Mobile Privacy and Big Data Analytics: Big Data for Social Good Considerations

November 2018

The GSMA represents the interests of mobile operators worldwide, uniting more than 750 operators with over 350 companies in the broader mobile ecosystem, including handset and device makers, software companies, equipment providers and internet companies, as well as organisations in adjacent industry sectors. The GSMA also produces the industry-leading MWC events held annually in Barcelona, Los Angeles and Shanghai, as well as the Mobile 360 Series of regional conferences.

For more information, please visit the GSMA corporate website at **www.gsma.com**

Follow the GSMA on Twitter: **@GSMA**

Contents

1 Introduction 2

2 'Big data analytics' 3

Scenario 1 4
Scenario 2 4
Scenario 3 5
Scenario 4 5

3 Privacy considerations 6

Personal Data 6
Privacy impact assessment, privacy-by-design 6
Accountability 7
Security Access 7
Cross-border transfers of data 8
Ethics 8

1 Introduction

Big data analytics and data driven services play a critical role in digital life and will continue to do so in the future. The mobile industry is harnessing big data to help public agencies and NGOs tackle infectious diseases, disasters, environmental impacts and climate change. Big data analytics depends on both the availability of data and on consumer trust.

The mobile industry is determined to help realise the economic and societal benefits of big data analytics through good digital responsibility practices, so that society can unlock the huge potential of big data analytics in a way that respects well established privacy principles and fosters an environment of trust.

The considerations in this document comprise safeguards to think about when engaging in big data analytics activities for social good. This document is based on the GSMA Mobile Privacy and Big Data Analytics[1] document published in February 2017 and applies considerations in this document to big data for social good. This document should be considered alongside existing GSMA material on data privacy topics.[2]

1. Mobile Privacy and Big Data Analytics, available at: https://www.gsma.com/publicpolicy/mobile-privacy-big-data-analytics
1. See The GSMA Mobile Privacy Principles; The GSMA Mobile Policy Handbook position on Privacy and Big Data; GSMA guidelines on the protection of privacy in the use of mobile phone data for responding to the Ebola outbreak; The GSMA Privacy Design Guidelines for Mobile Application Development

2

‘Big data analytics’

In this document we have used the term ‘big data analytics’ to mean the use by any organisation of big data analytics techniques or the services provided by them across a range of scenarios including:

- Mobile network operators (MNOs) using their in-house analytics services to provide insights into human movement patterns, enriched with third-party data sources, such as hospital intakes, death counts and weather data, to enable relief agencies to make decisions on when, where and how to deploy resources.
- Making selected data available to an analytics service provider for them to conduct analytics and return insights;
- Making selected insights available to an application provider for them to deliver an enriched service to subscribers opting into the enhanced service; and
- Several parties transmitting their data to an analytics service provider to conduct analytics and return insights on all the data.

The following examples are intended to illustrate some of the data privacy considerations that may be taken into account:

Scenario 1

A city wants to better combat the adverse health impact of air pollution. For example, it would like to know if it can study mobility patterns to better monitor and predict pollution levels. An MNO puts a contract in place with the city authority to provide insights developed through the MNO's machine learning algorithms, using anonymised data from the mobile network. Anonymising the data preserves the privacy of users- no identifiable data is shared or accessed by anyone outside of the MNO. The anonymised data can then be combined with other data sources, such as weather, traffic and pollution sensors, to predict pollution levels in advance. This information enables local authorities to take preventative steps if emissions are likely to reach dangerous levels.

Scenario 2

A country is experiencing an outbreak of a highly contagious disease. Mobile data can be used to provide detailed, up-to-date behavioural insights across large populations to help track and mitigate the spread of contagious disease. An MNO partners with an international health organisation in order to supply this data and help the health organisation to direct their resources more effectively. However, given that this data could potentially reveal information about individuals, the MNO will anonymise and aggregate the data before sharing any data with the international organisation. The MNO can work with the international organisation to analyse the anonymised and aggregated data, with the goal of identifying potential hotspots of infectious disease, which could then help target preventative intervention.

Scenario 3

A nation with high earthquake risk wants to ensure they are prepared in the case of a natural disaster. Mobile data can be used to show movement patterns before and after previous disasters, to better map routes. The government can work with the mobile network to identify the correct aggregated data sets, and the mobile network and other partners can combine this data with other data, such as weather, traffic and flood data, to provide unique insights to develop an early warning system before and during natural disasters.

Scenario 4

A government wants to better identify communities affected by climate change. Anonymised and aggregated mobility data generated by a mobile network can show the displacement of people, without identifying individuals. The MNO can use big data algorithms to study the changes in populations in particular areas over time. This mobile data can be combined with data about droughts and other weather conditions, to help the government predict future displacement and intervene accordingly.

3

Privacy considerations

In order to realise the potential societal and economic benefits of big data analytics in a way that is compatible with recognised data privacy principles, the following considerations may be taken into account:

Personal Data

Some of the data used in big data services is not personal data.[3] Readings from weather sensors, for example, would not constitute personal data. Big data analytics services should take into account that such non-personal data can become personal data if it is associated with a particular individual, for example, if the location of a connected car detected by a traffic management system is subsequently combined with the vehicle registration number and the vehicle ownership records.

Big data analytics services can consider guarding against the possibility of re-identification of individuals when the data is merged with other data sets. Where personal data is collected, for example, when a mobile phone user's location is recorded, this can be de-identified through the removal of data fields that enable identification and through reporting the analytic insights only in aggregate or approximated form.

Example

If an MNO is analysing mobility data, and has removed any identifying information about individuals, the MNO could consider whether that mobility data could lead to the inadvertent identification of individuals. Aggregating mobility data, and setting limits on data granularity, can prevent this de-identified data from being re-identified.

Privacy impact assessment, privacy-by-design

Through identifying new correlations across data sets, many big data analytics services hope to provide actionable insights that have a positive impact on society or individuals.

A key tool for recognising privacy impact on individuals is the data privacy impact assessment, which helps organisations to identify and mitigate privacy risks.[4]

Developers of big data analytics services can consider adopting a 'privacy-by-design'[5] ethos or methodology by which privacy and security safeguards are considered and designed into products, services, processes or projects at each stage of the lifecycle from cradle to grave.

Example

In Scenario 3, assuming the system is designed and built in the anticipation of future emergencies, a data privacy impact assessment could reveal the sort of impact that data disclosures may have. Rather than disclose identifiable data to the government, the system can be designed to only send aggregated insights to the government. If the government needs to communicate directly with individuals whose life is at risk, the MNO can send a message to the affected population.

2. Definitions vary, but generally speaking personal data is considered to be information relating to a living individual or from which an individual may be identified either from the information itself or when combined with other data that is likely to come into the possession of the organisation.
3. See ICO Code of Practice on conducting privacy impact assessments; Centre for Information Policy Leadership - A Risk-based Approach to Privacy: Improving Effectiveness in Practice
4. See Office of the Information and Privacy Commissioner of Ontario – Privacy by Design; Office of the Information and Privacy Commissioner of Ontario - Privacy Considerations at Each Stage of the Big Data Lifecycle;
5. Office of the Information and Privacy Commissioner of Ontario - Big Privacy: Bridging Big Data and the Personal Data Ecosystem through Privacy By Design

Accountability

As mentioned above, there are many different scenarios in which big data analytics services may be delivered.

Where one organisation makes their data or insights available to third parties, that organisation is in a position to set out clearly what they will be responsible for and what they expect of the third parties. For example, the organisation may want to keep a log of the API access events or ask third parties to report on their use of the data and/or insights.

Where several organisations are collaborating under a common framework or where they are pooling their data,[6] a responsible approach might be to set out clearly in an agreement or protocol the ground rules for access, for example:

- What data will be made available to whom and in what form?
- Will the data be pseudonymised?
- Will there be access to any raw data?
- Who has access to the findings?
- Who can determine the design/purpose of the algorithms?
- How and to whom will insights from the analytics be disseminated?
- Who will control actions taken based on the insights?

Example

Limiting access to data can help protect privacy. In the context of the GSMA Big Data for Social Good initiative, operators involved in trials limit access to data. If many parties are involved in a project, all need a clear understanding of who is responsible for limiting access to data.

Security Access

As with other personal data processing activities, security is a key safeguard to protect people's privacy. Big data analytics services can improve security by limiting access to data or insights to authorised users only and by securing the data or insights in transit, in rest during the analytics phase, and in the release of reports or insights.

When setting time limits for data retention, big data analytics services can consider the sensitivity of the data and whether it is possible to continue beyond a certain time with pseudonymised/aggregated data only.

Example

Providing service may require a mobile operator to retain identifiable data about consumers for a limited time, and thereafter they may need to keep de-identified data for a number of years to understand trends and correlations. Other big data analytics services where several parties operate under one framework or as one platform may want to establish themselves as a longer-term resource, in which case other security measures such as pseudonymisation, access control or encryption will take on a greater significance.

6. See ICO Data Sharing Code of Practice; ICO data sharing checklists

Cross-border transfers of data

Cross-border transfers of personal data are currently regulated by a number of international, regional and national instruments and laws intended to protect individuals' privacy, the local economy or national security.

Transmitting personal data across national boundaries can sometimes trigger additional duties. Big data analytics services can comply with cross-border transfer requirements and enable their services if they implement contractual privacy safeguards or if they embrace accountability mechanisms such as the APEC Cross-Border Privacy Rules or the EU's Binding Corporate Rules which allow organisations to transfer personal data generally under certain conditions.[7]

Example

The arrangements in Scenario 2 are likely to involve some transfers of data across national boundaries, particularly if the international health organisation, major NGOs and MNOs all agree to run the system on a common platform in anticipation of future outbreaks. The parties may want to consider entering into contractual terms that will protect the interests of future individuals whose data will be collected and use these terms to comply with data transfer restrictions.

Ethics

In addition to considering legal requirements, big data analytics services may also consider the overall fairness and the ethical dimension of what they, or the third parties accessing the insights, are proposing to do.

Organisations can incorporate ethical decision-making models into their business processes to help build better services and foster an environment of trust.

Example

If the disclosure of aggregated data under Scenario 3 were to show the movement of particular ethnic groups, but not individuals, in the aftermath of an emergency, the MNO might consider the proposed big data analytics services and the circumstances in which such information may be released from an ethical rather than just a legal point of view.

7. See The GSMA Mobile Policy Handbook position on Cross-Border Data Transfers

gsma.com/betterfuture/bd4sg

For more information, please visit the GSMA website at
www.gsma.com/betterfuture/bd4sg

GSMA HEAD OFFICE
Floor 2
The Walbrook Building
25 Walbrook
London EC4N 8AF
United Kingdom
Tel: +44 (0)20 7356 0600
Fax: +44 (0)20 7356 0601