Hayley Hung
Technical University of Delft

Teaching Privacy: Multimedia Making a Difference

Julia Bernd, Blanca Gordo, Jaeyoung Choi, Bryan Morgan, Nicholas Henderson, Serge Egelman, Daniel D. Garcia, and Gerald Friedland
International Computer Science Institute and University of California, Berkeley

"When does the information you fill in on a Web form get submitted to the server? Is it submitted when you click okay, or could it already be submitted before that?" A high school computer science teacher asked one of the coauthors these questions. A student had told the teacher a driving school seemed to have her contact information after she visited their website and started to fill in a form, even though she changed her mind and never clicked any buttons to submit that information.

As experts, we know that the answer to the teacher's questions depends on many factors, in particular whether the school used CGI or servlet technology to host the site or whether it used the more recent Ajax or HTML5. But even a relatively tech-savvy high school teacher might find this seemingly simple question hard to answer. Most high school teachers lack the technical background, and even if they did know the answer, how would they explain it to a teenager? Even if they successfully explained it, there is a further issue: Given that the average human would find it rather difficult to tell whether a site used HTML5 versus CGI scripting, what actionable suggestion could be made to help make Web browsing safer?

Much current multimedia research and development centers around applications with great potential to compromise the privacy of Internet users, directly or indirectly. Multimedia scientists and engineers are developing new methods to automatically identify the people depicted in an image or video, or even the person who uploaded it; detect what's happening in an image or video; and determine where it was recorded. As researchers, we often don't think about the privacy implications of such developments down the road. Of course, scientists and technologists want to continue pursuing these fruitful and interesting avenues for enhancing multimedia analysis and retrieval capabilities, but at the same time, we can mitigate the potential negative effects of those capabilities by using our expertise to educate the public about the effects of the new technology on their privacy.

In this article, we describe the Teaching Privacy project at the International Computer Science Institute (ICSI) and the University of California, Berkeley, in which an interdisciplinary team of researchers and educators are developing educational tools to empower K-12 students and college undergraduates in making informed choices about privacy. We describe our interdisciplinary approach to developing and disseminating engaging, interactive educational apps that demonstrate what happens to personal information on the Internet, with a particular focus on multimedia, and our approach to explaining the underlying social and technical principles in accessible terms.

From Research to Education

Teaching Privacy grew out of several strands of work at ICSI and UC Berkeley. These strands came together as researchers working in different areas realized that the explosion of multimedia content being shared on social media was giving rise to a new need for credible information about online privacy—information based on solid technical knowledge rather than panicked speculation.

One major motivation grew out of theoretical research at ICSI on the privacy implications of multimedia technology, including speaker-matching[1,2] and multimedia-retrieval techniques. For example, while working on multimodal location estimation—automatically identifying where an image or video without geotags was recorded according to its visual and acoustic similarity to geotagged media[3,4]—coauthor Gerald Friedland's Multimedia Research Group at ICSI became aware of how few Internet users (at that time) even realized that the images and

1070-986X/15/$31.00 © 2015 IEEE

Published by the IEEE Computer Society

Guidelines for Socially Responsible, Inclusive Privacy Education

By Blanca Gordo

The ongoing integration of technological innovations into social structures is resulting in substantial changes in privacy—and even our understanding of what constitutes privacy—that affect everyone in the global society of end users. Harmful effects arise both from a lack of comprehensive consumer privacy protections and a lack of scientifically based educational guidelines on how to teach end users about privacy and contextualize how it is changing.

These issues affect both long-term users of Internet technology and new entrants, who are frequently low-income, less-educated, older, and/or immigrant populations. One key difference between experienced and inexperienced users, however, is that new entrants are more vulnerable to potential dangers online because they lack the systematic knowledge that comes with time and continuous direct experience. New entrants are still learning new technical skills, new ways of doing things, and new social norms.

To address this disparity and to provide a framework for privacy decision makers, it is necessary to develop a comprehensive, cross-disciplinary theoretical concept of privacy. Such a grounded conceptualization must account for, among other things, the interconnectivity of social, institutional, economic, political, and technology systems; the dynamic effects of social context and governance; the function of network communication technology as transferring and connecting past, present, and future bits of information; the nature of information as a commodity; the limitations of the legal system with regard to corporations' treatment of personal information, including undisclosed tracking and third-party data use; the trade-offs between security and privacy in government surveillance; sophisticated advancements in analytics and the availability of big data; and the range of individuals' views and online experience.

Such a conceptual framework can help us build more universally applicable educational messages that can help anyone grasp the trade-offs inherent in Internet use, the common possibilities for harm, and the general ramifications of online behavior. We have much to learn about the most strategic, effective ways to explain how technology design, policy, corporate structure, social behaviors, and values lead to privacy-related outcomes. But understanding how privacy works in people's daily lives, and how a range of populations conceptualize those workings and process information, can inform the design of effective educational resources that resonate with everyone's daily experience.

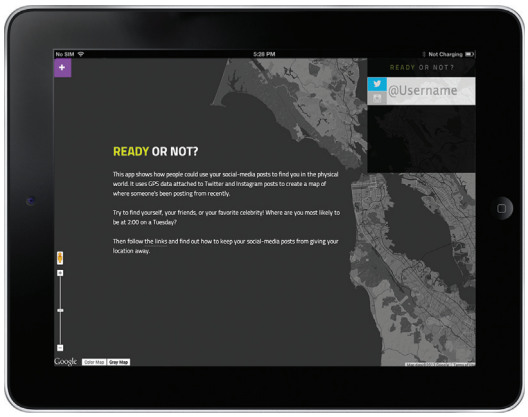
We need to build tools that help users grasp the mechanisms of transmission, collection, storage, and leakage of bits of information that make up someone's personal profile. To accomplish this, digital tools can simulate the vital components of privacy, including technology systems, social behavior, cultural norms, and governance, and illustrate (via metaphors as well as literal descriptions) how technology works within the frames of networked structures and of social-institutional systems.

Taking the experience of new entrants into account when building privacy-related educational tools will also help us to better teach long-term users, who also frequently have misperceptions about privacy and may tend to take it for granted that they can control their personal information. For example, even many experienced users believe that they can remain anonymous if they use an anonymization proxy or that secondary use of personal information must be authorized by the individual. The easier we can make it for the disconnected to grasp the workings of privacy, the easier it will be for us to educate everyone.

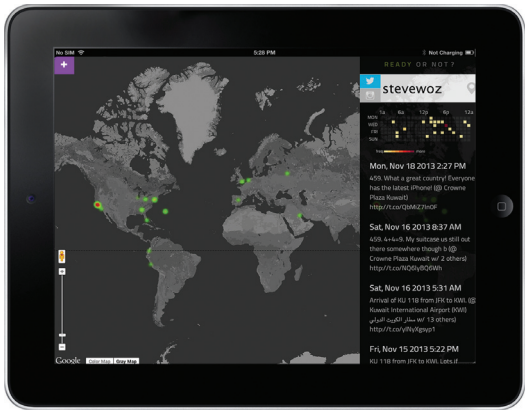
videos they uploaded often included GPS metadata in the first place, much less that it was possible to estimate the recording locations of nontagged media. The multimedia group began working with privacy and security researchers at ICSI, including Robin Sommer and Nicholas Weaver, to explore the potential privacy implications of so much multimedia data and metadata being constantly uploaded and shared.[5]

One of the initial sparking moments for the education project was when coauthor Daniel Garcia invited Friedland to talk about this multimedia-privacy research in a professional development session for high school computer science teachers. The teachers found it by far the most engaging topic of the day and asked a multitude of questions—not just because it was technically interesting, but because they were so eager for information about privacy that they could pass on to their students. They were well aware that their students needed more information and guidance about online privacy but felt unqualified to teach about it because they were not themselves sufficiently well-versed in the technical details.

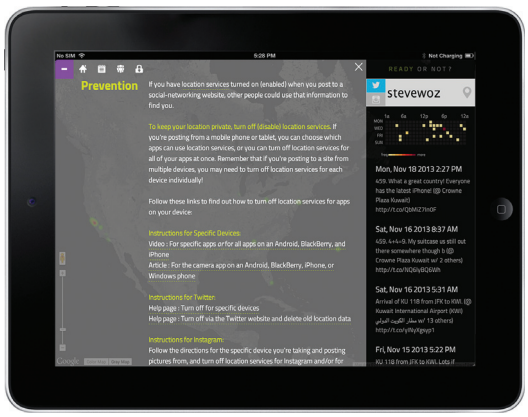
Similarly, coauthor Blanca Gordo's interest arose out of her social science research on developing a theoretical framework for



(a)



(b)



(c)

Figure 1. Ready or Not? interactive app. (a) Welcome screen, (b) a 2013 heatmap and timeline result showing the frequent coordinates of a not entirely random Twitter user, and (c) information about how to turn off location services.

understanding network technology and a pedagogy for teaching technology to people who are newly going online (see the "Guidelines for Privacy Education" sidebar). Those charged with teaching these new entrants about the technology know it is vital to explain the privacy implications of online activities, but they too do not have the necessary expertise and find few ready-made learning materials.[6] Meanwhile, coauthor Serge Egelman's research in human-computer interaction focuses on how people make decisions about their online privacy and how to help them make better ones.

Teaching Privacy also includes the Berkeley Foundation for Opportunities in Information Technology (BFOIT),[7] of which coauthor Nicholas Henderson is a program lead. BFOIT provides historically underrepresented ethnic minority and female middle and high school students with knowledge, resources, practical programming skills, and guidance in their pursuit of higher education and production of technology. BFOIT educators were aware of a deep need for accurate, in-depth privacy education among the program participants—in their current online activities, but especially if they were to go on to become technology designers.

Multimedia Apps and Learning Tools

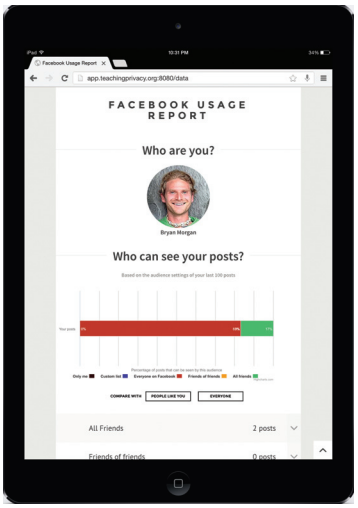
The unusual level of interest from K-12 and college teachers in their multimedia privacy research inspired ICSI's Multimedia Group to begin a small project in which they would work with educators to develop a set of interactive privacy-visualization apps and learning tools for classroom use. These hands-on learning tools were designed to help educators raise students' awareness about the privacy implications of social media use, especially of multimedia sharing. The project quickly grew, attracting an interdisciplinary team of researchers working on privacy.

Our first learning tool, a production-quality app called *Ready or Not?*, was intended to draw students' attention to the risks posed by geotagged social media posts. Given a username, it pulls recent Twitter or Instagram posts from the sites' APIs and uses attached GPS metadata to create a heatmap and timeline of where that user has been posting from recently. It then gives prevention tips for how to keep social media posts from giving away location information (see Figure 1).

In addition to being used in classrooms, Ready or Not? also inspired stories by several high-profile national and local news agencies, drawing the attention of a much larger audience to the risks of geotags.[8–10] As of November 2014, more than 25,000 unique users had tried the app.

To date, the Teaching Privacy project has also created two other interactive apps tailored to help young people visualize how

(a)



(b)

Figure 2. Teaching Privacy online learning tools. (a) Oh, the Places You (and Your Data) Will Go! and (b) Social Media Usage Report.

information persists and travels on the Internet along with a number of classroom activities and two videos. Here is a sampling of these learning tools:

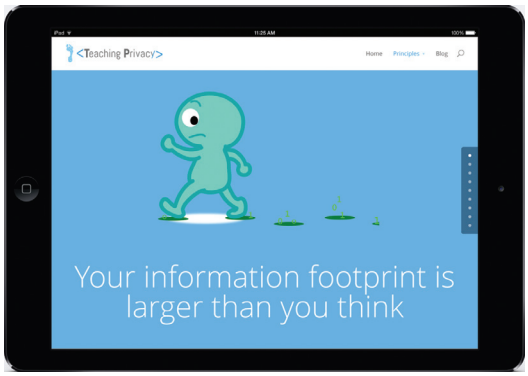
- *Oh the Places You* (*and Your Data*) *Will Go*!: This choose-your-own-adventure activity for the classroom illustrates principles of privacy we deal with in daily life. Students make privacy decisions for a hypothetical person and see what happens. The result is interspersed with teaching material about the implications of those choices and encourages users to actively engage with privacy issues (see Figure 2a).
- *Social Media Usage Report*: This app, which doubles as an evaluation tool, allows a user to visualize how many people see different types of information about them on Facebook and compare themselves with other people who have similar personalities. Along with the visualization, the app gives detailed instructions and advice on changing one's Facebook privacy settings (see Figure 2b).
- *Welcome to the Internet*: Aimed at college or advanced high school level classes, this stand-alone online lab demonstrates how the structure of the Internet affects students' privacy. It includes a technical introduction to concepts like routing, host redirects, and browser signatures as well as exercises based on the Teaching Privacy content and website.
- *Digital Footprints*: The first in our series of classroom-ready educational videos, Digital Footprints explores the many factors and activities that add to each person's ever-growing information footprint and touches on some of the strategies that can limit it. The narration is accompanied by live-drawn humorous visuals that turn each point into a memorable story.

In addition to the authors, Eungchan Kim, Arany Uthayakumar, and Ketrina Yim helped produce the tools described here. All of these resources are accessible via our Teachers' Portal at www.teachingprivacy.org/teachers-portal.

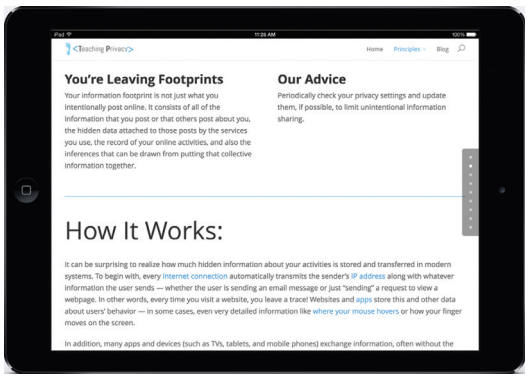
A Curriculum for Teaching Privacy

As we were developing the educational apps, we also began to develop a content base explaining how and why personal information travels around the Internet, along with practical guidance about how young people can better protect themselves online, given the facts on the ground. The project team together identified "Ten Principles for Online Privacy," a set of fundamental, but often counterintuitive, precepts around which to focus the more extensive explanations and suggestions. For example, our first principle, "Your information footprint is larger than you think," draws attention to important (but perhaps esoteric) technical concepts like metadata, device identifiers, and data-mining and inference techniques.

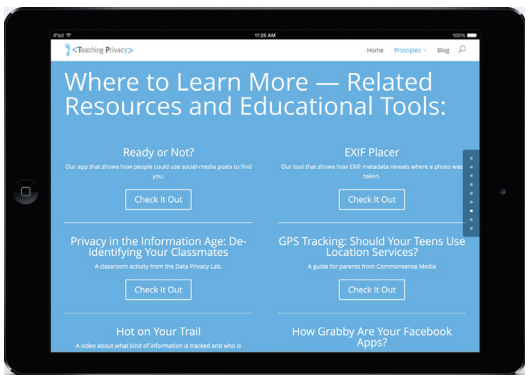
The apps and content base are hosted on the Teaching Privacy website, www.



(a)



(b)



(c)

Figure 3. Excerpts from the Teaching Privacy webpage explaining the privacy principle "Your information footprint is larger than you think." (a) Principle, (b) summary, and (c) resources. (Illustration by Ketrina Yim.)

teachingprivacy.org (see Figure 3). The website includes a page for each of the Ten Principles, with an easy-to-understand description of the underlying technical and social concepts; suggestions for actions people can take to better protect their privacy; "ignite" questions to provoke critical thinking; and links to related resources, learning tools, and guides.

To make the resources we are creating easier to integrate into classroom lessons, we are currently developing a set of flexible, classroom-ready teaching modules and a teachers' guide, together called TROPE (Teachers' Resources for Online Privacy Education). We are making these materials available at http://teachingprivacy.org/teachers-portal as we build them. We invite educators to use them and provide feedback on how to make them most effective as classroom tools across a variety of situations.

As public concern over online privacy in the United States has grown in the last few years,[11] more educational resources have become available. However, we believe Teaching Privacy is unique in our combination of accurate, accessible technical details; comprehensive coverage; and attention to practical strategies.

Despite increased public awareness of privacy issues, most people still do not have a good handle on the specific mechanisms involved, nor the steps they can take to protect their privacy online. In addition, a major challenge we have faced in designing learning tools is how to engage young people in actively thinking about their online privacy without resorting to scare tactics. Such tactics can unintentionally suggest that there's no point in even trying to manage one's online privacy, which is problematic given that, in our interactions with students, we found that many young people are already close to relinquishing the idea of controlling their privacy altogether.

Our approach therefore focuses on linking awareness of the specific privacy implications of social multimedia with knowledge of actionable, practical strategies for managing privacy. For example, in the case of the Ready or Not? app, we teach users about the inferences about daily habits that can be drawn from GPS-tagged images and then show them how they can control which apps use location services. We also encourage them to use those strategies proactively. One of our objectives in Teaching Privacy is for young people to understand that there is always something they can do to manage their online privacy, whether by changing their privacy settings, choosing different services, or communicating with friends and family about their privacy preferences.

Spreading the Word and Gaining Inspiration

Throughout the project, we have been sharing our learning tools and materials through multiple curriculum-resourcing and professional development channels associated with Berkeley's CS 10 course, The Beauty and Joy of Computing (BJC, http://bjc.berkeley.edu). BJC engages non–computer science majors with

Sharing our Curriculum with Teachers and Students, Worldwide

By Daniel D. Garcia

The Teaching Privacy materials have been incorporated at UC Berkeley through our The Beauty and Joy of Computing (BJC) nonmajor course. BJC draws in students at all levels, freshman through graduate (and staff), from every department on campus. The course has been chosen twice as a College Board Advanced Placement Computer Science Principles (AP CSP) pilot course, recognizing it as a model university course that covers the learning objectives in their curriculum framework.

We teach beginning programming using Snap! (http://snap.berkeley.edu/), a friendly, blocks-based language; cover several big ideas like abstraction, recursion, and higher-order functions; and discuss the social implications of computing. The Teaching Privacy materials fit perfectly into that last theme, where we highlight the trade-offs many people make with computing innovations, balancing convenience with the privacy implications.

For example, many smartphone users appreciate automated photo geotagging because it allows their photos to be automatically organized by location. They often don't realize (or do realize and find it worth the trade-off) that sharing geotagged photos on the Web can reveal the location of their home and even tell potential burglars when their home is vacant.

BJC reaches 700 UC Berkeley students per year. Since 2010, the National Science Foundation has provided us with funding so we can offer professional development (PD) to high school teachers for our BJC curriculum. To date, more than 200 high school teachers have been to our summer PD sessions, and many are now teaching this material in their high schools. We have just received another NSF grant to take the course to 100 more teachers in New York City, the nation's largest and most diverse school system. (See the Bringing BJC to New York City High Schools webpage, http://bjc.berkeley.edu/website/bjc4nyc.html, for more details.) We are currently developing a massive open online course (MOOC) version of our BJC course, entitled BJCx, that will launch on Labor Day 2015.

With the AP CSP exam beginning in the spring of 2017, we look forward to this course, and the Teaching Privacy material contained within it, reaching thousands of students across the country and the world.

technological concepts, including the social implications of computing (see the "Sharing Our Curriculum" sidebar for more details).

We receive valuable feedback by using the learning tools and materials in BJC. For example, we presented students with an early draft of the Ten Principles with brief explanations and asked which were most surprising, which explanations were confusing, and what we might be missing. This input helped us refine the principles and explanations. For example, many students pointed out that the seeming contradiction between the principles "Identity is not guaranteed on the Internet" and "There is no anonymity on the Internet" could be confusing, so we revised the materials to explain explicitly how both can be true, depending on the resources and technical knowledge of both the person trying to hide their identity and the person trying to figure it out.

We also discuss the resources with high school educators and pilot-test them with BFOIT students, who have helped us identify which approaches are likely to be most engaging. Interns from both BJC and BFOIT have provided guidance on what their peers in our target demographic do and don't know regarding online privacy issues, provided critical feedback on content and learning tools, and even led the creation of some learning tools.

Figure 4. Teaching Privacy demonstration at the Cal Day open house in April 2014. UC Berkeley student Madeeha Ghori demonstrates the Ready or Not? interactive app and explains the implications to prospective Berkeley students and their parents. (Photo by Bryan Morgan.)

In more general outreach, we have presented Teaching Privacy in public lectures and in a popular "What Does the Internet Know

Experts can provide the public with a realistic understanding of what individuals actually can and cannot do to protect their privacy, in practical terms.

About You?" interactive lab attended by hundreds of high school students and their parents during an open house at UC Berkeley (see Figure 4).

We will be introducing the TROPE materials in March 2015 at the annual gathering of the ACM's Special Interest Group on Computer Science Education (SIGCSE). As well as introducing the materials, this workshop will allow us to solicit feedback and on-the-ground stories, so we can gain a better understanding of specific problems faced by both teachers and students.

Engaging with people about online privacy at these public events not only improves and gains buy-in for the Teaching Privacy project, it also provides new inspiration and provokes new questions for our respective research programs.

Leveraging Expertise to Make a Difference

The current lively public discussion about online privacy provides a new opportunity for technical experts to contribute to an informed dialogue. In particular, experts can provide the public with a realistic understanding of what individuals actually can and cannot do to protect their privacy, in practical terms. The Teaching Privacy project provides one model for such contributions on a national scale (see the "Sharing Our Curriculum" sidebar), with multimedia researchers contextualizing practical advice in explanations of how multimedia content creation and distribution work in the context of Internet architecture.

As well as demonstrating the utility of contributions from multimedia experts, Teaching Privacy has shown the importance of cross-disciplinary collaboration in making those contributions most effective. The expertise of educators and social scientists—for example, Gordo's fieldwork with community college students and parents of school-age children—has grounded our work in a broad understanding of what different sectors of the public know (or don't know) about online privacy and how best to reach them. This understanding increases our potential to engage people about actively managing their privacy. In addition, interaction with computer science education also increases our potential to spark the interest of young people in developing and researching (not just using) multimedia technology.

We hope that the successes of this project so far will raise awareness among computer scientists and engineers like the readers of this magazine, not only of the essential need to consider privacy concerns in researching and designing multimedia applications, but also of the potential active contributions we can make by reaching out to nonexperts and supporting public understanding of the ideas and principles that engage us most.

Lastly, we are asking you, the readers, to help with your expertise. Contact us and tell us your privacy story or about your privacy-related multimedia application. We hope to link more and more interesting projects and stories from our website, making it an ever-growing resource that anyone can use as a basis for teaching or learning about privacy. **MM**

Acknowledgments

Teaching Privacy has been the work of a whole team of researchers, educators, and interns. In addition to the authors listed here, project contributors include Alexis Conway, Orpheus Crutchfield, Isha Doshi, Melia Henderson, Jeffrey Jacinto, Eungchan Kim, Itzel Martinez, Gerardo Sánchez, Robin Sommer, Arany Uthayakumar, and Ketrina Yim.

Teaching Privacy is supported in part by National Science Foundation grants CNS-1065240: Understanding and Managing the Impact of Global Inference on Online Privacy and DGE-1419319: Teachers' Resources for Online Privacy Education as well as by the National Telecommunications and Information Administration's Broadband Technology Opportunities Program, through the California

Connects program administered by the Foundation for California Community Colleges. Any opinions, findings, and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the funders.

References

1. H. Lei et al., "User Verification: Matching the Uploaders of Videos across Accounts," *Proc. IEEE Int'l Conf. Acoustic, Speech, and Signal Processing* (ICASSP), 2011, pp. 2404–2407.
2. X. Anguera Miro et al., "Speaker Diarization: A Review of Recent Research," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 20, no. 2, 2012, pp. 356–370.
3. G. Friedland, O. Vinyals, and T. Darrell, "Multimodal Location Estimation," *Proc. ACM Int'l Conf. Multimedia*, 2010, pp. 1245–1251.
4. M. Larson et al., "Automatic Tagging and Geotagging in Video Collections and Communities," *Proc. 1st ACM Int'l Conf. Multimedia Retrieval (ICMR)*, 2011, article no. 51.
5. G. Friedland and R. Sommer, "Cybercasing the Joint: On the Privacy Implications of Geotagging," *Proc. 5th USENIX Workshop on Hot Topics in Security (HotSec)*, 2010, pp. 1–8.
6. B. Gordo, "Developing a Framework for Understanding Online Privacy. Appendix K," *California Connects: Improving Digital Opportunities in Underserved California Communities*, Sept. 2013. http://digitalequality.net/understanding-online-privacy.
7. O.S.L. Crutchfield et al., "Berkeley Foundation for Opportunities in Information Technology: A Decade of Broadening Participation," *ACM Trans. Computing Education*, vol. 11, no. 3, 2011, article no. 15.
8. A. Truong, "Privacy App Pinpoints Your Exact Location Using Social Media," *Fast Company*, 3 Sept. 2013; www.fastcompany.com/3016662/fast-feed/privacy-app-pinpoints-your-exact-location-using-social-media.
9. K. Shubber, "Mapping Websites Reveal Just How Stupid It Is to Geotag Your Tweets," *Wired UK*, 4 Sept. 2013; www.wired.co.uk/news/archive/2013-09/04/twitter-geotagging.
10. J. Watts, "ConsumerWatch: Social Media Users May Be Revealing Too Much about Location," video, KPIX/CBS SF Bay Area, 4 Nov. 2013; http://sanfrancisco.cbslocal.com/video/9491297-consumerwatch-social-media-users-may-be-revealing-too-much-about-location/.
11. M. Madden, "Public Perceptions of Privacy and Security in the Post-Snowden Era," Pew Research Center, Nov. 2014; www.pewinternet.org/2014/11/12/public-privacy-perceptions/.

Julia Bernd is a linguist and social science researcher who is currently adding interdisciplinarity to the Audio & Multimedia Group at the International Computer Science Institute, a nonprofit research institute affiliated with the EECS Department at the University of California, Berkeley. Contact her at jbernd@icsi.berkeley.edu.

Blanca Gordo is a social scientist and senior researcher in the Artificial Intelligence Group at ICSI and a visiting scholar at the Institute for the Study of Societal Issues at UC Berkeley. Contact her at blanca@icsi.berkeley.edu.

Jaeyoung Choi is a computer scientist and staff researcher at ICSI, working on multimodal location estimation of videos. Contact him at jaeyoung@icsi.berkeley.edu.

Bryan Morgan is a multimedia designer who works on Teaching Privacy at ICSI as well as a master's student in information management and systems at UC Berkeley. Contact him at brynamo@icsi.berkeley.edu.

Nicholas Henderson is the program director for Science for Youth (SCI-FY), the middle school component of the BFOIT computing outreach program at ICSI.

Serge Egelman is a senior researcher in the Networking & Security Group at ICSI, and he holds an appointment in UC Berkeley's Electrical Engineering and Computer Sciences Department. Contact him at egelman@icsi.berkeley.edu.

Daniel D. Garcia is a senior lecturer SOE in the EECS Department at the University of California, Berkeley, the codeveloper of the BJC CS Principles curriculum, and an ACM Distinguished Educator. Contact him at ddgarcia@cs.berkeley.edu.

Gerald Friedland is the director of the Audio and Multimedia Group at ICSI and a part-time lecturer in the UC Berkeley EECS Department. Contact him at fractor@icsi.berkeley.edu.