EU Data Protection Law: An Ally for Scientific Reproducibility?

Mireille Hildebrandt
Vrije Universiteit Brussel, Belgium
Radboud University, Netherlands
m.hildebrandt@cs.ru.nl

ABSTRACT
This keynote will introduce some of the key concepts of European data protection law, and clarify how and why this is not equivalent with privacy law. Next, I will explain why and how EU data protection law could enhance the methodological integrity of machine learning applications, also in the domain of multimedia.
The question is, first, how the General Data Protection Regulation (GDPR) applies to inferences captured from multimedia data. This raises a number of questions. Does it matter whether such data has been made public by the person it relates to? Does processing personal data always require consent? What counts as valid consent? What if the inferences are mere statistics? What does the prohibition of processing ‘sensitive data’ (ethnicity, health) mean for multimedia analytics? This keynote will provide a crash course in the underlying ‘logic’ of the GDPR [3], with a focus on what is relevant for inferences based on multimedia content and metadata.
I will uncover the purpose limitation principle as the guiding rationale of EU data protection law, protecting individuals against incorrect, unfair or unwarranted targeting.
In the second part of the keynote I will explain how the purpose limitation principle relates to machine learning research design, requiring keen attention to specific aspects of methodological integrity [2]. These may concern p-hacking, data dredging, or cherry picking performance metrics, and connect with the reproducibility crisis in machine learning that is on the verge of destroying the reliability of ML applications [1].

CCS Concepts/ACM Classifiers
• Applied computing ~ Law, social and behavioral sciences

Author Keywords
General Data Protection Regulation; machine learnin

BIOGRAPHY
Mireille Hildebrandt is a Research Professor on ‘Interfacing Law and Technology’ at Vrije Universiteit Brussel (VUB), appointed by the VUB Research Council. She is co-director of the research group on Law Science Technology and Society studies (LSTS) at the Faculty of Law and Criminology. She also holds the part-time Chair of ‘Smart Environments, Data Protection and the Rule of Law’ at the Science Faculty, at the Institute for Computing and Information Sciences (iCIS) at Radboud University. She has been teaching law to master students of computer science for the past eight years, resulting in the first serious introduction into law for ‘computer scientists and other folk’, to be published in open access by OUP later this year. See the MIT pubpub site for the open review of the manuscript.

Her research interests concern the implications of automated decisions, machine learning and mindless artificial agency for law and the rule of law in constitutional democracies. Hildebrandt has published 4 scientific monographs, 22 edited volumes or special issues, and over 100 chapters and articles in scientific journals and volumes. Her most recent monograph is ‘Smart Technologies and the End(s) of Law’ (Edward Elgar 2015). In 2018 she received an ERC Advanced Grant for her project on ‘Counting as a Human Being in the era of Computational Law’ (2019-2024): www.cohubicol.com, which will bring together a legal team from VUB with a CS team from Radboud University to explore the mine-field of e.g. quantified legal prediction and self-executing contracts.
ACKNOWLEDGMENT

The research on which this keynote is based was partly funded by the European Research Council (ERC) under the HORIZON2020 Excellence of Science program ERC-2017-ADG No 788734.

REFERENCES

