Privacy in Human-Robot Interaction: Survey and Future Work

Matthew Rueben and William D. Smart

Robotics Group, School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University

****** Draft for We Robot 2016 – Please do not cite without permission ******

Robots can cause privacy concerns for humans, partly because robots can gather data and move around in the real world, and partly because they can be perceived by humans as social actors. This is a growing issue as robots are increasingly located near humans. We propose a new area of research—"privacy-sensitive robotics"—to deal with this issue. The first part of this paper lays a foundation for new research with several broad literature surveys: of privacy in philosophy, U.S. law, the social sciences, medicine, and technology; of existing robotics technologies that could be used to protect privacy; and of existing work in "privacy-sensitive robotics" as defined herein. The second part presents a roadmap of future research directions by which to develop this new sub-field.

Keywords: Privacy, social robots, robot ethics, human values, robot law, robot policy, social psychology

1. Introduction

Roboticists usually conceive of robot behavior as a list of "do"s: do pick up the cup, do go to the other room, do locate the red object. Along with each "do," or *goal*, however, comes a host of "do not"s, or *constraints*. Do not drop or crush the cup; do not hit anyone on your way to the door; do not stare at people or they'll feel uncomfortable. Our focus on "do"s—we call this *positive* design—is effective insofar as the incumbent "do not"s are clearly and strongly apparent in each "do." For example, it is obvious and urgent that a robot should not hit walls during a navigation task. Some particular "do not"s, however, demand to be examined by themselves. These include safety, privacy, politeness (e.g., Takayama et al., 2009), transparency (e.g., Takayama et al., 2011)[1], and honesty[2]. These "do not"s are likely to be overlooked by researchers because only a few current applications need them clearly, specifically, and urgently enough to justify the extra effort. This research gap is a problem because we believe these "do not"s will become much more important and relevant as robotics technology progresses. We call the specific focus on *constraining* robot behavior a sort of *negative* design: we focus on telling the robot what *not* to do.

This paper will focus on negative design for privacy-sensitive robots—that is, robots proficient in respecting human privacy expectations. Underpinning the entire paper are two human-robot interaction paradigms that we believe will be prevalent in the future. First is autonomous robot behavior, which especially intersects privacy to the degree that robots are anthropomorphized—that is, thought

[1] By *transparency* we mean an agreement between what the robot expresses externally and what is going on inside the robot.

[2] We would say a robot is *honest* if it does not try to deceive people with its speech, actions, or even with its appearance. One could perhaps say that a robot that looks and acts like a human is actually being *dishonest*; see Section 6.1.7.

of as human. Second is the possibility of remote human operators using robots as body surrogates, which we can call "robot-mediated presence." This also has privacy implications, since the remote operator can drive around and manipulate things in the local user's space. Furthermore, the mediation itself could have unforeseen social effects.

In the next section we present some background on both of these interaction paradigms to show why they are important and related to privacy. Section 3 reviews privacy research in a multitude of relevant fields, and Section 4 reviews techniques for building privacy-sensitive robots. Section 5 presents the research that we consider to fall within privacy-sensitive robotics. Section 6 discusses the future of privacy-sensitive robotics as a budding field of research including several proposed research directions. Section 7 summarizes this paper, which we intend to serve as the first motivating research survey for the emerging field of privacy-sensitive robotics.

2. Background

Privacy-sensitive robotics can be thought of as a subset of human-robot interaction. See Goodrich & Schultz (2007) for a survey of human-robot interaction and Fong et al. (2003) for a survey of socially-interactive robots. The focus in both of these surveys is on autonomous robot behaviors, although in some cases, autonomy is shared between the human and the robot.

How can a robot threaten somebody's personal privacy? This section examines this question in terms of two human-robot interaction paradigms: autonomous robot behavior and robot-mediated presence.

2.1 Autonomous Robots and Privacy

Why do autonomous robots pose a privacy concern for humans? One could object that privacy is fundamentally a problem between humans. Animals, much less inanimate objects, do not invade upon our privacy...do they? But some robots can record and transmit data in a human-readable format; animals cannot. Can robots also violate other kinds of privacy, such as personal space, territoriality, and solitude (defined below in Section 3.1.3)? They can, but only if people can *anthropomorphize* robots by attributing human characteristics to them. If this seems unlikely, the authors of *The Media Equation* would beg to differ (Reeves & Nass, 1996). In their 1996 book, Byron Reeves and Clifford Nass present results from 35 studies wherein images, videos, and computers are substituted for real people. Their findings are summarized in what they call "the media equation:"

$$Mediated\ Life = Real\ Life$$

The results are compelling. For example, one study was about interpersonal distance. Proximity between humans causes psychological arousal; it makes the interaction feel more intense. In the study, images of faces were substituted for real people, and image size was used to manipulate perceived closeness. The same effects were observed for the images as for real people. This directly relates to Remote Presence Systems (RPS). If the robot brings the screen with the remote user's picture closer (or makes it bigger), we predict that local users will react the same as if a real person came closer.

Another study examined whether people can think of devices as "specialists" in the same way as they do for people. In one condition, a single "generalist" television set showed both news and entertainment shows. In the other condition, two "specialist" televisions were used, one labeled "news" and the other "entertainment." Not only did people praise the programs more when viewing them on the "specialist" televisions, but they also denied that the labels had any effect on their judgments. This can also be applied to robots. Some robots are in fact designed for a single purpose, and others are more general-purpose. If Reeves and Nass are right, then people could potentially trust specialist robots more and expect better results from them, a typically human-human phenomenon.

Think about the last scary movie you watched. The scenes can be quite compelling, and you have to remind yourself that it's not real: "it's only a movie, it's only a movie." Even though you don't believe it's real, however, you respond in a natural and social way with fear and anxiety. It can be very hard to pull yourself out of the appearances and back to reality. This is the gist of the findings presented by Reeves & Nass (1996).

This is often summarized with the statement, "Computers Are Social Actors" (CASA). We now claim that autonomous robots may also be perceived as social actors and therefore participate in concepts that typically exist only between people, such as privacy. So in response to our initial question of whether autonomous robots can pose a privacy concern for humans, we answer "yes," they can.

The degree to which a robot is seen as a social actor can vary. HRI researchers have designed several studies to measure whether a particular robot is seen as a social actor in a particular situation. They operationalize social actorhood in several different ways, including the social facilitation effect (Schermerhorn et al., 2008), responses to a robot's greeting and other linguistic factors (K. Fischer, 2011), cheating deterrence (Hoffman et al., 2015), and the act of keeping a robot's secret (Kahn et al., 2015). Other HRI studies have also tapped constructs related to CASA, such as anthropomorphism (e.g., Bartneck et al., 2009) and theory of mind (e.g., Levin et al., 2008). Several such constructs might relate to privacy in HRI, and studying this connection clearly requires the expertise and close involvement of social scientists.

2.2 Robot-Mediated Presence and Privacy

In the robot-mediated presence paradigm, personification is natural because the robot really is (representing) a person. One simple form of robot-mediated presence is that of Remote Presence Systems (RPSs, already mentioned above). These robots present a video chat interface with the addition of a mobile base so the remote operator can look around and even drive from place to place. One can see how this is very much a remote presence, as one could even participate in a tour or meet-and-greet from afar. Examples of RPSs include the InTouch Telemedicine robots (InTouch Health, 2012), the Beam RPS (Suitable Technologies, 2012), and the VGo robot (VGo Communications, 2012). See Lee & Takayama (2011) and Beer & Takayama (2011) for two evaluations of RPSs in realistic scenarios. Robots could also be mediators, or even full-body surrogates, for persons with severe motor disabilities (see Chen et al., 2013).

It seems clear that remote presence systems, like autonomous robots, cause concerns about privacy. Without the mobile base, RPSs are essentially video media spaces, which have a slew of privacy problems themselves (see Boyle et al., 2009, for a review; this is discussed below in Section 3.6). Moreover, adding the mobile base adds new privacy concerns. RPSs can be driven into private spaces, or used to look around at things against the will of the local user(s). With video chat software like Skype, the local user controls the direction of the webcam; with RPSs, the remote operator has this control.

3. Privacy

There is a lot of literature on privacy. The conversation spans many disciplines and is difficult to summarize concisely. We have organized this survey of privacy beginning with abstract ideas and proceeding towards concrete applications in the real world; it is outlined as follows. First, much ink has been spilled over the *definition* of privacy. In particular, what specific concerns are included under its umbrella, and why are they important? Next, the *history* of privacy in philosophy and U.S. law is given as a sort of backdrop. Psychology, anthropology, and several other social sciences are discussed third along with medicine as the main sources of *scientific research* on privacy. Finally, technology is discussed as an *application domain* for privacy.

3.1 Definition: What is Privacy?

Perhaps the simplest definition of privacy is that of Judge Thomas Cooley: the right "to be let alone" (Warren & Brandeis, 1890). It becomes clear that privacy is easy to define, but not without generalizing or inviting criticism. What follows is an aggregation of many definitions and descriptions of what privacy is. We refrain from choosing just one or prescribing our own, but rather recommend being specific about using such a vague term (see Section 3.1.4 below). We also recommend the Stanford Encyclopedia of Philosophy article on privacy by Judith DeCew as a comprehensive guide to the definition of privacy (DeCew, 2013), especially in law and philosophy. Most of the references in this section we owe to the bibliography from that article. As we explore the various *definitions* of privacy in the literature, we will follow the threefold division offered by DeCew (2013): informational privacy, constitutional privacy (i.e., in making intimate decisions about oneself), and privacy in terms of *access* to the self, both physical and mental.

3.1.1 Informational Privacy. Informational privacy refers to privacy concerns about personal information. This was the definition of privacy held by Warren & Brandeis (1890) in their famous article. A number of other authors have presented informational definitions of privacy, although whether they believed this to be the *only* aspect of privacy is not in view here. Prosser (1960) divided (informational) privacy into four parts. His formulation continues to be referenced today. Briefly, Prosser divides (informational) privacy into intrusion, public disclosure, false light, and appropriation. These mean the following. First, intrusion into one's private affairs includes trespassing, search, and remote intrusion such as wire tapping. Second is public disclosure of private facts. Third is publicly portraying the victim in a false light, e.g., by misattributing to the victim a statement or opinion. Fourth is appropriation, or pretending to be the victim for one's own advantage. Solove (2008) has constructed a taxonomy of privacy concepts based on Prosser's formulation. It is shown in Figure 1 as a general overview of informational privacy concerns.

A number of other authors offer informational definitions of privacy that seem to gel with Prosser's. Fried (1970) defines privacy as control over knowledge about oneself. Parent (1983) defines privacy as the condition of others not possessing undocumented information about oneself. Parent also follows, to a certain extent, in the footsteps of so-called "privacy reductionists" like Thomson (1975) in that he dismisses other, non-informational aspects of privacy as being covered by other rights. For example, Parent considers constitutional privacy concerns about the freedom to make important personal decisions to be a matter simply of liberty. Privacy is not a branch of liberty, he argues, because they could be at odds; for example, when someone freely gives up their privacy they are simultaneously exercising their liberty! Parent argues against Thomson, however, by keeping the right to privacy as a distinctive idea; Parent believes that Thomson had to invent some far-fetched human rights in order to maintain her view. "Privacy reductionism" is discussed further in Section 3.2.

Moore (2003) defines privacy as, "control over access to oneself and information about oneself." This is a "control-based" definition of privacy, in which it doesn't matter *per se* whether somebody accesses you or your information, but rather whether you can control that access. Control-based definitions account for situations in which someone invites others into his close company, or willingly gives out personal information. These actions would violate privacy if privacy is the state of being let alone, or of having all your personal information kept to yourself. But authors holding to control-based definitions of privacy maintain that the person in question is still in control, so there's no violation; this especially makes sense in the legal context. Moore (1998) defines this control in terms of the inner spheres of personal information, about the "core self." He limits his definition to the "core self" to intentionally limit the First Amendment[3] freedom of speech in some cases. Moore

[3] The First Amendment to the U.S. Constitution reads in its entirety, "Congress shall make no law respecting an establish-

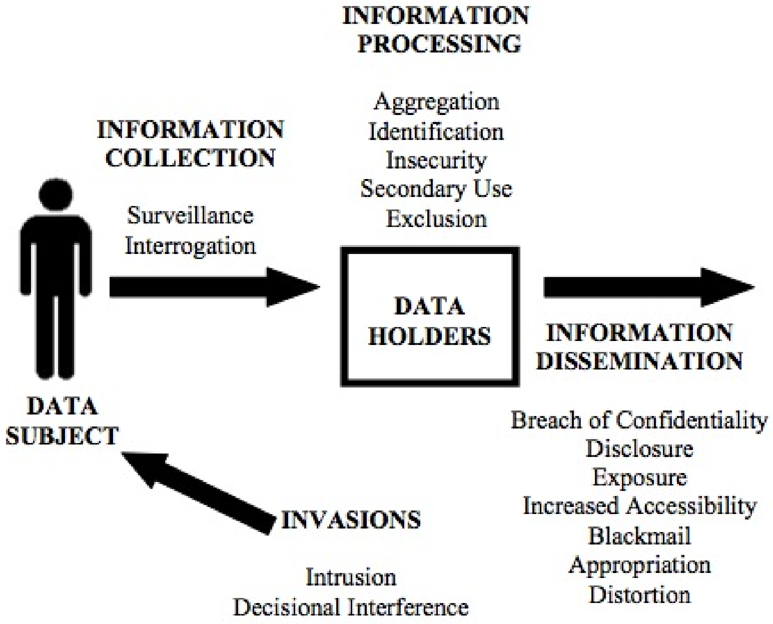

Figure 1. Daniel Solove's visual "model" of his taxonomy of (informational) privacy (Solove, 2008).

argues that sex offenders *should* be forced to reveal their criminal history to their neighbors, and that politicians, having chosen a very public profession, cannot complain when the public learns about some of their more outward characteristics. On the other hand, Moore says, it is suspicious whenever the press claims that the public has a "right to know the truth" about some controversial but very private incident. He holds that privacy may win out in such a situation.

Austin (2003) offers a more nuanced definition of privacy: freedom from "public gaze." She argues that this updated definition deals with the problem of new technologies to which older definitions of privacy do not object. In particular, Austin is concerned about cases wherein people know they are under surveillance, about the collection of non-intimate but personal information (e.g., in data mining), and about the collection of personal information in public. She claims that other, older definitions of privacy do not agree with our intuition that these technologies (could) invade our privacy by denying us our freedom from "public gaze."

3.1.2 Constitutional Privacy and Autonomy. A second legal interpretation of privacy emerged in 1965 with the *Griswold v. Connecticut* decision (Westin, 1967). This was named the "constitutional right to privacy," defined rather vaguely as, "a right protecting one's individual interest in independence in making certain important and personal decisions about one's family, life and lifestyle" (DeCew, 2013). Note that the protection is against governmental, not private, action. This has been used to overturn laws against certain sexual acts and against abortion. In law, this represents a gen-

ment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances." (*U.S. Constitution. Amend. I.*, 1791) eral shift from property-based definitions of privacy to definitions based more on personal liberty. The constitutional right to privacy is mentioned here as a nod to its importance in the conversation about defining privacy; since it applies very little to technology and robotics, it is largely ignored outside of this section and the section on privacy in U.S. law (Section 3.3).

3.1.3 Access Privacy: Intimacy, Solitude, and Space. Daniel Solove wrote that a society without privacy is a "suffocating society" (Solove, 2008). This section presents the broadest conception of privacy, which concerns the human need for being away from others for the sake of both the self and one's intimate relationships. Along these lines, Alan Westin lists four different states of privacy: solitude, anonymity, intimacy (i.e., being alone *with* someone), and reserve (i.e., keeping to oneself) (Westin, 1967). Similarly, Leino-Kilpi et al. (2001) divide privacy into (1) *physical privacy*, over personal space or territory; (2) *psychological privacy*, over thoughts and values; (3) *social privacy*, over interactions with others and influence from them; and (4) *informational privacy*, over personal information. This section reviews the definitions of privacy from the first three of those categories—physical, psychological, and social privacy—and from all four of Westin's private states. These conceptions of privacy are of special interest to anthropologists, psychologists, philosophers, and those in the medical profession.

Julie Inness wrote the book on privacy as it relates to intimacy (Inness, 1992). She disagrees with Fried's understanding of intimacy as a "commodity" which derives its value from its scarcity, and rather proposes that intimate interactions must be *motivated* by liking, love, or care in order to be intimate. As evidence she points to Supreme Court decisions wherein constitutional privacy protection was conferred to issues of the family and sexual health due to the personal, emotional impacts that made those issues intimate. In this way, Inness seems to define privacy as the protection of intimate matters, where intimacy comes from the motivation and not the behavior itself (e.g., kissing is not automatically intimate). She recognizes that this definition of intimacy is subjective, making legal rulings more difficult.

Privacy might also include solitude, i.e., being physically removed from other people. Solitude is more than a freedom from trespassing; one needn't be at home to desire solitude. Allen (2011) includes solitude in her article on privacy and medicine. In the medical setting, the sick often want to be comforted by company, but also to have some time alone. This could be especially true for patients with terminal illnesses, who might want to reflect on their lives and make some important decisions. In such cases we tend to respect their wishes. Allen (2011) also mentions "associational privacy," the ability to choose one's own company. She notes that patients do not merely desire intimacy, but rather "selective intimacy" with certain loved ones, and this is an aspect of privacy to consider.

Finally, privacy could be defined in terms of one's personal space or territory. These concepts are found readily in proxemics literature as well as in psychology and ethology (i.e., animal behavior studies) in general. Patricia Newell includes territoriality in her review of *Perspectives on Privacy* (Newell, 1995), although she also cites a study that separates between the two (Edney & Buda, 1976). We have already mentioned that Leino-Kilpi et al. (2001) define physical privacy as being over personal space and territory, and Westin also mentions it when he links human privacy ideas with animal behavior in Westin (1967).

3.1.4 Summary of Definitions. This concludes our overview of different definitions of privacy. One begins to see some common themes, but also the scattered nature of the literature. Newell (1995) in particular and also Leino-Kilpi et al. (2001) do a good job of synthesizing broad sections of what has been written (from psychological and medical perspectives, respectively). See Figure 2 for Newell's table of privacy definitions. Newell (1995) says, quoting an earlier privacy review, that, "theorists do not agree...on whether privacy is a behavior, attitude, process, goal, phenomenal state,

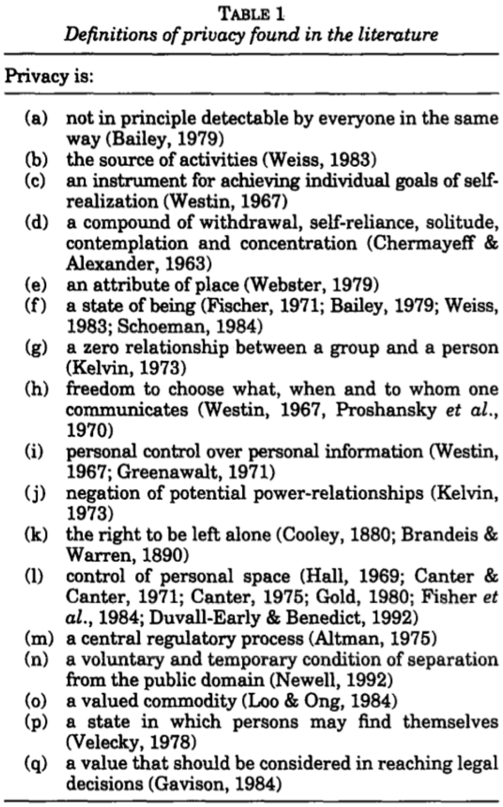

TABLE 1
Definitions of privacy found in the literature

Privacy is:

(a) not in principle detectable by everyone in the same way (Bailey, 1979)
(b) the source of activities (Weiss, 1983)
(c) an instrument for achieving individual goals of self-realization (Westin, 1967)
(d) a compound of withdrawal, self-reliance, solitude, contemplation and concentration (Chermayeff & Alexander, 1963)
(e) an attribute of place (Webster, 1979)
(f) a state of being (Fischer, 1971; Bailey, 1979; Weiss, 1983; Schoeman, 1984)
(g) a zero relationship between a group and a person (Kelvin, 1973)
(h) freedom to choose what, when and to whom one communicates (Westin, 1967, Proshansky *et al.*, 1970)
(i) personal control over personal information (Westin, 1967; Greenawalt, 1971)
(j) negation of potential power-relationships (Kelvin, 1973)
(k) the right to be left alone (Cooley, 1880; Brandeis & Warren, 1890)
(l) control of personal space (Hall, 1969; Canter & Canter, 1971; Canter, 1975; Gold, 1980; Fisher *et al.*, 1984; Duvall-Early & Benedict, 1992)
(m) a central regulatory process (Altman, 1975)
(n) a voluntary and temporary condition of separation from the public domain (Newell, 1992)
(o) a valued commodity (Loo & Ong, 1984)
(p) a state in which persons may find themselves (Velecky, 1978)
(q) a value that should be considered in reaching legal decisions (Gavison, 1984)

Figure 2. Patricia Newell's summary table of privacy definitions from the (largely psychological) literature (Newell, 1995).

or what." Privacy is mysterious. But we have covered a diverse array of definitions for privacy thus far; DeCew (2013) gives a cursory recapitulation:

> Privacy can refer to a sphere separate from government, a domain inappropriate for governmental interference, forbidden views and knowledge, solitude, or restricted access, to list just a few.

Here is our stance on the definition of privacy: "privacy" is a word. A given word can have multiple meanings or be used differently by different individuals or cultures. We should not abandon the word "privacy" because it is ambiguous, however, because it is used in everyday speech, and then to indicate real values and concerns. Rather, we must be clear about which sense of the word we mean, e.g., "privacy, by which I mean control over who can engage you in conversation," perhaps even operationalizing a construct into something measurable, e.g., "access privacy, which here means the fraction of daily conversations that are unwanted by the subject." The "fraction" in the latter example is relatively unambiguous compared to "privacy" or even "access privacy;" researchers should favor such language to avoid misunderstandings.

As a first definition, "privacy-sensitive robotics" we take to include all the constructs within "Informational Privacy" and "Access Privacy" as defined above.

3.2 History: Privacy in Philosophy

Privacy has a long but scattered history in the field of philosophy. Several philosophers, such as Solove, Thomson, and DeCew, have already been mentioned above as they contributed to defining privacy. This section includes some older thinkers, and presents some useful distinctions that go beyond just defining privacy. Here again, Judith DeCew's article in the Stanford Encyclopedia of Philosophy is our guide (DeCew, 2013).

Aristotle wrote the earliest extant philosophy of privacy. Remember, however, that modern usage of the word "privacy" is different than ancient usage. The Greek concept of privacy was a distinction between the affairs of the *oikos* (household) as separate from the *polis* (city-state) (Allen, 1998). The former situation had free males ruling collectively over the city-state, whereas in the latter they ruled "despotically" over the household. So when, in Book II of his *Politics*, Aristotle asks whether property ought to be held "privately" or "in common," he is asking whether each individual man should own it or whether all the men should own it together (Aristotle, 2004). To this question he answers, "privately." Aristotle argues that, if men own things privately, they will take better care of them and enjoy them more. Furthermore, sharing will take place naturally (an optimistic claim!), and all the evils commonly blamed on the paradigm of private property are actually due to human wickedness. Hence, Aristotle is an early proponent of our Western ideal of private ownership, which undergirds certain aspects of privacy.

John Locke was an Enlightenment philosopher who greatly influenced American political thought. In his liberal philosophy, Locke was a promoter of human rights. Adam Moore has written a book on the Lockean idea of intangible property, which includes things like intellectual property which we do not purchase but nonetheless possess by right (Moore, 1998). According to Locke, things that we create are ours as the right of the laborer, "at least where there is enough and as good left for others" (Moore, 1998). In other words, our ideas or self-expressions can be ours by property rights so long as they aren't significantly depriving anyone else of similar rights. Applying this rule can be tricky in the legal context.

Modern philosophers joined the conversation about privacy that started around the beginning of the 20th century. They especially help to subdivide the topic into clear categories. For example, they distinguish between *descriptive* and *normative* accounts of privacy, i.e., between describing what privacy covers and examining why it's important as a value or right. The rest of this section is divided according to that distinction; we start by looking at some *descriptions* of privacy, then move to some work on the *normative* side.

A key contribution by philosophers to descriptive privacy is the coherentist-reductionist debate (DeCew, 2013). Schoeman (1984) coined the words "coherentist" and "distinctivist" to represent separate groups of people, but we will take them together as a single group that talks about privacy as an independently important idea that cannot be *reduced* to other ideas—"reductionists" hold that other view (DeCew, 2013). We think that taking the reductionist view here would contradict popular opinion and shatter this one paper into many, so we shall remain coherentists. Nevertheless, it is valuable to understand the reductionist viewpoint, stated most famously by Judith Thomson.

In her 1975 article, *The Right to Privacy*, Thomson does not debunk privacy altogether but rather calls it a "derivative" right, i.e., a right based on other rights (Thomson, 1975). The argument runs as follows. First, we have rights over not just our property but (even more so) over our own bodies. What might these rights of the body include? Here Thomson cites what she believes to be the earliest statement of body rights: "A Definition of Privacy," written in 1974 by Professor Richard Parker. Parker says the following:

> Privacy is control over when and by whom the various parts of us can be sensed by others. By 'sensed,' is meant simply seen, heard, touched, smelled, or tasted. By

> 'parts of us,' is meant the parts of our bodies, our voices, and the products of our bodies. 'Parts of us' also includes objects very closely associated with us [e.g., possessions].

Thomson's examples of body rights are the right to not be looked at and the right not to be listened to. She is quick to state that these rights are not always *claimed*; rather, we usually *waive* them implicitly when we go out in public, perhaps whistling as we walk. But they become very apparent in some cases: Muslim women would surely claim the right not to have even a bare knee seen by a stranger, for example.

Surprisingly, Thomson does not throw privacy out completely here. Instead, she defines each privacy right as some other right with a personal component. For example, someone reading my book violates my right to it as my property, whereas someone reading my diary violates the *same* right, but with regard to my personal information, from which arises an additional right to privacy. Or consider the case of torture: torturing me to extract some random fact, like the capital of the state of Oregon, only violates my body right to go unharmed by others. Torturing me for *personal* information, however, violates the same body right *and* my further right to privacy. The key here is the *direction* of the logic: my right not to be harmed for my personal information does not come from my right to privacy, but rather the other way around.

Most authors, however, talk about privacy as distinct and not (wholly) derived from other rights. On that note, we now consider some *normative* accounts of privacy; that is, some different conceptions of why privacy is important. Why is it good or useful to value privacy? For Bloustein (1964), privacy is about the safety and freedom of each person's individuality and dignity. Fried notes its necessity for friendship and trust (Fried, 1970). Inness among others says it is necessary for intimacy with others (Inness, 1992). James Rachels, besides making several other cogent contributions that have been kept from this paper for brevity, addresses the question of why we need privacy if we aren't doing anything embarrassing (Rachels, 1975). He answers that the presence of other people influences our actions, so being unable to control one's company is actually a loss of autonomy. Hence, privacy is a concern even when we're behaving ourselves—in fact, inasmuch as we feel pressured to "behave," privacy includes the freedom to escape such a situation.

3.3 History: Privacy in U.S. Law

The story of privacy law in the United States is in many ways the crux of the history of privacy. What began as essentially a set of property rights has grown at both the state and federal levels to include rights to confidentiality and about intimate personal decisions. This section surveys the major contributions in chronological order, beginning with the ratification of the Bill of Rights in 1791.

The Fourth Amendment to the U.S. Constitution begins as follows:

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated... (*U.S. Constitution. Amend. IV.*, 1791)

. . . to which the reader should silently add, "by agents of the *government*." The Bill of Rights was drafted in order to protect the rights of U.S. citizens from an unlimited federal government; it was only by the doctrine of "incorporation" that, starting in the 1920s, courts began holding state governments accountable for these rights as well. History lessons aside, the Fourth Amendment has been one foundation, although not the only one, of court rulings about privacy rights in the United States.

In 1967, Alan Westin published *Privacy and Freedom*, the result of over four years of effort and study by a group of researchers (Westin, 1967). His book includes a chronology of privacy law in the United States which we will follow here with some interjections from other authors. Westin begins with America as a new nation, framed by the Constitution. Compared to Europe, America had much in the way of privacy law before the Civil War (Westin, 1967). American politics were founded on the worth of the individual, limited government, property rights, and the liberty to do what one will with his property. The First Amendment conferred the freedom to speak, but also to keep silent. Anonymous publication of one's opinions was allowed, and police surveillance of public meeting places, as was prevalent in Europe, was expressly forbidden. Justice Story held that the Third Amendment (against housing troops in private homes) was meant to secure, "that great right of the [English] common law, that a man's house shall be his own castle, privileged against all civil and military intrusion" (Westin, 1967). There was also judicial precedent (part of common law) that forbade nuisances, trespassing, and eavesdropping. Trademarks, corporate and government secrets, and letters all enjoyed general protection.

It was technology that began causing problems from the legal perspective (Westin, 1967). The telephone was invented in the 1880s, followed closely by wire-tapping within a decade. Similarly, the microphone was invented in the 1870s, followed by the dictograph recorder in the 1890s. For capturing images, "instantaneous photography" became possible with the Kodak camera in the 1880s, and by 1902 *The New York Times* reported on what we would call the first paparazzi: "kodakers" waiting to capture photos of important people (Westin, 1967). It is not a coincidence that this is when the Supreme Court and legal scholars began to treat seriously the expansion of American privacy rights.

We have already mentioned the true beginning of the privacy conversation in America: the short article written in 1890 by Samuel Warren and Louis Brandeis entitled, "The Right to Privacy" (Warren & Brandeis, 1890). The occasion was a humorous one. Warren's wife was among the social elite and her parties tended to be written up in the newspaper; eventually, Mr. Warren got fed up and joined with Brandeis to write the article (Prosser, 1960). Warren and Brandeis cite Judge Cooley's right "to be let alone" in the context of increasing invasions by journalists of Americans' private lives (Warren & Brandeis, 1890). They argued that both the intimate and the mundane details of one's life are protected, along with one's thoughts, sentiments, and emotions, until "published" to others. Violations are to be redressed by tort (i.e., monetary compensation) or by a court injunction (i.e., restraining order).

Prosser (1960) reports that the article by Warren and Brandeis was argued over for about 30 years. By the 1930s, its arguments began to be accepted by the courts, and most states had privacy regulations along the lines of Warren and Brandeis by 1960. Prosser reviewed the over 300 privacy cases on the books by 1960 and reports on the emergence of four separate torts over the years. We have already discussed them in Section 3.1.1: invasion, public disclosure of private facts, false light in the public eye, and appropriation of someone's name or likeness. We now continue with Westin's account (Westin, 1967). The right to privacy swelled in the late 1900s. The Supreme Court continued to try cases in which new technologies created privacy concerns beyond physical entry and tangible items. Other new protections included "associational privacy" over group memberships and "political privacy" over unfair questioning on account of political positions. Anonymity in public opinion was upheld in *Talley v. California, 1960* (Westin, 1967). Privacy of the body was upheld in a 1964 case wherein the "reasonableness" of certain compulsory medical examinations were in question (Westin, 1967). Most importantly, and as discussed above in Section 3.1.2, "constitutional privacy" emerged in 1965 with the *Griswold v. Connecticut* decision (Westin, 1967).

In 1948, the Universal Declaration of Human Rights was adopted by the United Nations General Assembly (The United Nations, 1948). Although it came before the rise of "constitutional privacy" in America, this document communicated an updated understanding of privacy rights to the world:

> No one shall be subjected to arbitrary interference with his privacy, family, home or

> correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks. (The United Nations, 1948, Article 12)

The words "interference" and "privacy" seem to be quite general here; we see this statement as a pretty broad dispensation of privacy rights.

A number of authors have written about privacy more recently (say, since 1980). We have already discussed several in Section 3.1, such as Parent (1983), Austin (2003), and Solove (2008), that contributed to *defining* privacy in the legal domain. Beginning in the 1970s, however, feminist criticism of traditional understandings of privacy arose (Allen, 1998). Anita Allen, for one, reminds privacy scholars that home life is not automatically good; privacy can hide abuse, subordination, and isolation, especially for women (Allen, 1998). Allen believes that we can salvage privacy as a society, however, which is good news seeing as constitutional privacy rights about abortion and contraceptive use are especially important for protecting women. Her argument seems a bit conflicted in places, but underlines that women can be especially affected by the way we implement privacy as a society.

3.4 Research: Privacy in the Social Sciences

Various social sciences—anthropology, sociology, psychology, economics—have studied privacy in humans. That being said, one could also argue that the roots of human privacy values can be seen in animal behavior—that is, in ethology. Alan Westin states that almost all animals seek privacy, either as individuals or as small groups (Westin, 1967). It is from animals that we get the idea of *territoriality*, or the defense of one area against intrusion by others of one's own species (which raises questions, by the way, about what species a robot might be—see Section 6.1.7). Westin reports three types of spacing observed between animals: personal distance between individuals (e.g., between birds on a wire), social distance between groups, and flight distance at which an intruder causes fleeing (Westin, 1967). At the same time, animals often gather in large crowds, and it seems that at least part of the reason is the mere social stimulation. Animals, like humans, seem to live in a tension between privacy and sociality.

The English words "private" and "privacy" come from the Latin adjective *privatus*, meaning "set apart, belonging to oneself (not to the state), peculiar, or personal." This is used in contrast to *publicus* and *communis*, which have the expected meanings. *Privatus* is itself the past participle of *privare*, a verb meaning "to separate or deprive," which in turn comes from the adjective *privus*, meaning "one's own, individual" (*Online Etymology Dictionary*, 2015). Several modern languages lack a word with the exact sense of English "privacy;" these include Arabic, Dutch, Japanese, and Russian (Newell, 1995). French has some related words, but none covers the concept in general. Japanese has four related concepts with similar meanings to "private life," "freedom," "solitude," and "secrecy" in English (Newell, 1995). Therefore, having a single word that is even roughly equivalent to the English "privacy" is *not* universal across languages.

Is privacy-regulating behavior peculiar to Western cultures, or even to American culture? Irwin Altman investigates in his study entitled, "Privacy Regulation: Culturally Universal or Culturally Specific?" (Altman, 1977). He hypothesizes that privacy is a universal human need, but that different cultures have different privacy regulation mechanisms. Altman surveys two types of cultures: those with apparently minimal privacy and those with apparently maximal privacy. If mechanisms exist to regulate privacy in those extreme cultures, he argues, then privacy must be universal despite all appearances. This was in fact the case. For example, the Mehinacu culture in Brazil appears to have almost no privacy; housing is communal, people enter huts without announcement, and each person's actions and whereabouts are generally known (Roberts & Gregor, 1971). Upon a closer look, however, we see mechanisms for regulating privacy. It is permissible to leave the village for days at a time, or to walk various secret paths alone and unaccounted for. Lying was used regularly to avoid revealing too much while living in such close quarters. Some privacy-related behaviors were humorous, especially that of the Sapps people of Northern Europe. These are reindeer herders who live in tents, and while anyone could make oneself welcome in anybody's tent, a disgruntled occupant would often feign sleep to signal for privacy (R. Paine, 1970).

We move now to studies of human privacy in the modern, developed world. Privacy during childhood is a good place to start. Zeegers et al. (1994) found that 58 of 100 three-, four-, and five-year-olds said they had a special place at the daycare center that belonged only to them. When confronted with reduced time and space for privacy, children in daycare can become territorial, more defensive of their personal space, and more attentive of their physical privacy (Jacobs, 1977). Newell (1995) mentions the importance of "refuge" to autistic children especially; she even reports that some children use poor behavior to get put in isolation.

What about privacy norms in adults? Patricia Newell's study of American, Irish, and Senegalese subjects (mostly students) found that they usually sought privacy when sad or tired, or to concentrate, and felt relaxed and refreshed afterwards (Newell, 1998). Friedman et al. (2009) studied the difference in privacy judgments between the so-called "watcher" and the person being "watched" in a public setting. The results revealed a gender difference: women were significantly more likely to be surprised or troubled by the idea of being watched via a video camera. Also, men were less privacy-sensitive as watchers than they were when watched.

Acquisti & Grossklags (2005) study the dichotomy between privacy *attitudes* and privacy *behaviors*. Sometimes, even people who value their privacy don't seem to act to protect it. The authors hypothesize three reasons for this dichotomy. First, people have incomplete information when they make privacy decisions. Second, humans have only "bounded rationality," so we can't make perfectly rational decisions if there is too much information to consider. Third, we have certain systematic biases. For example, we lack self-control and opt for instant gratification even when it's sub-optimal in the long run. The results of a 119-subject survey supported all three hypothesized factors. For example, 44% of subjects showed a time-inconsistency bias by discounting future payoffs. This suggests that people cannot be expected to make fully rational decisions about privacy, even if they value it dearly.

Another study by Acquisti and his colleagues assessed the *subjective value* of privacy (Acquisti et al., 2013). They draw an important distinction between the willingness to accept (WTA) payment to give up one's privacy and the willingness to pay (WTP) for privacy protection. Economists have found that people are willing to accept only a very high payment to give things up, perhaps due to "loss aversion" (Acquisti et al., 2013). The results gathered by Acquisti et al. (2013) showed five times higher WTA than WTP in a privacy-related context; subjects were unwilling to make even small monetary sacrifices to gain privacy protection. They also found that valuation of privacy depends on how you ask. It depends very much how much privacy protection a person is given to start with, and whether you propose to take some of it away or to provide more.

3.4.1 Personal Space, Territoriality, and Environment. Social psychologist Altman (1975) pulls together the related concepts of privacy, personal space, territoriality, and crowding. His book, along with Burgoon (1982) and Sommer (1969), discussed below, is a good foundation for *environmental and spatial* factors related to privacy. Altman's theory defines privacy as a *boundary regulation process* wherein people try to achieve their ideal privacy state by using certain *mechanisms* to regulate interaction with others. Notice how this definition allows privacy to sometimes mean *more* interaction with others, and sometimes *less* interaction; successfully switching between the two is the key. Along these lines, Altman calls privacy a *dialectic process*, i.e., a contest between two opposing forces—withdrawal and engagement—which alternate in dominance. Hence, privacy to

Altman is *dynamic* in that the desired level of engagement changes over time for a given individual. This theory is necessary for understanding Altman's discussion of personal space, territoriality, and crowding.

To Altman, personal space and territoriality are two *mechanisms* for achieving a desired privacy state. Another example is speech. Reviewing the literature at the time, Altman reports that there are individual differences in the use of personal space, but it's hard to confidently state clear differences between men and women or between different cultures. It does seem that people maintain less personal space with people they know and like, as well as in informal settings. Personal space is also relationship-dependent; being too close is an invasion for strangers and some interactants, but excessive *distance* can be undesirable between intimates. As for territoriality, Altman notes that the phenomenon in animals is related in some ways, but also markedly different in other ways. He lists three kinds of territories: primary (e.g., a bedroom or home), secondary (e.g., a local bar), and public (e.g., a bus seat). The three kinds vary in the extent to which *markers* and other strategies can be used to claim and defend space. Altman also distinguishes from outright *intrusions* both *obtrusions*—defined as "excessive use" of the space, such as being too noisy—and *contaminations* of the space. Altman recommends longitudinal studies in natural environments for future research on territoriality.

Altman distinguishes *crowding* from *density*. Crowding is when too much social interaction occurs, whereas density is a quantitative measure of persons per area, per room, per home, etc. The opposite extreme from crowding is social isolation; both are bad and indicate failure of the regulatory mechanisms that enforce a person's privacy. Altman presents results of laboratory experiments on the effects of crowding on people, as well as various mechanisms for coping with crowding.

Judee Burgoon presents a communication perspective on privacy, including territoriality, in a broad survey (Burgoon, 1982). She argues that more "physical" privacy could consist of blocking more communication channels, including sight, sound, and even smell (e.g., the smell of food being cooked next door). We would add further channels enabled by technology: phone calls, text messages, Facebook posts, and the like. Alternatively, Burgoon writes that to have more territory, higher-quality territory (e.g., better-insulated walls), and more unquestioned control over that territory is to enjoy more physical privacy. She goes on to list many aspects of physical privacy: in spatial intrusions, the probability of violation, humanness of the violator, and relationship with the violator; in social interactions, control over who the interactants are, the frequency and length of the interaction, and the content of the interaction; also, formality allows interactions to be kept short and impersonal, whereas a more private state is characterized by greater rule-breaking, "backstage" behavior, and freedom to engage in emotional release.

Burgoon cites Robert Sommer's book on personal space with respect to the possibility that an intruder would be treated as a "nonperson" (Sommer, 1969). Sommer observes that inanimate objects like trees and chairs, as well as pets in some circumstances, are not treated as intruders at all. This same phenomenon can occur with people. For example, a cubicle dweller might continue a sensitive phone conversation even as the janitor enters and begins emptying the trash. In certain social contexts, such as subway trains and sporting events, it can be socially normal to treat others as nonpersons.

Sommer also studied territorial markers. To effectively protect an area, markers need to either be an explicit sign (e.g., a "Reserved" sign) or to have intrinsic value (e.g., coat, purse, umbrella); litter doesn't count. Valid markers can be very effective at protecting a spot, especially during low room density. As more people enter the space, markers are more likely to be ignored, and others in the room might be asked to confirm whose territory that is.

We now return to Burgoon's list of privacy *mechanisms* in Altman's sense (see discussion of Altman, 1975, above), clearly influenced by her expertise in communication. She lists six categories of mechanisms. First is environment and artifacts, which is further broken into architectural design, artifacts and furnishings, and gatekeepers (e.g., receptionists effectively guard the rest of the building). Factors like the size of each area, the blockage of senses between areas, and other cues that say "socialize here" or "don't" all impact the privacy of a space. Second is personal space and touch, which includes interpersonal distance, seating positions at tables, body orientation, degree of sideways or backwards lean, and use of physical touch. Third is chronemics, or the usage of time. For example, people could use the same space but at different times, avoid social areas during peak usage hours, or even declare different functions for a shared space at different times. Fourth is kinesics and vocalics, or cues from the body and voice. Examples of "exclusion cues" include "body blocks" or closed postures in a dyad, saying "go away," silence, or avoiding eye contact so others can't begin visual communication. "Affiliation cues" include smiling, relaxed postures, greater mirroring, and warm vocal tones. Fifth is physical appearance and attire, and sixth is verbal mechanisms. Numerous verbal mechanisms exist. The linguistic features of speech, such as tense, use of personal pronouns, use of negation, and ambiguity can regulate privacy. So can the degree of self-disclosure and formality, changing the topic, brevity or verbosity, and direct references to one's possessions, territory, or rights. Finally, there is the idea of "linguistic collusion"—using in-group language to exclude others (Burgoon, 1982).

We will mention several field studies in this area. First, Becker & Mayo (1971) note the possibility of confusing the concepts of personal space and territoriality. They study whether what Sommer & Becker (1969) call territorial behavior in a cafeteria setting is actually closer to Hall's concept of personal distance (Hall, 1966). Is using markers to reserve a spot at a table really just a mechanism for maintaining a comfortable personal distance? If so, the person who marked the space will be just as likely to move to a new spot upon intrusion than defend that spot in particular. A study of 48 unsuspecting university students supported this hypothesis. The authors argue that the active construct in that scenario was personal distance, not territoriality.

Walden et al. (1981) study how incoming freshmen at an American university cope with different levels of crowding. The authors measure both objective and subjective crowding (what Altman calls density and crowding, respectively; see above). They also highlight the difference between one's *expectation* of achieving the desired privacy state and the *value* or pay-off of achieving that goal. For the study, students were assigned either a two-person or a three-person dormitory room and surveyed about their experiences. The results showed an apparent difference between the way males and females cope with crowding, but were limited by a low sample size, especially of male subjects.

Sebba & Churchman (1983) interview 45 Israeli families with two to four children living in apartments of almost identical size and layout. Based on their results, they classified territorial areas into four categories. The first three were called individual, shared, and public areas. Single-occupancy bedrooms were individual territorial areas, multiple-occupancy bedrooms were shared, and most living rooms were public. The final category they called jurisdiction areas, wherein one person has jurisdiction but the space is used by everybody. About half of the kitchens were defined as jurisdiction areas—they were used by everybody, but belonged to the mother. Additional research is needed to determine whether these findings generalize to other demographic groups and types of spaces.

3.4.2 Online Privacy. Privacy on the Internet is a bit of a different animal. The user can browse as an anonymous, disembodied agent, and is confronted with compelling advertisements and endless information. In a study of online interaction in general, subjects were guided by an online avatar named "Luci," who asked them personal questions to help them pick out some things to buy (Berendt et al., 2005). Some questions were purely relevant to the purchasing task, while others were more prying. In general, subjects gave away a lot of information, and again this was true even of those who self-reported as valuing their personal information.

C. Paine et al. (2007) studied *why* people (don't) act to protect their privacy online. Five hundred and thirty people from various countries responded to an open-ended survey administered by an automated instant-messaging bot. When asked why they were (not) concerned about privacy online, those concerned cited viruses, hackers, etc., whereas those not concerned cited their own competency, protective software, apathy, and a lack of valuable information to hide. Those who reported *not* taking privacy-protective actions cited as reasons apathy, a lack of felt need for protection, and (this was troubling) not knowing *how* to get effective privacy protection online. Again, the questions were open-ended and subject to respondents' interpretations of "privacy online," but the results enumerate some issues that need addressing in online privacy.

Of special interest to this paper is the development of privacy metrics and evaluation methods, since those could possibly be adapted for use in human-robot interactions (see Section 6). Buchanan et al. (2007) construct a new online privacy measure (namely, a questionnaire) and test it against two popular privacy measures from the previous literature. They propose three privacy metrics to be measured by a 28-item questionnaire: General Caution (when using the Internet), Technical Protection, and Privacy Concern. These metrics were validated both as constructs and against two other scales: Westin's Privacy Segmentation (Harris and Associates & Westin, 1998) and the IUIPC (Malhotra et al., 2004). Over 1000 respondents took all three measures, and results were significantly correlated in general.

3.5 Research: Privacy in Medicine

The field of medicine is aware of the importance of privacy. This is evidenced by the *Declaration on the Promotion of Patients' Rights in Europe*, which includes respect for privacy as 1 of its 6 human rights (The United Nations, 1948). The *Declaration* frowns on any medical treatment that cannot be performed with respect for the patient's privacy.

Allen (1995) lists three main uses of the term "privacy" in the healthcare domain: physical privacy, informational privacy, and decisional privacy. The third usage—decisional privacy—is especially salient in healthcare. Allen defines decisional "privacy" as the freedom to make one's own decisions (here, about medical treatment) and act on them without undue outside influence. Careful readers will notice that decisional "privacy" is the same as constitutional privacy, discussed above in Section 3.1.2. Allen places quotation marks around the word "privacy" in decisional "privacy" because of some controversy over whether decisional "privacy" should be called privacy at all, but rather liberty, freedom, or autonomy (Allen, 1995). In a different article, Allen (2011) lists *modesty* as an important physical privacy concern in medical settings, especially from the philosophical standpoints of Christian ethics and virtue ethics. Modesty may drive patients to request same-sex or even same-race doctors.

Things we consider common etiquette, such as knocking on a door before entering, can be supportive of physical privacy (Leino-Kilpi et al., 2001). Also, technology is causing problems in the medical as well as the legal realm; a survey in a Finnish hospital revealed that, "Only 30% of 166 respondents, however, believed that their data was safe in the hospital's computer system" (Leino-Kilpi et al., 2001). In a 2005 study of an Australian emergency department, 105 of 235 survey respondents reported a privacy breach, defined as either personal information being overheard or private body parts being seen (Karro et al., 2005). Influencing factors included the length of stay and whether patients were separated by solid walls or just curtains.

Applegate & Morse (1994) have published a study of privacy in a nursing home for Canadian war veterans. The authors focused on the relationships between the residents, between the staff members, and between residents and staff. Relationships were categorized by how relational others were treated: as friends, strangers, or inanimate objects. This last phenomenon is understood as *de-*

humanization, and took several different forms. Privacy was violated most often for those who were dehumanized. Dehumanization was more common towards the less mentally competent residents; other residents would sometimes push them out of the way, and staff members might administer their medicine forcibly and without verbal acknowledgement. This study highlights how privacy-relevant behavior is subject to other factors—here, the mental act of dehumanization—that might not be expected by the privacy researcher but nevertheless drastically affect the situation. Moreover, here we see that there are places wherein privacy is not a fringe concern, but rather an everyday concern, a central quality of life issue like it is in a prison. These may be edge cases, but they are nonetheless motivating.

3.6 Applications: Privacy in Technology

Video media spaces (VMS) connect people separated by distance with video channels. Mediated interactions differ from face-to-face interactions in several key ways; Boyle et al. (2009) present a vocabulary for understanding these differences. *Disembodiment* is the stripping of context (e.g., at home, hard at work) from the interaction. *Role conflict* is when a media space places someone in several disparate contexts at once, as in the familiar case of working from home. *Dissociation* is the decoupling of one's body and identity from one's actions, as with a remote operator of a robot. Social signals that are lightweight for in-person interactions are more difficult in a media space. For example, signalling availability (regulating *solitude*) is done with nuanced facial expressions, tones of voice, hand and posture signals, and environmental cues like a door leaned shut. This is all possible in video media spaces, but is currently awkward instead of natural and blunt instead of nuanced.

Boyle et al. (2009) continue by arguing that privacy risks are unavoidable with technology. A privacy *risk* has both *probability* and *severity*. Privacy risks are only worth it if counterbalanced by some reward; the technology must provide a comparable benefit. Just as for other personal rights, privacy rights need protection by *policing*, which includes real *punishments* for violations and warrants the formulation of privacy *rules*. This is especially true for situations with an imbalance of power, such as when one person can access (i.e., see and hear) another person's space via a one-way connection. We can restore balance through *reciprocity*, which is when person A can do to person B what B can do to A.

The authors also distinguish between *access control* over who can use the media space and *content control* over what users can see and hear (Boyle et al., 2009). Content control is often provided by filters. For example, the eigen-space filter by Crowley et al. (2000) ensures that only images from a socially-acceptable set will be displayed. Extreme examples of filtering include the availability indicators (e.g., green means available, red means busy) used in instant messaging applications. This brings up concerns about the minimum amount of *fidelity* needed for some task, and also about *data integrity*, i.e., that the video feed is faithfully modified before it reaches the recipient. Finally, the authors mention evidence that mandatory media space usage causes social changes among the users over time (Boyle et al., 2009). One could speculate that robot usage will cause even more drastic changes.

Privacy is also a topic of conversation within the subdiscipline of *ubiquitous computing*. Weiser (1993) defines ubiquitous computing, or "ubicomp," as, "the method of enhancing computer use by making many computers available throughout the physical environment, but making them effectively invisible to the user." Hong & Landay (2004) claim that "privacy is easily [ubiquitous computing's] most often-cited criticism." One major principle that has emerged from this conversation is to consider privacy issues during product design instead of just making rules to fit the products. Bellotti & Sellen (1993) introduce some "design for privacy" principles with respect to the RAVE media space environment at EuroPARC, including some specific design suggestions that demonstrate what design for privacy is all about. Langheinrich (2001) gives six principles for privacy in ubicomp: notice, choice and consent, proximity and locality, anonymity and pseudonymity, security, and access and recourse. Here, "proximity and locality" means using location information to enforce access rules based on where the accessor is; "access and recourse" means users should be able to access their personal information and see how it has been used by others (i.e., via usage logs).

Lederer et al. (2002) promote a so-called "situational faces" metaphor for privacy settings in ubicomp environments. Users, they argue, need adequate *notice* that describes the "situation" to the point where they can *consent* to the appropriate "face" (like a user profile; e.g., "secure shopper," "anonymous," "hanging out with friends"). Each "face" could cover multiple situations. Their focus on notice and consent comes from the fair information practices, which also include, e.g., access, security, and redress. The authors also cite the Boundary Principle, which says to place privacy notices[4] at the *boundaries* between different ubicomp environments. The same authors have explored two factors, *inquirer* and *situation*, via a questionnaire-based study (Lederer et al., 2003). Their study addressed the question, e.g., of whether a person would change "face" if the same inquirer requests access to personal information in two different situations.

Langheinrich (2002) and Hong & Landay (2004) propose specific architectures for handling privacy issues in ubicomp environments. The architecture in Langheinrich (2002) is based on the six principles given by Langheinrich (2001), and uses the "machine-readable format for privacy policies on the Web (P3P)" described by Reagle & Cranor (1999). Hong & Landay (2004) developed "Confab, a toolkit for building privacy-sensitive ubicomp applications." User needs were assembled from a variety of interviews, papers, and other sources, and then summarized into four categories, including plausible deniability (i.e., ability to give an excuse without the system proving you a liar) and special exceptions in emergency situations—namely, sacrificing privacy for safety.

Closely related to ubiquitous computing is the concept of the Internet of Things (IoT), which focuses more on objects (artifacts) imbued with computing power and network connectivity. Atzori et al. (2010) survey the topic and lists a few privacy-protection strategies being considered, including anonymization of data collected by sensor networks and use of privacy brokers between users and services. Weber (2010) gives a legal scholar's perspective of security and privacy issues with the IoT, including a short survey of technical requirements for a privacy protection system as well as some existing privacy enhancing technologies. The author also evaluates the role and actions of the European Commissions with respect to security and privacy in the IoT up until late in 2009.

3.6.1 Concerns Specific to Robots. Ryan Calo, a law professor, wrote the "Robots and Privacy" chapter of *Robot Ethics* (Calo, 2010). He notes that robots are rapidly trending towards ubiquity, and perhaps this isn't all good. Calo identifies three privacy dangers posed by robots: surveillance, access to living and working spaces, and social impact. The worry about access is worsened by the work by Denning et al. (2009) wherein the authors demonstrate security vulnerabilities on several toy robots. Calo divides his concerns about social impact into three parts. First, having robots everywhere could make solitude hard to find. There is some push for this to happen; the South Korean government, for example, officially intended for a robot to be in every household by 2015. Robots can also act as information extractors and persuaders, perhaps better than humans can. Over a decade ago, the ELLEgirlBuddy was deployed to advertise Elle Girl Magazine to teenagers, from whom it also harvested marketing information via instant messaging (Calo, 2010). Calo also introduces the idea of *settings privacy*, or concern that the way you personalize your robot could be sensitive information about you; what before was kept to yourself is now *datafied* and stored onboard the robot. This is especially concerning in light of David Levy's 2007 book *Love + Sex with Robots* (Levy,

[4] Some researchers are seeking to improve the privacy notices themselves. See Kelley et al. (2009) for an example inspired by the nutrition labels required by the U.S. Food and Drug Administration.

2009). As robots become increasingly amenable to customization, and also able to learn from their experiences and interactions, an increasing amount of personal information may become embedded in our robots. In this case, lawmakers should be concerned about the "third-party doctrine," under which individuals can lose Fourth Amendment protection for information they voluntarily give to some third party (Calo, 2010). As it stands in American law, you may lose your claim over any personal information that your robot learns about you.

Additional robot-specific privacy concerns have been proposed at We Robot, an annual conference about robotics law and policy. Thomasen (2012) discusses the possibility of robotic interrogation. Bankston & Stepanovich (2014) discuss the interception of email messages by the US National Security Agency (NSA), arguably by web-crawling robots that make decisions about the data. Hartzog (2015) points out that robots in particular can be unfair or deceptive to consumers. Barabas et al. (2015) discuss telepresence robots using a specific system developed by the authors as an example. The issues raised by all these authors demand two responses: study of the real privacy phenomena by HRI researchers and technologies that can be implemented on real robots to help protect user privacy. The latter is the topic of the next section.

4. Constraints for Privacy-Sensitive Robots

We now turn from our survey of privacy to a new topic. If robotics is still a young field, what we call privacy-sensitive robotics is even younger. Here, then, we must be content with a survey of robotics techniques that show promise for protecting the privacy of users. We have focused our search on techniques that can be used to *constrain* or limit robot behavior in some way, as privacy will mostly be a restraining and not a liberating force on robots. Specifically, we will need to restrict what the robot sees (Section 4.1), where it goes (Section 4.2), and also what it touches (Section 4.3). What a robot (or robot operator) actually *does* with the data it collects or in the spaces it occupies is not being considered here, although data usage and robot behavior will impact users' privacy expectations.

It should be remembered that constraining a robot can draw extra attention to private objects or areas. If a remote operator can see that an object is being redacted in a video stream, he or she might wonder what is being hidden; if the robot conspicuously avoids an area or object, local users might become curious. If a remote operator is malicious, or even just curious, he or she might move the robot to try to see past a filter or manipulate a filtered object. Perhaps additional, fake restrictions can be added to make the real ones less interesting by comparison. In any case, good privacy-protecting robot constraints should be tested against malicious users to ensure they are robust to tampering. If a constraint that is perceivable by remote operators or local users would attract too much attention, that constraint might have to be made imperceivable or removed altogether.

4.1 Constraining Perception

Many robots are equipped with cameras. Here we will consider how to limit what a robot can see. Some robots can hear, feel, or even smell, but it is more difficult for these senses to violate someone's privacy than it is for vision. We have said that privacy can be informational (which includes seeing and hearing private things) as well as spatial. Some of the most basic privacy issues are visual: seeing someone without their clothes, for example. Also, many informational privacy violations among humans are visual: intellectual property is stolen by reading a document or seeing a product. Even spatial privacy may have a significant visual portion, as indicated by the trope of blindfolding outsiders as they are led through a secret area.

It is important to note from the outset that image manipulation is only important if the vision system actually captures privacy-compromising data. For example, Zhang et al. (2012) present a computer vision system to detect if an older adult has fallen down. Their work points out that certain sensors—here, a depth camera—can function without compromising a person's identity. They also use an RGB camera, which, by collecting color image data, affords much less privacy protection. It becomes clear that, when a user is beyond the range of the depth camera, a decision needs to be made as to whether to use the RGB camera, which compromises the user's identity, to continue providing functionality. Once it is decided that privacy-compromising data will be collected, image manipulation techniques may be employed to mitigate the privacy violation.

Templeman et al. (2014) present a classifier for deciding whether an image was taken in a private area such as a bathroom or bedroom. The classifier attempts to match landmarks in the image stream with landmarks in sample images of the private areas. The authors focus on first-person cameras for humans, but their work probably extends to cameras onboard robots. This becomes unnecessary for robots that can localize themselves with high confidence in a map with the private areas labeled; the classifier remains useful for robots that do not localize or do so with too much uncertainty.

Forsyth et al. (1996) present a technique that detects whether one or more naked people are in an image and, if so, provides a mask of the offending region. Privacy-sensitive robots could use a nudity detector to *avoid* naked humans in some contexts (e.g., upon opening a bedroom door and seeing the occupant in the middle of changing his or her clothes, the robot could decide to leave immediately) and, in other (e.g., medical) contexts, to *cover* the naked bodies with an image filter (see next section for examples) before the images reach the remote operator. It seems that most people regard certain of their body parts to be private—i.e., off-limits for viewing by others—when uncovered. A nudity detector would help robots respect this aspect of privacy. It appears that the state of the art has a long way to go if robots are to reliably respect this form of privacy; difficult unsolved problems include automatically deciding whether a partially-dressed person is "decent" (i.e., has all the private body parts covered) and differentiating between, e.g., a swimsuit, underwear, and lingerie, each of which could evoke a different set of privacy rules.

4.1.1 Image Manipulation Techniques: Descriptions. Many techniques can be found in the graphics and animation literature for post-processing images. Note that these techniques only use the information available within the image itself; outside help from the artist or from additional (e.g., depth) sensors is not being considered. Some image manipulations are intended to obscure parts of the image. These methods include pixelating, blurring, redacting, and replacing either the entire image or just certain regions (see pictures in Zhao & Stasko, 1998; Boyle et al., 2000; Raval et al., 2014; Hubers et al., 2015). Only a few works are discussed here, as representative examples.

Perhaps the simplest way to simplify a digital image is to reduce its resolution. This makes the image appear blocky, hence the moniker, "pixelation." Naive pixelation preserves low-detail regions but obscures high-detail regions of an image. For example, in Figure 3b (reproduced from Gerstner et al., 2012) the subject's eyes, nose, and mouth are difficult to discern, but his clothing remains discernable as a suit and tie. With more intelligent pixel grouping, it becomes easier to discern the eyes, nose, and mouth, and arguably without making the subject's identity more recognizable (Figure 3e).

Blurring and redaction are also common methods for obscuring images, especially on post-processed television. Blurring smooths an image by allowing each pixel's value to be influenced by the values of the pixels around it. Including a larger neighborhood ("kernel") of pixels around each target pixel makes the image blurrier. Redaction is simply removing pixels from an image, yielding the familiar black box over the objectionable part of the image. Redaction benefits from an intelligent way to select objects in an image; one such method is GrabCut (Rother et al., 2004), which uses a partial classification of foreground and background pixels to intelligently divide the image in two along the object boundary. GrabCut has recently been extended to use depth as well as color information (Vaiapury et al., 2010).

Korshunov & Ebrahimi (2013) show the possibility of morphing images of faces so they are

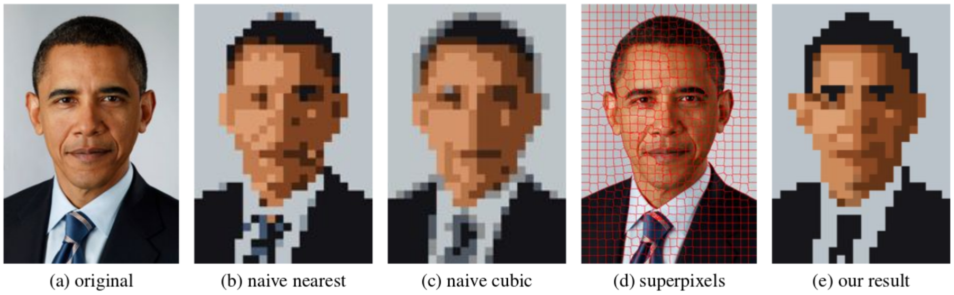

Figure 3. Reproduction of Gerstner et al. (2012), Figure 1. Original caption reads: *Pixel art images simultaneously use very few pixels and a tiny color palette. Attempts to represent image (a) using only 22 x 32 pixels and 8 colors using (b) nearest-neighbor or (c) cubic downsampling (both followed by median cut color quantization), result in detail loss and blurriness. We optimize over a set of superpixels (d) and an associated color palette to produce output (e) in the style of pixel art.*

unrecognizable and don't need to be blurred or redacted altogether. Building on that initial work, the method proposed by Nakashima et al. (2015) preserves facial expressions and is shown to work on a few realistic images.

Replacement is like redaction, but uses something more purposeful than a black box to cover the redacted areas. A familiar example is the use of a chroma key or "green screen" technique to replace the background of a movie set or news studio with a computer-generated, moving environment. If we know what sort of thing has been selected, we can do specialized replacements. For example, if a person detector returns the positions in the image of a person's joints, we could cover the person with a generalized cartoon in the same pose. Replacing an object with what "would be behind it" as described by Hubers et al. (2015) requires knowledge that a camera cannot give, since the object occludes that area. One way to circumvent this is to make educated guesses about what is behind the object based on patterns or colors in the unoccluded image. This is the strategy of so-called "inpainting" or "image completion" techniques like those by Sun et al. (2005), S. S. Cheung et al. (2006), Bugeau et al. (2010), and Herling & Broll (2012). Cheung et al. (2009) apply inpainting to surveillance video. The key idea here is that, in order to perform inpainting, one must somehow know what is behind the target object without actually seeing it. For certain scenarios, this will be impossible without more information.

That additional information could be recorded beforehand if the environment does not change much over time. Recently, sensors that fuse color and depth information have made it possible to build colored 3d maps; using a mapping framework like OctoMap (Hornung et al., 2013) with simultaneous localization and mapping as in RGB-D SLAM (demonstrated by Endres et al., 2012) could very well provide the knowledge for intelligent image replacement in semi-static environments.

Whereas our goal thus far has been to review methods for *obscuring* objects in images, most graphic artists have different goals in mind. They seek to make images simpler and less busy (Kyprianidis, 2011), or more attractive (J. Lu et al., 2010), or easier to understand (DeCarlo & Santella, 2002; DeCarlo et al., 2003). Nevertheless, the techniques they present could also be useful for protecting user privacy in robotic systems.

Such techniques are typically categorized as non-photorealistic rendering (NPR), as opposed to techniques that focus on fidelity to photorealism. So-called "painterly" techniques (e.g., J. Lu et al., 2010) use a variety of different brush-stroke effects to make images look more artistic, incidentally removing identifying details in the process. Image abstraction techniques such as that by Kypriani-

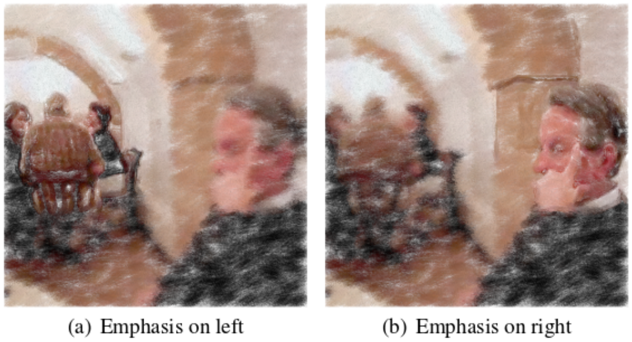

Figure 4. Reproduction of J. Lu et al. (2010), Figure 8. Original caption reads: *Placing emphasis via controlled stroke density.*

dis (2011) seek to retain the gist of an image and discard the distracting details. Kyprianidis (2011) uses the anisotropic Kuwahara filter, which aims to provide a consistent level of abstraction across different levels of detail while also being robust to even higher-contrast noise. Human faces are still quite recognizable (as intended) after applying this filter. DeCarlo & Santella (2002) add definition to an abstracted image by introducing certain of the detected edges that were removed during abstraction. The particular implementation by DeCarlo & Santella (2002) is not so practical for our purposes because it requires users to designate which areas need more definition. In fact, these areas are designated implicitly via eye tracking. One cannot help but wonder: would it be effective to do the opposite by starting with all the edges and *subtracting* the important ones in order to obscure features of interest? Conversely, perhaps removing color (and, therefore, textures) could also protect privacy. DeCarlo et al. (2003) present a way to represent 3d models of objects using only the surface contours and some "suggestive" extensions of those contours. The technique is automatic and can yield recognizable renderings of the models, but extending it to 2d images without depth information might yield results that are less satisfactory.

Artists, like magicians, are masters of attention control. They possess an arsenal of ways to make the viewer focus on something, or (germane to privacy concerns) to *not* focus on something else. Cole et al. (2006) present several methods for subtly manipulating which regions are emphasized in 3d drawings. This could be thought of as indirect obscuration, as the regions *not* emphasized experience dramatic reductions of detail. Of course, if obscuration is the goal, any user control over the "gaze direction" would have to be limited. Another compelling illustration (no pun intended) of emphasis control—here, by controlling stroke density in a painterly filter—is shown in Figure 4 (reproduced from J. Lu et al., 2010).

4.1.2 Image Manipulation Techniques: Evaluations. From the standpoint of privacy, the primary goal of image manipulation is to obscure private objects. Depending on the application, the secondary goal could be either to maximize the fidelity of the rest of the image (e.g., so the user can use the image feed for some purpose or task) or to hide the fact that manipulation has occurred, or both. From the standpoint of non-photorealistic rendering, however, the primary goal of image manipulation is to enhance or simplify the image; privacy is not necessarily being considered there, and the goals may conflict. Here we review some evaluations of how image manipulation impacts image perception by humans.

We frame this section with the model of (visual) privacy loss proposed by Saini et al. (2014) for multi-camera surveillance systems. They define privacy loss as the product of "identity leakage" (i.e., the probability of being identified) and a "sensitivity index" (i.e., how sensitive the information is that you're being identified with). Their model considers not just facial recognition, but also the "what" (clothes, gait, behavior), "when" (daily schedule), and "where" (location) as additional inference channels for identifying a person. The authors also discuss some privacy implications of *who* exactly is seen by a multi-camera surveillance system; e.g., if the same three people are always visible, it becomes easier to discern their identities because there are fewer possible identities to choose from.

Early evaluations of image manipulation to protect privacy are on simple filters in video media space applications. Zhao & Stasko (1998) presents five filtering techniques and evaluates them on short video segments in a privacy context. The filters were pixelization, edge detection, and three other techniques that were basically abstractions. All filters appeared to make the actors more difficult to identify without making it much more difficult to discern what they were doing. Subjects made an interesting comment: actors may be easily identified by shirt color even after heavy image filtering.

Boyle et al. (2000) are also concerned with filtering short video clips for privacy. Two filters—blur and pixelize—are each used at 9 different fidelity levels. It is shown that, for each filter type, there was an appropriate fidelity level for which privacy was protected and some basic awareness information (e.g., number of people in the room) was preserved.

Schiff et al. (2007) focus on reliable tracking of color markers that people wear in order to have their faces obscured in a video feed. They test their system at a construction site using high-visibility vests and hard hats for markers. The robustness of their system to lighting conditions and partial occlusions is examined.

Korshunov et al. (2012) present a crowdsourcing approach to evaluating privacy filters using Facebook. These online results are validated against an in-person, laboratory study. The study itself shows that privacy is most preserved by a redaction filter, followed by a pixelation, and least of all by a blur filter.

Halper et al. (2003) provide our introduction to the psychology of how people perceive NPR images. First, separating foreground from background is facilitated if everything in the foreground is rendered with one style of NPR and everything in the background is rendered in another. Perhaps private objects could be de-emphasized by rendering them like the background, and key objects could be intentionally emphasized by the reverse process. Second, people tend to perceive sketchy renderings as more open to change, and want to explore areas with a higher level of detail. These findings appear useful in mapping, both for visualizing (un)certainty about different regions and for guiding a user to key places (and away from irrelevant places, thereby protecting some privacy).

If we use NPR to remove private details or replace private objects seen by a RPS, will users still be able to complete their tasks? That is, will users still accurately interpret the scene? It is crucial to see the difference between artistic renderings, in which different users are expected to interpret the product in different ways, and functional renderings such as are found in technical manuals, which seek to communicate the same things to everyone (Gooch & Willemsen, 2002). NPR provides a "functional realism" that is well-suited for the latter purpose, but Gooch & Willemsen (2002) and Phillips et al. (2009) show how non-photorealistic environments can distort user perceptions of a scene. Gooch & Willemsen (2002) demonstrate that people typically underestimate distances in NPR immersive environments, e.g., while walking down a hallway. In a more extensive study, Phillips et al. (2009) compare a NPR virtual environment to a highly realistic one and confirms that a lack of photorealism in particular seems to cause the distance judgment error. The authors draw on earlier findings to make a non-intuitive point: it wasn't the *reduction of detail* that caused the misjudgments, but (so it appears) the *reduction of photorealism* itself. It seems, then, that using NPR to protect privacy could inhibit tasks for which distance judgments matter.

Perhaps the most important consideration for using NPR to protect privacy is how it affects the *discernability* of objects; that is, whether an object is there, what sort of object it is, and whether it's real or virtual. J. Fischer et al. (2006) address the last of these questions. The authors test a promising strategy: if an image containing both real and virtual objects is uniformly stylized, does it become harder to discern between them? Yes, the results showed that users had more trouble discerning virtual objects from real ones with this "stylized AR" technique. This is promising both for immersive, realistic AR and for convincing object replacement in a privacy protection system.

Several papers directly address using NPR to protect privacy. Erdelyi, Barat, et al. (2014) compare a cartooning (i.e., abstraction) filter to blur and pixelation filters in terms of privacy protection and utility. Abstraction provided the most utility and the least privacy, followed by blurring, then pixelation. The same authors present a modified version of their technique (Erdelyi, Winkler, & Rinner, 2014) as an entry into the MediaEval 2014 Visual Privacy Task (see Badii et al., 2013). Images containing people were abstracted once, then additional abstraction and pixelation filters were applied to just the faces. The filter performance was evaluated by crowdsourced survey responses. Output images were rated in terms of intelligibility, privacy, and pleasantness, and scored near the median performance for the competition.

4.2 Constraining Navigation

Robot navigation must be constrained if we are to prohibit the robot from entering certain areas. For example, bedrooms or children's play areas might be private and therefore off-limits for mobile robots. Several methods for enforcing these constraints are discussed below.

4.2.1 Motion Planning and Obstacles. Motion planning algorithms typically use a model of the obstacles in the world to restrict the valid planning space for the robot. In this framework, the "configuration space" includes all configurations of the robot in which it does not collide with an obstacle (LaValle, 2006, Ch. 4). This can be readily adapted to include areas that are private in the sense that the robot must not enter or pass through them; those areas can simply be designated as obstacles. This will have predictable effects on the plans produced by the algorithm. The new privacy "obstacles" blot out part of the search space and paths become either (a) less optimal, (b) undoable, or (c) the same as they would be otherwise. If the search space includes a temporal dimension, we can also model obstacles that toggle on and off or change size over time. For example, the bedroom may be off-limits only at night, or a morning person might desire more personal space in the early hours of the day. Besides the added dimension, the planning problem stays the same conceptually with the addition of time information; that is, as long as the time spent planning is trivial.

What if these private regions are not defined beforehand, or what if they change periodically? Here we can use algorithms for mapping and planning in an unknown environment. When the obstructed space changes (i.e., when a private region is turned on, turned off, or moved), the robot will have to use techniques summarized by Russell & Norvig (2009), Ch. 12.5–6, such as plan monitoring, replanning, and continuous planning. Private regions that move around, such as the personal spaces of people, might warrant the use of "differential" planning constraints (LaValle, 2006, Ch. 13), which regulate velocity and acceleration in addition to position. Obeying these constraints becomes more difficult for more complex robots, and frameworks such as the "whole-body control framework" for humanoid robots presented by Sentis & Khatib (2006) become necessary for planning natural, stable motions.

4.2.2 Semantic Maps. The maps used for constraining navigation in the above discussion were metric maps. Metric maps are to-scale, cartesian representations of a space. Time could be added as an additional dimension, but the map remains purely quantitative. The problem is, humans don't typically think about their environments in terms of numbers. With the exception of temporal information (e.g., “between the hours of 9 and 11 A.M.”), people typically use qualitative labels to refer to objects, not spatial coordinates. For example, “the document on my desk” is used in lieu of, “the 8-1/2in x 11in rectangular object of nominal RGB color value (113, 7, 24) located at (x,y,z) = (1.2m, -0.2m, 0.8m) from the...”. This suggests the utility of adding higher-level conceptual information to metric maps to form *semantic*, or meaning-laden, maps. The robot could store meaningful labels for persons, places, and things in the world. These labels could then be compared, grouped into hierarchical categories, and used to inform a robot’s decisions.

Galindo et al. (2005) present a semantic mapping framework. In the authors’ language, conceptual entities in the conceptual part of the map are “anchored” to locations, areas, or objects in the spatial part of the map. Some semantic labels can be assigned automatically, such as grouping spaces into “rooms” and “corridors” based on connectivity. The system can then make inferences such as, “this room with a bed in it must be a bedroom”. This framework makes robot navigation easier by allowing commands like, “go to the kitchen.” It also makes robot localization faster with inferences like, “I see a TV set, so I must be somewhere in the living room.” Galindo et al. (2008) show this same framework used in a series of robot task planning experiments. The semantic map framework allows the robot to make inferences from the state information and to shrink planning problems by raising the level of abstraction. Of special note is the robot’s ability to intelligently handle commands that are too specific; e.g., when told to go to the bedroom, the robot reasons that all the rooms it has found are candidate bedrooms and proceeds to search for discriminating evidence. This ability to generalize commands when unsure about their meaning is promising for privacy applications.

Semantic maps may be created without manual labeling by a human. Rusu et al. (2009) present a system for creating semantic maps automatically from 3d point clouds. That same system is implemented autonomously in a kitchen environment with a PR2 robot by Goron et al. (2011). The robot is able to explore, generate a hypothesis map, and then interact with appliances to verify and revise the map. These semantic mapping frameworks are promising for embedding privacy settings in the robot’s model of the world, which can then be honored by avoiding certain actions.

4.2.3 Rules for Moving amongst Humans and other Robots. Privacy can be thought of as including personal space, since violating one’s personal space violates both one’s solitude and one’s control over access to oneself. The robotics community has addressed personal space in mobile robot navigation via the notion of proxemics, introduced by Hall (1966) almost sixty years ago. Since proxemics was originally defined as being between humans, a new array of studies is needed to discover any differences in the human-robot interaction scenario. Such work exists; for example, J. T. Butler & Agah (2001) studied what types of approach behaviors make humans uncomfortable, and Takayama & Pantofaru (2009) included some human traits as factors when they studied this in more detail. Findings from such work will inform the ways we constrain robots to respect personal space. Several robot behaviors have already been implemented with personal space in mind: standing in line (Nakauchi & Simmons, 2002), following a person (Gockley et al., 2007), and passing a person in the hall (D. Lu & Smart, 2013; D. Lu et al., 2013, explained in detail below). Some other relevant but older studies include that by Asama et al. (1991) for robot-robot passing behaviors and another by Kato et al. (1992) for handling human-robot traffic through passageways and workspaces.

When a robot passes a person using the standard ROS navigation library, it often violates personal space by passing too close, possibly without slowing down. D. Lu & Smart (2013) change the standard costmap implementation in several ways (and further techniques are explained by D. Lu et al., 2013) to incentivize the path planner to give the human more space. Also, constraining the robot’s gaze direction matters in this scenario. D. Lu & Smart (2013) predict that the robot needs to look at the human at least once to acknowledge that his presence is accounted for, but a constant stare is "creepy." The user study concluded that the costmap manipulation increased passing speed for humans, but that an intermittent eye contact policy did not; gaze effects seem to be more complicated than they thought.

4.3 Constraining Manipulation

We have said that being sensitive of privacy restricts what the robot should see, where it should go, and also what it should touch. We now turn to that last element, perhaps the least-explored of the three: how can we restrict the robot from *touching* things in ways that would violate privacy? This could include touching personal possessions, objects in a person's territory, or even a person's body. Also, *touching* is too narrow here; some things may be inappropriate to pick up but OK to touch, whereas other things might be inappropriate even to reach towards or point to. All these actions fit within robotic manipulation, or at least motion planning for robot arms.

During autonomous operation, applying privacy constraints to robotic manipulation could mostly be handled either by modeling private objects as obstacles (see Section 4.2.1) or by labeling them appropriately (e.g., "NoTouch" or "NoGrab"; see Section 4.2.2). In the first case, a typical trajectory planner would plan around the private object. In the second case, a high-level, semantic planner would reject plans that include illegal actions (e.g., "Touch the Book that has label NoTouch").

During teleoperation, an obstacle-based approach could still be used to simply take away the user's control whenever some action would violate privacy. This does not, however, give much feedback to the user. Perhaps an improved method would use a haptic device to deliver force feedback to the user when the robot nears a restricted area. The privacy boundary could exert an increasing amount of normal force via the haptic device as the user nears the boundary's edge. Rydén presents a relevant framework called "forbidden-region virtual fixtures" in his doctoral dissertation (O. F. Rydén, 2013). His work enables remote touching of moving (F. Rydén & Chizeck, 2012) objects that considers both the position and orientation of the remote toucher (F. Rydén & Chizeck, 2013)—in our case, a robotic end effector. Perhaps this system could be used in assistive robotics, where humanlike robots might need to touch and move patients while dynamically maintaining appropriate hand positions.

5. Existing Work in Privacy-Sensitive Robotics

Some work has already been done towards making robots privacy-sensitive at the time of writing (January 2016). Here we present a brief overview of the papers we consider to be privacy-sensitive robotics work, i.e., that directly address an aspect of privacy within the larger field of robotics.

5.1 Visual Privacy

An initial concern has been the tradeoff between visual privacy and *utility*, or usefulness of an interface towards its intended purpose. Jana et al. (2013) present a privacy tool for simplifying videos to only the necessary information. Although no user study is conducted, the tool is shown to preserve utility and protect privacy through a simple analysis. Raval et al. (2014) present two more video privacy tools. One uses markers to specify private areas, and the other uses hand gestures for the same purpose. Hubers et al. (2015) and D. Butler et al. (2015) evaluate the privacy-utility tradeoff in the context of robots. Hubers et al. (2015) tested a patrol surveillance task with the Turtlebot 2 robot, and D. Butler et al. (2015) tested a pick-and-place task with the PR2 robot. Both studies found it feasible to complete the tasks with effective privacy filters in place.

Rueben et al. (2016) give an initial report of a controlled user study that compares three different interfaces for specifying visual privacy preferences to a robot. Their results suggest that usability differences between interfaces depends on the *scenario*, e.g., a point-and-click GUI is quick and easy until the robot needs to be teleoperated to a new viewpoint. The authors suggest that future work could include testing the impact of visual filter fidelity (i.e., whether a blur effect ever flickers) as well as of the apparent quality of the privacy-specifying interface (i.e., whether it appears to be professionally made) on how much users *trust* the privacy protection system.

Two additional works study privacy concerns for robots operating outside the laboratory. Lee et al. (2011) interview participants who have interacted with a social robot in the workplace. They found that participants had trouble inferring the robot's sensing capabilities from its appearance. Many participants expressed a desire to be *notified* when the robot was recording data about them, since they could not tell. The authors call for "privacy-sensitive design" of social robots. Caine et al. (2012) test whether older adults use privacy enhancing behaviors when monitored by a camera, a static robot, or a mobile robot. Some results suggested that mobile robots evoked *less* privacy enhancing behavior, which was counter to the authors' expectations.

5.2 Proxemics

We have already referenced several works in Section 4.2.3 that could be considered privacy-sensitive robotics work in the area of *proxemics* (e.g., Takayama & Pantofaru, 2009). Proxemics studies count as privacy-sensitive robotics when the goal is understanding how robots can move and position themselves to promote human comfort and acceptance. Positioning optimally for performing a co-operative task with a human does not count by itself. Mumm & Mutlu (2011) include *psychological distance* in their definition of proxemics in reference to Hall (1966).

The study by Mumm & Mutlu (2011) concerns both physical and psychological distancing. E.g., they showed that participants who disliked the robot maintained a greater physical distance when the robot was looking at them. The eye contact operationalized psychological proximity. The authors also propose a model of both physical and psychological distancing in HRI.

Okita et al. (2012) report initial findings with a legged humanoid robot, which may, they point out, be perceived differently than wheeled robots. The robot walked toward participants with differing verbal and non-verbal cues, and the reactions of both adults and children were measured in terms of movement and physiological arousal level.

Joosse et al. (2013) seek to validate and refine a measure of social normativeness for robot behaviors. The instrument measures both attitudinal and behavioral responses by humans. The validation experiment was a HRI in which a robot invaded someone's personal space.

Henkel et al. (2014) report on an experiment in which a tank-treaded robot approached participants acting as victims in a simulated disaster scenario. The robot's velocity, head movements, and headlight brightness were scaled based on proximity to the human. The results suggested that scaling behaviors according to certain physical and psychological rules increased participant comfort.

5.3 Territoriality

Territoriality is a construct of great interest in the social sciences (e.g., in Altman, 1975, discussed above in Section 3.4.1), but has been studied very little in the HRI domain. One work, by Satake et al. (2014), develops and tests a model of territory in front of shops. The authors show that people preferred chatting with a robot that understands which territory belongs to the store. Much more exploration of territoriality in HRI is warranted.

6. A Roadmap for Privacy-Sensitive Robotics

In this section, we propose a roadmap for privacy-sensitive robotics, setting out what we see as the most important research questions that need to be addressed. We begin with *basic research* that will be foundational for any privacy-sensitive robotics application, then list some *applied research* ideas that begin to approach real-world solutions that actually promote privacy in HRI.

6.1 Basic Research

6.1.1 What is Privacy? A Privacy Taxonomy for HRI. The first question that we need to answer for privacy-sensitive robotics to move forward, before we even begin to think about robots or the technology involved, is what we mean by the word "privacy." As we have said in Section 3.1.4, there are many different ideas that could be described as "privacy," from the common notion of having information you don't want revealed to the notions of personal space and solitude. We propose to identify a common vocabulary and create a taxonomy of privacy, in the context of robots and robotics technologies. Once we have this in place, we have a language to start talking about what privacy is and is not in a particular context. More importantly, we can be *precise* about what we mean in a given context, avoiding ambiguous interpretations of what we are trying to do (and the resulting confusion about whether or not we succeeded). This paper serves as a starting point for this vocabulary and taxonomy in that it collects ideas of privacy from a number of different areas. However, the next step is to systematize this information and to give it a structure, so that the relationships between the different definitions and understandings become explicit.

In order to study privacy-sensitive robotics, we must do human-subject experiments; tackling a human-robot interaction problem without consulting the humans is a doomed endeavor. Testing hypotheses about "privacy" is impossible, though, without being much more specific and choosing just a small part of "privacy" to work with. We need to break down "privacy" into smaller constructs (i.e., ideas we cannot directly observe) organized into a hierarchical taxonomy based on a review of the privacy literature. More specifically, we need constructs that are specific enough to be operationalized into unambiguous, concrete, observable measures. For example, perhaps the broad construct "privacy" can be divided into several narrower constructs, one of which is "informational privacy," which itself can be further divided into several even narrower constructs, one of which is "personal information collection." Each of these three constructs would warrant careful definition. This would result in a three-level, hierarchical taxonomy that, starting with a general construct, names and defines progressively more specific sub-constructs that are easier to operationalize into measures. Here, for example, one operationalization of "personal information collection" might be whether someone knows your social security number—a simple, binary measure. This process—constructing a valid taxonomy, operationalizing constructs into measures, and conducting careful experiments with an eye for validity—is a prerequisite for meaningful progress in understanding privacy in human-robot interactions.

6.1.2 How do People Think of Privacy? Identifying Key Factors that Mediate Perceived Privacy Violations in HRI. Armed with a way to talk about privacy, the next step is to determine what people think about privacy in the context of robots. What are their concerns? What types of privacy violations worry them? Is it even something that they have thought about?

Indeed that is the first question: we need to look at how much value people put on privacy protection, if they even want it at all. Many people are willing to trade privacy violations for convenience—using location services on their mobile phone, for example. Understanding the (perceived) cost of privacy violations in relation to the (perceived) benefits of a privacy-violating service will further help us understand what we need to work on, and how important it is to people. This will let us focus our finite resources where they will have the most impact.

When a person feels that his or her privacy has been violated when interacting with a robot, it is important to ask which aspects of the person, the robot, or the surrounding situation influenced that feeling. Here we are concerned with subjective or perceived privacy, not, if such a thing exists, objective privacy. There are probably multiple factors that influence this perception for each facet of privacy; for example, Takayama & Pantofaru (2009) give experimental evidence for several factors that impact personal space, including experience with pets and robots, robot gaze direction, and the sex of the subject. It is important to identify what key factors mediate perceived privacy violations so robot designers with good intentions can avoid offending users. Beneficent designers like these would also want to know what sorts of scenarios give people a *false sense of privacy* so as to avoid them; an informed awareness about this would also help policymakers and judges make better legal decisions about robots.

Answering this question would involve first taxonomizing the idea of "privacy" as described in the previous section, then enumerating and testing all the plausible mediators for each construct. Mediators could be part of the human, such as personality traits, experiences, or demographics; the robot, such as morphology, behavior, or appearance; or the surrounding context, such as the task at hand, the way the robot is introduced or framed, environmental factors such as ambient noise or temperature, and so on. The goal is to identify the controlling mediators for each construct under the "privacy" umbrella. We want to be able to accurately predict which types of privacy will be of concern for a given HRI scenario, or, conversely, to recommend ways to minimize a certain kind of privacy violation.

We predict that *framing* will be an especially important factor to study. Can privacy concerns be raised or lowered by telling the right story about the robot (explicit framing), or by designing it to look a particular way (implicit framing)? When is framing ethical or unethical, and how much should we disclose the capabilities of the robot system? How do a person's *a priori* privacy concerns change with framing? Once the scenario has been framed, how do people's perceptions of the robot system change as they interact with it? How quickly do people move through these different states of belief about the system?

It will be important to conduct long-term studies that look at how people's privacy concerns change as they become used to robots and robotic technologies, especially as these systems are integrated into their daily lives. While these studies are important, they are also problematic to carry out. Embedding a robot into someone's daily routine in an unobtrusive way will be tricky. Each robot will have to work flawlessly for months at a time. It will have to be integrated into the study participant's life in such a way as to pose a privacy risk but without getting in the way. We will also have to develop instruments to measure how each subject's privacy concerns change over time.

Determining how people think about privacy over the long-term and in the presence of framing and habituation will help us prioritize the technical work that will have to be done in this area. Privacy covers a lot of ground but there are, we claim, likely to be some areas that are of more pressing concern and we should work on these first. This is especially true as we see real robots, such as Jibo (*Jibo*, 2016), start to enter people's lives. If we continue to deploy these systems without understanding and thinking hard about the privacy implications, it will go poorly.

6.1.3 How do we Evaluate Privacy? Developing Standard Scenarios and Measures for Privacy-Sensitive Robotics. Existing work in privacy-sensitive robotics (see Section 5) has not included much collaboration. In particular, each of the user studies conducted thus far has featured a unique experimental design, custom-made for the specific purposes of that one study. We know of no *replications* of privacy-sensitive robotics studies. The standard taxonomy of privacy terms discussed above is crucial for ensuring that researchers are talking about the same things, but two further standardizations—of *scenarios* and of *measures* of privacy—will also help promote the collaboration between researchers that is so essential.

By *scenario* we mean both the physical environment and how participants in a user study understand it. This includes the briefing given to participants, the type of space (i.e., public or private, large or small, home or office or outdoors), the surrounding sights, sounds, and smells, and of course the type of robot being used and its behaviors. There are two purposes for developing standard *scenarios* to be reused between studies. First, in basic research, we want to perform valid replications to confirm findings. We also want to check whether results *generalize* to different types of people (part of *external validity*), which requires holding the scenario constant. Second, in applied research, we want to make valid comparisons between privacy protection systems. To say that system A is better than system B requires that we test them in comparable scenarios.

A standard privacy *scenario* could be any human-robot interaction in which privacy is a concern. Of course, it should be meaningful to real-world applications and reasonably easy to replicate. An example scenario would be for a janitorial robot to clean a bathroom that might have people in it. We could define a series of tasks to be completed in order (e.g., clean the toilets, polish the mirrors, take out the trash) and one or more configurations of human occupants (e.g., sitting in a stall, washing one's hands, walking in the door). A given scenario could be implemented in a natural setting, a laboratory setting, or in a computer simulation, perhaps with humans participating via virtual reality. These scenarios can used both as a series of challenges to motivate work in privacy-sensitive robotics, and as benchmarks for performance testing like toy domains are in the field of machine learning. We add as a warning, however, that privacy-sensitive robotics promises to be an especially hard area for making replicable scenarios. Certain kinds of privacy—territoriality and personal objects come to mind—seem to depend on the *relationships* people have with spaces and objects. Controlling these relationships between replications of an experiment or even between subjects in a single experiment seems especially difficult in this area of research.

This brings us to *measures* of privacy-sensitivity in robots, which we hold to be equivalent to privacy upheld and violations avoided for humans. These measures will be common, reusable operationalizations of the privacy constructs defined by the privacy taxonomy that we recommended above. It is important to establish and validate measures that can be reused by multiple researchers for replication or convenience. *Objective measures* will be concerned with the robot's actions and with the situation as it physically occurs, whereas *subjective measures* will be concerned with how real people think and feel. Both types are useful, but when constructing the latter we should first review the experimental social psychology literature (e.g., Nisbett & Wilson, 1977) for best practices and additional techniques that are not yet well-established in HRI research. For example, long-term studies might involve ethnographic observation and experimental settings that are less controlled than we are currently comfortable with as a field.

We will also want to pay attention to the distinction made by Friedman et al. (2009) between the "watcher" and the "watched," as well as a new, third role, the robot *operator*, who is probably a "watcher" but also actively controls the robot[5]. People in each of these three roles could experience privacy differently in a given human-robot interaction, so good measures will distinguish between them. It will also be important to control the users' level of understanding of the robot's abilities and lifelikeness. This is one area in which scenarios and measures converge, since the briefing given to participants will affect what is being measured, and with what validity. Uncontrolled framing and poorly-worded questions can easily skew the results, especially for experiment designers not intimately familiar with these instruments (like most roboticists).

6.1.4 What are our Tools? Implementing Privacy Protection on Real Robots. Once we understand people's privacy concerns and how they evolve, we can start to map these concerns onto technology that protects their privacy. If people are, for example, concerned with their image being seen by a remote robot operator, we can draw on the field of computer graphics to provide visual filters (redaction, blurring, or other more sophisticated techniques; see Section 4.1.1) to obscure the sensi-

[5] We owe the observation that a single individual could hold several of these three roles to Ross Sowell.

tive parts of the image while retaining enough information for the remote operator to complete their task. Similar techniques from other fields could be used to address other types of privacy concerns.

We will also have to implement all of these ideas and verify that they work as expected. As above, this will have to be done in long-term experiments to determine if they are actually useful in a real-world context, with all of the operational problems outlined above. Along the way, we will also have to integrate them with real robot systems, and deal with all of the problems that they introduce: limited computation, error-prone localization, imperfect object recognition, and all the rest.

6.1.5 Privacy-Sensitive Robotics and other Fields. A fundamental question to ask is how privacy-sensitive robotics interacts with robotics application areas such as telepresence and industrial automation. Is it similar to a field like computer security, where there is a core of technical research into algorithms and techniques as well as a specialized application of these techniques in different contexts like banking, internet browsing, and file encryption? At a first glance, this metaphor seems to be appropriate. There are core technical elements to privacy-sensitive robotics, such as the algorithmic blurring of faces in an image stream. These elements are then applied in a particular context, such as a telepresence system, as appropriate. Not all elements are appropriate for every context, just as with computer security. An understanding of the context and of potential privacy harms will influence which techniques we choose to apply similar, again, to computer security. We believe that this metaphor does not quite fit, however, since it suggests (perhaps implicitly) that one solution is right for all users. In computer security, most browsers use 128-bit encryption, even if the data are not particularly sensitive. When thinking about privacy, however, we believe that the individualization of privacy protections is much more nuanced.

With this in mind, we propose *accessibility* as a model on which to base our thinking about privacy-sensitive robotics. Accessibility refers to the design of artifacts and services that can be used by people with disabilities. It is not a single set of techniques, but a collection of designs, features, and approaches that can be combined to accommodate any particular set of disabilities. No two people with disabilities are exactly alike, just as no two people have the same set of privacy concerns. There are, however, broad classes of disabilities, just as (we expect that) there are broad classes of privacy concerns. Accessibility is improved when the accommodations line up with the particular disabilities, just as (we claim) privacy protections will be more satisfying when the particular technical measures align with an individual's privacy concerns. Finally, just as with accessibility, the final decision on whether or not a feature is useful lies with the person who is using it, not with some objective measure.

We believe that using accessibility as a model will help us think about privacy-sensitive robotics and, in particular, how it interacts with application domains. Privacy-sensitive robotics needs to be studied in its own right because, though it might look different when applied to different situations, there is core, general knowledge to be gained as well. This is why accessibility is treated as a standalone field, and privacy-sensitive robotics should be as well. On the other hand, privacy-sensitive robotics is barren until applied to a specific privacy construct and a specific context. In fact, *applying* the general knowledge and best practices in a context-aware way is an especially large part of privacy-sensitive robotics, just as it is for accessibility. We might have guessed these things without using accessibility as a template, but now we can use this template to answer questions about how privacy-sensitive robotics should *progress* as a field, which might be unclear due to its unique nature.

6.1.6 Challenge Problem Frameworks. Another question to consider is whether we should have large challenge problems in privacy-sensitive robotics. In other areas, challenge problems have helped guide the community, focus resources, and provide a common purpose. It is easy to imagine challenge problems being useful here for some of the technical details, such as building a faster face blurring system for image streams. Once we move past implementation details to whole system performance in real scenarios, however, things become harder. Privacy varies from person to person—does posing individual challenge questions even make sense? Does making the problem general enough to pose as a challenge also rob it of its usefulness in the real world?

Part of the problem lies in the evaluation of privacy protections. In many areas of robotics, there is an objective measure of performance by which you can compare two systems. In robot localization, you can measure how far your position estimate is from the actual position of the robot. In the DARPA Grand Challenge, you can measure how long it takes for the robotic vehicles to travel a pre-specified course. When assessing privacy protections, it is harder to have such a crisp, automatically-calculated metric. This makes designing challenges hard.

This puts the onus on us to rigorously define what we mean by privacy protection in a particular context, and to come up with ways to assess this. Depending on the case, assessment could be objective or, via the use of human judges, subjective. This highlights the need for a taxonomy of privacy terms, as we have already discussed above. Once this taxonomy is in place, we will be better-equipped to design challenge problems for privacy-sensitive robotics.

If challenge problems do prove useful in this field, they should account for the fact that privacy is inherently an adversarial notion. One person is trying to protect their privacy, while another is trying to violate it, although perhaps not intentionally. This naturally leads us to think of privacy challenges as being two-sided: one side trying to protect privacy in some limited, well-specified context and the other trying to violate it by overcoming the protections. In addition to making for a more exciting challenge, we believe that this sort of challenge will better test our technology by exposing the weaknesses in different approaches. As before, however, it relies on a rigorous way to evaluate both the protection and violation of privacy.

6.1.7 Re-thinking Intentionally Anthropomorphic Robots. We often program anthropomorphic behaviors into robots as aids to human understanding and interaction. A robot might point its head at you to signal attention or scratch its head while it is processing something. Those cues are easy for humans to understand, as well as attractive and compelling. Humans like to paint their humanity onto other things; we are artists, creators of meaning. But making robots humanlike also causes a class of problems grounded in the reality that they are fundamentally different from humans. Anthropomorphic behaviors encourage people to project human characteristics onto robots, even ones that are not the case. For example, imagine a robot vocalizing that it is planning a path with a thoughtful, "hmmmmm!" Just using its voice might make humans think it can speak, or even understand speech. Even robots that can understand and speak some words can only do so to a very limited extent, but new users don't know that, and might assume that it understands and responds to unstructured speech. This is just one example of how anthropomorphic behaviors—here, speaking—can cause users to *infer* things about the robot that aren't true, thereby stumbling into a misunderstanding. Richards & Smart (2012) call this mistake "The Android Fallacy," and add that it can also lead lawmakers to be too lenient towards robots that look or act like humans because of the illusion of free will or fallibility. The authors argue that we must remember that robots are machines, and ought to be held to machine standards.

"The Android Fallacy" also causes problems in settings where privacy matters. Anthropomorphic behaviors might trick users—perhaps intentionally—into believing that a robot's sensors have the same limitations as human sensors do, or that a robot is socially aware and privacy-sensitive when it isn't. For these reasons, privacy-sensitive roboticists should re-examine the practice of anthromorphization in robotics. Particularly, we are interested in the adverse effects for privacy of framing robots as humanlike (see Darling, 2015). What sorts of robot appearances, behaviors, and descriptions cause the false inferences that constitute "The Android Fallacy?" When does this become a privacy risk? In those cases, can we think of non-anthropomorphic ways to accomplish the goals of the HRI? Finally, do any themes emerge from seeking alternatives to anthropomorphic robot design?

This research direction can be expanded beyond the issue of anthropomorphization. In fact, we are concerned about any wrong conception of a robot formed by users. Perhaps the more general question is, what sort of robot appearance, behavior, and description encourages observers to form an accurate mental model of how the robot works? Is it a good idea to reuse pre-formed mental models (e.g., of a human, a dog, a calculator, a hammer) as local approximations of what robots are like in certain scenarios, or do we need to foster a mental model completely unique to robots? If the latter, how might that model look?

6.1.8 Can Privacy Protection Make Privacy Worse? Many of the privacy protection methods given in Section 4 and in the applied research directions below have the potential to call new attention to private objects, regions, or people. If a robot blurs out a particular person or avoids a certain room, onlookers and operators might begin to wonder why that person or room is so important. If the robot avoids an object that the remote operator can't see, that person might even be able to infer its location. What's worse is, these inferences of value, presence, or location all help a malicious user to try breaking through the privacy protection. Research is required to investigate when and how this phenomenon occurs, and how to prevent it when possible.

6.2 Applied Research

6.2.1 Privacy Warning Labels for Robot Transparency. We want privacy-sensitive robots to be *transparent* about their actions and inner workings. This means that what the robot *appears* to be doing and thinking matches what it is *actually* doing and thinking. We also want to make this clear and obvious to human observers, with a minimum of possible interpretations.

Pursuant to these goals, future research should investigate using a standardized labeling system to disclose privacy risks to people that are around robots. Sometimes it's hard to tell what a robot is capable of; labels on the robot's outer casing could indicate whether the robot can record video or sound, recognize peoples' faces, or connect to the Internet. If the robot's capabilities depend on which software packages are currently active, an outward-facing screen on the robot's body could indicate whether, e.g., the robot is programmed to respect personal space as it speeds down the hallway. These labels could also be broadcast to nearby devices—mobile devices of nearby users or devices fixed to the walls of robot-inhabited spaces. Users could also interact with the labels, disabling certain capabilities by touch either on the robot itself or on a separate device. Each label could light up when that particular capability is in use, e.g., when the robot is looking for faces it recognizes. This system might inspire the sort of easy dialogues we want in a privacy protection interface: users can be warned of risks, adjust preferences, and even receive pushback when preferences conflict or certain robot capabilities are too important to disable.

Research is required to develop a visual language for all the robot capabilities and possible harms relevant to privacy. Prototypes of this active privacy labeling system should be tested with real users in realistic scenarios for usability, understandability, and trust that privacy preferences are being honored. At first glance, *standardization* seems to be the key here for helping members of a diverse public understand a broad spectrum of robots.

6.2.2 Graphical Interface Elements and Behaviors for Robot Transparency. Besides an explicit labeling system, robots can also become more *transparent* through what they make apparent to humans through graphical interfaces and robot behaviors. This could increase user trust in a privacy protection system by allowing users to *see* their preferences visualized or acted out by the robot. Users could also see sensor data and data products to better grasp the robot's sensing modalities and limitations. For *visualization*, augmented and virtual reality (AR and VR) devices could help immerse users in the world as seen and processed by the robot. The robot could also project images into the environment, especially flat surfaces nearby, such as the ground around its immediate footprint. Regardless of display modality, visualizations help users confirm their privacy preferences and open up the robot to increased introspection.

Robot *behaviors* can also make privacy protection more transparent. For example, the robot can avoid private objects and regions more obviously, making it clear that the privacy restriction is changing what would be the robot's normal behavior. Here perhaps we can apply work on legible robot motion, e.g., by Dragan & Srinivasa (2013) for manipulation and as reviewed by Lichtenthäler & Kirsch (2013) for navigation. This might be especially important for respectful manipulation, since it may be inappropriate not only to touch something, but even to reach towards it. Research is required to evaluate the application of legible motion to privacy protection, however, since legible reaching and respectful reaching, for example, might not be perfectly correlated. Other transparency-promoting behaviors might include turning off lights on a sensor and pointing it in a useless direction to show that privacy settings are being obeyed. Research in graphical and behavioral transparency is required to find the best ways to help users understand the robot's actions enough to be confident about the privacy protection system, assuming that confidence is deserved.

6.2.3 Robust Visual Privacy Protection for RPS Robots. Another future research direction is to make a robust visual privacy protection system for Remote Presence Systems (RPSs). The goal is to allow users to mark people, objects, and areas to avoid, and the RPS will honor those privacy settings. Part of this research would be on the effectiveness of physical markers vs. GUIs for specifying private things. Rueben et al. (2016) give some initial, qualitative findings, but many questions remain unanswered. Which types of interfaces are easier to use over long periods of time? Which promote trust that the privacy protection system is working? How can interfaces make privacy settings clearer, and when are there opportunities for mistakes? If physical markers are used, can the robot robustly acquire and reacquire them to continuously honor those privacy settings?

This new research could also aim for the most complete privacy protection technique: image replacement. As introduced above in Section 4.1.1, OctoMap and RGB-D SLAM could be used to maintain a somewhat high-fidelity 3d color map of the robot's environment. This map could be built with private persons and objects absent, and then used to provide convincing replacements when such things are present.

It is impossible to *guarantee* privacy with any real system, and *probably private* might be a good enough setting for some users. A probabilistic framework could be used to provide these *quantified guarantees* based on user preferences. Such a framework would use the robot's certainty about which regions of its video feed are private to decide how aggressively to apply filters. If the position of a private object is modeled as a Gaussian probability distribution, a privacy filter could be centered at the mean position and inflated beyond the object's actual size to provide the user with more certainty. This would filter the entire screen whenever the robot is not localized within its map, which is a rational action given those circumstances.

User studies would be necessary to answer our research questions for the replacement filter and also the probabilistic privacy framework. First, do they work? Are they seamless, usable for various tasks, and convincing to the remote operator? Does the local user trust that privacy is being protected? How does the local user know when to *trust* the system—does this require extra feedback channels?

6.2.4 Respecting Privacy via Personal Space. This is more complex than keeping beyond a constant distance from each person; here we again refer to the work by Altman (1975) and Burgoon (1982), discussed above in Section 3.4.1. Personal space changes with situation. Elbows can touch in a crowded hallway, but not in a deserted alleyway. Certain activities, like exercising or doing delicate work, demand extra space, whereas others, like dancing or taking in a view, invite close company. Emotional states, too, communicate the need for space or openness to proximity. There is also evidence for various sex differences in personal space (see La France & Mayo, 1979, for a review). Territoriality probably matters here as well: the situation will change when it becomes *my* house, *my* room, *my* desk, or *my* computer. A personal space-sensitive robot should be able to perceive all these factors, reason about the situation, and move so as to respect the personal space of all humans present given the context.

How, then, would the spatial preferences be implemented when we need to plan a path to a goal? D. Lu & Smart (2013) has already been cited in Section 4.2.3 for modifying cost maps to account for personal space. More work is needed to account for all the contextual factors, several of which are mentioned in the previous paragraph. Since human actions are difficult to predict and the situation can change suddenly, the robot might need to replan midway through a movement; we will need an anytime planner. The robot also needs to deal with the inevitable awkward situations when personal space requirements conflict and violations occur. Perhaps gestures or utterances familiar from human social encounters would be useful here (pending the resolution of concerns in Section 6.1.7), such as averting the eyes or saying "excuse me" as appropriate. The research aspect of this idea answers the questions, "does it work?," "does it help?," and, "how can we make it better?"

6.2.5 Audio Privacy: Initial Exploration. Personal privacy can be violated just by listening. Some conversations are private, some sounds are embarrassing, and sometimes people don't want to be heard. In one sense, protecting audio privacy might seem easier than for visual privacy. Robots can probably operate without sound in more applications than without images, so the easy solution is to turn off the microphone in sensitive times or areas. When this is not possible, however, audio privacy may well prove to be more difficult than it seems to be. For instance, notice the *asymmetry*: a robot could be listening to a conversation from the next room without the conversants being able to hear (if it's quiet) or see the robot. Vision, on the other hand, is usually symmetric. Asymmetry certainly poses problems for the *transparency* protocols discussed above: from the next room, the conversants would not see the "audio recording enabled" warning label illuminate, nor would they see the robot's microphone extend towards the connecting doorway. Restoring symmetry here might cause other problems: e.g., a robot that hums so loudly that you can hear it from the next room would not only interfere with its microphone, but would also be loud and annoying. Short of shouting a warning ("I can hear you talking in there!"), we need transparent and socially acceptable ways to enforce audio privacy protection to minimize accidental or malicious violations.

7. Summary

The purpose of this paper has been to introduce the emerging subfield of privacy-sensitive robotics. Two in-depth surveys were conducted, one of the concept of privacy and one of robotics techniques that could be used for privacy protection. The survey of privacy began with definitions, then outlined the history of privacy in philosophy and U.S. law. Next, an array of studies in the social sciences were presented before closing with a review of privacy in the technology literature. The survey of robot constraints was divided into three parts—perception constraints, navigation constraints, and manipulation constraints—and was presented in light of a need for privacy-based restrictions on robot behavior. Future work in privacy-sensitive robotics was also suggested. This was divided into basic research, which addresses questions relevant to *any* concern within privacy-sensitive robotics, and applied research, which develops and tests concrete solutions in specific scenarios.

Several themes emerged. First, we saw that the word "privacy" is variously defined. There is no unanimously accepted theory of privacy, but most theorists acknowledge that "privacy" refers to more than one idea. Hence, it is very important for privacy-sensitive robotics researchers to give a specific definition for each privacy-related construct being used. Second, we saw that privacy research has been done in many different fields—e.g., law, psychology, economics, and computer science. Privacy-sensitive robotics researchers will benefit from connecting with several of these existing trees of research as they begin making their own contributions. Third, most privacy constructs are *subjective*; the same scenario might violate some people's privacy, but not others'. Thus, *user studies* are necessary, followed by careful analysis. Making broad generalizations is especially dangerous in privacy research. Fourth, privacy-sensitive robotics is only just beginning to be explored by researchers, and it appears that many well-defined and useful research projects can be started right away. We presented our own ideas in the previous section, including some considerations about how this unique sub-field should progress.

We feel privacy is a real and pressing problem in human-robot interaction. The more it is taken seriously, and soon, the better. The problem seems likely to continue growing as technology advances. Thus, an opportunity looms over privacy-sensitive roboticists: to make a good contribution here is to perform a great service for human society.

References

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*(1), 26–33.

Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, *42*(2), 249–274.

Allen, A. (1995). Privacy in Health Care. In W. T. Reich (Ed.), *Encylopedia of bioethics* (Vol. 4). New York: Simon & Schuster.

Allen, A. (1998). Coercing privacy. *Wm. & Mary L. Rev.*, *40*, 723.

Allen, A. (2011). Privacy and Medicine. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2011 ed.).

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, and crowding.* Monterey, CA: Brooks/Cole Publishing Company.

Altman, I. (1977). Privacy Regulation: Culturally Universal or Culturally Specific? *Journal of Social Issues*, *33*(3), 66–84.

Applegate, M., & Morse, J. M. (1994). Personal Privacy and Interaction Patterns in a Nursing Home. *Journal of Aging Studies*, *8*(4), 413–434.

Aristotle. (2004). *Politics* (B. Jowett, Ed.). NuVision Publications, LLC.

Asama, H., Ozaki, K., Itakura, H., Matsumoto, A., Ishida, Y., & Endo, I. (1991). Collision avoidance among multiple mobile robots based on rules and communication. In *Proceedings of IEEE/RSJ International Workshop on Intelligent Robots and Systems* (pp. 1215–1220). IEEE.

Atzori, L., Iera, A., & Morabito, G. (2010, October). The Internet of Things: A survey. *Computer Networks*, *54*(15), 2787–2805.

Austin, L. (2003). Privacy and the Question of Technology. *Law and Philosophy*, *22*(2), 119–166.

Badii, A., Einig, M., Piatrik, T., & others. (2013). Overview of the MediaEval 2013 Visual Privacy Task. In *MediaEval.*

Bankston, K. S., & Stepanovich, A. (2014). When robot eyes are watching you: the law & policy of automated communications surveillance. In *Proceedings of We Robot 2014.* University of Miami. (draft)

Barabas, C., Bavitz, C., Matias, J. N., Xie, C., & Xu, J. (2015). Legal and ethical issues in the use of telepresence robots: best practices and toolkit. In *Proceedings of We Robot 2015.* University of Washington. (draft)

Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2009, January). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, *1*, 71–81.

Becker, F. D., & Mayo, C. (1971). Delineating personal distance and territoriality. *Environment and Behavior*.

Beer, J. M., & Takayama, L. (2011). Mobile Remote Presence Systems for Older Adults: Acceptance, Benefits, and Concerns. In *Proceedings of the 6th International Conference on Human-robot Interaction* (pp. 19–26). New York, NY, USA: ACM.

Bellotti, V., & Sellen, A. (1993). Design for privacy in ubiquitous computing environments. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 1317 September 1993, Milan, Italy ECSCW93* (pp. 77–92). Springer.

Berendt, B., Gnther, O., & Spiekermann, S. (2005). Privacy in e-commerce: stated preferences vs. actual behavior. *Communications of the ACM*, *48*(4), 101–106.

Bloustein, E. J. (1964). Privacy as an aspect of human dignity: An answer to Dean Prosser. In F. D. Schoeman (Ed.), *Philosophical dimensions of privacy: An anthology.* Cambridge University Press.

Boyle, M., Edwards, C., & Greenberg, S. (2000). The effects of filtered video on awareness and privacy. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 1–10). ACM.

Boyle, M., Neustaedter, C., & Greenberg, S. (2009). Privacy factors in video-based media spaces. In *Media Space 20+ Years of Mediated Life* (pp. 97–122). Springer.

Buchanan, T., Paine, C., Joinson, A. N., & Reips, U.-D. (2007). Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology*, *58*(2), 157–165.

Bugeau, A., Bertalmo, M., Caselles, V., & Sapiro, G. (2010). A comprehensive framework for image inpainting. *Image Processing, IEEE Transactions on*, *19*(10), 2634–2645.

Burgoon, J. (1982). Privacy and communication. In M. Burgoon (Ed.), *Communication yearbook 6.* Routledge.

Butler, D., Huang, J., Roesner, F., & Cakmak, M. (2015). The Privacy-Utility Tradeoff for Remotely Teleoperated Robots. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Portland, OR.

Butler, J. T., & Agah, A. (2001). Psychological effects of behavior patterns of a mobile personal robot. *Autonomous Robots*, *10*(2), 185–202.

Caine, K., abanovic, S., & Carter, M. (2012). The Effect of Monitoring by Cameras and Robots on the Privacy Enhancing Behaviors of Older Adults. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 343–350). New York, NY, USA: ACM.

Calo, R. (2010). Robots and privacy. In G. B. Patrick Lin & K. Abney (Eds.), *Robot ethics: The ethical and social implications of robotics.* Cambridge: MIT Press.

Chen, T. L., Ciocarlie, M., Cousins, S., Grice, P. M., Hawkins, K., Hsiao, K., et al. (2013, March). Robots for Humanity: Using Assistive Robotics to Empower People with Disabilities. *IEEE Robotics and Automation Magazine*, *20*(1), 30–39.

Cheung, Venkatesh, Paruchuri, Zhao, & Nguyen. (2009). Protecting and managing privacy information in video surveillance systems. In *Protecting privacy in video surveillance* (pp. 11–33). Springer.

Cheung, S. S., Zhao, J., & Venkatesh, M. V. (2006). Efficient object-based video inpainting. In *Image Processing, 2006 IEEE International Conference on* (pp. 705–708). IEEE.

Cole, F., DeCarlo, D., Finkelstein, A., Kin, K., Morley, R. K., & Santella, A. (2006). Directing Gaze in 3d Models with Stylized Focus. *Rendering Techniques*, *2006*, 17th.

Crowley, J. L., Coutaz, J., & Bérard, F. (2000). Things That See. *Communications of the ACM*, *43*(3), 54–ff.

Darling, K. (2015). "Who's Johnny?" Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy. In *Proceedings of We Robot 2015.* University of Washington. (Available at SSRN: http://ssrn.com/abstract=2588669 or http://dx.doi.org/10.2139/ssrn.2588669)

DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., & Santella, A. (2003). Suggestive contours for conveying shape. *ACM Transactions on Graphics (TOG)*, *22*(3), 848–855.

DeCarlo, D., & Santella, A. (2002). Stylization and abstraction of photographs. In *ACM Transactions on Graphics (TOG)* (Vol. 21, pp. 769–776). ACM.

DeCew, J. (2013). Privacy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 ed.).

Denning, T., Matuszek, C., Koscher, K., Smith, J. R., & Kohno, T. (2009). A spotlight on security and privacy risks with future household robots: attacks and lessons. In *Proceedings of the 11th international conference on Ubiquitous computing* (pp. 105–114). ACM.

Dragan, A., & Srinivasa, S. (2013). Generating legible motion. In *Proceedings of robotics: Science and systems (r:ss).*

Edney, J. J., & Buda, M. A. (1976). Distinguishing territoriality and privacy: Two studies. *Human Ecology*, *4*(4), 283–296.

Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., & Burgard, W. (2012). An evaluation of the RGB-D SLAM system. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on* (pp. 1691–1696). IEEE.

Erdelyi, A., Barat, T., Valet, P., Winkler, T., & Rinner, B. (2014). Adaptive cartooning for privacy protection in camera networks. In *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on* (pp. 44–49). IEEE.

Erdelyi, A., Winkler, T., & Rinner, B. (2014). Multi-Level Cartooning for Context-Aware Privacy Protection in Visual Sensor Networks.

Fischer, J., Cunningham, D., Bartz, D., Wallraven, C., Beulthoff, H., & Strasser, W. (2006). Measuring the discernability of virtual objects in conventional and stylized augmented reality. In *12th Eurographics Symposium on Virtual Environments, Lisbon, Portugal, May 8th-10th, 2006* (p. 53). Transaction Publishers.

Fischer, K. (2011). Interpersonal Variation in Understanding Robots As Social Actors. In *Proceedings of the 6th International Conference on Human-robot Interaction* (pp. 53–60). New York, NY, USA: ACM.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3), 143–166.

Forsyth, D. A., Fleck, M., & Bregler, C. (1996). Finding naked people. *International Journal of Computer Vision*.

Fried, C. (1970). An anatomy of values. *Cambridge, Mass.*

Friedman, B., Kahn Jr, P. H., Hagman, J., Severson, R. L., & Gill, B. (2009). The Watcher and the Watched: Social Judgments about Privacy in a Public Place. In *Media Space 20+ Years of Mediated Life* (pp. 145–176). Springer.

Galindo, C., Fernndez-Madrigal, J.-A., Gonzlez, J., & Saffiotti, A. (2008). Robot task planning using semantic maps. *Robotics and Autonomous Systems*, *56*(11), 955–966.

Galindo, C., Saffiotti, A., Coradeschi, S., Buschka, P., Fernandez-Madrigal, J., & Gonzalez, J. (2005, August). Multi-hierarchical semantic maps for mobile robotics. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005. (IROS 2005)* (pp. 2278–2283).

Gerstner, T., DeCarlo, D., Alexa, M., Finkelstein, A., Gingold, Y., & Nealen, A. (2012). Pixelated image abstraction. In *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering* (pp. 29–36). Eurographics Association.

Gockley, R., Forlizzi, J., & Simmons, R. (2007). Natural Person-following Behavior for Social Robots. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction* (pp. 17–24). New York, NY, USA: ACM.

Gooch, A. A., & Willemsen, P. (2002). Evaluating space perception in NPR immersive environments. In *Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering* (pp. 105–110). ACM.

Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: a survey. *Foundations and trends in human-computer interaction*, *1*(3), 203–275.

Goron, L. C., Marton, Z.-C., Pangercic, D., Ruhr, T., Tenorth, M., & Beetz, M. (2011). Autonomous Semantic Mapping for Robots Performing Everyday Manipulation Tasks in Kitchen Environments. In *Proceedings of the International Conference on Robots and Systems (IROS)* (pp. 4263–4270).

Hall, E. T. (1966). *The Hidden Dimension*. Doubleday, Garden City.

Halper, N., Mellin, M., Herrmann, C. S., Linneweber, V., & Strothotte, T. (2003). Towards an understanding of the psychology of non-photorealistic rendering. In *Computational Visualistics, Media Informatics, and Virtual Communities* (pp. 67–78). Springer.

Harris and Associates, I., & Westin, A. F. (1998). E-commerce and privacy: What net users want. *Privacy and American Business, Hackensack, NJ*.

Hartzog, W. (2015). Focus on cyberlaw: Unfair and deceptive robots. *Maryland Law Review*, *74*, 785–1031.

Henkel, Z., Bethel, C., Murphy, R., & Srinivasan, V. (2014, June). Evaluation of Proxemic Scaling Functions for Social Robotics. *IEEE Transactions on Human-Machine Systems*, *44*(3), 374–385.

Herling, J., & Broll, W. (2012). Pixmix: A real-time approach to high-quality diminished reality. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on* (pp. 141–150). IEEE.

Hoffman, G., Forlizzi, J., Ayal, S., Steinfeld, A., Antanitis, J., Hochman, G., et al. (2015). Robot Presence and Human Honesty: Experimental Evidence. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 181–188). New York, NY, USA: ACM.

Hong, J. I., & Landay, J. A. (2004). An architecture for privacy-sensitive ubiquitous computing. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services* (pp. 177–189). ACM.

Hornung, A., Wurm, K. M., Bennewitz, M., Stachniss, C., & Burgard, W. (2013). OctoMap: An efficient probabilistic 3d mapping framework based on octrees. *Autonomous Robots*, *34*(3), 189–206.

Hubers, A., Andrulis, E., Stirrat, T., Tran, D., Zhang, R., Sowell, R., et al. (2015, March). Video Manipulation Techniques for the Protection of Privacy in Remote Presence Systems. In *HRI 2015 Extended Abstracts.* Portland, OR.

Inness, J. C. (1992). *Privacy, intimacy, and isolation*. Oxford University Press.

InTouch Health. (2012). *InTouch Telemedicine System*. (http://www.intouchhealth.com/)

Jacobs, E. (1977, November). The Need for Privacy and the Application of Privacy to the Day Care Setting.

Jana, S., Narayanan, A., & Shmatikov, V. (2013). A Scanner Darkly: Protecting user privacy from perceptual applications. In *Security and Privacy (SP), 2013 IEEE Symposium on* (pp. 349–363). IEEE.

Jibo. (2016). Available from `http://www.jibo.com`

Joosse, M., Sardar, A., Lohse, M., & Evers, V. (2013). BEHAVE-II: The revised set of measures to assess users attitudinal and behavioral responses to a social robot. *International journal of social robotics*, *5*(3), 379–388.

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., et al. (2015). Will People Keep the Secret of a Humanoid Robot?: Psychological Intimacy in HRI. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 173–180). New York, NY, USA: ACM.

Karro, J., Dent, A. W., & Farish, S. (2005, April). Patient perceptions of privacy infringements in an emergency department. *Emergency Medicine Australasia*, *17*(2), 117–123.

Kato, S., Nishiyama, S., & Takeno, J. (1992). Coordinating Mobile Robots By Applying Traffic Rules. In *IROS* (Vol. 92, pp. 1535–1541).

Kelley, P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009). A nutrition label for privacy. In *Proceedings of the 5th symposium on usable privacy and security* (p. 4).

Korshunov, P., Cai, S., & Ebrahimi, T. (2012). Crowdsourcing approach for evaluation of privacy filters in video surveillance. In *Proceedings of the ACM multimedia 2012 workshop on Crowdsourcing for multimedia* (pp. 35–40). ACM.

Korshunov, P., & Ebrahimi, T. (2013). Using face morphing to protect privacy. In *10th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 208–213).

Kyprianidis, J. E. (2011). Image and Video Abstraction by Multi-scale Anisotropic Kuwahara Filtering. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering* (pp. 55–64). New York, NY, USA: ACM.

La France, M., & Mayo, C. (1979). A review of nonverbal behaviors of women and men. *Western Journal of Speech Communication*, *43*(2), 96–107.

Langheinrich, M. (2001). Privacy by designprinciples of privacy-aware ubiquitous systems. In *Ubicomp 2001: Ubiquitous Computing* (pp. 273–291). Springer.

Langheinrich, M. (2002). A privacy awareness system for ubiquitous computing environments. In *UbiComp 2002: Ubiquitous Computing* (pp. 237–245). Springer.

LaValle, S. M. (2006). *Planning algorithms*. Cambridge University Press.

Lederer, S., Dey, A. K., & Mankoff, J. (2002). *A conceptual model and a metaphor of everyday privacy in ubiquitous computing environments* (Tech. Rep. No. UCB/CSD-2-1188). Computer Science Division, University of California, Berkeley.

Lederer, S., Mankoff, J., & Dey, A. K. (2003). Who wants to know what when? privacy preference determinants in ubiquitous computing. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems* (pp. 724–725). ACM.

Lee, M. K., & Takayama, L. (2011). "Now, I Have a Body": Uses and Social Norms for Mobile Remote Presence in the Workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 33–42). New York, NY, USA: ACM.

Lee, M. K., Tang, K. P., Forlizzi, J., & Kiesler, S. (2011). Understanding users' perception of privacy in human-robot interaction. In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on* (pp. 181–182). IEEE.

Leino-Kilpi, H., Valimaki, M., Dassen, T., Gasull, M., Lemonidou, C., Scott, A., et al. (2001). Privacy: A Review of the Literature. *International Journal of Nursing Studies*, *38*, 663–671.

Levin, D. T., Killingsworth, S. S., & Saylor, M. M. (2008). Concepts About the Capabilities of Computers and Robots: A Test of the Scope of Adults' Theory of Mind. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (pp. 57–64). New York, NY, USA: ACM.

Levy, D. (2009). *Love and sex with robots: The evolution of human-robot relationships*. New York, NY: Harper.

Lichtenthäler, C., & Kirsch, A. (2013). Towards legible robot navigationhow to increase the intend expressiveness of robot navigation behavior. In *Proceedings of international conference on social robotics – workshop on embodied communication of goals and intentions.*

Lu, D., Allan, D. B., & Smart, W. D. (2013). Tuning Cost Functions for Social Navigation. In *Proceedings of the International Conference on Social Robotics (ICSR).*

Lu, D., & Smart, W. D. (2013). Towards More Efficient Navigation for Robots and Humans. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).*

Lu, J., Sander, P. V., & Finkelstein, A. (2010). Interactive painterly stylization of images, videos and 3d animations. In *Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3d Graphics and Games* (pp. 127–134). ACM.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, *15*(4), 336–355.

Moore, A. D. (1998, October). Intangible Property: Privacy, Power, and Information Control. *American Philosophical Quarterly*, *35*(4), 365–378.

Moore, A. D. (2003). Privacy: its meaning and value. *American Philosophical Quarterly*, 215–227.

Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th international conference on Human-robot interaction* (pp. 331–338). ACM.

Nakashima, Y., Koyama, T., Yokoya, N., & Babaguchi, N. (2015). Facial expression preserving privacy protection using image melding. In *Ieee international conference on multimedia and expo (icme)* (pp. 1–6).

Nakauchi, Y., & Simmons, R. (2002, May). A Social Robot that Stands in Line. *Autonomous Robots*, *12*(3), 313–324.

Newell, P. B. (1995). Perspectives on Privacy. *Journal of Environmental Psychology*, *15*, 87–104.

Newell, P. B. (1998). A cross-cultural comparison of privacy definitions and functions: A systems approach.

Nisbett, R. E., & Wilson, T. D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, *84*(3), 231–259.

Okita, S. Y., Ng-Thow-Hing, V., & Sarvadevabhatla, R. K. (2012). Captain may I?: proxemics study examining factors that influence distance between humanoid robots, children, and adults, during human-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 203–204). ACM.

Online Etymology Dictionary. (2015). Available from `http://www.etymonline.com/`

Paine, C., Reips, U.-D., Stieger, S., Joinson, A., & Buchanan, T. (2007). Internet users' perceptions of "privacy concerns" and "privacy actions". *International Journal of Human-Computer Studies*, *65*, 526–536.

Paine, R. (1970). Lappish decisions, partnerships, information management, and sanctions: A nomadic pastoral adaptation. *Ethnology*, *9*(1), 52–67.

Parent, W. A. (1983, October). Privacy, Morality, and the Law. *Philosophy & Public Affairs*, *12*(4), 269–288.

Phillips, L., Ries, B., Interrante, V., Kaeding, M., & Anderson, L. (2009). Distance perception in NPR immersive virtual environments, revisited. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization* (pp. 11–14). ACM.

Prosser, W. L. (1960). Privacy. In F. D. Schoeman (Ed.), *Philosophical dimensions of privacy: An anthology.* Cambridge University Press.

Rachels, J. (1975). Why privacy is important. *Philosophy & Public Affairs*, 323–333.

Raval, N., Srivastava, A., Lebeck, K., Cox, L. P., & Machanavajjhala, A. (2014). MarkIt: Privacy Markers for Protecting Visual Secrets. In *In Proceedings of the Workshop on Usable Privacy and Security for Wearable and Domestic ubIquitous DEvices (UPSIDE).*

Reagle, J., & Cranor, L. F. (1999). The platform for privacy preferences. *Communications of the ACM*, *42*(2), 48–55.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press.

Richards, N. M., & Smart, W. D. (2012). How Should the Law Think About Robots? In *Proceedings of We Robot: The Inaugural Conference on Legal and Policy Issues Relating to Robotics.* Coral Gables, FL.

Roberts, J. M., & Gregor, T. (1971). Privacy: A cultural view. In J. R. Pennock & J. W. Chapman (Eds.), *Privacy and personality.* Transaction Publishers.

Rother, C., Kolmogorov, V., & Blake, A. (2004). Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (TOG)*, *23*(3), 309–314.

Rueben, M., Bernieri, F. J., Grimm, C., & Smart, W. D. (2016). User Feedback on Physical Marker Interfaces for Protecting Visual Privacy from Mobile Robots. In *Proceedings of the Eleventh Annual ACM/IEEE International Conference on Human-Robot Interaction, Late-Breaking Reports.* Christchurch, NZ. (in press)

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Prentice Hall.

Rusu, R. B., Marton, Z. C., Blodow, N., Holzbach, A., & Beetz, M. (2009). Model-based and learned semantic object labeling in 3d point cloud maps of kitchen environments. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 3601–3608). IEEE.

Rydén, F., & Chizeck, H. J. (2012). Forbidden-region virtual fixtures from streaming point clouds: Remotely touching and protecting a beating heart. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 3308–3313). IEEE.

Rydén, F., & Chizeck, H. J. (2013). A method for constraint-based six degree-of-freedom haptic interaction with streaming point clouds. In *2013 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2353–2359). IEEE.

Rydén, O. F. (2013). *Real-Time Haptic Interaction with Remote Environments using Non-contact Sensors*. Unpublished doctoral dissertation.

Saini, M., Atrey, P. K., Mehrotra, S., & Kankanhalli, M. (2014). W3-privacy: understanding what, when, and where inference channels in multi-camera surveillance video. *Multimedia Tools and Applications*, *68*(1), 135–158.

Satake, S., Iba, H., Kanda, T., Imai, M., & Saiki, Y. M. (2014). May I talk about other shops here?: modeling territory and invasion in front of shops. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (pp. 487–494). ACM.

Schermerhorn, P., Scheutz, M., & Crowell, C. R. (2008). Robot Social Presence and Gender: Do Females View Robots Differently Than Males? In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (pp. 263–270). New York, NY, USA: ACM.

Schiff, J., Meingast, M., Mulligan, D. K., Sastry, S., & Goldberg, K. Y. (2007). Respectful Cameras: Detecting Visual Markers in Real-Time to Address Privacy Concerns. In *Proceedings of the IEEE/RSJ International Conference on Robots and Systems (IROS)* (pp. 971–978). San Diego, CA.

Schoeman, F. D. (1984). *Philosophical dimensions of privacy: An anthology*. Cambridge University Press.

Sebba, R., & Churchman, A. (1983). Territories and territoriality in the home. *Environment and Behavior*, *15*(2), 191–210.

Sentis, L., & Khatib, O. (2006). A Whole-Body Control Framework for Humanoids Operating in Human Environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2641–2648).

Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Sommer, R. (1969). *Personal space: The behavioral basis of design*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Sommer, R., & Becker, F. D. (1969). Territorial defense and the good neighbor. *Journal of personality and social psychology*, *11*(2), 85.

Suitable Technologies. (2012). *Beam Remote Presence System*. (https://www.suitabletech.com/)

Sun, J., Yuan, L., Jia, J., & Shum, H.-Y. (2005). Image completion with structure propagation. In *ACM Transactions on Graphics (ToG)* (Vol. 24, pp. 861–868). ACM.

Takayama, L., Dooley, D., & Ju, W. (2011). Expressing thought: improving robot readability with animation principles. In *Proceedings of the 6th international conference on Human-robot interaction* (pp. 69–76). ACM.

Takayama, L., Groom, V., & Nass, C. (2009). I'm Sorry, Dave: I'm Afraid I Won't Do That: Social Aspects of Human-agent Conflict. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2099–2108). New York, NY, USA: ACM.

Takayama, L., & Pantofaru, C. (2009). Influences on proxemic behaviors in human-robot interaction. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 5495–5502). IEEE.

Templeman, R., Korayem, M., Crandall, D., & Kapadia, A. (2014). Placeavoider: Steering first-person cameras away from sensitive spaces. In *Network and distributed system security symposium (ndss).*

The United Nations. (1948). *The Universal Declaration of Human Rights.*

Thomasen, K. (2012). Liar, Liar, Pants on Fire! Examining the Constitutionality of Enhanced Robo-Interrogation. In *Proceedings of We Robot 2012.* University of Miami. (draft)

Thomson, J. J. (1975, July). The Right to Privacy. *Philosophy & Public Affairs*, *4*(4), 295–314.

U.S. Constitution. Amend. I. (1791).

U.S. Constitution. Amend. IV. (1791).

Vaiapury, K., Aksay, A., & Izuierdo, E. (2010). GrabcutD: Improved Grabcut using Depth Information. In *Proceedings of the 2010 ACM Workshop on Surreal Media and Virtual Cloning (SMVC)* (pp. 57–62).

VGo Communications. (2012). *VGo Robotic Telepresence for Healthcare, Education, and Business.* (http://www.vgocom.com/)

Walden, T. A., Nelson, P. A., & Smith, D. E. (1981). Crowding, privacy, and coping. *Environment and Behavior*, *13*(2), 205–224.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard law review*, 193–220.

Weber, R. H. (2010). Internet of ThingsNew security and privacy challenges. *Computer Law & Security Review*, *26*(1), 23–30.

Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM*, *36*(7), 75–84.

Westin, A. F. (1967). *Privacy and Freedom.* New York, NY: Athenaeum.

Zeegers, S. K., Readdick, C. A., & Hansen-Gandy, S. (1994, December). Daycare Children's Establishment of Territory to Experience Privacy. *Children's Environments*, *11*(4), 265–271.

Zhang, C., Tian, Y., & Capezuti, E. (2012). Privacy preserving automatic fall detection for elderly using rgbd cameras. In *Proceedings of the 13th international conference on computers helping people with special needs* (pp. 625–633).

Zhao, Q. A., & Stasko, J. T. (1998). Evaluating image filtering based techniques in media space applications. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work* (pp. 11–18). ACM.

Authors' names and contact information: Matthew Rueben, ruebenm@oregonstate.edu; William D. Smart, bill.smart@oregonstate.edu. Both authors are with the Robotics Group, School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University, Corvallis, OR, USA.