BIG DATA AND HUMAN RIGHTS: ETHICS, LAW, AND TECHNOLOGIES

Dr. Svetlana Maltseva,
National Research University Higher School of Economics,
Dean, Business Informatics Faculty
Dr. Mikhail Komarov
National Research University Higher School of Economics,
Dean, Business Informatics Faculty
Dr. Andrey A. Shcherbovich,
National Research University Higher School of Economics,
Lecturer, Department of the constitutional and municipal law

Today the Big Data sciences turn its age out. Some years pass, and there will be no need of data scientists, because all process of the big data collection will be automated. And this makes a big challenge to the scope of issues related to human rights of the subjects of personal data. This is a complex issue related to ethical, legal, and technological problems of human rights in Internet Governance.

Big data, as we now refer to enormous collections of facts, figures and unstructured information like metadata and tweets has helped us better understand crime rates and predict outbreaks of communicable diseases, and it radically improves our online shopping experiences. But imagine the potential benefits when such data science innovations are applied to the world of human rights. Rather than a digital hazard, computer technology that can handle big data can draw from

information about human sentiments and actions to predict potential atrocities reveal patterns of destructive human activities such as trafficking and help weigh prescriptive policies.

Supranational Level

For example, the Amnesty International creates a model of researching Big Data's effect on Human Rights. Rights group Amnesty International USA could soon use data analytics to predict which incidents are likely to escalate into larger human rights violations. If successful, this endeavor may enable those concerned about human rights to more effectively address situations before they reach crisis points.

We know the basic universal instruments related to the personal data are:

1. Universal Declaration of Human Rights, adopted at the third session of the UN General Assembly Resolution 217 A (III) of 10.12.1948 , which states that no one shall be subjected to arbitrary interference with privacy, family, everyone has the right to the protection of the law against such interference or attacks (Article 12);
2. International Covenant on Civil and Political Rights (New York, 19.12.1966);
3. Council of Europe Convention on the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS N 108) (concluded in Strasbourg, January 28, 1981).

The Convention establishes the procedure for the collection and processing of personal data, the principles of storage and access to these data, the methods of physical protection of data. Convention guarantees respect for human rights in the collection and processing of personal data, as well as prohibit the processing of data on race, political opinions, health, and religion without proper legal basis.

One of the most detailed European instruments is Directive 95/46/EC of the European Parliament and of the Council of 24.10.1995 on the protection of individuals with regard to the processing of personal data and the free circulation of such data.

Now we could outline different approaches took into account in studying big data regulation and legislations of different countries.

United States

The growth of smartphones and social media are giving the world instant, first-hand accounts of human suffering and political repression during events like the 2010 Haiti Earthquake, recent elections in Kenya, and the ongoing uprising in Syria.

To investigate how social media and big data analytics are changing human rights fact-finding, and to better understand the ways that these technologies can advance human rights protection in the future, the MacArthur Foundation recently awarded an 18-month,

$175,000 grant to Carnegie Mellon University's Center for Human Rights Science, directed by Jay D. Aronson.

Human rights organizations, governments and the general public are increasingly turning to this massive accumulation of images, video and text to investigate and understand the human impact of conflicts, disasters and political violence. It remains unclear, however, whether this huge amount of data actually improves the global community's ability to protect and promote the rights of vulnerable individuals around the world - particularly those who still lack reliable and secure access to the Internet or whose rights are violated in private, rather than in public view[3].

Technological progress should bring greater safety, economic opportunity, and convenience to everyone. And the collection of new types of data is essential for documenting persistent inequality and discrimination. At the same time, as new technologies allow companies and government to gain greater insight into our lives, it is vitally important that these technologies be designed and used in ways that respect the values of equal opportunity and equal justice. American approach is based on these five points.

1. **Stop High-Tech Profiling**. New surveillance tools and data gathering techniques that can assemble detailed information about any person or group create a heightened risk of profiling and discrimination.

[3] Information systems; Carnegie Mellon to study how social media and big data affect protection of human rights. (2013). Information Technology Newsweekly, 576.

Clear limitations and robust audit mechanisms are necessary to make sure that if these tools are used it is in a responsible and equitable way.

2. **Ensure Fairness in Automated Decisions**. Computerized decision-making in areas such as employment, health, education, and lending must be judged by its impact on real people, must operate fairly for all communities, and in particular must protect the interests of those that are disadvantaged or that have historically been the subject of discrimination. Systems that are blind to the preexisting disparities faced by such communities can easily reach decisions that reinforce existing inequities. Independent review and other remedies may be necessary to assure that a system works fairly.

3. **Preserve Constitutional Principles**. Search warrants and other independent oversight of law enforcement are particularly important for communities of color and for religious and ethnic minorities, who often face disproportionate scrutiny. Government databases must not be allowed to undermine core legal protections, including those of privacy and freedom of association.

4. **Enhance Individual Control of Personal Information**. Personal information that is known to a corporation — such as the moment-to-moment record of a person's movements or communications — can easily be used by companies and the government against vulnerable populations,

including women, the formerly incarcerated, immigrants, religious minorities, the LGBT community, and young people. Individuals should have meaningful, flexible control over how a corporation gathers data from them, and how it uses and shares that data. Non-public information should not be disclosed to the government without judicial process.

5. **Protect People from Inaccurate Data**. Government and corporate databases must allow everyone — including the urban and rural poor, people with disabilities, seniors, and people who lack access to the Internet — to appropriately ensure the accuracy of personal information that is used to make important decisions about them. This requires disclosure of the underlying data, and the right to correct it when inaccurate[4].

In Europe, there are strict rules about what companies can and can't do in terms of collecting, using, disclosing and storing personal information, and governments are pushing to make the regulations even stronger. That has prompted renewed debate about whether it is time for the U.S. to toughen its relatively lax privacy regulations.

In one camp are those who believe the U.S. government should refrain from meddling. They say the

[4] Civil Rights Principles for the Era of Big Data // CivilRights.org, a project of The Leadership Conference on Civil and Human Rights & The Leadership Conference Education Fund. URL: http://www.civilrights.org/press/2014/civil-rights-principles-big-data.html

lack of privacy restrictions in the U.S. has encouraged innovation in the online-marketing industry, which is still evolving, and they question whether a Congress that isn't capable of passing a budget can be trusted with crafting complex privacy legislation.

The U.S.'s experiment with self-regulation has been a failure; say those who believe Europe's approach to privacy is superior. By trusting industry to police itself, the U.S. has created a situation where consumers have little control over personal data and few remedies when they find their privacy has been invaded[5].

European Communities

In an age of "Big Data" (when data relating to our own actions are shared and/or exploited in aggregate form) and the "Internet of Things" (when more and more physical objects – things – are communicating over the Internet), it is becoming increasingly difficult to ensure true anonymisation: the more data there are the easier it becomes to identify a person. Moreover, the "mining" of the Big Data resources, in ever more sophisticated ways, tends to lead to the creation of "profiles". Although these profiles are being used to spot rare phenomena (e.g., to find a terrorist in a large set of data on thousands of people, such as airlines' Passenger Name Records), they are unreliable and can unwittingly lead to discrimination on grounds of race, gender, religion or nationality. Yet

[5] Big data (A special report) - should the U.S. adopt European-style data-privacy protections? (2013, March 11). Wall Street Journal.

these profiles are constituted in such complex ways that the decisions based on them can become effectively unchallengeable: even those implementing these decisions are unable to fully comprehend the underlying reasoning[6].

The data relating to our own actions and the data generated and reported on by "things", are also increasingly shared and/or exploited in aggregate form, as so-called Big Data. This can include medical data in supposedly de-identified formats, the number of crimes in a specific area, demographics and school results. Companies and governments are keen to exploit these data resources to the fullest extent.

the analyses and mining of the Big Data resources, in ever more sophisticated ways (to turn Big Data into Smart Data), tend to lead to the creation of "profiles": algorithms derived from the data that establish statistical correlations between often seemingly unrelated facts. Once created, these profiles are then applied to the real world and to individual people: to identify risk factors so that people susceptible to certain diseases can be called in for preventive checks; or to increase their insurance premiums; or to identify the effects of street design and lighting on crime levels, to improve planning; or to direct police resources; or indeed to identify people who may be wanting to commit

[6] The rule of law on the Internet and in the wider digital world. Issue Paper published by the Council of Europe, Commissioner for Human Rights. September 2014.

suicide by throwing themselves under a train (as is done in the London Underground) or who may be terrorists.

In this new environment, we – and the "things" around us – all generate extremely detailed personalized or quasi-personalized data trails, even if we are only half-aware of them. These data can be used to map social networks: the spiders' webs of contacts linked to contacts linked to further contacts. Combined with Big Data and profiles, they can show surprisingly revealing details of every man and woman's life, beliefs, inclinations, health and activities – at least with a high degree of probability. Just a few "likes" on Facebook suffice to predict religion, race or sexual orientation of the user with high degrees of accuracy; and just a few innocent purchases (e.g., of unscented body oils) have been used to identify women who were likely to be in the second trimester of pregnancy, but who had never revealed this fact.

Japan

Dr. Taro Komukai believes that the Personal Data Protection Law in Japan requires entities only to publicize the purpose of use. It doesn't require that entities get consent of the person like the regulation in EU. And there is no provision on the deceitful action in use of personal data like the regulation in the US. Therefore, even when many people think a particular purpose of use is unacceptable, the use cannot be stopped on the ground that the purpose is not

appropriate. These two points should be discussed for the reform.

Japanese data protection scheme is different from that of EU and U.S. in two points. One is lack of the privacy commissioner or some other type of authority that is responsible for the rule making and enforcement. The other is that there is no provision to ensure the entities make good the appropriate purpose of use.

There are many approaches of self-regulation for data protection in Japan. The Smartphone Privacy Initiative is one of the approaches for the privacy online. The initiative focused on the data protection associated with smart phones and proposed six principles. 1) Ensuring Transparency, 2) Securing the Opportunity of User Participation, 3) Ensuring Data Collection through Proper Means, 4) Ensuring Proper management of User Information, 5) Properly Handling Complaints and Requests for Advice, 6) Privacy by Design[7].

Russian Federation

Information systems have become an important and essential attribute of all spheres of human activity. Rapid evolution of ICT stimulates demand for new products in almost all directions. The development of this area will be primarily associated with the development of cloud computing, new architectures and principles of computing, problem solving very large

[7] Kshetri, Nir. The Expert Opinion // *Journal of Global Information Technology Management;* 2013; 16, 4; ABI/INFORM Global, pg. 68.

scale data (Big Data), the development of new analytical tools.

The key scientific and technological trends shaping the face of this priority primarily include:

– development of research in the field of a single management environment and a common information space of the transport infrastructure (environment unified exchange of information between vehicles); the development of this trend will help to cope with the constant increase in the density of traffic flows at complication of the organization by increasing the efficiency of supply chains;

– development of research in the field of new principles of algorithms, creation of computer architectures built on new paradigms, including neurons, biological, optical, quantum, self-locking, recurrence, which will increase the maximum clock frequency of an optical computer to 10^{12}- 10^{14} Hz (for 3 - 5 orders of magnitude higher than existing electronic analogues);

– development of research in the field of machine learning based on new methods and algorithms, the results of which have a very wide range of applications: the intellectualization of decision support, such as geographic information systems and decision-making in medicine, monitoring of financial and stock markets, and others;

– development of research in the field of communication infrastructures with terabyte information rate determines the future of the technological base of

network infrastructures and avoids restrictions on the organization of the main channels of universal broadband access, as well as significantly increase the size of the potential computing clusters;

– development of supercomputing through the development of new algorithms for applications with complex logic calculation process requiring processing of non-numeric data or data with complex representation, development languages and systems of parallel programming for inhomogeneous supercomputer systems (including distributed object-oriented systems) as well as expanding the range of specialized single-chip processors used in high-performance computing complexes with non-uniform architecture;

– development of cloud infrastructures, networks of personal computers and mobile devices will reduce the cost of maintaining the IT infrastructure, as well as lead to the creation of market infrastructure external remote location that has a direct impact on the appearance of the country specialization and global competition in this market;

– development of research into new interfaces (tactile sensors, 3D-printers, including bioprinting, built-in intelligent systems, interfaces "brain - computer" hardware clock monitoring critical physiological parameters) would go to a new level of integration network technology in everyday life and will be important for preventive medicine and healthy lifestyles;

– growth of mobile devices (tablets and smartphones), consisting of interface devices users of

information systems and services will form a new model of information systems and increase the mobility of both individual and corporate users that will lead to the spread of employment schemes remote employees;

– creation of separate hardware information and integrated systems with realigning terminal (sensor and actuators) modules in the design of man-made systems for the nodes of the address control spending their resources, maintaining high efficiency and reduce degradation caused by wear and tear, aging and extreme external factors;

– evolution of the Internet, which implies further development of the concept of distributed networks with independent and adaptive routing nodes between them in terms of working with content (Semantic Web - submission of information on the Internet in a form suitable for machine processing) and the inclusion of new classes of infrastructure objects (Internet of things - various items of information and integrate them into a network of networks).

Further development of the above scientific and technological trends will significantly strengthen the impact of ICT on social processes in society; there will be new forms of socialization and social interaction, change the character and way of employment of employees expected to offset development centers,

competences and production outside of developed countries[8].

On September 1, 2016 shall come into force the laws, which set a new duty of the operator of personal data relating to the collection of this information, including via the Internet.

So, recording, systematization, accumulation, storage, clarification (update, change), the extraction of personal data of citizens of the Russian Federation shall be carried out with the use of databases residing on the territory of Russia. Exceptions to this rule will make cases where the processing of personal data is required, for example, to reach the envisaged international treaty of the Russian Federation or the law purposes, as well as some of the other (n. 2, 3, 4, 8 h. 1, Art. 6 of the Law on Personal Data). Ensure that on the territory of the Russian Federation databases must be holders of information, information system operators.

Access to information resources on the Internet, including the network address, domain name, index pages, allowing the identification of the information processed with violations of the law, may be limited. This will be possible on the basis of which came into force a judicial act subject to the procedure provided for Art. 15.5 Information Act in the wording of the Act restricting access to a register of rights violators of

[8] Forecast of the long-term socio-economic development of the Russian Federation for the period up to 2030 (designed by Russian Ministry of Economic Development) // ConsultantPlus Legal Database System.

personal data. After the elimination of violations or the entry into force of the court decision to cancel the previously accepted judicial act domain name, index page or network address removed from the register.

In addition, the current legislation provides for administrative liability for violation of the collection, storage, use or dissemination of information about citizens. Such violation shall entail a warning or a penalty for legal entities from 5000 to 10000 Rubles.

Before the entry into force of the Act must be created the conditions necessary for compliance with the provisions analyzed. It seems that, for example, foreign companies operating with personal data of Russian citizens have to ensure the availability of databases on the territory of the Russian Federation. In addition, it is unclear as in the preparation of personal data over the Internet will be determined by whether a person is a citizen of the Russian Federation[9].

***

Dr. Mikhail Komarov states that we should think about "privacy by design" issues and probably special certification for systems dealing with personal data. He also supports initiative of "open interfaces to enable

[9] ConsultantPlus: Analytical review of the August 5, 2014. Personal data of Russian citizens as a general rule will be processed in Russia (Federal Law of 21.07.2014 N 242-FZ) // ConsultantPlus Legal Database System, 2014.

communications between members and non-members". There is a good example explaining how it works with terms and conditions and our privacy – a movie "Terms and conditions may apply" by Cullen Hoback.

We should not fear of the Big Data concept development and implementing new technologies in our life however, we should allow individuals being excluded from all the analytical and statistical processes at any time. Due to the fast growth in technologies area and in amount of data and types of data at the Internet there is slow reaction on it from the legal side of our life which leads to the lack in laws and policies protecting our privacy. It is the goal of international community to jointly update current laws regulating data and information dissemination policy (including at the Internet). How long it would take to arrange joint international actions? [10] We still in need the modern Instruments, which connected to the Internet Governance specificity.

References

1. Information systems; Carnegie Mellon to study how social media and big data affect protection of human rights. (2013). Information Technology Newsweekly, 576.

[10] Komarov, Michael. Big Data leads to new international data processing policies // MIND - Multistakeholder Internet Dialog. Editor-in-chief: W. Kleinwächter. Vol. 7: Privacy and Internet Governance. Berlin: Internet & Society Collaboratory, 2014. P. 64.

2. Civil Rights Principles for the Era of Big Data // CivilRights.org, a project of The Leadership Conference on Civil and Human Rights & The Leadership Conference Education Fund. URL: http://www.civilrights.org/press/2014/civil-rights-principles-big-data.html
3. Big data (A special report) - should the U.S. adopt European-style data-privacy protections? (2013, March 11). Wall Street Journal.
4. The rule of law on the Internet and in the wider digital world. Issue Paper published by the Council of Europe, Commissioner for Human Rights. September 2014.
5. Kshetri, Nir. The Expert Opinion // *Journal of Global Information Technology Management;* 2013; 16, 4; ABI/INFORM Global, pg. 68.
6. Forecast of the long-term socio-economic development of the Russian Federation for the period up to 2030 (designed by Russian Ministry of Economic Development) // ConsultantPlus Legal Database System.
7. ConsultantPlus: Analytical review of the August 5, 2014. Personal data of Russian citizens as a general rule will be processed in Russia (Federal Law of 21.07.2014 N 242-FZ) // ConsultantPlus Legal Database System, 2014.
8. Komarov, Michael. Big Data leads to new international data processing policies // MIND - Multistakeholder Internet Dialog. Editor-in-chief:

W. Kleinwächter. Vol. 7: Privacy and Internet Governance. Berlin: Internet & Society Collaboratory, 2014. P. 64.