

SÍNDIC
EL DEFENSOR
DE LES
PERSONES



IOI - Europe
International Ombudsman Institute
European Chapter

Artificial Intelligence & Human Rights

ombudsmanship
challenges, roles
and tools

2 & 3 March 2020
Disseny Hub Barcelona

Collaboration



Barcelona
Supercomputing
Center
Centre Nacional de Supercomputació

Sponsorship



Obra Social
Fundació 'la Caixa'

Monday, 2 March 2020

Opening	5
Rafael Ribó, European President of the IOI and Ombudsman of Catalonia	
Pablo Martín, Deputy Secretary of the Spanish Ministry of Science and Innovation	
Àngels Chacon, Catalan Government Minister for Business and Knowledge	
Laia Bonet, Third Deputy Mayor at Barcelona City Hall	
Peter Tyndall, President of IOI and Ombudsman of the Republic of Ireland	
Facts and limits of the artificial intelligence	15
Chairperson: Andreas Pottakis, Director of IOI and Greek Ombudsman	
Ulises Cortés, Professor and Senior Researcher at the Polytechnic University of Catalonia (UPC), Barcelona Supercomputing Center (BSC)	
Montserrat Pardo Bayona, Government Affairs Director Microsoft España	
Alfonso González Herrero, Director Government & Regulatory Affairs, IBM Spain	
Oliver Smith, Strategy & Ethics Director at Alpha Health	
Carina Lopes, Head of the Digital Future Society Think Tank (Mobile World Capital)	
From analogic to digital society	39
Chairperson: Genís Roca, Internet Expert	
José María Lassalle, Director of ESADE Technological Humanism Forum and Professor of Philosophy in Law at ICADE	
Daniel Innerarity, Professor of Political Philosophy at the University of the Basque Country; Ikerbasque Foundation for Science; European University Institute of Florence	
Renata Ávila, Executive Director of Fundación Ciudadano Inteligente	
David Cabo, Director at Fundación Civio	

Tuesday, 3 March 2020

Rights, challenges: Sectoral approach 53

Chairperson: Esther Paniagua, Technology Expert Freelance Journalist

Cybersecurity Ángel Gómez de Ágreda, Aviation Colonel, Chief of Geopolitical Analysis Area, DICOES (Secretary of Defense Policy, Ministry of Defense)

Health and databases research Josep M. Argimon, Managing Director of the Catalan Health Institute

Social and non-discrimination programs Liliانا Arroyo, Expert in digital transformation; Institute for Social Innovation (ESADE)

Bioethics Itziar de Lecuona, Deputy Director of the Bioethics and Law Observatory, UNESCO Chair in Bioethics at the University of Barcelona

Childhood and adolescence José María Sánchez Bursón, consultant

Rights, challenges: International treaties and definitions 75

Chairperson: Miquel Molina, Deputy Director of *La Vanguardia*

Gregor Stojin, president of the CAHA (Council of Europe)

Martha Stickings, Policy Analyst, Fundamental Rights Agency (FRA)

Geraldine Mattioli-zeltner, Adviser at European Commissioner for Human Rights, Council of Europe

Peter Bonnor, Senior Legal Officer, EU Ombudsman

Ombudsmanship challenges, roles and tools 91

Chairperson: Joan Manuel del Pozo, Girona University Ombudsman

Rob Behrens, UK Parliamentary and Health Service Ombudsman

Elisabet Rynning, Parliamentary Chief Ombudsman of Sweden

Quirine Eijkman, Commissioner and Vice President of the Netherlands Institute of Human Rights

Eric Houtman, Belgian Federal Energy Ombudsman and member of NEON

Conclusions 107

Chairperson: Catherine De Bruecker, IOI Europe Vice President & Federal Ombudsman of Belgium

Ulises Cortés, Professor and Senior Researcher at the Polytechnic University of Catalonia (UPC), Barcelona Supercomputing Center (BSC)

Daniel Marco, Catalan Government Director for Innovation & Digital Economy

Rafael Ribó, European President of the IOI and Ombudsman of Catalonia

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Monday, 2 March 2020

Opening

Pablo Martín, Deputy Secretary of the Spanish Ministry of Science and Innovation

Àngels Chacon, Catalan Government Minister for Business and Knowledge

Laia Bonet, Third Deputy Mayor at Barcelona City Hall

Peter Tyndall, President of IOI and Ombudsman of the Republic of Ireland

Rafael Ribó, European President of IOI and Ombudsman of Catalonia

Rafael Ribó

European President of IOI and Ombudsman of Catalonia

Welcome to the IOI Workshop on Artificial Intelligence and Human rights.

I confess that I am ignorant about artificial intelligence, but it scares me because I find few people being experts on human rights and artificial intelligence. And our mission, that of the ombudsmen, is human rights. The new Secretary of State for Digitalisation told me that we have to build a risk matrix. And others say that we have to look for ethical algorithms, and here I am starting to get lost. Others talk about illiberal algorithms. And above all, everyone tells me that algorithms provoke biases, especially gender equality biases. And therefore it is clear that we have to pressure for policies, measures, a series of concrete actions, which in terms of the more solemn legal frameworks are urgent.

Talking to the Spanish ambassador to the delegation of the Council of Europe, for us the most sacred body for the defence of human rights, with the most solid human rights tool on the planet, which is the European Convention, he said to me: "We have to reform the treaties". But the Council of Europe just created the Ad Hoc Committee on Artificial Intelligence in September 2019. It is here in the corner. September 2019. National human rights structures are also working on it. We really need something that allows us all, starting with the administrations, not only the risk matrix, but measures such as those announced by the Government of Catalonia, for example, the so-called CIDAI.

Ombudsmen are constantly discussing about best practices. And as you know, my colleagues, I don't like to stop myself in best practices. I prefer to talk about human rights. And in the middle, in these questions, appears the word *ethics*. The first time I met a big company of GAFAM and I was told: "We are creating a committee on ethics". And I said: "No, you should create a committee on human rights". We need to know what ethics is, because in a comparative way there are many approaches to ethics, and we need to approach ethics really from what is described, what is really in definition, even within the gates of human rights. Ombudsmen find problems like credit cards for social benefits, fiscal duties of those companies, essential service supplies through algorithms, smart meters, obligations for all the European citizens... A lot of bias through algorithms and especially intersectoral discrimination, multiple vulnerability through intersectoral discrimination. We should really reform treaties. We, ombudsmen, if we want to play a role, attacking maladministration, looking for cases of best practices, but especially with reference to human rights.

I would like to make three solemn appeals: first, to all my fellow ombudsmen. I believe that today we are taking the first step as a European and perhaps world collective of ombudsmen to enter this world of artificial intelligence and human rights. And the first appeal is: Ombudsmen everywhere, we cannot leave this issue ignored and unaware of its limits. It is a great challenge, which is the second appeal to the whole of society, it is a great democratic challenge. And we really need, as someone said, digitally literate citizens, that is to say, citizens who are digitally capable of thinking and criticising. We must

not computerise them just to put them on another ring in the chain. We need them to be able to think and criticise about this reality.

Therefore, a second appeal to all citizens, especially those, some of whom will speak here and tomorrow here, going in the forefront unmasking where there can be a harm to human rights. And the third appeal, as I have hinted before, policy makers, MPs, international organizations, hurry up, it is urgent to reform the treaties. A citizen of Catalonia, a citizen of Barcelona, tells you so. Barcelona boasts of being one of the most important digital hubs in Europe. I wish they would one day boast, very proudly, of being one of the most important digital human rights hubs.

Thank you so much.

Pablo Martín

Deputy Secretary of the Spanish Ministry of Science and Innovation

Good afternoon, ladies and gentlemen, first of all, I would like to apologise for the absence of Minister Pedro Duque. He was scheduled to attend this event today, this inauguration, and to be present at this very interesting conference, but unfortunately, last minute issues have prevented him from being here today and it falls to me to represent him in the best possible way.

It is no coincidence that it is being held in Barcelona, a city that is very committed to knowledge and technology, a city known internationally for the excellent science that is carried out and the enormous innovative capacities that it brings together. An example of this is the national supercomputing centre itself, the Barcelona Supercomputer Center, and the Alba Synchrotron scientific structure, to name but a few.

As we mentioned in the previous section, artificial intelligence is something that in just a few years has gone from being an area of research and computation to conditioning our entire lives, to being present in our lives in a clear and permanent way.

Without perhaps having adequately prepared society, it is already shaping the reality around us, our jobs, our way of relating and communicating, our health systems, the way we teach, the way we learn and many other aspects that escape me now.

There are well-known cases in history of technologies that have changed our societies. I can mention two, and there will undoubtedly be many more. Steam, initially introduced for transport, but quickly applied in many other sectors; or electrification, which automated industrial processes and changed our cities, social habits and services. These examples brought about enormous changes, altered the economic fortunes of the places where they developed and then of the entire planet. At least in their early stages they all generated winners and losers. In a very short time, universal and immediate access to knowledge became widespread, thanks to the interconnection of all computers through the invention of the appropriate protocols that we now call the Internet. The amount of available data has multiplied extraordinarily in the last few years, including enormous volumes of apparently trivial information, but which, treated with the appropriate techniques and with access to processing capacities, also enormous, allow us to draw disturbing, inexplicable conclusions. Perhaps whoever has access to data and processes and conclusions is in an enormous position of advantage.

If the conclusions of this massive process are accurate, they can help to make more and more money by pushing others out of the market. If they are biased, all sorts of injustices may occur without anyone appearing to be at fault, or responsible.

The current discussion therefore revolves around how the development and use of artificial intelligence should be regulated, what should be regulated and how it should be regulated so that its impact is for the collective benefit.

In recent months, there has been an increase in the number of forums raising these doubts, some advocating a commitment to self-regulation without delving too deeply into ethical principles or specific mechanisms, others discussing the ethical frameworks and principles that should govern them. An activity that has generated a multitude of statements promoting ethical artificial intelligence.

It is very necessary that we continue to create environments for public discussion, to encourage communication on the social and ethical implications of these new information storage and processing techniques, forums with different profiles and different sensibilities that allow us to facilitate collective agreements on the dangers, benefits and opportunities that they entail.

I am finishing because I suppose that the most interesting discussion is the one that follows, but I would like to reaffirm that I am sure that in this congress on artificial intelligence and human rights the discussions will be very intense and fruitful and that, after its conclusion, we will have a better idea of how to collectively regulate this new reality and how to manage together the interesting future that awaits us.

Thank you very much and, above all, best of luck, we will need it.

Thank you!

Àngels Chacón

Catalan Government Minister for Business and Knowledge

Good afternoon, Mr Rafael Ribó, Mr Peter Tyndall, Mr Pablo Martín, Ms Laia Bonet. I would also like to extend my greetings to Mr Carles Grau, director of Mobile; also to Mr David Ferrer, who is with us today, secretary of Digital Policies, a department with which we also work on all these issues; Ms Juana Barbany, director general of the Digital Society, and Ms Elisabet Nebrera, director general of Social Affairs. I would also like to thank the collaborators and sponsors, without whom this conference would not be possible, and of course the speakers.

The first thing Mr Ribó mentioned was that we are aware that we are incompetent in this area and I think that is the first step we have to take, to relearn.

Thank you very much for inviting me to be part of this inauguration, of the inauguration of this conference on artificial intelligence and human rights. And the same pleasure we feel at being invited as an authority is in turn an obstacle for us to be able to stay for what would really be a satisfaction for us, which would be to stay for the whole day and learn.

First of all, let me thank the Catalan Ombudsman and the other ombudsmen and ombudswomen, because we also need women, and here I especially salute Mrs. Catherine De Bruecker, since we also need women in this field, for their necessary work in defence of people and rights. And this concern for the well-being of all citizens must also include the reflections that are raised at events such as today's. What is being discussed at this conference is the need for women's involvement in the fight for human rights.

What is being discussed at this conference is how we can all work together (citizens, administrations, companies) to design a future in which new technologies are also our allies in making the world a fairer and more prosperous place. We know that technology must be an ally, but now we also need to know how to place it for the collective benefit, that beyond its attractive aesthetics, it has an ethic behind it.

We are interested and enthusiastic about technology, everything it brings us and what it allows us to build, but we should also be interested and enthusiastic about the brain of technology, the ethics and values on which it is built, how to make it open, accessible and beneficial for all, avoiding control and the concentration of power in the hands of a few. And to do that we need knowledge. Knowledge is the key element that will also mark the competitiveness of territories and companies in the coming years. That is why I am especially grateful to the Catalan Ombudsman for his willingness to increase and disseminate knowledge in such an interesting field as artificial intelligence. And above all for the interest in what lies beyond the simple technology, the interest in its impact on human rights and the ethics of its operation.

Today's conference focuses on artificial intelligence and human rights, and just a fortnight ago we presented Catalonia's artificial intelligence strategy at the Palau de la Generalitat. On that day, at the Palau, it was commented that this is a strategy for the country, a further step towards positioning Catalonia as a reference among knowledge societies, which is what we must aspire to if, as a country, we want a richer, more innovative society that is open to the world.

Artificial intelligence, therefore, is at the heart of this government's strategies. And once again, knowledge must be the central axis from which to grow our society and our economy. But this entire strategy would lose its meaning if it were not based on prior reflection, on aligning it with the Catalan Charter of Digital Rights, which are the rights of the 21st century. These include, for example, universal access to the Internet, an open, inclusive and plural governance model with a diversity of actors, freedom of expression and information, and ethics in the use of artificial intelligence and algorithms.

We must understand and know how to use new technologies such as artificial intelligence. We have increasingly sophisticated tools. But the aim must always be to endow these tools with the necessary ethics to help design solutions for the common good and solve the country's social challenges, such as mobility, the educational model, health and active ageing, climate change and economic growth.

Artificial intelligence is one of the main technologies impacting the daily lives of all citizens and revolutionising the competitiveness of companies. It is therefore particularly important that its learning and constant evolution should be based on parameters where everyone, everyone, can feel comfortable and, above all, where our rights are guaranteed. Genís Roca, who is also present here today, explained this very well in one of his magnificent articles. He said: We should not debate what an artificial intelligence is capable of, but rather how we humans should intend to live in the world we are creating. Who will do the tasks? On the basis of what data? Under what premises? To whom will they be accountable? How will wealth be distributed? On whom will responsibilities fall? And to confirm whether these technical developments will actually result in improvements for the livelihoods and comfort of the majority of the group and not just the few.

That is why it is important to make this effort, to understand their enormous potential, but also to put human rights at the centre. We cannot afford resignation and stop questioning how this artificial intelligence is generated, under what umbrella and with what learning model. I am sure that the contributions made at this conference will be very useful to further refine its application for the benefit of society as a whole. And to ensure that artificial intelligence becomes an agent for society, it is especially important for Catalonia to be a proactive subject in its development and not a mere spectator of its evolution.

Today Catalonia is one of the main European research poles, as many indicators show, and this is especially true in some areas of knowledge, such as artificial intelligence itself. Approximately two thirds of the scientific production in artificial intelligence in Spain comes from Catalonia. There are many technology and research centres in Catalonia that also work in this field. Currently, as indicated in a study on artificial

intelligence carried out by ACCIÓ, we have identified 179 companies in Catalonia in this field and companies that employ almost 9,000 people. We are therefore talking about a sector that is growing and will continue to do so in the coming years.

We must see technology as an ally and not as a threat, because only in this way will we be able to bring its benefits closer to the collective welfare.

Finally, I would also like to highlight the importance of public-private collaboration. Spaces such as these for reflection and consensus between all spheres of society also contribute to enriching the strategies that strengthen us as a country, especially in debates as important as the defence and integration of human rights in a world full of new technologies, such as artificial intelligence. Each of us, aware of our duties and responsibilities, must do our part, knowing how to add up to continue making Catalonia a country of opportunities.

Now I was thinking, as Catalan minister of Business and Knowledge, that we must develop different strategies from the Department to promote this productive, innovative, sustainable economy that is open to the world. Yes, should we all generate knowledge together? Yes. Do new technologies need to evolve? Yes, but above all, we must never lose our awareness. It is this awareness that we must all have. Spaces like today's help us to be more aware of who we are, of who we want to be, of our individual and collective rights and freedoms. Of our individual and collective rights and freedoms, because without this awareness we will never be a complete society. Thank you very much.

Laia Bonet

Third Deputy Mayor at Barcelona City Hall

Ladies and gentlemen, good afternoon and thank you very much for participating in this conference on artificial intelligence and human rights.

First of all, I would like to thank the Catalan Ombudsman and the International Ombudsman Institute for the invitation and, above all, for organising this conference.

Thanks also to the speakers who are with us today and who will undoubtedly contribute, over these two days, to extremely important and urgent discussions.

A few weeks ago, Daniel Innerarity, who is also taking part in this session today, asked in an article: Will artificial intelligence be our last invention? It is a question that sums up very well the meaning of this conference.

The discussion on artificial intelligence, and the effects it has and will have on our lives, is the most relevant debate of our times. And it is so because it brings us back to the essential question: What does it mean to be human today? As José María Lassalle, who is also participating in this afternoon's second panel, points out, the very horizon of our identity is exposed to the challenge of a new otherness. And this new otherness forces us to reinvent ourselves, while at the same time making us return to the basics and, if you will allow me, also to the classics.

We need not forget that throughout history we have only progressed when we have placed humanity at the centre of reflection and also at the centre of action. Humanism was characterised from the beginning by putting humanity at the centre of philosophical and political reflection. It was necessary to think about the problems of humanity as a whole. From this concern derived the several political ideologies that have accompanied us for the last three centuries. They were all based on the idea of progress as a tool to solve the problems of humanity.

For some years now, however, progress is no longer what it used to be, the future is not what we thought it would be and the dystopias of the great classics of literature, and also of cinema, are too close to reality.

That is why we need a new humanism. A new technological humanism, if you will allow me to call it that, as we like to do at Barcelona City Council. A technological humanism that puts humanity back at the centre of the conversation in a world that is already structurally technological and that is altering our interpretative frameworks.

Artificial intelligence is holding up a mirror to our fears, and also of course to our contradictions, and to overcome them we need to empower ourselves, we need to take responsibility for controlling the automation of the world.

We are in a moment of disruption, of profound change, and we must avoid two dangers: the danger of inaction and the danger of fear. Both are closely related because fear often paralyses us.

We are at a time when we must remember the words of President Roosevelt at his first presidential inauguration in 1933 when he said: "Let me assert my firm believe that the only thing we have to fear is fear itself". The unjustified and irrational terror that often paralyses us. A maxim that is valid for the United States in the crisis of the 1930s, but which is also valid for today's world with the fear of robots and artificial intelligence, or global contagions - as in the case of the coronavirus - that can paralyse us completely.

We must guard against this risk, we must avoid being paralysed by fear -often irrational and unjustified- and we must turn this disruption into an opportunity.

We have the challenge and the opportunity to put citizenship and digital rights at the centre of the technological debate. We are at a crossroads and, depending on the path we choose, we will end up building one future or another. If we choose the path of rights, we have to defend a new technological humanism, to think and to develop technology in the service of citizenship and general interests.

To speak of technological humanism, as we at Barcelona City Council understand it, means: increasing the level of responsibility of both companies and administrations, both developers and users; increasing our decision-making capacity to continue to maintain control over artificial intelligence; and, finally, regulating algorithmic models by thinking about their usefulness for society, introducing ethical criteria and criteria of public interest to avoid new discriminations.

Algorithms work on the basis of data, and bias in their collection can condition all the decisions we make. Therefore, we need to implement a good data collection policy.

In Barcelona we want to turn technological humanism into a city commitment, a commitment to the regulation of artificial intelligence and the massive use of data, and that is why we have promoted an alliance of cities for digital rights.

This conference that opens today allows us to deepen our reflection on how to guarantee the digital rights of citizens, which should be considered as a new generation of human rights. The foundation of this new generation of digital human rights must be the full protection of human dignity. We must avoid a digital age without rights. Indeed, for some experts, the dictatorship of the algorithm is so real that one could already speak of a digital emergency situation, as we do or have become accustomed to do in relation to the climate emergency.

We therefore need to respond proactively to this situation and the only really effective tool we have to regulate technology, and I want to say this very clearly, is politics. The political system that defends the general interest through laws. This is what I understand the Catalan Ombudsman was referring to in his speech when

he spoke of the need to regulate, to establish the regulatory frameworks to make this possible. Only public administrations, only governments if you will, on a local, national and global scale, will be able to harness the unbridled horse of technology.

The control of data is at stake in the big cities, and that is why we in the city of Barcelona are also very committed to intervening, participating and promoting this reflection. And that is why it is so important that we regulate the deployment of 5G in large urban concentrations. In cities there are millions of cameras that monitor what happens in our streets and we already have the technology to incorporate facial recognition, with all that this implies. European cities have to choose which model of public space we want.

Of the eight most monitored cities in the world, seven are Chinese. Do we want to go down this road, or do we want to go down another one? If we want to choose another path, we have to work to make technology a tool at the service of people and their rights, and not a tool at the service of governments and big corporations. Cities must be the first line of defence of digital rights.

Artificial intelligence should help us to respond to citizens' problems and facilitate the democratisation of the city, not the other way around. If we know how to use them correctly, artificial intelligence and the use of public data by administrations can be an opportunity to offer new services to citizens and, above all, to offer them in a much more proactive and direct way.

We need to define a new smart city model: a city that thinks for itself and gives meaning to the technology it develops and applies; a human city that puts humanity at the centre of technological development; and a city that is democratic, inclusive and oriented towards the common good. To ensure that technology is a tool for making the right to the city effective and guaranteeing what we like to call the right to the future, Barcelona City Council is linking the digital transition with the sustainable development goals of the 2030 Agenda.

We are convinced that we will not achieve the sustainable development goals in the economic, environmental and social spheres if we are not able to use technology and digital transformation to have cleaner energy, more efficient mobility, social policies that are better targeted at the different groups they serve, and more efficient and intelligent public services.

We don't just need more public services, but better public services: more effective, more efficient and smarter. And in Barcelona we are trying to do just that. We have launched several challenges for companies to propose technological and innovative solutions to respond to the specific problems that exist in the city and which are, unfortunately, increasingly present.

We believe that the development of 5G technology opens up an opportunity to influence its urban development model, making it green and neutral and putting it at the service of improving municipal public services. That is why we have established a 5G award for the best technological initiative in response to a social challenge.

In the first edition, which we awarded a few weeks ago, the prizes went to two innovative projects to provide assistance and support to elderly people living alone. The first is a small personal robot capable of moving and interacting with the user to provide assistance and support, and the second is a videoconferencing device optimised for very simple use, to facilitate the performance of remote activities by user groups. These are pilot projects that can be used to develop new, more personalised and adapted services, thanks to the use of this artificial intelligence that we have, that surrounds us and that we want to use for all these objectives. These projects also help the technological ecosystem that exists in our city to collaborate with the City Council on the basis of shared objectives.

Barcelona is already a city of reference in innovation and technology, but we also want it to be a city of reference in technological humanism. We want to promote reflection and, at the same time, provide solutions and concrete measures to humanise technology. We are working to turn Barcelona into the capital of technoethics by involving companies, social entities and universities.

Next December, for example, the city will host the Digital Future Society, the international forum that will address the challenges of technological transformation and seek a public and private commitment to advance in the good governance of the digital future. An initiative promoted by the Ministry of Economy and also by the Mobile World Capital with the aim of improving the impact of technology on society in order to build a fairer and more inclusive future in the digital era in which we live. We want this forum, which has already organised several discussions in our city, to have continuity and also allow us to move forward in concrete initiatives, of reflection and action, and of landing on the ground.

To make this possible, I am convinced that we can count on all of you who share this concern today and who are attending this conference. We are counting on all of you, we encourage you to participate as well, and with your proposals and reflections I am convinced that we will be able to carry out a very enriching discussion. I am also convinced that we all share the idea that the future is not waiting for us and that we need urgent answers. Thank you very much!

Peter Tyndall

President of IOI and Ombudsman of the Republic of Ireland

Good afternoon. What a pleasure it is to be back in Barcelona. I want to start by thanking the Síndic and President of IOI Europe for organizing this conference today. His office has often been at the cutting edge of ombudsmen work and has identified issues which will have a big impact on the way that the ombudsman institution develops in the future. He has regularly organised conferences around emerging themes such as this one, at important moments.

To say that we live in a changing world is a truism. Nothing remains the same for long. The work of the ombudsman isn't immune to this. The world we work in now is very different to the one in which the first ombudsman in Sweden operated more than two hundred years ago. The principles, though, which underpin our work remain the same. We are independent, we are free to access, we are objective. Our job is to balance the power difference between the administration and the individual. We are an integral part of democratic states. We provide redress for individuals who suffered injustices while also using the lessons from our work to improve public services and to stop errors from being repeated. The institution of the ombudsman has proved itself to be remarkably adaptable to a changing world. Without losing sight of its fundamental characteristics, it has taken on many roles and responsibilities, and learned to operate in a highly changed environment. The successful evolution of the ombudsman concept is reflected in its global spread. The IOI, for example, has almost two hundred members in more than a hundred countries across the globe.

What I want to do now is consider some of the current challenges faced by the institution, particularly that posed by AI. In the past, ombudsman offices generally received complaints in writing. People were given a form to fill in, or came to the office in person. It often took a long time for the complaint to be investigated, and for an outcome to be determined.

Nowadays people transact most of their business online: they shop, they do their banking, they book their travel and accommodation, they communicate and, most importantly, they deal with public services. They've come to expect not just speedy results, but very

high levels of interactivity. They also expect to comment on their interactions and to rate them. My generation has digital as a second language. However, many people are digital natives; they've grown up with smart phones and the Internet, and they can't conceive of the time before them. They expect to interact with public services as they do with Amazon or Facebook; to communicate as they do on WhatsApp. For many, the last time they wrote anything was in an exam. The pen, for them, is obsolete.

Ombudsman offices are also used to dealing with bodies who held their information on paper files. Now, however, more and more of it is in the cloud. We haven't been used to delivering instant results. We don't expect high levels of interactivity with complainants, and we don't expect our performance to be reviewed online and the subject of social media discussion. Our colleagues in the UK have had some experience of groups of complainants coming together to criticize their performance. It's now much easier to do this online. And the tenor of social media conversations can be very angry, or even abusive. People say online, things they hesitate to say to your face. Most ombudsman offices are responding to these new circumstances; they are making it easier to complain online; many now give status updates to complainants.

The e-People system in Korea is a long established example of a joined-up approach to complaint handling. The same complaint system is shared by public bodies and the ombudsman. This makes for a much speedier transfer of complaints, which are now resolved locally, and provides a rich source of learning from complaint data. Complainants will increasingly want to interact with our offices, as they do in other parts of their lives, whether by Facebook, Instagram or WhatsApp, to name the platforms most used in my part of the world. They also expect much quicker responses, and are quick to express their displeasure at any delay, or where we don't uphold their complaints. The need for speedy outcomes will further reinforce the trend to seek to resolve complaints through interacting with service providers, rather than to painstakingly investigate each complaint and to produce a report many months later. This in turn presents challenges for how ombudsman offices respond to trends in complaints data. It makes it harder to ensure that systemic issues are identified and dealt with. It isn't enough to resolve complaints. We must also make sure that we improve public services and avoid repeated errors. We must also consider how we can best upskill our staff to deal with evidence that is on systems and not on paper.

Just in passing, not everyone is engaged in the digital universe, perhaps because of age, although there are many silver surfers, or because of deprivation. So we still have to ensure that while engaging in this digital universe that we continue to make ourselves accessible to people who don't transact their business online. We must also look to instances where individuals are denied access to services because public bodies will only interact online.

We've adapted with various degrees of success to the digital environment in which we now work. Artificial intelligence will pose new challenges for us. Recently there had been reports, for instance, of a breakthrough in detecting cancer using artificial intelligence; machines can now read mammograms as well as humans; operations are being conducted using robot technology driven by AI. Some of you may have experienced artificial intelligence in dealing with businesses by phone, or through an app or website. Artificial intelligence is enabling self-driving cars. It soon will be endemic and people say it's likely to cause many job losses.

Many of the public services in our jurisdictions will embrace AI. Decisions that are now made by people will be made by computers. As we've heard, a new branch of ethics covering AI is being developed; previously unimaginable dilemmas are being contemplated. If the car driving itself comes across a parent with a baby in a pram crossing a road and can't stop, should it run them down and save the occupant of the car, or crash knowing that this will endanger the life of the occupant? These are ethical questions of a kind that were unimaginable before AI.

We'll need to evolve our analysis of complaints to determine whether injustice has arisen as a result of the use of artificial intelligence. Not every circumstance can be foreseen. Sometimes the use of discretion can be essential to take account of the particular circumstances of an individual. Can an artificial intelligence do this? The answer is probably yes if it was taken into account at the design stage.

We'll have to decide if a computer is capable of maladministration, and if so, how can it be rectified? How can we stop the same injustices occurring again? The black box problem is one that will undoubtedly be featured in our consideration in this conference. As machine learning becomes endemic, outcomes will be reached where the process of decision making is opaque. We are used to determining whether there has been maladministration by examining the decision-making process, and seeing whether the steps taken are justified. This will be potentially much more difficult in the future. It's easy to ask a person how they arrived at a decision. That won't always be the case with AI.

We must also consider how the ombudsman can use artificial intelligence. Can we use it for instance to determine if complaints are in jurisdiction, if they are in time, whether there's evidence of injustice? Can it interact with our complainants rather than having staff do this? Can it play a role in investigating complaints? The very least, it could play a role in identifying systemic problems by considering the data from complaints.

It's clear that the environment we work in will continue to change and bring new challenges, but also opportunities. I'm confident however that the ombudsman institution will continue to adapt but remain true to its core values and principles. Providing access to justice, improving public services, promoting the rule of law, and vindicating rights and fundamental freedoms have always been our guiding principles, and we'll continue to be guided by them in whatever circumstances evolve in the future. This conference is very timely, and will help us to understand the current and future evolution of AI and how our offices must adapt to seize the opportunities it offers and face the undoubted challenges. Thank you.

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Monday, 2 March 2020

Facts and limits of the artificial intelligence

Chairperson: Andreas Pottakis, Director of IOI and Greek Ombudsman

Ulises Cortés, Artificial Intelligence Professor at the Technical University of Catalonia (UPC), Barcelona Supercomputing Center (BSC)

Montserrat Pardo Bayona, Director of Corporate Affairs, Microsoft Spain

Alfonso González, Government & Regulatory Affairs Director, IBN Spain

Oliver Smith, Strategy & Ethics Director at Alpha Health

Carina Lopes, Head of the Digital Future Society Think Tank (Mobile World Capital)

Andreas Pottakis

Director of IOI and Greek Ombudsman

I'm delighted to be here. I'm delighted to be back in Barcelona. It has been some time; since 2017, again for an event organised by the Síndic.

An event that I think had a lot of relevance to what we are going to debate today and tomorrow, because part of the discussion back in 2017 had to do with manipulation of the internet and of different platforms in order to promote hate speech and xenophobia.

I mention this because I think that, as I will be making a couple of comments before I give the floor to our distinguished members of the panel, one of the major problems and one of the major threats and risks that we ombudsman institutions, in particular, and human rights institutions in general, have to address has to be the threat and the problems that artificial intelligence and the use or abuse of artificial intelligence may have on our civic rights and our civic consciousness.

So, first, I thank the Síndic for his invitation and for allowing me to chair this first panel. I think that I share with the Síndic the same ignorance on the subject matter. In fact, I have the suspicion that he wanted me to chair because he wanted me to expose my ignorance!

Having openly and unashamedly admitted my ignorance, it would be safer for me to limit my address to offering some concerns and, perhaps, to posing some further questions in the debate, although I am afraid that most of the points that I was considering to raise have already been addressed by Peter Tyndall, the President of the IOI. Clearly, as an ombudsman institution, we are dealing with human rights and we are also dealing with maladministration, but we have the human at the core of our work. Now, I think we are about to be addressing the so-called hackable animal and not the individual human being anymore.

The subject of the first panel, that I think that my distinguished friends here will help us address, has to do with where we are at present and probably where we are heading in the near future when it comes to artificial intelligence. So, I have a lot to learn personally. From what I understand, and you will correct me of course, we are still at the phase of this so-called narrow artificial intelligence sort of application, encompassing the most common use for us users of internet platforms, etc. The typical examples are content personalization or filter bubbles etc., which have to do with creating some sort of a virtual image of who we are or at least, according to the different algorithms, what we pretend to look like. I understand also that at some point what we will have to address is the so-called

point of singularity, where we will be reaching a different level of application when it comes to artificial intelligence. I hope you are smiling because I am not far from where we are, but you will clearly correct me.

While it is essential for us ombudsmen to seek and receive the assistance of the experts in order to understand first of all what artificial intelligence really is, what artificial intelligence can do, I also understand that what tops the list of our concerns is human oversight; to debate and to reflect upon the type of human oversight on artificial intelligence. I think this point has already been mentioned or at least insinuated by both my colleagues who preceded me.

I would like to conclude my first, very brief, I hope, comment by reflecting on a very recent piece of document that is of significance to us ombudsmen. I am referring to the European Commission White Paper on Artificial Intelligence, published just ten or so days ago. I note that one of the final recommendations or conclusions upon which the European Commission is focusing at its White Paper is the issue of governance when it comes to artificial intelligence. I take it to mean also a way of securing human oversight, a way of reflecting on the type of human oversight that we will have. Will it be again of the type of an ethical sort of code? Will it perhaps take the form of an enforceable set of legal instruments? Will this be in addition to the existing legal instruments that we are applying? I'm referring, for instance, to the legislation on anti-discrimination -as it has already been mentioned that artificial intelligence may provoke bias and discrimination- on consumer protection or the GDPR, the notorious GDPR with which most of us are having problems.

But my main concern here is not on ensuring protection of our consumer rights, it's protecting our civic rights, our political rights from possible threats that may arise from an unregulated application of artificial intelligence. As an ombudsman, my great fear, and this is how I will conclude, is that even present applications of artificial intelligence, uses that already form part of our daily routine, may not only constitute a breach to our freedom of expression, but may also present a serious risk to our freedom to form an opinion to begin with.

Although I promised to be brief, I fear I may have tired most of the audience!

So, now it is time to listen to something much more intelligent, and we'll start with Ulises Cortés, the professor and senior researcher at the Polytechnic University of Catalunya and Barcelona Supercomputing Centre.

Ulises Cortés

Artificial Intelligence Professor at the Technical University of Catalonia (UPC),
Barcelona Supercomputing Center (BSC)

Thank you very much for the invitation to this event. I will try to draw, in 15 minutes, the essential facts and technological limits of Artificial Intelligence (AI). The goal is to explain my view on AI and its actual limits¹, but first, a disclaimer: I am just a professor in Artificial Intelligence. I am not an ethicist. I may have significant errors in using some terms. I am not a lawyer. Also, I am not a futurist, therefore I am not trying to convince you that I have a clear view of the future. My idea is to clarify what we know now about the uses of AI and what we can build with the existing AI-tools and computers that we have at hand.

The AI impact on Society

Today, you are here because AI is massively impacting on society. There is no question about that. Big companies - like Google, Amazon, Facebook, Apple, Microsoft and Alibaba, contribute to the confusion by making much noise about their use AI to transform our way of living and earning a lot of money out of it. They are creating great marketing campaigns telling us which is the extent of the capabilities of their technologies and how helpful they are. Furthermore, what we know is that they are putting on the market, at breakneck speed, technologies that have to be adapted and corrected every day. Hence, they have at disposition new versions of that technology all the time, and none but them is capable of understanding the actual functioning of those programmes. At this point, we have different points of view about the possible impact of AI from the most extremist and not very positive views of AI, like the one of Elon Musk and Stephen Hawking preventing the misuse of AI-based technologies. On the other extreme, we find people trying to regulate its use to avoid misuse. However, maybe it is not only AI but also its combination with the evergrowing computer power, a lot of new uncensored algorithms and a lack of market regulation what is creating today's situation. It is evident that AI impacts us, but we have to differentiate the marketing noise from the real effect of AI, and this is not an easy task. We have to be clear on the fact that AI is just a part of the problem, the most visible one, but it is not the whole picture.

To use common terminology, let me start with a working definition of artificial intelligence. There are many, I take this that was coined by the European High-Level Expert Group on Artificial Intelligence:²

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.

¹ During the original talk I used, as a support, some slides. To clarify this text, I am including as section names some of the titles. Also, I included some images from to illustrate some concepts.

² <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

Most of my references will be to the European Union documents because Europe has its own position about the permissible uses of AI. Europe wants a trustworthy human-centric AI, and I adhere to it, even if have some doubts. AI-based systems are a type of software that can act both at the Internet, that is a virtual space, and in the physical world. Moreover, what is more interesting is that these machines not only accumulate information - people often only talk about the use of personal data, data is just a part - but also accumulate knowledge, and this is the exciting part of AI.

Those intelligent machines use data and knowledge; both are used to make decisions. Those machines are making decisions for us all the time, representing us in many situations. This use is changing the way we are living and interacting with the world. Most people are now giving-up this part of their responsibilities to a machine for them to make small decisions, for example, those made by a route optimiser when we are trying to go some location. Those parcels of machine autonomy are becoming bigger and bigger because they are aggregated to extend its number and its frequency. As an answer, the European AI community is trying to convince people that they need to get educated in technology to be able to understand what is behind these algorithms and their impact.

Broadly speaking, there are two kinds of AI: one that is based on the use of symbols to reason about a problem to solve it. Also, it can make a transparent explanation about how it uses knowledge and data. The other does not use symbols; it works as a black box to solve problems; those systems cannot explain how they solved the problem. Those systems are better known as deep learning or machine learning. The first thing you have to know is that deep learning is just a small part of machine learning and machine learning is a part of AI. Deep learning is popular today because of the marketing it is receiving and some spectacular results. Furthermore, we have to avoid to be fooled by the idea that AI, in general, or deep learning, in particular, can do everything overperforming humans.

How far or near are we from this promised land of AI?

How far or near are we from this promised land of AI? To find out this, let us examine some examples of AI applications (see Figure 1). Today, one of the most common expectations about AI-based systems is that soon we will have autonomous³ vehicles. Personally, I cannot understand why anyone wants to have an autonomous car because I want the car to be free and go elsewhere. I do not need a machine to drive me somewhere. I do not need a machine to communicate all the time where I am or which is my current speed. We are not close to cars with full autonomy, but there is very intense research to produce those cars. Meanwhile, as a result, today, we have useful some driving modules, for example, for people that may have some disability. Some cars can support parking or park themselves, etc.

Now, let us briefly touch military uses of AI. Armies in several countries have already semi-autonomous weapons. However, we are very far from seeing real autonomous weapons. Nevertheless, maybe not very far, because of course most of the AI technology has emerged from military research and armies still investing in this kind of research. Common sense invites to forbid those weapons able to make autonomous decisions to achieve their goals or even to change objective. If we are unable to forbid them at least, we need to regulate them.

³ Tim Smithers says that robots and software agents are autonomous if they are able effectively to deal with the kinds of environments in which humans live and work: environments which have significant large scale spatial and temporal invariant structure, but which also have large amounts of local spatial and temporal dynamic variation and unpredictability, and which lead to the frequent occurrence of previously unexperienced situations for the agents that interact with them.

Sesame Street

NEAR	FAR
Assisted Driving modules	Autonomous Car
Semi-autonomous weapons	Autonomous weapons
Decision Support Systems	Fully Autonomous Decision Systems
Transportation route optimization algorithms for project planning optimization	Autonomous Building Systems
Clinical image recognition	Autonomous medical diagnostic systems based on clinical images
...	...



Figure 1: Promised land of AI

Let me introduce some terms to classify weapons. A rough classification contains these classes: controlled, supervised, automated and autonomous weapons. Those define the various degrees of human supervision upon a weapon. For example, a drone that is controlled with radio control in all the actions it performs belongs to the first class. Those systems have supervisors, that is a machine that has a goal, and will perform it under the supervision of a man that is always in the loop. The automated machines, like the landmines -I do not know why we are still producing those- are prepared to blindly react to any input. Then you have the future semi-autonomous weapons like the robots produced by Boston Dynamics in the United States. There is an ongoing discussion here, either you like to automate things, or you want to extend the human capabilities using AI. The uncontrolled accumulation and spread of weapons of all categories has a devastating humanitarian impact. If you go to the automatization of a process, you are losing control over the device.

When we say that someone is extended? For example, if someone is extended if she uses lenses to extend her visual capability. Now we can extend someone's capabilities without that person having a problem or disability. For example, wearing infrared glasses to see in the dark and measure the distance to the target. In the same way, we can talk about AI-based tools that extend human capabilities. From those, we can distinguish good uses of AI from the wrong ones.

One example of double use are biometrical authentication tools. That is a beneficial and handy application. For example, when it is necessary to restrict access to critical zones, this system will only allow entrance to authorised individuals, no one else is permitted. A critical zone may be your telephone that uses face recognition to allow open it. This kind of tools becomes alarming if they used for permanent and remote biometric identification. In this case, individual rights are compromised. Especially when systems are used to continually sense people using their biometric information. Public security issues are complex and not always are black or white.

Another possible misuse is surveillance capitalism or surveillance tasks where AI is used to collect data from us, data is harvested elsewhere and then integrated to identify, classify or profile people. This already happens every day, in many times a day in any device that is connected to Internet. Every one of us who is not closing the localisation service of their mobile phones is giving information not only about themselves, but also about all the people in their surroundings. We have to be socially convinced that it is not good to give, without our explicit conformity, that information.

Among other wrong uses of AI, we have AI-empowered fake news and cyber-attacks. Those are part of the dark side of AI. One of the things that I want to recall is that AI is not harmful by itself. This is not the first time that research or scientific results have being used unethically. We have examples back in 1776 when Jensen performed the first vaccination experiment. Let me enumerate a few the Tuskegee Syphilis Study-USPHS, US Department of Energy studies on radioactive fallout, etc.

Ethics and AI

To answer the unethical uses of AI, in 2017, the AI community proposed the Asilomar principles. Also, in 2017, the European AI community proposed the Barcelona Declaration of Good Use of the AI, led by Professor Luc Steels and Professor López de Mántaras. The ethical guides proposed by the High-Level Expert Group are the European political position.

Nowadays, everybody is creating their principles that fit their interests. We have plenty of principles, but they are not respected at all. Even large companies are contributing to this confusion. There is an article in the journal Nature, published at the beginning of 2020, in which authors report more than ninety-four different declarations from associations, countries and companies. Ninety-four! They range from seven to twenty-seven principles of AI, and only four of them are common to all. Moreover, if we do not have enough, also the Church now is involved in this movement. Microsoft and IBM are looking for having an alliance with the Church for an ethical AI. I hope you now have a clearer idea of why AI is at the center of many social, political and economic discussions.

Let me talk briefly about what we are doing in Europe. As I said, we have an experts' group, which produced ethical guidelines for AI⁴. These guidelines establish three requirements for AI to be permissible in Europe. First, it has to be a lawful AI. This means that it complies with all laws. Mostly all those that are created by the Council of Europe and the GDPR. Second, this AI has to be ethically aligned, and third, it ought to be robust. This last characteristic is important because companies are rushing to put their technologies in the market before they are thoroughly tested. We have to rethink what is the exact meaning of *robust* in this context.

In the HLEG-AI guidelines, there are four ethical imperatives: the principle of autonomy, as I said before, it requires a human on-board all the time. We have the principle of non-maleficance; the action of AI, either software or hardware, do not harm humans. Principle of justice, it should be fair. We have to define the concept fairness, but,

⁴ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

basically one of the things that we need to have is machines that are bias-free, that means software creators have to prove their data and their programs before they go in the market. Moreover, the principle of explicability: we need machines able to explain why they are making decisions. If I go to a doctor and she diagnoses cancer to me, I want to know which is the evidence for that and the reasons behind. If I am not convinced, I go to get a second opinion. I do not like the situation in which if I have a diagnosis made by Google, and if I do not like it I go to Yahoo to get a new diagnosis. Replace humans in those tasks is not a good idea. We need in Europe to have an ethical human-centric approach. We need respect for human dignity; freedom for individuals; respect for democracy and justice, and we will see examples of how we can affect that; equality; non-discrimination and solidarity, and the civil rights have to be in the first place.

Data deluge or data economy

Conclusions

Now, I will describe the EU view on data and AI. There is a White Paper on AI that is focused on the new data economy. The objective is clearly to regulate the new data-based economy:

Europe's current and future sustainable economic growth and societal wellbeing increasingly draws on value created by data. AI is one of the most important applications of the data economy. Today most data are related to consumers and are stored and processed on central cloud-based infrastructure. By contrast a large share of tomorrow's far more abundant data will come from industry, business and the public sector, and will be stored on a variety of systems, notably on computing devices working at the edge of the network.

We all know that the data is one of those engines that are moving all our economy and Europe recognizes that we are far behind in the competition with the USA, China and the gigantic commercial platforms. However, we have an outstanding microelectronics industry in Europe. It is possible to use that technology to bring computation to the edge and use the data that is produced. The question is how to use it fairly?

Knowing that data is the most valuable raw material, let us talk about the avalanche of data in which we live immersed, let us talk about big data. I would not say that I like the term big data. I prefer to talk about the right data. To give an idea of the amounts of information we are talking about, see figure 2, so this will give a panoramic view of what we are talking about. For example, and if my memory does not fail, artificial intelligence was born in 1956. In the same year, IBM produced the first data storage disk with five megabytes. Back then each megabyte was 10.000 dollars.

Now, you have in your camera plenty of GBs. Everyday internet is producing 10 to 18 Zettabytes, and growing with augmenting use of the Internet of Things (IoT). It is a mindboggling amount of data. The European Union claims that this year (2020), we will produce 40 Zettabytes. Moreover, in the same White Paper, they are forecasting for the next five years. In five years there will be 175 Zettabytes a year. As a society, we are producing tons of data, see figure 3.

The Data Deluge

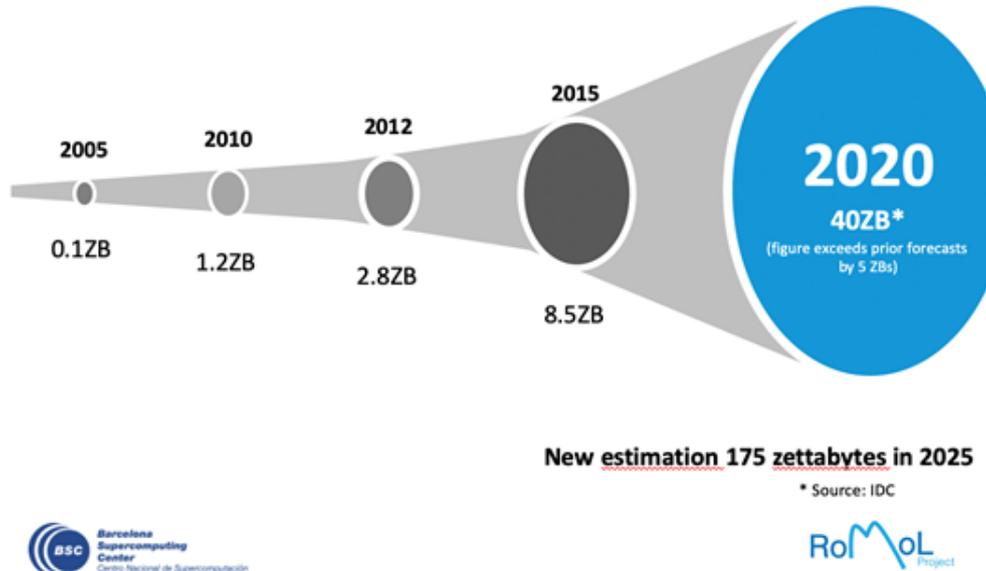


Figure 2: The data deluge.

Who controls my data?

One may wonder, where is all that data? What are we going to do with that data? Who assumes responsibility for the data? On an individual level, the question is, who controls my data?

Furthermore, it is clear that at this point, it is the big companies that are controlling our data. That is true now and will remain valid for an undetermined time. What's more, some companies say they don't want to have all the data. They do not want it, because they already know which is the relevant data to keep is.

We can ask ourselves, do you trust the Internet giants to protect your personal data? What is the profit from the use of our data? Ricardo Baeza-Yates wrote that each year large companies are producing \$ 180,000 per person. In the white paper, it talks about the differentiation between services and products, and all European legislation deals with products, but not with services. In Spanish, sounds better: "los servicios de Google más allá de la muerte"⁵.

Moreover, if you read, because none of us read all the app contracts, this is what the Google contract says about your data, and it says: this license continues even if you stop using our services. For example, when you are dead. They have all the data. They are classifying us. They are trying to predict the way we behave. They are using bubble filters to put our behaviour on a specific track, and if they are not able to make our profile, then they choose to induce us to change habits.

Do you trust your government to protect your personal data from the giants of the Internet? My answer is no; they are not doing that. Do you trust your government to protect your rights, your human rights and civil rights from the giants of the Internet? My answer still no, and if not, you can see what happened with Cambridge Analytica or the Compass

⁵ "Google services beyond death"

cases in the past. So, it is clear that there is a lack of regulation. We have all the means, we have all the research, we are strong in Europe to produce much scientific results, but we do not have the money to pay the researchers. The big companies are paying much money to achieve their goals; they bring all the brilliant minds to do, well, what they have to do. So, the biggest challenge is to avoid become useless disposable idiots just attached to a telephone. The director of the Barcelona Supercomputing Center, professor Mateo Valero, always says that in your mobile phones you have ten thousand times the power that put the man on the moon. And, what is more, I am sure that almost none of you have ever programmed a line of code in your mobile phones. So, why do you have all that power? Do you need it? Or who needs that computing power in all mobiles, to do what?

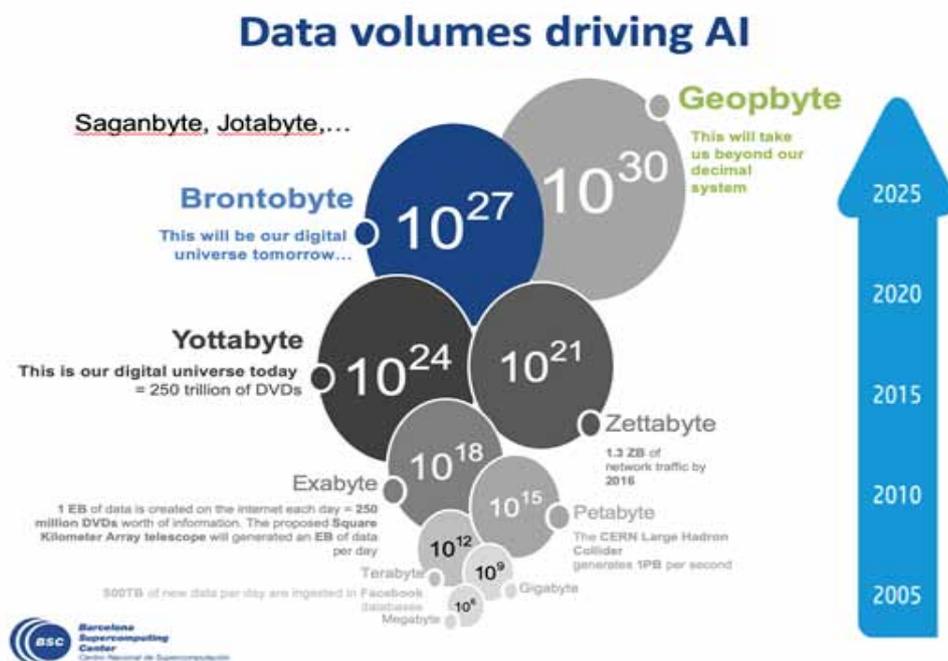


Tabla 3: Tierra prometida e IA.

A takeaway. It is clear that AI is influencing our societies. We are in Europe, let us try, as Europeans, to influence the way AI is marking the pace for society. We can achieve this goal by designing the AI software in a way that is aligned with a common set of European ethical, cultural, legal and socio-economic values that should be standing up there for us. It is obvious, but it is worth to remember that knowing ethics is not being ethical. So, it does not matter if we have all these declarations and regulations; we need to teach to educate our children at home with ethical values. In school, you can complement them, but it is at home where you learn all that.

Artificial intelligence needs to show accountability, responsibility and transparency. I want for the near future the government rules in order that the new Ai-based products have a similar life cycle as a chemical for pharmacy. Those drugs are developed and have some years of a clinical trial, social trial, and not only ruled by companies and government but also the non-governmental organisations should be there. If not, we will lose control. AI systems are artefacts that we build, and we design them. Let us be explicit in the way AI has to shape our future. Machines are not ethical, it is necessary to design them to be ethically aligned, and when they are fully autonomous, they may need another kind of ethics. So, let us be conscious, educate critically and technically of society. Only informed societies can claim their rights and claim for them to be respected.

Montserrat Pardo Bayona

Director of Corporate Affairs, Microsoft Spain

Hello, good afternoon everyone and thank you to Catalan Ombudsman for the invitation.

Today I would like to give you a little bit of our vision, from Microsoft. Our mission is to empower people to achieve more. And why a mission? At the beginning, I was quite anti-American system and I wanted to join Microsoft to really understand how to work in an American company and one of the things I have learned is that if you don't know where you are going, you won't find favourable winds. So, to know where we are going, to really know what each and every one of us wants at Microsoft, because, at the end of the day, at Microsoft, all we are is people. What we want is to empower. Whether it's people, organisations, public institutions. And how do we do that? Well, now that we are talking about the change of the new era, what artificial intelligence is really doing is pivoting and leading this change of era.

What we really believe is that artificial intelligence does not come to replace human intelligence, but rather to amplify it, absolutely, and what is happening is a series of circumstances, that we really have a huge amount of data, an enormous computing capacity and algorithms... and Jordi Ribes, who is our vice-president in artificial intelligence, also originally from here, from Manresa, always says: "We are already reaching results with algorithms that are really taking days, things that we thought would take years". So we are at a really important moment.

And what are we seeing? That countries are taking the lead. There is China, there is France... And where are we? We are also here. Already in Davos, last year, President Pedro Sánchez met and accelerated everything that is the strategy of artificial intelligence. And here, in Catalonia, last week, we were also presenting, together with the Generalitat, with the Barcelona City Council, with the Supercomputing Center, we were presenting what is really something innovative: the centre for innovation in data and artificial intelligence, whose ultimate objective is to really establish Catalonia as an excellent region. It really is a model that can be replicated around the world and that shows that there really is a super-powerful ecosystem here, of talent, people, companies, institutions that really want to increase the productivity and competitiveness of this region.

So there are challenges and there are responsibilities. And here... all these uncertainties, all these fears, what will happen? Will we lose jobs? All these biases, in terms of gender, in terms of race, which are really fundamental rights issues, human rights issues. What's going on? This artificial intelligence thing is really showing the world we live in, and it's making it transparent, and artificial intelligence is learning. That is why it is so important that, right now, machines are becoming more and more human-like. Humanism, ethics, human rights are becoming more and more important in history. Because, for this to work, we will have to work together and amplify that human ingenuity, keeping those human rights and ethics in mind.

And what are we doing at Microsoft? Well, we are trying to communicate better, because today it has been said: "I only know that I don't know anything". The session started today and it is true, but if there is one thing we have learned at Microsoft, it is to learn how to learn, and to learn that we do not know everything, but that we want to be the ones who know and we learn faster. But we are really communicating, trying to communicate to society that this is indeed about opportunities, about tools, but that there are also certain threats, there are certain challenges. I like to talk about challenges. And I encourage you to read this book, because it's really written by Brad Smith and Carol Ann, from our team, and they sit at a global level with all the governments, they listen, we try to build, and from

this book you will see that, and that is universal and timeless. In the end what we're talking about here, whether it's artificial intelligence, in the past we talked about other technologies, is the good use or the bad use we make of those technologies and that's why in the end if we put people, human rights, at the centre, we're going in the right direction.

And indeed, not all intelligences are the same. Artificial intelligences, I mean, because it is not what technology can do, but what it should do, and we have found ourselves many times, and in our case, at Microsoft, we have said no to certain institutions, because we have not considered ethically that the project should be done. Why? You will say Microsoft in artificial intelligence you will have many engineers, many female engineers... We have in that artificial intelligence team, in the whole company, many anthropologists, also, many philosophers, psychologists. Again, multidisciplinary will be more than ever fundamental for this to really succeed. Because artificial intelligence and innovation, what for? If not to make a better world, a world that is economically more competitive, that socially has and ensures that everyone is included and no one is left behind. If we don't make the world more sustainable, we won't be doing it right.

And what did we do, and there we were pioneers, in 2016, four years ago, we launched these ethical principles and they really are timeless, because today you will see that we are talking about these timeless ethical principles, because with that intention they were made: privacy by design; security, all of them as fundamental rights; equity, as a fundamental right; inclusion, as another fundamental right; transparency and accountability. And these are not words, they are deeds, because if we don't move from words to deeds, in the end what good is it to be a person if you don't have values, what good is it to be an institution if you don't have a culture that reflects them?

So, I will not go into each one of them, but you should know that we consider the issue of privacy by design as a fundamental right, and we were the first company that, beyond Europe, we expanded the issue of GDPR globally by design in all our countries, and it is essential to ensure that security. And you know that there are more and more attacks, cyber-attacks, and it is essential, not only at the institutional level, critical structures, but also companies, large and small, which are attacked every day and even us as citizens, and in our role, and that someone can be a politician and at the same time have a personal life, to avoid any kind of vulnerability in that sense, right? Transparency, because if we really, as a company, are not transparent and we do not make transparent all those decisions on which we base ourselves, we would not be able to gain that trust, that confidence of everybody. And of course accountability. And that is in our DNA. And fairness. Because there is a lot of talk about bias.

We do a lot of work to have male and female developers, precisely to avoid all those biases, which can come from gender, but also, not only gender, there are also accessibility issues. Our best developers, the ones who have made us better, have been blind people in our company, who have made technology much more accessible. And we return to the issue and the discussion on human rights. For the first time in Spain, we have hired a young female engineer. You will say: well, that's great! An engineer, young, top of her class... deafblind. I can assure you that this person, Sandra, has made us much better. In terms of improvements in our technology, but above all in the way we work today at Microsoft in Spain. And then, our head of diversity and accessibility is deaf. So we are touching all those axes and we understand diversity and equity in the broad sense. You will see that disability is recurrent in our case, because artificial intelligence is really allowing us to make incredible advances. Today it is possible for blind people to walk the Camino de Santiago, which was unthinkable before, right? And they can do it with the help of technology.

And we are just getting started. And this is not being done by Microsoft developers. This is being done by empathy. The empathy of people who know how to find an artificial intelligence solution, and who are listening to doctors or who are listening to those parents of that child with autism. And really the conjunction, the interaction, again, of the educational world, for example, with artificial intelligence, is what is going to do the magic that happens in these cases.

Why diversity, and why do we champion it so much at Microsoft? Because we have understood that there is no diversity, no innovation, without diversity, and that if our workforce is not a reflection of society we will be wrong. That's why we defend it so much. And I wouldn't want to stand here and not talk about that need to have more women in our sector, because really if we want artificial intelligence that reflects society, we have to get it. And that those women go from using the technology to actually developing it, working on it and developing it.

Here, and in Barcelona in particular, and thanks to all the work being done by the different institutions and by Mobile, and also through projects such as Barcelona Digital Talent or Digital Future Society, I would like to launch this call to action, it is going to be fundamental, here, in Barcelona, in Catalonia, in Spain, throughout Europe and the world. In this race to really make a better society and that no one is left behind, we must all be prepared. And when I say everyone, I mean starting with us, but at all levels and for all audiences. It is going to be a reskilling that is called upskilling and skilling in every sense. But I would go further. What we have to be trained in is learnability, the ability to learn continuously and I would say even more, skill to skill, to be ready to skill, right? That we should always be prepared, capable of training ourselves constantly, because we don't really know what the market is going to ask for, because this is moving very fast.

But we do know that jobs are going to disappear and that we are going to have to be prepared, because others are going to appear. And we talked about some time ago, didn't we? And there is no need to be afraid, because all revolutions and changes of era have generated the need to learn and to change models, but it will really be those societies that adopt these changes first, those that are going to be the most advanced. And that is why I was telling you before, I was showing you China, I was showing you Macron, but we are also there. And here, in Barcelona, there is an incredible ecosystem that is happening here. We have to take advantage of it, just a few examples. Clearly on LinkedIn we see it, because there is artificial intelligence there, that these are really the most in-demand skills and we see it because there is indeed all this data that, for the benefit of all, can be used precisely to train people in what the market is demanding for the first time. Because we have often trained these people in things that were not in demand, and here, simply by way of example, you can see the different audiences and we all have, from the blind to the most vulnerable people and groups, because our future really depends on it. And this is our ambitious, well, aspiration. To reach ten million people in the next five years and that no one is left behind in this skilling.

And this AI for good, again, I would like you to keep this message: artificial intelligence, for what? I would stop this world if technologies are not really going to help us to make a better world. And a better world on three axes. I would put, to put a focus today, apart from the health, which is super important, too, but for a better planet, for really accessibility and for the humanitarian issues. And artificial intelligence, and you have at your disposal, we have launched fifty million for institutions, NGOs, start-ups, which are actually working on projects in these four areas, in artificial intelligence for the planet. She is a person, a partner of ours, who has helped us a lot to generate this type of programmes that, really, in the next five years, we also want to dedicate to accessibility projects, forty million destined to institutions so that they develop. We want this artificial intelligence for the common good.

And why not? I don't want to leave this part out of the end, because there really is a very, very big responsibility in all these cases. We have more and more natural disasters with the issue of climate change. We have more and more children who are in very vulnerable positions and refugee issues. And artificial intelligence can help us. We can't do it from Microsoft. We are not as giant as we seem. In Catalonia, to give you a perspective, we are seventy people. Seventy people who really generate an ecosystem of SMEs, which are the ones that really innovate through our technology, which are forty thousand people here in Catalonia, and two hundred thousand in the case of Spain. These are the ones that make us great and these are the giants. These technological SMEs, these institutions that develop from this base technology.

And as I said, we are not going to win this if we do not do things like those that are being done here around Barcelona, where we are joining forces from the institutions, including also in many cases the central government itself together with the Generalitat, together with the innovation centres, together with companies and civil society, that we are really transforming and keeping in mind those ethical values, those principles, those human rights, that if we do not have them and do not put them at the centre, we are going to make a mistake. So, I invite you to create this future and the best way to predict the future is to create it together. So let's paint it all together and I leave here my contact details if you want to continue the conversation. Thank you very much!

Alfonso González Herrero

Government & Regulatory Affairs, IBM Spain

Thank you very much. Although I prepared the speech in Spanish, I am going to try to do it in English.

Artificial intelligence and human rights. Let me first give you a brief overview of what is IBM. Maybe you have in mind a picture of IBM that is not very current since the company is in constant change. Probably, if we would have had this conference five years ago, or ten years ago, IBM would have been much different to what it is today, related to artificial intelligence. IBM was founded 109 years ago and I want to put this as a first slide, because we are not a company founded ten or twenty- years ago. We are not a company founded sixteen years ago (like some others...). We have a much longer history, we made our mistakes, we learned from those mistakes and we know well what it takes to do things right.

For twenty-seven years in a row we have been leading the patents ranking in the US. We are still the number one in the Interbrand ranking in terms of business services. We are, essentially, a business-to-business company. We are not in the consumer market, although many of our clients are; this is, we sell to other companies. And today we are leading in several areas that, as I said before, we were not some years before, like cloud computing, artificial intelligence, internet of things, business analytics. We have been also in Europe for more than a hundred years. We have offices and centres all around Europe. We have a data centre here, in Barcelona. We have the European Headquarters in Madrid. We run from Madrid, not just the Spanish operations, but also the Portuguese, Greek and the Israeli operations. So, we know what the European values are. We are European and we have been in Europe for more than a hundred years and this is another key message that I want to give you. As I was saying, we are in Spain, we have not been a hundred years, but almost, both in Madrid and Barcelona around ninety-four years and we have centres of international responsibilities in Madrid and in Barcelona; offices, in Seville, Valencia, Bilbao, la Coruña, etc, etc.

So, this is about IBM, and I was saying before, our centre today, our focus, in terms of business is in big data and analytics, cloud computing and, of course, artificial

intelligence, which is what we are discussing here today. Now, we are talking about artificial intelligence, but many times we are referring to things that —and probably professor Cortés knows much better than me—, that are not the same. I mean all the terms, the different concepts that we are using for artificial intelligence when we are talking about different things. And here (in this slide) you have some of these concepts: machine learning, robotics. We prefer in IBM to talk about the *last terms you have in the slide: augmented intelligence and cognitive computing*. Why? Because these terms put the emphasis on artificial intelligence as a technology that is here to help humans, as opposed to the term "artificial", that it seems that it is something opposed to the "natural", like if there was a natural intelligence and an artificial intelligence, and these semantics are probably part of the problem when we approach artificial intelligence. That we are thinking that we are choosing something that is opposed to our natural intelligence.

Now, as professor Cortés said before, the term, the concept of artificial intelligence was first used in 1956 during a conference, I believe the person with the red shirt in the slide, John McCarthy, if I remember correctly, was the one that came up with this term. There were two IBMers at that time in that conference. Two people that worked for IBM. The one on the left, Trenchard More is one of them. This is a picture taken a few years ago to commemorate this summer course in Dartmouth College in 1956. So, IBM has been involved in artificial intelligence since the beginning. I mean, since the very first beginning in Dartmouth College. We know what we are talking about. You will remember some of the most iconic steps in artificial intelligence, like in 1997 when an IBM Deep Blue system defeated Kasparov. And this was a very iconic moment, because chess is believed to be one of the games that humans we thought play better than machines. 2011, Jeopardy, a very famous program in the US, that has been broadcasted for more than thirty years, also an IBM machine, Watson, defeated the two most successful contestants at that time. Today we are involved in what we call Project Debater, which involves teaching machines to debate humans on complex issues. This is where we are focusing today.

Now, in terms of what we are doing in the business? Just a few examples from the last three months: We have more than 20,000 projects on artificial intelligence in the world right now, in IBM. For example, we have an assistant for Orange in Spain helping clients to get answers through WhatsApp. A team from Barcelona, just a couple of months ago, won a contest, an international contest for IBM, called Call for Code, that is helping firemen to be safer when they go to put off a fire. We are helping Cerealto Siro Foods to put news products in the market and analysing consumer trends and putting in the market products that take into account these trends. And, of course, we are in health with our system Watson Health trying to help a better diagnosis for doctors. Helping doctors, not replacing doctors, to do their job

Artificial intelligence is already used in our everyday lives. We use it when we pay with our credit card; when we use our system in our car to drive from point A to point B; in our emails, the spam filters use artificial intelligence for that; when we use our phone, Siri for example, or when we do a search in Google for example. So, it is present already. We are not talking about the future; artificial intelligence is already here and we use it every day. Now, a word that has been already mentioned here: "trust". That it is the key for the adoption of artificial intelligence. If we don't trust the systems, this technology won't evolve. I mean, we trust our car because we know the car breaks when you press the pedal, and we go into an elevator because we know that if we push number five, the elevator stops at the fifth level, or we go to an automated teller-machine and get money from the teller machine, because we know that if we ask for a hundred euros, the machine is going to give us a hundred euros and not fifty. So, trust is behind all this technology. Now, some of the questions that we ask ourselves when we are talking about artificial

intelligence: well, first, is it fair? I mean, is there any bias? Is it easy to understand? Do we understand how this system works? Has anybody tampered with it? I want to make sure that this is a robust system and no one has done anything illegal or that put me in danger. And is it accountable? I mean, if something goes wrong, there is someone behind that is going to be responsible for it, right?

Now, yes, we also have our principles. Not sure if 2016 or before, I won't argue that, we have our principles, and these principles are around what has been always key for IBM, which is trust and transparency. These are basically the three principles (in the slide): first, to make sure that artificial intelligence is here to help humans, to augment human intelligence, not to replace us, and that there are some concepts behind it, as human oversight, etc. The data and insights belong to their creator. Our business model is quite different from some of the platforms that professor Cortés was mentioning and put on the screen before. We don't make money out of our client's data. We provide a technology, our clients built on that technology, but we make sure that they retain their rights on their data. And that is a very different business model to the ones of the platforms in the business to consumer market. Now, the third principle: new technology including, of course, artificial intelligence systems must be transparent and explainable. I mean, we cannot be asked as humans to trust blindly a black box. So, we need transparency, okay? Now, do we ask for regulation? Yes, we are asking for regulation. Not any kind of regulation. We are asking for "precision regulation". We think that not all companies are behaving the same, that there are some companies which have created a problem and that problem needs to be addressed. But not the whole digital ecosystem needs or has to suffer that regulation. So, we are asking for accountability, we are asking for transparency and we are asking for fairness and security. In other terms, precision regulation. The same technology used by different people in different contexts should be regulated differently. That is basically the idea behind it. We have been collaborating, so, it is not self-regulation, it is co-regulation.

So, that's why we have been collaborating in the Ethics Guidelines for Trustworthy Artificial Intelligence, published in April last year by the European Commission. We welcomed the OECD principles for the development and use of artificial intelligence, for example. There is a consultation open on the White Paper on Artificial Intelligence, from the European Commission, a consultation that will be open until the end of May, mid-May, and we will be contributing with our comments. And I think that if you are here debating how the Ombudsman can contribute to the discussion, here, there is an opportunity where you can contribute with your ideas to this debate.

As you can see in the next slide, we are involved in several other initiatives. We are members of the Future of Life Institute. We are also together with other companies, like Microsoft, etc. We are also founders of the Partnership on Artificial Intelligence, because showing leadership in an industry is not only to develop the most advanced technologies. I mean, leadership is also trying to lead the industry to behave in a correct way and that is why we are trying to get involved in all these initiatives. And, of course, the Vatican Initiative, that professor Cortés was mentioning before.

Now, we are doing this, but we are also walking the talk. We have already products in the market that help other companies to be fair and to make sure that the products that they launch in the market, the artificial intelligence products, don't have any bias and are transparent. We have two of them, a couple of examples, like Artificial Intelligence Fairness 360 and Watson OpenScale. The one on the left, more thinking on the developers and data scientists, and one on the right, Watson OpenScale, for business users. I had a video. I don't know if it is possible to launch it. The one on the right, Watson OpenScale. It's a two-three minutes video, not sure if it is ready to be launched. It seems not. It's not ready, okay. You can see it

on YouTube and you can see in practical terms how a product already can help companies to test the artificial intelligence solutions that they are marketing, to make sure that its transparency helps the companies to know if there is any bias and gives transparency in case a regulator asks why a certain decision has been taken, okay?

And I am going to end with just one phrase that our current CEO, Ginni Rometty, said last year in Davos and reflects what our company thinks about artificial intelligence and it is very clear: "Companies that can't explain the decisions made by the artificial intelligence in their products, should get off the market, should not be allowed to be in the market". That is IBM's position, and this is the end of my intervention. Thank you very much!

Oliver Smith

Strategy & Ethics Director at Alpha Health

Thank you, Andreas. Thanks to the Síndic and the IOI for inviting me today. A big thanks to the translators as well, who I think have the hardest job and are living proof that artificial intelligence is not there yet.

My big message for you today is that artificial intelligence ethics is really hard. Now, you probably knew that already. So that's not new, but it's not hard because of the technology bit, although that is challenging and difficult, and of course you need to learn about it, but it's hard because of the humans bit. Because it is the humans' values that we think about and we deploy. It's humans that create a new user technology. And I think you, as "ombudspeople" (I am not quite sure if that is a word, but "ombudspeople"), men and women together, you understand people and that is what you are used to dealing with. So, I think you can be more confident in what you are doing, as you are trying to tackle the challenges of the AI. So, I will try to talk much more about a sort of practical challenges that we have been facing as we build artificial intelligence into our products and then understand how they can be ethical or not.

So, first a bit of context about Alpha and our work, then I'll talk about what we are doing in ethics and then about some of those challenges. Alpha Health is part of Telefónica, so, for those of you who aren't from Spain, Telefónica is a big Spanish telecommunications company that is also in the UK and Germany and a lot of South America. And Alpha was created in 2016 to undertake really long-term innovation, to try and address some of the really big challenges that the world faces, and to come up with solutions. And we are based here in Barcelona. For those of you who live here, that's our building. On a very windy day, like today, it sways a bit, which is my first experience of that. It is fun.

In Alpha Health our mission is to give a billion people an extra five years of health by enabling them to adopt healthier habits. And to do that we want to build the personal health assistant for all of your health needs that really understands you, your health, your behaviour and what drives it, and then advises you and supports you to build an effective path so you can be happier and healthier. So, that's what we are trying to do, and to give you a sense of how we are doing that we have four full products or prototypes in our portfolio at the moment. And we have artificial intelligence built into all of these.

We start on the left here with a product, with Evermind, which is all about selfcare, really. This is for people who have stress and helping people address that stress. It's really about trying to prevent health problems in the first place. But, of course, we cannot just focus on prevention, so we also have a couple of products, Perspectives and Mindset, which are really about treatment. These use a technique called Cognitive Behavioural Therapy, which is very, very well evidenced in the health literature, and they deliver

that to you through your mobile phone. And then finally, on the right, we would also like to predict whether people are going to be ill. So, we have a product called Foresight, which we are creating with the National Health Service in the United Kingdom, and that is using in first instance electronic health record data, so patient record data, to actually train an algorithm to try and predict crisis. Because of course, the individual would like to not go into crisis. The hospital would like you to not go into crisis, because, of course, they can give you better care and avoid all that expense and pain and burden that you feel as an individual, and of course the systems feels, and citizens feel as well. And of course, we think that the most powerful versions of these tools are going to be ones that use artificial intelligence, so they can be personalized.

The great thing about technology is that it can operate at scale, which is great for health, particularly mental health, which we focus on a lot, because there is so much unmet need. But, if technology can scale, one of its big problems is that it is actually not very personal and, of course, there is nothing better than being in front of the doctor having that face to face experience and feeling really cared for, and having really personalized advice that is right for you, and that is what we want to get our technology to do. Take Evermind, for instance, we have artificial intelligence that tries to understand what is this person's emotional state right now. Are they stressed? What can we therefore give them? What intervention could we give them? What advice could we give them that would be best for them right now? Similarly, for Perspectives and Mindset because of what I just mentioned, we are trying to actually predict whether someone is going into crisis.

For all of these, we think that they will not work unless we can be ethical, unless we can be trusted, and there are three reasons for this. Firstly credibility. When we talked to the National health service clinicians that we work with, they said: "We know you can do all of sorts of clever things with artificial intelligence, but we are not going to use this unless you can explain to us why it's made that prediction. Why it's said that I should talk to these patients, because in two weeks they should have a crisis. Because if I talk to that patient, I cannot talk to another patient. So, you need explain to me why that is happening". The second reason is related, really, it's effectiveness. For instance, Evermind supports people with their stress. One of the interventions is people actually writing down how they feel, and this is a product that we offer to corporations to help their staff deal with stress. So, if that person feels that, if they are writing a message about their boss and they don't trust that the application is not going to send that data to their boss, then of course they won't use that journal application. They won't use that journal tool and it won't be as effective. So, we need to build that trust for them to use the interventions. And finally, of course, you would expect me to say this, you would expect anyone who works in artificial intelligence to say this: Data access. Of course, if we are not trusted, we won't have access to the data we need to train algorithms, and therefore we won't be able to have that personalization that we want for it to be really effective.

So, ethics is fundamental to everything that we do. In terms of how we do this, we also have an ethics framework, we've looked at many of the frameworks that are out there, whether it's the European Union or OECD and some other companies, because Microsoft and IBM do great work here as well. And we've also looked at the Universal Declaration of Human Rights and the European Charter of Human Rights, and we took all that together and we came up with five principles in our framework, which are about happiness and health for all, so not being biased; it is about user control and autonomy; it is about understandability and explainability; it is about security and privacy, and finally, it is about governance.

Rather than go into each of those in detail, what I wanted to do for you is to really go into some of the examples that we are putting into practice and then, as I said earlier, I will go into some of the challenges, that we faced on the way. So, just three examples which I have boldly labelled "early successes".

I hate terms and conditions and I hate reading them, having to think about reading them, and then feeling guilty that I don't read them because I just pressed "yes", and just go straight through, because I just want to get to the good stuff. So, we know that terms and conditions are awful, privacy policies are awful. And we must do better, and this is part of our work, to try and have new design patterns, new ways of actually getting this information across to people. The picture I've shown here is for a sleep tool that we have in Evermind. We wanted to convey to people much more clearly what data we were collecting. So, actually we said: "well, let's just do a visual, let's have a picture of a person in bed and what data the phone is collecting". And this is just one example of others that we are trying to put into place, where we are trying to have really simple language; using visuals where we can. Also layering the consent process or the terms of conditions process. You don't have to set it all out up front. The experience most of us have is that you launch your application and it immediately asks you to approve and consent to absolutely everything and then you never see that stuff again, whereas leading companies are trying to move beyond that.

The second example I want to talk about is explainable AI, and colleagues on the panel have already talked about this. I mentioned that in the NHS this was really important to them and Ulises was absolutely right in saying earlier that deep learning is not the only form of machine learning, so, don't only focus on that. But of course, the challenge that people will face, that we were facing with the NHS clinicians is: okay, well, that does take deep learning off the table. That was okay in that instance, we could use other forms of machine learning. But we didn't want to take deep learning off the table forever, because it is a very powerful technique. There is a reason why people talk about it a lot.

And I think for you, as ombudsmen and women, there is a challenge here that you need to wrestle with, is that you can have what is called local explanations of a neural network, of a deep learning neural network. And that's great if you are the end user and you say: "Well, why did it do this? Why did this system recommend that to me in this particular instance?" However, I think you probably want a more global model that understands, that can explain, the totality of how an algorithm is working. And this doesn't exist yet. There are systems that can do the local explanation, but they can be fooled. There are some papers recently in AI, Ethics and Society, in New York, that show that actually you can make people believe that what is actually a biased algorithm, is in fact not biased. So, one of the pieces of work that we are doing is saying: well, can you create a global model for deep learning neural networks? And we can to an extent, for some neural networks. We can't yet for convoluted neural networks, but that is something that we are working on. But I think this kind of work is really important for you, as ombudsmen and women.

The final example is external audits. And, again, this has been talked about, this notion of accountability and external accountability is massively important. And as we thought about it, we recognized that, well, we can't really mark our own homework here. So, what we've done is we commissioned a local company, actually Etics Research and Consultancy, who work closely with the university Pompeu Fabra, again here in Barcelona. And they audit our work, and one of the pieces that they did recently, last year really, was on one of our prototypes. They looked at it to understand: is it biased? And they came back and said: "Well, actually no, it is not biased, you are okay". And this is really important, but of course we need to push a bit more beyond it and we will be publishing those audits soon, but we can't just, as I said, mark our own homework.

I think what you really want to hear about is the things that haven't gone so well. And there are again three I want to talk about. The first is: principles can conflict. It turns out that actually ethical principles can conflict. And I'll go back to that example I just mentioned, actually. This audit on bias. And so, great! Etics and UPF say: "So, it's not biased". But actually what they initially said, on this prototype

that we built, was: "Actually we don't know if it is biased, because you have done such a good job of not collecting any personal information, thought we actually can't really directly give you evidence or whether it is biased or not". And, of course, as soon as they said that we thought: "oh, yeah! That is an incredible obvious point you just made, why did we not think about it from the beginning?" We didn't, because we were very much focused on privacy and ensuring that we were really delivering on that, as a principle, but then, when it came to bias, it meant that: Oh, we cannot directly do anything here. Fortunately, in this case, Eticas and UPF were able to look at some indirect evidence so they looked at comments that were made on Google Playstore, they interviewed us, and they looked at all of our paperwork and they said: "well, actually, we think in that instance, this application, this algorithm is not biased against particular groups. So, you can be comforted by that, but you do need to think about your different principles upfront". And they also said: "You can't blame GDPR". Because of course whenever I talk to our lawyers they will say: "Well, of course, we have to make sure that we are not collecting any personal data here". But GDPR also talks about importance of accuracy and the importance of fairness. So, we can't just say: "Oh well, just because we follow GDPR, that's why we weren't able to give any direct evidence". So, in reality we have to deal with the fact that principles conflict and you have to think about this upfront.

The second challenge we faced recently is the kind of "all or nothing nature" of trust and ethics. We can develop our principles, we can develop the different pieces of the plan that we need to put into place and of course you can imagine a nice Gantt chart that says, okay, this week privacy, next week bias, the week after that we will do something on explainability. But, of course, you can't really just try and do one thing perfectly then get to the next one, then get to the next one, because you find that you, when you are talking to people externally, they go: "Oh, hang on you, you've not done this". So to give you more concrete example, when we were doing our work with the NHS, we focused on privacy, we focused on explainability for the doctors and nurses within the hospital, but we weren't that great on transparency. We actually didn't really talk much about it outside of the local hospital or outside the local community. We could have done a better job of actually talking to people at the more national level. And this really hit home to us recently when a lot of media stories came up to talk about it in the UK and they were saying: "Well, what are you doing here?" and, of course, not talking about something upfront makes people think that you are hiding something. So we were playing catch-up and saying "oh, no, no, it's okay. We haven't got anything to hide here. And this is what we are doing, this is who we are talking to, we are talking to the regulator, we are not going to do anything until they have approved this next phase". But it just really hit home that you have to move forwards on all aspects at once.

And then, finally, "perfect versus good enough". And I think this is a challenge that just bedevils work in AI and ethics. I talked earlier about terms and conditions and privacy policy. We have a better set of policies than we did. They are not where I want them to be. I am having an ongoing debate with one of our privacy lawyers and Alpha is supported by Telefónica, so we are lucky to be able to use their privacy lawyers, and we were having a debate about the extent to which we can simplify the language within terms and conditions and privacy policies. His argument is: "Ah!, well, of course, there is some technical language that just needs to be there". And my argument is: "Well, I still think we need to be able to explain it to people, at an eleven years-old reading age, and surely that must be possible because, believe me, if you can explain it to me and I am probably at the level of an eleven years-old for legal contracts, then we need to be able to do that for everyone". And we haven't got that solved yet. But if I waited until we've got that completely solved and we had a nice understandable legal contract, then we would be stuck with the more traditional version.

So, should we move now and have something that is imperfect, or should we wait? And there isn't really a right answer to that. And there is a second part of "perfect versus not good enough". Sometimes there is not really a right answer, so when people talk about bias and having unbiased algorithms, that is an area to debate because, arguably, for a classification system, for a recommended system, you want some bias because that is kind of the point. You want to be able to distinguish some people from another, one group from another. In reality, you need to go a bit deeper, you need to actually be able to understand when this organisation, when this company is talking about bias, what do they really mean? Which groups do they really mean? And what harms did they anticipate that might happen as a result of a bias being there or not? So, I think we need to get beyond these notions of perfection and live in the real world.

And where I want to end really is by going back to what I said before, at the beginning, about AI ethics being hard. One of the things that really annoys me, one of the aspects that really grates with me in the current dialogue about ethics is that people have this notion that you... there's two things. Let's step back a bit. So, one is when I go to more academic conferences and when I talk to my dear computer science colleagues, and often with this debate on AI ethics they are just sort of discovering ethics and they are kind of discovering that it is quite hard, and what they really want is to have an algorithm for ethics, "just give us the maths". So, they will often come across a utilitarian or consequentialist approach and, this is fantastic! So, it means there is an end result I can optimize for. I love that and I know how that works and makes perfect sense. And then someone says: "aha! But have you read Kant? And have you thought about deontology?" If you thought about a more rules-based approach. Oh, right! So that means that... no, it's all quite complicated. So, there is a part of the debate that I think is just... what computer scientists are really just beginning to understand, which is that there isn't a perfect answer in ethics. We haven't had one for thousand of years, we are not going to suddenly find one just because we have introduced artificial intelligence. And the second frustrating part of the debate at the moment is something I think companies are particularly guilty of, and which is this idea that you can... just add ethics, right? You have got this great product and we are just going to add a little sprinkling of ethics and then that will be perfect. And that is just not the way it is. Every company, every organisation has a set of values, that it is deploying and it has ethics that it is using. So, I think the question for you is, ombudsmen and women, do they know? and have they thought about it? Thank you very much for your time.

Carina Lopes

Head of the Digital Future Society Think Tank (Mobile World Capital)

Thank you. After these wonderful presentations, I would like to take us back to the core of our session which is the people. We are here to talk about citizens and their rights, and what happens when artificial intelligence systems affect their daily lives. So, I would like us to do a collective exercise, because sometimes it is really challenging to talk about artificial intelligence and ethics and values without really understanding the consequences they have on people's lives. Because when someone comes to you with a problem, it's about something really personal that really impacts their daily life.

So, today I want us to think about Salvador, and I introduce you to Salvador. He is here having some drinks, he is celebrating that he actually just retired, he is a sixty-five years old male. He has been a taxi driver for all his life and is just having some drinks to celebrate the beginning of his new life. And he is very excited about it after so many years driving a car for hours and hours a day. So, after enjoying a glass of wine with his friends he goes back to his house. He actually lives in a social housing flat. He lives on his own. He does not have family, you know, most of his

connections are local friends, some are his neighbours. And he decides to look at his computer and he looks at a six-weeks-old email that actually states you have until the end of the month to update the social services so you can start receiving your new health and retirement benefits. And this is six weeks old, and then he remembers that he also received a letter saying that the system was going, all the social services were going digital, and he also had to transfer to a digital ID. A digital ID, something very simple that is happening in so many countries and with so many challenges: how you move all the population from an analogue system to a digital system, and with them understanding what the consequences of that introduction are. And I am sure you have some of these cases at your desks waiting for resolution.

So, Salvador calls the helpline and he gets online's support. We are talking about bots that some of your companies have been designing, and they are very useful sometimes, especially if you understand what is happening and that you are not really talking with a human, but you are talking with a robot. But, you know, he goes through the system, pressing the key buttons, trying to get to a human face and he understands that actually all his benefits have been cut out, because he did not do all the new administrative work on time. So, he's been red flagged, you know. Red-flagged, what does that mean? It just means that he is been highlighted in the system as not complying with the new rules and the digital rules.

So, what happens to Salvador? When we came up with this story we actually thought initially, you know, this is a case of 2030, this is very hypothetical. But then you start reading in the news, you start looking at what is happening in Europe, with many of the automatic decision systems in Europe, and you understand it is actually not hypothetical. And what happened with Salvador is that suddenly he lost his health benefits, his retirement fund. He did complain and he did file an appeal, because he felt that he was being unfairly treated, but that took months and months to get solved, because how do you stop this machine? How do you make the system to understand that actually you are being treated in an unfair way? So, you know, he spends his savings, he loses his housing, because he doesn't have a digital ID and can also not access the healthcare services and he has chronic heart condition, so, he ends up living on the street. It is like going through a spiral of bad decisions, overlooking emails or letters. And how can you blame a person for that to happen? It can happen to any of us. So, he ends up sleeping on the streets and he feels like a victim of bureaucracy and automation.

And I don't want to say that artificial intelligence cannot present benefits for all of us, and I am sure it can, but I want to also raise awareness to how automated decisions-making systems can increment and augment inequality in the digital public services. So, before I go deeper into this, I just want to say: do we really want a digital future where we are outsourcing the difficult decisions to machines? And I am asking this question, because when it comes to public services and the digital health, the most common reasons you will find for the digitalization of welfare devices, and this comes up and up on reports about the digital transformation of public services are about reducing costs; reducing the amount of people that you need processing these cases, detecting and reducing fraud. So, basically, we are highlighting, and we are starting the problem from the wrong perspective. We are looking at vulnerable collectives, we are talking about the people that are in most need of these services and assuming that the use they are doing of these services is not the correct one while we are ignoring the real impact that a bad decision can have on these people's lives, because so many times these public services are the safety net of these people and, when that safety net is not there, it's very hard to come out of it. And, you know, we always talk in the digital sector about the digital divide between different sectors of society and collectives of society. But I would like to tell you about the knowledge gap, because I think that this is where the challenge really is for all of us. So, you know, for all of us that

work in companies, or in the private sector, or citizens. How do we feel about this knowledge gap in understanding what these systems are doing? How do we know when we are red-flagged, and most of us at some stage of our lives will be red flagged by one of these systems. It can be a bank; it can be a public service. You know, there are many ways and many moments of our lives when we are constantly being evaluated and risk assessed. So, my question is: when we are red flagged, what happens? And how do we avoid these situations?

So, because today we are where we are, I want to go into more detail into three examples of automated decision-making systems that are being implemented in Europe. Because I think this is when we start to really understand the challenges that are at the table and understand what really means this knowledge gap. And the challenge that we have, for example, as a public service and the provider of a public service. So, the first example I bring today is the Gladsaxe example. This is a system being implemented in Denmark. It all started very well, with very good intentions. So, we are talking about a system where three small cities outside Copenhagen we asked to do a project where they created a risk assessment system, a points-based system that anticipates the risk that children face. The idea is that social services can identify children before they actually have the needs to receive their services. So, how does this points-based system work? So, basically, you know, it is almost like, in a very cruel way, like a shaming system. You get points for not behaving as the society thinks that you should behave. So, you get divorced, or you are divorced, you get one thousand points. You miss a paediatrician appointment for your child, you have another thousand points. You have a mental health issue; you have three thousand points. So, we are talking about the system that tries to red flag and identify children that might in a near future need the service of social services in Denmark. So, actually these children can be followed in a closer way. But what we are actually also doing is tagging and controlling the life of these families, and families that in so many ways are in already in a hard position. Think about the single mom that, you know, is caring for their children but also has to fit one or two jobs and she is being constantly penalised, because of not fitting into this system, this points-based system. And I am sure when it started it was all about good intentions. But it is only when we start looking at the impact that this has in people's lives that we really understand how good intentions can turn, can go really badly very quickly. So, this one is related with health, and this is an interesting one.

So, RiskEr is a system that is being developed in Italy and basically it's a statistical procedure that combines demographic and health variables in order to understand and predict the risk of hospitalization of certain people. So far so good, very interesting system, it's about making your health system more efficient, it's being piloted by twenty-five health centres in Italy. The challenge starts when you read the small letters. So, they are collecting five hundred elements, variables, about health and social context of these people. So, we are talking about five hundred variables. As a public servant that's already very hard to control. How can you talk about bias and if you are being fair when you are combining five hundred indicators? But the challenge here is that actually this system was developed in collaboration with a hospital in the US, where we had just a few weeks ago the news about the system that has intentionally no bias in terms of race. So, there were no race indicators in that algorithm. But what the system was doing in US was evaluating the type of care that each person should have based on their chronic conditions, and it ended up, you know, through research at university, it ended up finding that actually black people were indicated as in need of less care, less services by hospitals. But, you know, the system actually was designed to not have that bias, so how did that happen? Because historically black people do spend less, they go less to the doctor or they spend less money in doctors, or they also have, because it's the US system, they have lower insurance coverage. So, historically they actually use the public health services in a lower rate than their white counterparts. So, again we are talking

about the system that was designed to not include race data, race indicators, but because of the socio-economic conditions they were treating unfairly this collective and this is one of the challenges that these systems present. How do you notice, how do you flag that these systems are to treat these collectives in a fair way until actually they are already into action? Because the level of complexity is very very high.

So, the last one I want to present, and I am sure most of you will know about it, is the system that has been just stopped by the Netherlands Court, and this is a risk indication system. So, it talks about a welfare service system that has been indicated by the Dutch Court as violating human rights. And why did they get to that conclusion? Because this system was mostly being deployed at low-income neighbourhoods and they were collecting data across government. So, this was a system to detect fraud. So, the potential for a citizen to commit fraud in terms of employment law. So, we are talking about a very specific issue, probably a very low percentage of the population actually commits fraud in terms of employment. But then what we have is that data was being collected and data that was collected by the government was suddenly put together, mashed up, and we had a specific collective of population being under continuous surveillance. So, this is a really big challenge to understand what is happening because it is not just about what data is doing, using datasets that were not designed or were not built for this specific objective. So, it is very hard to assure actually the quality of that data.

So, you know, I think it is also very important to look at the timing, and today I am very happy that you organized this event in Barcelona for different reasons. We have been talking about technology and human centred technology in Barcelona for a while, but also there are different momentums at international level that I think helps putting this issue of artificial intelligence and human rights on everyone's agenda, and I hope that all of us, when we go back to our institutions and cities, we can put this issue on everyone's agenda.

So, this year is also a year of opportunities because, just back in October Philip Alston, that also visit Spain very recently, published a report exactly about these issues in the public sector, and how cost-effective and fraud detecting systems deployed across welfare services were constantly punishing the same people that were in most need of these services. It is like punishing twice these collectives, not only they are in poverty, we punish them again because we think that they need to be under control. Last week we had the publication of the EU Digital Strategy that sets many of the agendas of what the next four to five years will be. And these days, technology has to work for the people. My question is: what does that mean in our day-to-day jobs? How do we put technology at the service of the people, and we make sure that what we develop is not innovation for the sake of innovation, but it is technology at the service of the people, but also assuring that it is not always for the benefit of the same? You know, it is our duty to make sure that when we talk about technology at the service of people that actually is for everyone regardless of their context, of their social status or migrant condition.

But also, and I will tell you a bit more about our work at Digital Future Society, because this year we are celebrating twenty-five years of the Beijing Declaration. That talks about women's rights and, you know, like in the Declaration of Human Rights, this agenda does not really talk about technology. But what we are finding in our research is that these same systems that we are talking about today also discriminate in an exponential way women. Because historically the social economic condition of women has always been in a weaker position. We know that when it comes to pensions, when it comes to access to technology, to knowledge, to education. So, we really have to understand what are the technologies that we are developing and who are being targeted and affected by this type of design of technologies.

So, what I want you and I would like you to take home and also the conversations today, is really thinking about, you know, artificial intelligence beyond the technology and data collection. How do we assure these data quality? How do we ensure that we understand these biases when they are not so visible? We only understand them once they are being implemented. But also, I really think how this affects certain collectives: women, children, migrants, you always end up having the same people. Maybe it is time to bring them into these conversations, to listen to their experience of feeling continuously red-flagged by the same systems that should support them.

So, you know, there are different things that we can do from our institutions. But I think one is regulation, and it is very easy to talk about regulation, sometimes it is harder to implement. But I think the oversight, the relevance of oversight bodies, but also having specific process of audit. As a citizen, I want to know where I can go when I feel that I am being unfairly treated by a digital system and that someone has the power to stop that system doing that work quickly. We were talking that in the Netherlands social activists have been working on these issues for nearly a decade. How many lives have been destroyed by the system that was assessing this people for a decade? How many lives have to be put on through there, you know, the basic of a safety network before we can actually take action? You know, that is the question I always make: how much does it take for us to actually rethink how we design these systems and also the historical data? And also, one of the challenges, I think the biggest challenge we have at our hands is how we build capacity within the public sector, because the knowledge gap is very important and very wide. But also, we are talking about procurement. How do you do procurement to ensure that these situations do not happen and once these systems come into place, what are the tools that the public sector has to stop these systems? What are the legal aspects that actually provide safety to the citizens when the governments change suddenly and the data that I provided for one specific use is actually being reused for another objective? And lastly, how do we empower citizens to really understand what is happening here and how these systems become traceable?

But also, to remember that, for every digital transformation, for every digital transition, we should have an analogue, a human face there, because at the end of the day, and it is so good to say that artificial intelligence is designed by humans and accounts for human bias, but that is not enough. We need to make sure that we always have a human face on the other side and someone with power to stop that unfair decision.

So, you know, just to give you a bit of context, this is part of the some of the work we have been doing at the Digital Future Society, and our objective is to really set the agenda for the next five years for the different actors of the ecosystems: public authorities, but beyond this, the industry and so on. And last year we started publishing a report that I think is at the entrance, some examples of automated decision-making systems in the public sector and the challenges that emerging technologies offer to the public sector, but also opportunities. And this year we actually will be producing more work on this but looking at the impact that these technologies have also on gender issues. So, I invite you all of you to look at more of our work, you can find us at digitalfuturesociety.com. You are also welcomed to come and talk to us, and also we have Carlos Grau, the president of our foundation, here today with us. So, I hope I have given you some food for thought and to take home. Thank you!

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Monday, 2 March 2020

From analogic to digital society

Chairperson: Genís Roca, Internet Expert

José María Lassalle, Director of ESADE Technological Humanism Forum and Professor of Philosophy in Law at ICADE

Daniel Innerarity, Professor of Political Philosophy at the University of the Basque Country; Ikerbasque Foundation for Science; European University Institute of Florence

Renata Ávila, Executive director of Fundación Ciudadano Inteligente

David Cabo, Director at Fundación Civio

Genís Roca

Internet Expert

It is a pleasure for me to coordinate this round table, it is a luxury. The four of you are Spanish speakers, we have participants from Bilbao, Santander, Guatemala and Zaragoza, although life has brought you to live who knows where.

It's an interesting panel, classifying and labelling people is a bad habit, but here we could say that there are two philosophers and two activists, although I insist that these labels are a bit unfair.

Coming from the previous session, we have seen the academic and business world and the world of society discussing what artificial intelligence (AI) is. We could agree that AI is very good, that it is going to save the world, and that it needs ethics and that it is all very good. But the truth is that it seems difficult for AI to move in the right direction if we do not set rules. Today's speaker is an ombudsman who is asking for a legal framework to be able to set the rules. I am sure that many of the companies and actors that are going to be involved in all of this would appreciate it if there were rules to make it clear what is inside and what is outside the law and not have these ethical debates. In the end, the ethical debate is the absence of rules, but also, and I think Ulises has been quite clear, you trust that the government can guarantee that the rules are complied with, right?

We are dealing with a type of technology where the control mechanisms, of how all this is being developed, are quite difficult. David will be able to tell us how difficult it is to control what goes on behind an algorithm. The rules that are making decisions that can affect our lives are not clear and it is much less clear how to supervise and audit these rules, which means that, as much as we may want a legal framework, we may have the problem that we are putting freedoms at stake and, although there is a rule, the difficulty of complying with a rule is obvious. This could lead us to mechanisms where if we introduce self-control, if we introduce citizen surveillance, if we introduce a regulatory framework. Yes, we do all three things, we introduce the regulatory framework, we introduce self-control, and we introduce citizen surveillance, but where are the roles? And to what extent is this redefining democracy or the rules of the game of democracy? And to what extent will the mental frameworks we have had until now of what is a rule and what is compliance with the rule be sufficient to live in this world?

Someone once said that AI at the moment is very artificial and unintelligent, but there is no doubt that it is progressing well, and that what is embryonic today is going to have a strong and powerful impact on our daily lives.

We want to discuss four of these things with you. You have an initial time to set out your approach. Each of you brings a backpack and your point of view, and that is why

you are here, because we are interested in how all this complements each other and, as far as possible, to see how you cross paths, if there is some debate and if the public has any questions or any consideration of what you might say now.

José María Lassalle

Artificial Intelligence Professor at the Technical University of Catalonia (UPC),
Barcelona Supercomputing Center (BSC)

I am also a person involved in activism, and I believe that from politics, although I have already left politics, activism is still possible.

During my time as Secretary of State for the Digital Agenda, we set up the digital rights commission. The commission's work resulted in what has been the first statement contained in the Data Protection Act of 2019 and in the commitment, at least in the explanatory statement, to approach a digital constitution for our country.

Moreover, the commission was made up of experts in digital rights, but also activists and people from the private sector, and it was coordinated by a former socialist minister, Tomás de la Cuadra. Thus, I believe, we guarantee a relative independence and autonomy in its management.

Then, we implemented something that is now in the imaginary of public debates, which is the Digital Gender Gap Commission. When we launched it in January 2018, many people looked at us as if to say: "What are we talking about? And the truth is that it is one of the things I am most proud of, because it was difficult to make it progress, and the government that came after us took it on and has continued the work.

We wanted to set up the AI commission, unfortunately that was not possible due to the change of government and now it has been taken up again by a State Secretariat for AI. So, we therefore activated the embryo of what later became the Digital Forum Society, with the reflection we proposed on the Digital Davos and the budgetary contribution that made it possible to launch this marvellous initiative, which has generated what for me was a fundamental factor. This essential fact, along with rights and the search for equity, was to face a space for critical reflection that would make feasible what I believe would be the most profound debate that we should be considering at this time. We are not talking about ethics, we are talking about regulation, and we are talking about regulation for a simple reason, because when the other great technological revolution took place, which was the industrial revolution, it unleashed a potential for change that radically transformed European societies. It buried the societies of the ancient regime, changed the economic model, changed mentalities, but for many years it generated a democratic deficit, situations of extraordinary injustice and inequality which took democratic civilisation and European societies more than 100 years to correct.

In other words, we cannot understand the social and democratic rule of law as we understand it today. It should be seen as the consequence of an attempt to alleviate the deficit of equity generated by the industrial revolution. A revolution that revolved around two fundamental vectors: the vector of speed and the vector of capacity, which radically transformed the world and changed mentalities and led to the revolution of 1848, to the Soviet revolution, to a whole series of social awareness, of struggles that have generated a feed-back that we cannot throw away now, when we are facing another moment as important or much more important than the industrial revolution, which was an extraordinary change, because it was a change associated with progress, therefore associated with a fact. As Ortega said: man is a being with an innate epistemological deficit that needs to generate, through sources of information, instruments with which to cover his imperfection and thus arises the power of technology from ancient times to the present day. From *homo habilis* to the present day.

And that is the generation of a series of layers of institutional knowledge that have brought us this far, which we cannot renounce because technology is an instrument of progress, an instrument of change and an instrument of progress at all levels. But what has happened to us from probably the first decade of the 20th century to today is that, when there has been an explosion that has been disruptive of the capacity to generate speed and content that could circulate through the networks that have allowed the most plastic technological revolution, which has been the introduction of mobile telephony, that has generated an explosion of inequality, a situation of lack of democratic regulation, a critical deficit in terms of democracy really reflecting on what was happening. And ten years later, in this acceleration of time that we lived through, we have discovered that our world has changed.

Human beings are experiencing an anthropological revolution that is altering the key to the way they experience themselves. Human beings for the first time are making bodies withdraw from reality, that our identity is beginning to see how the body, which has been the base on which our identity has been built, is beginning to withdraw in its capacity to manage the information through which we construct our own life. When we articulate a new experience of identity, through digital identity, when we live with others through mobile telephony, when we live trapped in screens, we realise that we have lost our way, that we have lost not only ethics, we are losing the human being and we are losing the legal dimensions that have protected the human being as we have understood it up to now.

But there was not enough for us just to transplant the legal mentality of modernity to our technological post-modernity. It is no longer enough to imagine, in legal terms, how we are going to manage ourselves in the 21st century under the new experiences that we are going to face, and which are linked to a new sensitivity, to a new way of relating to each other, to a new anthropological experience of the human being that needs to be thought of in legal terms. It is not enough just to transplant the fundamental rights, conceived in the analogue society, to the digital sphere.

Law has the task of legally imagining a new generation of fundamental rights. Just as there was a first generation of rights linked to freedom, a second generation of rights linked to the democratic experience, a third generation of rights linked to how marginalities or negative externalities generated by democratic society and the market were mitigated. We must face a new generation linked to the negative externalities that the digital revolution is generating.

And for that purpose, we cannot think in analogue terms, the distinction between privacy and advertising is no longer useful. There is an intimacy, we are not only users and consumers of content. We are flawed as people when we work and operate as such on the networks, let alone when we are citizens. As citizens we lost our identity because there is no regulatory framework. Is there a regulatory framework that defines how to make an algorithm? Is there a regulatory framework that tells us who owns the data? What rights do we have when it comes to managing our data? For the first time capitalism faces a theory of the market around the platform economy and does not have a theory of ownership.

The *commodities* of the new platform economy, which are data, do not have owners. Yes, they have an owner: the one who introduces the algorithm and then obtains the externalities that add value and allow him or her to monetise his/her business model.

That is the new design we are building, and in that design, there has been no democratic reflection. That is the great shortcoming we have now and to address it we not only need rights, but we also need to change and evolve the institutions.

I am very obsessed with a topic. When we built a democratic society, which is the product of generations, revolutions, and very profound social changes. Those changes were faced because modernity understood that we wanted rights because society had an idea of progress and, therefore, there was a need for action to guide human efforts to change. Social change was identified as a possible reality through the law and, therefore, modernity thought about how to guide the efforts of human, individual and collective action.

Now we mainly need –and not much is said about this– the freedom to understand, we need not only the freedom to act, but also the freedom to understand what is going on. After all, what meaning do we want to give to the machines? What meaning do we want to give to everything that is happening to us?

There was a clear sense of technology being linked to progress and social change. Nowadays, we see that technology is altering the gravitational axes of human experience, not only on a collective level, but also on an individual level. And we need the capacity to think about this.

The great freedom now is to recover the time to be able to think, among other things, that our corporeality is in disuse for example. And we built up our whole emotional side, a very important part of our legal categories, for instance the fight against torture from the experience of knowing what a suffering body was.

This approach which has accompanied the development of law and political theory until today is being questioned, it is getting out of hand. And we don't have time to think about it, because, among other things, we are trapped by a daily acceleration where we are imbued. Moreover, we are trapped by an obsession with maximising the profitability of mental efforts, as if we were transferring the profit and loss account of the large technological corporations that seek to monetise exhaustively all the investments they are making, also in ethics now. Investments that they take to the human domain and to the domain of our everyday life, and we don't even have time to think.

I believe that this is the great revolutionary challenge that we should be able to demand from democratic societies, time to think about what is happening to us and to make sense of what we are doing.

Daniel Innerarity

Professor of Political Philosophy at the University of the Basque Country;
Ikerbasque Foundation for Science; European University Institute of Florence

I am a philosopher; sorry, nobody is perfect. And philosophers, when we do things right, which is not always, what we manage to do is to show a special ability to detect problems, even where people do not see them, even where you would say there aren't. And that has to do with the fact that we have a kind of addiction, like people who are addicted to drugs or alcohol. And that has to do with the fact that we have a kind of addiction, like people who are addicted to drugs or alcohol, we have an addiction to problems, and we are also used to being asked, what is a guy like you doing in a place like this? We are always being asked those kinds of questions and we go through life with our trouble detector. Wherever there is a disruption, a change, a mismatch, and everything is turned upside down, we rub our hands together because there is food for thought.

I have the impression that there is a lack of reflection on the political dimension of artificial intelligence and a surplus of ethical codes. Last year I had already collected 84. I do not know what the exact figure will be by now.

We have made great progress as a society, and in Europe especially, in what we could call private law thinking on these issues. We have made the GDPR, the system for the protection of property and privacy, which is a great advance from the perspective of habeas corpus, but we have made very little from the perspective of the common good.

Democracy is popular decision, it is people's self-government, but first and foremost it is a decision. What is going to happen when more and more things are automated? What really fascinates me conceptually is how we reconcile democracy, (*demos*) and (*cratos*), government of the people, decision of the people, with more and more automated systems.

In this respect, the ideological field is divided into two broad types of analysis; when we have no idea about artificial intelligence, the field is divided between the optimists and the pessimists. The optimists think that automation is fantastic because the old ideologized politics that we have had until now will be replaced by systems that will decide with much greater precision. The pessimists see something terrible in this, the dystopia that we are going to stop deciding, we are going to have to close down parliaments and intelligent systems are going to decide for us. And then comes the imagery of robots replacing us. Both optimists and pessimists think that politics is something that can be overcome, that it can be replaced by another instance of whatever kind, and they differ only in whether they consider it good or bad news.

My approach is that politics cannot be replaced by anything else at all, that even in those technologies that present themselves as lacking ideology there is always a politics.

What happens when we have an automated public sphere? What happens to democracy? How can we think about its compatibility with democracy? What sense can democratic politics make in a world in which more and more things are automated? A world that is increasingly "dehumanised", not in the ethical sense, but in which there are no humans in the ever more decision-making systems. Nowadays, there is no one in the lifts helping you up (elevator attendants), there are robotic crop fields, geostationary satellites, automated doors, in other words, a large number of automated processes, to the point that someone has said that we could end up with a world that would function without us. Or, as Minski said, where we would be very lucky if machines adopted us as pets.

And at the same time, it is said, this is full of black boxes, the society of black boxes, that society where there are decision-making devices whose logic we do not understand. Here, the magic solution has been transparency, then a little more nuanced with the idea of explainability, as we were told at the previous round table. This is very important and at the same time very complicated because, first of all, there are different types of intransparency, of opacity. There is an opacity that must do simply with the fact that in order to protect property for state secrets or patents or whatever, it is necessary to set up a system to mask certain things. There is another opacity that has to do with the complexity of the machine. No matter how much you explain an algorithm to us, we are not going to understand it. But there is a third opacity, which is even more complex, which is the opacity that has to do with the devices of the post-programming era in which we live, an era in which machines obey a logic that is not even in continuity with the will of the person who designed it.

What will democracy be like in that time when for the first time in human history we are able to design machines that we do not understand? Because this is what artificial intelligence means. How do we translate this in terms of explainability? There is a problem with explainability, which we can certainly talk about, but I think article 22 of the GDPR is very unrealistic (that we cannot be conditioned by automatic processes that are not somehow understandable to us), and it also has a private view of explainability, i.e., it says that explainability is something that

can be done individually, that can be managed like the one who signs an opt-in clause. The most eloquent example is when, before the economic crisis, you asked for a loan, went to bank and signed an agreement that no one read. Then Europe tightened the conditions so that the bank had to guarantee that it had been understood. When we demand individual intelligibility, we are applying schemes from a simple world that individuals could understand to a complex world that can only be understood collectively.

It is not about forcing people to finally read the credit conditions, which would be the private solution, but about creating the conditions for people to take on credit in a responsible way, so that there is no excessive incentive to take excessive risks. Therefore, counter-cyclical regulations should be in place, i.e., the important thing is the context, the regulatory framework.

And I would like to finish with a couple of things about what the future of democracy might look like in this context.

First of all, I think we are facing a risk that needs to be considered carefully, which is what we could call the risk of anti-democratic populism. What is happening in China, not the virus, but the fact that it is becoming more and more tempting to resort to exact machines as opposed to ideologized politics.

AI has many advantages, for example, in terms of the possibility of making public policies that are much more in line with reality and then, of course, measuring the impact of these public policies in practice. One could make an algorithmically tuned policy, if I may say so, taking advantage of precisely these different levels of detail (granularity) of AI mechanisms.

Therefore, what is the problem with this? Why algorithms and data have biases? Because there are evil people or because they have designed them wrongly? No, because life as such has biases and if AI is designed on the assumption that our future behaviour will be a continuation of our past behaviour, there is no way to give political track to experiences of innovation, change or rupture. How do you break with data that is sexist? Because our behaviour and social structure, which the algorithms reproduce, are also sexist. Thank you!

Renata Ávila

Executive director of *Fundación Ciudadano Inteligente*

Thank you very much. Although I prepared the speech in Spanish, I am going to try to do it in English.

Artificial intelligence and human rights. Let me first give you a brief overview of what is IBM. Maybe you have in mind a picture of IBM that is not very current since the company is in constant change. Probably, if we would have had this conference five years ago, or ten years ago, IBM would have been much different to what it is today, related to artificial intelligence. IBM was founded 109 years ago, and I want to put this as a first slide, because we are not a company founded ten or twenty- years ago. We are not a company founded sixteen years ago (like some others...). We have a much longer history, we made our mistakes, we learned from those mistakes, and we know well what it takes to do things right.

For twenty-seven years in a row, we have been leading the patents ranking in the US. We are still the number one in the Interbrand ranking in terms of business services. We are, essentially, a business-to-business company. We are not in the consumer market, although many of our clients are; this is, we sell to other companies. And today we are leading in several areas that, as I said before, we were not some years before, like cloud computing, artificial intelligence, internet

of things, business analytics. We have been also in Europe for more than a hundred years. We have offices and centres all around Europe. We have a data centre here, in Barcelona. We have the European Headquarters in Madrid. We run from Madrid, not just the Spanish operations, but also the Portuguese, Greek and the Israeli operations. So, we know what the European values are. We are European and we have been in Europe for more than a hundred years and this is another key message that I want to give you. As I was saying, we are in Spain, we have not been a hundred years, but almost, both in Madrid and Barcelona around ninety-four years and we have centres of international responsibilities in Madrid and in Barcelona; offices, in Seville, Valencia, Bilbao, la Coruña, etc, etc.

So, this is about IBM, and I was saying before, our centre today, our focus, in terms of business is in big data and analytics, cloud computing and, of course, artificial intelligence, which is what we are discussing here today. Now, we are talking about artificial intelligence, but many times we are referring to things that--and probably Professor Cortés knows much better than me--, that are not the same. I mean all the terms, the different concepts that we are using for artificial intelligence when we are talking about different things. And here (in this slide) you have some of these concepts: machine learning, robotics. We prefer in IBM to talk about the last terms you have in the slide: augmented intelligence and cognitive computing. Why? Because these terms put the emphasis on artificial intelligence as a technology that is here to help humans, as opposed to the term "artificial", that it seems that it is something opposed to the "natural", like if there was a natural intelligence and an artificial intelligence, and these semantics are probably part of the problem when we approach artificial intelligence. That we are thinking that we are choosing something that is opposed to our natural intelligence.

Now, as Professor Cortés said before, the term, the concept of artificial intelligence was first used in 1956 during a conference, I believe the person with the red shirt in the slide, John McCarthy, if I remember correctly, was the one that came up with this term. There were two IBMers at that time in that conference. Two people that worked for IBM. The one on the left, Trenchard More is one of them. This is a picture taken a few years ago to commemorate this summer course in Dartmouth College in 1956. So, IBM has been involved in artificial intelligence since the beginning. I mean, since the very first beginning in Dartmouth College. We know what we are talking about. You will remember some of the most iconic steps in artificial intelligence, like in 1997 when an IBM Deep Blue system defeated Kasparov. And this was a very iconic moment, because chess is believed to be one of the games that humans, we thought play better than machines. 2011, Jeopardy, a very famous program in the US, that has been broadcasted for more than thirty years, also an IBM machine, Watson, defeated the two most successful contestants at that time. Today we are involved in what we call Project Debater, which involves teaching machines to debate humans on complex issues. This is where we are focusing today.

Now, in terms of what we are doing in the business? Just a few examples from the last three months: We have more than 20,000 projects on artificial intelligence in the world right now, in IBM. For example, we have an assistant for Orange in Spain helping clients to get answers through WhatsApp. A team from Barcelona, just a couple of months ago, won a contest, an international contest for IBM, called Call for Code, that is helping firemen to be safer when they go to put off a fire. We are helping Cerealto Siro Foods to put news products in the market and analysing consumer trends and putting in the market products that take into account these trends. And, of course, we are in health with our system Watson Health trying to help a better diagnosis for doctors. Helping doctors, not replacing doctors, to do their job.

Artificial intelligence is already used in our everyday lives. We use it when we pay

with our credit card; when we use our system in our car to drive from point A to point B; in our emails, the spam filters use artificial intelligence for that; when we use our phone, Siri for example, or when we do a search in Google for example. So, it is present already. We are not talking about the future; artificial intelligence is already here, and we use it every day. Now, a word that has been already mentioned here: "trust". That it is the key for the adoption of artificial intelligence. If we don't trust the systems, this technology won't evolve. I mean, we trust our car because we know the car breaks when you press the pedal, and we go into an elevator because we know that if we push number five, the elevator stops at the fifth level, or we go to an automated teller-machine and get money from the teller machine, because we know that if we ask for a hundred euros, the machine is going to give us a hundred euros and not fifty. So, trust is behind all this technology. Now, some of the questions that we ask ourselves when we are talking about artificial intelligence: well, first, is it fair? I mean, is there any bias? Is it easy to understand? Do we understand how this system works? Has anybody tampered with it? I want to make sure that this is a robust system, and no one has done anything illegal or that put me in danger. And is it accountable? I mean, if something goes wrong, there is someone behind that is going to be responsible for it, right?

Now, yes, we also have our principles. Not sure if 2016 or before, I won't argue that we have our principles, and these principles are around what has been always key for IBM, which is trust and transparency. These are basically the three principles (in the slide): first, to make sure that artificial intelligence is here to help humans, to augment human intelligence, not to replace us, and that there are some concepts behind it, as human oversight, etc. The data and insights belong to their creator. Our business model is quite different from some of the platforms that Professor Cortés was mentioning and put on the screen before. We don't make money out of our client's data. We provide a technology, our clients built on that technology, but we make sure that they retain their rights on their data. And that is a very different business model to the ones of the platforms in the business to consumer market.

Now, the third principle: new technology including, of course, artificial intelligence systems must be transparent and explainable. I mean, we cannot be asked as humans to trust blindly a black box. So, we need transparency, okay? Now, do we ask for regulation? Yes, we are asking for regulation. Not any kind of regulation. We are asking for precision regulation. We think that not all companies are behaving the same, that there are some companies which have created a problem and that problem needs to be addressed. But not the whole digital ecosystem needs or has to suffer that regulation. So, we are asking for accountability, we are asking for transparency, and we are asking for fairness and security. In other terms, precision regulation. The same technology used by different people in different contexts should be regulated differently. That is basically the idea behind it. We have been collaborating, so, it is not self-regulation, it is co-regulation.

So, that's why we have been collaborating in the Ethics Guidelines for Trustworthy Artificial Intelligence, published in April last year by the European Commission. We welcomed the OECD principles for the development and use of artificial intelligence, for example. There is a consultation open on the White Paper on Artificial Intelligence, from the European Commission, a consultation that will be open until the end of May, mid-May, and we will be contributing with our comments. And I think that if you are here debating how the Ombudsman can contribute to the discussion, here, there is an opportunity where you can contribute with your ideas to this debate.

As you can see in the next slide, we are involved in several other initiatives. We are members of the Future of Life Institute. We are also together with other companies, like Microsoft, etc. We are also founders of the Partnership on Artificial

Intelligence, because showing leadership in an industry is not only to develop the most advanced technologies. I mean, leadership is also trying to lead the industry to behave in a correct way and that is why we are trying to get involved in all these initiatives. And, of course, the Vatican Initiative, that Professor Cortés was mentioning before.

Now, we are doing this, but we are also walking the talk. We have already products in the market that help other companies to be fair and to make sure that the products that they launch in the market, the artificial intelligence products, don't have any bias and are transparent. We have two of them, a couple of examples, like Artificial Intelligence Fairness 360 and Watson OpenScale. The one on the left, more thinking on the developers and data scientists, and one on the right, Watson OpenScale, for business users. I had a video. I don't know if it is possible to launch it. The one on the right, Watson OpenScale. It's a two-three minutes video, not sure if it is ready to be launched. It seems not. It's not ready, okay. You can see it on YouTube and you can see in practical terms how a product already can help companies to test the artificial intelligence solutions that they are marketing, to make sure that its transparency helps the companies to know if there is any bias and gives transparency in case a regulator asks why a certain decision has been taken, okay?

And I am going to end with just one phrase that our current CEO, Ginni Rometty, said last year in Davos and reflects what our company thinks about artificial intelligence and it is very clear: "Companies that can't explain the decisions made by the artificial intelligence in their products, should get off the market, should not be allowed to be in the market". That is IBM's position, and this is the end of my intervention. Thank you very much!

David Cabo

Director at Fundación Civio

Thank you very much for the fantastic introduction. Yes, I am the co-director at Fundación Civio. Our mission is basically to increase the transparency and accountability of Spanish public administrations. We have been carrying out this work for eight years. We do information requests, data journalism and, if necessary, litigation. We fight against the Government and sometimes we win, like recently in the Supreme Court we won a case against the Court of Auditors.

The first thing I want to say is that I am a computer engineer, and I am fully aware of the potential, and all the positive stories that are told about technology. I come from the technological side and when I co-founded Civio I did it because I believed that technology could be used to increase transparency in administrations. And I still believe that there are tools, for example anti-corruption tools, that can improve the public sector, which is very important, because the public sector belongs to all of us, we can make the public sector work better. But, at the same time, I also went from having a somewhat naïve vision of all this, a very utopian vision that information would make us free, to having a slightly more critical vision and realising to what extent not only is technology not neutral, but it is used intentionally by those who have the power, who are the ones who have the resources to create technology, design technology and apply it in one way or another.

So, it is very important for civil society or ombudsmen to be in the front line, a little bit, of the administration to fight against the excesses of power, such as attempts to discriminate or to maintain discrimination that may exist in society.

So, from the point of view of AI and technology, there is a concept that is always used, which is that of the digital welfare state, which, for example, the United Nations rapporteur on extreme poverty, mentioned in a previous talk. When we hear the word

digital, we may think of services that are more accessible, closer to the citizen, that everything is easier, that the administration is at the service of the citizen. But often the prevailing discourse, as Carina said in a previous talk, is a bit like fighting against the fraud of those who ask for a subsidy, isn't it?

This speech on efficiency and control of those who ask for unemployment benefits is causing, as the United Nations rapporteur denounced, an increasing number of areas within the administration to be opaque, which are not governed by human rights. That is somewhat what has happened in the case of the Netherlands, where the Dutch Government has created this mega system of control, crossing many sources of data, and applies it only to a specific type of aid or social benefits, which is that used by poor people or certain poor neighbourhoods.

Therefore, when you talk about the bias of the algorithm, sometimes it is not just the bias of the algorithm itself, but the bias in the application of this technology. The fact that you create this monitoring tool for something that affects poor citizens, and you deploy it in poor cities or poor neighbourhoods.

Within this shift from transformation to the digital welfare state, we are concerned about many things, but let us say three of them to be very brief:

One is the issue of the digital gap: there are people who do not have access to the Internet or, if they do, they are faced with extremely complex processes, as I will explain later with the issue of the social tariff. Complex both conceptually, in terms of notifications, multiple steps, forms, and technologically. If you have tried to use digital certificates, you must deal with Java versions, browsers.... We have been complaining for years that it is extremely difficult to make a request for information because of the technological barriers that exist, and all these are issues that add up, and that I am able to overcome because I have spent years struggling with these systems, but there are many people, especially vulnerable people, people who need this aid, who cannot deal with these barriers.

Another thing that concerns us is the systematisation of bias. The typical case that is often put forward is that of predictive policing but applied to other cases we have seen here. That is, you train a system with historical data that shows a bias, a racism, for example, of the police, which may be directed especially against poor people or black people. You train a system and use that output to say: This is neutral because I do not say so, the technology says so. There is a phrase that I like a lot, which is "machine learning is like money laundering for bias", because it allows you to say that you are no longer the one who is racist, that there is an objective system and the numbers do not lie, numbers say there is more crime in this neighbourhood.

This is leading to the elimination of the ultimate responsible for political decisions, which is a fundamental issue for the quality of democracy. In other words, we are seeing in many countries that when a system is implemented, for example, one that regulates public aids, in practice, what is being done is to reduce this aid. But there is no open political decision that can be discussed, but what is done is to say: No, this is a technical solution that is going to manage aid. But, in practice, it has a series of criteria that are not discussed, which means that many people are denied such aid.

And finally, another thing that concerns us is that new areas of opacity are being created, new excuses are being created. One of these excuses is that this is very complicated and that this is a discussion that belongs to the technicians or that belongs to the engineers, and this is the specific case we are working on, which I am going to explain very briefly.

One of the things we do, or that Eva Belmonte, journalist and co-director of Civio, does, is to read the BOE (Spanish Official Journal) every morning and see if there is relevant news for citizens, because part of our mission is to explain the decisions of

the administrations. In 2017 we saw that the system of the social tariff, the system that gives aid to vulnerable people who cannot pay their electricity bill, was going to be changed. We started to report that millions of people were being left out, that they were not applying for the new social tariff and that they were going to be left without aid when the previous system expired. We contacted the CNMC to create an online calculator to facilitate the whole process. It is an extremely complex process. To give you an idea of the complexity, the lawyers and the people in charge of the CNMC, people who are the regulators of the electricity sector, had a lot of doubts about the process to follow and we spent weeks fighting and asking questions to the Ministry, to understand exactly the cases that were not clear, where if you were a large family and at the same time you were a victim of terrorism, etc. It is a system that not even the CNMC lawyers can understand.

As a result of this, hundreds of thousands of people used the calculator. We received hundreds of calls from people who had problems, there was brutal misinformation, there were people who also complained that our calculator said yes, but then the electricity company that managed the aid had told them no, that they did not know what to do to defend themselves, there was an enormous lack of defence.

And, because of all this, we decided to make a request for information to try to monitor the system or to audit the system externally, which we believe is essential. And, very briefly, we asked for the technical specification of the application, the use cases, and the source code. The Ministry said no. We appealed, we went to the Transparency Council, we got the technical specification and the use cases, we showed that there were errors in the application. The Ministry acknowledged it and said they were going to fix it, which is part of our mission. We are glad that this has been fixed, but we are still fighting, and we have gone to court, and we are in court right now to get the source code. Because, just as there have been mistakes in transposition in the technical specification, there may be errors in coding.

And this is a case that technically is very simple, we are not talking about machine learning, we are not talking about neural networks. It is an algorithm that comes from a law and, even so, it is extremely complex to audit. We are extremely afraid of what it would be like with a more complex system. We are now in the first instance. The Ministry, on the one hand, the Ministry said that it is intellectual property and they do not want to give us the source code. We replied that it has been developed by the Government, so it must be public. And, furthermore, they are saying that there is a security issue, that if they explain to us how this works and give us the code, anyone can connect to the Tax Agency and get their neighbour's data and blackmail them, literally. We are fighting to ensure that the control mechanisms we have now, even if they are very inefficient, continue to be applied to new technologies and new systems.

Finally, I sometimes use the analogy that AI is a bit like the Wizard of Oz. If you have seen the movie, you know that everybody went to Oz and waited for him to solve all their problems, to give them a heart, to give them a brain.... And then you find out, well this is a bit of a spoiler, that the Wizard of Oz was actually a man behind the curtain. The AI is a bit like this. Our mission is to open this curtain, and what we are going to discover is that behind all this is not a magical system, but that there are humans or people who have developed these systems; people who, moreover, as Renata said, are white and are men, so this introduces biases.

And my goal is to try to fight against fatalism and this sense of fetishization of technology and the idea that technology is fantastic and that it is going to come and that it is going to be deployed everywhere and that it is already there. We must watch every single technological system that is deployed. We expect the ombudsmen to take on this role and to fight to monitor the implementation of each of these systems, and we need all the systems that regulate public rights to be able to explain their decisions,

to be externally audited and to have an ultimate responsible, a bit along the lines of what Carina said, that there is a person, that there is a face that can help you, that can defend you in the event that the system says no. We cannot be at the mercy of the system. We cannot be at the mercy of technology no matter how much it sometimes seems like progress and how fantastic it also seems.

And that would be my message. Thank you.

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Tuesday, March 3, 2020

Rights, challenges: Sectoral approach

Chairperson: Esther Paniagua, Technology Expert Freelance Journalist

Ángel Gómez de Ágreda, Colonel of Aviation, Head of the Geopolitical Analysis Area, DICOES

Josep M. Argimon, Managing Director of the Catalan Health Institute

Liliana Arroyo, Expert in digital transformation; Institute for Social Innovation (ESADE)

Itziar de Lecuona, PhD in Law and Associate Professor, Faculty of Medicine and Health Sciences; Deputy Director of the Observatory of Bioethics and Law-UNESCO Chair in Bioethics at the University of Barcelona

JM Sánchez Bursón, consultant

Esther Paniagua

Technology Expert Freelance Journalist

Thank you very much for being here again and I would like to thank Rafael Ribó and the whole team of the Barcelona Ombudsman for organising this conference on such an important issue and recognising that we are not experts, but we need to know how to deal with this because it is fundamental, it is already essential, and it will be crucial for citizens' rights.

Yesterday some examples were given, initially at the first-round table, on how artificial intelligence can help human rights to close gaps, etc., but in the end what the Ombudsman is responsible for, what is really important, is to see the opposite, how he can ensure that these rights are not eroded, right? The risks of this imbalance between administration and citizens that may exist because of the use of these technologies and how is this being done in different areas. This topic was also discussed a little bit yesterday.

Discrimination was discussed. Carina talked about how inequality can be automated, how discrimination can be automated. Renata also gave some examples of this. And we will discuss further about this today, approaching different areas and how this directly affects citizens. First we'll have a general overview of how this is affecting people and then we'll try to see how to address it, what the solutions could be or what the ombudsman can do.

Renata also made a call to action that seems to me, regardless of how it is done, to be necessary. Well, I don't want to get too long because yesterday we saw that there was little time for the debate, and I would like there to be time for the debate with the participants at this table. But above all I would like you to have the opportunity to ask questions about your concerns, about any doubts you may have, because in the end that is what we are here for, we are here for you. So, I would encourage you, please, do not be afraid to ask questions and please write them down. I would ask you to write them either in Spanish or in English; you would be doing the moderator a favour. And after a few brief presentations, we will ask some questions for the table and then we will move on to queries from the audience.

Well, I'm not going to delay any longer, we are going to talk about artificial intelligence and cybersecurity, and how cybersecurity and cybersecurity-related policies can affect, can have a serious impact on people's lives. And for this we have Ángel Gómez de Ágreda here, who is an Aviation Colonel, Chief of the

Geopolitical Analysis Area of the Defence Policy Secretariat of the Ministry of Defence and, moreover, author of an excellent book that we will not hear about today, because we have not come here to talk about your book, *Ángel. Mundo Orwell (Orwell World)* is the name of the book, isn't it? Well, we'll leave it at that.

We are also going to discuss about the impact of the use of artificial intelligence in research, data processing, databases, and health research, specifically in a very, very sensitive area, with Josep María Argimón, Managing Director of the Catalan Health Institute. We're going to talk, we're going to keep talking, because we've already started to give ourselves a glimpse yesterday of issues of social impact, gender discrimination, etc., and for this we have Liliana Arroyo here, who is an expert in digital transformation at ESADE's Institute for Social Innovation, who has also developed several reports on these very interesting issues and will be able to talk to us about them.

We are going to talk about bioethics with Itziar de Lecuona, who is Deputy Director of the Bioethics and Law Observatory of the University, of the UNESCO Chair in Bioethics at the University of Barcelona.

And last but not least, because this is a crucial fact, we are going to talk about the impact of the use of artificial intelligence on children and adolescents. For that we have here the adviser José María Sánchez Bursón. Welcome everyone and let's start in order, so that we can make life easier for the technicians up there with the presentations, for those who have them, and for those who don't, well in that case there's no problem.

Angel, what can you tell us about cybersecurity?

Ángel Gómez de la Ágreda

Colonel of Aviation, Head of the Geopolitical Analysis Area, DICOES

Well, I would like to begin by thanking the Catalan Ombudsman, Mr Ribó, and the IOI for inviting me to participate in the seminar and for welcoming me here in Barcelona. Thank you very much.

First of all, I am here in my private capacity, more as the author of the book and as an academic than as an aviation colonel, but obviously the role of an aviation colonel –my training is as a pilot and as a parachutist– gives you a more specific vision with regard to the issue of security.

I have some good news and above all some challenges in terms of cybersecurity and security in general in the processing of artificial intelligence data.

We are in Barcelona. We are in one of the most conscious, most innovative cities in terms of artificial intelligence, but we have to think that we are living in what some are calling a *phygital* world: a physical world on the one hand and a digital world on the other. Yesterday we were basically dealing with the digital world of artificial intelligence, of computing in general as a tool. I have been arguing for a long time that it is an ecosystem. We are living in a physical world and we are living in a logical world simultaneously; in other words, it is not a tool, it is a place or a nonplace in which we are living, and that has connotations that are really important.

Data, for example. Yesterday we were talking about the importance of preserving data because it is ours. Data is not ours, data is us. It is not the same thing. We are made of flesh and blood, but we are also made of ones and zeros in the digital world, and our digital self is now almost as important as our physical self. And we don't go around giving away noses, ears, arms, so we shouldn't go around giving away our data, our ones and zeros either. It's not that the data is ours, it's not a property, it's not a tool, it's an identity that we are defending, and that's a concept that I think is important.

We live in both realities. 94% of Spaniards use Google on a regular basis, the whole Google system. We are living in Barcelona, we are living in Catalonia, we are living in Spain and we are living in Google. I have a friend, Pilar Trucios, who always says: Where would you hide a dead person on the Internet? On the second page of Google because when we search for something we usually stick with what Google tells us on the first page. The first page is different for each one of us, and depending on whether we are connected with our user or we are on our computer.

So the universe that we are seeing depends on Google, it depends on Google's algorithm and what Google knows about us. The problem for those of us of a generation that is over 50 is that we still often think of the Internet as the Internet of the 90s, an Internet in which nobody knew you were a dog, because platforms did not yet exist, which is now the vehicle we use to connect to the Internet, and therefore we had direct access to content, more or less direct access to content, when now there is an accumulation of data about us that allows us to determine not only that we are a dog, but also the breed of dog with whom we live, where we are, where we come from and where we are going. The three fundamental questions: who we are, where we come from and where we are going.

Who we are is told to us by Google on the first page of searches, where we come from and where we are going is told to us by Google Maps every time we are geolocated, which is absolutely always. I tend to say in my lectures that when we disable the location function on our phone, we are telling the phone that we don't want to know where we are, not that the phone doesn't know where it is.

There is a very interesting experiment you can do. Just type in Google "sensor codes" and the model of your phone, and it will give you a number by calling up the page with all the sensors that your phone has active at any given moment, and you will see seventy-something/eighty different parameters that your phone knows about you at any given moment: location, acceleration, in all three axes, atmospheric pressure, temperature, longitude and latitude. A lot of parameters that we don't know, but our phone knows about us. So the data that is accumulating there is tremendously important.

This image suggests two things to me and I think both are important. In a way, we live in a mediated world. Until a few years ago we had a first-hand experience. We knew our neighbourhood, we knew if anything our city, a more or less close environment and the experience we had of things was very direct. The experience we have now is 90% of the time mediated by a screen, a screen through which our data is transparent to everybody. Well, not for everyone, but for those behind it, for the companies behind it, who use that data to construct the reality that they then return to us in the form of a concrete vision of the world aligned with their interests, often commercial ones. What does this mean? We are losing freedom and we are losing security. We are losing security the moment we are giving our data and freedom the moment we do not see reality as it is, but rather we construct our truth based on perceptions that come to us through the screen.

The other thing that the image suggests to me is with regard to artificial intelligence. What artificial intelligence does is to give us back our image magnified and optimised according to the parameters, the algorithms, and that is what scares us, that it makes us see ourselves as we are without any kind of filter and perhaps taken to the extreme.

In the end, the whole issue of security, the whole issue of autonomy, of machines or algorithms, has to do with human dignity, and only human dignity. It is with the feeling of superiority that human beings have with respect to the rest of the creatures or to the rest of the machines, or to the rest of our environment, and that makes us see ourselves as unique beings, and we are afraid of the way machines see us, the way we see ourselves.

I will not go into the details of the different types of cyber-attacks that can be carried out, not least because it is irrelevant. In the end, there is almost never what we think there is behind it. Right now, municipalities have a fundamental role to play in the security and protection of our data. In fact, we are seeing that the latest ransomware attacks, which encrypt your computer or data or files and leave you unable to access them, are mainly targeting local councils, hospitals and insurance companies; in other words, precisely those institutions, those entities that have a lot of data about us, and which have it organised in such a way that it can be used very quickly.

Does anyone really believe that they are going to attack your computer to charge you 300 euros in ransom? Or is it more lucrative, while you are worried about recovering your computer, to steal your data and then be able to sell it on the one hand and use it on the other?

Since 2007 the denial of service attack on Estonia that took Estonia offline for two and a half weeks, an attack that came from Russia, but there is no way to prove that it was an institutional attack. Since 2007 we have been aware of our dependence on the Internet.

Yesterday, we were discussing what would happen if the Internet were to end, if someone were to suddenly do away with the Internet.

Basically, I imagine the conclusion would be that we are no longer prepared to live without the Internet and that we are going to go back to the Neolithic era. We are exposed every day, every moment. The security of our data is exposed. Well,

it came out not long ago, you saw the news about people walking around the subway or airports with a credit card reader and reading your credit card details. Well I imagine that by now you all carry your cards in a Faraday cage so that the card can't be read while it's in your pocket until you take it out, don't you?

What this brings me to a little bit is the world that we have designed, the digital world that we have all designed, we are all connected to each other. Then the state, the institutions can somehow provide us with a certain level of security, but what is clear is that the direct connection that we each have with each other means that we all retain a responsibility for our own security, which we cannot delegate and which cannot be provided by any entity, by any external institution.

The ransomware attacks that I mentioned a moment ago are growing exponentially, as everything in cyberspace normally does, and they are growing precisely because there are people who have realised the value of our data, and above all because there are many of us who have not realised the value of our data.

Of course, a key chapter is probably going to be the protection of critical infrastructure, the need we have to secure access to our critical infrastructure. There are a couple of very revealing films in that sense that explain what would happen in the event of an attack.

A couple of years ago there was an attack, precisely in Ukraine, that disabled the electricity infrastructure for a few hours on 23 December, which in Ukraine is a very inconvenient time to be without electricity and gas at home, even causing at one point even some deaths from freezing in homes without electricity. Our finances right now is another area that needs to be protected, but, come on, urgently. Almost all funds are using fast-trading and financial algorithms to make trading more efficient and these systems are extremely easy to hack.

In fact, a few years ago the Syrian Electronic Army simply by hacking into the Twitter account of Europa Press, a news agency, induced the algorithms to think that there had been an attack on the White House and, therefore, there was a financial crash, a small financial crash that lasted five minutes. It took two minutes for the New York Stock Exchange to fall 142 points and another two minutes for it to rise again, once they regained control over the account. For the group that had initially gone short and capitalised on the stock market decline and then capitalised again on the rise, it was a very lucrative financial operation and very easy to do. Osama Bin Laden had done the same thing before 9/11, betting that the New York Stock Exchange, if the planes didn't stray two blocks, would go down. If it went two blocks off course, it was going to disappear directly because it was right next to the twin towers. And, indeed, it seems that he was financing Al Qaeda for a while with the financial operation he did of speculating on the New York Stock Exchange.

Another area where it is being used a lot, where security has to be strengthened because of the permeability and ubiquity of our data, is the issue of money laundering. A number of tax havens have somehow been created globally and one of the fundamental characteristics of attacks in cyberspace is that they also have to do with artificial intelligence, is that there is no longer the symmetry that there used to be of companies attacking companies, individuals attacking individuals, states attacking states. No, right now we are seeing an increasing number every year of states attacking companies.

North Korea finances itself in large part by robbing foreign banks. Obviously, for a company, the fact that a state attacks you, with all the power behind it, is particularly serious and relevant, so we have to be very careful with this type of attack, especially because they are attacks that are often not signed by the state, but are left a bit up in the air.

To conclude, in Spain we are doing quite well in terms of cybersecurity. This is a ranking that changes very frequently, among other things, because this is like the stock market, where past performance does not guarantee future performance. Today you're doing very well, you lose your track for two days and then you're in 200th place without realising it in a moment. I took one that favoured us, that suited us very well and gave us a certain optimism, which we need.

After these conferences we will surely need a touch of optimism, and we are doing well, among other things, because we have been doing our homework since 2013, when the National Cybersecurity Strategy was created. The Joint Cyber Defence Command was also created, where I was assigned for three years, and a series of committees that allow Moncloa (Spanish Government) to coordinate all the actors, both at company level, at private level, at state level and between the different levels of the Administration and at defence level. Creating links, creating ways of exchanging information at each of the strategic, operational and tactical levels. And I think that the technical level is also very important, so that actionable day-to-day information can be exchanged, as we say.

This also has its positive applications. We can generate a sustainable economy, we can generate (and there is not much to innovate on what Barcelona is already doing), we can generate initiatives that allow us to stimulate an improvement in the urban behaviour of citizens, we can create duplicates of our cities to do the experiments with gas, instead of doing them directly on the real city; we can do it on the duplicates that exist in it, although we could have done it by generating or using the data that we have taken from cyberspace.

In the end there are two aspects, going back to the beginning, that can be hacked. The network can be hacked, the technological part can be hacked and people can also be hacked, and we are seeing that the use, precisely, of the intelligence of bots worries me a lot.

I am doing my doctoral thesis on the ethics of artificial intelligence, and it seems to me that any artificial intelligence that is capable of altering your will, your freedom, is as dangerous as one that can threaten your life, because, at a given moment, this is the purpose of war. In the end it is to change your point of view, to change your freedom, and therefore it is as dangerous to hack the machine as it is to hack the end user, to hack the end user's data. Platforms have introduced us to echo chambers that tend to make us right, that tend to polarise our opinions as well, that tend to hold our attention by making us right, and that's where what I call in my book egorithms come into play. Algorithms group us together and then egorithms, that is, our desire to be relevant within the group, end up polarising those groups, they end up pushing us to the extremes and leaving us with almost no room for manoeuvre to speak, for dialogue, for interaction, which is so important. They end up almost leaving us with different languages, with the impossibility of communicating with those who do not think exactly like us.

With regard to fake news, which is not fake news –in reality it is a story in news format– the important thing is not whether it is true or a lie, it is its capacity to manipulate and direct the story. The important thing is not who kills the mammoth, the important thing is who tells the story, because twenty thousand years later, nobody remembers who ate the mammoth, what they remember is the painting it left behind.

This is used by terrorists. This is the cover of the magazine Rumiya, after the attacks here in Barcelona. 146 dead on a gothic cover, on a gothic façade. There were not 146 dead. It says dead and wounded.

Did anyone read "dead and wounded"? No, did they not? The important thing is the 146 in red, highlighting a large figure that gives the feeling of a great victory

for them, a great humiliation, a great defeat and a great loss for us. It is the control of the narrative that really matters. I was saying that we have lost security and we have lost freedom, and we have done so in exchange for gaining convenience. Right now we buy anything that is convenient, that is immediate and that is apparently free.

We have to reverse this trend, we have, and it was mentioned yesterday, the need for the users themselves to be educated and to make a sufficient effort to build our freedom. Freedom is not free. As the Korean War memorial in Washington says, "Freedom is not free". Freedom and security come with the effort of each of us. If we pretend to live without effort, we will usually live without freedom. We have to add that facet of responsibility to avoid becoming transparent, and a very important thing too, the last one:

We are in Barcelona. Remember that Barcelona, all cities in the end are but a part of what we are seeing, is also in cyberspace. We are in the physical world, but we are also in the digital world at the same time. In that Barcelona and in that Barcelona that social networks are creating for us. Thank you very much.

Josep M. Argimón

Managing Director of the Catalan Health Institute

Very quickly. In the health systems of most countries in the developed world, I believe there are two common characteristics.

The first is that investment in health is having less and less of a return. It is not like a century ago, when we could invest in vaccines, in preventive measures, in hygienic measures and, therefore, life expectancy or health indicators were increasing. Nowadays, this return is less, and although our children and adolescents are going to be or will be the 100-year generation, in other words, people who are 14, 15, 16 years old today are going to live 100 years, we all know that this is not exclusively, or even mainly, due to the health services they are going to receive. It is because there are a whole series of economic and social factors that are as important, if not more important, than the health system itself.

And, from this perspective, from these social and economic factors, there is also, and this is another characteristic of health systems, more and more inequality, and we can think of more inequality in countries such as the United States, where health coverage is not universal, and this is true, but also in European countries, where we have health systems with universal coverage, inequality is increasing.

The second important feature is that with digitisation a massive amount of data is being and has been generated and is growing exponentially.

Just to get an idea, and I am going to provide you with very, very local data, from a small country like Catalonia, with 7.5 million people, and which are therefore not at all comparable with the data provided yesterday by Ulises, the BCC researcher, which took you aback.

But, for example, with 7.5 million, we must think that every time someone in Catalonia goes to see their nurse, their doctor, or is admitted to a hospital, or goes to the pharmacy, they go with an identification card. Therefore, we have information on 150 million electronic prescriptions, on 50 million visits made in primary health care, and we have structured information on health and diagnosis, and a lot of other information that is not structured. Or we can have all the radiological imaging (X-Ray images), digitised for 10 years, and more or less well categorised.

All of this, and with this unique identifier, gives you a clear potential to be able to use in research, but also always thinking about whether this research is going to provide you with the benefits we mentioned before.

Firstly, to reduce inequalities, above all to reduce the diagnostic errors that we have, because there are some; the treatment errors, because we have them, there are some. And from these premises, in 2002 the European Union launched the report "Redesign health in Europe for 2020", which focuses precisely on the need to reuse the information that we have in the health systems, basically for research, innovation and health planning.

And these data can be used in a more or less traditional way, as we have been doing in research for decades. The only thing is that, instead of collecting information data by data, or person by person, we reuse our data, but looking for a similar architecture of studies. Let us do the cut-off studies, etc., and here we can very quickly find elements of improvement. For example, for pharmacovigilance, we can see the side effects of medicines that are widely used in our field. Surely in this room there is more than one person who yesterday, today, or tomorrow will be taking a gastric protector, omeprazole, with their accompanied risk, which it has, of femur fracture. This is difficult to do with traditional research because it means spending many years, gathering prospective information from many people, reusing that data in a safe, anonymous way. Then, if you want, we will talk about anonymisation, knowing that you are never going to be able to achieve 100% anonymisation. That is to say, the risk exists and the capacity to de-anonymise exists, we must make everything clear, but it allows you this speed in terms of obtaining results.

It is also true that all this data information can be used with a much more artificial intelligence perspective. Artificial intelligence in healthcare systems has been introduced for years. Now I could think of the robots that help us, that assist us in surgical interventions or something that is very important, and that drastically reduces the risks of a mistake in treatment when you are admitted to a hospital, which are all the robotic cabinets that allow you, therefore, to dispense medication.

This has been introduced years ago, but in recent years what is being introduced are decision support systems. For example, in the field of radiology or in the field of pathological anatomy or mental health, cardiology or in any of the fields.

In fact, on Saturday I spent some time searching in Pubmed. Pubmed is the most widely used database and the one with the largest collection of indexed medical articles in journals. Pubmed is peer reviewed and simply by searching the keywords: *artificial intelligence*, *deep learning*, *machine learning*, etc., the title and abstract, no further, year 2010, year 2019, their publication has multiplied by 17.

All these algorithms that are now on the market and that are being offered, in inverted commas, to most health centres are algorithms that can really help and that are going to help us. A very simple example can be algorithms that we are currently testing in the Vall d'Hebron hospital itself, to be able to diagnose or rule out 21 pathologies in a simple chest X-ray. To show you how useful it is, the chest X-ray is not reported, that is, when a doctor asks for a simple X-ray, it is not reported, normally they have the capacity, the skills to be able to say whether there is a pathology/injury or not, but at some moments there is doubt and, when in doubt, this report can obviously be requested, which will take a few days.

If we are able to make this tool available to primary care professionals, we will have gained something, as long as this tool is a valid tool, it is an unbiased tool, and I will come back to the biases, and it is a transparent tool.

The article that was also discussed yesterday in the first afternoon session of this symposium. So, the idea is that the field is very promising, it is still very much in the

research area, but already hospitals are introducing this: decision support algorithms. Another area that is vital where digitisation is lagging, at least here in Europe, and which we are now starting to digitise, is everything related to anatomical pathology. Although it may not seem to you, or it will seem to you a somewhat shocking fact, but the concordance today in analogue anatomical pathology systems is low. We have up to 25%, 30% of discordance. We are talking about diagnoses, in some cases of tumours, therefore, digitalisation will allow us to reduce this discordance, increase concordance and, in turn, associated with it, to be able to implement these decision support algorithms.

However, all that glitters is not gold. We must bear in mind that these algorithms are first derived from population data that sometimes we do not really know where they come from, and this is very different from a clinical trial. In a clinical trial you know very well who is included and who is not included, and clinical trials themselves are often too strict in including patients. They are too strict and, therefore, they are in a way, in inverted commas, criticisable, because when you are going to use them in clinical practice, that spectacular result that is given in the published clinical trial is reduced to a minimum expression when it is used in the real world. But we know that in the world of algorithms this is more complicated to know.

In a clinical trial, and I think it is also important to stress this, there is a whole series of highly protocolized rules on how they should be carried out, how they should be published, how they should be communicated and, if they have not been done in this way, a journal, or at least journals, where there is peer review, will not admit that clinical trial. As long as these rules of standardisation in the algorithms, applicable later to medical decision support, do not yet exist, this is a first task that needs to be carried out.

Secondly, when I refer to biases, we have clear examples, such as algorithms for dermatological analysis, and specifically in melanoma that do not include certain groups of patients of certain races, this is a first element of bias. There is always potentially a gender bias. In clinical trials it also exists, but it is clearly identifiable. So, these elements of standardisation are key, and then there is logically that when we get it wrong individually, the damage is small. If I make a mistake, I can harm few people. In the end I can visit 15, 20 people. If I get an algorithm wrong, the damage can be much higher. And then, when there are breaches, what Angel was talking about, well, if someone takes a medical record out of my office, it is a breach, and very punishable, but it is a case. If there is a breach in an algorithm, a breach with the intention to harm, the damage can be much higher.

Liliana Arroyo

Expert in digital transformation; Institute for Social Innovation (ESADE)

Bon dia a tothom, good morning and thank you very much for the invitation and the opportunity to share some of the ideas that are floating around in the offices of ESADE Institute for Social Innovation.

The first idea I would like to share is that today, in the press and in many forums, there is always talk of artificial intelligence and bias. Fortunately, we often talk about artificial intelligence and bias, but sometimes we forget that artificial intelligence or algorithms in particular were born to discriminate, and discriminate from the point of view of creating categories, identifying patterns and, on that basis, applying different solutions for group A, group B, group C.

It is surprising, then, that we are caught unaware by the fact that artificial intelligence discriminates. Not least because it is something we do every day. We discriminate when we open the fridge, depending on whether we want to follow

a healthy diet or not that day; we discriminate when we choose a school for our children, and so on and so forth, or we discriminate when we travel, and because we work with stereotypes that sociologists in sociology call the *social imaginary*.

Therefore, artificial intelligence is a huge mirror of 1,500 magnification of what the social imaginary is. The problem we are facing is that, right now, the social imaginaries that are developing this artificial intelligence are very few and too homogeneous. Having said this as a starting point, which would be a kind of sociological diagnosis of the problem we are discussing today, I would like to delve into examples of types of bias, especially towards vulnerable groups, with a gender perspective, race, but also classism, ableism, ageism, a series of axes of discrimination that are historically permanent. But before going into these examples, and I will also end with some good practice, in case we can take ideas from there, I would like to talk about three different types of biases. I'm sure there are more, but I think these three will give us a little background.

The first one is the one you find on the left (points to image), because it is the most common. It is the cognitive bias, which is basically the "glasses" we wear, the glasses we have built for ourselves through culture, education and direct experience. These are biases that are directly transferred to this mirror, this magnifying mirror. They are the cognitive biases that we carry with us, our baggage, our background, our way of looking at the world.

Then there is this one at the top right that says *code*. It would be a small code window, because there is the technical bias, which I take as a broad category not only of the bias that can be introduced in the code, but also something that was also pointed out here, which is the bias that the data may have. If we use the data we have at hand and the data we can use, because in the end to apply artificial intelligence, and experts know this better than me, the conditions of the information have to be very specific. It seems that we can put everything, but it is very different if the data is structured. If the information is not structured or if we do not know the origin of the data, if we do not know how much manipulation that database has suffered, etc., so the technical bias can come from the origin of the data, the composition of that information, data that is outdated...

Even human errors in writing the specific line of code, and the fact that there is no validation behind it, that would be another chapter. And then we have what are called *emerging biases*, which are those that are neither exactly cognitive nor technical. That is to say that they develop as this artificial intelligence learns, executes this automated learning and new interactions and new possibilities are generated. And here I bring an example, which is the Microsoft bot. Perhaps many of you know it. It was called Tai and it lasted 24 hours, because they put it on Twitter, I think it was in 2016. They didn't open the account, it tried to be a profile of a girl of about 18 or 19 years old who basically what she was doing was learning, feeding off what she was reading on Twitter with a lot of capacity. After 24 hours, they realised that what she had been able to absorb was hate speech, misogynistic speech and they ended up with a bot, what some headlines called a *misogynistic neo-Nazi bot*, when in fact it was simulating a 19-year-old teenager.

That reinforces this idea of saying, well, sometimes this particular bot perhaps didn't have a cognitive bias when it was designed, perhaps it didn't have a technical bias, but it presented us with an emergent bias, which in the end closes the circle a little bit of what we were saying. It is still a bit of the imaginary that a platform like Twitter feeds, although I must say, that there is not only hatred on Twitter, but there are many professionals who interact on this platform in a civilised way.

The point is that it's a clear example of emerging bias. Having said that, I would like to mention some classics quickly, as I do not want to go on too long. But going deeper into this repetition, in this recurring situation that if you are a person with

white skin, Microsoft Face Plas or IBM are going to be much more capable of identifying your gender than if you are dark-skinned. But not only if you are dark-skinned. The data tell us that these platforms identify you much worse, with a much higher error rate, if you are a woman, not only if you have dark skin, or non-white skin, but also if you are a woman. This is double discrimination.

We also have the case of Amazon. You know that Amazon designed a recruitment algorithm, precisely because they had open vacancies, very high levels of CVs received and difficult to process in the time they needed to fill those professional vacancies. So they said: "No problem, we are going to take a database where we collect the professional profiles of Amazon's staff for the last 10 years and we are going to design an algorithm that is basically inspired by that data". What happened? That during those ten years Amazon's professional workforce was basically men, it was basically male.

Therefore that algorithm, when the word *woman* appeared on your CV, not only did you identify yourself as a woman or female in the gender section, but if you would have said "I have been participating in technical development studies of the algorithm for menopausal women". I'm making that up as an example, okay? Just because the word *woman* appeared on your CV, somewhere on your CV, you would directly be ruled out.

Obviously, thanks to journalistic work, that came to light. They cancelled the algorithm. But notice that there was that ten-year bias. Another very well-known case, and it came from *ProPublica*, which did a very thorough investigation of the Compass case, which is a system that tries to predict the risk of recidivism of a person who has been arrested. You know very well that there is a problem of racism on the part of the police in the US, and that is obviously also reflected in the algorithm.

To give you an example. On the top left (points to image), you have two people who were arrested for drug possession. To the left is Dylan Fugett, who was stopped at the time and the algorithm diagnosed or predicted that he had a very low risk of re-offending. It gave him a number 3, whereas next to him a racialised person, with the same conditions, same arrest, no criminal record, etcetera, etcetera, had a risk of 10.

After a while, and this is what *ProPublica* provided with their investigation, they saw what had happened with these two people. Because in the case of Dylan, the non-racialised person who had a low risk, a risk of recidivism number 3, they arrested him three more times for drug possession, while the person who had risk number 10, the one on the right, Bernard Parker, they arrested him once and they never had to arrest him again because they didn't find that he was dealing drugs or involved in any other crime after that.

However, if you look at the two graphs below, the one on the left shows the distribution curve of the risk of recidivism for white people. You can see that there is a sort of slide that goes like this (she points to a curve with his hand), which would have to do with an expected and usual distribution. Whereas if you are a racialised person, it seems that they do it by rolling the dice, because it doesn't matter where you are placed, as in reality the proportion of people who can reoffend between one and ten times is practically the same, and that is where the discrimination of the case lies.

More examples. This was published a few days ago by *The Guardian*. For instance, nowadays there's a lot of interest in being able to read our emotions by using facial recognition, but, well, we also know that frowning doesn't mean the same thing here as it does in Papua New Guinea or China, or any other gesture. We all know that we, to indicate that we agree with something, we nod (shake our

head up and down) like this. If you have worked with people from India, they will do like this (shake their head sideways) and with this gesture they are not saying more or less, they are saying yes. Well, we can extrapolate this to many other examples. But again, if we assume that the way of expressing emotions is universal, we are making a mistake, because we will believe that we are eliminating the cognitive bias of this facial recognition algorithm. But in fact, we will be starting from an erroneous assumption. And I am already pointing out some things that David said yesterday, for the Ombudsmen and women: you must not think it is a technical problem, it is a problem of rights, and it is a problem that needs to be discussed.

More information. Well, there is a study by Bolukbasi and other researchers entitled "Men are to computer programmers what women are to housewives". Well, if you search in Google for nursing, medicine or web developer, you will see what kind of images you get.

I also wanted to talk about Joy Buolamwini's case, I don't know if you know her. Well, this girl is holding a white mask in her hand, three masks in the images, if you look at her image in the centre. She's a researcher, she's doing her PhD –I think she's finished it– in the US, on facial recognition. The problem was that the database she was using had no sample of racialised people, so in order to develop her own algorithm she needed to use a mask with the features of a non-racialised person in order to develop her research.

Imagine that you are developing an algorithm, the algorithm that is part of your doctoral thesis, and it does not recognise you. Imagine that. On a scientific level, that has a lot of implications, but on a psychological level, on a personal level, this person could have decided to give up his career, unless it had occurred to this person to say: Hey, what's going on? And it turns out that the algorithm did recognise a mask, which is not a person, which has no rights, etcetera, etcetera, but it didn't recognise her.

There is a documentary, the case of the Netherlands, which was discussed yesterday, and I would like to comment on three good practices, as I am already focusing the home straight. These good practices have to do, for example, with a study that has just been funded in Australia on diversity issues. We tend to think a lot about race and gender, or more precisely, we still think little about it but we are starting to talk about it more. But we think little in terms of capabilities and disabled people, access to technology, for example, and the ability to be represented in social imaginaries are very limited. The interesting aspect of this study is basically that it generates an alliance where there is co-design on the part of these people. That is, the algorithm is not developed separately, the solution is developed in a laboratory in an industrial context and then it is tested, but questions are asked from the outset about what is of concern, what is relevant, and so on and so forth.

In Finland, in Espoo, there is an artificial intelligence trial very similar to the one reported in the Netherlands a few weeks ago. But what it is doing is precisely to ensure that its algorithm does not intensify surveillance and data collection in those neighbourhoods with the poorest populations, because the economic issue is also important. Finland is also an example because it understands that it is necessary to know and comprehend a little about how artificial intelligence works, so that we can *opt in-opt out*, whichever option we choose, which is why they have created an open course, which any of you can take, on the basic elements of artificial intelligence, a little along the lines of digital literacy.

Then here, we have local examples, and this is from a report by Carme Peiró, also a few weeks ago. I don't know if she is here today, but she was here yesterday, she was participating. She has spoken to several administrations, and, for example, there is one called Riscanvi, from the Generalitat (Government of catalonia), which

is a protocol for the evaluation and management of the risk of violence in the prison population, similar to Compass. What it basically does is that instead of collecting, I don't know if Compass collected about 300 variables, this one only collects 43. If a person with a certain level of risk detected asks for an exit prison permit and do whatever, he or she is allowed to go out, but certain measures are activated. This algorithm is also updated every three years, improvements are added. Meaning the algorithm is being revised, and there is a treatment board that can validate it.

The area of social rights of Barcelona City Council also discriminates in deciding who is granted aid. And what this algorithm is allowing is, instead of the algorithm making the decision, to assist the professional in allocating these grants. So, one of the effects it is achieving is the homogenisation of the responses, with which there is a certain improvement in the response and the quality of this service.

Finally, my last idea would be: Why do we insist on achieving what would be zero bias when, in fact, as we said at the beginning, if we use artificial intelligence to discriminate. Then, the question or the issue is, how can we move towards a bias that is conscious, that is consensual, and, above all, that is responsible?

Thank you.

Itziar de Lecuona

PhD in Law and Associate Professor, Faculty of Medicine and Health Sciences;
Deputy Director of the Observatory of Bioethics and Law-UNESCO Chair in
Bioethics at the University of Barcelona

Thank you very much. Many thanks to the Catalan Ombudsman and to the organisation for having managed to hold this conference, which I find very appropriate and timely. And thanks to all the participants for their contributions from this interdisciplinary approach. I have followed the work of the Catalan Ombudsman, also through the figure of the Patient's Ombudsman, and I believe that it is a fundamental actor. One example is the appeal of unconstitutionality on the possibility of profiling to guide voting intentions. In this regard, you should already be aware of the implications of the Cambridge Analytica case. Let's reflect on the ethics or "bio-data-ethics" of artificial intelligence, its ethical, but also legal and social implications.

Ethics is always present, ethics as a support for law and ethics also leads us to reflect on political decisions. It worries me tremendously that we are here as if ethics and the regulation of artificial intelligence were at ground zero, as if it were the future, and it is not the future. I would say that artificial intelligence is already routinely applied, and in some cases I would go so far as to speak of the past. So I think that those of us who are sitting here, whether you are ombudsmen or we are different profiles, we have to take action on the matter, and especially bring the legislation down to earth, with what we already have, because what we have –which is a lot– has to be processed and it is time to apply it. This speech, this presentation, in 2017, would have been talking about the future, but since May 2018 the situation has changed because since then the General Data Protection Regulation has been applied, which I am going to refer to, and we have a political decision, such as the Digital Strategy and the White Paper on artificial intelligence. Moreover, Europe advocates "trusted" artificial intelligence, as the High Level Expert Group on Artificial Intelligence has made clear.

For the first time, Europe has decided to break its dependence on the US technology empire GAFAM, i.e. Google, Apple, Facebook, Amazon and Microsoft (excluding the Asian giants). Angela Merkel warned about this overdependence a few months ago. Neither the member states nor Europe have their own digital infrastructures. It is essential to understand the consequences of this situation, just as it is essential to realise

that, for the time being, we need third parties, the bigtechs. This dependence must be studied in detail, particularly in the field of health, in order to carry out research projects that are financed through tax payments and on which a result is expected, either directly or with an impact on future generations.

My contribution to this debate is academic, but it is also intended to be practical, due to my experience as a member of research ethics committees in public research centres and universities. Ethics committees in charge of assessing the methodological, ethical and social aspects of research and innovation projects based on the development of algorithms that feed on personal data and that apply artificial intelligence and other technologies such as big data or virtual reality, to give a few examples. From the philosophy of law, I understand bioethics as a bioethics of impact, which not only serves to rethink life and death and the implications of scientific and technological advances. In this sense, the intellectual debate is exciting, and I subscribe to the words of Daniel Innerarity and José María Lasalle at yesterday's round table.

Decisions have to be made, and we have to contribute to advancing the development and application of artificial intelligence for previously agreed purposes. My presentation is aimed, from a critical perspective, at understanding the problem and the issues we are facing, because we cannot expect citizens who are going to hand over their personal data to develop algorithms in the medical field and will not question it to be able to make free and informed decisions, if we do not first specify what we are talking about and identify the challenges and problems. The asymmetrical situation that is generated is logical: whoever has the data has the power, and whoever gives it away from good intentions loses it if an informed social debate is not first generated and if, having rules, they are not properly applied.

We are no longer anonymous but re-identifiable, and we have a problem if we do not accept that anonymity can no longer be guaranteed. Artificial intelligence is currently in vogue, but just recently there was talk of big data, when what we should be talking about is the convergence of the aforementioned technologies. We must understand that, in the data-driven digital society, we are the raw material, the providers of personal data and also the recipients of the results of applications and services based on the processing of personal data. Our data feeds artificial intelligence, so this is a matter of concern for all of us. We need to understand the implications, especially in terms of policy and regulatory design.

Given the research and innovation processes based on the application of technologies such as artificial intelligence, big data, biometrics, etc., especially in the field of health, I would like to reflect with you on the ethics of automated decisions, because that is what our times are all about. Automated decisions that are determined by an algorithm programmed by humans, which draws on data from reality and incorporates biases. Hence, algorithmic discrimination must be avoided. These are decisions that, according to the applicable rules, require human review before they are made. Imagine the implications this has in the medical field, and know that since 2018 there is a right to complain against these decisions. Artificial intelligence has arrived to provide support, to help improve decision-making, but it cannot replace humans, and even less so in matters as important as people's health. It is true that it is amusing to pose this dystopian future, as has always happened, let's think first of the power of genetics and now the power of artificial intelligence.

Let's continue to look at the intricacies of artificial intelligence such as the fairness of the algorithm and the fairness of the data, because the question is no longer about programming the algorithm fairly. The question is what data it is fed with. Then there is the responsibility of all actors. We have to work in an interdisciplinary, proactive and risk management approach to the processing of personal data. Problems are often raised as if they were ethical, not technical, when they are precisely that, technical

problems such as, for example, ensuring that no personality can be attributed to the set of personal data being handled through appropriate pseudonymisation techniques. If we do not first understand the technical problems, ethical reflection will be useless.

The quality of the personal data that feed algorithms is not a trivial issue. Neither is the management of this data and how it is shared. I alluded earlier to Europe's lack of infrastructure for storing and exploiting data, especially its interoperability and re-use. On the other hand, it is logical that there is interest in accessing health and other databases, since whoever has personal datasets today, well structured and meeting quality criteria, has the power. Power that is also achieved, with the necessary time and effort, by combining quality databases with others of relative quality, so to speak, to extract behavioural patterns, predict trends and thus improve decision-making. And this interest in access is shared by both public and private initiatives for a variety of purposes.

In particular, I am interested in discussing with you on the value and price of personal data in the digital society; and, above all, how we have to combine, and this is not to be denied, public and private initiative, the common good, with personal interests. Anyone who separates public and private is wrong. Separation is in many cases not possible, synergies between public and private initiative are necessary and are part of the research, innovation and development plans of the States. So, today, we should remember that Europe adopted in June 2019 the Directive on the re-use of public sector data. It is not an idea in the development phase, it is a political, scientific and economic commitment with an enforceable regulation. Thus, in order to make progress in the application of artificial intelligence, it is necessary for data to be findable, accessible, interoperable and reusable. In this sense, we should also analyse informed consent models for the exploitation of personal data for medical and research purposes, for example. Only if the individual consents to certain processing of personal data, these can be used. To this model we should add the possibility of presumed consent (opt out), i.e. that, by default, we are in databases that would be used for certain purposes such as biomedical research, among others. The aim would be to avoid secondary uses and covert discrimination based on who may have access to the data and therefore who would be authorised and under what conditions.

If you look at recent policy –well, recent so to speak, because big data has been talked about since the early 2000s– the European Union makes it very clear: what can big data do for you? And then, as I said, the European alliance on artificial intelligence. We recently read in the press that Margaret Vestager, one of the vice-presidents of the new European executive, had only 100 days, and she has done it, to present her new strategy, which I mentioned at the beginning.

Another issue I would like to bring to the debate is the need for research and innovation projects in artificial intelligence in the field of health to be endorsed by institutions of recognised prestige such as research centres and hospitals. What for? For what you said before, to validate those interventions, those innovations, to obtain the seal of quality and the necessary certifications. It provides security, and allows an ethical endorsement of the proposed intervention. This is not a minor issue, since the burden of these guarantees falls on the research ethics committees.

What are ethics for? Ethics is about happiness, about freedom, about the values a society strives to achieve and about the limits and threshold of risk we are willing to bear.

We really have to ask ourselves whether this digital society makes us more or less free than the analogue one; and whether we make the decisions ourselves or whether others make them for us. A cyber-attack can be identified, but not such

covert discrimination on the grounds of access to personal data because health data, for example, which, as you know, are special categories of personal data, have not been adequately protected. Let's see what I mean: When my son is looking for a job and is told that he does not fit the profile, maybe it is because he has a genetic predisposition to suffer from some disease, and maybe he does not even know it. Perhaps because of our ignorance and lack of understanding of the problem, these personal data have been released and are in the hands of third parties who should not have had access to them and so the concerned person is discriminated against in a covert way.

On the other hand, more efficient healthcare systems, personalised medicine, and the ability to predict the adverse effects of medicines in less time, or even COVID-19 prediction systems, are desirable. These issues, which are already a priority for the EU from a political point of view, must lead to interventions and applications that are proportionate to the aims pursued and respectful of the rights of those involved.

Therefore, given these scenarios, I stress again that we must stop referring to data being anonymised, because this is not the case. The combination of the enormous amount of information stored in databases containing personal data and the current state of scientific-technical knowledge lead to re-identification by default as a basis on which to determine what we want, what we prioritise and how we are going to do it so that it responds to the social challenges of our time, otherwise it would be maleficent. It would be about combining these datasets to determine patterns of behaviour, predict behaviours and improve decision-making, aided by artificial intelligence.

My lines of research focus on the study of the ethical, legal and social implications of biomedicine and biotechnology, and for several years now I have been analysing the application of artificial intelligence and big data, etc. as well as the development of digital devices and Apps in the field of health. In this context, and in particular, as far as health research and innovation projects are concerned, my contribution is to the development of guidelines to improve the evaluation processes of these projects by research ethics committees and, especially, to identify data markets disguised as research and innovation.

There are numerous proposals for access to databases, to health databases, for example, with interests that are not strictly speaking those of research, but which appear to be so. It is obvious that we live in a market society in which there is a clear tendency towards the commodification of the human body and its parts, and I would add the monetisation of personal data, which is not a desirable situation. But here we are already seeing tensions. We want personalised medicine and at the same time the European Union is committed to a digital, data-driven single market, which makes it more competitive. I will now show you some slides to show you what I mean:

The European project NESTORE, funded by the European Union's Horizon 2020 framework programme, in which I participate, is an example that is part of these scientific and economic policy decisions, which aims to develop a virtual companion for people of a certain age (elderly people, I would dare say) who are not dependent through monitoring and recommendation systems for active and healthy ageing. It is a virtual companion with not inconsiderable funding and developed by a public-private consortium. A year ago, on 4 March 2019, I took part in the presentation in Granada of the Spanish Artificial Intelligence Strategy for R&D&I, which is another example of a political decision in the field of research and innovation with economic impact.

If you travel to Amsterdam by plane, as you pass through the huge Schiphol airport you will have noticed, and if not here you can see from these photographs that I took on several occasions, that it is absolutely packed with advertisements from one of the big tech giants. These are advertisements about a cloud that is going to change the future of golf or DNA analysis capabilities exponentially. This cloud, and fundamentally

cloud-delivered services, which is where the value lies, will be able to help improve decision making. The personal problems that one of the best golfers of our time has are well known, and if that personal information is combined with the position and speed of the shot, the problem is obvious. In these databases that are going to be exploited by reputable third parties, we may have to start to cover up information to comply with the principle of minimisation (to which Dr. Argimon referred) and the need for data.

This example (points to slide) is interesting, isn't it, because nobody is interested in whether you have problems with sex, with alcohol, but the information is there, ready to be combined. Should we? And here's the thing, certain personal data cannot be available to everyone and is not necessary. Let's continue with photographs. Scanning 7 trillion DNA in less than 12 hours, this used to be unthinkable. But once again it is bigtech that offers this possibility.

Catalonia is a pioneer. Catalonia took the European guidelines very seriously, aligning itself with the opt-out system, which is based on presumed consent. As for this image (photo of the health card), it accumulates a lot of information: the children who depend on me, whether I pay social security contributions for them, my pharmaceutical consumption, my advance directives, if I have them, my postcode, my sex, my date of birth, which, as you know, allows the person to be re-identified.

Therefore, we have the obligation to protect people through the protection of their data, their dignity, honour and freedom are at stake depending on the uses. It should be added that the Catalan model of computerised and shared medical records (HC3) is a benchmark for data quality, reliability and security. For this reason, it is not possible to speak only of the future, because it comes from before. The situation should be reviewed in order to integrate ethical issues from the outset and take the right direction.

In this sense, the Data Analytics Programme for Research and Innovation (PADRIS), the programme formerly known as VISC+ (from its Catalan acronym "més valor a la informació en salut", i.e., more value to health information), was born with some problems and unresolved issues that led it to be paralysed by parliamentary decision precisely because, in short, ethics had not been integrated from the outset. It is important not to put aside the lessons learnt, and not to repeat mistakes that can be blunderous and impossible to solve when personal data are at stake.

There is a need to foster a societal debate on artificial intelligence and its ethical aspects, in order to make free and informed decisions, based on transparency and accountability. This is about determining which projects and initiatives are prioritised; whether it is an opt out system; how uses and purposes are reported in a clear and transparent way; and how systems of governance are developed in line with the responsible research and innovation that Europe advocates. In short: the more information and the more transparency, the less obsession with informed consent as a guarantee. In our system, which is based on solidarity, personal data cannot be conceived as a bargaining chip, nor can it be handed over to the highest bidder. This is a misconception of the established rules of the game.

It seems like a good idea to download an application connected to our medical records that warns us that we are going to have a heart attack before we have one, which prevents and saves time, resources and, above all, saves lives (see the following slide of a mobile health application advertised by The Economist). But such applications require for their creation and use access to personal data, health data, among others. Sensitive information needs to be released to third parties outside the health context who may have different interests in the datasets.

How is research done today? Hackathons are held to develop algorithms (e.g. for predicting certain diseases), as part of a competition involving programmers who may never have received training in medical ethics, or in confidential data processing.

Another example: direct-to-consumer genetics companies. These business models, which are proliferating today, consist of analysing DNA samples sent by post to predict the risk of suffering certain diseases and also (for an additional charge) provide information about your ancestry. Pharmaceutical companies then make agreements with these companies to access their databases (which include not only health information but also other information such as tastes, hobbies, etc.) to develop interventions and treatments. This model of access to personal data for research is not in the manuals, nor can it be categorically stated that those who sent the saliva sample to find out more about their DNA were thinking of being research participants.

What am I getting at? We should not think that our data is of no interest to anyone. We, people, in isolation and as a group, and the data we generate are of special interest to both private and public initiatives. Our digital society is fuelled by data, as we have seen, and ontologies need to be created in order to make more personalised decisions, which requires a lot of data.

Returning to the need to ground the existing regulation, I am particularly interested in highlighting here that together with the principles and rights that the European General Data Protection Regulation establishes, it also incorporates a novelty that is highly relevant from the perspective of risk management and proactive responsibility, which is the assessment of the impact of the processing of personal data of the people concerned. This assessment, which must be carried out by means of certain methodologies, is mandatory in certain circumstances, such as, for example, when processing health or genetic data, etc., and allows for a joint analysis of technical and security issues, together with the rights and ethical perspective already analysed. This assessment leads to an action plan to be implemented, which will be decisive for the adequate processing of the data.

The concrete proposal is, as I said at the beginning, that we stop thinking that regulation is yet to come and implement what we have. But there is a tendency to demand regulation: to regulate artificial intelligence, to regulate big data, and it is not possible to do so. The problem is that regulation is not being implemented, primarily because we do not identify the problem due to a lack of proper understanding of the not-so-new digital paradigm we are immersed in. It is complex, scientific and technological changes are happening at breakneck speed and it also implies changing the approach to ensuring data protection in the face of the intensive application of new technologies. This is based on having access to the technical and organisational measures of projects and initiatives, right from their design, in order to be able to analyse the impact of the proposed data processing on the individuals concerned.

These conferences are proving that we have legal and ethical references, that there is relevant work on the subject by different bodies and organisations. If we are not capable of processing, never better said, and applying these frameworks, we cannot expect to have citizens empowered by technology. Thank you very much for your attention.

José María Sánchez Bursón

Consultant

Many thanks to the Catalan Ombudsman's Office, and personally to its team, for allowing us to share with you all a set of general reflections on the context of children, childhood, in the new scenarios of artificial intelligence. In general, we talk about the risks and difficulties that artificial intelligence will bring to our society, to people. I would like to focus more on the opportunities that artificial intelligence will bring to children in the new social contexts governed by new technologies.

If I may, I will take two minutes to analyse the impact that the technological emergency, through artificial intelligence and other technologies, is really going to have on the new society we are joining. First, there is a huge technological emergency looming. I give you a first slide, based on the singularity theory, which explains how the potential of our personal computers and their relationship with human intelligence will grow. Our brain is also an information processor. Basically, we have a natural, chemical/biological processor, which processes information and ends up analysing around 10 to 16 operations per second. This is the intellectual processing capacity of the human brain.

If you look at the evolution of a personal computer, here we take a reference value of a thousand dollars, as a superior mobile phone, in 2010 it had the capacity to process the synthetic intelligence of the brain of a fly. Later, around 2020, it will reach the processing intelligence of a rat, and it is estimated that in the decade between 2030 and 2040 it will have the processing capacity of a human being. In other words, a thousand-dollar computer, like our mobile phone, will have the same intellectual capacity as a person, it will process a volume of information similar to a human brain. And it is expected that between 2050 and 2060 it will have the same capacity as the whole of humanity, the 10 billion people who live in the world. In other words, a thousand-euro computer will give us more intellectual capacity than humanity as a whole.

This is going to have a severe impact on all of our lives, how we live, what we do, what we work on, obviously, huge, inevitable. It's curious because all the processes of developing artificial intelligence are being built on the replication of the human brain itself. Basically, what we do is accumulate information and then, through a processor, we make decisions, discriminating, selecting the information we consider most relevant, and our artificial computers are doing the same, and that is why our artificial computers also learn, also deduce and will also be able to incorporate the ethical element, and will foreseeably soon be far superior to ourselves. I think that is important.

Moreover, all technological processes occur in an increasingly hurried manner. The next slide shows us the times of adherence of technologies to social culture. While electricity took 50 years to reach 25% of the population, today technology such as Pokémon games reached 300 million people in two days. In other words, society's capacity to embrace the new technological emergency is explosive. This will generate a permanent renewal of our knowledge.

It is estimated that by 2040 human knowledge, i.e. everything we know, all our encyclopaedic knowledge, will double every two years. Just look at the potential that this means and the great challenges that lie ahead. It is not so much knowledge, but what futurologists talk about as anti-knowledge, because a lot of what we think we know is actually false, it is anti-knowledge. Possibly, almost everything we studied at school has subsequently been revealed to be wrong or false, everything has been substituted. In chemistry studies at school, I stayed with the atom, today it is known that underneath the atom there are five smaller units. We study the seven planets, we no longer know very well what a planet or exoplanet, gas planet, etc. is. That is to say,

the knowledge that has been revealed is so powerful that it is really transforming in a very intense way the whole presence of man in the new society.

And it is in these new contexts of technological acceleration that our children find themselves. First of all, it should be noted that our societies have traditionally neglected children's rights. Historically and even today. Children are the most invisible and silent people in society. At the end, if we have time, I will give you a small reference of the invisibility suffered by our European, not African, children in Europe. They do not vote, they are not represented, they do not participate in the media, they are really socially invisible.

We subject them to endless rules of protection, risk prevention and safety without taking them into account. Moreover, when we boast in the West that children are the first in emergencies and disasters, this is also false, they are the last to be helped. The little research that is available on children is occasional, they are under-represented in mainstream research, and it is clear from what we deduce, that in mass catastrophic incidents it is actually children who die first and foremost. So even that is not true. They really are the last ones in this society, and that is why we are hopeful that in the new scenarios, opportunities will be created for the valorisation of children in the next societies, especially by strengthening the areas of education and health.

We need to define a new governance and commit to an ethic for the future, because we are the first generation of humanity that will have the capacity to impact the whole of human development, putting its own sustainability at risk. That's right, current generations have the capacity to impact the planet, through global warming, overexploitation, population growth, pollution, etc., putting human sustainability itself in crisis for recent generations.

There are many children in the world, but there are fewer and fewer of them, and Spain is a paradigmatic case of what we are talking about. A hundred years ago children represented 40% of the human population and now they only account for around 28% of the population, and they will continue to decline, and what is more, most of these children live in developing countries. In more advanced countries such as ours, the pre-eminence of children has declined dramatically. In fact, our population pyramids are truncated. Japan today represents the most extreme case. Although in Spain we are the third oldest in the world, according to the indicators we use, the average Spanish population is around 44 years old, and the voting population, which is very significant because they are the decision-makers, is between 52 and 53 years old. And that is why we estimate that children are going to be fewer and fewer, but in the new societies they are going to be very significant, with a series of variables that we cannot go into in depth now, but which are connected to the capacity for innovation and creativity, adherence to technological appropriation, the inverted knowledge pyramid, the importance of education, and other reasons.

On the other hand, in a way, artificial intelligence is going to make it possible for our children to be children with superpowers. They are children who are going to be born with what we call an enriched personality, an extended personality. We are going to enjoy a very augmented reality. There has never been such a severe biological and technological gap between us and our children or our grandchildren. It is totally different how they are born, grow up and live. Shortly after they are conceived, at three months they already have a 3D video of the mother's ultrasound scan. Before they are born, a significant percentage are already present on social networks, they will have immediate access to the most advanced technologies in a very short time and they are born in a process characterised by a radical biological gap between how our lives have been shaped and how our children's lives will be ordered.

The attainment of digital skills is leading the way. A study by antivirus company AVG looks at children's digital skills and finds that digital skills are more accelerated than manual and traditional ones. Children learn to surf the internet long before they

learn to swim in a pool, to handle searches before they learn to ride a bike, to use a mouse before they learn to tie their shoelaces, and so on. Our children's digital skills are far more advanced than even ordinary manual skills, and this will generate 21st century super-children with superpowers of information, communication, knowledge, connection, etc. The Catalan researcher Dolors Reig published a book a few years ago on hyper-connected young people, and she indicates that these are children who are connected 24 hours a day, faster in their communications and development of symbolic language, who multitask, who do not differentiate between the on-line and off-line (real-virtual) worlds, who handle new forms of expression and, paradoxically, are increasingly hyper-individualistic, but at the same time more sociable.

In short, it is foreseeable that technological development will eventually promote a great potential in our children, that they will be more intellectually and socially developed, more empowered, and that they will enjoy a great potential in life. The key element will be the education and learning process they undertake. In this area as well, trends and production procedures are changing.

From the outset, education is also becoming globalised. A report by UNICEF's International Children's Observatory in Florence explains how the educational capacity of children has also become globalised, it has become very extensive. Before, we educated within families, schools and close contexts, today education has opened up, the media, the Internet, television, etc. educate.... The great sporting leaders, the leaders of the game, incorporate more values than the personal relationship with parents. And, in general, a series of variables that are going to transcendently change the evolution of our children's educational models and the value of knowledge. Clearly, education is going to be vital in the new societies.

To put it another way, according to forecasts, children born today will live to be around 100 years old. As it has already been mentioned, in a world populated by 12 billion people, children will develop different personal, family and professional identities throughout their lives in a context where knowledge will double in a short period of time. In other words, we have no idea what the world will be like in 2040 or 2050. Imagine how a kindergarten or pre-school teacher approaches learning, the knowledge of these children, with the scenario of uncertainty they are given.

And artificial intelligence is going to shape the entire educational process, and this is one of the great challenges that those of us who are trying to activate children's rights are taking on. That is to say, how we integrate the new realities into the education system and not apply our traditional educational patterns to them. Pedagogy, cognitive diet, curricula are going to change. Even the very survival of the school, according to UNESCO, since among different trends, some are considering its disappearance.

The same is true in the world of health, which is going to change radically. Other colleagues who have spoken before me have already pointed this out, and I don't need to reiterate it. Likewise, all governance, everything that means the incorporation of children in public spaces and in the strategic planning of countries, the approval of budgets and other public affairs. Norway is currently the only country in the world that applies an analysis of the impact of public budgets on the younger generations. And this is a key element in the results of the society we build. I will try to explain this in the final part of the presentation.

And we also need to build a new ethic of the future, a future that provides for the long-term sustainability of our humanity, for the coexistence of the new generations and the generations that have lived, and to build a planet and a culture that is truly respectful, with a culture of learning, with the care and development of children and with children's rights.

And with regard to the challenges we face. To start with a brief reflection. We are witnessing a process, not of confrontation with technologies, but on the contrary, a

process of fusion, of absolute technological integration. In other words, we are now experiencing what is known as the black winter. We are thinking more about control, protection, about everything we have been talking about before. We are afraid of insecurity and uncertainty, but technologies are going to lead us to embrace humans, electrons are going to be incorporated into our own biology, and we are going to see processes of transhumanism, augmented reality, reinvented humanism, increased humanism, which is the scenario in which our children are going to live. And the great challenges, obviously with regard to childhood, are to break with the invisibility of our children in these new contexts, which is still very severe despite the circumstances and the evolution that has been taking place. The invisibility of children is still very strong in our society, and is faithfully reflected, for example, in our public budgets: 57% of public expenditure goes to the older classes, and only 7% to children. The dedication of our own public budgets is manifesting the sense of our current decisions, and this obviously needs to be corrected, and most of all with regard to the new generations on whom we have to rely for all our own social sustainability.

In the Spanish reality, the effect is perverse. We have a public debt of close to 100%, which we are financing with 20, 30 and even 40 years' credit at sight, which our children and grandchildren will have to pay, and part of the debt that our grandchildren will pay is owed to our current creditors, one of which, for example, is the Norwegian Sovereign Wealth Fund, which profits from the sale of the oil extracted in their country, which invests it for the benefit of their new generations in 25 years' time. Therefore, our grandchildren will have to finance the lives of the grandchildren of the Norwegians, because they will have to pay our debts, while exploiting the credits of the present Norwegians, because they have been applying the contrast of the analysis of budgets in the new generations and have considered generational sustainability.

Indeed, sustainability in our human development is built on the basis of how we treat our children, and the new social contexts will help to value childhood as an essential intellectual legacy for building the societies of the future. Without children there is no future, let alone progress. Today there are more than a billion children interacting in multiple social networks, in an open, free, uncritical and non-hierarchical way, and they do so despite the risks, lavishing behaviour based on collaboration and mutual cooperation. We hope that the future will provide a much more meaningful forecast of children's rights in new, more integrative and creative societies. I am at your disposal for any further questions you may have. Thank you!

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Monday, March 2, 2020

Rights, challenges: International treaties and definitions

Chairperson: Miquel Molina, Deputy Director of *La Vanguardia*

Gregor Stojin, president of the CAHAI

Martha Stickings, policy analyst, Fundamental Rights Agency (FRA)

Geraldine Mattioli, adviser at European Commissioner for Human Rights (Council of Europe)

Peter Bonnor, Senior Legal Officer, EU Ombudsman

Miquel Molina

Deputy Director of *La Vanguardia*

It is an honour for me to participate in this meeting. My name is Miquel Molina, I am Deputy Director of *La Vanguardia* and I usually write every Sunday about culture and cities. And in the framework of these topics that I write about, in harmony with some institutions and administrations, I am interested in, and that is why I really like to talk about this area, Barcelona becoming in some way one of the capitals of the debate on technological humanism or technoethics.

We have several things in our favour here: we have very important research centres in areas such as artificial intelligence, robotics, photonics, and supercomputing, among others; we have the Mobile World Congress, as you know, and we also have significant political energy that can be channelled in many ways, and one of them is for Barcelona to play this important role. So, apart from my newspaper, as I said, *La Vanguardia*, the Catalan Ombudsman is also committed to this idea of making Barcelona a benchmark, one of the world's benchmarks in this debate, and also institutions such as the Digital Photo Society, the Barcelona City Council and the Catalan and Spanish governments are all in favour of this goal.

So, what we are going to talk about today in this debate is above all the legal framework, about what this reality is, this future, as it was very well explained in the previous talk, this reality that is already the past, the reality, and the future of artificial intelligence. There is no need, at this point, after all the debates that have taken place since yesterday, to go into further detail about what we mean when we talk about the advantages of artificial intelligence and when we talk about the problems, the risks, the threats that artificial intelligence involves.

This debate today poses an important and interesting challenge: how the European Commission and the European Union in general, the member countries, should participate, how they should keep up with this technological race around artificial intelligence, avoiding the risk of creating a fragmented market that makes European companies less competitive, and at the same time maintaining these European standards, which are much higher than those of other advanced economies, European standards that guarantee their citizens' rights.

To answer these questions, we have Martha Stickings, Political Analyst and Head of the Institutional Cooperation at the Fundamental Rights Agency, a European agency; we have Geraldine Mattioli-Zeltner, who is Adviser at the European Commissioner for Human Rights; and we have Peter Bonnor, who is Senior Legal Officer at the European Ombudsman. We were also counting on the participation of Gregor Stojin, Chairman of the Ad-Hoc Committee set up by the European Commission on Artificial Intelligence, but he has excused his presence and has sent us a video of his first speech. We'll watch it first, so that we can then move on to the human debate.

Gregor Stojin

President of the CAHAJ

Hello and greetings from Slovenia.

I'm sorry that I'm not able to be at the workshop in person, to greet you in Barcelona, but I was supposed to fly through Venice Airport and, due to the measures because of the recent outbreak of coronavirus in Northern Italy, this could cause some complications when I want to return, and also when I want to go to some European institutions, because they are limiting access for travellers through Northern Italy.

Be that as it may, I am also relatively uncomfortable of speaking via a pre-recorded video message, due to not having the interactive possibility of discussing the matter with you. So I'm not sure what you have already heard on the artificial intelligence during today's discussions. I have looked at the agenda and it's very interesting, and I would have wished to have participated. I will try to present the outline of what we're doing at the Ad-Hoc Committee on Artificial Intelligence at the Council of Europe, why we are doing it, what we are trying to achieve by doing it, and what are the expected outputs.

Currently, it seems that much of the global development in the field of artificial intelligence is based on self-regulation. Self-regulation in turn is based on trust, and trust has been breached a number of times in recent past when it comes to artificial intelligence and its effects on human rights. We need to re-establish trust, also as a condition for a more willing and responsible use of the technology by the authorities, and through this, for the encouragement of further development of responsible and trustworthy technology and its applications that could enhance the benefits for the society.

But what do we mean when we say "regulation"? Lack of regulation represents regulation by itself. We could substitute the word with "policy" or "governance" or just limit it to legislation, but what is more important to take into consideration? We also need to understand the autonomous impact of market forces, social norms, and technology on the process.

Not everything that is feasible from a technological perspective can be considered acceptable from a human rights perspective. Currently, many national, international and supranational bodies are developing or assessing various types of regulatory mechanisms to address the issues posed by the emergence and proliferation of AI technologies.

At the moment, there are more than two hundred soft-law instruments. Is the current legal framework adequate? Can non-binding documents, such as recommendations and guidelines effectively protect the desired levels of human rights? There are some issues that are worth keeping in mind to achieve the most effective solutions that would adequately and appropriately balance the beneficial aspects of technological development with simultaneous protection of humans and the society from negative aspects, both intentional and inadvertent. Impacts that are contrary to the existing human rights standards already fall under the preventative responsibility of states. However, are they effectively addressed?

Regulatory powers are primarily a competence of states. Within technologies and activities with significant extra territorial or even the global effects common rules are desired in order to ensure clarity and to prevent the race to the bottom by putting too much emphasis on technological possibilities. International and

supranational organizations to which such work is often delegated can and do differ in their jurisdiction, competences and mandate and can produce documents and norms with different influence or binding strength, which is relevant, for example, to enforcement in individual cases. Council of Europe is one of the few that can produce legally binding instruments. Different approaches should not be considered in competition, but as complementary contributions to yet developing frameworks. For example: common definitions and classifications are most valuable to ensure clarity and potential legal certainty and to allow a common platform for assessment of risks and opportunities. Duplication, on the other hand, can cause confusion and potential divergence in implementation practices, and should be avoided unless well substantiated.

We are all probably familiar with the examples relating to criminal justice for assessment of offenders or recidivism, which have shown that the bias in data used for training models is replicated and scaled when applied by the technology to real life scenarios.

Problems can also arise from algorithms themselves, either due to the work of designers and programmers or due to the nature of the decision making process. This is probably the most AI relevant aspect which still looks for an adequate solution - black boxes not being transparent enough or not showing or not allowing adequate explainability and consequent legal scrutiny.

Sometimes, models are based on assumptions that have no or low scientific value, such as assessment of emotions based on image recognition. Sometimes, such tools are used in a private setting, such as in human resources for vetting candidates, either for employment or for termination of employment. Sometimes, it is also impossible to know if AI was used in the process at all. Obscurity can come on different levels.

Other types of problems arise from unscrupulous use of technology without adequate risk or impact assessment, with facial recognition technology being one prominent example. It might be that there are existing legal instruments that can protect us, but the question then becomes whether they are effective enough, whether they are responsive enough. Such were, for example, also findings regarding challenges to existing tort law regimes in a recent EU study on liability for artificial intelligence. And if you allow me, I'll just read and excerpt from the study. It says: "It is possible to apply existing liability regimes to emerging digital technologies, but in light of a number of challenges and due to the limitations of existing regimes, doing so may leave victims under- or entirely uncompensated. The adequacy of existing liability rules may therefore be questionable, considering in particular that these rules were formulated decades or even centuries ago, based on even older concepts and incorporating a primarily anthropocentric and monocausal model of inflicting harm".¹

Another example for application of existing legal mechanisms was also made by a recent Dutch decision on the so-called "SyRi" automated learning program, which was used by the Dutch authorities to predict which citizens are likely to commit some form of housing or welfare fraud. The ruling, which came out at the beginning of February, is landmark because it was based on the right to a private life that's set out by article 8 of the European Convention on the Human Rights. The court found problematic the lack of transparency about how the algorithmic risks scoring system functioned and determined that the SyRi legislation fails a balancing test, which requires that any social interest should be weighted against the violation of individuals private life with a fair and reasonable balance required.

¹ See Liability for Artificial Intelligence and other emerging digital technologies, Report from the Expert Group on Liability and New Technologies, European Commission, p. 19; <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>

The automated risk assessment system however failed this test in the court's view. It was a black box and, as such, not legal.²

Different stages of AI system life cycle can produce different risks and implications, with various levels of severity, both in terms of actors as well as stakeholders. Also, the definition, the technology, the impacts, the use of the technology, they are all a moving target. Clear understanding of effects, impact and risks requires both classification and categorization of AI systems and applications, their continuous monitoring and an interdisciplinary assessment of findings. This can eventually lead to either verification, validation, certification. Human rights impact and risk assessment seem almost necessary and would require not only a technical analysis but also various humanistic knowledge and input. There are various options that can be applied and it is clear that there cannot be one solution to fit it all.

So, how do we go about it? CAHAI, which is an acronym for Ad-Hoc Committee on Artificial Intelligence, has been tasked by the Committee of Ministers of the Council of Europe to prepare a feasibility study of a legal framework and its potential elements for design, development and application of AI in the context of existing Council of Europe standards for protection of human rights, democracy and the rule of law. It had its first plenary in November 2019 and it is expected to produce the feasibility study by the end of 2021.

During the first stage, work at CAHAI is being done on identification of risks and opportunities of artificial intelligence considering human rights standards and on mapping of binding and non-binding instruments. The goal is to produce a comprehensive overview of the existing legal framework and the evaluation of its capabilities to address the most pressing effects. This can eventually lead to identification of gaps and areas which need to be addressed anew.

Council of Europe has substantial mileage on the wider aspect of the subject. In addition to the European Convention on Human Rights, it is the home of the conventions 108 and 108+ on data protection, which are sometimes considered grandmothers of GDPR, and it's also home of the Budapest convention on Cybercrime. There are currently more than forty other initiatives at the Council of Europe, including, for example, the recommendation of the Commissioner for Human Rights, who produced the document "Unboxing AI: 10 steps to protect Human Rights".

Also CEPEJ, the European Commission for the Efficiency of Justice, was in 2018 the first to adopt the ethical charter on the use of AI in judicial systems.

What could also be considered partially relevant, and historically relevant? We have an example from 1964 with the convention that established the European Pharmacopoeia, which is an institution that verifies the suitability of medical products and procedures which are put on the European market. Prior to that, regulation of medical products was decentralised and it was in the public health interest to establish such a mechanism. It might be so that this is the path that we also have to take with modern technologies, such as artificial intelligence. Of course we are not dealing with a static product, such as tablets or pills or medical procedures. We are dealing with a dynamic system. But there are some aspects that are similar, for example, we are looking at: Does the system do what it says it does? Does it also do something else? What is it composed of? What are its effects? What are its side-effects? What are the indications? What are the counter-indications?

What we want to do is identify problematic artificial intelligence applications and effects from a perspective of ensuring respect for human rights, rule of law

² See "Blackbox welfare fraud detection system breaches human rights, Dutch court rules", Natasha Llomes, *Techcrunch*, 6 de febrer de 2020; <https://tcm.ch/376Jq8a>

and democracy. We want to identify the key values and principles which are arising in a bottom-up perspective from sectorial approaches and from ethical guidelines and then from top down: fundamental rights, democracy and rule of law requirements. There can be red lines which seem that should not be allowed to be crossed, for example: AI powered mass surveillance, social scoring as we know it from certain parts of the globe, deep fakes and their effects on democracy. How do we address that? And most likely prohibition or not allowing the discussion on legal personality of AI. And then, we want to set down possible options of a Council of Europe legal framework by taking into consideration the contents, addresses, added value and member states' and multistakeholder's expectations on the various options that can be used. And these can be: an update of existent legally binding instruments, it could be a new convention, it could be a new framework convention, it could be soft law instruments, such as: recommendations, guidelines. It could also be other types of support to member states. We also need to look at possible complementarity between the horizontal, transversal elements which could be part of a Convention-type instrument and the vertical, sectorial world which could give rise to specific instruments of a different nature. Also, what is of crucial importance: the feasibility study will have to address operational mechanisms to ensure effectiveness of the legal framework.

Work is done through broad multistakeholders participation. In addition to forty seven member states, we have a number of representatives from other bodies of Council of Europe, but also other countries that serve as observers to the Council of Europe, non-governmental organizations, civil societies and also from business.

It is my opinion that is important to build on the existing work, especially in terms of developing a comprehensive and mutually acceptable body of definitions and classifications. This is why we also include various international organizations as participants and observers. It is, however, too early to predict whether approaches will eventually converge, and to what degree, but we should prevent distraction and we should prevent delay. We should prevent race to the bottom by, if nothing else, trying for too long or, on the other hand, if we only try to search for a minimum consensus. This probably would not be enough for what our challenge is about.

What is required in the long run when it comes to AI impact on human rights is: First: acceptance of political responsibility for unregulated risks and impacts.; Second, awareness of possible policy instruments, and third, regulatory proportionality and agility when using them, which requires (fourth) constant, continuous and objective assessment of mechanisms' impact and effectiveness. So, this is the broad outline of what we are trying to do at CAHAI. We are still at the early stages of the mapping exercise, but hopefully we will be able to contribute to the development of the field and to the protection of human rights, rule of law and democracy.

With this, I conclude my message to you and I wish you very good work in the continuation of the workshop.

Martha Stickings

Policy analyst, Fundamental Rights Agency (FRA)

Thank you very much, chair, thank you also to the organizers for inviting the Fundamental Rights Agency to participate in this important event. It's really a great pleasure to be here and to be able to talk a bit about some of the work that we have been doing on artificial intelligence and the fundamental rights considerations that come into play as part of that discussion.

Just a few brief words on the organization that I work for. We are an independent agency of the European Union with a mandate to collect data and provide evidence and analysis to the EU institutions and the member states when they implement Union law. Like everyone else, I think, artificial intelligence and the broader digital agenda is an ever more important part of our work and we have several initiatives that I'll tell you a little bit about where AI in particular is a focus.

Before I go on, I should just say briefly that as you heard from my introduction we are an independent agency, so I can't speak to the European Union's position or the Commissions' position but what I'll focus on is some of our views from a fundamental rights perspective, in terms of some of the main concerns and considerations in this discussion.

To start with this, perhaps it is useful to reflect quickly on what the current state of play is. We've heard from lots of the speakers already at this event that there are many different policy developments that are on the way at the European level, and we see that fundamental rights considerations are coming into play as part of many of those discussions. There are far too many of these initiatives to mention and it's perhaps worth highlighting a few which give an idea of how discussions around AI came onto the policy agenda.

In 2017 the European Parliament adopted two resolutions which highlighted the importance of big data, AI and robotics and also referred to the importance of fundamental rights. There were also Council conclusions in 2017 after which the European Commission published the communication on AI for Europe and set up the high level expert group on AI that we've already heard a bit about. Both of these initiatives include a strong reference to fundamental rights and one of our heads of unit is a member of the high level group. Then, of course, most recently in February, there was the publication of the Commission's white paper on Artificial Intelligence which you have already heard quite a bit about. But it's worth also recalling that it was just one element of a broader package of documents that was released in February and there was other reports on liability issues, there was a European data strategy, and it's useful, I think, to put the AI paper in that wider context.

At the Council of Europe level you have already heard about the CAHAI Committee but there is also the work of CEPEJ and other committees, and I'm sure you will hear more about that later.

Given the number of initiatives that there are, we have been trying to keep up with this proliferation and we have a list on our websites of many of the different initiatives from the EU and beyond, which currently amounts to 216 odd, but it's increasing all the time and we do regular updates of that overview.

Well, some of the recent activities do make explicit reference to fundamental rights and we have to also acknowledge that many of these initiatives prefer softer ethical frameworks and we have heard a lot about ethics in the discussions so far here. It's, of course, quite attractive to focus on ethics and it has strong links to trust and again we've heard a lot about trust of consumers in products and it's seen

as encouraging innovation and being supportive of business models to take for advantage of the single market and, of course, we know that there's a competition between what is happening at the European level and other developments in the US and China for example.

From our perspective we would be a bit wary of the use of ethics, at least in isolation, on the basis that they give connotations of voluntariness and that they are not enforceable by legal regimes and that they are not necessarily standardised, so they might be sort of different ethical models that apply in different contexts. As a result from the perspective of the Fundamental Rights Agency, we very strongly encourage the use of fundamental rights as a framework as the starting point for the evolution of any law and policy concerning AI.

That's partly because, in contrast to ethics, human rights are legally binding instruments. Also the human rights put the clear responsibility on states, because states have the ultimate responsibility for protecting and promoting human rights, also because we have a great deal of experience in terms of using human rights to ensure protections and using that legal framework to support subsequent legislation but also redress. And here I think it's worth reflecting on the fact that human rights have already been used as a very useful framework in some related areas. So, for example, we can think of the guidance on business and human rights, and that some of those principles that have been established in that context can also be brought into the discussion around AI.

And lastly, the advantage of the human rights based approach is that it allows us to focus on the particular harm that is being done to individuals, and you see a reference to that also in the Commissions White Paper.

In terms briefly of some of our work on artificial intelligence, we are trying to support some of the policy processes that I mentioned earlier, with evidence and advice that follows this fundamental rights based approach. We published three papers so far, the first of which was looking at discrimination in the context of automated decision making, the second that was looking at the issue of data quality. Both issues that we have heard quite a lot about in the last day and most recently a paper on facial recognition technology, which is a particular use of artificial intelligence that can really cause quite significant fundamental rights concerns. These papers are part of a larger project on the use of artificial intelligence in EU member states and the results of that work will be published at the end of this year. The intention behind that research is to really be able to identify practical use cases of where AI is being used across the different sectors in the economy, so whether that be employment, retail, insurance, the public sector, specifically law enforcement, etc., and to really understand the extent to which fundamental rights are part of the discussions when AI tools are being developed and deployed.

Lastly, as I mentioned, we are also actively involved in some of the policy processes that are ongoing, both at the Commission level and at the Council of Europe.

The work that we've done so far really underlines the full range of fundamental rights that can be impacted by artificial intelligence and different related technologies. So we heard a lot in these discussions about data protection and privacy and also increasingly about non-discrimination, and of course both of those rights are extremely important in any such discussion. But it is also really important that we think of the full range of rights that can potentially be impacted, and really it is hard to think of rights that can't potentially be impacted by artificial intelligence, so that includes: freedom of expression, freedom of information, freedom of assembly, but also gender equality, the rights of particular groups and societies, the rights of the child, the rights of older persons and the specific situations that might pertain to those groups. I guess in the context of Ombuds organizations is

particularly important to think of access to justice rights that are involved in any discussion of artificial intelligence and also the right to good administration.

It is also important to recognize that there are going to be questions about balancing of rights, and we got some hints in that direction yesterday. So, for example, the question of the right to privacy, and data protection and the right to non-discrimination. So if an algorithm is checked to see where it discriminates against particular religious groups, then information on religious affiliation is going to need to be collected to determine whether that discrimination is taking place.

So, what does that mean for any potential future regulation of artificial intelligence? Our research suggests some elements that could be desired foundations for future law and policy? Here the first thing to say, and I think our colleague from CAHAL also reflected on this, is that in Europe we already have strong legal framework, so we have to acknowledge that we are not starting from zero, that there are already strong protections in place. The Charter of Fundamental Rights covers many different rights which you saw on the slide before that can be impacted, and of course those are binding on the EU and member states when implementing Union Law. Also, we have the GDPR and we have the existing quality directives. So for example: if the word 'discriminatory' outcomes in a recruitment tool using AI, then that would be illegal under EU law already. It is also really important to address the full range of rights affected, so we need to make sure that any discussion really looks to try to map the implications of any particular AI usage on the full range of rights.

Thirdly, I think it's very important that we have concrete use cases about how AI is actually being deployed in practice, and that's partly because a lot of the discussion at the moment can often remain at quite a theoretical level, and we don't necessarily have the knowledge and information about some of the practical ways in which AI technologies are already being used, and, in that respect, the session yesterday I think was really useful in terms of giving us a sense of some of the usages that are already in place. And this allows for a bit more of contextualised discussion about fundamental rights implications of any use of AI. So, for example, in our recent report on facial recognition technology, we are highlighting that there is actually non one type of this technology, but it depends on how it is used and the fundamental rights implications would change accordingly. So it's very different if facial recognition technology is used for verification purposes, so, for example, if you go through an automated passport gates at the airport, that is identification, so one to many, which will be the usage in terms of live facial recognition, for example, for security purposes.

Fourthly, the question of literacy for people involved in and affected by the use of artificial intelligence is particularly important. Often there is just not enough knowledge about the technology, about the way it might be used and about the particular concerns that it might raise. But that literacy goes both ways, so it's a question of people from the technical side being more literate also of fundamental rights standards and concerns, as well as people coming from a fundamental rights perspective being better able to understand the technological elements.

Lastly, it is also very important that we ensure cooperation across different areas and disciplines. Again I think that's a point that we've heard in various different ways at this session but really ensuring that we bring together these different voices is particularly important, and there perhaps it is also worth adding that, as these technologies are increasingly rolled out, the hearing of the voices of people who may have been negatively affected by decisions made by AI would be a particularly important part of that discussion.

And finally, I just want to close by thinking of some aspects that there will be useful for Ombuds organizations to consider as part of this discussion. It's interesting to

know that in the Commission's White Paper it does highlight risks of fundamental rights as some of the most significant challenges that are posed by artificial intelligence, and the White Paper also highlights that there are very real risks that it will be more difficult to seek redress to and to get compensation for victims, relative to victims of harms caused by more traditional products and services. And so, one of the issues is whether people are actually aware that the data is being processed by AI tools. Of course, the right effective remedy requires that you actually know what is happening, and that's not necessarily going to be the case. So, thinking about how people can be more aware of how that data is being processed is particularly important.

Secondly, we have heard about the challenges posed by the opacity of algorithms and automated decision making and that, of course, links also to questions of complexity and, again, the white paper addresses that in some detail, looking at not only whether we can actually understand what the algorithm is doing, but also whether we could, for example, be able to get an access the data that would be necessary to be able to build a case in the instance of harm being caused by an algorithm.

Thirdly, there is the question of the allocation of responsibility. There are different fundamental rights that can arise at different stages in the process of designing, deploying, using algorithms, and it's really important to kind of think about what those risks might be and what that would mean ultimately for who might be responsible for harms that are caused, and there we see that this accompanying report on liability from the Commission goes into that question in quite a lot of detail.

And lastly, there is a question of scale and the number of people that are potentially affected by the use of algorithms, and I think we heard a reference at the very beginning yesterday morning, of someone saying of the impact of the GDPR on resources for bodies that might be handling complaints related to the GDPR and given the scale of that application or potential application of AI technologies, having the resources to be able to ensure that people have redress is going to be particularly important.

So my final call there would be that perhaps this and some of the other discussions that we have had can be some food for thought for Ombuds organizations in terms of thinking about your contribution to the consultation on the Commissions' White Paper. It will be really important to get the perspective of national human rights bodies as part of that discussion, because I think you have very relevant experience and perspectives that really ought to be heard as a part of this process. Thank you very much.

Geraldine Mattioli

Adviser at European Commissioner for Human Rights (Council of Europe)

Thank you very much and thank you for inviting us. It's really a pleasure to be here to present the work of the Council of Europe Commissioner for Human Rights on artificial intelligence and human rights. For those of you who may not be very familiar with the Commissioner's institution, it is an independent and impartial non-judicial organ of the Council of Europe established to promote awareness and respect for human rights in the 47 Council of Europe Member States. Despite the "Commissioner" title, she is not a member of the European Commission at the European Union. It is important to differentiate between the European Union and the Council of Europe. The Commissioner's name is Dunja Mijatović and she took up the office in April 2018.

My presentation will have two parts. First, I will broadly present a Recommendation on Artificial Intelligence that the Commissioner published last year, which was mentioned

earlier by Mr. Stojin, and then I will zoom into the points in the Recommendation that specifically touch on the possible role of Ombuds institutions and other national human rights structures regarding artificial intelligence and human rights.

Since she started her mandate, Ms. Mijatovic identified Artificial Intelligence as one of her priorities. The reason for that was that she recognised very early on that Artificial Intelligence will impact each and every of the human rights she is mandated to monitor and protect throughout the forty seven member states of the Council of Europe. Thus, it was very clear to her that this development would be fundamental in all of the work of her Office. The approach of the Commissioner can be characterised as follows: it may indeed eventually be necessary to identify the gaps in the existing legal framework and to adopt new instruments to regulate Artificial Intelligence. But this is going to take time, and in the meantime the technology is being rolled out. Therefore we need to start now looking at the human rights implications of Artificial Intelligence, and we need to start with the tools we have. Let's not forget that human rights law, discrimination law, data protection law are very powerful tools, which are already at our disposal and which we should leverage to ensure that artificial intelligence is deployed in a human rights compliant way.

Given these considerations, the Commissioner published a Recommendation based on existing standards last summer. It is a very short document. I brought some copies and encourage you to take one. The Recommendation identifies just ten steps that national authorities should take to make sure that Artificial Intelligence is compliant with human rights. In other words, the Recommendation outlines a human rights-based approach to using Artificial Intelligence.

I am not going to mention all of these ten steps now, but let me illustrate a few particularly relevant ones. If a public authority is thinking of using an algorithm along the lines of the examples we heard yesterday, that is for social welfare benefits, facial recognition, police predicting, then there are a number of things that these authorities should undertake before doing so. The recommendation suggests, for example, to develop human rights impact assessments. If the use of the algorithm is likely to have a possible impact on the human rights of users, then the authorities should assess this possible impact beforehand. They should also repeat this impact assessment once the algorithm is being used to make sure that no human rights violations are taking place.

Another step suggested in the Recommendation is for public authorities to inform the public and be transparent about their use of algorithms. Currently, most of the time, people do not know that they are the subject of an automated decision, they are not told that is not a real person who actually reviews their application for social welfare benefits. This should not be the case, transparency is very important. In addition to their own duties, public authorities also have a positive obligation to encourage businesses to implement human rights standards when they are developing algorithms.

Another point mentioned in the Recommendation is the need for oversight. We heard a lot about that during the panels yesterday, too. Keeping human oversight over Artificial Intelligence systems is really important, and such oversight should be carried out independently from the businesses developing the algorithms and from the authorities using them. How exactly to develop this oversight? This is a big question we are going to have to tackle. Finally, the Recommendation also covers a point stressed by Martha mentioned a moment ago, states have the responsibility to develop AI literacy among the population. In conclusion, the Commissioner's Recommendation aims to be a road map for national authorities so that they can avoid that Artificial Intelligence is deployed in a wild, uncontrolled way.

The Recommendation mentions at several moments the possible role of National Human Rights Structures, by which we mean Ombudsman institutions, but also National Human Rights Institutions, Equality Bodies, Data Protection offices, i.e. all

the bodies at the national level that already now have the mandate to control that human rights are respected. This is the second part of my intervention: I am going to present the possible roles that are envisaged for Ombudsman institutions in relation to Artificial Intelligence in the Commissioner's Recommendation and I really look forward to hearing your reactions and your thoughts about these ideas.

At the outset, I want to underscore that we are aware of the fact that Ombudsman institutions can be very different, have different mandates depending on the countries. Of course, the possible functions would depend on what these Ombuds institutions do in each of the different countries.

But, in general, I would like to invite you to think about the role of Ombuds institutions in four broad different categories. First, what can they do in the field of "prevention" - to prevent human rights violations by Artificial Intelligence? Second, what can they do to increase sensitization of the public and educate people about these issues? Third, what can Ombuds institutions contribute in terms of oversight? And lastly, what can they in terms of access to justice and redress? For each of these four categories, we see two types of activities: what I would call the "classical" Ombuds activities and then some newer things to consider, specifically in the context of Artificial Intelligence.

By "classical" activities, I mean the things that Ombuds institutions are already doing every day in their role of controlling respect for human rights, but applied to the use of Artificial Intelligence. For example: Ombuds institutions regularly review draft legislation and public policies to ensure that these are human rights compliant. This can be done for legislation and policies concerning Artificial Intelligence. We heard yesterday that there is a charter of digital rights in Barcelona. Was the Síndic consulted? Did it review the charter? That is an important question. Right now, many states are preparing national strategies on Artificial Intelligence or are adopting new legislation in relation to Artificial Intelligence. It is essential to make sure that Ombuds institutions review these documents to ensure they are human rights compliant.

Other types of "classical" activities include: launching a public information campaign about Artificial Intelligence and human rights, or launching investigations into allegations of systemic violations that may be happening in the context of the use of Artificial Intelligence systems, publishing reports, publishing recommendations to the authorities, and then, of course, handling individual complaints, when they are filed – that is the core of the Ombuds institutions' mandate. Ombuds institutions could also provide legal assistance to victims, and conduct strategic litigation as it may not be possible to deal with each and every complaint if there are thousands of them.

What I just said may seem obvious but it is important to recognise that the work of Ombuds institutions remains the same, irrespective of whether the human rights violation is committed by an algorithm or by a person. It is crucial to recognise that Artificial Intelligence is not a science fiction concept, it is already in our societies, it is already happening. One first step to consider is for Ombuds institutions to conduct a mapping of Artificial Intelligence systems used in their countries. The Swedish and German Equality Bodies have launched such studies that will enable them to better understand the possible human rights impacts and risks of Artificial Intelligence at the national level.

Beyond these "classical" activities, I would like to mention briefly two specific roles. First, Ombuds institutions could be involved in the human rights impact assessments of Artificial Intelligence systems, which I mentioned at the beginning. Indeed, Ombuds institutions have a lot of relevant experience that can be harnessed to help the authorities design these impact assessments. Ombuds institutions can provide guidance about what questions should be asked, who should be consulted. Maybe Ombuds institutions should even be involved in conducting the impact assessments. This is something to consider.

The other idea concerns the issue of independent oversight. The Commissioner's Recommendation is purposefully vague about what such a mechanism could look like. There are a lot of questions about who should be responsible for such oversight. A certification mechanism? A public institution or private institution? Should different bodies be involved in the oversight, bringing diverse expertise and perspective? Or should we go towards the creation of a specific oversight mechanism for Artificial Intelligence, on the model of data protection offices? Whatever model is chosen, the Commissioner for Human Rights believes that it is crucial for the oversight mechanism to include human rights expertise, including through the involvement of Ombuds institutions.

I want to finish with two positive points. First, the Commissioner is very encouraged to see that several Ombuds institutions have already started working on Artificial Intelligence and human rights. There are Ombuds institutions that are leading the way and doing important work. I can give you some examples: in France, the French defender has taken several cases linked to the use of Artificial Intelligence, for example regarding recruitment in universities, or the use of algorithms in the judicial system; in Finland the Ombuds has won a strategic litigation case about age discrimination by algorithms; in Slovenia the Ombuds brought a constitutional case regarding automated license plate reading and privacy law. I also had a discussion with Ms. Catherine de Bruecker yesterday, who gave me several examples in Belgium where the Ombudsman is looking at automated decision-making processes. There is a lot going on and this conference is another demonstration that Ombuds institutions are seizing the challenge. That is very positive.

The second positive point I want to mention is to not forget that Artificial Intelligence can actually be used to advance human rights. Artificial Intelligence will help us to document and demonstrate instances of discrimination, for example, in a way that we were not able to prove before, through the use of big data. The Commission for National Human Rights in Ireland, for example, was able to use an algorithm to analyse hate speech online on Twitter in the specific context of Ireland and to better understand where it was coming from. There are important potentials for Artificial Intelligence to help in the human rights work of Ombuds institutions, which should be kept in mind.

The Commissioner is planning to continue her work on Artificial Intelligence and human rights. The next step will probably be to try and look at specific uses of Artificial Intelligence in several countries of the Council of Europe. The Commissioner already examined the impact of Artificial Intelligence on the human rights of older persons in Estonia in a report published in September 2018. Members of the Commissioner's office also conducted exploratory missions to Sweden, to look into the use of Artificial Intelligence in social welfare, and to Serbia, to examine the deployment of facial recognition technology. I want to close by encouraging everyone in the audience to be in contact with the Commissioner's office if you consider that the use of Artificial Intelligence in your country raises specific human rights issues. Thank you.

Peter Bonnor

Senior Legal Officer, EU Ombudsman

Thank you for giving me the floor and thanks for the invitation.

A few words about the European Ombudsman. The European Ombudsman is the Ombudsman for the public administration of the European Union, and what is that? That is the institutions that we know from Brussels, the European Commission, the European Parliament, the Council and the large number of agencies that we have all over Europe, specialised in different fields, like fundamental rights, or which are

involved in regulation of health and safety, safety of chemical products, safety of medical products and so forth. The size of the European Public Administration is somewhere between fifty and sixty thousand people. That is a very small number of persons for a public administration for the European Union. It's a number of people that more or less corresponds to the number of public servants of a mid-sized European city. The reason for that, of course, is that we don't have very much direct administration, we don't have police officers, and we don't have a European Commission hospital or anything like that.

The European Union Public Administration is a regulatory machinery. That is very important in this context.

The European Ombudsman office has around eighty staff. We are twenty-five to thirty people involved in the day-to-day investigation work. We have so far not received complaints related to artificial intelligence, we do not have initiative in this field, and I think I'm by now the most AI-knowledgeable member of staff. So, in what ways are we relevant?

I want just to flag two fields in which the European Ombudsman can be quite relevant in relation to artificial intelligence and probably will be.

The European Union, although it is relatively small, distributes a lot of money. It is without doubt going to pour plenty of money into this field of artificial intelligence. Through this activity, there will be call for tenders, there will be call for proposals, there will be contracts, there will be consortia with dozens of partners located around Europe. Experience shows that there will be problems, there will be misunderstandings, and there will be situations where the European Commission says to someone: "No, you can't have your money because you haven't given us the deliverables" and then we have a crisis for a project. And these are the situations where the European Ombudsman can then come in and work as a kind of can-opener and help people through assessments based not only on law but also on fairness and good administration. So, that is relevance at a practical but still relatively narrow level.

Then, on another level, remember that the current European Ombudsman, who was re-elected for five years period in December last year, has taken the office's approach to a level that very much takes into account the fact that the European Public Administration is a regulatory machinery. The European Ombudsman cannot examine and look into legislation as such, but it can get rather close to the regulatory processes. We've heard about ethics and AI. We are looking at ethics in regulation at EU level. We can look at whether there are situations of conflicts of interests in the processes that are leading up to the regulatory rules. We can look even at whether there is an appropriate balance in, for instance, expert groups, workshops, conferences, that kind of thing, if people then flag to us: "Listen, the European Commission is only inviting industry people to participate in this or that activity", then we can go and have a look at it. (To illustrate, I think the first conference panel we saw yesterday was quite balanced, so that was a good example of a balanced group on the issue.) The ombudsman has carried out such assessments in a number of fields, including agriculture, really going in and checking who are members of these expert groups that are going to feed into the policy process that will ultimately lead to EU legislation.

In the EU, we have two very powerful and competent commissioners in Brussels in charge of this area of artificial intelligence, digitalization, the whole digital strategy and things may soon be moving very fast.

I cannot but emphasise, like my colleagues in this panel, that, please, do take the current opportunity of feeding into the policy process by replying or contributing to the public consultation organised by the European Commissioner. You just have to

google it "European Union public consultations" and you will find a list of open public consultations on artificial intelligence and on data strategy. It requires a little bit of homework reading the two white papers, in particular on artificial intelligence and data strategy, but feeding into these processes is so important. It's not a box ticking exercise, it's not just a formality, it is something that is being taken really serious. And it's part of the culture in Brussels, it's part of the culture at the European Union level to do public consultations and carefully examining the input. We even do that as ombudsman office. We have done several public consultations on different issues and it has been interesting thing to see that we get contributions from just about any side, from individuals, from researchers, and even from governments or EU institutions. So, it's a really rich exercise that will influence what will be put in the legislation. One can even think of some of you getting together on an ad hoc basis, making a contribution as a group.

Then, I would like to say one or two words, and this is just my own thoughts to bridge with the next session: What kind of tools can we, the participants, bring back?

I think that the think-tank presentation that we saw yesterday really hit the nail on its head in one of the slides that contained a reference to public procurement. It is true that we have a general mandate as Ombudsman, we can investigate a lot of things, but let's not forget the limits that we also have: we cannot instruct the public administration to do something, we can only recommend, right? It is therefore very important for us to be in the field very early on, in order to influence things, this is just a basic rule of negotiation.

At the European Ombudsman office we don't have experts on artificial intelligence and I think most public administrations in Europe will not have real experts on artificial intelligence. They will be employed by big companies. Salaries are better, the opportunity for being creative will be greater, and so forth. So, there will be an issue of just having to follow the money. Where will the money go? There is an opportunity for the Ombudsman right now, to sit down, talk to the public administration and say: "I would like to know right now what you are planning to do" and "I want to go and have a look at what you are going to put it into the functional and technical specifications for the tenders that you are going to publish". And then, taking the standards that we have heard about during for instance this conference, examine the public tenders and the assessment points that are given to the companies that then reply to these public tenders. Must the level of AI-transparency good? Must the AI-related testing be at a very high level? etc. So, to some extent it's the old rule: follow the money trail when you want to do efficient investigations.

Going back to the public consultation that is currently ongoing here with the European Commission, here one or two ideas about what one could consider proposing.

We hear about lots of things that went wrong somewhere and that people inside knew about it. We know it from dieselgate, we know it from the scandal that Boeing currently has, that people from the inside knew about the things that were not going well. I haven't seen in the white papers any high standard being proposed for whistle-blowers, whistle-blowers for instance for software programmers. If there is a software programmer somewhere who knows that something is going the wrong way, it might be a good idea for that person to have a safe opportunity to go and blow the whistle to the right body and with the right protective mechanisms.

Then, to round off, I can only reinforce what has already been said, that AI is simply something that is happening, and that lots of the current standards already apply to this field. So again, going back to the tool box, what would be a concrete action to take? One is to make sure that we internalise regular checking of equal opportunities, data protection and so forth, and that we have the right networks that we can

access immediately. For instance, as an Ombudsman investigator, I would want to be able to access a network of AI experts, but if next week I get a complaint related to AI, I would want not to have to spend weeks or months doing a public tender to be able to draw on the relevant know-how. I can see in Barcelona you are very privileged, you are surrounded by people who know what you are talking about. I, as an investigator at the European Ombudsman office, I would at this point in time not really know whom to call if I have a case about AI. So creating networks that we can access immediately might be one important part of our tool-box.

Thank you.

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Tuesday, March 3, 2020

Ombudsmanship challenges, roles and tools

Chairperson: Joan Manuel del Pozo, Girona University Ombudsman

Robert Behrens, UK Parliamentary and Health Service Ombudsman

Elisabeth Rynning, Swedish Parliamentary Chief Ombudsman

Quirine Eijkman, Deputy President and Commissioner of the Netherlands Institute of Human Rights

Eric Houtman, Belgian Federal Energy Ombudsman and membre of NEON

Joan Manuel del Pozo

Girona University Ombudsman

Good afternoon ladies and gentlemen. We are going to start the table scheduled for this time, with a small delay for which we apologize, under the topic "Roles, challenges and opportunities for the ombudsmen".

First, a presentation of the table where we notice the involuntary absence of Mrs. Madise, Ombudsman of Estonia. Her contribution will be replaced by an initial video by Mrs. Elisabet Rynning, Parliamentary Chief Ombudsman of Sweden. And among the people present we have: Mr. Rob Behrens, UK Parliamentary and Health Service Ombudsman; Mrs. Quirinne Eijkman, Commissioner and Vice President of the Netherlands Institute of Human Rights, and Mr. Eric Houtman, Belgian Federal Energy ombudsman and member of NEON.

First of all, I would like to thank the Catalan Ombudsman for organising the conference and for inviting me to participate in this round table, and for the presence of all of you, which obviously gives meaning to the collective reflection that we are all engaged in.

In this table, we must summarise everything that has been reflected on, with a wealth of perspectives and with a very good contribution of information and criteria, and to try to talk about the roles we have to play, the challenges we face and the opportunities we should seize as ombudsmen, as a guideline towards the responsibility of the ombudsmen.

I would like to make a brief introductory remark on this, making it clear that, as it has been said in a previous intervention, I fully agree with it, I believe that we ombudsmen should never give up what is in our nature as ombudsmen. First of all, we are neither judges, nor legislators, nor rulers. We are defenders of citizens, and this is a counterbalancing function to the enormous weight of the administrations and the classic powers that we have just mentioned. And in this defence of rights, we have a democratically principal and essential role for which we feel proudly responsible.

It is true that, apart from all the possible violations of human rights that human history has unfortunately taught us, another problem is now emerging that can affect, as we have seen over the last two days, the human rights of the new societies. In this sense, I believe that we must keep very clear the role of restitution of any personal violation of human rights. This is our first and fundamental function. The citizen who comes to us and says: I feel disturbed, violated in some right, whatever it is, of privacy, of freedom, of equality because of, in this case, a misuse of artificial intelligence. This is our first essential mission.

The second one is to anticipate with initiatives, the so-called own initiatives, so that, aware as we are of the evolution of society, and in this case of the already very strong evolution of artificial intelligence, we can anticipate the possible infringement of these rights. In fact, that is what we are doing these two days. And I would go so far as to say that derived from this capacity for initiative that we have, we should take on another task, which has also been mentioned in a previous intervention, which I would call social pedagogy of the problem. This social pedagogy consists of raising awareness of the risk of artificial intelligence in relation to people's rights. We have a situation in which artificial intelligence does not look like an abominable beast that can be seen coming from afar, but artificial intelligence presents itself with a subtlety that is difficult for many citizens to perceive. So as ombudsman, we should contribute to explain to society that there is an ethic, that there is a need to protect our dignity, our freedom and our equality that could be affected by certain uses of artificial intelligence. And I would like to stress that in the ethical concern that we are having these days we should be aware that ethics is not imposed by anybody.

In other historical periods, ethics, or morality, was imposed by an authority, whether religious or political. No, today ethics is a necessity for our people and our societies. It is not a luxury, nor is it a fashion. In fact, since the 1960s, when one of the best-known ethics in its application to human life appeared, that is bioethics, applied ethics have been coming out in response to new social challenges. And so, for example, we hear talk of communication ethics, we even hear talk of business ethics. And I believe it is a good thing. But the fact that it is a fashion should not make us lose sight of the fact that underneath the fashion there is an ethical claim that comes with the history of humanity itself. There are values that are consubstantial to the human condition, that are also prior to our time and that obviously demand our permanent attention.

I would like to read to you very briefly fragments of two thinkers who are today interpreting our situation well, who tell us something interesting to reflect on. In this sense, I propose that our speakers give their opinion and say what they think about. The Israeli thinker Yuval Noah Harari, well known for *Homo Deus* or for the *21 lessons for the 21st century*, writes specifically in his last book a very short sentence that is a real challenge for our reflection: "When artificial intelligence decides more and better than us, our concepts of humanity and of life will have to change". In other words, he is telling us that the self-understanding of the human condition is in the process of change. Some people might, if they are pessimists, think of a change for the worse, they have the right to do so. Others, perhaps more optimistic, may think of change for the better.

In any case, we must be vigilant to change, because what as ombudsman we should not accept is that this change is made at the cost of some of the rights that we consider essential. And to put it in basic words, let's say: dignity, freedom, equality. If this change is to be made at the expense of these values, we should obviously offer some resistance.

The other thinker is Byung-Chul Han, of Korean origin but settled in Germany. You could practically say that he is a German thinker, of German philosophical inspiration. He says, with a quote that is a bit longer, but I think it is worth listening to it carefully. He says: "Intelligent, kind power does not operate openly against the will of the submitted subjects but directs this will in its own favour. It is more affirmative than denying, more seductive than repressive, it strives to generate positive emotions and to exploit them, it seduces rather than prohibits, it does not confront the subject, it facilitates. The present crisis of freedom consists in the fact that we are faced with a technique of power that does not deny or subdue

freedom but exploits it". In other words, we are faced, as I said before, not with a horrifying beast but with a subtle essence that enters our minds in such a way that we hardly notice it. Hence the call I made earlier for a certain social pedagogy, to raise awareness of the challenge facing everyone. We should not be complacent about certain misuses of artificial intelligence, we should not fall asleep, we should not be disoriented, we should not be clueless.

Well, in view of this, I believe that the challenges for the ombudsmen must go in many directions. I would like to mention some very important ones that have also been mentioned in previous speeches.

Firstly, attention, as ombudsman, to be aware of the biases in legislation and the emerging ethical codes. Bias is a risk, and we must be vigilant that in the construction of algorithms and in the construction of programming, there are no biases that end up distorting the human use of artificial intelligence.

And secondly, the challenge is to participate as ombudsman in processes of social reflection on the proper use of artificial intelligence. We could say that this meeting is an example of this, it is an example of a willingness to anticipate and participate in processes of ethical reflection.

And finally, and this is my last point, I think that the great opportunity we have, given that this meeting has been framed as a relationship between artificial intelligence and human rights, we have the opportunity to use human rights as a frame of reference that could be considered a kind of basic global code of ethics. Some thinkers have already referred to this and I believe it is useful to consider that, given the danger of a certain ethical relativism, that any ethics will do, and that one ethic can be changed by another ethic (as the comedian said: "Those are my principles, and if you don't like them, I have others"). In the face of this risk, let us have human rights as a kind of ethical claim, of a basic nature for all humanity.

This is the end of this modest moderator's presentation, and we now turn to the speeches of our speakers. First, Mr. Rob Behrens, the UK Parliamentary and Health Ombudsman. Excuse me, I am being informed about the Swedish Ombudsman's video.

Robert Behrens

UK Parliamentary and Health Service Ombudsman

I hope you can hear me, I am not a hologram. And this particular session will be participative, so expect to be involved in the course of what I want to say. I want to begin by thanking Rafael Ribó and his team for inviting me to contribute to this seminar. It's been a real learning experience.

I came here with a prepared speech and associated slides. I've listened very carefully over two days and I've therefore torn up both the speech and the slides, to your relief, because you would have heard what you have already heard and that would be boring and a waste of time. So, I've written a shorter presentation reflecting on what I've heard and focusing on the implications of artificial intelligence for my Ombuds office and others, the National Health Service in the United Kingdom and service users and practitioners in healthcare. And I want to end with a number of short emerging conclusions.

But I want to begin with a story about the world's first industrial city. So, who can name the world's first industrial city? Manchester, correct. How many people here have been to Manchester? Excellent. I knew it was a sophisticated, cultivated audience. More than half of you have been to Manchester, where my office is based. And the story I want to tell you is about what happened in 1830 when the private sector entrepreneur, George Stephenson, opened the first public intercity railway in the whole world. It was called the Liverpool to Manchester railway, and the station in Manchester has recently been restored from 1830 and is well worth a visit when you come to my office. The steam engine on this line was called the Rocket and it had an average speed, which was the fastest in the world, of twelve miles per hour and a top speed of thirty miles per hour. To celebrate the opening of the line, some dignitaries came along, including the celebrated rising, Government Minister, William Huskisson. Because it was new, there were no safety barriers and no one was used to the high speeds of the new train. Very tragically, Huskisson, overweight and recovering from serious ill health, couldn't move quickly enough and was knocked down and killed by the oncoming train. After he was killed, there were predictable calls in the newspapers for the banning of railways, the return of stage coaches powered by horses and for regulation for railway safety.

I suggest to you that this is a story of artificial intelligence in the United Kingdom, if not in the rest of Europe. The train, the Rocket, is a metaphor for artificial intelligence. George Stephenson is a representation of the companies that promote artificial intelligence, as we saw yesterday, reassuring us that everything is brilliant. And for me, the most important point is that the overweight, unfit politician who got knocked over is a metaphor for the Ombudsman at this time, the Ombudsman in 2020. In addition, after the tragedy no party talked to others about what had happened. They simply made assumptions.

In my view, and having sat here for two days, this is an existential crisis for the Ombuds. I am not a romantic. We cannot go back by denying the reality of artificial intelligence. We are captivated, enchanted by what artificial intelligence can do for us, just as Huskisson knew that a steam engine could revolutionise transport. The old world is dying and will not return. Artificial intelligence will not go away. And it has great advantages for the three sectors that I now want to briefly talk about.

My office, the Ombuds for the United Kingdom, has a budget of around 30 million pounds a year. We have 125,000 telephone enquiries from citizens about service failures, we get 30,000 complaints a year, and on the telephone we refer back 70,000 complaints. They are not in a condition for us to look at, because of the legislation, requires that they must be handled first by the front line body. 80% of what we get is

to do with our National Health Service, and avoidable deaths and serious failures in the health service. The mandate also allows for us to look at the clinical judgement of doctors in the health service, and so we commission independent advice from doctors and nurses.

So, artificial intelligence has huge potential to help us, as the Ombudsman in the United Kingdom, in terms of process, in terms of conversation, in terms of coming to judgements about clinical decisions and whether they are appropriate. But the truth is, that we in the United Kingdom are not yet comfortable with using it.

In terms of the evolution of the Ombudsman office, we have concentrated in the last three years on winning the benefits of not only updating our IT, but also face to face conversations with complainants after years of paper-based investigations. In this mode, people have said to us that 'you are not properly listening because you're not having a face to face conversation.'

One big problem for all Ombuds institutions is that their service users will not necessarily benefit from the introduction of artificial intelligence. This is particularly the case, and we've heard it already, for more vulnerable people, people with mental health challenges and old people, who are the least likely to use our service because we already requested them to use computers when they make a complaint to us. The challenge for all of us is to devise an Ombuds scheme which embraces those who are not comfortable with computers and information technology, and if we could do this, with the amount of time and resource we could save and involve people, this might be revolutionary for our service and our efficiency

We've thought about some real potential examples. Take the case of a colleague, Professor Stephen Perkins, who is a Board Director of the Local Government and Social Care Ombudsman in the United Kingdom. Two weeks ago, he was interviewed by a robot who was programmed to ask questions over half an hour and respond to Perkins' responses in the way that those who know say we could use with service users who either come to us or ring us up. The questions are: would that put people off? What value would it be and how would we know if it went wrong? What we do know, however, is that, in general, artificial intelligence speech and language capabilities, including translation, will speed up our understanding of what complainants are saying and help us to give a more efficient service.

We could also use handwriting recognition to convert the content of handwritten letters, removing the need for manual data entry. And there are lots of possibilities in artificial intelligence which we should not dismiss, merely because they are strange and they are new.

I turn secondly to the benefits to our health system, and we've heard some bold statements from the private sector about that. Some of that is true, and the point to make is that the status quo that we have at the moment is not ideal and needs dramatically reforming- we should not be romantic about it. Data management and mismanagement in the British National Health Service is notoriously uncertain. I deal all the time with documents which are lost before an investigation is carried out when people have died. This includes documents which disappear and documents which are falsified by clinicians after death of a patient to cover up service failure and avoidable death. The present is not ideal, and we should not be frightened of investing in artificial intelligence if it is going to deal with that situation.

Britain is spending on average about 1.8 eight billion pounds on investment in artificial intelligence, but we are not in the vanguard of European countries in finding a solution to this. We know from Freedom of Information research, which was done in the last couple of years, that perhaps only half of hospitals in the United Kingdom have actually invested in artificial intelligence so far. And they have been very reluctant to join the Government in thinking that this is going to be the way forward. But the

Government is right in understanding that more than one million patients every thirty-six hours have transactions with the health service, and the more we can use artificial intelligence to make sense of that, the better it would be.

There are in addition some very important clinical developments which are saving people's lives as a result of artificial intelligence. Inner eye scanning for patients in Cambridge is leading to a clear-up of prostate cancer in the UK more quickly than it was before. CT scans of patients who are suspected of having coronary heart disease are being used in a way which is much better than using angiograms. There are lots of innovations with Moorfields Eye Hospital in London, for example, and with GPs and practice nurses in London, which are saving lives concretely as a result of artificial intelligence. But there is a problem, and Elisabeth Rynning referred to it, and that is service users have to be convinced. Now, were would you be if you were invited to have major surgery performed by a robot? Would you be with the Nigerians in saying that 69% of them would be in favour of having the surgery, or would you be with the timid Britons where only 27% are likely to agree to robotic surgery? How many would be on the Nigerian side? Two, three, four. How many would be on the British side? We are after Brexit now, it doesn't matter, you can say what you think. Ok, more people would be sceptical. This was a survey by YouGov of 12,000 people. Not enough is being done to convince citizens that the use of artificial intelligence is safe and credible and efficient, and therefore, public trust, which is the key issue in all public services, is not as great as it should be. There is also an issue of clinicians understanding what value it has and how they need to be educated to go along with this, and that doesn't happen at the moment.

I want to end by talking about four emergent issues very quickly:

First of all, public trust is the key issue for the 21st century in terms of the delivery of public services. If the use of artificial intelligence is going to undermine public trust, we have more serious problems than we had before where public trust was already being eroded.

Secondly, there needs to be regulation. At the moment we have, we were told, perhaps between 84 and 85 regulatory frameworks in Europe. And that is simply not credible, particularly where you have a private sector which wants to make money out of what they are doing, rather than promoting the public interest. The problem for the public service is that it's not as intelligent or knowledgeable about artificial intelligence as the private sector. So, it has to get its act together in order to deal with the beliefs of the private sector. Otherwise rings are going to be run around the public service. We can't have the situation of Groucho Marx saying: "these are my principles and if you don't like them I have others". The public interest has to defend public trust in this case.

Third, there is an assumption in the academic world that ombudsman officers who just concentrate on being ombuds are more effective than those who gather other functions to them, in terms of being standards authority or being an information commissioner as well, which happens in a number of countries. In my view, that is turned on its head by artificial intelligence. It's the people, my colleagues, who have responsibility for data protection, and for human rights, who are ahead of the game in the Ombuds world than those of us who simply deal with complaints from citizens.

And the last point is this, and it's very important: the whole issue of artificial intelligence underlines the importance of Ombuds being professional in what they do. As Peter Bonnor and Rafael Ribó have pointed out with honourable exceptions, ombudsman officers are not prepared or skilled enough for engagement with artificial intelligence at the moment. If we don't address this, we will end up like William Huskisson, dead on the railway line. Thank you very much.

Elisabeth Rynning

Swedish Parliamentary Chief Ombudsman

Ladies and gentleman, dear colleagues,

First allow me to express how truly sorry I am for not being able to join you in Barcelona today. However, I'm sincerely grateful for this digital opportunity to share some of my thoughts with you. Since I'm obviously not with you in real life I must apologize beforehand if I should repeat any points already made in the discussions earlier today.

My name is Elisabeth Rynning, I'm the Chief Parliamentary Ombudsman of Sweden, and I have been asked to provide a few comments from the perspective of the Swedish Ombudsmen.

The Swedish Government, as well as many Swedish citizens, takes great pride in being very innovation and technology friendly. Some automated decision-making in the public sector was introduced already in the 1990s in the fields of tax and social insurance. Today we have a general provision in the administrative procedure act, which allows automated decisions but does not specify any additional safeguards. It has been questioned whether this really satisfies the requirements of the data protection regulation.

As you will already have discussed, the use of artificial intelligence could indeed have many positive effects, such as: cost efficiency, accessibility of public services, consistency and quality assurance of decisions, etc.

Two years ago, our Government set the aim that Sweden should be world leading in making use of the possibilities offered by artificial intelligence and introduced the Swedish Agency for Digital Government. Since September 2018 this agency has the task to coordinate and support the digitalization of the public administration. The agency is now mapping and analysing present applications of artificial intelligence and investigating future possibilities. Important aspects to observe obviously include both ethical and legality issues, various safeguards, and issues related to transparency, responsibility and accountability.

Many Swedish supervisory agencies and other bodies are also looking into different aspects of the use of artificial intelligence. The National Audit Office is investigating the efficiency of State agencies use of automated decision making without infringements of the rule of law, the Data Protection Agency is considering various personal data implications, and the National Board on Health and Welfare is studying the use of artificial intelligence in healthcare, just to mention a few.

It is thus fair to say that the Swedish public administration is on the move, albeit still quite cautiously, when it comes to artificial intelligence. But what do the Swedish citizens think? A study from 2018 shows that 80 % of the respondents were in fact, at that time, unaware that some decisions in the public sector are made by automated processing. The majority of the respondents believed such decisions to be on the one hand less biased, yet on the other hand less adapted to individual circumstances, less transparent, and no more accurate than decisions made by humans. I would assume that recent developments in this area and the increased public awareness of artificial intelligence in general might have affected public opinion since. Two years is quite a long time when it comes to artificial intelligence.

I will now turn into a description of a few examples of problems, or challenges if you will, that the development of artificial intelligence in public administration in Sweden has given rise to.

With a regard to the protection of personal data, the Data Protection Agency brought a case to the Supreme Administrative Court already fifteen years ago, concerning the use of biometric identification in a school. The Court found that requesting a fingerprint scan in order for pupils to access a plate in the school canteen was not lawful without consent to the data processing involved. Today, the Agency has a pending case on the issue of attendance checks by way of facial recognition scans, which the Agency considers to be unlawful regardless of any consent, since attendance could easily be checked by other less intrusive methods. We have yet to see how this case ends.

When it comes to accuracy or quality of automated decisions, we have seen examples of smaller as well as larger problems. It would thus seem that the automated road toll system identifying license plates cannot always distinguish Latvian number plates from Swedish, which may result in wrongful toll charges that still have to be paid while being contested.

Another example was provided by the Public Employment Agency in 2018, when it became clear that quite a large number (we are talking about thousands of individuals) had had their unemployment benefits wrongfully forfeited due to a faulty algorithm. The persons concerned were eventually reimbursed, but had of course suffered an injustice that ought to have been prevented.

My last example is a case handled by one of my colleagues at the Swedish Ombudsman Institution, which has to do with automated decisions on social welfare benefits in a Swedish municipality. The complaint does not concern the automated decision-making as such, but the lack of transparency concerning the algorithm that is used, which is of course a question of the utmost importance. The local government arguably has not been willing to disclose the algorithm, which has been developed by a private company. My colleague has yet not reached a decision on the matter. This last example leads me quite naturally to some closing remarks on the role of the Swedish Parliamentary Ombudsmen, both in general and in particular with regard to artificial intelligence.

The ombudsmen, as you may know, there are four of us, all independent and personally appointed by the Parliament, have the main task to supervise the application of laws and other statutes in the public administration. In particular, we are to ensure the constitutional requirements of objectivity and impartiality and that the fundamental rights and freedoms of citizens are not infringed upon in the public administration. The overall aim is to promote the rule of law. The ombudsmen handle complaints from the public, but also take initiatives and do inspections. During our investigations we have far-reaching powers laid down in the Constitution, but when it comes to our decisions, they are formally advisory and we cannot change the decision of a court or a public agency. It is in our power to propose amendments of legislation in order to remedy legislative shortcomings, but the Government is under no legal obligation to act on our proposals. When it comes to the use of artificial intelligence, the ombudsmen can supervise the public sector and provide guidance on good governance, transparency, access to justice and in general promote the rule of law. Furthermore, we can provide opinions on law proposals, where proposed implementations of artificial intelligence may cause risks to fundamental rights and we can propose changes in existing legislation when called for.

Artificial intelligence is, as we are all aware of, not a phenomenon that can be separated from, for example, fundamental rights. When used as a tool by the public administration, artificial intelligence forms an intrinsic part of the use of public powers and must be revised as such.

Our challenges involve, as always, prioritising the use of resources, but in this case, also finding the right competence for supervising technically complex digital tools.

With the constantly increasing privatization of public services, it is also a challenge to handle the public/private divide in our supervision.

We do indeed have challenges to face in our investigative role. At the same time, our internal activities also need to be up to date. Even very old institutions must adapt to changes in society, albeit still upholding important core principles. Hence, we must also try to become more efficient by using digital solutions and artificial intelligence tools within our own offices. Maybe by getting a robot to sort the incoming complaints, like our Norwegian colleagues have. Furthermore, we need to ensure that the skills of our staff fit the new needs.

In order to better meet the challenges and possibilities of a more digital future, I have suggested to the Swedish Parliament that the Ombudsmen might need a slightly adjusted mandate. My colleagues and I were thus very pleased when the Parliament recently decided to have the regulatory framework of the Swedish Ombudsman Institution revised, in this aspect as well as in a number of others. I feel certain that we could and will become better equipped to continue to promote the rule of law in our modern society, with artificial intelligence as part of it.

On that note I end and thank you for listening. I send my very best wishes for a fruitful discussion and I sincerely hope to see you in person next time.

Quirine Eijkman

Deputy President and Commissioner of the Netherlands Institute of Human Rights

Thank you, thank you very much. Thank you for the invitation to be here. I'm representing ENNHRI today, this is the European Network of National Human Rights Institutions, as well as the Netherlands Institute of Human Rights. Thank you very much. I really enjoyed your presentation, by the way, especially the part where you said that National Human Rights Institute or other colleagues of ombudsmen people might have a role to play, but maybe also collaborate with other institutions and I will say a little bit more about the Dutch situation in a minute.

But first ENNHRI. What does ENNHRI do? This European Network of National Human Rights Institutions, forty five members. Well, we support other national human rights institutes, for the establishment, exchange of good practices, capacity building and training, intervening on legal and policy documents, such as on AI, and engaging with regional mechanisms. And we are very happy to collaborate with IOI. It's very important for ENNHRI, in particular because half of our members are both National Human Rights Institute and Ombudsman mandates. So, it's not the same in every country. Well, thank you very much Mr. Rafael Ribó for organizing this event, we are very honoured.

I want to tell you a little bit about National Human Rights Institutes before I explain what we do in the Dutch context and why we are engaged in the digitalization and human rights as a strategic programme as well as artificial intelligence. So, National Human Rights Institutes are State mandated bodies, independent of the Government. We have broad mandates to promote human rights at a national level. We are regularly assessed on complaints with the United Nations about various principles that ensure independence, pluralism, accountability and effectiveness. And I don't know about you, but I used to work at Amnesty International and when I worked at Amnesty International in the Netherlands I never had to explain what Human Rights were or why we did to work, but as soon as I'm in the Netherlands talking to Dutch citizens they are always asking why should human rights be for our country. It's broadly assessed, it's something we do in our foreign policy much less than human rights in our country. So, sometimes certain issues are not really perceived as a human rights issue and this also applies to technology and the use of technology by our Government as well as private companies.

Well, National Human Rights Institutes address all different human rights, including civil, political, economic, social and cultural rights and we are confronted every day with the challenges but also the opportunities, such as Elisabeth Rynning just said, about opportunities that technology and artificial intelligence also bring for our own work, for Ombudsman's office work in the Netherlands, but also for citizens.

So ENNHRI is deeply interested in these issues and just like the Netherlands we are kind of new in this topic as well, we have much to learn, a lot of work to do and a lot of work to do in terms of catching up. Nonetheless, within our network we try to share our experiences, so if you have anything to add, please also come to the ENNHRI meetings, there is even a special committee now dealing with it, and I think it's only by sharing and be a collective that we can become more effective.

So, the Netherlands Human Rights Institute, as an ENNHRI member, we are delighted to be part of this workshop and we are participating as an expert, and at the same time, already this past one and a half days I've learnt so much that I can share with my whole team, so I'm very, very happy to be here. So, the Netherlands Human Rights Institute is not just a human rights institute, we are also an equal treatment body, so with my nine other commissioners, including our President, we have policy, we observe laws, policy, give expert advice, etc., but we also deal with individual complaints in relation to equal treatment. So, people can submit individual complaints and then we can provide a ruling, we have about one hundred and fifty per year, but I think just as important are the phone calls we have every day of people who are just stuck in system. And these phone calls usually start with people saying: "My rights are violated, please help me, I don't understand the system, I may be digitally illiterate", in the sense that what you were also saying, sometimes elderly people or sometimes also people who don't speak Dutch or have other limitations, find it quite hard to engage with the State or with private companies to just, you know, apply for a bank account, or etc. So we get many of these calls in our front office and we often try to help people manage or enter the system. Now we also have this part of the National Human Rights Institute, where we try to give, to advocate human rights, give expert advice on draft laws, and Parliamentary reports, etc.

Over the years, and also due to budget limitations, we've tried to create strategic programmes. For a couple of years we work on these programmes, so we try to be more effective. For instance, last couple of years I led a programme on human rights education, which is not compulsory in the Netherlands, and we tried to advocate, to get that into legislation and we worked with a team of about ten staff to get that done. One of our new programmes is Digitalization and Human Rights, and the reason why we selected this team is that we felt that we should complement the work that our Data Protection Authority is doing, because we feel the challenges posed by digitalization, and also of course including artificial intelligence, it's not just the privacy or data protection issue, but many other human rights are involved. Now, in this conference here, that is very evident, because we already discussed equal treatment, discrimination, access to justice, etc. But we must, I think, also realise that in many public and political debates, at least in my country, this issue is only related to privacy. So, we really have this urge to educate decision-makers and the general public that the threats from a human rights perspective do not primarily relate to privacy and data protection issues but are far broader than that.

A second reason why we've selected this program is that increasingly we are getting reports that equal treatment in the labour market is a very big issue, also in our educational, especially our vocational training issue, and we felt we had to do something about it because, whether we like it or not, recruitment of new colleagues or recruitments of people who want to work are increasingly going to be done by computers or by AI, algorithms or systems that are related, are semi-

automatic or automatic. And one of the things we want to engage in and talk to employers is whether these mechanisms are neutral, and whether everybody can be increased, because there is a huge risk of bias.

Now, if we talk about the second part of our programme, is that in the Netherlands on paper there is access to justice. So for instance the SyRI programme, which was discussed a couple of times during this conference and I'll tell you a little bit more about that in a minute. There is this believe, which is not true, that people can't resist the decision that was made or related to that system. This is not the case, but the problem is that people often are not aware of how the decision is made that affects them. So, when the people who are flagged by this automatic system on social fraud, is not as if they receive a letter from the Ministry or from the Municipality saying "we flagged you for this and this and this reason, and that's why we are investigating the fraud". They didn't get that kind of information. So, one of the things we want to achieve with our programme is to make access to justice in practise more transparent, that people become more aware of what is the basis of the decision to investigate them, because we feel that the system risk and negation system, these systems are used by many of public authorities, by the tax authority, etc. And for me, personally, I'm not saying this on behalf of the system, sorry, of my Institute, I think these systems are there to stay, whether we like it or not. The issue is more: are they human rights proof? Is it human rights by default? Is the system created in a fair way? And even though the District Court of The Hague ruled that the detection tool system should comply with other standards, I don't think we should for itself, it's very likely that the Dutch lawmakers will amend the law and this system in one form or the other will be used in the future, and I'm not saying that is good or bad, but I also think that we have to educate our members of Parliament and the lawmakers who make the decisions that these kind of systems are used, that they become aware that they have to be human rights proof rather than saying that the system as such should not be used.

The second issue is that we currently, and I think that's maybe the most interesting to finish off, there are a lot of debates now in the Dutch Parliament about public oversight for the use of artificial intelligence by public authorities. That is not just related to the social for detection indication systems, so the SyRI system, but it's also related to the tax authority. The Swedish Chief Ombudsman Elisabeth Rynning said something about AI and PR, so the police who make photos of cars and automatically decide if you are flagged or receive a fine or not, these systems are used, and now the Dutch members of Parliament are saying: "We should do something, maybe we should draw all these motions of Parliament saying we should have a new Commissioner, a new artificial intelligence commissioner to do something about it, they should register, they should do this, they should do that."

In November, the Minister of the Interior decided that there should be an investigation, an official investigation on the use of algorithms by public authorities, and the whole debate about how oversight in that context should be regulated or is already regulated. And I think I really enjoyed the research that was conducted by Valerie Frissen and Marlies van Eck. Marlies van Eck is a very interesting researcher in the Netherlands, who has done a lot of studies, she just finished her PhD on automated decision-making by the tax authority. And their advice to the Minister of the Interior is actually that we shouldn't create more or new commissioners or bodies, but that existing oversight mechanisms, such as the Ombudsman, such as the Data Protection Authority, such as the Netherlands Institute of Human Rights should use their mandates to investigate these kind of issues, but also should hold public authorities to account to explain more clearly and to vote them accountable for the fact that if you use these kind of systems for public use that are transparent, that are human rights proof, etc., and that is, the use of algorithms is part of your duty of care, just as it is with other systems that you use.

I also think, I'm not sure whether the Minister or the Cabinet will follow the advice of this report, but I do think the role of the Ombudsman, as well as the role of the Netherlands Human Rights Institute, as well as the National Audit Commission etc. should become much more focused on the use and oversight in relation to the use of artificial intelligence. And now suddenly everybody in the Netherlands is doing this topic, so the National Audit Commission has decided to do an investigation now on the use of algorithms by public authorities. I think that's a very positive step because we are not even sure who is using which algorithm when, so we're catching up. On the other hand, it is easier to say these things than to get them done, even in our own strategic program. I think we have to be very clear on how we define success and I think this is a real challenge and an opportunity for ombudsmen. One of the challenges is also: who do we hire? At our Netherlands Human Rights Institute we tend to have a lot of lawyers, not a lot of people with a technical background, people who have actually worked in the tech industry or are able to make an algorithm him or herself. And I think that's very important, because when I used to work at Amnesty there was the same problem. When we discussed these issues, we really had to work on recruiting people who actually understand the problem from a technical perspective.

Last but not least, when I was talking about this report on the use of algorithms by public authorities our institute was interviewed and we were talking with a lot of the public bodies who are actually responsible for providing welfare, responsible for all sorts of public services, more kind of the executive parts of our Government, a lot of the people involved there. The data scientist said: "We have a very big problem and one of the problems is that we have to make a lot of ethical decisions whereas we are middle management, but the most senior people in our organizations (the President, the directors, etc.) do not understand the problem always very well". And I think that's also something we have to take very seriously to educate the people, the decision makers, to really understand, take responsibility about the decisions they make based on artificial intelligence. When I was in that expert meeting, I really became aware that the data scientists are very uncomfortable in making these decisions for them, and I think we really have to become more aware of that, because if as a Government agency you use artificial intelligence you have to take responsibility for what you are doing and I think the ruling of the District Court of The Hague has made the Ministry of Social Affairs more aware of that.

Last but not least, we should do a lot more to educate our lawmakers when these laws are passed through the Government, because when SyRI was passed through our Parliament a couple of years ago, there wasn't even a vote by hand, it was just "the hammer piece" and nobody asked questions, and I think that's where the real problem is, because I like strategic litigation but it's not really an alternative for good political decision-making as well as laws that are being passed to Parliament in relation to artificial intelligence.

So, these are lot of challenges, not necessarily just for the Ombudsman offices, but I completely agree with what you were saying Rob, that I don't think that the Ombudsmen or the National Human Rights Institute can do these things alone, we really have to cooperate and have a joint strategic perspective in how we should ensure accountability, both in the private and in the public sector.

So thank you very much.

Eric Houtman

Belgian Federal Energy Ombudsman and member of NEON

Hello everybody, dear colleagues, I'm Eric Houtman, Belgian Energy Ombudsman, also member of NEON. And NEON is the National Energy Ombudsman Network and it currently consists of the Energy Ombudsman Services from Belgium, UK, Catalonia, Ireland, Italy, Poland and Georgia. And Belgium, UK and Catalonia, the Síndic, were the founding members of NEON. I'm a sectorial Ombudsman and with legal mission to deal with complaints and disputes between end users, citizens and private energy companies. So, I'm very honoured, of course, as a sectorial Ombudsman, to be invited to this event organised by IOI and I'm very glad and thank Mr. Rafael Ribó for the excellent hosting of this event.

Artificial intelligence is already defining our energy consumption patterns. Within a few years, home assistance and smart meters have become very widespread; we are surrounded by many web-connected devices, starting with our smartphones, and in a few years' intelligence, self-learning machines and software will be able to use the incredible quantity of data generated by the internet of things to make decisions and perform actions without human supervision. Maybe tomorrow our self-driving electric car will decide for itself when to charge its battery after having talked to our smart house.

So, I'll give you now an overview of the green and digital ambition of Europe in the energy sector, because digitalization is one of the cornerstones of the new European energy policies. It allows to decentralise sources of energy such as the energy produced by individual solar panels to be integrated into the grid, and it will be an excellent driver of decarbonisation and energy efficiency. For instance, AI and big data can detect energy needs more accurately resulting in more energy efficiency infrastructure and adapted consumption. In some countries, people already have access to dynamic prices matching the general flows and wholesale prices and aggregators will also act as intermediaries for consumers by combining demand and flexibility data, selling this in the wholesale market and sharing the benefits with end users.

Are there challenges for consumers? Of course, consumers will be confronted with more data, and all these data will be safely managed and stored. Actually, it's very easy, every data before the meter reading is managed by the DSO, the Distribution System Operator, and every data after the meter will be part of the private sector, will be the responsibility of the private sector. There will be more actors. Who will be responsible for what? The consumer needs to be active, needs to be activated, he will be a prosumer, he will have an agreement of contractual terms with the supplier, but also with other service providers. It exists already now, that suppliers offer products like woods, like pellets, they offer also other services like the maintenance and the repair of the heating installation or the solar panels, it's already the case.

The main question, of course, is: Should AI take all the decisions? But especially, what about the vulnerable people without the financial, technical, and digital capacities to become an active consumer or prosumer? And then another question is, of course, what about discrimination coming from data cross-referencing? This is happening already today in the energy sector; it's called a *mystery switch*. Your energy contract can be taken over by a new supplier and this supplier is the supplier of your neighbour, because he was using the identity code of your meter instead of the meter of your neighbour, and you will have a contract with a new supplier, and this is called a *mystery switch* or *kidnap by using data*, the mismatch of data. And all these questions are the reasons why the energy sector is one of the main targets of the white paper of the European Commission that was published fourteen days ago, I think.

Are there challenges for the Ombudsman? Of course. We have to deal with new relationships. And I mentioned the NEON because the work of the NEON was very, very good at the European level. Because last year a new electricity, the forth electricity directive, was published, and in article 26 there is a provision that also mentioned ombudsmen. And why is this important? Because there was a discussion between the DG Justice and DG Energy. DG justice was saying: "Oh, the directive of 2013 is existing already, so we can only make a reference to this ADR directive, but there was a problem and DG Energy understands this problem because they were aware that with all these new products, all these new services, the algorithms used by operators, aggregators, there will be not only a contractual relationship between a consumer and a supplier, but there will also be other relationships, other relationships between business to business, between non-profit organizations, for instance: schools with solar panels on their roof, they can have an issue with the supplier or with an aggregator or with an individual consumer, so it was important that also ombudsmen were mentioned in the provision of that directive because it is important that we not only deal with issues, problems, complaints, between an energy company and an end user but also, B to B relationships, C to C, even peer-to-peer platforms and also services that are indirectly linked with the supply of electricity and natural gas. And that's also the reason why at the end of 2019, last year, I wrote a memorandum, a political advice to the Federal Government, that is not really there, everybody knows we have some problems in Belgium to have a federal government with full powers, but I wrote it already for the next federal government, a political advice that makes it possible to extend the legal mission of the Ombudsman in order to tackle also these new relationships, new services and new products.

What is also important and challenging for the Ombudsman is that social, ecological and environmental public service obligations must also be strengthened. It's important that they need to be updated to ensure at least the same level of protection as at present, because I think it's difficult to explain to an end user that you can be cut off of the grid because you have a dispute with a new product or a new service and you're cut off of the grid for the electricity or your natural gas. I think this would be very difficult to explain to end users, but it can be the case. Apparently in Germany they have already had these cases. Another example is that now in Flanders they started with the rollout of the smart meters and it's clear that now the supplier wants to use the smart meters as prepayment meters. When people cannot pay their bill or they are discussing their bill, they can be cut off of the grid simply because the supplier can ask it to the DSO. Do we want this? Or do we want to have the same protection as now? So, we need to adapt our legislation and also our regulation.

It is also important that without clear boundaries, artificial intelligence can exploit consumer's vulnerabilities and can lead to harmful target advertising and unfair commercial practices. We have already a lot of problems with selling practices of some, not all but some, energy companies, so I think we need to avoid that people will be forced to switch because the computer was saying that they needed to switch. Because there are already service providers who give the possibility to people to offer the service that they will look two, three times a year for the best price, not the cheapest price, but the best price in function of the consumption behaviour of the client. But you need to give a go, of course, you need to give an opt-in before such a service provider can switch you from the former old supplier to your new supplier, but it's already a service that exists and that is even developed by a consumer organization, a private consumer organization in Belgium. So be aware of that. And of course about access to ADR justice and regulation. Especially regulation must become more agile and dynamic, in the energy sector. We are talking about dynamic regulation because we need to tackle all these challenges about new services, new products, new actors, aggregators, and other service providers.

I think that there should always be a human oversight also, and then, as ombudsmen, we should be able to monitor and verify the artificial intelligence as to correct the shortcomings and demand for immediate adjustments if necessary. But, of course, I agree with you that we need the right expertise, the technical expertise for doing that, because it is not so easy to have a real insight and oversight of all these mechanisms that are involved in these systems.

And also important, I think, because now in the electricity directive it is mentioned that every Member State has to need a system of financial compensation. So not only for financial compensation or individual redress for the victims of energy, electricity and natural gas related problems, but also for victims of these new products, services or service providers. As an Ombudsman, I think artificial intelligence creates a new landscape of opportunities to identify patterns and shortcomings, and we can feedback to the regulator and policy makers to adjust the market. I think this is a very important role for the Ombudsman.

And then, I'll jump to conclusions, so I think a robust legal framework is necessary to deal with AI-related issues. Ombudsmen must also assert their role of impartial watchdog in the green and digital transformations of the EU economy.

And last but not least, everyone deserves privacy, fairness, dignity and the right to have humane treatment and an answer by a human being.

Thank you.

Artificial Intelligence & Human Rights

ombudsmanship challenges, roles and tools

Tuesday, March 3, 2020

Conclusions

Chairperson: Catherine de Bruecker, IOI Europe Vice President & Federal Ombudsman of Belgium

Ulises Cortés, professor and Senior Researcher at the Universitat Politècnica de Catalunya (UPC), Barcelona Supercomputing Center (BSC)

Daniel Marco, director General for Innovation and the Digital Economy

Rafael Ribó, European President of IOI and Ombudsman of Catalonia

Catherine de Bruecker

IOI Europe Vice President & Federal Ombudsman of Belgium

First of all, I want to thank Rafael Ribó and his staff for organizing this workshop as it is a very important and actual topic we need to engage with. Artificial intelligence is no science fiction. As it was said, it is already present in our everyday life. What is absolutely sure is that we were ignorant maybe yesterday, but we cannot pretend anymore to be ignorant today, after the very intense and important debate we had during these two days. We now both have the difficult task to draw up conclusions. And because these two days were so intense, I thought we could have needed maybe some artificial intelligence system to help us summarize all the information we gathered. But fortunately I have much better than artificial intelligence to help me in this task and I'm very glad first to give the floor now to professor Cortés, who will help us grab these conclusions.

Ulises Cortés

Professor and Senior Researcher at the Universitat Politècnica de Catalunya (UPC), Barcelona Supercomputing Center (BSC)

I thank the organization for trusting me to present the conclusions of the event. The sessions have been very interesting and I have learned a lot from the different speakers and from their points of view on the subject at hand: the impact of AI on human rights. These two days here have been very fruitful and exciting for me because, as I said yesterday, I am just a typical artificial intelligence professor in a Computer Science school, I have my students, my research teams and I have almost no contact with legal institutions.

First some acknowledgements: I would like to thank all the speakers for their participation. Some of them helped me to draw these conclusions, let me mention in alphabetic order Atia Cortés (BSC), Catherine De Bruecker (IOI), Joan Manel del Pozo (UdG), Geraldine Mattioli-zeltner (ECHR) and Martha Stickings (FRA). Access Now and FRA produced fascinating documents, and I must confess that I was unable to read the document that Madame Mattioli-zeltner gave us today.

AI & Society

Let us go with the conclusions. We have spent two days at work trying to explain each other, from different points of view, the following questions: What is it that causes concern when we talk about the possible uses of AI? and how those uses can affect human rights?

There is no single answer but I believe that now it is clear that AI is not an independent and isolated entity. Many participants address to AI as something that exists, thinks, has intentions and, maybe, political will. But AI it is just an intellectual creation that has many successful applications that affect us on a daily basis in many aspects of daily life, from the most trivial to the most complex and confidential. What is evident is that AI has an important economic value, and therefore it impacts society. The AI disruption is transforming our countries, organizations, markets and lives. There is ample evidence of that, in fact if we compare the top ten companies just fifteen years ago and today, we will look at the depth of the change, only one remains: Microsoft, and in today's top ten only one is not tech: Saudi Aramco.

What is not so obvious for the general public is that this technology is almost unregulated mostly because it has built on the roots of war technology, which is in turn unregulated because of its own nature. Furthermore, AI is used on purpose, so humans assign intentions to machines using this technology to interact with us. One of those intentions is, of course, to make money, and companies try to apply AI in any field and with any intention because there is almost no regulation, they do not have limits. Moreover, in this market there is no self-regulation, it is a free land to do all kind of things that imagination and technology permit. Individuals, companies and governments do the same - there are governments that are using the AI technology for their own purposes - have almost no restrictions. In the world, you have a good collection of examples of countries that are using AI.

Having human oversight for AI-based systems is a must. It is an EU principle, if the action of an intelligent system affects living beings and/or the environment it has to be human-controlled. This notion of control ought to be clearly regulated, we cannot put the man outside the loop of the decision-making. That is simply a bad idea, and this may lead to many undesirable implications. Other artefacts and technologies are always supervised. This notion of control brings some questions: How close this control must be? How many agencies have to intervene? How often you control the system? The answer to those questions has legal and economic implications.

The role of GAFAMA

Governments, industrial societies, economic organizations and even GAFAMA,¹ which stands for Google, Amazon, Facebook, Apple Microsoft and Alibaba are in a rush to influence the good practices in the design and use of AI-Based systems. Please, do not forget that China represents, at least, 30% of all market, so they have to have to be considered in the picture. China is investing astronomical sums in AI research with the intention of leading it and applying it in its economic development. As we said yesterday, even those gigantic companies are all trying to establish ethic policies and guidelines, and we have to praise them. Still, the impact of AI applications on human rights needs to be assessed very precisely. Who will put the bell on that cat? Those are tigers, not small cats. Someone has to volunteer to put that bell because we need to hear –all the time - that they are approaching us.

The regulation game is not only about ethics and AI. We also have to talk about legal, socioeconomic, and cultural aspects - the ELSEC elements - that is, the ones promoted by the European Union. Among those, we have to include gender. We cannot go ahead in this task, without considering half of the population, half of the talent that has to be there in the first line, and that is often forgotten in our discourse. Once those aspects are taken into consideration, we have to ask ourselves why those companies attempt to modify those values. If Europe has agreed on these values, we have to fight for them. Why are we letting companies modify the way we behave?

¹ GAFAM is now FAANG. FAANG is an acronym referring to the stocks of the five most popular and best-performing American technology companies: Facebook, Amazon, Apple, Netflix and Alphabet

Dilemmas, values and videotapes

This title comes from a famous film. You may have watched it. In Figure 1 we show pairs of options that are presented when we think about the design of an application, in this case when we try to design an AI-based application. In our society, we enjoy freedom and independence. We have freedom of choice: for example, the freedom of not using Internet services or depending on them. We have free choice versus machine-based recommendation. It is your option. However, maybe machine-based recommendation should be regulated to protect citizens from be influenced.

What happens if you want to choose between Security and Privacy. If you want a service to be more private, then your security needs to be stronger, this implies to have less flexible interfaces. If we talk about Explainability against Performance. With the available computational power performance is easy to achieve, for example using Deep Learning. Today Deep Learning is unable to explain how it arrives to a solution.

You have other pairs like Accountability and Confidentiality; Prosperity and Sustainability. And then all this creates an ethical overload: you cannot have it all. Moreover, if you pretend have it all, you have to pay the price.

Ockham's razor

Here you have some ideas for your work as ombudsmen. Please do not reinvent the whole legal framework. We have, as we saw in the session before lunch that Europe has an important legacy, that should be explored and used. Nevertheless, you should be aware that technology that is both ELSEC compliant and made at home will be more expensive than the one that comes from China or the States. Also, we need citizens to be aware of that and they ought to be ready to pay the price. Having local AI-based technology means, for example, that the data will be inside Europe, we will control the servers, the persistency of the data, the data security and privacy, and when necessary the data will be erased. Also, citizens can, if needed, exercise the right to be forgotten on the Internet. I believe that all these characteristics are essential but may be expensive.

Please, get the appropriate technological assessment to understand each AI-based service on use. There are no two pieces of software that are identical. They may be implemented differently by two diverse teams using the same specification, but you have to be sure that they are compliant with the specification. Within Europe, you have the talent to do that. The universities have plenty of researchers, the European Union is funding research and the education of talent, but we are unable to retain that talent.

Technical education for our society is a must. To have the appropriate technical and critical education for the future generations maybe it has to be included in the syllabus, and perhaps complement the technological education of today's citizens.

Homework

I have some homework for you the ombudsmen. The first assignment is about transparency: you need to supervise transparency certificates of all AI services in use. Someone with recognised authority has to certify that those services are fair. For example, in public administration, all AI-based services must be transparent even if you try something straightforward as using labels of three colours: green, yellow and red, that will be enough to inform citizens. The reason for this request is the following. Although I can read and understand all the legal documents that accompany any computer application, I may not be capable of understanding the rights and obligations contained. It would be better for someone authorised, by the State, to read them and evaluate its legality and qualify it. She has to state if both the app and the contract are correct to be used, i.e. it behaves

inside our legality. If this is so the app will have to have a green seal. The ombudsmen, for example, has to supervise that all Apps in use have a seal.

Another task is what to do when a service has a red flag? What to do when people are using a red-flagged service? Is this admissible? How this affects other citizens. You have to reflect on it. Moreover, of course, you have to demand the EU Parliament and/or the Council of Europe the appropriate legislation to have those transparency certificates installed everywhere in Europe.

Another claim that we heard from David Cabo, from Civio, is that all software produced with public funds and used in the public administration has to be available for inspection. I agree that you cannot give access to that software to anyone to supervise. Maybe through the mediation of the ombudsmen services, third parties under all the certificates of non-disclosure can inspect that software. This, I believe, will augment the citizens' trust in those tools. We have to remember that if the government paid for the software, it is a citizens' property.

Regarding (machine) responsibility. What happens is a machine fails in a decision? Who is responsible?

Please, support and certify only apps that are fair and bias-free. I know that both are fuzzy terms. Nevertheless, it is necessary to start thinking about those terms and how to test them. It is necessary to have a public sandbox to test all those programmes and a regulation for the software life cycle. To answer questions like When it will be necessary to requalify a piece of software?

Furthermore, I believe that you have to supervise all algorithms used in the public administration, in particular, those that used to aggregate personal data to make decisions in, at least, the following fields: healthcare; human autonomy; freedom of movement; liberty, security, justice, equality before courts (we have talked already about Compass); privacy and data protection; equality, non-discrimination and political participation (the unfortunate example of Cambridge Analytica in the United States and other places in Europe, for example in Spain, are examples of how you can use those algorithms to modify people's opinion in voting, I mean, to elect their political representatives); education and environment.

Let me call your attention to the idea of letting Google provide services to schools, installing the Google Suite, I believe, is just an aberration. Kids do not need that. We arrived at this point of civilisation and all of you, most of you, never used Google Suite when you attended school. The most you had was a small calculator. Why kids now need to have all those services and tools? Kids do not need it. You need to develop critical thinking and then use technology. We are in this meeting because we are using AI-based tools without really understand how it works, and we fear that it is. As a society, we do not need the technology to be used just by default, without understanding it.

More about responsibility. Support the avoidance of the use in the public administration of inappropriate surveillance services and/or inadvertently harvest personal data. We have to avoid it. Nevertheless, of course, there are law enforcement agencies that may need those techniques in a given moment to provide society with the appropriate security level. I know that this a controversial issue and a source of discussion. The States have to be prepared to have an in-depth discussion about this because society needs to have those services. It may be necessary to use them in given circumstances to track a targeted person during a period that has been agreed within the security forces. This implies having them on hand for when they are needed and disconnecting them when they have fulfilled their function. It must be clear who supervises that this process is respectful of human rights. Less democratic societies already start using them to control their citizens.

Regarding accountability, it is necessary to certificate automated decision-making and avoid the accountability gap.

Trust is one central issue; we already heard the previous speaker talking about it. A way to build trust is to be accountable. It is the human responsibility to correctly design the use of data and decision-making performed by a machine. The machines will not do it. There must be someone identifiable who can be asked directly "why this machine is doing or performing like that?". It is necessary to create a methodology that allows identifying responsibility in the design of the AI-based system because that will imply the intention that was programmed in the system. Once a machine can answer how it arrived at a solution and can be audited and is accountable, will you use that machine to answer in front of a jury? It is a question.

The answer is yes; you can do it. There are already machines that can make arguments and argue like a lawyer moreover if they have enough evidence that they may be using. Are we prepared to do that? Because we are prepared to go into a surgical intervention by using a robot, so why not? Most of us fly in aeroplanes that mostly are flown with the support of three computers. The pilot is all most of the time but who is doing most of the calculation is the system of three computers. So, we are confident in machines that make critical decisions. Of course, the pilot is always there to supervise the loop.

This is the last one, and it is very tough. All data produced data citizens in a city, region, or country belong to the producers. Public administrations have to assure those producers can obtain a direct benefit. In Mexico City, authorities are fighting against Google and other companies to get all the data produced by the use of Waze and Google Maps to improve their infrastructures. Why? This shows how those companies work accumulating data in exchange for a service. Cities and governments are permitting them to operate; they are using national telecommunication infrastructure. The uses of citizens data have to be used in the benefit of the citizen.

I believe that we have to regulate the life cycle of AI-based services. Furthermore, I put here an example that is very easy to understand: if you want to create a chemical to cure any disease, this process may last years. The drug discovery may take several years, between finding the targets, designing the molecule. Then the clinical trials, which are six years long. Society does not need technology to change every year, because we need first to experiment with it, we have to be sure that that technology is giving the appropriate results, and it is not harmful. It is necessary to regulate the time and conditions of delivery in the market with the appropriate tests. We do with pharmaceutical companies, why we cannot do it with software companies or other telecoms?

I do not know if AI is the problem or the solution. It is a scientific area. Our research is trying to find out which are the mechanisms that are behind what we call animal intelligence. We want to reproduce it into a machine. One of the targets that we are aiming is to get some artificial intelligence that is different from human intelligence. At that point, a dialogue with that intelligence may help to solve problems having another point of view. AI is aiming to replace humans. This is not the idea. In my opinion, it is a perverse idea to create machines, any machine, to replace some work placement of any individual. If someone wants to replace a citizen in a line of work, she first has to provide means for those people to have a living and pay taxes for that, doing that.

"Mankind was my business; charity, mercy, forbearance, and benevolence, were, all, my business. The deals of my trade were but a drop of water in the comprehensive ocean of my business!"

This is a quote from A Winter's Tale, Charles Dickens. We are always talking about humanity, and we want to improve humankind's life. We do not want to use technology

to make them live in a virtual reality where they smile, and the non-smiling people are the ones that are controlling society.

Thank you very much.

Daniel Marco

Director General for Innovation and the Digital Economy

Good afternoon, hello everyone. My thanks and congratulations to the Catalan Ombudsman and the International Ombudsman Institute for organising this seminar and also to all the experts who have participated as speakers or as audience. It is an honour for me to be able to speak here today at this closing ceremony, which I will try to keep brief. The truth is that I have been able to follow the different talks and round tables online and it will be difficult to add anything that is not reiterative. Or worse, inconsequential. But I will try to contribute some conclusions and also from my personal perspective and on behalf of the Government of Catalonia on the next steps we should take.

What has become clear is that the digital revolution is transforming our society and that artificial intelligence is one of the technologies that will have the greatest effect due to its transformative capacity and impact on society as a whole. This transformative effect of the digital revolution will also affect changes, as we have seen, in what would be as we understand it, human relations and the rights of citizens in society itself. In the same way that the industrial revolution changed, modernised, society and also enabled the achievement of certain social rights, we find ourselves at a time when digital technologies will change, I would like to say radically, but at least they will certainly change social structures, and this means that we must take into account how we renew this social pact or how we manage in this digital world not to lose acquired social rights and, above all, to improve. In the end, a modernisation, a transformation of society, must be for the betterment of the common good. New concepts will be necessary, associated with a reflection and a debate on the limits of ethical development and the positive impact that this development must have on our society.

This is a global problem, and it has also been present here today, these two days we have been seeing this impact on an international level. But as has also been mentioned, the fight for rights and freedoms has had a need for local action, and the benefits acquired locally have been able to spread globally.

But both Barcelona and Catalonia, in this sense, I believe that they must play a relevant role locally and with global impact in this new challenge that has been discussed here. I personally believe this, for many reasons. In addition to Catalonia's experience in the field of artificial intelligence and the excellent scientific community that we have, in 2017 the scientific community was a pioneer in Barcelona, in the Barcelona Declaration on artificial intelligence, which in fact laid the foundations with experts from all over Europe, of what later became the ethical guide, on the ethical guidelines on artificial intelligence of the European Commission and other countries that have managed to have these guidelines. In 2018 we launched the Digital Future Society, which also seeks a forum for debate and reflection on the areas that were discussed here today and yesterday. In 2019, the Government presented the Charter of Digital Rights and Responsibilities, which seeks to incorporate these new digital rights into the ideology of the common good and of the administrations themselves. And very recently, three weeks ago, we approved the Catalonia.AI strategy, the artificial intelligence strategy of Catalonia, which has a specific focus on the field of ethics and society, on the development of artificial intelligence for the common good, transparent, centred on and thinking about people, and where the development and impact itself is based on consensus.

Thus, we have a history that can position us well in this environment, but we also have strong digital activist organisations, we have associations and an organised civil society in many areas, we have experts in research in the field of the ethics of artificial intelligence. Therefore, we have the ingredients to make proposals and articulate a political and democratic debate. What is missing, perhaps, and what we need to do, and what we need to do from the public administrations and society in general, is to bring this debate to the first level, both in the political sphere and in civic squares, neighbourhood organisations and society in general.

I believe that the discussion we had over the last two days with experts from different sides need to be taken to the whole of democratic politics in order to consolidate and develop this consensus, and the political administrations need to believe this in a more profound way. And here, from the Government of Catalonia, we will work, not from a solely and exclusively discursive perspective, but with actions. As I have said, the very strategy within the artificial intelligence of Catalonia will work in this regard. Not only by saying what we will and will not do, but also by carrying out actions regarding website closures, the digital decree or not being against the provision of autonomous lethal weapons in their development.

It is clear that seventy years ago, when the Universal Declaration of Human Rights was made, the digital world, the digital society, was unimaginable in that area and that after all this development, what we have to achieve are two fundamental aspects. Firstly, that the universal rights of human rights are also applied in the digital environment in the same way that we have achieved in many areas, in most of the territory and in most of society. Therefore, these rights must be maintained, secured and safeguarded. But we must also start working on new digital rights, new emerging rights of the 21st century. Aspects that were perhaps not included in that Charter and which we have begun to reflect on today.

As far as the Government of Catalonia is concerned, you should know that we maintain this as a priority. The rights, freedoms and equality of our society must also exist in a digital country, and in order for it to be a social country, we need the digital revolution.

Thank you all for these two days, to the Catalan Ombudsman for the organisation, and also for being able to monitor the coming months and years so that, at the end of the day, we can have a freer, more egalitarian and equitable society.

Thank you very much.

Rafael Ribó

European President of IOI and Ombudsman of Catalonia

First of all, my thanks to all of the translators for the fantastic work you did during these two days. But also thank you very much to all the Síndic team that since many weeks ago were going to work to prepare this event. Thank you very much to all my colleagues. Thanks to all IOI members. The last IOI workshop was a very interesting one in Riga. Today we have with us the Riga ombudsman. And also, thanks to all the other ombudsmen, local and university ombudsmen, who are accompanying us during these days, to all the entities that are working on those matters all around. Thanks to the Hub Museum, that provided us all the facilities, to the speakers that from different angles were in our discussions, to the sponsors, especially to "La Caixa". But I want to make a very special expression of thanks to two people that since weeks ago, months ago, when we started discussing on this workshop, started immediately, even maybe not being acquainted with, pitching us on artificial intelligence. Dr. Ulises Cortés, a fantastic man, as you have heard in the last intervention, that knows a lot, and with a very objective and critical

approach to artificial intelligence. Ulises, thank you, thank you very much. And also, to Carina, from Digital Future Society, who was the last speaker in the first-round table, and who also helped us a lot with the presentation and introduction of the different fields we should approach.

We have been webcasting live the whole workshop through streaming, you have it also in YouTube and the complete interventions are going to be posted in our website in a matter of days. And I consider that as a first step, at least for us that as I said at the beginning, we were ignorant on those matters. As a first step and as the IOI ex-president Catherine De Bruecker said just now, I am sure the IOI is going to continue in this same direction. At least to continue, in my humble opinion, learning.

Should we trust or not? Should we trust on AI or not? I am using again the magic word: *trust*. I would say: do we know or not? Are we informed or not? We still are, in Catalonia, maybe in general terms in Spain, figuring out what was the main reason to suspend the Mobile World Congress of last week. There are several interpretations, I do not want to choose any of those, but there are several interpretations that prove we lack the complete information on very important decisions. And I am very proud that the Mobile World Congress happens in Barcelona, of course.

Do we know exactly about how to validate algorithms? And how to get, as Ulises was saying, a biasfree information? And that is very important. Do we know how to make a constant work of auditing, of accountability, on all those matters? Very few days ago, weeks ago, we all knew in Spain that there was supposed to be approved a tax for some of these big companies. The tax was approved but the application was delayed, *sine die*, maybe at the end of 2020 or maybe next year. What happened? Again, lack of information or transparency of the real factors and the trust.

Do we know about, really, our rights? And I stick very strongly to the word *rights*. Let me tell it in a brutal way. I am fed up with listening "codes of ethics". I think it was Daniel Innerarity or maybe Lasalle who said: I counted 84 different codes of ethics. The so-famous self-regulation. For me, it is very important to stick with the concept of *human rights*. And remembering that the rights are always beyond the laws. You never conquer rights only in the strict legal frame; you always go further, as humanhood has done during our history. In the second-round table today we saw how rights were defined as binding matters, in charge of public authorities: state, regional, local... of public authorities and protecting. These rights are binding, not recommendations, they are a binding public responsibility and protecting. We have seen a long coverage of these rights. The last one presented by Ulises, about our right to privacy, our right to equality of gender, etc. And, of course, human rights will take care of the huge work of education, especially when we all agree there is a problem of literacy or what years ago was called the digital gap. Now we are talking about literacy or illiteracy on all those matters.

That is why I am very happy that as ombudsman, and as I see through my colleagues' ombudsmen, the IOI is now in shape to start approaching artificial intelligence and human rights continuing learning and taking actions. Why not presenting all these also in the next Dublin Conference? I say that without any compromise, because I am not the organiser, and I do not know if I am going to be there, but it is the World Conference. Maybe somebody will present those things there.

I am trying to apply all the homework we saw during those two days, especially today or what should be our duties. And I remember the three appeals that I was doing in my first speech. Appeals to all the ombudsmen, as Catherine was saying, we, members of the board, asked the ingoing board and those who are going to rule IOI from Dublin on to continue with this work. I appeal to the social

implication. Somebody said: Why don't you ombudsmen open a window for social participation, beginning with the complains of so many entities that are far before us, working for those rights and those matters? And of course, a third appeal, not only to those who came here representing international bodies, as Marta, Geraldine, and Mr Bonnor, not only you, all of us, should we... I would not say change, but should we improve or reform, but at least improve, our treaties in order to include all the different sides of this phenomena?

Look, I am going to end with a simple thought: we are in a neighbourhood that historically is close to what we call Poblenou, the new town, the small new town. But now we call it, especially the whole thing, 22@, referring to a concentration of start-ups, of tech enterprises, etc. I want to remember with a lot of pride that centuries ago Catalonia was the takeoff of capitalism in Spain. And here started the real industrialisation, and that is why we are more or less the machine of the economy. I would also like to remember that, at least two centuries ago. I want to remember that in 1851 in Barcelona there was the first general strike in all over the Spanish state, and it is natural, because it is where all the factories and industrial was. 1851. But the banner of the strike was: association or death. A very deep human right, the right to association, for unions. And now, also with pride, I say that here, around us, is the machine of the whole digitalisation, and more than that, of the artificial intelligence developing.

I take my compromise, in a very humble way, to fight so that besides Mobile World Congress, Digital Future Society, besides all those fantastic things, some day in Barcelona there is going to be a big human rights hub for artificial intelligence.

Thank you very much.



SÍNDIC
EL DEFENSOR
DE LES
PERSONES


IOI - Europe
International Ombudsman Institute
European Chapter

Collaboration:

 Digital
Future Society

 Barcelona
Supercomputing
Center

Sponsorship:

 Obra Social
Fundació 'la Caixa'