"That's Not An Architecture Problem!": Techniques and Challenges for Practicing Anticipatory Technology Ethics

Katie Shilton, University of Maryland College Park

Abstract
Anticipatory technology ethics is the practice of analyzing how emerging technologies will be built, imagining how they might be used, and interpreting what their consequences might be. Information researchers and professionals are increasingly interested in understanding the possible social impacts of technologies before wide-spread deployment, and anticipatory technology ethics provides one lens through which to do so. This paper reports on two anticipatory ethics projects. These contrasting projects tested techniques for practicing anticipatory ethics, and also illuminated challenges that stem from diverse engineering work practices and the design of user-facing versus infrastructural technologies. In particular, the paper focuses on the problem of *at what layer* ethical issues can be found in technology design, and the ways that the material realities of design impact developers' willingness and readiness to participate in anticipatory ethics.

1   Introduction

Recent scholarship has suggested that anticipatory ethics be incorporated into the design of emerging technologies (Brey, 2012; Johnson, 2007). Anticipatory ethics is the practice of using the design phase to reflect upon how a system or technology's affordances will impact their use and potential consequences. This paper reports on two projects to illustrate techniques and challenges for practicing anticipatory ethics during the design of emerging information technologies. The first project involved three years spent as a participant observer in the Center for Embedded Networked Sensing (CENS), a technology research center developing data collection systems using mobile phones. The second project consisted of three years as a participant observer with the Named Data Networking (NDN) project, a research project centered on redesigning the underlying architecture of the Internet. While both projects point to successful techniques for practicing anticipatory ethics, they also illustrate distinct challenges faced by information professionals and researchers who intervene to promote values and ethics within design. These challenges stem from the material and practical differences inherent in the design of user-facing versus infrastructural technologies. Developers in the NDN project posed meaningful challenges to the practice of anticipatory ethics based upon questions of, to borrow a metaphor from network architecture, *at what layer* ethical consequences appear in a technology.

The paper begins with a review of anticipatory ethics drawn from the information studies, technology ethics, and values in design literatures. It then provides a high-level summary exploring successful techniques for practicing anticipatory technology ethics gathered from two contrasting participant observations of academic technology design labs. (Greater detail on both projects can be found in (Shilton, 2013) and (Shilton, 2014)). The paper then discusses the ways that both work practices and technological realities complicated the practice of anticipatory ethics in these projects. This discussion of challenges is intended to generate discussion about how information researchers might broach these challenges as a community.

2   Anticipatory Technology Ethics

Anticipatory ethics is an emerging branch of technology ethics which seeks to highlight ethical challenges in emerging technology and encourage conversations about ethics during design. Brey (2012) has suggested a systematic method for conducting anticipatory ethics by analyzing properties of, and techniques used within, technologies while they are still in the research and development stage. Brey

suggests that technology can be decomposed into techniques and applications to analyze during research and development. Analyzing the processes technologies bring about can determine whether those processes are good or bad, right or wrong. He argues that the problem of uncertainty in an emerging technology's adoption and future use can be overcome through precise (and varying) methods of analysis at each level of decomposition.

Johnson (2007) takes anticipatory ethics a step further by advocating direct intervention in design. She believes it is important for ethics researchers to intervene early in technology development, writing, "the ideas that circulate during the early stages of development of a new technology influence the construction of meaning as well as the material design of the technology" (Johnson, 2011, p. 63). Johnson leaves open, however, the challenge of how best to conduct anticipatory ethics. As she writes, "The hard part is to figure out how to bring ethical notions and practices and ethicists explicitly, intentionally, and effectively into the fray" (Johnson, 2011, p. 67).

Although it uses a vocabulary of values rather than ethics, a similar approach to proactive analysis of design and systems for ethical challenges can be found in values in design, an ongoing project in the fields of information studies (Knobel & Bowker, 2011), human-computer interaction (Friedman, Kahn, & Borning, 2006), media studies (Flanagan, Nissenbaum, & Howe, 2008; Jarvenpaa & Leidner, 1998), and science and technology studies (Johnson, 1997, 2011; Sclove, 1995; Winner, 1989). In this literature, values are understood as identifiable entities with ethical import that appear in technologies, embedded consciously or unconsciously by developers and concretized through affordances (Friedman & Nissenbaum, 1997; Johnson, 2000). This research posits that the process of designing something is about interpreting meaning (Latour, 2008). As an individual or group designs an artifact, they partially construct its uses and thereby its meanings. Through this process, designers broach questions of values: as Latour puts it, "good versus bad design" (2008, p. 5). Much of the VSD and VID work has been proactive, introducing particular values such autonomy, privacy, or sustainability into design settings (Friedman & Hendry, 2012; Manders-Huits & Zimmer, 2009). Approaches such as reflexive design (Sengers, Boehner, David, & Kaye, 2005) and critical technical practice (Agre, 1997) have been proposed to explicitly introduce discussions of values and ethics—rather than proscribed values—into design practices.

This paper explores how to, as Johnson (2011) put it, bring ethical notions into the fray of emerging technology design. It uses data gathered as a participant-observer on two very different emerging information technology projects: an academic laboratory focused on the development of tracking software for mobile phones to collect data about people and their habits; and a multi-campus academic research effort to redesign the backbone technologies of the Internet. This paper interrogates the ways that design practices and processes intersect with, and conflict with, anticipatory ethics. It advocates for engaging technology developers in anticipatory ethics, while pointing to some difficult challenges to this engagement.

3   Method

From 2008-2011, the author worked as a participant-observer at the Center for Embedded Networked Sensing (CENS), an NSF-funded science and technology center at the University of California, Los Angeles. CENS engineers developed software to run on mobile phones to collect data about people and their activities. The Center's work included deciding what kinds of data (e.g. location, motion, images) to collect, at what granularity, how long to keep it, and how to inform users about data collection. Ethical challenges discussed included privacy, consent, equity, and forgetting (Shilton, 2010). The author recorded over 200 hours of field notes over two years of participant-observation, and interviewed thirty members of the team.

The author studied the Named Data Networking (NDN) team from 2011-2014, and remains engaged in the project. The NDN team was comprised of researchers from ten US universities. As a result, much of the team's work involved distributed collaboration. The author attended dozens of team videoconferences and teleconferences, and traveled to all-hands meetings to engage with the development teams in person, recording over 50 hours of field notes. The author also conducted and recorded interviews with nine project PIs, two students, and two staff members.

By analyzing field notes and interviews from both projects, the author tracked what she came to call *values levers*: work practices which opened ethical discussions within design, and encouraged ethics to be named and framed as important to the work of development (Shilton, 2013). Values levers enable successful openings for anticipatory ethics. Successful openings were defined as conversations or deliberations about values, ethics, and the ethical implications of design choices. The next section explores these values levers in more detail.

4   Techniques for Anticipatory Technology Ethics

CENS participants reported discovering privacy, consent and equity concerns while testing prototypes of their applications and those of their colleagues. As in many development labs, it was common practice to test prototype systems internally before conducting testing with outside users. Experiencing using the system, and in particular, contributing the kinds of data under request (for example, location as well as questions about eating, sleeping and exercise habits) allowed the participants to imagine what inferences might be made about their behavior, and made ethical concerns concrete. Participants' prototype testing experiences contributed to a group consensus around privacy as a design principle for CENS systems.

Working on interdisciplinary teams was another design practice that enabled discussions of privacy, consent, equity and forgetting. The majority of CENS participatory sensing designers had undergraduate degrees in computer science (CS) or electrical engineering (EE). However, a few had backgrounds in statistics, design/media arts, and information studies. Statisticians, for example, led discussions about what data were useful and meaningful to collect. These conversations encouraged ethical debate about data representation, sharing, and security. These debates helped participants reach consensus about the importance of data legibility, a design principle that stemmed from concerns about equity. Being challenged to talk across disciplinary boundaries helped the design team articulate and agree upon values like legibility and, more broadly, equity.

Testing prototype applications and working on interdisciplinary teams were both work practices intrinsic to CENS: they existed without outside intervention. This suggests that there may be existing practices in development workplaces which can cultivate anticipatory ethics. To further test the theory of values levers and look for intrinsic work practices which encourage anticipatory ethics, the author turned to a second case study of a very different form of technology development. The author approached this case study to challenge or confirm previously-noticed levers such as interdisciplinary work and testing prototypes, as well as to observe additional levers. In the Named Data Networking (NDN) project, researchers strove to reinvent the backbone of the Internet, replacing the TCP/IP protocols with a new, content-centric method of data exchange. NDN researchers worked on diverse technical objects ranging from routing algorithms to Internet applications that could maximize the potential of the new data transport protocols. The NDN architecture diverges from TCP/IP by requiring semantically-meaningful data names, tracking provenance of data, enabling pervasive caching of data, and facilitating decentralized communication. These changes from TCP/IP could impact social debates about privacy and autonomy, intellectual property, cybersecurity, and fairness and neutrality in the provision of Internet services. Ethical challenges discussed in this project ranged from concerns about security to equity and democratization of information (Shilton & Koepfler, 2013).

One work practice that served as a values lever within NDN was imagining users and use cases. NDN developers were redesigning a familiar technology: the Internet. As a result, much discussion of use cases was grounded in familiar (to American network architects) uses of the Internet: checking the *New York Times* webpage, sending email, conducting conference calls, using Facebook, and streaming video. Discussion of these familiar social contexts tended to elicit particular values salient to those contexts. For example, discussing NDN applications for email or conference calls, or retrieving the *New York Times*, tended to generate concerns about provenance and trust in content. Research on vehicular networks, or communications between cars, led to discussions of constrained resources for sharing data. Efficiency became a value salient to this context, as did equity, because researchers had to decide which data was the most important to send in a constrained situation. Equity became part of conversations about data exchange in constrained environments, including data congestion management. Imagining the constraints imposed by the context of vehicular networks helped bring this value to light.

As at CENS, interdisciplinary work was another practice that functioned as a values lever in NDN development. For example, a legal scholar attached to the research team challenged the NDN developers to consider the ethical context of their work by making a presentation on current US government methods for wiretapping the Internet. This presentation led to debates among the team about whether they should support current US wiretapping legislation through design decisions. This interdisciplinary work helped highlight social concerns such as privacy, power, and surveillance, and several participants cited these interdisciplinary discussions as pivotal to their awareness of these social concerns.

Another similarity to CENS was that constructing and testing prototypes served as a values lever. For example, a prototype file sharing service raised challenges for verifying identity, and in particular, who has the rights to use a particular identity. As a result of experimenting with file sharing, the team discussed how best to handle identity management, and participated in nuanced conversations about identity in the digital age. A second example was using NDN to implement theater lighting control.

Lighting control provided a new technical challenge, because commands to the lights had to be authenticated to prevent unwanted interference. NDN was initially designed with built-in security for disseminated data, but not necessarily for two-way communication. Building prototypes for lighting control helped the developers experience the security limitations of the architecture, and helped foster new discussions about the security requirements of diverse uses of the Internet. Finally, as at CENS, using prototypes could result in developer realizations about ethical issues. When staff deployed NDN-enabled webcams in their office, they realized that pervasive caching inherent to NDN, in which data is replicated indiscriminately across servers, raised privacy concerns when combined with office webcams.

Comparing results from both projects highlights successful techniques for conducting anticipatory ethics. The most successful techniques were work practices already embedded in the design context. These included imaging use cases, interdisciplinary work, and prototype testing. A common thread among the most successful values levers is that they helped developers to imagine their technologies in relevant social contexts. Many ethical issues (for example, privacy, consent, and equity) are impacted by social and cultural contexts which dictate norms (Nissenbaum, 2009; Schwartz, 1994). As evidenced by both CENS and NDN observations, privacy concerns were raised by contexts in which data was shared in new and surprising ways; equity concerns were raised by contexts in which constrained resources dictated limitations on what could be accessed. Imagining use contexts helped engineers discuss the social norms and ethics relevant to their design.

Successful techniques for anticipatory ethics also help engineers embody the experience of using their technology (Dourish, 2001). Beyond thinking about social contexts, embodying the technology helped developers experience ethical concerns as personal and relevant. Techniques ranging from prototype self-testing to interdisciplinary work encouraged engineers to feel and act like a user, and embody some of the concerns future users might have.

Software engineers often face a tension between firm grounding in particular contexts, and a more abstracted model of design (Rosenberg, 2008). Values levers are work practices push back against abstraction, helping to ground discussions about ethics and social impact. This insight can help researchers interested in technology ethics think about values levers in diverse engineering workplaces. Though workplaces may have different work practices and technological foci, finding methods to counter abstraction can help introduce values levers. Looking for the work practices which ground technologies in a social context and help developers embody their technologies can improve the practice of anticipatory technology ethics.

5   Challenges for Anticipatory Technology Ethics

Though practices which ground technology development and help developers experience technologies may improve the success of anticipatory technology ethics, there are serious challenges and barriers for anticipatory ethics raised by both diversity of work practices among development teams, and the material nature of technologies under development. Comparing anticipatory ethics work in CENS and NDN highlights challenges that stem from differences in engineering work practices, as well as the design of user-facing vs. infrastructural technologies. This second difference – dissimilarity in the nature of the object being developed – introduces the challenge of *at what layer* ethical issues can be found in technology design.

One major difference between the two projects was their mode of collaboration. CENS employees worked in the same physical location; NDN employees spanned ten campuses throughout the US. Talking about ethical challenges and social impacts during design requires trust among team members. Groups build such trust over time, and trust is built more quickly when working groups have frequent face-to-face interactions (Hinds & Bailey, 2003). Because of its relationship to building trust, co-location mattered to the anticipatory ethics effort. While CENS engineers regularly had conversations about the social impacts of their design, NDN engineers had fewer. NDN engineers were challenged to find the right time and place for such conversations; face-to-face retreats were dominated by formal presentations, and didn't foster the same sorts of casual conversations hosted in the CENS lab (Shilton & Koepfler, 2013). Because they have fewer opportunities for building trust and having informal conversations, distributed teams may face a greater challenge when practicing anticipatory ethics. More studies of ethical discussions and decision-making on distributed teams could help to add details to this challenge and explore modes of interaction which might help overcome it.

A major finding of these contrasting projects is that the nature of the technology being designed matters to the practice of anticipatory ethics. CENS was designing a user-facing technology: a system designed to be operated directly by consumers. NDN was developing an infrastructural technology: an underlying layer designed not to be noticed except by web application developers. User-facing

technologies were much easier for developers to imagine in a social context, with more obvious norms and information rules. Though CENS developers could not imagine all future uses of their systems, they regularly participated in discussions of how their systems would fit into real-world scenarios such as community mapping and storytelling, public health and wellness research, environmental advocacy, and citizen science. Scenarios such as community mapping helped them discuss ethical decisions such as data representation and literacy. Scenarios such as public health research evoked values such as privacy and doctor-patient confidentiality.

Computer ethics researchers such as Albrechtslund (2007) have pointed to the difficulty of predicting future use during anticipatory ethical analyses. While this problem exists for all technologies to some extent, it is heightened by the development of infrastructure, which is meant to have almost infinite interpretive flexibility. This sort of ethical future-casting was, as a result, much more challenging for the NDN team. Imagining users and use contexts for NDN was a difficult task because users could conceivably be anyone, and the technology was meant to perform the diverse duties of the current Internet, plus unknowable future functions. This challenge was reflected in many team disagreements. For example, the team struggled to decide whether to build standardized time stamps into data packets. Doing so would enhance security, but would lock NDN into an assumption that every machine on the network could access standardized time. Similarly, the team disagreed about whether data names should derive from current real-world institutions. Doing so would enhance data provenance and trust, but would lock NDN into a naming system that could challenge network neutrality by making it easy to prioritize data from powerful institutions. Though imagining users and use cases was more challenging in NDN, these disagreements proved fruitful to *discussing* values in design. But it was often hard to make ethical *decisions* – e.g. deciding to prioritize security over flexibility in the case of universal time stamps – because the future context remained speculative.

Working on the NDN project also emphasized a problem of technology *layers* previously unaddressed in the research at CENS. A frequent comment when discussing issues such as privacy, intellectual freedom, or equity with the NDN team was, as one NDN PI put it in an email: "This is not an NDN problem; it's an applications problem." (This was also expressed by other network architecture colleagues in variations such as "That's not an architecture problem!") There was often fundamental skepticism among team members that the network layer could embody non-technical values. Though infrastructural technologies are not neutral and certainly have ethical consequences (Bowker & Star, 2000; Winner, 1980), features of their design frustrate a number of the techniques that might ordinarily encourage engineers to consider social contexts and social implications. It required long and complex conversations to settle upon what, exactly, were ethical problems relevant to the architectural layer. The high interpretive flexibility of infrastructure technologies provide more ethical "outs," encouraging engineers to believe that social problems could and should be solved by policy or social action at a much later stage. The technology pushes back; the material constraints imposed by the kind of technology being developed can make the practice of anticipatory technology ethics more difficult.

A final challenge to practicing anticipatory technology ethics within NDN was a material one. The realities of designing an architecture meant that huge amounts of design and planning had to be completed before even rudimentary applications could be prototyped in NDN. Prototype applications such as the file-sharing service described above arrived in the third year of NDN funding. As discussed above, embodied experience through prototype testing proved particularly good for making values issues relevant to a research team's work. But critical, values-laden decisions about routing protocols and data names had to be made before any prototypes could be built. This meant that some values were concretized in the architecture before developers could experience values concerns through prototype testing. Though these decisions may be changed in the future, a powerful status quo has already been established. This observation presents a challenge of timing for information researchers practicing anticipatory technology ethics. What new interventions might be targeted to earlier stages of infrastructural design?

6   Conclusion

Comparing the CENS and NDN projects illustrates that varying work practices impact anticipatory technology ethics by changing the salience and effectiveness of values levers during development. And further, the comparison illustrates that the nature of the technology under development matters to considering the social contexts in which a technology will be deployed, and therefore the values which developers can or should attempt to build into their design decisions. Practicing anticipatory ethics is more challenging for infrastructural technologies than those that are user-facing. Awareness of these

challenges can help the information research community improve the current practice of anticipatory ethics.

First, finding and strengthening existing development practices that cultivate ethics conversations can help those interested in technology ethics find a closer fit "into the fray" of technology development. Taking advantage of these practices can help ethics advocates be more successful within design, and can improve ethical development.

Second, understanding the limitations placed on anticipatory ethics by the logistical and material constraints of design indicate new research questions for anticipatory ethics. We must understand that technologies at different layers of abstraction or distance from users will face different challenges for ethical analysis during design. If the information research community believes that infrastructures have ethical consequences just as user-facing technologies do, we must take this challenge seriously. Developing new techniques for anticipatory ethics for emerging infrastructural technologies is a critical challenge for the ethical computing community.

References

Agre, P. E. (1997). Toward a critical technical practice: lessons learned in trying to reform AI. In G. C. Bowker, L. Gasser, S. L. Star, & B. Turner (Eds.), *Social Science, technical systems, and cooperative work: beyond the great divide* (pp. 131–158). Hillsdale, NJ: Erlbaum.

Albrechtslund, A. (2007). Ethics and technology design. *Ethics and Information Technology*, *9*(1), 63–72.

Bowker, G. C., & Star, S. L. (2000). *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA and London: The MIT Press.

Brey, P. A. E. (2012). Anticipating ethical issues in emerging IT. *Ethics and Information Technology*, *14*(4), 305–317.

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. Cambridge, MA and London: The MIT Press.

Flanagan, M., Nissenbaum, H., & Howe, D. C. (2008). Embodying values in technology: theory and practice. In J. van den Hoven & J. Weckert (Eds.), *Information Technology and Moral Philosophy* (pp. 322–353). Cambridge: Cambridge University Press.

Friedman, B., & Hendry, D. G. (2012). The envisioning cards: a toolkit for catalyzing humanistic and technical imaginations. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 1145–1148). New York, NY, USA: ACM. doi:10.1145/2207676.2208562

Friedman, B., Kahn, P. H., & Borning, A. (2006). Value sensitive design and information systems. In D. Galletta & P. Zhang (Eds.), *Human-Computer Interaction and Management Information Systems: Applications* (Vol. 6). New York: M.E. Sharpe.

Friedman, B., & Nissenbaum, H. (1997). Bias in computer systems. In B. Friedman (Ed.), *Human values and the design of computer technology* (pp. 21–40). Cambridge and New York: Cambridge University Press.

Hinds, P. J., & Bailey, D. E. (2003). Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams. *Organization Science*, *14*(6), 615–632.

Jarvenpaa, S. L., & Leidner, D. E. (1998). Communication and Trust in Global Virtual Teams. *Journal of Computer-Mediated Communication*, *3*(4), 0–0.

Johnson, D. G. (1997). Is the global information infrastructure a democratic technology? *SIGCAS Comput. Soc.*, *27*(3), 20–26. doi:10.1145/270858.270865

Johnson, D. G. (2000). *Computer Ethics* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Johnson, D. G. (2007). Ethics and Technology "in the Making": An Essay on the Challenge of Nanoethics. *NanoEthics*, *1*(1), 21–30.

Johnson, D. G. (2011). Software Agents, Anticipatory Ethics, and Accountability. In G. E. Marchant, B. R. Allenby, & J. R. Herkert (Eds.), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight* (pp. 61–76). Springer Netherlands. Retrieved from http://link.springer.com.proxy-um.researchport.umd.edu/chapter/10.1007/978-94-007-1356-7_5

Knobel, C. P., & Bowker, G. C. (2011, July). Values in design. *Communications of the ACM*, *54*(7), 26–28.

Latour, B. (2008). A cautious prometheus? A few steps toward a philosophy of design (with special attention to Peter Sloterdijk). In F. Hackne, J. Glynne, & V. Minto (Eds.), *Proceedings of the 2008 Annual International Conference of the Design History Society* (pp. 2–10). Falmouth, UK: Universal Publications.

Manders-Huits, N., & Zimmer, M. (2009). Values and pragmatic action: the challenges of introducing ethical intelligence in technical and design communities. *International Review of Information Ethics*, *10*, 37–44.

Nissenbaum, H. (2009). *Privacy in context: technology, policy, and the integrity of social life*. Stanford, CA: Stanford Law Books.

Rosenberg, S. (2008). *Dreaming in code: two dozen programmers, three years, 4,732 bugs, and one quest for transcendent software*. New York: Three Rivers Press.

Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues*, *50*(4), 19–45. doi:10.1111/j.1540-4560.1994.tb01196.x

Sclove, R. E. (1995). *Democracy and Technology* (1st ed.). New York: The Guilford Press.

Sengers, P., Boehner, K., David, S., & Kaye, J. "Jofish." (2005). Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility* (pp. 49–58). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1094562.1094569

Shilton, K. (2010). Participatory sensing: building empowering surveillance. *Surveillance & Society*, *8*(2), 131–150.

Shilton, K. (2013). Values levers: building ethics into design. *Science, Technology & Human Values*, *38*(3), 374 – 397.

Shilton, K. (2014). Anticipatory ethics for a future Internet: analyzing values during the design of an Internet infrastructure. *Science and Engineering Ethics*. doi:10.1007/s11948-013-9510-z

Shilton, K., & Koepfler, J. A. (2013). Making space for values: communication & values levers in a virtual team. In *Proceedings of the 6th International Conference on Communities and Technologies (C&T 2013)*. Munich, Germany: ACM.

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

Winner, L. (1989). *The Whale and the Reactor: A Search for Limits in an Age of High Technology* (1st ed.). Chicago: University Of Chicago Press.