

Teaching AI, Ethics, Law and Policy

Asher Wilk[†]
Tel Aviv University
Tel Aviv Israel

ABSTRACT

The cyberspace and the development of intelligent systems using Artificial Intelligence (AI) created new challenges to computer professionals, data scientists, regulators and policy makers. For example, self-driving cars raise new technical, ethical, legal and policy issues. This paper proposes a course **Computers, Ethics, Law, and Public Policy**, and suggests a curriculum for such a course. This paper presents ethical, legal, and public policy issues relevant to building and using software and artificial intelligence. It describes ethical principles and values relevant to AI systems.

KEYWORDS

Artificial Intelligence, Ethics, Law, Regulation, Public Policy, Machine Learning, Autonomous Systems, Education, Privacy, Security, Computers and Society.

1 Introduction

The Web, Internet and new technologies, such as intelligent systems and cognitive robots using Artificial Intelligence (AI), help humanity but they also affect human behaviors, induce social and cultural changes, and create new moral and legal problems related to safety, privacy and human dignity. Therefore, computer professionals and decision makers need a better understanding of ethical and legal issues. People perform moral evaluations regularly. Studying ethics can improve our decision-making and “improve how we live” [1]. By being moral, we enrich our lives and the lives of those around us. I suggest teaching ethics and law using examples and case studies, and to demonstrate ethical and legal decision-making. Making an ethical decision can be a complex task and involves knowing the relevant field (profession) and considering other fields such as law, psychology, sociology, and politics.

This paper is based on my experiences teaching ethics and law to computer science students and to students studying security. The structure of this paper is as follows. **Section 2** describes the promise and danger of AI. **Section 3** on “Autonomous Systems” addresses ethical and legal problems of autonomous cars and weapons. It deals with issues relevant to algorithmic decision-making. **Section 4** is about “Social Media, Fake News and

Journalism”. **Section 5** deals with Big Data and Privacy. **Section 6** describes justifications and initiatives for ethics and legal education. **Section 7** on “Law: What to Teach?” suggests what legal subjects to teach, in relation to ethics, such as human rights, privacy and freedom of speech. **Section 8** on “Ethics: What to Teach?” suggests teaching ethical theories, code of ethics, and ethical dilemmas. **Section 9** discusses regulation. **Section 10** deals with programming ethics and laws. **Section 11** concludes with a summary.

2 AI: The Promise and the Danger

For many years, the field of AI had limited success. It took many years to develop good expert systems, but they had limited capabilities and addressed specific areas. However, since 2012 dramatic changes occurred because of advancements in Machine Learning (ML) and in particular Deep Learning (DL). The developments in ML allowed rapid progress in a variety of domains including speech recognition and image classification. Stanford researchers recently created a facial recognition analysis program that can detect sexual orientation [9]. The start-up Faception says its ML technology can spot character traits in a person's face. Such applications raise ethical and legal questions. Engineers, scientists, and policy makers should be aware of societal implications of computer applications.

AlphaGo Zero is the first computer program to defeat a world champion at the ancient Chinese game of Go. It learns to play simply by playing games against itself, starting from a completely random play [47]. Scientists had thought it would take many more years to get to that stage. Scholars are worried how to ensure that computers do not acquire a mind of their own, threatening human existence. It is unknown how quickly we can develop human level Machine Intelligence. Scholars are afraid that together infotech and biotech will have major impacts on liberal democracies [42].

3 Autonomous Systems

Machines equipped with AI now make decisions without direct human input. Soon AI may be able to control machines such as automobiles and weapons. Autonomous weapons and cars will affect human safety, privacy, human dignity, and autonomy. They raise new legal and ethical challenges. For instance, who will be held accountable when fully autonomous systems produce unwanted results.

[†] The views and opinions expressed here are those of the author. Version: 24 May 2019. © 2019 Copyright held by the author. Email: wilk.asher@post.idc.ac.il

3.1 Autonomous Weapon Systems

The military interest is to increase the autonomy of weapon systems in order to have greater military capability, minimize risks, and reduce costs. Weapon systems with significant autonomy already exist; however, they are not capable of understanding complex and dynamic situations, and perform complex decision-making performed by humans.

There are different views on whether in the future “fully autonomous weapon systems” with AI will be in use. There is recognition of the importance of maintaining human control although it is not so clear what would constitute ‘meaningful human control’. The UK policy is that “operation of weapon systems will always be under human control” [2]. Many countries have not yet developed their policy on this or have not discussed it openly. When discussing the ethical and legal problems of using autonomous weapons, international humanitarian law (IHL) must be addressed. Presently programming a machine to make qualitative judgments requires applying “the IHL rules of distinction, proportionality and precautions in attack” [2].

Two central principles from the Just War tradition are the principle of discrimination (distinction) of military objectives and combatants from non-combatants, and the principle of proportionality of means, where acts of war should not yield damage disproportionate to the ends that justifies their use. There are views that current technology is not capable of acting with proportionality and discrimination, and it is unlikely to be possible in the near future. As AI in military robots advances, the meaning of warfare may change and the existing international laws might not be suitable and new laws will be needed.

AI researchers called for a ban on offensive autonomous weapons beyond meaningful human control [14]. A campaign to stop killer robots argued that robots and AI lack moral capability and decisions of significant moral weight ought to remain humans [18]. According to present laws, “states, military commanders, manufacturers and programmers may be held accountable for unlawful ‘acts’ of autonomous weapon systems under a number of distinct legal regimes: state responsibility for violations of IHL and international human rights law; international criminal law; manufacturers or product liability; and corporate criminal liability” [2]. Developing autonomous weapon system raises questions of accountability, responsibility, compliance with IHL rules, ethical challenges, and the human right to dignity.

3.2 Autonomous Vehicles (AV)

Self-driving cars will reduce fatal road accidents and the number of cars and parking lots. It will have many impacts on society, economy, mobility, the environment, and more. For instance, positive economical impacts of AV are “the reduction in the transportation costs and the reduction in the value of travel time” [38]. “Society will only take advantage of self-driving vehicles if they behave cooperatively ... both vehicle-to-infrastructure and

vehicle-to-vehicle communications will play an essential role” [38]. Self-driving cars will have to deal with potentially life-threatening situations and make moral decisions. Such decisions can be described as an application of the trolley problem. The trolley problem can be regarded as an abstraction of many dilemmas involving AI systems. I recommend teaching the trolley problem while looking at the MIT Media Lab Moral Machine experiments.

Since automated driving will not be able to avoid all accidents on roads, there is a need to devise laws and rules for what to do in various cases. Presently, liability in the case of an accident is a controversial topic, in which ethics and economic interests play a role. According to the U.S. National Highway Traffic Safety Administration, states should consider how to allocate liability among owners, operators, passengers, manufacturers, and others when a crash occurs. According to the German Ethics Commission for automated and connected driving, the protection of individuals takes precedence over all other utilitarian considerations. “The licensing of automated systems is not justifiable unless it promises to produce at least a diminution in harm compared with human driving” [11]. Anytime a driverless car makes a decision, it has to make a trade-off between safety and usefulness. Car manufacturers expect regulators to determine what reasonable decision-making is. Ethics and regulation will determine when fully AVs come to the market. Data management is another important issue that requires an ethical consideration. In particular, handling cyber attacks and protecting user privacy.

3.3 Algorithmic Decision Making

Automated decision-making algorithms are now in use, and they have the potential for significant societal impact. For instance, COMPAS was used to decide whether defendants awaiting trial are too dangerous to be released on bail. “The investigative news organization ProPublica claimed that COMPAS is biased against black defendants” [22]. The company that created the tool released a report questioning ProPublica’s analysis. Scholars reported algorithmic discrimination and the need for improved transparency and algorithmic fairness. The IEEE Standards Association launched the IEEE Global Initiative on Ethics for Autonomous and Intelligence Systems [58] aimed at improving fairness of algorithmic decision-making systems.

Scholars suggested introducing ethical considerations into algorithms, opening algorithmic black boxes [5, 39], regulating algorithms, and **Fairness, Accountability, and Transparency (FAT)** in algorithmic decision-making. Making code and data public may not help a user to understand an algorithm and its effects. Scholars suggested to monitor the outputs of codes to ensure they do not discriminate or cause harm. Another suggestion is to ensure that algorithmic decisions can be explained. According to [13] “without being able to explain decisions taken by autonomous systems, it is difficult to justify them: it would seem inconceivable to accept what cannot be

justified in areas as crucial to the life of an individual as access to credit, employment, accommodation, justice and health”. Hinton claims that explainable AI is problematic. “People have no idea how they do that ... Neural nets have a similar problem”, and you should regulate AI algorithms based on how they perform [32].

Algorithmic fairness is very important, but to define fairness and to design fair algorithms is difficult. Fairness can be a data problem. Learning algorithms that will learn from biased data will create biased results. Data should be unbiased, but in some cases, it is complicated to determine if data is biased.

Responsibility is very important, because in spite of our best efforts, things will go wrong at times. Who should be responsible for decisions made by an algorithm? Who will be responsible to take any necessary actions? Developers should take steps to check and validate the accuracy of the algorithm and the data it uses.

4 Social Media, Fake News and Journalism

Web sites already influence us. From a cognitive psychological perspective, media affect our minds. Targeted social media advertising based on user profiling has emerged as an effective way of reaching individuals. In the case of political advertising, this may present a democratic and ethical challenge, and policy and regulation questions. “Social media are systematically exploited to manipulate and alter public opinion” [21]. According to [24] “There is worldwide concern over false news and the possibility that it can influence political, economic, and social well-being” and “Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information”. To ensure that false content is not amplified across platforms, scholars claim that there is a need for some “regular auditing of what the platforms are doing” [26] and “a new system of safeguards” [26].

Nowadays, news organizations such as AP, Reuters, and others are generating thousands of automated stories a month, a phenomenon called “robot journalism”. Presently most uses of robot journalism have been for formulaic situations: stock market summaries and sports stories. Fully automated journalism is going to be very limited for quite a time. However, news companies will be using automatic news writing more and more on challenging subjects. This raises new ethical issues, for instance, the accuracy of the data, and the legal rights to the data. Will the fact that a story was automatically produced be disclosed? Will a human editor check every story before it goes out? Can we program ethical concerns into robot journalism algorithms?

5 Big Data and Privacy

Corporations and governments are collecting, analyzing, and sharing detailed information about individuals over long periods. Novel methods for big data analysis are yielding deeper understandings of individuals’ characteristics, behavior, and relationships. These developments can advance science, benefit

individuals and our society, but they are also “privacy risks that multiply as large quantities of personal data are collected over longer periods of time” [25]. ML operating on big data produces results that may violate the privacy of individuals, and can have discriminatory outcomes. “Existing regulatory requirements and privacy practices in common use are not sufficient to address the risks associated with long-term, large-scale data activities” [25]. Notice and consent or deidentification are not enough [25].

5.1 Internet of Things (IoT) and Ethics

With the expected increase of using the Internet of Things (IoT) more and more data will be collected. According to [19] AI based on big data and combined with IoT might eventually govern core functions of society and it is necessary to apply the principles of **rule of law, democracy and human rights** in AI. With IoT, trust is a major issue. When everything is connected to the Internet, it is hard to trust that information is only shared with those declared. Scholars suggested that people should know about their rights, and that they should be able to control their information and accesses. With IoT, “Informed consent, privacy, information security, physical safety, and trust is foundational” [18].

5.2 Privacy: What to Teach?

The Computer Science Curricula 2013 (CS2013) [20] recommends teaching philosophical foundations of privacy rights, legal foundations of privacy protection, privacy implications of wide spread data collection and technology-based solutions for privacy protection [20, 12]. I taught the approach of Warren and Brandeis [44] on privacy as a “right to be let alone”, and the approach of Alan Furman Westin seeing privacy as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” [45]. I pointed out constitutions and bills of rights addressing privacy, for instance, the European Convention on Human Rights (ECHR).

There are many laws dealing with privacy, and different states understand privacy differently. The EU General Data Protection Regulation (GDPR) [10] is probably the best privacy regulation, but it is quite complex. I suggest discussing mainly principles and concepts that are in the GDPR, such as: (1) Lawfulness, fairness and transparency (2) Purpose limitations (3) Data minimization (4) Accuracy (5) Storage limitation (6) Integrity and confidentiality. GDPR includes important rights for data subjects that should be addressed, e.g., the right to be forgotten, the right to object to profiling, the right to data portability, and the right to be informed about the collection and use of personal data. I recommend addressing Privacy by Design (PbD) [3]. The idea is to integrate technical privacy principles in system design.

6 Ethics and Law Education

The curricula of many engineering schools includes humanities and social science subjects to broaden the scope of education of

students. Many firms realized that a company has to be concerned with **corporate social responsibility (CSR)** and business schools are teaching Business Ethics and CSR. New technologies, present new ethical, legal, societal, and policy issues. Students who will become computer professionals or policy makers should be able to practice their profession in a way that integrates legal and ethical concerns into their work. Recommendations of ethics and law studies to computer students appear in various publications [12, 15, 20]. According to the Computer Science Curricula 2013 (CS2013) “Graduates should recognize the social, legal, ethical, and cultural issues inherent in the discipline of computing” [20].

6.1 Why to Teach Ethics?

Following are justifications to teach ethics: (1) Some professions, such as medicine and the law, have traditional codes of ethics, and ethics is part of the professional education. This should be the case for AI and computer professionals. (2) “Being ethical is simply good business” [55]. (3) Employers are not looking only for technical skills but they prefer to hire persons with good soft skills that include work ethics and responsibility. (4) “Ethics has become a major consideration for young people in their selection of work and career” [55]. (5) Many cases of unethical behavior suggest that ethics education is important. (6) Law develops slower than new technologies and may not provide solutions to ethical dilemmas. (7) Acts can be legal, but not appropriate from an ethical point of view (8) Ethics provides tools helping to make better decisions for the benefit of a person and society.

Academic institutions are launching new courses on computing ethics and cyberethics. According to Tavani [40] cyberethics is the study of moral, legal, and social issues involving cyber technology. Cyber technology comprises computing and communication systems. Harvard and MIT are offering a course on AI ethics. Some are integrating ethics across their computing curricula. “The idea is to train the next generation of technologists and policymakers to consider the ramifications of innovations - like autonomous weapons or self-driving cars - before those products go on sale” [16]. The course “The Ethics and Governance of Artificial Intelligence” [17] includes concepts such as: Governance, Explainability, Accountability, Transparency, Discrimination, Fairness, Algorithmic Bias, Autonomy, Agency, and Liability, which are important to teach. According to WIPO [34] legal and ethical concerns linked to AI “include the transparency, verifiability and accountability of AI, the right to privacy, the right to equal treatment and avoidance of bias, and the mitigation of negative impacts on employment”.

Recently scholars advocate that ethics be part of the data science curriculum [46, 59]. Data science ethics should address the quality of data and evidence, data accuracy and validity. Data scientists need to address security, privacy and anonymity of data.

6.2 Why to Teach Law?

My view is that basic legal knowledge is important for many reasons, including: (1) Legal cases provide good examples on

ethical decision-making. (2) It helps to understand the limitations of law and the need to act responsibly. (3) Basic legal education enables computer professionals to interact with law professionals and regulators. (4) Certain areas of law are relevant to computer professionals (e.g., privacy and intellectual property). (5) To understand the relationship between law and ethics.

7 Law: What to Teach?

For students to engage fully in the study of any subject they need to find it interesting. There are various alternatives of what subjects of law to teach. Security professionals might be more interested in cyber war and crime, while computer professionals might be more interested in privacy and intellectual property (IP). CS2013 [20] suggests teaching Intellectual Property, Privacy and Civil Liberties, and Professional Ethics. The Computer Engineering Curricula 2016 [15] recommends teaching computer engineers contract law, intellectual property, privacy and secrecy. Quinn in “Ethics for the Information Age” [28] besides dealing with ethics has chapters on Intellectual Property, Information Privacy, Privacy and the Government, and Computer Reliability. Spinello in [27] analyzes moral dilemmas and social problems arising in cyberspace, and looks at content control, free speech, intellectual property, privacy and security, and social networking.

In courses I taught on ethics and law, I choose: (1) Teaching law by case examples and choosing examples that will interest the students; (2) Concentrate on principles (3) Teach legal subjects that relate to ethics, such as subjects related to human rights and civil liberties, privacy, the need to balance rights (e.g., freedom of speech vs. defamation) and proportionality; (4) Use legal case examples that refer to ethical dilemmas; and (5) Introduce students to law fundamentals, which include the difficulties in legislation, problems of law interpretation, and problems that arise from law being local while cyber and technology are global.

8 Ethics: What to Teach?

Nowadays ethics or moral philosophy is being concerned mainly with what is morally good or bad, right or wrong. Ethical rules and principles differ from one culture or society to another, and have changed from one generation to the next. Therefore, there is place to address various schools and philosophies of ethics. Having a basic understanding of the major ethical theories will help us in decision-making. In the western philosophy, the following approaches are well known, and I recommend addressing them: (1) **Ethics of consequences** including utilitarianism proposed by Jeremy Bentham and John Stuart Mill [51], act utilitarianism and rule utilitarianism. (2) **Ethics of Duty (Deontological ethics)**, including formulations by Immanuel Kant [50]. (3) **Virtue Ethics** focuses on virtues (behaviors) that will allow a person to achieve well being. Most known is the ethics of Aristotle [52] that asks what is the character or personality of an ethical person. (4) **Justice** as in ideas of John Rawls [53], and **Social Contract Ethical Perspectives**. (5) **Principlism**, that focuses on the moral principles of autonomy, beneficence,

nonmaleficence, and justice [54], and nowadays provides an ethical framework applicable to medical practices.

Burton [23] advocate that ethics should be part of the AI curricula, and that it is necessary to be familiar with deontological ethics, utilitarianism and virtue ethics. There are ethical dilemmas that do not have optimal solutions and different ethical theories will lead to different solutions. “Consider, for example, a utilitarian doctor who harvests one healthy (but lonely and unhappy) person’s organs to save five other people. Those five people could go on to experience and create more happiness combined than that one person ever could alone. Although this action is morally problematic, utilitarianism would seem to justify it if it best maximizes utility” [48]. However, there are ethical theories that forbid taking someone’s life.

It is important to look at ethical codes of organizations, firms, and professions. They provide practical guidelines. Ethic codes are important, but to be effective, training, practice, and examples are needed. Lafollette [1] theorizes that ethics is like most everything else that we strive to be good at; it requires practice and effort. A code of ethics of a profession needs to include a description of the profession, principles, and rules. When teaching either ethics or law, it is important to point out the difference between principles and rules while demonstrating them.

8.1 Ethics and Law

Ethics tells us how to behave in the absence of directions in the law. The relationship between law and ethics should be addressed. It is a controversial issue whether morality (ethics) is the ultimate foundation of law. After World War II, there were claims that there can be a conflict between the obligation to obey the law and the obligation to act morally. Should one consider as “law” all the rules that were formally valid by the Nazi legal system, regardless of their amoral content or not. In a democratic society where laws are generally taken to be legitimate, something being legal usually is also moral.

8.2 Values, Principles, and Rules

Rules provide solutions to very particular situations, while principles are more general and they set standards. For instance, the principle of equal access to justice is part of many legal systems, and it expressed by particular rules.

An ethical decision process involves going from a set of cultural norms, values, and beliefs to principles and from principles to rules, policies and procedures. For example, as part of the human dignity value, conclude the principle of transparency, stating that entities should always take steps to be transparent about their use of data. This principle can be translated into rules and procedures. Principles and rules are used to form a code of ethics. An effective legal or ethical system must build also on principles and cannot be constructed entirely on a rulebook that provides answers to ethical dilemmas or to legal questions. Professional codes of ethics, for instance, the ACM and IEEE, include both rules and principles. I suggest discussing the four principles of healthcare ethics: Autonomy, Beneficence, Nonmaleficence, and Justice [54].

Autonomy is about the right to decide what happen or does not happen to us. It is about an informed consent, which is fundamental in bioethics [49]. Informed consent appears in privacy laws such as the GDPR [10]. Justice is about fairness and equality. Beneficence is about “Doing good” and acting in the best interests of others. Nonmaleficence is about “Do no harm”, which is an important principle in the ACM code of ethics. Each profession has particular values and norms. For example, military ethics emphasizes human dignity, loyalty to the country, respecting the rule of law, courage, responsibility, reliability, friendship, dedication, discipline, and personal example [56].

In the late 1940s, Norbert Wiener created the field of “computer ethics”. He viewed “The Principle of Freedom”, “The Principle of Equality” and “The Principle of Benevolence” as “great principles” upon which society should be built [43]. The World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) of UNESCO proposed a technology-based ethical framework on robotics ethics [6], which includes principles and values of: (1) human dignity; (2) autonomy; (3) privacy; (4) “Do not harm”; (5) responsibility; (6) beneficence; and (7) justice. Microsoft outlined ten principles to approach AI [7] which clearly state that AI must prevent bias. Google’s CEO Sundar Pichai [8] has outlined principles that will guide Google. Among them are: avoid creating or reinforcing unfair bias, be built and tested for safety, be accountable to people, and incorporate privacy design principles. Yi Zeng [29] Lists 27 proposals of AI principles. The proposals address concepts, such as, dignity, human rights, equality, fairness, justice, bias, discrimination, transparency, privacy, data protection, security, safety, validation, accountability and responsibility. Such concepts should be addressed in an ethics and law course.

8.3 Teaching Strategies

Ethics does not always provide a right answer to moral problems. For many ethical issues there is not a single right answer. Teaching ethics should include discussing ethical dilemmas and how to make ethical decisions. Burton [23] suggests teaching by presenting case studies, discussing ethical issues and analyzing them using utilitarianism, deontological, and virtue theories. A methodology similar to the one used by the National Association of Social Workers [30] can be used: (1) Determine the ethical dilemma. (2) Identify the key values and principles involved. (3) Rank the values or principles that are most relevant to the dilemma. (4) Develop an action plan that is consistent with the ethical priorities that have been determined as central to the dilemma. (5) Implement your plan, utilizing the most appropriate practice skills. (6) Evaluate the consequences for those involved.

Another strategy that can be used is Brey’s method [41, 40]: (1) Identify a practice or a technological feature that is controversial from a moral perspective. Determine whether there are any specific guidelines (professional codes) that can help resolve the issue. (2) Analyze the ethical issue by clarifying concepts and

situating it in a context. (3) Deliberate on the ethical issue by applying one or more ethical theories and justifying the position you reached by using logic and critical thinking. Brey [41] suggests addressing four key values as starting points for studies in computer ethics: justice, autonomy, democracy and privacy.

9 Ethics and Regulation

Although many firms have ethical codes, there are cases where firms acted not only unethically but also illegally. The recent case of Volkswagen regarding compliance with emission control standards highlights the critical importance of oversight, an ethical culture, and a compliance system. In the fallout of Enron, laws like Sarbanes Oxley were formed for the purpose of protecting the public and the business from fraud or errors. The solution to reducing car accidents was not ethics training for drivers, but public policy and laws addressing the safety of cars, the legal duty to put on seatbelts, the safety of roads, licensing of drivers, drunk-driving laws, and the like. Targeted advertising on the Internet is very profitable and it is unlikely companies will abandon this business model because of ethical reasons. The solution is in regulations such as the GDPR. The Information Technology (IT) industry has successfully lobbied for decades against any attempt to regulate IT claiming, “regulation stifles innovation.” Regulation that uses strict rules can impede innovation, might limit technology, and interfere with the development process. However, regulation is sometimes necessary, especially to ensure safety. According to AI now 2018 [5] “There is an obvious need for accountability and oversight in the industry, and so far the move toward ethics is not meeting this need. This is likely in part due to the market-driven incentives”.

Arguments that are now being presented against legislation for AI have been presented against legislation for data protection (GDPR) [19]. Those arguments include that the law is not able to develop as fast as technology, the law is not precise enough to regulate complex technology, and that the law is not providing sufficient legal certainty now and is not sufficiently open to provide flexibility for the future. Scholars claim that such arguments are a way of saying what corporations have always said: we want no obligations by law as with laws we could be held accountable through enforcement. Business has no problem with the fact that any ethics code lacks democratic legitimacy and cannot be enforced. The law has democratic legitimacy and it can be enforced, even against powerful mega corporations.

Creating new laws sometimes takes many years because of difficulties to understand the problems, priorities, and political disagreements. Regulating intelligent autonomous systems is also difficult because (1) it is difficult to know how future intelligent systems will work and (2) technology is global, while regulation is usually local. There is a need for global regulation of AI, but past efforts at global regulation of other matters indicate that it is

very difficult. Since regulation has its problems, it is also important to educate people to be responsible and behave ethically.

10 Programming Ethics and Laws

A major obstacle to automating ethical decisions is the disagreement on ethical principles among moral philosophers [33]. Human ethical decisions are affected by emotions, social upbringing, maturity, gut instinct, and philosophical worldviews [36] and this is impossible to reduce to fixed codes. Such ability is built up by years of parenting, socialization, and involvement in cultures, societies, and communities.

Some scholars suggested that autonomous systems should only be deployed in situations where there is a consensus on the relevant ethical issues [57]. According to [35], when ground truth ethical principles are not available, we must use “an approximation as agreed upon by society”. For self-driving cars, surveys such as the Moral Machine can help to come up with a social consensus. How self-driving cars should be programmed to handle situations in which harm is likely either to passengers or to others outside the car? A possible strategy, following the ideas in [4], is to use technical ways to avoid a dilemma while obeying legal constraints, and if not possible then deciding based on learning from preferences of voters (e.g., using preference data collected from millions of people through the Moral Machine website).

It is questionable if ML can arrive at ethical or legal rules. This is a too difficult task for current ML technology. In addition, if we want legal or ethical rules to be used by machines, humans still want to control the deciding of what the rules should be. In the near future, most likely autonomous system will use rules created by people or rules controlled and verified by people. These rules will be specific for particular situations.

Intel developed a model called Responsibility-Sensitive-Safety (RSS) [31]. RSS is designed to achieve three goals: (1) the interpretation of the law should be sound in the sense that it complies with how humans interpret the law (2) the interpretation should lead to a useful driving policy (3) the interpretation should be efficiently verifiable [31]. RSS is constructed by formalizing the following five “common sense” rules: (1) Do not hit someone from behind. (2) Do not cut-in recklessly. (3) Right-of-way is given, not taken. (4) Be careful of areas with limited visibility. (5) If you can avoid an accident without causing another one, you must do it. Each of these general rules need to be programmed for different scenarios, involving mathematical computations. Consider the principle “Do not hit someone from behind”. What if by hitting someone from behind you can save your life and the lives of four passengers while killing one person in front of you.

It was suggested to include ethical modules in software. They should be transparent and “it should be possible to verify that it respects the values it reasons about, particularly ... where safety is

concerned” [37]. What presently can be done is devising specific rules for specific situations using limited information and techniques involving calculations, probabilities, and formal logic. This is far from ethical decisions by humans. The decisions made by algorithms will have to be tested and verified using techniques such as simulation, testing and formal verification, taking into consideration many possible real world scenarios.

11 Summary

This paper explained the importance of teaching ethics and law fundamentals to computer professionals and decision makers that includes legal aspects, ethical aspects, and professional responsibility. Future jobs for computing graduates will require not only technical knowledge, but also ethical and legal awareness. We described relevant topics and challenges computer professionals and decision makers should address. New technologies will require developing ethics and law.

REFERENCES

- [1] H. LaFollette. 2007. *The Practice of Ethics*. Blackwell Publishing.
- [2] Autonomous weapon systems: Technical, military, legal and humanitarian aspects. Expert meeting, Geneva, Switzerland, 26-28 March 2014.
- [3] A. Cavoukian, 7 Foundational Principles, <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>
- [4] R. Noothigattu, N.S. Gaikwad, E. Awad, S. Dsouza, I. Rahwan, P. Ravikumar, and A.D. Procaccia, A Voting-Based System for Ethical Decision Making.
- [5] AI now reports, 2017, 2018. AI Now Institute, New York University. <https://ainowinstitute.org/>
- [6] UNESCO, “Report of World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) on Robotics Ethics” SHS/YES/COMEST-10/17/2 REV. Paris, 2017.
- [7] A. P. Chaudhury, “Microsoft CEO Satya Nadella rolls out 10 rules that will define Artificial Intelligence”. <https://analyticsindiamag.com/microsoft-ceo-satya-nadella-rolls-10-rules-will-define-artificial-intelligence/>
- [8] S. Pichai, “AI at Google: our principles”. <https://blog.google/technology/ai/ai-principles> Published Jun 7, 2018
- [9] Heather Murphy, Why Stanford Researchers Tried to Create a ‘Gaydar’ Machine, The New York Times, Oct. 9, 2017.
- [10] European Union. 2018. The EU General Data Protection Regulation. <https://eugdpr.org/the-regulation/>
- [11] Federal Ministry of Transport and Digital Infrastructure - Ethics Commission. 2017. Automated and Connected Driving. Jun. 2017.
- [12] Asher Wilk. 2016. Cyber Security Education and Law. 2016 *IEEE International Conference on Software Science, Technology and Engineering*. SWSTE 2016.
- [13] C. Villani. 2018. For a Meaningful Artificial Intelligence. Towards a French and European strategy.
- [14] <https://spectrum.ieee.org/robotics/military-robots/do-we-want-robot-warriors-to-decide-who-lives-or-dies>
- [15] Computer Engineering Curricula 2016, Curriculum Guidelines for Undergraduate Degree Programs in Computer Engineering (Interim Report), ACM and IEEE Computer Society, Oct. 2015.
- [16] N. Singer. 2018. Tech’s Ethical ‘Dark Side’: Harvard, Stanford and Others Want to Address It. *The New York Times*. Feb. 12, 2018.
- [17] Joi Ito, Jonathan Zittrain, The Ethics and Governance of Artificial Intelligence, Syllabus, MIT, Spring 2018.
- [18] F. Allhoff, A. Henschke, The Internet of Things: Foundational ethical issues.
- [19] Paul Nemitz. 2018. Constitutional democracy and technology in the age of artificial intelligence, <https://doi.org/10.1098/rsta.2018.0089>
- [20] Computer Science Curricula 2013, Curriculum Guidelines for Undergraduate Degree Programs in Computer Science, The Joint Task Force on Computing Curricula, ACM, IEEE Computer Society, Dec. 2013.
- [21] E. Ferrara. 2017. Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election.
- [22] Avi Feller, Emma Pierson, Sam Corbett-Davies and Sharad Goel, “A computer program used for bail and sentencing decisions was labeled biased against blacks. It’s actually not that clear”, *The Washington post*, October 17, 2016.
- [23] E. Burton, J. Goldsmith, S. Koenig, B. Kuipers, N. Mattei, T. Walsh. Ethical Considerations in Artificial Intelligence Courses. [arXiv:1701.07769v1](https://arxiv.org/abs/1701.07769v1).
- [24] Soroush Vosoughi, Deb Roy and Sinan Aral, The spread of true and false news online. *Science* 359 (6380), 1146-1151.
- [25] M. Altman, A. Wood, D. O’Brien, and U. Gasser. 2018. *Practical Approaches to Big Data Privacy Over Time*. Oxford University Press.
- [26] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, et al. 2018. The science of fake news. *Science*. 9 Mar 2018, Vol. 359, Issue 6380, pp. 1094-1096.
- [27] R. Spinello. 2016. *Cyberethics: Morality and Law in Cyberspace*. 6th edition.
- [28] Michael J. Quinn. *Ethics for the Information Age*. 7th Edition.
- [29] Yi Zeng, Enmeng Lu, Cunqing Huangfu. 2019. Linking Artificial Intelligence Principles. AAAI-Safe AI 2019. <https://arxiv.org/abs/1812.04814>
- [30] <https://www.socialworkers.org/pubs/code/>
- [31] S. Shalev-Shwartz, S. Shammah, A. Shashua. 2017. On a Formal Model of Safe and Scalable Self-driving Cars. <https://arxiv.org/pdf/1708.06374.pdf>
- [32] Hesse Jones, “Geoff Hinton Dismissed The Need For Explainable AI: 8 Experts Explain Why He’s Wrong”, *Forbes*, Dec 20, 2018.
- [33] B. Williams. *Ethics and the Limits of Philosophy*. Harvard Univ. Press, 1986.
- [34] WIPO Technology Trends 2019: Artificial Intelligence. Geneva: World Intellectual Property Organization.
- [35] C. Dwork, M. Hardt, T. Pitassi, et al. 2012. Fairness through awareness. In *Innovations in Theoretical Computer Science Conference*, pages 214-226.
- [36] Neil McBride, Robert R. Hoffman. Bridging the Ethical Gap: From Human Principles to Robot Instructions. *IEEE Intelligent Systems*, 2016.
- [37] P. Bremner, L.A. Dennis, M. Fisher, and A.F. Winfield. On Proactive, Transparent, and Verifiable Ethical Reasoning for Robots, *Proc. of the IEEE*.
- [38] Margarita Martínez-Díaz, Francesc Soriguera, Ignacio Pérez. 2019. Autonomous driving: a bird’s eye view. *Proc. of the IEEE. IET Intell. Transp. Syst.*, Vol. 13 Iss. 4, pp. 563-579.
- [39] F. Pasquale. 2016. *The black box society: the secret algorithms that control money and information*. Cambridge, MA: Harvard University Press. 320 pages.
- [40] H. Tavani. 2016. *Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing*. 5th Edition. Wiley, 2016.
- [41] P. Brey. 2000. Disclosive computer ethics. *ACM SIGCAS Computers and Society*, Volume 30 Issue 4, December 2000, Pages 10-16.
- [42] Y.N. Harari. 2018. *21 Lessons for the 21st Century*. 1st edition. Spiegel & Grau.
- [43] N. Wiener. N. 1950. *The Human Use of Human Beings: Cybernetics and Society*, Boston: Houghton Mifflin, Second Edition Revised, New York, NY: Doubleday Anchor 1954.
- [44] S.D. Warren, L.D. Brandeis. The Right to Privacy. *Harvard Law Review*, Vol. 4, No. 5, pages 193-220, Dec. 15, 1890.
- [45] Alan F. Westin. 1967. *Privacy And Freedom*.
- [46] J.S. Saltz, N.I. Dewar, and R. Heckman. 2018. Key Concepts for a Data Science Ethics Curriculum. In *ACM Technical Symposium on Computer Science Education*, Feb. 21–24, 2018, Baltimore, MD, USA.
- [47] D. Silver et al. Mastering the game of Go without human knowledge, *Nature* 550, 354–359 (19 October 2017).
- [48] Christopher Grau. 2016. There Is No “I” in “Robot”: Robots and Utilitarianism. *IEEE Intelligent System*. July/August 2006, pp. 52-55, vol. 21.
- [49] UNESCO. 2006. Universal Declaration on Bioethics and Human Rights. <http://unesdoc.unesco.org/images/0014/001461/146180E.pdf>
- [50] Immanuel Kant. 1785. Fundamental Principles of the Metaphysic of Morals. <https://www.gutenberg.org/ebooks/5682>
- [51] <https://www.utilitarianism.com/>
- [52] Aristotle. 350 B.C.E. *Nicomachean Ethics*. Translated by W. D. Ross. <http://classics.mit.edu/Aristotle/nicomachaen.html>
- [53] J. Rawls. 1999. *A Theory of Justice*. Revised edition (1st edition 1971).
- [54] T.L. Beauchamp, J.G. Childress. *Principles of Biomedical Ethics*. Oxford University Press.
- [55] <https://opentextbc.ca/businessethicsopenstax/>
- [56] CF Code of Values and Ethics. <http://www.forces.gc.ca/en/about/code-of-values-and-ethics.page>
- [57] M. Anderson and S.L. 2011. *Machine Ethics*. New York: Cambridge University Press.
- [58] <https://ethicsinaction.ieee.org/>
- [59] <https://www.coursera.org/learn/data-science-ethics>