

NASA's AI-Generated Metadata Tags — Machine-Learning Auto-Tagging for Federal Open Data Discovery

AI in the Public Sector · Good practice · United States of America

Quality score: 80/100

Evidence strength: Strong

Summary

Since around 2019, NASA's Scientific and Technical Information Program has used a machine-learning model — trained on 3.5 million tagged documents — to auto-tag NASA's open data against a taxonomy of roughly 20,000 standardized keywords, and open-sourced the tooling.

Key dimensions

Dimension	Score
Evidence of impact / measured public value	2/3
Transparency, fairness & accountability	3/3
Transferability / demonstrated replication	3/3
Scalability beyond pilot	2/3
Governance, capability & sustainability	2/3

Evaluated by C-NAPSE

Data sources

[C-NAPSE web research](#) — accessed 2026-07-25

[NASA Scientific and Technical Information \(STI\) Program / NASA OCIO Data Analytics Team](#) — accessed 2026-07-25

[Federal Data Strategy case study](#) — accessed 2026-07-25

[NASA concept-tagging-api \(GitHub\)](#) — accessed 2026-07-25

Cite this — Evidoria. *NASA's AI-Generated Metadata Tags — Machine-Learning Auto-Tagging for Federal Open Data Discovery*. Persistent ID: 101d2614-30cc-4196-a1c4-b916819bcc2.