



Centre d'expertise
international de Montréal
en intelligence artificielle



GPAI / THE GLOBAL PARTNERSHIP
ON ARTIFICIAL INTELLIGENCE

TECHNOLOGY TO SUPPORT DATA AVAILABILITY FOR RESPONSIBLE AI

*Demonstrating practical use of Privacy Enhancing and adjacent technologies for
well-governed data access for AI*
Request for Proposal (01/03/2022)

This Request for Proposal (RFP) provides guidance on deliverables that the Global Partnership on AI's ("GPAI") Data Governance Working Group is expected to deliver by the end of July 2022. Formal requirements for tender responses are specified under Annex 1.

1. Context

This project, from the GPAI data governance working group, will demonstrate how technologies used to enhance privacy, justice or rights can help increase the availability / usability of AI systems. The project will demonstrate technical means to safely develop and use data sets while preserving privacy, sovereignty and IP rights for public good projects, specifically in the context of climate action, better health, and/or the future of work.

The project as a whole is expected to last until the end of 2023, by which time it is expected to deliver a practical demonstration of such technologies at work, as well as practical guidance for data scientists, an outreach plan to raise awareness of, and confidence in, such technologies, and guidelines to feed into international standards and/or minimum technology specifications. The project is intended to help decision-makers in making appropriate technology choices to meet their needs in this space.

This RFP relates to the initial phase of activity for the project (Scoping, Design and Planning; April-July 2022), during which the GPAI aims to identify use case(s) for candidate demonstration(s) and produce an implementation plan.

The project is being developed in collaboration between GPAI and the Singapore government, a founding member of GPAI. Singapore's project engagement is being led by the [Infocomm Media Development Authority](#) ("IMDA") - the selected delivery partner will be expected to work with the IMDA on an implementation plan that could be adopted by Singapore and 1-2 other countries in cross-border collaboration. The 1-2 countries may be identified through engagement with complementary international initiatives (e.g. those in the UK and the US).

About GPAI

The [Global Partnership on AI](#) ("GPAI") has been established with a mission to *"support and guide the responsible adoption of AI that is grounded in human rights, inclusion, diversity, innovation, economic growth, and societal benefit, while seeking to address the UN*

Sustainable Development Goals".¹ Driven by a multistakeholder approach, it is supported in its mission by four Working Groups made up of leading international experts, and 19 member governments and counting.

The Data Governance Working Group, co-chaired by Jeni Tennison and Maja Bogataj Jančič, supports GPAI's mission with its mandate to *"collate evidence, shape research, undertake applied AI projects and provide expertise on data governance, to promote data for AI² being collected, used, shared, archived and deleted in ways that are consistent with human rights, inclusion, diversity, innovation, economic growth, and societal benefit, while seeking to address the UN Sustainable Development Goals."*

The Data Governance Working Group is supported by a Centre of Expertise (one of two supporting GPAI as a whole), [the CEIMIA](#). The Centre is led by an Executive Director, Sophie Fallaha.

About the project

Why this topic?

Using AI for public good and/or to address multi-stakeholder problems (e.g. health and pandemic response, climate change, the future of work) requires access to data (for training, testing and evaluation) from multiple sources - within and between organisations, and across geographic or jurisdictional boundaries. Individuals, researchers, corporations and governments do not necessarily have the incentives to make such data available, due to concerns around:

1. Privacy: Identifiability or potential identifiability of individuals represented in data. While not typically considered in scope for 'privacy', there are similar concerns around the identifiability of organisations or of groups of individuals as well.
2. Sovereignty: Respecting company, government, community, individual and Indigenous rights to govern and control use of data
3. IP protection: Recognition of and respect for any property rights inherent in data
4. Data security: Ensuring safe handling and housing of data throughout the data life cycle
5. Data travel and localisation: Legal requirements, size of data sets, and other considerations may limit the movement of data from their originating location

These requirements are informed by operational needs and objectives, laws, ethical principles, and values of users and data subjects. Technology is likely to be an essential part of any solution supporting broader AI development, with the objective of enabling desired or approved uses of data while preventing misuse. Within the current GPAI context, projects focused on climate action, better health and the future of work can all benefit from the use of such technologies to improve data availability for training AI models. To be useful, selected technologies must be trusted, which requires integration testing and transparency.

What technologies might be relevant in this context?

There are a range of inter-connected technologies, that could help meet this challenge

¹ Global Partnership on AI Terms of Reference

² The Mandate draws upon the definitions set out within the [OECD Recommendation on Artificial Intelligence](#) for this purpose

outlined above, including but not limited to the following:

- Differential privacy: techniques that make it difficult to identify, by looking at a data set, if the data on a single individual or any given record was included in the input data set;
- Federated learning/ secure multi-party computing: techniques to train an AI algorithm across a decentralised set of data samples, without having to exchange those samples. This can help with privacy, sovereignty, IP and data travel concerns;
- Homomorphic encryption: a form of cryptography that enables computation on data encrypted using an algorithm, so that the generated encrypted result exactly matches the result of operations that would have been performed on unencrypted text;
- Data wallets or equivalent: tools that enable people or organisations with controlling authority to be able to manage access to data under their control (e.g., to include individual management of data, organisational management of IP or sovereign control of data about indigenous people)
- Distributed consent management: tools that enable organisations to reliably know whether they have permission to use a particular piece of data, for a particular purpose, without having to directly approach the data subject or steward.

The intended effect of the use of these technologies extends beyond a focus on privacy to the broader and emerging concept of [“structured transparency”](#). This concept is defined as systems that conform *“with a set of predetermined standards about who should be able to know what, when they should be able to know it, and what they should [be] able to do with this knowledge.”*

Why this project?

There has been significant recent work to bring together a view of the current ‘state of the art’ in these technologies, particularly Privacy Enhancing Technologies (PETs), and the ways in which they can potentially support real-life use cases. For example, on PETs, the Centre for Data Ethics and Innovation (CDEI) in the UK has published a [good introductory resource](#) with a good set of use cases and [example projects](#). There has also been some limited real-life application of these technologies - for example, progress with the UN's PET Lab for official statistical organisations.

However, actual integration, deployment, and adoption of such technologies still remains relatively limited. Where they are being used, they tend to be ad-hoc and piecemeal in nature. There are few publicly available details on real-life implementation of these technologies in multi-organisation settings. As a result, there is a lack of shared context for how all the parts work together in an effective implementation. In addition, there is a need for a better understanding of the legal issues involved in multi-country, cross-jurisdiction systems, and mechanisms to address these in a scalable manner.

We are also aware from the Working Group’s work on [advancing research and practice in data justice](#) of the ways in which these technologies can be used to perpetuate, as well as challenge, inequities. In accordance with [GPAI’s mission](#), we are particularly interested in examining and advancing the use of PETs and adjacent technologies in ways that advance human rights, diversity and inclusion.

The project's outputs should therefore be of interest to at least three sets of stakeholders: (a) commercial and public sector organisations or other groups that are struggling to overcome the challenges around data availability / usability / sharing when building AI applications; (b) government / inter-governmental agencies tasked with encouraging responsible innovation in AI (e.g. to encourage competition, meet SDGs); and (c) technology vendors (start-ups and established) that have research/ IP/ early products in this space but are struggling to see a path to commercialisation.

2. Objectives

The overall objective of this project (**across all its phases**) is to demonstrate how a set of technologies including PETs can be deployed at scale in a real-world "AI for good" use case to make more data available for development of AI systems, particularly when involving commercial and governmental stakeholders in cross border use cases. This will be achieved through four deliverables

- i. A practical demonstration of how such technologies can help improve data availability for AI use cases beneficial to humanity, probably in collaboration with GPAI initiatives around climate action, better health and/or the future of work.
- ii. Practical guidance to analysts/ data scientists on how to work with such technologies - for example how to ensure data quality, or how to address the legal issues involved in cross-country systems.
- iii. An outreach plan that yields greater awareness of, and (where appropriate) confidence in, technology solutions to address privacy/ IP/ sovereignty concerns - among data owners/ custodians, AI developers and adopters, and regulators
- iv. Guidelines on further development and adoption of such technologies (which could translate into international standards); OR Minimum specifications for technologies that are deemed to address structured transparency.

The objective of the first phase (Scoping, Design and Planning) is to identify use case(s) for candidate demonstration(s) and produce an implementation plan to deliver the candidate demonstration.

3. Outputs expected from the first phase

The following outputs are expected from the first phase:

- One or more AI system use cases identified for technology demonstration in future phases, taking into account
 - The extent to which the use case supports one of the three "AI for good" topics of climate action, better health and the future of work
 - The extent to which data availability, data sovereignty/ localisation considerations and/or quality acts as a barrier to the use case, and therefore makes it suitable to test/ demonstrate the value of PETs and adjacent technologies
 - Feasibility of completing the demonstration project within 15-18 months, taking into account the likely technological, data and legal/IP challenges

- Availability and commitment of a partner GPAI expert group (pandemic response and drug discovery, climate action, future of work) or a potential partner initiative to provide the necessary engagement for the demonstration
- Detailed implementation plan for use case(s) including timeline, costs, stakeholders and requirements of resources, staffing, legal & compliance, etc. The delivery partner will be expected to collaborate with the IMDA and the Project Steering Group/CEIMIA so that the implementation plan could be adopted by Singapore and 1-2 other countries in cross-border collaboration

These will be presented in a written report and presentation to the **Project Steering Group** and the **Data Governance Working Group**.

For the avoidance of doubt, the first phase does **not** include implementation of the demonstration project, the development of practical guidance/ standards or outreach efforts.

4. Resource

The Centre of Expertise will provide programme funding (of up to 355760 Canadian dollars in total for all outputs specified in this ToR (Phase 1 only) to appoint expert consultancy/ system integrator support for the development and production of these outputs.

The project will be supported by a **Project Steering Group**, co-led by Shameek Kundu (TruEra) and Kim McGrail (University of British Columbia). It will include representatives from GPAI's Working Groups including Data Governance, Pandemic Response, the Future Work, and GPAI's committee on climate and biodiversity. It will also include representation from the Singapore government, the IMDA, who are collaborating with GPAI on this project.

From this group, the Centre will establish an **Evaluation Panel** for responses to this Request for Proposals (absolving those with a potential conflict of interest at this stage).

5. Governance

The **Consultancy/ System Integrator Partner** will be responsible for the development and production of outputs specified in the Sections above, and be responsible for the final product satisfying the standards and expectations of the Working Group. Once appointed, the Partner will:

- engage GPAI workgroups and relevant external or international stakeholders (e.g. UN PETs Lab, UK CDEI) to identify high-impact opportunities;
- translate identified opportunities into specific use cases – commercial/social outcomes, participants, data items, AI system, suitable technology type, etc;
- finalise use case(s) based on feasibility study of social and/or economic value, ability & commitment of participants, relevance of PET or adjacent technologies in use case, engineering & infrastructure requirements, cost and time (<18 months) of implementation, legal constraints, etc; and
- develop an implementation plan for use case(s) including timeline, costs, stakeholders and requirements of resources, staffing, legal & compliance, etc.

In developing content for the report, the Partner is expected to work closely with the **Project Steering Group** members, who will meet with the Partner on a regular basis. The Partner is also expected to map and, where appropriate, coordinate with relevant international

initiatives. This is important both in the context of expected governmental involvement in the project itself (Singapore, US and UK) and other (non-GPAI) existing and planned initiatives in related areas.

The **Centre** (CEIMIA) will run the administrative process for the selection of the Partner, provide advice to the Project Steering Group, and provide contractual and day-to-day engagement with the Partner for the duration of the project. It will be responsible for graphic design (in line with GPAI's visual identity) and translation of written outputs.

The **Evaluation Panel** (made up of members of the Project Steering Group) will lead on the scoring of bids.

The **Data Governance Working Group** will review outputs at key milestones (e.g drafts) and be invited to workshops organised and run by the **Partner**. This review may be further supported by selected **external reviewers** who will provide additional peer review and quality assurance of the final output.

6. Methodology

The **Centre** (CEIMIA) will lead on organising and running the process to select a **consultancy/ system integrator partner** and will apply the evaluation criteria under Annex 1. A comprehensive, appropriate and robust approach to delivering the full scope of the work is one of the evaluation criteria for the Partner (see methodology under Annex 1).

7. Timeline

This phase of the project is expected to be completed by the end of July 2022.

Key dates are set out in the table below.

Note on evaluation: the Data Governance Working Group as a whole currently meets on a monthly basis (in the middle of the month) for up to two hours. The Working Group Co-Chairs will meet with Project Co-Leads on a fortnightly basis. The Project Steering Group will be available to meet on a weekly basis. As part of the proposed methodology, the Evaluation Panel will expect the proposal to account for review and feedback from both the Project Steering Group and the Data Governance Working Group.

Milestone	Date
Request for Proposals published	1 March 2022
Deadline for Proposals	Midnight (Anywhere on Earth), 22 March 2022
Evaluation and selection of proposal	28 March 2022
Notification of tenderers	29 March 2022
Contract signed	4 April 2022

Start of consultancy engagement (including mobilisation)	11 April 2022
Interviews and workshops with GPAI workgroups completed	To be proposed by successful bidder
Feasibility study completed and initial proposal on candidate use case(s) ready for review	To be proposed by successful bidder
Draft implementation plan completed	To be proposed by successful bidder
Candidate use case(s) and implementation plan finalised	22 July 2022
Submission of final outputs	29 July 2022

Annex 1: Tender response and evaluation criteria

By midnight (anywhere on Earth) on 22 March 2022, interested parties should submit a costed proposal (in English) to sophie.fallaha@ceimia.org and edward.teather@ceimia.org , which includes:

- the project title “**Demonstrating practical use of Privacy Enhancing and adjacent technologies to overcome data barriers to “AI for public good”**” in the email subject line
- a proposal that meets the requirements of the evaluation criteria specified below

If you have any questions about the tender, please contact sophie.fallaha@ceimia.org (Executive Director, CEIMIA) and edward.teather@ceimia.org (seconded to CEIMIA) no later than the 10 March 2022. CEIMIA reserves the right to make both anonymised questions and answers public or shared with other organisations having stated their interest (*please see Annex 2 - questions from interested parties*).

The proposal must come from a single organisation that will be contractually accountable and responsible for the project outputs. However, such an organisation may choose to include up one or more partners as part of the proposal (e.g., a system integrator or consulting firm may choose to partner with specialist PET vendors or privacy-focused legal firms). In such instances, the proposal will be evaluated on the combined credentials of the lead and partner organisations.

Evaluation criteria

The evaluation panel will review proposals as below:

Criteria	Weight	Page limit	Guidance
Skills, expertise and experience	40%	5 pages (plus CVs in appendix)	<p>Please demonstrate the relevant skills and expertise of your organisation and staff, to undertake this work.</p> <p>To evidence this, please provide summary CV details (relevant experience and qualifications) for the people who will be responsible for delivering this work.</p> <p>Please include an organogram showing a summary of roles and responsibilities, and the amount of each person’s time to be dedicated to this project.</p> <p>Required: Document examples of previous projects in commercial, public or social sectors that - involved the design, planning and implementation of systems projects involving the use of the technologies relevant to the objectives of this project (<i>overcoming barriers to data usability and availability for “AI for good” projects, including those around privacy, IP rights and data localisation/ sovereignty - safely and responsibly</i>)</p>

			<ul style="list-style-type: none"> - involved one or more of climate change, public health and future of work - demonstrate a history of team members' ability to deliver on the scope of this request. - demonstrate a good global understanding of the AI & data market and legal / compliance requirements for AI & data technical projects, and understand both of these within the broad international context and cross-border applications. <p>Describe the team's ability to work together to collectively produce meaningful and impactful outputs, particularly where team members will be spread across locations and/or organisations</p>
Ability to include diverse perspectives	10%	2 pages	<p>Please describe how the team will be able to bring in diverse perspectives, including those from the Global South, in the selection of candidate use case(s) and the development of the implementation plan</p> <p>Please also outline how your team will bring in adequate understanding of the legal/ compliance requirements and the data ecosystem in different jurisdictions, including Singapore</p>
Proposed methodology (including delivery plan plus commitment to work collaboratively and in the open)	10%	2 pages	<p>Please provide details of your Delivery Plan and methodology for the project, from contract commencement to contract completion.</p> <p>The Delivery Plan should:</p> <ul style="list-style-type: none"> - Include a proposed methodology for how you will carry out the work including key stages, tasks and activities - Demonstrate how you will work collaboratively and in the open both with GPAL's Data Governance Working Group of Experts and the wider external community - Demonstrate a clear and appropriate methodology for delivering the outputs and scope of the project, with clear rationale - Provide demonstrable evidence that your approach will deliver the required outputs at the promised project scope.
Project management, timeline and milestones	20%	2 pages	<p>Please provide details of your methodology for management of the project, from contract commencement to contract completion.</p>

			<p>Please provide project management proposals (including management structures to be established, delivery timetable & identification and management of risks that may arise and strategy for how to overcome these).</p> <p>Please provide:</p> <ul style="list-style-type: none"> - Project Plan setting out key milestones and dependencies; - How you will meet the timescales required, and ensure the timings and contract price are not exceeded; and - Key risks you feel may arise and how you will overcome these. <p>Your response should provide confidence that:</p> <ul style="list-style-type: none"> • your approach and ability to deliver the scope of the project is appropriate; • the project plan is comprehensive, realistic and achievable; • the programme will be managed effectively and with flexibility, and that the management structure is robust; • you will meet the timescales required, and ensure the timings and contract price are not exceeded; • any key risks are identified, considered and appropriate mitigation strategies are proposed; • that key roles/responsibilities have been identified and appropriate identified individuals are assigned; and • you have demonstrated sufficient capacity within the tenderers organisation to deliver the requirements.
Pricing	20%	1 page	<p>Please include a costing breakdown of day rates and time allocated by project team members.</p> <p>The maximum marks available for this part of the proposal will be 20% and will be awarded to the Tenderer submitting the lowest price. The remaining Tenderers will receive marks on a pro-rata basis from the lowest to the highest price.</p> <p>The calculation used is the following:</p>

			<p>Score = $\frac{\text{Lowest Tender Price}}{\text{Tender Price}} \times 20$ (Maximum available marks)</p> <p>However, if a proposal is judged as having serious concerns in any of the three preceding categories, it will be excluded from the 'lowest tender price' consideration</p>
--	--	--	--

Each part of the criteria will be evaluated and marked on a scale of 0-4 where:

- 0 – Serious concerns:** e.g. does not meet requirements, and/or raises serious concerns
- 1 – Minor concerns:** e.g. meets some requirements but with gaps and/or some minor concerns
- 2 – Adequate confidence:** e.g. meets most/all requirements, but lacks sufficient detail or evidence in some areas
- 3 – Good confidence:** e.g. meets all requirements and provides a detailed response but lacks evidence in minor areas
- 4 – Excellent confidence:** e.g. meets all requirements, provides a detailed response and evidence which demonstrates a particularly strong understanding of the requirements

Your score will be determined by the marks awarded for each question (out of 4), in accordance with the applicable weighting.

For example, if the weighting for a question is 10%, a mark of 4 for that question would lead to a score of 10%. A mark of 3 would lead to a score of 7.5%, a mark of 2 would lead to a score of 5%, a mark of 1 would lead to a score of 2.5%, and mark of 0 would lead to a score of 0%.

Annex 2: Questions from interested parties

As noted under Annex 1, the CEIMIA reserves the right to make both anonymised questions and answers public. A log is kept as below:

- A number of interested parties have contacted the CEIMIA to request an extension to the deadline. The deadline has been extended by one week to 22 March, by Midnight "Anywhere on Earth"
- An interested party contacted the CEIMIA to confirm that CEIMIA would be the contracting entity. To confirm: CEIMIA is the contracting entity, it is a Canadian organisation based in Quebec, and subject to Canadian law.
- An interested party contacted the CEIMIA to enquire whether the Delivery Partner would need to find the entities that would be willing to implement the recommended use case. We would expect the delivery partner to work with the Singapore government - specifically the IMDA - on this, and include an assessment of viable partners who could implement the solution as part of its wider feasibility assessment.