Response of The MITRE Corporation to the OSTP RFI on Advancing Privacy-Enhancing Technologies

July 8, 2022

For additional information about this response, please contact:
Duane Blackburn
Center for Data-Driven Policy
The MITRE Corporation
7596 Colshire Drive
McLean, VA 22102-7539

policy@mitre.org
(434) 964-5023

<<This page is intentionally blank.>>

About MITRE

MITRE is a not-for-profit company that works in the public interest to tackle difficult problems that challenge the safety, stability, security, and well-being of our nation. We operate multiple federally funded research and development centers (FFRDCs); participate in public-private partnerships across national security and civilian agency missions; and maintain an independent technology research program in areas such as artificial intelligence, intuitive data science, quantum information science, health informatics, policy and economic expertise, trustworthy autonomy, cyber threat sharing, and cyber resilience. MITRE's 9,000-plus employees work in the public interest to solve problems for a safer world, with scientific integrity being fundamental to our existence. We are prohibited from lobbying, do not develop or sell products, have no owners or shareholders, and do not compete with industry. Our multidisciplinary teams (including engineers, scientists, data analysts, organizational change specialists, policy professionals, and more) are thus free to dig into problems from all angles, with no political or commercial pressures to influence our decision-making, technical findings, or policy recommendations.

MITRE has extensive privacy experience supporting federal, state, local, and international government agencies. MITRE's demonstrated privacy capabilities include conducting research, development, test and evaluation (RDT&E) activities that help government agencies better manage privacy risk, meet privacy compliance requirements, and strategically address privacy policy and technology challenges. RDT&E activities include investigating and reviewing privacy-enhancing technologies (PETs) and shaping privacy best practices to maximize the value of new and emerging technologies.

Additionally, MITRE recently established the Center for Data Privacy and Protection (CDP2) to better streamline the demands on the institution's privacy capabilities and corporate compliance efforts. The mission of CDP2 is to build privacy considerations into business operations and engagements by implementing privacy policies that reduce risk and foster trust, accountability, and transparency. The establishment of CDP2 further illustrates MITRE's commitment and value to privacy and data and protection.

Introduction and Overarching Recommendations

Protecting sensitive data is more involved than simply removing personal information from datasets. Modern PETs offer the potential to protect sensitive data while also helping government agencies achieve their mission goals. PETs represent one set of a series of tools that can be used to protect data and minimize legal, privacy, and ethical risks. To ensure proper understanding and use of PETs, MITRE recommends the following three overarching activities:

1. Conduct an independent review and analysis of existing PET products and services. Test and evaluate how well PETs perform in different scenarios, identify the technical expertise required to implement and maintain PETs, document the potential risks and rewards, estimate financial cost, and determine which solutions can be adopted and implemented in the near term.

2. Ensure the use of project management and systems engineering best practices. Systems designed via a "solution looking for a problem" approach rarely succeed and are not recommended. The first step instead should be to thoroughly define the problem/use case and then design an appropriate solution, which could include PETs as a component.
3. Conduct pilot use cases and document the benefits, limitations, successes, and areas for improvement.

Responses to Selected Questions Posed in the RFI

1. *Specific research opportunities to advance PETs*: Information about Federal research opportunities that could be introduced or modified to accelerate the development or adoption of PETs. This includes topics for research, hardware and software development, and educational and training programs. This also includes information about specific techniques and approaches that could be among the most promising technologies in this space.

Advancing PETs into proper application requires late-stage research centered on specific use cases, which can often be overlooked while developing overarching research strategies. Recommended use cases for this portion of a research strategy include:

Social Security Number (SSN) and personal information anonymization – Customer and personnel information systems frequently contain large collections of sensitive personal information, such as SSNs, bank account numbers, and vaccination records. Data processors frequently use manual processes to anonymize or mask sensitive data. PETs may speed up the process by automating anonymization and allowing the underlying, non-sensitive data to be used for intended purposes.

Zero Trust Architecture (ZTA) – Executive Order 14028 on Improving the Nation's Cybersecurity instructs federal agencies to adopt ZTA. PETs using homomorphic encryption, multiparty computation, or zero-knowledge proofs may help ZTA technologies to perform better at protecting confidentiality, integrity, and accessibility of data.

Digital assets – Executive Order 14067 on Ensuring Responsible Development of Digital Assets addresses privacy and data security throughout the Order. Securing and protecting data is critical to the stability and trustworthiness of any digital assets ecosystem. PETs promise to play a key role in digital assets privacy and security protections.

Public data – Public data from social media sites such as Twitter and Facebook may serve as an early warning indicator for federal, state, and local first responders. This might include information about fires, floods, and tornados or missing persons alerts. PETs may allow government first responders to use public data in a privacy-preserving manner, if sufficiently consistent with social norms.

Synthetic data – Artificial intelligence synthetic data generators evaluate real-world data and then generate statistically accurate synthetic datasets that mimic real-world data. This allows for accurate data analytics without disclosing personal information and identities. Use cases may include census data, taxpayer filings, healthcare records, and immigration trends.

2. *Specific technical aspects or limitations of PETs*: Information about technical specifics of PETs that have implications for their development or adoption. This includes information about specific PET techniques that are promising, recent or anticipated advances in the theory and practice of PETs, constraints posed by limited data and computational resources, limitations posed by current approaches to de-identification and deanonymization techniques, limitations or tradeoffs posed when considering PETs as well as technical approaches to equity considerations such as fairness-aware machine learning, security considerations based on relevant advances in cryptography or computing architecture, and new or emerging privacy-enhancing techniques. This also includes technical specifications that could improve the benefits or privacy protections, or reduce the risks or costs of adopting PETs.

PETs are most effectively deployed based on a holistic view of a use case, the environment in which they are embedded, and the nature of the relevant PETs. Absent this kind of broad systems approach, PETs may enable ethically and/or societally problematic use cases, alleviating surface concerns while simultaneously undermining more fundamental privacy norms. From a systems engineering standpoint, PETs are not Band-Aids that can be simply dropped onto system designs to render those that are privacy problematic less so. PETs should be considered one set of tools in a larger toolbox of privacy and risk management tools and strategies, to be applied as appropriate in an integrated fashion throughout the systems engineering life cycle.

PETs provide a structured approach to protecting data. However, there are a wide range of subjective requirements that must be addressed. These include legal authorities to collect and use the data for specific use cases; guarding against known and unknown biases such as age, race, and gender discrimination; and ethical considerations. Objective PETs solutions do not always address subjective risks. They often require human subject matter experts to analyze the risks and develop and implement appropriate protections in conjunction with PETs.

This is particularly true for PETs grounded in cryptography and/or theoretical computer science, which offer certain kinds of mathematical guarantees. How such guarantees relate to actual privacy requirements and objectives is not necessarily straightforward, and the work PETs do (or don't do, as the case may be) must be properly situated within the larger socio-technical system. Mathematical guarantees have little intrinsic value outside of their disciplinary contexts; their value is a function of the real-world requirements they support and the conditions under which they hold.

Promoting trust is a core privacy principle, and transparency is critical to promoting trust. The transparency process should work to inform underserved and marginalized groups that do not have time or resources to read privacy notices, privacy impact assessments, and system of records notices that their personal information is protected using privacy preserving data sharing and analytics technologies.

Finally, most PETs products and services lack benchmarks and metrics. MITRE recommends an independent entity conduct test and evaluation benchmarking and propose standards and metrics that allow government agencies to examine their options.

3. *Specific sectors, applications, or types of analysis that would particularly benefit from the adoption of PETs:* Information about sectors, applications, or types of analysis that have high potential for the adoption of PETs. This includes sectors and applications where data are exceptionally decentralized or sensitive, where PETs could unlock insights or services of significant value to the public, where PETs can reduce the risk of unintentional disclosures, where PETs might assist in data portability and interoperability, and sectors and applications where the adoption of PETs might exacerbate risks, including in the areas of privacy, cybersecurity, accuracy of data analysis, equity for underserved communities, and economic competition. This topic covers opportunities to improve the effectiveness of data sharing among specific Federal agencies and between specific Federal agencies and entities outside the Federal Government, including the goals outlined in Section 5 of Executive Order 14058: Transforming Federal Customer Experience and Service Delivery To Rebuild Trust in Government.

Key Areas

Government

Consumer protection – Each year the Federal Trade Commission and state consumer protection agencies receive millions of identity theft and fraud reports. PETs may help agencies process the records in a privacy-preserving manner that also identifies patterns leading to perpetrators.

Tax records – The IRS reported $2.3 billion in tax fraud for fiscal year 2020. PETs may help identify instances of taxpayer fraud while also preserving the privacy of law-abiding taxpayers.

Homeland Security, law enforcement, and national security records and data – PETs can help ensure that data has been collected lawfully, is being used and maintained in accordance with regulatory and policy requirements, and is shared in a privacy-respecting manner.

Private Industry

Banking, financial, and payment systems, and tax records – PETs allow industry members to exchange data in a secure and privacy enhanced way, as well as comply with state, national, and international data protection regulations. PETs may also help identify potential financial criminal activities such as money laundering and payments for illicit goods.

Healthcare records and data – Data is critical to quality healthcare, medical research, and artificial intelligence/machine learning research and development. Protecting patient data is also critical and required by statutes and regulations. PETs have the potential to enhance privacy protections beyond the existing statutory and regulatory requirements such as the Health Insurance Portability and Accountability Act (HIPAA). In turn, this would open the door to more advanced medical research and development.

Insurance industry data – Automobiles generate "telematics" data that records information such as mileage, fuel, geolocation, speed, and engine diagnostics. The auto insurance industry could use this data to improve automobile and driver

safety, reduce accidents, and set more accurate premiums. However, privacy regulations limit the use of this data. PETs may offer a solution that resolves those limitations.

Marketing data – Big data and social media companies generate revenues based on advertising and marketing data, matching their users to specific products and services. This involves extensive collection of personally identifiable information (PII) that, when combined, could lead to disclosures of sensitive information. PETs may allow users and social media companies to enhance protections of personal information.

Statistical Organizations

Census and statistical bureau/organization data – PETs such as differential privacy can be used to inject "noise" into datasets in a manner that sufficiently preserves the accuracy and privacy of the underlying data.

Trustworthy artificial intelligence (AI) and machine learning (ML) – Trustworthy AI involves building a series of elements and protections into AI/ML algorithms and models. Elements include accuracy, explainability, privacy, security, and mitigation of differential performance. PETs may provide a pathway to achieving some trustworthy AI goals.

PETs were originally developed to protect individual privacy. However, many of the government and industry use cases noted above involve organizational data that may not specifically contain PII. PETs may be adapted to protect sensitive organizational data.

5. *Specific laws that could be used, modified, or introduced to advance PETs:* Information about provisions in U.S. Federal law, including implementing regulations, that could be used, modified, or introduced to accelerate the development or adoption of PETs. This includes provisions, safe harbors, and definitions of use, disclosure, safeguards, and breaches. Information may also include comments on how to advance PETs as part of new or proposed legislation, such as that which would create a National Secure Data Service. Information may also include comments on State law or on international law as it applies to data sharing among international entities.

Safe Harbors assume that a technical privacy solution can resolve all privacy risks and should therefore allow the organization implementing PETs to escape any responsibility or liability pertaining to a privacy breach. The challenge with this approach is that PETs are one part of a multifaceted solution. PETs are an objective approach to privacy risks. But there are subjective approaches and analyses that also need to be conducted to understand the full scope of risks and mitigation strategies. Moreover, the protections afforded by PETs are not themselves absolute, and some residual risk will usually remain.

As stated in the response to Question 2, PETs provide an objective approach to protecting data. However, there is a wide range of subjective requirements that must be addressed. These include legal authorities to collect and use the data for specific use cases; guarding against known and unknown biases such as age, race, and gender discrimination; and ethical considerations. Objective PETs solutions rarely address subjective risks. They often require human subject

matter experts to analyze the risks and develop and implement appropriate protections in conjunction with PETs.

6. *Specific mechanisms, not covered above, that could be used, modified, or introduced to advance PETs:* This includes the development of open-source protocols and technical guidance, the use of public-private partnerships, prize challenges, grants, testbeds, standards, collaborations with foreign countries and nongovernmental entities, the Federal Data Strategy, and data sharing procedures with State, local, tribal, and territorial governments. This also includes interpretations and modifications of standard non-disclosure agreements, confidentiality clauses, data use or sharing agreements, etc.

There can be considerable confusion both on the part of potential PETs adopters and on the part of PETs developers that hinders effective design and use.

Potential PETs adopters often struggle to understand the relevant technical and operational characteristics of particular PETs, while PETs developers are often unclear about the characteristics and exigencies of real-world use cases. One way of addressing the first issue is development of standard design patterns for distinct types of PETs, especially those that are cryptographically based, while an approach to the second issue is the development of structured use case specifications.

Design patterns are structured solution templates for addressing recurring problems and have a long history in software development. They are highly adaptable, including with respect to the amount of technical detail. Appropriately configured design patterns that, among other facets, convey trust relationships and processing states could help potential PETs adopters better understand the key operational characteristics of different types of PETs. This would enable more accurate assessments of their applicability to specific use cases. While design patterns could facilitate better understanding of PETs functionality on the part of potential adopters, structured use case specifications could facilitate better understanding on the part of PETs developers of the types of problems for which solutions are sought. Such documentation would also benefit potential PETs adopters, as it would force them to articulate problems with sufficient granularity to enable meaningful analysis of the applicability of different types of PETs.

7. *Risks related to PETs adoption:* Identification of risks or negative consequences resulting from PETs adoption as well as policy, governance, and technical measures that could mitigate those risks. This includes risks related to equity for underserved or marginalized groups, the complexity of implementation and resources required for adoption, as well as from conceptual misunderstandings of the technical guarantees provided by PETs. This also includes recommendations on how to measure risk of PETs adoption and conduct risk-benefit analyses of use.

For reasons articulated in the response to Questions 2 and 5, sufficiently expansive risk analysis becomes more, rather than less, necessary for appropriate PETs deployment.

This applies to both the analytical methods and the risk models employed. This should leverage methodologies, where appropriate, beyond the typical privacy impact assessment, such as System Theoretic Process Analysis for Privacy.[1,2] More specialized forms of assessment should also be considered where appropriate. For example, MITRE has developed a Supplemental Technology Assessment methodology for a federal agency that is specifically intended to assess the privacy implications of using prosaic technologies in unusual ways or under atypical circumstances. These methodologies, in turn, must entail the use of sufficiently rich risk models that go beyond the standard ones revolving around Fair Information Practice Principles, such as Solove's taxonomy of privacy problems,[3] as well as synthetic consequences.[4]

Privacy risk models, though, need to expand beyond consequences (which most of them focus on) to model threats and vulnerabilities as well. Contextual integrity[5] is one way of conceptualizing privacy vulnerabilities. MITRE is currently developing a Privacy Attack Taxonomy that will provide a standard structure for mapping privacy attacks that can be used to model privacy threats.

Privacy threats are currently not well understood, and privacy threat modeling is not actively included in risk management processes within many organizations. It is important to effectively assess privacy threats and use privacy threat information as input for PETs selection. Otherwise, organizations may not select the privacy-enhancing technologies that are appropriate for their environment. MITRE's Privacy Attack Taxonomy will provide a standard structure for mapping privacy attacks that can be used to model privacy threats and facilitate privacy risk management, including PETs selection. Security risk modeling typically focuses on confidentiality-based threats to information about individuals (e.g., data breaches). However, the Privacy Attack Taxonomy will enable identification of threats beyond those typically addressed in security risk modeling (e.g., threats related to consent, notice, and inappropriate use, sharing, or retention of information about individuals). This expansion in focus will enable consideration of a broader set of PETs for potential implementation.

Risk-appropriate PETs deployments may be undermined by poor implementation. This is particularly the case with PETs based on cryptography. As discussed in the response to Question 6, standardized PET descriptions, such as PETs-specific design patterns, can help guard against this, as well as against the application of particular PETs to use cases for which they are ill suited. Poor implementation of the right solution or selecting a misaligned solution in the first place may lead to greater problems.

De-identification is an aspect of PETs in which organizations are often challenged with selecting the appropriate methodology and properly implementing it. De-identification reduces the ability

---

[1] S. Shapiro. Privacy Risk Analysis Based on System Control Structures: Adapting System-Theoretic Process Analysis for Privacy Engineering. 2016. IEEE, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7527748. Last accessed July 1, 2022.

[2] R.J. Cronk. Strategic Privacy by Design, 2nd edition. 2022. International Association of Privacy Professionals, https://iapp.org/resources/article/strategic-privacy-by-design/. Last accessed July 1, 2022.

[3] D. Solove. Understanding Privacy. 2010. Harvard University Press, https://www.hup.harvard.edu/catalog.php?isbn=9780674035072. Last accessed July 1, 2022.

[4] S. Shapiro. Deriving and Using Synthetic Consequences for Privacy Risk Modeling. In ICT Systems Security and Privacy Protection. 2022. Springer, https://link.springer.com/book/10.1007/978-3-031-06975-8. Last accessed July 1, 2022.

[5] H. Nissenbaum. Privacy in Context: Technology, Policy, and the Integrity of Social Life. 2009. Stanford University Press, https://www.sup.org/books/title/?id=8862. Last accessed July 1, 2022.

to associate information with an identifiable individual, thereby supporting data privacy and security. However, de-identification is typically used in an all-or-nothing fashion, acting as the sole privacy risk control for a dataset. In principle, though, de-identification should be usable as one of a set of privacy risk controls. For example, MITRE is developing a Data De-Identification Process Architecture that appropriately guides the application of de-identification as a privacy risk control by aligning the extent of de-identification with utility requirements (i.e., intended or projected uses) via quantitative models, assessing residual privacy risk, and indicating additional controls to mitigate the residual risk. ***This approach is distinct from typical approaches that either prioritize addressing risk to enable dataset release or default to maximal security protection of minimally de-identified data.***

Privacy, ethics, and civil liberties risks are normally addressed on a use case-by-use case basis. Existing privacy regulations focus on protecting personal information. However, data analytics generates different risks. How will the use of PETs be integrated with existing, subjective privacy, ethics, and civil liberties reviews involving religion, ethnicity, gender, age, and disabilities? As noted earlier, legal, privacy, civil liberties, and ethics subject matter experts will still be needed to identify and mitigate these risks.

8. *Existing best practices that are helpful for PETs adoption:* Information about U.S. policies that are currently helping facilitate adoption as well as best practices that facilitate responsible adoption. This includes existing policies that support adoption, including in the areas of privacy, cybersecurity, accuracy of data analysis, equity for underserved communities, and economic competition. This also includes information about where and when PETs can be situated within tiered access frameworks for accessing restricted data, ranging from publicly accessible to fully restricted data.

There are several existing privacy best practices and frameworks that can be used to facilitate PETs adoption. These include:

1. **NIST SP 800-53, Rev. 5, Security and Privacy Controls for Information Systems and Organizations**[6] – Provides a catalog of security and privacy controls for information systems and organizations to protect organizational operations and assets, individuals, and other organizations from a diverse set of threats and risks, including hostile attacks, human errors, foreign intelligence entities, and privacy risks. The controls are flexible and customizable, and are implemented as part of an organization-wide process to manage risk. Consideration should be given in the next revision to including additional PETs-related controls beyond the current ones.
2. **NIST Privacy Framework**[7] – Voluntary tool intended to help organizations identify and manage privacy risk to build innovative products and services while protecting individuals' privacy.

[6] Security and Privacy Controls for Information Systems and Organizations. 2020. National Institute of Standards and Technology, https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final. Last accessed July 6, 2022.

[7] Privacy Framework. 2022. National Institute of Standards and Technology, https://www.nist.gov/privacy-framework. Last accessed July 6, 2022.

3. **Fair Information Practice Principles**[8] – Widely accepted as a general framework for privacy requirements that is reflected in numerous privacy statutes and regulations in the U.S. and internationally. The principles serve as the basis for analyzing privacy risks and determining appropriate mitigation strategies. In particular, the principle of Data Minimization calls for organizations to collect only personal information directly relevant and necessary to accomplish the specified purpose and retain data only for as long as is necessary.
4. **Privacy by design** – Originated from PETs development and implementation, incorporating privacy principles into system and business process development and operation.
5. **Privacy engineering** – Supports the operationalization of privacy by design by applying systems engineering principles and approaches to the development of socio-technical systems. MITRE's Privacy Engineering Framework provides high-level guidance regarding fundamental privacy engineering activities, including how to map them to different types of life cycles (e.g., agile).
6. **MITRE Privacy Maturity Model**[9] – Framework for developing, implementing, maintaining, and evaluating privacy programs within organizations.
7. **MITRE Supplemental Technology Assessment** – Enhanced method of identifying privacy risks and mitigation strategies to minimize risks and maximize rewards in specific contexts, beyond what traditional privacy impact assessments normally identify.
8. **MITRE ATT&CK Framework**[10] – Curated knowledge base that tracks cyber adversary tactics and techniques, many of which frequently impact the confidentiality of PII and sensitive data.

MITRE has extensive experience with these and other privacy best practices and frameworks. This working knowledge has been applied to support federal, state, local, and international government agencies' adoption and implementation of privacy and security policies and procedures. MITRE has also supported the testing, evaluation, and implementation of PETs at government agencies with privacy best practices incorporated.

9. *Existing barriers, not covered above, to PETs adoption*: Information about technical, sociotechnical, usability, and socioeconomic barriers that have inhibited wider adoption of PETs, such as a lack of public trust. This includes recommendations on how such barriers could be overcome. Responses that focus on increasing equity for underserved or marginalized groups are especially welcome.

[8] The Fair Information Practice Principles: Framework for Privacy Policy at the Department of Homeland Security. 2008. Department of Homeland Security, https://www.dhs.gov/sites/default/files/publications/privacy-policy-guidance-memorandum-2008-01.pdf.

[9] Privacy Maturity Model, Version 1. 2019. MITRE, https://www.mitre.org/sites/default/files/publications/pr-19-3384-privacy-maturity-model.pdf. (Note that Version 2 has been completed and will soon be published to www.mitre.org/privacy.)

[10] MITRE ATT&CK. 2022. MITRE, https://attack.mitre.org/. Last accessed July 6, 2022.

Privacy and security overlap in various areas. However, there are unique aspects to privacy that are not addressed by security, particularly regarding notice, consent, individual participation, and collection and use limitation. Privacy and security are mutually supportive, and privacy and security teams should work closely together to protect information about individuals. Better integration between privacy, cybersecurity, and systems/technology development and acquisition is needed to successfully implement PETs.

Many organizations do not have mechanisms in place whereby these different areas can regularly engage. PETs implementation should include formal mechanisms that allow engagement across different PETs stakeholders in an organization. For example, a PETs advisory board can be used that is composed of representatives from areas such as privacy, security, legal, information technology, and data management so that inputs regarding PETs selection and implementation are provided from all relevant stakeholders.

More education regarding PETs, privacy engineering, and technical aspects of privacy is needed. Privacy professionals have historically been more focused on legal, regulatory, and compliance issues, and do not typically have the technical skills needed to manage privacy risks regarding the use of technology. Individuals working in PETs stakeholder areas besides privacy frequently do not have the right level of knowledge of privacy needed for engagement regarding the use of technology that handles information about individuals.

Organizations considering use of PETs should assess privacy workforce needs and identify privacy-related skillset gaps. NIST is currently leading development of a NIST Privacy Workforce Taxonomy, which will contain task, knowledge, and skill statements that are aligned with the NIST Privacy Framework and the National Initiative for Cybersecurity Education Workforce Framework for Cybersecurity. Considerations regarding PETs selection and implementation should be included in the NIST Privacy Workforce Taxonomy. Privacy training should be enhanced to include PETs selection and implementation considerations, and privacy certifications, such as the Certified Information Privacy Technologist certification available from the International Association of Privacy Professionals, should include knowledge of PETs areas as a requirement.

Conclusion

Privacy-enhancing technologies, in conjunction with other privacy and security risk mitigation methodologies, have the potential to substantially enhance PII and sensitive data protections, reduce privacy and security risks, and allow authorized users access to data in a secure manner. MITRE recommends an independent study and review of existing PETs be conducted. The review should include PETs not grounded in mathematical formalisms. Improved capabilities for detecting PII, tracking data flows, specifying and enforcing policies, and measuring privacy risk posture, to note just a few examples, are as important from a utility standpoint as those that leverage cryptographic protocols.