WORLD **PRIVACY** FORUM

Comments of the World Privacy Forum to the Office of Management and Budget regarding Request for Information on Privacy Impact Assessments

Sent via https://www.regulations.gov

Office of Management and Budget
725 17th St. NW
Washington, DC 20503

March 29, 2024

The World Privacy Forum is pleased to submit comments in response to the Office of Management and Budget's *Request for Information on Privacy Impact Assessments*. OMB published the request in the Federal Register on January 30, 2024, 89 Federal Register 5945, https://www.govinfo.gov/content/pkg/FR-2024-01-30/pdf/2024-01756.pdf.

The World Privacy Forum is a non-partisan public interest research group focused on conducting research and analysis in the area of privacy and complex data ecosystems and their governance, including in the areas of identity, AI, health, and others. WPF works extensively on privacy and governance across multiple jurisdictions, including the U.S., India, Africa, Asia, the EU, and additional jurisdictions. WPF has written in-depth, influential studies, including groundbreaking research regarding systemic medical identity theft, India's Aadhaar identity ecosystem —peer-reviewed work which was cited in the landmark Aadhaar Privacy Opinion of the Indian Supreme Court — and The Scoring of America, an early and influential report on machine learning and consumer scores. Recently, WPF published *Risky Analysis: Assessing and Improving AI Governance Tools*, a global analysis of the implementation layer of trustworthy AI. This report, published in 2023, defines AI governance tools and benchmarks this area of critically important work. WPF co-chairs the UN Statistics Data Governance and Legal Frameworks working group, and is a special advisor to the WHO's HDC board. At OECD, WPF researchers participate in the OECD.AI Expert Groups, among other activities. WPF participated in the core group of AI experts that worked on the OECD Recommendation on Artificial Intelligence, now widely viewed as normative principles regarding AI. Recently, WPF participated in the

2024 update to the AI Recommendation. WPF research on complex data ecosystems governance has been presented at the National Academies of Science and the Royal Academies of Science. See World Privacy Forum for more information, https://www.worldprivacyforum.org.

The World Privacy Forum has some modest experience with the Privacy Act of 1974. Over the years, we have often commented on System of Records Notices (SORNs) and Routine Uses (RUs) published by federal agencies. We were also the administrative sponsor of a recent report proposing a complete revision of the Privacy Act of 1974. See Robert Gellman, *From the Filing Cabinet to the Cloud: Updating the Privacy Act of 1974* (2021), https://www.worldprivacyforum.org/2021/05/from-the-filing-cabinet-to-the-cloud-updating-the-privacy-act-of-1974/.

In 2023, OMB published its *Request for Comments on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence Memorandum.* The request was published in the Federal Register on November 3, 2023, https://www.govinfo.gov/content/pkg/ FR-2023-11-03/pdf/2023-24269.pdf. WPF filed comments in response to this request, https://www.worldprivacyforum.org/wp-content/uploads/2024/03/WPF_Comments_OMB_AgencyUse_of_AI_3Dec2023_fs.pdf.
In our response to the Nov. 3 OMB draft memorandum, we included a section on PIAs and the Privacy Act of 1974, which is again relevant here for this new RFI. While there is some overlap, we made suitable adjustments for the new request and have added substantive additional material.

Part I. Use Existing Privacy Requirements to Assist with New AI Obligations: An Overview

Existing law provides a useful method for informing the public about and for seeking comments on agency activities affecting privacy. We refer, of course, to the Privacy Act of 1974, 5 U.S.C. § 552a. The Privacy Act of 1974 directs agencies to use Systems of Records Notices (SORNs) to describe their activities involving the use of personal information. Agency activities that involve personal information are described through notices in the Federal Register with the solicitation of public comment, and the notices remain available publicly to inform the public of agency activities. In many ways, publications by agencies under the Privacy Act of 1974 are unique. Other countries that have more advanced and more comprehensive privacy laws mostly abandoned requirements for publications that include descriptions of personal data systems. In our view, the Privacy Act of 1974's publication provisions are one of the Act's most successful obligations. This is true not only because the publications inform the public of agency processing of personally identifiable information (PII) but because it obliges each agency to think about and provide written descriptions of its processing activities. This is particularly true where new technologies, systems, or platforms are being considered for deployment.

The Privacy Act of 1974 became law long before anyone envisioned the need for a privacy impact assessment (PIA). A later law, the E-Government Act of 2002, 35 U.S.C. § 3501 note, imposed a requirement for PIAs. The statutory requirement was always inadequate for the purpose, but it was supplemented by useful OMB guidance, the OMB *Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002* (September 26, 2003) (M-03-22).

To a significant extent, many of the features of both of the PIA requirements are simply out of date and trail current technology and information practices by decades. A focus on Artificial Intelligence (AI) and PIAs is most welcome, but there are other shortcomings with PIAs and with the Act itself that still need improvement. To some degree, success with AI impact assessments will require attending to some of the more significant aspects of the Privacy Act that need to be addressed in regards to modernization.

We note, for example, that this Administration is aware of the need to focus more attention on the use by federal agencies of data brokers and other commercial data providers who provide personal data and other information for the use of federal agencies, as well as attention on data brokers who receive personal information from federal agencies.[1] WPF applauds that effort with great enthusiasm. Many of these activities are hidden from the American public because existing privacy laws and rules do not expressly direct agencies to disclose their use of data brokers as sources or recipients. Data broker data and other commercially available information resources are integral to some AI activities, and the need for data will only increase, even with lean data techniques.[2]

It is possible without a major effort to fold the AI and data broker initiative together with a limited reform of existing Privacy Act of 1974 and PIA requirements. An initiative sponsored by WPF already includes ideas and fully drafted language that will accomplish all of these tasks. The report from that initiative, *From the Filing Cabinet to the Cloud: Updating the Privacy Act of 1974* (2021), is cited above. We note that the report was completed before the current wave of attention to the use of AI.

The goal of that report is a complete revision of the Privacy Act of 1974. That goal will not be accomplished administratively, but some administrative changes can still achieve the same ends. For example:

[1] *Readout of White House Roundtable on Protecting Americans from Harmful Data Broker Practices*, August 16, 2023. https://www.whitehouse.gov/briefing-room/statements-releases/2023/08/16/readout-of-white-house-roundtable-on-protecting-americans-from-harmful-data-broker-practices/ Note: WPF was a participant in this Roundtable.

[2] For example, see: Yejin Kim, Scott Rome, Kevin Foley et al., *Improving Content Recommendation: Knowledge Graph-Based Semantic Contrastive Learning for Diversity and Cold-Start Users*, GraphLab, George Washington University and Applied AI Research, Comcast, March 2024. https://arxiv.org/pdf/2403.18667.pdf .

1.) Better publications: The report proposes on page 85 to expand existing requirements to describe categories of sources of information in a Privacy Act system of records. We note that AI is likely to be a new category of information for many AI activities covered by the OMB memorandum. Language in the report – which can work just as well in an OMB memorandum as in a statute – expands upon the existing requirement to describe the categories of sources of records in the system.

> "New language [proposed in the draft legislation] makes it explicit that sources include commercial, governmental, and other sources that the agency routinely reviews, consults, or uses. It is especially important for agencies to inform the public when using commercial sources. For example, if an agency has a contract with a consumer reporting agency ("credit bureau") to use credit records, it must so state. If there is a reasonable prospect that the particular source may change but not the category of sources, the agency may choose to identify the category (e.g., "credit bureau") rather than identifying which specific credit bureau it uses. If an agency routinely uses Internet search engines to find information on individuals, it must so state. If an agency routinely seeks information from social media, the agency should identify at least the major social media used. All the information about sources will help individuals figure out how particular information about them ended up in agency records. This is especially important when the agency uses the information to make decisions about individuals. It is even more important if an agency consults but does not retain a copy of information held by a third party." (*From the Filing Cabinet to the Cloud*, p. 85)

This type of disclosure addresses current Administration priorities for AI and for data brokers. OMB could make an adjustment in its Privacy Act of 1974 and PIA *administrative requirements* to expand public disclosure on data broker activities. We note that not only may agencies themselves use AI tools that need to be disclosed and explained, but data brokers and others who provide support and data to agencies may also independently use AI in providing that support. OMB should tell agencies expressly that there must be disclosure of the AI's use, terms, and controls if there is use of AI anywhere in the chain of federal PII processing. The need for disclosure is the same whether an AI activity is internal to an agency, external through a contractor, or even deeper in the background through resources or data used by an agency contractor or by a cooperating state agency.

In particular, we are concerned about the use of AI in connection with computer matching activities. One issue is that the use of AI may be hidden from public view by the currently inadequate disclosures that occur when agencies undertake matching. A second issue relates to the due process requirements that the matching process requires. Those who have been identified by computer matching as meeting certain criteria may not be aware that AI based standards used to identify individuals are biased, discriminatory, or simply insufficient for the purpose.

In short, we suggest that OMB should take a closer look at computer matching in an AI context. As a minimum, any AI component in computer matching should be disclosed and vetted. The Federal Privacy Council might be useful in vetting AI activities that involve multiple agencies or common private sector sources.

2.) Better PIAs: Formal impact assessments can be useful to address many different policy concerns. Everyone seems aware of the overlaps between some AI assessments and some privacy assessments. The same report that proposed revising the Privacy Act of 1974 also suggested revising the existing privacy impact assessment requirement. We will not repeat all of the details here. We refer you to pages 115-123 of the report.

We summarize here, however, by noting that the report emphasized the need for a PIA *process* and just not a flat, one-time, one-size-fits-all assessment. This need will be just as true for an AI assessment as it is for a privacy assessment.

We state expressly that:

- Some assessments will require more attention, more consultation, and a long time than others;
- Some assessments will require regular reviews over time because consequences are not static and because agency programs change over time; and
- Some assessments will require lesser efforts because the risks are smaller.
- Responsible agency personnel should be allowed to make determinations about which activities need more assessment than others. OMB may be able to propose standards for making these distinctions.

We specifically recommend that:

A. OMB expand existing requirements to describe AI as a new category of sources of information in a Privacy Act system of records.

B. Each PIA expressly state if any PII processing activity covered by the PIA involves (or does not involve) the use of AI, whether the use of AI is accomplished directly by the agency or indirectly by agency contractors or data vendors;

C. Each PII processing activity covered by an existing PIA be revised before the agency adds any new use of AI to the activity, whether the use of AI is accomplished directly by the agency or indirectly by agency contractors or data vendors;

D. Each system of records notice should state whether the notice covers (or does not cover) any processing of PII that involves the use of AI; and

E. That any processing that involves AI that is added to an activity covered by a SORN be added to the SORN in the usual fashion so that the public will have notice and the opportunity to comment.

Part II. Responses to Specific Questions

In this section we have directed our responses to specific questions in the RFI. The relevant question(s) are repeated in this document, followed by our response.

Responses to Question 3

> *3. What guidance should OMB consider providing to agencies to help reduce any duplication that may arise in preparing PIAs along with other assessments focused on managing risks (e.g., security authorization packages or the AI impact assessments proposed in OMB's Draft Memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence) and to support these assessments' different functions?*

Regarding question 3, given the significant overlap with some AI assessments and some privacy assessments, we suggest that an agency's Privacy Officer and the agency's AI Officer be directed to identify those overlaps and to work jointly on the assessments that they determine need substantial attention from both perspectives. OMB should not allow overlapping and intersecting policy concerns to be evaluated in a stovepipe manner by entirely separate offices.

OMB should be express in telling agencies that joint assessments and recommendations are required when appropriate. The WPF report proposing a revision to the Privacy Act (cited above) offers specific ideas on assessments that will also be relevant to AI assessments.

We repeat again here that the report emphasized the need for a PIA *process* and not a flat, one-time, one-size-fits-all assessment. This is especially true for AI activities because they are relatively new and can have intensive overlaps with traditional Privacy Act PIAs. We note again that while the Privacy Act revision proposed legislative changes, much can be accomplished administratively by OMB in its memoranda. There is no need to wait for legislation for making major changes to assessments.

We have an additional concern that is best discussed here in relation to PIAs and AI Impact Assessments being a ***process that is continuous***. We can't stress this need for an ongoing process enough. Allow us provide a concrete example as to why. First and foremost, a looming issue is that AI models require constant updating. We have concerns regarding how the pace for AI assessments intersects with the administrative tools currently available. For example, consider the serious problems attached to data drift.

Exemplar — Data drift monitoring: There is a need to monitor and update data and various aspects of data distribution in AI models due to *data drift*. Data drift happens as ML models shift over time due to a number of factors. When the models are being used in a healthcare context, there can be serious consequences unless data drift is monitored and addressed. The U.S. FDA articulated this best in a recent paper on the topic:

> "Machine learning (ML) methods often fail with data that deviates from their training distribution. This is a significant concern for ML- enabled devices in clinical settings, where data drift may cause unexpected performance that jeopardizes patient safety." [3]

There are data drift monitoring tools and processes available for AI models, and a great deal of innovation and growth is occurring in this area. Data drift monitoring and adjustment is not something that is done once a year; it is a process that is ongoing, and part of essential AI hygiene. We urge you to consider the U.S. FDA Office of Science and Engineering Laboratories paper on data drift which articulates this issue in the health / AI context.[4] This is exactly the kind of use case that reveals the challenges of where PIAs and AI Impact Assessments need to be reformed. How can the U.S. government move to a more process-oriented approach to allow these impact assessments to impact the implementation layer of AI systems? This is worth deep consideration and a great deal of work. WPF has engaged in this work looking directly at the tools - metrics - implementation layer of AI systems where all of this activity is taking place. We have done this because in AI systems privacy is effectuated or not at the implementation layer. Ensuring patient safety (and privacy) in AI systems will be densely and intricately automated, and this will require significant intellectual, technical, legal, and procedural adjustments based on data, metrics, and analysis gathered from monitoring. In thinking about clinical systems that have a PIA and an AI Impact Assessment, both need to be process-inclusive and reach into the implementation layer, which could require automation and further work to ensure the automation is validated and fit for purpose.

[3] Ghada Zamzmi, Kesavan Venkatesh, Brandon Nelson, Smriti Prathapan, Berkman Sahiner, Paul Yi, Jana G. Delfino, *Out-of-distribution detection and data drift monitoring using statistical process control*, preprint, U.S. Food and Drug Administration Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, 12 February 2024. https://arxiv.org/pdf/2402.08088.pdf .

[4] Ghada Zamzmi, Kesavan Venkatesh, Brandon Nelson, Smriti Prathapan, Berkman Sahiner, Paul Yi, Jana G. Delfino, *Out-of-distribution detection and data drift monitoring using statistical process control*, preprint, U.S. Food and Drug Administration Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, 12 February 2024. https://arxiv.org/pdf/2402.08088.pdf . *See also*: Dipak Wani, Samuel Ackerman, et al., *Data drift monitoring for log anomaly detection pipelines*, IBM Research, 17 October 2023. https://arxiv.org/pdf/2310.14893.pdf.

Recommendations regarding Question 3:

A. The agency's Privacy Officer and the agency's AI Officer should identify overlaps between PIAs and AI Impact Assessments and to work jointly on the assessments that they determine need substantial attention from both perspectives.

B. Ensure that the AI Impact Assessment and PIA are ongoing processes that are continuous.

C. OMB should study and document the technical, legal, and procedural adjustments needed to conduct privacy and AI assessments of AI systems using AI governance tools, or other metrics, techniques, and procedures alone or in combination with current approaches such as standardized PIAs, etc.

Responses to Questions 4, 5 A-B, and 6

> *4. What role do PIAs play in your search for information about how agencies handle PII and address privacy risks? For what purpose(s) do you read agencies' PIAs?*
> *5. What improvements to PIAs would help you better understand agencies' assessment of privacy impacts and risk mitigation strategies?*
>
> > *a. What improvement(s) would you recommend to make it easier to find and access agencies' PIAs?*
> > *b. What improvement(s) would you recommend to make it easier to read and understand agencies' PIAs?*
> > *6. How can agencies increase awareness of PIAs among stakeholders?*

Lack of awareness among stakeholders is a significant problem for the Privacy Act of 1974. With all of the current attention to privacy policy matters, the Act still remains obscure. Further, we note that there are quite a few Privacy Act publications in the Federal Register. While preparing these comments, we conducted a search for the number of publications that mentioned the Act. In the year preceding our search, there were 487 documents. Of course, while not all documents represent major Privacy Act matters, there is still a lot to review.

While some public interest groups pay some attention to agencies' Privacy Act activities in narrow ways, most Privacy Act notices are ignored and receive little, if any, public response. The WPF tries to review Privacy Act matters, but we do not have the resources to be comprehensive. Review of a new SORN or a revised SORN can be relatively simple, but reviewing a major new activity and its PIA requires much more attention. To the extent that the activities involve AI, the work will be even harder. We suspect that anything involving AI will get more attention generally, and to the extent that privacy matters are interwoven with AI, that may help to attract more attention to privacy as well.

We admit to some frustration because when we do respond (often to proposed routine uses that are significantly overbroad), agencies typically pay no attention. Whether this is at least in part a consequence of the absence of formal Administrative Procedure Act notice and rulemaking is hard to say, but it may be a contributing factor.

We do not think that this lack of public attention is likely to change much without more effort by agencies. We have a few ideas to help improve the situation.

First, for major Privacy Act activities (and especially for those involving AI), agencies might be directed to specifically identify those individuals, public interest groups, and other stakeholders that are most likely to be affected. That notice could be part of a SORN.

Second, OMB might establish a calendar or resource that highlights significant Privacy Act and AI assessments. It should not be difficult to establish a central webpage for this purpose and to update it regularly. The Federal Privacy Council could play a role here, but it is a low-profile organization. OMB, after the creation of a centralized resource, could convene civil society, business, and other stakeholders to bring awareness to any new resources.

Third, for major assessments – and especially for new SORNs, significant AI activities involving processing of PII, multi-agency activities, or major adjustments to existing activities – agencies could be directed or encouraged to convene a meeting with stakeholders at the earliest possible stage of development. Across government, we guess that there might be no more than a dozen assessments annually that would warrant a stakeholder meeting of this type. This will assist in meeting the goals of the minimum practices for rights – impacting AI as discussed in OMB Memorandum M-24-10, *Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence* (March 28, 2024), which specifies that in some circumstances government agencies should reach out proactively to stakeholders for feedback.

Fourth, there is a real need for an independent privacy organization (regulatory or otherwise) to review federal government privacy activities. Proposals for the creation of a federal privacy agency have been around for decades, and this is not something that OMB can do on its own. The only non-statutory approach would be for OMB itself to step up its own review of Privacy Act matters. We would support more attention from OMB to privacy and to AI.

Recommendations for Questions 4 and 5:

A. Agencies could identify stakeholders relevant to a an agency AI activity, and provide a Federal Register notice of this, for example, as part of a SORN.
B. Establish a centralized resource announcing relevant AI - related Privacy Act activities.
C. Identify stakeholders for major Privacy Act activities and convene them early, including civil society.

Responses to Question 7 A-B

> *7. AI and AI-enabled systems used by agencies can rely on data that include PII, and agencies may develop those systems or procure them from the private sector.*
>
> > *a. What privacy risks specific to the training, evaluation, or use of AI and AI-enabled systems (e.g., related to AI system inputs and outputs, including inferences and assumptions; obtaining consent to use the data involved in these activities; or AI-facilitated reidentification) should agencies consider when conducting PIAs?*
> > *b. What guidance updates should OMB consider to improve how agencies address and mitigate the privacy risks that may be associated with their use of AI?*

Regarding 7a, This question raises a number of substantive issues. We are focusing here on issues where we have depth of knowledge and have conducted field research, though we acknowledge many additional challenges exist.

First, regarding inferences and the use of AI system outputs, we refer to you to *The Scoring of America*, a report we published in 2014. It was among the very first reports that addressed the precise intersection of privacy, fairness, consumer scoring, and AI. Yes, it is old. However, the report was deeply researched and quite early, and the policy analysis in the report regarding how to address bias and improper uses of AI outputs is still quite pertinent today as a matter of policy. We refer here to that report and incorporate its recommendations as a part of these comments. See: Pam Dixon and Robert Gellman, *The Scoring of America*, World Privacy Forum, April 2014. https://www.worldprivacyforum.org/2014/04/wpf-report-the-scoring-of-america-how-secret-consumer-scores-threaten-your-privacy-and-your-future/ .

Second, regarding consent in biometric systems, we refer you to Pam Dixon's extensive field research on India's Aadhaar biometric system, which contains the biometrics of 1.4 billion people. This research contains a detailed discussion of biometrics and consent in the Indian, European, and U.S. context and this discussion would be helpful in response to this question. We incorporate this research by reference into these comments as a response to 7a. We note that this research was directly cited by the Supreme Court of India in its landmark Aadhaar decision regarding biometrics and privacy. See: Pam Dixon, *A Failure to Do No Harm: India's Aadhaar biometric ID program and its inability to protect* *privacy in relation to measures in Europe and the U.S., S*pringer Nature, Health Technology. DOI 10.1007/s12553-017-0202-6. http://rdcu.be/tsWv. Open Access via Harvard- Based Technology Science: https://techscience.org/a/2017082901/.

Third, OMB has largely and appropriately focused on individual consent. However, we would be remiss if we did not mention the issue of collective consent in the AI context. As OMB likely is aware, AI systems can be at odds with indigenous socio-technical approaches, which often stress privacy as a collective issue and not as an individual issue. There is a critically important policy literature written by indigenous peoples regarding data held at the tribal level which includes persuasive legal arguments that in some jurisdictions, tribal governments possess the authority to enact data privacy laws at the tribal level, and that this would help define what constitutes “tribal data.” [5] The boundaries of what is and is not tribal data is a central question in the development of inclusive AI standards. An additional important issue is the idea of collective data ownership, and collective privacy rights, as well as the collective application of ethical principles in AI development. These types of approaches can be seen, for example, in the U.S. Indigenous Data Sovereignty Network [6] and the Māori Data Governance Model, Te Kāhui Raraunga, among other indigenous governance frameworks, such as the First Nations Principles of OCAP. [7]

Regarding 7b, there are enough issues here to fill many books. We will be brief and focus on guidance regarding the evaluation of AI systems for privacy and trustworthiness, including compound AI systems, because this sits at the very core of the questions OMB has posed here. We briefly discussed evaluative issues in AI systems in our response to Question 3 in our discussion of data drift as an exemplar.

To reiterate the example briefly here, data drift (which is just one example of a type of problem that can occur) is dangerous in AI systems, and can have deleterious impacts. AI systems used in the health context, for example, must monitor in an ongoing fashion for data drift or risk patient safety.[8] Something that OMB has largely not addressed is that most AI systems have already become complex enough to require automated checking for many areas of AI trustworthiness, including for issues such as data integrity, privacy and bias. WPF calls such automated evaluative tools AI governance tools, which we

[5] Tsosie, *Tribal Data Governance and Informational Privacy: Constructing 'Indigenous Data Sovereignty,*' 80 Montana Law Review 229 (2019)

[6] U.S. Indigenous Data Sovereignty Network, https://usindigenousdatanetwork.org/ .

[7] *The First Nations Principles of OCAP,* First Nations Information Governance Centre. https://fnigc.ca/ocap-training/

[8] Ghada Zamzmi, Kesavan Venkatesh, Brandon Nelson, Smriti Prathapan, Berkman Sahiner, Paul Yi, Jana G. Delfino, *Out-of-distribution detection and data drift monitoring using statistical process control*, preprint, U.S. Food and Drug Administration Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, 12 February 2024. https://arxiv.org/pdf/2402.08088.pdf .

define as “Socio-technical tools for mapping, measuring, or managing AI systems and their risks in a manner that operationalizes or implements trustworthy AI.”[9]

These tools are already in widespread use and are relied upon by the private sector. They are likely relied upon by Federal agencies, though this has not been surfaced. AI governance tools sit at the implementation layer and are used to check on the status of implementation. It is one thing for an agency to state that it wants to check AI systems for bias and for privacy. It is another to address the reality of the implementation for AI systems, which will need to use checks that go well beyond the current conception of AI Impact Assessments and PIAs. A large range of vetted and high-functioning AI governance tools will be necessary. WPF interviews in the private sector for the forthcoming update of *Risky Analysis* indicate strongly that large multinational companies are having to build and adapt automated AI governance tools routinely in order to evaluate their AI systems for privacy impacts.

The privacy risk here is that current privacy guidance is not able to accommodate how privacy is assessed in AI systems at the actual implementation context. Currently, as we are at the beginning of the installation and use of today’s advanced AI systems, there is still some overlap between what we are used to by way of impact assessments and what we are moving into. Unless privacy guidance for Federal agencies is adapted to older methods as well as newer methods, there will be an increasing disconnect in guidance and the reality on the ground. This can still be avoided.

PIAs are important, and WPF supports their use. However, the current iteration of PIAs alone will not be enough to address the full range of challenges that AI presents. Additional tools and approaches will be needed.

We add here that assessments for privacy, human rights, certain aspects of governance, and other assessments and validation in relation to today’s AI activities are far from mature. And even the most perfect assessment tools will not be enough to address the full range of challenges through the AI lifecycle. More guidance will be needed to address all of the aspects of the lifecycle, including implementation through automated AI governance tools. This work has barely begun in regards to addressing advanced forms of AI.

OMB can develop additional guidelines that address these issues. A good beginning would be to provide guidance regarding how to navigate the transition as a variety of types of assessment are still in play.

Another good step would be to provide guidance regarding quality control for the use of AI governance tools in the context of AI systems and privacy. We are aware that this is not the way most agencies think about privacy. However, as we have mentioned multiple

[9] Kate Kaye, Pam Dixon, Robert Gellman ed., John Emerson, data visualization, *Risky Analysis: Improving and assessing AI governance tools*, World Privacy Forum, 15 December 2023. https://www.worldprivacyforum.org/2023/12/new-report-risky-analysis-assessing-and-improving-ai-governance-tools/ .

times, privacy checks in AI systems are increasingly being automated. The larger the system, the more this becomes the case. If the tool is faulty, the measurement results will be faulty, and the end result in AI systems can be poor privacy measurements and outcomes. To mitigate these and related issues, WPF has suggested several key early mitigations. We have discussed these mitigations in depth in Risky Analysis.

Recommendations for Question 7:

A. **Require documentation and validation of the AI governance tools used to perform checks on privacy and other aspects of trustworthy AI.** First and foremost, AI governance tools often lack documentation or validation. Our research found high variability in the documentation and labeling of AI governance tools. This suggests that developing norms regarding documentation and labeling of AI governance tools could produce meaningful levels of initial quality improvements. For example, it would be helpful if tools routinely include information about the developer, date of release, results of any validation or quality assurance testing, and instructions on the contexts in which the methods should or should not be used. A privacy and data policy is also important and should be included in the documentation of AI governance tools.

Additional items can be provided in the documentation, for example:

- Appropriate performance metrics for validity and reliability

- Documentation should provide the suggested context for the use of an AI governance tool. AI systems are about context, which is important when it comes to applicable uses, environment, and user interactions. A concern is that tools originally designed for application in one use case or context may potentially be used in an inappropriate context or use case or “off-label” manner due to lack of guidance for the end user.

- Documentation should give end users an idea of how simple or complex it would be to utilize a given AI governance tool.

- Cost analysis for utilizing the method: How much would it cost to use the tool and validate the results?

- A data policy: A detailed data policy should be posted in conjunction with each AI governance tool. For example, if applicable, this information could include the kind of data used to create the tool, if data is collected or used in the operation of the tool, and if that information is used for further AI model training, analysis, or other purposes.

- Complaint and feedback mechanism: AI governance tools should provide a mechanism to collect feedback from users.

- Cycle of continuous improvement: Developers of AI governance tools should maintain and update the tools at a reasonable pace.

- Conflict of interest: The identities of those who financed, resourced, provided, and published AI governance tools should be made public in a prominent manner in conjunction with publication or distribution of the tool.

B. **Address how to capture the automation of the implementation layer of AI systems in the AI Impact Assessment process.** The more complex and compound the AI system, the more it will require multiple tools, metrics, and approaches to asses AI systems for trustworthy AI, including privacy. This process will likely need to be an ongoing process at the implementation level.

C. **Ensure that agencies have a mechanism and procedure to find or make and then test AI governance tools.** If such a mechanism or procedure is not in place in the public sector, it will be difficult if not impossible for agencies to ensure privacy in AI systems.

D. **Ensure there is training for agencies to adjust their processes to AI tools that are used to determine trustworthiness of AI systems** as well as what frequency of updating and checks will be needed for each context.

E. I**ncorporate U.S. indigenous views in OMB's AI privacy work.** Reach out to U.S. Tribal leaders to understand and address views on collective privacy and collective consent vis à vis AI systems. There is almost no work being done in this area.

Responses to Question 8 A-B:

> *8. What role should PIAs play in how agencies identify and report on their use of commercially available information (CAI) that contains PII?*
>
> *a. What privacy risks specific to CAI should agencies consider when conducting PIAs?*
>
> *b. OMB M–03–22 requires PIAs "when agencies systematically incorporate into existing information systems databases of information in identifiable form purchased or obtained from commercial or public sources, "while noting that "[m]erely querying such a source on an ad hoc basis using existing technology does not trigger the PIA requirement." What guidance updates should OMB consider to improve how agencies address and mitigate the privacy risks thatmay be associated with their use of CAI that contains PII?*

First, we will discuss commercially available information — often bought and sold by data brokers — that contains PII. There should be meaningful procurement safeguards that ensure that the commercial data is:

A. **Consensually collected** with meaningful opt in (double opt-in) and prominent notice.

B. **Dated**; in other words, the **collection date** for the CAI should be noted in whatever is being used by Agencies.

C. **Verifiably accurate and tested independently to be accurate** to agreed upon and reasonable cutoff points which provide consumers with protection from the risks of being adversely impacted by incorrect data.

   1. Cutoff points for levels of accurate data should be no lower than 98 percent. The risks stemming from inaccurate consumer data are too high at this point.

D. **Data must be checked against activities that have polluted the data**; for example, if a person is a victim of identity theft, multiple aspects of CAI that exists about them may be entirely incorrect. Any predictive analytics based on this data will be skewed.

E. **There must be a means for consumers to opt out of CAI data and to object to data that is old, out of date, inaccurate, or discriminatory**.

Second, we would like to discuss commercially available information that has been massaged, cleansed, aggregated, and stripped of identifiers. This data does not objectively contain PII per the government's definition of PII. However, due to significant advances in AI and analytical capacity, it is very possible today to use a variety of cleaned, aggregated data to predict the actions, probable economic future, risks, and other aspects of individuals. The aggregate data analysis has become so sophisticated today that PII is not necessarily needed for individual members of the public to be negatively impacted by purportedly neutral or deidentified CAI.

As the Privacy Act revision report (*From the Filing Cabinet to the Cloud*, WPF) suggests, the collection of PII from private sector sources is a major issue that the Act fails to address in any meaningful manner. Whether agencies use private sector PII for AI activities or in other ways, there is a tremendous need for greater disclosure of sources and uses. With AI, there is the real possibility that the private sector will supply data that is technically not PII but that is used to affect how an agency treats individuals. This and other technical workarounds have become increasingly problematic.

For example, a private company could collect and analyze PII and then report to an agency its findings of what classes of individuals are most likely to engage in fraud or be eligible for assistance. These examples would not expressly trigger Privacy Act disclosure or other obligations, but they can be troubling. The point here is that focusing just on private sector activities that involve the sharing of PII is too narrow. For AI purposes, the use of PII anywhere in the process warrants both attention and disclosure.

Another activity that needs more attention is the use of PII within an agency. The Privacy Act provision on internal use (5 U.S.C. § 552a(e)(1)) has an open-ended standard for internal uses ("to those officers and employees …who have a need for the record in the performance of their duties"). In practice, we believe that this provision imposes absolutely no barrier to internal uses. Fixing this problem is beyond the scope here, but OMB can take steps to avoid making things worse as AI systems are implemented.

Our concern here is that agency records can be analyzed internally using AI (or otherwise) to make decisions about individuals without any meaningful public notice and without any possibility of due process. OMB needs to impose standards for the internal agency use of PII in AI activities that have any effect on the rights, benefits, or privileges of individuals. We suggest an overt requirement for identifying these activities in SORNs that provide source material for the AI activities as well as in the SORNs for the AI activity itself. In effect, we propose that OMB direct agencies to establish the equivalent of *routine uses* for any PII sharing within an agency that uses AI in any way that would an individual's rights, benefits, or privileges. Agencies should be obliged to notice the public and to accept public comments on AI-related internal uses.

We also want to consider the intersection of AI with existing routine use requirements. The Act imposes a standard for the definition of routine uses, which are disclosures to anyone outside an agency. The statutory standard is both broad and vague, and it rarely prevents an agency from making disclosures that the agency wants to make. Even if the statutory standard would prohibit an external disclosure, agencies often promulgate an overly broad routine use or simply ignore the limits of the law.

Our concern with the impact of AI systems here is that weakening of the limits of the law as a result of agency practice over decades will make matters worse. Agencies may use vague routine uses to allow for the sharing of PII for contributing to the training or implementation of an AI-based function. Without an express requirement for disclosure of the specifics of any AI-related external disclosure, the public may have no idea how one agency may contribute to another agency's AI training or even to the training or other activities of a private sector AI function. In short, whether internal or external, any AI-related use or disclosure must be expressly identified.

Recommendations for Question 8:

A. Enact meaningful safeguards regarding the collection and use of commercially available data, whether it be personal data, PII, or aggregate data.

B. Agencies should be obliged to provide notice to the public and to accept public comments on AI-related **internal uses.**

C. Set precise rules around the setting of Routine Uses regarding AI — do not allow agencies to stuff meaningful AI activities into obscure RU notices that over time lead to a weakening of the law at the ground level.

Responses to Question 9

> *9. What else should OMB consider when evaluating potential updates to its guidance on PIAs?*

We repeat the comment we made above regarding the Privacy Act of 1974 being obscure in comparison to flashier privacy statutes, and we add here, neglected. We will not go over OMB's history with its Privacy Act oversight other than to say that it has been haphazard. Technology and other factors are highly changeable, and PIA guidance, like other privacy and AI guidance, needs to be nimble. It needs to be updated regularly, and PIAs need to be attached to a fulsome process and commitment to ongoing improvement.

We reiterate here our responses to questions 3 and 7. OMB will need to address the significant emergence today of automated privacy and trustworthy AI assessment tools. This stands to become a major factor in privacy assessment, and OMB needs to begin addressing the issues of integrating validation of these tools and creating guidance around their use by the government as soon as practicably possible. Otherwise, poorly functioning tools could over time deeply undermine all of the work OMB is doing now on adjusting for increased AI usage in the government.

Certainly, OMB can do more to highlight examples of noteworthy PIAs. Perhaps OMB – working together with the Federal Privacy Council and perhaps even some organizations outside the federal government – could give recognition to the best of the agency PIAs on an annual or biennial basis. If agencies compete with each other for recognition, the result could be across-the-board improvements. In addition, the best ideas would provide a basis for updating guidance on a regular basis.

We note here that the *Privacy Impact Assessment for the Body Worn Camera Program (May 2019 DOJ PIA Template)*, is a good exemplar of a responsible PIA that was conducted on a very controversial issue. The PIA is thorough, and informative for the public. If more PIAs were like this, the public would be much better informed about government systems. U.S. Bureau of Alcohol, Tobacco, Firearms and Explosives, September 16, 2022, https://www.justice.gov/d9/2023-08/atf_bwc_pia_final.pdf .

Recommendations for Question 9:

A. **Establish nimble processes** for PIA and AI guidance.

B. **Address the emergence of automated tools that assess privacy and trustworthy AI at scale.** This is a completely new animal, and this entire area needs focused and informed guidance.

C. **Highlight worthy instances of agency PIAs.**

Conclusion

We are living through an extraordinary time in the history of privacy. There are few who would question that modern forms of AI are already transforming key sectors of the economy and are well on their way to permeating major (and minor) digital ecosystems.

The issue is not *if,* but *how*, and *what* are the contours of the best and most effective response(s) to the changes. We acknowledge with gratitude that OMB is working as quickly as possible to understand and address how privacy and trustworthiness is operating in this new environment at the agency level.

It is tempting for us to write endlessly about all of the issues we see. Here, we have sought to provide OMB with comments based on empirical research. We are very aware that some of what we are saying about new forms of automation of privacy assessment and trustworthy AI assessment at scale is easy to dismiss unless you have lived through exploring the data, the AI, the privacy impacts, and interviewing entities large and small, public and private sector, regarding what is happening in the ground reality of AI implementation. WPF has done this, and we would welcome an opportunity to discuss the deeper research with you. In the trenches of implementation, it is fairly easy to see the trendlines and where they are leading. We have articulated these hard-won observations in our comments.

We conclude these comments with a final comment and recommendation. We note that OMB Memorandum M-24-10, published March 28, 2024, contained a significant recommendation regarding hiring AI talent within agencies, in (4) c. AI Talent. We agree with this recommendation, and would encourage an addendum relating to OMB's efforts regarding the intersection of privacy and AI. That is, AI expertise needs to be joined with legal, privacy, governance, and human rights skills and knowledge. Otherwise, there could be meaningful cultural differences as experts learn to adapt to new forms of AI while complying with legal structures and governance requirements pre-dating advanced AI. If this combined skill set is not contained in one person, then perhaps it can be brought together in a cooperative team. This kind of socio-technical-legal plus AI skillset will be indispensable at OMB and will be essential in working through the challenging intersections of AI, privacy, and the need for effective identification of and mitigations for problems. AI cannot be ignored, but neither can the Privacy Act. Both are important. There will need to be a careful and respectful navigation of the interests and tradeoffs made in this negotiation between legal and AI ecosystems.

We look forward to engaging with OMB on the important issues regarding how AI is intersecting with privacy, inclusion, and fairness, among many other areas of importance.

Respectfully submitted,

Pam Dixon
Executive Director, World Privacy Forum
www.worldprivacyforum.org