Federal Register Notice 87 FR 35250, https://www.federalregister.gov/documents/2022/06/09/2022-12432/request-for-information-on-advancing-privacy-enhancing-technologies, June 9, 2022

Request for Information (RFI) on

Advancing Privacy Enhancing Technologies

Carnegie Mellon University

Carnegie Mellon University Response to Office of Science and Technology Policy (OSTP) RFI on Advancing Privacy-Enhancing Technologies

Alessandro Acquisti
Professor, Heinz College of Information Systems and Public Policy

Travis Breaux
Associate Professor, Institute for Software Research, School of Computer Science

Jason Hong
Professor, Human-Computer Interaction Institute, School of Computer Science

Norman Sadeh
Professor/Co-Director, CMU Privacy Engineering Program
Institute for Software Research, School of Computer Science

Steven Wu
Assistant Professor, Institute for Software Research, School of Computer Science

Introduction

This document is a response to OSTP's Request for Information on Advancing Privacy-Enhancing Technologies (87 FR 35250), by several faculty at Carnegie Mellon University's CyLab Security and Privacy Institute. The authors' areas of expertise include AI, HCI, privacy, cybersecurity, software engineering, fairness, economics, and more.

We have organized our response below by the questions posed in the RFI.

1. *Specific research opportunities to advance PETs:*

Studies to Understand What Influences End-Users to Adopt PETs
Support more studies as to what technical, usability, social, and other factors lead end-users to adopt (or not) different kinds of PETs, as well as what kinds of interventions are useful. There are already a vast number of PETs, but there also seems to be limited adoption of them by end-users. For example, in 2018, Google reported that less than 10% of their users used two-factor authentication, despite the practice being widely advocated by security experts. A deeper understanding of why people do or don't adopt PETs could help lead to better designs and interventions.

For example, Das et al [DKDH14] found that about half of changes in security behaviors were due to a social influence. In follow-up work, Das et al [DKDH14] found that simple social interventions (e.g. "108 of your friends use extra security settings") increased people's likelihood of adopting features. In related work, Krsek et al [KWDH+22] found that seeing what strangers recommended for Facebook settings as well as self-reflection were also effective in getting people to change their privacy settings.

In another line of work, Das et al [DKDH15] analyzed large-scale data about use of security features from Facebook users and found that social influence affects one's likelihood to adopt a security feature, varying based on the observability of the feature, its adoption rate among one's friends, and the number of distinct social circles from which those feature-adopting friends originate.

Researchers have also examined how people naturally come up with surprising ways of managing their own privacy. For example, Jin et al [JGRY+22] found that, in the context of their own smart homes, people would unplug devices or cover up cameras so that they would not be monitored while at home. Having a better understanding of these kinds of naturalistic approaches could help in terms of understanding people's mental models, as well as why people choose these relatively simple techniques over more sophisticated ones.

One potentially promising avenue of research here is to develop stage models of adoption for privacy. For example, in health care, some people need more general awareness of their health problem while others are already active in preventative behavior and only need support in maintaining that behavior. By having a better model of behavior change, researchers and practitioners can develop better kinds of systems, user interfaces, and interventions to help foster adoption. Other relevant examples of the type of research that could help promote the adoption of PETs include the development and evaluation of different types of nudges. For instance, Story et al. have recently reported on their exploration of Protection Motivation Theory (PMT), Action Planning (AP), and Coping Planning (CP) bases for designing interventions designed to promote adoption of the Tor browser [SSC+22].

Studies to Understand What Influences Developers to Adopt PETs
Support more studies as to what technical, usability, social, and other factors lead developers to adopt (or not) different kinds of PETs in their products, as well as what kinds of interventions are useful. Developers are one of the greatest points of leverage for improving privacy, and so it's worth investigating more as to how to best educate, influence, and support this group.

For example, Witschey et al [WZWM+15] interviewed 42 developers and quantified why they did or did not adopt different security tools, using Diffusion of Innovations as a framework for organizing the results. Both Li et al [LLDH21] and Tahaei et al [TLV22] studied how developers talk about privacy, analyzing the /r/AndroidDev subreddit and StackOverflow respectively, generally finding that developers have low understanding of privacy and that app store requirements have a much greater influence than privacy laws.

Tahaei et al [TRLH+22] also examined the developer documentation for managing privacy on popular ad networks, and conducted user studies with developers to understand their perceptions. Tahaei et al found that documentation about privacy regulations was scattered across several pages, and used terms and language developers found hard to understand. There was also a mismatch in expectations, with ad networks putting the burden of complying with regulations on developers, while developers felt ad networks should be responsible for ensuring compliance with regulations.

Privacy Annotations to Improve the Privacy Ecosystem
Many programming languages support annotations, which are metadata about a program in the source code added by a developer. This metadata can be used by the compiler, the IDE, and other tools to manage workflow, documentation, and so on. Li et al first proposed *privacy annotations* in 2018 [LAH18], where developers specify the purpose of data collection, data retention, where the data will be sent, and so on. These annotations can be used by the IDE to offer hints about alternative APIs to use and to help with auditing of Android apps. In 2021, Li et al [LRAC+22] extended this work so that privacy annotations could be used to auto-generate useful user interfaces for informing end-users of data access by apps and offering user interfaces for configuring privacy settings.

More broadly, we believe privacy annotations can greatly improve the entire privacy ecosystem. Developers can use them at programming time, to help with many basic privacy-related development tasks, such as choosing appropriate APIs, adding privacy user interfaces, and filling out privacy nutrition labels. These same annotations can also be used by other members of a development team to audit the data collection behaviors of apps. These annotations can also be embedded into compiled apps, making it easier for app stores and other third parties to audit the behavior of apps and ensure that they are doing what they say they will do. One long term and ambitious idea would be to have annotations across all components of a system, including front-end as well as back-end cloud computing, enabling full end-to-end support for developers and for auditors.

App Manifests for Declaring and Enforcing App Behaviors
A common problem with almost all apps is *over access*. A sleep monitor app might use microphone data, but in reality it only needs loudness. A weather app might use GPS data, but only really needs city granularity. A smart TV app might access the raw viewing logs, but only needs a summary of most viewed channels and duration rounded to the nearest hour. Overaccess happens because there is a mismatch between what APIs there are for accessing data and what data the app actually needs.

Jin et al [JLHK+22] proposed three key ideas to address overaccess. First, all apps (smartphone, IoT, browser plugins, etc) are required to declare their data collection behaviors in a short and human-readable app manifest. Second, behaviors are specified using a small and pre-defined set of operators for gathering and transforming the data. Third, the manifest is enforced by a trusted hub which also runs the operators. Together, these form a trusted data flow that can help with minimizing the granularity of data shared with others.

An example flow for a sleep monitor might be “get access to microphone every minute, transform it to loudness, send to sleep.com”. While an auditor might not fully know how the data will be used, they can easily understand that only loudness granularity will be sent out. This approach also makes it possible to auto-generate a privacy nutrition label from a manifest, track who knows what about you, modify the data flow based on any kinds of rules (e.g. no face data should be sent out), and auto-generate consistent user interfaces across apps.

Jin et al investigated this approach with manifests in the context of IoT. One promising angle for research here is to extend the approach to other platforms, e.g. browser plugins, smartphones, cloud data, etc. Another angle is to investigate how to combine this approach with privacy annotations, making it so that there is a single unified approach that offers the best of both worlds. This combined approach could also make it possible to have a full end-to-end approach that works across distributed components, e.g. front-end and cloud backend.

Improving Auditing of Apps and Devices by Developers and Third Parties
Many PETs focus on preventing unauthorized access or disclosure of sensitive data. However, relatively little work has focused on making it easier to detect the same, which would help with ensuring that apps and devices comply with any stated behaviors and privacy regulations.

For example, an auditor internal to the company might want to do white box testing, tracing data from collection, transmission to cloud storage, processing, and reporting. Right now, this is a labor intensive process, with many disparate tools currently only usable by people with a high level of technical expertise.

Similarly, third parties such as researchers, consumer advocates, and journalists might also be interested in auditing apps and devices, to gather data and develop new techniques for checking for compliance at scale. However, there are a lack of standards and tools, making it hard even for experts to understand what data is being collected in the first place. Furthermore,

encryption techniques like TLS, while a best practice for privacy and security, unintentionally make it hard to conduct legitimate research.

PETs to Support Developers, Ecosystem Operators and Regulators and Assist with Compliance Analysis

Research has shown that software developers, both large and small, struggle to ensure that their products and services comply with applicable regulations. For instance, in their analysis of over a million Android Apps in the Google Play Store, researchers found that apps had an average of over 3 potential compliance issues [ZSS+19] (see also [ZWZ+17]). Research combining code analysis, privacy policy text analysis, as well as wizards designed to prompt developers to carefully reflect on their data practices can go a long way in helping developers ensure that their products and services are compliant (e.g., [ZSS+19,GFR+22]). We have only started to scratch the surface of what is possible in this space. Research has consistently shown that developers generally lack adequate expertise and support to properly disclose their data practices and ensure that their code is compliant. Similar technologies can also be used to help ecosystem operators (e.g., app stores, browser extension stores, IoT ecosystem operators) ensure that technologies developed by third parties and made available by the ecosystems to their user communities comply with relevant regulations. Those same technologies can also be used to conduct at-scale compliance analysis, something that regulators badly need if they are to keep pace with the emergence of ecosystems such as app stores that feature several million apps (e.g. [ZWZ+17,ZSS+19])..

PETs for the Enterprise to Study Data Flows Within and Across an Organization

Research on “PETs for the Enterprise,” which moves beyond web and mobile applications to consider how data is processed within and across an organization’s enterprise architecture. This focus includes how services built on Kubernetes, and Apache NiFi, Spark, and Hadoop for data processing, integrate with technical privacy controls for revocable authorization, confidentiality and disassociation. PETs for the enterprise can aid decision makers in planning privacy-at-scale by rationalizing how data flows within their service infrastructure, and verifying that against their corporate privacy policy on data protection. Service-oriented architectures that allow companies to decentralize data management are frequently used to scale performance to meet demand across government services. This places agencies at risk of over-collection and repurposing data [BSH15], particularly, when the volume of data represents hundreds of thousands of data subjects with thousands of data types.

Enhancements to formal methods for modeling data flow could allow enterprise planners to pinpoint where within a large system, and how to apply PETs as agencies increasingly look to share information across services. The ability to design systems to grant and revoke consent as individual privacy preferences and data use policies change over time, without degrading ongoing data processing and analytics is an open area of research [RBP+23].

Empirical Social Science Research on Impact of PETs

Support more social science research on the impact of PETs, and in particular empirical studies, including natural and field experiments, on the downstream implications and outcomes of the

deployment of PETs, including 1) the analysis of the economic trade-offs associated with the adoption of PETs in industry, government, or by consumers; 2) the investigation of how the potential burdens of adoption or potential losses from decreased granularity of data are allocated among different stakeholders, and how those burdens can be mitigated and minimized; 3) the study of how the usage of PETs affects or alters the allocation, across different stakeholders, of the value accrued from data analytics.

2. *Specific technical aspects or limitations of PETs:*

Generalizable techniques for tracing information flows

Restrictions on data sharing, including access control and de-identification techniques, rely on developer claims and assertions made at design time that data can only be accessed through specific endpoints, otherwise data leaks can arise that circumvent these restrictions. The efficacy of these restrictions is no greater than the reliability of these claims and assertions. In source code analysis, tools have been developed to trace information flow from collection points to sharing and use points to detect leaks and privacy policy violations [SWH+16]. Due to the wide variety of programming languages and software development frameworks or Application Programmer Interfaces (APIs), technical challenges exist in tracing information flow throughout filesystems, graphical user interfaces (GUIs) [WQH+18, QNG+19], and networked services [ZWS+20]. Solutions have largely been one-off examples tied to specific operating systems and APIs. Due to the availability of application binaries, advances in information flow tracing have largely accrued on mobile applications. Techniques that are generalizable beyond specific operating systems, programming languages and frameworks, and that can be passively integrated into developer toolchains, are needed to realize end-to-end information flow tracing to detect privacy leaks at scale.

Risk metrics for sensitive data

Confidentiality and disassociation techniques introduce cost to software development both in terms of software design, implementation and maintenance effort, as well as a perceived loss of data utility in certain applications. Techniques are needed to assess the privacy risk associated with sharing data to enable the application of PETs to the most sensitive data. Evidence exists that data subjects perceive privacy risk differently based on the ambiguity of how data is used [BBR+16], and their experience and social and physical distance to the privacy harm [BB18]. African Americans and Hispanics, for example, perceive greater risk from data sharing by law enforcement [BRA+19]. Bhatia and Breaux have designed and evaluated reliable privacy risk metrics that can be used to measure the perceived risk by data subjects to inform design decisions [BB18]. However, their work discovered two unaddressed technical challenges in privacy risk measurement: (1) risk changes as new datasets are composed from other datasets, yet the change in risk due to composition is unknown, including both the direction and magnitude of such risks; and (2) risk is subject to environmental effects, including the recency and severity of known privacy harms. Addressing these limitations of risk metrics would allow developers to combine PETs in ways that result in predictable, measurable impacts to data subject privacy as risks evolve.

Development of Human-AI Interaction Technologies for PETS

While new regulations such as CCPA/CPRA are requiring more detailed data practice disclosures and the availability of privacy choices (e.g., opt-ins and opt-outs) that were not previously required, these new regulations also have highly unrealistic expectations when it comes to what lay users can realistically be able to do (e.g., [LLS14,LLS+14]). No one reads privacy policies and no one has the time to analyze all the privacy choices and configure all the privacy settings now available to users. What is required is the development of intelligent interfaces that can help people take advantage of these regulatory advances and effectively regain control of their data (e.g. [LAS+16,RBW+19]). Recent advances in the development of Natural Language Processing and Machine Learning techniques to automatically analyze the text of privacy policies and help answer people's privacy questions offer the promise of enabling people to take advantage of more detailed data practice disclosures without requiring people to actually read the text of privacy policies (e.g., see CMU's Usable Privacy Policy Project [UPPP22,WSL+18]). Similarly machine learning techniques that can model people's privacy policies and help provide recommendations to users and reduce the burden required to configure an otherwise unmanageable number of privacy decisions could also go a long way in helping users (e.g, see work on personalized privacy assistants [PPA22,LAS+16, DDS+18]). Yet significantly more research is required to configure these technologies and ensure that users retain effective control over their privacy decisions while benefiting from the predictive power of these technologies.

Development of APIs and Infrastructure for PETs

In addition to developing intelligent interfaces to assist users with the management of their privacy, it is important to recognize that new technologies such as the Internet of Things also lack basic standards and APIs to communicate privacy policies and expose privacy settings to users. It is therefore critical to also promote the development of infrastructures, APIs and standards to support such communication. A particularly prominent effort in this regard is the development at Carnegie Mellon University of a Privacy Infrastructure for the Internet of Things (e.g. see [IoTPI22a, IoTPI22b,DDS+19,FYS21]). Additional research is required to promote the development and adoption of such infrastructure and promote the adoption of standards necessary for the adoption of such infrastructures.

3. *Specific sectors, applications, or types of analysis that would particularly benefit from the adoption of PETs:*

Emergency Response

Both immediate emergencies (e.g. earthquakes, flash floods, or terrorist attacks) as well as longer running ones (e.g. fires, hurricanes) could benefit greatly from more real-time data, which is possible from the smartphones that people carry as well as from the large number of cameras and other sensors in neighborhoods and cities. Some examples might include using location data to estimate how many people are still in a building for recovery, running private queries on people's smartphone photos to look for a specific individual's face (think Boston Marathon bombing), and using cell phone call data records to infer a major spike in activity in a localized area (possible accident or terrorist attack).

Urban Analytics
It is possible to use large-scale location traces to understand people's behaviors in cities (see for example [CSHS12]). Using geotagged social media, cell phone call data records, or smartphone tracking libraries embedded in popular apps, it is possible to acquire large-scale traces of where people go in a city. This kind of data can be helpful for urban planning, traffic analysis, understanding economic impacts of road closures, how people use parks, the walkability of different neighborhoods, and more. However, the sensitive nature of this kind of data leads to a large number of privacy concerns.

Public Health
Access to public health suffers from disparities in economic status that model disparities in trust in government data collection, processing and surveillance, including public health surveillance.

Telemedicine is one area where PETs could improve adoption. Telemedicine appointments provide unprecedented convenience to patients with inflexible work and childcare schedules, as well as unprecedented access to a patient's physical space outside the clinic or hospital, whether at home or at work. Privacy and security risks in telehealth are well known [HK14], but research into this vulnerable population and how PETs can improve access is lagging, and health apps and connected devices continue to risk privacy harm due to data breach [FTC21].

Contact tracing is another area that could benefit from PETs. Contact tracing seeks to understand who an infected individual may have been in contact with, and can help with early isolation measures. Despite its potential, and despite the fact that there are a number of tracking mechanisms (e.g. Bluetooth, GPS on smartphones, QR codes), electronic-based contact tracing measures have seen only limited adoption. Privacy is one of several barriers to widespread adoption of electronic contact tracing mechanisms [LCYB+21].

Distributed community sensing
Many people are already starting to deploy sensors in their homes, but these sensors tend to be isolated from one another. Having a programmable substrate that links many sensors in a neighborhood or across an entire city together would enable the creation of many new kinds of apps. For example, by linking together many smart speakers, one could create a noise map of the entire city, which could be used by public health researchers as well as individuals wanting to purchase a home. By linking together smart doorbells, one could create a smarter Amber Alert spotter, a lost dog spotter, a neighborhood parking space estimator, a pedestrian counter, a car traffic counter, a snow plow spotter, and more. However, privacy is a major barrier here, as such a system could be easily abused to stalk individuals or to infer if someone is home or not.

Covert Spying in Everyday Situations
Here, we mention two cases where PETs could help. The first is mitigating intimate partner violence. Freed et al [FPML+18] showed how malicious individuals use known spyware as well as rather innocuous technologies to track and monitor intimate partners. Better PETs could help victims detect misuse of these technologies or hide their activities.

The second is with temporary rentals, such as Airbnb and VRBO. A number of media articles have reported on how people have found undisclosed surveillance devices in these rentals, cameras in particular. Song et al [SHCH20] presented the results of a survey of 192 participants who have stayed in an Airbnb before, with 50% stating that they searched for devices, and 8 participants actually finding undisclosed devices. Assisting people in finding what sensing devices are nearby, where exactly those devices are, whether they are active or not, and how the data will be used would greatly help in this rental scenario, as well as many other smart home and smart building scenarios.

4. *Specific regulations or authorities that could be used, modified, or introduced to advance PETs:*

The role of the FTC, FDA and HHS in regulating health apps can be improved to cover services at the “margin” of medical devices and electronic health records. The FTC released a recent 2021 press release highlighting the privacy risks of health apps and connected devices [FTC21], whereas the FDA has moved to establish rules for Software as Medical Device. While many health apps do not meet the eligibility criteria to be classified as medical devices, nor are they included under the umbrella of the HIPAA Privacy Rule, they are collecting sensitive health data, which has become a gap in regulation and oversight. Regulators should work with industry to deploy lightweight PETs that can increase privacy assurances while supporting innovation in this area.

Regulations Requiring Standards and APIs to Make Privacy Usable
While new regulations such as CCPA/CPRA are requiring more detailed data practice disclosures and the availability of privacy choices (e.g., opt-ins and opt-outs) that were not previously required, these new regulations also have highly unrealistic expectations when it comes to what lay users can realistically be able to do (e.g., [LLS14,LLS+14]). No one reads privacy policies and no one has the time to analyze all the privacy choices and configure all the privacy settings now available to users. An example is the cookie management prompts that pop up on nearly every website one visits, forcing users to repeatedly enter their choices over and over again as they move from one website to the next. If users could specify their choices once and for all in their browsers and let their browser communicate these choices to the websites they visit, user burden could significantly be reduced (e.g., see Smullen et al. for a discussion of this and related issues [SYF+21]). Regulation that would require websites to support simple APIs that allow browsers to communicate these preferences would go a long way to reduce unnecessary burden and allow people to actually exercise choices that are otherwise only theoretical. The same is true for the communication of privacy policies and the presentation of privacy choices in the context of the Internet of Things (e.g., see Zhang et al. for a discussion of this issue in the context of video analytics scenarios [ZFB+21]). In short, APIs and standards to support these scenarios would go a long way towards helping people truly benefit from new regulations such as CCPA/CPRA (as well as similar regulations introduced in a number of other states).

Moving from Opt-Out to Opt-Ins
Beyond the above, US consumers would significantly benefit from regulations that move away from opt-out choices and make opt-in the default approach to exercising one's privacy choices [CS19]. Opt-out choices have been shown to be consistently gamed by industry, with these choices often buried deep in the text of privacy policies and/or websites and users being required to jump through an unrealistic number of hoops [HPW+19]. Opt-in choices provide a level playing field where consumers have to first provide consent before companies can engage into potentially invasive data practices and tend to much better reflect people's privacy preferences (e.g., [SYZ+21]).

5. *Specific laws that could be used, modified, or introduced to advance PETs:*
See comments provided under 4.

6. *Specific mechanisms, not covered above, that could be used, modified, or introduced to advance PETs:*

Data Sets for Evaluating Different PETs
Having access to large-scale data sets could help with evaluating the effectiveness of different PETs and the tradeoffs involved. For example, location data of many individuals could help in developing and comparing new kinds of algorithms. Note that large-scale location data sets do exist, but tend to be proprietary and/or expensive, making them out of reach of the vast majority of researchers. Given the sensitive nature of location data, there would also need to be other kinds of safeguards in place to prevent accidental release as well as misuse of the data.

See also comments already provided under 4 on regulations to require the adoption of standards and APIs to make privacy usable.

7. *Risks related to PETs adoption:*
No comment from our team.

8. *Existing best practices that are helpful for PETs adoption:*
No comment from our team.

9. *Existing barriers, not covered above, to PETs adoption:*

A High Burden on Developers for Compliance
In an analysis of a popular discussion board for Android developers, Li et al [LLDH21] found that developers often viewed privacy as a burden with little benefit for themselves. For example, developers felt that they had to learn new kinds of APIs, keep up to date with vague documentation and changes to the Android operating system, and comply with new policies on Google Play app store, all without any clear benefit to their app. Using a colloquialism, for developers, privacy is currently no carrots and all sticks. Li et al recommended looking for ways to reward developers for their effort, for instance prioritizing their app on search results or highlighting apps that were especially good with respect to privacy.

This same paper also found that discussions of privacy amongst Android developers rarely happened due to new privacy laws or regulations, and that developers rarely talked about privacy issues when proposing new apps or getting feedback from other devs. Instead, privacy was brought up primarily because of external events, such as changes to the app store or the operating system.

Encryption Makes it Hard for Researchers and Third-Party Auditors to Evaluate Privacy
Encryption techniques like TLS are a best practice for privacy and security, making it hard for attackers to eavesdrop on potentially sensitive communications. However, these same techniques make it hard for legitimate researchers and third parties to assess and to audit what data an app or device is sending out. Whether it is supported by an operating system, library, device, or app, having some kind of explicit setting that lets legitimate researchers and auditors examine data would be incredibly helpful in improving transparency and might also facilitate adoption of many new kinds of PETs.

10. *Other information that is relevant to the adoption of PETs:*

Division of Labor for Privacy Across the Entire Ecosystem
The division of labor as to who is responsible for what aspects of managing privacy is currently unclear. Hong [Hong17] argued that the burden of privacy is currently too heavy for end-users, and that instead we should consider how other parts of the ecosystem can help. Using an analogy with spam email, it used to be the case that end-users had to manually delete all spam themselves. Over time, however, email service providers started using machine learning to filter emails, network providers started deploying protocols for blocking bad emails, and law enforcement took down egregious offenders, with the end result being vastly fewer spam emails.

However, it is unclear how best to achieve a similar result for privacy, partly because there are vastly more stakeholders involved. Furthermore, there are many misconceptions and disagreements as to who should be responsible for what aspects of privacy. For example, as noted earlier in [TRLH+22], ad networks placed the responsibility of complying with privacy on developers, but developers felt that ad networks should be responsible instead.

In all cases, we should aim to avoid policies and regulations that place undue burden on end-users, as that has repeatedly been shown to be ineffective in practice.

References

[ABL20] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. "Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age." *Journal of Consumer Psychology* 30.4. (2020): 736-758. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3688497

[ATW16] Alessandro Acquisti, Curtis Taylor, and Liad Wagman. "The economics of privacy." *Journal of Economic Literature* 54.2. (2016): 442-92. https://www.aeaweb.org/articles?id=10.1257/jel.54.2.442

[BB18] Jaspreet Bhatia and Travis D. Breaux. “Empirical Measurement of Perceived Privacy Risk.” *ACM Transactions on Computer-Human Interaction.* 25, 6, Article 34. (December 2018) 47 pages. https://doi.org/10.1145/3267808

[BBR+16] Jaspreet Bhatia, Travis D. Breaux, Joel R. Reidenberg and Thomas B. Norton. "A Theory of Vagueness and Privacy Risk Perception." In Proceedings of the *IEEE 24th International Requirements Engineering Conference (RE)*. (2016) pp. 26-35, doi: 10.1109/RE.2016.20. https://ieeexplore.ieee.org/abstract/document/7765508

[BRA+19] Brooke Auxier et al. “Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information.” PEW Research Center. (November 15, 2019) https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[BSH15] Travis D. Breaux, Daniel Smullen, and Hanan Hibshi. "Detecting repurposing and over-collection in multi-party privacy requirements specifications." *2015 IEEE 23rd International Requirements Engineering Conference (RE)*. (2015) pp. 166-175, doi: 10.1109/RE.2015.7320419. https://ieeexplore.ieee.org/document/7320419

[CS19] L Cranor and N Sadeh, “Congress, Make Privacy the Rule - Not the Exception.” Op-Ed in the Hill. (February 1, 2019) https://thehill.com/opinion/cybersecurity/428121-congress-make-privacy-the-rule-not-the-exception/

[CSHS12] J Cranshaw, R Schwartz, JI Hong, N Sadeh. “The livehoods project: Utilizing social media to understand the dynamics of a city.” ICWSM 2012. (2012) https://ojs.aaai.org/index.php/ICWSM/article/view/14278

[DDS+18] A Das, M Degeling, D Smullen, and N Sadeh. "Personalized Privacy Assistants for the Internet of Things," 2018 IEEE Pervasive Computing: Special Issue - Securing the IoT. (April 2018) https://www.privacyassistant.org/media/publications/IEEE_magazine_2018.pdf

[DKDH14] S Das, ADI Kramer, LA Dabbish, JI Hong. “Increasing security sensitivity with social proof: A large-scale experimental confirmation.” CCS 2014. (2014) https://dl.acm.org/doi/10.1145/2660267.2660271

[DKDH14] S Das, THJ Kim, LA Dabbish, JI Hong. “The Effect of Social Influence on Security Sensitivity. SOUPS 2014. (2014) https://dl.acm.org/doi/10.5555/3235838.3235851

[DKDH15] S Das, ADI Kramer, LA Dabbish, JI Hong. “The role of social influence in security feature adoption.” CSCW 2015. (2015) https://dl.acm.org/doi/10.1145/2675133.2675225

[FPML+18] D Freed, J Palmer, D Minchala, K Levy, T Ristenpart, N Dell. ““A Stalker’s Paradise”: How Intimate Partner Abusers Exploit Technology.” CHI 2018. (2018) https://dl.acm.org/doi/10.1145/3173574.3174241

[FTC21] FTC Warns Health Apps and Connected Device Companies to Comply With Health Breach Notification Rule. (September 2021) https://www.ftc.gov/news-events/news/press-releases/2021/09/ftc-warns-health-apps-connected-device-companies-comply-health-breach-notification-rule

[FYS21] Yuanyuan Feng, Yaxing Yao, Norman Sadeh, "A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. (May 2021) https://www.privacyassistant.org/media/publications/chi21-design-space.pdf

[GFR+22] Jack Gardner, Yuanyuan Feng, Kayla Reiman, Zhi Lin, Akshath Jain and Norman Sadeh, "Helping Mobile Application Developers Create Accurate Privacy Labels." IWPE'22. (May 2022) https://usableprivacy.org/static/files/Gardner_IWPE_2022.pdf

[HK14] Timothy M. Hale, Joseph C. Kvedar. “Privacy and Security Concerns in Telehealth.” AMA Journal of Ethics, 16(12): 981-985. (2014) doi://10.1001/virtualmentor.2014.16.12.jdsc1-1412 https://journalofethics.ama-assn.org/article/privacy-and-security-concerns-telehealth/2014-12

[Hong17] “The privacy landscape of pervasive computing.” IEEE Pervasive Computing 16 (3), 40-48. (2017) https://ieeexplore.ieee.org/document/7994573

[HPW+19] H Habib, S Pearman, J Wang, Y Zou, A Acquisti, LF Cranor, N Sadeh, "" It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. (April 2020) https://usableprivacy.org/static/files/habib_chi_2020.pdf

[IoTPI22a] Carnegie Mellon University’s “Privacy Infrastructure for the Internet of Things” portal. (2022) https://www.iotprivacy.io/login

[IoTPI22b] “Personalized Privacy Assistants for the Internet of Things” project website. (2022) https://privacyassistant.org/iot/

[JGRY+22] H Jin, B Guo, R Roychoudhury, Y Yao, S Kumar, Y Agarwal, JI Hong. “Exploring the Needs of Users for Supporting Privacy-Protective Behaviors in Smart Homes.” CHI 2022. (2022) https://dl.acm.org/doi/abs/10.1145/3491102.3517602

[JLHK+22] H Jin, G Liu, D Hwang, S Kumar, Y Agarwal, J Hong. “Peekaboo: A Hub-Based Approach to Enable Transparency in Data Processing within Smart Homes.” IEEE Security and Privacy 2022. (2022) https://www.computer.org/csdl/proceedings-article/sp/2022/131600b571/1CIO8pmx6jm

[KWDH+22] I Krsek, K Wenzel, S Das, JI Hong, L Dabbish. “To Self-Persuade or be Persuaded: Examining Interventions for Users’ Privacy Setting Selection.” CHI 2022. (2022) https://sauvikdas.com/uploads/paper/pdf/38/file.pdf

[LAH18] T Li, Y Agarwal, JI Hong. “Coconut: An IDE plugin for developing privacy-friendly apps.” Ubicomp 2018. (2018) https://dl.acm.org/doi/10.1145/3287056

[LAS+16] B Liu, MS Andersen, F Schaub, H Almuhimedi, S Zhang, N Sadeh, A Acquisti, and Y Agarwal, "Follow My Recommendations: A Personalized Assistant for Mobile App Permissions", Symposium on Usable Privacy and Security (SOUPS '16). (June 2016) https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-liu.pdf

[LCYB+21] T Li, C Cobb, JJ Yang, S Baviskar, Y Agarwal, B Li, L Bauer, JI Hong. “What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention.” Pervasive and Mobile Computing. (2021) https://www.sciencedirect.com/science/article/pii/S1574119221000833

[LLDH21] T Li, E Louie, L Dabbish, JI Hong. “How Developers Talk About Personal Data and What It Means for User Privacy: A Case Study of a Developer Forum on Reddit.” CSCW 2021. (2021) https://dl.acm.org/doi/10.1145/3432919

[LLS14] B Liu, J Lin, N Sadeh, "Reconciling Mobile App Privacy and Usability on Smartphones: Could User Privacy Profiles Help?" 23rd International Conference on the World Wide Web (WWW 2014). (July 2014) https://www.cs.cmu.edu/~bliu1/Bin_Liu_WWW2014_Reconciling.pdf

[LLS+14] J Lin, B Liu, N Sadeh, and JI Hong, "Modeling Users’ Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings." 2014 ACM Symposium on Usable Security and Privacy (SOUPS 2014). (July 2014) https://www.usenix.org/system/files/conference/soups2014/soups14-paper-lin.pdf

[LNAH21] T Li, EB Neundorfer, Y Agarwal, JI Hong. “Honeysuckle: Annotation-guided code generation of in-app privacy notices.” Ubicomp 2021. (2021) https://dl.acm.org/doi/10.1145/3478097

[LRAC+22] T Li, K Reiman, Y Agarwal, LF Cranor, JI Hong. “Understanding Challenges for Developers to Create Accurate Privacy Nutrition Labels.” CHI 2022. (2022) https://dl.acm.org/doi/10.1145/3491102.3502012

[QNG+19] Xue Qin, Robert Neuhaus, Diego Gonzales, Xiaoyin Wang, Travis Breaux, and Jianwei Niu. “Taming web views in the detection of Android privacy leaks.” In Proceedings of the 6th Annual Symposium on Hot Topics in the Science of Security (HotSoS '19). Association for Computing Machinery, New York, NY, USA, Article 23, 1–2. (2019) https://doi.org/10.1145/3314058.3317732

[RBW+19] Abhilasha Ravichander, Alan W Black, Shomir Wilson, Thomas Norton and Norman Sadeh, "Question Answering for Privacy Policies: Combining Computational and Legal Perspectives." 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019), Hong Kong, China. (November 2019) https://usableprivacy.org/static/files/ravichander_emnlp_2019.pdf

[RBP+23] Marco Robol, Travis D. Breaux, Elda Paja, Paolo Giorgini. “Consent verification monitoring.” *ACM Transactions on Software Engineering Methodology (TOSEM)*. (January 2023) https://arxiv.org/pdf/2206.06406

[SHCH20] Y Song, Y Huang, Z Cai, JI Hong. “I'm All Eyes and Ears: Exploring Effective Locators for Privacy Awareness in IoT Scenarios.” CHI 2020. (2020) https://dl.acm.org/doi/abs/10.1145/3313831.3376585

[SSC+22] Peter Story, Daniel Smullen, Rex Chen, Yaxing Yao, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, Florian Schaub. ”Increasing Adoption of Tor Browser Using Informational and Planning Nudges.” Proceedings on Privacy Enhancing Technologies 2022 no. 2, pp. 1–32. (2022) https://peterstory.me/publications/story_popets_2022.pdf

[SWH+16] Rocky Slavin, Xiaoyin Wang, Mitra Bokaei Hosseini, James Hester, Ram Krishnan, Jaspreet Bhatia, Travis D. Breaux, and Jianwei Niu. “Toward a framework for detecting privacy policy violations in android application code.” In Proceedings of the *38th International Conference on Software Engineering (ICSE '16).* Association for Computing Machinery, New York, NY, USA, 25–36. (2016) https://doi.org/10.1145/2884781.2884855

[SYF+21] D Smullen, Y Yao, Y Feng, N Sadeh, A Edelstein, and R Weiss. ”Managing Potentially Intrusive Practices in the Browser: A User-Centered Perspective.” Proceedings on Privacy Enhancing Technologies 2021 (4), 500-527. (October 2021) https://sciendo.com/it/article/10.2478/popets-2021-0082

[TLV22] M Tahaei, T Li, K Vaniea. “Understanding Privacy-Related Advice on Stack Overflow.” PETS 2022. (2022) https://petsymposium.org/2022/files/papers/issue2/popets-2022-0038.pdf

[TRLH+22] M Tahaei, KM Ramokapane, T Li, JI Hong, A Rashid. “Charting App Developers’ Journey Through Privacy Regulation Features in Ad Networks.” PETS 2022. (2022) https://mohammad.tahaei.com/publication/pets-2022-charting-regulations-adnetworks/pets-2022-charting-regulations-adnetworks.pdf

[UPPP22] Usable Privacy Policy Project website.(2022) https://usableprivacy.org/

[WQH+18] Xiaoyin Wang, Xue Qin, Mitra Bokaei Hosseini, Rocky Slavin, Travis D. Breaux, and Jianwei Niu. “GUILeak: tracing privacy policy claims on user input data for Android applications.” In Proceedings of the 4*0th International Conference on Software Engineering (ICSE '18)*. Association for Computing Machinery, New York, NY, USA, 37–47. (2018) https://doi.org/10.1145/3180155.3180196

[WSL+18] Shomir Wilson, Florian Schaub, Frederick Liu, Kanthashree Mysore Sathyendra, Daniel Smullen, Sebastian Zimmeck, Rohan Ramanath, Peter Story, Fei Liu, Norman Sadeh, Noah A. Smith, "Analyzing Privacy Policies at Scale: From Crowdsourcing to Automated Annotations." ACM Transactions on the Web, 13, 1. (December 2018) https://usableprivacy.org/static/files/swilson_tweb_2018.pdf

[WZWM+15] J Witschey, O Zielinska, A Welk, E Murphy-Hill, C Mayhorn, T Zimmerman. “Quantifying developers' adoption of security tools.” 2015 Joint Meeting on Foundations of Software Engineering. (2015) https://dl.acm.org/doi/10.1145/2786805.2786816

[ZFB+21] Shikun Zhang, Yuanyuan Feng, Lujo Bauer, Lorrie Faith Cranor, Anupam Das, and Norman Sadeh, ""Did you know this camera tracks your mood?": Understanding Privacy Expectations and Preferences in the Age of Video Analytics." Proceedings on Privacy Enhancing Technologies, 2021, 1. (April 2021) https://www.petsymposium.org/2021/files/papers/issue2/popets-2021-0028.pdf

[ZFD+20] Shikun Zhang, Yuanyuan Feng, Anupam Das, Lujo Bauer, Lorrie Faith Cranor, Norman Sadeh, "Understanding People’s Privacy Attitudes Towards Video Analytics Technologies." School of Computer Science Technical Report CMU-ISR-20-114 Carnegie Mellon University. (December 2020) http://reports-archive.adm.cs.cmu.edu/anon/isr2020/CMU-ISR-20-114.pdf

[ZSS+19] Sebastian Zimmeck, Peter Story, Daniel Smullen, Abhilasha Ravichander, Ziqi Wang, Joel Reidenberg, N. Cameron Russell, and Norman Sadeh, "MAPS: Scaling Privacy Compliance Analysis to a Million Apps." Privacy Enhancing Technologies Symposium (PETS 2019), 3. (July 2019) https://usableprivacy.org/static/files/popets-2019-maps.pdf

[ZWS+20] Xueling Zhang, Xiaoyin Wang, Rocky Slavin, Travis Breaux, and Jianwei Niu. “How does misconfiguration of analytic services compromise mobile privacy?” In Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering (ICSE '20). Association for Computing Machinery, New York, NY, USA, 1572–1583. (2020) https://doi.org/10.1145/3377811.3380401

[ZWZ+17] S Zimmeck, Z Wang, L Zou, R Iyengar, B Liu, F Schaub, S Wilson, N Sadeh, SM Bellovin, JR Reidenberg, "Automated Analysis of Privacy Requirements for Mobile Apps." NDSS'17: Network and Distributed System Security Symposium. (February 2017) https://usableprivacy.org/files/news/NDSS17.pdf