

Department of Legislative Services
Maryland General Assembly
2025 Session

FISCAL AND POLICY NOTE
First Reader

House Bill 823
Economic Matters

(Delegate Wu, *et al.*)

Generative Artificial Intelligence - Training Data Transparency

This bill requires, beginning January 1, 2026, that developers that release or substantially modify a generative artificial intelligence (AI) system post on the developer’s website specified information detailing the data and datasets used to train the generative AI system.

Fiscal Summary

State Effect: The bill is not anticipated to affect State governmental operations or finances, as discussed below.

Local Effect: The bill is not anticipated to affect local governmental operations or finances, as discussed below.

Small Business Effect: Potential meaningful.

Analysis

Bill Summary:

Relevant Definitions

“Generative AI” means AI that can generate derived synthetic content, such as text, images, video, and audio, that emulates the structure and characteristics of the data used to train the AI.

“Developer” means a person or a unit of State or local government that designs, codes, produces, or substantially modifies a generative AI system.

“Substantially modifies” means to release a new version of or update to a generative AI system that materially changes the functionality, performance, or training data of a generative AI intelligence system.

Applicability

The bill’s requirements apply to a generative AI intelligence system that was released on or after January 1, 2022, for use by the general public in the State, regardless of whether the terms of that use require compensation.

The bill does not apply to a generative AI system that is exclusively made for (1) use by hospital medical staff; (2) use by an affiliate of the developer of the generative AI system; (3) ensuring physical safety of an individual; (4) protecting confidential personal information; (5) detecting, resisting, or prosecuting malicious, deceptive, fraudulent, or illegal actions; (6) the operation of aircraft in the national airspace; (7) use by a federal entity to support national security, military, or other defense purposes.

Information to Be Posted

By January 1, 2026, and before each time a developer releases or substantially modifies a generative AI system, the developer must post on the developer’s website documentation detailing the data and datasets used to train the generative AI system, including:

- the sources or owners of the data;
- a description of how the data furthers the intended purpose of the generative AI system;
- specified data if a dataset is used for training;
- specified data about the labels that may or may not be used by the generative AI system;
- whether any data are in the public domain or are protected by copyright, trademark, or patent;
- whether the developer purchased or licensed the data;
- whether the data include any personal information;
- whether the data include aggregate consumer information;
- whether the developer cleaned, processed, or modified that data and why the developer cleaned, processed, or modified the data;
- the time period during which the data were collected and whether the data collection is still ongoing;
- whether the data were first used during the development of the generative AI system; and
- whether the generative AI system used or uses synthetic data generation.

Current Law: Chapter 496 of 2024 expanded the responsibilities of the Secretary of Information Technology and Department of Information Technology (DoIT) as they relate to the procurement and use of AI by State agencies and codified the Governor’s AI Subcabinet that was established by [Executive Order 01.01.2024.02](#). Broadly speaking and among other things, the Act:

- requires DoIT to adopt policies and procedures, in consultation with the Governor’s AI Subcabinet, concerning the development, procurement, deployment, use, and ongoing assessment of systems that employ high-risk AI by a unit of State government;
- prohibits units of State government from procuring or deploying a new system that employs AI unless the system complies with the policies and procedures adopted by DoIT;
- requires each unit of State government to conduct a data inventory to identify data that meets criteria established by the Chief Data Officer and that is (1) necessary for the operations of the unit or otherwise required to be collected as a condition to receive federal funds or by federal or State law; and (2) in a form prescribed by the Chief Data Officer, including when the data is used in AI; and
- requires each unit of State government to conduct an inventory of systems that employ high-risk AI.

State/Local Expenditures: Although the Department of Legislative Services (DLS) is aware of various State agencies researching issues related to AI, including DoIT and the State’s public universities, DLS is unaware of any State or local agency that is actively developing a generative AI system that is intended to be publicly released. As such, the bill is not anticipated to affect State or local operations or finances. Nevertheless, a State or local agency that develops a generative AI system in the future may incur substantial costs to collect, organize, and report the required information in the manner required by the bill.

Small Business Effect: DoIT advises, and DLS concurs, that the bill’s requirement to post the required information each time a generative AI system is “substantially modified” can be interpreted to require substantial ongoing effort and costs for developers, many of which begin as small start-up companies. Generative AI systems evolve constantly by receiving new data, documents, and other information on which to base the content they generate. Based on the definition of “substantially modified” established by the bill, each additional new piece of information could be interpreted to have materially changed the AI’s training data. As a result, the developer (which could also include the State or a local government in the future) would, under this interpretation, likely need permanent staff to regularly update the information provided on the website to meet the bill’s requirements.

Additional Information

Recent Prior Introductions: Similar legislation has not been introduced within the last three years.

Designated Cross File: None.

Information Source(s): Department of Information Technology; Maryland Department of Emergency Management; Maryland Municipal League; Maryland State Department of Education; University System of Maryland; Department of Budget and Management; Maryland Department of Health; Maryland Department of Labor; Maryland Department of Planning; Maryland Department of Transportation; Department of Legislative Services

Fiscal Note History: First Reader - February 14, 2025
km/mcr

Analysis by: Richard L. Duncan

Direct Inquiries to:
(410) 946-5510
(301) 970-5510