August 22, 2019

The Office of Management and Budget
725 17th Street NW
Washington, DC 20503
ATTN: Russell T. Vought, Acting Director

Re: Identifying Priority Access or Quality Improvements for Federal data and models for Artificial Intelligence Research and Development (R&D), and Testing

Dear Acting Director Vought,

Inpher appreciates the opportunity to advise the Office of Management and Budget ("OMB") on federal data safeguards for research and development ("R&D") and testing efforts on agency uses of artificial intelligence ("AI") systems.[1] We make the following recommendations pursuant to the Executive Order on *Maintaining American Leadership in Artificial Intelligence*,[2] and the OMB's mandate under Section 5.a.i of the order to identify needs for additional access to, or improvements in the quality of, federal data used in AI R&D.

As agencies are tasked with the increasing demands of modern governance, it has become the priority of the federal government to establish a national strategy to invest in emerging technologies that can improve public administration. With the appropriate oversight, advances in AI and machine-learning ("ML") can augment and correct deficiencies in human capital with quantifiably more efficient and accurate processes.

Innovation at the federal level increases the capacity of agencies to regulate and respond to rapid technological advancements in the private sector. Harnessing these benefits is therefore an urgent goal for the U.S. government, especially for the purposes of evidence-based policymaking, civic engagement, and competent enforcement.

Yet, the urgency to seize these opportunities must be balanced with proper safeguards for civil liberties and individual interests implicated in the development and deployment of AI in government. Public accountability is key to implementing a national AI strategy that is

[1] The Office of Management and Budget, *Request for Information: Identifying Priority Access or Quality Improvements for Federal Data and Models for Artificial Intelligence Research and Development, and Testing* (Fed Reg. 2019-14618), https://www.regulations.gov/document?D=OMB-2019-0003-0001

[2] United States, Executive Office of the President Donald Trump, *Executive Order on Maintaining American Leadership in Artificial Intelligence* (Feb. 11, 2019), https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

sustainable and internationally recognized. Therefore, American leadership in AI R&D must fundamentally protect and prioritize data privacy.

We strongly urge the OMB to make the following recommendations: (1) agencies should transition from outdated legacy IT systems which can no longer securely store, share, and process personal data, and (2) the federal government should facilitate regulatory sandboxes for privacy-enhancing technologies ("PETs") to improve institutional capacity-building in AI.

Inpher Background

We are a US-based cryptography and machine-learning company with the conviction that encryption and privacy are foundational to the future of computing and commerce. Inpher applies years of academic research on Fully Homomorphic Encryption ("FHE") and secure Multi-Party Computation ("MPC") into commercially-ready applications that financial institutions are using in production today.[3]

Inpher's customers include some of the world's largest multinational financial institutions that use our software platform for privacy-preserving analytics and computation with mathematical guarantees of data security and sovereignty. This 'secret computing' technology enables compliant data processing across siloed departments, cross-jurisdictional and cross-industry information sharing, and zero-knowledge cloud computing, as the host never 'sees' the data nor has access to the keys. Our legal and public policy department facilitates public education on privacy-preserving technologies and advocates for data protection by design, global privacy, and algorithmic accountability.

Issues and Gaps in Current AI Policy

Building on civil society advocacy and incremental academic literature on the importance of algorithmic accountability, policymakers are attempting to animate these values with regulatory and legislative reform. The past two years have witnessed a surge of proposals and new regulations that provide greater access to personal data and its downstream uses in predictive modelling and insight analytics. The implementation of the EU General Data Protection Regulation ("GDPR") in May 2018 has also provided a catalyst for U.S. policymakers to draft a national AI policy[4] and propose an omnibus privacy legislation.[5]

---

[3] Inpher, *Case Studies*, https://www.inpher.io/case-studies-1#case-studies

[4] U.S. National Science and Technology Council, *THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN* (Oct. 2016), https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf

[5] House Committee on Energy & Commerce, HEARING ON "PROTECTING CONSUMER PRIVACY IN THE ERA OF BIG DATA" (Feb. 26, 2019),

Although the federal government has made great strides in addressing the evolving impact of automation, a careful review of the current legal landscape indicates several issues inhibiting effective regulation of data protection in AI.

First, existing AI governance frameworks are broadly and vaguely worded, in contrast to the categorical precision requisite in data science. When laws mandate the provision of an explanation about the basis of an automated decision and how different factors weighed into the final output, the extent of this disclosure is unclear. Presently there is no decisive authority on whether "a right to explanation" contemplates the disclosure of a source code that executed the function; labelling categories for input data; training models; audit logs—or if a high-level explanation akin to a credit score analysis would suffice.[6]

Second, algorithmic transparency rights are rarely invoked to achieve the checks and balances they aim to provide against the societal threats of emerging technology.[7] Parallel to the failures of privacy policies in actuating real, informed consent for data processing, disclosure requirements for automated tools like facial recognition, employment filtering, and risk assessments have minimal real-time impact. Requirement of transparency does little to alleviate the information and power asymmetry between the system developers and individuals impacted by automated decisions. Lack of judicial precedent in both national and international courts exacerbate the confusion on how algorithmic transparency laws will be interpreted in practice to determine accountability for adverse automated decisions.

Third, most legislation targets commercial deployments of AI for consumer protection against unfair or deceptive practices. This creates a regulatory vacuum for government-deployed

---

https://energycommerce.house.gov/committee-activity/hearings/hearing-on-protecting-consumer-privacy-in-the-era-of-big-data; Senate Committee on Commerce, Science, and Transportation, Hearing on Policy Principles for a Federal Data Privacy Framework in the United States (Feb. 27, 2019), https://www.commerce.senate.gov/public/index.cfm/2019/2/policy-principles-for-a-federal-data-privacy-framework-in-the-united-states

[6] Fair Credit Reporting Act, 15 U.S.C. § 1681 et seq.; As of December 2004, the Fair Credit Reporting Act as modified by the Fair and Accurate Credit Transactions Act, or FACTA, ended score secrecy formally, and required consumer reporting agencies to provide consumers with more extensive credit score information, upon request. Also made available to the public was the context of the score (its numeric range), the date the score was created, some of the key factors that adversely affected the score, and some other items.

[7] In 2015, EY conducted an article by article analysis of Regulation EC No. 45/2001 (now repealed and replaced by Regulation EU 2018/1725 to reflect the GDPR), a supplementary legislation for data processing by EU institutions. It operated analogously with the 1995 Data Protection Directive and contained a provision on automated decision-making. The report states that "The right not to be subject to individual automated decisions is rarely applied but is entirely relevant in the context of the development of big data," indicating that individuals rarely invoked their right not to be subject to automated decision-making by EU public authorities. See, Evaluation regulation 45-2001 - European Commission – Europa, https://ec.europa.eu/newsroom/just/document.cfm?action=display&doc_id=40697

automated systems that are being integrated into administrative functions in all three levels of government: federal, state, and municipal.

There is a dearth of guidelines to constrain public sector uses of AI. Search queries on 'public sector', 'government', 'administrative agency', and 'artificial intelligence', 'machine-learning', and 'automation' fetch more information on national AI policies aimed at economic competition and job market growth, than accountability for the deployment of new technologies within government. This imbalance likely indicates the lack of public knowledge about the various sandbox initiatives for automation in government agencies, and how commonplace such tools have become in everyday public administration.

Current legal frameworks take a necessary step forward in AI disclosures, but represent insufficient safeguards for government accountability for federal data breaches and unauthorized data processing during the deployment of AI/ML. Therefore, federal agencies should embed privacy engineering requirements and validation standards for personal data used as input for AI systems.

Institutional Capacity Building for Sustainable AI

In 2016, the U.S. Government Accountability Office ("GAO") released a report: 'Federal Agencies Need to Address Aging Legacy Systems.'[8] The GAO identified an alarming number of federal agencies that are currently using obsolete legacy IT systems that were implemented from 31 to 56 years ago.

Due to a lack of incentive for technological capacity-building, federal agencies often neglect to conduct necessary risk assessments on information systems that are used to make significant decisions on enforcement, rulemaking, and adjudication. Public sector inertia for innovation is not only bad policy, but also extremely detrimental to the collective privacy rights of individuals whose sensitive information is stored in vulnerable legacy systems.

Responsible innovation will require the federal government to continually engage in multi-stakeholder dialogue with the private sector, academia, international organizations, and civil society. In particular, consultations on the use of federal data should focus on and facilitate sandboxes in privacy-preserving AI and ML.

Privacy-Preserving Technologies for Federal Data

Inpher believes that governmental collaboration with experts in applied PETs will be a necessary step to addressing the vulnerabilities of legacy information systems in federal government, and to

[8] U.S. Government Accountability Office, *Federal Agencies Need to Address Aging Legacy Systems* GAO-16-696T (May 25, 2016), https://www.gao.gov/products/GAO-16-696T

pave the way for federal data utility and compliant inter-agency data sharing. Cryptographic technologies can provide a solution to traditional tradeoffs in privacy and analytical precision (for example, with differential privacy methods), and allow secure collaboration across agency data silos for greater coordination and scalability.

Advances in MPC and FHE allow functions to be performed on encrypted data without revealing the underlying information. Therefore, cryptographic PETs such as MPC and FHE offer incorruptible *ex ante* privacy safeguards against unauthorized access by intermediaries and third parties.[9] The regulatory focus on data training and analysis for AI should shift to implementing PETs that can keep data securely encrypted in storage, transit, and *in-use* (while being processed), so that sensitive plaintext information is not exposed to those who may violate their data-sharing agreement or fiduciary obligations to engage in misconduct.

Collaborative information-sharing on advanced privacy-preserving technologies such as MPC and FHE will be critical to investing in AI that is scalable and interoperable across various agencies and departments. MPC and FHE, which keep data encrypted in use, can reduce barriers to the use of AI technologies to "promote their innovative application while protecting American technology, economic and national security, civil liberties, privacy, and values."[10] These guidelines would protect data provenance and enforce appropriate use policy in AI R&D. They would moreover instill organizational accountability by requiring agencies to implement better technological safeguards and protective measures for privacy.

Conclusion and Recommendations

Public-private partnerships via hackathons, tech sprints, and regulatory sandboxes are integral to understanding, supporting, and applying new technologies that enable (1) secure and intelligent data insight, (2) multimodality machine-learning on heterogeneous data sources,[11] and (3) integrating models and ontologies in the training process.

At the international level, the UK Financial Conduct Authority ("FCA")[12] and Information

---

[9] Yehuda Lindell & Benny Pinkas, *Secure Multiparty Computation for Privacy-Preserving Data Mining*, The Journal of Privacy and Confidentiality (2009), http://jpc.cylab.cmu.edu; *ING Belgium Sees Opportunities for 'Secret' Sharing of Encrypted Data*, The Wall Street Journal (Jun. 1, 2017), https://blogs.wsj.com/cio/2017/06/01/ing-belgium-sees-opportunities-for-secret-sharing-of-encrypted-data/

[10] Executive Order on Maintaining American Leadership in Artificial Intelligence (Feb. 11, 2019), at Sec. 6(a)(ii).

[11] NATIONAL SCIENCE & TECHNOLOGY COUNCIL, *A Report by the SELECT COMMITTEE ON ARTIFICIAL INTELLIGENCE: THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN: 2019 UPDATE* (Jun. 2019), at 9 (Advancing data-focused methodologies for knowledge discovery).

[12] UK Financial Conduct Authority, *2019 Global AML and Financial Crime TechSprint* (Held on Jul. 29, 2019 to Aug. 2, 2019), https://www.fca.org.uk/events/techsprints/2019-global-aml-and-financial-crime-techsprint

Commissioner's Office ("ICO")[13] have recently hosted sandbox sessions designed to support the development of products and services that aid compliance with data protection laws. Inpher participated in the FCA TechSprint to demonstrate the application of secure MPC to enable knowledge sharing between financial regulators to detect trade anomalies and monitor financial crimes whilst adhering to privacy law requirements.[14]

These partnerships create room for technologically-driven regulation and create demonstrable public value with responsible innovation. We encourage the OMB and the federal government to support privacy-preserving technologies in AI R&D. We believe that regulatory sandboxes in PETs will facilitate valuable interdisciplinary collaboration that is foundational to the long-term development of AI that is interoperable, multifunctional, and most importantly, safe for individuals.

Thank you for the opportunity to comment on this important consultation. If you have any questions regarding our comments, or if Inpher could be of any assistance, please do not hesitate to contact me at sunny@inpher.io.

Sincerely,

Sunny Seon Kang
Senior Privacy Counsel, Head of Policy
Inpher, Inc.

[13] Information Commissioner's Office, *ICO opens Sandbox beta phase to enhance data protection and support innovation* (Mar. 29, 2019), https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/03/ico-opens-sandbox-beta-phase-to-enhance-data-protection-and-support-innovation/

[14] Inpher, *Inpher Wins People's Choice Award at FCA TechSprint* (Aug. 9, 2019), https://www.inpher.io/news/2019/8/9/inpher-wins-peoples-choice-award-at-financial-conduct-authoritys-2019-tech-sprint.