

Enabling data sharing for social benefit through data trusts

Terms of reference (04/06/21)

This Terms of Reference provides guidance on deliverables that the Global Partnership on AI's ("GPAI") Data Governance Working Group will deliver for GPAI's 2021 Summit (expected to take place November 11-12th 2020). Formal requirements for tender responses are specified under Annex 1.

1. Context

"Enabling data sharing for social benefit through data trusts" is a new project from the GPAI data governance working group, which will explore the role that data trusts can play in enabling effective data stewardship. This project will seek to build common understandings of the core functions of data trusts, identify best practice in the creation and operation of data trusts, and create an environment that supports their development.

This document relates to an initial phase of activity for the project. In this phase, GPAI hopes to lay the foundations for an ambitious research agenda and set of practical actions to create data trusts in areas of need, by synthesising the 'state of the art' in data trust methods, identifying best practice and understanding where further action may be required.

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 Interim Executive Director, Jacques Rajotte, and the Working Group is supported by Ed Teather (on secondment from the UK's Office for Artificial Intelligence).

For Summit 2020, the Data Governance Working Group produced two projects:

1. [A Framework for GPAI's work on Data Governance](#) - setting the stage for all future Working Group projects, serving as an overview over the most relevant terms and defining the understanding of the Working Group of data governance in the context of AI. The Consultant may wish to familiarise themselves with this paper and

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

its terms, and, acknowledging that it is a living document of the Working Group, suggest updates as relevant in light of this project.

2. [An investigation into the Role of Data in AI](#) - to complement and dig into topics in the Framework in more depth, this situates the importance of data to AI development and identifies areas both where more data would be useful - such as specific, open, datasets that could be worthy of national support or international collaboration - and where harms arise due to the collection of, use of or access to data.

Following a process of ideation, concept development, and prioritisation, the Working Group has determined that it will focus on two cross-domain projects for the next 18 months: (1) *“Advancing data justice research and practice”*, and (2) *“Enabling data sharing for social benefit through data trusts”*. In addition, the Working Group will be collaborating with other GPAI projects to advise on the data governance aspects of their projects; these include climate change and biodiversity, drug discovery and open science, social media governance, and intellectual property.

This Terms of Reference focuses on *“Enabling data sharing for social benefit through data trusts”* and the production of key deliverables in support of GPAI’s 2021 Summit, scheduled for 11th-12th November 2021. The purpose of the Summit is to bridge the gap between theory and practice on AI by supporting cutting-edge research and applied activities on AI-related priorities.

About the project

Why this topic?

Effective data stewardship is essential for a variety of public policy goals and for realising the wider economic value of data. Recent years have brought numerous examples of the potential benefits from innovative uses of data, and examples are emerging of areas in which communities are seeking to promote data use for certain purposes. At the same time, evidence of the harms associated with poor data governance practices is mounting across sectors from recruitment to criminal justice, leading to calls for new data institutions that provide careful stewardship of data, through independent oversight and mechanisms for accountability. In response, new forms of collaborative data stewardship are emerging, and data trusts are a promising tool to empower individuals and communities.

What is a data trust?

A data trust is a data institution in which individuals pool their data or data rights, creating an independent organisation that is tasked with stewarding the data rights it holds for the benefit of its members. Data trusts perform the following three core functions:³

- Provide independent stewardship of data or data rights on behalf of the trust’s beneficiaries to enable data use for agreed desirable purposes and to protect against vulnerabilities arising from data use;
- Create robust institutional safeguards and strong fiduciary responsibilities surrounding the use of data or rights held in trust, with the aim of preventing mis-use and supporting remedial action if it does occur.
- Enable collective action, supporting its members to secure more favourable terms and conditions of data use, based on the enhanced negotiating power secured by bringing together data or rights in the trust.

³ In June 2021, GPAI will be hosting a workshop to review these core functions and assess whether there are further functions distinct to data trusts - in comparison to other data institutions - that should be added to this list. Outputs from this workshop will be provided to the delivery partner to support their analysis.

Data trusts fill a gap in the current data governance landscape by empowering individuals to exercise their data rights; supporting individuals and communities to share data for social benefit; and providing stewardship that anticipates the vulnerabilities resulting from data (mis)use and takes action to prevent these digital harms emerging. In these respects they offer a new form of data institution.

Different legal frameworks exist in different jurisdictions to support the creation of data trusts. In many common law countries, trust law can be used as a basis for data trusts (and the three core functions listed above); in civil law jurisdictions, other mechanisms exist to fulfil the core functions of a trust.

Why this project?

Important challenges must be addressed to establish data trusts:

- There is a need to create consensus around the core features of a data trust, which – as a starting point – this project takes to be: independent stewardship, strong institutional safeguards, and bottom-up collective action.
- There is a range of operational considerations that must be taken into account in the implementation of data trusts – the nature of the data rights being managed, long-term financial sustainability, technical architectures, and mechanisms for accountability, for example – and further work is needed to understand these and identify best practice.
- Different legal models will be more (or less) suited to operationalising data trusts in different areas. Better understanding of the jurisdictional issues associated with their establishment are necessary, alongside better sharing of information about lessons learned from projects across the world.
- The process of creating data trusts also seems likely to expose the limitations of current data governance frameworks, identifying areas where action is needed from policymakers; this might include creating new legal principles, for example in rights arising from co-generated data, or adapting current policies to allow innovation in data governance while providing appropriate safeguards.

2. Objectives

In this first stage of activity, this project's primary goals are to:

- Consolidate international perspectives on 'what is a data trust?'
- Identify current practice in developing data trusts from across the world and better understand the incentives and ways of working that can make data trusts – or data institutions that fulfil equivalent functions – accessible to all in society;
- Identify what legislative frameworks may be required to support the development of data trusts; and
- Highlight gaps in the operational or governance environment that need to be addressed to support the wider use of data trusts for social benefit.

Based on the above, GPAI will identify interventions to support pilot projects that trial data trusts methods in practice.

3. Outputs

There are two final required outputs that must be delivered by this project - proposals may be submitted for one or both:

1. **A synthesis of the 'state of the art' in the design and implementation of data trusts, based on a review of existing relevant data institutions from across the**

world that explores how different projects have implemented the core features of a data trust (hereafter ‘Output 1: Data trust survey’). This will include: a description of the core functions of data trusts, an analysis of how different data institutions have attempted to fulfil the functions of a data trust, and case studies to explore how data trusts can work in practice. This will take the form of a report (30-40 pages in length), presentation summarising key findings and an interactive toolkit that will enable readers to explore how different projects have addressed different operational challenges.

2. **A review of the legal and legislative frameworks that are in place or emerging for the governance of data institutions, to analyse the current landscape of data rights, understand the legal and legislative frameworks that are required to develop data trusts, and identify areas of uncertainty** (for example, rights relating to co-generated data) **or need** (for example, structures for accountability or safeguarding in the event of institutional failure). (Hereafter ‘Output 2: Legal review’). This will take the form of a report (20-30 pages in length) and presentation summarising key findings and areas for action.

In addition to these final outputs, an **interim report** should be produced to be presented at GPAI’s 2021 summit. This interim report should present a description of the core functions of data trusts, an analysis of how different data institutions have fulfilled these functions, case studies, and a review of the legislative frameworks surrounding data institutions. It should also present emerging findings or areas for action. After the interim report, there will be an opportunity to refine findings and create the interactive toolkit as part of the production of final outputs.

For the **Interim Report** presented at the Summit (see ‘Timelines’ under Section 8), our expectations are that outputs should be approved as meeting the standards and expectations of the Working Group and GPAI’s Steering Committee, and be ready for publication for the purpose of consultation. The finalised outputs, refreshed in light of discussion at Summit 2021 and wider consultation and external review, will then be published by 31 January 2022.

4. Scope

Output 1 (Data trusts survey) seeks to (a) provide clarity on the core functions of data trusts (b) generate insights into the ways of working that can deliver those functions and issues associated with those ways of working and (c) identify gaps in understanding or practice. This output should:

- Consider the role of data trusts in data stewardship (in the context of the range of available data governance interventions).
- Set out the core features of a data trust, as distinct from other forms of data institution. This will be informed by a draft consensus statement on ‘what is a data trust?’ to be produced by GPAI working group members and provided to the consultancy partner.
- Survey the current landscape of data trust-like initiatives to evaluate how these current projects fulfil the core features of a data trust. This survey would include:
 - o A description of survey methodology;
 - o A summary that shows how projects identified by the survey demonstrate the different features of a data trust;
 - o Descriptions of the ways of working employed by different projects, including:
 - The purposes for which data is used (and how these purposes relate to the expressed preferences of individuals and groups involved in the trust).
 - The legal basis for data sharing and the form of legal relationship that is the foundation of the data sharing activity.

- The technology approach being used to enable data sharing.
- The business or funding model that enables the data sharing activity.
- The decision-making process through which relevant stakeholders agree on the form and function of data sharing in the project (for example: what type of consultation process is involved).
- The organisational capabilities required to enable effective operation of a data institution.
- Other factors relevant to the establishment of a data trust.
- Provide case studies of specific projects that allow the reader to understand how data trusts operate. These may include success stories, where data trusts have enabled effective data sharing (and the factors contributing to success); and war stories, where projects have not achieved their goals (and the reasons for this).
- Combine insights from the survey and case studies into a toolkit that will enable individuals/groups seeking to set up a data trust to explore the ways of working that other projects have operationalised to fulfill the core functions of a data trust.
- Summarise the core messages emerging, by:
 - Describing the data governance challenges that data trusts have been able to address.
 - Describing the current 'state of the art' in data trusts design and implementation;
 - Identifying of gaps in current practice;
 - Commenting on where further guidance or best practice could play a role in the development of data trusts;
 - Providing summary visuals or analysis to explore which practices or projects are emerging from across the world.

Output 2 (Legal review) seeks to (a) assess what new types of data right might be emerging from new patterns of data use (b) characterise the role that different legal frameworks can play in helping data trusts and data institutions manage data rights and (c) identify where there may be gaps in current legal and legislative frameworks in relation to data rights and data trusts. This output should:

- Set out which rights currently exist in relation to data use and data sharing, and the legal basis for these rights.
- Analyse:
 - What data rights already exist (and for whom);
 - Which patterns of data use might challenge the boundaries of these existing data rights, seeking to understand whether current data rights are sufficient to address current or emerging practices in data sharing (for example, in relation to third party access and co-generated data);
 - What legal mechanisms or institutional frameworks (for example, trust law, in the case of data trusts) are emerging as ways of governing these new types of data use.
 - Where there are gaps in current legislative frameworks, arising either from:
 - The need for new types of data right (for example, around co-generated data);
 - The need for legislative interventions to support the development of data institutions that influence how individuals or groups exercise their data rights (for example, structures for accountability or safeguards surrounding the development of data trusts); or
 - Other issues emerging from the review.
- Give an overview of what legal frameworks exist to enable the creation of data trusts in different jurisdictions. This would be supplemented by:
 - Description of research methods, including justification for any countries selected as specific case studies.

- o Case studies that illustrate different frameworks in more detail. The jurisdictions to be explored in such case studies should be proposed as part of the research methodology (Annex 1).
- Provide a summary of core messages from the review, including:
 - o Description of how different legal models can support data stewardship;
 - o Identification of gaps in current legal or legislative frameworks;
 - o Lessons that emerge from across jurisdictions with regards to the development of data trusts.

5. Resource

The Centre of Expertise will provide programme funding (of up to \$147,000 Canadian dollars in total for all outputs specified in this ToRs) to appoint expert consultancy support for the development and production of these outputs.

Note that if a proposal is submitted for only one output, then the programme funding is up to \$73,500 Canadian dollars in that instance. Proposals from consortia will be welcomed.

Working Group members that have volunteered for this project currently includes:

- Neil Lawrence (Project Co-Chair), University of Cambridge (United Kingdom)
- Seongtak Oh (Project Co-Chair), National Information Society Agency (South Korea)
- Jeni Tennison (Working Group Co-Chair), Open Data Institute (UK)
- Maja Bogataj Jančič (Working Group Co-Chair), Intellectual Property Institute (Slovenia)
- Jess Montgomery, University of Cambridge (United Kingdom)
- Matija Damjan, University of Ljubljana (Slovenia)
- Carole Piovesan, INQ Data Law (Canada)
- Kim McGrail, University of British Columbia (Canada)
- Bertrand Monthubert, Occitanie Data/EkitIA (France)
- Nicolas Mialhe, The Future Society (France)
- Paul Dalby, Australian Institute of Artificial Intelligence (Australia)
- Christiane Wendehorst, University of Vienna / The European Law Institute (Austria)
- Yeong Zee Kin, Personal Data Protection Commission (Singapore)
- Teki Akuetteh Falconer, Africa Digital Rights' Hub (Ghana)
- Alison Gillwald, Research ICT Africa (South Africa)
- Josef Drexl, Max Planck Institute for Innovation and Competition (Germany)
- Alejandro Pisanty Baruch, National Autonomous University of Mexico (Mexico)
- Iris Plöger, Federation of German Industries (Germany)
- Ricardo Baeza-Yates, NTENT (Spain)
- Aleksandra Przegalińska, Kozminski University (Poland)
- Zümrüt Müftüoğlu, Yildiz Technical University (Turkey)

From these, the Centre will establish an **Evaluation Panel** for bids (absolving those with a potential conflict of interest at this stage), and form the **Project Steering Group**.

6. Governance

The **Consultancy 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 Consultancy Partner will:

- Undertake literature reviews and survey activities to scope and deliver the outputs in section 2.

- Collate feedback on the draft consensus statement on ‘what is a data trust’ as part of producing output 1.
- Manage consultation with the Project Group as the project progresses, by:
 - Contributing to regular project team meetings to report on progress in developing these outputs.
 - Convening red-teaming sessions to review emerging insights and outputs:
 - At project initiation;
 - When initial findings/conclusions are available;
 - At ‘heads of report’ stage;
 - When draft outputs are available.
- Convene expert workshops on specific themes to gather insights from the GPAI community.
- Produce draft reports, summaries and toolkits for consideration by GPAI.
- Coordinate review and collate feedback from the Project Group, Working Group, external reviewers and Steering Committee to produce a final draft.

In developing content for the report, the consultancy partner would be expected to work closely with the Working Group members listed under section 5, who will act as a **Project Volunteer Group** for the project.

A sub-set of the Project Volunteer Group, the **Project Steering Group** will lead on project direction. The Group consists of: Jess Montgomery, Neil Lawrence, Seongtak Oh and YeJee Parks. The Project Steering Group will meet with the Consultancy Partner on a regular basis.

The **Centre** (CEIMIA) will run the administrative process for the selection of consultancy, provide advice to the **Project Steering Group**, and provide contractual and day-to-day engagement with the **consultancy 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** 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 **Consultancy Partner**. This review will be further supported by selected **external reviewers** who will provide additional peer review and quality assurance of the final output.

The **GPAI Steering Committee** will have final sign off on outputs pre-publication.

7. Methodology

The **Centre** (CEIMIA) will lead on organising and running the process to select a **consultancy 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 Consultancy Partner (see methodology under Annex 1). Once appointed, the Consultancy Partner will:

- undertake a **literature review and survey activities** structured to align to the scope of the project (building on the background materials that the Project Steering Group will have collated) - following consultation with the Project Steering Group on the best reference points
- in consultation with the Project Steering Group, convene **expert workshops**.

- incorporate a **set of review points** (as referenced under section 6) so that the outputs align with the perspective of the Project Steering Group and wider Working Group.

8. Timeline

The project will be a critical contribution to GPAI's November 2021 plenary meeting. Bidding consultancy partners are required to set out a timescale for meeting the full scope of the project in their proposals (see Annex 1 under Project Management).

Key dates are set out in the table below:

Milestone	Date
Request for Proposals published	4 June 2021
Deadline for Proposals	Midnight (Anywhere on Earth), 18 June 2021
Evaluation and selection of proposal	22 June 2021
Notification of tenderers	24 June 2021
Contract signed	27 June 2021
Start of consultancy engagement	28 June 2021
Research proposal reviewed by Project Steering Group (meeting)	1 July 2021
Emerging findings and heads of report reviewed by Project Group (meeting / workshop)	28 July 2021
Red-teaming session with Project Group on survey results and updated heads of report.	7 September 2021
Draft Summit outputs provided for review by Working Group (and selected external reviewers)	10 September 2021
Feedback provided by Working Group and reviewers	22 September 2021
Updated draft of summit outputs provided for approval by Working Group and CEIMIA.	24 September 2021
Summit outputs as approved by Working Group and CEIMIA submitted to GPAI's Steering Committee	30 September 2021

Summit presentation of interim report	11-12 November 2021
Interim output published for external consultation.	11 November 2021 - 8 December 2021
Review session to agree updates to the interim output.	10 December 2021
Drafts of final outputs and toolkit provided to Project Group and Working Group for review.	10 January 2021
Feedback provided by Project Group and Working Group on final drafts.	14 January 2022
Submission of outputs to Project Steering Group for final review	19 January 2022
Submission of final outputs	31 January 2022

The Working Group as a whole currently meets on a monthly basis for up to two hours. The Working Group Co-Chairs will meet with Project Co-Chairs on a fortnightly basis. The Project Steering Group will be available to meet on a weekly basis.

Annex 1: Tender response and evaluation criteria

By midnight (anywhere on Earth) on 18 June 2021, interested parties should submit a costed proposal (in English) to jacques.rajotte@ceimia.org and edward.teather@ceimia.org, which includes:

- the project title “Enabling data sharing for social benefit through data trusts” 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 jacques.rajotte@ceimia.org (Interim Executive Director, CEIMIA) and edward.teather@ceimia.org (secondee to CEIMIA) no later than the 14th of June. CEIMIA reserves the right to make both anonymised questions and answers public or shared with other organisations having stated their interest.

Evaluation criteria

GPAI has a mission statement that states that it should “*facilitate international project-oriented collaboration ... taking into particular account the interests and contributions from emerging and developing countries*”. This is reflected within our criteria under “ability to access and include perspectives from low and middle income countries”. The Working Group strongly encourages proposals from teams that can represent these perspectives.

The evaluation panel will review proposals as below:

Criteria	Weight	Page limit	Guidance
Skills, expertise and experience	20%	2 pages	<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, qualifications and professional accreditations) 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 previous reports, papers, books, articles, op-eds, talks, or other deliverables that demonstrate a history of team members’ ability to deliver on the scope of this request. Describe the team’s ability to work together to collectively produce meaningful and impactful outputs.</p>
Proposed methodology (including delivery plan plus	20%	3 pages	<p>Please provide details of your Delivery Plan and methodology for the project, from contract commencement to contract completion.</p>

commitment to work collaboratively and in the open)			<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.
Ability to access and include perspectives from low and middle income countries	20%	2 pages	Please demonstrate how your work will include perspectives from low and middle income countries.
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

			<p>and with flexibility, and that the management structure is robust;</p> <ul style="list-style-type: none"> • 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><i>Score = $\frac{\text{Lowest Tender Price}}{\text{Tender Price}} \times 20$ (Maximum available marks)</i></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%.