Faculty of Law
Chair for Legal Sociology and Media Law
Prof. Christoph B. Graber, PhD
Treichlerstrasse 10
CH-8032 Zurich
Phone +41 44 634 15 05
www.rwi.uzh.ch/graber
lst.graber@rwi.uzh.ch

Philosophical Questions about AI, Law and Governance

Faculty of Law, University of Zurich

Berkman Klein Center of Internet and Society, Harvard University

Workshop objectives

Adaptive persistent tracking systems using artificial intelligence are evolving rapidly and demand that law- and policymakers pay closer attention to the effects that these technologies are having on society. Facial recognition algorithms are considered to be the leading candidate among these technologies. They have been deployed in different areas and by different actors serving private as well as public policy interests. Their broad spectrum of application and intrinsic potential is accompanied by pressing concerns that the unprecedented disclosure of personal information and availability of such information will allow to further refine an individual's profile and manipulate her behavior.

The workshop starts from the premise that coping with the disruptive force of persistent tracking technologies requires overcoming the existing gap between Science and Technology Studies (STS) and legal studies in the analysis of legal and political issues raised by AI. Its aim is to explore novel ways of interdisciplinary research to better understand normative implications of smart technologies. Ultimately, we expect to identify an agenda for future research. The workshop is based on two case studies, drawing from two articles of the Economist (summarized on pp.4-5), and will be divided into the following three sessions:

1. The communicative capabilities and possible discriminatory potential of AI systems integrated with persistent tracking technologies.
2. Implications of using adaptive persistent tracking systems for mass surveillance purposes.
3. Private applications of persistent tracking technologies and the perceived risks for individuals/consumers.

Each session will be led by two workshop participants in their role as expert moderators who will briefly (5-8 minutes) introduce the topic of the segment, structure the discussion, and finally provide a summary of the session results. Please note that we are not planning to have formal presentations. However, we will complete a list of materials and provide hand-outs should you like to share such with the workshop participants.

The workshop is supported by the University of Zurich's Faculty of Law and Digital Society Initiative (DSI). The DSI has been set up as an academic platform for the purpose of promoting critical, interdisciplinary reflection on all aspects of the digitalization of society and the sciences. This workshop intends to contribute to these aims.

The discussions during the workshop will be recorded for internal use.

Agenda of the Workshop

Villa Hatt, Freudenbergstrasse 112, 8044 Zürich

15th February 2018

Persistent Tracking, Sensitive Data and Discrimination
Case study 1: Advances in AI are used to spot signs of sexuality

9.30	**Welcome Coffee**
10.00	**Welcome address** *Christian Schwarzenegger, Vice President University of Zurich*
10.15 – 11.00	**Introduction to workshop and case study** *Christoph Graber*
11.00 – 12.30	**(Sensitive) Data and persistent tracking. The era of communicative machines?** *Moderation of debate: Christoph Graber* The processing of sensitive data, such as facial traits, biological properties, behavioral and genetic characteristics, reveals deeper and even unconscious information about the data's owner. How does this feeling of being constantly observed impact individuals and society? Technical studies have shown that facial recognition/persistent tracking technologies can improve the communicative capacities of algorithms, making them aware of human non-linguistic communicative episodes. Can this be considered communication? Does this make the human-machine interaction as communicative as the interaction between people?
12.30	**Lunch**
13.45 – 15.45 (coffee break at 14.30)	**Algorithmic biases, discrimination and smart regulatory responses** *Moderation of debate: Eszter Hargittai* If the accuracy rate of a system is a function of the quality of data and the predictive capacities of the algorithm, how can we best define data quality? And on the basis of what methodological approach? If data merely provides for representations of the real world which are not necessarily good representations they can give rise to acts of discrimination. How should we take this issue into account when regulating AI? Is there a way to tackle algorithmic biases? Sensitive information can be used to set up predictive models, which can be utilized in improving the effectiveness of political campaigns and administrative functions. How does this interfere with power dynamics in western constitutional democracies? How can the consequences of algorithmic biases that emerge in the sphere of public functions be limited?
15.45 – 16.00	**Conclusions** *Juan Carlos De Martin*
16.00 – 18.00	**Visit to the Giacometti Collection (Kunsthaus Zürich)**
18.00	**Apéro followed by dinner** Wirtschaft Neumarkt, Neumarkt 5, 8001 Zurich.

16th February 2018

Persistent Tracking, Mass Surveillance and Commercial Purposes
Case study 2: Even better and cheaper, face recognition technology is spreading

9.15 – 9.30	**Introduction to case study** *Malavika Jayaram*
9.30 – 12.30 (coffee break at 11.00)	**Artificial Intelligence, persistent tracking and mass surveillance** *Moderation of debate: Malavika Jayaram* Persistent tracking technologies are currently used for security and surveillance purposes in order to identify suspects, or for various police initiatives. It is not a case that in certain states (e.g. China), AI and persistent tracking is connected to pervasive forms of censorship. A series of intriguing questions arise in this respect. If privacy is the shield against the exploitation of AI for mass surveillance, how is the function of privacy predicted to evolve? How can the trade-off between reinforcing an individual's privacy and stifling innovation in AI be handled by nation states? What role would AI and persistent tracking play in shaping the essence of law and governance? What is regulation likely to become, for which benefits and at what costs? How could a perfectly tailored regulation impact the power dynamics?
12.30	**Lunch**
13.30 – 15.30 (coffee break at 14.15)	**Artificial Intelligence for commercial purposes. An unprecedented global governance?** *Moderation of debate: John Palfrey* The main development of persistent tracking systems using AI is being driven by the private sector. In particular, social networks, the smartphone industry, the digital advertising industry and the entertainment industry are the main innovators in this field. Different questions emerge in this regard. How could an understanding of data as 'economic resource' be implemented in regulating data-driven technologies? Is sensitive information and persistent tracking relevant from a competition law perspective? Is there a need to reinvigorate the essential facility doctrine and to strengthen the concept of interoperability? How will technological regulation by private corporations redefine the current power dynamics? Is this behavioral shaping likely to impact the rule of law? How should the use of persistent tracking be regulated to preserve human self-determination and autonomy? What are the main consequences of AI and persistent tracking on individuals' contractual freedom? Is the technological regulation likely to evolve into a form of global governance?
15.30 – 16.00	**General conclusions** *Ryan Budish*

CASE STUDY 1

Advances in AI are used to spot signs of sexuality

Machines that read faces are coming

The Economist, 9 September 2017[1]

SUMMARY

In the past decade, Artificial Intelligence has been used to predict rare diseases, people's ages or to map poor regions from satellite images by spotting hidden patterns in large volumes of data, which no human could have done before. But this represents just the tip of the iceberg. A recent research study at Stanford University has found that facial recognition technologies can infer the sexual orientation of a person by analyzing that person's face and picking up on subtle differences of the facial structure. The software achieved a 91% accuracy rate when analyzing men's faces and respectively 83% for women's faces. The facial images were first taken from a popular dating website and then fed into a piece of software, which would produce 'face prints', long strings of numbers representing each person. Subsequently, a predictive model was set up to find correlations between the features of those face prints and their owners' sexuality (as declared on the dating website).
The study has some limitations though. In fact, it has been observed that images from a dating site are likely to be particularly revealing of sexual orientation; hence outside the lab – in the real world – the accuracy rate would be much lower. Despite these weaknesses, however, the experiment has shown that creating a piece of software capable of detecting intimate personal information (such as sexual orientation) is now a reality. This may potentially lead researchers in this field to train their systems to reveal other intimate personal traits, such as IQ or political views. Privacy violations, in this respect, will be an inevitable future consequence thereof. After all, a similar system has already been deployed to target voters during the last U.S. presidential campaign. And, not surprisingly, it has received a lot of criticism.

[1] *See* The Economist, 9 September 2017, https://www.economist.com/news/science-and-technology/21728614-machines-read-faces-are-coming-advances-ai-are-used-spot-signs.

CASE STUDY 2

Even better and cheaper, face recognition technology is spreading

China's Megvii has used government-collected data to lead the sector

The Economist, 9 September 2017[2]

SUMMARY

The headquarters of Megvii in Beijing resembles Big Brother's engine room: Cameras are used to recognize visitors in the firm's lobby in the blink of an eye while other such devices are deployed around the office. Since its founding in 2011, the company has made significant progress in developing cutting-edge face-recognition technology ("FRT") called Face++, which is widely used by more than 300,000 companies and individuals, making Megvii the first billion-dollar startup in the "facial-industrial complex". Even though the market is still small, FRT has started to permeate the wider business landscape due to a massive improvement in its accuracy rate and it is therefore poised to follow in the footsteps of speech recognition, which gained wide popularity as its accuracy improved.

FRT can be separated into two categories: The underlying capability and the applications that make use of it. Megvii's Face++ falls into the first category, as do similar offerings from Amazon, IBM and Microsoft, which provide face recognition as a cloud-computing service. Having access to the Chinese government's image database of 700m citizens, Megvii's service is able to rely on good data. Chinese state agencies have taken a particular interest in this technology. For instance in Shenzhen, government agencies use FRT to identify jaywalkers, whilst in Beijing the municipality has started to use the technology to catch thieves stealing toilet paper in public restrooms. FRT applications, on the other hand, are spreading even faster and are mainly being used by private corporations. Ant Financial, a subsidiary of Alibaba used its "Smile to Pay" system for the first time in a physical store. Facebook's algorithms can recognize tags on photos. Google uses FRT to group pictures that people have uploaded to its photo service. Amazon is following suit by incorporating a camera in its home speaker, Echo Look. US airlines have taken initial steps to match passengers' faces to passport photos with the aim of eliminating boarding passes. FRT also has the potential to lift sales by recognizing loyal customers and VIPs who deserve special treatment or by detecting dissatisfaction on shoppers' faces. The spread of such facial recognition services has already prompted efforts to thwart them. An Israeli startup, for instance, has developed software that slightly alters photos so that algorithms cannot recognize them. Whilst it seems that a struggle between opponents and supporters of FRT will ensue, it is, however, unlikely that such efforts will keep FRT from being widely used.

[2] *See* The Economist, 9 September 2017, https://www.economist.com/news/business/21728654-chinas-megvii-has-used-government-collected-data-lead-sector-ever-better-and-cheaper.