IGF2020

Best Practice Forum

Data and New Technologies in an Internet Context

'A dialogue on how users' data is collected, analysed and used,

and best practices

to ensure that this data is used to bring benefit and not to harm users.'

Draft BPF Report

October 2020

Acknowledgements

The *Best Practice Forum Data and New technologies in an Internet Context* is a community driven multistakeholder intersessional activity of the *Internet Governance Forum (IGF)*. This report is intended as input for the BPF session during the virtual Internet Governance Forum (IGF 2020) . The report is the result of the collaborative work of many involved in the BPF's discussions and contributing to its calls for input.

The BPF Data and New Technologies is inviting Community feedback on this draft report!

How?

Please send your feedback to **bpf-IoT-BD-AI@intgovforum.org**

Format?

Feedback can be sent in an email or as a word or pdf document attached to an email.

If a comment is on a specific section or paragraph, please indicate this clearly.

Deadline?

It is possible to submit feedback on this document until the last day of the IGF2020 meeting. However, we would appreciate your feedback by Tuesday 10 November, as this would allow us to take your feedback into account during the BPF workshop during the IGF meeting.

Publication?

Received feedback will be posted on the BPF webpage (link) - unless the author indicates that he/she prefers that it is not published - and feed into the final BPF output report.

Table of contents

I. Introduction

1. The Internet Governance Forum

The *Internet Governance Forum (IGF)* is a global forum, convened by the United Nations Secretary General[1], where governments, civil society, the Internet technical community, academia, the private sector, and independent experts discuss Internet governance and policy issues.[2]

The *IGF 2020*, the fifteenth annual meeting of the IGF, and due to the COVID-19 pandemic the first virtual IGF, is hosted by the United Nations Department of Economic and Social Affairs (UN DESA) from from 2 to 17 November 2020 under the overarching theme 'Internet for human resilience and solidarity'.

2. IGF Best Practice Forums

The IGF *Best Practice Forums (BPFs)[3] [4]* provide a platform for experts and stakeholders to exchange experiences in addressing Internet policy issues, and to discuss and identify emerging and existing good practices. BPFs are expected to be open, bottom-up and collective processes, and their outputs to be community-driven.

BPFs prepare their work in a series of intersessional discussions that culminate in a BPF session at the annual IGF meeting and a report published as part of the IGF outputs.

The objective is to collect from community experience, not to develop new policies or practices. BPF outputs intend to contribute to an understanding of global good practice, and to serve as a resource to inform policy discussions, standards development, business decisions, as well as public understanding, awareness, and discourse.

---

[1] The [resolution adopted by the UN General Assembly on 16 December 2015, (70/125)](#) 'Outcome document of the high-level meeting of the General Assembly on the overall review of the implementation of the outcomes of the World Summit on the Information Society', extended the mandate of the IGF as set out in paragraphs 72 to 78 of the Tunis Agenda.

[2] IGF website: http://www.intgovforum.org - the IGF is one of the key outcomes of the World Summit for the Information Society (WSIS).

[3] BPFs were re-introduced in 2014 as part of the intersessional programme to complement the work of the IGF community's activities and develop more tangible outputs to 'enhance the impact of the IGF on global Internet. This intersessional programme was designed in accordance with the recommendations of a 2012 report by the Commission on Science and Technology for Development (CSTD)'s Working Group on IGF Improvements.
https://www.intgovforum.org/multilingual/index.php?q=filedepot_download/4586/588

[4] BPF Outputs and activities are archived on the IGF webpage:
https://www.intgovforum.org/multilingual/content/bpfs-outputs .

3. The IGF 2020 Best Practice Forums

BPFs are organised under the supervision of the *IGF Multistakeholder Advisory Group (MAG)* and receive substantive IGF Secretariat support. The MAG selected the following four topics for the 2020 BPF cycle[5]:

- Data and New Technologies in an Internet Context[6]

- Exploring Best Practices in Relation to International Cybersecurity Agreements

- Gender Impact on Shaping Internet Policy

- The Protection, preservation and remuneration of creative work and collective wisdom from a local content perspective.

---

[5] https://www.intgovforum.org/multilingual/content/best-practice-forums-bpfs
[6] MAG agreed Proposal BPF Data and New Technologies in an Internet Context
https://www.intgovforum.org/multilingual/filedepot_download/9655/1965

II. Data and New Technologies in an Internet Context

1. A dialogue on how users' data is collected and used

The Internet connects people, connects devices, and connects people and devices. The fast-growing connected world generates a vast variety and high volume of data from devices, networks, platforms and applications used by billions of users around the world. This data is being used in multiple ways and expectations run high about how data can and will fundamentally change and improve people's daily lives, and how data contributes to solving complex problems, addressing societal issues, and facing global policy challenges.

The number of applications that make use of new technologies to collect and analyse data is booming and will continue to grow exponentially in the coming years and decades. It doesn't take long to find examples of applications that use data and new technologies to facilitate people's lives, or help to address environmental, healthcare, transportation, and many other challenges. These applications may improve existing solutions, make them more efficient, or make it possible to approach issues in a totally new and more effective way. They can also empower people who today, for a variety of reasons, may have limited possibilities to act or influence.[7]

There's a vast variety of data and information that can be collected in the connected world. This BPF is focussing on users' data generated and collected in an Internet context. The data helps making progress in addressing environmental, healthcare, transportation, and many other challenges. The user will, directly or indirectly, benefit as the data makes life more convenient or contributes to creating a better world. At the same time, however, concerns are growing about how the same data can be used and abused by companies and governments, in ways that intentionally or unintentionally put users in a weaker and powerless position or even harm them. Therefore, the BPF on Data and New Technologies in an Internet Context intends to establish a dialogue on how users' data is collected, analysed, and used, and best practices to ensure that this data is used to bring benefit and not to harm users.[8]

The BPF on Data and New Technologies in an Internet Context

aims to establish

a dialogue on how users' data is collected, analysed, and used,

and best practices to ensure that this data is used to bring benefit and not to harm users.

---

[7] The 2019 BPF IoT, Big Data, AI showcased examples and opportunities of how data and new technologies may contribute to achieving the SDGs, improving cybersecurity, making the Internet more useful and accessible to people, addressing challenges in the field of urbanisation, civil protection, security, financing, crisis management, etc.
https://www.intgovforum.org/multilingual/filedepot_download/8398/1915
[8] Proposal for a BPF on Data and New Technologies In an Internet Context,
https://www.intgovforum.org/multilingual/filedepot_download/9655/1965

2. Context and methodology

The BPF developed and carried out its work plan through a series of virtual meetings[9], online discussions on the BPF mailing list[10] and a session as part of IGF 2020 virtual meeting[11]. All meetings were open to all interested in the topic as was subscription to the dedicated BPF mailing list.

As part of its activities, the BPF compiled a 'Data and New Technologies Issues Card' which was published for community review and feedback on the IGF website.[12]

The BPF launched a 'Call for Case Studies'[13] via the IGF website to collect examples of how stakeholders address the challenges that emerge when collecting and using users' data. Examples could relate to applications (deployed or under deployment) that use users' data or frameworks, guidelines and policies that aim to ensure that users can benefit and not risk to be harmed. The call explicitly mentioned that the BPF was interested in case studies related as well as unrelated to the COVID-19 pandemic.

A draft version of this output report was published on the IGF website on 30 October 2020 as part of the BPF inputs for IGF 2020, and the community was invited to review and provide feedback.

The BPF will organise a session at the virtual IGF 2020 on Thursday 12 November at 9:00-11:30 UTC[14]. Substantive input from this session will be reflected in the final version of this report.

Throughout its work the BPF built on the work of the BPF on the Internet of Things (IoT), Big Data, and Artificial Intelligence (AI). In 2018, the BPF IoT, Big Data, AI, identified best practices to facilitate stakeholder dialogue on issues pertaining to IoT, Big Data, AI in an Internet context[15], and in 2019 it discussed best practices to face the policy challenges that arise when using IoT, Big Data, AI to contribute to solving societal challenges[16].

---

[9] The BPF had its kick-off meeting on 20 February 2020 and met regularly through to end October 2020. Summaries and recording of the meetings are archived on the BPF webpage https://www.intgovforum.org/multilingual/content/best-practice-forum-on-data-and-new-technologies-in-an-internet-context .

[10] BPF mailing list https://intgovforum.org/mailman/listinfo/aiiotbd_intgovforum.org

[11] BPF Data and New Technologies, 12 November 2020, https://www.intgovforum.org/multilingual/content/igf-2020-bpf-data-and-new-technologies-in-an-internet-context

[12] The BPF Data and New Technologies Issue Card was published on 9 July 2020 for community feedback at https://www.intgovforum.org/multilingual/content/data-new-technologies-issues-card .

[13] https://www.intgovforum.org/multilingual/content/bpf-data-new-technologies-case-studies

[14] Session details at https://www.intgovforum.org/multilingual/content/igf-2020-bpf-data-and-new-technologies-in-an-internet-context

[15] Output report of the 2018 BPF IoT, Big Data, AI, https://www.intgovforum.org/multilingual/filedepot_download/6733/1438

[16] Output report of the 2019 BPF IoT, Big Data, AI, https://www.intgovforum.org/multilingual/filedepot_download/8398/1915

3. Data in the year of the global COVID-19 pandemic

Data and statistics play an important role in the fight against the global COVID-19 pandemic. Around the world connected technologies are being used to collect information on the spread of the virus as part of strategies to monitor and manage the crisis. Different initiatives and applications are being developed, tested and launched to help governments and health services, academics and industry with mapping, fighting and predicting the evolution of the COVID-19 pandemic.

The COVID-19 crisis has led to an increased and intensive use of smart-working and other apps and solutions that collect and share data. Some applications intensively collect personal data, other initiatives do not rely on personal data. As a side effect, the pandemic has so far been an important accelerator for discussions on the challenges and the boundaries of using online generated users data for public policies.[17]

However, many of the concerns and challenges that are expressed in relation to the COVID-19 apps are the same or similar regardless of whether data is collected and used to fight COVID-19 or the data is collected and analysed for other purposes. Therefore, the BPF intends to collect case studies related as well as un related to the COVID-19 pandemic. In addition, as part of its discussion at IGF 2020, the BPF intends to explore whether the pandemic has changed views and concerns related to the use of users' data.

4. Importance of clear definitions and concepts

Data?
Data is a broad concept that covers a vast range of types of data and information, both in the on- and offline world. This BPF report is limiting its focus to users' data generated and collected in an Internet context. Yet, this 'users' data' and related concepts frequently used in discussions on the topic, are often used without specifying them. This may create confusion or result in discussions going into different directions instead of maintaining focus.[18]

The BPF IoT, Big Data, AI in 2018 discussed best practices to foster and facilitate multistakeholder discussion on issues related to the use of the new technologies and recommended to define terms narrowly so that it is clear for policy makers and stakeholders what aspects of a technology they are discussing. Not doing so can lead to sweeping generalizations and measures that come with a range of unintended consequences.

Also, policy discussions should strive to be technology-neutral. Because technologies are changing quickly and potential problems may or may not develop (or may be solved rapidly), it is dangerous and may prove unproductive to try to write laws and regulations that cover one specific type of technology or one specific type of application. Best practice is usually to focus on what an application

---

[17] BPF Data and New Technologies, Virtual Meeting I, 20 February 2020, meeting summary, https://www.intgovforum.org/multilingual/filedepot_download/9655/2043
[18] BPF Data and New Technologies, Virtual meeting II, 17 April 2020, meeting summary, https://www.intgovforum.org/multilingual/filedepot_download/9655/2069

DOES not on how the technology DOES IT.[19] It was felt that these recommendations are also valid for discussions in the context of the current BPF on data and new technologies.

> *Be aware of terminology and concepts.*
>
> *Do not hesitate to pause and clarify definitions.*
>
> *It will help to avoid confusion and maintain focus.*

In many discussions on data and new technologies terms and concepts are used without specifying what they refer to. This can create confusion and misunderstandings, and participants involved in a discussion can be talking about different things without realising it. To illustrate this, BPF participants, after a short brainstorming, compiled a list of key concepts often heard in discussions related to data and new technologies, and possible meanings. The list is not exhaustive, and is intended as a warning to raise awareness of the potential issue. Stakeholders or any other party should not hesitate to pause and clarify terminology.

> *Terms and concepts are often not well defined what can lead to misunderstandings, confusion and ineffective discussions related to data and new technologies.*
>
> *To illustrate this, the BPF listed possible interpretations of frequently used terms.*
>
> Data
> - Data generated by devices
> - Data intentionally collected through technology
> - Data collected from surveys and other means
> - Internet data
> - User's data
>
> User's data
> - Data on the user's (online) activities
> - Personal identifiable information
> - Personal data
> - Private data
>
> Internet data
> - Data collected through controlled measures such as online surveys and cookies
> - Data collected through web scraping
> - Data resulting from the an analysis of the online behaviour of connected people
>
> The Role of the Internet
> - The Internet as a TOOL to enable data collections
> - The Internet as MEDIUM through which data is retrieved
> - The Internet as TECHNOLOGY to obtain data
> - The Internet as STORAGE for large amounts of data

---

[19] IGF 2018 BPF IoT, Big Data, AI, output report, December 2018, https://www.intgovforum.org/multilingual/filedepot_download/6733/1438 .

Data governance?

Data governance, a term used by some to refer to how data is handled or to a set of rules for how this is done, requires some caution. There exist different definitions and understanding of what data governance is and is not. For example, data governance can, depending on the context, refer to:
- A the set of rules and procedures to collect, store, process and share data;
- A company's or organization's rules and procedures to manage internal data flows;
- Rules and techniques used to statistically validate and clean up data;
- Government policies to, for example, safeguard privacy, restrict the flow or data, or censor data.

BPF Participants provided examples of how data governance is used with a different meaning:

- Wikipedia[20] notes that '*Data governance is a term used on both a macro and a micro level. The former is a political concept and forms part of international relations and Internet governance; the latter is a management concept and forms part of corporate governance'.*

- Data was one of the main themes of the IGF 2019[21] and is a one of the four thematic tracks of the IGF 2020[22]. The theme's narrative contains the following description for the data governance track: '*The Data Governance track will provide for discussions on the fundamental challenge of ensuring the benefits of the data revolution to contribute to inclusive economic development while protecting the rights of people.*

  The global nature of the Internet and the transfer of digital information across borders brings an international dimension to discussions around data. The generation, collection, storage, transfer and processing of data (including personally identifiable data) have enabled new social, cultural, and economic opportunities than ever previously imagined. At the same time, the massive collection, transfer and processing of data (in particular through the application of algorithms/AI/machine learning) by public as well as private entities pose challenges around privacy, freedom of expression and the exercise of other human rights.

  The Data Governance track will contribute to identifying best approaches to ensure the development of human-centric data governance frameworks at national, regional and international levels. It will enable an exchange of views on how to support and operationalize the exercise of human rights and the empowerment of individuals in their digital identity in current uses and development of data-driven technologies. And it will consider how to create the conditions needed to facilitate data-driven innovation, to ensure competition, and to foster trust in the development of services and new technologies, including through the use of inclusive data and the fulfillment of the UN's 2030 Agenda for Sustainable Development.'

---

[20] https://en.wikipedia.org/wiki/Data_governance
[21] https://www.intgovforum.org/multilingual/content/igf-2019-themes
[22] https://www.intgovforum.org/multilingual/content/igf-2020-thematic-tracks

- Article 4 of the GDPR[23] Regulation describes what activities, within the scope of the regulation, are considered processing of data: *(2) 'processing' means any operation or set of operations which is performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;*

- Data collaboratives[24] - *Data Collaboratives are a new form of collaboration, beyond the public-private partnership model, in which participants from different sectors - in particular companies - exchange their data to create public value.*

Some expressed the concern that a too narrow definition of data governance risks ignoring several worries people have about their data and how it is handled, while others warned that data governance can be too blurry as a concept. Some BPF participants expressed a preference for alternative concepts such as data processing, practices related to data, data operation policies, data policy, or data collaboratives.

What is a 'best practice' ?

On Wikipedia[25] a best practice is defined as '*a method or technique that has been generally accepted as superior to any alternatives because it produces results that are superior to those achieved by other means or because it has become a standard way of doing things, e.g., a standard way of complying with legal or ethical requirements.*'

It was proposed to define a best practice within the scope of the BPF as '*a practice that helps to ensure that data is used to bring benefit and not to harm users*'. Some commented that this is too narrow, focuses on practical issues and does not sufficiently include policy best practices (e.g. how governments deal with data). Others remarked that practices often have different sides - what is good for one, might harm another - and that expecting that a 'best' practice benefits all, without considering possible trade-offs, is a too idealistic approach. A consensus was found in the definition of a best practice in the context of this BPF as a data processing activity that is less harmful to individuals and their liberties and compliant with applicable data protection frameworks and principles. [26]

In the context of this BPF, a best practice is defined as follows:

'a best practice is a data processing activity that is less harmful to individuals and their liberties and compliant with applicable data protection frameworks and principles'.

---

[23] EU Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data; https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN
[24] https://datacollaboratives.org
[25] https://en.wikipedia.org/wiki/Best_practice
[26] BPF Data and New Technologies Virtual Meeting III, 8 May 2020, https://www.intgovforum.org/multilingual/filedepot_download/9655/2096

4. Data and New Technologies Issues Card

The BPF developed a Data and New Technologies Issues Card, which maps potential issues and challenges related to the use of users' data by new technologies.

The Issues Card is intended as a tool to foster discussion on data and new technology applications. It provides questions that stakeholders can use when assessing how data is collected, analysed and used and to question what decisions and choices are made to ensure that the data is used to bring benefit and not to harm users.

The Issues Card is a tool, framework or checklist that can be used by all kinds of stakeholders to structure their discussions on the subject matter. The Issue Card raises awareness about potential issues and challenges and could also help individual users to exercise caution and ask critical questions about what happens with their data when using an application or tool, when sharing their data and clicking on a consent button.

Note:
The Issues Card is an evolving document. We still appreciate your feedback and suggestions:
https://www.intgovforum.org/multilingual/content/data-new-technologies-issues-card .

Data & New Technologies Issues Card

Introduction:

The Issues Card is developed by the [IGF Best Practice Forum on Data and New Technologies in an Internet Context](). It maps potential issues and challenges related to the use of users' data by new technologies, formulated as four sets of questions.

How to use the Issues Card?

The Issues Card is intended as a tool to foster discussion on data and new technology applications. It provides questions that stakeholders can use when assessing how data is collected, analysed and used and what decisions and choices are made to ensure that the data is used to bring benefit and not to harm users.

Data collection

- What type of data is collected ?

- How is the data collected ?

- Where is the data collected / where does it come from?

- Is collected data anonymized? Aggregated?

- What is the legal basis used for the data collection?

- Are data producers required to give their explicit consent?

- Can individuals request to delete, correct or update his/her/their already collected data and what is the procedure?

- In what context are the data collected? (e.g. democratic and free society vs repressive system)

- How is 'data in transit' protected? (e.g. what security measures or security framework is in place?; what measures avoid leaking?)

Data storage

- How/where is the data stored? (e.g. centralised or decentralised?)

- How is the data protected? (e.g. what security measures or security framework is in place?; what measures avoid leaking?)

- How long is data kept? what after this period? (e.g. is data automatically deleted?)

- Do data producers have sovereignty over their data? (e.g. are they able to delete their data directly or demand that it is deleted upon their request ('right to be forgotten')

Labelling and unlocking value

- What purpose is the data serving? Is this useful? Necessary?

- Was the purpose under which the data was collected specified and legitimate?

- Who decides which are legitimate purposes and aims of data collection? (e.g. democratic decision-making processes versus top-down approaches)

- What measures are taken to assure that data is not used for purposes different than the ones that were specified on initiating the collection?

- How is the quality of data assured? How is bias avoided?

- What duty do policymakers (on national and international levels) have to make sure that data collection purposes meet certain legal and ethical standards, including human rights?

Data sharing

- How can the data be shared in order to create value for all parties?

- What model(s) of data sharing is (are) used?

- Who is benefiting from the data sharing?

- What restrictions apply to sharing the data ? (e.g. data localisation laws)

- Is the data producer getting any benefit from its data?

- Can the data be reused (for similar or different purposes?)

- To what extent is the data shared with a government and third parties?

- What possibilities do individuals (the people whose data are being collected) have trace the sharing of their data to other parties?

5. Addressing the challenges - case studies and best practices

The BPF launched a Call for Case Studies and examples of how stakeholder address the challenges that emerge when collecting and using users' data in an Internet context. The case studies are expected to help with identifying existing practices that contribute to ensuring that users' data is used in such a way that the users benefit from their data and that their data is not used to harm them.

The BPF looked for examples of

- applications (deployed or under development) that use users' data to provide benefits to the users and the measures that should avoid that the collected data may be used to harm users;

- frameworks, guidelines, and policies that address challenges and aim to ensure that users whose data is being collected and used can benefit from their data and do not risk to be harmed.

The BPF invited case studies related as well as unrelated to the COVID-19 pandemic, and asked submitters to refer in their case studies to the challenges identified on the Data and New Technologies Issue Card.

Received case studies are summarised on the following pages and archived on the BPF webpage.

Note:

The BPF still welcomes additional case studies and examples of how stakeholders address challenges related to the collection and use of users' data.

For instructions, check the Call for Contributions at
https://www.intgovforum.org/multilingual/content/bpf-data-new-technologies-case-studies .

Overview of Case studies

<div align="center">

Case study

Microsoft's privacy framework

</div>

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2266

Microsoft's approach to privacy is grounded in the belief that, for technology to succeed, people need to be in control of their data, and be empowered with information that explains how their data will be collected and used. Furthermore, companies need to be accountable and responsible for this data.

Six key privacy principles:

- ❏ **Control:** We will put our customers in control of their privacy with easy-to-use tools and clear choices.

- ❏ **Transparency:** We will be transparent about data collection and use so our customers can make informed decisions.

- ❏ **Security:** We will protect the data entrusted to us through strong security and encryption.

- ❏ Strong legal protections: We will respect local privacy laws and fight for legal protection of our customer's privacy as a fundamental human right.

- ❏ **No content-based targeting:** We will not use a customer's email, chat, files or other personal content to target ads to them.

- ❏ **Benefits to the user:** When we do collect data, we will use it to benefit our customer and to make their experiences better.

More information, including an explanation of the different types of data that Microsoft collects, can be found at https://privacy.microsoft.com/en-us/privacy.

Case study

Dos and Don'ts for COVID-19 contact tracing apps (AccessNow)

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2269

Contact tracing is important to monitor and manage the spread of viruses. The COVID-19 pandemic prompted multiple initiatives around the world, to facilitate and improve this traditionally resource and time intensive process, by developing tracing and reporting apps that make use of new technologies.

In May 2020, AccessNow published a list of 8 dos and 7 don'ts for COVID-19 contact tracing Apps. These are aimed at governments, companies and scientists involved in developing and using the applications, and takes into account concerns such as their effectiveness and impact on human rights.

Do's	Don'ts
❏ Voluntary use and sunset policy clause	Sell or monetize the data
❏ Privacy, data protection, and security by design and ensure accessibility	Apps that give users access to, or enable sharing of, personal information about other users
❏ Informative user experience	Apps that interfere with or reconfigure devices automatically
❏ Open source protocols and code available for auditing	Apps that enable unauthorized disclosure of data
❏ Adhere to data protection laws	Allow repurposing of the the app or data
❏ Limit data storage	Enable or allow targeted ads
❏ Decentralization protocols	World with companies with a track record for facilitating human rights abuse
❏ Ensure clarity, transparency and remedy for potential breaches or abuse	

https://www.accessnow.org/privacy-and-public-health-the-dos-and-donts-for-covid-19-contact-tracing-apps/

https://www.accessnow.org/privacidad-y-salud-publica-que-hacer-y-que-no-en-el-desarrollo-de-aplicaciones-de-rastreo-de-contactos-para-combatir-el-covid-19/

Case study

Microsoft's privacy principles related to the collection of data to tackle COVID-19

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2266

In relation specifically to COVID-19, Microsoft shared seven privacy principles to consider as governments, public health authorities and companies develop and implement technical solutions to address the COVID-19 pandemic. These principles are designed to apply to any COVID-19 technological solutions that involve the collection and use of personal data such as health data, precise geolocation data, proximity or adjacency data, and identifiable contacts.

> Principles:
>
> ❏ Obtain meaningful consent by being transparent about the reason for collecting data, what data is collected and how long it is kept.
>
> ❏ Collect data only for public health purposes.
>
> ❏ Collect the minimal amount of data.
>
> ❏ Provide choices to individuals about where their data is stored.
>
> ❏ Provide appropriate safeguards to secure the data.
>
> ❏ Do not share data or health status without consent, and minimize the data shared.
>
> ❏ Delete data as soon as it is no longer needed for the emergency.

More detail about these principles and Microsoft's approach to privacy in the context of COVID-19 can be found in an April 2020 blog from Microsoft's Chief Privacy Officer at https://blogs.microsoft.com/on-the-issues/2020/04/20/privacy-covid-19-data-collection/.

Case study

CLAIRE Task Force on AI and COVID-19

___

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2268

In March 2020 CLAIRE (Confederation of Laboratories in Artificial Intelligence Research in Europe) launched a volunteer effort to help tackle the pandemic and established a task force to coordinate the contribution of the AI experts supporting it. This task force collected information on leveraging AI techniques in the context of COVID-19 and supported the development of new projects, connecting the European network of AI experts together with health institutions and governments. By the end of March, the task force had enrolled 150 volunteers, covering the full spectrum of AI methods, tools and technologies.

The task force was able to assemble a vast catalogue of AI & COVID-19 related resources covering Funding opportunities (21 calls), Datasets (30 resources), Hackathons challenges and webinars (7), and other initiatives (71). Moreover 7 thematic groups of volunteers achieved important outcomes in several areas, from bioinformatics to robotics, and some of these outcomes have been already released to the public.

Findings and observations from CLAIR research groups

❏ **Data collection:** Almost all the research groups had difficulties in collecting data due to the privacy regulation, missing data collection policies, lack of an integrated data acquisition and storage pipeline at health institutions and at governments level.

❏ **Data Storage:** We identified the lack of a common European Cloud to share in a safe, legal and interoperable way health data for research purposes. Moreover, the few already existing public-private partnerships (e.g. to collect data on medical imaging) lacked transparency on the management of the data and future accessibility of products derived from such (eg. whether derived models will be publicly available).

❏ **Labelling and unlocking value:** We identified the difficulty to have enough labelled data to present to AI systems when a new dynamic and volatile non-deterministic environment emerges such as a pandemic. This means that we need to start addressing the development of AI self-learning and adaptation to an unlabeled world with the ability to self-build knowledge to represent a particular situation. Furthermore, in emergency, allowing secondary uses for already collected data can be beneficial. Clear and expedite ways for granting for exceptions (and oversight mechanisms) could be considered.

More information: https://covid19.claire-ai.org/ .

Case study

Setting out principles and launching initiatives to harness the power of open data (Microsoft)

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2266

In April 2020, Microsoft announced the launch of the Open Data Campaign to close the "data divide" and ensure that organizations of all sizes have access to the data they need to innovate with artificial intelligence (AI). To demonstrate the importance of being more open with data and the need to share data to address pressing issues, we committed to the development of 20 data collaborations by 2022. Through these collaborations, we will work with partners to address issues that are "top of mind" and require urgent action.

When armed with the right data, organizations are empowered to make decisions that positively impact their employees, customers and the communities they serve. In many ways, data is taking center stage in the response from governments and companies to fuel promising solutions and ideas. The central information point for all of Microsoft's work on open data is news.microsoft.com/opendata.

Microsoft is collaborating in three open data trials with various partners in the areas of climate change, COVID- 19, and digital access and education.

- **Climate change**: On September 1, Microsoft joined with Allianz, Amazon and S&P Global to announce plans to launch the **Climate Finance Foundation**, led by the Linux Foundation. The ability to leverage high-quality, open, corporate sustainability data will be critical for enabling the investment community to make informed decisions based on accurate and reliable economic models around corporate climate- related risk and opportunity. Microsoft is investing heavily in sustainability, and we have made a commitment to share relevant sustainability data to the open Data Commons supported by this effort.

- **COVID-19**: The Alan Turing Institute partnered with the Greater London Authority, supported by Microsoft and the London Data Commission, to demonstrate the value of data sharing to help support London's response and recovery to COVID-19. This pilot looks at London's "busyness" – or movement around the city – through multiple data sources as COVID-19 restrictions are relaxed to monitor how people are responding to the changes. Microsoft provided Azure AI and cloud infrastructure and services to support the COVID-19 pilot. Transport for London is already operationalizing the outputs from this pilot.

- **Digital access and education**: An initial priority of Microsoft's Open Data Campaign was to work with the Open Data Institute and BroadbandNow to help address the issue of broadband availability. With COVID-19 taking a significant toll on students' ability to access face-to-face education, this need is more urgent than ever. As governments, policymakers, nonprofits, and organizations around the world are looking at ways to target resources that serve students more effectively, Microsoft will be launching
an Open Data Challenge in October 2020 to look at the impact of digital access and COVID-19 on young students' education.

Case study

IoT-based scenario case study on the collection and use of users' data and best practices (China-FIOT Open Lab; Fujian Normal university)

Link: https://www.intgovforum.org/multilingual/filedepot_download/9655/2249

The case study highlights current weaknesses in some IoT-based applications (for example due to technical incompatibilities or 'grey' blind spot areas), explaining how users' data could be compromised despite conscious attempts by the data collectors to maintain data privacy.

Example 1: Centralized data conduits by providers

Today, many people trust centralized services that are backed by state-owned or authoritative institutions (*e.g.* banks, telecom operators). However, there is a tradeoff between centralization and market power -- inherent in this trust, people have also ceded their data privacy and bargaining power to these central providers [2]. In the burgeoning IoT industry, central providers *e.g.* telecom operators and local wireless access network providers, offer subscription access to wireless internet (4G, NB-IoT), as well as cloud-based management platforms.

One of the challenges inherent in the use of centralized service in the possibility of censorship or data monitoring by the service providers. Within China, applications such as Facebook, Twitter, Google cannot be accessed, partly due to IP addresses blocked by the telecom operators. In Singapore, the Infocomm Media Development Authority regulates websites that are flagged as illegal vice, deviant pornographic practices *etc.*; again, network operators are sanctioned to intercept messages to assist law enforcement, judicial or government agencies [3]. Another common practice in the IoT industry is the unnecessary routing of data uploaded by IoT devices to an additional intermediary, before the data is finally routed to the end goal server/cloud platform. This intermediary can be a middle-man platform that simply exists because users are using a service by a specific provider; an analogy to traditional systems would be the "calls dispatch centers". Once user volume is accrued, these intermediaries can become powerful inherent to their access to users' data and the analysis of big data derived thereof, and data privacy can quickly become an issue.

The fine print clauses in user agreement essentially dictate users to agree to the providers' terms if they want to continue to use the services. While user agreements are a socially responsible tool for informing users, in reality users may find it difficult to understand the legal jargon, may not have time to read through the extensively long document, or find themselves in the "all or nothing" situation where users' need to use the service override their discomfort in how their data privacy are handled.

Example 2: State organized expert database

An interesting case study is the "Special Commissioned Scientific Experts (科技特派员Ke Ji Te Pai Yuan)" database, first initiated in 1999 by the government in Nanping city, Fujian province, P.R. China. The intention is to select and dispatch scientific experts, with the goals of using new technology to rejuvenate rural villages and alleviate poverty.Gallup World Poll conducted on citizen trust in governments reveal that confidence in national government in BRIICS countries is relatively higher than OECD countries [4].

As such, this expert database hold authority and esteem both from political and technical perspectives. As of 2016, 729,000 technical persons are listed as experts [5], deployed in 51,400 for-profit agriculture-related entities, serving 60 million farmers, helping to bootstrap 15,900 new enterprises and establish 16,000 "Scientific Experts" workstations [6].

There is a public database featuring the scientific experts' background, expertise area, phone contact number and a work diary recording technical interactions with local farmers. The database manager/provider has attempted to maintain the scientific experts' privacy by representing the phone number in the database with a series of "X"s. However, if a user tries to call the expert's phone number, the actual string of numbers (expert's actual phone number) will then appear on the mobile phone unit's local Android or Apple iOS system. This weakness in the system can be attributed to a common problem – a technical issue of compatibility of data protocols between different systems and how this affects data hand-over between IoT devices and systems.

Current problems

In the first example, users' data privacy can be compromised if there are insufficiencies in policies to govern how centralized service providers, government agencies, regulatory bodies, and intermediaries in the data transmission chain access users' data.

In the second example, data protocols between two different systems are incompatible, resulting in insecure data handover between the Experts Database and local phone operating system. Even though the database manager/provider has taken care to anonymise the scientific expert's phone number, this phone number is still exposed during a call. The database manager-provider and phone operating system companies have the issue of data not in sync.

Best Practices

(the best practices are in the form of practical solutions, both currently deployed and under research)

❏ There is a need for **social responsibility of data collectors** to ensure users' data privacy and data usage within permission limits. These centralized providers and/or intermediaries should offer open channels of communication and feedback, actively listen to users and respond to their feedback. Furthermore, the users' agreement and policies stating how data will be used, should be open for public inspection. The feedback mechanism need to work both ways for a successful outcome; users should also play an active role in providing feedback and defining who access which parts of their data.

❏ There is a need towards **more unified data protocol standards**, especially in the highly fragmented IoT industry. Vendors of IoT devices, systems and databases each develop their own data transfer protocols, so it is common to experience incompatibility of data transferred from one system to another. Ideally a neutral, independent party like FIOT Open Lab, can lead industry efforts to standardize IoT security and data protocols, however in reality, it is often difficult to achieve consensus amongst the many stakeholders involved.

❏ There are **emerging technological solutions** such as blockchain that can increase trust between users and the service providers, as well as allowing users to pre-determine permissions and decide who gets to access which parts of their data. An example of such a blockchain is zkLedger [7], a new blockchain protocol developed at MIT that keeps

---

> contents of data transactions private, while allowing a third party to access certain parts of the data or query the system, without fully revealing contents of data transactions.

Society today is becoming more digital as people consume more information and video streaming, as well as exchange more data, thanks to the rapid progress in 5G telecommunication, electronic products and IoT technologies. There is an urgent need to protect data privacy, come up with policies that regulate how data is used and managed, as well as novel technologies to solve weakness that are present in the current system.

References

[2] S. Athey, C. Catalini, C. Tucker, *"The Digital Privacy Paradox: Small Money, Small Costs, Small Talk",* 2018.
[3]https://thelawreviews.co.uk/edition/the-technology-media-and-telecommunications-review-edition-10/1211270/singapore
[4] OECD (2013), "Trust in government, policy effectiveness and the governance agenda", in *Government at a Glance 2013*, OECD Publishing, Paris. DOI: https://doi.org/10.1787/gov_glance-2013-6-en
[5]http://www.scio.gov.cn/m/xwfbh/xwbfbh/wqfbh/2014/20140213/zy30387/Document/1362969/1362969.htm
[6] http://finance.china.com.cn/roll/20160301/3608406.shtml
[7]https://static1.squarespace.com/static/59aae5e9a803bb10bedeb03e/t/5aa1b35ce4966bd538d3f1d2/1520546653653/zkledger.pdf

III. Discussion and way forward

The BPF session at the virtual IGF 2020 will further discuss existing concerns and challenges related to the use of users' data generated, collected and used in an internet context and best practices to address the challenges to assure that users benefit from their data and not risk to be harmed.

The session will also discuss next steps and a way forward, and in particular, invites views on whether there's a need for new models and mindsets, and which role Internet governance and the IGF can play.

The highlights and key findings of this discussion will be included in the report.

Sessions details:

IGF 2020

Best Practice Forum Data and New Technologies in an Internet Context

Session at the virtual IGF 2020

Thursday 12 November

9:00 - 10:30 UTC

Session page https://www.intgovforum.org/multilingual/content/igf-2020-bpf-data-and-new-technologies-in-an-internet-context

Questions

1. What are current challenges and concerns related to the collection and use of users' data by new technologies and what best practices exist to ensure that data is used to bring benefit and not to harm users?

2. Are views and concerns changing and is there a need for new mental models and mindsets?

3. What is the way forward and what can be the role of Internet Governance and the IGF ?