



Request for Proposal (April 6, 2023) GPAI - Data Governance Working Group

From co-generated data to generative AI - new rights and governance models in digital ecosystems (CO-GEN)

**** EXTENDED DEADLINE:** To better facilitate and encourage a competitive process, CEIMIA has elected to extend this RFP submission deadline from its original date of April 30, 2023 (midnight, anywhere on Earth) **to a revised date of May 7, 2023 (midnight, anywhere on Earth)**. Please refer to the 'Milestones' table on pages 7 and 8, and take note of impacted changes to the project's timeline. All changes have been highlighted in blue text for ease of visibility. **

This Request for Proposal (RFP) provides guidance on deliverables that the Data Governance Working Group of the Global Partnership on AI ("GPAI") is expected to deliver by the end of March 2024. Formal requirements for tender responses are specified under Annex 1.

1. Context

The term 'co-generated data' (originally coined by the American Law Institute, ALI, and the European Law Institute, ELI)¹ refers to situations where data is generated by a number of actors, often with different roles, as is common in the digital economy. The co-generation of data often gives rise to concrete rights and obligations for the parties involved, such as access and portability rights, which are increasingly recognised. The question of who has rights to data, on what basis, and what these rights might mean is a fundamental one, and there is a need for transnational coordination and convergence to support data access and sharing for AI. Data rights are also granted to parties that have not contributed to the generation of the relevant data, i.e. third parties, often motivated by public interest and/or within open data frameworks. The latter are not part of the concept of rights in co-generated data, but we are seeing the boundaries blur recently, for example with some data portability rights, so the project will also address third party data rights.

Rights in co-generated data, as currently recognised, may be just one aspect of a much larger phenomenon, which is the need to introduce new types of ownership and enable co-determination in highly complex and interconnected digital ecosystems. By its very nature, AI is often based on the contributions of many parties, sometimes billions of parties, such as

¹ ALI-ELI Principles for a Data Economy (2021), approved drafts available at https://europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Projects/Data_Economy/Principles_for_a_Data_Economy_Final_Council_Draft.pdf and at <https://ali.org/projects/show/data-economy/>, print publication forthcoming.

generative AI trained on what generations of people have created on the World Wide Web. Does this have implications for the development and use of the technology and the content it creates? Does this call for new types of ownership and control?

This call is for research commissioned and supervised by GPAI. The project, which will be carried out in three phases, is expected to last until the end of March 2024, at which time a research report is expected to be delivered that will answer the research questions and provide the necessary guidance. The preliminary results of this research will be published and presented at the GPAI Summit in November 2023.

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”*.² Driven by a multistakeholder approach, it is supported in its mission by four Working Groups made up of leading international experts, and 29 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 AI³ 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

The characteristics of value creation in the digital economy are very different from those in the analogue economy. While the latter is characterised by value chains, in the digital economy value is generated within network structures of multiple actors who contribute in different ways to the generation and processing of data. These considerations are particularly relevant to data governance in the context of AI, where access to and use of data for the purpose of training AI is critical.

It is now more or less accepted that rights in data do not necessarily resemble the rights that have developed in relation to tangible property or intellectual property (IP), to name but two potentially relevant protection regimes. The main reason why rights to data do not resemble rights to tangible property is that data is a non-rival resource that can be used by multiple people at the same time. Data sharing therefore has the potential to increase economic welfare. Similarly, universal recognition of IP-like protection for all data would risk restricting access to

² Global Partnership on AI Terms of Reference

³ The Mandate draws upon the definitions set out within the [OECD Recommendation on Artificial Intelligence](#) for this purpose

data and could undermine third party access and use of data. This explains why ALI and ELI initially developed the concept of "rights in co-generated data", which are rights to access or port data, rights to demand that data not be used, to delete or correct data, or rights to share in the profits generated by the use of data.⁴ The concept as such has already been widely recognised, including by the German Data Ethics Commission,⁵ by the UNCITRAL Secretariat⁶ and by the European Commission in its 2020 data strategy⁷ and 2022 proposal for a Data Act⁸. It has also been recognized in the GPAI's framework paper for the Data Governance Working Group.⁹

However, this may be only one aspect of a much broader phenomenon. More generally, the question is how to allocate rights to access and use what has been generated by multiple parties in digital ecosystems, in particular data, content or AI itself. 'Co-generation' is a very broad concept and difficult to capture in general terms. Co-generation can refer to contributions by individual parties, but also to contributions by groups of people or even the public at large. Depending on the context, the legal relationships between different co-generators are often very complex, and co-generators may be numerous and their exact identities unknown, as is typically the case with AI, especially generative AI such as ChatGPT. Recognising collective rights to co-generated data, content or AI could help to prevent phenomena of 'digital colonialism' and promote systems of empowerment, such as indigenous digital sovereignty. 'Co-generation', in the sense that the public has contributed to the generation of data with taxpayers' money, is also one of several justifications for open data in the public sector.

Objectives

This project will examine how existing concepts of co-generation have been developed so far by existing initiatives, such as the recent EU proposal for a Data Act, how far these initiatives address an AI-specific context, and make suggestions on how they could be further developed to provide optimal benefits in the context of AI. As noted in the Working Party's framework paper, particular attention should be paid to the question of whether rights to co-generated data, content and AI should belong not only to individuals, but also to groups of individuals, such as those defined by language (e.g. for speech data), ethnic origin (e.g. for genetic data), location (e.g. satellite or sensor data) or activities (e.g. data from connected vehicles). Also, most attention has been paid to access rights, while rights to request desistance from collection,

⁴ See Part III of the ALI-ELI Principles for a Data Economy (2021), https://europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Projects/Data_Economy/Principles_for_a_Data_Economy_Final_Council_Draft.pdf.

⁵ Opinion of the Data Ethics Commission (2019), pp. 85 ff., available at https://www.bmj.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN_lang.pdf?__blob=publicationFile&v=3. The commission was a body advising the German Federal Government on issues of digital policy, and a whole range of the recommendations made have meanwhile been taken up by the EU or German legislators.

⁶ United Nations Commission on International Trade Law (UNCITRAL), Legal Issues Related to Digital Economy – Proposal for Future Work on Data Transactions (2022, A/CN.9/1117) available at <https://uncitral.un.org/en/commission>.

⁷ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A European Data Strategy, COM(2020) 66 final.

⁸ Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), COM(2022) 68 final.

⁹ GPAI, Framework Paper for GPAI's work on data governance, version 2.0, 2022, available at <https://gpai.ai/projects/data-governance/>

access or use have been notoriously neglected. Perhaps even more importantly in the context of AI, the focus has been on data and not so much on the content created using data or the AI trained on the data itself. It would therefore be essential to conduct research into the underlying concepts and principles once what has been co-created goes beyond data. Finally, more research would be needed to find out what 'co-generation' should mean, for example, whether it should include economic co-generation, and whether a party who did not actually contribute to the generation of the data, content or AI, but is now affected by the use of the data because that party shares common characteristics with the parties who contributed, could have similar rights.

In terms of impact, the project aims to contribute to the transnational convergence of national and regional policies on the design and allocation of digital rights and obligations. The project should draw on the values and objectives of the GPAI, i.e. human rights, inclusion, diversity, innovation, economic growth and societal benefits, as well as the UN Sustainable Development Goals, to guide the formulation of guidelines on data rights that are closely aligned with the GPAI's mandate. In doing so, the project aims to promote these values and goals beyond policy convergence as such. The interdisciplinary mix of GPAI's multi-stakeholder expert group provides the ideal institutional framework for the ambitious goals of the project.

2. Outputs expected

The required outputs from this project are:

- a research report (minimum 50 pages) that answers the following research questions and addresses the gaps identified below, and;
- a policy brief (5 to 10 pages) that distils the findings of the study into an easily digestible format and highlights the key (actionable and measurable) recommendations for policy makers.

The consultancy will be expected to carry out the research and share the results with the working group in 3 phases. The report and policy brief will be presented to the Project Steering Group and the Data Governance Working Group. Finally, the report will be published under GPAI and in accordance with GPAI communication guidelines. The consultancy will be responsible for delivering final written content of such outputs. Ceimia will manage the graphic design and layout of the documents.

Research Questions (What we need to know):

Phase 1 Q2 2023

- Collect and analyse different situations related to 'co-generation' in digital ecosystems and related governance regimes, including those of a legal, technical or organisational nature. Given the blurring of boundaries and in order to better identify the specificities of co-production, research should also cover data access rights and similar rights granted to third parties who have not contributed to the generation of the data, including in open data frameworks.

- Examine whether existing concepts and regimes can be made fruitful not only for individual parties but also for groups of parties, leading to collective data rights, with a particular view to combating phenomena of "digital colonialism" and digital exploitation of marginalised groups. Consider also the extent to which existing concepts and regimes adequately reflect the multi-relational nature of data and AI, for example with respect to parties affected by inferences drawn in relation to them.

Phase 2 Q3 2023

- Determine the extent to which existing concepts and regimes take into account the specific requirements of AI as a technology and the interest in developing and using AI from a public interest perspective, and if not, how they could be adapted.
- Explore the extent to which existing concepts and regimes are appropriate for dealing not only with data, but more broadly with content (including user-generated content) and technologies created with the help of data, in particular AI. Again, collective data rights and lessons from the open data debate should be drawn upon.

Gaps (What we don't have):

Phase 3 Q4 2023 - Q1 2024

- Develop principles that integrate different debates, in particular those on co-generated data, open data in the public sector, and existing sector-specific approaches, identifying cross-cutting issues and overarching patterns.
- Emit policy options and recommendations, taking into account the results of previous GPAI work, in particular on data justice, data institutions and PETs, as well as the state of the art on data sharing technologies and data sharing facilitators.
- Formulate guiding principles for dealing with 'co-opetition' situations in digital ecosystems that can inform and inspire a wide range of policy makers and decision makers, including legislators, in the GPAI participating states and globally, thus serving as a basis for more transnational coherence.

3. Resources

The Centre of Expertise will provide programme funding of up to a total of **146,500 Canadian dollars, including taxes**, for all outputs specified in this Request for Proposals (RFP) to appoint expert consultant/research partners to support the development and production of these outputs.

The project will be supported by a Project Steering Group led by Professor Christiane Wendehorst (University of Vienna). The Steering Group will include other members of the GPAI Working Group on Data Governance, and will seek to include representatives from the Working Groups on Responsible AI, Future of Work, and Innovation and Commercialisation. It may also include representatives from various member governments or external organisations wishing to collaborate with GPAI on this project.

From this group, the Centre will form an evaluation panel for responses to this call for proposals (excluding those with a potential conflict of interest at this stage).

4. Governance

The Consultancy/Research Partner will be responsible for the development and production of the outputs specified in the sections above and will be accountable for ensuring that the final product meets the standards and expectations of the Working Group.

In developing the content of the report, the Partner will be expected to work closely with the members of the Project Steering Group, who will meet with the Partner on a regular basis. The Partner is also expected to be aware of relevant international initiatives and to coordinate with them as appropriate. This is important both in the context of the expected involvement of governments or external organisations in the project itself and in other (non-GPAI) existing and planned initiatives in related areas.

The Centre of Expertise (CEIMIA) will manage the administrative process for the selection of the Partner, advise the Project Steering Group and provide contractual and day-to-day management of 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 the evaluation of bids and apply the evaluation criteria in Annex 1. A comprehensive, appropriate and robust approach to deliver the full scope of work is one of the evaluation criteria for the Partner (see methodology in Annex 1).

The Data Governance Working Group will review deliverables at key milestones (e.g. drafts) and will 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.

5. Timeline

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	April 6, 2023
Deadline for Proposals	Midnight (Anywhere on Earth), May 7, 2023
Evaluation and selection of proposal	May 11, 2023
Notification of successful tenderer	May 12, 2023
Contract signed and notification of remaining tenderers	May 19, 2023
Start of contractual engagement	May 22, 2023
<p>Preliminary round of interviews with the 4 GPAI working groups completed (Data Governance, Responsible AI, Future of Work, Innovation and Commercialization)</p> <p>Note: This milestone has been moved and is now captured under the completion of Phase 1 activities to allow sufficient time for its completion under the revised timeline.</p>	May 31st, 2023
<p>Phase 1 Completed:</p> <ul style="list-style-type: none"> ● Conducted preliminary round of interviews (and summarised key findings from) with experts from the 4 GPAI Working Groups (Data Governance, Responsible AI, Future of Work, Innovation and Commercialization) ● Completed research tasks foreseen for Phase 1 of the project (see above) ● Completed consultation and feedback workshop with GPAI DG Working Group and Project Advisory Group 	July 5, 2023
<p>Phase 2 Completed:</p> <ul style="list-style-type: none"> ● Completed research tasks foreseen for Phase 2 of the project (see above) ● Completed consultation and feedback workshop with GPAI DG Working Group and Steering Group ● Submitted a report containing the preliminary results of phase 1 and 2 (with the intention to be publicly shared at the 2023 GPAI summit) 	September 28, 2023

Milestone	Date
<p>Phase 3 Completed:</p> <ul style="list-style-type: none"> Completed research tasks foreseen for Phase 3 of the project (see above) Completed consultation and feedback workshop with GPAI DG Working Group and Project Advisory Group 	February 15, 2024
<ul style="list-style-type: none"> Submit drafts of the full scope of outputs Conducted a feedback workshop with the Project Advisory Group and the GPAI Data Governance Working Group to collect comments and feedback. 	March 1, 2024
<ul style="list-style-type: none"> Submit final outputs Final presentation of the full scope to the Data Governance Working Group 	March 24, 2024

Annex 1: Tender response and evaluation criteria

By midnight (anywhere on Earth) on **May 7, 2023**, interested parties should submit a costed proposal (in English) to Sophie Fallaha (sophie.fallaha@ceimia.org) and Arnaud Quenneville-Langis (arnaud.quenneville@ceimia.org), which includes:

- the project title “**From co-generated data to generative AI - new rights and governance models in digital ecosystems (CO-GEN)**” 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), arnaud.quenneville@ceimia.org (AI Initiatives Lead, CEIMIA) and stephanie.king@ceimia.org (Director of AI Initiatives, CEIMIA) no later than **April 30, 2023**. 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 one or more partners as part of the proposal. 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 the commercial, public or social sector that demonstrate the team’s ability to work together to produce meaningful and impactful results, particularly where team members will be spread across different locations and/or organisations. Legal expertise in the field (expertise in policy translation and integration) will be an essential element. The team</p>

Criteria	Weight	Page limit	Guidance
			should have a demonstrated ability to engage productively by providing actionable information to policy makers.
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</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 GPAI'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>

Criteria	Weight	Page limit	Guidance
			<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> $\text{Score} = \frac{\text{Lowest Tender Price}}{\text{Tender Price}} \times 20 \text{ (Maximum available marks)}$

Criteria	Weight	Page limit	Guidance
			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

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:

-