european association of innovation consultants

EAIC Position on the Responsible and Governed Use of AI in Proposal Evaluation

May 2026

www.eaic.eu

Context

In February 2026, following a formal complaint, the European ombudswoman opened an inquiry into the rules governing evaluators' use of AI systems in the context of EU programmes implemented by EISMEA as well as into how transparent the agency is about this use of AI. This inquiry follows the formal complaint from a company based in Poland. The complainant, a company that submitted a proposal under the Horizon Europe EIC Accelerator programme, raised concerns about how the European Innovation Council and SMEs Executive Agency (EISMEA) had evaluated its proposal. Among other things, they were concerned that the external experts evaluating its proposal for EISMEA had used third-party artificial intelligence (AI) systems in a manner that had made the evaluation unfair. The complainant also argued that clearer rules should be established governing the use of AI systems by project evaluators.

This complaint and associated inquiry raise legitimate interest since the European Commission published in April 2025 the "Living guidelines" on the responsible use of generative AI in research, warning evaluators and reviewers to refrain from using AI in activities like peer review or proposal evaluation to avoid unfair assessments and protect unpublished ideas.

Also, in March 2026, the ERC Scientific Council adopted new guidelines on the use of artificial intelligence in the evaluation of grant proposals. The guidelines aim to safeguard the integrity of peer review while allowing limited use of AI where it does not compromise responsibility or trust. It relies on two key principles: (1) non-delegation of the evaluation task, and (2) privacy and confidentiality. As such, reviewers may not:

- Summarise the proposal to avoid reading it in full;
- Provide any form of assessment, explicit or implicit, of the merit of the proposal;
- Upload a proposal, or part of it, to an online AI tool such as ChatGPT.

EAIC Positioning

The EAIC recognises both the complexity and the increasing volume of project proposal evaluations within EU funding programmes. Evaluation processes require the systematic analysis of highly technical, multidisciplinary and commercially sensitive information under strict time constraints. In this context, digital tools, including Artificial Intelligence systems, may offer efficiency gains and structured analytical support.

However, the use of AI in evaluation processes raises legitimate concerns regarding confidentiality, fairness, transparency, and accountability. EAIC therefore considers that the question is not whether AI will be used, but under which governance framework it may be used responsibly.

EAIC supports a model of AI-assisted evaluation under strict governance and full human authority, based on the following principles:

1. Principle of Complementarity

EU funding programmes such as EIC and other centrally managed instruments regularly assess thousands of highly technical, multidisciplinary and commercially sensitive project proposals per

year. Evaluations are conducted under strict timelines and require the systematic analysis of complex scientific, technological, financial and market information.

Ensuring efficiency, consistency and quality in such a demanding context requires modernisation of evaluation processes. The responsible use of digital tools, including Artificial Intelligence systems, can enhance analytical capacity, reduce administrative burden, and allow expert evaluators to focus on substantive judgement rather than procedural processing.

For this reason, EAIC considers that AI tools may legitimately serve as supportive instruments to assist evaluators in managing information, structuring analysis, or identifying formal inconsistencies. AI may contribute to tasks such as:

- structured summarisation of proposals,
- cross-referencing against published evaluation criteria,
- identification of missing elements or inconsistencies,
- Or comparison of factual data points across sections, among others.

However, AI systems must remain strictly auxiliary in nature. Under no circumstances should AI systems:

- autonomously determine scores,
- replace expert judgement,
- substitute panel deliberation,
- generate binding evaluative conclusions,
- or exercise any form of decision-making authority.

Evaluation under EU programmes is inherently a qualitative, expert-driven process that requires contextual interpretation, sectoral understanding, and balanced judgement. These functions cannot be delegated to automated systems. **The evaluative judgement, scoring, and final recommendation must remain exclusively human responsibilities.**

Take the case of CHIPS JU proposals. Applicants are required to demonstrate how their work maps onto the CHIPS JU Strategic Research and Innovation Agenda (SRIA) – a document running to several hundred pages across multiple research pillars. Evaluators are expected to verify this alignment for every proposal they assess, under tight deadlines and across a large panel. In practice this is one of the most time-consuming parts of the evaluation, and one where the risk of inconsistency between evaluators is highest. The same challenge applies to Horizon Europe proposals, where alignment with the relevant Work Programme topic must be established and evidenced.

This is exactly the kind of task where AI assistance makes sense and stays clearly within bounds. An AI tool – running in a closed, sovereign environment – could map each section of a proposal to the corresponding SRIA pillar or Work Programme topic, flag where that alignment is missing or unconvincing, and produce a simple matrix the evaluator can use as a reading guide. It could also check for internal inconsistencies: budget figures in the narrative that do not match the financial tables, TRL claims that sit oddly against the methodology, market data that looks implausible against publicly available sources. None of this involves scoring or forming a view on the merit of the proposal. It is structured fact-checking against published documents. The evaluator decides what to do with the findings.

2. Principle of Confidentiality, Trade Secrets Protection and Digital Sovereignty

The European Union has long established itself as a global standard-setter in the protection of personal data, commercially sensitive information, and fundamental rights in the digital environment. Through instruments such as the GDPR, the Trade Secrets Directive, and the emerging AI Act, **the EU has built a regulatory architecture based on trust, legal certainty, and technological sovereignty**.

EU funding programmes operate within this framework. Project proposals submitted to EIC, and other centrally managed instruments routinely contain unpublished scientific results, proprietary technologies, patentable knowledge, strategic roadmaps, financial structures, and protected trade secrets. In many cases, these documents reflect years of R&D investment and constitute the core competitive assets of European innovators. **The integrity and credibility of the EU innovation ecosystem depend on the absolute protection of this information.**

As Artificial Intelligence tools become more capable and widely accessible, their integration into evaluation processes must be governed by the same high standards that define Europe's approach to digital governance. The pursuit of efficiency must not weaken confidentiality, intellectual property protection, or Europe's digital sovereignty. In this context:

- Project proposals, in whole or in part, must never be uploaded to open, public, or externally hosted AI systems that fall outside controlled and sovereign environments.
- Any AI tools used in evaluation must operate within secure, closed infrastructures under EU jurisdiction or equivalent legally enforceable safeguards.
- Systems must guarantee strict non-retention, non-reuse, and non-training of proposal data.
- Explicit technical measures must prevent data extraction, replication, secondary processing, or indirect model training.
- Governance frameworks must ensure full compliance with GDPR, the Trade Secrets Directive, the AI Act, and applicable EU cybersecurity standards.
- Digital sovereignty is not an abstract objective; it is a prerequisite for maintaining trust in EU funding mechanisms. Confidentiality must therefore be technically enforced, institutionally supervised, and demonstrably auditable, not merely contractually declared.

EDF proposals present a specific challenge. The technologies involved frequently have dual-use dimensions – potential applications beyond their stated defence purpose – and evaluators are expected to assess whether applicants have addressed this adequately. The difficulty is that EDF calls cover an increasingly wide technology landscape: quantum communications, autonomous systems, directed-energy, advanced materials. No single evaluator can be a domain expert across all of it, which creates a real risk that dual-use implications go unnoticed. At the same time, these proposals contain some of the most sensitive material in the EU funding landscape – unpublished research, proprietary technologies, strategic roadmaps representing years of investment.

It is this combination – the evaluator expertise gap on one side, the extreme sensitivity of the content on the other – that makes the principle of confidentiality and digital sovereignty a practical necessity rather than an abstract requirement. An AI tool could help by cross-referencing the technology description against publicly available dual-use control lists (the EU

Dual-Use Regulation Annex I, the Wassenaar Arrangement categories) and producing a flag list for the evaluator to work through. It would not make a compliance determination. That judgement stays with the evaluator. But the tool could only ever be used this way inside an infrastructure that guarantees the proposal data is never retained, never reused, never used to train a model. Anything less is not a risk worth taking with this category of material. This is also the kind of purpose-built, approved agent that Section 5 of this paper has in mind.

3. Principle of Traceability, Transparency and Auditability

Trust in EU evaluation procedures depends not only on confidentiality, but also on procedural integrity. If AI tools are permitted as auxiliary instruments, their use must be governed by a framework that ensures full traceability, structured transparency, and effective auditability. The legitimacy of evaluation outcomes requires that any AI-assisted activity be reconstructible, reviewable, and attributable. This entails safeguards at three complementary levels:

A. Ex-ante Transparency and Policy Clarity	Before any AI system is used in evaluation: • Funding bodies must publish clear internal policies defining the scope of permissible AI use. • Evaluators must receive formal guidance and training on acceptable and prohibited uses. • The distinction between auxiliary support and decision-making functions must be explicitly codified. • Conflict-of-interest and confidentiality declarations should be updated to include AI usage clauses.
B. Operational Traceability During Evaluation	Where AI tools are used: • Evaluators must formally declare their use of AI in the evaluation report. • Interactions with AI systems should be logged in a secure and proportionate manner, including time of use and functional purpose. • Where technically feasible, prompts and generated outputs should be retained within the secure evaluation environment. • Systems should ensure attribution of AI use to identifiable evaluators to preserve accountability. • Traceability does not imply bureaucratic burden, but rather the ability to reconstruct how analytical support was used in case of review.
C. Ex-post Auditability and Procedural Safeguards	To preserve procedural fairness: • Funding bodies should establish mechanisms allowing independent audit of AI-assisted evaluations. • In case of formal complaints, appeals, or investigations (including by oversight bodies), relevant AI usage records must be accessible under controlled conditions. • Periodic internal reviews should assess compliance with AI usage policies. • Governance frameworks should include sanctions for undeclared or improper use.

	Auditability is particularly important in light of ongoing scrutiny regarding the role of AI in public decision-making. Robust oversight mechanisms strengthen institutional credibility and reduce litigation risk

Transparency mechanisms are not optional enhancements; they are structural safeguards. Where AI enters the evaluation process, accountability must remain clearly identifiable, human, and reviewable.

Only through structured traceability and independent oversight can AI-assisted evaluation maintain the trust of innovators, evaluators, and European institutions alike.

The February 2026 Ombudswoman inquiry arose from a straightforward problem: evaluators had used AI tools, and there was no way to know how, or to what extent, because nothing had been logged and nothing had been declared. The inquiry did not find that using AI was wrong in principle. It found that the absence of any governance framework made it impossible to verify whether the evaluation had been conducted appropriately. That is the gap Principle 3 addresses. The question is what a well-designed evaluation management system would actually look like if it took this seriously.

Before an evaluator can access any AI-assisted feature within the platform, the system should require them to complete AI usage training and sign an updated confidentiality declaration that includes an explicit AI clause. This is the ex-ante layer: the evaluator knows the rules before they start, and the platform enforces them. During evaluation, every time the evaluator uses an AI tool within the platform, the interaction is logged automatically – which tool, for what purpose, at what time, and where the technology allows it, the prompt and the output. These logs sit inside the secure evaluation environment. Other evaluators cannot see them; applicants cannot see them. Before submitting their Individual Evaluation Report, the evaluator confirms in a structured declaration whether they used AI assistance, for what, and that all scores and written commentary are their own. After evaluation closes, those logs and declarations are retained. If a complaint is filed – as happened in the case that triggered this paper – investigators have something to look at. That is the difference between a system that is auditable in principle and one that is auditable in practice.

4. Principle of Human Oversight and Exclusive Human Responsibility

In line with the European Union's regulatory approach to Artificial Intelligence, particularly the requirement of meaningful human oversight as reflected in the AI Act, the use of AI in proposal evaluation must remain strictly subordinate to human authority.

EU funding decisions involve qualitative judgement, contextual interpretation, risk assessment, and strategic evaluation of innovation potential. These functions require sectoral expertise, experience-based reasoning, and deliberative judgement that cannot be delegated to automated systems.

Where AI tools are used for auxiliary support, they must operate within a clear human-in-the-loop framework:

AI outputs must be critically reviewed and independently validated by the evaluator:

- Evaluators must retain full discretion to accept, modify, or disregard AI-generated content.
- No automated scoring, ranking, or evaluative weighting may be implemented without human intervention.
- Final scoring, written assessments, and panel recommendations must be drafted, reviewed, and approved by human experts.

Responsibility for the evaluation outcome must remain clearly attributable to identifiable individuals. The use of AI cannot dilute, redistribute, or obscure accountability. Evaluators should therefore:

- Formally confirm that their assessment reflects independent expert judgement.
- Declare compliance with applicable AI usage rules.
- Assume full professional and procedural responsibility for the final evaluation report.

Human oversight must be substantive, not symbolic. The authority to interpret, balance, and decide must remain exclusively human at every decisive stage of the evaluation process.

Only by preserving clear human responsibility can EU funding programmes ensure legitimacy, fairness, and legal robustness in an increasingly digital administrative environment.

EAIC does not advocate for an unregulated use of AI in evaluation processes, nor for a blanket prohibition that ignores technological realities and operational needs. Instead, EAIC calls for a structured, transparent and controlled governance framework that enables efficiency gains while safeguarding confidentiality, fairness and human authority.

Such an approach reinforces the credibility of EU funding programmes, protects innovators, and ensures that technological tools serve the evaluation system, not replace it.

An evaluator sitting down with a Horizon Europe EIC Accelerator Stage 2 full proposal has a specific task in front of them: completing an Individual Evaluation Report (IER) structured around three award criteria. Under Excellence, they assess the deep-tech and breakthrough nature of the innovation, whether TRL 5 has been credibly demonstrated, and the strength of the IP strategy. Under Impact, they look at the market opportunity and business model, the realism of financial projections, and whether the innovation contributes to European technological leadership or reduces strategic dependencies. Under Level of Risk, Implementation and Need for Union Support, they consider the team, the risk management approach, the workplan with its milestones and deliverables, and the investment case. The proposal in front of them may be 80 pages long. The challenge is not understanding the criteria – it is locating the relevant content efficiently enough to assess it properly.

This is where a reading navigation aid – running within the governed platform described under Principle 3 – can help without crossing any line. The tool points the evaluator to the sections of the proposal most relevant to each IER criterion, with page references. It flags criteria where it cannot locate substantive content, which the evaluator then verifies. That is all it does. It does not summarise the proposal, it does not characterise whether the content is strong or weak, and it produces nothing that could plausibly end up in the IER as written text. The evaluator still reads the full proposal. The scores and the commentary are still entirely theirs. This is the direct opposite of the failure modes described in the Context section of this paper – the parroting of proposal content, the TRL misreading, the mid-sentence cut-offs that suggest text was copied

rather than written. Before submitting, the evaluator completes the AI usage declaration in the platform. The record is there if it is ever needed.

5. Going beyond: Establish an EU made AI-based evaluation tool for experts appointed by the European Commission

As appears in almost all organisations, AI represents an opportunity, not only a threat. In this context, AI can offer to improve many elements of evaluation by providing efficient, accurate, standardised, and unbiased support to evaluators while also supporting administrators to manage evaluation processes and applicants to understand their fit for the program.

In this regard and recognising that the use of AI in evaluations is widespread, there is an opportunity for the European Commission to establish qualified AI agents that can be used by evaluators. Establishing these agents inside an evaluation platform, where the evaluator is only allowed to use approved agents, will improve transparency, reduce problems of confidentiality and will allow the Commission to monitor the performance of the evaluator and govern their use of AI. Applicants could also have access to the agent to be able to understand their possibilities of being financed and identify potential hallucinations that could occur with their application.

Annex

The European Association of Innovation Consultants (EAIC) also gathered some testimonies from members, suspecting the use of AI in the evaluation process, for example:

- frequent occurrences of the em dash (-) in evaluation reports may raise doubts about the possible use of generative AI in the evaluation writing process. This mark is quite often linked to specific patterns used by Large Language Models (LLMs) but this patterning could also happen when generative AI is used to reformulate or correct a report written by humans.
- similarly, the frequent use of arrows that are rarely used in text written by humans. Real example: “encode→embed→recover→decode validation”;
- frequent parroting of the proposal, either verbatim or summarised, without any analysis or added insight or explanation of the selected score.

About EAIC

The European Association of Innovation Consultants (EAIC) brings together active innovation consultants in the field of European research and innovation projects. The association aims to facilitate the exchange and promotion of best practices, as well as to uphold professional skills and expertise in European RDI funding. To date, the group has assembled more than 50 members who are active throughout Europe.

Contact

EAIC – European Association of Innovation Consultants

Square de Meeûs, 35

1000 Brussels

Belgium

www.eaic.eu