Date: Jul 2, 2026 at 15:19

Source: https://aipolicylab.se/aipex/

AI Policy Lab
at Umeå University

Question Zero: Why Responsible AI Begins Before AI Adoption

Tatjana Titareva (AI Policy Lab, Umeå University), Jason Tucker (Institute for Futures Studies & AI Policy Lab, Umeå University), Rachele Carli (AI Policy Lab, Umeå University), Viktoriia Movchan (AI Policy Lab, Umeå University), Virginia Dignum (AI Policy Lab, Umeå University)

Abstract

The Question Zero (Q0) Self-Assessment Tool for Responsible AI, developed by the AI Policy Lab at Umeå University, supports organisations in posing foundational questions before adopting AI. Grounded in the concept of *Question Zero*: *"Under what conditions should an AI system be adopted, if at all?"*, the tool offers support for cross-functional team discussions covering motivation, stakeholder mapping, system type, adoption process and infrastructure. It is designed for public institutions, civil society organisations, policymakers and other actors looking for responsible AI decision-making approaches. With the help of the tool, we argue that purposeful, problem-led assessment must precede any procurement and deployment decisions.

Keywords: Question Zero, Q0, Responsible AI, AI Self-Assessment, AI Governance, EU AI First Strategy, AI Policy, AI Adoption Motivation, Stakeholder Mapping, Explainability, AI Policy Lab, Umeå University

The Problem with Starting from the Answer

Public and private organisations face growing pressure to adopt artificial intelligence (AI). Governments across the EU and beyond are committing vast public resources to AI acceleration. Organisations, large and small, feel urged to integrate AI into their operations for fear of being perceived as falling behind. The European Commission's Apply AI Strategy (European Commission, 2025) has formalised this pressure into policy direction, framing AI as the default first response to organisational and societal challenges. The same trend can be seen at a national level, as well as across sectors and organisations.

However, this urgency leads to a fundamental disruption of sound decision-making. Rather than beginning with a problem and asking what solutions are available, an “AI First” logic begins with the technology and asks where it can be applied. As we have argued elsewhere, the result is that the question of whether AI is appropriate in a given context is hardly ever

properly posed (Dignum et al., 2025). The consequences of skipping this step are not abstract.

The Dutch childcare benefit scandal, in which an algorithmic system wrongly flagged tens of thousands of parents as committing fraud, disproportionately affecting ethnic minority and low-income families (Amnesty International, 2021), is one of the examples of what might happen when AI is deployed without adequate prior deliberation about necessity, appropriateness and who might be harmed. Similarly, the UK's A-level grading algorithm reinforced existing inequalities, putting students from smaller schools and lower socio-economic backgrounds at a disadvantage by relying on historical data to determine results (Kolkman, 2020). In the United States, ProPublica's investigation into the COMPAS recidivism algorithm revealed significant racial biases, with Black defendants nearly twice as likely as their white counterparts to be wrongly flagged as high risk (Larson et al., 2016). Australia's Robodebt scandal exposed the dangers of deploying flawed automated decision-making systems at scale, showing how vulnerable citizens can be unjustly penalised without adequate oversight, transparency and safeguards (Royal Commission into the Robodebt Scheme, 2023).

Together, these cases illustrate how without robust AI governance, such systems can deepen discrimination against marginalised groups and undermine public trust in key institutions. While these cases differ in context and technical form, they share an important feature: technology was introduced without adequately addressing Question Zero (Q0), meaning the prior scrutiny of necessity, appropriateness, accountability and social impact was weak or absent.

The question ***Q0 ("Under what conditions should an AI system be adopted, if at all?")*** *we pose*, is a call for rigour and strategic foresight. It asks organisations to define the problem before considering technological solutions, to assess alternatives, to map who benefits and who bears risks and to ensure that any decision on how to proceed is genuinely justified. Only sometimes, not always, is AI the right answer. Thus, the purpose of the Q0 Self-Assessment Tool is to facilitate a structured pause to allow for better decision making about responsible AI.

What is the Q0 Self-Assessment Tool?

The Q0 Self-Assessment Tool is a free assessment tool designed not as a compliance checklist or technical audit, but as a structured self-reflection instrument for cross-functional groups within organisations. Its primary goal is to support decision-makers in working through the foundational questions of responsible AI governance before procurement or deployment decisions are finalised (Titareva et al., 2026).

The tool is organised into five sections: *Why, Who, What, How* and *Where*, each targeting a distinct dimension of responsible AI governance. These sections are not a linear sequence; organisations are encouraged to move between them iteratively, as answers in one area frequently raise additional questions in others.

Section A: Why (Motivation) asks organisations' decision-makers and employees to articulate the problem they are attempting to solve, the reasons they are considering AI, the alternatives they have considered (including human and non-AI technical solutions) and why and whether AI is the best solution for the existing problem. This is the core of Q0: encouraging explicit justification before any further investment of time or resources into AI procurement and deployment.

Section B: Who (Stakeholder Mapping) asks organisations to identify the stakeholders who may be directly or indirectly affected by the adoption of an AI system. This includes colleagues, employees, customers, and members of marginalised or underrepresented groups who are often overlooked in initial assessments (as well as at later stages). The section also prompts organisations to consider who currently performs the tasks the AI system is intended to undertake, who is likely to benefit, who could be harmed and whether meaningful opt-out options exist. By mapping the full stakeholder landscape, organisations can assess how different groups may experience the adoption of AI and ensure that no group is disproportionately burdened, excluded or deprived of meaningful choice.

Section C: What (Type of AI System) asks organisations to define and describe the specific AI system, method or tool under consideration (e.g., predictive, generative, categorising or hybrid), as well as whether it will be procured off the shelf, developed in-house or built for specific needs of an organisation. This section encourages organisations to ensure that the chosen system fits the problem they want to solve and that the system’s complexity, development approach and flexibility match its intended use.

Section D: How (Adoption Process) asks organisations to consider how the AI system would be deployed, how existing workflows would change, how outputs and performance would be monitored, how data security would be ensured and how affected users would be able to understand the basis of the system's conclusions. It also prompts organisations to define how complaints and errors would be handled and who would be accountable for oversight throughout the system's lifecycle. This section emphasises that accountability must be embedded from the beginning rather than added retrospectively, and that explainability should be treated as an ongoing organisational practice rather than merely a compliance requirement.

Section E: Where (Infrastructure and Control) asks organisations to consider where the AI system will operate, where data will be stored, where the provider is based and where the training data originates. These questions address issues of digital sovereignty, data

jurisdiction and the extent to which organisations can maintain meaningful control over the AI systems they adopt. This section emphasises that effective governance and accountability depend on understanding where the system operates, who controls it and how data is collected, stored and managed. It therefore encourages organisations to ensure transparency regarding infrastructure, models and training data in order to meet legal, security and ethical responsibilities and requirements.

Each section concludes with a confidence rating on a five-point scale, allowing organisations to reflect on how confidently they can answer the questions within that dimension. Rather than measuring compliance or assigning a level of AI maturity, the ratings are intended to stimulate discussion, identify areas of uncertainty and highlight where further information, stakeholder engagement or organisational preparation may be needed. The aggregated scores are visualised in a radar diagram, providing an overview of confidence across all five dimensions. This visualisation enables organisations to compare results across teams, over time or between different AI use cases, thus supporting internal learning, reflection and continuous improvement. The ratings are not intended for external accreditation, benchmarking or audit, but as a self-assessment tool that helps organisations recognise strengths, identify gaps and guide responsible AI governance.

A final section of concluding questions invites decision-makers and employees of organisations to reflect on the insights gained throughout the self-assessment process by evaluating any pilot experiences, considering the expected benefits of the proposed AI system alongside its possible risks and harms and identifying appropriate next steps. Rather than prescribing decisions or producing a pass-fail outcome, the tool is designed to support informed, context-sensitive judgement, recognising that responsible AI adoption depends on an organisation's objectives, values, legal obligations and operational context. As a decision-support instrument, the Q0 tool helps structure deliberation, surface assumptions and strengthen transparency, but responsibility for interpreting the findings and deciding how to act upon them remains with the organisation and its employees as part of its internal AI governance.

Figure 1
Question Zero Conceptual Operationalisation

Positioning the Q0 Self-Assessment Tool Among Existing AI Assessment Frameworks

The Q0 self-assessment tool builds on a growing ecosystem of AI governance and assessment instruments. At the same time, it addresses a different stage of the AI lifecycle. Existing frameworks such as the European Commission's Assessment List for Trustworthy Artificial Intelligence (ALTAI), UNESCO's Ethical Impact Assessment (EIA), the UNESCO Readiness Assessment Methodology (RAM), the UNDP Human Rights Impact Assessment Toolkit and similar international and national frameworks support responsible AI through structured reflection on legal, ethical, technical and organisational issues. Although these tools have different purposes and target different audiences, they share a common goal:

They help governments or organisations identify, manage and govern the risks associated with AI systems and promote transparency, accountability, human rights and public trust.

The Q0 tool shares many of these principles. Like ALTAI, it addresses human oversight, transparency, accountability, fairness and data governance. Similar to UNESCO's Ethical Impact Assessment, it encourages multidisciplinary participation and considers the wider social impacts of AI alongside technical issues. The UNDP Human Rights Impact Assessment Toolkit also emphasises the importance of identifying rights holders and affected communities. UNESCO's Readiness Assessment Methodology takes a broader perspective by examining institutional capacity, governance, infrastructure and organisational readiness. Q0 tool follows a similar approach by treating responsible AI governance as an institutional challenge rather than only a technical one.

The main difference is not the *questions* themselves, but *when they are asked*. The assessment tools mentioned above often assume that an organisation has already decided to develop, procure or deploy an AI system. Their purpose is to assess whether that system is trustworthy, legally compliant or ethically governed. Some frameworks do include questions about necessity or proportionality during the pre-adoption stage. However, these questions usually form one part of a broader impact or compliance assessment when the AI system’s adoption project has already been initiated.

The Q0 tool starts one step earlier. It asks whether AI should be adopted at all, and if yes, under what conditions. This is the central purpose of Question Zero. The tool encourages organisations to define the problem first, consider non-AI alternatives and examine governance, stakeholder impacts and organisational readiness before resources are committed or procurement begins. It also places greater emphasis on collective deliberation and participation. It recommends cross-functional participation, the inclusion of affected stakeholders where possible, iterative use of the tool, documentation of disagreements and integration with existing governance processes. The goal is not a single assessment but an ongoing process of organisational learning and reflection.

A second difference is the intended use of the tool. Many existing assessment instruments support regulatory compliance, formal impact assessment or risk management. The Q0 tool has a different purpose. It is a self-assessment instrument. It does not generate a compliance score or certify that an AI system or case is trustworthy. The confidence ratings are intended to stimulate discussion, reveal uncertainty and identify gaps in knowledge. The radar diagram allows organisations to compare discussions across teams, projects or points in time. This helps identify differences in understanding and areas that require further attention.

The Q0 Self-Assessment Tool is intended to complement, not replace, existing governance frameworks. Organisations that decide AI is appropriate can then carry out more detailed

assessments, such as ethical impact assessments, human rights impact assessments, data protection impact assessments or assessments required under the national or regional legislation (e.g., the EU AI Act and others). Q0 tool therefore supports an earlier stage of decision-making. It aims to improve the quality of the decision that determines whether these later assessments will be necessary.

The Q0 tool also has limitations. It is a self-assessment tool. Its value depends on organisations' employees engaging with the questions and including a diversity of perspectives. It cannot remove organisational bias or guarantee that all relevant stakeholders are represented. It cannot ensure that recommendations will be implemented. It is also not a substitute for legal, technical or human rights assessments where these are required. Finally, no questionnaire can capture the full complexity of every organisational context. The Q0 tool should therefore be seen as a governance aid rather than a decision-making mechanism. It supports better judgement, but responsibility for the final decision always remains with the organisations' leadership and employees.

Example of the Q0 Self-Assessment Tool in Practice

A university department is considering introducing a generative AI assistant to answer students' questions about courses, deadlines and administrative procedures. Instead of immediately assuming that an AI chatbot is the right solution, the department first uses the Q0 Self-Assessment Tool. Addressing the questions in the section ***"Why"***, university employees from diverse departments and functional areas define the problem, for example, long response times during busy periods. They also consider other options, such as improving online guidance, extending office hours or employing more student assistants. AI is treated as one possible solution, not the starting point.

The remaining Q0 tool's sections help to broaden the discussion. Under the section ***"Who"***, the participants identify as many categories of stakeholders as possible, who may be affected, including students, lecturers, administrative staff, IT services and others. It also considers students with disabilities, international students and others who may be disadvantaged by digital-only support. Under ***"What"***, participants compare different AI systems and ask whether a simpler non-AI or rule-based solution could meet the need. Under the sections ***"How"*** and ***"Where"***, they discuss governance, accountability, transparency, data security, hosting arrangements and whether the university keeps appropriate control over its data.

At the end of the process, the group gives a confidence rating for each section. The radar diagram shows where participants feel confident and where more work is needed. In this example, confidence could be high for the motivation behind the project but lower for infrastructure, data governance and long-term accountability. These results do not

determine whether AI should be adopted. Instead, they help the department identify where further discussion, expert advice or pilot testing is needed before making a final decision.

Conclusion

AI is not inevitable. Its adoption is not inherently beneficial. Responsible AI governance begins before procurement, development or deployment. It begins by asking whether AI is the best solution for your current problem, if so, under what conditions. This is the purpose of Question Zero.

The Q0 Self-Assessment Tool provides organisations with a practical way to support this early reflection. It does not replace existing governance frameworks or assessments. Instead, it complements them by addressing a stage of decision-making that is often overlooked. It focuses on motivations, stakeholders, system choice, adoption, infrastructure and organisational control.

The value of the Q0 Self-Assessment Tool lies not in providing answers, but in improving the quality of the questions organisations ask before adopting AI systems. It creates space for discussion, makes assumptions explicit and helps identify issues that require further attention. In the end, the responsibility for deciding whether to adopt an AI system remains with the organisations' employees. The Q0 process helps ensure that this decision is more informed, transparent and accountable.

Access the Tool

The Q0 Self-Assessment Tool (Version 3, April 2026) is available: HERE

References

Ala-Pietilä, P., Bonnet, Y., Bergmann, U., Bielikova, M., Bonefeld-Dahl, C., Bauer, W., ... & Van Wynsberghe, A. (2020). *The assessment list for trustworthy artificial intelligence (ALTAI)*. European Commission.

Amnesty International (2021, October 2025). *Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal*. https://www.amnesty.org/en/documents/eur35/4686/2021/en/

Dignum, V., Carli, R., Ericson, P., Titareva, T., & Tucker, J. (2025). 'AI First' to 'Purpose First': Rethinking Europe's AI Strategy. *AI Policy Exchange Forum (AIPEX)*. https://doi.org/10.63439/LPOU6506

Dignum, V., Carli, R., Dahlgren Lindström, A., Ericson, P., Titareva, T., & Tucker, J. (2026, June). Question Zero for Explainability and Vice Versa: The Case of the EU's AI First Strategy. In *International Conference on Human-Computer Interaction* (pp. 17-31). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-29456-2_2

European Commission (2026, March 27). *Apply AI Strategy. Shaping Europe's Digital Future*. https://digital-strategy.ec.europa.eu/en/policies/apply-ai

Kolkman, D. (2020, August 26). "f**k the algorithm?": What the world can learn from the UK's A-level grading fiasco. *The London School of Economics and Political Science (LSE) Impact*. https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/#:~:text=%E2%80%9CF**k%20the%20algorithm%E2%80%9D?:%20What%20the%20world%20can,address%20through%20making%20their%20algorithms%20more%20explainable.

Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016, May 23). How we analyzed the COMPAS recidivism algorithm. *ProPublica*. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Royal Commission into the Robodebt Scheme (2023, July 7). *The Report of the Royal Commission into the Robodebt Scheme to the Governor-General, His Excellency, General the Honourable David Hurley AC DSC (Retd)*. https://robodebt.royalcommission.gov.au/publications/report

Titareva, T., Carli, R., Dahlgren Lindström, A., Dignum, V., Fabris, B., Ericson, P., & Tucker, J. (2026). *Q0 Self-Assessment Tool with Guiding Questions for Responsible AI Approach by AI Policy Lab (AIPL)*. Test version 3 from April 2026. Umeå University.

UNESCO. (2023). *Ethical Impact Assessment: A tool of the Recommendation on the Ethics of Artificial Intelligence*. https://doi.org/10.54678/YTSA7796

UNESCO. (2023). *Readiness Assessment Methodology: a Tool of the Recommendation on the Ethics of Artificial Intelligence*. https://doi.org/10.54678/YHAA4429

United Nations Development Programme. (2025). *Human Rights Impact of AI Assessment Tool*. https://www.undp.org/eurasia/publications/human-rights-impact-ai-assessment-toolkit