Data Governance
Working Group Report

November 2022 - GPAI Tokyo Summit

GPAI / THE GLOBAL PARTNERSHIP
ON ARTIFICIAL INTELLIGENCE

CONTENT

Co-Chair's Welcome

Dr. Jeni Tennison
Executive Director
Connected by Data

Dr. Maja Bogataj Jančič
Founder and Head
Intellectual Property
Institute

The increasing use and collection of data has a direct influence on our daily activities. We still underestimate the impact and influence of artificial intelligence (AI) on our daily choices. An action as simple as selecting the best route for our daily commutes or choosing the music that will accompany us is intrinsically linked to the data transmitted on the various digital platforms we use.

Our Working Group mandate aims to "collate evidence, shape research, undertake applied AI projects and provide expertise on data governance, to promote data for AI 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."

Since the Paris Summit 2021, we have achieved considerable advancement in the field of Data Governance. We are deeply proud of the progress we have made with our wonderful Working Group members who are passionate about responsible use and governance of data as we are!

Throughout the last 12 months we have delivered two great projects and we have also launched a third one in April 2022. We've worked on those exciting projects with fantastic partners and we were also pleased to secure additional funding from the UK's Office for AI and the Singapore Infocomm Media Development Authority (IMDA) to scale our up from the "seed funding" initially provided by GPAI's Montreal Centre of Expertise (the "CEIMIA"):

(1) **Enabling data sharing for social benefit through data institutions**: supporting the creation of real-world data trusts that enable safe and equitable data sharing for social benefit and empower individuals to enact their data rights

(2) **Advancing research and practice on data justice**: providing a framework for data justice research and practice and include considerations of justice in terms of access to, representation and transparency in data used in AI development

(3) **Technology to support data availability for AI**: demonstrating practical use of Privacy Enhancing and adjacent technologies for well-governed data access for AI

As an important first step on Data Trusts, the Working Group produced an interim report that explored real-world use cases and operationalisation strategies where data trusts could offer social benefit, with a specific focus on the GPAI's Council priority: AI and climate action. The Working Group has collaborated with the Open Data Institute, the Aapti Institute and Cambridge University on this research.

On Data Justice, we published in March 2022: A Repository of Case Studies, An Integrated Literature Review, An Interim Annotated Bibliography and three Guides for policymakers, developers and impacted communities. The Working Group has collaborated with the Alan Turing Institute and 12 pilot organisations representing Low and Middle Income Countries. Following those publications, the Working Group decided to distilled those reports into two primers and a policy brief.

In April 2022, we launched our Privacy-Enhancing Technologies (PETs+) project and we have finalised the first phase by the selection of a use case on "Better Health – Modelling the effects of human movement during a pandemic", for which the demonstration can seek to integrate with the existing Responsible AI Working Group's project on Immediate Response (now re-termed as Pandemic Resilience). This use case enriches existing pandemic models that use anonymized and aggregated data, by adding data that describes (in detail) an individual's movements.

These three projects kept us busy and we have many more ideas (demonstrated by the Applied Research Agenda we published last year), and are excited to now be bringing forward two new projects in 2023.

We go into 2023 full of excitement for what we can achieve, and appreciation for all the commitment, dedication and of our Working Group experts, who have been a joy to collaborate with.

We also want to acknowledge the tremendous work of our project co-leads who's term as an expert will end by the end of 2022. Very many thanks to Alison Gillwald, Dewey Murdick, Seongtak Oh, Shameek Kundu, Neil Lawrence, and Teki Akuettah Falconer for your passion, dedication and all the fruitful discussions we had over the last months.

Introducing the Data Governance Working Group

The Data Governance Working Group (DG for short) brings together 43 experts, including six observers, from 22 countries with experience in technical, legal and institutional aspects of data governance. All our experts have shown tremendous collaboration, creativity, commitment and, if we may say so, great humour as we advanced our work over the past year.

We have also welcomed the new energy, expertise and insights from new self-nominated experts. The interdisciplinary and intercultural diversity within the Working Group continues to make the research fresh and exciting, and we look forward to building on this again next year.

40% of DG''s members are women, a number which we'll work to increase in the future.

Most members (65%) come from the science sector, 23% are from the civil society and 12% are from the industry. A better balance should be achieved 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.

Members were based in 27 countries and territories: Austria, Australia, Belgium, Brazil, Canada, Czech Republic, Denmark, France, Germany, Ghana, India, Italy, Ireland, Israel, Japan, Korea, Mexico, the Netherlands, New Zealand, Poland, Singapore, Slovenia, South Africa, Spain, Taïwan, the United Kingdom, and the USA.

These members have either been designated by the members of GPAI or through the self-nomination process. It's worth mentioning that each member acts with full independence inside the Working Group.

Finally, six additional specialists take part in RAI's activities as observers. One of them is a representative of the OECD, a strategic partner of GPAI, and another one a representative of UNESCO.

The next page presents the Working Group's experts and observers

DG's Members

Experts of GPAI's Data Governance Working Group

Jeni Tennison (Co-Chair) – Connected by Data (UK)
Maja Bogataj Jančič (Co-Chair) – Intellectual Property Institute (Slovenia)
Allan Feitosa – Summ.link (Brazil)
Alejandro Pisanty Baruch – National Autonomous University (Mexico)
Aleksandra Przegalińska – Kozminski University (Poland)
Alison Gillwald – Research ICT Africa (South Africa / UNESCO)
Alžběta Krausová – Institute for State and Law (Czech Republic)
Anderson Soares – Artificial Intelligence Center of Excellence (Brazil)
Asunción Gómez – Technical University of Madrid (Spain)
Bertrand Monthubert – Ekitia (France)
Carlo Casonato – University of Trento (Italy)
Carole Piovesan – INQ Data Law (Canada)
Ching-Yi Liu – National Taïwan University (Taïwan)
Christiane Wendehorst – European Law Institute / University of Vienna (Austria / EU)
Dani Chorin – Israeli Government (Israel)
Dewey Murdick – Center for Security and Emerging Technology (USA)
Hiroki Habuka – Kyoto University (Japan)
Iris Plöger – Federation of German Industries (Germany)
Jhalak Mrignayani Kakkar – Centre for Communication Governance (India)
Jeremy Achin – hOS / Neo Cybernetica (USA)
Josef Drexl – Max Planck Institute (Germany)
Kim McGrail – University of British Columbia (Canada)
Marc Rotenberg – Centre for AI and Digital Policy (USA)
Matija Damjan – University of Ljubljana (Slovenia)
Mikael Jensen – D-Seal (Denmark)
Neil Lawrence – University of Cambridge (UK)
Nicolas Miailhe – The Future Society (France)
Oreste Pollicino – University of Bocconi (Italy)
Paola Villerreal – Snap Inc. (USA)
Paul Dalby – Australian Institute of Machine Learning (Australia)
P. J. Narayanan – International Institute of Technology, Hyderabad (India)
Radim Polčák – Masaryk University (Czech Republic)
Ricardo Baeza-Yates – Universitat Pompeu Fabra & Northeastern University (Spain)
Robert Kroplewski – Minister for Digitalisation of the Information Society (Poland)
Seongtak Oh – National Information Society Agency (South Korea)
Shameek Kundu – TruEra (Singapore)
Takashi Kai – Hitachi (Japan)
Teki Akuetteh Falconer – Africa Digital Rights Hub (Ghana / UNESCO)
Te Taka Keegan – University of Waikato (New Zealand)
Ulises Cortés – Barcelona Supercomputing Center & Universitat Politècnica de Catalunya (Spain)
V. Kamakoti – International Institute of Technology, Madras (India)
Yeong Zee Kin – Infocomm Media Development Authority (Singapore)
Yoshiaki Nishigai – University of Tokyo (Japan)

Observers
Claudia Juech – The Patrick J. McGovern Foundation
Christian Reimsbach-Kounatze – OECD
Jaco Du Toit – UNESCO
Nagla Rizk – American University in Cairo
Naoto Ikegai – Hitotsubashi University, Tokyo
Zümrüt Müftüoğlu – Yildiz Technical University

Mandate of the Data Governance Working Group

As mentioned in the foreword, the Working Group's work aligns closely with GPAI's overall mission. Our Working Group aims to "collate evidence, shape research, undertake applied AI projects and provide expertise on data governance, to promote data for AI 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."

There are interactions between data governance and the remits of the other Working Groups. Our Working Group is pleased to be continuing its collaboration with the Innovation and Commercialisation Working Group on intellectual property. We also have good collaboration happening with both the Responsible AI and Future of Work Working Group on the Privacy-Enhancing Technologies (PETs+) project with one of the chosen use cases on "Pandemic Resilience".

The Working Group also has the chance to help coordinate GPAI's applied AI ambitions, shape projects carried out by or funded by GPAI's members and across its wider partnerships, and to influence the policy recommendations created by the OECD through its work. Our goal is that our work is also useful more widely, amongst those researching, thinking about and implementing data governance practices in AI.

Working Group Timeline

JANUARY

First Working Group meeting (12th): welcoming new self-nominated experts and project updates.

FEBRUARY

Second meeting of the Working Group (15th) – presentations in the content to be included in the project outputs (Data Justice & Data Trusts) for the AI:UK Launch Event.

MARCH

AI UK Launch Event - Official Launch Event of the (21st & 22nd)

Third meeting of the Working Group (30th) – working session on the next phases of the ongoing projects and the nex proposals for the 2023 work plan.

APRIL

Privacy-Enhancing Technologies Project Launch (6th)

JUNE

Fifth and sixth meeting of the Working Group (15th) – project updates and presentation of the proposals to be included in the work plan 2023.

JULY

Working Group members to meet one-on-one with the Montreal Center of Expertise Coordinator during the summer period. Those meetings allowed us to discuss the work field of each expert and also gather useful comments on the projects, GPAI structure and what ideas could be explored to deliver more impact.

AUGUST

Seventh meeting of the Working Group (24tht) - special presentation on "technocoloniality" by CEIMIA's Researcher-in-Residence, Thomas Nkoudou, followed by project updates and timeline up to the plenary meeting at the Tokyo Summit.

SEPTEMBER

Eight meeting of the Working Group (14tht) - working session on the priorities and policy recommendations to be included into the Multi-Stakeholder Plenary Report.

OCTOBER

Ninth meeting of the Working Group (19th) – finalisation of the session that will be presented at the plenary session at the Summit followed by a welcoming presentation by two experts of the working group.

NOVEMBER

Tenth meeting of the Working prior to the Summit (2nd)

Presentation of finalized outputs and open workshop on next projects at the Summit (21st & 22nd)

Progress Report

Building on a detailed investigation into the role of data in AI and committed to push forward cross-domain projects and develop cross-working groups collaborations, the working group has developed a research agenda with seven detailed concept notes. Based on these the working group has worked trough out 2022 on three projects combining elements of those concept notes that were approved to move forward:

(1) **Enabling data sharing for social benefit through data trusts**: established to support the creation of real-world data trusts that enable data sharing for social benefit. It will support new institutions that empower individuals and communities to enact their data rights, ensuring that data sharing activities reflect the diverse interests of all in society. The end goal is to help GPAI realise the potential of data trusts as a tool to promote the safe, fair, legal and equitable sharing of data, in service of the UN Sustainable Development Goals.

(2) **Advancing research and practice on data justice**: established to fill a gap in Data Justice research and practice that provides a frame to help policy makers, practitioners and users to move beyond understanding data governance narrowly as a compliance matter of individualised privacy or ethical design, to include considerations of equity and justice specifically as it relates to redressing the uneven distribution of opportunities, and harms, associated with AI and ML currently. The objective is to make significant progress in getting more equitable access to, greater visibility and fairer representation of those individuals and communities marginalised from data used in the development of AI/ML systems, through the adoption of more just principles into AI policy and practice.

(3) **Privacy-enhancing technologies project**: established to demonstrate the viability of AI systems in helping achieve the UN SDGs such as global health, climate action, and the future of work in harmony and with flourishing of human dignity, by providing a means to safely develop, use and share data while preserving privacy, sovereignty, personal integrity, IP rights, and security. The project also aims to overcome challenges to data usability commonly faced when working with PETs by publishing practical guidance and lessons learnt from the demonstration system. This can support innovation by helping smaller organisations or corporations to compete more effectively with large (and sometimes monopolistic) data-rich organisations that have access to massive datasets within their organisational boundaries.

Enabling data sharing for social benefit through data institutions

Building on our work on data trusts in 2021 with an exploration of a practical context: climate action

In 2022, the Working Group has explored real-world use cases and operationalisation strategies where data trusts could offer social benefit, with a specific focus on the GPAI's Council priority: AI and climate action[1]. The Working Group has been supported in this work by the Open Data Institute, Aapti Institute, and the Data Trusts Initiative, with special funding by the UK Government.

This built upon important foundational work by the Working Group in 2021, including the first international definition of a data trust[2], a synthesis of the 'state of the art' in the design and implementation of data trusts[3], and a review of the legal and legislative frameworks that are in place to support the operationalisation of data trusts[4].

Determining where exploration of bottom-up data institutions could make the most impact on climate action

To begin exploration of potential use cases, the team consulted with experts in data stewardship, climate, data science and analysis, and artificial intelligence to create a long-list of climate challenges including city mobility, energy use, agriculture, community resilience, climate migration, water use, wildlife conservation, and air quality.

To narrow down to three priority domains for data institutions, the team considered the following questions:

- Can data play a critical role in tackling the challenges faced in this domain?

- Can data from this domain be used to develop or deploy artificial intelligence based systems?

- Are there opportunities in this domain for bottom-up approaches that empower individuals to play a role in stewarding data?

- Would focusing on this domain help us to explore privacy (and other rights) enhancing technologies?

From this initial exercise, the Working Group agreed identified the following three use cases as good areas of focus for bottom-up data institutions. Whilst each use case focused on a particular localised community and ecosystem, they could obviously apply more broadly:

1. City cycling in London: to provide the cycling community with a platform to contribute data to affect positive changes to the design of cycling infrastructure in their city.

2. Small shareholder farming in India: to help farmers gain value from their data in the form of more tailored advisories for improving the efficiency of their practices, and also provide hyperlocal agricultural information to other stakeholders.

3. Climate migration in Peru: to act as steward of qualitative data sets related to indigenous climate displacement from across Peru, ensuring that the data is shared with trustworthy organisations, and importantly, accessible to displaced communities themselves.

---

[1] See the full report here
[2] See the Working Group statement on 'Understanding Data Trusts' here
[3] See Part 1 of the Summit 2021 report here
[4] See Part 2 of the Summit 2021 report here

Developing a feasibility assessment to test data trusts against these climate challenges; a roadmap, and broader findings

As the Working Group's 'state of the art' review in 2021 found, data trusts remain a novel, experimental approach (distinctive for their independent trustees with fiduciary obligations) within the wider family of data institutions.

As an additional contribution and tool for the exploration of data trusts real-world deployment, **a feasibility assessment approach was developed to test data trusts** in the context of climate migration, small shareholder farming, and city cycling.

The feasibility assessment has a strong focus on community, and includes criteria on: incentives; capacity; leadership; data demand and clearly bound use cases; legal instruments in the specific jurisdiction; available technical means for data control by the data trust, and financial sustainability.

In its findings on the feasibility assessment, the team found a **positive assessment on the potential of a city cycling data trust to enable communities to use data to advocate for, and inform the design of, sustainable transport infrastructure.** A design and practical roadmap has been articulated for this use case that could be adopted by city authorities, and we would welcome approaches from parties interested in exploring implementation.

The assessment of **small shareholder farming** and **climate migration** indicated challenges for the broader implementation of data trusts (especially regarding the data rights necessary for data trusts being present in local jurisdictions), yet highlighted important challenges in the Global South with clear opportunities to improve how data is collected, used and shared.

It was recommended that the Working Group and policymakers motivated by climate action more broadly should consider wider, existing bottom-up data institutions[5] where communities are empowered around their data without the need for trustees with fiduciary obligations, developing new forms of data stewardship where needed and documenting existing forms of trustworthy practices where evident, whilst deepening community engagement as they do so.

Next steps: a deeper and broader exploration of climate migration

The Working Group is now carrying forward the recommendation to undertake a deeper, broader exploration of how data institutions and AI applications could make a difference on climate-induced migration, placing local communities at the heart of their governance.

Our primary focus is on **Lake Chad Basin** (Nigeria, Niger, Chad and Cameroon), a region grappling with a complex humanitarian crisis with over 3.2 million people displaced, largely because the lake has shrunk by 90% (scarce water supplies, food insecurity, degraded farmlands, farmers-herders conflicts).

We plan to engage with local organizations and consult with affected communities to co-design a framework for trustworthy data exchanges within the climate migration data ecosystem, improving how data is being collected, stewarded, shared and used to better serve their needs and empowering them to play an active role in the data value chain.

The research is currently being undertaken at the time of writing; we intend to share a final report and recommendations on this final phase of our work on data institutions this Winter.

---

[5] Examples are included under Section 2 of the report here

Advancing research and practice on data justice

Understanding Data Justice

Data Justice has been understood by the Working Group as fairness in the way that groups are made visible, represented, and empowered as beneficiaries in the collection and use of data for the development of AI systems. It promotes a broader lens than a narrower conception of data governance focused on compliance and individualised privacy; it does so by providing a framework to account for collective identities in view of the particular impact that AI decision-making systems can have at the community level.

The Working Group's goal and contribution

Since 2021, the Working Group has been working to help advance data justice research and practice, by seeking to advance the field from its theoretical foundations and towards practical contexts, in particular for three target audiences: policymakers, developer communities, and marginalised communities.

To support this work in its first exploratory phase, the Working Group partnered with the Alan Turing Institute and 12 'Policy Pilot Partners' from Low and Middle Income Countries spanning Latin America, Asia, Africa, and Oceania, representing one of the widest and deepest global surveys on data justice to date. This collaboration produced a deep body of work exploring the state of the art in data justice and offering practical guidance for the wider community, including:

- an Annotated Literature Review, Repository of Use Cases, Annotated Bibliography and Table of Organisations;

- guides for Impacted Communities, Policymakers, and Developer Communities;

- 12 Policy Pilot Partner reports, drawing from consultations from across Asia, Latin America, Africa, and Oceania

- a video guide: 'Introducing Data Justice'

An expansion of the work was supported by the UK Government Office for AI, with the outputs above launched at the Alan Turing Institute's AI:UK event in March 2022 by the UK Minister for Tech and the Digital Economy, and a live audience of 600 members of the AI community.

Since sharing this work, the Working Group is now pleased to share three further outputs for Summit 2022 as short, easily digestible resources:

(1) A Data Justice Policy Brief of 10 pages, condensing the greater body of work into a short, easily digestible, resource focused on clear actions for decision makers. The Working Group has developed this resource with the support of Research ICT Africa.

(2) Two short 'Primers' on Data and Social Justice, and Data and Economic Justice - again offering short introductions on these topics for policymakers considering data governance, with the economic primer produced in collaboration with IT for Change and social primer in collaboration with Research ICT Africa

Recommendations

The Working Group commends these resources as inputs for those in the AI community considering and undertaking reviews of their approach to data governance.

In particular, the Policy Brief and Primer have already started to help shape wider GPAI projects - including the Responsible AI Working Group's roadmap on AI and biodiversity (where indigenous

populations play a critical role in conservation) and our own Working Group's project on data institutions (with its current focus on refugee communities in the Lake Chad Basin).

Whilst we recommend reading the Policy Brief and Primers in full, we would like to draw attention to ten high-level policy recommendations. The Working Group would welcome approaches from policy makers interested in exploring these recommendations further, and thanks all our partners who have helped shape them, as well as the tireless commitment of the Working Group co-leads, Alison Gilwald and Dewey Murdick:

1. **Basing data regulation on rights:** Data today is so closely intertwined with our social and economic organisation and outcomes, that the need to base data regulation on human and community rights has to be recognised as a key political imperative. This requires developing actionable data rights frameworks, which include economic and collective rights to data. A basic data rights framework should include: the right to benefit from one's data, and to not be harmed by data collection and use; the right to access and port one's data; the right to appropriate representation in data, including to remaining invisible; the right to participate in governing one's data, and the data systems based on it; the rights to alternative and collective forms of data stewardship.

2. **Democratic participation of affected communities**: Data justice requires policymakers to identify the full set of stakeholders who might be impacted by data collection and use, and data-driven activities. Individual and collective data subjects, as well as primary data generators, are essential stakeholders. Their participation must be built-in democratically to the design, development and deployment of data-intensive systems, including AI.

3. **Contextualisation and localisation**: While larger frameworks of data rights, and data justice—and transversal laws and policies based on them—are necessary and useful, data justice also requires policymakers to move beyond one-size-fits-all solutions. Data justice when put into practice looks different in different places and contexts, and these nuances and differences should be locally developed through appropriate participatory exercises, and in consideration of the political economies and institutional endowments of countries.

4. **Equitable access to resources**: Material inequality and structural exclusion can prevent marginalised groups from sharing in the benefits of data-driven systems. To mitigate this, policy needs to ensure equitable access to skills development and digital infrastructure including connectivity and computing resources, as well as data assets—especially where communities have contributed to the generation of these data assets.

5. **Preventing anti-competitive data practices**: Trade and competition and antitrust law should be updated to respond to the role of data and platforms in economic structures. This includes preventing data hoarding, and requires data sharing to address unfair data practices by market gatekeepers.

6. **Enabling alternative forms of data sharing/stewardship**: Equitable access to data can be achieved through responsible data sharing models such as access to data, or data commons, in managed safe conditions. Vehicles such as data trusts or data cooperatives can be empowered to manage data in the collective interest.

7. **National data sovereignty and global governance**: Sovereignty as a right of a national community to manage its affairs, including its resources, is enshrined in various international human rights covenants. Data (and the ability to utilise it) is so central a resource to our social and economic organisation today that principles of national sovereignty must refer to data and data skills and infrastructure. This applies at personal, enterprise, community and national levels, but must also be informed by the needs of the global digital society and economy—particularly as some of the most intractable governance issues relating to the globalised and cross-border flows of data can only be addressed through international cooperation and solidarity.

8. **Workers' data rights**: Alongside the need for the expansion of employment and social protections to platform workers, regulation is needed to elaborate and advance workers' data rights. Specific data rights for workers should include limitations on the ability of platforms or employers to collect workers' data without their consent, and to monetise workers' data without ensuring workers share in its value, and possibly some participation in governing the corresponding data-based systems. In addition regulation should ensure that platform workers can own and port their work profiles, ratings and reviews off-platform.

9. **Transparency in data practices and systems**: Those with power in processes of collection use of data and data-driven innovation should be obliged to make information publicly available about what data is collected and how it is used, including information about AI/ML inputs, and algorithms, and to provide this information directly to impacted individuals and communities.

10. **Appropriate frameworks for redress**: Where data collection and use of data, and data-driven systems result in breaches of individual or collective rights, and injustices, there must be clear, institutionalised processes for individual and collective redress. These should be encoded in local, national and global legal frameworks, and place obligations on powerful data users.

Overcoming data barriers via trustworthy privacy-enhancing technologies

In 2022, the Working Group kicked-off a new technically-focused project seeking to explore possible applications of privacy-enhancing technologies (PETs+) in AI-for-social-good contexts, specifically across three priority topics: Better Health, Future of Work, and Climate Action. The Working Group has been supported in this work by Capgemini, and the Infocomm Media Development Authority[6] (IMDA), a Singapore government agency.

This project seeks to build upon the learnings of both the data justice and data trust projects. The primary objective is to increase the availability and/or usability of AI systems by providing a means to safely develop and use data sets while preserving privacy, sovereignty, IP rights, and security.

Setting the foundations for conducting a practical demonstration

In 2022, the Working Group focused on identifying and evaluating possible use cases across three topics: Better Health, Future of Work, and Climate Action. GPAI subject matter experts representing each area were interviewed to identify possible use cases that may exist within the topic's domain. Throughout the course of the interview process, over 20 use cases were identified as possible candidates for a PET demonstration project. Evaluation criteria were then developed to assess the use cases' true potential, falling into three categories:

1. Feasibility: is there an identified owner, a possible timeline, data availability, and what is its security and sensitivity?

2. Relevance and dissemination potential: does it address one or more of the elements of concern (privacy, sovereignty, etc.), does it have a positive societal impact, is it relevant across multiple geographies, and can it demonstrate processes and technologies at scale, with real data?

3. Exploitability and reusability: can it produce generalised and reusable knowledge, and can it relate to the topics of data trusts or data justice?

Upon evaluating the use cases against the gating criteria, it was shown that most use cases could not pass, primarily due to an issue of lack of ownership. While it was easy to identify "categories" of owners (e.g., a municipality, a hospital…) the lack of specific ownership was a common showstopper. Based on this finding, the use cases were re-categorized to help the Working Group prioritise those use cases with viable potential to be taken forward into additional planning stages:

- **Category 1**: Use cases that have existing owners AND can address one or more elements of concern limiting access to data

- **Category 2**: Use cases that have potential owners AND can address one or more elements of concern limiting access to data

- **Category 3**: Use cases that do not have potential owners at this moment OR cannot address one or more elements of concern limiting access to data.

The use cases which fell into Category 3 were deemed to fail the gating criteria, and were subsequently not pursued to be developed into a demonstration project use case. The results of the re-categorization, as outlined in a report[7] prepared by Capgemini, were as follows:

Category 1:

- Better Health - "Modelling the effects of human movement during a pandemic"

Category 2:

- Climate Action – "Hyper-personal journey planning"

- Climate Action – "Hyper-personal city planning"

- Future of Work – "Smart cameras for ethical surveillance/monitoring"

The Working Group then held a use case selection workshop, having an open discussion on each of the above-listed use cases, including their scope, performance against the original evaluation criteria, and their demonstration potential. Each attendee was then asked to rank their top two use case selections, which culminated in a group-level recommendation to the Project Steering Committee for which use case to pursue.

The top-ranked use case (and, use case selected by the Project Steering Committee for the demonstration) was "**Better Health – Modelling the effects of human movement during a pandemic**", for which the demonstration will integrate with the current Responsible AI Working Group's project on Immediate Response (now re-termed as Pandemic Resilience). This use case enriches existing pandemic models that use anonymised and aggregated data, by adding data that describes (in detail) an individual's movements. The connectivity of people, or their location information, could be derived from mobility data, inferred from the use of mobile devices, transactions at points of sale, or public transport. These data will be used to build an important input to the existing models, commonly referred to as the contact/transmission network. The Working Group envisions that PETs will enable the use of this individual-level mobilisation data without any violation of rights or privacy of those involved, and, hopefully improve the accuracy of the contact/transmission networks.

Similarly, it is worth noting that the second-ranked and third-ranked use cases, "**Climate Action – Hyper-personal city planning**", and "**Climate Action – "Hyper-personal journey planning**", respectively, also focus on the use of data describing an individual's movements. These two use cases assess the possibility of improving the effectiveness of urban planning activities, as well as minimising the environmental impact of private transport by promoting personalised/customised eco-friendly forms of commuting (ex: via public transport, carpooling, etc.). The Working Group envisions that PETs could play the same role for these two use cases as for the "Better Health - Modelling the effects of human movement during a pandemic" use case, enabling the use of this individual-level mobilisation data without any violation of rights or privacy of those involved, and, hopefully, improve upon existing methods of urban planning and commuting, further promoting eco-friendly modes of transport.

Next steps: preparing and conducting the PET demonstration

The Working Group has partnered with Singapore's Infocomm Media Development Authority (IMDA) to conduct the "**Better Health – Modelling the effects of human movement during a pandemic**" demonstration. Singapore's Digital Trust Centre (DTC) will act as the delivery partner, and scoping is currently ongoing to determine the precise data sources and models that will make up the demonstration. The next phase of the project is expected to kick off in October 2022, in close collaboration with the Responsible AI Working Group and Pandemic Resilience project teams.

Additionally, the Working Group will also explore the possibility of conducting a second demonstration for the "**Climate Action – Hyper-personal city planning**", and "**Climate Action – "Hyper-personal journey planning**" use cases. Ideally, conducting the two PETs+ demonstrations in tandem would allow for shared learnings between the two project streams, further enabled by the overlapping focus on individual-level mobility data. The Working Group endeavours to establish a partnership to conduct this scope moving into 2023.

Forward look

For 2023, we are excited to pursue our work on privacy-enhancing technologies and data institutions. We are also looking forward to starting two new projects based on our 2020 Applied Research Agenda.

The Working Group has proposed the following two new projects for 2023 subject to GPAI Council approval at the Tokyo Summit:

(1) **Formulating transnational legal guidelines governing rights in co-generated data and third-party data**
(2) **The role of government as a provider of data for AI.**

For "**Formulating transnational legal guidelines governing rights in co-generated data and third-party data"**, the Working Group plans to focus on data rights within the context of AI with regard to two contexts: (1) 'co-generated' data, for which principles have been designed to recognise many different players have contributed to the generation of data in many different ways; and (2) third-party data rights, where data rights are vested in parties that have not contributed, for example access to public sector data for commercial re-use.

This project would review how these concepts have been developed to date by existing initiatives, such as the recent EU Proposal for a Data Act, ascertain how far those initiatives address an AI-specific context, and what further protections (legal, technical and institutional) may therefore be required. As part of this project, collective data rights would be explored - and whether, and under what conditions, they should exist. As was noted in the Working Group's Framework Paper, there is a question as to whether rights in co-generated data should be vested not only in individuals but also in groups of individuals, such as defined by language (e.g for speech data), ethnic origin (e.g for genetic data), or activities (e.g data from connected vehicles).

The project seeks to contribute to the transnational convergence of national laws as regards the design of data rights and their allocation which is highly pragmatic. While it is true that globally jurisdictions rely on very different value judgments and objectives for regulating the digital sector, the project should rely on the values and goals of GPAI, i.e. human rights, inclusion, diversity, innovation, economic growth and societal benefits, as well as the UN Sustainable Development goals as guidance for the formulation of guidelines concerning data rights which aligns closely with GPAI's mandate. Thereby, beyond legal convergence as such, the project is designed to promote said values and goals. By also addressing the technological prerequisites of data sharing, the project would also help guarantee that data governance structures for the purpose of AI development promoting said values and goals will develop transnationally.

For "**The role of government as a provider of data for AI",** the Working group aims to support governments to make decisions about whether and how to share data they steward with AI developers. The intended impact is to increase the availability of publicly held data for AI grounded in the principles of human rights, inclusion, diversity, innovation and economic growth by helping governments to prioritize their efforts and to reduce their concerns about the risks of sharing public data for AI by providing clear guidance, use cases and examples that demonstrate how it can be done safely and responsibly. This project aligns closely with the GPAI's mandate whilst focusing on applied AI issues and aiming at practical resorts and assessments. It also responds to GPAI's criteria to identify key projects : it's impactful, practical and ambitious.

As noted earlier in this report, the Working Group will also focus on completing two of its current projects:

(1) For the **Privacy-enhancing technologies project**. The Working Group plans to demonstrate the viability of AI systems in helping achieve the UN SDGs in congruence with the OECD AI Principles, by providing a means to safely and securely develop, use and share data while preserving privacy, sovereignty and IP rights. The project also seeks to overcome challenges to data usability commonly faced when working with PETs by publishing practical guidance and lessons learnt from the demonstration system.

(2) For the **Data Institutions** workstream, our Working Group aims at moving towards practical interventions by taking a deeper exploration of how data institutions and AI applications could make a difference on climate-induced migration. Our focus is on Lake Chad Basin and we plan to engage with local organizations and consult with affected communities to help improve how data is being collected, stewarded, shared and used to better serve their needs.

We're thrilled by these projects' work and we're hopeful that our future research agenda will guide the next steps on opportunities to go further and deeper in advancing research and practice on data governance.

Participation across our Working Group is a big part of what makes these projects true international collaborations. We would like to invite those who are interested to make a personal contribution to these projects by joining our project steering groups to help shape direction, give feedback and review research. You can express your interest to contribute by connecting with the Montreal Centre of Expertise (the CEIMIA) at info@ceimia.org

Annex1

Project advisory group on Enabling data sharing for social benefit through data institutions

Co-Leads

Neil Lawrence – University of Cambridge
Seongtak Oh – National Information Society Agency
Teki Akuetteh Falconer – Africa Digital Rights Hub

Members

Jeni Tennison (Working Group Co-Chair) – Connected by Data
Maja Bogataj Jančič (Working Group Co-Chair) – Intellectual Property Institute
Matija Damjan – University of Ljubljana
Carole Piovesan – INQ Data Law
Kim McGrail – University of British Columbia
Bertrand Monthubert – Ekitia
Nicolas Miailhe – The Future Society
Paul Dalby – Australian Institute of Machine Learning
Christiane Wendehorst – European Law Institute / University of Vienna
Yeong Zee Kin – Infocomm Media Development Authority
Alison Gillwald – Research ICT Africa
Josef Drexl – Max Planck Institute
Alejandro Pisanty Baruch – National Autonomous University
Iris Plöger – Federation of German Industries
Ricardo Baeza-Yates – Universitat Pompeu Fabra & Northeastern University
Aleksandra Przegalińska – Kozminski University
Zümrüt Müftüoğlu (Observer) – Yildiz Technical University
Claudia Juech (Observer) – Patrick McGovern Foundation (Observer)

Invited specialists

Jess Montgomery – Data Trusts Initiative / University of Cambridge

Project advisory group on Advancing Data Justice Research and Practice

Co-Leads

Alison Gillwald – Research ICT Africa (South Africa / UNESCO)
Dewey Murdick – Center for Security and Emerging Technology (USA)

Members

Jeni Tennison (Working Group Co-Chair) – Connected by Data
Maja Bogataj Jančič (Working Group Co-Chair) – Intellectual Property InstituteTakashi
Kai – Hitachi
Shameek Kundu – TruEra
Nagla Rizk – American University in Cairo
Zümrüt Müftüoğlu – Yildiz Technical University
Teki Akuettah Falconer – Africa Digital Rights Hub
Te Taka Keegan – University of Waikato
PJ Narayanan – International Institute of Technology, Hyderabad (India)
Kim McGrail – University of British Columbia
Allan Feitosa – Summ.link
Anderson Soares – Artificial Intelligence Center of Excellence
Ricardo Baeza-Yates – Universitat Pompeu Fabra, Spain & Northeastern University, USA
Nicolas Miailhe – The Future Society
Bertrand Monthubert – Ekitia
Jaco Du Toit (Observer) – UNESCO

Invited specialists

Kelle Howson – Research ICT Africa
Parminder Jeet Singh – IT For Change
Anita Gurumurthy – IT For Change

Project advisory group on Privacy-Enhancing Technologies

Co-Leads

Kim McGrail – University of British Columbia
Shameek Kundu – TruEra

Members

Jeni Tennison (Working Group Co-Chair) – Connected by Data
Maja Bogataj Jančič (Working Group Co-Chair) – Intellectual Property InstituteBertrand
Monthubert – Ekitia
Christiane Wendehorst – European Law Institute / University of Vienna
Dewey Murdick – Center for Security and Emerging Technology (USA)
Lee Tiedrich – Duke University
Michael Justin O'Sullivan – University of Auckland (New Zealand)
Nicolas Miailhe – The Future Society
Robert Kroplewski – Minister for Digitalisation of the Information Society (Poland)
Zümrüt Müftüoğlu (Observer) – Yildiz Technical University

Invited specialists

Adhiraj Saxena – IMDA
Chein Inn Lee - IMDA
Wan Sie Lee – IMDA