

# EVA - European Verified Agents

*Trusted Industrial AI Teammates*

## Industry Background Document

Companion to the thirteen-slide presentation, with additional use cases and illustrative value calculations.

18 October 2026

We invite industrial companies to help develop a shared AI teammate for practical tasks in their operations. The aim is to move from individual AI assistance to context-aware Verified Agents that support entire teams. This document explains the problem EVA addresses, what a pilot would investigate and how potential value can be assessed. Participation starts with an exploratory conversation, with no pilot or budget commitment.

### SLIDE 01

## Expertise that stays with the team

EVA aims to help teams retain and apply practical expertise in shared tasks. Good decisions in production, maintenance and service often depend on several people combining their knowledge. The proposed agent would maintain a shared view of the task, responsibilities and current status, ask for missing explanations and provide relevant guidance for each role. This is the team context that separate personal AI assistants may not share.

For an industrial partner, the intended outcome is an evidence-based decision: does this approach improve a selected task enough to justify implementation and operating costs? We are preparing a Horizon Europe research proposal to investigate this. Funding has not yet been awarded.

### SLIDE 02

## New shift. Same problem. Context lost.

The packaging example makes the problem concrete. Rejects increase after a changeover. One shift adjusts a setting and carries out checks. At handover, however, the reasons for some steps, the significance of the results or the outstanding questions are missing. The next shift reconstructs the case and may repeat work already completed.

In this situation, EVA would aim to preserve the links between observations, checks and reasoning for the whole team. The purpose is to prepare the next shared decision, rather than simply create more documentation. The packaging line is an illustrative example, not a requirement for the eventual pilot.

SLIDE 03

# Technical readiness is not production clearance

“Ready” needs a clear reference point. In this simplified example, the technical setting has been corrected, but an agreed quality check is still outstanding. Maintenance is reporting that its own work is complete, not that production is authorised to resume.

| Role        | Statement in the example    | Meaning for the shared case                                  |
|-------------|-----------------------------|--------------------------------------------------------------|
| Production  | “Can we resume production?” | Asks for clearance to operate.                               |
| Maintenance | “Technically ready.”        | The setting is corrected. Another role still needs to check. |
| Quality     | “Not yet cleared.”          | The quality check is still outstanding.                      |

The misunderstanding arises when “technically ready” is interpreted as “cleared for production”. EVA would link the different statements to the same case and identify the missing condition. An appropriate follow-up would be: “The setting has been corrected. Quality clearance is still pending. Which check is outstanding?”

EVA would neither grant clearance nor start equipment. The responsible people check, decide and act under their site procedures. The research question is whether a shared agent makes such dependencies more visible than separate personal AI assistants. It needs the relevant context and must not treat missing information as evidence that work is complete.

SLIDE 04

# Turn expertise into usable guidance

For this support to help beyond a single situation, the case must become reusable knowledge. A video may show what someone does. EVA would also need experts to explain why a check is necessary, when it applies and which exceptions matter. Authorised videos, test results, manuals and expert interviews are possible inputs.

We propose different levels of detail for the same task: a short visual guide for use on the job, a fuller explanation for learning and structured detail an agent can process. Experts would review the content before reuse. Every version must retain critical conditions, warnings and limits. Shortening the text must not remove an essential prerequisite.

## SLIDE 05

## Answers, process support and proactive suggestions

EVA's level of initiative should fit the task and the team. On request, it would answer a specific question. In process support, it would take an agreed role in the dialogue, for example during handover. Proactively, it could flag a relevant gap: "We have conflicting status updates. Shall we clarify them?"

More initiative is not automatically better: suggestions may help or distract. Employees should help shape the role, timing and interaction. Voice, text, images or video may suit different work environments; no avatar is required. In the planned pilot, EVA remains advisory. It does not execute process steps, assign work autonomously or control equipment. Any later extension would require a separate scope, authorisation and assessment.

## SLIDE 06

## Measure team performance and operational value

The pilot would assess team performance and operational value against agreed baselines, comparable tasks and quality criteria. Measures could include avoidable repeat checks, time to an agreed next step, the correctness of joint decisions and later application of learned rules. Workload and the effort required to capture, review and maintain knowledge and operate EVA also count.

The following five cases are starting points for discussion. They show where a shared agent might help and which sources of value should be measured. All volumes, time savings and hourly rates are hypothetical assumptions, not benchmarks, quotations or EVA forecasts. The values must not be added together until overlaps have been ruled out.

### A Shift handover after a changeover

Two employees reconstruct the task status during handover. EVA would bring together completed checks, their rationale and outstanding issues.

**Value drivers:** Less time searching and coordinating; fewer unnecessary repeat activities.

**Model assumptions:** 3 handovers per production day, 220 days per year, 2 participants, each saving 10 minutes per handover.

**$3 \times 220 \times 2 \times 10 \div 60 = 220$  person-hours per year**

At an assumed internal hourly rate of EUR 55, this represents capacity valued at EUR 12,100. It says nothing about equipment downtime and does not automatically reduce cash costs.

**Assess in the pilot:** Actual time per participant and handover, information completeness and unchanged decision quality.

## Use cases and value drivers

Continued from slide 6. All figures are hypothetical assumptions.

### B Joint troubleshooting of recurring faults

Maintenance and production assess the same fault from different perspectives. EVA would keep previous checks and suspected causes together and ask targeted questions about missing information.

**Value drivers:** Less duplicated diagnostic work and less time coordinating between roles.

**Model assumptions:** 80 suitable cases per month, 12 months, 2 participants, each saving 12 minutes of diagnostic and coordination effort per case.

$$80 \times 12 \times 2 \times 12 \div 60 = 384 \text{ person-hours per year}$$

At EUR 60 per hour, this represents capacity valued at EUR 23,040. Any reduction in the duration of the fault is measured separately. Staff time saved must not simply be treated as equipment downtime avoided.

**Assess in the pilot:** Agreement on the next diagnostic step, repeat checks, total staff time including EVA use and follow-up review, and technical correctness.

### C Connect field service and technical support

A field service technician hands a case to technical support. EVA would organise customer context, equipment condition and completed tests so the second team can continue with a clear focus.

**Value drivers:** Fewer follow-up questions and less reconstruction of the case. Avoided extra visits may provide a separate benefit that reduces cash expenditure.

**Model assumptions:** 40 handovers per month, 12 months, 2 participants, each saving 20 minutes of reconstruction effort.

$$40 \times 12 \times 2 \times 20 \div 60 = 320 \text{ person-hours per year}$$

At EUR 70 per hour, this represents capacity valued at EUR 22,400. Separate additional scenario: 6 genuinely avoided visits, each with EUR 900 in avoidable travel or external costs, yield EUR 5,400. Staff time already valued must not be counted again within those visit costs.

**Assess in the pilot:** Follow-up questions per case, total handling time, solution quality and genuinely avoidable additional visits compared with similar cases.

## Use cases and value drivers

Continued from slide 6. The calculations describe potential benefits, not proven effects.

### D Team-based onboarding and reskilling

New employees or people changing roles learn a task alongside experienced colleagues. EVA would explain reasons, limits and exceptions, and revisit the learning in later tasks.

**Value drivers:** Less learning and mentoring time with at least the same demonstrated competence.

**Model assumptions:** 10 people per year; each requires 16 fewer hours of learning and 6 fewer hours of support from experienced colleagues.

**10 × 16 = 160 learning hours; 10 × 6 = 60 mentoring hours**

160 × EUR 40 + 60 × EUR 65 = EUR 10,300 in valued capacity. The calculation concerns different people's time. Mandatory training or required assessments are not shortened simply because a task is completed faster.

**Assess in the pilot:** Time to the agreed evidence of competence, workload for experts and correct application of critical rules in later tasks.

### E Translate engineering changes into operational practice

Engineering, production and quality teams need to understand changes to a product variant or work instruction. EVA would connect current versions with their implications for each role and make unresolved questions visible.

**Value drivers:** Less searching, coordination and rework caused by unclear or outdated information.

**Model assumptions:** 30 changes per year, 3 roles with one person in each, each saving 1.5 hours of avoidable additional effort per change.

**30 × 3 × 1.5 = 135 person-hours per year**

At EUR 65 per hour, this represents capacity valued at EUR 8,775. Quality defects, scrap or liability risks are not assigned a blanket monetary value. Approval remains with the responsible people.

**Assess in the pilot:** Use of the correct version, follow-up questions and documented rework. The baseline includes only avoidable additional effort, not the technical assessment that is needed anyway.

## What the calculations mean for the business case

Continued from slide 6

Person-hours measure capacity. An internal hourly rate assigns a value to that capacity for a business decision. A cash benefit arises only when expenditure such as overtime, external services or avoidable travel actually falls. With unchanged staffing, the benefit may instead be more time available for other work.

Case B illustrates how sensitive the result is to the effect actually achieved. Keeping case volume, two participants and the EUR 60 hourly rate unchanged gives:

| Time saved per person per case | Capacity per year | Annual value of capacity |
|--------------------------------|-------------------|--------------------------|
| 6 minutes                      | 192 hours         | EUR 11,520               |
| 12 minutes                     | 384 hours         | EUR 23,040               |
| 18 minutes                     | 576 hours         | EUR 34,560               |

If only 50% of the time released in the middle scenario can be usefully redeployed, the valued benefit is EUR 11,520 per year. This is a reference value for the business case, not a proposed budget or a promise of cash savings.

Benefits must be weighed against one-off setup, integration, knowledge capture and adoption costs, plus recurring review, maintenance, IT and operating costs. The first year includes the agreed one-off costs. If review time is already included in the measured task time, it must not be deducted again. Benefits from different cases may be combined only where their time and effects do not overlap.

### SLIDE 07

## Check the evidence. People stay in control

For EVA, “Verified” means traceable checks for a specific task. The agent would consider the source, version, scope and access rights, and explain the evidence behind a recommendation. If it identifies a critical gap, it should make uncertainty visible and ask the responsible person. This is neither a guarantee that every error will be detected nor a general certification.

The proposed safeguards keep data under the industrial partner’s control, with no cross-company sharing by default. Existing intellectual property stays with its owner; rights to new results are agreed. Employees participate on an informed basis, data is limited to what is necessary and individual performance monitoring is not planned.

People remain accountable for decisions and actions. In the planned pilot, EVA controls neither workflows nor equipment. Access, hosting, information security, responsibilities and review procedures must fit the site and be agreed before testing. In regulated environments, traceable recommendations and human review must complement the required validation and approval processes. Time savings cannot compensate for a breach of quality or safety requirements.

## An OEE-inspired assessment of agent use as a research goal

Continued from slide 7

An OEE-inspired measure for AI agent use in teams.

Overall Equipment Effectiveness combines availability, performance and quality. [3] The analogy for EVA is an understandable assessment built from several dimensions, not reuse of the OEE formula. Employees and decision-makers should be able to see which task an agent was assessed for, how well it supports the team and where its limits lie.

Possible dimensions include the technical quality of results, operational reliability and demonstrated contribution to team performance. The proposed Agent Passport would record these results together with the task, agent and knowledge versions, test conditions, scope and date, limits of use and assessment responsibility. Any aggregate measure would remain traceable to this evidence, rather than becoming an opaque “trust score”.

**Important:** Effectiveness is not the same as trustworthiness. Permissions and safety requirements must be met separately. Strong time savings must not offset critical errors. EVA needs to research which measures and forms of aggregation are useful. Any score would apply only to the context assessed.

### SLIDE 08

## When does a shared AI teammate add more value?

The central question is when a shared AI teammate improves coordination, learning and decision quality compared with individual AI assistance. We plan to compare current practice, personal AI assistants and a shared agent using comparable tasks and evidence. Their underlying AI capabilities should be comparable. The outcome is open; individual assistance remains a serious alternative.

For operations, the research informs practical design choices. Shared understanding concerns how people develop an accurate view of the task, roles and current status. Role and initiative concern when advice, learning support or proactive suggestions help and when they interrupt. Evidence and trust concern whether people accept sound recommendations and recognise incorrect ones, rather than simply relying on AI more often.

We also ask whether short guides preserve the same critical conditions as detailed and agent-readable versions, and whether people apply that knowledge correctly later. The findings would inform the level of detail, interface and human review appropriate to a task. Employees would help shape the work; workload, accessibility and potential bias, including gender-related bias, form part of the assessment.

SLIDE 09

## Existing building blocks. Focused research for EVA

EVA would build on existing foundations, but the shared industrial agent still needs to be developed and tested in operations. Potential components include ZHAW PROMISE for stateful dialogue management and WIT AINNOV8 for voice, chat and image AI in telecommunications. Their suitability, reuse and integration into EVA remain to be agreed.

| Partner                    | Proposed contribution                             |
|----------------------------|---------------------------------------------------|
| WIT Software               | Technical implementation and agent platform       |
| Schopf Meta Consult        | Capture practical expertise and support adoption  |
| ZHAW                       | Human–AI teamwork and interaction design          |
| UPM                        | Evaluate AI advice and its use in teams           |
| FAU                        | Industrial integration                            |
| Open Industry 4.0 Alliance | Industry network and access to suitable use cases |

Industrial partners would help define the relevant task, the evidence they can provide and the measures of success. The allocation of work and resources remains subject to joint planning. Existing components alone do not demonstrate operational value for EVA.

SLIDE 10

## Industrial standards as building blocks for EVA

EVA would combine existing industrial standards and technical components to suit the operational task, rather than develop a new industrial standard. A pilot could start with a shared AI assistant. The use case and the industrial partner’s ambitions would determine which data and systems to add.

### Start with shared knowledge

Approved instructions, videos and practical expertise can provide the foundation. EVA would answer questions while keeping information connected for several participants. A pilot could therefore begin with a limited knowledge base before more extensive equipment integration becomes worthwhile.

### Structured asset context through the AAS

The Asset Administration Shell (AAS) describes an asset or component in a standardised, machine-readable structure. Suitable submodels can provide identifiers, technical data and document references, helping EVA associate information with the correct asset and variant. Documents can remain in existing systems. The AAS is one possible foundation, not a prerequisite for every pilot. [4]

## Current operating data through existing interfaces

Where a task needs current equipment status, we would assess additional connections. Options include OPC UA or MQTT and existing interfaces to production, quality or service systems. OPC UA supports industrial communication and information models; MQTT supports messaging. These components can enrich asset context and enable situation-specific guidance. [5, 6]

## Hosting and AI models to meet your requirements

We would assess on-site operation, for example on edge systems or company servers, hosting in Europe, and hybrid options where appropriate. Model selection follows the task; European and locally deployable models are possible options. Model origin and hosting location are separate decisions. Data flows, access rights and operating responsibility would be agreed with the industrial partner.

## Let operational value guide integration

Relevant, maintained data can expand the useful context. More data alone does not guarantee better recommendations. We would weigh the expected benefit of each connection against integration, review and maintenance effort. Authorised access connects the selected sources with EVA. Guidance remains advisory in the planned pilot; people decide and act.

### SLIDE 11

## 45 minutes to explore the fit

The first step is a 45-minute conversation with someone responsible for operations, maintenance or service. We discuss a recurring problem, a useful success measure and whether a joint pilot outline is worth pursuing. No preparation, files, NDA or budget commitment are required. Confidential details should be kept out of this initial conversation.

Only if both sides are interested would we define a bounded task. Process owners, experts and, where needed, IT/OT colleagues would then clarify data, interfaces and effort. Before any pilot commitment, we would agree person-days by role, user hours, costs and any company contributions, data rights and stop criteria. A limited pilot would aim to deliver a prototype, assessed evidence and a basis for a decision, not a promised production-ready solution.

The integration scope would follow the options in slide 10. A focused start with existing knowledge is possible, as is the targeted addition of asset context and operating data.

### SLIDE 12

## Help shape the pilot early

Early industrial partners can shape the problem, assessment criteria and research. EVA targets HORIZON-CL4-2027-02-DIGITAL-EMERGING-52-two-stage, “New approaches for Human/AI collaboration for the workforce of the future”. It connects co-learning and knowledge capture with human–AI teamwork. The call requires new research approaches; implementing existing software alone is insufficient. [1, 2]

## Call procedure and funding

Continued from slide 12

The deadlines are 2 February 2027 for Stage 1 and 2 September 2027 for Stage 2, both at 17:00 Brussels time. Stage 1 uses a 10-page proposal and blind evaluation, with partial randomisation for a subset of qualifying proposals. Meeting quality thresholds does not guarantee an invitation to Stage 2. Funding has not been awarded; a funded pilot depends on selection, the grant agreement and partner agreements. [2, 7]

The call is a Research and Innovation Action with lump-sum funding. The RIA rate is up to 100% of eligible estimated project costs used to establish the grant. Each partner's activities, resources and funding allocation would be agreed in the proposal. Payments depend on proper implementation of agreed work packages. This is not a blanket commitment to cover all company expenditure. [2, 7, 8]

Each site would gain a basis for its own decision. European industry would benefit from design and assessment methods that transfer to other tasks and sites. Approved or abstracted findings can support this; confidential operational data is not automatically shared across sites.

A compelling business need could also lead to a separately commissioned implementation independently of EU funding. Scope, effort, acceptance and operation would be agreed separately, with research and implementation kept distinct in scope and financing. The first conversation commits a company to neither the application nor a commercial contract.

### SLIDE 13

## Find the right starting point for your team

Start with a recurring task where several roles depend on one another and spend time searching, asking follow-up questions or reconstructing knowledge. Experts should be able to judge a correct outcome. The five examples offer starting points; your operations determine the choice.

For the first conversation, a short description is enough: where does the effort arise, who is involved and how would you recognise an improvement? Together, we can assess whether EVA fits and whether further scoping is worthwhile.

**Contact:** Dr. Peter Schopf · Schopf Meta Consult GmbH  
Peter@schopf-meta-consult.de

## Sources on funding, standards and OEE

[1] European Commission / CORDIS: Human/AI collaboration topic

[2] Horizon Europe Work Programme 2026–2027, Part 7, DIGITAL-EMERGING-52-two-stage

[3] Vorne / OEE.com: Calculating OEE and its three factors

[4] IDTA: Asset Administration Shell specifications

[5] OPC Foundation: OPC UA Overview and Concepts

[6] OASIS: MQTT Version 5.0

[7] European Commission: Horizon Europe 2026–2027 General Annexes

[8] European Commission: How to manage your lump sum grants

Call information checked against official EU sources on 18 October 2026. The value calculations are explicitly hypothetical models.