Chapter 8: Upholding Democratic Values: Privacy, Civil Liberties, and Civil Rights in Uses of AI for National Security

Blueprint for Action

The U.S. needs an approach for adopting AI domestically for national security that upholds and bolsters respect for democratic values, including privacy, civil liberties, and civil rights. Such an approach must strengthen, provide, and show leadership with regard to: 1) transparency; 2) approaches for AI system development and testing; 3) the ability to contest AI decisions; 4) oversight over AI development and use; and 5) legislative and regulatory controls on how AI is used. Our recommendations include immediate actions that the President, the Congress, and agencies should take; a comprehensive assessment by a Task Force that leads to reforms for AI governance and oversight; and areas for continued work. The recommendations are aimed at assuring that AI systems used by national security agencies uphold democratic values. Secondarily, the adoption of these recommendations can earn and inspire public confidence, both domestically and abroad, in uses of AI by national security agencies.

Recommendation

Recommendation Set 1: Increase Public Transparency about AI Use through Improved Reporting

Actions for Congress:

- **For AI systems that involve U.S. persons, require AI Risk Assessment Reports and AI Impact Assessments to assess the privacy, civil liberties and civil rights implications for each new qualifying AI system or significant system refresh.**
    - The Commission proposes Congress require elements of the Intelligence Community (coordinated by the Office of the Director of National Intelligence (ODNI)) as well as the Department of Homeland Security (DHS) and the Federal Bureau of Investigation (FBI), to prepare and publish an AI Risk Assessment Report and conduct AI Impact Assessments to assess the privacy, civil liberties, and civil rights implications of each new qualifying AI system or significant system refresh. The Commission recognizes the current requirements for privacy impact

assessments and civil liberties impact assessments done at agencies as required by current statute. AI-related technologies may be reviewed by these, but are not fully/adequately captured by the current assessments. The Commission's recommendation intends to augment these requirements.

- The AI Risk Assessment Report and AI Impact Assessment would be required for "new qualifying AI systems" and for "significant system refreshes." The Commission proposes that the Task Force described later in this Blueprint be charged with determining the decision procedures for identifying which AI systems and significant system refreshes would require AI Risk and Impact Assessment Reports.
- The intent of the AI Risk Assessment Report and AI Impact Assessment is to ensure potential impacts are considered and mitigated while avoiding an unnecessary increase in compliance burdens.
  - Legislated frameworks for ensuring effective and pragmatic risk mitigation (with the ability to categorize systems per risk and determine the appropriate mitigations if any) exist in other models that can be used as a template (e.g. FISMA).
- The *AI Risk Assessment Report* should include a detailed analysis of system implications for, and *steps to mitigate and track risks (e.g., through metrics) to:*
  - Freedom of expression (e.g., is the AI-enabled surveillance targeting people because of their First Amendment protected activity or is the AI-enabled government surveillance causing or may potentially cause a chilling effect?);
  - Equal protection (e.g., is the AI-enabled surveillance biased toward a protected class? What are the likely effects the new technology or program will have on key demographics such as race, gender, or disability?);
  - Privacy (e.g., is a warrant required for the government action? Are minimization and query processes sufficient/satisfactory?);
  - Redress and due process (e.g., what mechanisms exist, or limitations have been accepted, for providing redress for adverse government actions taken based on information generated by the AI system?); and
  - The assessment should account for the environment in which the AI system will be deployed, including its interactions with other AI tools and programs that collect personally identifiable information (PII).
- *AI Impact Assessment* should be made available periodically, but no less than annually, to the agency's Privacy and Civil Liberties (PCL) Office to determine the degree to which a qualifying AI system remains compliant with the constraints and metrics established in the Risk Assessment. AI Impact Assessments should be based on outcomes, impacts, and metrics collected during system use, and determine if the existing validation processes should be improved.
- *Resources and staffing.* PCL Offices should assess the resources, including staff, needed to adequately complete the above. Agency heads should support additional resourcing for PCL Offices as part of the annual budget process.
- *Disclosure notices.* Congress should require ODNI, DHS, and the FBI to review non-public and/or classified AI programs once the program is shut down for declassification and/or disclosure.

Action for the President:

- **Should Congress not require new privacy, civil liberties, and civil rights reporting (as identified above), conduct AI Risk Assessment and AI Impact Assessment Reports as described above.**

Actions for DHS and the FBI:

- **DHS and the FBI should impose new obligations for System of Record Notices (SORNs) and Privacy Impact Assessments (PIAs) specific to AI systems to ensure that they provide richer information.**
  - o SORNs and PIAs should provide a holistic picture about the collection, use, and storage of personal information by any AI system, including its connections to existing systems and accounting for the layering of different surveillance technologies where applicable. Agency practices do not sufficiently support the production of SORNs and PIAs that adequately depict how AI systems collect, use, and store personal information.[1]
  - o DHS and the FBI should require that all PIAs include description of the algorithm(s) used and purpose of the algorithm(s); the potential for inferring additional information about individuals from the aggregation of multiple data sources; and importantly, the measures that will be used to address these risks.
  - o DHS and the FBI should require that SORNs provide more specificity in describing types of data collected, data sources and the connections between data sources, and who will use such data and why.
- **DHS and the FBI should take steps to increase public transparency about the AI systems they employ.**
  - o DHS has recently started an effort to improve transparency, and those efforts should be prioritized and assessed as they are implemented.[2]
  - o The FBI should implement similar reforms to improve transparency.
- **DHS and the FBI should make their websites easier for the public to navigate and ensure the websites are regularly updated.** Privacy, Civil Liberties, and Civil Rights Risk and Impact Assessment Reports, related semiannual reports, PIAs, and SORNs should be located in a central place; have clearly marked dates next to the title, and chronologically ordered, and published in a timely manner. DHS and the FBI should seek public comments annually about the navigability of their websites and potential improvements.

Recommendation

Recommendation Set 2: Develop & Test Systems per Goals of Privacy Preservation and Fairness

Actions for the President:

- **Through Executive Order, the President should require the Director of National Intelligence, the Secretary of Homeland Security, and the Director of the FBI to take the following actions:**

- **Implement steps to mitigate privacy, civil liberties, and civil rights risks associated with any AI system on an iterative basis and require documentation of all accepted risks.**
    - In implementing steps to achieve this objective, the Commission recommends that ODNI, DHS and the FBI adopt practices from the Key Considerations. In particular:
        - Use privacy protections such as robust anonymization that can withstand sophisticated reidentification attacks, and when possible, privacy-preserving technology such as differential privacy, federated learning, and machine learning (ML) with encryption of data and models.[3]
        - Mitigate bias in development and testing. For development, conduct stakeholder engagement to establish consensus on the definition of fairness metrics to be used for the specific development and identify necessary constraints on system behavior to protect civil rights and avoid inequitable outcomes.[4] In testing, confirm that identified constraints are enforced.[5] Testing to expose unintended bias should include testing for and documentation of different types of error rates (e.g., differences in false positive or false negative rates) or disparate outcomes across demographics.[6]
        - Use AI-tools to support assessing fairness (e.g., industry tools cited in the *Key Considerations*).[7]
        - Ensure the MLOps toolchains include routine calibration of agreed-upon fairness metrics throughout continuous development and integration.[8]
        - Assess model performance and system impact during fielding on an ongoing basis, including emergent behavior, to ensure compliance with privacy, civil rights, and civil liberties objectives.[9]
- **Designate an office, committee, or team in each agency to conduct a pre-deployment review of AI technologies that will impact privacy, civil liberties, and civil rights, including relevant documentation.**
    - This should include review in advance of their deployment and for compliance over the life span of the system.[10] An office in each Intelligence Community agency, DHS, and the FBI should be equipped to assess data, model, and system documentation, and testing results of technologies per their intended use.
    - In undertaking this review, the Commission recommends the designated office use the *Key Considerations*.

Actions for Congress:

- **Establish third-party testing center(s) to allow independent, third-party testing of national-security-related AI systems that could impact U.S. persons.**
    - Congress should fund NIST to create a Third-Party AI Testing Lab program under the NIST National Voluntary Laboratory Accreditation Program.[11]

- o The third-party test mechanism's mandate should be to:
    - ▪ Tailor metric assessment per agency mission and authorities;
    - ▪ Develop an approach for testing both software products that can be installed in a test facility and cloud-based services;
    - ▪ Establish binding data dissemination agreements with stakeholders of the system to be tested (e.g., the agency requesting testing and relevant vendors and data owners);
    - ▪ Collaborate with the agency seeking testing to reach consensus on how to handle the test data provided and the test results and analyses.[12]
- o Third-party test center(s) should allow government vendors to share proprietary data without fear of it being exposed to competitors; and offer the benefits of an aggregated view of performance across a sector or collection of corporations and aggregated best practices.
- o Third-party test center(s) should be used by agencies prior to procuring or fielding high-consequence systems that impact U.S. persons, and use should be considered to overcome in-house testing limitations.

- **Require the Department of Justice (DOJ), in consultation with the Privacy and Civil Liberties Oversight Board (PCLOB), to develop binding guidance for the use of third-party testing (e.g., thresholds for high-consequence systems or unprecedented factors) of AI systems.[13]**
    - o This should include criteria for when an AI system may pose high enough risk for privacy, civil liberties, and civil rights that it would trigger a testing requirement by a third party. In forming such guidance, PCLOB and the DOJ should consult with PCL Officers in federal agencies.

Acknowledgment of continued work for the technical community and legal experts.

There are significant unresolved tensions between various technical approaches to preserving civil rights and civil liberties and current and anticipated legal frameworks. For example, scholars have expressed concern "that technical and legal approaches to mitigating bias will diverge so much that laws prohibiting algorithmic bias will fail in practice to weed out biased algorithms and technical methods designed to address algorithmic bias will be deemed illegally discriminatory."[14] Continued work in the technical, legal, and policy domains is required to find a consensus balance that addresses technical approaches to preserving privacy, civil liberties, and civil rights and evolving policy.

Recommendation

Recommendation Set 3: Strengthen the ability of those aggrieved by AI to seek redress and have due process.

Actions for FBI and DHS:

- **The FBI and DHS should each conduct a review of its respective policies and practices related to AI technology to ensure that parties aggrieved by government**

action involving the use of AI, including through system actions or misuse, can seek redress and clearly know how to do so. At least annually, the FBI and DHS shall assess if updates or changes are required to their respective reviews.

- o This review should determine whether notice of AI use in decision-making is adequately provided to aggrieved parties to enable redress, as well as the degree of auditability and interpretability needed to contest.
- o The FBI and/or DHS review team—which must include the Offices of Privacy and Civil Liberties—should submit recommendations to their respective agency heads for any regulatory and/or policy changes necessary to amend existing redress mechanisms to reflect issues raised by the use of an AI-enabled system.
- o The Attorney General, working with the Director of the FBI, and the Secretary of Homeland Security, respectively, should direct appropriate actions to ensure that each agency:
    - ▪ provides adequate redress, based on the recommendations of the review; and
    - ▪ provides the public with clear, updated guidance on how to seek redress in situations covered by the review, including by posting relevant information on their websites.

Actions for the Attorney General:

- **Issue federal guidance on AI and due process. This guidance should describe how relevant agencies should safeguard the due process rights of U.S. persons when AI use may lead to a deprivation of life or liberty.** This should include what obligations agencies have to disclose on its use of AI[15] to a criminal defendant in a timely manner prior to trial or hearing (i.e., notice obligations), including the role that AI played leading to an arrest, charge, or criminal prosecution. Such guidance should be incorporated into agency operational guidelines.

Acknowledgment of continued work by the judicial and/or legislative branches:

The above actions should ensure that agencies receive clear guidance on AI-related redress and due process[16] in the interim as Congress and/or the courts weigh in on federal requirements. Continued work will be needed to provide baseline guidance with the evolution of AI capabilities and their application,[17] and to address open questions on the federal rules of evidence and criminal procedure as they relate to AI.[18]

Recommendation

Recommendation Set 4: Strengthen Oversight and Governance Mechanisms to Address Current and Evolving Concerns

Actions for Congress:

- **Strengthen the Privacy and Civil Liberties Oversight Board's (PCLOB) ability to provide meaningful oversight and advice to the federal government's use of AI-enabled technologies for counterterrorism purposes.** To achieve this, Congress should provide for a targeted expansion of PCLOB's authorities and appropriations as follows:

- o *Awareness of AI programs.* As part of PCLOB's authority to access all relevant material from agencies, agencies should be required to provide PCLOB notice prior to the fielding or repurposing of an AI system, as well as any associated privacy, civil liberties, and civil rights impact assessments.
- o *Visibility into technology.* Agencies should be required to provide to PCLOB, upon PCLOB's request, specific information about technology used in any AI system, including: the data used for AI systems (e.g., documentation regarding the data collection processes for AI-enabled tools and programs, including disclosure and consent processes); models used (and supporting model documentation regarding training and testing); and model repurposing (beyond that context for which it was trained/approved).
- o *Resources and other organizational requirements.* PCLOB requires an increase to its resources, both in terms of talent and funding, to achieve its mission and manage its portfolio as AI adoption increases. PCLOB should provide Congress with a self-assessment of its resources and organizational structure given the expected increase of AI-related programs that fall under its current mandate and responsibilities.

- **Empower DHS Offices of Privacy and Civil Rights and Civil Liberties.** Congress should bolster the roles of DHS' Office of Privacy and Office of Civil Rights and Civil Liberties by requiring the Chief Civil Rights and Civil Liberties Officer, in coordination with the Privacy Officer, to play an integral role in the legal and approval processes for the procurement and use of AI-enabled systems, including associated data of machine learning systems in DHS. As part of this legislation, the Privacy and Civil Rights and Civil Liberties offices should report back to Congress concerning additional staffing or funding resources that are required to satisfy this mandate.

Action for the Secretary of Homeland Security:

- **Ensure the Privacy Officer and the CRCL Officer receive permanent seats in the new DHS enterprise-wide AI Coordination and Advisory Council.** Such appointments are needed in order to meaningfully satisfy the DHS AI Strategy objective titled, "Formalize AI Governance Processes at DHS."[19]

Actions for the President:

- **Through Executive Order, require stronger coordination and alignment among oversight and audit organizations through creation of an interagency working group focused on oversight and audit.** Voluntary compliance by agencies with AI documentation and testing requirements should be supported by rigorous, technically informed oversight. To achieve this and overcome current auditing impediments, a standing body (e.g., an interagency working group) should align and coordinate to enhance AI oversight and audit with respect to privacy, civil liberties, and civil rights. This includes system auditability such that the government can monitor and trace the steps that produced a system's output,[20] and auditing to ensure systems are not being misused.

- o *Composition:* Organizations should include the Department of Justice Intelligence Oversight Section; Office of the Inspector General of the Intelligence Community; the Government Accountability Office; the Privacy & Civil Liberties Oversight Board; Civil Liberties and Privacy Offices of national security agencies; the National Security Council, and the Office of Science & Technology Policy.
- o *Function:* The interagency working group should provide a forum for members to substantively and regularly address and share information. The working group should:
    - ▪ Develop an inventory of the types of AI-relevant oversight and audit currently performed by and anticipated by the participant organizations.
    - ▪ Develop an inventory of specific capabilities developed in each organization to address AI oversight and audit.
    - ▪ Assess available AI-enabled tools that can be adapted to support more effective and efficient oversight and audit.
        - Tools that support financial audit[21] and model risk management[22] are examples of advances in applying AI to improve the efficiency and scalability of audits that should be reviewed for adoption.
    - ▪ Identify priority investment requirements for each organization to address current needs.
    - ▪ Identify priority research topics for open S&T gaps in supporting AI oversight and audit, including research gaps in applications of AI in support of privacy and civil liberties (e.g., ML techniques for classification, recommendation, anomaly detection, and other applications)[23] and extending tools such as those that support financial audits and model risk management;
    - ▪ Recommend policy or legislative changes for specific authorities granted to the individual organizations.
    - ▪ Address mission and focus overlap among representative organizations.
    - ▪ Issue reports, at a minimum annually, on key oversight and audit activities as well as S&T gaps.

Action for the President or Congress:

- **Establish a task force to assess the privacy and civil rights and civil liberties implications of AI and emerging technologies.**

The goal of the task force would be to identify gaps and make recommendations to ensure that uses of AI and associated data in U.S. government operations comport with U.S. law and values, and to study organizational reforms that would support this goal. Specifically, it should assess existing policy and legal gaps for current AI applications and emerging technologies, and make recommendations for:

- legislative and regulatory reforms on the development and fielding of AI and emerging technologies;[24] and

- institutional changes to ensure sustained assessment and recurring guidance on privacy and civil liberties implications of AI applications and emerging technologies.

As mentioned in Chapter 8 of this report, the advancement of AI requires a forward-looking approach to oversight that anticipates the continued evolution and adoption of new technologies, and better positions the government to manage their employment responsibly well into the future. The Commission assesses that, to achieve this goal, a new task force is needed.

The Commission recommends that the President or Congress create a task force with the proposed membership, structure, function, and priorities identified below.

For expediency, the President should:

- **Issue an Executive Order that creates a task force charged with recommending reforms for AI governance and oversight.**
  - o *Membership and structure.* The President should create a task force in the Executive Office of the President to develop recommendations on ensuring adequate AI governance and oversight. The President should designate a senior official to lead the task force. Members should include the heads of OMB, NIST, PCLOB, and the GAO; and Chief Civil Liberties and Privacy Officers and Inspectors General of all national security agencies. In addition, the task force should include representatives from civil society (including organizational leaders with expertise in privacy, civil liberties, and civil rights), industry, and academia. The National AI Advisory Committee Subcommittee on AI and Law Enforcement should also be represented.[25]
  - o *Function.* The task force should be charged with the following responsibilities:
    - ▪ Conducting a macro assessment of the privacy and civil rights and civil liberties implications of the capabilities of AI and emerging technologies;
    - ▪ Making recommendations for legislative and regulatory reforms on the development and fielding of AI and emerging technologies, including associated data, in the following key areas:
      - *Privacy, Civil Liberties, and Civil Rights (P/CLCR) reporting.* Binding guidance on when the IC, DHS, and FBI should prepare and publish an AI Risk Assessment Report and AI Impact Assessments, specifically what constitutes a qualifying AI system or significant system refresh (as discussed in the first recommendation of Chapter 8 of this report).
      - *Biometric technologies.* This should include baseline standards for federal government use of biometric identification technologies, including but not limited to, facial recognition.
        - o To address the urgent need for baseline standards and safeguards regarding facial recognition, this should include assessing gaps in federal legislation, gathering input from agency stakeholders (and their legal counsel) currently using facial recognition for national security missions; privacy, civil liberties, and civil rights experts inside and outside of government, including PCLOB; and from the public at large in order to make facial recognition legislation recommendations.

        - o Beyond facial recognition, guidance will be needed regarding other biometric identification tools including voiceprints.

    - *Government procurement of commercial AI products.* This should include contractual best practices for ensuring industry AI products (including associated data) procured by the government uphold privacy, civil liberties, and civil rights expectations (including privacy, information security, fairness/non-discrimination, auditability, and accountability). This should include third-party requirements that should be incorporated into procurement terms that speak to responsible AI objectives, including for testing validation.[26] Consideration should be given to both government-off-the-shelf and commercial-off-the-shelf (COTS) procurement.[27]

    - *Data privacy and retention.* Updates to and reforms of government data privacy and retention requirements to address AI implications.

- Making recommendations for institutional changes to ensure sustained assessment and recurring guidance on privacy and civil liberties implications of AI applications and emerging technologies.

    - Evolving AI capabilities are poised to challenge existing expectations for privacy, civil liberties, and civil rights.[28] In light of this, the task force should assess the utility of a new entity within the federal government to regulate and provide government-wide oversight of AI use by the federal government.

    - In evaluating options for a new entity, the task force should consider the following:

        - o Authorities and resources necessary for the new entity to provide ongoing guidance and baseline standards for:

            - The federal government's development, acquisition, and fielding of AI technologies to ensure they comport with privacy, civil liberties, and civil rights law and values, and to include guardrails for their use and disallowed outcomes[29] to be incorporated in policy and embedded in system development; and

            - Transparency to oversight entities and the public regarding the Federal Government's use of AI systems and the performance of those systems.

        - o Existing interagency and intra-agency efforts to address AI oversight; and

        - o The unique needs of national security, law enforcement, and other government missions with respect to AI systems and potential implications for privacy, civil liberties, and civil rights, and civil liberties.

    - After considering the potential utility of a new organization, make recommendations on organizational placement and structure, composition, authorities, and resources needed.

- Assessing ongoing efforts to adapt regulation of the private sector's AI adoption,[30] and as appropriate, consider and recommend institutional or organizational changes to facilitate adequate regulation of commercial development and fielding of AI and associated data.

- o *Reporting.* The task force should issue a report to the President with its legislative and regulatory recommendations on a rolling basis, but no later than within 180 days of its creation. It should issue a report to the President with its recommendations for organizational changes within one year of its creation. The Commission recommends that the report be provided to Congress to ensure transparency and assist Congress in examining these critical issues.

- **In the alternative, Congress should mandate the existence of this task force as outlined above.**

Acknowledgment of continued work to update and clarify legal frameworks on key issues in data protection and data privacy:

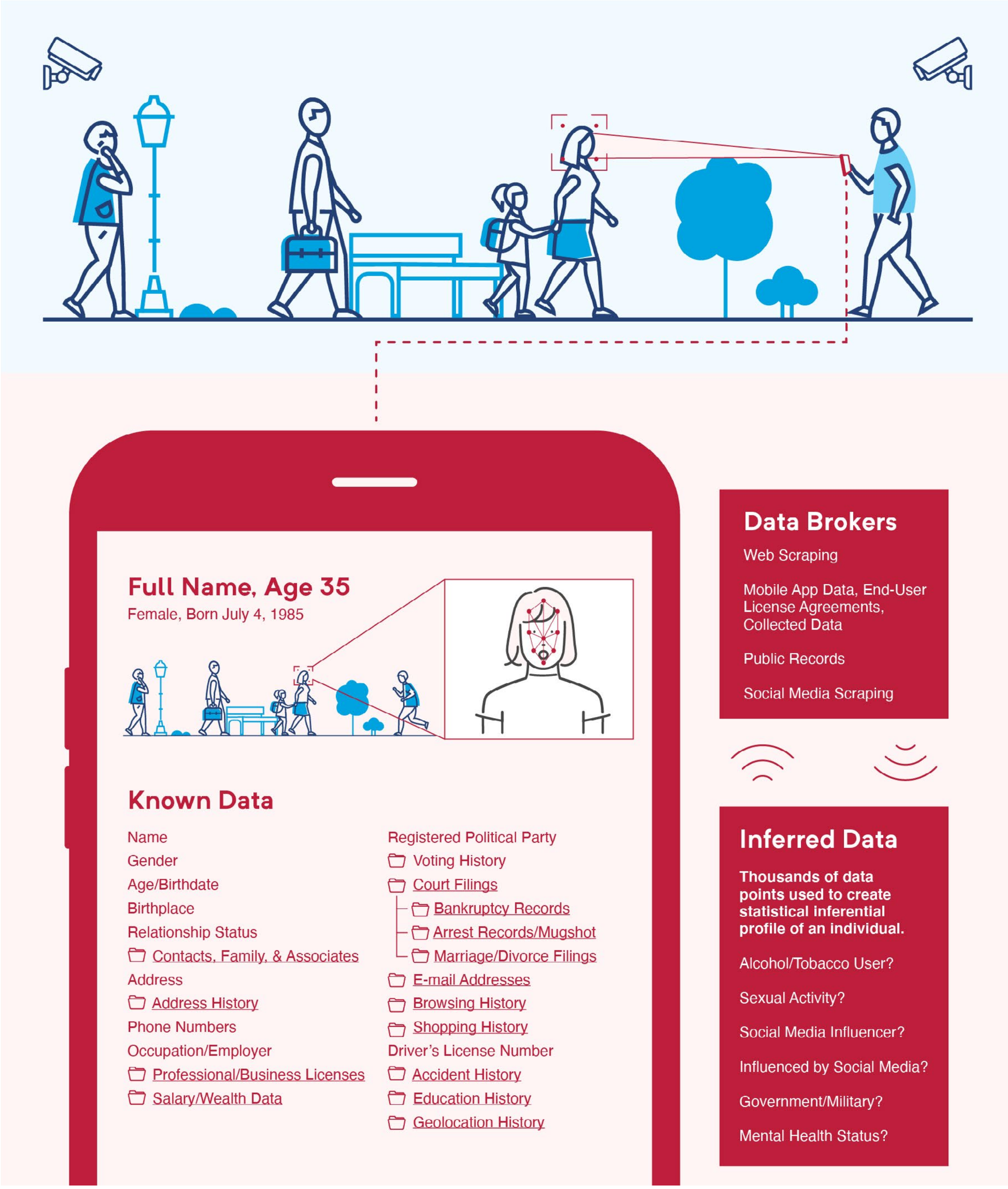
A comprehensive approach to upholding privacy and civil liberties in the AI era requires tackling several large, unresolved policy and legal questions regarding data protection and data privacy. Detailed recommendations on these issues would extend beyond the scope of this Commission's mandate, but we identify them here in order to urge further study and congressional action.

- *Legal concerns over federal use of third-party data.* Congress and/or the Judiciary should assess the adequacy of current legal constraints over the federal government's obtainment and use of third-party data, including data acquired from data brokers. Either through evolving case law or legislation, agencies would benefit from clarity surrounding the Fourth Amendment's application on third-party data.[31] In the meantime, agencies should provide transparency on their respective policies and legal basis for accessing and using commercial data.[32]

- *National data protection standards.* Data privacy policies and standards that apply to government agencies alone will be inadequate, and in some cases may harm national security interests.[33] This is particularly important considering how adversaries (both state and non-state actors) can access and use data collected about U.S. persons. As Congress considers proposals for national data security and privacy protection, it should ensure any future legislation addresses the issue of microtargeting. As noted in Chapter 1 of this report, AI systems will create new capabilities for state actors to target individuals with precision as well as numerous aspects of our society like cities, supply chains, universities, corporations, infrastructure, and financial transactions. Strong data privacy protections will be necessary for a multitude of reasons, including to shield the United States from this new phenomenon.

- *National framework for use of biometric technologies.* In the absence of federal legislation regulating the use of facial recognition, the existing patchwork of state and local laws and regulations creates a number of difficulties for government officials, industry, and the public. This has led to actions including: companies prohibiting the sale of facial recognition to law enforcement,[34] and local government bans on the use of facial recognition have emerged from coast to coast.[35] The lack of a consistent federal approach is also a liability for national security agencies when best practices are not used locally.[36] In developing regulation, it will be critical that policy and legislation account not only for facial recognition, but other types of biometric identification that, when combined with other AI technology, can introduce additional concerns.[37]

Unregulated and Legal Data Collection & Brokering for AI-enabled Predictions and Identification



Full Name, Age 35

Female, Born July 4, 1985

Known Data

- Name
- Gender
- Age/Birthdate
- Birthplace
- Relationship Status
- Contacts, Family, & Associates
- Address
- Address History
- Phone Numbers
- Occupation/Employer
- Professional/Business Licenses
- Salary/Wealth Data
- Registered Political Party
- Voting History
- Court Filings
  - Bankruptcy Records
  - Arrest Records/Mugshot
  - Marriage/Divorce Filings
- E-mail Addresses
- Browsing History
- Shopping History
- Driver's License Number
- Accident History
- Education History
- Geolocation History

Data Brokers

Web Scraping

Mobile App Data, End-User License Agreements, Collected Data

Public Records

Social Media Scraping

Inferred Data

Thousands of data points used to create statistical inferential profile of an individual.

Alcohol/Tobacco User?

Sexual Activity?

Social Media Influencer?

Influenced by Social Media?

Government/Military?

Mental Health Status?

Blueprint for Action: Chapter 8 - Endnotes

[1] For instance, a recent DHS IG report criticizes the DHS Privacy Office for not establishing controls to ensure that privacy compliance documentation is complete and submitted as required, and specifically noted DHS had not performed required periodic reviews for new and evolving privacy risks. DHS IG, DHS Privacy Office Needs to Improve Oversight of Department-wide Activities, Programs, and Initiatives, OIG-21-06, (Nov. 4, 2020). Civil society members have noted that PIAs and SORNs are often too opaque to be helpful, and that agencies sometimes try to shoehorn new data collections under older SORNs and PIAs. See *Comments of the Electronic Frontier Foundation Regarding System of Records Notices 09-90-2001,09-90-2002*, Electronic Frontier Foundation (Aug. 17, 2020), https://www.eff.org/files/2020/08/17/2020-08-17_-_eff_comments_re_hhs_regs_re_covid_data.pdf (criticizing two SORNs issued by the Department of Health and Human Services during the pandemic, as "overly vague in describing the categories of data collected, the data sources, and the proposed routine uses of the data").

[2] The Commission acknowledges DHS' steps to improve public records as noted in the DHS AI Strategy: "Future AI systems implemented by DHS will require a public release of system information with appropriate exceptions for certain sensitive military and intelligence systems, and some exceptions for law enforcement activities. DHS will produce a framework for releasing AI system information and a process for public comment." See *U.S. Department of Homeland Security Artificial Intelligence Strategy*, U.S. Department of Homeland Security at 14 (Dec. 3, 2020), https://www.dhs.gov/publication/us-department-homeland-security-artificial-intelligence-strategy.

[3] To support agencies in this goal, federal R&D investment should continue to advance the state of the art for preserving personal privacy. For information regarding the critical AI research areas the Commission recommends OSTP prioritize, see the Chapter 11 Blueprint for Action. Agencies should also assign responsibility for assessing the state of the practice and encouraging federated learning and anonymization pilots for government databases used in machine learning developments (e.g., to Chief Data Officers at each agency).

[4] Development practices should also include documenting trade-offs made, including optimizations that cause a trade-off in performance across fairness metrics.

[5] For instance, constraints about proxies for national origin or protected classes used for rules-based system predictions.

[6] These include: 1) Documenting operating thresholds including those that yield different true positive and false positive rates or different precision and recall across demographics; (2) Assessing the representativeness of data and model for the specific context at hand; (3) Using tools to probe for unwanted bias in data, inferences, and recommendations; (4) Testing for fairness and articulating the approach, performance, and metrics used. For an extensive list of practices, see the Appendix of this report containing the abridged version of NSCAI's *Key Considerations for Responsible Development & Fielding of AI*. For additional details on the Commission's recommendations to mitigate bias in development and testing, see the *Key Considerations for Responsible Development & Fielding of Artificial Intelligence: Extended Version*, NSCAI (2021) (on file with the Commission).

[7] Examples of tools available to assist in assessing and mitigating bias in systems relying on machine learning include Aequitas by the University of Chicago, Fairlearn by Microsoft, AI Fairness 360 by IBM, and PAIR and ML-fairness-gym by Google. Microsoft's AI Fairness checklist provides an example of an industry tool to support fairness assessments. See Michael A. Madaio et al., *Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI*, CHI 2020 (April 25-30, 2020), http://www.jennwv.com/papers/checklists.pdf.

[8] A widely used Industry example of a fairness metric is Equality of Opportunity (EEO), defined in *Machine Learning Glossary: Fairness*, Google Developers (Feb. 11, 2020), https://developers.google.com/machine-learning/glossary/fairness. Note that EEO is suited for some contexts and a poor fit for others—this is why careful deliberation of the operational metrics for fairness must be established early in the development process.

[9] Select practices include: 1) Assessing statistical results for performance over time to detect emergent bias; 2) recurrent testing and validation at strategic milestones, especially for new deployments and classes of tasks; and 3) Continuously monitoring AI system performance, including the use of high-fidelity traces to determine if a system is going outside of acceptable parameters (e.g., for fairness and privacy leakage) pre-deployment and in operation. For an extensive list of practices, see the Appendix of this report containing the abridged version of NSCAI's *Key Considerations for Responsible Development & Fielding of AI*. For additional details on the Commission's recommendation for maintenance and deployment, see the section on "System Performance" in *Key Considerations for Responsible Development & Fielding of Artificial Intelligence: Extended Version*, NSCAI (2021) (on file with the Commission).

[10] ML systems in particular require ongoing assessments of privacy and fairness assurances, including the specific definition of fairness being assumed.

[11] This requires the creation of an AI TEVV handbook, a culmination of applied research, to create the testing protocols that should be carried out by third-party testing lab(s) and the accreditation procedures by which labs can become certified.

[12] In some cases, exposure of test results could reveal weaknesses in a national security system that could be exploited by an adversary.

[13] As noted in *Ethical Considerations for Commercial Use of AI*, "rigorous testing is particularly important for high-risk applications, and standards should be established to determine the nature of those applications." See Frances Duffy, *Ethical Considerations for Use of Commercial AI*, Johns Hopkins Applied Physics Laboratory (Dec. 2020).

[14] Alice Xiang, *Reconciling Legal and Technical Approaches to Algorithmic Bias*, Tennessee Law Review at 7 (July 13, 2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3650635. See also Zachary Lipton, et al., *Does Mitigating ML's Impact Disparity Require Treatment Disparity?*, arXiv (Jan. 11, 2019), https://arxiv.org/abs/1711.07076. (Some approaches to mitigate disparate outcomes explicitly make use of membership in protected classes such as race or gender, and are demonstrably more equitable than comparable algorithms that are "blind" to protected classes.)

[15] Disclosure requirements should be specific to each application of AI. See Frances Duffy, *Ethical Considerations for Use of Commercial AI*, Johns Hopkins Applied Physics Laboratory at 31 (December 2020). ("Appropriate disclosure requirements should be created for the purposes of traceability in a court case or for the government's own internal use.")

[16] As noted in the *Key Considerations*, existing policies for contestability should be assessed and updated as needed to ensure accountability and to mitigate errors though feedback loops. See the Appendix of this report containing the abridged version of NSCAI's *Key Considerations for Responsible Development & Fielding of AI*. For additional details on the Commission's recommendation to adopt policies to strengthen accountability and governance, see the section on "Accountability and Governance" in *Key Considerations for Responsible Development & Fielding of Artificial Intelligence: Extended Version*, NSCAI (2021) (on file with the Commission).

[17] Due process rights require that individuals have the ability to meaningfully challenge a decision made against them. In federal criminal trials, this includes having the government's explanation of how an unfavorable decision was reached, so it can be contested. In cases where AI-assisted or AI-enabled decisions are made, certain AI techniques will be less conducive to due process. See Danielle Keats Citron, *Technological Due Process*, Washington University Law Review (2008), https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview; see also Ryan Calo & Danielle Keats Citron, *The Automated Administrative State: A Crisis of Legitimacy*, Emory Law Journal (April 3, 2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3553590.

[18] For instance, evidentiary standards for admitting AI evidence in court have yet to be developed and are not encompassed in current *Daubert* standards guidance.

[19] DHS's Artificial Intelligence Strategy, dated December 2020, includes the establishment of a DHS enterprise-wide AI Coordination and Advisory Council composed of internal subject matter experts to monitor and support the adoption of AI technology by DHS Components. See *U.S. Department of Homeland Security Artificial Intelligence Strategy*, U.S. Department of Homeland Security at 10 (Dec. 3, 2020), https://www.dhs.gov/publication/us-department-homeland-security-artificial-intelligence-strategy.

Blueprint for Action: Chapter 8 - Endnotes

[20] For issues relevant to AI system audits, see *Global Perspectives and Insights: The IIA's Artificial Intelligence Auditing Framework Part*, Institute of Internal Auditors (2018), https://na.theiia.org/periodicals/Public%20Documents/GPI-Artificial-Intelligence-Part-II.pdf.

[21] See e.g., Audit Map (last accessed Jan. 3, 2021), https://auditmap.ai/; *The Next Generation of Internal Auditing–Are You Ready?*, Protiviti (2018), https://www.protiviti.com/sites/default/files/united_states/insights/next-generation-internal-audit.pdf.

[22] See e.g., Bernhard Babel, et al., *Derisking Machine Learning and Artificial Intelligence*, McKinsey & Company (Feb. 19, 2019), https://www.mckinsey.com/business-functions/risk/our-insights/derisking-machine-learning-and-artificial-intelligence; Saqib Aziz & Michael Dowling, *Machine Learning and AI for Risk Management*, Disrupting Finance at 33-50 (Dec. 7, 2018), https://link.springer.com/chapter/10.1007/978-3-030-02330-0_3.

[23] Xuning (Mike) Tang & Yihua Astle, *The Impact of Deep Learning on Anomaly Detection*, Law.com (Aug. 10, 2020), https://www.law.com/legaltechnews/2020/08/10/the-impact-of-deep-learning-on-anomaly-detection/.

[24] Examples include baseline AI standards and policy guidance for biometric identification technologies; for government procurement of commercial AI products; and for federal data privacy standards.

[25] In the FY2021 NDAA, Congress directed the Secretary of Commerce, in consultation with other senior Executive branch officials, to establish the National AI Advisory Committee, including a Subcommittee on AI and Law Enforcement. The Subcommittee is tasked to "provide advice to the President on matters relating to the development of artificial intelligence relating to law enforcement." Pub. L. 116-283, sec. 5104 William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021, 134 Stat. 3388 (2021).

[26] These should seek to encourage contracts with companies that have transparent policies and practices in support of traceability and auditability and those that share information about how their technology works and how it performs in independent testing.

[27] "Federal government acquisition regulations require that agencies procure software commercially off-the-shelf whenever possible, due to their cost effectiveness. Only when no comparable systems exist are agencies permitted to develop government off-the-shelf solutions." See Frances Duffy, *Ethical Considerations for Use of Commercial AI*, Johns Hopkins Applied Physics Laboratory at S-1 (December 2020). As standards and requirements for system development and testing evolve, it may be helpful for the government to "establish and maintain a list of COTS AI technologies that have been vetted and approved for micro-purchasing, based on their consistency with government security and testing standards, as well as their transparency." This could facilitate both rapid procurement and proper assessment of a vendor's consistency with Responsible AI practices. See Frances Duffy, *Supplement to Ethical Considerations for Commercial Use of AI: Implications of Acquisition Scale*, Johns Hopkins Applied Physics Laboratory (forthcoming).

[28] For example, policymakers and legislators will need to direct future attention to policies to preserve PCL as technological capabilities for ubiquitous sensing grow, e.g., in smart cities. In the future, ubiquitous sensing may make it impossible to distinguish U.S. persons' data versus non-U.S. persons' data for AI analytics. Another example for continued consideration includes the role of AI in filtering to remove U.S. persons' information from bulk data and conversely using AI to reveal such information, as minimization and de-minimization guidance may evolve based on AI efficacy relative to the status quo.

[29] Disallowed outcomes and guidance will need to be updated over time as community norms and technical capabilities change.

[30] See, for example, *Remarks of Commissioner Rebecca Kelly Slaughter: Algorithms and Economic Justice*, FTC (Jan. 24, 2020) https://www.ftc.gov/system/files/documents/public_statements/1564883/remarks_of_commissioner_rebecca_kelly_slaughter_on_algorithmic_and_economic_justice_01-24-2020.pdf; *Artificial Intelligence and Machine Learning in Software as a Medical Device*, U.S. Food and Drug Administration (January 2021), https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.

[31] See Byron Tau, *Homeland Security Watchdog to Probe Department's Use of Phone Location Data*, Wall Street Journal (Dec. 2, 2020), https://www.wsj.com/articles/homeland-security-watchdog-to-probe-departments-use-of-phone-location-data-11606910402 (reporting that "DHS's general counsel began examining [the agency's use of location tracking data] after concerns were raised by several offices within the department that use of the technology wasn't compatible with [Carpenter]," and that the DHS IG planned to investigate the matter).

[32] In ODNI Director Avril D. Haines' confirmation hearing, she was asked about the IC's use of commercially available location data. She testified that she would "try to publicize, essentially, a framework that helps people understand the circumstances under which we do that and the legal basis that we do that under. . . I think that's part of what's critical to promoting transparency generally so that people have an understanding of the guidelines under which the intelligence community operates." Charlie Savage, *Intelligence Analysts Use U.S. Smartphone Location Data Without Warrants, Memo Says*, New York Times (Jan. 22, 2021), https://www.nytimes.com/2021/01/22/us/politics/dia-surveillance-data.html.

[33] Investigative reporting and opinion pieces have underscored the national security threats involved with smartphone location data. Charlie Warzel & Stuart A. Thompson, *They Stormed the Capitol. Their Apps Tracked Them*, New York Times (Feb. 5, 2021), https://www.nytimes.com/2021/02/05/opinion/capitol-attack-cellphone-data.html?referringSource=articleShare; Stuart A. Thompson & Charlie Warzel, *How to Track President Trump*, (Dec. 20, 2019), https://www.nytimes.com/interactive/2019/12/20/opinion/location-data-national-security.html; Stuart A. Thompson & Charlie Warzel, *Twelve Million Phones, One Dataset, Zero Privacy*, New York Times (Dec. 19, 2019), https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.

[34] See Larry Magid, *IBM, Microsoft And Amazon Not Letting Police Use Their Facial Recognition Technology* (June 12, 2020), https://www.forbes.com/sites/larrymagid/2020/06/12/ibm-microsoft-and-amazon-not-letting-police-use-their-facial-recognition-technology/?sh=34b473dc1887; Asa Fitch, *Microsoft Pledges Not to Sell Facial-Recognition Tools to Police Absent National Rules*, Wall Street Journal (June 11, 2020), https://www.wsj.com/articles/microsoft-pledges-not-tosell-facial-recognition-technology-to-police-absent-national-rules-11591895282.

[35] See *Ban Facial Recognition*, Fight for the Future (last accessed Feb. 4, 2021), https://www.banfacialrecognition.com/map/.

[36] The Department of Defense, the Drug Enforcement Administration, Immigrations and Customs Enforcement, the Internal Revenue Service, the Social Security Administration, the U.S. Air Force Office of Special Investigations, and the U.S. Marshals Service have all had access to one or more state or local face recognition systems. See Clare Garvie, et al., *The Perpetual Line-up: Unregulated Police Face Recognition in America*, Georgetown Law Center on Privacy & Technology (Oct. 18, 2016), https://www.perpetuallineup.org/.

[37] Such types of identification aided by AI include voice recognition and gait detection. An example of additional risks includes when biometric identification is coupled with other advancing capabilities; for instance, for identity recognition or for emotion recognition. See *Emotional Entanglement: China's Emotion Recognition Market and its Implications for Human Rights*, Article 19 (January 2021), https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf. See also Drew Harwell & Eva Dou, *Huawei Tested AI Software that Could Recognize Uighur Minorities and Alert Police, Report Says*, Washington Post (Dec. 8, 2020), https://www.washingtonpost.com/technology/2020/12/08/huawei-tested-ai-software-that-could-recognize-uighur-minorities-alert-police-report-says/; Parmy Olson, *The Quiet Growth of Race Detection Software Sparks Concerns Over Bias*, Wall Street Journal (Aug. 14, 2020), https://www.wsj.com/articles/the-quiet-growth-of-race-detection-software-sparks-concerns-over-bias-11597378154.