Billing Code 3110-05

OFFICE OF MANAGEMENT AND BUDGET

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

AGENCY: Office of Management and Budget (OMB), Executive Office of the President.

ACTION: Notice of request for information: Identifying Priority Access or Quality Improvements for Federal data and models for Artificial Intelligence Research and Development (R&D), and Testing.

SUMMARY: Under the *Executive Order on Maintaining American Leadership in Artificial Intelligence* (Section 5.a.i), https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/, the Office of Management and Budget is inviting the public to identify needs for additional access to, or improvements in the quality of, Federal data and models that would improve the Nation's artificial intelligence (AI) research and development (R&D) and testing efforts.

DATES: [INSERT DATE OF PUBLICATION IN THE *FEDERAL REGISTER*].

ADDRESSES: Submissions are due 30 days from publication of this notice through www.regulations.gov.

SUPPLEMENTARY INFORMATION:

The National Artificial Intelligence Research and Development Strategic Plan discusses fundamental challenges, novel ideas for human and AI collaboration, and creating AI that is

more trustworthy (e.g., AI techniques that address challenges of bias and fairness, transparency and explainability, and robustness, security, and safety).

Beyond fundamental advances in AI, open challenges also include the application of AI to key domains, such as those highlighted on ai.gov that include:

- Transportation

- Healthcare

- Manufacturing

- Financial Services

- Agriculture

- Weather Forecasting

- National Security & Defense

Depending on the R&D goal and application domain, different data sets and models may be needed to accelerate AI advances. Additionally, the use of these data sets and models could stimulate new developments that would enhance the transparency and explainability of the AI application, and illuminate ways to ensure the robustness, security and safety of AI applications.

Over the years, a number of data sets have already been made available via data.gov. Some of these datasets are fully publically available, while others have restricted use (see restricted use

data sets). However, these data sets may or may not be useful or suitable for AI R&D and testing.

The following lists of topics cover the major areas for which information is sought. These lists are not intended to limit the topics that may be addressed by respondents, who may provide information about any topic that would inform the objective of this action.

To developing requests for additional accesses for data and models to improve AI R&D and testing, input to the following questions is sought:

- What Federal data and models are you seeking to use that are available to the public with no use restrictions, but which have technical issues inhibiting data access? Specifically, what are the technical issues (e.g., is it too big to be downloaded, is it not optimally formatted)? What types of AI R&D and testing would be accelerated with increased access to this data?

- What Federal data and models are you seeking to use that are restricted to the public, i.e., the data asset is available under certain use restrictions? What types of AI R&D and testing would be accelerated with increased access to this data?

- What Federal data and models are you seeking to use that are private and not at all available to the public? Describe the agency that has the data and what, if any, attempts you are aware of that have been made to increase access to the data or model. What types of AI R&D and testing would be accelerated with increased access to this data?

- What are key gaps in data and model availability that are slowing progress in AI R&D and testing? Which areas of AI R&D and testing are most impacted?

In developing requests for quality improvements to accessible data and models to improve AI R&D and testing, input to the following questions is sought:

- As agencies review their data and models, what are the most important characteristics they should consider? Stated differently, what characteristics of data sets or models make them well-suited for AI R&D?

- Which models are most important for agencies to focus on, and why?

- What characteristics should the Federal Government consider to increase a data set or model's utility for AI R&D (e.g., documentation, provenance, metadata)?

- What data ownership, intellectual property, or data sharing considerations should be included in federally-funded agreements (including, but not limited to, federal contracts and grants) that results in production of data for R&D?

- What research questions and applications are you trying to solve with AI, that require specific types and/or quantities of Federal data and models, and how might the Federal Government reduce barriers to discovery and access?

- Accelerating the application of AI can be enabled with pre-trained models (e.g., ResNet trained on ImageNet) that facilitate transfer learning. What research questions and applications would benefit most from the transfer learning?

Respondents to this RFI may define "data", "data set", and "model" as they desire, indicating clearly what they mean when using the term.

Instructions for Written Responses

Interested parties should provide written responses to the questions outlined in the Purpose of This Federal Register notice section. Submissions are due 30 days from publication of this notice through www.regulations.gov.

Please include the below in your response, *limiting this portion of your response to one page:*

- The name of the individual(s) and/or organization responding. Anonymous responses will also be accepted.

- A brief description of the responding individual(s) or organization's mission and/or areas of expertise, if the responder feels appropriate.

- A contact for questions or other follow-up on your response if desired.

Comments submitted in response to this notice are subject to FOIA. OMB may also make all comments available to the public. For this reason, please do not include in your comments information of a confidential nature, such as sensitive personal information or proprietary information. If you send an email comment, your email address will be automatically captured and included as part of the comment that is placed in the public docket and made available on the internet. Please note that responses to this public comment request containing any routine notice about the confidentiality of the communication will be treated as public comments that may be made available to the public notwithstanding the inclusion of the routine notice.

Next Steps

For background information on the areas of AI activity in the Federal government, please visit: AI.gov.

Russell T. Vought,
Acting Director, OMB.
[FR Doc. 2019-14618 Filed: 7/9/2019 8:45 am; Publication Date: 7/10/2019]