



**AI-Ready Federal Statistical Data:  
An Extension of Communicating Data Quality**

**FCSM 25-03**

**May 2025**

**Recommended Citation:** Travis Hoppe, Jeffrey Gonzalez, Lisa Mirel, and Rolf Schmitt. 2025. *AI-Ready Federal Statistical Data: An Extension of Communicating Data Quality*, FCSM 25-03. Federal Committee on Statistical Methodology. May 2025.

# Federal Committee on Statistical Methodology

## 2025 FCSM Members

**E. Ann Carson**

Bureau of Justice Statistics

**Chris Chapman**

National Center for Education Statistics

**Jennifer Hunter Childs**

U.S. Census Bureau

**Elise Christopher**

National Center for Education Statistics

**Morgan Earp**

National Center for Health Statistics

**Tala Fakhouri**

U.S. Food and Drug Administration

**John Finamore**

National Center for Science and Engineering Statistics

**Ellen Galantucci**

U.S. Department of Veterans Affairs

**Jeffrey Gonzalez**

Bureau of Labor Statistics

**Michael Hawes**

U.S. Census Bureau

**Travis Hoppe**

National Center for Health Statistics

**Brandon Kopp**

Bureau of Labor Statistics

**Carla Medalia**

U.S. Census Bureau

**Lisa B. Mirel**

National Center for Science and Engineering Statistics

**Jennifer Nielsen**

National Center for Education Statistics

**Jennifer M. Ortman**

U.S. Census Bureau

**Mark Prell (Co-Chair)**

Economic Research Service

**Michael Ratcliffe**

U.S. Census Bureau

**Rolf Schmitt**

Bureau of Transportation Statistics

**Bob Sivinski (Chair)**

Office of Homeland Security Statistics

**Matt Spence**

U.S. Census Bureau

**Ed Strocko**

Bureau of Transportation Statistics

**Jennifer Truman**

Bureau of Justice Statistics

**G. David Williamson**

Centers for Disease Control and Prevention

**Linda Young**

National Agricultural Statistics Service

**Lori Zehr (project manager)**

U.S. Census Bureau

## AI-Ready Federal Statistical Data: An Extension of Communicating Data Quality

### Vision

Statistical Application Programming Interfaces (APIs) enhanced with metadata and context to support accurate and trusted generative AI results for large language models.

### Challenge & Opportunity

Generative Artificial Intelligence (AI) is redefining how people interact with public information and shaping how public data are consumed. Recent advances in large language models (LLMs) mean that more Americans are getting answers from AI chatbots and other AI systems, which increasingly draw on public datasets. The federal statistical community can take action to advance the use of federal statistics with generative AI to ensure that official statistics are front-and-center, powering these AI-driven experiences.

The Federal Committee on Statistical Methodology (FCSM) developed the [Framework for Data Quality](#) to help analysts and the public assess fitness for use of data sets. AI-based queries present new challenges, and the framework should be enhanced to meet them. Generative AI acts as an intermediary in the consumption of public statistical information, extracting and combining data with logical strategies that differ from the thought processes and judgments of analysts. For statistical data to be accurately represented and trustworthy, they need to be machine understandable and be able to support models that measure data quality and provide contextual information.

FCSM is working to ensure that federal statistics used in these AI-driven interactions meet the data quality dimensions of the Framework including, but not limited to, accessibility, timeliness, accuracy, and credibility. We propose a new collaborative federal effort to establish best practices for optimizing APIs, metadata, and data accessibility to support accurate and trusted generative AI results.

### Approach

By enriching data access points with high-quality metadata and exposing them through standardized, AI-friendly APIs, agencies can significantly improve how LLMs discover, interpret, and relay publicly available federal statistics. The Department of Commerce's recent ["AI-ready data" initiative](#) exemplifies this approach. The guidance is explicitly aimed to *"improve accuracy of AI responses and increase the prioritization of its authoritative data in AI tools over nonauthoritative sources."* In practice, this could be interpreted as ensuring data elements and AI-friendly APIs are optimized for machine consumption. However, this interpretation only considers the format of the data and does not consider the meaning of the data.

One guideline from the [Commerce report](#) recommends adding *"comprehensive variable-level*

*metadata for machine understandability*” – in other words, giving AI detailed context about each data element. The [FCSM framework](#) provides three domains to evaluate these metadata improvements: Utility, Objectivity, and Integrity. Each data element should be assessed for data quality across these three domains and corresponding dimensions including accessibility, timeliness, and accuracy.

Creating AI-ready data builds on established open-data principles. Title II of Public Law No: 115-435 (the OPEN Government Data Act) already requires agencies to publish data in machine-readable formats *“while ensuring no semantic meaning is lost.”* Today’s AI era demands federal agencies go further – from **machine-readable to machine-understandable** data.

### **Impact & value**

OMB Statistical Policy Directive 4 and 44 USC 3563(1)(a) instructs agencies to strive for the widest, most accessible, and appropriate dissemination of its statistical products. In the age of AI assistants, this translates into making sure **AI systems can find the appropriate data and use the data correctly**. By carefully implementing improvements to metadata and APIs and benchmarking against the domains of the FCSM Framework, agencies can improve the data consumed through generative AI.

### **Vision/Roadmap for implementation/testing**

To do this, agencies can engage in **structured experiments** which build on the Commerce report to pave a path forward for making public data more machine understandable. The goal of these experiments is to empirically test how enhanced metadata and API standards impact AI-driven data access. Multiple experiments should be launched simultaneously and then results could be widely disseminated across the federal system with the goal of developing tools to assist agencies in enhancing AI-consumable metadata. Key components include:

1. **Baseline and follow-up assessments:** First, evaluate how current generative AI tools are using existing data. For example, perform test queries on a range of AI chatbots (e.g., asking for recent unemployment rates, population statistics, etc.) and observe the sources and accuracy of responses. This provides a baseline of AI performance with the status quo. Indeed, a cross-agency landscape analysis of “machine interpretability” of public data is [already being explored](#) through a National Secure Data Service (NSDS) project. Agencies could document where AI answers correctly, where it errs, and what might be causing confusion (e.g., ambiguous metadata, lack of API access, outdated data in the AI’s training set, poor models that hallucinate). The baseline assessment should include a high-level summary of how LLMs and other AI tools consume data (e.g. how variables are weighted) or prompt engineering strategies when compared to processes and judgments of analysts, particularly in the approach of using metadata. After implementation, agencies should plan on an assessment to evaluate the effectiveness of AI readiness practices.

2. **Implement API & Metadata Enhancements:** Next, select one or two high-value data products to improve. This might involve adopting or upgrading it to ensure that it returns comprehensive metadata (descriptions of variables, units, update timestamps, etc.). Use common schemas or standards like the [DCAT metadata standard](#) used in data.gov catalogs or [SDMX](#) so that AI developers have a consistent experience. Consider the domains and dimensions as outlined in the FCSM Framework for Data Quality to inform enhancements. Additionally, consider creating **AI-optimized data structures** – for example, simplified JSON responses that are easy to parse, or providing summary endpoints that an AI can call for quick context like “describe this dataset” or “give me the latest value”. The goal is to develop and document standards that, when incorporated, minimize the effort for any AI or human developer to interpret the data correctly by providing supporting information with each query.

As part of these pilots, agencies should explore new integration protocols designed for AI. One promising example are Model Context Protocols (MCPs), open-source standard defined by <https://github.com/modelcontextprotocol>. By setting up a compatible service by either wrapping around an existing API or as a gateway to your data, agencies can allow AI assistants to retrieve data in a controlled, standard way. Experiment with tools or sandbox environments where an AI agent (like a chatbot) uses the MCP to query the data – measure whether this leads to more relevant and accurate answers or reduced developer time

3. **Cross-Agency Learning:** Statistical and other federal agencies, working committees (including FCSM), state and local governments, academics and data innovators outside the government, will facilitate documenting approaches and sharing of lessons learned from these experiments. Part of this sharing may include success stories created by restructuring a dataset and providing an API or adding metadata tags that provide information about the data quality. A collaborative “AI-ready data” working group for Federal agencies under FCSM could periodically review progress, compare findings, and refine best practices. The outcome could be a practical **toolkit for transforming statistical data products into machine-understandable, AI-ready formats**, applicable across government. Notably, efforts are already underway to develop such toolkits and ranking methods for AI readiness of data and efforts will be made to engage with these efforts to incorporate successes that are currently underway.

The cycle of three steps can then be repeated to continually improve the APIs for machine understandability and other emerging applications. Once tested and incorporated into the use of public data, statistical agencies can then explore the application (and additional privacy considerations) related to making confidential data more machine understandable.