



**By Karuana Gatimu, Principal Manager  
Microsoft 365 Customer Advocacy Group**

Disclaimer: This document has been prepared by Microsoft to provide insights into AI governance best practices for generative AI experiences including, but not limited to Microsoft Copilot and Copilot for Microsoft 365.

This document is provided on an “as is” basis and to the maximum extent permitted by law Microsoft disclaims all conditions, warranties, and guarantees, express or implied, including but not limited to any warranty or guarantee that the use of the framework set out in this document will not infringe any rights or any warranty or guarantee of merchantability or fitness for a particular purpose. The document is intended to be used to supplement and inform your Microsoft 365, Microsoft Copilot and/or Copilot for Microsoft 365 adoption approach for your organization. Content specified in the document should be used as an example only. Before using components set out in this document, you should evaluate their suitability for your organization. In particular, if you choose to act upon the recommendations of the approach, then you do so at your own risk.

Apart from any use permitted under the Copyright Act 1968, and the rights explicitly granted above, all rights are reserved.

v2.1

# Creating an AI Council - A Guide for Responsible and Ethical AI Governance

## Introduction

AI technology is transforming every aspect of business and society, creating new opportunities and challenges for organizations. However, AI also poses significant risks and ethical dilemmas, such as bias, privacy, accountability, transparency, and social impact. Therefore, it is essential for organizations to establish a governance framework that ensures the responsible and ethical use of AI across their operations and stakeholders.

One way to achieve this is to create an *AI Council*, a **cross-functional and multidisciplinary body** that oversees and guides the development, deployment, and evaluation of AI solutions within the organization. An AI Council can help to align the organization's AI strategy with its vision, values, and goals, as well as to identify and mitigate potential risks and harms of AI. An AI Council can also foster a culture of trust, collaboration, and innovation among the AI practitioners, users, and beneficiaries within and outside the organization.

In this document, we will provide some guidance on how to create an AI Council in your organization, based on the best practices and recommendations from various sources, such as the OECD Principles on AI, the EU Guidelines for Trustworthy AI, and the Microsoft AI Governance Framework. We will cover the following topics:

- The objectives and functions of an AI Council
- The composition and structure of an AI Council
- The roles and responsibilities of an AI Council
- The processes and tools for an AI Council
- The challenges and benefits of an AI Council

## The Objectives and Functions of an AI Council

The main objective of an AI Council is to ensure that the organization's AI solutions are aligned with its ethical principles, legal obligations, and social responsibilities. To achieve this, an AI Council performs the following functions:

- **Define and communicate the organization's AI vision**, values, and policies, based on the relevant ethical frameworks and standards, such as the OECD Principles on AI, the EU Guidelines for Trustworthy AI, and the Microsoft AI Governance Framework.
- **Review and approve the AI use cases** and projects proposed by the different business units and teams within the organization, based on the criteria of feasibility, desirability, and viability, as well as the potential benefits and risks of AI.
- **Model trusted leadership behavior for driving change** including, but not limited to curiosity, inclusiveness, encouragement and empathy. Authenticity while on the shared learning journey is a key aspect of understanding risks, issues, opportunities, and gaps in depth.
- **Monitor and evaluate the performance and impact of the AI solutions** deployed by the organization, using the appropriate metrics and indicators, such as accuracy, fairness, transparency, accountability, and sustainability.
- **Provide guidance and support to the AI practitioners** and users within the organization, such as data scientists, engineers, developers, analysts, and managers, on how to apply the AI principles, policies, and best practices in their work.
- **Engage and collaborate with the external stakeholders** of the organization, such as customers, partners, regulators, civil society, and academia, on the issues and opportunities related to AI, and solicit their feedback and input on the organization's AI strategy and solutions.
- **Promote and facilitate the learning and innovation of AI** within the organization, by creating and sharing the knowledge, resources, and tools for AI development and use, and by encouraging the experimentation and exploration of new AI possibilities and applications.

## The Composition and Structure of an AI Council

The composition and structure of an AI Council should reflect the diversity and complexity of the AI domain, as well as the organization's culture and context. An AI Council should include the following elements:

- A **core team of members** who are responsible for the day-to-day operations and decision-making of the AI Council, such as setting the agenda, conducting the reviews, providing the guidance, and reporting the outcomes. The core team should consist of representatives from the distinct functions and levels of the organization,

such as senior management, legal, compliance, risk, ethics, data, technology, business, and human resources.

- A **network of experts** who are consulted by the core team on specific AI topics and issues, such as technical, legal, ethical, social, and domain-specific aspects. The network of experts should consist of internal and external experts, such as data scientists, engineers, developers, lawyers, ethicists, sociologists, psychologists, and domain specialists.
- A **steering committee of sponsors** who are accountable for the overall direction and oversight of the AI Council, such as defining the vision, values, and policies, approving the budget and resources, and endorsing the recommendations and actions. The steering committee should consist of senior executives and leaders from the organization, such as the CEO, CTO, CDO, CIO, CISO, CFO, and CMO or other senior roles as appropriate.
- A **stakeholder group of participants** who participate in the implementation and evaluation of AI solutions, such as proposing, developing, deploying, and using AI solutions. The stakeholder group should consist of the AI practitioners and users from the different business units and teams within the organization, as well as the customers, partners, regulators, civil society, and academia outside the organization.

The composition and structure of an AI Council can vary depending on the size, scope, and maturity of the organization and its AI activities. However, it is important to ensure that the AI Council has the following characteristics:

- **Diversity:** The AI Council should include members and experts from diverse backgrounds, perspectives, and experiences, to ensure a comprehensive and balanced view of AI.
- **Independence:** The AI Council should operate with autonomy and authority, to ensure a credible and objective assessment of AI.
- **Transparency:** The AI Council should communicate its processes, criteria, and outcomes, to ensure a clear and consistent understanding of AI.
- **Accountability:** The AI Council should be accountable for its decisions and actions, to ensure responsible and ethical use of AI.

# The Roles and Responsibilities of an AI Council

The roles and responsibilities of an AI Council can be divided into four phases in alignment with the overall [Microsoft Cloud Adoption Framework](#) and specifically the [Microsoft 365 Implementation Framework](#): Getting Ready, Onboard & Engage, Deliver Impact and Extending and Optimizing business value.



These phases incorporate the three distinct journeys your organization and its people will embark upon: the leadership journey, user enablement and technical readiness. The creation of an AI Council is an activity within the leadership journey in your organization.

In each phase, the AI Council coordinates with user enablement and technical readiness teams as well as external experts to understand the risks, issues and opportunities presented by AI in the organization. It often performs the following roles and responsibilities by phase:

**Getting Ready:** In this phase, the AI Council defines and communicates the organization's AI vision, values, and policies, and reviews and approves the AI use cases and projects proposed by the different business units and teams within the organization.

- Define and communicate the organization's AI vision, values, and policies, based on the relevant ethical frameworks and standards, such as the OECD Principles on AI, the EU Guidelines for Trustworthy AI, and the Microsoft AI Governance Framework.
- Establish and communicate the criteria and process for reviewing and approving the AI use cases and projects, such as the feasibility, desirability, and viability, as well as the potential benefits and risks of AI.
- Review and approve the AI use cases and projects submitted by the different business units and teams within the organization, using the criteria and process established by the AI Council.

**Onboard & Engage:** In this phase, the AI Council provides guidance and support to the AI practitioners and users within the organization on how to apply the AI principles, policies,

and best practices in their work. Opportunities to engage in this phase include, but are not limited to:

- Provide guidance and support to user enablement and technology leaders on how to apply the AI principles, policies, and best practices in their work, such as the data quality, security, privacy, fairness, transparency, accountability, gathering feedback and sustainability of AI.
- Provide feedback and recommendations to the proposers of the AI use cases and projects, such as the alignment, compliance, and improvement of the AI solutions.
- Consult and collaborate with the network of experts on specific AI topics and issues, such as technical, legal, ethical, social, and domain-specific aspects.
- Monitor and supervise the development and deployment of AI solutions, using the appropriate tools and methods, such as audits, tests, validations, and certifications. This includes active participation in AI Service Health Reviews as conducted by a joint User and Technical Enablement team.
- Identify and mitigate the potential risks and harms of the AI solutions, using the appropriate measures and actions, such as risk assessments, impact assessments, mitigation plans, and contingency plans.

**Deliver Impact:** In this phase, the AI Council monitors and evaluates the performance and impact of the AI solutions deployed by the organization, using the appropriate metrics and indicators, such as accuracy, fairness, transparency, accountability, and sustainability. It may conduct activities such as:

- Fostering a culture of trust in sharing AI enablement experiences. Ensure transparency in all forms of projects including those deemed as learning experiences rather than demonstrable successes.
- Engage and collaborate with the external stakeholders of the organization, such as customers, partners, regulators, civil society, and academia, on the issues and opportunities related to AI, and solicit their feedback and input on the organization's AI strategy and solutions.
- Report and communicate the outcomes and learnings of the AI solutions, using the appropriate channels and formats, such as dashboards, reports, presentations, and publications.
- Promote and facilitate the learning and innovation of AI within the organization, by creating and sharing the knowledge, resources, and tools for AI development and use, and by encouraging the experimentation and exploration of new AI possibilities and applications.

**Extend and Optimize:** In this phase, the AI Council evaluates future opportunities based on current performance, identified business opportunities, competitive differentiators, and emerging technologies.

- Monitor and evaluate emerging technologies for new capabilities.
- Review with internal/external stakeholders and experts opportunities for competitive differentiation or business value acceleration through the use of AI capabilities.
- Review recommendations made through regular Service Health Reviews as conducted by the joint implementation team.
- Understand employee sentiment with existing AI services through established feedback channels. Evaluate for capability adjustments or to identify gaps.
- Provide recommendations for further investment in AI capabilities for the organization based on the analysis of the above.

## The Processes and Tools for an AI Council

The processes and tools for an AI Council are the methods and instruments that enable the AI Council to perform its functions effectively and efficiently. The processes and tools for an AI Council can be categorized into four types: communication, collaboration, evaluation, and documentation. In each type, the AI Council can use the following processes and tools:

**Communication:** The communication processes and tools are the methods and instruments that enable the AI Council to communicate its vision, values, policies, criteria, processes, outcomes, and learnings to the different stakeholders of the organization and its AI solutions, such as the senior management, the business units, the AI practitioners, the AI users, the customers, the partners, the regulators, the civil society, and the academia. The communication processes and tools can include the following:

- **Meetings:** The meetings are the regular and ad hoc gatherings of the AI Council members and experts, as well as the other stakeholders, to discuss, review, approve, monitor, evaluate, and report the AI use cases, projects, and solutions. Select meetings may be open to the general population of the organization.
- **Forums & Communities:** Regular engagement throughout internal communities of practice regarding AI related topics. The forums are online platforms that are used to collect the feedback and input of the different stakeholders on the organization's AI

vision, values, policies, criteria, processes, outcomes, and learnings, as well as on specific AI use cases, projects, and solutions, through moderated discussion.

- **Events:** Conduct general awareness town halls or large-scale meetings to aid in the user enablement and understanding of AI capabilities and their use in the organization.
- **Newsletters:** The newsletters are the periodic publications that are used to communicate the AI Council's vision, values, policies, criteria, processes, outcomes, and learnings to the different stakeholders.
- **Websites:** The websites are online platforms that are used to communicate the AI Council's vision, values, policies, criteria, processes, outcomes, and learnings to the different stakeholders, as well as providing access to the resources and tools for AI development and use.
- **Presentations\Video:** The presentations are the oral and visual displays that are used to communicate the AI Council's vision, values, policies, criteria, processes, outcomes, and learnings to the different stakeholders, such as in conferences, workshops, seminars, and webinars.
- **Publications:** The publications are written and digital documents that are used to communicate the AI Council's vision, values, policies, criteria, processes, outcomes, and learnings to the different stakeholders, such as in reports, papers, articles, and books.

**Collaboration:** The collaboration processes and tools are the methods and instruments that enable the AI Council to collaborate with the different stakeholders of the organization and its AI solutions, such as the senior management, the business units, the AI practitioners, the AI users, the customers, the partners, the regulators, the civil society, and the academia, on the issues and opportunities related to AI, and to solicit their feedback and input on the organization's AI strategy and solutions. The collaboration processes and tools can include the following:

- **Surveys:** The surveys are the questionnaires that are used to collect the feedback and input of the different stakeholders on the organization's AI vision, values, policies, criteria, processes, outcomes, and learnings, as well as on the specific AI use cases, projects, and solutions.
- **Interviews:** The interviews are the conversations that are used to collect the feedback and input of the different stakeholders on the organization's AI vision, values, policies, criteria, processes, outcomes, and learnings, as well as on the specific AI use cases, projects, and solutions.



- **Research & focus groups:** The focus groups are the small groups of stakeholders that are used to collect the feedback and input of the different stakeholders on the organization's AI vision, values, policies, criteria, processes, outcomes, and learnings, as well as on the specific AI use cases, projects, and solutions, through a moderated discussion.
- **Workshops:** The workshops are the interactive sessions that are used to collect the feedback and input of the different stakeholders on the organization's AI vision, values, policies, criteria, processes, outcomes, and learnings, as well as on the specific AI use cases, projects, and solutions, through a facilitated activity.
- **Communities:** The communities are the groups of stakeholders that are used to collect the feedback and input of the different stakeholders on the organization's AI vision, values, policies, criteria, processes, outcomes, and learnings, as well as on the specific AI use cases, projects, and solutions, through a shared interest and engagement.
- **Evaluation:** The evaluation processes and tools are the methods and instruments that enable the AI Council to evaluate the feasibility, desirability, and viability, as well as the potential benefits and risks of the AI use cases, projects, and solutions, and to monitor and measure the performance and impact of the AI solutions deployed by the organization, using the appropriate metrics and indicators, such as accuracy, fairness, transparency, accountability, and sustainability. The evaluation processes and tools can include the following:
  - **Audits:** The audits are the systematic and independent examinations of the AI use cases, projects, and solutions, to verify their compliance with the organization's AI vision, values, policies, criteria, and processes, as well as with the relevant ethical frameworks and standards, such as the OECD Principles on AI, the EU Guidelines for Trustworthy AI, and the Microsoft AI Governance Framework.
  - **Tests:** The tests are empirical and experimental assessments of the AI use cases, projects, and solutions, to measure their technical and functional performance, such as the accuracy, reliability, robustness, and efficiency of AI.
  - **Validations:** The validations are the logical and theoretical assessments of the AI use cases, projects, and solutions, to verify their conceptual and methodological soundness, such as the validity, consistency, coherence, and completeness of AI.
  - **Certifications:** The certifications are the formal and official recognitions of the AI use cases, projects, and solutions, to attest to their quality and trustworthiness, such as the conformity, interoperability, security, and safety of AI.
  - **Risk assessments:** The risk assessments are systematic and comprehensive analyses of AI use cases, projects, and solutions, to identify and evaluate their potential risks and harms, such as bias, privacy, accountability, transparency, and social impact of AI.

- **Impact assessments:** The impact assessments are the systematic and comprehensive analyses of the AI use cases, projects, and solutions, to identify and evaluate their potential benefits and opportunities, such as the value, utility, innovation, and sustainability of AI.
- **Mitigation plans:** The mitigation plans are the strategies and actions that are designed and implemented to reduce or eliminate the potential risks and harms of the AI use cases, projects, and solutions, such as the corrective, preventive, and protective measures for AI.
- **Contingency plans:** The contingency plans are the strategies and actions that are designed and implemented to cope with or recover from the potential risks and harms of the AI use cases, projects, and solutions, such as the backup, recovery, and emergency measures for AI.
- **Metrics:** Metrics are the quantitative and qualitative measures that are used to monitor and evaluate the performance and impact of AI solutions deployed by the organization, such as the accuracy, fairness, transparency, accountability, and sustainability of those solutions.
- **Indicators:** The indicators are the signals and signs that are used to monitor and evaluate the performance and impact of the AI solutions deployed by the organization, such as the feedback, ratings, reviews, and complaints of AI.
- **Documentation:** The documentation processes and tools are the methods and instruments that enable the AI Council to document and record the AI use cases, projects, and solutions, as well as their criteria, processes, outcomes, and learnings, for the purposes of transparency, accountability, and learning. The documentation processes and tools can include the following:
  - **Data sheets:** The data sheets are the documents that describe the characteristics and properties of the data that are used for the AI use cases, projects, and solutions, such as the source, type, format, size, quality, security, and privacy of data.
  - **Model cards:** The model cards are documents that describe the characteristics and properties of the models that are used for the AI use cases, projects, and solutions, such as the architecture, parameters, inputs, outputs, performance, and limitations of models.
  - **Fact sheets:** The fact sheets are the documents that describe the characteristics and properties of the AI use cases, projects, and solutions, such as the purpose, scope, functionality, and impact of AI.
  - **Explanation sheets:** The explanation sheets are the documents that explain the logic and reasoning behind the AI use cases, projects, and solutions, such as the assumptions, methods, algorithms, and outcomes of AI.

- **Consent forms:** The consent forms are the documents that obtain the consent and permission of the data subjects and the AI users for the AI use cases, projects, and solutions, such as the data collection, processing, and sharing, and the AI functionality, and impact.
- **Dashboards:** The dashboards are visual and interactive displays that show the status and progress of the AI use cases, projects, and solutions, as well as their performance and impact, using metrics and indicators, such as the accuracy, fairness, transparency, accountability, and sustainability of AI.
- **Reports:** The reports are the written and digital documents that summarize and communicate the AI use cases, projects, and solutions, as well as their criteria, processes, outcomes, and learnings, using the data sheets, model cards, fact sheets, explanation sheets, and dashboards, such as the feasibility, desirability, and viability, as well as the potential benefits and risks of AI.
- **Knowledge bases:** The knowledge bases are the repositories and databases that store and organize the AI use cases, projects, and solutions, as well as their criteria, processes, outcomes, and learnings, using the data sheets, model cards, fact sheets, explanation sheets, consent forms, dashboards, and reports, for the purposes of transparency, accountability, and learning.

## The Challenges and Benefits of an AI Council

The creation and operation of an AI Council can pose some challenges and difficulties for the organization and its AI solutions, such as the following:

**Complexity:** The AI domain is complex and dynamic, involving multiple dimensions, aspects, and factors, such as technical, legal, ethical, social, and domain-specific aspects, which require a comprehensive and balanced understanding and assessment of AI.

**Pace of Change:** AI capabilities are changing at a rapid pace. Staying up to date requires ongoing collaboration with internal and external experts to understand the evolving landscape.

**Authority:** The AI Council should have a clear mandate and authority to oversee and guide the AI strategy, governance, and implementation of the organization. This can ensure the alignment of AI with the organization's vision, mission, and values, but also requires the support and trust of the senior management and other relevant parties.

**Impact:** The AI Council must avoid staying in the analysis phase for too long and generate iterations of guidance at a pace that supports the evolving landscape and needs of the business.

**Collaboration:** To be effective the AI Council must include feedback and insights from the core workforce, not just its executive management, to have a deeper understanding of risks, issues, and opportunities for cross-organizational process improvement. This feedback can be gathered through the user enablement workstream and from insights presented from Service Health Reviews and other sentiment gathering tools.

## Conclusion

The creation of an AI Council can enhance and accelerate your journey to becoming an AI powered organization while managing risk and addressing operational challenges. It provides an opportunity for learning across your leadership team in how to approach leveraging this transformational and important emerging technology.

Most importantly, it creates an opportunity for leaders to expand their skills of leading in the era of AI by creating an environment of trust with employees, customers, and partners about the intended use of AI within the business. By modelling transparent and consistent behavior as AI solutions are evaluated trust throughout the organization can translate to deeper employee engagement, competitive advantages, and a readiness to optimize the opportunities for growth presented by AI.