BIG DATA, PANDEMIC AND DATA PROTECTION: AN ITALIAN PERSPECTIVE

FABRIZIO CUGIA DI SANT'ORSOLA and **SILVIA GIAMPAOLO** on the need for a holistic approach to data regulation in the ever-more connected global society

The spread of the COVID-19 pandemic and the need to access data to develop adequate public health responses such as quarantine, social distancing and tracking of virus origination has fed a general debate on the need to regulate big data as a key social resource. Big data has already revealed its importance in relation to sustainable growth, urban transportation, environmental programs and for health research however its use for COVID-19 purposes has renewed dialogue on topics such as privacy, security and consumer protection.

In some ways, the debate mingles with more general questions related to the use of big data extracted from the online behaviours and social media interactions of individuals. The revelation from the Cambridge Analytica scandal has been that political brokers and program insiders may easily misuse data for political purposes. Data usage to address the social and medical needs of the population, on the other hand, may be seen as the positive side of the same coin.

From the international regulatory perspective, "big data" still must be defined. While it appears clear that regulating the phenomenon globally on a general scale – as is done for instance for the internet through the Internet Governance Forum (IGF) – would certainly assist in addressing collective topics such as individual rights, proprietary limitations on re-elaborated data or the like, so far there appears to have been no particular "need" for a cutting edge shared definition and regulation of big data on the whole.

In relation to the protection of personal data, with both the General Data Protection Regulation (GDPR)[1] and ePrivacy Directive,[2] lately enacted, the EU has set a solid and trusted legal framework on a global scale. On November 25, 2020, the EU Commission also adopted the proposed regulation on data governance ("Data Governance Act")[3] with the aim of ensuring data processing and data sharing via a new set of identified data intermediaries.

The correct framing and valid expression of individual consent for the use of personal data for the purposes of further data processing or elaboration should also involve a general protection standard on essential individual rights. Big data is a phenomenon which involves a plethora of players, as well as third party use and re-elaboration of data. Under the EU GDPR general principles, the threshold set in favor of individuals is an informed consent involving the direct use and elaboration by the data controller. Yet it appears clear that the aggregation of data performed during big data (re)elaboration necessarily interrupts the link of informed consent The same principles should apply in the case of third party processing or further elaborations of data, yet there appears to be little public awareness of knowledge of the data processing that is being done.

Big data is also a process in which personal data is generally matched with non-personal information (the latter not strictly subject to consent requirements). Processing and further development (such as profiling) generates further data not referred back to individual subjects but to collective samples; a refined and elaborated data product of algorithms, proprietary software or applications. The need for definition and regulation of a new set of responsibilities on data processors under the general accountability principles would appear to be mature.

Following the conclusions drawn from the recent "Triple Report" of the three independent Italian regulatory authorities for privacy, communications and antitrust, this article intends to shed some light on what big data effectively may mean today, and how its regulation appears to be of utmost importance in the ever-more connected global society.

A CRISIS OF THE "NOOSPHERE"?

Confinement by virtue of the global pandemic has provided new evidence that all societies are part of one same ecosystem, physically, socially and digitally. Experience has also further shown that big data is not simply the "aggregation" of multiple personal data, but rather a new product which may be elaborated by proprietary algorithms and artificial intelligence (AI), which is

the reason why a new frame of mind is needed to regulate it from a holistic perspective.

The big data ecosystem, relying fully on the internet as its fundamental source and generating a constant flux of new information, recalls the idea of the "noosphere",[4] the identified layer of knowledge and information "surrounding" the Earth, such as the concept of the "biosphere" first envisaged by Edouard Le Roy and developed by Vladimir Vernadsky and philosopher Teilhard de Chardin.[5] As a "common layer" in which "human and machine super-intelligence" abounds, transforming reality itself, the noosphere may represent the "fossil" (or the remains or impression) of big data.

While the (still important) implications of the noosphere fall outside the scope of this paper, a key point is that the data layer – reflective of individuals and collectives, living and non- living elements, both human and automated – has real-world, interrelated and ever-increasing implications for the Earth and its inhabitants. While the extent and impact of the two-way relationship between the noosphere and biosphere is debatable,[6] policy and regulatory frameworks do not yet conceive of data as an essential resource, for instance, or grapple with the rise of the internet of things (IoT) and machine to machine (M2M) interaction.

Content (in the form of data information) is already self-generated by means of algorithms and self-processing activities. Today only partially do people effectively "create" content over the internet. Technological development beats the drum, and from a general standpoint human activity over the internet is progressively shrinking. Non-human interaction (IoT and M2M) are destined to generate autonomous data spontaneously, up to the point of extending the mining of "personal" data to non-personal data, vitalizing information itself.

The issue lies in ensuring a solid and forward-looking policy and regulatory framework able to protect personal rights in a world of digitalised interactions. Collective wisdom must ensure respect of individual rights in data processing activities and big data regulation should probably differentiate between pro-social needs and self-interested economic entrepreneurship. We are already greatly relying on the Internet as a kind of artificial limb or "prosthesis" of our memories, so avoidance of big data is certainly not the solution: we should take a step further and address what activities regarding big data are useful for ensuring pro-social distribution of knowledge and outcomes, and which activities must be regulated differently in light of their final productive scopes.

This does not mean limiting in any way free entrepreneurship, nor ferrying mankind to a supposed "digital Maoism".[7] Rather, regulation should address and tackle issues prior to the generation of regulatory gaps. We need to understand the phenomenon from all points of view, in order to secure a proactive approach.

> **“A clear-cut general regulatory definition of big data has still not been put forward.”**

BIG DATA ECOSYSTEM & STATE OF THE ART

A clear-cut general regulatory definition of big data has still not been put forward. Part of the difficulty lies in its mutable frontiers: data has become progressively an end product of online connection and is self-generated spontaneously by online terminals. According to the International Data Corporation

⬅ (IDC) the volume of data produced in the world is an unrestrainable flow growing rapidly from 33 zettabytes in 2018 to an expected 175 zettabytes in 2025.[8] Part of the current data produced is not of human origin but relates to terminal interaction (M2M and IoT).

The increased use of the internet by individuals gives way to an unlimited source of data, including geo-localised data, across photo sharing, the posting of comments, payments, emails and other real time activities, which produce digital footprints. Datafication represents a net contributor to the general proliferation of random data circulating on the web.

Digitalised data related to individual activity may be increasingly collected through sensor systems that are pervading daily activities. The tip of the iceberg of this new data proliferation is video surveillance and facial recognition. With the advent of 5G mobile devices containing embedded sensors, we should expect an amplification of side activities on data processing, such as tracking, storage and automation. Such information may fall within an appropriate data definition when linked to accelerometers, gyroscopes, magnetometers, proximity detectors and other sensors able to ensure data subject matching (such as fingerprint readers and facial recognition apps, light sensors, thermometers).[9] Such information is also capable of generating a new generation of automated "big data", strategic in profiling and addressing key targets of commercial campaigns.

New infrastructures and network services are destined to support big data, enabling online collection, transport and storage of data, also produced by third party sources (e.g. IoT, M2M, smart metering, applications and services). This allows for the autonomous generation of data at the network and third-party level, as well as in application/service and operating procedures/processing levels.[10]

In future, 5G mobile networks will lead the collection of big data, working in synergic interaction with AI applications allowing the efficient processing of data related to IoT transmissions. In combination with AI techniques, big data will offer valuable solutions in new areas such as:

- network development: analysis of data quality, traffic and use, and complaints can allow optimisation of network planning and creation processes;
- active network maintenance: collection and analysis of events and alarms coming from networks allowing preventive identification of faults or malfunctions and the carrying out of maintenance before the occurrence of the inefficiency;
- network security management: the collection and analysis of events and alarms coming from the network and their comparison with threshold values to allow identification and management of potential attacks on network security.

PRIVACY AND THE ROLE OF INNOVATION

Italy was of the first nations hard-hit by the COVID-19 pandemic. In a matter of days, in a still unknown medical and epidemiological scenario (still today, the origin of the virus is debated), the Italian Government was compelled to introduce a series of public health and social measures based principally on tracking and positioning of individuals. The adoption of social distancing, quarantine and lockdown were fundamental protective measures, albeit somewhat medieval. For the purposes of limiting the spread of the virus, different layers of information available online were scrutinised to understand protective methodologies. Whilst it appears clear that quarantine helped to curb the spread of the disease, the exercise also demonstrated the crucial role of accessible data for general purposes.

Amidst the outbreak of the virus, Italy introduced innovative tools and measures with regard to the tracking and utilisation of personal data for social and medical needs.[11] On June 1, 2020, the Italian Privacy Authority authorised the use the "Immuni" contact tracing app in response to COVID-19, considering it indispensable to set out protective measures in order to enhance security of data related to the individuals downloading the app, to mitigate risks from processing. The Authority required operators to inform users adequately about operation of the algorithm used to assess exposure risks.[12]

This built on earlier advice, including a clarification from the Italian privacy authority aimed at protecting the fundamental rights of interested parties, that certain GDPR provisions apply to big data, aimed at addressing the potential risks deriving from profiling[13] and decisions based solely on automated processing.[14] In order to mitigate risks occurring from big data activity, data controllers and processors must implement privacy by design and by default set out by article no.25 of GDPR, and adopt security measures processing such as the pseudonymisation and encryption of personal data. Such principles were also re-affirmed by the European Data Protection Board (EDPB) in the Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, adopted on 21 April 2020, which pointed out that the GDPR and Directive 2002/58/EC (the "ePrivacy Directive") both contain specific rules allowing for the use of anonymous or personal data to support public authorities and other actors at national and EU levels in monitoring and containing the spread of COVID-19.

What these developments demonstrate is the need, from a regulatory standpoint and for accountability purposes, for a better understanding of and distinction between personal data, non-personal data and general side-information not necessarily associated with data subjects, but rather self-generated by automated devices connected to the internet.

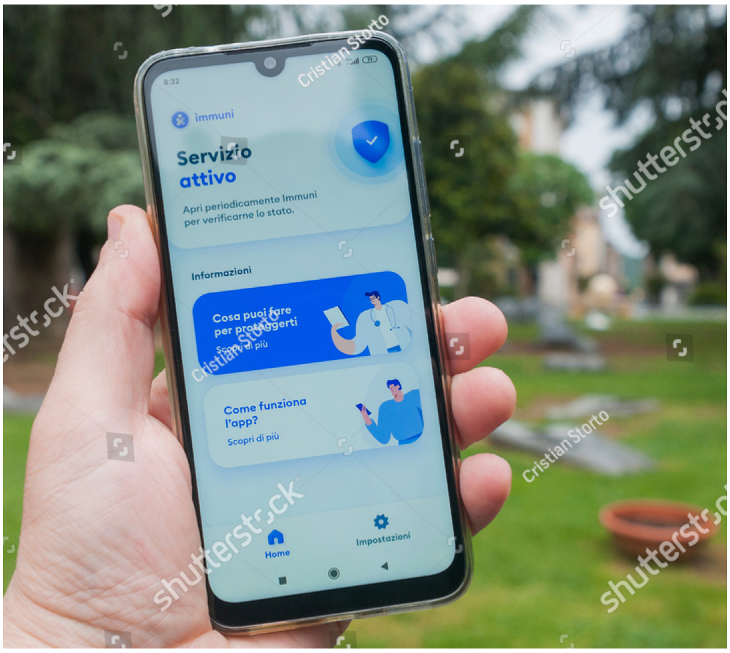
PROBLEMS ASSOCIATED WITH LACK OF DEFINITION

The European Commission states that "big data" has become a key asset for the economy and society, since "good" processing of data may lead to "innovations in technology", "bring opportunities to more traditional sectors such as transport, health or manufacturing" and "transform Europe's service industry" among other things.[15] This rather naïve view of the big data phenomenon appears to beckon allegiance to a laissez-faire, "light touch" regulatory approach, mixing self-fulfilling targets with the current state of the art. Implications for consumer protection, privacy and antitrust are less clear as purposes and objectives may evolve in the same elaboration of data.

Very often data is used by third parties or in a secondary mode, where new data is added and made richer, or following requests of services, or because processing is necessarily performed via other operators in order to address different questions and issues involving the use of the same data. The more removed from the original transaction (where consent may be provided in an informed way) the use of data becomes, the more regulatory issues arise.

Rights. Processing of data may collide with individual rights if utilised inconsistently with the original consent provided, yet it may happen that further processing is not performed by the original data controller. For instance, profiling and online targeting for the needs of political campaigns have fed fake news as hooks and instruments for achieving consensus using internet trends identified by means of AI systems. In ecommerce and online trading, big data possession already represents a key element of competitive advantage, ensuring privileged access to online behaviours capable of being used to target marketing campaigns.[16] A more consistent and universally shared regulatory definition of the utilisation of big data appears appropriate, for the purpose of ensuring a framework of clear and transparent scope, a level playing field among competitors and equal access to resources.

Actors and activities. Industrial processing of large amounts of data implies different types of activities, the most preeminent of which consist in aggregation of data sources, retrieval, processing, treatment, profiling, storage and re-elaboration of data. Such activities differ substantially, implying the need to differentiate applicable regulation for actors involved. The general use of cloud computing in data storage and the intermediate role of "software as a service" (SAAS) players diffuses the ambiguity around different actors, such that "better, clearer and unambiguous rules are desperately needed on applicable law", with regards to data protection.[17] Big data activities may differ according to the diverse sources and retrieval procedures adopted by operators, for instance M2M or IoT information that may (or may not) be linked with personal data. The possession of matching side information on data subjects may determine data re-elaboration which may require prior informed consent. Also, personal data and subject information may be retrieved, processed and re-assembled by means of proprietary AI systems, capable of identifying reliable data sources and developing matches. Data is thus transformed, subject to "ownership" by the elaborator under intellectual property principles.



Italy, June, 2020: The contact tracer app Immuni, developed for the Italian government, to monitor COVID-19 infections in the area

Antitrust. From an antitrust perspective, keeping in mind that big data processing generally implies the use of personal and non-personal information, the issue is not so much the aggregation of data as such, but rather targeting effectiveness and proprietary use in commercially aggressive practices based on profiling or behavioral advertising of data subjects. Following the definition of Gartner of 2001 (according to which big data is "high-volume, high-velocity and/or high-variety information assets requiring new forms of processing to enable enhanced decision making, insight discovery and process optimization"[18]), regulation should take into account whether big data may determine anticompetitive advantages or abusive practices.

Privacy. From a privacy perspective, big data processing involves the collection, analysis and accumulation of large quantities of personal and non-personal data, yet only the personal information of relating to an identified or identifiable natural person (EU "data subject") falls within the protections afforded under the GDPR.

Non-personal data as such falls outside GDPR and within the regulation on the free flow of non-personal data (FFD).[19] With regards to "non-personal data" and following the expansion of IoT, AI and machine learning, big data processing may imply generation of content by means of matching data with non-personal information. The proliferation of non-personal data means the traditional distinction between "personal data" and "non-personal data" appears more and more obsolete. Modern elaboration and aggregation of data renders it extremely difficult to establish in fact on a preliminary basis (ex ante) whether information pertaining to an individual may effectively fall within the "personal" or "non-personal" category, being the nature of such information dependent on the amount of data stored and aggregated, as well as the context used and reference to relevant technologies. Some psychometric techniques can easily gather sensitive individual information (such as political orientation, drug addiction, etc.) from a simple set of non-personal data.[20] In its retrieval and aggregation activities, non-personal data may transform itself into personal, and thus fall within different regulation.

THE TRIPLE REPORT

By coincidence, a few months before the outbreak of COVID-19, the Triple Authority Report on big data online activities was published, following an ambitious joint inquiry by three Italian regulators into the general big data phenomenon: the competition authority Autorita' Garante della Concorrenza e del Mercato (AGCM); the communications authority Autorità per le Garanzie nelle Comunicazioni (AGCOM); and the personal data protection authority, Garante per la protezione dei dati personali (GPDP).

In July 2019, AGCM, AGCOM and the GPDP reached a common view on the big data phenomenon and in February 2020 the final document "Indagine conoscitiva sui Big Data" (Big Data fact finding survey), was published. Preceding this was the June 2018 publication of two preliminary surveys of significant importance: a "Big Data Interim Report"[21] and the report of a survey carried out by the AGCM alone, aimed at understanding online consumers' propensity to allow the use of personal data in exchange of online services.[22]

In the Triple Report, the big data ecosystem is defined and characterised by the interacting activities of a plethora of differentiated actors, such as:

- subjects generating data (data providers);
- suppliers of technological equipment, typically in the form of data management platforms;
- utilisers, i.e. operators utilising and processing big data to create added value;
- data brokers, i.e. organisations collecting data from a set of sources, both public and private, offering them, upon payment, to third party organisations;
- companies and research organisations, whose activities are fundamental for developing new technologies and new algorithms by exploring data and extracting value;
- public bodies, both market regulators and entities involved in public administration activities, focused on improving products and services offered to the citizens in view of increasing public interest.

Interactions among such operators determine a market structure in which (few) large multinational companies (such as OTTs), with a high degree of vertical, diagonal and horizontal integration in all (or almost all) phases of the ecosystem, operate alongside a myriad of small specialised businesses that often, after the period of "start-up", tend to be acquired by larger ones. The ecosystem appears characterised by the presence of several forms of incomplete contracting, implicit markets (i.e. in which the bargaining of the asset takes place in spurious manners) as well as notional areas characterised by perfect vertical integration and potential market demand.

According to the survey, the ecosystem is characterised by a fundamental market failure, capable of undermining the social, static and dynamic efficiency of the entire big data production chain, as well as by the existence of barriers to entry, in particular at the first stage of data collection and acquisition, due to technological, regulatory or economical pitfalls. Barriers are typically found in the form of operating systems, search engines and social networks, along with data storage activities performed by data centres. In this context, the general data market appears to be converging towards market concentration.

The analysis also highlights a clear interaction between online pluralism and competition: the traditional approach to pluralism is to work on the offer side, promoting pluralism by means of pluralist offer. The conclusion drawn by the Triple Report is that the web economy has now changed this framework: scarcity comes from the demand side on the web, as there is an overload of information and algorithms select what information people receive.

In its fact-finding survey, the AGCM also highlighted a fundamental lack of awareness among digital users about the use of their data. The survey showed that four out of ten users are unaware of the fact that their online actions generate data used to analyse and predict online behaviours.

The Report further underlines the frequent inverse correlation between application (app) pricing and authorisation requests to users, stressing the need for users to be made aware, during purchase decisions and data transfers, of the connection between consent mechanisms and the need for further authorisations.

An ex ante approach is suggested in the Triple Report with regards to possible regulation of algorithms, taking into account moments and methods of data acquisition (data gathering & storage), functioning of algorithms (algorithm accountability), methods of conservation and analysis (data analytics) and deriving (primary and secondary) uses of data, so to pursue consumer welfare with the aid of antitrust law tools against anticompetitive practices of major digital firms facilitated by software and proprietary algorithms.

The recommendation is for new legislation able to coordinate with regulation on new technologies – including AI and machine learning, IoT and M2M – given that data collection, storage and analysis are now activities embedded in terminal equipment. With regards to the assessment of market power, the Triple Report clearly identifies the special importance of vertical and conglomerate integration. Data collection, management, processing and profiling are separate and critical new monetary tools, and privacy has become a qualitative service, able to limit market power. Regulating the bargaining activity between platforms and intermediaries (possibly ex ante) may be the correct approach.

EUROPEAN STRATEGY

The European strategy for data and White Paper on artificial intelligence are the first pillars of the new strategy of the European Commission (EC) on big data.[23] They are founded on the assumption that data is a crucial resource for economic growth, competitiveness, innovation, job creation and social progress, driving productivity and resource efficiency across all sectors of the economy, and allowing for more personalised products and services to improve health and wellbeing as well as enabling better policy making.

The strategy aims at creating a single market by 2030 that will ensure Europe's global competitiveness and data sovereignty. In order to fulfil this ambition, the EU will build on a strong legal framework – in terms of protection of personal and non-personal data, fundamental rights, safety and cybersecurity – protecting its internal market. The EU cybersecurity certification framework[24] and the EU Agency for Cybersecurity (ENISA)[25] are expected to play an important role towards such endeavor. The EU data strategy is based on four pillars:

i) a cross-sectoral governance framework for data access and use;

ii) investments in data and strengthening Europe's capabilities and infrastructures for hosting, processing and using data;

iii) competences: empowering individuals, investing in skills and in SMEs;

iv) common European data spaces in strategic sectors.

The European Commission's proposed Digital Governance Act of 25 November 2020,[26] while intending to facilitate data sharing across the EU and ensuring data brokerage via intermediaries of different forms (data trusts, data cooperatives and data stewardships), still leaves apparently unresolved the issue of accountability and transparency on such operators.

The Digital Governance Act has the merit of setting common goals and general standards and clarifying the general interests involved in processing big data. However, the final approach must necessarily take into account, and be integrated with, the aim of promoting the application of AI systems and mitigating the risks that AI may entail, such as opaque decision-making and gender-based or other kinds of discrimination. The use of AI can impinge on the values on which the EU is founded and lead to breaches of fundamental rights, including the right to freedom of expression, freedom of assembly, human dignity and non-discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.

The European Commission's proposed Artificial Intelligence Act of 21 April 2021[27] correctly identifies in this respect a set of clear requirements and obligations regarding specific uses of AI.

The new framework is destined to adjust the existing legislative framework relating to AI applications linked to the use of big data.

CONCLUSION

Big data has already evolved as a crucial resource and the pandemic has given greater urgency to the need for a global regulation of this phenomenon. Data generation per se is not the problem: when purchasing behaviour is used by an online counterpart, or when a commercial online transaction transfers personal data with consent, there appears to be no underlying problem with personal rights. When it comes to collective rights however, and in view of the evolution of the internet economy, regulation of online pluralism may be the solution.

As the Triple Report has confirmed, the key to regulating big data is in the correct interplay between online pluralism and competition. Voluntary actions of platforms against disinformation, fake news and in providing fact-checking tools on the demand side with respect to content are also a solution for ensuring that big data processing is performed under transparent ethical standards and mandatory codes of conduct. Yet a more holistic approach is needed: regulation should attempt to limit the predatory use of data and monetisation models which appear to conflict with the transparent use of this critical resource.

***FABRIZIO CUGIA DI SANT'ORSOLA** and **SILVIA GIAMPAOLO** are partners at Cugia Cuomo & Associati, Rome, Italy.*

REFERENCES **1** Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). **2** Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications. **3** (COM(2020) 767 final European Commission (2020). Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 25 November. https://wayback.archive-it.org/12090/20210728140404/https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0767 **4** Oleg G, Suleimanov I, and Vituleva E (2019). Artificial intelligence in the context of noosphere studies. Conference: SCTCMG 2019 – Social and cultural transformations in the context of modern globalism 2019, December. DOI:10.15405/epsbs.2019.12.04.130 **5** "Noosphere" is a term first introduced by Édouard Le Roy (1870–1954), Vladimir Vernadsky (1863-1945) and Teilhard de Chardin (1881-1955) referring to the sphere of human thought (Levit, in 2000, further developed the concept, ensuring a factual consistency and autonomy to noosphere). Edouard Le Roy was the first who used the notion "noosphere" in a publication entitled "L'exigence idéaliste et le fait l'évolution" published in 1927. **6** Lahoz-Beltra R. The 'crisis of noosphere' as a limiting factor to achieve the point of technological singularity. https://arxiv.org/pdf/1405.3378.pdf; Kurzweil (2005). The singularity is near. Viking, 1-652; Le Roy E (1927). L'exigence idéaliste et le fait de l'évolution; Lovelock JE (1979). Gaia A new look at life on Earth; Teilhard de Chardin P (1959). The Phenomenon of Man; Teilhard de Chardin P (1964). The Future of Man. **7** Lanier J (2006). Digital Maoism: the hazards of the new online collectivism. https://www.edge.org/conversation/digital-maoism-the-hazards-of-the-new-online-collectivism. **8** IDC (2018). Data age 2025, November; European Commission (2020). A European strategy for data, 19 February. https://wayback.archive-it.org/12090/20210727031936/https://digital-strategy.ec.europa.eu/en/policies/strategy-data **9** AGCOM (2018). Cognitive survey concerning the development prospects of wireless and mobile systems towards the fifth generation (5G) and the use of new spectrum portions above 6 GHz, 5 March. **10** AGCOM, GPDP, AGCM (2020) Indagine conoscitiva sui Big Data, p.43. **11** A series of Prime Minister's Decrees were adopted between March and May 2020. **12** GPDP (2020). Green light to the 'Immuni' contact tracing app by the Italian SA. 1 June. https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9356588#english **13** Article 1, par1 lett 4 of GDPR: profiling means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements. **14** Article 22, par 1 of GDPR: The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. **15** European Commission. Shaping Europe's digital future. https://ec.europa.eu/digital-single-market/en/big-data **16** Note that re-elaboration of data may in itself be the by-product of a registered competitive advantage, perfectly legitimate if developed in a transparent fashion. In some cases, consumer prediction and behavioural advertising may be the product, not of elaboration of data or data poaching as such, but rather derived from commercial insight or experience, or even from a commercial bargain agreed among consumers and operators. **17** LRPD Kantor Ltd in association with Centre for Public Reform. New Challenges to Data Protection – EU Commission 2010 in Millad C (ed) (2013). Cloud computer law. **18** As early as 2001, the company META Group, later become Gartner, highlighted in a report the critical aspects related to data management by focusing on three dimensions: volume, speed and variety: Laney D (2001). 3D data management: controlling data volume, velocity, variety. META group report, file 949. Subsequently, in 2012, the definition was coined in a new report. Beyer MA and Laney D (2012). The importance of big data: a definition. Gartner, analysis report ID: G00235055. **19** Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union. **20** Agcom clearly identifies the possibility of data modifying its own regulatory nature (personal or non-personal) by means of the same elaboration made by AI systems, in a somewhat peculiar online mutation. **21** AGCOM, AGCM, GPDP (2018). Big data Interim report as part of the fact-finding survey pursuant to resolution no. 217/17/CONS, 8 June. **22** AGCM (2018). Fact-finding Survey on Big Data, 8 June. https://en.agcm.it/en/media/detail?id=6a0face7-79ce-44dc-ab72-eab9623970af&parent=Press%20releases&parentUrl=/en/media/press-releases **23** European Commission (2020). A European strategy for data, 19 February. https://wayback.archive-it.org/12090/20210727031936/https://digital-strategy.ec.europa.eu/en/policies/strategy-data and https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf; European Commission (2020). White paper: On artificial intelligence - a European approach to excellence and trust, February. https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf. **24** See https://digital-strategy.ec.europa.eu/en/policies/cybersecurity-certification-framework. **25** See https://www.enisa.europa.eu/topics/standards/certification. **26** European Commission (2020). Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 25 November. https://wayback.archive-it.org/12090/20210728140404/https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0767 **27** European Commission (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, 21 April. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206