI-assisted Software Development* (2025) – dora.dev/dora-report-2025 (download via form)
- US Joint Chiefs of Staff: *Joint Publication 2-0, Joint Intelligence* – jcs.mil/Doctrine (public edition)
- Atlantic Council: *How NATO can integrate AI to prevail in future algorithmic warfare* – atlanticcouncil.org
- NATO: *Summary of the NATO Artificial Intelligence Strategy / Principles of Responsible Use* (2021) – nato.int
- Parasuraman/Sheridan/Wickens: "A Model for Types and Levels of Human Interaction with Automation". *IEEE Trans. SMC–A* 30(3), 2000 – doi.org.

Additionally evaluated (new in v1.1): Duke, Annie: *Thinking in Bets. Making Smarter Decisions When You Don't Have All the Facts.* Portfolio, 2018 — Part A2.7; full appraisal: ToT memorandum (A2.5).

The judgment block (read and appraised by the curator, v2.0): Agrawal, Ajay / Gans, Joshua / Goldfarb, Avi: *Prediction Machines – The Simple Economics of Artificial Intelligence.* Harvard Business Review Press, 2018; Updated and Expanded 2022. · Kahneman, Daniel / Sibony, Olivier / Sunstein, Cass R.: *Noise – A Flaw in Human Judgment.* Little, Brown Spark, 2021 (German ed. *Noise – Was unsere Entscheidungen verzerrt*, Siedler). · Mollick, Ethan: *Co-Intelligence – Living and Working with AI.* Portfolio, 2024 (German ed. *Co-Intelligenz*, Redline 2025).

Legal frame (as of 2026-08-18; information without guarantee, have it legally reviewed at introduction): Regulation (EU) 2024/1689 (AI Act) · eur-lex.europa.eu/eli/reg/2024/1689/oj · European Commission: guidelines on the transparency obligations under Art. 50 (July 2026) and *Code of Practice on Transparency of AI-Generated Content* · digital-strategy.ec.europa.eu · Digital Omnibus amendment of the AI Act (2026; postponed high-risk deadlines, transitional period for Art. 50(2)) · nationally: BetrVG (German Works Constitution Act), GDPR.

Internal references: “Team of Teams & Staff Training” · “Enterprise Architecture & Solution Situational Picture” · “Architecture Thought Leaders” · decision record in the curator's working archive (decision state 2026-08-13) · situational-awareness reading list (in the curator's working archive) · `docs/portfolio_roadmap_solution-lagebild.md` (Phase D).

Version history

Version	Date	Changes
1.0	2026-08-13	Initial edition per the decision state of the discussion paper of 2026-08-13: dual guiding image (staff member/situation service), four-level taxonomy, Part B military intelligence + NATO/RAND + SOC, seven provision patterns, six AI principles (incl. P6 “teams, not individuals”), automation dial, competency field 7, 90-day pilot, curriculum module 7, software Phase D (pilot: delta briefing). Planned v2.0: deepening after evaluation of *Prediction Machines*, *Co-Intelligence*, *Noise*.
1.1	2026-08-13	Minor addition (no reorientation): Annie Duke, *Thinking in Bets*, worked in as A2.6 – resulting literacy for probabilistic situational information (C.5), the truthseeking cross-check against automation bias (C.5, D.2), the decision-point wake-up call (M4) classified as a Ulysses contract (C.3); sister-document reference updated to ToT v4.0. The deepened v2.0 remains planned.
1.2	2026-08-14	Reference maintenance for the online publication: cross-references to the series link the online overview, freely available sources carry real URLs (DOI/official pages), working-archive paths removed; cross-references without version numbers. Content unchanged.
2.0	2026-08-15	**Deepening after the curator's reading** of *Prediction Machines*, *Noise* and *Co-Intelligence*: Part A2 recast as the **judgment block** (A2.1 decomposition into six components/AI Canvas; A2.2 bias ≠ noise + decision hygiene; A2.3 jagged frontier + human in the loop – each with the curator's position; Duke integrated as A2.7). New: **principle 7 decision hygiene** (C.4), chapter **C.6 assignment of software/AI/human per decision component**, competency field 7 extended by hygiene and decomposition (C.5), risk “noise amplification” (D.2), **noise audit** in the 90-day pilot (D.3), decomposition kata and mini noise audit in module 7 (D.4), Phase D extended by D7 noise-audit support and a hygiene provision for D3 (roadmap PR). Executive summary updated accordingly.
2.1	2026-08-18	Minor addition (no reorientation): **legal frame of the AI Act** brought up to date as of August 2026 – legal review box in D.2 recast (entry into application after the Digital Omnibus, deployer obligations Art. 4/Art. 50 including the exemption for editorial responsibility, high-risk trap Annex III, point 4, open-source exemption does not apply to Art. 50), risk “regulatory misclassification” (D.2), references in B.3, principles 2/6 (C.4) and competency field 7 (C.5), **legal guardrails for Phase D** (memo items; adopted into the roadmap), bibliography section “legal frame”. Also documents the clarification of the B.3 sentence on NATO drawing the line (“AI serves on the staff, it does not lead”), made on 2026-08-18 still within v2.0. Transparency notice per the series standard (versioning ground rules 5/6).