



*Ethics, research and the data-driven society*  
*Speech to ESOMAR European Insights Summit*  
*24 January 2017*  
*Giovanni Buttarelli*

First of all thank you to all at ESOMAR for the invitation to be with you today for this closing session in the middle of ‘data protection week’.

Information is the raw material and the output of research activities.

It is true that, more than ever, there is a great premium on personal information, personal data.

At the same time, there is a lot of concern in the research community that rules on data processing impede cutting-edge research.

An urgent debate is taking place in all sectors, commercial, government, academic and non-for profit.

This debate concerns how to harness the potential power of big data for the benefit of individuals and wider society.

And, in fact, I would say that ethical use of personal data for scientific research lies at the heart of the matter.

I believe that the new gold standard for such laws is the EU’s General Data Protection Regulation, which will be fully applicable from 25 May 2018.

I believe that the GDPR gives research an opportunity to harness big data for the common good.

Big data is responsible for bringing people closer together in some ways.

But it has also alienated individuals in other ways.

It has certainly widened the gulf between individuals and powerful tech companies and even governments.

People feel disempowered when it comes to what data about them is being processed, who is processing the data and for what purposes.

This sense of lack of control is the Digital Condition for most people, and it is, in my view, unsustainable in the longer term.

This gulf, this lack of control, will only become greater as we move towards ever more connected everyday objects, and towards machine learning and AI.

When we talk about the ethics of research, the most obvious precedents are initiatives like the Helsinki Declaration, which concerns medical experiments conducted on humans.

The interaction between research and individual undergoing such research, in such contexts, is generally direct and unmediated.

Consent, therefore, is a routine requirement.

But with data research, the process is heavily mediated by technology.

Data can be understood in many different ways.

For instance, the Internet Accountability Foundation has developed a taxonomy of personal data, ranging from data which we consciously and willingly provide, for example when we complete a form, through data which is observed or derived to data which is inferred about us.

It is like a spectrum of the degree to which individuals are informed and in control of the data which concerns them.

Much big data consists of 'publicly available' or 'open source' data.

But even here, rights are still at stake.

According to EU law, even disseminating via a search engine old news about the debts of a Spanish gentleman amounts to an interference with the rights of the individual.

That was the finding of the CJEU in the *Costeja* case.<sup>1</sup>

There is, in my view, a clear need for a common ethical framework for personal data research.

The controversy about the Facebook emotion experiment in 2014 is an excellent example of why this framework is needed. In this case, a reputable US scholar conducted research on behalf of a corporation, but the individual user account holders

---

<sup>1</sup> CJEU Judgement C-131/12 [2014] *Google Spain v AEPD and Mario Costeja González*.

had no idea what was taking place. Even though the results of the experiment were published, the raw data remained the property of the company.

Allow me to spend a few minutes discussing some key aspects of the new legal framework - principally the General Data Protection Regulation.

First of all, the GDPR aims to encourage innovation, as long as organisations implement the appropriate safeguards.

Scientific research was one of the areas of the GDPR which, quite correctly, required long and hard reflections.

The GDPR follows five guiding principles in this regard:

1. Though there is no formal definition of scientific research, the GDPR, like its predecessor Directive 95/46, acknowledges the need to facilitate research.
2. It addresses the specific needs of different types of research - scientific and historical research, statistical research, and archiving in the public interest.
3. It recognizes the need for ethical and responsible research, with appropriate limitations (Recital 33)
4. It stresses the importance of safeguards for data subject's rights
5. It takes (Recital 159) a broad view of what research is, mentioning explicitly technological development, fundamental research, applied research and privately funded research, and 'studies conducted in the public interest in the area of public health'. So there is no arbitrary division between public and commercial research.

The GDPR is long document, 99 articles, with more detailed obligations compared to the previous Directive.

But there is clearly a lighter regime for research.

For example, there are special, lighter rules on:

1. data retention limits;
2. purpose limitation so that researchers to process personal data beyond the purposes for which they were first collected;
3. whether consent is required as a legal basis for data processing as part of research;
4. allowing the processing of sensitive data;
5. transfers personal data to third countries including those that have not been judged 'adequate' in the level of protection.

There is also a general effort by the lawmaker to encourage EU member states to enact greater protections for the processing of sensitive data for health-related purposes, especially where the controller is responsible for processing genetic, biometric, or health data.

This more flexible regime depends on the provision of safeguards for the individual concerned by the data.

It provides a basis for a new ethical standard, to be derived from new principle in the GDPR of accountability.

Accountability, perhaps the most important innovation in the GDPR, means taking steps to ensure you, as a data controller, comply with the rules. And taking steps to demonstrate that you are complying with the rules.

These steps include :

- 'data protection by design and by default' (Art. 25),
- data protection impact assessments (Art. 35),
- keeping records on data processing activities (Art. 30), and
- reporting on data breaches (Art. 33-34).

It is clear that much research has moved out of the publicly-funded university and into corporate settings.

This suggests that there may be a need for new principles and approaches to the ethics of research activity that relies on large volumes of personal data.

There have been a number of different and interesting suggestions.

For example, Consumer Subject Review Boards for corporate data research (Ryan Calo University of Washington School of Law) and an agreed system for weighing benefits and risks, and risk mitigation,<sup>2</sup> and a new standard for trust and a duty of loyalty<sup>3</sup>

Certainly a collective effort is needed to determine when certain data research activities may raise ethical considerations.

And should, furthermore. the ethical assessment take place *ex ante* - before the experiments -or after? What would be the effect on innovation?

There is surely a role for independent data protection authorities, and for the European Data Protection Board in considering appropriate guidance for the research community.

---

<sup>2</sup> Ryan Calo \*\*, Tene and Polonetsky

<sup>3</sup> Richards and Hartzog

Later this year, I intend to develop these ideas, with your input I hope, in a joint paper with a leading privacy scholar based in the United States.

So in conclusion, data protection laws in the EU have always aimed to work with the grain of innovation and social progress.

Personal information, which is so closely connected to an individual's privacy and freedom to develop her own personality, is now ubiquitous.

People are beginning to realise that they need dependable safeguards in the digital age.

So this is moment to consider need norms for data research, just as we have had - for a long time - with medicine.

The GDPR is potentially an ally for the research community in your basic remit to advance generalisable knowledge.

Thank you for listening.