A Framework for Ethics in Cyber-Physical-Human Systems [⋆]

Pramod P. Khargonekar [*] **Meera Sampath** [**]

[*] *Department of Electrical Engineering and Computer Science,*
University of California, Irvine, CA 92697
(e-mail: pramod.khargonekar@uci.edu).
[**] *Office of Research and Economic Development,*
State University of New York, Albany, NY,
(e-mail: meera.sampath@suny.edu)

Abstract: This paper proposes a conceptual framework for consideration of ethical issues in the emerging category of smart cyber-physical systems. Cyber-physical systems (CPS) that bring together controls, communications, computing, and physical systems are being developed in a wide variety of application domains ranging from transportation, energy, and manufacturing, to biomedical and agriculture. Smart CPS are already being and will increasingly be deployed to work with humans, in workplaces, homes, or public spaces, resulting in the creation of cyber-physical human systems (CPHS). Ethical issues in smart CPS and CPHS can be examined within the larger frameworks of ethics of technology and ethics of artificial intelligence. We begin with a description of trends and visions for the future development of smart CPS. We next outline fundamental theories of ethics that offer foundations for thinking about ethical issues in smart CPHS. We argue that it is necessary to fight the tendency toward technological determinism. We argue that in analyzing ethics of smart CPHS, we need to anticipate increasing capabilities and the future deployment of such systems. Ultimately, if these systems are widely deployed in society, they will have a very significant impact, including possible negative consequences, on individuals, communities, nations, and the world. Our framework has two main dimensions: (i) stage of development of CPHS domain from early stage research to mature technologies; and (ii) locus of decision making: individual, corporate, and government settings. We illustrate the framework with some specific examples.

Keywords: Cyber-physical-human systems, ethical issues, social impact of automation,

1. INTRODUCTION

In the United States, the National Society of Professional Engineers has a code of ethics for engineers. It states, “Accordingly, the services provided by engineers require honesty, impartiality, fairness, and equity, and must be dedicated to the protection of the public health, safety, and welfare. Engineers must perform under a standard of professional behavior that requires adherence to the highest principles of ethical conduct.” Its canons include:

- Hold paramount the safety, health, and welfare of the public.
- Act for each employer or client as faithful agents or trustees.
- Conduct themselves honorably, responsibly, ethically, and lawfully so as to enhance the honor, reputation, and usefulness of the profession.

In a similar vein, the Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct proclaims, “Computing professionals’ actions change the world. To act responsibly, they should reflect upon the wider impacts of their work, consistently supporting the public good. The ACM Code of Ethics and Professional Conduct (‘the Code’) expresses the conscience of the profession.”

This paper explores ethical issues that emerge in the setting of cyber-physical-human systems (CPHS). The ongoing integration of controls, communications, and computing into various physical systems with sensors and actuators combined with increasing levels of automation enabled by machine learning and artificial intelligence is creating smart cyber-physical systems (CPS). The increasing deployment of such systems in open, dynamic environments is transforming how humans interact with smart CPS. Transportation, manufacturing, energy, medicine, and agriculture are all going through this transformative change as one can observe in the development of smart and connected cars, cobots on the factory floor, robotic surgery, precision agriculture, and so on. It is only likely that these types of changes will continue and quite possibly accelerate in the coming years and decades.

With such transformative CPHS innovations that affect individuals and society, we should expect that we will face important ethical issues. In this paper, we present a framework for examining current and possible future

---

[⋆] Sponsor and financial support acknowledgment goes here. Paper titles should be written in uppercase and lowercase letters, not all uppercase.

ethical issues in smart CPHS. The proposed framework has two major dimensions — (i) stage of development of the CPHS technological domain and (ii) individual and organizational setting for decision making. The framework is proposed for articulating and anticipating ethical and moral considerations that go beyond regulatory and legal requirements as well as traditional engineering standards and guidelines.

A few authors have examined ethics specifically in the context of cyber-physical systems. In particular, we mention the work of Thekkilakattil and Dodig-Crnkovic (2015), in which the authors develop a framework for responsibility attribution, an important ethical issue. There is a considerably larger literature on the ethics of safety-critical systems. Ethics of technology and artificial intelligence provide rich background for our discussion. These are discussed in greater detail in Section 3.

In the final analysis, individuals, be they technologists, scientists, engineers, or policy makers, make choices and decisions that have ethical implications. Our goal is for this paper to provide readers with a useful background, analysis and decision framework for examining ethical issues and making good decisions in accordance with their normative ethical values. It is to be hoped that such decisions will result in enhanced individual lives and ethical use of CPS in human society.

2. CYBER-PHYSICAL-HUMAN SYSTEMS: TRENDS AND VISIONS

Cyber-physical systems (CPS) bring together control, computing, communications, networks, sensors, and actuators with physical systems to create integrated technological systems [Baheti and Gill (2011)]. The CPS concept allows fundamental control systems knowledge to be leveraged and integrated into modern technological systems. Applications of CPS range from manufacturing, energy, transportation, aerospace, and defense, to biomedical and healthcare, and beyond. For example, autonomous and connected vehicles on smart roads and networks are creating the transportation systems of tomorrow. Industry 4.0 and smart manufacturing are being made possible by CPS. Integration of renewable electricity from solar and wind generation is a key objective for smart electric grids [Annaswamy and Amin (2013)] enabled by CPS. Quality, efficiency, patient outcomes, and cost of health care systems can be positively impacted by incorporating CPS in diagnostics, therapeutics, monitoring, home health, and telemedicine applications, see Chen et al. (2018). For a very comprehensive forward looking vision of the future of control and cyber-physical systems, we refer the reader to Lamnabhi-Lagarrigue et al. (2017).

In almost all cases, cyber-physical systems involve deep interactions with humans at various levels, as individuals, as members of various organizations and as communities, as inhabitants of cities, states, and nations, and society as a whole. In recent years, the term cyber-physical-human systems (CPHS) has been used to capture the totality of CPS interacting with and embedded in human society [IFAC (2018)].

As we look forward, it is most probable that breakthrough progress in data science, machine learning (ML) and artificial intelligence (AI) will be integrated into CPHS to create what we might call ‘smart CPHS’. While initial applications of AI amd ML are in e-commerce, information processing, and computer vision, CPHS are natural targets for applications of ML and AI [Bughin et al. (2018)]. Indeed, smart electric grids, smart manufacturing, smart health are already major topics of research and development, and we can expect broader commercial development of smart CPHS in the coming years and decades.

As the complexity of CPHS as well as their prevalence in industry and society increase, [see Törngren and Grogan (2018)] for a detailed discussion of increasing complexity of CPHS)], a wide range of ethical and moral issues will come to the fore. These issues will span across safety and security, transparency, bias, and fairness arising from integration of AI/ML, human rights issues, and jobs. They will also include potential loss of autonomy and empowerment with increased levels of automation, broader problems of equitable access to technology, and socio-economic considerations. To the best of our knowledge, there are few comprehensive studies of the ethics of cyber-physical systems, and certainly not to the level of discourse and debate as in the AI/ML domains.

These considerations motivate the central question for this paper: What conceptual frameworks can be used to consider ethical issues as we research, develop, commercialize, and deploy smart CPHS?

The scope of our considerations extends through all levels of technology development and maturity, from basic research to commercial deployment of products and services. Therefore, it is important to clarify the terminology. All CPHS-based products and services result from suitably designed combinations of sensors, communications, controls, computing (hardware and software), and other generic enabling technologies with application domain technologies such as transportation, aviation, manufacturing. Thus, a modern aircraft is an integrated system (of systems) designed, manufactured, and operated for providing air transportation services. Much, although not all, of CPHS research focuses on component technologies such as controls, communications, or computing as relevant to the final application domain need or opportunity. However, ethical issues become most pressing and visible when humans interact with the product or service forming CPHS.

3. BROADER CONTEXT: ETHICS OF TECHNOLOGY AND AI

In this section, we provide a high-level overview of key concepts, frameworks, and approaches related to the study and practice of ethics in technology. We start with a brief discussion about the notion of technological determinism, which we believe is core to any treatment of ethics in technology. We then discuss approaches to examining ethical issues in the special case of emerging technologies, followed by a look at ethics in the specific context of AI and autonomous systems.

3.1 *Technological Determinism*

Technological determinism is the idea that technologies evolve exogenously, either solely due to scientific advances or following an autonomous development path of their own, and then impact society. While still very ingrained in popular thinking and practice, this notion is increasingly contested by researchers and thinkers in the field of science and technology studies. MacKenzie and Wajcman (1999) argue why technological determinism is an inherently flawed concept, an oversimplification, or at best a "partial truth." Technology development in any given domain is driven not only by scientific and engineering advances, but also by various other social factors, including other existing technologies, anticipated future costs and profits, state sponsorship, and the very process of usage and adoption by society. Likewise, social, political, economic, and even other, unseen technological outcomes, are driven by choices made by individuals and organizations, consciously or otherwise.

Noble (1978) writes, "there is always a range of (technology) possibilities or alternatives that are delimited over time — as some are selected and others denied — by the social choices of those with the power to choose, choices that reflect their intentions, ideology, social position, and relations with people in society," . . . "social impacts issue not so much from the technology of production as from the social choices that technology embodies." In a similar vein, Winner (2010) cautions that technologies can be designed, consciously, or otherwise, to "open certain social options and close others"; "some technologies are, in given social circumstances, more compatible with some social relations than with others."

A key message that MacKenzie and Wajcman (1999) and others drive is that predicting the future consequences of technology choices may not be easy, but the "difficulty of the task is not reason for avoiding it." A passive attitude toward technology development that accepts the deterministic view focuses our mind on how to adapt to it not how to shape it, and thereby robs us of the opportunity to influence the evolution of these technological changes. This understanding is a necessary first step in exploring ethics of technology.

3.2 *Ethics and Emerging Technologies*

Next we look at ethics in the context of emerging, "revolutionary" technologies. We discuss two papers on this topic, the first one by Moor (2005), and the second by Brey (2012). Moor argues that, in this era of rapid technological progress and the convergence of many technologies of potential transformational social impact, ethics-as-usual will not work. Based on a tripartite maturity model for revolutionary technologies that is comprised of an introductory stage, a permeation stage, and a power stage - each of which is characterized by progressively increasing availability and standardization of the technology, increasing number of users, decreasing costs, and most importantly increasing social impact, Moor posits that as technological revolutions mature and increase their social impact, so will ethical problems. It is therefore necessary, at each stage of technology development, starting from the early introductory phase and through the power stage, to continually anticipate how technology could be used and what consequences may arise, and to proactively develop ethical guidelines and policies. Such ethical analysis should not come as an afterthought, and certainly not in the later stages of technology maturity, when the complexity of the ethical issues and their ramifications increases significantly, and where lack of adequate policies may leave one confronted with "policy vacuums."

Brey builds on Moor's work and proposes an approach that he calls anticipatory technology ethics (ATE), focusing specifically on the R&D (research and development) phase of new technology development. Ethics in the R&D phase is different from that of the introductory phase (which is further downstream) because R&D focuses on creation of basic techniques and methods that may eventually result in concrete applications, whereas early applications already exist in the introductory phase. Ethical analysis, in the R&D phase, is largely speculative and hence subject to various challenges, including the need to engage in forecasting (which is central to the ATE approach) without becoming too speculative and needing to integrate technology forecasting with normative ethical analysis. ATE breaks down the problem into analysis at three levels: technology (collection of core techniques), artifact (that arises out of the technology), and application (use of the artifact for a particular purpose). At each level, the objects of analysis are first identified by applying various methods of forecasting. Ethical analysis is then determined by identifying and examining moral values and issues associated with each of these objects, with the ultimate goal to guide technology design or to inform policymaking.

3.3 *Ethics in ML and AI*

Automation, machine learning and AI are likely enablers of next-generation smart CPHS. Ethics has become a central topic in AI and there are many papers and volumes - e.g., Winfield et al. (2019), Winfield and Jirotka (2018), in this field. We focus here on the recent report "Ethically Aligned Design (EAD)," published by the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (A/IS) [IEEE (2019)]. The conceptual framework for ethical analysis proposed in this report is based on three pillars that capture the anthropological, political, and technical aspects of ethics and design: (i) universal human values, (ii) political self-determination and data agency, and (iii) technical dependability. These three pillars form the basis of eight general principles that are considered as imperatives for the ethical design of A/IS: human rights, well-being, data agency, effectiveness, transparency, accountability, awareness of misuse, and competence. Various ethical issues under each of these eight topics are examined in depth. Detailed recommendations on how to address these issues are provided with the overall goal of helping A/IS creators reduce to practice relevant principles in the context of their own specific product or service.

In a chapter specifically devoted to classical ethics methodologies, EAD provides several examples of how these age-old and established traditions can guide today's A/IS creators in addressing ethical issues and dilemmas. For instance, Aristotle's virtue ethics that emphasizes the goal of “eudaimonia,” or human flourishing, provides a framework to balance against excesses, a common tendency in economically motivated environments. Kant’s duty-based deontological ethics would guide us to build intelligent systems that respect humanity and human dignity, such as for example, by preserving privacy of personal information, and may require us to limit functions and capabilities of A/IS, in order to achieve these goals. Utilitarian ethics, or consequentialist ethics, which espouses maximizing utility for the greatest number of people, would guide A/IS developers to consider the negative impact of automation technologies on employment and design systems that supplement human capabilities and benefit all stakeholders.

4. ETHICAL FRAMEWORK FOR CPHS

We propose a two-dimensional ethical framework to examine ethical considerations in CPHS research, design, development, deployment, and commercialization. The first dimension considers the stage of development of the CPHS technological domain. The second dimension considers the individual, organizational, and government contexts in which decisions are made. Figure 1 illustrates our proposed framework. We believe that an ethical issue in current or future CPHS can be effectively examined by considering it along these two dimensions. This analysis can assist individuals, organizations, and society to make individual and collective decisions on important CPHS ethical issues.

The reader may find our illustrative examples very diverse and unconnected. This is intentional. We wish to demonstrate the generality of the proposed framework by showcasing its applicability in a wide variety of CPHS application domains.

4.1 Ethical Considerations and the Stage of Development

Like many other socio-technological systems, CPHS products and services go through stages of evolution and maturity. These range from early-stage research, prototypes, and small-scale development to large-scale commercial and societal deployment.

Mature CPHS Technological Domains: For mature CPHS technological domains, ethical considerations are often the clearest and also have the largest consequences. Examples of such domains include incumbent aerospace, automotive, manufacturing, agriculture, transportation, and energy technologies. (Although we designate these as mature, we do recognize that there are continuing advances in one or more of the component technologies.) From a systems viewpoint, ethical issues in mature CPHS arise from from considerations of safety, health, environment, security, etc. Ethical guidelines and protocols are generally well-defined and the role of individuals, organizations and governments becomes one of ensuring ongoing compliance, vigilance, transparency, and refinement of policies.

Let us consider modern aviation, a major exemplar of CPHS. This technology has matured over the last century into wide societal deployment. Indeed, one can argue that it has shaped the development of current human civilization. Modern aircrafts increasingly incorporate advanced communications, controls, sensing, and computing technologies. In this domain, the ethical considerations include passenger safety from accidents, transparency in the certification processes, and ownership of responsibility in decision making across all levels of the organization and the employee base, among others.

A particular contemporary and enlightening example is the safety issues in the Boeing 737 Max. Full consideration of ethical issues in this case would involve the role of various people at Boeing, at the Federal Aviation Agency (FAA), and policymakers in the Congress. Here we focus on the role of engineers at Boeing. We quote from a recent article in The New York Times (dated October 2, 2019) by Kitroeff et al. (2019):

> *A senior Boeing engineer filed an internal ethics complaint this year saying that during the development of the 737 Max jet the company had rejected a safety system to minimize costs, equipment that he felt could have reduced risks that contributed to two fatal crashes . . . The engineer who filed the ethics concerns this year, Curtis Ewbank, went a step further, lodging a formal complaint and calling out the chief executive for publicly misrepresenting the safety of the plane. . . . According to Mr. Ewbank’s complaint, Ray Craig, a chief test pilot of the 737, and other engineers wanted to study the possibility of adding the synthetic airspeed system to the Max. But a Boeing executive decided not to look into the matter because of its potential cost and effect on training requirements for pilots.*

It is clear that robust mechanisms and frameworks, such as safety and robustness checklists for engineers and mechanisms for reporting and escalating issues up the management chain, were and are necessary for employees and leaders to systematically and comprehensively consider ethical implications of their decisions, . It is also a *teachable moment* to other mature industries and companies.

Developing CPHS Technological Domains: In this category, we can include CPHS that have not yet reached broad societal deployment but are significantly beyond research and prototyping phases. Examples include: unmanned civilian aircraft and drones, smart and connected vehicles, smart electric grids, precision agriculture, and wearable sensors. In each case, there are companies and government organizations that are developing systems with hopes of large-scale societal deployment. In these scenarios, ethical considerations become most compelling when examined in the framework of anticipatory technology ethics developed in the work of Moor and Brey. More specifically, the questions for such developing CPHS technological domains become: What will be the social, environmental and economic impacts and what could be unintended consequences of large-scale, power stage deployment if the technology under development is extremely successful? What issues of human rights violations might arise? How do we ensure shared prosperity and inclusion? Who bears accountability and responsibility for ethical violations? How could the given developing technology become part of other existing or developing technologies and what would the resulting impacts be? What actions and precautions should be taken by individuals and organizations? What are the guidelines and regulations to ensure that the final outcomes are in accord with their ethical values?

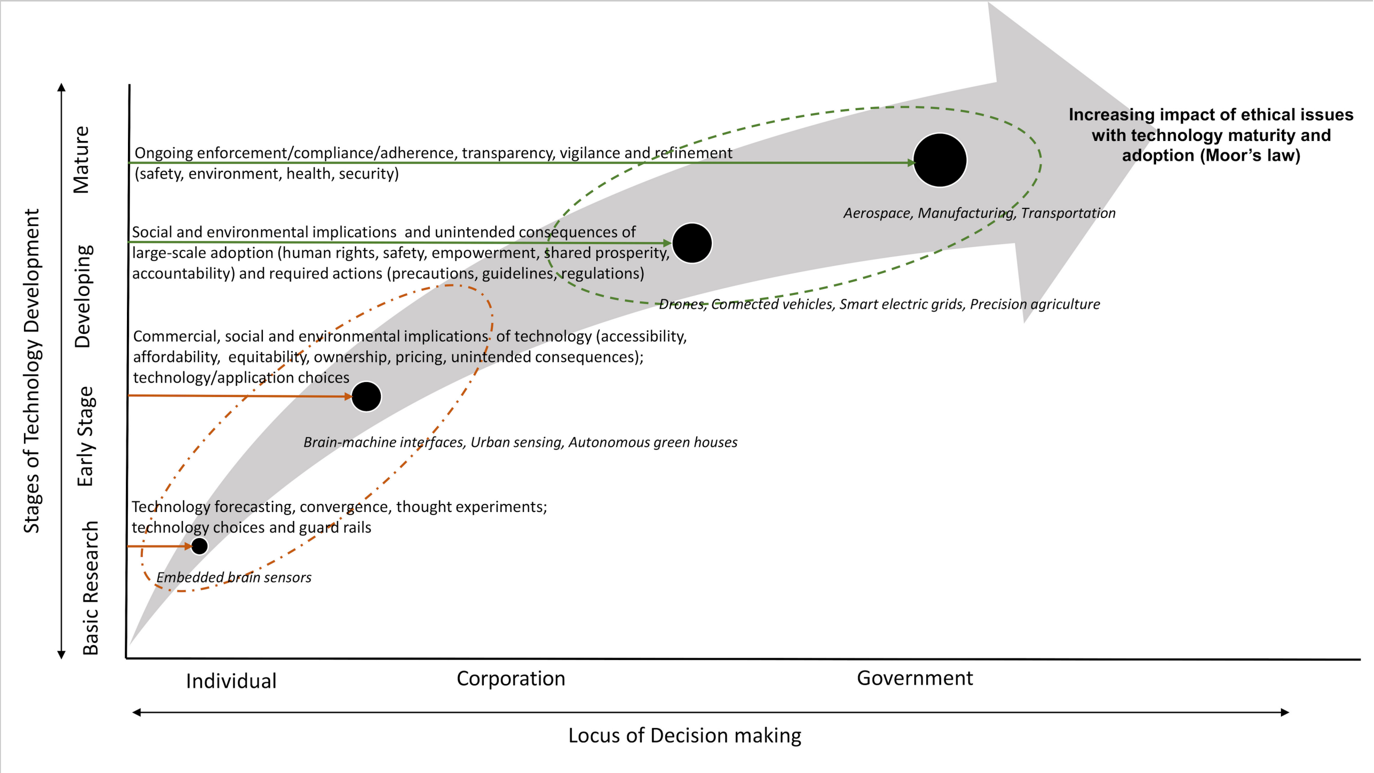

Fig. 1. Framework for Ethics in CPHS

As an illustration, let us consider the extremely important field of autonomous and connected cars (see, for example, Shladover (2018)). This is clearly a developing technology that is beyond prototypes. Lower levels of autonomy have been realized, but full level 5 autonomy and large-scale deployment are still in the future. By now, there are several papers that discuss a wide variety of ethical issues in this domain (see for example, McBride (2016); Goodall (2016); Luetge (2017)). Ethical issues that range from safety and responsibility, to displacement of human drivers will gain in prominence and importance as self-driving cars become more ubiquitously adopted and integrated into society. As we anticipate this future, and as this technology achieves the power stage of deployment, many questions arise. One can get a glimpse of this future already from the following questions in the report on ethics of automated and connected cars by the Ethics Commission of the Federal Ministry of Transport and Digital Infrastructure in Germany (see Di Fabio et al. (2017)). "At the fundamental level, it all comes down to the following question. How much dependence on technologically complex systems — which in the future will be based on artificial intelligence, possibly with machine learning capabilities — are we willing to accept in order to achieve, in return, more safety, mobility and convenience? What precautions need to be taken to ensure controllability, transparency and data autonomy? What technological development guidelines are required to ensure that we do not blur the contours of a human society that places individuals, their freedom of development, their physical and intellectual integrity and their entitlement to social respect at the heart of its legal regime?"

Early-stage CPHS Domains: In this category, we include CPHS concepts that are beyond exploratory fundamental research with prototypes but not yet sufficiently developed for real-world, commercial-scale deployment. Typically, these would be at lower levels of technology readiness. Some examples to illustrate this category include: brain-machine interfaces [Courtine et al. (2013)], urban sensing [O'Keeffe et al. (2019)], autonomous greenhouses [Ko et al. (2014)], etc.

In case of this class of technologies, there is tremendous potential benefit in identifying and addressing ethical issues before the technology becomes cost-effective and commercially attractive leading to wide scale adoption. Ethical considerations may include issues of uniform accessibility, affordability, equitability, ownership, pricing of technology, as well as unintended consequences. Key actions driven by such ethical considerations may include specific technology and application choices made by individual researchers, technologists, or R&D organizations.

As an illustration, consider the case of autonomous greenhouse technologies. If successful, these technologies have the potential to reduce agricultural resource use and costs, and increase productivity. Assuming great technological success, there might be issues of access to these technologies in developing countries and/or to poorer farmers. Would intellectual property protections compete with wide global access? One can imagine parallels with access to expensive pharmaceuticals. In another dimension of the problem, how would the critically important weather data, generated by public investments in weather prediction systems, be priced? What are the key principles for ownership of such critical data? In yet another dimension, would the posited environmental benefits accrue at very high penetration levels of such farming practices?

Basic Research in CPHS: In this category, we include fundamental research that is very far from commercial development or even prototypes. Typically, much of academic research would fall in this group. Also, in case of CPHS, these would be in the core disciplines such as controls, communications, machine learning, algorithms, and security. Ethical issues in this category are the most difficult to identify and understand. The approach based on anticipatory technological forecasting is not easy to apply. Nevertheless, we propose a few ideas that may be helpful. As was mentioned earlier in this paper, most CPHS applications involve combinations of technologies. So consider a basic research project in a particular field, e.g., control systems engineering. Now imagine this research project is completely successful. What if the results were to be combined with existing or future technologies from other complementary fields in some application domain? Would there be new ethical issues as a result of such a combination? Careful consideration of such questions and thought experiments may lead to technology choices by individuals and establishment of policies by organizations and governments.

To illustrate the above approach, consider the case of nano-sensors and nano-actuators with advanced communications technologies. Suppose that such sensors and actuators can be injected into the human brain (see, Hong et al. (2018)). Could such devices be manipulated from outside to give control over a human being to some other human being? Could such devices compromise human autonomy and self-control? As another thought experiment, consider the possibility that advanced signal processing and control combined with powerful non-invasive remote sensors and actuators might enable external manipulation of human emotions and thoughts. What are the ethical responsibilities of researchers working on such research projects? Should certain research projects or research goals in smart CPHS be deemed to carry so much potential for harm that they be considered unethical? How should the public participate in the governance of such research projects and programs? All these questions require serious ethical analysis.

4.2 Ethical Considerations for Individuals, Corporations, and Government

In the ultimate analysis, each individual makes decisions that have ethical dimensions. While individuals bear the ultimate responsibility and authority over their decisions, they are often members of groups. A person may be a member of a family, an employee of an organization, a member of a social group, and a resident of a state or nation. For example, a female controls engineer may be a mother, a professor at a university, a member of an environmental nonprofit organization, and a citizen of a certain developed nation. As another example, consider a male computer engineer, CEO of a start-up robotics company with potential customers in the agricultural sector in many nations across the world. As a third example, consider a senior engineer who is an elected member in government with the role and responsibility of crafting legislation. How might the ethical considerations of CPHS be taken into account in decisions made by these individuals?

Individual Ethical Frames: Individuals play a key role in driving ethical decisions and outcomes in all four stages of technology maturity discussed above. An individual, as a researcher, technologist, business leader, or policymaker may adopt one or more ethical frames for weighing their choices and decisions as they consider various alternatives and their ethical implications. Classical ethical frames, such as those discussed in Section 3, include: virtue ethics, moral character, duty-based deontological ethics, and consequentialist ethics. Depending on their role, one or more of these frames might be applicable regardless of whether the individual's decisions focus on human-centric, technical, or political aspects. For example, an engineer working on a self-driving car may consider issues of safety in its algorithms and systems. She might decide that it is her duty that she should insist on developing rigorous safety protocols. And that before the product is approved for sale, it undergoes rigorous testing, meets and exceeds all the applicable standards, etc. On the other hand, an engineer developing next-generation unmanned vehicles might consider the societal consequences of their misuse for surveillance or terrorism. A researcher working on smart cities might decide that it is virtuous to focus their services on the disabled or underprivileged members of the city. On the other hand, she may ensure that the sensors and communications contain switches to disable their misuse for illegal or unethical surveillance of the citizens. University professors of CPHS may decide that it is their duty to include discussions and debates on ethical issues in their classrooms or research groups (see, e.g., Meckl (2003)). Regardless of an individual's role, the issue under study and the framework they choose to adopt, it is key for "those with the power to choose" to remember that their decisions will shape future individual and societal outcomes.

Ethics for Corporations: Corporations involved in developing and commercializing CPHS products and services will need to play a critical role in proactively formulating and addressing ethical issues, given their role as the primary entities responsible for technology adoption and success. The essential systems aspect of CPHS calls for the collaboration of numerous individuals within and external to a corporation, including suppliers and vendors, in taking a CPHS product or service to the market. This raises critical issues of principles and processes for ethical behavior where knowledge, activities, and responsibilities are broadly distributed and interdependent. Most corporations have a code of ethics, although one can find a wide range of views and empirical data on the actual effectiveness of business codes of ethics [Kaptein and Schwartz (2008)]. At a minimum, employees and stakeholders should not only be aware of the relevant codes but also be encouraged to actively adopt them in their day-to-day work. The example of Boeing 737 Max is compelling in this regard. The Boeing Code of Conduct "establishes behavioral expectations for Boeing employees at all levels of the company wherever they are in the world. Along with Boeing's Enduring Values, the code serves as the foundation for our workplace culture . . . ". Integrity, quality, and

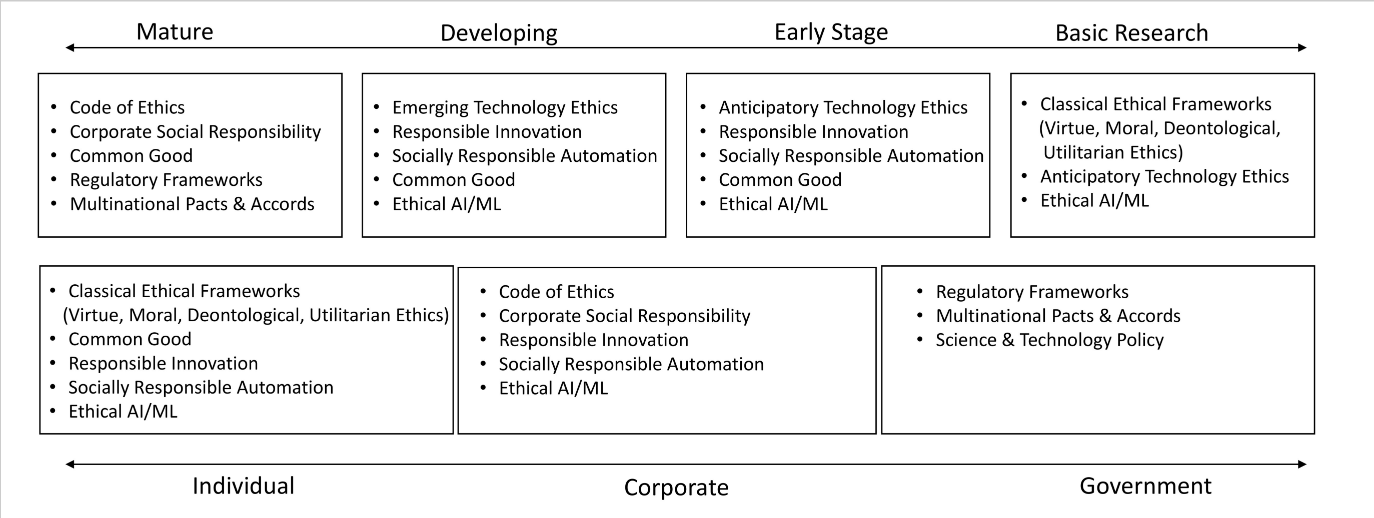

Fig. 2. Summary of Framework for Ethics in CPHS

safety are among its Enduring Values. We quote from the New York Times article [Kitroeff et al. (2019)]: “Boeing’s chief executive, Dennis A. Muilenburg, said in a speech on Wednesday that “it is critical we take a step back to humbly look at our culture.””

As recommended by IEEE (2019), corporations may identify checkpoints along the product lifecycle from early research and development to early customer adoption and testing to large-scale deployment where “ethical” filters are in place. It may be especially important to have these checkpoints when products move from one development team to another to preempt ethical considerations from getting lost in the transition. EAD also recommends the establishment of an ethics review board, cultivation of ethics leaders, clear direction from leadership to innovation and engineering teams regarding which values and norms are to be promoted, and empowerment of staff to raise and voice ethical concerns. Finally, the tight integration of sociologists and ethicists in design and engineering teams, as well as routine and meaningful stakeholder engagement are critical to ensure that ethical considerations remain at the forefront throughout the development process.

Ethics for Governments: The principal role for government entities is in creating and enforcing forward-looking policies and legislation. It is generally believed that in rapidly advancing technologies, legislation lags behind technological capabilities; i.e., the role of government in driving ethical technology development is most observed in the case of mature and to some extent, developing technologies. A contemporary relevant example, although not specific to CPHS, is the EU General Data Protection Regulation (GDPR). From the developing CPHS domains, the case of autonomous cars presents a current example. In the United States, a bipartisan group of senators are working to pass legislation for federal regulations on self-driving cars [Miller (2019)]. However, there are serious disagreements from elected leaders as well as consumer groups. The main disagreements surround consumer safety and cybersecurity. There are also disagreements on local and regional effects of federal laws. The CPHS research community has an important role to play in these policy discussions so that the enacted policies and regulations lead toward desirable social outcomes and induce ethical behavior. A policy maker (with or without CPHS background) may consider it her ethical duty to proactively engage the relevant R&D community as well as affected social groups in the process of developing legislative proposals.

Examples of ethical frameworks that one can draw upon at each stage of technology maturity and at each level of decision making are depicted in Figure 2.

5. COMMON GOOD AND RESPONSIBLE RESEARCH AND INNOVATION

The “common good” principle [Velasquez et al. (1992)] is a powerful paradigm that is especially useful in the context of ethical dilemmas and their resolution. With roots in the writings of philosophers such as Plato, Aristotle, and Cicero, a contemporary definition of common good comes from the political and moral philosopher Rawls (2009): “maintaining conditions and achieving objectives that are similarly to everyone’s advantage.” Applying this principle would not only help technologists to consider the values that may be supported or compromised by the choices they make, but would also force them to formulate and articulate the rationale for their decisions, which is key for transparency in systems design. Finally, we recently proposed the “socially responsible automation” (SRA) framework [Sampath and Khargonekar (2018)] in the specific context of automation. SRA as well as the broader but closely related concepts of responsible research and innovation [Jirotka et al. (2016); van den Hoven (2013)] will be potentially useful in examining several of the above stages of development of CPHS and may serve as useful frameworks for guiding many CPHS applications in business contexts.

6. CONCLUSIONS

Human beings develop their values and expectations regarding ethical behavior through their family, community, education, religious, and work experiences. While the fundamental canons of ethics go back thousands of years in

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

various societies and civilizations, and have a strong basis in the pro-social nature of human beings, accelerating socio-technological changes in a globally connected world create novel situations that require us to be much more agile and forward-looking. Smart CPHS will be significant drivers of these changes. We hope that the framework proposed in this paper will provide a basis for discussions and development of ethics of CPHS and will be useful to the members of the CPHS engineering community in discharging our ethical responsibilities in our various roles.

REFERENCES

A. M. Annaswamy and M. Amin. *IEEE Vision for Smart Grid Controls: 2030 and Beyond*. IEEE, 2013.

R. Baheti and H. Gill. Cyber-physical systems. *The Impact of Control Technology*, 12(1):161–166, 2011.

P. A. E. Brey. Anticipating ethical issues in emerging IT. *Ethics and Information Technology*, 14(4):305–317, 2012.

J. Bughin, J. Seong, J. Manyika, M. Chui, and R. Joshi. Notes from the AI frontier: Modeling the impact of AI on the world economy. *McKinsey Global Institute, Brussels, San Francisco, Shanghai, Stockholm*, 2018.

J. Chen, M. Chen, J. Qu, H. Chen, and J. Ding. Smart and connected health projects: Characteristics and research challenges. In *International Conference on Smart Health*, pages 154–164. Springer, 2018.

G. Courtine, S. Micera, J. DiGiovanna, and J. del R Millán. Brain–machine interface: closer to therapeutic reality? *The Lancet*, 381(9866):515–517, 2013.

U. Di Fabio, M. Broy, R. J. Brüngger, et al. Ethics commission: automated and connected driving. *Federal Ministry of Transport and Digital Infrastructure of the Federal Republic of Germany*, 2017.

N. J. Goodall. Can you program ethics into a self-driving car? *IEEE Spectrum*, 53(6):28–58, 2016.

G. Hong, X. Yang, T. Zhou, and C. M. Lieber. Mesh electronics: a new paradigm for tissue-like brain probes. *Current opinion in neurobiology*, 50:33–41, 2018.

IEEE. *Ethically Aligned Design*. IEEE Press, 2019.

IFAC. Second IFAC Conference on Cyber-Physical & Human Systems, `http://www.cphs2018.org/frontiers.php`, 2018.

M. Jirotka, B. Grimpe, B. Stahl, G. Eden, and M. Hartswood. Responsible research and innovation in the digital age. *Communications of the ACM*, 2016.

M. Kaptein and M. S. Schwartz. The effectiveness of business codes: A critical examination of existing studies and the development of an integrated research model. *Journal of Business Ethics*, 77(2):111–127, 2008.

N. Kitroeff, D. Gelles, and J. Nicas. Boeing 737 max safety system was vetoed, engineer says. *New York Times*, 2019. `https://nyti.ms/2nLjFda`.

M. H. Ko, B.-S. Ryuh, K. Chul Kim, A. Suprem, and N. P. Mahalik. Autonomous greenhouse mobile robot driving strategies from system integration perspective: Review and application. *IEEE/ASME Transactions On Mechatronics*, 20(4):1705–1716, 2014.

F. Lamnabhi-Lagarrigue, A. Annaswamy, S. Engell, A. Isaksson, P. Khargonekar, R. M. Murray, H. Nijmeijer, T. Samad, D. Tilbury, and P. Van den Hof. Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control*, 43:1–64, 2017.

C. Luetge. The German ethics code for automated and connected driving. *Philosophy & Technology*, 30(4):547–558, 2017.

D. MacKenzie and J. Wajcman. *The Social Shaping of Technology*. Open University Press, 1999.

N. McBride. The ethics of driverless cars. *ACM SIGCAS Computers and Society*, 45(3):179–184, 2016.

P. H. Meckl. Integrating ethics into an undergraduate control systems course. In *Proceedings of the ASEE Annual Conference*, volume 8, pages 8.736.1–8.736.7. ASEE, 2003.

M. Miller. New push to regulate self-driving cars faces tough road. *The Hill*, 2019. `https://bit.ly/2IZF5u7`.

J. H. Moor. Why we need better ethics for emerging technologies. *Ethics and information technology*, 7(3): 111–119, 2005.

D. F. Noble. Social choice in machine design: The case of automatically controlled machine tools, and a challenge for labor. *Politics & Society*, 8(3-4):313–347, 1978.

K. P. O'Keeffe, A. Anjomshoaa, S. H. Strogatz, P. Santi, and C. Ratti. Quantifying the sensing power of vehicle fleets. *Proceedings of the National Academy of Sciences*, 116(26):12752–12757, 2019.

J. Rawls. *A Theory of Justice*. Harvard University Press, 2009.

M. Sampath and P. P. Khargonekar. Socially responsible automation. *Bridge*, 48(4):45–72, 2018.

S. E. Shladover. Connected and automated vehicle systems: Introduction and overview. *Journal of Intelligent Transportation Systems*, 22(3):190–200, 2018.

A. Thekkilakattil and G. Dodig-Crnkovic. Ethics aspects of embedded and cyber-physical systems. In *2015 IEEE 39th Annual Computer Software and Applications Conference*, volume 2, pages 39–44. IEEE, 2015.

M. Törngren and P. Grogan. How to deal with the complexity of future cyber-physical systems? *Designs*, 2(4):40, 2018.

J. van den Hoven. Value sensitive design and responsible innovation. *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, pages 75–83, 2013.

M. Velasquez, C. Andre, T. Shanks, and M. J. Meyer. The common good. *Issues in Ethics*, 5(2):45–60, 1992.

A. F. Winfield, K. Michael, J. Pitt, and V. Evers. Machine ethics: the design and governance of ethical AI and autonomous systems. *Proceedings of the IEEE*, 107(3): 509–517, 2019.

A. F. T. Winfield and M. Jirotka. Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133):20180085, 2018.

L. Winner. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. University of Chicago Press, 2010.

40 pt
0.556 in
14.1 mm

40 pt
0.556 in
14.1 mm

68 pt
0.944 in
24 mm