The Heart of the Matter: Patient Autonomy as a Model for the Wellbeing of Technology Users

Emanuelle Burton
Dept of Computer Science
University of Illinois at Chicago

Kristel Clayville
Philosophy and Religion Dept
Eureka College

Judy Goldsmith
Dept of Computer Science
University of Kentucky

Nicholas Mattei
TJ Watson Research Center
IBM

Abstract

We draw on concepts in medical ethics to consider how computer science, and AI in particular, can develop critical tools for thinking concretely about technology's impact on the wellbeing of the people who use it. We focus on patient autonomy—the ability to set the terms of ones encounter with medicine—and on the mediating concepts of informed consent and decisional capacity, which enable doctors to honor patients' autonomy in messy and non-ideal circumstances. This comparative study is organized around a fictional case study of a heart patient with cardiac implants. Using this case study, we identify points of overlap and of difference between medical ethics and technology ethics, and leverage a discussion of that intertwined scenario to offer initial practical suggestions about how we can adapt the concepts of decisional capacity and informed consent to the discussion of technology design.

Introduction

Machines will be making life-and-death decisions for individuals in the near future, as well as decisions that have a profound impact on the quality of human lives. Not only will they drive vehicles and deliver aid, they may triage disaster victim rescues and hospital admissions, they will control thermostats, schedule emergency services, help farmers predict weather and timing of growing seasons, work with food processing plants' supply chains, adjudicate insurance and parole claims, and decide who has access to emergency shelters in the wake of natural disasters. In ways both large and small, current and in-development applications of AI are altering the basic conditions of ordinary human experience, from the imminent availability of self-driving cars to robot companions for the elderly (Sabanovic et al. 2013) or the robophilic (Danaher and McArthur 2017).

All of these AI-driven decisions are necessarily predicated on comparative value judgments about human worth and human goods: the importance of children's lives vs. seniors' lives in a natural disaster, or the value of students' security vs. their personal dignity at a high-risk high school, or the appropriate course of medical care for a terminally ill patient who is physically and emotionally suffering. These are the same value judgments that transplant teams make every time they prepare to operate. Whether those values are predetermined by developers or companies or are learned by example through machine learning algorithms, these mechanized decisions—and the substrate of comparative values that structure the automated decision-makers—will have a profound impact on people's lives and wellbeing.

But what exactly makes a human life valuable and distinctive? What qualities of internal self or external environment need to be in place for a person to be able to live and act as a person? How do particular changes to their environment enhance, or circumscribe, their ability to be a version of themselves that they recognize and prefer? For most technologists who understand their work as a way to improve human lives, the importance of those lives and the reasons why they matter have been largely a product of moral intuition rather than of carefully-defined principles. Such intuitions are difficult to formalize in a way that can be programed directly or entrusted to an algorithm to learn by example, particularly in the absence of a conceptual language that can identify them or draw distinctions between them.

The proliferation of AI in daily life makes it ever more vital and pressing that technologists can think specifically about those aspects of the person that make them recognizable and distinct as people, and furthermore how those human qualities are amenable to improvement, or vulnerable to harm, through specific changes in the conditions of daily life (Burton, Goldsmith, and Mattei 2015; Burton et al. 2017). Furthermore, *it is imperative that AI ethics develop its own conceptual tools that can account for the particular ways in which AI can impact those conditions of daily life.* So equipped, technologists will be able to discuss the parameters and significance of the interventions that their designs are making, and to think more concretely about how design and programming choices can protect and enhance the lives of individuals and societies.

This aim—to enhance, rather than diminish, human lives through technology—is made particularly difficult by the knowledge gap between those who build and maintain the technologies and those who use them without the technological expertise to understand how they work. Such non-specialists users face several disadvantages, even with respect to technologies and platforms that are designed for non-specialist use such as a smartphone or Twitter. Not only are these users less likely to be aware of potential security

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

breaches, the signs of such breaches, or the steps they might take to prevent them; they are also far less likely to be aware of any modifications that would enable these users to fine-tune their experience for their own personal comfort and convenience. Thus, *even at the level of everyday personal technology use, there exists a significant power imbalance between technology experts and non-experts*. The depth and scope of that power imbalance grows exponentially if one also considers those experts' professional work designing, building and maintaining the systems that other users rely on but lack the expertise to understand.

This expertise-based power imbalance, while particularly pressing in technology ethics (and perhaps AI in particular), is not unique. A similar power imbalance has long existed in medicine, a field whose practitioners need extensive specialist knowledge even as they serve a user base of patients who mostly lack that knowledge. Because of the power imbalance implicit in the vast majority of patient-practitioner relationships, patients are often prevented from making choices about their own care even when doctors or nurses are at pains to leave the choice in the patient's hands (Henderson 2003). *To mitigate this problem, medical ethics has developed a family of concepts and practices to help its expert practitioners to navigate the inevitable imbalance in power and knowledge* (Quill and Brody 1996). As this paper will demonstrate, these concepts can be usefully imported (with some significant revision) into technology ethics (Johnson 2009), and can be used to identify specific technology design practices that preserve non-expert technology users' capacity for self-determination.

Contribution. In this paper, we described the concept of patient autonomy from medical ethics, as well as the corollary concepts of informed consent and decisional capacity. We use a fictional case study to highlight the both the points of intersection and points of divergence between traditional medical ethics concerns and technology ethics concerns. We then, on the basis of case study discussion, develop working definitions of informed consent and decisional capacity that are attuned to the central problems facing technology ethics. Finally, on the strength of these newly-adapted concepts, we will offer some concrete examples of how current projects in AI and technology are working to support human autonomy, or how they could be adapted to better support it.

Autonomy in Medical Ethics

Most western medical practitioners would identify **autonomy** as the central tenet of medical ethics. Autonomy is the principle that mandates **respect for persons**, meaning that individuals have free exercise with regard to whether and what kind of treatment to receive, and honoring this independence is central to contemporary medical ethics (Jonsen, Siegler, and Winslade 2015) . Patient autonomy as a governing concept in medical ethics is relatively recent; the shift toward it and away from medical paternalism was fueled both by broader social movements that sought to empower the individual and by the development of a more consumerist model of medicine as physicians sought to protect themselves from malpractice (Billings and Krakauer 2011).

In practical medical ethics, the term autonomy has two distinct uses, which are related but which also operate independently of each other. The first usage is to affirm that the patient deserves autonomy, the power to exert influence over what happens to them; the second usage concerns the question of whether the patient is able to exercise that autonomy. Because people frequently seek medical care at a moment when they are mentally and physically compromised, it is not enough to affirm that a patient deserves autonomy. It is necessary for medical providers to take deliberate steps in order to protect the patient's autonomy, and ensure that the patient is able and empowered to make decisions that reflect their wishes.

Neither dimension of autonomy—autonomy-as-recognition or autonomy-as-exercise—simply exists as a given. Because of the systemic power imbalance between expert care providers and their nonexpert patients, two important constraints have been put in place to ensure that the patient's autonomy is honored not only in principle but in practice. They are **informed consent** and **decisional capacity**. In the United states, when a patient undergoes a medical procedure, that patient must consent to it, and that consent must follow a conversation in which the doctor explains the procedure's risks, benefits to the patient, as well as other treatment options. After this conversation has happened, the patient signs a document acknowledging that this conversation took place, and the patient is thereby giving *informed* consent to the procedure. Because informed consent documents a conversation, *it is approached as a process rather than a one-time event*. Patients can change their minds at any point leading up to or during the procedure.

No medical procedures or treatments should be undertaken without informed consent, but only patients who have decisional capacity can give informed consent. In general, adult patients are presumed to have decisional capacity, but there are categories of patients who lack it. Patients can lack decisional capacity due to age (children), medical status (dementia patients), temporary states (sedated), or institutional status (prisoners). But this absence of decisional capacity is not permanent; children will age into being decisional and able to give informed consent, sedated patients will wake up, and prisoners may be freed, thus enabling them to make decisions free of coercion.

Paradoxically—or so it seems at first—these limits on a patient's decision-making were instituted precisely to preserve the patient's autonomy, because they place limits on a doctor's ability to manipulate patients into undergoing treatments. The constraints were developed in response to abuses of paternalism, and were designed to constrict doctors' freedom by preventing them from taking advantage of patients who were, for whatever reason, unable to exercise their own autonomy.

As medical culture has evolved toward being more patient-centered, the language and conceptual framework of autonomy have likewise been enhanced to focus more on how patients can exercise autonomy, rather than on the constriction of the doctor's. Patients can, in fact, prepare for a future in which they are non-decisional, by creating legal

documents that spell out their wishes, should they be incapacitated. They can also cede decision-making power to specified others, for such an eventuality. In the absence of such explicit and legally binding instructions, it is assumed in most societies that a surrogate decision maker from the family can speak for the patient's wishes.

As we will argue, the concept of patient autonomy—and its concepts of informed consent and decisional capacity—offer a useful model for technology ethics in thinking about how to preserve and enhance the wellbeing of technology users. As the above discussion illustrates, however, the core problems in medicine are not identical to those in technology. *In order for these imported concepts of autonomy, decisional capacity, and informed consent to be useful to technology ethics, they need to be adapted, but in a way that preserves the element that makes them useful*. We use the following fictional case study to illustrate points of overlap and divergence.

Case Study

Consider a heart patient, Joe, who has two implants to help with his heart: a pacemaker, which regulates his heartbeat, and an implantable cardioverter defibrillator (ICD), which can restart his heart if it stops. This is a common case in the US with over 947 heart related implants per million people (Mond and Proclemer 2011). Some years ago, in consultation with his doctor (as is legally required), Joe requested and was granted Do Not Resuscitate (DNR) status. At a recent doctor's visit, Joe was told that restarting his heart would be intensely painful, and that in such an event, his heart would likely fail and need to be restarted repeatedly. Given his DNR status, Joe's doctor asked whether Joe wants the ICD turned off.

Joe's case raises a set of questions that are common to many medical ethics case studies, most of which center around autonomy.

1. Does Joe have the right to make these decisions? If he is in pain, can his judgment be trusted?
2. Do Joe's previous decisions express a state of mind that is still binding for the present?
3. For Joe's doctor, is there a meaningful difference between Joe refusing aggressive CPR (an external treatment) and refusing an ICD?
4. For Joe, is there a meaningful difference between refusing an ICD and turning off an ICD that is already implanted?
5. For both Joe and his doctor, would turning off the ICD be comparable to euthanasia?

The framing of these questions presumes the concept of autonomy: that Joe deserves the right to determine what happens to him, and that this right to self-determination must be preserved in balance with medicine's broad imperative to preserve and extend life whenever possible. Joe's right to refuse treatment is recognized, but so is the fact that the very conditions of his treatment may mean that he is not decisional, and thus not fit to make decisions that may harm his person.

But as technologists and those thinking about technology ethics will immediately recognize, this slate of questions excludes some important issues, including issues that might be understood in terms of autonomy. Other questions should be raised pertaining to the security of Joe's personal information and self-direction that are directly influenced by the specific technologies that are now part of his body.

1. Who is responsible for implanting and maintaining Joe's machines?
2. What risks are there to Joe in having his cardiac data possibly transmitted by WiFi and stored online?
3. What risks are there in allowing an off-site monitor to control the pacemaker?
4. Should any of the defibrillator itself, a control system, or a human monitor be able to decide to not resuscitate Joe?

Like the medically-oriented questions, these technological questions also recognizably concern Joe's autonomy as a patient/technology user. The underlying premises of the technology ethicist's questions recognize Joe as an entity deserving of the same sort of autonomy accorded to him by the medical ethics list. But there are two key differences between them. The first is that these questions expand the sphere of Joe's autonomy (in the autonomy-as-recognition sense) to include concerns about his personal information and to consider a wider range of possible agents who might impact Joe's wellbeing. The second difference is that, while these questions broaden the scope of Joe's autonomy as something for professionals to worry about, they constrict its actual exercise by the patient himself (in the autonomy-as-exercise sense). In focusing—appropriately and necessarily—on systems-level concerns such as information security and encryption of medical data, *these questions leave little room for Joe's ability to make decisions for himself, or even to understand what is at stake in the decisions he might make.* Although the questions are about the sphere of Joe's autonomy, they do not create or identify an opportunity for him to exercise it.

The contrast between these sets of questions highlights both how medical ethics could refine its notion of autonomy in conversation with technology ethics, and how technology ethics stands to benefit from an imported version of autonomy from medical ethics. With respect to the first dimension of autonomy—recognizing what the patient deserves as a person—technology ethics usefully broadens the sphere of Joe's autonomy insofar as it broadens the scope of things in the world that are not only *his* but *him:* his pacemaker and defibrillator, perhaps even his data. In an age when medicine is increasingly reliant on networked technology and data, medical ethics would do well to learn from technology ethics' reconfiguration of autonomy.

Yet technology ethics is less well equipped than medical ethics to attend to the second aspect of autonomy, the patient's right to determine what happens to him. A concern for Joe's right to exercise his own particular preferences might lead to questions such as the following: Does Joe understand the capabilities and risks (either to his body or his data) of the devices that have been implanted within him, to a degree

that he can make an informed decision about them? Is he aware of the experiences of other patients with similar implantations? Does he feel able to ask his doctors to shut off the implanted devices, to opt out after opting in?

We argue that these are the sorts of questions technologists need to be asking, i n particular, the designers of AI technologies that can manage the content of a user's online experience or automatically transmit sensitive medical data to doctors. Because technology is necessarily systems-oriented in its approach, the challenges in making room for users' autonomy-as-self-direction are different—and, arguably, even more difficult to overcome—than those in medicine. Therefore, it is not helpful for technology ethics to simply adopt the concept of autonomy from medical ethics unmodified. And yet, if the human wellbeing of technology users—technology ethics' equivalent of patients—is not to fade from view, it is crucial to identify and clarify a notion of autonomy that technologists *can* use, a definition that is analogous to that in medical ethics but more closely keyed to the problems faced in technology ethics. As technology increasingly sets the conditions for human life, not only in medicine but in the public and private sphere, this sort of working definition will prove crucial for technologists who wish to preserve a space for the exercise of autonomy.

Reframing Autonomy for Technology

As our case study indicates, the notion of patient/user autonomy is relevant for technology as well as for medicine, even though the precise contours are different. As human lives are increasingly managed at both macro- and micro-level by smart technologies —and as medical technology itself advances—it becomes pressing for technologists to consider how to enhance (or at least to preserve) users' autonomy. To do so, technologists must consider not only users' right to make decisions for themselves (the first aspect of autonomy), but the conditions that enable them to exercise that autonomy (the second aspect).

In addition, technology ethics also faces some particular hurdles in incorporating user autonomy into existing frameworks of inquiry. As is seen when we compare the two sets of questions in our case study, the very nature of technological work is already an impediment to conceiving of persons in terms that recognize and extend their ability to exercise their autonomy. These hurdles are particularly difficult to overcome in the case of AI, which outsources both large- and small-scale decision-making to programmed learners—and sometimes in ways that are designed to "solve" the idiosyncrasies of users' exercise of their self-directing autonomy (Rapoport 2013).

A further challenge to technology ethics is that there is rarely an appointed human mediator between the user and the technological establishment as there is in medicine. Medical ethics is structured around the relationship between patient and care provider, and this can invest the individual care provider with particular duties and responsibilities. Any useful adaptation of patient/user autonomy needs to assign responsibility in a manner that is both ethically and practicably plausible.

The concept of user autonomy can be rendered more manageable when we approach it by way of of informed consent and decisional capacity. As discussed above, these two concepts were developed in medical ethics as a means to preserve the patient's autonomy when her capacity to exercise that autonomy is in some way compromised. Informed consent and decisional capacity function essentially as "sluice gates" to make sure that the patient/user's autonomy is maintained even in the presence of disruptive or distorting factors.

Informed Consent

In a medical context, *informed consent* helps to preserve the patient/user's autonomy by requiring the doctor to keep the patient apprised of relevant information, and permitting the patient to rescind consent at any point. Informed consent presumes a user who never develops expertise of her own, and is not penalized for it; the burden remains on the expert-provider to communicate clearly and consistently with the user, to ensure she understands and that her wishes are being honored. While this is not the norm in technology we are starting to see ideas like this appear. For example, the Android operating system's reliance on *permissions* for apps which can be granted or revoked from an easy to find screen (Andriotis, Sasse, and Stringhini 2016).

Informed consent presents deep challenges to the basic design principles of technology, because it is deliberately inefficient and resistant to closure. First, it prioritizes certainty that the patient/user understands over the efficient delivery of information. Second, by allowing the patient/user to opt out at any point, it mandates a structure in which processes are begun but never completed, both because patient/users sometimes withdraw consent partway and because even consenting patients/users retain the option to withdraw consent.

But this inefficiency is absolutely vital if the patient/user's autonomy is to be preserved. Because efficiency requires that certain decisions or functions take place en masse for a group of entities without stopping to consult each one, some kinds of efficiency cannot coexist with informed consent. The smarter and more seamless a technology becomes, the more deliberate the technology designer has to be about maintaining space within it for this sort of inefficiency. For example, a massive push update to a high-tech medical implant will be much easier to accomplish if the manufacturers assume that the patient/users have already consented simply by having the device implanted. If, however, a patient's condition or wishes have changed, she might not want her implant to be updated.

It is important to note that not all kinds of efficiency are necessarily at odds with informed consent. Many forms of automation increase the efficiency with which the user's goals are achieved without eclipsing her ability to revise her goals or judgments. There is no need for a given technology to build in opportunities for ongoing consent when that technology executes tasks the users already understand and intend to perform, such as washing dishes or taking depth or temperature measurements.

Whenever technological efficiency is achieved by eliminating the need for the user's input, there is a real risk that the user's autonomy could be compromised. Any technol-

ogy that makes decisions for its users—even when those decisions are based on prior expressions of consent or preference—is one that has the potential to violate users' autonomy. Although the efficiency of self-monitoring thermostats and smart surveillance technologies is one of their main selling points to users, that very ease of use is what makes it possible for those users' autonomy to be compromised, when their personal data is transmitted in a manner they are not comfortable with or their home monitoring systems do something they dislike. Indeed, for a device or platform to incorporate informed consent in a meaningful way that it must preserve some kinds of inefficiency. The fact that this notion may present a challenge to the normal way technology developers think underscores the need for a concrete concept by which technology ethicists can assert why it is necessary to constrict some kinds of efficiency in order to preserve or enhance users' wellbeing.

In considering what kinds of inefficiency are important for maintaining informed consent, it is helpful to look back to the original concept in medical ethics. In medicine, the deliberate inefficiency of informed consent affords the patient the time to consider (and reconsider) her options in terms of her values and goals. It also forces the care provider to support the patient in this process, rather than imposing decisions upon her. Because the patient's goals or preferences might shift over time or due to changes in her circumstances, the efficient option—taking the patient's initial goals and decisions as a presumptive guide to the future—would undermine her autonomy. Such changes in goals or preferences can be understood as "human" inefficiencies: inefficient or unpredictable movements of character or goals that are essential to a person's autonomy and crucial to preserving their wellbeing. In a medical context, informed consent protects the patient's autonomy by preserving ongoing ability to express her preferences, even when it renders her overall program of care more inefficient. The efficiency of the treatment process is valuable as long as it preserves or enhances the autonomy of the patient/user, and is potentially damaging to her autonomy insofar as it imposes efficiency on the messy and inefficient processes of self-determination.

Therefore, a usable concept of informed consent for technology ethics is one that enables technologists to consider the specific ways in which a given technology creates efficiency. Does it smooth the user's path to a goal she understands and wants? Does it equip her to understand which sort of determinations are being made for her by automated processes, and to single out the determinations that matter to her for further scrutiny and input? Does it create space for her to revise her engagement with it, should her goals or preferences change? With such questions in mind, a technologist is better prepared to evaluate which kinds of efficiency might categorically interfere with a user's autonomy, which ones require ongoing user input of some kind, and which functions can best serve the user in silent efficiency.

Decisional Capacity

Like informed consent, the notion of *decisional capacity*—the recognition that autonomous users are sometimes not in a state to exercise their own autonomy—can be adapted to technology ethics as a means to preserve and enhance user autonomy. As noted above, medical doctors use a range of criteria to determine whether a patient is decisional, but those criteria have two common denominators: they expect the decisional patient/user to make choices in a manner consistent with their previous character and preferences, and they expect any departures from that prior consistency to be "reasonable"—that is, in line with socially-determined ideas.

Decisional capacity in medical settings is typically binary in nature, because the patient/user's role in the relevant medical process is widely understood to be one of consent, rather than execution. (See, for instance, (Jeste, Palmer, and et al. 2007).) If heart patient Joe decides that he wants his ICD turned off, his decisional capacity depends only on whether he is currently capable of making the decision: a medical expert (either Joe's doctor or an ICD specialist) will implement the decision. Joe will be the one to live with the consequences of his choice—which is why he must be decisional in order to make the choice—but his capacity to execute that decision is not a relevant factor. if Joe's judgment is sufficiently consistent with himself, and/or with what is "reasonable," to make what his doctor deems to be a clear-headed decision, then his decision is medically legitimate.

Technology complicates this notion of decisionality because, in most cases, users are also in charge of implementing their decisions. While technology use is not (usually) as complicated as a surgical procedure, some binding End User Agreements (EULAs) are. Additionally, it can require some deftness of body and mind to manipulate a device with an injured hand, or to craft a rejoinder tweet while in a state of righteous outrage that will not cause regret in an hour. Like medicine, technology is a sphere that can magnify the consequences of a given decision; but unlike medicine, technology empowers users to act *without* the mediation of an expert practitioner who can clarify the scope or the stakes of the user's action.

In most cases, the fact that technology extends the scope of its users' ability to act is the primary virtue. The fact that users are able to take these actions instantly, or near-instantly, is further evidence of the quality of a piece of technology. But these same qualities make users particularly vulnerable to undertaking actions whose technologically-augmented scope exceeds the user's capacity to assess the consequences in the moment of decision. It therefore seems not only helpful but necessary to adapt the notion of decisional capacity for use in technology ethics.

In order to be optimally useful for technology ethics, the notion of decisional capacity needs to be expanded to account for the user's role in implementing their own decisions. It can be helpfully recast for technology ethics as *decisional-executive capacity*, incorporating a second layer that raises the question of whether the user is fit, in a given moment, to undertake an action in a manner that they will be happy with later. Examples of at least checking for this include automatic tone alerts for angry emails and Slack warnings before a message is sent to everyone at the workplace.

Decisional capacity creates an opportunity for AI to enhance the autonomy of technology users and medical pa-

tients. As noted above, decisional capacity is quite imperfectly realized in a medical context, as doctors are far more likely to deem a patient decisional if the patient agrees with them. An AI, however, is less likely to succumb to this bias (Hurst 2004). While a doctor's ingrained biases can compromise her assessment of whether her patient is decisional, the doctor-patient relationship is nonetheless a useful model for the AI-user relationship in one key respect. While consistency (the first criterion for determining decisionality) is best judged only with respect to the patient himself, the reasonableness of his wishes (the second criterion) is more broadly culturally determined; what seems like a good reason in one society may seem bizarre in another. Because the human doctor will be influenced by the same broad cultural norms, she is well-positioned to assess whether the patient's expressed wishes fit within those cultural norms, though she is also less likely to be sympathetic to reasons that do not fit those norms. In contrast, an AI that determines decisionality could be structured on universal terms, the ideal approach might call for an AI to learn primarily from local data in order to better assess the reasonableness of expressed wishes.

Conclusion

There is a pervasive societal disease about artificial intelligence that ranges from fears of loss of jobs for humans to terror that we will be displaced entirely by self-aware, higher-functioning AIs. One strain of this anxiety is that the machines will be programmed with more concern for efficiency than for the wellbeing of the humans they are designed to serve. But these things are not determined yet. What is necessary to balance the drive toward efficiency is a focus on how AI can support the distinctively human qualities of its users. We believe that engineers and computer scientists can learn from medical ethicists, and provide a vital viewpoint to the field of medical ethics itself. Through this and broader communication throughout the industries and domains where AI is applied will ensure that AI can live up to the potential envisioned by its boosters, and become a vital part of the architecture of a better human future.

References

Andriotis, P.; Sasse, M. A.; and Stringhini, G. 2016. Permissions snapshots: Assessing users' adaptation to the android runtime permission model. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, 1–6.

Billings, J. A., and Krakauer, E. L. 2011. On patient autonomy and physician responsibility in end-of-life care. *Arch Intern Med* 171(9):849–853.

Burton, E.; Goldsmith, J.; Koenig, S.; Kuipers, B.; Mattei, N.; and Walsh, T. 2017. Ethical considerations in artificial intelligence courses. *AI Magazine* Summer.

Burton, E.; Goldsmith, J.; and Mattei, N. 2015. Teaching AI ethics using science fiction. In *1st International Workshop on AI, Ethics and Society, Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Danaher, J., and McArthur, N., eds. 2017. MIT Press.

Henderson, S. 2003. Power imbalance between nurses and patients: a potential inhibitor of partnership in care. *Journal of Clinical Nursing* 12(4):501–508.

Hurst, S. A. 2004. When patients refuse assessment of decision-making capacity: How should clinicians respond? *ARCH Internal Medicine* 164(16):1757–1760.

Jeste, D.; Palmer, B.; and et al., P. A. 2007. A new brief instrument for assessing decisional capacity for clinical research. *Archives of General Psychiatry* 64(8):966–974.

Johnson, D. G. 2009. *Computer Ethics*. Pearson, 4th edition.

Jonsen, A. R.; Siegler, M.; and Winslade, W. J. 2015. *Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical Medicine. 8th ed.* McGraw Hill Education.

Mond, H. G., and Proclemer, A. 2011. The 11th World Survey of Cardiac Pacing and Implantable Cardioverter-Defibrillators: Calendar Year 2009–A World Society of Arrhythmia's Project. *Pacing and Clinical Electrophysiology* 34(8):1013–1027.

Quill, T., and Brody, H. 1996. Physician recommendations and patient autonomy: Finding a balance between physician power and patient choice. *Annals of Internal Medicine* 125(9):763–769.

Rapoport, M. 2013. Being a body or having one: automated domestic technologies and corporeality. *AI & Society* 28(2):209–218.

Sabanovic, S.; Bennett, C. C.; Chang, W.-L.; and Huber, L. 2013. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*, 1–6. IEEE.