Can Ethics Be Standardized? Creating Modern Standards for Ethical Autonomous and Intelligent Systems

By Robert S. Fish, President, IEEE Standards Association

Technologists often react with an eye roll when the subject of ethics comes up. One interpretation of this is that they may think ethics is the purview of philosophy and religion and not really within their domain. An engineer may think of it as a "soft" subject, not the hard science and engineering of designing and building engineering artifacts for which they were trained.

Actually, technology has a long history of interacting with ethics. Many thousands of years ago, the technology of the moment was agriculture, that is, the deliberate planting and harvesting of crops and the domestication of animals that allowed humans to adopt a settled lifestyle. This seems simple enough. However, history suggests that the emergence of agriculture introduced ethical rules about how this technology should be used to address societal concerns other than those of food production. For instance, in an ancient source like the Bible (Leviticus 19:9), farmers are admonished not to harvest the corners of their fields, as these should be left for the poor to harvest. This is an early example of the intersection of a man-made technology (sickle-based grain harvesting) with a requirement for ethical behavior (taking care of the poor).

Today we are presented with new technologies, particularly in the area of Autonomous and Intelligent Systems (A/IS), for which ethical issues take center stage when we consider the deployment and use of these technologies. These issues cannot be avoided. However, we need to ask the following questions:
- What kind of ethical behavior can we reasonably expect from A/IS systems?
- What ethical principles can we use to design those behaviors into the A/IS systems we engineer?

An approach to addressing these questions is to think that we are really trying to create systems that exhibit the same ethical behavior that we would expect from our fellow humans. This is something like the golden rule applied to human-machine interaction, rather than human-to-human interaction.

Of course, safety guidelines have existed for decades to instruct roboticists and programmers on how to prioritize safety concerns in what they build. But, as with the introduction of any new technology, A/IS have introduced new societal issues engineers must account for in the design and proliferation of their work. Specifically, A/IS deeply affect human emotions, agency, and identity (via the sharing of human data) in ways that technology has rarely done before. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems was created to help address key ethical issues like accountability, transparency, and algorithmic bias in A/IS and to help recommend ideas for potential standards based on these technologies.

Inspired by the work being done by The Global Initiative, a series of IEEE standards projects have emerged over the past 18 months. Known as the IEEE P7000™ series of standards projects, the projects under development represent a unique addition to the collection of more than 1300 global IEEE standards and projects. Whereas more traditional standards have a focus on technology interoperability, safety, and trade facilitation, the IEEE P7000 series addresses specific issues at the intersection of technological and ethical considerations. Like their technical standards counterparts, the IEEE P7000 series empower innovation across borders and enable societal benefit. Standards provide a form of "soft governance" that can be utilized for policy as well as technology design and manufacture. This means that where these (or similar) standards are being launched by the engineering or technological community, it is imperative to have engineers and technologists, along with social scientists and philosophers, join the standards working groups. In addition, the IEEE P7000 Standards Working Groups include corporate and policy leaders to help facilitate discussions on how to move forward on these issues with pragmatic, values-design-driven standards.

Today, we see increasing ethical and regulatory challenges in many "emerging" technologies, even though they may have been around for many years and in many forms. Some examples include additive manufacturing methods, artificial intelligence, autonomous systems, and robotics. A difference now is that these technologies are entering into the domain of highly visible consumer goods and services. In the communications domain, we see the emergence of smart networks paired with smart edge systems and terminal devices that both start to make decisions for us as well as collect vast amounts of data about our behaviors and predilections. The potential for abuse of these systems is enormous, and ethical consequences must be considered as a core element in their development.

Autonomous and intelligent systems will continue to evolve, converge, expand, and forge new specializations and applications. In addition to "smart" machines like autonomous vehicles of various sorts, "smart" materials are being integrated into additive manufacturing. Smart materials may have the capacity to change, adapt, interact, and/or respond to environmental conditions. 4-D printing, for example, involves transformation over time and may produce artifacts that have the ability to adapt, self-repair, or even disintegrate. These examples of innovation are engineering marvels; however, there are obvious ethical aspects to each of them.

The standards projects in the IEEE P7000 series will give both engineers and technologists a better understanding and awareness of the issues at stake, the potential for unintended consequences, and the need to design ethical principles into engineering processes, products, and services in communications systems and in other areas. These standards aspire to equip engineers with guidance on addressing the risks and opportunities that accompany these new technologies and attempt to preempt some of the negative impacts on our world. Good governance that evaluates and balances outcomes against risk while including ethical considerations such as sustainability, inequality, human dignity, safety, security, and inclusiveness is needed. And of course, these concerns are very much within the domain of engineering.

The IEEE P7000 Standards Series currently includes the following:
- IEEE P7000™ - Draft Standard for the Model Process for Addressing Ethical Concerns During System Design
- IEEE P7001™ - Draft Standard for Transparency of Autonomous Systems
- IEEE P7002™ - Draft Standard for Data Privacy Process
- IEEE P7003™ - Draft Standard for Algorithmic Bias Considerations
- IEEE P7004™ - Draft Standard on Child and Student Data Governance
- IEEE P7005™ - Draft Standard on Employer Data Governance
- IEEE P7006™ - Draft Standard on Personal Data AI Agent Working Group
- IEEE P7007™ - Draft Ontological Standard for Ethically driven Robotics and Automation Systems
- IEEE P7008™ - Draft Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems
- IEEE P7009™ - Draft Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems
- IEEE P7010™ - Draft Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems
- IEEE P7011™ - Draft Standard for the Process of Identifying and Rating the Trustworthiness of News Sources
- IEEE P7012™ - Draft Standard for Machine Readable Personal Privacy Terms
- IEEE P7013™ - Draft Inclusion and Application Standards for Automated Facial Analysis Technology