

Information Technology and Law Series

Volume 36

Editor-in-Chief

Simone van der Hof, eLaw (Center for Law and Digital Technologies),
Leiden University, Leiden, The Netherlands

Series Editors

Bibi van den Berg, Institute for Security and Global Affairs (ISGA),
Leiden University, The Hague, The Netherlands

Gloria González Fuster, Law, Science, Technology & Society Studies (LSTS),
Vrije Universiteit Brussel (VUB), Brussels, Belgium

Eva Lievens, Faculty of Law, Law & Technology, Ghent University,
Ghent, Belgium

Bendert Zevenbergen, Center for Information Technology Policy,
Princeton University, Princeton, USA

More information about this series at <https://link.springer.com/bookseries/8857>

Alessandro Mantelero

Beyond Data

Human Rights, Ethical and Social Impact
Assessment in AI



ASSER PRESS



Springer

Alessandro Mantelero
DIGEP
Politecnico di Torino
Torino, Italy



ISSN 1570-2782 ISSN 2215-1966 (electronic)
Information Technology and Law Series
ISBN 978-94-6265-530-0 ISBN 978-94-6265-531-7 (eBook)
<https://doi.org/10.1007/978-94-6265-531-7>

Published by T.M.C. ASSER PRESS, The Hague, The Netherlands www.asserpress.nl
Produced and distributed for T.M.C. ASSER PRESS by Springer-Verlag Berlin Heidelberg

© The Editor(s) (if applicable) and The Author(s) 2022. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

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.

This T.M.C. ASSER PRESS imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature.

The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

To Bruna and Giuseppe

Foreword

It is probably safe to say that at the time of writing¹ more than 99% of the world's population do not yet understand what a game-changer AI can be...or is already proving to be. Much news coverage, for example, is still given to efforts which aim to prevent states like Iran or North Korea from developing nuclear weapons and increasingly sophisticated means of delivering them. Yet relatively little news coverage is given to the fact that, in reality, AI has made nuclear weapons obsolete. Why would a state—or indeed a terrorist—wish to deploy or acquire a very expensive and relatively unstable nuclear weapon when it can instead deploy much cheaper AI-controlled devices which do not create a radioactive crater or destroy so many valuable assets in a target zone?

In one of the saddest unintentional puns to emerge about the endemic inability of the world's nations to agree and deploy sufficient safeguards and remedies in international law, AI powers LAWs—Lethal Autonomous Weapons. These can take many shapes and sizes but perhaps none more sinister than “killer drones” capable of facial recognition thus being able to single out human targets to which they can deliver an explosive device. These drones can not only be easily and cheaply mass produced to the extent that a million of them can be transported in a standard shipping container but they can be released in swarms so numerous which make it well nigh impossible for air defense systems to shoot down enough of them to make adequate defence a plausible option. In this way these cheap² devices, all capable of individually or collectively using AI to select and identify individual human beings as their targets, are well on the way to becoming weapons of mass destruction.

Killer drones and drone swarms do not only exist in the fertile imagination of some or in science fiction. They have been deployed in combat for at least the best part of two years. A panel of UN experts in March 2020 reporting about the conflict

¹November 2021-January 2022.

²Current best estimates for costs of killer-drone LAWs range from between 10–30 dollars each if produced in sufficient quantities though those already available such as the Turkish-made Karga 2 understandably command a higher premium.

in Libya stated that “Logistics convoys and retreating [Haftar-affiliated forces] were subsequently hunted down and remotely engaged by the unmanned combat aerial vehicles or the lethal autonomous weapons systems such as the STM Kargu-2 ... and other loitering munitions.”³ The U.N. report goes on: “The lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true ‘fire, forget and find’ capability.”⁴ This was not an isolated incident. More recently, during operations in Gaza in mid-May 2021, “the Israel Defense Forces (IDF) used a swarm of small drones to locate, identify and attack Hamas militants. This is thought to be the first time a drone swarm has been used in combat.”⁵

The potential for harm in a device which can take actions which can infringe human rights by, e.g. discriminating on grounds of gender, age, ethnicity or political opinion should be immediately apparent. The fact that we already have devices such as AI-driven drones that could be programmed to identify a given individual off a list of politically inconvenient people and seek out and destroy such a person or be instructed to seek out and kill all people who look like Jews or dark-skinned people or all males in a city who are between the ages of 12 and 65 should have alarm bells ringing across all sectors of society. That they are not is a serious cause for concern in itself.

One of the many problems with LAWs is that the world currently does not have the right type of international law to cover this type of AI-driven technology. While, over the past 50 years, progress had been made on arms control in the form of the Treaty on the Non-Proliferation of Nuclear Weapons (1968–1970), the Chemical Weapons Convention (1997) and, most recently, the Biological Weapons Convention, the development and deployment of LAWs is characterised by lawlessness. Governments such as that of New Zealand have, in November 2021, taken a clear policy stance moving for a new international treaty to be made on the issue but these latest efforts were stunted during the 6th Review Conference of the Conventional on Conventional Weapons (CCW). Although, States agreed to continue the work of the Group of Governmental Experts related to emerging technologies in the area of lethal autonomous weapon systems for another year, with a renewed mandate for the group agreed to hold ten days of meetings in 2022, there is no guarantee that this will produce results better than those of 2021.

LAWs is just one example of why Alessandro Mantelero’s study is an important book about an important subject. Although formally titled *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI*, it could equally have been titled *Beyond Law: Human Rights, Ethical and Social Impact Assessment in AI*. For although Mantelero is a legal scholar with a growing pedigree in Technology Law, his book is an explicit plea to go beyond law and instead embrace a more holistic approach to Artificial Intelligence. There can be no doubting Mantelero’s

³<https://undocs.org/S/2021/229>.

⁴Ibid.

⁵<https://www.newscientist.com/article/2282656-israel-used-worlds-first-ai-guided-combat-drone-swarm-in-gaza-attacks/>.

commitment to human rights law, but he is fundamentally right in his position that a legal approach alone is not enough. Instead he advocates adoption of the HRESIA model. Today, especially with the advent of the GDPR, a growing number of people are familiar with the need to carry out an impact assessment in many of those cases where one intends to introduce a technology which deals with personal data. But, as Mantelero points out, HRESIA—Human Rights Ethical Social Impact Assessment—is a hybrid model taking into account the ethical as well as the social impact of a technology together with the legal dimensions such as those of human rights.

Mantelero is, through HRESIA, offering us a conceptual framework within which we can think about AI and also decide what to do about it from a policy point of view. The main components of HRESIA are the analysis of relevant human rights, the definition of relevant ethical and social values and the targeted application to given AI cases, thus combining the universality of human rights with the local dimension of societal values. In doing so Mantelero advocates a multi-stakeholder and human-centred approach to AI design. Participation and transparency form part of the mix promoted by HRESIA while retaining elements of more traditional risk management models such as the circular product development models.

Building on his knowledge of the most recent developments in data protection law, Mantelero walks the reader through the advantages and disadvantages of impact-assessment solutions in the field of data-centred systems such as PIA/DPIA, SIA and EtIA. He is at pains to point out that “the recent requirements of the GDPR—according to the models offered by the DPAs fail to offer a more satisfactory answer—by explaining that “Despite specific references in the GDPR to the safeguarding of rights and freedoms in general as well as to societal issues, the new assessment models do nothing to pay greater attention to the societal consequences than the existing PIAs.” Mantelero makes the point that “HRESIA fills this gap, providing an assessment model focused on the rights and freedoms that may be assessed by data use offering a more appropriate contextualisation of the various rights and freedoms that are relevant to data-intensive systems. The latter are no longer limited to data protection and should therefore be considered separately rather than absorbed in a broad notion of data protection”. Mantelero’s advocacy of HRESIA is part of his apparent agreement with the mood of those legal scholars who have highlighted “how the application of human rights is necessarily affected by social and political influences that are not explicitly formalised in court decisions” in a perspective wherein “HRESIA may be used to unveil the existing interplay between the legal and societal dimensions”.

Much as I deem privacy to be important, I am delighted that the HRESIA methodology extends to all human rights and not just privacy. This is very much in line with the approach I explicitly advocated as UN Special Rapporteur on Privacy in my report to the UN’ Human Rights Council in March 2016 as reflected in the HRC’s resolution of March 2017 *Recognizing the right to privacy also as an enabling right to the free development of personality and, in this regard, noting with concern that any violation to the right to privacy might affect other human*

rights, including the right to freedom of expression and to hold opinions without interference, the right to freedom of peaceful assembly and association. While also holding out the promise of significant benefits, AI has the potential to infringe or otherwise interfere with many or all of these human rights, hence the need for in-depth and constant detailed evaluation such as that inherent to a proper implementation of HRESIA.

Now, it is impossible in a work of relatively modest length to go in-depth through a comprehensive list of examples which would demonstrate beyond reasonable doubt that HRESIA is useful in all cases related to AI technology but it certainly promises to be a better start than most. Indeed, this is why I opened this preface with reference to just one example of AI-driven technology, i.e. LAWs. For the latter is clearly yet another instance where looking to existing rules or legal precedent may be helpful but certainly not enough. The societal impact of LAWs—including the potential use of such technologies against one’s own civilian population and not exclusively against a foreign enemy—as well as the multifarious ethical dimensions should provide a perfect case-study for the advantages—and practical difficulties—involved in applying HRESIA.

Indeed I look forward to other scholars—and possibly even Mantelero himself—rising to the challenge and methodically applying the HRESIA approach to the catalogue of problems that AI brings with it. For the use of AI in weaponry such as LAWs is just one of many issues we should be paying attention to. The misuse of AI, including racial and gender bias, disinformation, deepfakes and cybercrime is as much a part of a long TO DO LIST as the very standard programming that goes into AI itself. Given that AI involves specifying a fixed objective’ and since the programmer cannot always specify objectives completely and correctly, this results in a situation where having fixed but imperfect objectives could lead to an uncontrollable AI that stops at nothing to achieve its aim. What novel or useful solutions would HRESIA produce in Stuart Russell’s oft repeated and now classic “children and the cat”⁶ example? Likewise, what can HRESIA offer to an analysis of the impact that AI will have on jobs, making many obsolete and many workers redundant? What real benefits would the policy maker obtain from using HRESIA when faced with the decision of supporting, regulating or banning AI-powered robots designed to provide care to the elderly? How would HRESIA help resolve “privacy by design, privacy by default” issues in such cases not to mention the ethical and legally correct approaches to euthanasia, dementia, terminal illness, etc.?

Some analysts will no doubt spend much time over the coming years trying to pick holes in HRESIA. Eventually somebody may possibly also come up with an even better way of solving problems related to AI but, until that happens, Mantelero’s work offers some of the insights into the theoretical underpinnings of why it could be a useful approach when doing so. It is also a sign of the times. For

⁶Wherein a domestic robot programmed to look after children, tries to feed the children but sees nothing in the fridge. “And then... the robot sees the cat... Unfortunately, the robot lacks the understanding that the cat’s sentimental value is far more important than its nutritional value. So, you can imagine what happens next!”.

the best part of forty years, we have been gradually moving away from a mono-disciplinary approach in problem-solving to a multi-disciplinary approach, often coupled with an inter-disciplinary approach. The perspective obtained at the intersection of several disciplines can also be one which is profoundly more accurate and more practical/pragmatic than one which is constrained by the knowledge and practices of any single discipline. Indeed, the very notion of HRESIA implies taking into account the perspective of other disciplines outside Human Rights Law, ethics and social impact. Computer science, applied technologies, economics and social psychology are only a few of the other disciplines that immediately come to mind which need to be deeply and constantly involved in the way that society needs to think about AI. Speaking of “a holistic approach” has become something of a cliché yet it is difficult to think of a context which requires it more than AI...and that basically is the nub of the message in Mantelero’s current work. It is also an encouraging start on the fiendishly difficult task of regulating AI and producing sensible policy decisions outside the field of law which are however required to ensure that mankind reaps more benefits from AI and avoids the serious dangers inherent in the uncontrolled development and deployment of such technologies.

Tal-Qroqq, Malta
January 2022

Joe Cannataci

Joe Cannataci was appointed as the first ever [UN Special Rapporteur on Privacy](#) in 2015, following the Snowden revelations about mass surveillance. His UN mandate was renewed in 2018 (until August 2021). He is head of the [Department of Information Policy & Governance](#) at the Faculty of Media & Knowledge Sciences of the University of Malta. He also co-founded and continues as Co-director (on a part-time basis) of [STeP, the Security, Technology & e-Privacy Research Group](#) at the University of Groningen in the Netherlands, where he is Full Professor, holding the [Chair of European Information Policy & Technology Law](#). A Fellow of the British Computer Society (FBCS) and UK Chartered Information Technology Professional (CITP), his law background meets his techie side as a [Senior Fellow and Associate Researcher](#) at the CNAM Security-Defense-Intelligence Department in Paris, France and the [Centre for Health, Law and Emerging Technologies at the University of Oxford](#). His past roles include Vice-Chairman/Chairman of Council of Europe’s (CoE) Committee of Experts on Data Protection 1992–1998, Working Parties on: Data Protection and New Technologies (1995–2000); Data Protection & Insurance (1994–1998); CoE Rapporteur on Data Protection and Police (1993; 2010; 2012).

Preface

*As you set out for Ithaka
Hope the journey may be long,
Full of adventures, full of discovery*
Κωνσταντίνος Π. Καβάφης
Constantine Cavafy, Edmund Keeley, and Philip Sherrard,
Voices of Modern Greece: Selected Poems (Princeton
University Press 1981).

As in Cavafy's poem, this is the story of a journey lasting several years. It began in 2012, when, after several studies on data protection, my first investigation of the impact of large-scale data-intensive systems appeared in an article on Big Data and the risks of digital information power concentration published in an Italian law review.⁷

A few years after the Aspen Institute's report on The Promise and Peril of Big Data⁸ and several months after the provocative paper presented by danah boyd and Kate Crawford at the Oxford Internet Institute,⁹ Big Data became my new field of enquiry for two spring terms as a visiting fellow there in 2013 and 2014.

As a privacy scholar, I was concerned about the imbalance of power created by large-scale concentration of data and predictive power in the hands of a limited number of big players. Recognising the limits of the traditional individual

⁷Mantelero A (2012) Big Data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo [Big Data: the risks of digital information power concentration and oversight tools]. *Il diritto dell'informazione e dell'informatica* 2012 (1), 135–144.

⁸Bollier D (2010) *The Promise and Peril of Big Data*. Aspen Institute, Washington, DC http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf. Accessed 27 February 2014.

⁹boyd d and Crawford K (2011) Six Provocations for Big Data. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, Oxford Internet Institute, 21 September 2011 <https://papers.ssrn.com/abstract=1926431>. Accessed 3 August 2021; boyd d and Crawford K (2012) Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *15 Information, Communication & Society* 662.

consent-based model,¹⁰ I began to explore the collective dimension of data protection.¹¹

Antoinette Rouvroy was working at the time on her report on Big Data for the Council of Europe¹² and the peculiar circumstances of new scientific research practices in the digital era brought an unexpected consequence. After reading the draft of her report online, I posted several comments that led to my involvement as an adviser to the Council of Europe's Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, a collaboration that remains ongoing, first on Big Data,¹³ and later on AI regulation.¹⁴

These brief autobiographical notes, at a time when a concurrence of social and technological factors gave rise to a wave of AI development, explain the genesis of this book.

An interest in the theoretical limits of the existing legal framework—centred on data protection law and models established in the 1970s and early 1990s—plus my direct experience of the international regulatory demands and dynamics were the two driving forces behind extending the initial scope of my research to cover the new algorithmic society.

¹⁰Mantelero A (2014) The future of consumer data protection in the EU Re-thinking the “notice and consent” paradigm in the new era of predictive analytics. 30(6) *Computer Law & Security Review* 643–660; Mantelero A (2014) Toward a New Approach to Data Protection in the Big Data Era. In Urs Gasser, Jonathan Zittrain, Robert Faris, Rebekah Heacock Jones (eds) *Internet Monitor 2014: Reflections on the Digital World* (Berkman Center for Internet and Society, Harvard University 2014) <https://dash.harvard.edu/handle/1/13632937>. Accessed 13 August 2021.

¹¹Mantelero A (2017) From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era. In Taylor L., Floridi L., and van der Sloot, B. *Group Privacy New Challenges of Data Technologies*. Springer International Publishing, Chm, pp. 139–158.

¹²Rouvroy A (2015) “Of data and men”. Fundamental rights and freedoms in a world of Big Data. Council of Europe–Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, T-PD-BUR(2015)09REV, Strasbourg, 11 January 2016 <https://rm.coe.int/16806a6020>. Accessed 4 August 2021.

¹³Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (2017) Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data, T-PD(2017)01, Strasbourg, 23 January 2017 <https://rm.coe.int/16806ebe7a>. Accessed 4 February 2017.

¹⁴Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (2019) Guidelines on Artificial Intelligence and Data Protection, Strasbourg, 25 January 2019, T-PD(2019)01 <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>. Accessed 13 February 2019; Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (2019) Report on Artificial Intelligence Artificial Intelligence and Data Protection: Challenges and Possible Remedies, T-PD(2018)09Rev. Rapporteur: Alessandro Mantelero <https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6>. Accessed 13 February 2019.

In 2017, with the launch of the H2020 Virt-EU project on Values and Ethics in Innovation for Responsible Technology in Europe,¹⁵ a more structured examination of the impact of Big Data yielded an assessment model that looked beyond data protection, including its collective dimension. This was the PESIA (Privacy, Ethical and Social Impact Assessment) model, which broadened the traditional privacy impact assessment to include ethical issues for society raised by the new data-intensive applications.¹⁶

The legal component of the PESIA, however, remained largely focused on data protection. A turning point in my research came in 2018 when I presented my work on PESIA at an Expert Workshop on the Right to Privacy in the Digital Age organised by the Office of the UN High Commissioner for Human Rights in Geneva, where Joe Cannataci encouraged me to look beyond data protection and consider the broader human rights scenario. This suggestion together with discussions during an EU Agency for Fundamental Rights expert meeting a few days later altered my perspective, spawning the idea of the HRESIA (Human Rights, Ethical and Social Impact Assessment) which is the focus of this book.

As is customary in academia, this initial seed was subsequently refined in conferences and seminars around Europe, as well as publications. It was also fed by my direct field experience in various ERC Executive Agency ethics committees, the Ada Lovelace Institute Rethinking Data Regulation Working Group (2019–21) and not least in the work of the Council of Europe’s Ad hoc Committee on Artificial Intelligence (CAHAI).¹⁷

After three years’ investigation of the topic—plus several periods of research in Spain, at the Universitat Oberta de Catalunya and the Universidad de Murcia, free of daily academic commitments—I hope in this book to provide a theoretical and concrete contribution to the debate on the impact of AI on society from a legal and regulatory point of view.

¹⁵Values and ethics in Innovation for Responsible Technology in Europe (IT University of Copenhagen, London School of Economics and Political Science, Uppsala Universitet, Politecnico di Torino, Copenhagen Institute of Interaction Design, and Open Rights), project information available at <https://cordis.europa.eu/project/id/732027>. Accessed 15 August 2021.

¹⁶Virt-EU Values and ethics in Innovation for Responsible Technology in Europe (2018) Deliverable 4.3. Second Report: Report to the internal members of the consortium on the PESIA methodology and initial guidelines <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c0587e55&appId=PPGMS>. Accessed 15 August 2021. See also Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data 2017 (“2.3 Since the use of Big Data may affect not only individual privacy and data protection, but also the collective dimension of these rights, preventive policies and risk-assessment shall consider the legal, social and ethical impact of the use of Big Data”).

¹⁷Council of Europe (2020) Towards regulation of AI systems. Global perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe’s standards on human rights, democracy and the rule of law. DGI (2020)16 <https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a>. Accessed 4 January 2021. See also Chap. 4.

While interest in the impact of AI on human rights and society has grown in recent years and is now explicitly mentioned in several hard and soft law AI proposals, some approaches remain focused on data protection even at the cost of stretching its boundaries. Examples include an extended interpretation of fairness, and the call for a broad use of the data protection impact assessment, reshaped as a human rights impact assessment.

Against this background, Chap. 1 looks at the limitations of data protection law in addressing the challenges of data-intensive AI, stressing how a genuinely human-oriented development of AI requires that the risks associated with AI applications be managed and regulated.

Following this recognition of the limitations and challenges, Chap. 2 develops the human rights impact assessment (HRIA),¹⁸ the central component of the HRESIA model. Although HRIAs are already in place in several contexts, the chapter emphasises the peculiarity of AI applications and the need to rethink the traditional human rights assessment.

It also aims to close the existing gap in the current regulatory proposals that recommend the introduction of HRIA but fail to furnish a methodology in line with their demands, since the quantification of potential impact that risk thresholds entail is either lacking or not fully developed in HRIA models.

Chapter 3 builds on the initial idea of the PESIA model, focusing on the ethical and societal impacts of AI, but without taking a questionnaire-based approach. The new assessment model is centred on the role of expert committees building on experience in the field of biomedicine and research.¹⁹ Such expert assessment is key to an evaluation that is necessarily contextual in the case of ethical and social issues.

Having outlined all the components of the HRESIA and their interaction, Chap. 4 compares the proposed model with the chief risk management provisions of the two European AI proposals from the Council of Europe and the European Commission. Highlighting their differences and weaknesses with respect to standard impact-assessment models, the chapter shows how the HRESIA can complement these proposals and act as an effective tool in their implementation.

The novelty of the issues at stake, the continuing debate on AI regulation, and the range of possible tools (sandboxes, auditing, certifications, etc.), as well as the recent theoretical contributions in the fields of human rights and digital technology, inevitably leave open questions on future implementations, which are discussed in the concluding chapter.

¹⁸An early version of this model appeared in Mantelero A and Esposito MS (2021) An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems. *Computer Law & Sec. Rev.* 41, doi:[10.1016/j.clsr.2021.105561](https://doi.org/10.1016/j.clsr.2021.105561), Sections 1–3, 5, and 6 (all authored by Alessandro Mantelero).

¹⁹I am grateful to María Belén Andreu Martínez (Universidad de Murcia) for comments on medical ethics provided to the draft of this chapter.

With its focus on *ex ante* risk analysis and human rights-oriented design, the book does not discuss the *ex post* remedies to harms caused by AI based on product liability and liability allocation.²⁰

As in Cavafy's poem, my research has taken me on a long journey of varied experiences, combining academic work, drafting policy and empirical analysis. I have had many travelling companions within the international community of privacy scholars. Growing year by year, it is still a small and close-knit community made up of research centres across Europe, formal and informal meetings, and leading law journals.

The book has involved me in a marvellous voyage into the global dimension of data regulation and human rights. The reader will be the judge of this work, but the closing stanzas of Cavafy's poem reflect my feelings of gratitude to all those who made some contribution, however small, to the journey and shared the experience with me:

*Without her you wouldn't have set out.
She has nothing left to give you now.*

*And if you find her poor, Ithaka won't have fooled you.
Wise as you will have become, so full of experience,
you'll have understood by then what these Ithakas mean.*

Turin, Italy
October 2021

Alessandro Mantelero

²⁰European Parliament, Policy Department for Citizens' Rights and Constitutional Affairs and Directorate-General for Internal Policies, 'Artificial Intelligence and Civil Liability' (2020) [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU\(2020\)621926_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf). Accessed 24 July 2021; European Commission (2020) Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en. Accessed 5 May 2020; European Parliament–Directorate General for Parliamentary Research Services (2020) Civil Liability Regime for Artificial Intelligence: European Added Value Assessment <https://data.europa.eu/doi/10.2861/737677>. Accessed 3 July 2021; Council of Europe, Expert Committee on human rights dimensions of automated data processing and different forms of artificial intelligence (MSI-AUT) (2019) Responsibility and AI. A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility within a Human Rights Framework. Rapporteur: Karen Yeung <https://rm.coe.int/responsability-and-ai-en/168097d9c5>. Accessed 11 July 2021; European Commission–Expert Group on Liability and New Technologies and New Technologies Formation (2019) Liability for Artificial Intelligence and Other Emerging Digital Technologies https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Accessed 3 July 2021; Lohsse S, Schulze R, and Staudenmayer D (eds) (2019) Liability for Artificial Intelligence and the Internet of Things: Münster Colloquia on EU Law and the Digital Economy IV. Baden-Baden, Nomos, Hart Publishing.

Contents

1	Beyond Data	1
1.1	Introduction	2
1.2	Rise and Fall of Individual Sovereignty Over Data Use	3
1.3	Reconsidering Self-determination: Towards a Safe Environment	10
1.4	A Paradigm Shift: The Focus on Risk Assessment	13
1.5	HRESIA: A Multi-layered Process	15
1.6	The Role of Experts	19
1.7	Assessing the Impact of Data-Intensive AI Applications: HRESIA Versus PIA/DPIA, SIA and EtIA	20
1.8	The HRESIA and Collective Dimension of Data Use	27
1.9	Advantages of the Proposed Approach	30
1.10	Summary	30
	References	32
2	Human Rights Impact Assessment and AI	45
2.1	Introduction	46
2.2	A Legal Approach to AI-Related Risks	48
2.3	Human Rights Impact Assessment of AI in the HRESIA Model	51
2.3.1	Planning and Scoping	52
2.3.2	Data Collection and the Risk Analysis Methodology	54
2.4	The Implementation of the Model	60
2.4.1	A Case Study on Consumer Devices Equipped with AI	61
2.4.2	A Large-Scale Case Study: Smart City Government	76
2.5	Summary	83
	References	85

3 The Social and Ethical Component in AI Systems Design and Management 93

3.1 Beyond Human Rights Impact Assessment 94

 3.1.1 The Socio-ethical Framework: Uncertainty, Heterogeneity and Context Dependence 96

 3.1.2 The Risk of a ‘Transplant’ of Ethical Values 97

 3.1.3 Embedding Ethical and Societal Values 101

 3.1.4 The Role of the Committee of Experts: Corporate Case Studies 104

3.2 Existing Models in Medical Ethics and Research Committees 110

 3.2.1 Clinical Ethics Committees 110

 3.2.2 Research Ethics Committees 112

 3.2.3 Ethics Committees for Clinical Trials 117

 3.2.4 Main Inputs in Addressing Ethical and Societal Issues in AI 119

3.3 Ad Hoc HRESIA Committees: Role, Nature, and Composition 121

3.4 Rights-Holder Participation and Stakeholder Engagement 127

3.5 Summary 130

References 132

4 Regulating AI 139

4.1 Regulating AI: Three Different Approaches to Regulation 140

4.2 The Principles-Based Approach 142

 4.2.1 Key Principles from Personal Data Regulation 144

 4.2.2 Key Principles from Biomedicine Regulation 152

 4.2.3 A Contribution to a Future Principles-Based Regulation of AI 158

4.3 From Design to Law – The European Approaches and the Regulatory Paradox 159

 4.3.1 The Council of Europe’s Risk-Based Approach Centred on Human Rights, Democracy and Rule of Law 161

 4.3.2 The European Commission’s Proposal (AIA) and Its Conformity-Oriented Approach 166

4.4 The HRESIA Model’s Contribution to the Different Approaches 174

4.5 Summary 176

References 177

- 5 Open Issues and Conclusions** 185
 - 5.1 Addressing the Challenges of AI 186
 - 5.2 The Global Dimension of AI 188
 - 5.3 Future Scenarios 191
 - References 195

- Index** 199

About the Author

Alessandro Mantelero is Associate Professor of Private Law and Law & Technology at the Polytechnic University of Turin, Italy, where he holds the Jean Monnet Chair in Mediterranean Digital Societies and Law. He is Council of Europe scientific expert on AI, data protection and human rights and has served as an expert on data regulation for several national and international organizations, including the United Nations, the EU Agency for Fundamental Rights, and the European Commission. He is Associate Editor of *Computer Law & Security Review* and member of the Editorial Board of *European Data Protection Law Review*.