RESPONSIBLE AI WORKING GROUP REPORT
GPAI – MONTREAL SUMMIT 2020

EXECUTIVE SUMMARY

i. Introducing the Working Group

RAI has 30 members. Its international experts come from various fields, something which favors robust discussions and the emergence of diverse viewpoints. More precisely, 14 members of RAI come from the technical world (e.g. machine learning, information technologies), whereas 16 come from the social and human sciences sector and fields like communications, management, psychology, philosophy or international development.

Most members (63%) come from the academic sector, but 17% work in the private sector, 13% for non-profits and 7% in the Public Sector. We would like to achieve a better balance in coming months and years as we believe that the collaboration of *all* stakeholders will be necessary to ensure AI is produced and used in a responsible manner.

ii. Mandate

RAI's mandate aligns closely with that vision and GPAI's overall mission, that is, RAI strives to foster and contribute to the responsible development, use and governance of human-centered AI systems, in congruence with the UN Sustainable Development Goals (SDGs).

It is worth noting that RAI, as all other GPAI Working Groups, does not operate in silo within GPAI. Indeed, it intends to collaborate with other working groups whenever indicated. For instance, RAI will interface with the Data Governance Working Group when their respective projects share common dimensions. We also recognize the special role that partnerships - in particular between those with domain expertise combined with our own AI and technical expertise - will play in achieving our mission. Partnerships will help us harness the funding, knowledge, and implementation at the scale necessary to meet our ambitions.

iii. Work process

RAI has met regularly to develop its recommendations for the Summit, and, following a competitive tender, partnered with The Future Society (TFS) to assist in their preparation.

iv. Preliminary recommendations and outputs for the Summit

To inform where RAI could maximise its impact on the responsible development and use of AI systems in congruence with the UN Sustainable Development Goals, TFS's report:

- catalogues projects led by various stakeholders to promote the responsible research and development of beneficial AI systems and applications;
- analyzes promising initiatives that had great potential to contribute to the development and use of beneficial AI systems and applications that could benefit from international and cross-sectoral collaboration;

- and recommends new initiatives and how they could, in practice, be implemented and contribute to promote the responsible development, use and governance of human-centered AI systems.

Of the recommendations TFS has produced, RAI considers three to be especially pertinent:

1.   AI has wide-ranging applicability and hence has the potential to influence many of the most pressing issues humanity is facing; it can be a force for good to mitigate climate change or predict the next pandemic, and it can also exacerbate global challenges as evidenced by the rise of misinformation. The breadth of potential applications of Responsible AI creates a prioritization challenge. **Thus, RAI should create focused committees to address identified pressing issues.**

2.   To build an ecosystem with the ability to support and stimulate change, there is a need for governance tools and frameworks that promote transparency and alter incentives and behaviors throughout society to help the adoption of Responsible AI practices. There is also a need for systematic collaboration and cooperation across the ecosystem as well as a mechanism to connect cross-cutting initiatives on the domain level. Finally, for governments to implement these tools and frameworks at scale, there is a need to build capacity amongst policymakers as well as feedback loops between governments and other actors in the ecosystem. **Thus, RAI should create a focused committee on governance issues and governance means.**

3.   Many initiatives in the Responsible AI ecosystem have struggled to collect representative input to inform their activities. This lack of inclusiveness points to a lack of capacity by initiatives, stakeholders and governments to involve a wider group in the technological transition and, hence, to co-shape innovative solutions for addressing the opportunities and the risks. Ultimately, this lack of inclusiveness risks undermining the effectiveness and credibility of many Responsible AI initiatives as well as their ability to scale. **Thus, RAI should develop and disseminate good Diversity & Inclusion practices.**

(v) Priorities for H1 2021
In the next phase of work, RAI will focus on areas where current market structure and current government policies are insufficient to achieve SDGs or other key objectives. To achieve this, RAI will create as many as five internal committees:

1. **The Committee on Drug Discovery & Open Science** (linked to SDG 3: Good health and well-being)
2. **The Committee on Climate Change** (SDG 13: Climate action)
3. **The Committee on AI & Education** (SDG 4: Quality education)
4. **The Committee on Governance & Transparency of Social Media** (SDG 16: Peace, Justice and Strong Institutions)
5. **A Transversal Committee on Issues and Means of Governance** (that could work on the certification, assessment and audit mechanisms used to evaluate AI systems)

(vi) Longer term vision

Over the longer term, RAI would like to address Diversity & Inclusion, continue its work on governance in coordination with the OECD, work with international partners on Global South representation and challenges, and take on other SDGs beyond those that the five committees will start addressing shortly.