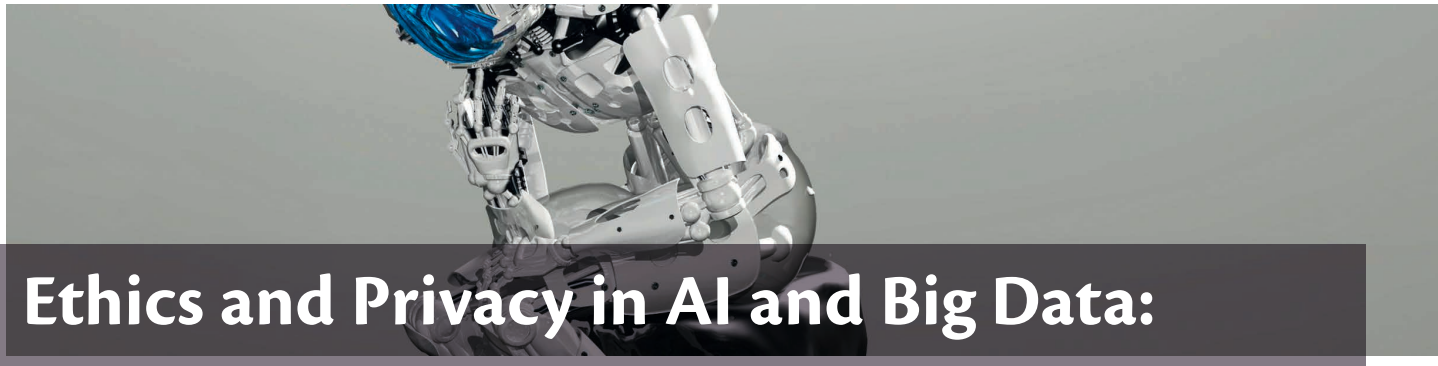

Ethics and Privacy in AI and Big Data:

Implementing Responsible Research and Innovation

Bernd Carsten Stahl | De Montfort University
David Wright | Trilateral Research

Emerging combinations of artificial intelligence, big data, and the applications these enable are receiving significant attention concerning privacy and other ethical issues. We need a way to comprehensively understand these issues and find mechanisms of addressing them that involve stakeholders, including civil society, to ensure that these technologies' benefits outweigh their disadvantages.

Information and communication technologies (ICTs) have long been recognized as having significant social and economic impact, such that they require regulatory supervision and call for ethical and social assessment. At present, one can observe two interlinking developments that have the potential to add greatly to the benefits of ICTs but that can also have undesirable impacts on ethics and human rights. These are an accelerated rate of production and collection of big data and novel ways of analyzing and using this data. Artificial intelligence (AI) and big data analytics are the key technological drivers of what we call "smart information systems" (SIS). Examples of such intelligent sociotechnical systems abound—Google's search engine, Google Translate, Amazon's recommendation system, Amazon's Alexa home assistant, Facebook's likes, smartphones with GPS tracking, predictive policing systems, automated share dealing, healthcare and surgery robots, personal fitness applications, virtual and augmented reality, and many others, ranging from social network data analysis for advertising to traffic data prediction for energy conservation.

The reason for the current prominence of SIS is that they promise solutions to many social problems and challenges. The European Commission,[1] for example, formulated a strategy promoting smart, sustainable, and inclusive growth as a means of overcoming the financial crisis of 2008, the after-effects of which are still visible and driving European policy decisions. In addition to these ongoing effects, Europe faces various challenges, ranging from demographic change to migration, social inclusion, healthcare, skills, and education. It is part of the accepted public narrative that many of these challenges require intelligent and specific solutions. SIS appear to be one way of achieving these policy goals. They have the potential to generate new sources of income, improve processes, and provide bespoke solutions that will improve workflow and user satisfaction.

At the same time, these novel SIS raise significant concerns. Privacy and data protection are the most obvious issues, but they are far from the only ones. Concerns range from questions of fairness and hidden biases in big data all the way to the possibility of truly autonomous machines that may harm or kill people, but that may also be subjects of ethical rights. A key topic of debate is the social consequence that SIS may have in the future of work and employment.

 Copublished by the IEEE Computer and Reliability Societies

In this article, we put forward the thesis that the fragmented debate around the advantages and downsides of SIS would profit from integrating various viewpoints and concerns. We suggest that the concept of responsible research and innovation (RRI) offers a perspective that allows for such an integration. We contend that RRI possesses the scope and flexibility needed to address a broad range of current and emerging ethical and social issues. By highlighting some of these issues, we demonstrate their breadth and show that an approach that is too focused on one issue, even one as encompassing and well-discussed as privacy and data protection, is not sufficient. To support this thesis, we describe how a large European project that has the potential to significantly advance both AI and big data analytics has implemented RRI.

Smart Information Systems: The Combination of AI and Big Data Analytics

There are many largely overlapping debates about the ethical and social consequences of artificial intelligence—that is, the use of hardware, software, and applications to perform big data analytics and to mimic human cognitive capabilities. These developments have pushed progress in ICT in recent years. We decided to use the term *smart information systems* as a shorthand for technologies that involve artificial intelligence, machine learning, and big data because these are all relevant for the understanding of the social consequences of these technologies.

SIS are formed and developed in an environment of what we call "enabling technologies" that generate and collect data and act on the world and interact with humans. SIS should not be seen in isolation as they tend to be enveloped in the broader technical infrastructure.

The social and ethical relevance of these technologies arises in a broader socio-economic environment that one needs to understand in order to appreciate why SIS merit broader reflection and oversight. The political context frames science and technology research and development as a means of addressing social challenges such as economic growth, environmental sustainability, demographic developments, security, and social inclusion. The political rhetoric is generally accepted and replicated by companies and research organizations. It sets the scene for positive expectations of SIS, which, nevertheless, have dark shadows relating to ethical issues and concerns.

Figure 1 gives an overview of the institutional and societal ecosystem in which SIS exist, some of the challenges they are meant to address, and examples of the ethical concerns they raise. The figure summarizes the socio-technical environment in which SIS are relevant. The top rectangle shows that SIS contain key technical drivers that provide their core capabilities, namely AI and big data. We take these to be mutually interdependent (at least in the current form of AI based on machine learning, which requires large datasets). The key technical drivers acquire their social relevance only in conjunction with other technologies that provide the required data or enable functionality. Examples of such enabling technologies are social networks and IoT (Internet of Things)-enabled sensor networks that provide large datasets as well as affective devices or autonomous machines that interact with humans. The set of enabling technologies shown in the figure is indicative. Not all SIS will use all of them, and there will likely be other technologies with enabling functions. These SIS are integrated in a broader environment where they cause tensions between the grand societal challenges they are meant to address (for example, smart growth and sustainability) and the ethical and human rights concerns they raise (notably privacy but also many others such as fairness, justice, consent, and so on).

The Ethics of Smart Information Systems

Here we discuss the ethical concerns in more detail as well as current proposals and established mechanisms for addressing them.

Ethical Issues

There is a long history of discussions of ethics and information technology, going back to the early days of modern computing technology.[2] A recent comprehensive review of ethics in ICT[3] found numerous issues discussed in the literature. SIS are a subset of ICT, and many of the ethical issues applicable to ICT are also applicable to SIS. In a review of 809 papers discussing ethics in ICT, we found that 177 addressed the issue of privacy and data protection, which makes this the most prominent issue. However, numerous other issues are also frequently discussed. These include autonomy of users, their agency, trust, consent, identity, inclusion and digital divides, security, harm, misuse, and deception, to name just a few.

In addition to the ethical issues arising from ICT more generally, there is currently much debate about specific SIS issues such as artificial intelligence, machine learning, and big data. One such forum is the Fairness, Accountability, and Transparency in Machine Learning (FAT ML) forum, which has developed "Principles for Accountable Algorithms and a Social Impact Statement for Algorithms."[4] Another takes place annually in Davos, Switzerland: the World Economic Forum, which has raised nine ethical questions to ask about AI systems.[5]

Will Knight, in an article in the *MIT Technology Review*, confronts what he calls "mind-boggling" questions when even the engineers who build AI apps

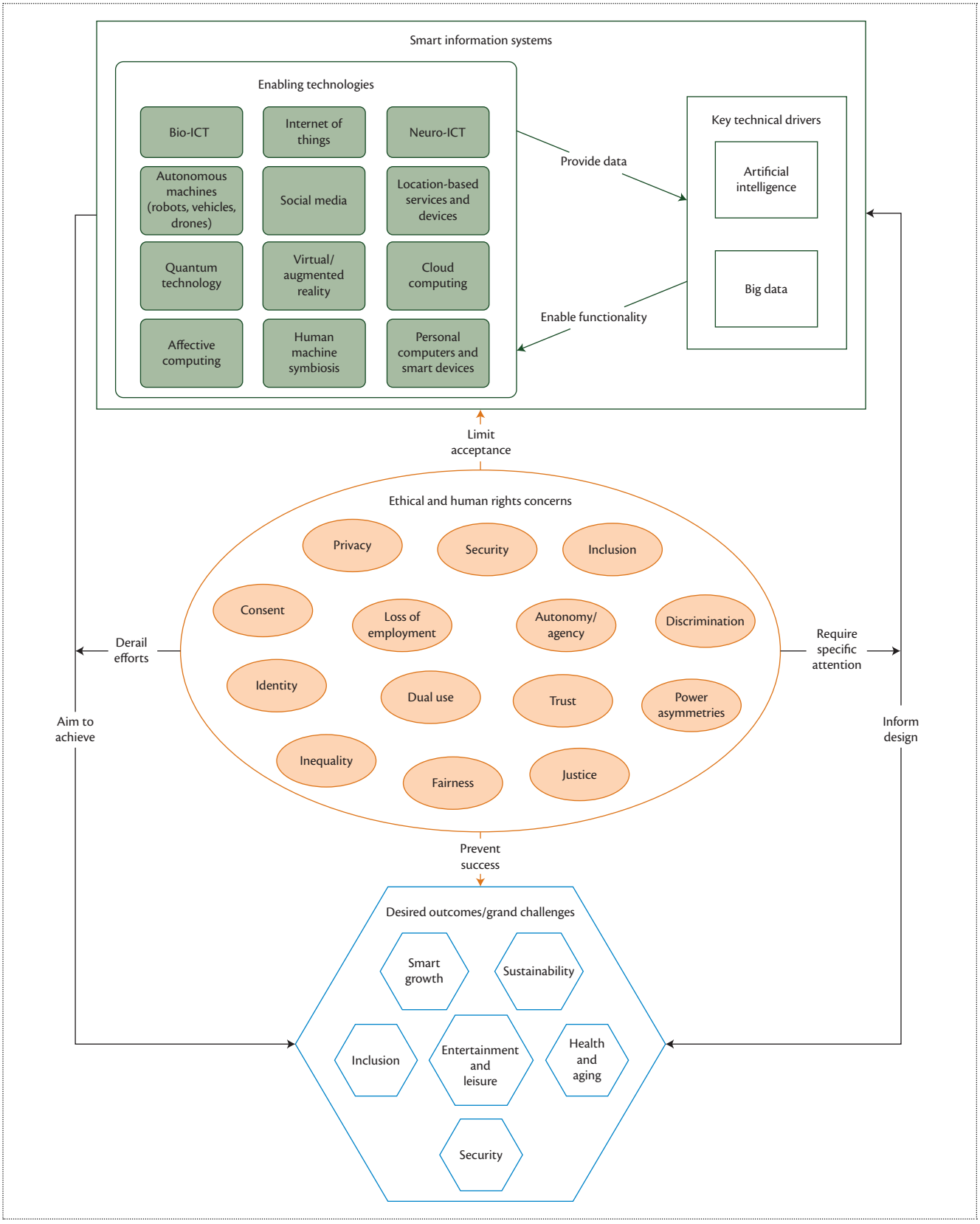

Figure 1. The ecosystem of smart information systems.

cannot fully explain their behavior.[6] How well, he asks, can we expect to communicate—and get along with—intelligent machines that could be unpredictable and inscrutable? Knight is not alone in asking questions about the ethics of AI.

To shed some light on the future of AI, a Presidential Panel, cosponsored by the Association for the Advancement of Artificial Intelligence (AAAI),[7] undertook a study on the long-term futures of AI. It came to be known as the Asilomar study after the venue in California where the panel met to review concerns about control of AI and to consider actions that could enhance societal outcomes. That Asilomar study was a decade ago. More recently, the 2017 Asilomar conference developed 23 principles applicable to AI.[8] And even more recently, the AAAI collaborated with the ACM on a conference on AI, Ethics, and Society in New Orleans in February 2018. Former US President Obama also had an interest in the future of AI.[9]

Different measures are being proposed to address the ethical issues arising from AI. Recently, Irish researchers have proposed an "Ethics by Design" methodology for AI research projects.[10] American researchers have proposed a way of detecting bias in black-box models.[9,11] Others argue in favor of ethics training. "Ethical training for AI practitioners and students is a necessary part of the solution. Ideally, every student learning AI, computer science, or data science would be exposed to a curriculum and discussion on related ethics and security topics."[9]

There seems to be growing recognition that those who create algorithms should be accountable for their consequences.[12] Accountability works if it can be enforced. Enforcement is typically a function of regulators. There have been calls to regulate AI systems by academics[13–15] and parliamentarians[16] because current ethical and legal principles are deemed insufficient in this area.

Ethical issues often are complex with unintended consequences. For example, predictive policing algorithms focus on "hot spots" or, if you like, street crime. They identify where crimes have occurred and at what times, so that police forces can deploy officers to those areas at those times in an effort to prevent crimes before they occur. The use of such algorithms raises several ethical issues. One is that they reinforce an existing spiral—that is, the police will arrest more people from those hot spot neighborhoods because more police have been deployed in those areas rather than other areas. A broader concern is that such approaches to policing can change the overall use of law enforcement resources in ways that may be undesirable or suboptimal.

Another example refers to the league tables of universities such as the one created by *US News & World Report*. While the league tables might have been created as a way of boosting the magazine's circulation, the algorithm that powered the weighting of those universities did not take into account some factors that it arguably should have (for example, which universities had the lowest tuition fees) and took into account other factors that it could have ignored (which had the best sports facilities). It has been argued that the algorithm had unwanted consequences such as increasing the cost of university education in the US and the associated student debt. The result is an education system that favors the privileged.[17] In other words, the US education system contributes to inequality—and all because of an algorithm that could have been designed better.

Many experts have noted that algorithms reflect the biases and mindsets of their creators, even if those biases and mindsets were not intended. Facebook carried out a notorious experiment in manipulating the news feeds of its subscribers—feeding some predominantly negative news and others predominantly positive news—to see how it affected the content that users then shared in response. It has been rightly castigated for not having informed its subscribers in advance.[18]

Algorithms are not only manipulating consumer behavior with targeted advertising based on extensive profiling, but they are manipulating voter intentions too, using data analysis to build detailed profiles of individuals to predict how they might vote and telling different things to different electors. The Cambridge Analytica scandal is a case in point.

There are concerns that algorithms and artificial intelligence, despite their undeniable benefits, pose negatives that not only raise ethical issues, but threaten democracy itself. Cathy O'Neil's book is subtitled "How Big Data Increases Inequality and Threatens Democracy."[17] An article in the February 2017 *Scientific American* raises similar concerns: "Will Democracy Survive Big Data and Artificial Intelligence?"[19] There are many different ways of destroying democracy, and the article points to the threat posed by personalized information systems that build a "filter bubble" around us, a kind of digital prison for our thinking.

Current technical developments such as SIS can exacerbate existing issues, but they also pose fundamentally novel ones. Among new issues, one can find hidden biases in algorithmic decision making[20] or more generally the ethical impact of increasingly autonomous machines. Policymakers are worried about questions of liability as well as practical implications of large numbers of people being made redundant by the new technologies.[21]

This brief discussion of some ethical issues raised by SIS highlights the range of issues, as indicated in Figure 1. We touched on these various issues to demonstrate the

potentially pervasive influence of SIS on society and, consequently, the need for an approach that addresses not only data protection and privacy but also other ethical and social issues. We believe that responsible research and innovation can provide such a broad approach. However, before we develop this argument in more depth, it is helpful to understand how ethics in SIS is currently addressed.

Current Responses to Ethical Issues

While we have coined the term SIS to point to a novel development, the technologies involved have a long history, as do some of the ethical questions they raise, such as privacy and data protection. Many of the SIS use personal data. Protecting such data is therefore an important step to avoid negative consequences. Many of the novel features of the European General Data Protection Regulation (GDPR) explicitly address the impact of SIS. Among the novel features relevant to SIS are breach notifications, hefty financial penalties, data protection impact assessments,[22] privacy by design,[16] and the so-called *right to be forgotten*.

Some issues raised by SIS are difficult to predict, so specific action targeting privacy or security can form only part of the solution. Aimed at the technical level, value-sensitive design[23,24] is meant to address ethical questions. The idea of value-sensitive design is starting to become regulatory reality in support of privacy by design in the GDPR.

Another way of dealing with some of the issues raised by SIS is standardization, which is increasingly used to address ethical and social issues, as is the case with the ISO 27000 family of standards. One of the most advanced standards is BSI8611 on "Robots and Robotic Devices: Guide to the Ethical Design and Application of Robots and Robotic Systems."[25] Complementing this standard is the recent IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems.[26] This effort aims to develop a family of four standards around general design principles (P7000), transparency (P7001), bias (P7002), and privacy (P7003).

Traditionally, an important way of dealing with ethics and human rights issues in ICT has been professionalism. Bodies such as the British Computer Society, the Association for Computer Machinery (ACM), and the IEEE have developed guidelines and codes of ethics aimed at addressing ethical issues. The emergence of SIS has provided the impetus for some of these bodies to rethink their stances. The ACM, for example, is currently consulting on an updated version of its general code,[27] which is supported by more specific initiatives such as the ACM Statement on Algorithmic Transparency and Accountability.[28] At the same time, researchers are adopting bottom-up approaches to develop guidelines dealing with the capabilities of emerging smart information systems, notably the recent Asilomar principles.[8]

These numerous activities and attempts to deal with ethics and human rights in SIS demonstrate that they are perceived as a significant challenge calling for a concerted effort to address them. This insight has spread among policymakers as various governments and political bodies have produced reports and recommendations on the constituent parts of smart information systems, notably AI[21,29,30] and robotics.[31] The latter report by the European Parliament may be the most far-reaching, proposing the introduction of a new regulator, a registration system, and other measures.

Combining Viable Approaches: Responsible Research and Innovation

So far, we have argued that SIS raise a complex mix of ethical and social issues and that there are numerous initiatives addressing such issues. What seems to be missing is a way of aligning the various remedies and ensuring they are joined up and create synergies. This is clearly an ambitious endeavor, given that SIS are developed in a decentralized and globalized way. It will therefore require many stakeholders and decision-makers working together toward this aim.

The concept of RRI and the discourses and practices that it has spawned offer a promising and realistic way of dealing with these challenges. RRI is a relatively novel concept that has gained prominence since about 2010. It is an attempt to rethink research and innovation governance with a view to ensuring that processes as well as outcomes of research and innovation are acceptable, desirable, and sustainable.[32] While the term RRI is relatively new, it builds on older streams of activities including technology ethics, technology assessment, science and technology studies, and philosophy of technology. There are different interpretations and foci of RRI. One key proponent of RRI is the European Commission, which emphasizes six keys of RRI: public engagement, ethics, science education, gender equality, open access, governance. The UK Engineering and Physical Science Research Council has adopted an alternative view of RRI based on anticipation of future outcomes, reflection of the research, engagement of stakeholders, and action and responsiveness. The debate about the exact definition of RRI and ways to implement it is ongoing and unlikely to lead to consensus anytime soon. However, these different positions should not be misunderstood as implying irreconcilable differences among them. All proponents of RRI agree that stakeholders need to be engaged at an early stage of research and innovation. RRI requires openness and transparency

and a willingness to be flexible and responsive to concerns, and needs to be integrated into projects as well as the broader funding and support environment.

Rather than discussing these questions in abstract terms, we now move on to a practical example of the implementation of RRI in a high-profile project that has the potential to move SIS in several ways.

Implementing RRI in SIS Research

We demonstrate the integration of RRI into current research with the Human Brain Project (HBP), one of the European Union's two current ICT flagship research projects. Its consortium includes more than 120 partner organizations. The project has a duration of 10 years and core EU funding worth more than €400 million. The HBP aims to develop an ICT infrastructure for neuroscience.[33] It is a cutting-edge SIS project in several ways. It uses massive amounts of neuroscientific data for the generation of new insights about the brain. It is developing novel computational architectures based on neuroscience that may overcome some of the limitations of current AI, such as the reliance on large training datasets and limited learning abilities. This may usher in the next generation of AI with capabilities that surpass current technologies significantly. It is therefore important to understand its social and ethical consequences.

From the time the project idea was first mooted, many people expected that the HBP would raise ethical and social questions. Hence, one of the 12 HBP subprojects is dedicated to a program of RRI under the heading of ethics and society.[34] The work in this subproject is divided into four main components: The foresight lab of the HBP investigates possible outcomes and consequences of the work undertaken by the HBP. It is accompanied by a work package that undertakes philosophical research on complex issues, such as the nature of consciousness or simulation. The third major component is public engagement that reaches out to stakeholders and the general public to discuss issues of relevance to the HBP and the public. Finally, the HBP has a work package on ethics management, that is, the administration of ethics-related issues. This includes the support of additional structures that help with ethics-related issues and structures, notably an external Ethics Advisory Board and a network of Ethics Rapporteurs who draw attention to ethics issues in the various subprojects. The HBP's RRI work includes a researcher awareness program, which reaches out to scientists and engages with them on specific issues, such as privacy by design and data governance.

HBP partners have already identified various ethical issues and are dealing with them. These include traditional biomedical research ethics issues covering both humans and animals. This means that the project has to comply with European rules on informed consent of research subjects. The project is subject to research integrity principles and intellectual property regulation. As a data integration project, the HBP needs to develop suitable data governance principles to ensure that data is treated appropriately. This covers the protection of personal data as well as ensures that data is collected in ways that are compliant with European ethical principles.

Among the broader social issues covered by the ethics and society section of the HBP, engagement of communities is especially important. This includes scientific communities that will make use of the HBP infrastructures as well as patient communities. The project is addressing the social consequences of the technologies to be developed, notably the changes to employment, the impact they may have on the role of medical doctors and neuroscientists as well as the military use of the technologies. Finally, there are fundamental and philosophical questions, such as what sets humans apart from other animals, what constitutes consciousness, and how our understanding of the brain affects our view of ourselves and our identity.

For important issues that require a more detailed understanding and specific guidance, the ethics and society group writes Opinions that are supported and implemented via specific action plans. So far, the group has produced an Opinion on data protection and is currently preparing a second one on dual use.

HBP has demonstrated that it is possible to integrate principles of RRI in complex research structures. The HBP is one example of a project that has the potential to significantly influence future developments of SIS. By integrating RRI into its structure, the HBP has embarked on a journey to explicitly reflect on ethical and social issues and not purely leave them to chance and react to them when they arise. This shows that integrating RRI into SIS research is possible.

We chose the example of the HBP for this article to show how RRI can be put into practice. The scale of the HBP project, the large number of participants, and its status as an EU flagship project have the potential to significantly influence the next stages of SIS development. Its group on ethics and society working on developing and implementing RRI can serve as an inspiration for other large SIS projects and provide lessons for other stakeholders, notably public funding bodies and research projects funded by industry.

RRI as a Response to Ethics and Privacy in Smart Information Systems

SIS raise issues related to the long-term impacts of their applications. High-powered SIS applications raise questions about the status of artificial versus natural intelligence and at what point a technical artefact acquires

moral rights and obligations. In between these very specific questions and the broad philosophical challenges, there are a plethora of issues, examples of which were mentioned above, including the prevention of algorithmic biases and the transparency of AI decisions.

The complexity of SIS technologies and their potential social and ethical impacts are such that an approach is needed that is able to learn, include external voices, incorporate reflection, and bring together different stakeholder groups. RRI offers such an approach, as its use in the HBP shows. We believe the RRI approach in the HBP can be adapted to other similar projects. The authors are partners in the EU-funded SHERPA project (SHERPA is the acronym for "Shaping the Ethical Dimensions of Information Technologies—a European Perspective"), which focuses on the social and ethical issues arising from SIS. In SHERPA, the partners will apply the RRI approach to a wide range of AI-embedded technologies.

Having said this, we do not claim that RRI is a panacea. At this point, the RRI discourse is very much focused on universities and publicly funded research, whereas innovations in SIS are being driven more by the big technology companies than by universities. However, it is notable that some of the structures of ethical reflection that are well-established in universities are finding their way into tech companies. Recent announcements that large AI investments in industry will be supervised by governance boards or ethics panels are promising.[35] The Partnership on AI (www.partnershiponai.org) represents most of the big industry players and aims to ensure that new developments are societally beneficial.

Research and innovation in SIS can bring enormous societal and economic benefits. R&I can also lead to a large number of undesirable consequences. Researchers, funders, and policymakers as well as the public at large are well aware of this. What is missing is a more unified understanding of the capabilities of emerging SIS, the impacts they may have, how they are viewed by relevant stakeholders, and what responses are needed to address them.

To contribute to this unification of the debate, we will, as part of the SHERPA project, undertake a representation of the ethical and human rights challenges of SIS through case studies, scenarios, and artistic representations. Working with stakeholders, we identify their concerns and preferred solutions (via interviews, a large-scale online survey, a Delphi study, and a stakeholder board). This will include an exploration of scientific, technical, and regulatory options. All of this work will be brought together in a workbook for the development of SIS. Following validation and prioritization exercises, the most promising proposals will be advocated to policymakers.

We will extract lessons from projects such as the HBP and other research and commercial initiatives on how to deal with SIS ethical issues. Our aim is to come to a broader consensus on what it would mean to RRI in the field of SIS. We do not believe that there will be an algorithmic solution to the ethics of SIS. The range of technologies, applications, and uses is simply too broad.

Both the HBP and SHERPA have the goal to develop a culture of responsibility where stakeholders are ready and willing to accept responsibility for the processes and outcomes of their research and innovation work.

A final observation may be that none of the RRI activities is completely failsafe. The development of SIS will lead to a number of perfectly predictable ethical concerns. Many of these can be addressed. SIS are likely to have some unexpected consequences that an approach like RRI can help identify and deal with. However, the future is not fully predictable, and we will need to live with the realization that some of the consequences of current research and innovation activities cannot be foreseen. There may thus be situations where, despite our best efforts, novel technologies including AI and big data analytics will cause ethical problems that we will have to deal with retrospectively. Despite this lack of perfection, we argue that it is advisable to accept the challenge and proactively engage with privacy and ethics of SIS, given that the alternative is to simply sit back and react to whatever happens. RRI offers a way to gain the initiative and help us shape AI and big data rather than wait for them to shape us. ■

Acknowledgments

The Human Brain Project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 720270 (HBP SGA1). The SHERPA project also benefits from EU funding under grant agreement No. 786641. The views expressed in this article are those of the authors and in no way are intended to reflect those of the European Commission.

References

1. "COM(2010) 2020: Europe 2020—A Strategy for Smart, Sustainable and Inclusive Growth," European Commission, Mar. 2010.
2. N. Wiener, *The Human Use of Human Beings*, Doubleday, 1954.
3. B.C. Stahl, J. Timmermans, and B.D. Mittelstadt, "The Ethics of Computing: A Survey of the Computing-Oriented Literature," *ACM Comput. Surv.*, vol. 48, no. 4, 2016, pp. 55:1–55:38.
4. "Principles for Accountable Algorithms and a Social Impact Statement for Algorithms," FAT/ML; https://www.fatml.org/resources/principles-for-accountable-algorithms.

5. J. Bossmann, "Top 9 Ethical Issues in Artificial Intelligence," *World Economic Forum*, 21 Oct. 2016; https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence.
6. W. Knight, "The Dark Secret at the Heart of AI," *MIT Technology Review*, 11 Apr. 2017; https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai.
7. "AAAI Presidential Panel on Long-Term AI Futures," AAAI, 2009; https://www.aaai.org/Organization/presidential-panel.php.
8. "Asilomar AI Principles," Future of Life Institute, Asilomar Conference, 2017; https://futureoflife.org/ai-principles.
9. C. Sandvig et al., "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms," *Data Discrim. Convert. Crit. Concerns Product. Inq.*, 2014, pp. 1–23.
10. M. d'Aquin et al., "Towards an 'Ethics in Design' Methodology for AI Research Projects," *AAAI/ACM Conf. Artificial Intelligence, Ethics, and Society*, 2018.
11. S. Tan et al., "Detecting Bias in Black-Box Models Using Transparent Model Distillation," Oct. 2017; https://arxiv.org/abs/1710.06169.
12. J.A. Kroll et al., "Accountable Algorithms," *Univ. Pa. Rev.*, vol. 165, 2016, pp. 633–705.
13. O.J. Erdélyi and J. Goldsmith, "Regulating Artificial Intelligence Proposal for a Global Solution," *AAAI/ACM Conference on Artificial Intelligence, Ethics and Society*, 2018.
14. A. Mantelero, "Regulating Big Data. The Guidelines of the Council of Europe in the Context of the European Data Protection Framework," *Comput. Law Secur. Rev.*, vol. 33, no. 5, 2017, pp. 584–602.
15. K. Yeung, "Algorithmic Regulation: A Critical Interrogation," *Regul. Gov.*, 2017.
16. "Privacy by Design," Information Commissioner's Office, 2008.
17. C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Penguin, 2016.
18. C. Flick, "Informed Consent and the Facebook Emotional Manipulation Study," *Res. Ethics*, vol. 12, no. 1, 2016.
19. D. Helbing et al., "Will Democracy Survive Big Data and Artificial Intelligence," *Sci. Am.*, vol. 25, Feb. 2017.
20. V. Mayer-Schonberger and K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think*, John Murray, 2013.
21. "Preparing for the Future of Artificial Intelligence," Executive Office of the President National Science and Technology Council Committee on Technology, Oct. 2016.
22. "Privacy Impact Assessment (PIA) Good Practice," CNIL, 2015.
23. B. Friedman, P. Kahn, and A. Borning, "Value Sensitive Design and Information Systems," *The Handbook of Information and Computer Ethics*, K. Himma and H. Tavani, eds., Wiley Blackwell, 2008, pp. 69–102.
24. N. Manders-Huits and J. van den Hoven, "The Need for a Value-Sensitive Design of Communication Infrastructures," *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments*, P. Sollie and M. Düwell, eds., Springer, 2009, pp. 51–62.
25. "BS8611—Robots and Robotic Devices: Guide to the Ethical Design and Application of Robots and Robotic Systems," BSI Standards Publication, BS8611:2016, Apr. 2016.
26. "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems," IEEE, 2017; https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.
27. B. Brinkman et al., "Listening to Professional Voices: Draft 2 of the ACM Code of Ethics and Professional Conduct," *Commun. ACM*, vol. 60, no. 5, 2017, pp. 105–111.
28. "Statement on Algorithmic Transparency and Accountability," ACM US Public Policy Council, Washington, Dec. 2017.
29. "Artificial Intelligence, Automation, and the Economy," Executive Office of the President National Science and Technology Council Committee on Technology, Dec. 2016.
30. "Robotics and Artificial Intelligence," House of Commons Science and Technology Committee, Sept. 2016.
31. *REPORT with Recommendations to the Commission on Civil Law Rules on Robotics*, 2015/2103(INL), Committee on Legal Affairs, European Parliament, A8–0005/2017, Jan. 2017.
32. J. Stilgoe, R. Owen, and P. Macnaghten, "Developing a Framework for Responsible Innovation," *Res. Policy*, vol. 42, no. 9, 2013, pp. 1568–1580.
33. K. Amunts et al., "The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain," *Neuron*, vol. 92, no. 3, 2016, pp. 574–581.
34. C. Aicardi, M. Reinsborough, and N. Rose, "The Integrated Ethics and Society Programme of the Human Brain Project: Reflecting on an Ongoing Experience," *J. Responsible Innov.*, Jun. 2017, pp. 1–25.
35. D. Etherington, "Microsoft Creates an AI Research Lab to Challenge Google and DeepMind," TechCrunch, 2017.

Bernd Carsten Stahl is professor of critical research in technology and director of the Centre for Computing and Social Responsibility at De Montfort University, Leicester, UK. His interests cover philosophical issues arising from the intersections of business, technology, and information. This includes ethical questions of current and emerging of ICTs, critical approaches to information systems, and issues related to responsible research and innovation. Contact at bstahl@dmu.ac.uk.

David Wright is director of Trilateral Research, a London-based company he founded in 2004. He has published more than 60 articles in peer-reviewed journals, and coedited and coauthored books, including *Privacy Impact Assessment* (Springer 2012) and *Surveillance in Europe* (Routledge 2015). Contact at david.wright@trilateralresearch.com.