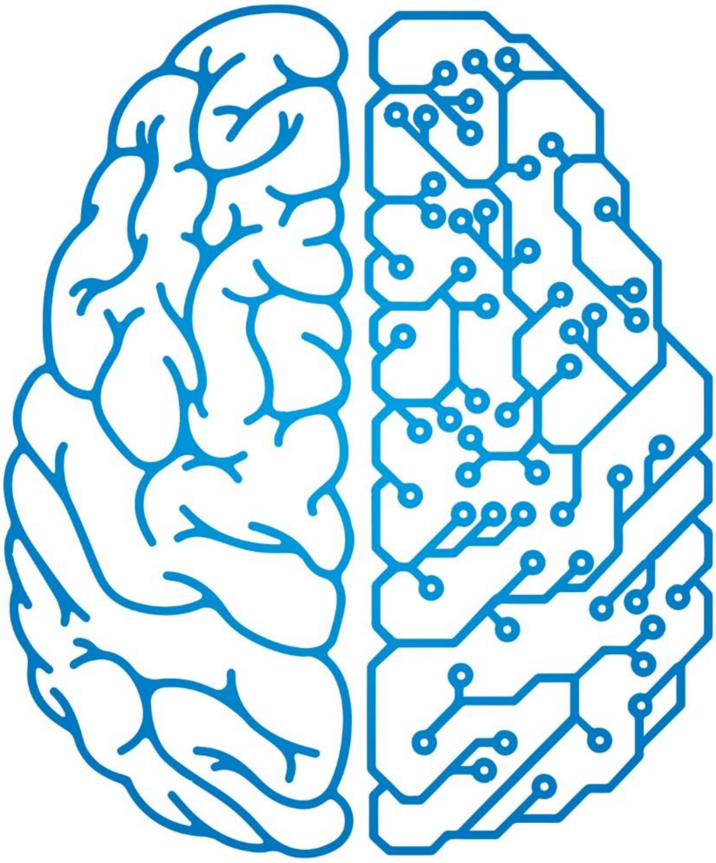
38TH INTERNATIONAL PRIVACY CONFERENCE Marrakech 2016

EDPS



Artificial Intelligence, Robotics, Privacy and Data Protection

Room document for the 38th International Conference of Data Protection and Privacy Commissioners

October 2016

1. Executive summary

Artificial intelligence and robotics are increasingly a reality and also present in the political agenda. Due to the interest of the topic, it is being discussed in the ICDPPC 2016.

Some people may claim it is too early to start discussing artificial intelligence and robotics but we can see the applications of artificial intelligence and robotics already and the data protection authorities need to start discussing them and developing a position.

We need to adopt a realistic approach, neither luddite[1] nor evangelical. A proper consideration will not slow down innovation but provide sound foundations for these technological developments.

By default the data protection framework is mainly applicable to controllers, few provisions apply to the whole artificial intelligence/robotics ecosystem: Data protection by design and by default can only be a reality if all the actors involved apply data protection principles.

This is a background document to stimulate and contribute to the discussion within the closed session of the ICDPPC 2016.

2. Topics

The following sections explore some popular artificial intelligence and robotics topics which are relevant for a discussion about the impact of artificial intelligence and robotics on data protection and privacy. Each topic is introduced briefly and some questions for reflection and discussion are presented.

(This page intentionally left blank.)

2.1. Big data, profiling and automatic decision making

'Big data'[2] refers to the practice of combining huge volumes of diversely sourced information and analysing them, often using artificial intelligence – machine learning solutions, to provide insight. One of the greatest values of big data is derived from the monitoring of *human* behaviour, collectively and individually, and its predictive potential[3].

The relation between artificial intelligence and big data is bi-directional: Artificial intelligence, through machine learning, needs a vast amount of data to learn: data in the realm of big data considerations. On the other direction, big data uses artificial intelligence techniques to extract value from big datasets.

One of the main issues regarding big data is information to individuals: transparency. Unless individuals are provided with appropriate information and control, they '*will be subject to decisions that they do not understand and have no control over*'[4]. Having that appropriate information can be complicated by two different factors: organisations claiming secrecy over *how* data is processed on grounds of trade secrets and the intrinsic difficulty in providing an explanation for a prediction when that prediction is based on an artificial intelligence algorithm that has been created using machined learning: the logic behind the machine reasoning may not be *expressible* in human terms.

Another very important concern with regards to artificial intelligence and big data is the *bias induced* via the input dataset provided for training the artificial intelligence. As the machine learns from the information provided and has no means to contrast that information with a bigger picture, whatever bias is contained in the training set will influence the predictions made. If those predictions are used to take decisions, a vicious circle of self-fulfilling prophecies can be created where the feedback the machine receives reinforces the bias present in the first place.

When machine learning is used to process big data, for DPAs to be able to look into the *black box*[5] of the algorithm is not useful. The analysis needs to be done on the machine learning process itself and on the data feed (to detect a possible bias). Indeed, we may come to the situation where for analysing a particular *algorithm* or artificial intelligence, a DPA may need to create another one, or use some analytics tool, to confirm that the model created in the first place is *fair.* This could lead to creating a self-referencing problem difficult to solve.

Elements for reflection

- How can DPAs support the right to information from the data subject when confronted with big data, artificial intelligence and machine learning?
- How to evaluate the bias in automated decisions when artificial intelligence and machine learning is used?
- How can DPAs supervise appropriately an organisation using intensively big data, artificial intelligence and machine learning?
- Should DPAs create their own pool of artificial intelligence experts and resources to be able to re-create and analyse the models used by the organisations under supervision?

References

- FTC *Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues* (6 January 2016) - https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report
- EDPS *Meeting the challenges of big data, A call for transparency, user control, data protection by design and accountability* (19 November 2015) - https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-11-19_Big_Data_EN.pdf
- EDPS *Towards a new digital ethics: Data, Dignity and Technology* (11 September 2015) - https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-09-11_Data_Ethics_EN.pdf
- WP29 *Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU* (16 September 2014) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf
- Norwegian Data Protection Authority *The Chilling Effect in Norway* (January 2014), - http://www.datatilsynet.no/Global/04_planer_rapporter/Nedkj%C3%B8ling%20i%20norge_eng_.pdf
- Office of the Australian Information Commissioner (OAIC) *Consultation draft: Guide to big data and the Australian Privacy Principles* (May 2016) - https://www.oaic.gov.au/resources/engage-with-us/consultations/guide-to-big-data-and-the-australian-privacy-principles/consultation-draft-guide-to-big-data-and-the-australian-privacy-principles.pdf

2.2. Image recognition

Determining which *objects* are present in an image, whether a static one (photo) or in a sequence video, is a classical problem in computer vision. This capability can be further developed as being able to recognise a particular object through several images or in video, detection of certain objects or circumstances, etc. Some actual practical applications are photo tagging, counting people in public spaces or, a very interesting one, facial recognition.

Currently, the best algorithms for image recognition are based on convolutional neural networks which in turn are a specific implementation of machine learning. The performance of these algorithms is now close to that of humans. As these algorithms are based on machine learning, they depend on the amount of training data (pictures tagged) that they can use. Fortunately for them, tagged pictures on the Internet are quite abundant.

Using facial recognition is possible to identify a person from a digital image or a video. This is achieved by detecting a *face* in the image or video and comparing it with a database containing both face pictures and *metadata* associating the picture with a person. Our face, like our fingerprints, is a biometric identifier: our facial characteristics and the proportions of our head do not change. As for fingerprints, specific characteristics are extracted (minutiae), for face recognition, the same process is applied (i.e. measuring nodal points on the face, such as the distance between the eyes or the shape of the cheekbones).

Currently there is an ongoing debate on the privacy implications and surveillance possibilities. With the widespread use of CCTV, the amount of video sources for identifying people is increasing constantly. For example, some airports are considering installing one of these systems to improve their security, although previous experiences have not produced the expected good results. Recently, the focus has shifted to its use in border control.

Elements for reflection

- What should be the policy on using publicly available information for training image recognition algorithms based on machine learning?
- The combination of face recognition with camera-equipped drones and easily accessible tagged photos makes a very powerful surveillance system available for everyone; are we reaching the limits of the *household exception*?
- How to supervise face recognition used for security or intelligence purposes?

References

- WP29 *Opinion on developments in biometric technologies* (27 April 2012) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp193_en.pdf
- WP29 *Opinion on facial recognition in online and mobile services* (22 March 2012) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp192_en.pdf
- EPIC *Facial recognition* - https://epic.org/privacy/facerecognition/
- CDT *Seeing Is ID'ing: Facial Recognition & Privacy* (22 January 2012) - https://cdt.org/files/pdfs/Facial_Recognition_and_Privacy-Center_for_Democracy_and_Technology-January_2012.pdf
- Dutch DPA *Letter to Google regarding Google Glass* (18 June 2013) - (https://autoriteitpersoonsgegevens.nl/sites/default/files/downloads/int/med_20130618_letter-to-google-regarding-glass.pdf
- Office of the Privacy Commissioner of Canada *Automated Facial Recognition in the Public and Private Sectors* (13 March 2014) - http://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2013/fr_201303/
- Privacy Commissioner for Personal Data (Hong Kong) *Collection and Use of Biometric Data* (13 April2016) - https://www.pcpd.org.hk/english/news_events/speech/files/HKGCC_Luncheon_20160413.pdf

2.3. Natural language processing

Natural language processing is a research area aiming to provide computers with the capacity to interact with persons using natural languages. Natural language processing involves natural language understanding and generation.

The first natural language processing systems were based on complex sets of rules designed by researchers. Since the late 1980s the approach changed with the introduction of machine learning algorithms. Current solutions implement a type of machine learning which needs vast amounts of data which is not a problem thanks to the Internet.

There are many products on the market using natural language processing. Some of the most popular are smartphone assistants like Google Now, Apple Siri or Microsoft Cortana[24], or automated translation services like Google Translate or Bing Translator[6].

Natural language processing systems may have the following capabilities (the list is far from exhaustive):

- Machine translation: Automatically translate text from one human language to another. This is quite a complex problem requiring all possible capabilities from a natural language processing system.
- Natural language understanding: Converting (written) natural language into a formal representation easier to manipulate by a computer.
- Question answering: Being able to answer a question formulated in natural language. Questions can be specific or open-ended.
- Information extraction: The extraction of semantic information from a text.
- Sentiment analysis: Extract subjective information usually from a discourse, written or oral (through speech recognition). This is a quite controversial area of research with a big impact on marketing or political science.
- Speech recognition: Extract the textual representation of an oral speech. This is also a very difficult problem due to the particularities of oral language.

IBM Watson[24] is a very good example of a system combining several of the capabilities mentioned above: natural language understanding, question answering and information extraction.

Besides those impacts on personal data and privacy common to all machine learning technologies, natural language processing also opens a possibility for data protection authorities to use these new technologies when performing their supervision responsibilities.

Elements for reflection

- Natural language processing opens the possibility of processing *unstructured data*, will we see an erosion on purpose limitation for personal data stored in (old) documents?
- Natural language processing uses like question answering and *customer-like* interactions, will they foster automated decisions?
- How will natural language processing alter the balance between *metadata* and actual data?
- Could DPAs use this technology? Some examples: Interactions with data subjects, pre-analysis of mandatory reporting by controllers (DPIAs, data breach notifications), law and jurisprudence analysis and query...

References

- D. Hovy , S.L. Spruit, *The Social Impact of Natural Language Processing* http://www.dirkhovy.com/portfolio/papers/download/ethics.pdf
- N. Kasch, *Text Analytics and Natural Language Processing in the Era of Big Data* (24 October 2014) https://blog.pivotal.io/data-science-pivotal/features/text-analytics-and-natural-language-processing-in-the-era-of-big-data

2.4. Autonomous machines

An autonomous machine (or autonomous robot) is one that is able to operate with a high degree of autonomy. This makes these machines particularly desirable, e.g. in dangerous or inhuman environments or for performing taxing tasks. For a machine to be autonomous it needs to: (1) perceive and react to its environment; (2) plan and realise pre-planned tasks; (3) operate without human intervention (including supplies and maintenance) and (4) be able to navigate a, sometimes human, environment. An autonomous machine may also learn from its own experience or through reprogramming.

One of the most important requirements for autonomous machines is to avoid hurting people or objects in their operational environment (unless on purpose.)

Probably the most widespread example is the home cleaning robot Roomba by iRobot although some other uses are being tested like home delivery or the upgrade of the factory robot (Baxter[7]).

All features of artificial intelligence may be applied in autonomous machines: natural language processing allows the direct interaction between humans and machines, image recognition is a powerful tool which allows robots to understand their environment and all this is supported by machine learning.

Autonomous machines can be considered artificial intelligences with physical bodies able to interact physically with their surrounding world. As such, from an impact or consequences point of view, they represent the apex of the artificial intelligence discussion. An extreme example could be an autonomous weapons system[8] capable of operating autonomously even up to the point of selecting targets: the combination of that capability with face recognition could create the ultimate assassin.

Elements for reflection

- When implemented in autonomous machines *automated decisions* can have an even greater impact. How could the data protection/privacy framework for automated decisions be applied to autonomous machines?
- Who is the data controller for an autonomous machine with self-learning capabilities?
- Should the data protection/privacy community translate the legal framework into *machine readable* law?

References

- U. Pagallo, *Robots in the cloud with privacy: A new threat to data protection?* (October 2013) https://www.researchgate.net/publication/259123308_Robots_in_the_cloud_with_privacy_A_new_threat_to_data_protection
- U. Pagallo, *What Robots Want: Autonomous Machines, Codes and New Frontiers of Legal Responsibility* (12 March 2013) http://link.springer.com/chapter/10.1007/978-94-007-6314-2_3
- L. Edwards and A. Winfield, *Regulating Robots* (15 November 2011) https://www.strath.ac.uk/media/faculties/hass/law/cilp/strath_robot_launch.pdf

2.5. Self-driving cars

Self-driving cars are probably the most popular example of an autonomous machine. Also, it is one of the best cases to reflect on the ethical dimensions of artificial intelligence and robotics. Self-driving vehicles will change the way individual travel is used and organised, and may blur the difference between private and public transport. The artificial intelligences steering the cars will govern decisions which may directly concern the physical integrity and even the life or death of individuals.

More formally, a self-driving car (also called driverless car or autonomous car) is a vehicle that is capable of navigating its environment according to a predefined objective without human input.

There are many potential advantages like:

- Reduction in collisions caused by human error.
- Improved capability to manage traffic flow and the possibility to eliminate certain externalities to *human vehicle driving* like traffic police or road signals.
- Changes in vehicles interior as no driving interface would be needed anymore.
- No more time consuming driving for both professional and personal reasons. (As a consequence professional driver jobs will be lost.)
- Higher speed limits and roadway capacity.
- New business models: car fleets self-operated; extinction of the *privately-owned car;* automatically shared private cars when not in use, etc.

For the possibilities to materialise several obstacles need to be overcome. Among them:

- The absence of a specific legal framework, also regarding liability and insurance. The new legal framework will need to take into account new ownership models.
- Substitution of the old *human-driven* fleet of vehicles to fully benefit from the technology.
- Individuals not wanting to relinquish their cars or driving them.
- As artificial intelligences existing on a substrate of computing power, these machines will be open to attack and intromission.
- The road infrastructure will need to be adapted to the specific requirements of self-driving cars to fully exploit their advantages.

From a privacy or data protection point of view, as we see continuously through this document, the biggest implication is the use of data, most of the time personal data. Self-driving cars need as perfect as possible cartography to operate but also as much information as possible on other vehicles and their trips (e.g. for congestion management.) If we consider this need for information together with new business models and the sensitivity of geolocation information in certain cases we are providing certain economic actors with an incredible profile of our daily whereabouts.

Elements for reflection

- *All questions applicable to autonomous machines apply to self-driving cars as well.*
- How to regulate self-learning machines (self-driving cars) processing huge amounts of geolocation data?
- What will be the impact of new business and ownership models on data subject/data controller/data processor relations?
- There are plenty of ethical considerations on self-driving cars (as the most current and popular example of autonomous machines), how are self-driving cars going to impact fundamental rights close to privacy and data protection like freedom of expression or freedom of association?

References

- US Department of Transportation *Federal Automated Vehicles Policy* (21 September 2016) - https://www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf
- Shearman & Sterling, *Connected Cars and Self-Driving Cars: Not on Auto Pilot in Terms of Legal Risks* (11 July 2016) http://www.shearman.com/~/media/Files/NewsInsights/Publications/2016/07/Connected-Cars-and-SelfDriving-Cars-Not-on-Auto-Pilot-in-Terms-of-Legal-Risks-PDP-071116.pdf
- Freshfield Bruckhaus Deringer, From Connected to Self-Driving Vehicles: the Regulatory roadmap (l.a. 2016) http://www.freshfields.com/en/global/auto/regulatory_roadmap/?LangType=2057
- B. Camarda, *New guidelines: cybersecurity, privacy and your self-driving car* (21 September 2016) https://nakedsecurity.sophos.com/2016/09/21/new-guidelines-cybersecurity-privacy-and-your-self-driving-car/

2.6. (Semi-)autonomous/unmanned aircraft systems

Popularly called d*rones*, although more appropriately called *unmanned aircraft systems* (UAS) or *remotely-piloted aircraft system* (RPAS)[9] depending on their autonomy, these are aircraft systems that can fly without requiring an on-board pilot. Currently drones serve mainly military purposes, but are increasingly used for purposes of surveillance, mapping, transportation, logistics and public security thanks to the sensors they carry such as cameras, microphones, GPS, which may allow the processing of personal data.

Whether *manned* or *unmanned*, drones can be used for several tasks:

- By companies, public authorities and professionals to monitor large-scale infrastructures such as bridges, energy plants (including nuclear ones), railways; apply pesticides on agricultural land; inspect electricity networks; carry out aerial mapping; monitor a concert zone; secure an area; deliver pizzas or books ordered; take wedding pictures or report on an event.
- Law enforcement uses, such as search and rescue; disaster response; border control/protection; civil protection; aerial surveillance; traffic monitoring; observation and pursuit of criminal suspects or observation of civil unrest.
- Military ones able to carry out missions such as surveillance, reconnaissance and airstrikes.
- Private uses by citizens as a hobby, such as model aircraft activities, photography, information technology.

Drones should be distinguished from aeroplanes and CCTV because their *mobility and discretion* enable them to be used in many more circumstances. Besides, when combined with the different sensors mentioned previously they become potentially powerful surveillance tools.

The intersection between drones and artificial intelligence may occur at several levels: drones may collect the information to be processed by an artificial intelligence algorithm remotely, drones already implemented *intelligent reflexes* to make themselves easier to control for their human pilots; or drones can equip an autonomous *artificial intelligence* themselves rendering human intervention, besides providing general instructions,[10] unnecessary (an example of an autonomous machine).

Elements for reflection

- What are the most pressing points with regard to drones from a data protection and privacy point of view?
- How to effectively control these *flying surveillance machines*?
- Linked to *automated decisions*, how to supervise autonomous drones taking their own decisions on the basis of the personal data they have collected?
- Should DPAs have their own fleet of drones for surveillance of other drones? Drones *anti-drones*?

References

- WP29 *Opinion 01/2015 on Privacy and Data Protection Issues relating to the Utilisation of Drones* (16 June 2015) - http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2015/wp231_en.pdf
- EDPS *Opinion on the Communication from the Commission to the European Parliament and the Council on "A new era for aviation - Opening the aviation market to the civil use of remotely piloted aircraft systems in a safe and sustainable manner"* (26 November 2014) - https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2014/14-11-26_Opinion_RPAS_EN.pdf
- Office of the Privacy Commissioner of Canada *Privacy Implications of the Spread of Unmanned Aerial Vehicles (UAVs) In Canada* (18 September 2014) - http://www.priv.gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/completed-contributions-program-projects/2013-2014/p_201314_10/
- Irish Data Protection Commissioner *Guidance on the use of Drones* (December 2015) - https://www.dataprotection.ie/docs/Guidance-on-the-use-of-Drone-Aircraft/1510.htm
- Office of the Privacy Commissioner for Personal Data (Hong Kong) *Guidance on CCTV Surveillance and Use of Drones* (March2015) - https://www.pcpd.org.hk//english/resources_centre/publications/files/GN_CCTV_Drones_e.pdf

3. What to do beyond current law?

There is a lot of power in artificial intelligence and robotics, creating endless possibilities for the best and for the worst. However, technology cannot dictate our values and rights. In today's environment, adherence to the law is not enough; we have to consider the ethical dimension of technologies like the ones presented in this document which are very dependent on the processing of data, most of the time personal data. Regardless of the regulatory framework, there are deep questions as to the impact of these technologies on dignity and individual freedom.

3.1. The ethical dimension

The Universal Declaration of Human Rights takes as its starting point the inviolability of human dignity. The dignity of a person is not only a fundamental right in itself but also is the foundation for subsequent freedoms and rights, including the rights to privacy and to the protection of personal data. Violations of dignity may include objectification, where a person is treated as a tool serving someone else's purposes.

In the early 21st century, individuals are increasingly disclosing, and being required to disclose, much more personal information over the Internet in order to participate in society. Digital profiles can be created thanks to artificial intelligence techniques and shared in microseconds without the individual's knowledge, and, applying again artificial intelligence, used as the basis for important decisions.

The use of artificial intelligence to predict people's behaviour risks stigmatisation, reinforcing existing stereotypes, social and cultural segregation and exclusion[11], subverting individual choice and equal opportunities.

Meanwhile, the combination of artificial intelligence and robotics and a continued state of exception on grounds of *security* provides multiple layers of intrusive and intelligent techniques for monitoring individuals' activity[12]. Understanding this *surveillance ratchet* and its relation with artificial intelligence, both as an enabler and a driver, requires a long-term perspective on the overall effects on society and individuals' behaviour.

All parties need to look hard at how to ensure that these values are not merely respected on paper while effectively being neutralised in cyberspace. With regard to artificial intelligence and robotics, we now have a *critical window* to build the right values into them before the mass adoption of these technologies[13] happens. This requires a new assessment of whether their potential benefits really depend on the collection and analysis of the personally-identifiable information of millions of individuals. Such an assessment could challenge researchers to design solutions on the basis of a different paradigm than machine learning or to restrict the use of personal data.

The changes we can expect from artificial intelligence and robotics will make the existing framework fail if we do not approach the future with innovative thinking. Truly independent and knowledgeable data protection authorities have a crucial role in preventing a future where the life of individuals is determined by artificial intelligences living above us in the cloud.

3.2. The technical dimension

Human innovation has always been the product of activities by specific social groups and specific contexts, usually reflecting the societal norms of the time[14]. However technological design decisions should not dictate our societal interactions and the structure of our communities, but rather should support our values and fundamental rights.

We should develop and promote engineering techniques and methodologies that permit artificial intelligence and robotics to fully respect the dignity and rights of the individual. Not only engineers but also researchers need to start considering privacy engineering principles like *privacy by default* and *privacy by design* in new research, products and services.

As artificial intelligence, through machine learning, needs vast amounts of data to be effective, researchers should explore the design and implementation of algorithms that conceal identities and aggregate data in order to protect the individual at the same time as harnessing the predictive power of that same data.

We must today lay the foundation for addressing these tasks by bringing together researchers, developers and data protection experts from different areas in broad networks, such as the Internet Privacy Engineering Network (IPEN)[15], which contribute to a fruitful inter-disciplinary exchange of ideas and approaches.

An ethical framework needs to underpin the building blocks of the artificial intelligence and robotics ecosystem.

4. Background information

> *'I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.' by A. M. Turing*[16]

4.1. Artificial intelligence prospects

Artificial intelligence and robotics enjoy great development and popularity: personal assistants like Apple Siri, Google Now or Microsoft Cortana, home robots like iRobot Roomba or the, soon-to-be-real?, self-driving cars like the fleet Uber is already testing.

Many public and private entities show interest in artificial intelligence and robotics. The European Parliament Committee on Legal Affairs has a working group on artificial intelligence and robotics[17] which has produced a 'Draft report with recommendations to the Commission on Civil Law Rules on Robotics'[18] where several requests to the European Commission are formulated, e.g.: legally define *smart autonomous robots* and their categories, create a register of *advanced robots* or the creation of an European regulator for robotics. Europe is definitely not alone in thinking about artificial intelligence and robotics: in the US, the White House plans to publish a roadmap soon[19]. This roadmap will contain proposals on regulating these new technologies along lines quite common to data protection and privacy: transparency, but also about the impact on jobs. Meanwhile, five of the most important information companies: Alphabet, Amazon, Facebook, IBM and Microsoft are already discussing about a new industry group to ensure that artificial intelligence and robotics is beneficial for everybody[20].

The future of artificial intelligence is bright and promising and all kind of capabilities and functionalities are expected, some actually researched. But also some voices, like Stephen Hawking, Elon Musk, Steve Wozniak or Bill Gates, have raised the alarm on future scenarios where *superintelligent* artificial intelligences may on purpose or by chance be an existential threat to humanity.

Without considering such extreme possibilities, artificial intelligence and robotic technologies are quite present and will be even more in the near future. Due to its reliance on huge amounts of data for their creation and operation, and the increased autonomy they are getting, the privacy/data protection community needs to start reflecting on this phenomenon and its implications on individuals.

4.2. Definition of artificial intelligence

There are some traditional or standard definitions of artificial intelligence. For example, the one provided by the Oxford English Dictionary: '*artificial intelligence: The theory and development of* *computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages*.'[21] or this other one provided by Wikipedia: '*Artificial intelligence [...] In computer science, an ideal "intelligent" machine is a flexible rational agent that perceives its environment and takes actions that maximize its chance of success at an arbitrary goal*[22]*.'* The first one is a definition oriented toward the task, while the second, more technical, one also provides a very pertinent list of the components of any artificial intelligence understanding: *agent, perception, environment, (re)action, goal*...

Wikipedia also provides a very interesting reflection on the public understanding of artificial intelligence and how it shrinks over time when certain capabilities considered *artificial intelligence* become normal and so, they are not artificial intelligence anymore (emphasis added): *'Colloquially, the term "artificial intelligence" is likely to be applied when a machine [...] competently perform or mimic "cognitive" functions that we intuitively associate with human minds, such as "learning" and "problem solving". [...]* ***[The] subjective borderline around what constitutes "artificial intelligence" tends to shrink over time; for example, optical character recognition is no longer perceived as an exemplar of "artificial intelligence" as it is nowadays a mundane routine technology***.'[23]

4.3. Artificial intelligence and machine learning

Machine learning is the most researched aspect of artificial intelligence at the moment. Some of the most well know developments in the field, IBM Watson, Google/Deepmind Alphago, Apple Siri or Microsoft Cortana[24], to name a few, all of them leverage the use of machine learning to advance their capabilities.

Machine learning was defined as a field of study that gives computers the ability to learn without being explicitly programmed, by Arthur Samuel in 1959[25]. It is a subfield in computer science related to or included in artificial intelligence research. It studies the construction of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from training data in order to make data-driven predictions or decisions rather than following domain based logic pre-programmed into them. As such, this programming paradigm is especially useful when a domain-based algorithm is unfeasible. Examples of applications are spam filtering, optical character recognition (OCR) or computer vision.

When the data contained in the training set is not labelled, there exists no predefined output for any of the inputs, we are referring to unsupervised learning which aims at creating a model to describe the hidden structure of a dataset. Since the examples given to the learner are unlabelled, neither right nor wrong, there is no reward signal to evaluate a potential solution.

Within the field of big data, machine learning is a method used to extract models from big datasets to make predictions: predictive analytics.

It is important to remark that the models created by machine learning will not be human understandable in most cases. The criteria a machine learning algorithm may find to classify input data, or the memory as weights in a neural network, will most probably lack expressivity as correct as they may be. This has a big impact when discussing *algorithmic transparency*.

4.4. Definition of robotics

The word robot comes from the Czech word *robota*, which means *corvée*, *serf labor*. It was made popular through the play R.U.R (Rossum's Universal Robots) (1920) by Karel Čapek. In the play a factory makes artificial people called robots as the perfect workers, *serfs*.

But Mr Čapek has not been the only one, or the first, in imagining *autonomous machines*. One of the earliest descriptions of automata appears in China. An encounter between King Mu of Zhou (1023–957 BC) and a mechanical engineer known as Yan Shi, an *artificer* who presented the king with a human figure. Also, the Greek designed and constructed automata. In the first century A.D, Heron of Alexandria described more than 100 machines and automata including a steam-powered engine, in Pneumatica and Automata.

From those automata and literary beginnings, nowadays, and again according to the Oxford English Dictionary: '*robotics: The branch of technology that deals with the design, construction, operation, and application of robots*' and '*robot: A machine capable of carrying out a complex series of actions automatically, especially one programmable by a computer.*'[26]

Even if *robotics* were present in the mechanical mind of humanity since classical times, it is only in the 20th century that robotics, as we understand it, has grown substantially. Today, robotics is a rapidly growing field, as technological advances continue; researching, designing, and building new robots which serve various practical purposes, whether domestically, commercially, or militarily.

4.5. Today's robots

Robots are used today for many different functions. Japan, definitely an early adopter of robotics, already has a hotel served completely by robots (with some human help) and also robots working in restaurant kitchens. We also have robots in our home cleaning the floor. These are some examples of the real capacities of robots, which, as anecdotic as they may seem, reflect deep research and the technological evolution of the last years: Advances in computer vision, proprioception, location and navigation of the physical space, natural language processing, human interaction, emotion analysis synthesis, etc.

Besides the more conventional use of industrial robots in factories we have many projects, and realities, in other fields like security (bomb disposal, rescue missions), military (weapon systems but also, for example, load carrying robots), or scientific research (probes sent to space or to hazardous environments like deep ocean or a volcano.)

[1] https://en.wikipedia.org/wiki/Luddite

[2] 'Big data refers to the exponential growth both in the availability and in the automated use of information: it refers to gigantic digital datasets held by corporations, governments and other large organisations, which are then extensively analysed (hence the name: analytics) using computer algorithms'; WP29 Opinion 3/2013 on purpose limitation. A White House report in 2014 described Big Data as 'The growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data', see Big Data: Seizing Opportunities, Preserving Values, Executive Office of the President ('Podesta-report'), May 2014.

[3] See for example the speech from United States Federal Trade Commission Chairwoman in 2014: 'The proliferation of connected devices, the plummeting cost of collecting, storing, and processing information, and the ability of data brokers and others to combine offline and online data means that companies can accumulate virtually unlimited amounts of consumer information and store it indefinitely. Using predictive analytics, they can learn a surprising amount about each of us from this;' Opening Remarks FTC Chairwoman Edith Ramirez, 'Big Data: A Tool for Inclusion or Exclusion?', Washington, DC September 15, 2014. According to Sandy Pentland, 'Social physics is a quantitative social science that describes reliable, mathematical connections between information and idea flow on the one hand and people's behaviour on the other… it enables us to predict the productivity of small groups, of departments within companies and even of entire cities'. This 'is what is required to build better social systems' (pp. 4, 7) and to 'allow (government officials, industry managers, and citizens) to use the tools of social network incentives to *establish new norms of behaviour'* (p. 189) (our italics); Pentland, *Social Physics: How Good Ideas Spread: The Lessons from a New Science*.

[4] WP29 Opinion 3/2013 on purpose limitation, Annex 2.

[5] On the concept of 'black boxes' and the importance of transparency, see, for example, 'The Black Box Society, The Secret Algorithms That Control Money and Information' by Frank Pasquale (Harvard University press, 2015).

[6] https://translate.google.com/; https://www.bing.com/translator.

[7] http://www.rethinkrobotics.com/baxter/

[8] Like the Korea's Dodam autonomous robotic turret called the Super aEgis II (http://www.dodaam.com/eng/sub2/menu2_1_4.php) or several other developments by armies and weapon developers all around the world (http://www.nytimes.com/2014/11/12/science/weapons-directed-by-robots-not-humans-raise-ethical-questions.html): '*Britain's "fire and forget" Brimstone missiles, for example, can distinguish among tanks and cars and buses without human assistance, and can hunt targets in a predesignated region without oversight. The Brimstones also communicate with one another, sharing their targets.*'

[9] According to the definitions given by the International Civil Aviation Organization in the Cir 328/190 (available at http://www.icao.int/Meetings/UAS/Documents/Circular%20328_en.pdf), an *unmanned aircraft system* (UAS) is an aircraft and its associated elements which are operated with no pilot on board whereas a *remotely-piloted aircraft system* is an aircraft where the flying pilot is not on board the aircraft. This is a subcategory of unmanned aircraft. A remotely-piloted aircraft system is a set of configurable elements consisting of a remotely-piloted aircraft, its associated remote pilot station(s), the required command and control links and any other system elements as may be required, at any point during flight operation.

[10] A ludic example of this possibility are the myriad of self-flying drones able to do *video selfies* of their owners; a very different one is the concept of military drones able to choose targets on their own.

[11] See European Group on Ethics in Science and New Technologies, Opinion on Ethics and Surveillance, p. 75. A study has suggested that an ad-targeting algorithm was discriminatory, with searches on average returning ads for higher paid jobs for men compared with women visiting job sites; Carnegie Mellon University and the International Computer Science Institute. On the tendency of digital assistants to be given by default a female voice, see for example Judy Wajcman, Feminist theories of technology. Cambridge Journal of Economics, 34 (1). pp. 143-152, 2010.

[12] Giorgio Agamben, *State of Exception*, 2005.

[13] Neil Richards, Neil and Jonathan King, *Big Data Ethics* (May 19, 2014), Wake Forest Law Review, 2014.

[14] 'Behind the technology that affects social relations lie the very same social relations', David Noble, 'Social Choice in Machine Design: The Case of Automatically Controlled Machine Tools', in *Case Studies in the Labour Process*, ed. Andrew Zimbalist, 1979. See also Judy Wacjman*, Pressed for Time: The Acceleration of Life in Digital Capitalism*, 2014 pp. 89-90.

[15] http://engineeringprivacy.eu/

[16] Computing Machinery and Intelligence by A. M. Turing (Turing, A.M. (1950). Computing machinery and intelligence. Mind, 59, 433-460.) (http://loebner.net/Prizef/TuringArticle.html)

[17] http://www.europarl.europa.eu/committees/en/juri/subject-files.html?id=20150504CDT00301

[18] http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONSGML%2bCOMPARL%2bPE-582.443%2b01%2bDOC%2bPDF%2bV0%2f%2fEN

[19] http://www.politico.eu/pro/white-house-eyes-regulations-for-artificial-intelligence-privacy-security-data-facebook-google/

[20] http://www.nytimes.com/2016/09/02/technology/artificial-intelligence-ethics.html?_r=0

[21] http://www.oxforddictionaries.com/definition/english/artificial-intelligence

[22] According to Wikipedia this definition, in terms of goals, actions, perception and environment, is due to Russell & Norvig Russell, Stuart J.; Norvig, Peter (2003), Artificial Intelligence: A Modern Approach (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall.

[23] https://en.wikipedia.org/wiki/Artificial_intelligence

[24] http://www.ibm.com/watson/what-is-watson.html; https://deepmind.com/alpha-go ; https://en.wikipedia.org/wiki/Siri; https://www.microsoft.com/en-us/cloud-platform/why-cortana-intelligence-suite

[25] '*[...] it is necessary to specify methods of problem solution in minute and exact detail, a time-consuming and costly procedure. Programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort.*' Samuel, A. L. (1959), "Some Studies in Machine Learning Using the Game of Checkers" in IBM Journal of Research and Development (Volume:3, Issue: 3).

[26] http://www.oxforddictionaries.com/definition/english/robotics and http://www.oxforddictionaries.com/definition/english/robot#robot__2