Artificial Intelligence and Robotics Sources

No.	Title	Author/Source	Date	Abstract/Summary	Keywords
Artificial Intelligence					
AI1	Machine learning: the power and promise of computers that learn by example	Royal Society	04/25/17	Machine learning is a branch of artificial intelligence that allows computer systems to learn directly from examples, data, and experience. Through enabling computers to perform specific tasks intelligently, machine learning systems can carry out complex processes by learning from data, rather than following pre-programmed rules. Link: https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf	Machine Learning, Big Data
AI2	Big data, artificial intelligence, machine learning and data protection	Information Commissioner's Office	03/01/17	Big data, artificial intelligence (AI) and machine learning are becoming widespread in the public and private sectors. Data is being collected from an increasing variety of sources and the analytics being applied are more and more complex. While many benefits flow from these types of processing operations, when personal data is involved there are implications for privacy and data protection. In our view though, these implications are not barriers. There are several tools and approaches that not only assist with data protection compliance but also encourage creativity, innovation, and help to ensure data quality. So it's not big data *or* data protection, it's big data **and** data protection. The benefits of big data, AI and machine learning will be sustained by upholding key data protection principles and safeguards. Link: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf	Big Data, Machine Learning, Data Protection, Compliance
AI3	The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term	Kate Crawford and Meredith Whittaker (Co-Chairs) AI Now	09/22/16	The *AI Now 2016 Report* provides an overview of the four focus areas, summarizes key insights that emerged from discussions at the Symposium, and offers high-level recommendations for stakeholders engaged in the production, use, governance, and assessment of AI in the near-term.	Healthcare, Social Inequality, Labor, Ethics

				Link: https://ainowinstitute.org/AI_Now_2016_Report.pdf	
AI4	The National Artificial Intelligence Research and Development Strategic Plan	National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee	10/2016	This *National Artificial Intelligence R&D Strategic Plan* establishes a set of objectives for Federally-funded AI research, both research occurring within the government as well as Federally-funded research occurring outside of government, such as in academia. The ultimate goal of this research is to produce new AI knowledge and technologies that provide a range of positive benefits to society, while minimizing the negative impacts. To achieve this goal, this AI R&D Strategic Plan identifies the following priorities for Federally-funded AI research: ● Strategy 1: Make long-term investments in AI research. ● Strategy 2: Develop effective methods for human-AI collaboration. ● Strategy 3: Understand and address the ethical, legal, and societal implications of AI. ● Strategy 4: Ensure the safety and security of AI systems. ● Strategy 5: Develop shared public datasets and environments for AI training and testing. ● Strategy 6: Measure and evaluate AI technologies through standards and benchmarks. ● Strategy 7: Better understand the national AI R&D workforce needs. Link: https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf	Research & Development
AI5	Preparing for the Future of Artificial Intelligence	Executive Office of the President, National Science and Technology Council Committee on Technology	10/2016	As a contribution toward preparing the United States for a future in which Artificial Intelligence (AI) plays a growing role, we survey the current state of AI, its existing and potential applications, and the questions that are raised for society and public policy by progress in AI. We also make recommendations for specific further actions by Federal agencies and other actors. A companion document called the National Artificial Intelligence Research and Development Strategic Plan lays out a strategic plan for Federally-funded research and development in AI.	Regulation, Research & development, Fairness, Governance, International Cooperation

				Link: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf	
AI6	Accountable Algorithms	Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu 165 U. PA. L. REV. 633 (2017).	2017	We challenge the dominant position in the legal literature that transparency will solve these problems. Disclosure of source code is often neither necessary (because of alternative techniques from computer science) nor sufficient (because of the issues analyzing code) to demonstrate the fairness of a process. Furthermore, transparency may be undesirable, such as when it discloses private information or permits tax cheats or terrorists to game the systems determining audits or security screening. The central issue is how to assure the interests of citizens, and society as a whole, in making these processes more accountable. This Article argues that technology is creating new opportunities—subtler and more flexible than total transparency—to design decision-making algorithms so that they better align with legal and policy objectives. Doing so will improve not only the current governance of automated decisions, but also—in certain cases—the governance of decision-making in general. The implicit (or explicit) biases of human decision-makers can be difficult to find and root out, but we can peer into the "brain" of an algorithm: computational processes and purpose specifications can be declared prior to use and verified afterward. Link: http://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=9570&context=penn_law_review	Transparency, Governance, Accountability
AI7	Do Choice Architects Dream of Programmable Humans?	Brett Frischmann and Evan Selinger Chapter 4 of Upcoming Book	?	So far, we've defined techno-social engineering, explained how techno-social engineering pervades our long history of tool use, and presented several examples of techno-social engineering that illustrate its power and sometimes subtle qualities. In this chapter, we deepen the conversation by discussing the contemporary *nudge agenda*. When we	

				refer to the nudge agenda, we aren't implying that something nefarious is happening. Instead, we're using "agenda" as shorthand to refer to the fact that nudging is a widespread approach to problem-solving and is becoming more popular with every passing year.	
AI8	Finding a Voice	Lane Greene The Economist, Technology Quarterly	01/05/2017	Computers have got much better at translation, voice recognition and speech synthesis, says Lane Greene. But they still don't understand the meaning of language. Link: https://www.economist.com/technology-quarterly/2017-05-01/language	Machine Translation, Speech Recognition, Language Technology
AI9	Values	Bertram F. Malle & Stephan Dickert The Encyclopedia of Social Psychology	2007	Discussion of "Values": Definition, Nature of Values, Taxonomies, Function, Historic and Cultural Differences Link: http://www.academia.edu/27751671/Values_2010_	
AI10	Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research	Danah Boyd, Emily F. Keller, Bonnie Tijerina Data & Society Working Paper	08/04/2016	This report provides valuable insights into the current state of collaboration between librarians and computer science researchers on issues of "big data" ethics. Statements and assertions represent information provided by participants, in combination with a literature review and additional formal and informal research. This report is not meant to be conclusive or comprehensive about all data science research, as we purposefully limited the scope of our work to a narrow band of institutions and actors. Yet, our findings do offer important insights that open up challenging questions and require future exploration. Link: https://www.datasociety.net/pubs/sedr/SupportingEthicsDataResearch_Sept2016.pdf	Ethics, Big Data, Research & Development, Data Sharing, Data Management
AI11	Will Democracy Survive Big Data and Artificial Intelligence?	Dirk Helbing, Bruno S. Frey, Gerd Gigrenzer, Ernst Hafen, Michael Hagner, Yvonne Hofstetter, Jeroen	02/25/2017	It can be expected that supercomputers will soon surpass human capabilities in almost all areas—somewhere between 2020 and 2060. Experts are starting to ring alarm bells. Technology visionaries, such as Elon Musk from Tesla Motors, Bill Gates from Microsoft and Apple co-founder Steve Wozniak, are warning that super-intelligence is a	Big Data, Politics, Society

		van den Hoven, Roberto V. Zicari, Andrej Zwitter Scientific American		serious danger for humanity, possibly even more dangerous than nuclear weapons. Link: https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/	
AI12	Ethical Considerations in Artificial Intelligence Courses	Emanuelle Burton, Judy Goldsmith, Sven Koenig, Benjamin Kuipers, Nicholas Mattei, and Toby Walsh	01/26/2017	The recent surge in interest in ethics in artificial intelligence may leave many educators wondering how to address moral, ethical, and philosophical issues in their AI courses. As instructors we want develop curriculum that not only prepares students to be artificial intelligence practitioners, but also to understand the moral, ethical, and philosophical impacts that artificial intelligence will have on society. In this article we provide practical case studies and links to resources for use by AI educators. We also provide concrete suggestions on how to integrate AI ethics into a general artificial intelligence course and how to teach a stand-alone artificial intelligence ethics course. Link: https://arxiv.org/pdf/1701.07769.pdf	Ethics, Educational Materials
AI13 R5	A Code of Ethics for the Human-Robot Interaction Profession	Laurel D. Riek and Don Howard We Robot	2014	As robots transition into human social environments, a new range of technical, ethical, and legal challenges are arising. This paper discusses the unique ethical challenges facing HRI practitioners designing robots for these spaces, and proposes a code of ethics for the profession. We argue that the affordance of all rights and protections ordinarily assumed in human-human interactions apply to human-robot interaction, and discuss various social, legal, and design considerations to facilitate this. Link: https://www3.nd.edu/~dhoward1/a-code-of-ethics-for-the-human-robot-interaction-profession-riek-howard.pdf	Ethics, Human-Robot Interactions, Social Robots, Design
AI14	How the machine 'thinks': Understanding opacity in machine learning algorithms	Jenna Burrell Big Data & Society	2016	This article considers the issue of opacity as a problem for socially consequential mechanisms of classification and ranking, such as spam filters, credit card fraud detection, search engines, news trends, market segmentation and advertising, insurance or loan qualification, and credit scoring. These mechanisms of classification all frequently	Machine Learning, Transparency

				rely on computational algorithms, and in many cases on *machine learning* algorithms to do this work. In this article, I draw a distinction between three forms of opacity: (1) opacity as intentional corporate or state secrecy, (2) opacity as technical illiteracy, and (3) an opacity that arises from the characteristics of machine learning algorithms and the scale required to apply them usefully. The analysis in this article gets inside the algorithms themselves. I cite existing literatures in computer science, known industry practices (as they are publicly presented), and do some testing and manipulation of code as a form of lightweight code audit. I argue that recognizing the distinct forms of opacity that may be coming into play in a given application is a key to determining which of a variety of technical and non-technical solutions could help to prevent harm. Link: http://journals.sagepub.com/doi/pdf/10.1177/2053951715622512	
AI15	Towards Moral Autonomous Systems	Vicky Charisi, Louise Dennis, Michael Fisher, Robert Lieck, Andreas Matthias, Marja Slavkovik, Janina Sombetzki, Alan F. T. Winfield, Roman Yampolskiy	03/16/2017	Both the ethics of autonomous systems and the problems of their technical implementation have by now been studied in some detail. Less attention has been given to the areas in which these two separate concerns meet. This paper, written by both philosophers and engineers of autonomous systems, addresses a number of issues in machine ethics that are located at precisely the intersection between ethics and engineering. We first discuss different approaches towards the conceptual design of autonomous systems and their implications on the ethics implementation in such systems. Then we examine problematic areas regarding the specification and verification of ethical behavior in autonomous systems, particularly with a view towards the requirements of future legislation. We discuss transparency and accountability issues that will be crucial for any future wide deployment of autonomous systems in society. Finally we consider the, often overlooked, possibility of intentional misuse of AI systems and the possible dangers arising out of deliberately	Ethics, Autonomous Systems, Transparency, Accountability, Intentional Misuse

				unethical design, implementation, and use of autonomous robots. Link: https://arxiv.org/pdf/1703.04741.pdf	
AI16	Slave to the Algorithm? Why a 'Right to an Explanation' is Probably Not the Remedy You are Looking For	Lilian Edwards and Michael Veale SSRN	05/23/2017	However, we argue that a right to an explanation in the GDPR is unlikely to be a complete remedy to algorithmic harms, particularly in some of the core "algorithmic war stories" that have shaped recent attitudes in this domain. We present several reasons for this conclusion. First (section 3), the law is restrictive on when any explanation-related right can be triggered, and in many places is unclear, or even seems paradoxical. Second (section 4), even were some of these restrictions to be navigated, the way that explanations are conceived of legally — as "meaningful information about the logic of processing" — is unlikely to be provided by the kind of ML "explanations" computer scientists have been developing. ML explanations are restricted both by the type of explanation sought, the multi-dimensionality of the domain and the type of user seeking an explanation. However "subject-centric" explanations (SCEs), which restrict explanations to particular regions of a model around a query, show promise for interactive exploration, as do pedagogical rather than de-compositional explanations in dodging developers' worries of IP or trade secrets disclosure. Link: https://ssrn.com/abstract=2972855	Transparency, Right to Explanation, GDPR, Machine Learning, Data Protection
AI17	Privacy by Design in Machine Learning Data Collection: A User Experience Experimentation	Jonathan Vitale, Meg Tonkin, Suman Ojha, Mary-Anne Williams Association for the Advancement of Artificial Intelligence (AAAI)	2017	Designing successful user experiences that use machine learning systems is an area of increasing importance. In supervised machine learning for biometric systems, such as for face recognition, the user experience can be improved. In order to use biometric authentication systems, users are asked for their biometric information together with their personal information. In contexts where there is a frequent and large amount of users to be enrolled, the human expert assisting the data collection process is often replaced in favour of software with a step-by-step user interface. However, this may introduce limitations to the	Machine Learning, Design, Transparency

				overall user experience of the system. User experience should be addressed from the very beginning, during the design process. Furthermore, data collection might also introduce privacy concerns in users and potentially lead them to not use the system. For these reasons, we propose a privacy by design approach in order to maximize the user experience of the system while reducing privacy concerns of users. To do so we suggest a novel experiment in a Human-Robot interaction setting. We investigate the effects of embodiment and transparency on privacy and user experience. We expect that embodiment would enhance the overall user experience of the system, independently from transparency, whereas we expect that transparency would reduce privacy concerns of the participants. In particular, we forecast that transparency, together with embodiment, would significantly reduce privacy considerations of participants, thus maximising the amount of personal information provided by a user. Link: https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/viewFile/15305/14583	
AI18 R12	Law and Regulation of Artificial Intelligence and Robots – Conceptual Framework and Normative Implications	Nicolas Petit SSRN	03/09/2017	Law and regulation of Artificial Intelligence (“AI”) and robots is emerging, fueled by the introduction of industrial and commercial applications in society. A common thread to many regulatory initiatives is to occur without a clear or explicit methodological framework. In light of the many challenges that affect attempts to devise law and regulation in a context of technological incipiency, this paper seeks to offer a methodology geared to the specific fields of AIs and robots. At bottom, the paper addresses the following normative question: should a social planer adopt specific rules and institutions for AIs and robots or should the resolution of issues be left to Hume’s three “fundamental laws of nature,” namely ordinary rules on property and liability, contract laws and the courts system? To explore that question, the analysis is conducted under a public interest framework.	Regulation, Ethics

				Section 1 reviews the main regulatory approaches proposed in the existing AI and robotic literature, and stresses their advantages and disadvantages. Section 2 discusses identifiable regulatory trade-offs, that is the threats and opportunities created by the introduction of regulation in relation to AIs and robotic applications. Section 3 focuses on the specific area of liability as a case-study. Finally, Section 4 proposes a possible methodology for the law and regulation of AIs and robots. In conclusion, the paper proposes to index the regulatory response upon the nature of the externality—positive or negative—created by an AI application, and to distinguish between discrete, systemic and existential externalities. Link: https://ssrn.com/abstract=2931339	
AI19	Data, privacy, and the greater good	Eric Horvitz and Deirdre Mulligan Science	07/17/2015	Large-scale aggregate analyses of anonymized data can yield valuable results and insights that address public health challenges and provide new avenues for scientific discovery. These methods can extend our knowledge and provide new tools for enhancing health and wellbeing. However, they raise questions about how to best address potential threats to privacy while reaping benefits for individuals and society as a whole. The use of machine learning to make leaps across informational and social contexts to infer health conditions and risks from nonmedical data provides representative scenarios for reflections on directions with balancing innovation and regulation. Link: http://erichorvitz.com/data_privacy_greater_good.pdf	Machine Learning, Privacy, Discrimination, Transparency
AI20	An exploration on artificial intelligence application: From security, privacy and ethic perspective	Xiuquan Li and Tao Zhang IEEE	06/19/2017	Artificial intelligence is believed as a disruptive technology, which will change our economy and society significantly in the near future. It can be employed to replace human labors in completing many dangerous and tedious tasks, providing us with more convenient and efficient life. We can benefit a lot from the wide application of this emerging technology. However, there are also potential risks and threats in application of artificial intelligence, which need	Ethics, Privacy, Security

				to be handled in a proper way before extensive usage. In the paper, we make discussions on the security, privacy and ethnic issues in artificial intelligence applications and point out the potential risks and threats. Countermeasures in research, regulation and supervision are suggested and our expectation for artificial intelligence development is given out. Link: http://ieeexplore.ieee.org/abstract/document/7951949/	
AI21	Recoding Privacy Law: Reflections on the Future Relationship Among Law, Technology, and Privacy	Urs Gasser Harvard Law Review Forum	12/2016	Reflecting across centuries and geographies, one common thread emerges: advancement in information and communication technologies have largely been perceived as *threats* to privacy and often led policymakers to seek, and consumers to demand, additional privacy safeguards in the legal and regulatory arenas. Link: https://harvardlawreview.org/2016/12/recoding-privacy-law-reflections-on-the-future-relationship-among-law-technology-and-privacy/	Regulation, Privacy, Privacy-Enhancing Technology
AI22	Introduction to special issue on computational methods for enforcing privacy and fairness in the knowledge society	Sergio Mascetti, Annarita Ricci, and Salvatore Ruggieri Artificial Intelligence and Law Vol. 22, Issue 2, pp 109-111	02/11/2014	This volume presents four papers which address, in an extensive and thorough way, several problems of privacy violation or discrimination with a multi-disciplinary approach. Each paper was reviewed by (at least) two computer scientists and (at least) one legal expert.	Privacy, Fairness, Society
AI23	Smart Policies for Artificial Intelligence	Miles Brundage and Joanna Bryson	08/29/2016	We argue that there already exists de facto artificial intelligence policy - a patchwork of policies impacting the field of AI's development in myriad ways. The key question related to AI policy, then, is not whether AI should be governed at all, but how it is currently being governed, and how that governance might become more informed, integrated, effective, and anticipatory. We describe the main components of de facto AI policy and make some recommendations for how AI policy can be improved,	Regulation

				drawing on lessons from other scientific and technological domains. Link: https://arxiv.org/pdf/1608.08196.pdf	
AI24	Written evidence submitted to the UK Parliamentary Select Committee on Science and Technology Inquiry on Robotics and Artificial Intelligence	A. F. Winfield University of the West of England	07/26/2016	This paper was submitted in response to question 4 of the Parliamentary Science and Technology Committee Inquiry on Robotics and Artificial Intelligence* on: 'The social, legal and ethical issues raised by developments in robotics and artificial intelligence technologies, and how they should be addressed'. The paper was drafted at the request of EPSRC and the UK Robotics and Autonomous Systems (RAS) Network, and an abridged version is incorporated into the UK RAS response to the inquiry. Link: http://eprints.uwe.ac.uk/29428/1/STC_RASinquiry_Winfield.pdf	Ethics, Regulation
AI25	Rethinking the Fourth Amendment in the Age of Supercomputers, Artificial Intelligence, and Robots	Melanie Reid West Virginia Law Review	03/16/2017	Law enforcement currently uses cognitive computers to conduct predictive and content analytics and manage information contained in large police data files. These big data analytics and insight capabilities are more effective than using traditional investigative tools and save law enforcement time and a significant amount of financial and personnel resources. It is not farfetched to think law enforcement's use of cognitive computing will extend to using thinking, real-time robots in the field in the not-so-distant future. IBM's Watson currently uses its artificial intelligence to suggest medical diagnoses and treatment in the healthcare industry and assists the finance industry in improving investment decisions. IBM and similar companies already offer predictive analytics and cognitive computing programs to law enforcement for real-time intelligence and investigative purposes. This article will explore the consequences of predictive and content analytics and the future of cognitive computing, such as utilizing "robots" such as an imaginary "Officer Joe Roboto" in the law enforcement context. Would our interactions with Officer Joe Roboto trigger the same Fourth Amendment concerns and protections as those when	Fourth Amendment, Law Enforcement, Privacy

				dealing with a flesh-and-blood police officer? Are we more afraid of a “robotic” Watson, its capabilities, and lack of feeling and biases, compared to a human law enforcement officer? Assuming someday in the future we might be able to solve the physical limitations of a robot, would a “robotic” officer be preferable to a human one? What sort of limitations would we place on such technology? This article attempts to explore the ramifications of using such computers/robots in the future. Autonomous robots with artificial intelligence and the widespread use of predictive analytics are the future tools of law enforcement in a digital age, and we must come up with solutions as to how to handle the appropriate use of these tools. Link: https://works.bepress.com/melanie_reid/20/download/	
AI26 R11	Machines Without Principals: Liability Rules And Artificial Intelligence	David C. Vladeck 89 Wash. L. Rev. 117	03/2014	The introduction of highly sophisticated autonomous machines may be literally around the corner. Truly autonomous machines may be driving cars through our neighborhoods or piloting drones that fly above our heads sooner than we think. So long as we can conceive of these machines as "agents" of some legal person (individual or virtual), our current system of products liability will be able to address the legal issues surrounding their introduction without significant modification. But the law is not necessarily equipped to address the legal issues that will start to arise when the inevitable occurs and these machines cause injury, but when there is no "principal" directing the actions of the machine. How the law chooses to treat machines without principals will be the central legal question that accompanies the introduction of truly autonomous machines, and at some point, the law will need to have an answer to that question. Link: http://digital.law.washington.edu/dspace-law/bitstream/handle/1773.1/1322/89WLR0117.pdf?sequence=1	Tortious Liability
AI27	Improving the Realism of Synthetic Images	Apple Machine Learning Journal	07/2017	Training machine learning models on standard synthetic images is problematic as the images may not be realistic	Machine Learning

				enough, leading the model to learn details present only in synthetic images and failing to generalize well on real images. One approach to bridge this gap between synthetic and real images would be to improve the simulator which is often expensive and difficult, and even the best rendering algorithm may still fail to model all the details present in the real images. This lack of realism may cause models to overfit to 'unrealistic' details in the synthetic images. Instead of modeling all the details in the simulator, could we learn them from data? To this end, we developed a method for refining synthetic images to make them look more realistic.	
AI28	Robotics and Artificial Intelligence	House of Commons, Science and Technology Committee	10/12/2016	After decades of somewhat slow progress, a succession of advances have recently occurred across the fields of robotics and artificial intelligence (AI), fueled by the rise in computer processing power, the profusion of data, and the development of techniques such a 'deep learning'. Though the capabilities of AI systems are currently narrow and specific, they are, nevertheless, starting to have transformational impacts on everyday life: from driverless cars and supercomputers that can assist doctors with medical diagnoses, to intelligent tutoring systems that can tailor lessons to meet a student's individual cognitive needs. Such breakthroughs raise a host of social, ethical and legal questions. Our inquiry has highlighted several that require serious, ongoing consideration. These include taking steps to minimize bias being accidentally built into AI systems; ensuring that the decisions they make are transparent; and instigating methods that can verify that AI technology is operating as intended and that unwanted, or unpredictable, behaviors are not produced. While the UK is world-leading when it comes to considering the implications of AI, and is well-placed to provide global intellectual leadership on this matter, a coordinated approach is required to harness this expertise. A standing Commission on Artificial Intelligence should be established with a remit to identify principles to govern the	Labor, Ethics, Governance, Discrimination, Privacy, Research & Development

				development and application of AI, provide advice to the Government, and foster public dialogue. Link: https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf	
AI29	#SocialEthics: A guide to embedding ethics in social media research	Harry Evans, Steve Ginnis, and Jamie Bartlett Innovate UK	11/12/2015	One of the focuses of the Wisdom of the Crowd project is to examine the ethical landscape surrounding aggregated social media research. In spring 2015, the first publication of this ethics strand contained a review of the legal and regulatory framework for using social media in market research. This second and final report builds on these findings, presenting our conclusions from quantitative and qualitative primary research with stakeholders and social media users, and outlining our recommendations for how the research industry should look to proceed if it is to be at the forefront of using social media data in an ethical way. Link: https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Publications/im-demos-social-ethics-in-social-media-research-summary.pdf	Research & Development, Ethics
AI30 R10	Averting Robot Eyes	Margot E. Kaminski, Matthew Rueben, William D. Smart, Cindy M. Grimm 76 Md. L. Rev. 983	2017	Home robots will cause privacy harms. At the same time, they can provide beneficial services - as long as consumers trust them. This Essay evaluates potential technological solutions that could help home robots keep their promises, avert their eyes, and otherwise mitigate privacy harms. Our goals are to inform regulators of robot-related privacy harms and the available technological tools for mitigating them, and to spur technologists to employ existing tools and develop new ones by articulating principles for avoiding privacy harms. We posit that home robots will raise privacy problems of three basic types: (1) data privacy problems; (2) boundary management problems; and (3) social/relational problems. Technological design can ward off, if not fully prevent, a number of these harms. We propose five principles for home robots and privacy design: data minimization,	Privacy, Social Robots, Human-Robot Interaction, Design

				purpose specifications, use limitations, honest anthropomorphism, and dynamic feedback and participation. We review current research into privacy-sensitive robotics, evaluating what technological solutions are feasible and where the harder problems lie. We close by contemplating legal frameworks that might encourage the implementation of such design, while also recognizing the potential costs of regulation at these early stages of the technology. Link: From Lexis.	
AI31	Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence	Dr. Michael Guihot, Anne Matthew, and Dr. Nicolas Suzor We Robot Conference at Yale University March 2017 (Note: Please do not cite without the authors' consent)	03/2017	There is a pervading sense of unease that artificially intelligent machines will soon radically alter our lives in ways that are still unknown. Advances in AI technology are developing at an extremely rapid rate as computational power continues to grow exponentially. Even if existential concerns about AI do not materialise, there are enough concrete examples of problems associated with current applications of artificial intelligence to warrant concern about the level of control that exists over developments in AI. Some form of regulation is likely necessary to protect society from risks of harm. However, advances in regulatory capacity have not kept pace with developments in new technologies including AI. This is partly because regulation has become decentered; that is, the traditional role of public regulators such as governments commanding regulation has been dissipated and other participants including those from within the industry have taken the lead. Other contributing factors are the dwindling of resources in governments on the one hand and the increased power of technology companies on the other. These factors have left the field of AI development relatively unregulated. Whatever the reason, it is now more difficult for traditional public regulatory bodies to control the development of AI. In the vacuum, industry participants have begun to self-regulate by promoting soft law options such as codes of practice and standards. We argue that, despite the reduced authority of public regulatory agencies, the risks associated with runaway AI	Regulation

				require regulators to begin to participate in what is largely an unregulated field. In an environment where resources are scarce, governments or public regulators must develop new ways of regulating. This paper proposes solutions to regulating the development of AI ex ante. We suggest a two-step process: first, governments can set expectations and send signals to influence participants in AI development. We adopt the term nudging to refer to this type of influencing. Second, public regulators must participate in and interact with the relevant industries. By doing this, they can gather information and knowledge about the industries, begin to assess risks and then be in a position to regulate those areas that pose most risk first. To conduct a proper risk analysis, regulators must have sufficient knowledge and understanding about the target of regulation to be able to classify various risk categories. We have proposed an initial classification based on the literature that can help to direct pressing issues for further research and a deeper understanding of the various applications of AI and the relative risks they pose. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3017004	
AI32	Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies	Matthew U. Scherer Harvard Journal of Law & Technology, Vol. 29, Number 2	2016	Artificial intelligence technology (or AI) has developed rapidly during the past decade, and the effects of the AI revolution are already being keenly felt in many sectors of the economy. A growing chorus of commentators, scientists, and entrepreneurs has expressed alarm regarding the increasing role that autonomous machines are playing in society, with some suggesting that government regulation may be necessary to reduce the public risks that AI will pose. Unfortunately, the unique features of AI and the manner in which AI can be developed present both practical and conceptual challenges for the legal system. These challenges must be confronted if the legal system is to positively impact the development of AI and ensure that aggrieved parties receive compensation when AI systems cause harm. This	Regulation, Tortious Liability

				article will explore the public risks associated with AI and the competencies of government institutions in managing those risks. It concludes with a proposal for an indirect form of AI regulation based on differential tort liability. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609777&rec=1&srcabs=2586570&alg=1&pos=4	
AI33	The Inadequate, Invaluable Fair Information Practices	Woodrow Hartzog 76 Md. L. Rev. 952	2017	A sea change is afoot in the relationship between privacy and technology. FIPs-based regimes were relatively well-equipped for the first wave of personal computing. But automated technologies and exponentially greater amounts of data have pushed FIPs principles like data minimization, transparency, choice, and access to the limit. Advances in robotics, genetics, biometrics, and algorithmic decision-making are challenging the idea that rules meant to ensure fair aggregation of personal information in databases are sufficient. Control over information in databases isn't even the half of it anymore. The mass connectivity of the "Internet of Things" and near ubiquity of mobile devices make the security and surveillance risks presented by the isolated computer terminals and random CCTV cameras of the "80s and "90s seem quaint. But we've come too far with the FIPs to turn back now. The FIPs model of privacy regulation has been adopted by nearly every country in the world that has decided to take data protection seriously. Normatively, the FIPs have been with us so long that in many ways they have become synonymous with privacy. At this point, abandoning the FIPs is out of the question. Even tinkering with them requires true urgency and a good plan. But modern privacy problems require more than just the FIPs. Hence, the pickle. Link: From Lexis.	Privacy, Fair Information Practices
AI34	Regulating Inscrutable Systems	Andrew D. Selbst and Solon Barocas	2017	This Article takes seriously the calls for regulation via explanation to investigate how existing laws implementing such calls fare, and whether interpretability research can	Regulation, Transparency, Machine Learning

				fix the flaws. Ultimately, it argues that while machine interpretability may make compliance with existing legal regimes easier, or possible in the first instance, a focus on explanation alone fails to fulfill the overarching normative purpose of the law, even when compliance can be achieved. The paper concludes with a call to consider where such goals would be better served by other means, including mechanisms to directly assess whether models are fair and just. Link: www.werobot2017.com/wp-content/uploads/2017/03/Selbst-and-Barocas-Regulating-Inscrutable-Systems-1.pdf	
AI35	An Education Theory of Fault for Autonomous Systems	William D. Smart, Cindy Grimm, Woodrow Hartzog	3/22/2017	We think that part of the problem with our discussion of fault is that we have yet to settle on the best approach and language to use to specifically target culpable behavior in the design and deployment for automated systems. The purpose of this paper is to offer an additional structured and nuance way of thinking about the duties and culpable behavior of all the relevant stakeholders in the creation and deployment of autonomous systems. In this article, we argue that some of the most articulable failures in the creation and deployment of unpredictable systems lie in the lack of communication, clarity, and education between the procurer, developer, and users of automated systems. In other words, while it is hard to exert meaningful "control" over automated systems to get them to act predictably, developers and procurers have great control over how much they test and articulate the limits of an automated technology to all the other relevant parties. This makes testing and education one of the most legally relevant point of failures when automated systems harm people. Link: http://www.werobot2017.com/wp-content/uploads/2017/03/Smart-Grimm-Hartzog-Education-We-Robot.pdf	Tortious Liability

AI36	Artificial Intelligence Policy: A Primer and Roadmap	Ryan Calo	08/2017	Talk of artificial intelligence is everywhere. People marvel at the capacity of machines to translate any language and master any game. Others condemn the use of secret algorithms to sentence criminal defendants or recoil at the prospect of machines gunning for blue, pink, and white-collar jobs. Some worry aloud that artificial intelligence will be humankind's "final invention." This essay, prepared in connection with UC Davis Law Review's 50th anniversary symposium, explains why AI is suddenly on everyone's mind and provides a roadmap to the major policy questions AI raises. The essay is designed to help policymakers, investors, technologists, scholars, and students understand the contemporary policy environment around AI at least well enough to initiate their own exploration. Topics covered include: • Justice and equity • Use of force • Safety and certification • Privacy (including data parity); and • Taxation and displacement of labor In addition to these topics, the essay will touch briefly on a selection of broader systemic questions: • Institutional configuration and expertise • Investment and procurement • Removing hurdles to accountability; and • Correcting mental models of AI Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3015350	Privacy, Labor, Fairness, Regulation
AI37 R9	Who's Johnny? Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy	Kate Darling	08/29/2016	People have a tendency to project life-like qualities onto robots. As we increasingly create spaces where robotic technology interacts with humans, this inclination raises ethical questions around use and policy. A human-robot-interaction experiment conducted in our lab	Privacy, Human-Robot Interactions, Social Robots

				indicates that framing robots through anthropomorphic language (like a personified name or story) can impact how people perceive and treat a robot. This chapter explores the effects of encouraging or discouraging people to anthropomorphize robots through framing. I discuss concerns about anthropomorphizing robotic technology in certain contexts, but argue that there are also cases where encouraging anthropomorphism is desirable. Because people respond to framing, framing could help to separate these cases. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2588669	
AI38	Privacy Self-Management and the Consent Dilemma	Daniel Solove	10/18/2015	The current regulatory approach for protecting privacy involves what I refer to as "privacy self-management" — the law provides people with a set of rights to enable them to decide how to weigh the costs and benefits of the collection, use, or disclosure of their information. People's consent legitimizes nearly any form of collection, use, and disclosure of personal data. Although privacy self-management is certainly a necessary component of any regulatory regime, I contend in this Article that it is being asked to do work beyond its capabilities. Privacy self-management does not provide meaningful control. Empirical and social science research has undermined key assumptions about how people make decisions regarding their data, assumptions that underpin and legitimize the privacy self-management model. Moreover, people cannot appropriately self-manage their privacy due to a series of structural problems. There are too many entities collecting and using personal data to make it feasible for people to manage their privacy separately with each entity. Moreover, many privacy harms are the result of an aggregation of pieces of data over a period of time by different entities. It is virtually impossible for people to weigh the costs and benefits of revealing	Regulation, Privacy

				information or permitting its use or transfer without an understanding of the potential downstream uses, further limiting the effectiveness of the privacy self-management framework. In addition, privacy self-management addresses privacy in a series of isolated transactions guided by particular individuals. Privacy costs and benefits, however, are more appropriately assessed cumulatively and holistically — not merely at the individual level. In order to advance, privacy law and policy must confront a complex and confounding dilemma with consent. Consent to collection, use, and disclosure of personal data is often not meaningful, and the most apparent solution — paternalistic measures — even more directly denies people the freedom to make consensual choices about their data. In this Article, I propose several ways privacy law can grapple with the consent dilemma and move beyond relying too heavily on privacy self-management. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2171018	
AI39	Artificial Intelligence and Public Policy	Adam D. Thierer, Andrea Castillo, Raymond Russell	08/22/2017	There is growing interest in the market potential of artificial intelligence (AI) technologies and applications as well as in the potential risks that these technologies might pose. As a result, questions are being raised about the legal and regulatory governance of AI, machine learning, "autonomous" systems, and related robotic and data technologies. Fearing concerns about labor market effects, social inequality, and even physical harm, some have called for precautionary regulations that could have the effect of limiting AI development and deployment. In this paper, we recommend a different policy framework for AI technologies. At this nascent stage of AI technology development, we think a better case can be made for prudence, patience, and a continuing embrace of "permissionless innovation" as it pertains to modern digital technologies. Unless a compelling case can be made that a new invention will bring serious harm to society, innovation	Labor, Fairness, Regulation, Research and Development

				should be allowed to continue unabated, and problems, if they develop at all, can be addressed later. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3021135	
AI40	Averting Robot Eyes (Repeat of AI30)				
AI41	AI, Ethics and Enhanced Data Stewardship	Information Accountability Foundation (IAF)	09/20/2017	The terms data ethics and ethical processing are in vogue. The popularity of these concepts stems from the rapid growth of innovative data-driven technologies and the application of these innovations to areas that can have a material impact on people's daily lives. The sheer volume of data that is observable and where inferences can be made as the product of analytics has and will continue to impact many facets of people's lives, including new health solutions, business models, personalization for individuals and tangible benefits for society. Yet those same data and technologies can have an inappropriate impact and even harm on individuals and groups of individuals and cause negative impact on societal goals and values. An evolved form of accountability, ethical processing, applicable to advanced analytics, is needed to help enable the realization of the benefits of this use of data but address any resulting risks. The Information Accountability Foundation (IAF) has established an Artificial Intelligence (AI) and Ethics Project to tackle these issues. The Project's objective is to begin the global discussion of how organisations might address the application of ethical data processing to new technologies. The IAF thinks this work is particularly necessary where data enabled decisions are made without the intervention of people. In these circumstances, corporate governance takes on added importance and ethical objectives need to be built into data processing architecture. The IAF further believes the governance structures being suggested are also applicable where data from observational technologies, such as sensors,	Ethics, Data Protection, Accountability

				inferences from analytics, and data synthesized from other data sets, are used to drive advanced analytics. Link: http://informationaccountability.org/wp-content/uploads/Artificial-Intelligence-Ethics-and-Enhanced-Data-Stewardship.pdf	
AI42 R8	Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being	Jason Borenstein and Ron Arkin	03/04/2015	Robots are becoming an increasingly pervasive feature of our personal lives. As a result, there is growing importance placed on examining what constitutes appropriate behavior when they interact with human beings. In this paper, we discuss whether companion robots should be permitted to "nudge" their human users in the direction of being "more ethical". More specifically, we use Rawlsian principles of justice to illustrate how robots might nurture "socially just" tendencies in their human counterparts. Designing technological artifacts in such a way to influence human behavior is already well-established but merely because the practice is commonplace does not necessarily resolve the ethical issues associated with its implementation. Link: In Dropbox	Ethics, Human-Robot Interaction
AI43 R2	Robots in American Law	Ryan Calo	03/2016	This article closely examines a half century of case law involving robots—just in time for the technology itself to enter the mainstream. Most of the cases involving robots have never found their way into legal scholarship. And yet, taken collectively, these cases reveal much about the assumptions and limitations of our legal system. Robots blur the line between people and instrument, for instance, and faulty notions about robots lead jurists to questionable or contradictory results. The article generates in all nine case studies. The first set highlights the role of robots as the objects of American law. Among other issues, courts have had to decide whether robots represent something "animate" for purposes of import tariffs, whether robots can "perform" as that term is understood in the context of a state tax on performance	Law of Robots

				halls, and whether a salvage team "possesses" a shipwreck it visits with an unmanned submarine. The second set of case studies focuses on robots as the subjects of judicial imagination. These examples explore the versatile, often pejorative role robots play in judicial reasoning itself. Judges need not be robots in court, for instance, or apply the law robotically. The robotic witness is not to be trusted. And people who commit crimes under the robotic control of another might avoid sanction. Together these case studies paint a nuanced picture of the way courts think about an increasingly important technology. Themes and questions emerge that illuminate the path of robotics law and test its central claims to date. The article concludes that jurists on the whole possess poor, increasingly outdated views about robots and hence will not be well positioned to address the novel challenges they continue to pose. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2737598	
AI44	Artificial Intelligence Policy: A Primer and Roadmap (Repeat of AI36)				
AI45	Equality of Opportunity in Supervised Learning	Mortiz Hardt, Eric Price, Nathan Srebro	10/2016	We propose a criterion for discrimination against a specified sensitive attribute in supervised learning, where the goal is to predict some target based on available features. Assuming data about the predictor, target, and membership in the protected group are available, we show how to optimally adjust any learned predictor so as to remove discrimination according to our definition. Our framework also improves incentives by shifting the cost of poor classification from disadvantaged groups to the decision maker, who can respond by improving the classification accuracy.	Discrimination, Fairness

				In line with other studies, our notion is oblivious: it depends only on the joint statistics of the predictor, the target and the protected attribute, but not on interpretation of individual features. We study the inherent limits of defining and identifying biases based on such oblivious measures, outlining what can and cannot be inferred from different oblivious tests. We illustrate our notion using a case study of FICO credit scores. Link: https://arxiv.org/abs/1610.02413	
AI46 R7	Robots Should Be Slaves	Joanna Bryson Chapter from Close Engagements with Artificial Companions: Key social, psychological, ethical and design issue, Yorick Wilks (ed.), John Benjamins	2010	Robots should not be described as persons, nor given legal nor moral responsibility for their actions. Robots are fully owned by us. We determine their goals and behaviour, either directly or indirectly through specifying their intelligence or how their intelligence is acquired. In humanising them, we not only further dehumanise real people, but also encourage poor human decision making in the allocation of resources and responsibility. This is true at both the individual and the institutional level. This chapter describes both causes and consequences of these errors, including consequences already present in society. I make specific proposals for best incorporating robots into our society. The potential of robotics should be understood as the potential to extend our own abilities and to address our own goals. Link: http://www.cs.bath.ac.uk/~jjb/ftp/Bryson-Slaves-Book09.html	Human-Robot Interaction, Society
AI47	Artificial Intelligence: the Public Policy Opportunity	Intel	10/18/2017	Intel powers the cloud and billions of smart, connected computing devices. Due to the decreasing cost of computing enabled by Moore's Law1 and the increasing availability of connectivity, these connected devices are now generating millions of terabytes of data every day. Recent breakthroughs in computer and data science give us the ability to timely analyze and derive immense value from that data. As Intel distributes the computing capability of the data center across the entire global	Research & Development, Labor, Data Protection, Privacy, Ethics, Transparency, Accountability

				network, the impact of artificial intelligence is significantly increasing. Artificial intelligence is creating an opportunity to drive a new wave of economic progress while solving some of the world's most difficult problems. This is the artificial intelligence (AI) opportunity. To allow AI to realize its potential, governments need to create a public policy environment that fosters AI innovation, while also mitigating unintended societal consequences. This document presents Intel's AI public policy recommendations. Link: https://blogs.intel.com/policy/files/2017/10/Intel-Artificial-Intelligence-Public-Policy-White-Paper-2017.pdf	
AI48	The Criminal Liability of AI Entities	Gabriel Hallevy	03/04/2010	In 1981, a 37-year-old Japanese employee of a motorcycle factory was killed by an artificial-intelligence robot working near him. The robot erroneously identified the employee as a threat to its mission, and calculated that the most efficient way to eliminate this threat was by pushing him into an adjacent operating machine. Using its very powerful hydraulic arm, the robot smashed the surprised worker into the operating machine, killing him instantly, and then resumed its duties with no one to interfere with its mission. Unfortunately, this is not science fiction, and the legal question is: Who is to be held liable for this killing? Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1564096&download=yes	Criminal Law, Law of Robots, Law Enforcement
AI49	Methodologies to Guide Ethical Research and Design	IEEE	2017	To ensure autonomous and intelligent systems (A/IS) are aligned to benefit humanity A/IS research and design must be underpinned by ethical and legal norms as well as methods. We strongly believe that a value-based design methodology should become the essential focus for the modern A/IS organization. Link: https://standards.ieee.org/develop/indconn/ec/ead_methodologies%20to%20guide%20ethical%20research.pdf	Ethics, Research & Development, Design

				https://standards.ieee.org/develop/indconn/ec/ead_methodologies_research_v2.pdf	
AI50	Artificial Intelligence and Privacy	Datatilsynet The Norwegian Data Protection Authority	01/2018	The Norwegian Data Protection Authority (DPA) believes itto be imperative that we further our knowledge about the privacy implications of artificial intelligence and discuss them, not only in order to safeguard the right to privacy of the individual, but also to meet the requirements of society at large. Link: https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf	Privacy, GDPR, Fairness, Transparency
AI51	Artificial Intelligence and the 'Good Society': the US, EU and UK Approach	Corinne Cath et al.	03/28/2017	In October 2016, the White House, the European Parliament, and the UK House of Commons each issued a report outlining their visions on how to prepare society for the widespread use of artificial intelligence (AI). In this article, we provide a comparative assessment of these three reports in order to facilitate the design of policies favorable to the development of a 'good AI society'. To do so, we examine how each report addresses the following three topics: (a) the development of a 'good AI society'; (b) the role and responsibility of the government, the private sector, and the research community (including academia) in pursuing such a development; and (c) where the recommendations to support such a development may be in need of improvement. Our analysis concludes that the reports address adequately various ethical, social, and economic topics, but come short of providing an overarching political vision and long-term strategy for the development of a 'good AI society'. In order to contribute to fill this gap, in the conclusion we suggest a two-pronged approach. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2906249	Ethics, Society
AI52	Artificial Intelligence and Consumer Privacy	Ginger Zhe Jin	01/29/2018	Thanks to big data, artificial intelligence (AI) has spurred exciting innovations. In the meantime, AI and big data are reshaping the risk in consumer privacy and data security. In	Privacy, Data Protection,

				this essay, I first define the nature of the problem and then present a few facts about the ongoing risk. The bulk of the essay describes how the U.S. market copes with the risk in current policy environment. It concludes with key challenges facing researchers and policy makers. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3112040	Consumer Protection, Big Data
AI53	Machine Ethics: Creating an Ethical Intelligent Agent	Michael Anderson and Susan Anderson	2007	The newly emerging field of machine ethics (Anderson and Anderson 2006) is concerned with adding an ethical dimension to machines. Unlike computer ethics—which has traditionally focused on ethical issues surrounding humans' use of machines—machine ethics is concerned with ensuring that the behavior of machines toward human users, and perhaps other machines as well, is ethically acceptable. In this article we discuss the importance of machine ethics, the need for machines that represent ethical principles explicitly, and the challenges facing those working on machine ethics. We also give an example of current research in the field that shows that it is possible, at least in a limited domain, for a machine to abstract an ethical principle from examples of correct ethical judgments and use that principle to guide its own behavior. Link: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.430.1790&rep=rep1&type=pdf	Ethics
AI54	Accountability of AI Under the Law: The Role of Explanation	Finale Doshi-Velez, Mason Kortz		The ubiquity of systems using artificial intelligence or "AI" has brought increasing attention to how those systems should be regulated. The choice of how to regulate AI systems will require care. AI systems have the potential to synthesize large amounts of data, allowing for greater levels of personalization and precision than ever before – applications range from clinical decision support to autonomous driving and predictive policing. That said, our AIs continue to lag in common sense reasoning [McCarthy, 1960], and thus there exist legitimate concerns about the intentional and unintentional negative consequences of AI	Accountability, Transparency

				systems [Bostrom, 2003, Amodei et al., 2016, Sculley et al., 2014]. How can we take advantage of what AI systems have to offer, while also holding them accountable? In this work, we focus on one tool: explanation. Questions about a legal right to explanation from AI systems was recently debated in the EU General Data Protection Regulation [Goodman and Flaxman, 2016, Wachter et al., 2017a], and thus thinking carefully about when and how explanation from AI systems might improve accountability is timely. Good choices about when to demand explanation can help prevent negative consequences from AI systems, while poor choices may not only fail to hold AI systems accountable but also hamper the development of much-needed beneficial AI systems. Below, we briefly review current societal, moral, and legal norms around explanation, and then focus on the different contexts under which explanation is currently required under the law. We find that there exists great variation around when explanation is demanded, but there also exist important consistencies: when demanding explanation from humans, what we typically want to know is whether and how certain input factors affected the final decision or outcome. These consistencies allow us to list the technical considerations that must be considered if we desired AI systems that could provide kinds of explanations that are currently required of humans under the law. Contrary to popular wisdom of AI systems as indecipherable black boxes, we find that this level of explanation should generally be technically feasible but may sometimes be practically onerous – there are certain aspects of explanation that may be simple for humans to provide but challenging for AI systems, and vice versa. As an interdisciplinary team of legal scholars, computer scientists, and cognitive scientists, we recommend that for the	

				present, AI systems can and should be held to a similar standard of explanation as humans currently are; in the future we may wish to hold an AI to a different standard. Link: https://arxiv.org/pdf/1711.01134.pdf	
AI55	Counterfactual Explanations Without Opening The Black Box: Automated Decisions and the GDPR	Sandra Wachter, Brent Mittelstadt, Chris Russell	10/06/2017	There has been much discussion of the "right to explanation" in the EU General Data Protection Regulation, and its existence, merits, and disadvantages. Implementing a right to explanation that opens the 'black box' of algorithmic decision-making faces major legal and technical barriers. Explaining the functionality of complex algorithmic decision-making systems and their rationale in specific cases is a technically challenging problem. Some explanations may offer little meaningful information to data subjects, raising questions around their value. Data controllers have an interest to not disclose information about their algorithms that contains trade secrets, violates the rights and freedoms of others (e.g. privacy), or allows data subjects to game or manipulate decision-making. Explanations of automated decisions need not hinge on the general public understanding how algorithmic systems function. Even though interpretability is of great importance and should be pursued, explanations can, in principle, be offered without opening the black box. Looking at explanations as a means to help a data subject act rather than merely understand, one can gauge the scope and content of explanations according to the specific goal or action they are intended to support. From the perspective of individuals affected by automated decision-making, we propose three aims for explanations: (1) to inform and help the individual understand why a particular decision was reached, (2) to provide grounds to contest the decision if the outcome is undesired, and (3) to understand what could be changed to receive a desired result in the future, based on the current decision-making model. We assess how each of these goals finds support in the GDPR, and the extent to which they hinge on opening	Machine Learning, Transparency, Right to Explanation, GDPR, Accountability, Data Protection, Machine Learning

				the 'black box'. We suggest data controllers should offer a particular type of explanation, 'unconditional counterfactual explanations', to support these three aims. These counterfactual explanations describe the smallest change to the world that would obtain a desirable outcome, or to arrive at a "close possible world." As multiple variables or sets of variables can lead to one or more desirable outcomes, multiple counterfactual explanations can be provided, corresponding to different choices of nearby possible worlds for which the counterfactual holds. Counterfactuals describe a dependency on the external facts that lead to that decision without the need to convey the internal state or logic of an algorithm. As a result, counterfactuals serve as a minimal solution that bypasses the current technical limitations of interpretability, while striking a balance between transparency and the rights and freedoms of others (e.g. privacy, trade secrets). Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3063289	
AI56	Emergent AI, Social Robots and the Law: Security, Privacy and Policy Issues	Ramesh Subramanian	2017	The rapid growth of AI systems has implications on a wide variety of fields. It can prove to be a boon to disparate fields such as healthcare, education, global logistics and transportation, to name a few. However, these systems will also bring forth far-reaching changes in employment, economy and security. As AI systems gain acceptance and become more commonplace, certain critical questions arise: What are the legal and security ramifications of the use of these new technologies? Who can use them, and under what circumstances? What is the safety of these systems? Should their commercialization be regulated? What are the privacy issues associated with the use of these technologies? What are the ethical considerations? Who has responsibility for the large amounts of data that is collected and manipulated by these systems? Could these systems fail? What is the recourse if there is a system failure? These questions are but a small subset of possible	Privacy, Ethics, Social Robots, Security

				questions in this key emerging field. In this paper, we focus primarily on the legal questions that relate to the security, privacy, ethical, and policy considerations that emerge from one of these types of technologies, namely social robots. We begin with a history of the field, then go deeper into legal issues, the associated issues of security, privacy and ethics, and consider some solutions to these issues. Finally, we conclude with a look at the future as well as a modest proposal for future research addressing some of the challenges listed. Link: http://scholarworks.lib.csusb.edu/jitim/vol26/iss3/4/	
AI57	Algorithmic Transparency for the Smart City	Robert Brauneis and Ellen Goodman	09/21/2017	Emerging across many disciplines are questions about algorithmic ethics – about the values embedded in artificial intelligence and big data analytics that increasingly replace human decision-making. Many are concerned that an algorithmic society is too opaque to be accountable for its behavior. An individual can be denied parole or denied credit, fired or not hired for reasons she will never know and cannot be articulated. In the public sector, the opacity of algorithmic decision-making is particularly problematic both because governmental decisions may be especially weighty, and because democratically-elected governments bear special duties of accountability. Investigative journalists have recently exposed the dangerous impenetrability of algorithmic processes used in the criminal justice field – dangerous because the predictions they make can be both erroneous and unfair, with none the wiser. We set out to test the limits of transparency around governmental deployment of big data analytics, focusing our investigation on local and state government use of predictive algorithms. It is here, in local government, that algorithmically-determined decisions can be most directly impactful. And it is here that stretched agencies are most likely to hand over the analytics to private vendors, which may make design and policy choices out of the sight of the client agencies, the public, or both. To see just how	Ethics, Transparency, Big Data

				impenetrable the resulting “black box” algorithms are, we filed 42 open records requests in 23 states seeking essential information about six predictive algorithm programs. We selected the most widely-used and well-reviewed programs, including those developed by for-profit companies, nonprofits, and academic/private sector partnerships. The goal was to see if, using the open records process, we could discover what policy judgments these algorithms embody, and could evaluate their utility and fairness. To do this work, we identified what meaningful “algorithmic transparency” entails. We found that in almost every case, it wasn’t provided. Over-broad assertions of trade secrecy were a problem. But contrary to conventional wisdom, they were not the biggest obstacle. It will not usually be necessary to release the code used to execute predictive models to dramatically increase transparency. We conclude that publicly-deployed algorithms will be sufficiently transparent only if (1) governments generate appropriate records about their objectives for algorithmic processes and subsequent implementation and validation; (2) government contractors reveal to the public agency sufficient information about how they developed the algorithm; and (3) public agencies and courts treat trade secrecy claims as the limited exception to public disclosure that the law requires. We present what we believe are eight principal types of information that records concerning publicly implemented algorithms should contain. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499	
AI58	When Enough Is Enough: Location Tracking, Mosaic Theory and Machine Learning	Steven M. Bellovin et al.	2014	In five parts, this article advances the conclusion that the duration of investigations is relevant to their substantive Fourth Amendment treatment because duration affects the accuracy of the predictions. Though it was previously difficult to explain why an investigation of four weeks was substantively different from an investigation of four hours,	Machine Learning, Privacy, Law Enforcement

				we now have a better understanding of the value of aggregated data when viewed through a machine learning lens. In some situations, predictions of startling accuracy can be generated with remarkably few data points. Furthermore, in other situations accuracy can increase dramatically above certain thresholds. For example, a 2012 study found the ability to deduce ethnicity moved sideways through five weeks of phone data monitoring, jumped sharply to a new plateau at that point, and then increased sharply again after twenty-eight weeks. More remarkably, the accuracy of identification of a target's significant other improved dramatically after five days' worth of data inputs. Experiments like these support the notion of a threshold, a point at which it makes sense to draw a Fourth Amendment line. Link: http://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=2379&context=fac_pubs	
AI59	People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship	Ryan Calo	08/24/2009	This article updates the traditional discussion of privacy and technology, focused since the days of Warren and Brandeis on the capacity of technology to manipulate information. It includes a novel dimension around the impact to privacy of anthropomorphic or social design. Technologies designed to emulate people - through voice, animation, and natural language - are increasingly commonplace, showing up in our cars, computers, phones, and homes. A rich literature in communications and psychology suggests that we are hardwired to react to such technology as though a person were actually present. Social interfaces accordingly capture our attention and improve interactivity, and can free up our hands for other tasks. At the same time, technologies that emulate people have the potential to implicate long-standing privacy values. One of the well-documented effects of interfaces and devices that emulate people is the sensation of being observed and	Privacy, Society

				evaluated. Their presence can alter our attitude, behavior, and physiological state. Widespread adoption of such technology may accordingly lessen opportunities for solitude and chill curiosity and self-development. These effects are all the more dangerous in that they cannot be addressed through traditional privacy protections such as encryption or anonymization. The unique properties of social technology also present an opportunity to improve privacy, particularly online. Careful use of anthropomorphic design might one day replace today's ineffective privacy policies with a direct, visceral notice that lines up our experience with actual information practice. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1458637	
AI60	The Future Computed: Artificial Intelligence and its role in society	Microsoft	2018	Beyond our personal lives, AI will enable breakthrough advances in areas like healthcare, agriculture, education and transportation. It's already happening in impressive ways. But as we've witnessed over the past 20 years, new technology also inevitably raises complex questions and broad societal concerns. As we look to a future powered by a partnership between computers and humans, it's important that we address these challenges head on. How do we ensure that AI is designed and used responsibly? How do we establish ethical principles to protect people? How should we govern its use? And how will AI impact employment and jobs? To answer these tough questions, technologists will need to work closely with government, academia, business, civil society and other stakeholders. At Microsoft, we've identified six ethical principles – fairness, reliability and safety, privacy and security, inclusivity, transparency, and	Society, Labor, Ethics

				accountability – to guide the cross-disciplinary development and use of artificial intelligence. The better we understand these or similar issues — and the more technology developers and users can share best practices to address them — the better served the world will be as we contemplate societal rules to govern AI. We must also pay attention to AI's impact on workers. What jobs will AI eliminate? What jobs will it create? If there has been one constant over 250 years of technological change, it has been the ongoing impact of technology on jobs — the creation of new jobs, the elimination of existing jobs and the evolution of job tasks and content. This too is certain to continue.	
Robotics					
R1	Inventing Japan's Robotics Culture: The repeated assembly of science, technology, and culture in social robotics	Sema Sabanovic	06/2014	Using interviews, participant observation, and published documents, this article analyzes the co-construction of robotics and culture in Japan through the technical discourse and practices of robotics researchers. Three cases from current robotics research--the seal-like robot PARO, the Humanoid Robotics Project HRP-2 humanoid, and 'kansei robotics' - show the different ways in which scientists invoke culture to provide epistemological grounding and possibilities for social acceptance of their work. These examples show how the production and consumption of social robotic technologies are associated with traditional crafts and values, how roboticists negotiate among social, technical, and cultural constraints while designing robots, and how humans and robots are constructed as cultural subjects in social robotics discourse. The conceptual focus is on the repeated assembly of cultural models of social behavior, organization, cognition, and technology through roboticists' narratives about the development of advanced robotic technologies. This article provides a picture of robotics as the dynamic construction of technology and culture and concludes with a discussion of the limits and possibilities of this vision in promoting a culturally situated understanding of technology and a multicultural view of science.	Social Robots, Human-Robot Interaction, Culture

				Source: 44(3) Social Studies of Science 342-367	
R2 AI43	Robots in American Law (See AI List)				
R3	Using Embodied Design Improvisation as a Design Research Tool	David Sirkin and Wendy Ju	10/2014	Embodied design improvisation is a generative and evaluative technique to elicit tacit knowledge about embodied experience. It incorporates storyboarding, Wizard of Oz prototyping, domain expert improvisation, video prototyping and crowdsourced experimentation. We have been developing this technique to design physical interactions with expressive, robotic everyday devices, eliciting the tacit rules and behavior patterns that comport with the social expectations established by human-human interactions. By codifying and providing an example of this technique, we hope to encourage its adoption in other design domains. Link: http://www.wendyju.com/publications/HBiD2014SirkinJu.pdf	Design
R4	Unfair and Deceptive Robots	Woodrow Hartzog	2015	Robots, like household helpers, personal digital assistants, automated cars, and personal drones are or will soon be available to consumers. These robots raise common consumer protection issues, such as fraud, privacy, data security, and risks to health, physical safety and finances. Robots also raise new consumer protection issues, or at least call into question how existing consumer protection regimes might be applied to such emerging technologies. Yet it is unclear which legal regimes should govern these robots and what consumer protection rules for robots should look like. The thesis of the Article is that the FTC's grant of authority and existing jurisprudence make it the preferable regulatory agency for protecting consumers who buy and interact with robots. The FTC has proven to be a capable regulator of communications, organizational procedures, and design, which are the three crucial concepts for safe consumer robots. Additionally, the structure and history of	Human-Robot Interactions, Social Robots, Consumer Protection, Regulation

				the FTC shows that the agency is capable of fostering new technologies as it did with the Internet. The agency generally defers to industry standards, avoids dramatic regulatory lurches, and cooperates with other agencies. Consumer robotics is an expansive field with great potential. A light but steady response by the FTC will allow the consumer robotics industry to thrive while preserving consumer trust and keeping consumers safe from harm. Link: http://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=3675&context=mlr	
R5 AI13	A Code of Ethics for the Human-Robot Interaction Profession (See AI List)				
R6	How Humans Respond to Robots: Building Public Policy Through Good Design	Heather Knight Center for Technology Innovation at Brookings	07/2014	My purpose in this paper is not to provide detailed policy recommendations but to describe a series of important choices we face in designing robots that people will actually want to use and engage with. Design considerations today can foreshadow policy choices in the future. Much of the current research into human-robotic teams seeks to explore plausible practical applications given improved technological knowhow and better social understandings. For now, these are pre-policy technical design challenges for collaborative robots that will, or could, have public policy implications down the road. But handling them well at the design phase may reduce policy pressures over time. Link: https://www.brookings.edu/wp-content/uploads/2014/07/HumanRobot-PartnershipsR2.pdf	Human-Robot Interactions, Social Robots, Design
R7 AI46	Robots Should Be Slaves (See AI List)				
R8 AI42	Robotic Nudges: The Ethics of Engineering A More Socially Just Human Being (See AI List)				

R9 AI37	Who's Johnny? Anthropomorphic Framing in Human-Robot Interaction, Integration and Policy (See AI List)				
R10 AI30	Averting Robot Eyes (See AI List)				
R11 AI26	Machines Without Principles: Liability Rules and Artificial Intelligence (See AI List)				
R12 AI18	Law and Regulation of Artificial Intelligence and Robots – Conceptual Framework and Normative Implications (See AI List)				
R13	Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots	C. Bartneck, D. Kulic, E. Croft, S. Zoghbi	01/02/2009	This study emphasizes the need for standardized measurement tools for human robot interactions (HRI). If we are to make progress in this field then we must be able to compare the results from different studies. A literature review has been performed on the measurements of five key concepts in HRI: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. The results have been distilled into five consistent questionnaires using semantic differential scales. We report reliability and validity indicators based on several empirical studies that used these questionnaires. It is our hope that these questionnaires can be used by robot developers to monitor their progress. Psychologists are invited to further develop the questionnaires by adding new concepts, and to conduct further validations where it appears necessary. Link: http://www.bartneck.de/publications/2009/measurementInstrumentsRobots/	Human-Robot Interaction

R14	A Review of Verbal and Non-Verbal Human-Robot Interactive Communication	N. Mavridis	01/2015	In this paper, an overview of human–robot interactive communication is presented, covering verbal as well as non-verbal aspects. Following a historical introduction, and motivation towards fluid human–robot communication, ten desiderata are proposed, which provide an organizational axis both of recent as well as of future research on human–robot communication. Then, the ten desiderata are examined in detail, culminating in a unifying discussion, and a forward-looking conclusion. Link: https://www.sciencedirect.com/science/article/pii/S0921889014002164	Human-Robot Interaction
R15	Robot Social Intelligence	Mary-Anne Williams	2012	Robots are pervading human society today at an ever-accelerating rate, but in order to actualize their profound potential impact, robots will need cognitive capabilities that support the necessary social intelligence required to fluently engage with people and other robots. People are social agents and robots must develop sufficient social intelligence to engage with them effectively. Despite their enormous potential, robots will not be accepted in society unless they exhibit social intelligence skills. They cannot work with people effectively if they ignore the limitations, needs, expectations and vulnerability of people working in and around their workspaces. People are limited social agents, i.e. they do not have unlimited cognitive, computational and physical capabilities. People have limited ability in perceiving, paying attention, reacting to stimuli, anticipating, and problem-solving. In addition, people are constrained by their morphology; it limits their physical strength for example. People cannot be expected to and will not compensate for social deficiencies of robots, hence widespread acceptance and integration of robots into society will only be achieved if robots possess the sufficient social intelligence to communicate, interact and collaborate with people. In this paper we identify the key cognitive capabilities robots will require to achieve appropriate levels of social intelligence for safe and effective engagement with people. This work serves as a	Social Robots

				proto-blueprint that can inform the emerging roadmap and research agenda for the new exciting and challenging field of social robotics. Link: https://link.springer.com/chapter/10.1007/978-3-642-34103-8_5	
R16	Effects of a Social Robot's Autonomy and Group Orientation on Human Decision-Making	P.P. Rau, Y. Li, J. Liu	2013	Social attributes of intelligent robots are important for human-robot systems. This paper investigates influences of robot autonomy (i.e. high versus low) and group orientation (i.e. ingroup versus outgroup) on a human decision-making process. We conducted a laboratory experiment with 48 college students and tested the hypotheses with MANCOVA. We find that a robot with high autonomy has greater influence on human decisions than a robot with low autonomy. No significant effect is found on group orientation or on the interaction between group orientation and autonomy level. The results provide implications for social robot design. Link: https://www.hindawi.com/journals/ahci/2013/263721/	Human-Robot Interaction, Social Robots
R17	Explorations in Engagement for Humans and Robots	C. Sidner, C. Lee, C. Kidd, N. Lesh, C. Rich	2005	This paper explores the concept of engagement, the process by which individuals in an interaction start, maintain and end their perceived connection to one another. The paper reports on one aspect of engagement among human interactors—the effect of tracking faces during an interaction. It also describes the architecture of a robot that can participate in conversational, collaborative interactions with engagement gestures. Finally, the paper reports on findings of experiments with human participants who interacted with a robot when it either performed or did not perform engagement gestures. Results of the human–robot studies indicate that people become engaged with robots: they direct their attention to the robot more often in interactions where engagement gestures are present, and they find interactions more appropriate when engagement gestures are present than when they are not.	Human-Robot Interaction

				Link: https://www.sciencedirect.com/science/article/pii/S0004370205000512	
R18	"Robots and Privacy," in Robot Ethics: The Ethical and Social Implications of Robotics	Ryan Calo	05/04/2010	Robots are commonplace today in factories and on battlefields. The consumer market for robots is rapidly catching up. A worldwide survey of robots by the United Nations in 2006 revealed 3.8 million in operation, 2.9 million of which were for personal or service use. By 2007, there were 4.1 million robots working just in people's homes (Singer 2009, 7-8l; Sharkey 2008, 3). Microsoft founder Bill Gates has gone so far as to argue in an opinion piece that we are at the point now with personal robots that we were in the 1970s with personal computers, of which there are now many billion (Gates 2007). As these sophisticated machines become more prevalent—as robots leave the factory floor and battlefield and enter the public and private sphere in meaningful numbers—society will shift in unanticipated ways. This chapter explores how the mainstreaming of robots might specifically affect privacy. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1599189	Privacy
R19	Understanding Users' Perception of Privacy in Human-Robot Interaction,	Min Kyung Lee, Karen Tang, Jodi Forlizzi, Sara Kiesler	03/2011	Previous research has shown that design features that support privacy are essential for new technologies looking to gain widespread adoption. As such, privacy-sensitive design will be important for the adoption of social robots, as they could introduce new types of privacy risks to users. In this paper, we report findings from our preliminary study on users' perceptions and attitudes toward privacy in human-robot interaction, based on interviews that we conducted about a workplace social robot. Link: https://www.cs.cmu.edu/~mklee/materials/Publication/2011-HRI-robot%20privacy.pdf	Privacy, Human-Robot Interaction

R20	The Grand Challenges of *Science Robotics*	Guang-Zhong Yang et al.	01/31/2018	One of the ambitions of *Science Robotics* is to deeply root robotics research in science while developing novel robotic platforms that will enable new scientific discoveries. Of our 10 grand challenges, the first 7 represent underpinning technologies that have a wider impact on all application areas of robotics. For the next two challenges, we have included social robotics and medical robotics as application-specific areas of development to highlight the substantial societal and health impacts that they will bring. Finally, the last challenge is related to responsible innovation and how ethics and security should be carefully considered as we develop the technology further. Link: http://robotics.sciencemag.org/content/3/14/eaar7650	Research & Development, Ethics
R21	The Other Question: Can and Should Robots Have Rights?	David J. Gunkel	10/17/2017	This essay addresses the other side of the robot ethics debate, taking up and investigating the question "Can and should robots have rights?" The examination of this subject proceeds by way of three steps or movements. We begin by looking at and analyzing the form of the question itself. There is an important philosophical difference between the two modal verbs that organize the inquiry—*can* and *should*. This difference has considerable history behind it that influences what is asked about and how. Second, capitalizing on this verbal distinction, it is possible to identify four modalities concerning social robots and the question of rights. The second section will identify and critically assess these four modalities as they have been deployed and developed in the current literature. Finally, we will conclude by proposing another alternative, a way of thinking otherwise that effectively challenges the existing rules of the game and provides for other ways of theorizing moral standing that can scale to the unique challenges and opportunities that are confronted in the face of social robots. Link: https://link.springer.com/article/10.1007/s10676-017-9442-4	Ethics, Social Robots

R22	Robots: Ethical by Design	Gordana Crnkovic	03/2012	Among ethicists and engineers within robotics there is an ongoing discussion as to whether ethical robots are possible or even desirable. We answer both of these questions in the positive, based on an extensive literature study of existing arguments. Our contribution consists in bringing together and reinterpreting pieces of information from a variety of sources. One of the conclusions drawn is that artifactual morality must come in degrees and depend on the level of agency, autonomy and intelligence of the machine. Moral concerns for agents such as intelligent search machines are relatively simple, while highly intelligent and autonomous artifacts with significant impact and complex modes of agency must be equipped with more advanced ethical capabilities. Systems like cognitive robots are being developed that are expected to become part of our everyday lives in future decades. Thus, it is necessary to ensure that their behavior is adequate. In an analogy with artificial intelligence, which is the ability of a machine to perform activities that would require intelligence in humans, artificial morality is considered to be the ability of a machine to perform activities that would require morality in humans. The capacity for artificial (artifactual) morality, such as artifactual agency, artifactual responsibility, artificial intentions, artificial (synthetic) emotions, etc., come in varying degrees and depend on the type of agent. As an illustration, we address the assurance of safety in modern High Reliability Organizations through responsibility distribution. In the same way that the concept of agency is generalized in the case of artificial agents, the concept of moral agency, including responsibility, is generalized too. We propose to look at artificial moral agents as having functional responsibilities within a network of distributed responsibilities in a socio-technological system. This does not take away the responsibilities of the other stakeholders in the system, but facilitates an understanding and regulation of such networks. It should be pointed out that the process of development must assume an evolutionary form with a number of iterations because the emergent properties of	Ethics, Design

				artifacts must be tested in real world situations with agents of increasing intelligence and moral competence. We see this paper as a contribution to the macro-level Requirement Engineering through discussion and analysis of general requirements for design of ethical robots. Link: http://www.academia.edu/1006598/Robots-Ethical_by_Design	
R23	Service Robots, Care Ethics and Design	A van Wynsberghe	08/22/2016	It should not be a surprise in the near future to encounter either a personal or a professional service robot in our homes and/or our work places: according to the International Federation for Robots, there will be approx 35 million service robots at work by 2018. Given that individuals will interact and even cooperate with these service robots, their design and development demand ethical attention. With this in mind I suggest the use of an approach for incorporating ethics into the design process of robots known as Care Centered Value Sensitive Design (CCVSD). Although this approach was originally and intentionally designed for the healthcare domain, the aim of this paper is to present a preliminary study of how personal and professional service robots might also be evaluated using the CCVSD approach. The normative foundations for CCVSD come from its reliance on the care ethics tradition and in particular the use of care practices for: (1) structuring the analysis and, (2) determining the values of ethical import. To apply CCVSD outside of healthcare one must show that the robot has been integrated into a care practice. Accordingly, the practice into which the robot is to be used must be assessed and shown to meet the conditions of a care practice. By investigating the foundations of the approach I hope to show why it may be applicable for service robots and further to give examples of current robot prototypes that can and cannot be evaluated using CCVSD.	Ethics, Design, Social robots

				Link: https://link.springer.com/article/10.1007/s10676-016-9409-x	
R24	Regulatory Challenges of Robotics: Some Guidelines for Addressing Legal and Ethical Issues	Ronald Leenes et al.	03/07/2017	Robots are slowly, but certainly, entering people's professional and private lives. They require the attention of regulators due to the challenges they present to existing legal frameworks and the new legal and ethical questions they raise. This paper discusses four major regulatory dilemmas in the field of robotics: how to keep up with technological advances; how to strike a balance between stimulating innovation and the protection of fundamental rights and values; whether to affirm prevalent social norms or nudge social norms in a different direction; and, how to balance effectiveness versus legitimacy in techno-regulation. The four dilemmas are each treated in the context of a particular modality of regulation: law, market, social norms, and technology as a regulatory tool; and for each, we focus on particular topics – such as liability, privacy, and autonomy – that often feature as the major issues requiring regulatory attention. The paper then highlights the role and potential of the European framework of rights and values, responsible research and innovation, smart regulation and soft law as means of dealing with the dilemmas. Link: http://www.tandfonline.com/doi/full/10.1080/17579961.2017.1304921	Regulation, Privacy, Research & Development
R25	A Spotlight on Security and Privacy Risks with Future Household Robots: Attacks and Lessons	Tamara Denning et al.	2009	Future homes will be populated with large numbers of robots with diverse functionalities, ranging from chore robots to elder care robots to entertainment robots. While household robots will offer numerous benefits, they also have the potential to introduce new security and privacy vulnerabilities into the home. Our research consists of three parts. First, to serve as a foundation for our study, we experimentally analyze three of today's household robots for security and privacy vulnerabilities: the WowWee Rovio, the Erector Spykee, and the WowWee RoboSapien V2. Second, we synthesize the results of our experimental	Social Robots, Privacy, Security, Design

				analyses and identify key lessons and challenges for securing future household robots. Finally, we use our experiments and lessons learned to construct a set of design questions aimed at facilitating the future development of household robots that are secure and preserve their users' privacy. Link: https://dl.acm.org/citation.cfm?id=1620564	
R26	Regulating Healthcare Robots: Maximizing Opportunities While Minimizing Risk	Drew Simshaw, Nicolas Terry, Dr. Kris Hauser, Dr. M.L. Cummings	02/24/2016	This paper will focus on the issues of patient and user safety, security, and privacy, and specifically the effect of medical device regulation and data protection laws on robots in healthcare. First, it will examine the demand for robots in healthcare and assess the benefits that robots can provide. Second, it will look at the types of robots currently being used in healthcare, anticipate future innovation, and identify the key characteristics of these robots that will present regulatory issues. Third, it will examine the current regulatory framework within which these robots will operate, focusing on medical device regulation and data protection laws. 22 Rich. J.L. & Tech. 3 (In Dropbox)	Privacy, Regulation, Data Protection, Social Robots
R27	Robots in the Home: What Have We Agreed To?	Margot Kaminski	2015	This essay begins by identifying the legally salient features of home robots: the aspects of home robots that will likely drive the most interesting legal questions. It then explores how current privacy law governing both law enforcement and private parties addresses a number of questions raised by home robots. First, how does privacy law treat entities that enter places (physically, or through sense-enhancing technologies) where they are not invited? Second, how does privacy law treat entities that are invited into a physical space, but were not invited to record in that space? How does privacy law treat consent, both express and implied? Fourth, how does privacy law address entities that lull--or deceive--people into revealing more than they intend to? And finally, in the private actor context, will robotic recording be considered to be speech? 51 Idaho L. Rev. 661 (In Dropbox)	Privacy, Law Enforcement

R28	Social Robotics: Proceedings of the Third International Conference, ISCR 2011	Bilge Mutlu, Christoph Bartneck, Jaap Ham, Vanessa Evers, Takayuki Kanda (Eds.)	2011	In this volume, you will find the papers presented at the Third International Conference on Social Robotics (ICSR), held during November 24–25, 2011, in Amsterdam, The Netherlands. In the new and rapidly growing and evolving research area of social robotics, building a community that is collegial, supportive, and constructive is crucial for fostering the scientific qualities needed to answer the questions that this strongly interdisciplinary field poses. The diversity of backgrounds and the sheer number of the Chairs involved in organizing this conference characterizes that enterprise. Likewise, the diversity of the papers in these proceedings and of the research discussed at the conference is an indication of the growing interest in social robotics research from a multitude of perspectives.	Social Robots, Human-Robot Interaction