Big data mining, fairness and privacy

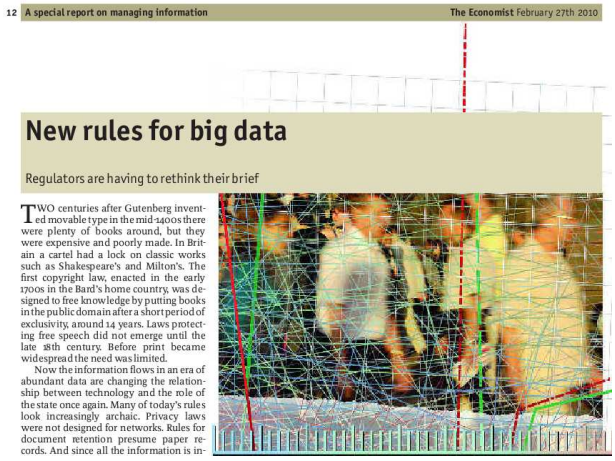
A vision statement towards an interdisciplinary roadmap of research

Dino Pedreschi, *Knowledge Discovery and Data Mining Lab, University of Pisa*
Toon Calders, *Information Systems Group, Eindhoven University of Technology*
Bart Custers, *eLaw - Centre for Law in the Information Society, Leiden University*
Josep Domingo-Ferrer, *UNESCO Chair in Data Privacy, Universitat Rovira i Virgili*
Giusella Finocchiaro, *Department of Law, University of Bologna*
Fosca Giannotti, *Knowledge Discovery and Data Mining Lab, ISTI-CNR, Pisa*
Morag Goodwin, *Tilburg Institute for Law, Technology, and Society, Tilburg University*
Mireille Hildebrandt, *Erasmus School of Law, Rotterdam and Centre for Law Science Technology and Society Studies, Vrije Universiteit Brussel*
Stan Matwin, *School of Information Technology and Engineering, University of Ottawa*
Yucel Saygin, *Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul*
Bart Schermer, *eLaw - Centre for Law in the Information Society, Leiden University*
Tal Zarsky, *Faculty of Law, Haifa University*

We live in times of unprecedented opportunities of sensing, storing and analyzing micro-data on human activities at extreme detail and resolution level, at society scale. Wireless networks and mobile devices record the traces of our movements. Search engines record the logs of our queries for finding information on the web. Automated payment systems record the tracks of our purchases. Social networking services record our connections to friends, colleagues, collaborators.

Ultimately, these **big data of human activity** are at the heart of the very idea of a

knowledge society: a society where decisions – small or big, by businesses or policy makers or ordinary citizens – can be informed by reliable knowledge, distilled from the ubiquitous digital traces generated as a side effect of our living. Increasingly sophisticated data analysis and data mining techniques support knowledge discovery from human activity data, enabling the extraction of models, patterns, profiles, simulation, what-if scenarios, and rules of human and social behavior – a steady supply of knowledge which is needed to support a knowledge-based society. The *Data Deluge* special report in "The Economist" in February 2010 witnesses exactly how this way of thinking is now permeating the entire society, not just scientists.



12 A special report on managing information The Economist February 27th 2010

New rules for big data

Regulators are having to rethink their brief

TWO centuries after Gutenberg invented movable type in the mid-1400s there were plenty of books around, but they were expensive and poorly made. In Britain a cartel had a lock on classic works such as Shakespeare's and Milton's. The first copyright law, enacted in the early 1700s in the Bard's home country, was designed to free knowledge by putting books in the public domain after a short period of exclusivity, around 14 years. Laws protecting free speech did not emerge until the late 18th century. Before print became widespread the need was limited.

Now the information flows in an era of abundant data are changing the relationship between technology and the role of the state once again. Many of today's rules look increasingly archaic. Privacy laws were not designed for networks. Rules for document retention presume paper records. And since all the information is in-

The capability to collect and analyze massive amounts of data has already transformed fields such as biology and physics, and now the human activity data cause the emergence of a data-driven "computational social science": the analysis of our digital traces can create new comprehensive pictures of individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies.

The paradigm shift towards human knowledge discovery comes, therefore, with unprecedented opportunities and risks: *we the people are at the same time the donors of the data that fuel knowledge discovery and the beneficiaries – or the targets – of the resulting knowledge services*. The paradoxical situation we are facing today, though, is that we are fully running the risks without fully grasping the opportunities of big data: on the one hand, we feel that our private space is vanishing in the digital, online world, and that our personal data might be used without feedback and control; on the other hand, the very same data are seized in the databases of global corporations, which use privacy as a reason (or excuse?) for not sharing it with science and society at large.

In the CNN show *Fast Future Forward*, the anchor-woman asked the panel of futurists: "Which major challenge are we going to have to deal with 10 years out?" The answer was: "We'll have to completely reverse our orientation to privacy. The reality is that we don't have privacy anymore: you use your cell phone, you drive your car, you go on-line, and it's gone." Although the debate on the vanishing privacy is going on for years, it is striking that now it is posed as a major problem in a popular prime-time show on the future trends of society and technology. The message is clear: privacy, once a human

right, is becoming a chimera in the digital era.

In the other extreme, knowledge discovery and data science run the risk of becoming the exclusive and secret domain of private companies – Internet corporations such as Google, Facebook, Yahoo, or big telecom operators – and government agencies, e.g. national security. For these data custodians, privacy is a very good excuse to protect their interests and not share the data, while users are not really aware how the data they generated are used. Alternatively, there might emerge a cast of privileged academic or industry researchers who are granted access over private big data from which they produce results that cannot be assessed or replicated, because the data they are based on cannot be shared with the scientific community. Neither scenario will serve the long-term public interest of accumulating, verifying, and disseminating knowledge.

Should we really give up and surrender to a wild digital far west, where people are exposed to risks, without protection, transparency, and trust, while society as a whole and science get little reward with respect to the opportunities offered by the secluded big data? We believe that another knowledge technology is possible, a fair knowledge discovery framework can be devised, where opportunities are preserved and risks are kept under control. A **technology to set big data free for knowledge discovery, while protecting people from privacy intrusion and unfair discrimination**.

In order to fully understand the risks, we should consider that the knowledge life-cycle has two distinct, yet intertwined, phases: *knowledge discovery and knowledge deployment*. In the first step, knowledge is extracted from the data; in the second step, the discovered knowledge is used in support of decision making; the two steps may repeat over and over again, either in off-line or in real-time mode. For instance, knowledge discovery from patients' health records may produce a model which predicts the insurgence of a disease given a patient's demographics, conditions and clinical history; knowledge deployment may consist in the design of a focused prevention campaign for the predicted disease, based on profiles highlighted by the discovered model. Hence, people are both the data providers and the subjects of profiling. In our vision, the risks in each of the two steps in the knowledge life-cycle are:

- *Privacy violation*: during knowledge discovery, the risk is unintentional or deliberate intrusion into the personal data of the data subjects, namely, of the (possibly unaware) people whose data are being collected, analyzed and mined;
- *Discrimination*: during knowledge deployment, the risk is the unfair use of the

discovered knowledge in making discriminatory decisions about the (possibly unaware) people who are classified, or profiled.

Continuing the example, individual patient records are needed to build a prediction model for the disease, but everyone's right to privacy means that his/her health conditions shall not be revealed to anybody without his/her specific control and consent. Moreover, once the disease prediction model has been created, it might also be used to profile the applicant of a health insurance or a mortgage, possibly without any transparency and control. It is also clear, from the example, how the two issues of profiling and privacy are strongly intertwined: the knowledge of a health risk profile may lead both to discrimination and to privacy violation, for the very simple fact that it may tell something intimate about a person, who might be even unaware of it.

Privacy intrusion and discrimination prevent the acceptance of human knowledge discovery: if not adequately countered, they can undermine the idea of a fair and democratic knowledge society. The key observation is that they have to be countered together: focusing on one, but ignoring the other, does not suffice. Guaranteeing data privacy while discovering discriminatory profiles for social sorting is not so reassuring: it is just a polite manner to do something very nasty. So is mining knowledge for public health and social utility, if, as a side effect, the personal sensitive information that feeds the discovery process is disclosed or used for purposes other than those for which it has been collected, putting people in danger. On the contrary, protecting data privacy and fighting discrimination help each other: methods for data privacy are needed to make the very sensitive personal information available for the discovery of discrimination. If there is a chance to create a trustworthy technology for knowledge discovery and deployment, it is with a holistic approach, not attempted so far, which faces privacy and discrimination as two sides of the same coin, leveraging on inter-disciplinarity across IT and law. The result of this collaboration should enhance trust and social acceptance, not on the basis of individual ignorance of the risks of sharing one's data, but on a reliable form of risk

measurement. By building tools that provide feedback and calculated transparency about the risk of being identified and/or discriminated, the idea of consent and opt-in may become meaningful once again.

In summary, a research challenge for the information society is the definition of a theoretical, methodological and operational framework for fair knowledge discovery in support of the knowledge society, where fairness refers to privacy-preserving knowledge discovery and discrimination-aware knowledge deployment. The framework should stem from its legal and IT foundations, articulating data science, analytics, knowledge representation, ontologies, disclosure control, law and jurisprudence of data privacy and discrimination, and quantitative theories thereof. We need novel, disruptive technologies for the construction of human knowledge discovery systems that, **by design**, offer native techno-juridical safeguards of data protection and against discrimination. We also need a new generation of tools to support legal protection and the fight against privacy violation and discrimination, powered by data mining, data analytics, data security, and advanced data management techniques.

The general objective should be the reformulation of the foundations of data mining in such a way that privacy protection and discrimination prevention are embedded the foundations themselves, dealing with every moment in the data-knowledge life-cycle: from (off-line and on-line) data capture, to data mining and analytics, up to the deployment of the extracted models. We know that technologies are neither good nor bad in principle, but they never come out as neutral. Privacy protection and discrimination prevention have to be included in the basic theory supporting the construction of technologies. Finally, the notions of privacy, anonymity and discrimination are the object of laws and regulations and they are in continuous development. This implies that the technologies for data mining and its deployment must be flexible enough to embody rules and definitions that may change over time and adapt in different contexts.

The debate around data mining, fairness, privacy, and the knowledge society is going to become central not only in scientific research, but also in the policy agenda at national and supra-national level. We conclude our discussion by proposing a list of the top ten

research questions, as a contribution to set the research roadmap around fair knowledge discovery and data mining.

1. How to define fairness: how to actually measure if privacy is violated, identity is disclosed, discrimination took place?

2. How to set data free: how to make human activity data available for knowledge discovery, in specific contexts, while ensuring that freed data have a reliable measure of fairness?

3. How to set knowledge free, while ensuring that mined knowledge has been constructed without an unfair bias? How to guarantee that a model does not discriminate and does not compromise privacy?

4. How to adapt fairness in different contexts, which raise different legitimate expectations with regard to privacy and non-discrimination? To what extent should discrimination on the basis of a higher risk be defined as discriminatory? Is this a legal issue or an ethical issue?

5. How to make the data mining process parametric w.r.t a set of constraints specifying the privacy and anti-discrimination rules that should be embedded in freed data and mined knowledge? How to take into account in the process the analytical questions that will emerge after having mined the data?

6. How to prove that a given software is fair, if it is claimed to be fair? How can the relevant authorities check whether e.g. a privacy policy or a non-discrimination policy is complied with?

7. What incentive does the industry have to opt for fair data mining technologies? What incentive do individual consumers have to opt for service providers that employ such technologies?

8. How to provide transparency about services that employ profiling and are potentially discriminatory? Which effective remedies do citizens have if they suspect unfair discrimination on the basis of data mining technologies (cf. art. 12 Directive 95/46/EC and the Anti-Discrimination Directives 2000/43/EC 2000/78/EC, 2004/113/EC, 2006/54/EC)?

9. Which incentives could be provided by the European legislator to employ fair data mining technologies (both on the side of the industry and on the side of individual consumers)? E.g., compulsory certification, new legal obligations, technical auditability schemes, new individual rights, new data protection competences?

10. How to specify and implement fairness on-line, so as to guarantee privacy and discrimination freedom at the very moment of data acquisition or capture?