EUROPEAN DATA
PROTECTION SUPERVISOR

Big data, big data protection: challenges and innovative solutions

ERA Conference on Recent Developments in Data Protection Law

Keynote speech

Brussels, 11 May 2015

Giovanni Buttarelli

European Data Protection Supervisor

I

I'm delighted to have the opportunity to contribute today. ERA's annual workshops are a tremendously valuable resource for anyone who needs to stay abreast of the fast-moving data protection debate.

II

The 1995 Data Protection Directive became law in a very different world.

In 1994 there was allegedly the first online transaction for a pizza from Pizza Hut. Around the same time Mark Butler published a book called How to Use the Internet which included advice such as:

- (On joining mailing lists) '*Although it is polite to say "please" and "thank you" to a human, do not include these words in the messages you send to a listserv. They may confuse the machine*.'

and

- (on searching the Internet): '*If a particular search yields a null result set, check carefully for typing errors in your search text. The computer will not correct your spelling, and transposed letters can be difficult to spot.'*

But Butler's book also contained advice which is as true today as it ever was:

- *'Surfing the Internet is a lot like channel surfing on your cable television. You have no idea what is on or even what you want to watch.'*

and

- '*Never forget that electronic mail is like a postcard. Many people can read it easily without your ever knowing it. In other words, do not say anything in an e-mail message which you would not say in public*.'

1995 was a watershed year for technology:

- the removal of the last restrictions on the use of the Internet to carry commercial traffic (NSFNET decommissioned and replaced by backbones operated by several commercial ISPs);
- the number of websites reached 10,000 and the number computers connected to the Internet reached two million;

- The 'second generation' mobile phone systems was still emerging, using digital instead of analog transmission. And more people starting using mobile phones thanks to the advent of prepaid services.

If you read the recitals of 1995 Directive, the EU legislators knew something big was happening in how we communicate. The words internet and computer do not appear at all, and ‘electronic’ appears only once, but the lawmakers were conscious of the emergence of an ‘information society’ making it easier to communicate across borders.

III

Fast forward twenty years, and today there are 45 billion web pages and roughly three billion web users.

We find ourselves in the ‘global village’ predicted by Canadian philosopher of communications Marshall McLuhan: the globe contracted into a village by electric technology and the instantaneous movement of information from everywhere to everywhere all the time.

Data moves around via mobile devices – phones mainly – but increasingly with other things that can be worn on your person: watches, SmartBands, glasses.

Algorithms based on neural networks aim to emulate the human brain and already can understand language and recognise images, and they are being used to analyse medical data.

The vast amount of data for these algorithms to establish patterns are now available, thanks to us, the data subjects – emails and web searches, and photos and videos which we have helpfully pre-labelled.

Our data is now an asset for big tech companies, analysed and monetised. To illustrate: WhatsApp is a company which in size and budget might be considered equivalent to the EDPS, the smallest institution in the EU. Yet it is an app whose number of users is approaching 1bn. And last year it was acquired by Facebook for around 19bn USD.[1]

Meanwhile, the Snowden revelations of 2013 indicated that the internet has been exploited to create a global surveillance state.

In both the business and government spheres, there is a worrying drift towards thinking that, with regards to personal information, whatever is possible is also desirable: if personal data are available, they should be collected and stored indefinitely and exploited for any expedient purpose.

IV

As European Data Protection Supervisor, I am five months into my mandate and two months into implementing our strategy for the next five years. It's an ambitious but realistic agenda and that reflects the basis for the appointment by the EP and Council of myself and the Assistant Supervisor, Wojciech Wiewiórowski.

V

[1] Facebook Press Release 19.4.2015.

We were challenged to use our mandate to do three important things:

- First, develop and communicate a global vision, think in global terms and propose concrete recommendations and practical solutions
- Second, provide policy guidance to meet new and unforeseen challenges in data protection
- Third, represent at the highest levels and develop and maintain effective relationships with diverse community of stakeholders in other EU institutions, Member States, non EU countries and other national or international organisations

Our strategy is focused on three broad objectives which address the changing landscape which I've just outlined. I'd like to share with you today a flavour of those priorities and would welcome further discussion at the end of my talk.

VI

First objective – data protection going digital. We need to find new ways for applying data protection principles to the latest technologies, be they big data, the internet of things, cloud computing, artificial intelligence, drones or robotics.

This means placing the individual more firmly at the heart of technological development, through transparency, user control and accountability.

We've seen a lot of headlines devoted to the 'right to be forgotten'. But in fact, in its judgment on Google Spain in May last year, the European Court of Justice did not invent a new right. It rather confirmed that if you process personal data (and, it ruled, search engines certainly do process and make decisions on processing personal data) then you have a responsibility to treat those data in a way that respects the rights and interests of the individual. Part of that responsibility is enabling the individual to challenge what you do with the information which relates to him or her.

There are imbalances in the market, raising questions of fairness of competition and consumer protection which we have been discussing with experts since publishing a Preliminary Opinion on the subject a year ago.

Data protection principles, such as fairness, purpose limitation and data minimisation, can guide software developers as well as legislators and judges in ensuring that the interests of the individual are always paramount.

We aim to provide guidance on this, starting this month with the development of smart devices and mobile apps for delivering health-related services.

VII

Second, we need global partnerships on the big questions posed by these technologies, and by the social and economic changes which accompany them.

If big data is characterised by velocity, volume and variety, then ‘big data protection’ means a dynamic and multi-disciplinary approach with concerted action on an international scale.

I believe in interoperability between different approaches to privacy and data protection around the world, if such interoperability is genuinely two-way, and both sides in the discussion respect the other's values in practice, not just in words.

This applies to international agreements like Safe Harbor, Transatlantic Trade and Investment Partnership (TTIP), The Trade in Services Agreement (TiSA), and to law enforcement like the EU-US Umbrella Agreement, Passenger Names Records and the Terrorist Finance Tracking Program. On the one hand these are unique opportunities for broad visions and constructive cooperation with our global partners. But at the same time we need to ensure that there are no back doors for weakening the protection of hard-fought rights.

As Bruce Schneier said in his book this year on surveillance: this is the cyber sovereignty moment, with the real threat of Balkanisation of the internet at great cost to individual freedom and social progress. So we need to build bridges between Europe and other regions which share our values.

VIII

Our third objective addresses the need for a new deal on data protection in the EU. The new data protection regulation, as Jan-Phillip Albrecht earlier today explained, is just the beginning. It needs to be properly implemented in all sectors.

But for now, we need to help the Parliament and the Council steer towards a set of rules which are simpler and easier to apply. Three years after our initial opinion on the Commission's data protection reform proposals, we intend to publish further advice to complement the trilogue negotiations.

IX

Last Wednesday, the Commission published a communication on the Digital Single Market. It's about trying to harness technology and enhance the EU's competitiveness: 'Big data, cloud services and the Internet of Things,' it states, 'are central to the EU's competitiveness.'

One of the proposals we are told to expect will aim to remove unjustified restrictions on the 'free flow of data', including questions of ownership, interoperability, usability and access to data in situations such as business-to-business, business to consumer, machine generated and machine-to-machine data. This echoes the G8´s 2013 Open Data Charter whose principle of ´open by default´ aims to make data more freely and openly re-usable.

So one of the first ways we will put the EDPS strategy into action is by publishing guidance for data controllers and policy makers on exploiting big data in the interests of the individual and their rights.

If the processing becomes more complex, data controllers have the responsibility to ensure users and consumers are properly informed.

This is not a new paradigm. We have been here before.

Previous generations of data protection rules have addressed the transition manual to automated processing, from analogic to digital networks, from the pioneering development of e-commerce to the Information Society, from silos to interconnected large-scale data systems.

Fundamentally this is a question of scale.

Existing principles must be applied creatively to safeguard against the harm to the individual of intrusive profiling and unfair discrimination.

We must keep the concept of identifiability under careful review in the light of computational capabilities allowing unstructured sets of data, amassed for different purposes and in different contexts, to be used to reidentify individuals in unpredictable ways.

We must address the challenge to data quality, because the inferences made by Big Data are not and probably never will be 100% reliable, accurate and trustworthy.

Big Data - used well - can be used to change the world positively without compromising our fundamental rights.

X

Our mandate runs to 2019. Technology is not going to wait for the EU to update its data rules.

You may have heard about smart cars - that's just one example of the internet of things, devices talking to each other and transmitting personal data about us, usually without the user being aware of it.

By 2019, there will be billions of these devices. The scale on which they gather and treat personal information will rival the industrial and agricultural revolutions. Let's be conscious of the side effects of those two earlier technology-enabled revolutions, which can be devastating, and devise normative regulation and legislation that fosters innovation and minimises harm to the individual.

XI

Thank you.