

## Classical Ethics in A/IS

We applied classical ethics methodologies to considerations of algorithmic design in autonomous and intelligent systems (A/IS) where machine learning may or may not reflect ethical outcomes that mimic human decision-making. To meet this goal, we drew from classical ethics theories and the disciplines of machine ethics, information ethics, and technology ethics.

As direct control over tools becomes further removed, creators of autonomous systems must ask themselves how cultural and ethical presumptions bias artificially intelligent creations. Such introspection is more necessary than ever because the precise and deliberate design of algorithms in self-sustained digital systems will result in responses based on such design.

By drawing from over two thousand years' worth of classical ethics traditions, we explore established ethics systems, including both philosophical traditions (utilitarianism, virtue ethics, and deontological ethics) and religious and culture-based ethical systems (Buddhism, Confucianism, African Ubuntu traditions, and Japanese Shinto) and their stance on human morality in the digital age.<sup>1</sup> In doing so, we critique assumptions around concepts such as good and evil, right and wrong, virtue and vice, and we attempt to carry these inquiries into artificial systems' decision-making processes.

Through reviewing the philosophical foundations that define autonomy and ontology, we address the potential for autonomous capacity of artificially intelligent systems, posing questions of morality in amoral systems and asking whether decisions made by amoral systems can have moral consequences. Ultimately, we address notions of responsibility and accountability for the decisions made by autonomous systems and other artificially intelligent technologies.

## Classical Ethics in A/IS

# Section 1—Definitions for Classical Ethics in Autonomous and Intelligent Systems Research

### Issue: Assigning Foundations for Morality, Autonomy, and Intelligence

#### Background

Classical theories of economy in the Western tradition, starting with Plato and Aristotle, embrace three domains: the individual, the family, and the *polis*. The formation of the individual character (*ethos*) is intrinsically related to the others, as well as to the tasks of administration of work within the family (*oikos*). Eventually, this all expands into the framework of the *polis*, or public space (*poleis*). When we discuss ethical issues of A/IS, it becomes crucial to consider these three traditional economic dimensions, since western classical ethics was developed from this foundation and has evolved in modernity into an individual morality disconnected from economics and politics. This disconnection has been questioned and explored by thinkers such as Adam Smith, Georg W. F. Hegel, Karl Marx, and others. In particular,

Immanuel Kant's ethics located morality within the subject (see: [categorical imperative](#)) and separated morality from the outside world and the consequences of being a part of it. The moral autonomous subject of modernity became thus a worldless isolated subject. This process is important to understand in terms of ethics for A/IS since it is, paradoxically, the kind of autonomy that is supposed to be achieved by intelligent machines as humans evolve into digitally networked beings.

There lies a danger in uncritically attributing classical concepts of anthropomorphic autonomy to machines, including using the term "artificial intelligence" to describe them since, in the attempt to make them "moral" by programming moral rules into their behavior, we run the risk of assuming economic and political dimensions that do not exist, or that are not in line with contemporary human societies. While the concepts of artificial intelligence and autonomy are mainly used metaphorically as technical terms in computer science and technology, general and popular discourse may not share in the same nuanced understanding, and political and societal discourse may become distorted or

## Classical Ethics in A/IS

misleading. The question of whether A/IS and the terminology used to describe them will have any kind of impact on our conception of autonomy depends on our policy toward it. For example, the commonly held fear that A/IS will relegate humanity to mere spectators or slaves, whether realistic or not, is informed by our view of, and terminology around, A/IS. Such attitudes are flexible and can be negotiated. As noted above, present human societies are being redefined in terms of digital citizenship via online social networks. The present public debate about the replaceability of human work by “intelligent” machines is a symptom of this lack of awareness of the economic and political dimensions as defined by classical ethics, reducing ethical thinking to the “morality” of a worldless and isolated machine.

There is still value that can be gained by considering how Western ethical traditions can be integrated into either A/IS public awareness campaigns or supplemented in engineering and science education programs, as noted under the issue “Presenting ethics to the creators of A/IS”. Below is a short overview of how four different traditions can add value.

- **Virtue ethics:** Aristotle argues, using the concept of *telos*, or goal, that the ultimate goal of humans is “*eudaimonia*”, roughly translated as “flourishing”. A moral agent achieves “flourishing”—since it is an action, not a state—by constantly balancing factors including social environment, material provisions, friends, family, and one's own self. One cultivates the self through habituation, practicing and strengthening virtuous action as the “golden mean” (a principle of rationality). Such cultivation requires an appropriate
- balance between extremes of excess and deficiency, which Aristotle identifies as vices. In the context of A/IS, virtue ethics has two immediate values. First, it provides a model for iterative learning and growth, and moral value informed by context and practice, not just as compliance with a given, static ruleset. Second, it provides to those who develop and implement A/IS a framework to counterbalance tendencies toward excess, which are common in economically-driven environments.
- **Deontological ethics:** As developed by 18th century German philosopher, Immanuel Kant, the basic premise of deontological ethics addresses the concept of duty. Humans have a rational capacity to create and abide by rules that allow for duty-based ethics to emerge. Rules that produce duties are said to have value in themselves without requiring a greater-good justification. Such rules are fundamental to our existence, self-worth, and to creating conditions that allow for peaceful coexistence and interaction, e.g., the duty not to harm others; the duty not to steal. To identify rules that can be universalized and made duties, Kant uses the categorical imperative: “Act only on that maxim through which you can at the same time will that it should become a universal law.” This means the rule must be inherently desirable, doable, valuable, and others must be able to understand and follow it. Rules based merely on personal choice without wider appeal are not capable of universalization. There is mutual reciprocity in rule-making and rule adherence; if you “will” that a rule should become universal law, you not only contribute

## Classical Ethics in A/IS

to rule creation but also agree to be bound by the same rule. The rule should be action-guiding, i.e., recommending, prescribing, limiting, or proscribing action. Kant also uses the humanity formulation of the categorical imperative: “Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end.” This produces duties to respect humanity and human dignity, and not to treat either as a means to an end.

- In the context of A/IS, one consideration is to wonder if developers are acting with the best interests of humanity and human dignity in mind. This could possibly be extended to A/IS whereby they are assisting humanity as an instrument of action that has an impact on decision-making capabilities, despite being based on neural machine learning or set protocols. The humanity formulation of “the categorical imperative” has implications for various scenarios. The duty to respect human dignity may require some limitations on the functions and capability of A/IS so that they do not completely replace humans, human functions, and/or “human central thinking activities” such as judgment, discretion, and reasoning. Privacy and safeguarding issues around A/IS assisting humans, e.g., healthcare robots, may require programming certain values so that A/IS do not divulge personal information to third parties, or compromise a human’s physical or mental well-being. It may also involve preventing A/IS from deceiving or manipulating humans.
- Potential benefits and financial incentives from exploiting A/IS may provide ends-means

justifications for their use, while disregarding the treatment of humanity as an end in itself, e.g., cutting back on funding rigorous testing of A/IS before they reach the market and society. Maintaining human agency in human-machine interaction is a manifestation of the duty to respect human dignity. For example, a human has the right to know when they are interacting with A/IS, and may require consent for any A/IS interaction.

- **Utilitarian ethics:** Also called consequentialist ethics, this code of ethics refers to the consequences of one’s decisions and actions. According to the utility principle, the right course of action is the one that maximizes the utility (utilitarianism) or pleasure (hedonism) for the greatest number of people. This ethics theory does, however, warn against superficial and short-term evaluations of utility or pleasure. Therefore, it is the responsibility of the A/IS developers to consider long-term effects. Social justice is paramount in this instance, thus it must be ascertained if the implementation of A/IS will contribute to humanity, or negatively impact employment or other capabilities. Indeed, where it is deemed A/IS can supplement humanity, it should be designed in such a way that the benefits are obvious to all the stakeholders.
- **Ethics of care:** Generally viewed as an instance of feminist ethics, this approach emphasizes the importance of relationships which is context-bound. Relationships are ontologically basic to humanity, according to Nel Noddings, feminist and philosopher of education; to care for other human beings is one of our basic human attributes. For such

## Classical Ethics in A/IS

a theory to have relevance in this context, one needs to consider two criteria: 1) the relationship with the other person, or entity, must already exist or must have the potential to exist, and 2) the relationship should have the potential to grow into a caring relationship. Applied to A/IS, an interesting question comes to the foreground: Can one care for humans and their interests in tandem with non-human entities? If one expects A/IS to be beneficial to humanity, as in the instance of robots assisting with care of the elderly, then can one deduce the possibility of humans caring for A/IS? If that possibility exists, do principles of social justice become applicable to A/IS?

### Recommendations

By returning to classical ethics foundations, expand the discussion on ethics in A/IS to include a critical assessment of anthropomorphic presumptions of ethics and moral rules for A/IS. Keep in mind that machines do not, in terms of classical autonomy, comprehend the moral or legal rules they follow. They move according to their programming, following rules that are designed by humans to be moral.

Expand the discussion on ethics for A/IS to include an exploration of the classical foundations of economy, outlined above, as potentially influencing current views and assumptions around machines achieving isolated autonomy.

### Further Resources

- J. Bielby, Ed., "[Digital Global Citizenship](#)," *International Review of Information Ethics*, vol. 23, pp. 2-3, Nov. 2015.
- O. Bendel, "Towards Machine Ethics," in *Technology Assessment and Policy Areas of Great Transitions: Proceedings from the PACITA 2013 Conference in Prague*, PACITA 2013, Prague, March 13-15, 2013, T. Michalek, L. Hebáková, L. Hennen, C. Scherz, L. Nierling, J. Hahn, Eds. Prague: Technology Centre ASCR, 2014. pp. 321-326.
- O. Bendel, "[Considerations about the Relationship between Animal and Machine Ethics](#)," *AI & Society*, vol. 31, no. 1, pp. 103-108, Feb. 2016.
- N. Berberich and K. Diepold, "[The Virtuous Machine - Old Ethics for New Technology?](#)" arXiv:1806.10322 [cs.AI], June 2018.
- R. Capurro, M. Eldred, and D. Nagel, *Digital Whoness: Identity, Privacy and Freedom in the Cyberworld*. Berlin: Walter de Gruyter, 2013.
- D. Chalmers, "[The Singularity: A Philosophical Analysis](#)," *Journal of Consciousness Studies*, vol. 17, pp. 7-65, 2010.
- D. Davidson, "Representation and Interpretation," in *Modelling the Mind*, K. A. M. Said, W. H. Newton-Smith, R. Viale, and K. V. Wilkes, Eds. New York: Oxford University Press, 1990, pp. 13-26.
- N. Noddings, *Caring: A Relational Approach to Ethics and Moral Education*. Oakland, CA: University of California Press, 2013.
- O. Ulgen, "Kantian Ethics in the Age of Artificial Intelligence and Robotics," *QIL*, vol. 43, pp. 59-83, Oct. 2017.

## Classical Ethics in A/IS

- O. Ulgen, “The Ethical Implications of Developing and Using Artificial Intelligence and Robotics in the Civilian and Military Spheres,” House of Lords Select Committee, Sept. 6, 2017, UK.
- O. Ulgen, “Human Dignity in an Age of Autonomous Weapons: Are We in Danger of Losing an ‘Elementary Consideration of Humanity’?” in *How International Law Works in Times of Crisis*, I. Ziemele and G. Ulrich, Eds. Oxford: Oxford University Press, 2018.

### Issue: The Distinction between Agents and Patients

#### Background

Of particular concern when understanding the relationship between human beings and A/IS is the uncritically applied anthropomorphic approach toward A/IS that many industry and policymakers are using today. This approach erroneously blurs the distinction between moral agents and moral patients, i.e., subjects, otherwise understood as a distinction between “natural” self-organizing systems and artificial, non-self-organizing devices. As noted above, A/IS cannot, by definition, become autonomous in the sense that humans or living beings are autonomous. With that said, autonomy in machines, when critically defined, designates how machines act and operate independently in certain contexts through a consideration of implemented order generated by laws and rules. In this sense, A/IS can, by definition, qualify as

autonomous, especially in the case of genetic algorithms and evolutionary strategies. However, attempts to implant true morality and emotions, and thus accountability, i.e., autonomy, into A/IS blurs the distinction between agents and patients and may encourage anthropomorphic expectations of machines by human beings when designing and interacting with A/IS.

Thus, an adequate assessment of expectations and language used to describe the human-A/IS relationship becomes critical in the early stages of its development, where analyzing subtleties is necessary. Definitions of autonomy need to be clearly drawn, both in terms of A/IS and human autonomy. On one hand, A/IS may in some cases manifest seemingly ethical and moral decisions, resulting for all intents and purposes in efficient and agreeable moral outcomes. Many human traditions, on the other hand, can and have manifested as fundamentalism under the guise of morality. Such is the case with many religious moral foundations, where established cultural mores are neither questioned nor assessed. In such scenarios, one must consider whether there is any functional difference between the level of autonomy in A/IS and that of assumed agency—the ability to choose and act—in humans via the blind adherence to religious, traditional, or habitual mores. The relationship between assumed moral customs, the ethical critique of those customs, and the law are important distinctions.

The above misunderstanding in definitions of autonomy arises in part because of the tendency for humans to shape artificial creations in their own image, and our desire to lend our human

## Classical Ethics in A/IS

experience to shaping a morphology of artificially intelligent systems. This is not to say that such terminology cannot be used metaphorically, but the difference must be maintained, especially as A/IS begin to resemble human beings more closely. It is possible for terms like “artificial intelligence” or “morality of machines” to be used as metaphors without resulting in misunderstanding. This is how language works and how humans try to understand their natural and artificial environment.

However, the critical difference between human autonomy and autonomous systems involves questions of free will, predetermination, and being (ontology). The questions of critical ontology currently being applied to machines are not new questions to ethical discourse and philosophy; they have been thoroughly applied to the nature of human *being* as well. John Stuart Mill, for example, is a determinist and claims that human actions are predicated on predetermined laws. He does, however, argue for a reconciliation of human free will with determinism through a theory of compatibility. Millian ethics provides a detailed and informed foundation for defining autonomy that could serve to help overcome general assumptions of anthropomorphism in A/IS and thereby address the uncertainty therein (Mill, 1999).

### Recommendations

When addressing the nature of “autonomy” in autonomous systems, it is recommended that the discussion first consider free will, civil liberty, and society from a Millian perspective in order to better grasp definitions of autonomy and to address general assumptions of anthropomorphism in A/IS.

### Further Resources

- R. Capurro, “[Toward a Comparative Theory of Agents.](#)” *AI & Society*, vol. 27, no. 4, pp. 479-488, Nov. 2012.
- W. J. King and J. Ohya, “The Representation of Agents: Anthropomorphism, Agency, and Intelligence,” in *Conference Companion on Human Factors in Computing Systems*. Vancouver: ACM, 1996, pp. 289-290.
- W. Hofkirchner, “[Does Computing Embrace Self-Organisation?](#)” in *Information and Computation: Essays on Scientific and Philosophical Understanding of Foundations of Information and Computation*, G. Dodig-Crnkovic and M. Burgin, Eds. London: World Scientific, 2011, pp. 185-202.
- [International Center for Information Ethics, 2018.](#)
- J. S. Mill, *On Liberty*. London: Longman, Roberts & Green, 1869.
- P. P. Verbeek, *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. University Park, PA: Pennsylvania State University Press, 2005.

## Classical Ethics in A/IS

---

### Issue: The Need for an Accessible, Classical Ethics Vocabulary

#### Background

Philosophers and ethicists are trained in vocabulary relating to philosophical concepts and terminology. There is an intrinsic value placed on these concepts when discussing ethics and A/IS, since the layered meaning behind the terminology used is foundational to these discussions and is grounded in a subsequent entrenchment of values. Unfortunately, using philosophical terminology in cross-disciplinary instances, i.e., a conversation between technologists and policymakers, is often ineffective since not everyone has the education to be able to encompass the abstracted layers of meaning contained in philosophical terminology.

However, not understanding a philosophical definition does not detract from the necessity of its utility. While ethical and philosophical theories should not be over-simplified for popular consumption, being able to adequately translate the essence of the rich history of ethics will go a long way in supporting a constructive dialogue on ethics and A/IS. With access and accessibility concerns intricately linked with education in communities, as well as secondary and tertiary institutions, society needs to take a vested interest in creating awareness for government officials, rural communities, and school teachers. Creating a more “user-friendly” vocabulary raises awareness on the necessity and application of classical ethics to digital societies.

Identifying terms that will be intelligible to all relevant audiences is pragmatic, but care should be taken not to dilute or misrepresent concepts that are familiar to moral philosophy and ethics. One way around this is to engage in applied ethics; illustrate how a particular concept would work in the A/IS context or scenario. Another way is to understand whether terminology used across different disciplines actually has the same or similar meaning and effect which can be expressed accordingly.

#### Recommendations

Support and encourage the efforts of groups raising awareness for social and ethics committees, whose roles are to support ethics dialogue within their organizations, seeking approaches that are both aspirational and values-based. A/IS technologists should engage in cross-disciplinary exchanges whereby philosophy scholars and ethicists attend and present in non-philosophical courses. This will both raise awareness and sensitize non-philosophical scholars and practitioners to the vocabulary.

#### Further Resources

- R. T. Ames, *Confucian Role Ethics: A Vocabulary*. Hong Kong: Chinese University Press, 2011.
- R. Capurro, “[Towards an Ontological Foundation of Information Ethics](#),” *Ethics and Information Technology*, vol. 8, no. 4, pp. 175-186, 2006.
- S. Mattingly-Jordan, R. Day, B. Donaldson, P. Gray, and L. M. Ingram, “[Ethically Aligned Design, First Edition Glossary](#),” Prepared for The IEEE Global Initiative for Ethically Aligned Design, Feb. 2019.

## Classical Ethics in A/IS

- B. M. Lowe, *Emerging Moral Vocabularies: The Creation and Establishment of New Forms of Moral and Ethical Meanings*. Lanham, MD: Lexington Books, 2006.
- D. J. Flinders, "[In Search of Ethical Guidance: Constructing a Basis for Dialogue](#)," *International Journal of Qualitative Studies in Education*, vol. 5, no. 2, pp. 101-115, 1992.
- G. S. Saldanha, "[The Demon in the Gap of Language: Capurro, Ethics and Language in Divided Germany](#)," in *Information Cultures in the Digital Age*. Wiesbaden, Germany: Springer Fachmedien, 2016, pp. 253-268.
- J. Van Den Hoven and G. J. Lokhorst, "Deontic Logic and Computer Supported Computer Ethics," *Metaphilosophy*, vol. 33, no. 3, pp. 376-386, April 2002.

---

### Issue: Presenting Ethics to the Creators of Autonomous and Intelligent Systems

#### Background

The question arises as to whether or not classical ethics theories can be used to produce meta-level orientations to data collection and data use in decision-making. Keeping in mind that the task of philosophical ethics should be to examine good and evil, ethics should examine values, not prescribe them. Laws, which arise from ethics, are entrenched mores that have been critically assessed to prescribe.

The key is to embed ethics into engineering in a way that does not make ethics a servant, but instead a partner in the process. In addition to an ethics-in-practice approach, providing students and engineers with the tools necessary to build a similar orientation into their inventions further entrenches ethical design practices. In the abstract, this is not so difficult to describe, but is very difficult to encode into systems. This problem can be addressed by providing students with job aids such as checklists, flowcharts, and matrices that will help them select and use a principal ethical framework, and then exercise use of those devices with steadily more complex examples. In such an iterative process, students will start to determine for themselves what examples do not allow for perfectly clear decisions, and, in fact, require some interaction between frameworks. Produced outcomes such as videos, essays, and other formats—such as project-based learning activities—allow for a didactic strategy which proves effective in artificial intelligence ethics education.

The goal is to provide students a means to use ethics in a manner analogous to how they are being taught to use engineering principles and tools. In other words, the goal is to help engineers tell the story of what they are doing.

- Ethicists should use information flows and consider at a meta-level what information flows do and what they are supposed to do.
- Engineers should then build a narrative that outlines the iterative process of ethical considerations in their design. Intentions are part of the narrative and provide a base to reflect back on those intentions.

## Classical Ethics in A/IS

- The process then allows engineers to better understand their assumptions and adjust their intentions and design processes accordingly. They can only get to these by asking targeted questions.

This process, one with which engineers are quite familiar, is basically Kantian and Millian ethics in play.

The aim is to produce what is referred to in the computer programming lexicon as a *macro*. A macro is code that takes other code as its input(s) and produces unique outputs. This macro is built using the Western ethics tradition of virtue ethics.

This further underscores the importance of education and training on ethical considerations relating to A/IS. Such courses should be developed and presented to students of engineering, A/IS, computer science, and other relevant fields. These courses do not add value *a posteriori*, but should be embedded from the beginning to allow for absorption of the underlying ethical considerations as well as allowing for critical thinking to come to fruition once the students graduate. There are various approaches that can be considered on a tertiary level:

- Parallel (information) ethics program that is presented together with the science program during the course of undergraduate and postgraduate study;
- Embedded (information) ethics modules within the science program, i.e., one module per semester;
- Short (information) ethics courses specifically designed for the science program that can be attended by the current students, alumni, or professionals. These will function as either introductory, refresher, or specialized courses.

Courses can also be blended to include students and/or practitioners from diverse backgrounds rather than the more traditional practice of homogenous groups, such as engineering students, continuing education programs directed at a specific specialization, and the like.

### Recommendations

Find ways to present ethics where the methodologies used are familiar to engineering students. As engineering is taught as a collection of techno-science, logic, and mathematics, embedding ethical sensitivity into these objective and non-objective processes is essential. Curricula development is crucial in each approach. In addition to research articles and best practices, it is recommended that engineers and practitioners come together with social scientists and philosophers to develop case studies, interactive virtual reality gaming, and additional course interventions that are relevant to students.

### Further Resources

- T. W. Bynum and S. Rogerson, *Computer Ethics and Professional Responsibility*. Malden, MA: Wiley-Blackwell, 2003.
- E. G. Seebauer and R. L. Barry, *Fundamentals of Ethics for Scientists and Engineers*. New York: Oxford University Press, 2001.

## Classical Ethics in A/IS

- C. Whitbeck, "[Teaching Ethics to Scientists and Engineers: Moral Agents and Moral Problems](#)," *Science and Engineering Ethics*, vol. 1, no. 3, pp. 299-308, Sept. 1995.
- B. Zevenbergen, et al. "[Philosophy Meets Internet Engineering: Ethics in Networked Systems Research](#)," GTC Workshop Outcomes Paper. Oxford: Oxford Internet Institute, University of Oxford, 2015.
- M. Alvarez, "[Teaching Information Ethics](#)," *International Review of Information Ethics*, vol. 14, pp. 23-28, Dec. 2010.
- P. P. Verbeek, [Moralizing Technology: Understanding and Designing the Morality of Things](#). Chicago, IL: University of Chicago Press, 2011.
- K. A. Joyce, K. Darfler, D. George, J. Ludwig, and K. Unsworth, "[Engaging STEM Ethics Education](#)," *Engaging Science, Technology, and Society*, vol. 4, no. 1-7, 2018.

---

### Issue: Accessing Classical Ethics by Corporations and Companies

#### Background

Many companies, from startups to tech giants, understand that ethical considerations in tech design are increasingly important, but are not sure how to incorporate ethics into their tech design agenda. How can ethical considerations in tech design become an integrated part of the agenda of companies, public projects, and research consortia? Corporate workshops and exercises will need to go beyond

opinion-gathering exercises to embed ethical considerations into structures, environments, training, and development.

As it stands, classical ethics is not accessible enough to corporate endeavors in ethics, and as such, are not applicable to tech projects. There is often, but not always, a big discrepancy between the output of engineers, lawyers, and philosophers when dealing with computer science issues; there is also a large difference in how various disciplines approach these issues. While this is not true in all cases—and there are now several interdisciplinary approaches in robotics and machine ethics as well as a growing number of scientists that hold double and interdisciplinary degrees—there remains a vacuum for the wider understanding of classical ethics theories in the interdisciplinary setting. Such an understanding includes that of the philosophical language used in ethics and the translation of that language across disciplines.

If we take, for instance, the terminology and usage of the concept of "trust" in reference to technology, the term "trust" has specific philosophical, legal, and engineering connotations. It is not an abstract concept. It is attributable to humans, and relates to claims and actions people make. Machines, robots, and algorithms lack the ability to make claims and so cannot be attributed with trust. They cannot determine whether something is trustworthy or not. Software engineers may refer to "trusting" the data, but this relates to the data's authenticity and veracity to ensure software performance. In the context of A/IS, "trust" means "functional reliability"; it means there is confidence in the technology's predictability, reliability, and security against hackers or impersonators of authentic users.

## Classical Ethics in A/IS

### Recommendations

In order to achieve multicultural, multidisciplinary, and multi-sectoral dialogues between technologists, philosophers, and policymakers, a nuanced understanding in philosophical and technical language, which is critical to digital society from Internet of Things (IoT), privacy, and cybersecurity to issues of Internet governance, must be translated into norms and made available to technicians and policymakers who may not understand the nuances of the terminology in philosophical, legal, and engineering contexts. It is therefore recommended that the translation of the critical-thinking terminology of philosophers, policymakers, and other stakeholders on A/IS be translated into norms accessible to technicians.

### Further Resources

- A. Bhimani, "[Making Corporate Governance Count: The Fusion of Ethics and Economic Rationality](#)," *Journal of Management & Governance*, vol. 12, no. 2, pp. 135-147, June 2008.
- A. B. Carroll, "A History of Corporate Social Responsibility," in *The Oxford Handbook of Corporate Social Responsibility*, A. Chisanthi, R. Mansell, D. Quah, and R. Silverstone, Eds. Oxford, U.K.: Oxford University Press, 2008.
- W. Lazonick, "Globalization of the ICT Labor Force," in *The Oxford Handbook of Information and Communication Technologies*, A. Chisanthi, R. Mansell, D. Quah, and R. Silverstone, Eds. Oxford, U.K.: Oxford University Press, 2006.
- IEEE P7000™, [IEEE Standards Project for Model Process for Addressing Ethical Concerns During System Design](#) will provide engineers and technologists with an implementable process aligning innovation management processes, IT system design approaches, and software engineering methods to minimize ethical risk for their organizations, stakeholders and end users.

### Issue: The Impact of Automated Systems on the Workplace

#### Background

The impact of A/IS on the workplace and the changing power relationships between workers and employers requires ethical guidance. Issues of data protection and privacy via big data in combination with the use of autonomous systems by employers are increasing, where decisions made via aggregate algorithms directly impact employment prospects. The uncritical use of A/IS in the workplace, and its impact on employee-employer relations, is of utmost concern due to the high chance of error and biased outcome.

The concept of responsible research and innovation (RRI) is a growing area, particularly within the EU. It offers potential solutions to workplace bias and is being adopted by several research funders, such as the Engineering and Physical Sciences Research Council (EPSRC), who include RRI core principles in their mission statement. RRI is an umbrella concept that draws on classical ethics theory to provide tools to address ethical concerns from the outset of a project, from the design stage onwards.

## Classical Ethics in A/IS

Quoting Rene Von Schomberg, science and technologies studies specialist and philosopher, “Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, **sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society).**”<sup>2</sup>

When RRI methodologies are used in the ethical considerations of A/IS design, especially in response to the potential bias of A/IS in the workplace, theoretical deficiencies are then often exposed that would not otherwise have been exposed, allowing room for improvement in design at the development stage rather than from a retroactive perspective. RRI in design increases the chances of both relevance and strength in ethically aligned design.

This emerging and exciting new concept aims to also push the boundaries to incorporate relevant stakeholders whose influence in responsible research is on a global stage. While this concept initially focuses on the workplace setting, success will only be achieved through the active involvement from private companies of industry, AI Institutes, and those who are at the forefront in A/IS design. Responsible research and innovation will be achieved through careful research and innovation governance that will ensure research purpose, process, and outcomes that are acceptable, sustainable, and even desirable. It will be incumbent on RRI experts to engage at a level where private companies will feel empowered

and embrace this concept as both practical to implement and enact.

### Recommendations

It is recommended, through the application of RRI as founded in classical ethics theory, that research in A/IS design utilize available tools and approaches to better understand the design process, addressing ethical concerns from the very beginning of the design stage of the project, thus maintaining a stronger, more efficient methodological accountability throughout.

### Further Resources

- M. Burget, E. Bardone, and M. Pedaste, “Definitions and Conceptual Dimensions of Responsible Research and Innovation: A Literature Review,” *Science and Engineering Ethics*, vol. 23, no. 1, pp. 1-9, 2016.
- European Commission Communication, “[Artificial Intelligence for Europe](#),” COM 237, April, 2018.
- R. Von Schomberg, “Prospects for Technology Assessment in a Framework of Responsible Research and Innovation,” in *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methode*. Wiesbaden, Germany: Springer VS, 2011, pp. 39-61.
- B. C. Stahl, G. Eden, M. Jirotko, M. Coeckelbergh, “From Computer Ethics to Responsible Research and Innovation in ICT: The Transition of Reference Discourses\_Informing Ethics-Related Research in Information Systems,” *Information & Management*, vol. 51, no. 6, pp. 810-818, September 2014.

## Classical Ethics in A/IS

- B. C. Stahl, M. Obach, E. Yaghmaei, V. Ikonen, K. Chatfield, and A. Brem, "[The Responsible Research and Innovation \(RRI\) Maturity Model: Linking Theory and Practice](#)," Sustainability, vol. 9, no. 6, June 2017.
- IEEE P7005™, [Standards Project for Transparent Employer Data Governance](#) is

designed to provide organizations with a set of clear guidelines and certifications guaranteeing they are storing, protecting, and utilizing employee data in an ethical and transparent way.

## Section 2—Classical Ethics from Globally Diverse Traditions

### Issue: The Monopoly on Ethics by Western Ethical Traditions

#### Background

As human creators, our most fundamental values are imposed on the systems we design. It becomes incumbent on the global community to recognize which sets of values guide the design, and whether or not A/IS will generate problematic, i.e., discriminatory, consequences without consideration of non-Western values. There is an urgent need to broaden traditional ethics in its contemporary form of "responsible innovation" (RI) beyond the scope of "Western" ethical foundations, such as utilitarianism, deontology, and virtue ethics. There is also a need to include other traditions of ethics in RI, such as those inherent to Buddhism, Confucianism, and Ubuntu traditions.

However, this venture poses problematic assumptions even before the issue above can be explored. In classifying Western values, we group together thousands of years of independent and disparate ideas originating from the Greco-Roman philosophical tradition with their Christian-infused cultural heritage and then the break from that heritage with the Enlightenment. What is it that one refers to by the term "Western ethics"? Does one refer to philosophical ethics (ethics as a scientific discipline) or is the reference to Western morality?

The "West", however it may be defined, is an individualistic society, arguably more so than much of the rest of the world, and thus, in some aspects, should be even less collectively defined than "Eastern" ethical traditions. Suggest instead: If one is referring to Western values, one must designate which values and to whom they belong. Additionally, there is a danger in the field of intercultural information ethics, however

## Classical Ethics in A/IS

unconsciously or instinctively propagated, to not only group together all Western traditions under a single banner, but to negatively designate any and all Western influence in global exchange to representing an abusive collective of colonial-influenced ideals. Just because there exists a monopoly of influence by one system over another does not mean that said monopoly is devoid of value, even for systems outside itself. In the same way that culturally diverse traditions have much to offer Western tradition(s), so, too, do they have much to gain from them.

In order to establish mutually beneficial connections in addressing globally diverse traditions, it is of critical importance to first properly distinguish between subtleties in Western ethics as a discipline and morality as its object or subject matter. It is also important to differentiate between philosophical or scientific ethics and theological ethics. As noted above, the relationship between assumed moral customs, the ethical critique of those customs, and the law is an established methodology in scientific communities. Western and Eastern philosophy are very different, just like Western and Eastern ethics. Western philosophical ethics use scientific methods such as the logical, discursive, and dialectical approach (models of normative ethics) alongside the analytical and hermeneutical approaches. The Western tradition is not about education and teaching of social and moral values, but rather about the application of fundamentals, frameworks, and explanations. However, several contemporary globally relevant community mores are based in traditional and theological moral systems, requiring a conversation around how best to collaborate in

the design and programming of ethics in A/IS amidst differing ethical traditions.

While experts in Intercultural Information Ethics, such as Pak-Hang Wong, highlight the dangers of the dominance of “Western” ethics in A/IS design, noting specifically the appropriation of ethics by liberal democratic values to the exclusion of other value systems, it should be noted that those same liberal democratic values are put in place and specifically designed to accommodate such differences. However, while the accommodation of differences are, in theory, accounted for in dominant liberal value systems, the reality of the situation reveals a monopoly of, and a bias toward, established Western ethical value systems, especially when it comes to standardization. As Wong notes:

Standardization is an inherently value-laden project, as it designates the normative criteria for inclusion to the global network. Here, one of the major adverse implications of the introduction of value-laden standard(s) of responsible innovation (RI) appears to be the delegitimization of the plausibility of RI based on local values, especially when those values come into conflict with the liberal democratic values, as the local values (or, the RI based on local values) do not enable scientists and technology developers to be recognized as members of the global network of research and innovation (Wong, 2016).

It does, however, become necessary for those who do not work within the parameters of accepted value monopolies to find alternative methods of accommodating different value systems. Liberal values arose out of conflicts

## Classical Ethics in A/IS

of cultural and subcultural differences and are designed to be accommodating enough to include a rather wide range of differences.

RI enables policymakers, scientists, technology developers, and the public to better understand and respond to the social, ethical, and policy challenges raised by new and emerging technologies. Given the historical context from which RI emerges, it should not be surprising that the current discourse on RI is predominantly based on liberal democratic values. Yet, the bias toward liberal democratic values will inevitably limit the discussion of RI, especially in the cases where liberal democratic values are not taken for granted. Against this background, it is important to recognize the problematic consequences of RI solely grounded on, or justified by, liberal democratic values.

In addition, many non-Western ethics traditions, including the Buddhist and Ubuntu traditions highlighted below, view “relationship” as a foundationally important concept to ethical discourse. One of the key parameters of intercultural information ethics and RI research must be to identify main commonalities of “relationship” approaches from different cultures and how to operationalize them for A/IS to complement classical methodologies of deontological and teleological ethics. Different cultural perceptions of time may influence “relationship” approaches and impact how A/IS are perceived and integrated, e.g., technology as part of linear progress in the West; inter-generational needs and principles of respect and benevolence in Chinese culture determining current and future use of technology.

### Recommendations

In order to enable a cross-cultural dialogue of ethics in technology, discussions on ethics and A/IS must first return to normative foundations of RI to address the notion of “responsible innovation” from a range of value systems not predominant in Western classical ethics. Together with acknowledging differences, a special focus on commonalities in the intercultural understanding of the concept of “relationship” must complement the process.

### Further Resources

- J. Bielby, “Comparative Philosophies in Intercultural Information Ethics,” *Confluence: Journal of World Philosophies*, vol. 2, 2016.
- W. B. Carlin and K. C. Strong, “A Critique of Western Philosophical Ethics: Multidisciplinary Alternatives for Framing Ethical Dilemmas,” *Journal of Business Ethics*, vol. 14, no. 5, pp. 387-396, May 1995.
- C. Ess, “[Lost in translation??: Intercultural dialogues on privacy and information ethics \(introduction to special issue on privacy and data privacy protection in Asia\)](#),” *Ethics and Information Technology*, vol. 7, no. 1, pp. 1-6, March 2005.
- S. Hongladarom, “[Intercultural Information Ethics: A Pragmatic Consideration](#),” in *Information Cultures in the Digital Age*. Wiesbaden, Germany: Springer Fachmedien, 2016, pp. 191-206.
- L. G. Rodríguez and M. Á. P. Álvarez, *Ética Multicultural y Sociedad en Red*. Madrid: Fundación Telefónica, 2014.

## Classical Ethics in A/IS

- P. H. Wong, "[What Should We Share?: Understanding the Aim of Intercultural Information Ethics](#)," ACM SIGCAS Computers and Society, vol. 39, no. 3 pp. 50-58, Dec. 2009.
- S. A. Wilson, "[Conformity, Individuality, and the Nature of Virtue: A Classical Confucian Contribution to Contemporary Ethical Reflection](#)," The Journal of Religious Ethics, vol. 23, no. 2, pp. 263-289, 1995.
- P. H. Wong, "[Responsible Innovation for Decent Nonliberal Peoples: A Dilemma?](#)" Journal of Responsible Innovation, vol. 3, no. 2, pp. 154-168, July 2016.
- R. B. Zeuschner, Classical Ethics, East and West: Ethics from a Comparative Perspective. Boston, MA: McGraw-Hill, 2000.
- S. Mattingly-Jordan, "[Becoming a Leader in Global Ethics](#)," IEEE, 2017.

---

### Issue: The Application of Classical Buddhist Ethical Traditions to A/IS Design

#### Background

According to Buddhism, the field of ethics is concerned with behaving in such a way that the subject ultimately realizes the goal of liberation. The question, "How should I act?" is answered straightforwardly; one should act in such a way that one realizes liberation (nirvana)

in the future, achieving what in Buddhism is understood as "supreme happiness". Thus Buddhist ethics are clearly goal-oriented. In the Buddhist tradition, people attain liberation when they no longer endure any unsatisfactory conditions, when they have attained the state where they are completely free from any passions, including desire, anger, and delusion—to name the traditional three, that ensnare one's self against freedom. In order to attain liberation, one engages oneself in mindful behavior (ethics), concentration (meditation), and what is deemed in Buddhism as "wisdom", a term that remains ambiguous in Western scientific approaches to ethics.

Thus ethics in Buddhism are concerned exclusively with how to attain the goal of liberation, or freedom. In contrast to Western ethics, Buddhist ethics are not concerned with theoretical questions on the source of normativity or what constitutes the good life. What makes an action a "good" action in Buddhism is always concerned with whether the action leads, eventually, to liberation or not. In Buddhism, there is no questioning why liberation is a good thing. It is simply assumed. Such an assumption places Buddhism, and ethical reflection from a Buddhist perspective, in the camp of mores rather than scientifically led ethical discourse, and it is approached as an ideology or a worldview.

While it is critically important to consider, understand, and apply accepted ideologies such as Buddhism in A/IS, it is both necessary to differentiate the methodology from Western ethics, and respectful to Buddhist tradition, not to require that it be considered in a scientific

## Classical Ethics in A/IS

context. Such assumptions put it at odds with the Western foundation of ethical reflection on mores. From a Buddhist perspective, one does not ask why supreme happiness is a good thing; one simply accepts it. The relevant question in Buddhism is not about methodological reflection, but about how to attain liberation from the necessity for such reflection.

Thus, Buddhist ethics contain potential for conflict with Western ethical value systems which are founded on ideas of questioning moral and epistemological assumptions. Buddhist ethics are different from, for example, utilitarianism, which operates via critical analysis toward providing the best possible situation to the largest number of people, especially as it pertains to the good life. These fundamental differences between the traditions need to be, first and foremost, mutually understood and then addressed in one form or another when designing A/IS that span cultural contexts.

The main difference between Buddhist and Western ethics is that Buddhism is based upon a metaphysics of relation. Buddhist ethics emphasizes how *action* leads to achieving a *goal*, or in the case of Buddhism, the final goal. In other words, an action is considered a good one when it contributes to the realization of the goal. It is relational when the value of an action is relative to whether or not it leads to the goal, the goal being the reduction and eventual cessation of suffering. In Buddhism, the self is constituted through the relationship between the synergy of bodily parts and mental activities. In Buddhist analysis, the self does not actually exist as a self-subsisting entity. Liberation, or nirvana, consists in realizing that what is known to be the

self actually consists of nothing more than these connecting episodes and parts. To exemplify the above, one can draw from the concept of privacy as often explored via intercultural information ethics. The Buddhist perspective understands privacy as a protection, not of self-subsisting individuals, because such do not exist ultimately speaking, but of certain values that are found to be necessary for a well-functioning society to prosper in the globalized world.

The secular formulation of the supreme happiness mentioned above is that of the reduction of the experience of suffering, or reduction of the metacognitive state of suffering. Such a state is the result of lifelong discipline and meditation aimed at achieving proper relationships with others and with the world. This notion of the reduction of suffering is something that can resonate well with certain Western traditions, such as epicureanism ataraxia, i.e., freedom from fear through reason and discipline, and versions of consequentialist ethics that are more focused on the reduction of harm. It also encompasses the concept of phronesis or practical wisdom from virtue ethics.

Relational ethical boundaries promote ethical guidance that focuses on creativity and growth rather than solely on mitigation of consequence and avoidance of error. If the goal of the reduction of suffering can be formulated in a way that is not absolute, but collaboratively defined, this leaves room for many philosophies and related approaches as to how this goal can be accomplished. Intentionally making space for ethical pluralism is one potential antidote to dominance of the conversation by liberal thought, with its legacy of Western colonialism.

## Classical Ethics in A/IS

### Recommendations

In considering the nature of interactions between human and autonomous systems, the above notion of “proper relationships” through Buddhist ethics can provide a useful platform that results in ethical statements formulated in a relational way, instead of an absolutist way. It is recommended as an additional methodology, along with Western-value methodologies, to address human/computer interactions.

### Further Resources

- R. Capurro, “[Intercultural Information Ethics: Foundations and Applications](#),” *Journal of Information, Communication & Ethics in Society*, vol. 6, no. 2, pp. 116-126, 2008.
- C. Ess, “[Ethical Pluralism and Global Information Ethics](#),” *Ethics and Information Technology*, vol. 8, no. 4, pp. 215-226, Nov. 2006.
- S. Hongladarom, “[Intercultural Information Ethics: A Pragmatic Consideration](#),” in *Information Cultures in the Digital Age*, K. M. Bielby, Ed. Wiesbaden, Germany: Springer Fachmedien Wiesbaden, 2016, pp. 191-206.
- S. Hongladarom, J. Britz, “[Intercultural Information Ethics](#),” *International Review of Information Ethics*, vol. 13, pp. 2-5, Oct. 2010.
- M. Nakada, “Different Discussions on Roboethics and Information Ethics Based on Different Contexts (Ba). Discussions on Robots, Informatics and Life in the Information Era in Japanese Bulletin Board Forums and Mass Media,” *Proceedings*

Cultural Attitudes towards Communication and Technology, pp. 300-314, 2010.

- M. Mori, *The Buddha in the Robot*. Suginami-ku, Japan: Kosei Publishing, 1989.

---

## Issue: The Application of Ubuntu Ethical Traditions to A/IS Design

### Background

In his article, “African Ethics and Journalism Ethics: News and Opinion in Light of Ubuntu”, Thaddeus Metz frames the following question: “What does a sub-Saharan ethic focused on the good of community, interpreted philosophically as a moral theory, entail for the duties of various agents with respect to the news/opinion media?” (Metz, 2015, 1). In applying that question to A/IS, it reads: “If an ethic focused on the good of community, interpreted philosophically as a moral theory, is applied to A/IS, what would the implications be on the duties of various agents?” Agents, in this regard, would therefore be the following:

- Members of the A/IS research community
- A/IS programmers/computer scientists
- A/IS end-users
- A/IS themselves

## Classical Ethics in A/IS

Ubuntu is a sub-Saharan philosophical tradition. Its basic tenet is that a person is a person through other persons. It develops further in the notions of caring and sharing as well as identity and belonging, whereby people experience their lives as bound up with their community. A person is defined in relation to the community since the sense of being is intricately linked with belonging. Therefore, community exists through shared experiences and values. It is a commonly held maxim in the Ubuntu tradition that, “to be is to belong to a community and participate.” As the saying goes, *motho ke motho ka batho babang*, or, “a person is a person because of other people.”

Very little research, if any at all, has been conducted in light of Ubuntu ethics and A/IS, but its focus will be within the following moral domains:

1. Among the members of the A/IS research community
2. Between the A/IS community/programmers/computer scientists and the end-users
3. Between the A/IS community/programmers/computer scientists and A/IS
4. Between the end-users and A/IS
5. Between A/IS and A/IS

Considering a future where A/IS will become more entrenched in our everyday lives, one must keep in mind that an attitude of sharing one’s experiences with others and caring for their well-being will be impacted. Also, by trying to ensure solidarity within one’s community, one

must identify factors and devices that will form part of their lifeworld. If so, will the presence of A/IS inhibit the process of partaking in a community, or does it create more opportunities for doing so? One cannot classify A/IS as only a negative or disruptive force; it is here to stay and its presence will only increase. Ubuntu ethics must come to grips with, and contribute to, the body of knowledge by establishing a platform for mutual discussion and understanding. Ubuntu, as collective human dignity, may offer a way of understanding the impact of A/IS on humankind, e.g., the need for human moral and legal agency; human life and death decisions to be taken by humans rather than A/IS.

Such analysis fleshes out the following suggestive comments of Desmond Tutu, renowned former chair of South Africa’s Truth and Reconciliation Commission, when he says of Africans, “(We say) a person is a person through other people... I am human because I belong” (Tutu, 1999). As Tutu notes, “Harmony, friendliness, and community are great goods. Social harmony is for us the *summum bonum*—the greatest good. Anything that subverts or undermines this sought-after good is to be avoided” (2015:78).

In considering the above, it is fair to state that community remains central to Ubuntu. In situating A/IS within this moral domain, they will have to adhere to the principles of community, identity, and solidarity with others. On the other hand, they will also need to be cognizant of, and sensitive toward, the potential for community-based ethics to exclude individuals on the basis that they do not belong or fail to meet communitarian standards. For example,

## Classical Ethics in A/IS

would this mean the excluded individual lacks personhood and as a consequence would not be able to benefit from community-based A/IS initiatives? How would community-based A/IS programming avoid such biases against individuals?

While virtue ethics question the goal or purpose of A/IS and deontological ethics question the duties, the fundamental question asked by Ubuntu would be, "How does A/IS affect the community in which it is situated?" This question links with the initial question concerning the duties of the various moral agents within the specific community. Motivation becomes very important, because if A/IS seek to detract from community, they will be detrimental to the identity of this community when it comes to job losses, poverty, lacks in education, and lacks in skills training. However, should A/IS seek to supplement the community by means of ease of access, support systems, and more, then it cannot be argued that they will be detrimental. In between these two motivators is a safeguarding issue about how to avoid excluding individuals from accessing community-based A/IS initiatives. It therefore becomes imperative that whoever designs the systems must work closely both with ethicists and the target community, audience, or end-user to ascertain whether their needs are identified and met.

### Recommendations

It is recommended that a concerted effort be made toward the study and publication of literature addressing potential relationships between Ubuntu and other instances of African ethical traditions and A/IS value design. A/IS

designers and programmers must work closely with the end-users and target communities to ensure their design objectives, products, and services are aligned with the needs of the end-users and target communities.

### Further Resources

- D. W. Lutz, "[African Ubuntu Philosophy and Global Management](#)," *Journal of Business Ethics*, vol. 84, pp. 313-328, Oct. 2009.
- T. Metz, "[African Ethics and Journalism Ethics: News and Opinion in Light of Ubuntu](#)," *Journal of Media Ethics: Exploring Questions of Media Morality*, vol. 30 no. 2, pp. 74-90, April 2015.
- T. Metz, "[Ubuntu as a moral theory and human rights in South Africa](#)," *African Human Rights Law Journal*, vol. 11, no. 2, pp. 532-559, 2011.
- R. Nicolson, *Persons in Community: African Ethics in a Global Culture*. Scottsville, South Africa: University of KwaZulu-Natal Press, 2008.
- A. Shutte, *Ubuntu: An Ethic for a New South Africa*. Dorpspruit, South Africa: Cluster Publications, 2001.
- D. Tutu, *No Future without Forgiveness*. London: Rider, 1999.
- O. Ulgen, "Human Dignity in an Age of Autonomous Weapons: Are We in Danger of Losing an 'Elementary Consideration of Humanity'?" in *How International Law Works in Times of Crisis*, I. Ziemele and G. Ulrich, Eds. Oxford: Oxford University Press, 2018, pp. 242-272.

## Classical Ethics in A/IS

### Issue: The Application of Shinto-Influenced Traditions to A/IS Design

#### Background

Alongside the burgeoning African Ubuntu reflections on A/IS, other indigenous techno-ethical reflections boast an extensive engagement. One such tradition is Japanese Shinto indigenous spirituality, or, *Kami no michi*, often cited as the catalyst for Japanese robot and autonomous systems culture, a culture that naturally stems from the traditional Japanese concept of *karakuri ningyo* (automata). Popular Japanese artificial intelligence, robot, and video-gaming culture can be directly connected to indigenous Shinto tradition, from the existence of *kami* (spirits) to puppets and automata.

The relationship between A/IS and a human being is a personal relationship in Japanese culture and, one could argue, a very natural one. The phenomenon of “relationship” in Japan between humans and automata stands out as unique to technological relationships in world cultures, since the Shinto tradition is arguably the only animistic and naturalistic tradition that can be directly connected to contemporary digital culture and A/IS. From the Shinto perspective, the existence of A/IS, whether manifested through robots or other technological autonomous systems, is as natural to the world as rivers, forests, and thunderstorms. As noted by Spyros G. Tzafestas, author of *Roboethics: A Navigating Overview*, “Japan’s harmonious feeling

for intelligent machines and robots, particularly for humanoid ones,” (Tzafestas, 2015, 155) colors and influences technological development in Japan, especially robot culture.

The word “Shinto” can be traced to two Japanese concepts: *Shin*, meaning spirit, and *to*, the philosophical path. Along with the modern concept of the android, which can be traced back to three sources—the first, to its Greek etymology that combines *andras* (“άνδρας”), or man, and *gynoids/gyni* (“γυνή”), or woman; the second, via automatons and toys as per U.S. patent developers in the 1800s; and the third to Japan, where both historical and technological foundations for android development have dominated the market since the 1970s—Japanese Shinto-influenced technology culture is perhaps the most authentic representation of the human-automaton interface.

Shinto tradition is an animistic religious tradition, positing that everything is created with, and maintains, its own spirit (*kami*) and is animated by that spirit—an idea that goes a long way to defining autonomy in robots from a Japanese viewpoint. This includes, on one hand, everything that Western culture might deem natural, including rivers, trees, and rocks, and on the other hand, everything artificially (read: *artfully*) created, including vehicles, homes, and automata (robots). Artifacts are as much a part of nature in Shinto as animals, and they are considered naturally beautiful rather than falsely artificial.

A potential conflict between Western and Japanese concepts of nature and artifact arises when the two traditions are compared

## Classical Ethics in A/IS

and contrasted, especially in the exploration of artificial intelligence. While in Shinto, the artifact as “artificial” represents creation and authentic being, with implications for defining autonomy, the same artifact is designated as secondary and often times unnatural, false, and counterfeit in Western ethical philosophical tradition, dating back to Platonic and Christian ideas of separation of form and spirit. In both traditions, culturally presumed biases define our relationships with technology. While disparate in origin and foundation, both Western classical ethics traditions and Shinto ethical influences in modern A/IS have similar goals and outlooks for ethics in A/IS, goals that are centered in “relationship”.

### Recommendations

Where Japanese culture leads the way in the synthesis of traditional value systems and technology, we recommend that people involved with efforts in A/IS ethics explore the Shinto paradigm as representative, though not necessarily as directly applicable, to global efforts in understanding and applying traditional and classical ethics methodologies to A/IS.

### Further Resources

- R. M. Geraci, "Spiritual Robots: Religion and Our Scientific View of the Natural World," *Theology and Science*, vol. 4, no. 3, pp. 229-246, 2006.
- D. F. Holland-Minkley, "[God in the Machine: Perceptions and Portrayals of Mechanical Kami in Japanese Anime.](#)" Ph.D. dissertation, University of Pittsburgh, Pittsburgh, PA, 2010.
- C. B. Jensen and A. Blok, "[Techno-Animism in Japan: Shinto Cosmograms, Actor-Network Theory, and the Enabling Powers of Non-Human Agencies,](#)" *Theory, Culture & Society*, vol. 30, no. 2, pp. 84-115, March 2013.
- F. Kaplan, "[Who Is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots,](#)" *International Journal of Humanoid Robotics*, vol. 1, no. 3, pp. 465-480, 2004.
- S. G. Tzafestas, *Roboethics: A Navigating Overview*. Cham, Switzerland: Springer, 2015.
- G. Veruggio and K. Abney, "22 Roboethics: The Applied Ethics for a New Science," in *Robot Ethics: The Ethical and Social Implications of Robotics*. Cambridge, MA: MIT Press, 2011, p. 347.

## Classical Ethics in A/IS

# Section 3—Classical Ethics for a Technical World

### Issue: Maintaining Human Autonomy

#### Background

A/IS present the possibility for a digitally networked intellectual capacity that imitates, matches, and supersedes human intellectual capacity, including, among other things, general skills, discovery, and computing functions. In addition, A/IS can potentially acquire functionality in areas traditionally captured under the rubric of what we deem unique human and social ability. While the larger question of ethics and A/IS looks at the implications of the influence of autonomous systems in these areas, the pertinent issue is the possibility of autonomous systems imitating, influencing, and then determining the norms of human autonomy. This is done through the eventual negation of independent human thinking and decision-making, where algorithms begin to inform through targeted feedback loops what it is we *are* and what it is we should decide. Thus, how can the academic rigor of traditional ethics speak to the question of maintaining human autonomy in light of algorithmic decision-making?

How will A/IS influence human autonomy in ways that may or may not be advantageous to the good life, and perhaps—even if advantageous—may be detrimental at the same time? How do these systems affect human autonomy and decision-making through the use of algorithms when said algorithms tend to inform (“in-form”) via targeted feedback loops?

Consider, for example, Google’s autocomplete tool, where algorithms attempt to determine one’s search parameters via the user’s initial keyword input, offering suggestions based on several criteria including search patterns. In this scenario, autocomplete suggestions influence, in real-time, the parameters the user phrases their search by, often reforming the user’s perceived notions of what it was they were looking for in the first place, versus what they might have actually originally intended.

Targeted algorithms also inform, as per emerging IoT, applications that monitor the user’s routines and habits in the analog world. Consider for example that our bioinformation is, or soon will be, available for interpretation by autonomous systems. What happens when autonomous systems can inform the user in ways the user is not even aware of, using one’s bioinformation in targeted advertising campaigns that seek to influence the user in real-time feedback loops based on the user’s biological reactions such as

## Classical Ethics in A/IS

pupil dilation, body temperature, and emotional reaction, whether positive or negative, to that very same advertising, using information about our being to in-form and re-form our being? On the other hand, it becomes important not to adopt dystopian assumptions concerning autonomous machines threatening human autonomy.

The tendency to think only in negative terms presupposes a case for interactions between autonomous machines and human beings, a presumption not necessarily based in evidence. Ultimately, the behavior of algorithms rests solely in their design, and that design rests solely in the hands of those who designed them. Perhaps more importantly, however, is the matter of choice in terms of how the user chooses to interact with the algorithm. Users often don't know when an algorithm is interacting with them directly or their data which acts as a proxy for their identity. Should there be a precedent for the A/IS user to know when they are interacting with an algorithm? What about consent?

The responsibility for the behavior of algorithms remains with the designer, the user, and a set of well-designed guidelines that guarantee the importance of human autonomy in any interaction. As machine functions become more autonomous and begin to operate in a wider range of situations, any notion of those machines working for or against human beings becomes contested. Does the machine work *for* someone in particular, or for particular groups but not others? Who decides on the parameters? Is it the machine itself? Such questions become key factors in conversations around ethical standards.

### Recommendations

A two-step process is recommended to maintain human autonomy in A/IS. The creation of an ethics-by-design methodology is the first step to addressing human autonomy in A/IS, where a critically applied ethical design of autonomous systems preemptively considers how and where autonomous systems may or may not dissolve human autonomy. The second step is the creation of a pointed and widely applied education curriculum that spans grade school through university, one based on a classical ethics foundation that focuses on providing choice and accountability toward digital being as a priority in information and knowledge societies.

### Further Resources

- B. van den Berg and J. de Mul, "Remote Control. Human Autonomy in the Age of Computer-Mediated Agency," in *Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology*, M. Hildebrandt and A. Rouvroy, Eds. London: Routledge, 2011, pp. 46-63.
- L. Costa, "[A World of Ambient Intelligence](#)," in *Virtuality and Capabilities in a World of Ambient Intelligence*. Cham, Switzerland: Springer International, 2016, pp. 15-41.
- P. P. Verbeek, "[Subject to Technology on Autonomic Computing and Human Autonomy](#)," in *The Philosophy of Law Meets the Philosophy of Technology: Autonomic Computing and Transformations of Human Agency*, M. Hildebrandt and A. Rouvroy, Eds. New York: Routledge, 2011.

## Classical Ethics in A/IS

- D. Reisman, J. Schultz, K. Crawford, and M. Whittaker, "[Algorithmic Impact Assessments: A practical Framework for Public Agency Accountability](#)," AI NOW, April 2018.
- A. Chaudhuri, "[Philosophical Dimensions of Information and Ethics in the Internet of Things \(IoT\) Technology](#)," EDPACS, vol. 56, no. 4, pp. 7-18, Nov. 2017.

### Issue: Implications of Cultural Migration in A/IS

#### Background

In addition to developing an understanding of A/IS via different cultures, it is crucial to understand how A/IS are shaped and reshaped—how they affect and are affected by—human mobility and cultural diversity through active immigration. The effect of human mobility on state systems reliant on A/IS impacts the State structure itself, and thus the systems that the structure relies on, in the end influencing everything from democracy to citizenship. Where the State, through A/IS, invests in and gathers big data through mechanisms for registration and identification of people, mainly immigrants, human mobility becomes a foundational component in a system geared toward the preservation of human dignity.

Traditional national concerns reflect two information foundations: information produced for human rights and information produced for national sovereignty. In the second foundation, State borders are considered the limits from which political governance is defined in terms of

security. The preservation of national sovereignty depends on the production and domination of knowledge. In the realm of migratory policies, knowledge is created to measure people in transit: collecting, treating, and transferring information about territory and society.

Knowledge organization has been the paramount pillar of scientific thought and scientific practice since the beginning of written civilization. Any scientific and technological development has only been possible through information policies that include the establishment of management processes to systematize them, and the codification of language. For the Greeks, this process was closely associated with the concept of *arete*, or the excellence of one's self in politics as congregated in the *polis*. The notion of *polis* is as relevant as ever in the digital age with the development of digital technologies and the discussions around morality in A/IS. Where the systematization of knowledge is potentially freely created, the advent of the Internet and its flows are difficult to control. Ethical issues about the production of information are becoming paramount to our digital society.

The advancement of the fields of science and technology has not been followed by innovations in the political community, and the technical community has repeatedly tabled academic discussions about the hegemony of technocracy over policy issues, restricting the space of the policy arena and valorizing excessively technic solutions for human problems. This monopoly alters conceptions of morality, relocating the locus of the Kantian "Categorical Imperative", causing the tension among different social and political contexts to become more pervasive.

## Classical Ethics in A/IS

Current global migration dynamics have been met by unfavorable public opinion based in ideas of crisis and emergency, a response vastly disproportionate to what statistics have shown to be the reality. In response to these views, A/IS are currently designed and applied to measure, calculate, identify, register, systematize, normalize, and frame both human rights and security policies. This is largely no different of a process than what has been practiced since the period of colonialism. It includes the creation and implementation of a set of ancient and new technologies. Throughout history, mechanisms have been created firstly to identify and select individuals who share certain biological heritage, and secondly to individuals and social groups, including biological characteristics.

Information is only possible when materialized as an infrastructure supported by ideas in action as a “communicative act”, which Habermas (1968) identifies in Hegel’s work, converging three elements in human-in-the-world relationships: symbol, language, and labor. Information policies reveal the importance and the strength in which technologies influence economic, social, cultural, identity, and ethnic interactions.

Traditional mechanisms used to control migration, such as the passport, are associated with globally established walls and fences. The more intense human mobility becomes, the more amplified are the discourses to discourage it, restricting human migrations, and deepening the need for an ethics related to conditions of citizenship. Together with the building of walls, other remote technologies are developed to monitor and surveil borders, buildings, and streets, also impacting ideas and

moral presumptions of citizenship. Closed Circuit Television(CCTV), Unmanned Aerial Vehicles (UAVs), and satellites allow data transference in real time to databases, cementing the backbone that A/IS draws from, often with bias as per the expectations of developed countries. This centrality of data sources for A/IS expresses a divide between developed and underdeveloped countries, particularly as relevant to the refugee.

Information is something that links languages, habits, customs, identification, and registration technologies. It provokes a reshaping of the immigrants’ and refugees’ citizenship and their value as people in terms of their citizenship, as they seek forms of surviving in, and against, the restrictions imposed by A/IS for surveillance and monitoring in an enlarged and more complex cosmopolis.

An understanding of the impact of A/IS on migration and mobile populations, as used in state systems, is a critical first step to consider if systems are to become truly autonomous and intelligent, especially beyond the guidance of human deliberation. Digital technology systems used to register and identify human mobility, including refugees and other displaced populations, are not autonomous in the intelligent sense, and are dependent on the biases of worldviews around immigration. In this aspect, language is the locus where this dichotomy has to be considered to understand the diversity of morals when there are contacts among different cultures.

## Classical Ethics in A/IS

### Recommendations

Is it recommended that the State become a proactive player in the globalized processes of A/IS for migrant and mobile populations, introducing a series of mechanisms that limit the segregation of social spaces and groups, and consider the biases inherent in surveillance for control.

### Further Resources

- I. About and V. Denis, *Histoire de l'identification des personnes*. Paris: La Découverte, 2010.
- I. About, J. Brown, G. Lonergan, *Identification and Registration Practices in Transnational Perspective: People, Papers and Practices*. London: Palgrave Macmillan, 2013, pp. 1-13.
- D. Bigo, "Security and Immigration: Toward a Critique of the Governmentality of Unease," in *Alternatives*, Special Issue, no. 27. pp. 63-92, 2002.
- R. Capurro, "[Citizenship in the Digital Age](#)," in *Information Ethics, Globalization and Citizenship*, T. Samek and L. Schultz, Eds. Jefferson NC: McFarland, 2017, pp. 11-30.
- R. Capurro, "[Intercultural Information Ethics](#)," in *Localizing the Internet: Ethical Aspects in Intercultural Perspective*, R. Capurro, J. Frühbauer, and T. Hausmanninger, Eds. Munich: Fink, 2007, pp. 21-38.
- UN High Commissioner for Refugees (UNHCR), [Policy on the Protection of Personal Data of Persons of Concern to UNHCR](#), May 2015.

### Issue: Applying Goal-Directed Behavior (Virtue Ethics) to Autonomous and Intelligent Systems

#### Background

Initial concerns regarding A/IS also include questions of function, purpose, identity, and agency, a continuum of goal-directed behavior with function being the most primitive expression. How can classical ethics act as a regulating force in autonomous technologies as goal-directed behavior transitions from being externally set by operators to being internally set? The question is important not just for safety reasons, but for mutual productivity. If autonomous systems are to be our trusted, creative partners, then we need to be confident that we possess mutual anticipation of goal-directed action in a wide variety of circumstances.

A virtue ethics approach has merits for accomplishing this even without having to posit a "character" in an autonomous technology, since it places emphasis on habitual, iterative action focused on achieving excellence in a chosen domain or in accord with a guiding purpose. At points on the goal-directed continuum associated with greater sophistication, virtue ethics become even more useful by providing a framework for prudent decision-making that is in keeping with the autonomous system's purpose, but allows for creativity in how to achieve the purpose in a way that still allows for a degree of predictability. An ethics approach that does not rely on a decision

## Classical Ethics in A/IS

to refrain from transgressing, but instead to prudently pursue a sense of purpose informed by one's identity, might provide a greater degree of insight into the behavior of the system.

### Recommendations

Program autonomous systems to be able to recognize user behavior for the purposes of predictability, traceability, and accountability and to hold expectations, as an operator and co-collaborator, whereby both user and system mutually recognize the decisions of the autonomous system as virtue ethics-based.

### Further Resources

- M. A. Boden, Ed. *The Philosophy of Artificial Life*. Oxford, U.K.: Oxford University Press, 1996.
- C. Castelfranchi, "Modelling Social Action for AI Agents," *Artificial Intelligence*, vol. 103, no.1-2, pp. 157-182, 1998.
- W. D. Christensen and C. A. Hooker, "Anticipation in Autonomous Systems: Foundations for a Theory of Embodied Agents," *International Journal of Computing Anticipatory Systems*, vol. 5, pp. 135-154, Dec. 2000.
- K. G. Coleman, "Android Arete: Toward a Virtue Ethic for Computational Agents," *Ethics and Information Technology*, vol. 3, no. 4, pp. 247-265, 2001.
- J. G. Lennox, "Aristotle on the Biological Roots of Virtue," *Biology and the Foundations of Ethics*, J. Maienschein and M. Ruse, Eds. Cambridge, U.K.: Cambridge University Press, 1999, pp. 405-438.
- L. Muehlhauser and L. Helm, "The Singularity and Machine Ethics," in *Singularity Hypotheses*, A. H. Eden, J. H. Moor, J. H. Soraker, and E. Steinhart, Eds. Berlin: Springer, 2012, pp. 101-126.
- D. Vernon, G. Metta, and G. Sandini, "[A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents](#)," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 2, pp. 151-180, April 2007.

---

## Issue: A Requirement for Rule-Based Ethics in Practical Programming

### Background

Research in machine ethics focuses on simple moral machines. It is deontological ethics and [teleological ethics](#) that are best suited to the kind of practical programming needed for such machines, as these ethical systems are abstractable enough to encompass ideas of non-human agency, whereas most modern ethics approaches are far too human-centered to properly accommodate the task.

In the deontological model, duty is the point of departure. Duty can be translated into rules. It can be distinguished into rules and metarules. For example, a rule might take the form "Don't lie!", whereas a metarule would take the form of Kant's categorical imperative: "Act only according to that maxim whereby you can, at the same time, will that it should become a universal law."

## Classical Ethics in A/IS

A machine can follow simple rules. Rule-based systems can be implemented as formal systems, also referred to as “axiomatic systems”, and in the case of machine ethics, a set of rules is used to determine which actions are morally allowable and which are not. Since it is not possible to cover every situation by a rule, an [inference engine](#) is used to deduce new rules from a small set of simple rules called axioms by combining them. The morality of a machine comprises the set of rules that is deducible from the axioms.

Formal systems have an advantage since properties such as decidability and consistency of a system can be effectively examined. If a formal system is decidable, every rule is either morally allowable or not, and the “unknown” is eliminated. If the formal system is consistent, one can be sure that no two rules can be deduced that contradict each other. In other words, the machine never has moral doubt about an action and never encounters a deadlock.

The disadvantage of using formal systems is that many of them work only in closed worlds like computer games. In this case, what is not known is assumed to be false. This is in drastic conflict with real world situations, where rules can conflict and it is impossible to take into account the totality of the environment. In other words, consistent and decidable formal systems that rely on a closed world assumption can be used to implement an ideal moral framework for a machine, yet they are not viable for real world tasks.

One approach to avoiding a closed world scenario is to utilize self-learning algorithms, such as case-

based reasoning approaches. Here, the machine uses “experience” in the form of similar cases that it has encountered in the past or uses cases which are collected in databases.

In the context of the *teleological model*, the consequences of an action are assessed. The machine must know the consequences of an action and what the action’s consequences mean for humans, for animals, for things in the environment, and, finally, for the machine itself. It also must be able to assess whether these consequences are good or bad, or if they are acceptable or not, and this assessment is not absolute. While a decision may be good for one person, it may be bad for another; while it may be good for a group of people or for all of humanity, it may be bad for a minority of people. An implementation approach that allows for the consideration of potentially contradictory subjective interests may be realized by decentralized reasoning approaches such as agent-based systems. In contrast to this, centralized approaches may be used to assess the overall consequences for all involved parties.

### Recommendations

By applying the classical methodologies of deontological and teleological ethics to machine learning, rules-based programming in A/IS can be supplemented with established praxis, providing both theory and a practicality toward consistent and determinable formal systems.

## Classical Ethics in A/IS

### Further Resources

- C. Allen, I. Smit, and W. Wallach, "Artificial Morality: Top-Down, Bottom-Up, and Hybrid Approaches," *Ethics and Information Technology*, vol. 7, no. 3, pp. 149-155, 2005.
- O. Bendel, [Die Moral in der Maschine: Beiträge zu Roboter-und Maschinenethik](#). Heise Medien, 2016.
- O. Bendel, Oliver, *Handbuch Maschinenethik*. Wiesbaden, Germany: Springer VS, 2018.
- M. Fisher, L. Dennis, and M. Webster, "[Verifying Autonomous Systems](#)," *Communications of the ACM*, vol. 56, no. 9, pp. 84-93, Sept. 2013.
- B. M. McLaren, "[Computational Models of Ethical Reasoning: Challenges, Initial Steps, and Future Directions](#)," *IEEE Intelligent Systems*, vol. 21, no. 4, pp. 29-37, July 2006.
- M. A. Perez Alvarez, "[Tecnologías de la Mente y Exocerebro o las Mediaciones del Aprendizaje](#)," 2015.
- E. L. Rissland and D. B. Skalak, "Combining Case-Based and Rule-Based Reasoning: A Heuristic Approach." *Proceedings of the 11th International Joint Conference on Artificial Intelligence, IJCAI 1989*, Detroit, MI, August 20-25, 1989, San Francisco, CA: Morgan Kaufmann Publishers Inc., 1989. pp. 524-530.

## Thanks to the Contributors

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

### The Classical Ethics in A/IS Committee

- **Jared Bielby** (Chair) – President, Netizen Consulting Ltd; Chair, International Center for Information Ethics; editor, *Information Cultures in the Digital Age*
- **Soraj Hongladarom** (Co-chair) – President at The Philosophy and Religion Society of Thailand
- **Miguel Á. Pérez Álvarez** – Professor of Technology in Education, Colegio de Pedagogía, Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México
- **Oliver Bendel** – Professor of Information Systems, Information Ethics and Machine Ethics, University of Applied Sciences and Arts Northwestern Switzerland FHNW
- **Dr. John T. F. Burgess** – Assistant Professor / Coordinator for Distance Education, School of Library and Information Studies, The University of Alabama
- **Rafael Capurro** – Founder, International Center for Information Ethics
- **Corinne Cath-Speth** – PhD student at Oxford Internet Institute, The University of Oxford, Doctoral student at the Alan Turing Institute, Digital Consultant at ARTICLE 19
- **Dr. Paola Di Maio** – Center for Technology Ethics, ISTCS.org UK and NCKU Taiwan
- **Robert Donaldson** – Independent Computer Scientist, BMRILLC, Hershey, PA

## Classical Ethics in A/IS

- **Rachel Fischer** – Research Officer: African Centre of Excellence for Information Ethics, Information Science Department, University of Pretoria, South Africa.
- **Dr. D. Michael Franklin** – Assistant Professor, Kennesaw State University, Marietta Campus, Marietta, GA
- **Wolfgang Hofkirchner** – Associate Professor, Institute for Design and Technology Assessment, Vienna University of Technology
- **Dr. Tae Wan Kim** – Associate Professor of Business Ethics, Tepper School of Business Carnegie Mellon University
- **Kai Kimppa** – University Research Fellow, Information Systems, Turku School of Economics, University of Turku
- **Sara R. Mattingly-Jordan** – Assistant Professor Center for Public Administration & Policy, Virginia Tech
- **Dr Neil McBride** – Reader in IT Management, School of Computer Science and Informatics, Centre for Computing and Social Responsibility, De Montfort University
- **Bruno Macedo Nathansohn** – Perspectivas Filosóficas em Informação (Perfil-i); Brazilian Institute of Information in Science and Technology (IBICT)
- **Marie-Therese Png** – PhD Student, Oxford Internet Institute, PhD Intern, DeepMind Ethics & Society
- **Derek Poitras** – Independent Consultant, Object Oriented Software Development
- **Samuel T. Segun** – PhD Candidate, Department of Philosophy, University of Johannesburg. Fellow, Philosophy Node of the Centre for Artificial Intelligence Research (CAIR) at the University of Pretoria and Research fellow at the Conversational School of Philosophy (CSP)
- **Dr. Ozlem Ulgen** – Reader in International Law and Ethics, School of Law, Birmingham City University
- **Kristene Unsworth** – Assistant Professor, The College of Computing & Informatics, Drexel University
- **Dr. Xiaowei Wang** – Associate professor of Philosophy, Renmin University of China
- **Dr Sara Wilford** – Senior Lecturer, Research Fellow, School of Computer Science and Informatics, Centre for Computing and Social Responsibility, De Montfort University
- **Pak-Hang Wong** – Research Associate, Department of Informatics, University of Hamburg
- **Bendert Zevenbergen** – Oxford Internet Institute, University of Oxford & Center for Information Technology Policy, Princeton University

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](https://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](#).

### Endnotes

<sup>1</sup> This edition of “Classical Ethics in A/IS” does not (and could not) aspire to universal coverage of all of the world’s traditions in the space available to us. Future editions will touch on several other traditions, including Judaism and Islam.

<sup>2</sup> R. Von Schomberg, “Prospects for Technology Assessment in a Framework of Responsible Research and Innovation” in *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methode*. Wiesbaden, Germany: Springer VS, 2011, pp. 39-61.