Human–Computer Interaction and the Ethical Computing Protocol

Teresa Gao

Massachusetts Institute of Technology

STATEMENT

This project extends my Spring 2022 SERC Scholars project, an annotated bibliography of references relating to the Ethical Computing Protocol (ECP). In my SERC Scholars project, I examined the ECP as applied to the field of Human–Robot Interaction; the ECP provides a contextual basis with which to evaluate the future implications of HRI. Previously, I investigated smart home technologies, AI foreign policy, and driving automation systems, respectively: I considered smart home technologies with respect to the potential of their future developments, AI foreign policy with respect to current actual government action, and driving automation systems with respect to forward-looking framework.

The purpose of this project is to broaden the audience of my previous work so that the ECP and annotated bibliography examples presented may be understood by not just engineers and academics but also members of the general public. To produce a cohesive end product, by the recommendation of my teaching fellow and classmates, I focus on only HRI in the future rather than including past and current scenarios and also concentrate on going deep into one case study rather than briefly touching on several.

Background (Transcript)

Hi there! This is Teresa Gao, and you're listening to a mini podcast episode on Human–Robot Interaction, the Ethical Computing Protocol, and the future of smart home devices.

When many of us think of technology, we consider it a tool or set of tools. For instance, we might use our cell phones as a map or a means for messaging — all cases where we, the humans, are in control of the technology we're using. But what if that cell phone in our hands were capable of interacting with us as autonomously and intelligently as another person would be? What if technology weren't a tool but a collaborator, a companion, an equal?

Though these technologies themselves may remain science fiction for now, the development of technology that can truly think for itself, by itself, in the near future is a very real prospect. In fact, many in the scientific community, from software engineers to philosophers and ethicists, are trying to figure out our relationship with our technology as it gets smarter. These research efforts are sometimes grouped under the term "Human–Robot Interaction," or HRI, which is the study of humans' relation to artificially intelligent (or AI) agents — and when we say "agents," we mean computers, robots, machines, or any technology which is capable of learning, understanding, and acting on its own.

HRI is a pretty forward-looking field because few of the technologies it studies exist currently. However, it only takes a little imagination to realize the full importance of considering the kinds of issues we humans might encounter with our increasingly intelligent AIs before they arise. For instance: How might we regulate the technological and social transition from humans to driving automation systems, otherwise known as self-driving cars? Or how might we reconsider camera surveillance given better facial-recognition algorithms, especially in the light of known biases and the potential for governmental oppression? These are big questions with no easy answers, and they're exactly the kind that HRI researchers love to think about.

One big concern of HRI is that the engineers of AI and the ethicists of AI might not have any overlap. In short, we might have engineers who are building technologies without recognizing their full ethical implications, and we might have ethicists who are analyzing rights and wrongs without the ability to affect the actual end product! Thankfully, several methods such as the Ethical Computing Protocol can help us incorporate ethics into project-planning and decision-making. The Ethical Computing Protocol, or ECP, was created by Abby Everett Jaques and Milo Phillips-Brown of the MIT Schwarzman College of Computing and is a forward-looking process encompassing four steps: Envision Futures, Identify Values, Map Values to Decisions, and Justify and Decide. Together, these steps help us evaluate our engineering endeavors and iterate until we arrive at something that is both practically and ethically responsible.

Today, we'll be focusing on the third step of the ECP: Map Values to Decisions. Map Values to Decisions connects possible futures to current choices through the investigation of Key Junctures, Paths and Decisions, and Tradeoffs. To demonstrate the power of mapping Values to Decisions, we'll walk through this step of the ECP with the case study of smart home technologies.

PODCAST PRODUCTION

The format of my project is a mini educational podcast episode. The structure of my project follows that of much other educational entertainment like Crash Course: lighthearted, explanation-dense content. Each section is also delineated by a musical separator that can hopefully help the reader better hear topic changes. To ensure that the content as well as medium for my project is broadly accessible as possible, I have limited jargon and intend to display transcripts of the episode alongside the raw audio content.

The mini podcast episode provides some background into the field of HRI, introduces the ECP. Additionally, it deep-dives into the third step of the ECP, Map Values to Decisions, which connects possible futures to technological design choices that can be made. Finally, this step of the ECP is explored in detail to examine a case study of smart home technologies.

In creating my mini podcast episode, I produced everything from scratch. I wrote out a full script before recording each section and splicing them together. As an amateur composer, I also came up with the original interlude music that plays between spoken sections. As a host of an educational radio show on WMBR 88.1 FM, I have some experience putting together audio segments.

REFERENCES

Jaques, Abby Everett and Milo Phillips-Brown. *The Ethical Computing Protocol.* MIT Schwarzman College of Computing.

Radhika Garg and Hua Cui. 2022. *Social Contexts, Agency, and Conflicts: Exploring Critical Aspects of Design for Future Smart Home Technologies.* ACM Trans. Comput.-Hum. Interact. 29, 2, Article 11 (April 2022), 30 pages. https://doi.org/10.1145/3485058.

Human–Computer Interaction and the Ethical Computing Protocol

Teresa Gao

Massachusetts Institute of Technology

STATEMENT

This project extends my Spring 2022 SERC Scholars project, an annotated bibliography of references relating to the Ethical Computing Protocol (ECP). In my SERC Scholars project, I examined the ECP as applied to the field of Human–Robot Interaction; the ECP provides a contextual basis with which to evaluate the future implications of HRI. Previously, I investigated smart home technologies, AI foreign policy, and driving automation systems, respectively: I considered smart home technologies with respect to the potential of their future developments, AI foreign policy with respect to current actual government action, and driving automation systems with respect to forward-looking framework.

The purpose of this project is to broaden the audience of my previous work so that the ECP and annotated bibliography examples presented may be understood by not just engineers and academics but also members of the general public. To produce a cohesive end product, by the recommendation of my teaching fellow and classmates, I focus on only HRI in the future rather than including past and current scenarios and also concentrate on going deep into one case study rather than briefly touching on several.

Case Study (Transcript)

When we refer to smart home technologies, we mean AI agents such as Siri, Alexa, Cortana, or Google Assistant that are used by the everyday person in a domestic environment. These can be embodied in any form from tall speakers to small cell phones, but they have one thing in common: you can talk to them, and they can talk back. At the moment, none of these smart home devices are capable of much: sure, you can ask them to tell a joke or play a song and, in a soothingly calm voice, they will usually oblige, but they can't learn to do things that they aren't pre-programmed to do. That might all soon change: as smart home devices are increasingly being adopted in the home environment, their capabilities — and the consequences of their capabilities — will likely also grow. Therefore, in designing future smart home technologies, we have plenty of things to think about in terms of ethical impact.

In order to map values to decisions, we start by identifying Key Junctures. If our possible futures are like paths, then junctures are the forks in the road — and key junctures are those forks between especially good paths and especially bad ones. One suggested exercise for determining key junctures is drawing a timeline, a map of the paths of our possible futures; then, color coding or otherwise indicating good branches from bad ones; and finally, circling those nodes where the path splits between a much better branch and a much worse one.

Let's walk through a few of these key junctures for future smart home technologies. When we build smart home technologies, we don't just "build" them: we design them, construct them, deploy them, and maintain them. Each of these subprocesses contributes greatly to the overall implementation workflow and could be a hugely important branching point in terms of possible futures. For example, let's consider design: in one possible future, our smart home AIs are embodied in seven-foot-tall robots which are helpful in moving furniture but less accessible and even intimidating to young children in the home; in another future, our smart home AIs are embodied in inch-tall figurines which are more approachable to children but are too easily lost or broken. Or consider maintenance: in one possible future, maintenance is smooth and continuous, with automatic software updates requiring no effort or input from the user; in another future, maintenance is choppy or nonexistent, providing a buggy or very non-frictionless experience for the user. In our consideration of possible futures, these types of engineering decisions are examples of key junctures, important branching points between good and bad paths.

Next, we identify our Paths and Decisions. For each key juncture, we ask ourselves what choices we can make that will put us on the best path at that juncture. Continuing our case study of future smart home technologies, how might we ensure that we end up designing helpful rather than harmful home AI? In their paper on critical aspects of design for future smart home technologies, Radhika Garg and Hua Cui of Syracuse University collected feedback from participants on their ideal at-home AI helper. The recommendations that they received called for greater adaptability, awareness of social contexts within a house, flexible agency, and conflict resolution abilities in smart home devices. The inclusion and relative emphases on each of these areas can help us end up making better choices at those key junctures we previously identified.

PODCAST PRODUCTION

The format of my project is a mini educational podcast episode. The structure of my project follows that of much other educational entertainment like Crash Course: lighthearted, explanation-dense content. Each section is also delineated by a musical separator that can hopefully help the reader better hear topic changes. To ensure that the content as well as medium for my project is broadly accessible as possible, I have limited jargon and intend to display transcripts of the episode alongside the raw audio content.

The mini podcast episode provides some background into the field of HRI, introduces the ECP. Additionally, it deep-dives into the third step of the ECP, Map Values to Decisions, which connects possible futures to technological design choices that can be made. Finally, this step of the ECP is explored in detail to examine a case study of smart home technologies.

In creating my mini podcast episode, I produced everything from scratch. I wrote out a full script before recording each section and splicing them together. As an amateur composer, I also came up with the original interlude music that plays between spoken sections. As a host of an educational radio show on WMBR 88.1 FM, I have some experience putting together audio segments.

REFERENCES

Jaques, Abby Everett and Milo Phillips-Brown. *The Ethical Computing Protocol.* MIT Schwarzman College of Computing.

Radhika Garg and Hua Cui. 2022. *Social Contexts, Agency, and Conflicts: Exploring Critical Aspects of Design for Future Smart Home Technologies.* ACM Trans. Comput.-Hum. Interact. 29, 2, Article 11 (April 2022), 30 pages. https://doi.org/10.1145/3485058.

Human–Computer Interaction and the Ethical Computing Protocol

Teresa Gao

Massachusetts Institute of Technology

STATEMENT

This project extends my Spring 2022 SERC Scholars project, an annotated bibliography of references relating to the Ethical Computing Protocol (ECP). In my SERC Scholars project, I examined the ECP as applied to the field of Human–Robot Interaction; the ECP provides a contextual basis with which to evaluate the future implications of HRI. Previously, I investigated smart home technologies, AI foreign policy, and driving automation systems, respectively: I considered smart home technologies with respect to the potential of their future developments, AI foreign policy with respect to current actual government action, and driving automation systems with respect to forward-looking framework.

The purpose of this project is to broaden the audience of my previous work so that the ECP and annotated bibliography examples presented may be understood by not just engineers and academics but also members of the general public. To produce a cohesive end product, by the recommendation of my teaching fellow and classmates, I focus on only HRI in the future rather than including past and current scenarios and also concentrate on going deep into one case study rather than briefly touching on several.

Conclusion (Transcript)

Finally, to finish mapping values to decisions, we identify tradeoffs. Tradeoffs are what we call the segments of paths on our timeline of possible futures where things go well in some ways for some stakeholders and less well in other ways for other stakeholders. According to Garg and Cui, there are various roles that a person might hope that their smart home device might be able to play: a Mediator, or neutral conflict resolver; a Scaffolder, or source of suggestions; a Consoler, or emotional supporter; an Executor, or independent task completer; a Collaborator, or team member; or an Encourager, or motivator. We can see that there are many tradeoffs to be had in deciding which role a smart home device should occupy: what are the possible futures that could arise if we programmed, say, an AI able to execute tasks independently but unable to empathize with human emotions? This might benefit a user who wants their smart home device to be nothing more than a tool for telling jokes or playing music but fail to support a user in moments of extreme distress and need. These types of tradeoffs must be balanced between various stakeholder groups to arrive at our final design decisions.

In this mini podcast episode, we've introduced Human–Robot Interaction and the Ethical Computing Protocol and also deep-dived into the third step of the ECP, Map Values to Decisions, for the case study of smart home technologies. Map Values to Decisions is only one step in the ECP; if you're interested in learning more, be sure to check out the full ECP by Abby Everett Jaques and Milo Phillips-Brown of the MIT Schwarzman College of Computing.

This mini podcast episode was produced by yours truly, Teresa Gao, as part of Experiential Ethics, a summer experiential learning program, and SERC Scholars, a community of students, researchers, and faculty advancing Social and Ethical Responsibilities of Computing in the MIT Schwarzman College of Computing. Thank you so much for listening!

PODCAST PRODUCTION

The format of my project is a mini educational podcast episode. The structure of my project follows that of much other educational entertainment like Crash Course: lighthearted, explanation-dense content. Each section is also delineated by a musical separator that can hopefully help the reader better hear topic changes. To ensure that the content as well as medium for my project is broadly accessible as possible, I have limited jargon and intend to display transcripts of the episode alongside the raw audio content.

The mini podcast episode provides some background into the field of HRI, introduces the ECP. Additionally, it deep-dives into the third step of the ECP, Map Values to Decisions, which connects possible futures to technological design choices that can be made. Finally, this step of the ECP is explored in detail to examine a case study of smart home technologies.

In creating my mini podcast episode, I produced everything from scratch. I wrote out a full script before recording each section and splicing them together. As an amateur composer, I also came up with the original interlude music that plays between spoken sections. As a host of an educational radio show on WMBR 88.1 FM, I have some experience putting together audio segments.

REFERENCES

Jaques, Abby Everett and Milo Phillips-Brown. *The Ethical Computing Protocol.* MIT Schwarzman College of Computing.

Radhika Garg and Hua Cui. 2022. *Social Contexts, Agency, and Conflicts: Exploring Critical Aspects of Design for Future Smart Home Technologies.* ACM Trans. Comput.-Hum. Interact. 29, 2, Article 11 (April 2022), 30 pages. https://doi.org/10.1145/3485058.