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

Work, Paul R.

Response to the Office of Science and Technology Policy (OSTP) Request for Information (RFI) on Advancing Privacy-Enhancing Technologies (PETs)

From: Paul R. Work

Organization: Member of the Public

Date: 7 July 2022

Topics Addressed: 3. *Specific sectors, applications, or types of analysis that would particularly benefit from the adoption of PETs;* 8. *Existing best practices that are helpful for PETs adoption;* 9. *Existing barriers, not covered above;* 10. *Other information that is relevant to the adoption of PETs*

Background: I am a retired person who spent almost 50 years working in the field of computer systems applied to everything from life and healthcare insurance, to toll road technologies, to logistics, to the defense sector, and other national security spaces. I spend the last 15+ years working closely with academia and small businesses on research and development efforts for organizations like the Defense Advanced Research Projects Agency (DARPA), Intelligence Advanced Research Projects Activity (IARPA), Office of Naval Research (ONR), Air Forces Research Laboratory (AFRL), and the Semiconductor Research Corporation (SRC), just to name a few. That said, while I have been a systems engineer and certified systems architect for the latter half of those 50 years, espousing holistic, multi-discipline, systems thinking and development, focused on "Chips to Data Centers, and all the Algorithms There On", and while my credentials in PETs are minimal, I do have a general understanding of some of and insight into the techniques and technologies of PETs, such as Multi-Party Computation (MPC) and Differential Privacy (DP).

I want to thank the leaders and members of the Fast Track Action Committee (FTAC) on Advancing Privacy-Preserving Data Sharing and Analytics (APPDSA) for their Subgroup Roundtables held virtually back in early June. There were a good number of contributors on the Zoom calls and a great discussion with researchers and organizations involved in the development and fielding of PETs. This was very helpful in catching up with a sense of the state of the art and the state of the research in PETs and the thoughts and concerns that they shared with the implementation and fielding of PETS.

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

As it may come out in the other responses below, my current interest in the development of PETs is focused on applications to folks like me, senior citizens on fixed incomes, and of an age where the majority of my peers, while users today of technologies like iPhones, iPads, Ring doorbells, Roku streaming devices, and Amazon Prime, were adults during the development of these tools. We have seen the maturation of these technologies, while our younger generations literally grew up with them. While some of us may have been involved in the "computer business", and a group of us are skilled practitioners in the arts and science of cybersecurity and data privacy, I would venture that the majority are simply users of the technology without a deep understanding of the nuances of it. Back in 2019, the Consumer Technology Association (CTA) conducted a study on "Active Aging: Consumer Perceptions and

Attitudes"[1] where the found that while "virtually all seniors want to live independently and stay home as they age," seniors view "cost and concern over learning new technologies could be barriers to adoption." The study also found that "Over two-thirds of both seniors and caretakers have concerns about privacy with health monitoring devices for seniors. With the plethora of news regarding cybersecurity breaches and how they effect nearly everyone, it's easy to understand this fear."

We depend on the good will and best efforts of researchers and practitioners who develop these sorts of technologies and tools. An example of this is the encouragement from the cybersecurity community to leverage the use of SSL (Secure Sockets Layer) encryption, and its more modern and secure replacement, TLS (Transport Layer Security) encryption to protect data sent over the Internet. For us that means simply use links which have the prefix: "https://" and depend upon our browsers to check for this and signal to use when we are not using it, to ensure our safety and security on the various sites we may go to.

As for the privacy of our data, the current approach seems to be the use of "pop-ups" which ask us (for every site!) what sort of "tracking" we want to allow (if they are nice enough to detail any of it.) Why can't we have browser and app technologies that are (1) required to honor already defined "answers" to these tracking questions and (2) if we do want to "share" our data with them, do so using PETs to ensure that our data is still protected beyond simply sharing in the "raw", even if it is encrypted during transmission from our devices? It appears that most websites and apps make use of upward to dozens of ad tracking tools, understandably to ensure they have a way to pay for their "free" services. If Amazon or CVS are already getting our monies for products or services we procure via their sites and apps, why we do we have to allow third parties to have access to our data (including our IP address) without our permission.

One of the most egregious offenders are the companies who are willing (by asking us) to put us on their emailing list, then send us an email with tantalizing headlines along with a sentence or two to entice us to follow a link and then when we do follow the link, we are "given to opportunity" to pay to read the rest, behind a "pay wall." All of this because we are using their email service for free. So, we have freely given us our email address, which they will sell to others, who will then gather more of our data and preferences when we hit their websites. We need to have basic infrastructure in the technology stack to ensure that only the details we truly wish to share are done so, and without having to make the decision again and again for every site we "touch." Can we have operating system hosted tools to configure our level of data sharing and does not simply block us from getting to a site or service or product (like Chrome will do if we don't use a https:// address and if it does the site actually has a valid certificate)?

My *recommendation* is the development of an Internet style (e.g., Internet Engineering Task Force (IETF)) standard that is implemented in every operating system (covering all devices from servers to mobile and Internet of Things (IoT) devices) and is contained in both human and machine-readable format, along with contextual semantics to ensure that there are sufficient opportunities to share our data, as well as, to secure it from sharing. This should also be extended to other tech stacks (e.g.,

[1] CTA Study: "Active Aging: Consumer Perceptions and Attitudes", https://www.cta.tech/Resources/i3-Magazine/i3-Issues/2019/May-June/Active-Aging-Perceptions-and-Attitudes, 5 June 2019, last accessed: 2 July 2022.

containerization, and Application Programming Interface (API) management) to ensure completeness of implementation and scalability.)

This sort of capability, to configure my "default" privacy settings and have all sites and apps honor them, through the use of standard interfaces, also applies, and maybe even more so, to artificial intelligence (AI) / machine learning (ML), where our private data may be consumed and used without our knowledge. There needs to a basic requirement for all AI/ML applications to have an "explainability" (think: XAI[2]) feature such that the privacy settings we configure can be understood by us (humans) and in a way that we can configure our data use to the benefit of the AI/ML tools. This will be essential for healthcare and financial applications, to ensure that traceability of our data is controlled by us. Existing laws and rules, such as the Health Insurance Portability and Accountability Act (HIPAA) still allow for a great deal of our personal data to be seen and shared so that we can receive medical services and products, yet currently with many more people than is really necessary able to see it. If we had mandatory PET use for exchange of this data, we would have higher confidence to share it, e.g., for medical or pharmacology purposes, to achieve the gains promised by the users of our data. Say if we have the privacy standard configuration and context data stored on mobile devices, we could share it with our medical providers, researchers, and application developers, with a high confidence of correct and proper use of our data, to the benefit of all. And to work with companies like Apple, Firefox, and DuckDuckGo to pilot these PET enabled capabilities, since they already have very well-known privacy advocation activities within their organizations.

Another *recommendation* is to have end users to be able to leverage PETs to be able to conduct business when talking with human customer service representatives (reps) (and which are being replaced with automated "intelligent" chat avatars). We often have to respond with either publicly available data which the customer reps have access to, or with "security" questions which again the human reps have access, so not actually very secure or private. The use of Authentication tools (such as Authy) provides a two-factor sort of service, yet this is only used in more automated interfaces (e.g., login) and isn't used in the verbal exchanges with customer reps. The intent here being to go beyond "authentication" in the exchange processes, so that information exchanges requiring Personally Identification Information (PII) (to include credit card information, driver's license number and IP addresses), so that any sort of "phishing" like approach or "man-in-the-middle" action can be reduced. Developing a solution along these lines should also aide in government -to -government exchanges (whether they be Federal, State or local) to minimize the amount of private data (even if it available in public databases) from being shared in the "raw" and to support the exchanges since not only is the PII protected, the actual data needed (e.g., law enforcement data) can support an inquiry being made without sharing it in all of its details.

Response to Topic 8: *Existing best practices that are helpful for PETs adoption.*

Recommendation: Work on ways to support PET implementation by leveraging/mirroring Data Labeling efforts done in the Machine Learning (ML) community for supervised and reinforcement learning, in order to collect and use these for context capture and development necessary for PETs to work well.

[2] See Samek, Wojciech, et al., eds. Explainable AI: interpreting, explaining and visualizing deep learning. Vol. 11700. Springer Nature, 2019.

Recommendation: Leverage organizations like the American Association of Retired Persons (AARP) and their safeguarding efforts like AARP BankSafe and the AARP Fraud Watch Network Helpline[3] programs. One such activity could be to approach AARP to propose to do a joint pilot program with them to roll out a customer service interactions approach with PETs in play to reduce the exposure of private data during interactions over the phone or in person. This pilot could also address the lack of basic understanding of PET capabilities, so training and other learning materials for all ages, similar to those developed for the cybersecurity "hygiene" materials, could be developed and disseminated, as well as work with one or two technology developers to prototype products that can then be tested.

Response to Topic 9: *Existing barriers, not covered above.*

Recommendation: Incorporate PET capabilities into Computer Science curriculum, at all age levels, so that we go from awareness to practice as a "normal" vs "exception" approach to incorporation of data privacy technologies. Understanding the challenge of curriculum changes from elementary to post-secondary, this is no small task, yet data privacy is just as important as cybersecurity and will need the development of "champions" to achieve this result.

Recommendation: Address the lack of tool suites for review of code and coding practices that enforce a "data privacy first" approach to developing both software, firmware, and hardware. This can be done by first identifying what commercial tools are being developed (and maybe already being used) as well as to continue the funding of both academic research as well as public-private partnerships to develop the tools and work with NIST on the development of the standard processes, similar to what is currently done in the cybersecurity space. AND then pilot these both in government and in private sector efforts to "wring them out." Leveraging an "open source" approach may be a good way to get greater adoption faster. Also incorporating this effort with the recommendation above, will certainly achieve multipliers (no pun intended) in results.

Recommendation: There is need for hardware accelerators, not just for cryptological purposes, to support the computational resource needs of PETs at both the "edge" (e.g., IOT or mobile devices) and the datacenters (cloud servers) and possibly even within the communications devices (e.g., routers & switches). These can be in general purpose computing devices, like CPUs as well as purpose build hardware accelerators, similar to Graphics Processing Units (GPUs) (from Nvidia), Tensor Processing Units (TPUs) (from Google), Adaptive Compute Acceleration Platform (ACAPs) (from Xilinx), or simply purpose built chiplets (from Intel and AMD). Again, a pilot program with one or more of the hardware vendors (e.g., at both the chip and the application level) would not just address development of prototypes, it would also cover the development of business models necessary to make these viable, as well as provide something with which to feed to the technology social media (e.g., Wired, TechCrunch, Ars Technica, and others) to raise both interest and awareness.

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

In the field of systems and software (also applicable to firmware and hardware) is the use of architectural design languages (e.g., SAE International standard AS5506D Architecture Analysis & Design

[3] AARP .org Bulletin, June 2022, p33 "AARP Expands Safeguards" and p34 "Hear these words? HANG UP!", https://press.aarp.org/2022-6-3-AARP-Bulletin-Exclusive-AARP-National-Geographic-Second-Half-of-Life-Research, last accessed: 2 July 2022.

Language (AADL)[4] are used to address the holistic treatment of systems development. *Recommendation*: Perhaps development of a Privacy Protection Analysis and Design Language (PPADL, pronounced "paddle" :) would be useful for both researchers and practitioners, for a holistic approach to privacy preservation. This would address APIs at both the hardware as well as the software.

Recommendation: Application of PETs to Digital Twins (DTs) and other Intellectual Property exchanges to allow for both additional levels of design as well as manufacturing while ensuring privacy of data in the models or DTs. This would allow the models to be shared with multiple suppliers in such a fashion as to ensure the information needed by the individual organizations was available to complete their tasks (e.g., complete additional design work or manufacture a particular part) while keeping other information protected. This might leverage the work done in the area of "Cognitive Digital Twins"[5] in order to capture the contextual model information to be incorporated into the PET exchanges.

Lastly, *Recommendation:* Develop of a standard lexicon of terms to include and help establish the concept of "Data Ownership" which can be monetized by both government / private industry business purposes as well as "consumers" of their business services and products. This last bit being difficult given the current state of development and implementation of Web 2.0 & 3.0 efforts, yet still in need for both communities to deal with fairness to both the individual and the corporations. Have these terms defined and agreed upon, then allow us to develop the technologies suggested in the prior responses.

[4] https://www.sae.org/standards/content/as5506d and https://en.wikipedia.org/wiki/Architecture_Analysis_%26_Design_Language; both last accessed 2 July 2022

[5] Ali, Patel, Breslin, Harik and Sheth, "Cognitive Digital Twins for Smart Manufacturing", IEEE Intelligent Systems, Volume: 36, Issue: 2, March-April 1 2021, DOI: 10.1109/MIS.2021.3062437.