European Commission

European Group on Ethics in Science and New Technologies

Statement on

Artificial Intelligence, Robotics and 'Autonomous' Systems

Research and Innovation

European Group on Ethics in Science and New Technologies
Artificial Intelligence, Robotics and 'Autonomous' Systems

European Commission
Directorate-General for Research and Innovation
Unit RTD.01 — Scientific Advice Mechanism

Contact Jim Dratwa, Head of the EGE Office
E-mail EC-ETHICS-GROUP@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Printed by OP in Luxembourg.

Manuscript completed in March 2018.

This is an EGE Statement, adopted by the members of the EGE: Emmanuel Agius, Anne Cambon-Thomsen, Ana Sofia Carvalho, Eugenijus Gefenas, Julian Kinderlerer, Andreas Kurtz, Jonathan Montgomery, Herman Nys (Vice-Chair), Siobhán O'Sullivan (Vice-Chair), Laura Palazzani, Barbara Prainsack, Carlos Maria Romeo Casabona, Nils-Eric Sahlin, Jeroen van den Hoven, Christiane Woopen (Chair). Rapporteur: Jeroen van den Hoven.

More information on the European Union is available on the internet (http://europa.eu).

Luxembourg: Publications Office of the European Union, 2018

Print	ISBN 978-92-79-80328-4	doi:10.2777/786515	KI-04-18-224-EN-C
PDF	ISBN 978-92-79-80329-1	doi:10.2777/531856	KI-04-18-224-EN-N

EUROPEAN COMMISSION

Statement on

Artificial Intelligence, Robotics and 'Autonomous' Systems

European Group on Ethics
in Science and New Technologies

Brussels, 9 March 2018

2018 Directorate-General for Research and Innovation

Table of Contents

Summary

Advances in AI, robotics and so-called 'autonomous' technologies[1] have ushered in a range of increasingly urgent and complex moral questions. Current efforts to find answers to the ethical, societal and legal challenges that they pose and to orient them for the common good represent a patchwork of disparate initiatives. This underlines the need for a collective, wide-ranging and inclusive process of reflection and dialogue, a dialogue that focuses on the values around which we want to organise society and on the role that technologies should play in it.

This statement calls for the launch of a process that would pave the way towards a common, internationally recognised ethical and legal framework for the design, production, use and governance of artificial intelligence, robotics, and 'autonomous' systems. The statement also proposes a set of fundamental ethical principles, based on the values laid down in the EU Treaties and the EU Charter of Fundamental Rights, that can guide its development.

[1] This statement pertains to a set of smart digital technologies that are rapidly converging and are often interrelated, connected or fully integrated, e.g. classical Artificial Intelligence, Machine Learning algorithms, Deep Learning and connectionist networks, generative adversarial networks, mechatronics and robotics. Self-driving cars and robotic weapon systems, chat bots and speech and image recognition systems are among some of the well-known exemplifications of combinations of these technologies.

Background

The first two decades of the 21st century have brought us striking examples of what is commonly referred to as 'autonomous technology' and 'artificial intelligence'. Self-driving cars and drones, robots in deep sea and space exploration, weapon systems, software agents, such as bots in financial trade, and deep learning in medical diagnosis, are among the most prominent, but certainly not the only examples. Artificial intelligence (AI), especially in the form of machine learning, and the increasing availability of large datasets from various domains of life are important drivers of these developments. The confluence of these digital technologies is rapidly making them more powerful, they are applied in an increasing number of new products and services, in public and private sectors, and can have both military and civilian application. The AI lodged in these systems can redefine work or improve work conditions for humans and reduce the need for human contribution, input and interference during operation. It can help to assist or replace humans with smart technology in difficult, dirty, dull or dangerous work, and even beyond.

Without direct human intervention and control from outside, smart systems today conduct dialogues with customers in online call-centres, steer robot hands to pick and manipulate objects accurately and incessantly, buy and sell stock at large quantities in milliseconds, direct cars to swerve or brake and prevent a collision, classify persons and their behaviour, or impose fines.

It is unfortunate that some of the most powerful among these cognitive tools are also the most opaque. Their actions are no longer programmed by humans in a linear manner. Google Brain develops AI that allegedly builds AI better and faster than humans can. AlphaZero can bootstrap itself in four hours from completely ignorant about the rules of chess, to world champion level. It is impossible to understand how exactly AlphaGo managed to beat the human Go World champion. Deep learning and so-called 'generative adversarial network approaches' enable machines to 'teach' themselves new strategies and look for new evidence to analyse. In this sense, their actions are often no longer intelligible, and no longer open to scrutiny by humans. This is the case because, first, it is impossible to establish how they accomplish their results beyond the initial algorithms. Second, their performance is based on the data that have been used during the learning process and that may no longer be available or accessible. Thus, biases and errors that they have been presented with in the past become engrained into the system.

When systems can learn to perform these tasks without human direction or without supervision, they are now often called 'autonomous'. These so-called 'autonomous' systems can manifest themselves as high-tech robotic systems or as intelligent software such as bots. Many of them are released into the world unsupervised and may accomplish things which are not foreseen by their human designers or owners.

We thus see the following relevant developments in technology:

(1) Artificial Intelligence in the form of machine learning (especially 'deep learning'), fuelled by Big Data, is rapidly becoming more powerful. It is applied in an increasing number of new digital products and services in public and private sectors and can have both military as well as civilian application. As noted, AI's inner workings can be extremely hard - if not impossible - to track, explain and critically evaluate. These advanced capabilities are accumulating in large part with private parties and are for a large part proprietary.

(2) Advanced mechatronics (a combination of AI and deep learning, data science, sensor technology, Internet of Things, mechanical and electrical engineering) is providing a wide range of increasingly sophisticated robotic and high-tech systems for practical applications in service and production industry, health care, retail, logistics, domotics (home automation) and security and safety. Two domains of application that stand out in public debates are robotic weapons systems and 'autonomous' vehicles.

(3) Ever smarter systems are produced that exhibit high degrees of what is often referred to as 'autonomy', which means that they develop and can perform tasks independently from human operators and without human control.

(4) There seems to be a push for ever higher degrees of automation and 'autonomy' in robotics, AI and mechatronics. Investments of countries and large companies in this field are enormous and a leading position in AI research is among the prominent goals of superpowers in the world.

(5) There is development towards ever closer interaction between humans and machines (co-bots, cyber-crews, digital twins and even the integration of smart machines into the human body in the form of computer-brain interfaces or cyborgs). Similar developments can be seen across the AI realm. Well aligned teams of AI systems and human professionals perform better in some domains than humans or machines separately.

Moral Reflections

Key questions

The advent of high-tech systems and software that can function increasingly independently of humans and can execute tasks that would require intelligence when carried out by humans, warrants special reflection. These systems give rise to a range of important and hard moral questions.

First, questions about safety, security, the prevention of harm and the mitigation of risks. How can we make a world with interconnected AI and 'autonomous' devices safe and secure and how can we gauge the risks?

Second, there are questions about human moral responsibility. Where is the morally relevant agency located in dynamic and complex socio-technical systems with advanced AI and robotic components? How should moral responsibility be attributed and apportioned and who is responsible (and in what sense) for untoward outcomes? Does it make sense to speak about 'shared control' and 'shared responsibility' between humans and smart machines? Will humans be part of ecosystems of 'autonomous' devices as moral 'crumple zones', inserted just to absorb liability or will they be well placed to take responsibility for what they do?

Third, they give rise to questions about governance, regulation, design, development, inspection, monitoring, testing and certification. How should our institutions and laws be redesigned to make them serve the welfare of individuals and society and to make society safe for this technology?

Fourth, there are questions regarding democratic decision making, including decision making about institutions, policies and values that underpin all of the questions above. Investigations are carried out across the globe to establish the extent to which citizens are taken advantage of by the use of advanced nudging techniques based on the combination of machine learning, big data and behavioural science, which make possible the subtle profiling, micro-targeting, tailoring and manipulation of choice architectures in accordance with commercial or political purposes.

Finally, there are questions about the explainability and transparency of AI and 'autonomous' systems. Which values do these systems effectively and demonstrably serve? Which values underpin how we design our policies and our machines? Around which values do we want to organise our societies?

And which values are we letting to be undermined – openly or silently – in the technological progress and utility trade-offs? AI driven 'optimisation' of social processes based on social scoring systems with which some countries experiment, violate the basic idea of equality and freedom in the same way caste systems do, because they construct 'different kinds of people' where there are in reality only 'different properties' of people. How can the attack on democratic systems and the utilisation of scoring systems, as a basis for dominance by those who have access to these powerful technologies, be prevented?

Key considerations

From an ethical perspective it is important to bear in mind that:

The term 'autonomy' stems from philosophy and refers to the capacity of human persons to legislate for themselves, to formulate, think and choose norms, rules and laws for themselves to follow. It encompasses the right to be free to set one's own standards and choose one's own goals and purposes in life. The cognitive processes that support and facilitate this are among the ones most closely identified with the dignity of human persons and human agency and activity *par excellence*. They typically entail the features of self-awareness, self-consciousness and self-authorship according to reasons and values. Autonomy in the ethically relevant sense of the word can therefore only be attributed to human beings. It is therefore somewhat of a misnomer to apply the term 'autonomy' to mere artefacts, albeit very advanced complex adaptive or even 'intelligent' systems. The terminology of 'autonomous' systems has however widely gained currency in the scientific literature and public debate to refer to the highest degree of automation and the highest degree of independence from human beings in terms of operational and decisional 'autonomy'. But autonomy in its original sense is an important aspect of human dignity that ought not to be relativised.

Since no smart artefact or system - however advanced and sophisticated - can in and by itself be called 'autonomous' in the original ethical sense, they cannot be accorded the moral standing of the human person and inherit human dignity. Human dignity as the foundation of human rights implies that meaningful human intervention and participation must be possible in matters that concern human beings and their environment. Therefore, in contrast to the automation of production, it is not appropriate to manage and decide about humans in the way we manage and decide about objects or data, even if this is technically conceivable. Such an 'autonomous'

management of human beings would be unethical, and it would undermine the deeply entrenched European core values. Human beings ought to be able to determine which values are served by technology, what is morally relevant and which final goals and conceptions of the good are worthy to be pursued. This cannot be left to machines, no matter how powerful they are.

The ability and willingness to take and attribute moral responsibility is an integral part of the conception of the person on which all our moral, social and legal institutions are based. Moral responsibility is here construed in the broad sense in which it may refer to several aspects of human agency, e.g. causality, accountability (obligation to provide an account), liability (obligation to compensate damages), reactive attitudes such as praise and blame (appropriateness of a range of moral emotions), and duties associated with social roles. Moral responsibility, in whatever sense, cannot be allocated or shifted to 'autonomous' technology.

In recent debates about Lethal Autonomous Weapons Systems (LAWS) and Autonomous Vehicles there seems to exist a broad consensus that *Meaningful Human Control* is essential for moral responsibility. The principle of Meaningful Human Control (MHC) was first suggested for constraining the development and utilisation of future weapon systems. This means that humans - and not computers and their algorithms - should ultimately remain in control, and thus be morally responsible.[2]

Beyond a narrow ethical framing

Two areas where the development of 'autonomous' systems has already led to high-profile ethical debates are self-driving cars and Lethal Autonomous Weapons Systems (LAWS). Although fully driverless cars are not yet on the market, several countries around the world are preparing for the legal possibility of allowing 'autonomous' vehicles on public roads. In 2016, moral controversy stirred up when the first person was killed in a car crash while driving in 'autonomous' mode. Moral debates are now often limited to discussion of exceptional use cases concerning so-called 'Trolley Problem' thought experiments. These cases are concerned with dilemmas of unavoidable accidents in which the only available choice is between options associated with the loss of human lives. This narrow construal of ethical problems invites a calculating approach and implies an often overly simplistic metrics in human affairs. Central questions in that framing mainly

[2] NGO Article 36, 2015

seem to concern the responsibility of 'autonomous' systems, their effects and how they should be programmed so that their deployment leads to morally acceptable outcomes in terms of lives lost respectively lives saved. This neglects broader questions such as 'which design decisions were taken in the past that have led up to this moral predicament', 'which values should inform design', 'how should values in design be weighed in case of conflict, and by whom', 'what is the status of the massive empirical findings that are accumulating concerning how people actually decide in Trolley cases and being transposed to automated vehicle settings?'

A second field of contestation and controversy are 'autonomous' weapon systems. These military systems can carry lethal weapons as their payload, but as far as the software is concerned they are not very different from 'autonomous' systems that we could find in a range of civilian domains close to home. A large part of the debate takes place at the Conference on Certain Conventional Weapons in Geneva concerning the moral acceptability of 'autonomous' weapons and legal and moral responsibility for the deployment of these systems. Now attention needs to turn to questions as to what the nature and meaning of 'meaningful human control' over these systems is and how to institute morally desirable forms of control.

A third important area of application is 'autonomous' software including bots. Trade, finance and stock markets are largely run by algorithms and software. Without human intervention and control from outside, smart systems today conduct dialogues with customers in online call-centres; speech recognition interfaces and recommender systems of online platforms, e.g. Siri, Alexa and Cortana, make suggestions to users. Beyond the straightforward questions of data protection and privacy, we may ask whether people have a right to know whether they are dealing with a human being or with an AI artefact. Moreover, the question arises whether there should be limits to what AI systems can suggest to a person, based on a construction of the person's own conception of their identity.

While there is growing awareness of the need to address such questions, AI and robotics are currently advancing more rapidly than the process of finding answers to these thorny ethical, legal and societal questions. Current efforts represent a patchwork of disparate initiatives. There is a clear need for a collective, wide-ranging and inclusive process that would pave the way towards a common, internationally recognised ethical framework for the design, production, use and governance of AI, robots and 'autonomous' systems.

This statement calls for the launch of such a process and proposes a set of fundamental ethical principles and democratic prerequisites that could also guide reflection on binding law. The EGE is of the opinion that Europe should play an active and prominent role in this. Overseeing the debates on moral responsibility for AI and so-called 'autonomous' technology, the EGE calls for more systematic thinking and research about the ethical, legal and governance aspects of high tech-systems that can act upon the world without direct control of human users, to human benefit or to human detriment. This is a matter of great urgency.

Towards a shared Ethical Framework for Artificial Intelligence, Robotics and 'Autonomous' Systems

Some of the most prominent initiatives towards the formulation of ethical principles regarding AI and 'autonomous' systems have stemmed from industry, practitioners and professional associations, such as the IEEE's (Institute of Electrical and Electronics Engineers) policy paper on 'Ethically Aligned Design',[3] ITU's (International Telecommunication Union) Global Summit 'AI for Good'[4] in summer 2017, and the ACM's (Association for Computing Machinery) work on the issue, including a major AAAI/ACM 'Conference on AI, Ethics, and Society'[5] in February 2018. Within the private sector, companies such as IBM, Microsoft and Google's DeepMind have established their own ethic codes on AI and joined forces in creating broad initiatives such as the 'Partnership on AI'[6] or 'OpenAI'[7], which bring together industry, non-profit and academic organisations.

One of the leading initiatives calling for a responsible development of AI has been launched by the Future of Life Institute and has culminated in the creation of the 'Asilomar AI Principles'. This list of 23 fundamental principles to guide AI research and application has been signed by hundreds of stakeholders,[8] with signatories representing predominantly scientists, AI researchers and industry. A similar participatory process has been launched upon the initiative of the Forum on the Socially Responsible Development of Artificial Intelligence held by the University of Montreal in November 2017, in reaction to which a first draft of a potential 'Declaration for a Responsible Development of Artificial Intelligence' has been developed. It is now publicly accessible on an online platform where all sectors of society are invited to comment on the text.[9]

A worldwide debate on the military use of AI has been initiated by the UN and the meetings for the Convention on Certain Conventional Weapons (CCW, Geneva), where a majority of the High Contracting Parties endorsed the so-called principle of 'meaningful human control for LAWS' stating that 'Autonomous Weapons Systems that require no meaningful human control

3. http://standards.ieee.org/news/2016/ethically_aligned_design.html
4. https://www.itu.int/en/ITU-T/AI/Pages/201706-default.aspx
5. http://www.aies-conference.com/
6. https://www.partnershiponai.org/
7. https://openai.com/
8. https://futureoflife.org/ai-principles/
9. http://nouvelles.umontreal.ca/en/article/2017/11/03/montreal-declaration-for-a-responsible-development-of-artificial-intelligence/

should be prohibited' (General Assembly UN, 2016). The UN has also established a special research institute in The Hague to study the governance of Robotics and AI (UNICRI). Several initiatives and NGOs that aim at AI and 'autonomous' systems 'for good' respectively campaign for a ban on 'autonomous' weapons, e.g. the Foundation for Responsible Robotics.

Meanwhile, at the national level initiatives are uneven, with some countries prioritising the development of rules for robots and artificial intelligence and going so far as to adopt legislation (e.g. to regulate self-driving cars on public roads), whereas other countries are yet to deal with the matter. This lack of a harmonised European approach has prompted the European Parliament to call for a range of measures to prepare for the regulation of advanced robotics,[10] including the development of a guiding ethical framework for the design, production and use of robots.

Against this backdrop, the EGE draws attention to the risks inherent to uncoordinated, unbalanced approaches in the regulation of AI and 'autonomous' technologies. Regulatory patchworks may give rise to 'ethics shopping', resulting in the relocation of AI development and use to regions with lower ethical standards. Allowing the debate to be dominated by certain regions, disciplines, demographics or industry actors risks excluding a wider set of societal interests and perspectives. Current discussions sometimes also lack an overview of 'autonomous' technologies that are likely to be studied, developed and implemented in the next decade, leaving a blind spot when it comes to regulatory foresight.

10. European Parliament, Committee on Legal Affairs 2015/2103 (INL) Report with Recommendations to the Commission on Civil Law Rules on Robotics, Rapporteur Mady Delvaux.

The EGE calls for a wide-ranging and systematic public engagement and deliberation on the ethics of AI, robotics and 'autonomous' technology and on the set of values that societies choose to embed in the development and governance of these technologies. This process, in which the EGE stands ready to play its part, should provide a platform for joining together the diverse global initiatives outlined above. It should integrate a wide, inclusive and far-reaching societal debate, drawing upon the input of diverse perspectives, where those with different expertise and values can be heard. The EGE urges the European Union to place itself at the vanguard of such a process and calls upon the European Commission to launch and support its implementation.

As a first step towards the formulation of a set of ethical guidelines that may serve as a basis for the establishment of global standards and legislative action, **the EGE proposes a set of basic principles and democratic prerequisites, based on the fundamental values laid down in the EU Treaties and in the EU Charter of Fundamental Rights.**

Ethical principles and democratic prerequisites

(a) Human dignity

The principle of human dignity, understood as the recognition of the inherent human state of being worthy of respect, must not be violated by 'autonomous' technologies. This means, for instance, that there are limits to determinations and classifications concerning persons, made on the basis of algorithms and 'autonomous' systems, especially when those affected by them are not informed about them. It also implies that there have to be (legal) limits to the ways in which people can be led to believe that they are dealing with human beings while in fact they are dealing with algorithms and smart machines. A relational conception of human dignity which is characterised by our social relations, requires that we are aware of whether and when we are interacting with a machine or another human being, and that we reserve the right to vest certain tasks to the human or the machine.

(b) Autonomy

The principle of autonomy implies the freedom of the human being. This translates into human responsibility and thus control over and knowledge about 'autonomous' systems as they must not impair freedom of human beings to set their own standards and norms and be able to live according to them. All 'autonomous' technologies must, hence, honour the human ability to choose whether, when and how to delegate decisions and actions to them. This also involves the transparency and predictability of 'autonomous' systems, without which users would not be able to intervene or terminate them if they would consider this morally required.

(c) Responsibility

The principle of responsibility must be fundamental to AI research and application. 'Autonomous' systems should only be developed and used in ways that serve the global social and environmental good, as determined by outcomes of deliberative democratic processes. This implies that they should be designed so that their effects align with a plurality of fundamental human values and rights. As the potential misuse of 'autonomous' technologies poses a major challenge, risk awareness and a precautionary

approach are crucial. Applications of AI and robotics should not pose unacceptable risks of harm to human beings, and not compromise human freedom and autonomy by illegitimately and surreptitiously reducing options for and knowledge of citizens. They should be geared instead in their development and use towards augmenting access to knowledge and access to opportunities for individuals.

Research, design and development of AI, robotics and 'autonomous' systems should be guided by an authentic concern for research ethics, social accountability of developers, and global academic cooperation to protect fundamental rights and values and aim at designing technologies that support these, and not detract from them.

(d) Justice, equity, and solidarity

AI should contribute to global justice and equal access to the benefits and advantages that AI, robotics and 'autonomous' systems can bring. Discriminatory biases in data sets used to train and run AI systems should be prevented or detected, reported and neutralised at the earliest stage possible.

We need a concerted global effort towards equal access to 'autonomous' technologies and fair distribution of benefits and equal opportunities across and within societies. This includes the formulating of new models of fair distribution and benefit sharing apt to respond to the economic transformations caused by automation, digitalisation and AI, ensuring accessibility to core AI technologies, and facilitating training in STEM and digital disciplines, particularly with respect to disadvantaged regions and societal groups. Vigilance is required with respect to the downside of the detailed and massive data on individuals that accumulates and that will put pressure on the idea of solidarity, e.g. systems of mutual assistance such as in social insurance and healthcare. These processes may undermine social cohesion and give rise to radical individualism.

(e) Democracy

Key decisions on the regulation of AI development and application should be the result of democratic debate and public engagement. A spirit of global cooperation and public dialogue on the issue will ensure that they are taken in an inclusive, informed, and farsighted manner. The right to receive

education or access information on new technologies and their ethical implications will facilitate that everyone understands risks and opportunities and is empowered to participate in decisional processes that crucially shape our future.

The principles of human dignity and autonomy centrally involve the human right to self-determination through the means of democracy. Of key importance to our democratic political systems are value pluralism, diversity and accommodation of a variety of conceptions of the good life of citizens. They must not be jeopardised, subverted or equalised by new technologies that inhibit or influence political decision making and infringe on the freedom of expression and the right to receive and impart information without interference. Digital technologies should rather be used to harness collective intelligence and support and improve the civic processes on which our democratic societies depend.

(f) Rule of law and accountability

Rule of law, access to justice and the right to redress and a fair trial provide the necessary framework for ensuring the observance of human rights standards and potential AI specific regulations. This includes protections against risks stemming from 'autonomous' systems that could infringe human rights, such as safety and privacy.

The whole range of legal challenges arising in the field should be addressed with timely investment in the development of robust solutions that provide a fair and clear allocation of responsibilities and efficient mechanisms of binding law.

In this regard, governments and international organisations ought to increase their efforts in clarifying with whom liabilities lie for damages caused by undesired behaviour of 'autonomous' systems. Moreover, effective harm mitigation systems should be in place.

(g) Security, safety, bodily and mental integrity

Safety and security of 'autonomous' systems materialises in three forms: (1) external safety for their environment and users, (2) reliability and internal robustness, e.g. against hacking, and (3) emotional safety with respect to human-machine interaction. All dimensions of safety must be

taken into account by AI developers and strictly tested before release in order to ensure that 'autonomous' systems do not infringe on the human right to bodily and mental integrity and a safe and secure environment. Special attention should hereby be paid to persons who find themselves in a vulnerable position. Special attention should also be paid to potential dual use and weaponisation of AI, e.g. in cybersecurity, finance, infrastructure and armed conflict.

(h) Data protection and privacy

In an age of ubiquitous and massive collection of data through digital communication technologies, the right to protection of personal information and the right to respect for privacy are crucially challenged. Both physical AI robots as part of the Internet of Things, as well as AI softbots that operate via the World Wide Web must comply with data protection regulations and not collect and spread data or be run on sets of data for whose use and dissemination no informed consent has been given.

'Autonomous' systems must not interfere with the right to private life which comprises the right to be free from technologies that influence personal development and opinions, the right to establish and develop relationships with other human beings, and the right to be free from surveillance. Also in this regard, exact criteria should be defined and mechanisms established that ensure ethical development and ethically correct application of 'autonomous' systems.

In light of concerns with regard to the implications of 'autonomous' systems on private life and privacy, consideration may be given to the ongoing debate about the introduction of two new rights: the right to meaningful human contact and the right to not be profiled, measured, analysed, coached or nudged.

(i) Sustainability

AI technology must be in line with the human responsibility to ensure the basic preconditions for life on our planet, continued prospering for mankind and preservation of a good environment for future generations. Strategies to prevent future technologies from detrimentally affecting human life and nature are to be based on policies that ensure the priority of environmental protection and sustainability.

Artificial intelligence, robotics and 'autonomous' systems can bring prosperity, contribute to well-being and help to achieve European moral ideals and socio-economic goals if designed and deployed wisely. Ethical considerations and shared moral values can be used to shape the world of tomorrow and should be construed as stimulus and opportunities for innovation, and not impediments and barriers.

The EGE calls upon the European Commission to investigate which existing legal instruments are available to effectively deal with the problems discussed in this statement and whether new governance and regulatory instruments are required.

The EGE calls for the launch of a process that paves the way towards a common, internationally recognised ethical and legal framework for the design, production, use and governance of artificial intelligence, robotics, and 'autonomous' systems.

Getting in touch with the EU

IN PERSON

All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: **http://europa.eu/contact**

ON THE PHONE OR BY E-MAIL

Europe Direct is a service that answers your questions about the European Union. You can contact this service

– by freephone: **00 800 6 7 8 9 10 11** (certain operators may charge for these calls),

– at the following standard number: **+32 22999696** or

– by electronic mail via: **http://europa.eu/contact**

Finding information about the EU

ONLINE

Information about the European Union in all the official languages of the EU is available on the Europa website at: **http://europa.eu**

EU PUBLICATIONS

You can download or order free and priced EU publications from EU Bookshop at: **http://bookshop.europa.eu.** Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact)

EU LAW AND RELATED DOCUMENTS

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: **http://eur-lex.europa.eu**

OPEN DATA FROM THE EU

The EU Open Data Portal (**http://data.europa.eu/euodp/en/data**) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

Advances in AI, robotics and so-called 'autonomous' technologies have ushered in a range of increasingly urgent and complex moral questions. Current efforts to find answers to the ethical, societal and legal challenges that they pose and to orient them for the common good represent a patchwork of disparate initiatives. This underlines the need for a collective, wide-ranging and inclusive process of reflection and dialogue, a dialogue that focuses on the values around which we want to organise society and on the role that technologies should play in it.

This statement calls for the launch of a process that would pave the way towards a common, internationally recognised ethical and legal framework for the design, production, use and governance of artificial intelligence, robotics, and 'autonomous' systems. The statement also proposes a set of fundamental ethical principles, based on the values laid down in the EU Treaties and the EU Charter of Fundamental Rights, that can guide its development.

Research and Innovation policy

Publications Office