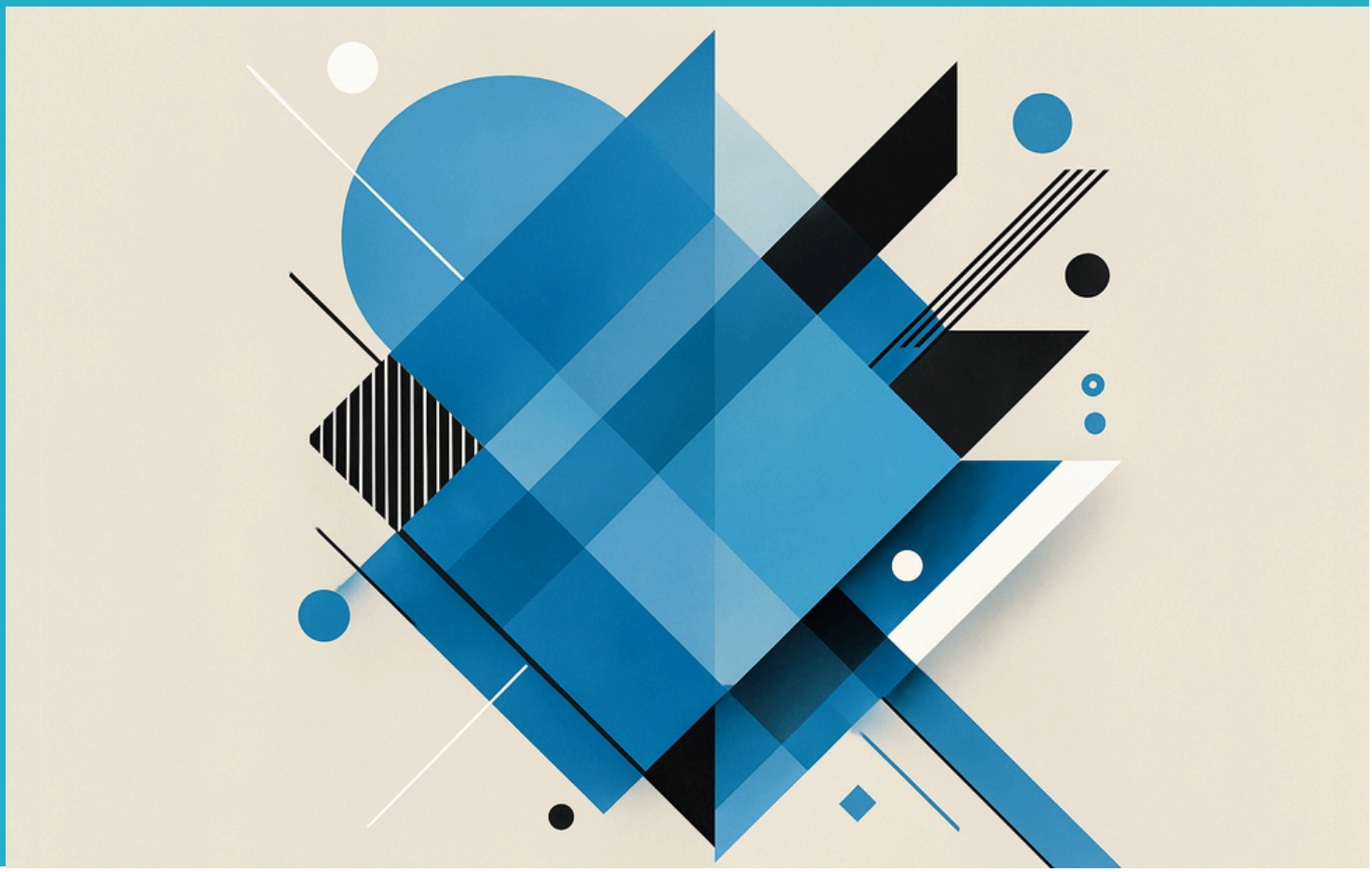
WHITE PAPER
April 2025

AI GOVERNANCE INITIATIVE | OXFORD MARTIN SCHOOL | UNIVERSITY OF OXFORD

# PROPOSAL FOR UN SCIENTIFIC PANEL ON AI: BALANCING RIGOR AND LEGITIMACY

Authors: Julia Morse, Robert Trager & Ranjit Lall



# AIGI Proposal for UN Independent Scientific Panel on AI: Balancing Rigor and Legitimacy

Julia C. Morse[1+], Robert Trager[2], and Ranjit Lall[3]

**28 April 2025[4]**

[1] Oxford Martin School AI Governance Initiative; University of California, Santa Barbara.
[+] Primary author who contributed most significantly to the direction and content of the paper
[2] Blavatnik School of Government, University of Oxford; Centre for the Governance of AI; Oxford Martin School AI Governance Initiative
[3] University of Oxford
[4] *This version of the white paper has been modified to reflect the 19 March 2025 zero draft.*

We are grateful to Amanda Craig, Sam Daws, Anna Jahn, Kevin Kohler, and Howard Wachtel for their valuable feedback on this project.

Corresponding author: julia.morse@oxfordmartin.ox.ac.uk

## Executive Summary

This white paper makes recommendations on a possible mandate and modalities for the proposed UN independent international scientific panel on artificial intelligence (AI). Our recommendations are designed to combine the goals of scientific rigor and credibility with political legitimacy and policy impact. We build on the 19 March 2025 zero draft's vision that the panel provide "multidisciplinary, independent and evidence-based scientific assessments of the opportunities, risks, capabilities, and impacts of Artificial Intelligence…through timely and adaptive outputs that address rapid advancements in the field" (OP 2). To achieve this objective, the paper expands on the zero draft's proposed governance framework and suggests fleshing it out as follows: an AI Advisory Council (40 member states), an AI Expert Committee (15-20 technical experts, with some assigned to specific topics and areas), and 4 working groups (15-20 technical experts) covering the topics of AI capabilities and risks, AI impacts, AI accessibility, and AI and the sustainable development goals (SDGs). We suggest each working group be chaired by a member of the AI Expert Committee and also have an organizational vice-chair with significant technical knowledge and expertise in a given area.

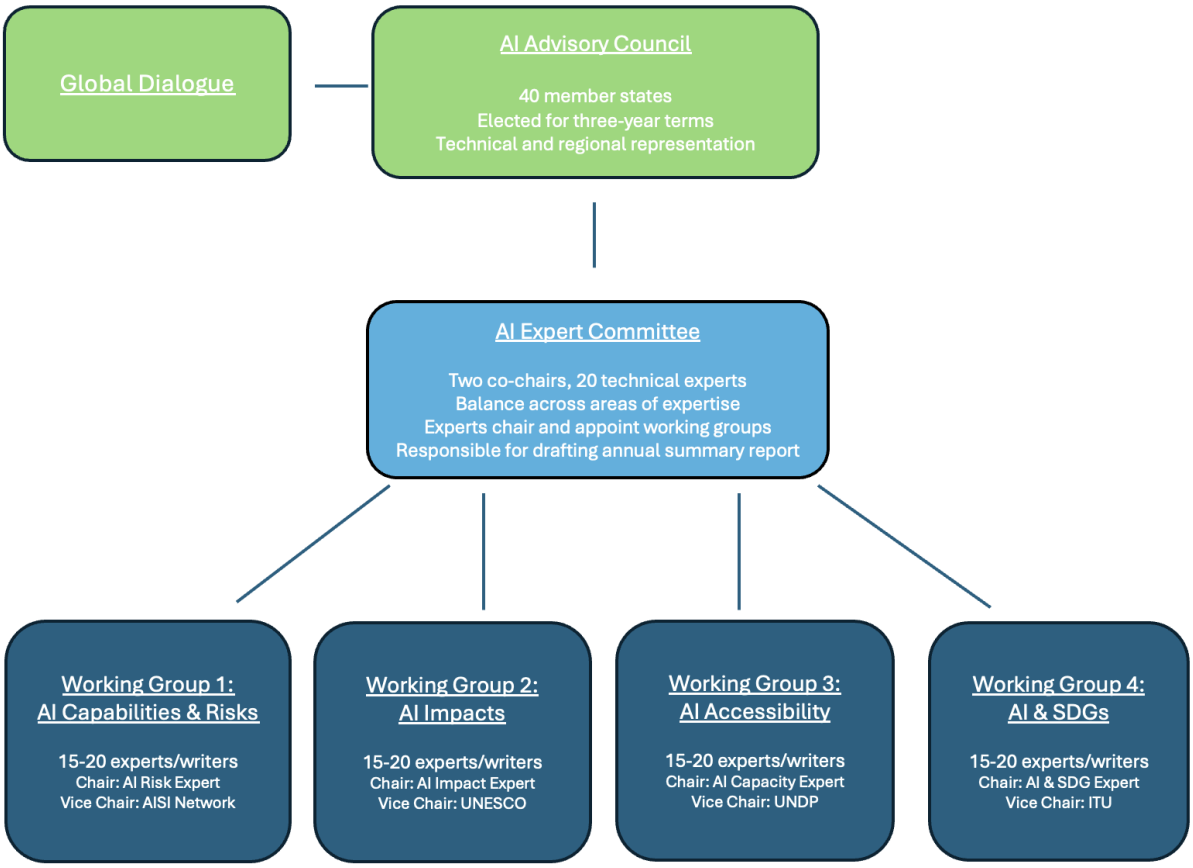
In line with the zero draft's emphasis on evidence-based scientific assessments, we recommend the AI Expert Committee be staffed with scientists, scholars, and technical experts with no direct ties to either industry or government. We support an open nomination process and propose that the UN Secretary-General put together a slate of proposed nominees to be approved by the AI Advisory Council. In recognition of the importance of multi-stakeholder engagement, we recommend that the AI Expert Committee maintain an industry advisory council, which would include representatives from both large developers and smaller firms, including general purpose developers, industry use-case application developers, and AI deployers. This would allow the council to provide up-to-date insights on developing risks and trends. We also recommend a second advisory council to incorporate input from civil society actors. An additional recommendation proposes allowing AI working groups to adopt a more inclusive approach to expertise by permitting the inclusion of subject-matter experts with ties to industry (no more than 3 per working group).

In line with the zero draft's call for the panel to produce "timely and responsive outputs, including one annual report synthesizing key AI research trends" (OP 11), we recommend a three-pronged approach. We suggest each working group produce a semi-annual report and a subsequent six-month update on its assigned area; this approach would allow the AI Expert Committee to begin drafting the annual report based on initial findings. Working groups could also release special ad hoc reports on timely topics of interest. We suggest that both semi-annual and ad hoc reports be released publicly as preliminary drafts upon working group approval and as final reports upon consultation with the AI Expert Committee.

In line with the zero draft's request that the panel "present its findings to the Dialogue," we recommend that the panel release the annual comprehensive report to correspond with the timing of the Global Dialogue. We suggest that the AI Expert Committee draft the annual report, basing the summary document on the working group semi-annual reports and updates.

In an effort to balance scientific independence with political legitimacy, we recommend the AI Advisory Council have the opportunity to review the report and add footnotes of dissenting opinions. To encourage political buy-in, we suggest the Advisory Council also have the ability to revise report language (subject to ⅔ majority approval on each amendment). The Advisory Council would also approve the final report, aiming for consensus but subject to ⅔ majority if necessary. While this voting approach might differ from scientific panels like the Intergovernmental Panel on Climate Change (IPCC), it may nevertheless be appropriate as the AI panel will touch on a broad range of issues beyond the scope of any traditional scientific panel. Moreover, a supramajority voting procedure is in line with procedures at the UN General Assembly, where the most important questions require a ⅔ majority.

# Organizational Structure



**Section I: Panel Mandate**

**In line with objectives laid out in UNGA resolutions A/78/L.49 and A/78/311, the UN Global Digital Compact, and the zero draft, we support a panel designed to provide a comprehensive, objective, and timely assessment of the scientific, technical, and socio-economic opportunities and risks of AI.**

We base this proposed mandate on the zero draft, which recognizes the need for "evidence-based scientific assessments" (OP 2), suggesting the panel's work products should be objective and technocratic rather than political or value-laden. The zero draft also recognizes the need for "timely and adaptive outputs that address rapid advancements in the field" (OP 2), which we note makes the panel's mandate quite different from most independent scientific panels where science moves slowly.

We recommend the mandate include language on the opportunities and risks of AI to align with the policy goals laid out in the two recent UN General Assembly resolutions on AI (A/78/L.49 and A/78/311) and in the UN Global Digital Compact. These documents share three core goals: bridging the AI divide, applying AI technologies to accelerate progress toward the SDGs, and promoting safe, secure, and trustworthy AI systems.

**To realize the goals of bridging the digital divide, applying AI toward the SDGs, and promoting safe and trustworthy AI systems, we suggest expanding the language in zero draft to pre-specify four working groups: AI capabilities and risks, AI impacts, AI accessibility, and AI and sustainable development.** More details on each working group's proposed mandate are provided below.

- *The AI Capabilities and Risks Working Group* could provide a comprehensive scientific assessment of the capabilities and risks of advanced AI systems. A key objective of this group could be to generate an agreed-upon definition of risk and to provide possible thresholds for evaluating capabilities and risk. We recommend a broad definition of risk that would include malicious use and "loss of control" risks, as well as risks related to malfunction. Risk assessment might also include factors like deployment security and context, as well as attempts to learn from risk management practices in other industries and governance contexts. In collecting information on risk, we recommend that

the working group draw on existing expertise and assessments, such as work done by AISIs and on regional bodies like the OECD.

- *The AI Impacts Working Group* could provide a comprehensive analysis of high-level trends, highlighting how AI technologies are reshaping the economy and society. It could provide science-based assessments of technological transformations to local, national, and transnational economies and discuss the ramifications of such changes for society. The mandate could also include discussing the systemic risks of AI technologies, including topics like changes to the labor market, climate impact, and risks to privacy.

- *The AI Accessibility Working Group* could focus on collecting and disseminating information designed to build capacity and upskill workers and governments to engage with AI technologies. The working group could provide updates on the state of AI technology research and the latest capabilities, identify adoption and use trends, and analyze best practices for training workforces and integrating AI technologies into work flows. It could also focus on capacity building and highlight opportunities for domestic innovation and AI engagement.

- *The AI and Sustainable Development Goals (SDGs) Working Group* could engage with concrete real-world applications of AI that offer current and prospective progress for helping governments achieve the SDGs. This working group could focus on collecting information about how a variety of actors, including governments, non-governmental organizations, and international organizations (IOs), are currently using AI to achieve SDG-related policy objectives. Possible topics might include AI for vaccine development, AI for energy management, or AI for optimizing irrigation and land use. The working group could also identify areas of high-value future applications.

Discussion: *The four working groups cover topics that are of interest to a wide range of states. By specifying different lines of research, the working group structure would allow for a more inclusive and representative panel with greater balance across regions, disciplines, and gender. Working groups on AI risk and AI and sustainable development, for example, would require different types of scientific expertise and disciplinary training. The working group approach also helps build epistemic*

*communities in each issue area, which will facilitate AI accessibility and help bridge the digital divide across a range of topics.*

## Section 2: Governance Structure

The zero draft proposes a governance structure that includes an advisory committee of 40 members (elected by the General Assembly but serving in their personal capacity), an expert committee of twenty members appointed by the Secretary General, and working groups (yet to be decided).

**We recommend diverging from the zero-draft structure by creating an advisory council of member states (elected by the General Assembly for three-year terms). We support the proposed expert committee but suggest certain expert positions be predetermined to align with working group roles. Finally, we suggest the resolution text stipulate four working groups.**

Our proposed governance structure and distribution of responsibilities across these bodies are designed to balance the need for scientific independence and timely reporting with political relevance.

### 2.1 AI Advisory Council

**We recommend an AI Advisory Council composed of 40 member states, elected for 3-year terms by the UN General Assembly.** Elections should be designed to balance the goals of regional representation and legitimacy with the need for membership from countries with the most advanced AI systems. One option would be to allocate 3-4 spots per region and to make the remaining spots open for general election. We recommend also establishing a norm of including leading AI countries on the Advisory Council, as excluding such countries would significantly undermine the Panel's credibility and legitimacy.

**We suggest the mandate of the AI Advisory Council be three-fold: approving the slate of experts for the AI Expert Committee, making recommendations on budget or financing (if relevant), and approving the final comprehensive report.** This recommendation follows similar leadership structures and responsibilities in other highly technical issue areas such as civil aviation and nuclear weapons. The

International Civil Aviation Organization (ICAO), for example, elects a 36-member-state governing council that appoints its expert body (the Air Navigation Commission), which oversees various working groups. The International Atomic Energy Agency (IAEA) has a 35-member-state Board of Governors, which approves updates to safety standards, and makes recommendations on budgetary matters.

**In line with decision-making procedures in the IAEA and the UN General Assembly where a ⅔ majority is required for important matters, we recommend the Advisory Council adopt a ⅔ supermajority decision-making procedure for approving the annual comprehensive report.** In line with the zero draft, we suggest the AI Expert Committee produce the annual comprehensive report, and its proposed text would be the default for adoption. We suggest that the Advisory Council have two pathways to alter the report's text: individual member states could add footnotes to the reporting noting dissent or Council members could amend the report's language with ⅔ approval from members. This procedure would provide a pathway for overcoming political roadblocks and is similar to the treaty reservation system, whereby member states can express reservations about individual paragraphs of an international treaty.

Summary:

- Responsibilities:
    - Approval of slate of experts for AI Expert Committee
    - Budgetary approval (if applicable)
    - Approval of comprehensive report

- Decision-making procedures:
    - ⅔ majority vote (consensus preferred but not required)
    - Comprehensive report - any member state can add a footnote of dissent but to change language in the report requires ⅔ majority support to overrule the AI Expert Committee

- Meeting schedule: 3 times per year

*Discussion: The UN Global Digital Compact calls for a balanced, inclusive and risk-based approach that includes full representation of all countries and meaningful*

*participation of stakeholders. A rotating AI Advisory Council that considers both technological advancement and regional representation balances inclusiveness with policy stakes. This is a model that has worked well in other issue areas with cutting-edge technologies such as nuclear weapons and civil aviation. Allowing member states to rotate onto the board provides an opportunity for countries to have meaningful engagement in this policy issue while also accounting for differences in state capacity and political priorities.*

*We recommend limiting the AI Advisory Council's mandate to two core responsibilities: approving the AI Expert Committee and approving the final report. By introducing mechanisms of indirect political influence, the proposal increases the probability that the panel will produce work that countries deem politically relevant. Member state support for a report will increase publicity and the report's impact on policy.*

*Our proposed voting structure for report adoption is different from the voting structure adopted by other scientific panels, such as the Intergovernmental Panel on Climate Change, which relies on consensus. We note, however, that a one-year report-and-approval cycle is much faster than the IPCC or similar panels on biodiversity or global resources. Given the rapid rate of change in this issue area, we believe the priority should be to create a structure that allows political intervention on topics of broad concern but otherwise prioritizes the publication of timely reports.*

2.2 AI Expert Committee

The zero draft proposes an AI Expert Committee composed of 20 experts appointed "with due consideration of the candidates' outstanding expertise, diverse multidisciplinary representation and geographical and gender balance" (OP 4b). The nomination process for this body is open to all relevant stakeholders and would proceed through the Secretary-General's office (OP 4).

**Building on this proposal, we recommend that the UN Secretary-General's Office for Digital and Emerging Technologies (ODET) stipulate criteria for identifying experts, including closing the application process to individuals with ongoing**

**direct or significant indirect ties to industry or government.**[5] ODET could process applications and present a recommended slate of candidates to the AI Advisory Council. In coordination with ODET, the AI Advisory Council could then approve the final set of experts to serve on the AI Expert Committee for two-year terms, with the possibility of renewal.

The zero draft suggests the Secretary-General appoint two chairs: one from the developing country and one from a developed country (OP 8). It does not offer any additional guidance on the appointment of experts as it does not discuss specific working groups or a more detailed mandate.

**We recommend revising the draft resolution to stipulate that the AI Expert Committee include at least some positions tied directly to specific issue areas and at least some positions designed to be general in nature.** Specific issue-area experts could include the chairs of each working group (e.g., AI Risk Expert, AI Impact Expert, AI Accessibility Expert, and AI and SDG Expert). Additional overlapping areas of expertise could include topics like AI innovation, AI and the economy, and AI safety. Generalists might be tasked with drafting the comprehensive report, which would require integrating information from all four working groups. Defining such tasks prior to appointment could be important given the fact that experts will be serving in their personal capacities and will have additional professional responsibilities related to their full-time employment.

The zero draft suggests that the AI Expert Committee "shall oversee all outputs from the Panel, define research areas; establish working groups to prepare assessments and technical briefs according to research priorities; appoint Chairs to each working group from among the members of the Expert Committee; assign members of the Expert and Advisory Committees to each working group; and invite external experts to engage with the working groups as necessary" (OP 9).

**Building on the zero draft, we recommend that the AI Expert Committee have two primary responsibilities: overseeing working groups and drafting the comprehensive report.** The Expert Committee could identify experts to serve on the

[5] This process would exclude individuals who are currently employed by governments or the private sector, or who receive significant financial compensation from such outlets through working as consultants. It would not exclude former employees.

four working groups, define the general mandate for each working group, and recommend possible topics for ad hoc reports. Individual experts from the Committee could also be responsible for chairing respective working groups.

**We recommend the AI Expert Committee be responsible for producing an annual comprehensive report that integrates semi-annual reports from all four working groups.** The Committee could begin working on the annual report with the publication of the semi–annual working group reports; information from subsequent working group reports could be integrated to update the comprehensive report. To facilitate the Committee's ability to produce a single report, we recommend it adopt a norm of procedures whereby the comprehensive report includes all credible positions on an issue. This will also provide valuable information about underlying uncertainty.

The zero draft does not specify a precise role for industry or civil society in the panel process, except to say that experts must fully disclose "financial, professional, and personal interests that may affect impartiality or independence" (OP 7). Paragraph 6 of the UN Global Digital Compact lays out a more precise vision of broad engagement: "As Governments, we will work in collaboration and partnership with the private sector, civil society, international organizations, the technical and academic communities and all other stakeholders, within their respective roles and responsibilities, to realize the digital future we seek." In discussing AI specifically, the Global Digital Compact notes the need for "meaningful participation of all stakeholders" (OP 50).

**Building on the vision suggested by the UN Global Digital Compact, we suggest that the Committee's third function be to maintain regular dialogues with industry and civil society via respective advisory councils.** The industry advisory council could be staffed with individuals representing both large developers and smaller firms, as well as general purpose developers and industry use-case application developers. Additional consideration for geographic diversity (as feasible) and gender might also be important. These forums could meet at least three times per year and provide an opportunity for the Committee to receive input on its working draft of the comprehensive report. Industry and civil society dialogue could begin early in the drafting process so as to encourage information exchange between all parties. We also suggest that both councils have the opportunity to add footnotes to the

comprehensive report, if such actions are approved unanimously by council members.

Summary:

- Responsibilities:
    - Appointing experts for individual working groups
    - Leading working groups
    - Identification of possible topics for ad hoc reports
    - Integrating information from across working groups into comprehensive draft report
    - Dialogue with industry and civil society

- Decision-making procedures:
    - Norm of including all credible positions
    - Consensus for comprehensive report
    - ⅔ majority for approval of working group reports

- Meeting schedule: Monthly

*Discussion: To establish scientific credibility, we recommend that the primary body responsible for drafting the comprehensive report and assembling working groups be composed of independent experts with no ties to industry or government. This structure allows governments some influence over the selection of experts but builds safeguards to ensure the panel is composed of scientific experts who are acting independently from member state or industry influence. Allowing the Secretary-General's office to designate experts for specific positions leverages ODET expertise and success with the recent High-Level Advisory Board on AI. It also ensures that the individual responsible for leading each working group is qualified and credible as a scientific expert. Finally, assigning specific roles provides advance information to prospective members of the AI Expert Committee on the depth of time and engagement required for participation.*

*We recommend that the AI Expert Committee maintain regular dialogue with industry and civil society to allow the Committee increased insight into new trends and technological developments. Industry professionals have unparalleled information on AI capabilities and risks, and may be more inclined to share this information if the*

*process is structured at regular intervals. Civil society actors play important monitoring roles in highlighting the impact of AI technologies on marginalized populations. Such processes of multi-stakeholder engagement are common across other institutional contexts. The World Health Organization, for example, has collaborated with national authorities, civil society, and academia to provide detailed guidance on health and sustainable development. The Intergovernmental Panel on Climate Change routinely convenes multi-stakeholder dialogues that solicit input from indigenous peoples, scientists, NGOs, businesses, and others. Even highly technical issue areas like financial regulation have regular forums for engagement with the private sector and civil society.*

2.3 Technical Working Groups

The zero draft recognizes the importance of working groups for preparing "assessments and technical briefs according to research priorities" (OP 9); however, it leaves it up to the AI Expert Committee to define relevant research areas and establish these working groups.

**We recommend building out the zero draft's language on working groups to pre-specify the establishment of four groups: AI Capabilities and Risks, AI Impacts, AI Accessibility, and AI and SDGs. We further recommend that each working group be chaired by an associated expert from the AI Expert Committee and have a vice-chair from an organization with substantive background and institutional resources.** For example, the International Telecommunication Union (ITU) might be well-placed to assist with the AI and SDGs Working Group due to the ITU's experience running the AI for Good program and global summit series. UNESCO might support the AI Impacts Working Group and UNDP could assist with the AI Accessibility Working Group.

With respect to AI Capabilities and Risks, we suggest a co-chair directly tied to the AISI network. The UK AISI, for example, has significant experience organizing the international AI safety report, and therefore could be well-placed to engage with this working group on behalf of the broader AISI community. Appointing the UK AISI as the vice chair of the working group, on behalf of the broader network of AISIs, would ensure harmony between the two processes and would facilitate the integration of the two reports, if so desired. This would also provide opportunities for the working

group publication schedule to align with the AI summit series, which has to-date included three successful high-profile meetings (Bletchley Park in Nov 2023, Seoul in May 2024, and Paris in February 2025) and a planned meeting in India.

The zero draft proposes that the AI Expert Committee "assign members of the Expert and Advisory Committees to each working group" and "invite external experts to engage with the working groups as necessary" (OP 9).

**Here, we diverge from the zero draft and suggest instead that the AI Expert Committee appoint 15-20 outside experts with specific expertise in a given area to serve on individual working groups.** Because we propose staffing the AI Advisory Council with representatives of member states, we recommend that Advisory Council members not play a role in writing the report. It may also be difficult for members of the AI Expert Committee to staff all working groups, write working group and ad hoc reports, and produce an annual comprehensive report in a single year. With this in mind, we suggest instead that the AI Expert Committee have the responsibility of appointing additional subject-matter experts to serve for two-year terms on individual working groups. We recommend that the Committee aim for broad and inclusive representation in terms of gender, region, discipline across the working groups as a whole rather than within each individual working group.

In recognition of the need for the panel to have access to the most up-to-date information, we suggest that the Committee consider allowing the appointment of international experts with ties to industry and/or civil society if such individuals also have substantive expertise and credentials to serve on individual working groups. Such appointments should be limited to no more than 3 individuals per working group.

Appointed experts could be responsible for collecting, analyzing, and summarizing scientific information related to the specific mandate of the respective working group. We recommend working groups adopt a norm of representing all credible positions within a report to facilitate the writing process. Each working group could produce semi-annual reports, which would be updates on the latest developments and scientific findings related to the issue area.

**We recommend that working groups also be responsible for producing ad hoc reports on special topics of interest.** While the AI Advisory Council and the AI Expert Committee could request such reports, final decisions about topics of interest would be left up to the chair of the working group. Ad hoc reports would be designed to be timely and targeted assessments produced in a shorter time span than the typical six-month cycle.

Summary:

- Responsibilities:
    - Collecting, analyzing, and summarizing scientific information on topics relevant to the working group's mandate, drawing in particular from the work of other multilateral and regional organizations with expertise in this area
    - Drafting related semi-annual reports
    - Identifying priority issues for ad hoc reports
    - Collecting, analyzing, and summarizing scientific information for ad hoc reports and drafting ad hoc reports in timely manner

- Decision-making procedures:
    - Norm of representing all credible positions in reports
    - Consensus

- Meeting schedule: Bi-weekly

*Discussion: Technical working groups are an essential part of all scientific panels, and will be particularly important for a topic as wide-ranging and quickly changing as* AI. *The working group structure allows for the appointment of experts with specialized knowledge, who can work more efficiently to process relevant information.* As *experts in their assigned issue areas, working groups are also best positioned to identify important trends that would necessitate an ad hoc report.*

*The proposed leadership structure for working groups is designed to facilitate the flow of information between the* AI *Expert Committee and working groups, and also to facilitate the building of institutional expertise. While experts are appointed for renewable two-year terms, the demands of the workload may lead some experts to opt-*

*out after a short period. For this reason, it is critical that each working group has the support of an institutional actor that can develop expertise and learn from the previous report writing experience. The UK AISI, for example, facilitated the writing of the first International AI Safety Report, which was published in January 2025 in advance of the Paris AI Action Summit. Because of its role in this process, the UK AISI now has significant expertise to draw on for the next safety assessment. Building out similar institutional support across the four issue areas will allow for learning and efficiencies to develop over time.*