

Law, Governance and Technology Series 20

Serge Gutwirth
Ronald Leenes
Paul de Hert *Editors*

Reforming European Data Protection Law

 Springer

Law, Governance and Technology Series

Issues in Privacy and Data Protection

Volume 20

Series Editors

Pompeu Casanovas
UAB Institute of Law and Technology
Bellaterra, Barcelona, Spain

Giovanni Sartor
European University Institute, Florence, Italy

Sub Series Editor

Serge Gutwirth
Law, Science, Technology and Society (LSTS)
Vrije Universiteit Brussels (VUB)
Brussel, Belgium

The *Law, Governance and Technology Series* is intended to attract manuscripts arising from an interdisciplinary approach in law, artificial intelligence and information technologies. The idea is to bridge the gap between research in IT law and IT-applications for lawyers developing a unifying techno-legal perspective. The series will welcome proposals that have a fairly specific focus on problems or projects that will lead to innovative research charting the course for new interdisciplinary developments in law, legal theory, and law and society research as well as in computer technologies, artificial intelligence and cognitive sciences. In broad strokes, manuscripts for this series may be mainly located in the fields of the Internet law (data protection, intellectual property, Internet rights, etc.), Computational models of the legal contents and legal reasoning, Legal Information Retrieval, Electronic Data Discovery, Collaborative Tools (e.g. Online Dispute Resolution platforms), Metadata and XML Technologies (for Semantic Web Services), Technologies in Courtrooms and Judicial Offices (E-Court), Technologies for Governments and Administrations (E-Government), Legal Multimedia, and Legal Electronic Institutions (Multi-Agent Systems and Artificial Societies).

More information about this series at <http://www.springer.com/series/13087>

Serge Gutwirth • Ronald Leenes • Paul de Hert
Editors

Reforming European Data Protection Law

 Springer

Editors

Serge Gutwirth
Paul de Hert
Law, Science, Technology & Society (LSTS)
Faculty of Law and Criminology
at the Vrije Universiteit Brussel
Brussels, Belgium

Ronald Leenes
Tilburg Institute for Law, Technology,
and Society (TILT)
Tilburg University
Tilburg, The Netherlands

ISSN 2352-1902

ISSN 2352-1910 (electronic)

ISBN 978-94-017-9384-1

ISBN 978-94-017-9385-8 (eBook)

DOI 10.1007/978-94-017-9385-8

Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2014954796

© Crown Copyright 2015

© Springer Science+Business Media Dordrecht 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The year 2014 is destined to be an important year for European Data protection. After lengthy discussions in the various committees, involving almost 4,000 amendments to the Commission's 2014 proposal, the European Parliament on 12 March 2014 adopted the proposal prepared by the committee chaired by MEPs Jan-Philipp Albrecht and Dimitrios Droutsas in the first reading. The waiting at the time of writing this foreword (June 2014) is for the position of the Council of Ministers on the Regulation. Once this is available, the European Parliament has to negotiate with the Council and the Commission on the final text.

The seventh annual Computers, Privacy and Data Protection (CPDP) conference was held in Brussels on 22, 23 and 24 January 2014, and was sharply influenced by the European Commission's new proposals and the discussions that led up to the almost 4000 amendments that were tabled by stakeholders within and outside Europe (e.g. the USA). The conference took place during a sort of 'interbellum'. At the time when contributors to the conference were preparing their papers and panels, the text of the draft Regulation was in flux. In October 2013, the European Parliament's influential LIBE Committee (Civil Liberties, Justice, and Home Affairs) had decided on the proposal to be forwarded to the Parliament. The LIBE version introduced changes to the original Commission proposal, for instance regarding the controversial 'right to be forgotten' provision (art. 17). As a result the legal reality at the conference differed from the one on which some authors based their texts.¹ Uncertainty regarding the final text of the Regulation also existed at the time of the conference itself; the European Parliament adopted the LIBE version in its first reading after the conference. And even when this volume will appear in print, we still may not know what the final Regulation will look like. This book volume reflects this state of affairs. It provides a reflection on the proposed changes in the data protection landscape that may appear outdated at the time of reading. The value of the contributions however remains because many of them extend beyond the actual regulation and hence have more principled value.

¹This is nothing new. Legal reality constantly changes due to changing legislation and case law.

The present book is one of the results of the seventh edition of the annual Brussels based International Conference on Computers, Privacy and Data Protection: *CPDP2014 Reforming Data Protection: The Global Perspective*. The conference welcomed almost 850 participants at ‘our’ venue – the magnificent *Les Halles* – while another 500 people were reached through free public events organized in the evenings, also in Brussels. The 3-day conference offered participants 60 panels, several workshops and special sessions, with 343 speakers from academia, the public and private sectors, and civil society.

Under a slightly adapted title – *Reforming European Data Protection Law* – this volume brings together 16 chapters offering conceptual analyses, highlighting issues, proposing solutions, and discussing practices regarding privacy and data protection. The first part of the book contains two chapters on one of the prominent recurring CPDP themes: profiling. The second part focuses on one of the important new directions in the Regulation: a focus on preventing privacy risks and harms through impact assessments. It contains discussions on the tools and methodologies for impact assessments, as well as case studies. The third part contains three chapters on the controversial Right to be Forgotten. It addresses the history of the proposed right, ten reasons why it should be forgotten and explores one of the important dimensions in forgetting: time. The fourth part contains two chapters on the purported trade-off between privacy and security. The final, fifth, part deals with ways to support privacy and data protection. It contains a chapter discussing the nature of the Data Protection reform and a chapter on people’s knowledge and awareness of privacy protection strategies. It furthermore offers three chapters on privacy by design and how to implement this in practice.

The chapters in this volume stem from two tracks. Six chapters (Chaps. 8, 9, 11, 13, 15 and 16) originate from responses to the conference’s call for papers and have thus already been presented during the conference. The remaining chapters (Chaps. 1, 2, 3, 4, 5, 6, 7, 10, 12 and 14) were submitted by some of the conferences’ invited speakers in the months following the conference.

All the chapters of this book have been peer reviewed and commented on by at least two referees with expertise and interest in the subject matter. Since their work is crucial for maintaining the scientific quality of the book, we would explicitly take the opportunity to thank them for their commitment and efforts: Rocco Bellanova, Colin Bennett, Paul Bernal, Laurent Beslay, Jean-François Blanchette Caspar Bowden, Ian Brown, Roger Brownsword, Peter Burgess, Denis Butin, Lee Bygrave, Jan Camenisch, Johann Cas, Roger Clarke, Claudia Diaz, Niels van Dijk, Simone Fischer-Hübner, Michael Friedewald, Lothar Fritsch, Raphael Gellert, Marieke de Goede, Seda Gürses, Rob Heyman, Mireille Hildebrandt, Dennis Hirsch, Joris van Hoboken, Chris Hoofnagle, Gerrit Hornung, Patrick Humblet, Paulan Korenhof, Eleni Kosta, Christopher Kuner, Marc Langheinrich, Marc van Lieshout, Gary T. Marx, Irma van der Ploeg, Charles Raab, Kjetin Rommetveit, Arnold Roosendael, Ronny Saelens, Joseph Savirimuthu, Jean Marc Van Ghyseghem, Diane Whitehouse, Brian Wynne and Tal Zarsky.

May this book meet the reader's expectations and contribute to the quality of the continuing debate about the future of privacy and data protection.

Brussels, Belgium
Tilburg, The Netherlands
Brussels, Belgium
30 June 2014

Serge Gutwirth
Ronald Leenes
Paul de Hert

Contents

Part I Profiling: A Persistent Core Issue of Data Protection and Privacy

- 1 Profiling Technologies and Fundamental Rights and Values: Regulatory Challenges and Perspectives from European Data Protection Authorities..... 3**
Francesca Bosco, Niklas Creemers, Valeria Ferraris, Daniel Guagnin, and Bert-Jaap Koops
- 2 On-line Behavioral Tracking: What May Change After the Legal Reform on Personal Data Protection 35**
Georgia Skouma and Laura Léonard

Part II Taming the Future: Assessments of Risks in the Sphere of Privacy and Data Protection

- 3 A Systematic Approach to the Legal Evaluation of Security Measures in Public Transportation 63**
Christian Ludwig Geminn and Alexander Roßnagel
- 4 Models and Tools for the Computational Support of Technology Impact Assessments, Applied in the Context of Mass Transportation 91**
Ronald R. Grau
- 5 Impact Assessments as Negotiated Knowledge 125**
Leon Hempel and Hans Lammerant
- 6 Data Processing in Employment Relations; Impacts of the European General Data Protection Regulation Focusing on the Data Protection Officer at the Worksite 147**
Clara Fritsch

Part III To Forget or Not to Forget? Or Is the Question: How to Forget?

- 7 Timing the Right to Be Forgotten: A Study into “Time” as a Factor in Deciding About Retention or Erasure of Data** 171
Paulan Korenhof, Jef Ausloos, Ivan Szekely, Meg Ambrose, Giovanni Sartor, and Ronald Leenes
- 8 The ‘Right to Be Forgotten’: Ten Reasons Why It Should Be Forgotten**..... 203
Christiana Markou
- 9 Tracing the Right to Be Forgotten in the Short History of Data Protection Law: The “New Clothes” of an Old Right** 227
Gabriela Zanfir

Part IV Does It Take Two to Tango: Privacy and Security?

- 10 Privacy Versus Security: Problems and Possibilities for the Trade-Off Model** 253
Govert Valkenburg
- 11 Privacy and Security – On the Evolution of a European Conflict**..... 271
Matthias Leese

Part V Designing and Supporting Privacy and Data Protection

- 12 Evolving FIPPs: *Proactive Approaches to Privacy, Not Privacy Paternalism*** 293
Ann Cavoukian
- 13 Evolution or Revolution? Steps Forward to a New Generation of Data Protection Regulation**..... 311
Attila Kiss and Gergely László Szőke
- 14 Do People Know About Privacy and Data Protection Strategies? Towards the “Online Privacy Literacy Scale” (OPLIS)** .. 333
Sabine Trepte, Doris Teutsch, Philipp K. Masur, Carolin Eicher, Mona Fischer, Alisa Hennhöfer, and Fabienne Lind
- 15 LEAP: The LEAP Encryption Access Project** 367
Elijah Sparrow and Harry Halpin
- 16 Enabling Privacy by Design in Medical Records Sharing** 385
Jovan Stevovic, Eleonora Bassi, Alessio Giori, Fabio Casati, and Giampaolo Armellin

Contributors

Meg Ambrose is an Assistant Professor in the Communication, Culture and Technology department at Georgetown University. Her research interests focus on the governance of emerging technology and cover a wide range of technology policy issues including comparative censorship and privacy law, engineering and information ethics, robotics law and policy, and the legal history of technology. She earned her J.D. at the University of Illinois and her Ph.D. in Technology, Media and Society from the University of Colorado, Engineering and Applied Science. Email: megLeta@gmail.com.

Giampaolo Armellin is in charge of the Research Unit at CRG – Centro Ricerche GPI, the research centre of the GPI company, co-located at the EIT ICT Labs in Trento. He received his M.Sc. in Computer Science at the University of Milan. He has experience in the design and development of systems for industrial automation, process control, CRM, contact centres and ERP for healthcare. Currently, he manages the research team at CRG, working on research projects in eHealth and eWelfare domains. He has been one of the industrial representatives in the ALLOW project on pervasive computing and collaborates with the Information Engineering and Computer Science department of the University of Trento and the Fondazione Bruno Kessler. He is a member of the Advisory Board of the Design Thinking Centre (Doctoral Training Centre) on ICT for Quality of Life in Trento. Email: giampaolo.armellin@cr-gpi.it.

Jef Ausloos is a Doctoral Researcher at the University of Leuven, Faculty of Law (iMinds – ICRI). He has worked both in academia and civil society organisations in Hong Kong and the USA and holds law degrees from the University of Namur (B.A.), University of Leuven (M.A.) and University of Hong Kong (LL.M). Jef Ausloos has written on a variety of topics in the area of privacy law and media law. Currently, his research focuses on issues that lie at the intersection of Privacy and Data Protection, Freedom of Expression and Intermediary Liability on the Internet. In his Ph.D., Jef Ausloos is looking more closely at the distribution of control over personal data between online corporate entities and data subjects. Email: Jef.ausloos@law.kuleuven.be.

Eleonora Bassi is a Research Fellow of the Department of Information Engineering and Computer Science of the University of Trento and a Fellow at the Nexa Center for Internet and Society of the Polytechnic University of Torino. She holds a Degree in Law and a Ph.D. in Philosophy of Law at the University of Torino. After her Ph.D., she focused her interest on Information Law, Fundamental Rights and Data Protection Issues, and later on Public Sector Information European legal framework and regional policies. Currently, her research follows two main directions. First, the focus is on the new European Data Protection framework that will have a strong impact on privacy rights in digital environments and the circulation of personal data within the information market. Second, her work is on policy-oriented research on Open Data and Big Data. Email: bassi@disi.unitn.it.

Francesca Bosco is Project Officer within the Emerging Crimes Unit in UNICRI, the United Nations Interregional Crime and Justice Research Institute. She earned a law degree in International Law and joined UNICRI in 2006 as a member of the Emerging Crimes Unit. In her role in this organization, Bosco is responsible for fundamental rights protection projects and for cybercrime prevention and cybersecurity projects. She is currently acting as Programme Manager of the project 'PROFILING: Protecting Citizens' Rights against Illicit Profiling' funded by the European Union's Fundamental Rights and Citizenship program. She is one of the founders of the Tech and Law Center and she is currently a Ph.D. candidate at the University of Milan. Email: bosco@unicri.it.

Fabio Casati is a Professor of Computer Science at the University of Trento. He received his Ph.D. from the Politecnico di Milano and then worked for over 7 years in Hewlett-Packard, USA, where he was technical lead for the research programme on business process intelligence. He has also contributed (as software and data architect) to the development of several HP commercial products and solutions in the area of web services and business process management. In Trento, he is leading or participating in five FP7 projects, is active in many industry-funded projects, both local and international, and has over 20 patents. His passions are now in social informatics, or, informatics at the service of the community. His latest efforts are on IT for better living, on collaborative programming, and on models for scientific disseminations that can help scientists work in a more efficient way. Email: casati@disi.unitn.it.

Ann Cavoukian is recognized as one of the leading privacy experts in the world. Noted for her seminal work on Privacy Enhancing Technologies (PETs) in 1995, her concept of *Privacy by Design* seeks to proactively embed privacy into the design specifications of information technology and accountable business practices, thereby achieving the strongest protection possible. Among a number of her roles and awards, Dr. Cavoukian serves as the Co-Chair of the OASIS Privacy by Design Documentation for Software Engineers Committee whose goal is to enable software organizations to embed privacy into the design and architecture of IT systems, without diminishing system functionality. In 2011, she was honoured

with the prestigious *Kristian Beckman Award* for her pioneering work on *Privacy by Design* and privacy protection in modern international environments. Email: Michelle.Chibba@ipc.on.ca.

Niklas Creemers is researcher at the Centre for Technology and Society (CTS) at Technische Universität Berlin (TUB). After graduating from University Duisburg-Essen, he worked in several research projects in security and privacy research, including projects like ‘SuSI-Team: Passengers’ Perceptions of Security in Public Transport Systems’, ‘SIMKAS-3D: Simulation of Cascading Crisis in Urban Critical Infrastructure Systems as a 3-D Model’, and ‘Dynamic Arrangements of Urban Cultures of Security – DynASS’, all funded by the German Federal Ministry of Education and Research (BMBF). Furthermore Creemers participated in the EU-funded project ‘PATS: Privacy Awareness through Security Organisation Branding’. He is currently part of CTS’s research team on the project ‘PROFILING: Protecting Citizens’ Rights against Illicit Profiling’ funded by the European Union’s Fundamental Rights and Citizenship program. Email: creemers@ztg.tu-berlin.de.

Paul de Hert is Professor of Law at the Faculty of Law and Criminology of Vrije Universiteit Brussel. He is the Director of the research group on Fundamental Rights and Constitutionalism (FRC) and senior member of the research group on Law, Science, Technology & Society (LSTS). Paul de Hert is also Associated-Professor Law and Technology at the Tilburg Institute for Law and Technology (TILT). Email: paul.de.hert@vub.ac.be.

Carolyn Eicher is a bachelor student of Communication Studies at the University of Hohenheim, Germany, in her 5th semester. During winter term 2013/2014 she was part of a student research project within the Humboldt Reloaded 3.0 program at the Department of Media Psychology at the University of Hohenheim and thereby contributed to the development of the Online Privacy Literacy Scale (OPLIS).

Valeria Ferraris, Ph.D., is Research Fellow and Adjunct Professor in Criminology at the University of Turin. She is also a member of the board of Amapola, an agency that carries out research and interventions on security and urban liveability, liberties and rights. In 2008 she earned her doctorate in Criminology at the Catholic university in Milan. She is the author of several research reports and essays on immigration and crime, trafficking in human beings, urban security and human rights. She is currently working in the project ‘PROFILING: Protecting Citizens’ Rights against Illicit Profiling’ funded by the European Union’s Fundamental Rights and Citizenship program. Email: valeria.ferraris@amapolaprogetti.org.

Mona Fischer is currently pursuing a master’s degree in communication management from University of Hohenheim, Germany, specializing in media psychology. In addition to her formal academic program, she also works as a student assistant at the Department of Media Psychology. Before arriving at University of Hohenheim, Mona Fischer completed her undergraduate degree in communication science and psychology at the Ludwig-Maximilians-Universität Munich, Germany.

Clara Fritsch cooperated on several national and international research on industrial relations and working conditions after studying sociology and political science. Since 2007 she is employed at the Austrian Union for Private Sector Employees, Graphical Workers and Journalists, mostly dealing with employees' data protection, ICT as a management instrument and blurring boundaries of work and privacy. Email: clara.fritsch@gpa-djp.at.

Christian Ludwig Geminn is a research assistant at the Project Group Constitutionally Compatible Technology Design (provet) at Kassel University. Having studied German and English Law (at Johannes Gutenberg-Universität Mainz and De Montfort University Leicester), his research interests lie inter alia with technology law and the relationship between liberty and security. Email: c.geminn@uni-kassel.de.

Alessio Giori received his M.Sc. degree in Computer Science, with specialization in 'Data, Media and Knowledge Management' from the University of Trento in Italy. He also received a professional master degree in 'Technologies for e-Government' and a bachelor degree in Computer Science from the same university. He is currently working at Fondazione Graphitech since on two European research projects called eENVplus and LIFE+IMAGINE. eENVplus is a Pilot Type A project, funded by European Union under the Competitiveness and Innovation Framework Programme – Information and Communication Technologies Policy Support Programme (CIP-ICT-PSP). LIFE+IMAGINE is a project co-funded by European Union under the LIFE+ Programme Environmental Policy and Governance in the framework of the objective 'strategic approach'. Email: alessio.giori@graphitech.it.

Ronald R. Grau is a Visiting Research Fellow at the University of Sussex where he obtained his D.Phil. in Computer Science and Artificial Intelligence in 2009. His research focuses on knowledge systems and representation design, complex process modelling, diagrammatic representation and reasoning, and the computational support of scientific discovery. For many years, Ronald Grau has also worked in commercial contexts, involving various technologies and applications, including knowledge integration, e-learning, graphics virtualisation, SaaS, databases, enterprise software, and business intelligence. In his role as Senior Researcher at Kingston University London, he worked on the EU FP7 project SIAM (Security Impact Assessment Measures) as a co-investigator. The project aimed to establish measures for evaluating technological solutions and their impact on society, and to make these measures accessible through a computational assessment support system. Email: r.r.grau@kingston.ac.uk.

Daniel Guagnin, M.A., is researcher at the Centre for Technology and Society (CTS) at Technische Universität Berlin (TUB). He is currently working in the PROFILING Project funded by the Fundamental Rights and Citizenship Scheme of the European Union. Before PROFILING, Guagnin worked in EU projects in the field of Privacy, Data Protection Accountability and Security Technologies,

namely PATS and SIAM. His doctoral project focuses on the role of source code for the exchange of knowledge between software experts and lays in Free/Open Source Communities and is funded by a scholarship of the Konrad Adenauer Stiftung. Guagnin received his M.A. in Sociology from Albert-Ludwigs-University in Freiburg (Germany). Email: guagnin@ztg.tu-berlin.de.

Serge Gutwirth is a Professor of Human Rights, Legal Theory, Comparative Law and Legal Research at the Faculty of Law and Criminology of the Vrije Universiteit Brussel (VUB), where he studied law, criminology and also obtained a postgraduate degree in technology and science studies. Gutwirth founded and still chairs the VUB-research group *Law Science Technology & Society* (<http://www.vub.ac.be/LSTS>). He publishes widely in Dutch, French and English. Serge Gutwirth is particularly interested both in technical legal issues raised by technology (particularly in the field of data protection and privacy) and in more generic issues related to the articulation of law, sciences, technologies and societies. Email: serge.gutwirth@vub.ac.be.

Harry Halpin is a World Wide Web Consortium (W3C) Team member and a Research Scientist at the Computer Scientist and AI Lab at MIT, and a visiting researcher at Centre Pompidou. His technical and philosophical work is aimed at evolving the Web towards becoming a secure platform for free communication in order to enable collective intelligence. He received a Ph.D. in Informatics at the University of Edinburgh and author of the book *Social Semantics*, with over 40 publications in areas ranging from search engines to the philosophy of cognitive science. He is also President of the Board of LEAP (LEAP Encryption Access Project). Email: hhalpin@w3.org.

Leon Hempel, Ph.D., studied Comparative Literature and Political Science. He is a senior researcher at the Centre for Technology and Society (CTS) at the Technical University of Berlin since 1999. His research areas are sociology of technology, risk and innovation as well as technology assessment. Currently he is carrying out two ethnographic studies focusing on control room practices in large infrastructure as well as on surveillance practices in the course of football matches. In the last 10 years he continuously built up the social science security research at the CTS on the basis of national, European as well as international research projects including URBANEYE, PATS and SIAM. Email: hempel@ztg.tu-berlin.de.

Alisa Hennhöfer is a bachelor student of Communication Studies at the University of Hohenheim (Germany) in her 5th semester. During winter term 2013/2014, she was part of a student research project within the Humboldt Reloaded 3.0 program at the Department of Media Psychology at the University of Hohenheim and thereby contributed to the development of the Online Privacy Literacy Scale (OPLIS).

Attila Kiss is a doctoral student and young researcher at the University of Pécs, Faculty of Law, Research Center for Information and Communication Technology Law, Hungary. Since 2011 he has participated in the Ph.D. programme specialized

in IT and IP Law, and his research centres on the legal protection of the individual's appearance and the regulation of surveillance cameras in public places from the personal data protection aspect. Email: kiss.attila@ajk.pte.hu.

Bert-Jaap Koops is Professor of Regulation and Technology at the Tilburg Institute for Law, Technology, and Society (TILT), the Netherlands. His main research fields are cybercrime, cyber-investigation, privacy, and data protection. He is also interested in topics such as DNA forensics, identity, digital constitutional rights, 'code as law', and regulatory implications of human enhancement, genetics, robotics, and neuroscience. With a personal postdoc (1999), VIDI (2003) and VICI (2014) grant, Koops is one of the few Dutch researchers who received all three stages of NWO's (Netherlands Organisation for Scientific Research) personal research-grant scheme. Email: e.j.koops@uvt.nl.

Paulan Korenhof is a Ph.D. student at the Tilburg Institute for Law, Technology, and Society (TILT, Tilburg University) and at Privacy & Identity Lab (PI. Lab, a collaboration between Radboud University, SIDN, Tilburg University and TNO). In these institutes, Paulan works together with legal scholars, computer scientists, philosophers and social scientists. Both the interdisciplinary character of these institutes and the interdisciplinary scientific roots of Paulan herself (she holds a masters degree in both Philosophy and Public Law) shape her Ph.D. research in which she explores the problems caused for individuals by the 'memory' of the World Wide Web. She approaches this topic from an applied philosophy of technology perspective with an eye on the mutual shaping of law, technology and society. Email: P.E.I.Korenhof@uvt.nl.

Hans Lammerant has studied philosophy (VUB, 1996) and law (VUB/UIA, 2004). He also has a candidate degree in industrial engineering (1992) and is currently finalizing a Master in Statistics (UGent). Previously he has worked in civil society organisations on peace and human rights issues. In his research, he focuses on the effect of new developments in data science and statistics on surveillance and privacy. More generally he is interested in how technological developments and globalization influence the development of new forms of exercising power and how this impacts law. Email: hans.lammerant@vub.ac.be.

Ronald Leenes is Full Professor in Regulation by Technology at a the Tilburg Institute for Law, Technology, and Society (TILT), the Netherlands. His primary research interests are privacy and identity management, techno-regulation, applied data protection regulation, big data, conceptual analysis of privacy and data protection, and robotics and human enhancement. Currently his work focuses on accountability and transparency in big data and the cloud. He was responsible for TILT's research in several EU projects, such as PRIME, PRIMELIFE, ENDORSE, Robolaw and A4Cloud, and has contributed extensively to NoE FIDIS. Email: R.E.Leenes@tilburguniversity.edu.

Matthias Leese is a researcher at the International Centre for Ethics in the Sciences and Humanities (IZEW), University of Tuebingen. His primary research interests lie in the fields of critical security studies, privacy/data protection, surveillance studies and STS. Email: matthias.leese@izew.uni-tuebingen.de.

Laura Léonard is Consultant in the Department of Security and Privacy of *Deloitte Bedrijfsrevisoren/Réviseurs d'Entreprises*, Belgium, and assistant in commercial law at the University of Namur (UNamur). Laura is actively involved in the development and implementation of the Trademark Clearinghouse in the field of domain name industry and in various privacy and data protection projects in view of assessing privacy risks and helps companies reach compliance with data protection rules. Email: lauleonard@deloitte.com.

Fabienne Lind is a bachelor student of Communication Studies at the University of Hohenheim (Germany) in her 5th semester. During winter term 2013/2014, she was part of a student research project within the Humboldt Reloaded 3.0 program at the Department of Media Psychology at the University of Hohenheim and thereby contributed to the development of the Online Privacy Literacy Scale (OPLIS).

Christiana Markou is a Lecturer at the European University Cyprus and a practicing lawyer at Markou-Christodoulou & Polycarpou LLC, a law office based in Cyprus. She graduated from the University of Sheffield with an LL.B (Hons) in 1998 and an LL.M in 2000. She also holds a Ph.D. from the University of Lancaster awarded in 2011. Her research interests are mainly EU law, consumer protection law and information technology law. She currently is writing a book with Ashgate Publishing on agent technology in e-commerce and EU consumer protection law. Email: c.markou@euc.ac.cy.

Philipp K. Masur is a research assistant at the Department of Media Psychology at the University of Hohenheim, Germany. He is currently pursuing a Ph.D. in the field of media psychology. He studied communication science, business and philosophy at the University of Mainz, German, and the Macquarie University of Sydney, Australia. In his research, Philipp K. Masur focuses on how people manage their privacy in social media as part of their everyday practices. In his dissertation, he addresses how users of social media perceive situational factors of privacy and how far this perception influences privacy regulations and self-disclosure in the social web.

Alexander Roßnagel holds the Chair for Public Law with an emphasis on the law of technology and ecology at the faculty of economics at Kassel University, where he is also a Director of the Interdisciplinary Research Center for Information System Design (ITeG), Managing Director of the Competence Centre for Climate Change Mitigation and Adaptation (CliMA) and Scientific Director of the Project Group Constitutionally Compatible Technology Design (provet). Furthermore, he is a Fellow at the German Informatics Society (GI). Alexander Roßnagel has published

extensively on all aspects of technology law and environmental law, and has served as a consultant for the German Federal Government on several occasions. Email: a.rossnagel@uni-kassel.de.

Giovanni Sartor is part-time Professor in Legal Informatics at the University of Bologna and part-time Professor in Legal informatics and Legal Theory at the European University Institute of Florence. He obtained a Ph.D. at the European University Institute (Florence), worked at the Court of Justice of the European Union (Luxembourg), was a researcher at the Italian National Council of Research (ITTIG, Florence), held the chair in Jurisprudence at Queen's University of Belfast, and was Marie-Curie Professor at the European University of Florence. He has been President of the International Association for Artificial Intelligence and Law. He has published widely in legal philosophy, computational logic, legislation technique, and computer law. He is Co-director of the *Artificial Intelligence and Law* journal and Co-editor of the *Ratio Juris* journal. His research interests include legal theory, logic, argumentation theory, modal and deontic logics, logic programming, multiagent systems, computer and Internet law, data protection, e-commerce, law and technology, aviation law, and human rights. Email: sartor@cirfid.unibo.it.

Georgia Skouma has practised as lawyer specialised in Information Technology, Communications and Privacy Law since 2001. She is currently Director in the Department of Security and Privacy of *Deloitte Bedrijfsrevisoren/Réviseurs d'Entreprises*, Belgium. Her role as business legal adviser with Deloitte is to help corporate clients in developing and implementing risk-based solutions and procedures to meet their legal obligations in a number of areas, primarily in data protection, information security, privacy and e-government. Email: gskouma@deloitte.com.

Elijah Sparrow is a longtime anti-surveillance activist and co-founder of riseup.net, a provider of secure communication alternatives for social movements. He is the primary inventor and lead technical architect of the LEAP Encryption Access Project software. He is also the lead programmer for the free software project Crabgrass, a social collaborative space for activist groups with a focus on security. His research areas include digital surveillance and the communication repertoires of social movements. Email: elijah@leap.se.

Jovan Stevovic is a researcher at Centro Ricerche GPI, a research centre of the GPI Company where he is working on different research and industrial projects in the healthcare domain. He received his Ph.D. in Computer Science from the University of Trento in 2014 with the thesis on methodologies and technologies to enable privacy-aware data sharing in healthcare. As past experiences he participated in the development of electronic health record and data warehousing systems in the Province of Trento, solving problems related to medical record sharing such as data and service integration, privacy and security. His current research activity is focused on solving regulatory compliance and privacy issues related to the healthcare data management in software as a service environment. Email: jovan.stevovic@cr-gpi.it.

Ivan Szekely, social informatist, senior research fellow of the Open Society Archives at Central European University, Associate Professor at the Budapest University of Technology and Economics, and board member of the Eotvos Karoly Policy Institute. His research interests are focused on information autonomy, openness and secrecy, privacy, identity, memory and forgetting, and archivistics. Email: Szekelyi@ceu.hu.

Gergely László Szőke is a researcher at the University of Pécs, Faculty of Law, Research Center for Information and Communication Technology Law, Hungary. He also works as the internal data protection officer of the University of Pécs. His research areas are data protection and freedom of information, e-commerce law, online dispute resolution and copyright law. Besides his teaching activity he participates in national and European research projects (e.g. PAW project). From 2010 to 2011 he worked as a part-time data protection expert at the Office of Hungarian Parliamentary Commissioner for Data Protection and Freedom of Information. Email: szoke.gergely@ajk.pte.hu.

Doris Teutsch is a research assistant and Ph.D. candidate at the Department of Media Psychology at the University of Hohenheim, Germany. She completed her studies with a Master of Science degree in Empirical Communication Studies at the University of Hohenheim. Prior to this she graduated from Ludwig-Maximilians-University Munich, Germany, with a Bachelor degree in Communication Science, with Sociology as minor subject and a semester abroad at Babes-Bolyai-University Cluj-Napoca, Romania. Her research focus is on privacy and self-disclosure in the social web.

Sabine Trepte is Professor for Media Psychology at the University of Hohenheim in Stuttgart, Germany. Her research focuses on online disclosures and privacy from a psychological perspective. Sabine currently investigates online privacy with two large-scale projects funded by the German Federal Ministry of Education and Research. Specifically she conducts a longitudinal study investigating privacy, disclosures, social support, and authenticity over time. With a network of international scholars that she founded in 2008, Sabine conducted a trans-cultural study on privacy and online disclosures in Germany, the USA, the UK, the Netherlands and China. In 2011, Sabine Trepte published the book *Privacy Online*. She holds an M.A. in Psychology and received her Ph.D. in Media Psychology. Email: sabine.trepte@uni-hohenheim.de.

Govert Valkenburg is an academic researcher working in the interdisciplinary field of *Science and Technology Studies*. He has investigated and published on a variety of technological fields, including human genomics, sustainable energy and privacy and security technologies. Across these fields of interest, he has attended to issues in the democratization of technological society, the functioning of expertise in democracy, and the theme of citizenship in technological societies. He is currently based at Maastricht University. Alongside his academic career, he works as a professional singer in classical and early music. Email: g.valkenburg@maastrichtuniversity.nl.

Gabriela Zanfir obtained her Ph.D. title in 2014 from the Faculty of Law of the University of Craiova, with the thesis ‘The rights of the data subject regarding personal data protection’. Her research focuses on a legalistic approach to privacy and data protection, characterizing the two concepts in a functional way by applying to them classical legal concepts and by adapting the latter where necessary. She is also interested in the relationship between cloud computing and data protection. In 2012 she was a visiting researcher at the Tilburg Institute for Law and Technology, for 3 months. She won the Junior Scholar Award at CPDP 2014. Currently, she is working for the European Data Protection Supervisor. (She submitted the paper in this volume as an independent researcher, therefore she is the sole responsible for the content). Email: gabriela.zanfir@edps.europa.eu.

Part I
Profiling: A Persistent Core Issue of Data
Protection and Privacy

Chapter 1

Profiling Technologies and Fundamental Rights and Values: Regulatory Challenges and Perspectives from European Data Protection Authorities

Francesca Bosco, Niklas Creemers, Valeria Ferraris, Daniel Guagnin, and Bert-Jaap Koops

Abstract This paper aims to map the field of profiling, its implications for fundamental rights and values, and the measures which are or can be taken to address the challenges of profiling practices. It presents a working definition of profiling and elaborates a typology of its basic methods. In the second section the paper gives an overview of the technological background of profiling to display how fundamental rights and values of European societies are endangered by the use of profiling. Finally the paper presents the findings of a questionnaire addressed to European DPAs on the current and future legal framework, the domains of application, the complaints and remedies procedures regarding the use of profiling techniques, the main risks and benefits for the fundamental rights, and citizens' awareness on this topic. These findings contribute important insights for the ongoing discussion on the regulation of profiling in Europe.

F. Bosco (✉)
Emerging Crimes Unit, UNICRI, Turin, Italy
e-mail: bosco@unicri.it

N. Creemers • D. Guagnin
Centre for Technology and Society (CTS), Technische Universität Berlin (TUB), Berlin, Germany
e-mail: creemers@ztg.tu-berlin.de; guagnin@ztg.tu-berlin.de

V. Ferraris
Law Department, University of Turin, Torino, Italy

Amapola Progetti per la sicurezza delle persone e delle comunità, Torino, Italy
e-mail: valeria.ferraris@amapolaprogetti.org

B.-J. Koops
Tilburg Institute for Law, Technology, and Society (TILT), Tilburg, The Netherlands
e-mail: e.j.koops@uvt.nl

1.1 Introduction

The term “Big Data” is grounded in socio-technological developments, which began with the invention of the computer and has unfolded a rapidly growing dynamic over the past decades. Technological advancement has fueled the digitization of our societies by increasingly powerful infrastructures, basing on digital devices and software. Mediated communication today has mostly become digital communication, and information has consequently become easy to process and store as data, and is at the same time fluid and persistent. New potentials of gathering data raise hopes for developing more advanced ways to manage societies. The more we know the better we can control social processes and steer societal progress. At least that is what we are promised by “Big Data” proponents. “Big Data” appears to be a fetish, a crystal ball which allows those who use it to not just look into the future but to gain information which enables them to shape it at their needs.¹

However, big data itself is not information but still mere data.² The more data we gather the harder it is to extract usable information as the huge amounts of data exceed human capabilities of consideration. Consequently data needs powerful tools to be utilized as a marketable resource. These tools are considered to be found in technologies such as data mining. They are supposed to turn “Big Data” into the new oil.³

Profiling can be understood as a specific data mining method. In this perspective profiling is regarded as an (semi-)automated process to examine large data sets in order to build classes or categories of characteristics. These can be used to generate profiles of individuals, groups, places, events or whatever is of interest. Profiles structure data to find patterns and probabilities. Using actuarial methods in this context is supposed to generate prognostic information to anticipate future trends and to forecast behavior, processes or developments. The aim is to develop strategies in order to manage uncertainties of the future in the present. In this regard, the

¹See Fraunhofer. IAIS, *Big Data – Vorsprung durch Wissen. Innovations potenzial analyse*, http://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/FraunhoferIAIS_Big-Data-Analyse_Doku.pdf, last accessed 01 April 2014. The programs of the world’s largest ICT fair CeBIT 2014, the Big Data Days 2013, and the European Data Forum and the presentations given there, draw an interesting picture of the potentials the ICT industry attributes to “Big Data” and big data analytics: <http://www.cebit.de/home>, last accessed 03 April 2014, <http://www.big-data-days.de>, last accessed 03 April 2014, and <http://2014.data-forum.eu/>, last accessed 03 April 2014.

²Sasa Baskarada and Andy Koronios, “*Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension*,” in *Australasian Journal of Information systems*, Vol 18, No 1 (2013): 5–24.

³Karl-Heinz Streibich, “*Big Smart Data. Mehrwert für Unternehmen*” (paper presented at the Big Data Days, Berlin, Germany, November 11–12, 2013).

ideology of “Big Data” and analytical tools such as profiling can be understood as an important facilitator and part of a preventive paradigm which can be found in diverse societal contexts.⁴

Even though the reality of profiling might not live up to the expectations of its prophets,⁵ the assumed potentials of gathering and processing data spawn the dream of overcoming human deficiencies with technology, these new technologies also draw fears and skepticism as they impose threats on some of the core values and principles of European societies. Key challenges which have been identified by scholars include infringements of democratic principles and the rule of law: Data gathering, exchange, and processing potentially harm central values like individual autonomy and informational self-determination as well as the fundamental rights of privacy, data protection, and non-discrimination.

This paper aims to map the field of profiling. It focuses on its implications for fundamental rights and values in different fields of application and on the assessment of the existing countermeasures to address the challenges of profiling practices. In the following section this paper proposes a working definition of profiling. The third section gives an overview of the technological evolution building the ground for the emergence of profiling, afterwards it is demonstrated how fundamental rights and values of European societies are endangered by the application of profiling in various contexts (Sect. 1.4). In Sect. 1.5 the legal regulation of profiling is sketched. Finally the paper presents the first findings of a questionnaire carried out by the project PROFILING,⁶ in order to gain knowledge about European Data Protection Authorities’ awareness, attitudes, and activities regarding profiling and its societal impacts.

1.2 Profiling: Towards a Definition

Profiling is a highly evocative term with multiple meanings, used in both specialist and non-specialist contexts. Whereas the literature on statistics does not pay specific attention to definitions and tends to focus on technical aspects (e.g. data mining

⁴See Susanne Krasmann, “*Der Präventionsstaat im Einvernehmen. Wie Sichtbarkeitsregime stillschweigend Akzeptanz produzieren,*” in *Sichtbarkeitsregime: Überwachung, Sicherheit und Privatheit im 21. Jahrhundert*, ed. Leon Hempel, Susanne Krasmann and Ulrich Bröckling (Wiesbaden: VS Verlag, 2010), 53–70 and Pat O’Malley, “*Risk, power and crime prevention,*” *Economy and Society* 21/3 (1992): 252–275.

⁵For some of the technical problems which harm the reliability of profiling results, see Daniel Guagnin, Leon Hempel and Justin Jung, “Evolution of Technologies in Profiling”, Working Paper, http://profiling-project.eu/wp-content/uploads/2013/08/Evolution-of-Technologies-in-Profiling_0208.pdf, last accessed 02 April 2014.

⁶The PROFILING project is funded from the European Union’s Fundamental Rights and Citizenship programme. The 2 year project started in November 2012. More information on the project can be found on the website <http://profiling-project.eu>.

techniques and predictive models), providing a definition appears an issue among socio-legal scholars and policy makers. However a widely shared definition has not yet emerged.

Gary T. Marx gave one of the oldest definitions of profiling in a paper that analyses systems of data searching. Profiling (defined by the author in contrast to “matching”) is defined by stressing the logic behind it: “the logic of profiling is more indirect than that of matching. It follows an inductive logic in seeking clues that will increase the probability of discovering infractions relative to random searches. Profiling permits investigators to correlate a number of distinct data items in order to assess how close a person or event comes to a predetermined characterization or model of infraction”.⁷ According to the author’s background, this definition is strictly related to the law enforcement domain.

Almost 10 years later, Roger Clarke defined profiling as a “dataveillance technique (. . .) whereby a set of characteristics of a particular class of person is inferred from past experience, and data-holdings are then searched for individuals with a close fit to that set of characteristics”.⁸

A legal scholar, Bygrave again stressed: “profiling is the inference of a set of characteristics (profile) about an individual person or collective entity and the subsequent treatment of that person/entity or other persons/entities in the light of these characteristics”.⁹

Later on, Mireille Hildebrandt was the one who put the best effort to precisely define profiling and its distinctive features and the working definition proposed here has built on her work. She defines profiling as “the process of ‘discovering’ patterns in data in databases that can be used to identify or represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent an individual subject or to identify a subject as a member of a group (which can be an existing community or a discovered category).”¹⁰

Profiling creates a new form of knowledge that makes visible patterns that are otherwise “invisible to the naked human eye”.¹¹ They are based on correlations found in data sets, and cannot be “equated with causes or reasons without further

⁷Marx, Gary and Reichman Nancy. “*Routinizing the Discovery of Secrets: Computers as Informants*,” in *American Behavioral Scientist*, 27, 4 (1984): 429.

⁸Clarke, Roger, “*Profiling: A Hidden Challenge to the Regulation of Data Surveillance*,” in *Journal of Law and Information Science* 4, 2 (1993): p. 403.

⁹Bygrave, Lee A., *Data protection law: Approaching its rationale, logic and limits* (The Hague: Kluwer Law International, 2002), 301.

¹⁰Mireille Hildebrandt, “*Profiling and AML*,” in *The Future of Identity in the Information Society. Challenges and Opportunities*, ed. Kai Rannenberg, Denis Royer and Andre Deuker (Heidelberg: Springer, 2009a), 275.

¹¹Mireille Hildebrandt, “*Who is Profiling Who? Invisible Visibility*” in *Reinventing Data Protection?*, ed. Serge Gutwirth et al. (Dordrecht: Springer, 2009c), 241.

inquiry; they are probabilistic knowledge.”¹² Profiling represents a shift from the idea that knowledge is the result of tested hypothesis. It generates hypotheses: “the correlations as such become the ‘pertinent’ information, triggering questions and suppositions”.¹³ Consequently profiling fosters new forms of generating and applying knowledge. Due to the growing capacities of databases, and capabilities of advanced analysis profiling procedures become increasingly complex. In this context the human role in interpreting data changes significantly.

As pointed out by Hildebrandt, profiling can be categorized into non-automated, automated and autonomic profiling. Non-automated profiling is a form of reasoning that does not rely on any process of automation. Automated profiling is based on “automated functions that collect and aggregate data” and develop into “automation technologies that can move beyond advice on decision-making, taking a load of low-level and even high-level decisions out of human hands.”¹⁴ Differently, autonomic profiling describes the process whereby the human role is minimized and the decision making process is entirely driven by the machine.¹⁵ Autonomic profiling “goes one step further than automated profiling.”¹⁶ The machines drive the decision making process, providing for a readjusted environment based on their profiling and without calling for human intervention. Besides their degree of automation profiling methods can be distinguished by their object and application. Profiling can be applied as group profiling or individual profiling: the techniques that identify and represent groups can also focus on individuals.¹⁷ Moreover profiling relies on data collected from one single person or group to apply the information derived from data processing to the same person or group – direct profiling – or it relies on categorization and generalisation from data collected among a large population to apply it to certain persons or groups – indirect profiling. Group profiling can also

¹²Gloria González Fuster, Serge Gutwirth and Ellyne Erika, “Profiling in the European Union: A high-risk practice,” in *INEX Policy Brief* 10 (2010): 2.

¹³Gloria González Fuster, Serge Gutwirth and Ellyne Erika, “Profiling in the European Union: A high-risk practice,” in *INEX Policy Brief* 10 (2010): 2.

¹⁴Mireille Hildebrandt, “Defining profiling: a new type of knowledge?,” in *Profiling the European Citizens. Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer, 2008), 28.

¹⁵See Mireille Hildebrandt, “Profiling: from Data to Knowledge. The challenges of a crucial technology,” in *DuD Datenschutz und Datensicherheit* 30(9) (2006): 548–552 and Mireille Hildebrandt, “Defining profiling: a new type of knowledge?,” in *Profiling the European Citizens. Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer, 2008), 17–47.

¹⁶Mireille Hildebrandt, “Profiling: from Data to Knowledge. The challenges of a crucial technology,” in *DuD Datenschutz und Datensicherheit* 30(9) (2006): 550.

¹⁷See Anton, Vedder, “KDD: The challenge to individualism,” in *Ethics and Information Technology* (1999): 275–281 and Arnold Roosendaal, *Digital Personae and Profiles in Law. Protecting Individuals’ Rights in Online Contexts*, Oisterwijk: Wolf Legal Publishers.

be classified as distributive group profiling or non-distributive group profiling.¹⁸ A distributive group profile identifies a certain number of people having the same attributes. All the members of the group share the same characteristics. In contrast, a non-distributive group profile identifies a certain number of people who do not share all the attributes of the group's profile.

These distinctions give an idea of the different types of profiling and their application. The forms of profiling, which are subject of this article are automated and autonomic profiling and their various forms and fields of application.

The following proposed definition takes into account the preceding evolution of technologies in which profiling is embedded and focuses on the purpose profiling is being used for. It will be the basis for this paper:

Profiling is a technique of (partly) automated processing of personal and/or non-personal data, aimed at producing knowledge by inferring correlations from data in the form of profiles that can subsequently be applied as a basis for decision-making.

A profile is a set of correlated data that represents a (individual or collective) subject.

Constructing profiles is the process of discovering unknown patterns between data in large data sets that can be used to create profiles.

Applying profiles is the process of identifying and representing a specific individual or group as fitting a profile and of taking some form of decision based on this identification and representation.

1.3 Societal Consequences of Digitization

Advanced data analysis tools have established new social practices of knowledge production and have created new types of knowledge. We argue that the practices of profiling have facilitated and are part of a broader societal paradigm of prevention. We will elaborate on the societal implications of changing social practices through emerging profiling technologies as a ground for the examination of threats for fundamental rights and values of European societies in Sect. 1.4.

Observations made by human beings need to be written down to be made explicit. The written documentation of observations can be regarded as a first step to enable a generalized and objectified way of keeping information and exchanging it between individuals and institutions.¹⁹ Digitized information, however, can be processed and analysed automatically so that information is easier and cheaper to store, process and analyse. An illustrative example of how exhaustive and expansive

¹⁸See Anton, Vedder, "KDD: The challenge to individualism," in *Ethics and Information Technology* (1999): 275–281.

¹⁹The important role the implementation of written files played as storage and medium for information but also as a symbol of power for the Inquisition trials in Italy is displayed by Thomas Scharff, "Erfassen und Erschrecken. Funktionen des Prozeßschriftguts der kirchlichen Inquisition in Italien im 13. und frühen 14. Jahrhundert." in *Als die Welt in die Akten kam. Prozeßschriftgut im europäischen Mittelalter*, ed. Susanne Lepsius and Thomas Wetzstein (Frankfurt a.M.: Vittorio Klostermann, 2008), 255–274.

the detailed documentation of people's activities and behaviour has been, is the comparison between digital data the NSA stores with the amounts of files the Stasi – German Democratic Republic's domestic secret service – produced. All the information captured throughout the Stasi history would fill about 48.000 cabinets covering approximately 0,019 km². The NSA's planned data centre in Utah will host about 5 zettabytes of data which could roughly be converted in about 42 quadrillion file cabinets covering 17 million km² – bigger than the European continent.²⁰ The example also shows the differing efforts needed to collect and archive data depending on whether using analog or digital data processing. While the Stasi needed to install microphones, hire staff to monitor and document people's behaviour to gain information about their habits, attitudes and social networks, in a digitized world a lot of that information can be monitored and stored on the fly through sensors, log data or user generated content. This shows that the digitization of communication and transactions does not only produce more data but also provides new kinds of information²¹ which can be used to extract knowledge about individuals: their social relations, interests and activities. Once stored and made accessible via computer networks, data becomes easily exchangeable worldwide. At the same time it becomes hard to grasp how data is exchanged, which information is gained and by whom. Furthermore the specific mediums can store specific data. Certain elements which can be archived on paper cannot be archived digitally and vice versa. Moreover certain information can hardly be digitized respectively digitally analyzed, e.g. hand-written information, and smells. By that, archives have a filtering function which shapes the accessibility of information as data. But simplified storage and exchange of data are only one aspect of the ongoing process of digitization of everyday life. Beyond that advanced methods of data analysis have fundamentally changed the procedures of knowledge production through automation.

Another effect of the digitization of data becomes evident when we think of the different haptic and cognitive perceptions of digital versus analog files and folders. Different items and elements can be put in an analog or digital file, and at the same time, the availability of and the access to certain kinds of information fundamentally changes. In other words: accessing information at a (real) desktop is very different from accessing information when sitting in front of a computer screen. Paper folders can be touched and felt, digital files are browsed on a screen and can be searched by keywords. Consequently, the way of reasoning changes, as first findings of one of the case studies conducted in PROFILING show.²² More interaction of the analyst is

²⁰Open Data City, *Stasi versus NSA*, accessed February 27, 2014, <http://apps.opendatacity.de/stasi-vs-nsa>.

²¹Bert-JaapKoops, "Technology and the Crime Society: Rethinking Legal Protection," in *Law, Innovation & Technology*, 1, 1 (2009): 93–124.

²²Technische Universität Berlin conducted a case study about the transformation of policing practices due to the application of data processing technologies. Expert interviews were conducted with scholars, civil rights activists, directors of security technology companies, a police representative, and a lawyer. Police as well as technology providers mentioned changes in the workflow and the

oriented towards computer interfaces and thus influenced by the way user interfaces are designed, information is presented, and how searches can be conducted.²³ The transformation of the human role in knowledge production processes is even more significant when it comes to examining large-scale databases. Learning algorithms are trained on specific data sets to build categories or to find patterns in the data. Assumptions or hypotheses made by the analyst play a minor role during data processing, they are to a certain degree hidden in the process of writing algorithms and training the algorithms. Finally, hypotheses are derived “from the material”.²⁴ As a consequence implicit assumptions driving the actors during the selection of training data, preprocessing target data and suitable algorithms become invisible and the outcomes produced by “the data” seem objectified. Subjective assumptions and social norms are hidden in the technology during the process of automatization, while outcomes based on computed models and databases are often perceived as solid statistics and thus more objective than human interpretation.²⁵ This perception as objectified knowledge of computer-generated models supports the thesis of a general tendency of technology to make social norms more durable²⁶ and more specifically the thesis that social sorting becomes strengthened if mediated through technology.²⁷ Profiles, as mentioned above, can be seen as hypotheses. These hypotheses are inductive as they are not necessarily developed on the basis of a theory or a common sense expectation, but often emerge in the process of data mining. This can be regarded as a shift from a more traditional, rather assumption-

construction of these from digitally stored information. The report of the case study’s final results will be available at <http://profiling-project.eu/>.

²³See Nina Degele, *Einführung in die Techniksoziologie* (Stuttgart, UTB, 2002), p. 167–168.

²⁴The results software can draw from data material are dependent on the quality of the data sets, which are examined, including the selection and pre-processing of data. Major problems, especially regarding large-scale data sets which combine data from various sources, are poor data quality, data incompatibility, and biased data sets which corrupt data mining outcomes. Furthermore operators might not be familiar with such reliability problems. Consequently operators might not act properly upon these problems. See Ana Canhoto and James Blackhouse, “General Description of Behavioural Profiling,” in *Profiling the European Citizens. Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer, 2008), 47–63 and Bernhard Anrig, Will Brown, and Mark Gasson, “The Role of Algorithms in Profiling,” in *Profiling the European Citizens, Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer, 2008), 65–87.

²⁵See Toon Calders and Indrė Žilobaitė, “Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures,” in *Discrimination and Privacy in the Information Society*, ed. Bart Custers et al. (Berlin: Springer, 2013), 43–57.

²⁶See Bruno Latour, “Technology is Society Made Durable,” in *A Sociology of Monsters: Essays on Power, Technology and Domination*, ed. John Law (London: Routledge, 1991), 103–131.

²⁷See Michaelis Lianos and Douglas Mary, “Dangerization and the End of Deviance: The Institutional Environment,” in *British Journal of Criminology* 40, 2 (2000): 261–278 and Rosamunde van Brakel and Paul De Hert, “Policing, surveillance and law in a pre-crime society: Understanding the consequences of technology based strategies,” in *Cahier Politistudies* 2011–3 no. 20 (2011): 163–192.