Noname manuscript No.
(will be inserted by the editor)

Introduction to Special Issue on Computational Methods for Enforcing Privacy and Fairness in the Knowledge Society

Sergio Mascetti · Annarita Ricci · Salvatore Ruggieri

Received: date / Accepted: date

We live in times of unprecedented opportunities for sensing, storing and analyzing micro-data on human activities at extreme detail and resolution, at a societal level. 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, and 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 business or policy makers - can be taken on the basis of reliable knowledge, distilled from the ubiquitous digital traces generated as a side effect of our living. Increasingly sophisticated intelligent systems support knowledge discovery and deployment from human activity data, enabling the extraction and the (often automatic) use of models, patterns, profiles, and rules of human behavior. This paradigm shift towards the knowledge society comes, however, with critical risks for human rights:

- *Privacy violation*: during knowledge extraction, the risk is of unintentional or deliberate intrusion into the personal data of 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 people who are classified, or profiled.

S. Mascetti
Department of Computer Science, University of Milan, Italy
E-mail: sergio.mascetti@di.unimi.it

A. Ricci
Department of Legal Studies, University of Bologna, Italy E-mail: annarita.ricci@unibo.it

S. Ruggieri
Department of Computer Science, University of Pisa, Italy, E-mail: ruggieri@di.unipi.it

People feel that private space is vanishing in the digital world, and that personal data can be used without feedback and control, possibly resulting in unfair decisions. To combat these threats, it becomes of primary importance to develop multi-disciplinary approaches, in which Law and Computer Science should influence each other reciprocally, each one guiding the other's research. Indeed, on one side Computer Science solutions for enforcing data protection and for preventing discrimination, run the risk of becoming a dead end if they are not modeled according to the regulations in force while, on the other side, Law may become a weak protection in itself, if legal norms do not take into consideration the rapid evolution of ICT techniques and methods. The Call for Papers of this special issue reflected a number of challenging research topics in the area of Artificial Intelligence and Law, including:

- methods for enforcing data privacy and anonymity
- methods for data portability, and for the right to oblivion
- methods for data protection and law enforcement
- privacy by design in intelligent systems
- privacy-preserving data mining
- privacy policies in social networks
- context-aware location privacy
- methods for unbiased data collection and processing
- methods for enforcing fairness in profiling and targeting
- methods for discrimination discovery from data
- statistical measures of discrimination
- methods for discrimination prevention in data mining
- computational argumentation in discrimination analysis
- design of (quasi-)experimental methods
- computational models of segregation in social networks
- computational models of evidential reasoning
- tools and systems, with case studies.

This volume presents four papers which address, in an extensive and thorough way, several problems of privacy violation or discrimination with a multi-disciplinary approach. Each paper was reviewed by (at least) two computer scientists and (at least) one legal expert.

Kwecka et al. propose an innovative perspective: to consider privacy as a public good, instead of an individual right. This view has clearly many legal and technical implications. On the technical side, the authors show how to adapt this approach in the field of data distribution among public authorities. From the legal point of view, the central topic of this paper is the concept of privacy: the meaning of this value and the variables of its meaning in a society where information, even from an economic point of view, is gaining an increasing importance.

Monreale et al. tackle the problem of disclosing sequential information, like transactional or location data. The problem with this kind of information is that it can reveal the identity of the data respondents. The proposed solution is an extension of the k-anonymity principle that guarantees formal safety

in the considered model while making it possible to preserve some relevant knowledge discovery analyses of these data. From the legal point of view, the authors consider an aspect of great interest: the concept of privacy by design. Such a concept is very topical also considering the recent "Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data," dated 25 January 2012.

Berendt and Preibusch fill a gap in the emerging domain of discrimination-aware data mining (DADM), where existing approaches take a constraint-oriented perspective by removing or preventing some (discriminatory) patterns to be used in decision making. They argue for complementing such approaches with exploratory DADM, where discriminatory patterns are discovered and flagged rather than suppressed. The paper discusses the relative merits of these two perspectives, both conceptually and in an empirical case study on loan applications. From the legal point of view, the paper focuses on the general principle of equality, on the related need to prevent unjust discrimination and on the need to balance these fundamental rights with automated decisions based on collected personal data.

Mancuhan and Clifton deal with the problem of social discrimination detection and protection in data mining. They devise an original approach, using Bayesian network models to detect individuals who have been discriminated against, which is able to tackle both direct and indirect discrimination. The approach is applied both to discover individuals who have been discriminated against in past decision records, and to correct Bayesian network classification algorithms in order not to learn to discriminate. From the legal point of view, the paper, moving from the prohibition of every use of a protected attribute in making a decision about an individual, focuses on some jurisprudencial cases, in order to reflect about the need to prevent discrimination.

We congratulate the authors of the above papers for their excellent work in advancing the state-of-the-art of the field. We are grateful to the journal's editors in chief for the opportunity to work on this exciting special issue, and to the reviewers for their helpful and constructive comments.

Acknowledgements

The guest editors gratefully acknowledge the support of the project *Enforce: Computer science and legal methods for enforcing the personal rights of non-discrimination and privacy in ICT systems* (2010-2014) funded by the Italian Ministry for University and Research under the Italian Basic Research Fund (FIRB) call *Futuro in Ricerca* 2008 (Project ID: RBFR081L58). The views expressed are those of the authors and should not be taken as representative of the project.