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

Meta

DISCLAIMER: Please note that the RFI public responses received and posted do not represent the views or opinions of the U.S. Government. We bear no responsibility for the accuracy, legality, or content of the responses and external links included in this document.

July 8, 2022

Via Electronic Mail

Office of Science and Technology Policy
1650 Pennsylvania Avenue
Washington, DC 20504

RE: Request for Information on Advancing Privacy-Enhancing Technologies

Thank you for the opportunity to respond to this Request for Information. Meta is investing heavily in privacy-enhancing technologies ("PETs"), and we enthusiastically support a national strategy to advance the adoption of these technologies.

We have been investing in PETs for several years across our organization, and based on that experience we share the Government's optimism that PETs—if widely adopted—will enhance data's benefits while preserving people's privacy. Meta's Data for Good program, for example, uses differential privacy to assist global responses to COVID-19 by providing public health experts and researchers with Movement Range Maps.[1] These maps visualize aggregated, county-level changes in population mobility rates, which, for example, allow our partners to assess how lockdowns impact income inequality and to steer public health messaging. Differential privacy has also been critical to the work of Meta's Facebook Open Research & Transparency ("FORT") team, which supports independent research into social media's impact on society.[2]

PETs also hold potential to advance civil rights and increase equity for marginalized groups, as the Request points out. At Meta, we are exploring secure multiparty

[1] *See* A. Hergağdelen et al., *Protecting Privacy in Facebook Mobility Data During the COVID-19 Response*, Meta Research (June 2, 2020), https://research.facebook.com/blog/2020/06/protecting-privacy-in-facebook-mobility-data-during-the-covid-19-response/. For more information on PETs in similar initiatives, *see* E. Finkle, *How Facebook Protects Privacy in Our COVID-19 Initiatives*, Meta Newsroom (Apr. 29, 2021), https://about.fb.com/news/2021/04/how-facebook-protects-privacy-in-our-covid-19-initiatives.

[2] *See* FORT, https://fort.fb.com/. As one example, the URL Shares dataset includes differentially private counts of how many people viewed, clicked, liked, commented, shared, or reacted to URLs on Facebook. G. King & N. Persily, *Unprecedented Facebook URLs Dataset Now Available for Academic Research Through Social Science One*, Social Science One (Feb. 13, 2020), https://socialscience.one/blog/unprecedented-facebook-urls-dataset-now-available-research-through-social-science-one; *see also* Meta, FORT Researcher Datasets, https://fort.fb.com/researcher-datasets.

computation (“SMPC”) as a tool to help measure whether people’s experiences with our products differ across racial and ethnic groups—without revealing individuals’ sensitive information to Meta.[3]

Meta’s investment in PETs also highlights their promise for driving the next generation of personalized ad experiences. Businesses—especially small ones—depend on personalized advertising to reach relevant audiences on limited budgets. Personalization does not have to be at odds with privacy, and PETs will help reinforce why. For instance, Meta is exploring SMPC to help advertisers understand how ads perform without disclosing to advertisers or Meta who sees an ad and who makes a purchase.[4] We have open sourced this framework so that anyone in the industry can develop similar products,[5] and we are working with industry partners like Mozilla to advance the use of PETs in personalization.[6]

Despite these promising applications of PETs, industry adoption has been limited. We agree with the Request that a host of challenges has contributed to this outcome. There are technical limitations and tradeoffs to many PETs, which must be both accounted for in policy and addressed with greater funding for research and development. There is also a lack of awareness around PETs and their benefits. Finally, entities using PETs face great regulatory uncertainty around how privacy concepts like deidentification are scoped, and what that means for how PETs and other protections can help meet obligations.

We elaborate on these and other challenges—and their potential solutions—in our responses to the Request’s topics below. The fulcrum of our responses is that PETs must be an essential part of *all* conversations around data and privacy. The work of the Fast Track Action Committee will be invaluable for achieving this, and we hope that our responses will serve as a springboard for future government collaboration with Meta.

1. Specific research opportunities to advance PETs

As our response to Topic 2 details, PETs are rapidly evolving, and some PETs are relatively new and less mature. Although newer PETs hold great promise, they face technical

[3] *See* R. L. Austin, *Race Data Measurement and Meta’s Commitment to Fair and Inclusive Products*, Meta Newsroom (Nov. 18, 2021), https://about.fb.com/news/2021/11/inclusive-products-through-race-data-measurement/.
[4] *See* Meta, *What Are Privacy-Enhancing Technologies and How WIll They Apply to Ads?*, Meta Newsroom (Aug. 11, 2021), https://about.fb.com/news/2021/08/privacy-enhancing-technologies-and-ads/ (also noting how on-device learning and differential privacy could enhance ads privacy).
[5] *See* G. Mudd, *Privacy-Enhancing Technologies and Building for the Future*, Meta Business News (Aug. 11, 2021), https://www.facebook.com/business/news/building-for-the-future.
[6] *See* M. Thomson, *Privacy Preserving Attribution for Advertising*, The Mozilla Blog (Feb. 8, 2022), https://blog.mozilla.org/en/mozilla/privacy-preserving-attribution-for-advertising/.

limitations and challenges. Additional federal funding for technical research programs to address these limitations would be immensely beneficial. Research could also help illuminate ways of strengthening privacy in artificial intelligence, such as assessing the vulnerability of models to data extraction attacks, which Meta is actively exploring.[7] This Administration has already recognized the importance of research, and we applaud efforts like its collaboration with the United Kingdom on prize challenges to advance PETs tackling financial crime.[8] We encourage further research opportunities and funding.

Regardless of the specific contours of research opportunities, they should strive to promote collaboration between a variety of stakeholders to the greatest extent possible. PETs will be advanced most effectively and responsibly if the Government, industry, academia, and civil society partner to make progress toward shared goals.

2. Specific technical aspects or limitations of PETs

Some PETs, such as at-rest encryption techniques, have existed in some form for several decades. Others are much newer, and this relative lack of maturity carries important implications for what PETs can and cannot achieve.[9] There might be limitations on the particular computational problems to which a less mature PET could be applied. Partial- and somewhat-homomorphic encryption, for instance, are limited to additive and/or multiplicative functions. Less mature PETs might also be more computationally intensive, potentially limiting their ability to be used at scale. Fully-homomorphic encryption and SMPC, for example, both demand significant computing resources.

Another, related limitation is the costs that applying PETs often incurs. As referenced earlier, some PETs demand particularly high computing resources. Applying PETs might

[7] Data extraction attacks are attempts by adversaries to discern whether a datapoint was in an AI model's training data. Meta recently simulated adversarial attacks to assess the privacy of its SEER computer vision model. *See* P. Goyal et al., *SEER 10B: Better, Fairer, Computer Vision Through Self-Supervised Learning on Diverse Datasets*, Meta AI Research (Feb. 28, 2022), https://ai.facebook.com/blog/seer-10b-better-fairer-computer-vision-through-self-supervised-learning-training-on-diverse-datasets/.

[8] *See* The White House, *U.S. and U.K. Governments Collaborate on Prize Challenges To Accelerate Development and Adoption of Privacy-Enhancing Technologies*, Briefing Room (June 13, 2022), https://www.whitehouse.gov/ostp/news-updates/2022/06/13/u-s-and-uk-governments-collaborate-on-prize-challenges-to-accelerate-development-and-adoption-of-privacy-enhancing-technologies/.

[9] For two explanations of some technical limitations of PETs, *see* The Royal Society, Protecting Privacy in Practice: The Current Use, Development, and Limits of Privacy Enhancing Technologies in Data Analysis (2019), *available at* https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-report.pdf; Centre for Data Ethics and Innovation, PETs Adoption Guide (2021), https://cdeiuk.github.io/pets-adoption-guide/.

also require reconfiguring how data is stored, itself a very costly endeavor. And there are the human resources required to implement technically complex PETs. These costs vary by situation, but they carry implications for what is practicable. As a result, some applications of PETs might be possible for only the largest, most well-resourced companies. Even the Government had to expend significant effort to apply differential privacy to the 2020 Census data, changing system designs multiple times.[10]

In addition to these technical limitations, there are also technical tradeoffs. First, there is a fundamental tradeoff with utility. Detailed data is valuable and has enabled wide swaths of data-driven ecosystems that benefit people and society. But, in some instances, applying PETs can reduce the informational value of data and, as a result, limit these benefits. Navigating this tradeoff requires PETs adopters "to have a clear idea of what information or value they are trying to protect, and they need to determine the potential benefits and costs of different PETs."[11] In some cases, further research and development could make this tradeoff less severe; technical advances could permit marginal increases in privacy resulting from the use of a PET to carry lower marginal costs to utility.

Another technical tradeoff—one the Request highlights—sometimes exists between applying PETs and achieving equity or fairness. For one, applying PETs can make analyses of data less accurate for smaller, historically marginalized communities represented in it.[12] Applying PETs may also be in tension with developing fairness-aware machine learning, which may require having access to sensitive information about individuals to measure and correct bias.[13] That said, there may be instances—exemplified by our research on SMPC[14]—in which PETs can actually enable privacy-protective insights into equity.

These technical limitations and tradeoffs will have to be carefully considered by all stakeholders in conversations around privacy, from technical conversations about the

[10] *See* M. Hawes, *Implementing Differential Privacy: Seven Lessons from the 2020 United States Census*, 2 Harv. Data Sci. Rev. (2020), https://hdsr.mitpress.mit.edu/pub/dgg03vo6/release/4.

[11] The Royal Society, *supra* note 9, at 29.

[12] This tradeoff was recently discussed in the context of applying differential privacy to U.S. Census data. *See* H. L. Wang, *For the U.S. Census, Keeping Your Data Anonymous and Useful Is a Tricky Balance*, NPR (Aug. 2, 2021), https://www.npr.org/2021/05/19/993247101/for-the-u-s-census-keeping-your-data-anonymous-and-useful-is-a-tricky-balance.

[13] *See, e.g.*, Federal Reserve Bank of San Francisco, Privacy Enhancing Technologies: Categories, Use Cases, and Considerations (2021) 16, *available at* https://www.frbsf.org/economic-research/wp-content/uploads/sites/4/Privacy-Enhancing-Technologies-Categories-Use-Cases-and-Considerations.pdf ("[T]esting for representativeness in a dataset may be important. . ., but that may not be possible if only a portion of the data can be viewed."); R. Cummings et al., *On the Compatibility of Privacy and Fairness*, Proceedings of the 27th Conference on User Modeling, Adaptation and Personalization 309 (2019), https://dl.acm.org/doi/abs/10.1145/3314183.3323847.

[14] *See* Austin, *supra* note 3.

direction of research to policy conversations about properly protecting data. In doing so, stakeholders should appreciate that the exact contours of limitations and tradeoffs will be highly dependent on the specific use cases and datasets to which PETs are applied.

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

PETs have the potential to be transformative across sectors and applications. Two stand out as particularly amenable to further innovation: data sharing with researchers, and advertising. PETs have great potential—as the Request notes—to enable sharing data with researchers to drive insights and innovation in a host of fields. We discussed previously how Meta's Data for Good and FORT teams leverage PETs to this end, and we encourage industry, academia, and federal agencies to do the same.[15] At the same time, parties should be mindful that any guidelines or frameworks used for sharing data will need to be flexible, accounting for limitations that PETs might impose in certain contexts on the representativeness of data and the generalizability of research findings.[16]

PETs will also be critical for building the next generation of personalized experiences and advertising. Consumers and businesses benefit from advertising relevant to their interests and needs, and industry and policymakers are increasingly working to develop even greater privacy protections. As discussed, Meta is exploring how a wide range of PETs—including SMPC, on-device learning, and differential privacy—can advance this goal.[17] These opportunities should not be overlooked, and we are eager to collaborate with policymakers, industry peers, and other stakeholders to realize PETs' full potential.

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

Regulatory uncertainty is one of the most pressing challenges facing the development and adoption of PETs. The White House has cited research from the intergovernmental Financial Action Task Force finding "that uncertainty about the regulatory implications of

[15] *See supra* notes 1-2.

[16] *Cf.* European Digital Media Observatory, Report of the European Digital Media Observator's Working Group on Platform-to-Researcher Data Access 11 (2022), *available at* https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf (highlighting challenges that differential privacy might pose to research outcomes and replicability in some instances).

[17] *See supra* notes 5-6.

using these technologies is a significant barrier to adoption."[18] And the Federal Reserve Bank of San Francisco has emphasized that PETs are often extremely costly and thus must be properly incentivized.[19] Regulatory uncertainty does the opposite.

The regulatory landscape in the United States is rapidly evolving at both the state and federal levels, and the Federal Trade Commission has stated its intent to commence a privacy-focused rulemaking.[20] As this landscape takes shape, we encourage regulators to consider the kinds of ambiguity and uncertainty that exist in current data protection schemes. Novel approaches should avoid these pitfalls, and incentivize innovation while protecting privacy. Without clear guidance for what legal obligations exist and how PETs can help comply with them, investing in PETs is a risky enterprise. We outline some proposals below, recognizing that implementation will vary depending on the legal regime being introduced or modified.

One fundamental uncertainty is the lack of agreement on key concepts relating to identifiability. Many data protection schemes impose requirements on "personal data," carving out data that have been "deidentified" or "anonymized." But schemes vary widely in how they define "personal data," and they use "deidentification" and "anonymization" in different, often overlapping or contradictory ways.

This discord stems from a deeper framework misalignment in many policy schemes - treating the risk of identifiability as a binary instead of a spectrum. Data have either been absolutely anonymized—reducing the risk of identifying an individual to zero—or not, and legal obligations are pinned to this binary. In practice, though, achieving absolute anonymization is often difficult, technically complex, and resource intensive.

But it can be possible to render data *effectively* anonymized—to reduce the risk of identification to a non-zero, but sufficiently negligible level. Data protection schemes should move toward this conception of effective anonymization, which the United Kingdom's Information Commissioner's Office has embraced.[21] In doing so, they should recognize not only that identifiability is a spectrum, but also that it is not a purely technical question. PETs can reduce the risk of identification, but so can other

[18] The White House, *supra* note 8. *See also* Financial Action Task Force, Stocktake on Data Pooling, Collaborative Analytics and Data Protection 32-33 (2021), *available at* https://www.fatf-gafi.org/media/fatf/documents/Stocktake-Datapooling-Collaborative-Analytics.pdf.

[19] *See* Federal Reserve Bank of San Francisco, *supra* note 13, at 6.

[20] *See* FTC Trade Regulation Rule on Commercial Surveillance, RIN 3084-AB69 (2022), https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=202204&RIN=3084-AB69.

[21] *See* Information Commissioner's Office, Chapter 2: How Do We Ensure Anonymisation Is Effective? (2021), *available at* https://ico.org.uk/media/about-the-ico/documents/4018606/chapter-2-anonymisation-draft.pdf.

protections, such as access controls, policies, procedures, and legal mechanisms. In some cases, effective anonymization might be achieved by using PETs and these other protections in tandem. In other words, although PETs may not always be a silver bullet, they are an important part of a suite of privacy-protective tools, methods, and practices.

This complexity and nuance around identifiability means that data protection schemes pegged to effective anonymization will have to be flexible. Prescriptive, one-size-fits-all tests for determining effective anonymization will not work. Rather, policies should provide a clear and consistent set of factors to balance when assessing risk, recognizing that assessing risk is highly use case-specific. Policies could help provide adopters certainty by incorporating concepts of reasonableness and industry best practices. In other words, policies should derisk beneficial uses of data by ensuring that good faith, considered efforts to deidentify data with PETs do not face undue legal burdens.

Policies should also consider how to incentivize privacy-enhancing data practices that are beneficial even if they do not achieve effective anonymization. For one, there might be contexts in which—given the nature of the data, the parties involved, and other factors—pseudonymized data are particularly valuable. Some schemes already exempt pseudonymized data from certain obligations.[22] There might also be ancillary privacy benefits of PETs other than reducing the risk of identification that are worth promoting. For example, some PETs, such as SMPC, could reduce the risk that a party to a computation can make inferences about individuals—an outcome that is valuable even if it does not relate to identifiability per se. These ancillary positive effects could be incentivized through tools like safe harbors and presumptions of good faith.

Finally, policies should explicitly permit data processing done in service of beneficial innovation. For example, processing should be allowed to deidentify data and to test the risk of identification.[23] Entities should also be permitted to process data for the purpose of measuring or increasing fairness, an end that the Request rightfully identifies as valuable and possibly in tension with strict, monolithic conceptions of privacy.

Exactly how these suggestions could be implemented in regulatory regimes will depend on their details and issuing authorities. But we encourage all agencies in the Federal Government to help bring greater nuance to the topic of deidentification. This could emerge in rules governing nongovernmental entities, those around interagency data

[22] *See, e.g.*, Virginia Consumer Data Protection Act § 59.1-577(D).

[23] The Federal Reserve Bank of San Francisco has highlighted the challenge of legally testing deidentification as a source of uncertainty. *See* Federal Reserve Bank of San Francisco, *supra* note 13, at 5 ("[P]rohibitions against the re-identification of data make the testing of the strength of de-identification techniques difficult.").

sharing, or—ideally—both. The Government has an opportunity to set examples for other regulators—not just in the United States, but around the world.

5. Specific laws that could be used, modified, or introduced to advance PETs

The principles laid out in our response to Topic 4 apply equally here. Just as the regulatory landscape in the United States is very dynamic, so is the federal statutory landscape. As Congress considers privacy-related bills, we encourage all Government stakeholders to evaluate how federal law can recognize the complexity of deidentification and the tradeoffs involved. A comprehensive federal privacy law that draws on our proposals above has the potential to incentivize innovation and reap the full value of data while protecting privacy and advancing PETs. We also encourage stakeholders to consider the ways in which the use and adoption of PETs may drive new and different thinking around how data is regulated in any proposed privacy legislation.

Outside of substantive privacy laws, Meta also encourages statutes that enable further research and development into PETs. Laws like the draft "Promoting Digital Privacy Technologies Act,"[24] for example, could help drive efforts needed to address some of the technical limitations and tradeoffs of PETs discussed in our response to Topic 2.

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

Our responses to the previous topics underscore the importance of consensus building on what are very complex issues without clear answers. The Federal Government can play an invaluable role in building this consensus through two additional mechanisms: developing technical standards and guidance, and fostering regulatory co-design.

In our response to Topic 4, we describe how regulatory uncertainty stems from a lack of agreement on foundational concepts like how to define and assess risk. We also suggest that, if policies shift their focus from absolute anonymization to effective anonymization, assessing effective anonymization should take into account industry best practices. Both of these challenges could benefit from Government leadership. Agencies like the National Institute of Standards and Technology could put forward standards and guidance addressing exactly these foundational questions. These materials could include, for example, a taxonomy that provides consistent vocabulary for talking about PETs and what they accomplish. In doing so, this vocabulary should allow stakeholders to understand how different PETs mitigate different discrete, people-centered privacy risks.

[24] *See* Promoting Digital Privacy Technologies Act, H.R. 847, 117th Cong. (2022).

Standards and guidance could also sketch a risk framework that treats identifiability as a spectrum and recognizes the multifaceted nature of risks and associated tradeoffs.

The Government could also create avenues for stakeholders to be intimately involved in co-designing regulations. Workshops building on the insights gained from this Request—on topics like effective anonymization mentioned above—would be valuable. Opportunities could also encompass tools like policy prototyping and regulatory sandboxing, which are useful in problem spaces—like deidentification—that are technically complex and not amenable to bright-line rules. At Meta, we are harnessing the potential of co-design through our Open Loop program.[25] Open Loop is helping policymakers in Europe, Latin America, and Asia craft artificial intelligence regulations, testing prototype policies in real-world conditions. We recently kicked off our largest prototyping program to date, with more than forty AI startups providing qualitative and quantitative feedback on provisions of the European Union's AI Act.

In these efforts, collaborating with a wide range of stakeholders will be key. Involving industry from different technical sectors will ensure that potential standards or policies are practicable, and that relevant use cases of PETs are not overlooked. The Government should also consult academic and civil society stakeholders representing marginalized and underrepresented groups, who have valuable perspectives on advancing equity. Finally, partnering with other governments—state, local, and foreign—will help prevent regulatory fragmentation. Data sharing and analysis are global activities that demand global consensus to fully incentivize privacy innovations and further their adoption.

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

Widespread embrace of PETs is currently hindered by a lack of public trust, understanding, and acceptance of PETs as a critical tool for protecting privacy. Without this trust and acceptance, it is harder for companies, particularly ones without significant resources, to justify the costs of implementing PETs. One challenge to building this trust is the difficulty of explaining PETs' benefits, in part due to their technical complexity. Many PETs are highly advanced technologies that make complex modifications to data or computations. It is difficult to explain these nuances to non-technical privacy stakeholders using intuitive language. For example, explaining how on-device learning works requires communicating the concept of "aggregated model summaries." That concept challenges even those with university-level degrees and high digital skills.

[25] *See* Open Loop, https://openloop.org/; N. Andrade, *Introducing Open Loop, a Global Program Bridging Tech and Policy Innovation*, Meta AI Blog (Jan. 11, 2021), https://ai.facebook.com/blog/introducing-open-loop-a-global-program-bridging-tech-and-policy-innovation/.

Another barrier to explaining PETs' benefits is that—even assuming simpler ways of communicating how they work—it is difficult to connect *what* a PET does in a technical sense to *how* it mitigates discrete privacy risks in real-world scenarios. Research by TTC Labs, a cross-industry effort we founded to create innovative design solutions, supports this. TTC Labs found that explanations of a PET are more beneficial when they focus on the PET's impacts on people's personal experiences than when they detail its inner workings.[26] Providing these more people-centered explanations might be easier for PETs like differential privacy, whose benefits are better understood, but harder for PETs like SMPC, that strengthen privacy in more nuanced ways.

Meta encourages further research into best practices for explaining PETs and their benefits, and we would welcome the opportunity to collaborate. The Government could also play a pivotal role in fostering a shared stakeholder understanding of PETs' benefits that would in turn bolster trust in PETs and prompt wider PETs acceptance and adoption. This could be achieved by, among other things, the taxonomy we suggest in Topic 6.

Conclusion

Meta is grateful for the opportunity to comment on this Request. PETs hold great potential for ushering in transformative, beneficial ways of using and sharing data while preserving privacy. But there are significant barriers. Research is required to advance the technical capabilities of PETs. There need to be more concerted efforts to explain the benefits of PETs to all relevant stakeholders. And data protection policies should move away from rigid conceptions of privacy that are incompatible with both the nuances of privacy risks and the complexities of how PETs mitigate them.

We are eager to lend our expertise in developing, applying, and explaining PETs to these efforts, and we welcome further conversations and collaborations. If you have any questions, please contact Brad Weltman.

[26] *See* L. Kaplan, *How Can Companies Help People Understand Privacy-Enhancing Technologies Like On-Device Learning?*, TTC Labs Research (2021), https://www.ttclabs.net/research/how-to-help-people-understand-privacy-enhancing-technologies.