DIGITAL ENLIGHTENMENT YEARBOOK 2013

Digital Enlightenment Yearbook 2013

The Value of Personal Data

Edited by

Mireille Hildebrandt

Radboud University Nijmegen, Vrije Universiteit Brussel, Erasmus University Rotterdam

Kieron O'Hara

University of Southampton

and

Michael Waidner

Technische Universität Darmstadt

IOS Press

Amsterdam • Berlin • Tokyo • Washington, DC

ISBN 978-1-61499-294-3 (print)
ISBN 978-1-61499-295-0 (online)
Library of Congress Control Number: 2013946636

Publisher
IOS Press BV
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

Distributor in the USA and Canada
IOS Press, Inc.
4502 Rachael Manor Drive
Fairfax, VA 22032
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

PRINTED IN THE NETHERLANDS

Digital Enlightenment Yearbook 2013
M. Hildebrandt et al. (Eds.)
IOS Press, 2013

Foreword

Robert MADELIN
Director General of DG Connect of the Euorpean Commission

I welcome the effort of the Digital Enlightenment Forum to harness the humanism, rationality and optimism of the Age of Enlightenment to better shape and to enhance the benefits to society of our evolving Digital Age. I write this foreword in the midst of a warm global debate on the value of Data – a theme for the DEF and for its Davos cousin the WEF – as well as on the balance between Westphalian security needs and personal privacy aspirations.

This yearbook's focus on the 'Value of Personal Data' is apt, because the explosion of data and of computing power over the last few years represents a further step-change in the development of the Internet. The statistics of the data explosion bears repetition: by some estimates, 90% of the world's current data has been produced in the last two years; in turn, this is more than was generated in the previous 2,000 years. Meanwhile, the world generates 1.7 billion bytes every minute. Big data (high volume, high velocity, high variability) is here to stay.

Beyond its purely commercial uses, the societal benefits of Big Data will be more slowly understood and developed. For the public sector, better data allows services that are more efficient, transparent and personalised. In the health sector, open results and open data permit whole new fields of research. For example, scientists at Columbia and Stanford universities analysed millions of online searches to learn about the symptoms and conditions of certain drugs. This led to the unexpected medical discovery that the combination of two drugs – paroxetine, an antidepressant, and pravastatin, a cholesterol-lowering drug – cause high blood sugar. Meanwhile, the Global Viral Forecasting Initiative (GVFI) uses advanced data analysis on information mined from the Internet to identify comprehensively the locations, sources and drivers of local outbreaks before they become global epidemics. Such techniques offer guidance up to a week ahead of previous indicators.

Controversy can arise. TomTom, a Dutch manufacturer of satellite navigation devices, ran into problems with the anonymised data that it collected from its users about individual driving behaviour, which it provided to the Dutch government to help improve the national road system. However, it transpired that data was being used in part to identify the most appropriate sites for speed cameras. Users complained that they were unaware of this application of their data, and were concerned that the police would be able to identify individual speeding violations from the data. TomTom assured consumers that the data had indeed been fully anonymised and that the company would prevent such use in the future. The lesson for all Big Data companies is to focus on *perception* of data use as much as on actual use.

As that example goes to show, user trust is key to Big Data success. Neelie Kroes, Commission Vice-President for the Digital Agenda, has put this succinctly in many public statements on ensuring such confidence: "Privacy is not just about technical features. Without privacy, consumers will not trust the online world. And without trust,

the digital economy cannot reach its full potential". She goes on to identify her three requirements for privacy in the digital age: *transparency* so that citizens know exactly what the deal is; *fairness* so that citizens are not forced into sharing their data; and *user control* so that citizens can decide – in a simple and effective manner – what they allow others to know. These concepts underpin much of the material presented in this Yearbook.

In an emergent field such as Big Data, the Forum's work can inform EU-wide innovation, the EU research agenda and our vision of our common future.

On the innovation front, I believe that the right standards can accelerate demand growth. Without interoperability and harmonised formats, large datasets can be too difficult to fit together and use in practice. In this respect the Commission has engaged with stakeholders in the European public sector information ecosystem to forge lightweight agreements and standards that are needed to enable interoperability and integration of Public Sector Information (PSI). It is also promoting standardisation of data formats on our EC Open Data portal and one of the goals of its Pan-European Open Data Portal is to drive the harmonisation of data-formats and licensing conditions in Europe.

On *research and innovation,* the Commission has provided on average €76 million p.a. for data and language technologies. In Horizon 2020, it intends to continue to fund innovation in the area of data products and services, and has also set up the Research Data Alliance, to help scientific data infrastructure become fully interoperable. Open Data standards are also considered set to continue as part of the Horizon 2020 activities from 2014.

As to the longer-term vision, this is where "Digital Futures" comes in, a foresight project launched by DG Connect to prepare for the world beyond 2030. The project taps into the collective wisdom and aspirations of stakeholders to co-create long term visions (on a time horizon 2040–50) and fresh ideas for policies that can inspire the future strategic choices of DG Connect and the Commission. It draws inspiration from the long term advances at the intersection between ICT and society, economy, environment, politics, humanities and other key enabling technologies and sciences. This is why we are co-hosting with the Forum a Digital Futures Workshop on the "Future of Personal Data and Citizenship".

In terms of the privacy implications of Big Data, these are just some of the issues that could be addressed by Digital Futures: Can we achieve better metadata management so that potential re-users understand better what uses of the data are covered by consent and/or the statutory law grounds on the basis of which data were collected? Can we build in features in data-management systems that allow a level of anonymisation of personal data compatible with legal requirements? How can we develop privacy-enhancing technologies facilitating the process of giving consent to new uses of personal data? Can we establish "data banks" – dedicated digital spaces for the management of the personal information for each data subject? What kind of role should self-regulation and co-regulation have in ensuring compliance with privacy rules? Regulators do not have all the answers, but we can at least ask the right questions.

Contents

Part IV. Other Sources of Data

Part V. Personal Data Management: Examples and Overview