

Methods to Guide Ethical Research and Design

Autonomous and intelligent systems (A/IS) research and design must be developed against the backdrop that technology is not neutral. A/IS embody values and biases that can influence important social processes like voting, policing, and banking. To ensure that A/IS benefit humanity, A/IS research and design must be underpinned by ethical and legal norms. These should be instantiated through values-based research and design methods. Such methods put human well-being at the core of A/IS development.

To help achieve these goals, researchers, product developers, and technologists across all sectors need to embrace research and development methods that evaluate their processes, products, values, and design practices in light of the concerns and considerations raised in this chapter. This chapter is split up into three sections:

Section 1—Interdisciplinary Education and Research

Section 2—Corporate Practices on A/IS

Section 3—Responsibility and Assessment

Each of the sections highlights various areas of concern (issues) as well as recommendations and further resources.

Overall, we address both structural and individual approaches. We discuss how to improve the ethical research and business practices surrounding the development of A/IS and attend to the responsibility of the technology sector vis-à-vis the public interest. We also look at that what can be done at the level of educational institutions, among others, informing engineering students about ethics, social justice, and human rights. The values-based research and design method will require a change of current system development approaches for organizations. This includes a commitment of research institutions to strong ethical guidelines for research and of businesses to values that transcend narrow economic incentives.

Methods to Guide Ethical Research and Design

Section 1—Interdisciplinary Education and Research

Integrating applied ethics into education and research to address the issues of A/IS requires an interdisciplinary approach, bringing together humanities, social sciences, physical sciences, engineering, and other disciplines.

Issue: Integration of ethics in A/IS-related degree programs

Background

A/IS engineers and design teams do not always thoroughly explore the ethical considerations implicit in their technical work and design choices. Moreover, the overall science, technology, engineering, and mathematics (STEM) field struggles with the complexity of ethical considerations, which cannot be readily articulated and translated into the formal languages of mathematics and computer programming associated with algorithms and machine learning.

Ethical issues can easily be rendered invisible or inappropriately reduced and simplified in the context of technical practice. For the dangers of this approach see for instance, Lipton and Steinhardt (2018), listed under “Further Resources”. This problem is further compounded by the fact that many STEM programs do not

sufficiently integrate applied ethics throughout their curricula. When they do, often ethics is relegated to a stand-alone course or module that gives students little or no direct experience in ethical decision-making. Ethics education should be meaningful, applicable, and incorporate best practices from the broader field.

The aim of these recommendations is to prepare students for the technical training and engineering development methods that incorporate ethics as essential so that ethics, and relevant principles, like human rights, become naturally a part of the design process.

Recommendations

- Ethics training needs to be a core subject for all those in the STEM field, beginning at the earliest appropriate level and for all advanced degrees.
- Effective STEM ethics curricula should be informed by experts outside the STEM community from a variety of cultural and educational backgrounds to ensure that students acquire sensitivity to a diversity of robust perspectives on ethics and design.
- Such curricula should teach aspiring engineers, computer scientists, and statisticians about the relevance and impact of their decisions in designing A/IS technologies. Effective

Methods to Guide Ethical Research and Design

ethics education in STEM contexts and beyond should span primary, secondary, and postsecondary education, and include both universities and vocational training schools.

- Relevant accreditation bodies should reinforce this integrated approach as outlined above.

Further Resources

- [IEEE P7000™ Standards Project for a Model Process for Addressing Ethical Concerns During System Design](#). IEEE P7000 aims to enhance corporate IT innovation practices by providing processes for embedding a values- and virtue-based thinking, culture, and practice into them.
- Z. Lipton and J. Steinhardt, [Troubling Trends in Machine Learning Scholarship](#). ICML conference paper, July 2018.
- J. Holdren, and M. Smith. "[Preparing for the Future of Artificial Intelligence](#)." Washington, DC: Executive Office of the President, National Science and Technology Council, 2016.
- Comparing the UK, EU, and US approaches to AI and ethics: C. Cath, S. Wachter, B. Mittelstadt, et al., "[Artificial Intelligence and the 'Good Society': The US, EU, and UK Approach](#)." *Science and Engineering Ethics*, vol. 24, pp. 505-528, 2017.

Issue: Interdisciplinary collaborations

Background

More institutional resources and incentive structures are necessary to bring A/IS engineers and designers into sustained and constructive contact with ethicists, legal scholars, and social scientists, both in academia and industry. This contact is necessary as it can enable meaningful interdisciplinary collaboration and shape the future of technological innovation. More could be done to develop methods, shared knowledge, and lexicons that would facilitate such collaboration.

This issue relates, among other things, to funding models as well as the lack of diversity of backgrounds and perspectives in A/IS-related institutions and companies, which limit cross-pollination between disciplines. To help bridge this gap, additional translation work and resource sharing, including websites and Massive Open Online Courses (MOOCs), need to happen among technologists and other relevant experts, e.g., in medicine, architecture, law, philosophy, psychology, and cognitive science. Furthermore, there is a need for more cross-disciplinary conversation and multi-disciplinary research, as is being done, for instance, at the annual ACM Fairness, Accountability, and Transparency (FAT*) conference or the work done by the Canadian Institute For Advanced Research (CIFAR), which is developing Canada's AI strategy.

Methods to Guide Ethical Research and Design

Recommendations

Funding models and institutional incentive structures should be reviewed and revised to prioritize projects with interdisciplinary ethics components to encourage integration of ethics into projects at all levels.

Further Resources

- S. Barocas, Course Material for Ethics and Policy in Data Science, Cornell University, 2017.
- L. Floridi, and M. Taddeo. "What Is Data Ethics?" *Philosophical Transactions of the Royal Society*, vol. 374, no. 2083, 1–4. DOI [10.1098/rsta.2016.0360](https://doi.org/10.1098/rsta.2016.0360), 2016.
- S. Spiekermann, Ethical IT Innovation: A Value-Based System Design Approach. Boca Raton, FL: Auerbach Publications, 2015.
- K. Crawford, "[Artificial Intelligence's White Guy Problem](#)", *New York Times*, July 25, 2016. [Online]. Available: http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html?_r=1. [Accessed October 28, 2018].

Issue: A/IS culture and context

Background

A responsible approach to embedding values into A/IS requires that algorithms and systems are created in a way that is sensitive to the variation of ethical practices and beliefs across cultures. The designers of A/IS need to be mindful of cross-cultural ethical variations while also respecting widely held international legal norms.

Recommendation

Establish a leading role for [intercultural information ethics](#) (IIE) practitioners in ethics committees informing technologists, policy makers, and engineers. Clearly demonstrate through examples how cultural variation informs not only information flows and information systems, but also algorithmic decision-making and value by design.

Further Resources

- D. J. Pauleen, et al. "[Cultural Bias in Information Systems Research and Practice: Are You Coming From the Same Place I Am?](#)" *Communications of the Association for Information Systems*, vol. 17, no. 17, 2006.
- J. Bielby, "[Comparative Philosophies in Intercultural Information Ethics](#)," *Confluence: Online Journal of World Philosophies* 2, no. 1, pp. 233–253, 2016.

Methods to Guide Ethical Research and Design

Issue: Institutional ethics committees in the A/IS fields

Background

It is unclear how research on the interface of humans and A/IS, animals and A/IS, and biological hazards will impact research ethical review boards. Norms, institutional controls, and risk metrics appropriate to the technology are not well established in the relevant literature and research governance infrastructure. Additionally, national and international regulations governing review of human-subjects research may explicitly or implicitly exclude A/IS research from their purview on the basis of legal technicalities or medical ethical concerns, regardless of the potential harms posed by the research.

Research on A/IS human-machine interaction, when it involves intervention or interaction with identifiable human participants or their data, typically falls to the governance of research ethics boards, e.g., institutional review boards. The national level and institutional resources, e.g., hospitals and universities, necessary to govern ethical conduct of Human-Computer Interaction (HCI), particularly within the disciplines pertinent to A/IS research, are underdeveloped.

First, there is limited international or national guidance to govern this form of research. Sections of IEEE standards governing research on A/IS in medical devices address some of the issues related to the security of A/IS-enabled devices. However, the ethics of testing those devices for the purpose of bringing them

to market are not developed into policies or guidance documents from recognized national and international bodies, e.g., U.S. Food and Drug Administration (FDA) and EU European Medicines Agency (EMA). Second, the bodies that typically train individuals to be gatekeepers for the research ethics bodies are under-resourced in terms of expertise for A/IS development, e.g., Public Responsibility in Medicine and Research (PRIM&R) and the Society of Clinical Research Associates (SoCRA). Third, it is not clear whether there is sufficient attention paid to A/IS ethics by research ethics board members or by researchers whose projects involve the use of human participants or their identifiable data.

For example, research pertinent to the ethics-governing research at the interface of animals and A/IS research is underdeveloped with respect to systematization for implementation by the Institutional Animal Care and Use Committee (IACUC) or other relevant committees. In institutions without a veterinary school, it is unclear that the organization would have the relevant resources necessary to conduct an ethical review of such research.

Similarly, research pertinent to the intersection of radiological, biological, and toxicological research—ordinarily governed under institutional biosafety committees—and A/IS research is not often found in the literature pertinent to research ethics or research governance.

Methods to Guide Ethical Research and Design

Recommendation

The IEEE and other standards-setting bodies should draw upon existing standards, empirical research, and expertise to identify priorities and develop standards for the governance of A/IS research and partner with relevant national agencies, and international organizations, when possible.

Further Resources

- S. R. Jordan, "The Innovation Imperative." *Public Management Review* 16, no. 1, pp. 67–89, 2014.
- B. Schneiderman, "[The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight.](#)" *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 48, 13538–13540, 2016.
- J. Metcalf and K. Crawford, "[Where are Human Subjects in Big Data Research? The Emerging Ethics Divide.](#)" *Big Data & Society*, May 14, 2016. [Online]. Available: SSRN: <https://ssrn.com/abstract=2779647>. [Accessed Nov. 1, 2018].
- R. Calo, "[Consumer Subject Review Boards: A Thought Experiment.](#)" *Stanford Law Review Online* 66 97, Sept. 2013.

Methods to Guide Ethical Research and Design

Section 2—Corporate Practices on A/IS

Corporations are eager to develop, deploy, and monetize A/IS, but there are insufficient structures in place for creating and supporting ethical systems and practices around A/IS funding, development, and use.

Issue: Values-based ethical culture and practices for industry

Background

Corporations are built to create profit while competing for market share. This can lead corporations to focus on growth at the expense of avoiding negative ethical consequences. Given the deep ethical implications of widespread deployment of A/IS, in addition to laws and regulations, there is a need to create values-based ethical culture and practices for the development and deployment of those systems. To do so, we need to further identify and refine corporate processes that facilitate values-based design.

Recommendations

The building blocks of such practices include top-down leadership, bottom-up empowerment, ownership, and responsibility, along with the need to consider system deployment contexts and/or ecosystems. Corporations should identify stages in their processes in which ethical considerations, “ethics filters”, are in place before products are further developed and deployed.

For instance, if an ethics review board comes in at the right time during the A/IS creation process, it would help mitigate the likelihood of creating ethically problematic designs. The institution of an ethical A/IS corporate culture would accelerate the adoption of the other recommendations within this section focused on business practices.

Further Resources

- [ACM Code of Ethics and Professional Ethics](#), which includes various references to human well-being and human rights, 2018.
- Report of UN Special Rapporteur on [Freedom of Expression. AI and Freedom of Expression](#), 2018.
- The [website of the Benefit corporations](#) (B-corporations) provides a good overview of a range of companies that personify this type of culture.
- R. Sisodia, J. N. Sheth and D. Wolfe, [Firms of Endearment](#), 2nd edition. Upper Saddle River, NJ: FT Press, 2014. This book showcases how companies embracing values and a stakeholder approach outperform their competitors in the long run.

Methods to Guide Ethical Research and Design

Issue: Values-based leadership

Background

Technology leadership should give innovation teams and engineers direction regarding which human values and legal norms should be promoted in the design of A/IS. Cultivating an ethical corporate culture is an essential component of successful leadership in the A/IS domain.

Recommendations

Companies should create roles for senior-level marketers, engineers, and lawyers who can collectively and pragmatically implement ethically aligned design. There is also a need for more in-house ethicists, or positions that fulfill similar roles. One potential way to ensure values are on the agenda in A/IS development is to have a Chief Values Officer (CVO), a role first suggested by Kay Firth-Butterfield, see “Further Resources”. However, ethical responsibility should not be delegated solely to CVOs. They can support the creation of ethical knowledge in companies, but in the end, all members of an organization will need to act responsibly throughout the design process.

Companies need to ensure that their understanding of values-based system innovation is based on *de jure* and *de facto* international human rights standards.

Further Resources

- K. Firth-Butterfield, “[How IEEE Aims to Instill Ethics in Artificial Intelligence Design](#),” The Institute. Jan. 19, 2017. [Online]. Available: <http://theinstitute.ieee.org/ieee-roundup/blogs/blog/how-ieee-aims-to-instill-ethics-in-artificial-intelligence-design>. [Accessed October 28, 2018].
- United Nations, [Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework](#), New York and Geneva: UN, 2011.
- Institute for Human Rights and Business (IHRB), and Shift, [ICT Sector Guide on Implementing the UN Guiding Principles on Business and Human Rights](#), 2013.
- C. Cath, and L. Floridi, “[The Design of the Internet’s Architecture by the Internet Engineering Task Force \(IETF\) and Human Rights](#).” *Science and Engineering Ethics*, vol. 23, no. 2, pp. 449–468, Apr. 2017.

Issue: Empowerment to raise ethical concerns

Background

Engineers and design teams may encounter obstacles to raising ethical concerns regarding their designs or design specifications within their organizations. Corporate culture should incentivize technical staff to voice the full range of ethical questions to relevant corporate actors throughout the full product lifecycle, including the design, development, and deployment

Methods to Guide Ethical Research and Design

phases. Because raising ethical concerns can be perceived as slowing or halting a design project, organizations need to consider how they can recognize and incentivize values-based design as an integral component of product development.

Recommendations

Employees should be empowered and encouraged to raise ethical concerns in day-to-day professional practice.

To be effective in ensuring adoption of ethical considerations during product development or internal implementation of A/IS, organizations should create a company culture and set of norms that encourage incorporating ethical considerations in the design and implementation processes.

New categories of considerations around these issues need to be accommodated, along with updated Codes of Conduct, company value-statements, and other management principles so individuals are empowered to share their insights and concerns in an atmosphere of trust. Additionally, bottom-up approaches like company “town hall meetings” should be explored that reward, rather than punish, those who bring up ethical concerns.

Further Resources

- [The British Computer Society \(BCS\)](#), Code of Conduct, 2019.
- C. Cath, and L. Floridi, “[The Design of the Internet’s Architecture by the Internet Engineering Task Force \(IETF\) and Human Rights](#),” *Science and Engineering Ethics*, vol. 23, no. 2, pp. 449–468, Apr. 2017.

Issue: Ownership and responsibility

Background

There is variance within the technology community on how it sees its responsibility regarding A/IS. The difference in values and behaviors are not necessarily aligned with the broader set of social concerns raised by public, legal, and professional communities. The current makeup of most organizations has clear delineations among engineering, legal, and marketing functions. Thus, technologists will often be incentivized in terms of meeting functional requirements, deadline, and financial constraints, but for larger social issues may say, “Legal will handle that.” In addition, in employment and management technology or work contexts, “ethics” typically refers to a code of conduct regarding professional behavior versus a values-driven design process mentality.

As such, ethics regarding professional conduct often implies moral issues such as integrity or the lack thereof, in the case of whistleblowing, for instance. However, ethics in A/IS design include broader considerations about the consequences of technologies.

Recommendations

Organizations should clarify the relationship between professional ethics and applied A/IS ethics by helping or enabling designers, engineers, and other company representatives to discern the differences between these kinds of ethics and where they complement each other.

Methods to Guide Ethical Research and Design

Corporate ethical review boards, or comparable mechanisms, should be formed to address ethical and behavioral concerns in relation to A/IS design, development and deployment. Such boards should seek an appropriately diverse composition and use relevant criteria, including both research ethics and product ethics, at the appropriate levels of advancement of research and development. These boards should examine justifications of research or industrial projects.

Further Resources

- HH van der Kloot Meijberg and RHJ ter Meulen, "[Developing Standards for Institutional Ethics Committees: Lessons from the Netherlands](#)," *Journal of Medical Ethics* 27 i36-i40, 2001.

Issue: Stakeholder inclusion

Background

The interface between A/IS and practitioners, as well as other stakeholders, is gaining broader attention in domains such as healthcare diagnostics, and there are many other contexts where there may be different levels of involvement with the technology. We should recognize that, for example, occupational therapists and their assistants may have on-the-ground expertise in working with a patient, who might be the "end user" of a robot or social A/IS technology. In order to develop a product that is ethically aligned, stakeholders' feedback is crucial to design a system that takes ethical and social issues into account. There are successful user experience (UX) design concepts, such

as accessibility, that consider human physical disabilities, which should be incorporated into A/IS as they are more widely deployed. It is important to continuously consider the impact of A/IS through unanticipated use and on unforeseen interests.

Recommendations

To ensure representation of stakeholders, organizations should enact a planned and controlled set of activities to account for the interests of the full range of stakeholders or practitioners who will be working alongside A/IS and incorporating their insights to build upon, rather than circumvent or ignore, the social and practical wisdom of involved practitioners and other stakeholders.

Further Resources

- C. Schroeter, et al., "[Realization and User Evaluation of a Companion Robot for People with Mild Cognitive Impairments](#)," *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2013)*, Karlsruhe, Germany 2013. pp. 1145–1151.
- T. L. Chen, et al. "[Robots for Humanity: Using Assistive Robotics to Empower People with Disabilities](#)," *IEEE Robotics and Automation Magazine*, vol. 20, no. 1, pp. 30–39, 2013.
- R. Hartson, and P. S. Pyla. *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Waltham, MA: Elsevier, 2012.

Methods to Guide Ethical Research and Design

Issue: Values-based design

Background

Ethics are often treated as an impediment to innovation, even among those who ostensibly support ethical design practices. In industries that reward rapid innovation in particular, it is necessary to develop ethical design practices that integrate effectively with existing engineering workflows. Those who advocate for ethical design within a company should be seen as innovators seeking the best outcomes for the company, end users, and society. Leaders can facilitate that mindset by promoting an organizational structure that supports the integration of dialogue about ethics throughout product life cycles.

A/IS design processes often present moments where ethical consequences can be highlighted. There are no universally prescribed models for this because organizations vary significantly in structure and culture. In some organizations, design team meetings may be brief and informal. In others, the meetings may be lengthy and structured. The transition points between discovery, prototyping, release, and revisions are natural contexts for conducting such reviews. Iterative review processes are also advisable, in part because changes to risk profiles over time can illustrate needs or opportunities for improving the final product.

Recommendations

Companies should study design processes to identify situations where engineers and researchers can be encouraged to raise and resolve questions of ethics and foster a proactive environment to realize ethically aligned design. Achieving a distributed responsibility for ethics requires that all people involved in product design are encouraged to notice and respond to ethical concerns. Organizations should consider how they can best encourage and facilitate deliberations among peers.

Organizations should identify points for formal review during product development. These reviews can focus on “red flags” that have been identified in advance as indicators of risk. For example, if the datasets involve minors or focus on users from protected classes, then it may require additional justification or alterations to the research or development protocols.

Further Resources

- A. Sinclair, “[Approaches to Organizational Culture and Ethics](#),” *Journal of Business Ethics*, vol. 12, no. 1, pp. 63–73, 1993.
- Al Y. S. Chen, R. B. Sawyers, and P. F. Williams. “[Reinforcing Ethical Decision Making Through Corporate Culture](#),” *Journal of Business Ethics* 16, no. 8, pp. 855–865, 1997.
- K. Crawford and R. Calo, “[There Is a Blind Spot in AI Research](#),” *Nature* 538, pp. 311–313, 2016.

Methods to Guide Ethical Research and Design

Section 3—Responsibility and Assessment

Lack of accountability of the A/IS design and development process presents a challenge to ethical implementation and oversight. This section presents four issues, moving from macro oversight to micro documentation practices.

Issue: Oversight for algorithms

The algorithms behind A/IS are not subject to consistent oversight. This lack of assessment causes concern because end users have no account of how a certain algorithm or system came to its conclusions. These recommendations are similar to those made in the “General Principles” and “Embedding Values into Autonomous and Intelligent Systems” chapters of *Ethically Aligned Design*, but here the recommendations are used as they apply to the narrow scope of this chapter .

Recommendations

Accountability: As touched on in the General Principles chapter of *Ethically Aligned Design*, algorithmic transparency is an issue of concern. It is understood that specifics relating to algorithms or systems contain intellectual property that cannot, or will not, be released to the general public. Nonetheless, standards providing oversight of the manufacturing process of A/IS technologies need to be created to avoid harm and negative consequences. We can look to other technical domains, such as biomedical, civil, and aerospace engineering, where commercial

protections for proprietary technology are routinely and effectively balanced with the need for appropriate oversight standards and mechanisms to safeguard the public.

Human rights and algorithmic impact assessments should be explored as a meaningful way to improve the accountability of A/IS. These need to be paired with public consultations, and the final impact assessments must be made public.

Further Resources

- F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press, 2016.
- R. Calo, “Artificial Intelligence Policy: A Primer and Roadmap,” *UC Davis Law Review*, 52: pp. 399–435, 2017.
- ARTICLE 19. “Privacy and Freedom of Expression in the Age of Artificial Intelligence,” Privacy International, April 2018. [Online]. Available: <https://www.article19.org/wp-content/uploads/2018/04/Privacy-and-Freedom-of-Expression-In-the-Age-of-Artificial-Intelligence-1.pdf>. [Accessed October 28, 2018].

Methods to Guide Ethical Research and Design

Issue: Independent review organization

Background

We need independent, expert opinions that provide guidance to the general public regarding A/IS. Currently, there is a gap between how A/IS are marketed and their actual performance or application. We need to ensure that A/IS technology is accompanied by best-use recommendations and associated warnings. Additionally, we need to develop a certification scheme for A/IS which ensures that the technologies have been independently assessed as being safe and ethically sound.

For example, today it is possible for systems to download new self-parking functionality to cars, and no independent reviewer establishes or characterizes boundaries or use. Or, when a companion robot promises to watch your children, there is no organization that can issue an independent seal of approval or limitation on these devices. We need a ratings and approval system ready to serve social/automation technologies that will come online as soon as possible. We also need further government funding for research into how A/IS technologies can best be subjected to review, and how review organizations can consider both traditional health and safety issues, as well as ethical considerations.

Recommendations

An independent, internationally coordinated body—akin to ISO—should be formed to oversee whether A/IS products actually meet ethical criteria, both when designed, developed, deployed, and when considering their evolution after deployment and during interaction with other products. It should also include a certification process.

Further Resources

- A. Tutt, “An FDA for Algorithms,” *Administrative Law Review* 69, 83–123, 2016.
- M. U. Scherer, “[Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies](#),” *Harvard Journal of Law and Technology* vol. 29, no. 2, 354–400, 2016.
- D. R. Desai and J. A. Kroll, “[Trust But Verify: A Guide to Algorithms and the Law](#).” *Harvard Journal of Law and Technology*, Forthcoming; Georgia Tech Scheller College of Business Research Paper No. 17-19, 2017.

Issue: Use of black-box components

Background

Software developers regularly use “black box” components in their software, the functioning of which they often do not fully understand. “Deep” machine learning processes, which are driving many advancements in autonomous and intelligent systems, are a growing source of black box software. At least for the foreseeable future, A/IS developers will likely be unable to build systems that are guaranteed to operate as intended.

Methods to Guide Ethical Research and Design

Recommendations

When systems are built that could impact the safety or well-being of humans, it is not enough to just presume that a system works. Engineers must acknowledge and assess the ethical risks involved with black box software and implement mitigation strategies.

Technologists should be able to characterize what their algorithms or systems are going to do via documentation, audits, and transparent and traceable standards. To the degree possible, these characterizations should be predictive, but given the nature of A/IS, they might need to be more retrospective and mitigation-oriented. As such, it is also important to ensure access to remedy adverse impacts.

Technologists and corporations must do their ethical due diligence before deploying A/IS technology. Standards for what constitutes ethical due diligence would ideally be generated by an international body such as IEEE or ISO, and barring that, each corporation should work to generate a set of ethical standards by which their processes are evaluated and modified. Similar to a flight data recorder in the field of aviation, algorithmic traceability can provide insights on what computations led to questionable or dangerous behaviors. Even where such processes remain somewhat opaque, technologists should seek indirect means of validating results and detecting harms.

Further Resources

- M. Ananny and K. Crawford, "[Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability](#)," *New Media & Society*, vol. 20, no. 3, pp. 973-989, Dec. 13, 2016.
- D. Reisman, J. Schultz, K. Crawford, and M. Whittaker, "Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability," AI NOW 2018. [Online]. Available: <https://ainowinstitute.org/aiareport2018.pdf>. [Accessed October 28, 2018].
- J. A. Kroll "[The Fallacy of Inscrutability](#)," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, C. Cath, S. Wachter, B. Mittelstadt and L. Floridi, Eds., October 15, 2018 DOI: 10.1098/rsta.2018.0084.

Issue: Need for better technical documentation

Background

A/IS are often construed as fundamentally opaque and inscrutable. However, lack of transparency is often the result of human decision. The problem can be traced to a variety of sources, including poor documentation that excludes vital information about the limitations and assumptions of a system. Better documentation combined with internal and external auditing are crucial to understanding a system's ethical impact.

Methods to Guide Ethical Research and Design

Recommendation

Engineers should be required to thoroughly document the end product and related data flows, performance, limitations, and risks of A/IS. Behaviors and practices that have been prominent in the engineering processes should also be explicitly presented, as well as empirical evidence of compliance and methodology used, such as training data used in predictive systems, algorithms and components used, and results of behavior monitoring. Criteria for such documentation could be: auditability, accessibility, meaningfulness, and readability.

Companies should make their systems auditable and should explore novel methods for external and internal auditing.

Further Resources

- S. Wachter, B. Mittelstadt, and L. Floridi. "[Transparent, Explainable, and Accountable AI for Robotics.](#)" *Science Robotics*, vol. 2, no. 6, May 31, 2017. [Online]. Available: DOI: 10.1126/scirobotics.aan6080. [Accessed Nov. 2017].
- S. Barocas, and A. D. Selbst, "[Big Data's Disparate Impact.](#)" *California Law Review* 104, 671-732, 2016.
- J. A. Kroll, J. Huey, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu. "[Accountable Algorithms.](#)" *University of Pennsylvania Law Review* 165, no. 1, 633–705, 2017.
- J. M. Balkin, "[Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation.](#)" *UC Davis Law Review*, 2017.

Methods to Guide Ethical Research and Design

Thanks to the Contributors

We wish to acknowledge all of the people who contributed to this chapter.

The Methods to Guide Ethical Research and Design Committee

- **Corinne Cath-Speth** (Co-Chair) – PhD student at Oxford Internet Institute, The University of Oxford, Doctoral student at the Alan Turing Institute, Digital Consultant at ARTICLE 19
- **Raja Chatila** (Co-Chair) – CNRS-Sorbonne Institute of Intelligent Systems and Robotics, Paris, France; Member of the French Commission on the Ethics of Digital Sciences and Technologies CERNA; Past President of IEEE Robotics and Automation Society
- **Thomas Arnold** – Research Associate at Tufts University Human-Robot Interaction Laboratory
- **Jared Bielby** – President, Netizen Consulting Ltd; Chair, International Center for Information Ethics; editor, *Information Cultures in the Digital Age*
- **Reid Blackman, PhD** – Founder & CEO Virtue Consultants, Assistant Professor of Philosophy Colgate University
- **Tom Guarriello, PhD** – Founding Faculty member in the Master’s in Branding program at New York City’s School of Visual Arts, Host of RoboPsyc Podcast and author of RoboPsyc Newsletter
- **John C. Havens** – Executive Director, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems; Executive Director, The Council on Extended Intelligence; Author, *Heartificial Intelligence: Embracing Our Humanity to Maximize Machines*
- **Sara Jordan** – Assistant Professor of Public Administration in the Center for Public Administration & Policy at Virginia Tech
- **Jason Millar** – Professor, robot ethics at Carleton University
- **Sarah Spiekermann** – Chair of the Institute for Information Systems & Society at Vienna University of Economics and Business; Author of the textbook “Ethical IT-Innovation”, the popular book “Digitale Ethik—Ein Wertesystem für das 21. Jahrhundert” and Blogger on “The Ethical Machine”
- **Shannon Vallor** – William J. Rewak Professor in the Department of Philosophy at Santa Clara University in Silicon Valley and Executive Board member of the Foundation for Responsible Robotics
- **Klein, Wilhelm E. J., PhD** – Senior Research Associate & Lecturer in Technology Ethics, City University of Hong Kong

For a full listing of all IEEE Global Initiative Members, visit standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ec_bios.pdf.

For information on disclaimers associated with EAD1e, see [How the Document Was Prepared](#).