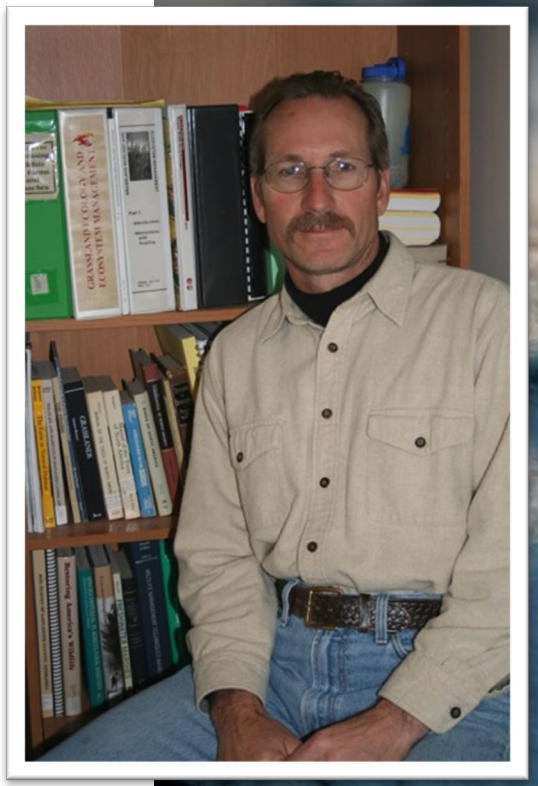


U.S. DEPARTMENT OF THE INTERIOR

Building capacity to communicate data quality

Actions at a non-statistical agency

OCTOBER 24, 2024

Jacob Malcom, Director, Office of Policy Analysis,
and Statistical Official

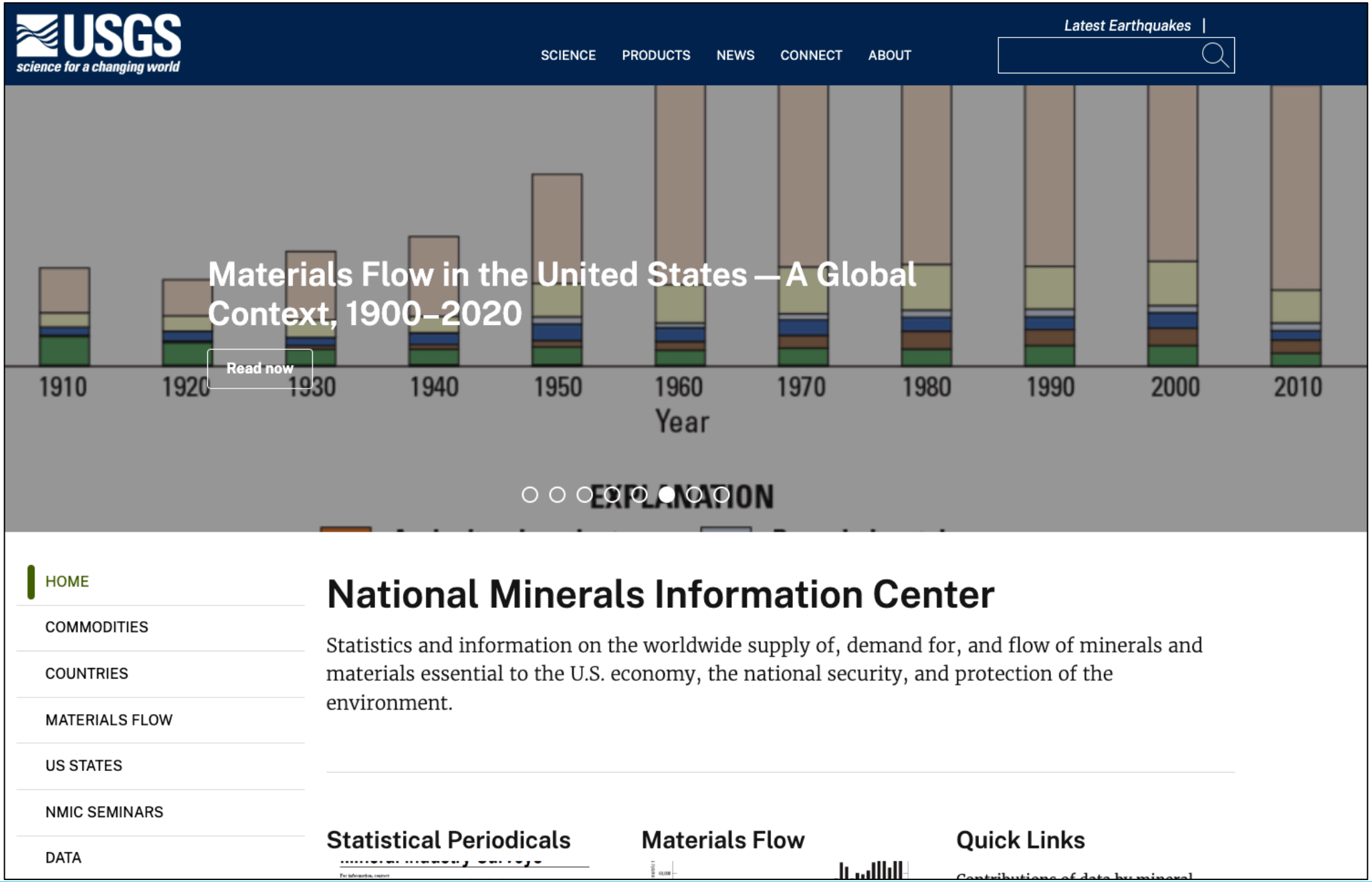
Hello from Interior

Relevant experience

- Field biologist with NPS and FWS, 1997-2008
- PhD/postdoc (ecological genomics, bioinformatics), 2008-2014
- Non-profit research and policy analysis, 2015-2022
- Department of the Interior, 2022-present
  - Director, Office of Policy Analysis
  - Statistical Official

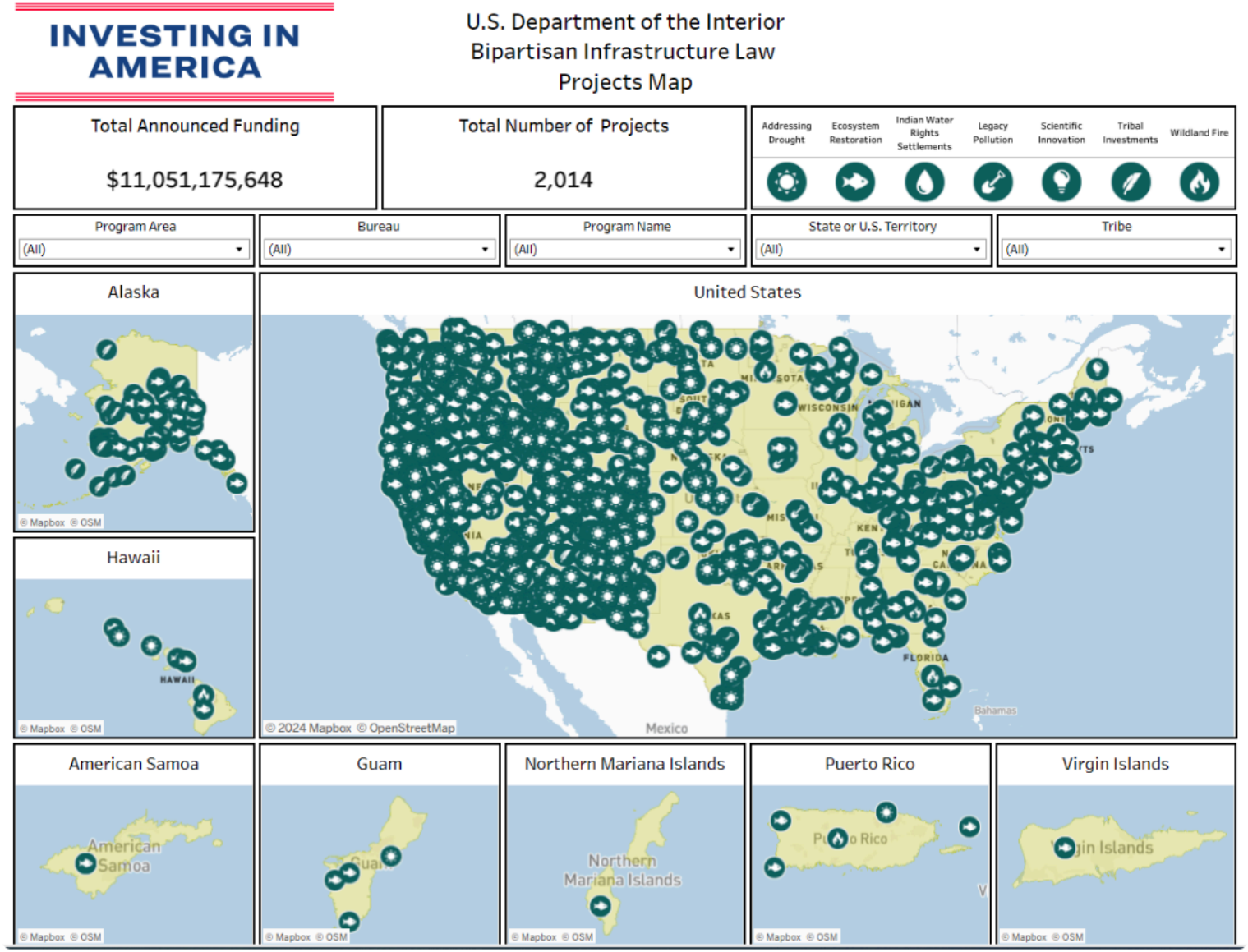
About Interior*

* The Department of Everything Else


USGS
science for a changing world
SCIENCE
PRODUCTS
NEWS
CONNECT
ABOUT
Latest Earthquakes |
Materials Flow in the United States — A Global Context, 1900–2020
Read now
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Year
EXPLANATION
HOME
COMMODITIES
COUNTRIES
MATERIALS FLOW
US STATES
NMIC SEMINARS
DATA
National Minerals Information Center
Statistics and information on the worldwide supply of, demand for, and flow of minerals and materials essential to the U.S. economy, the national security, and protection of the environment.
Statistical Periodicals
Materials Flow
Quick Links

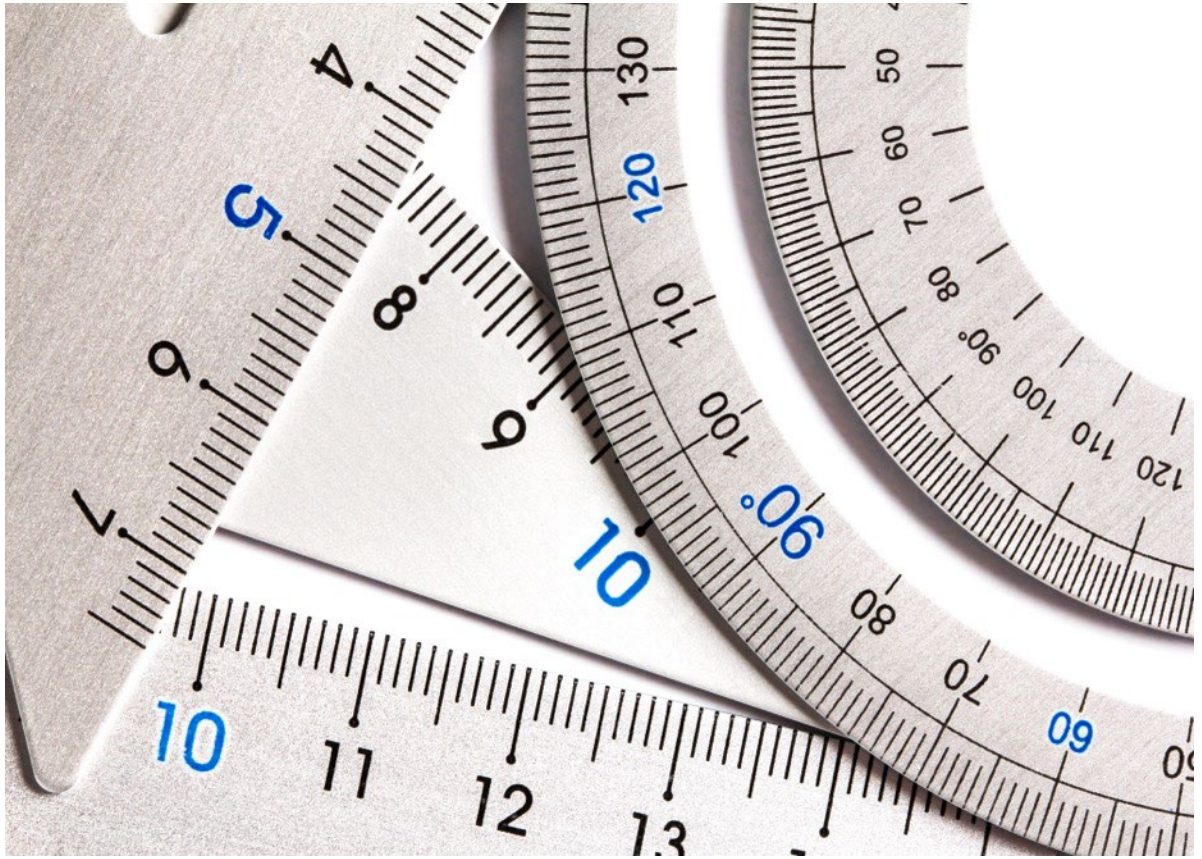
BIL & IRA

- $28.1B (BIL) + $6.4B (IRA)
- M-22-12 and direction to build capacity for evidence (and risk)



Is/are there a standardized method/s to measure success within programs across DOI?

- No.
- But we're working on it!



Foundational: Data quality

All data quality frameworks are wrong, but some are useful.

~Jacob Malcom

(Not George Box)

- FCSM Framework: useful!
- Big intro document
- Change management: meet people where they are
  - Brief guide with examples
  - Checklist

A Framework for Data Quality

FCSM-20-04

September 2020

A Brief Guide to Data Quality for Evidence-Building

U.S. Department of the Interior | Statistical Official | Version 1.0 | September 2024

Overview

Evidence about the effectiveness, efficiency, and equity of policies, programs, and offices must be built on high quality data. The Federal Committee on Statistical Methodology produced *A Framework for Data Quality* in 2020[1] to evaluate data quality in the context of statistical use and decision making. The Framework unites numerous lines of evidence and practice to ensure a strong, consistent approach to data quality (Figure 1). The *Framework* is built from the practices and requirements of Federal statistical agencies or statistical units. Although the Department of the Interior (Department) is not a Federal statistical agency and does not include any recognized statistical units, maximizing the utility, objectivity, and integrity of data is essential. The eleven quality dimensions are highly relevant to meeting this requirement and conducting rigorous evidence-based decision making.

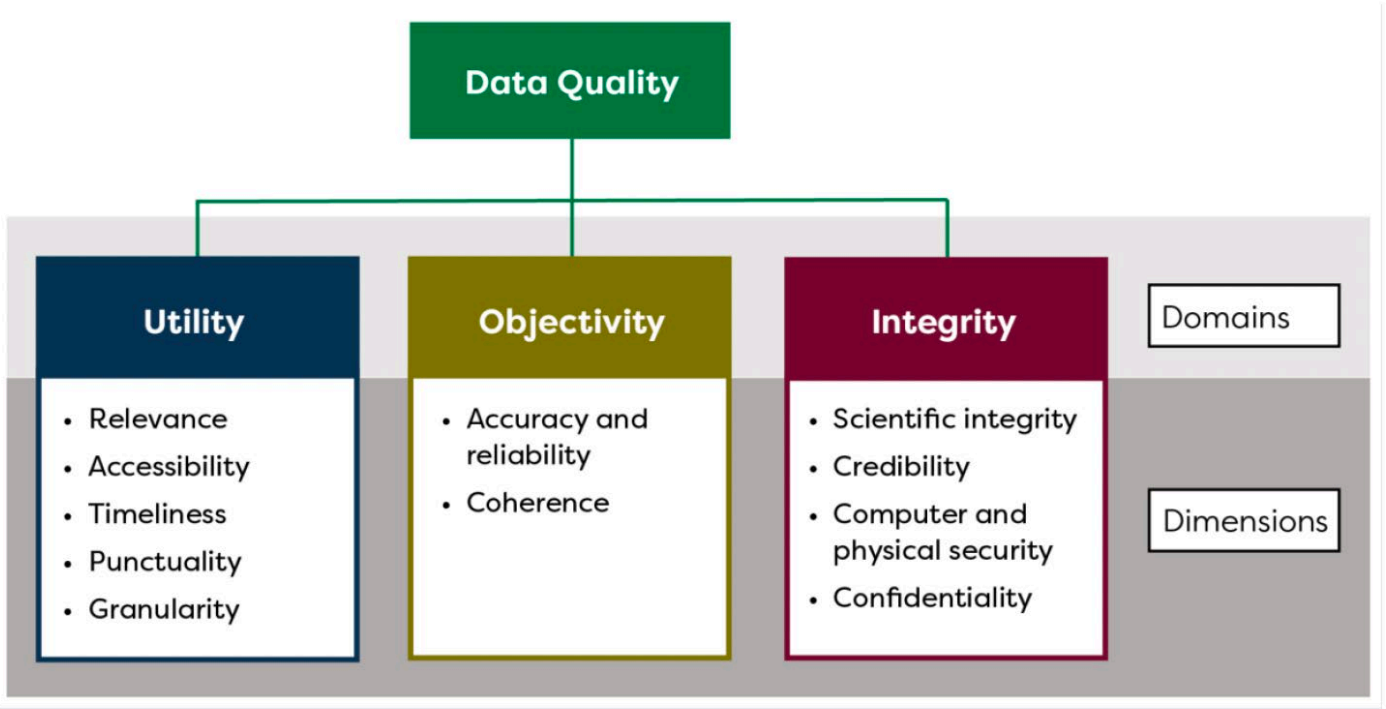


Figure 1. The "domains" and "dimensions," or characteristics, of the Federal Committee on Statistical Methodology framework for data quality.

Example application

Relevance [Domain: UTILITY]

Relevant data meet users' information needs. Data are relevant if they can be connected—for example through a logic model—to the outcomes[4] of interest. Determining relevance entails first understanding what is of interest, and then identifying measurements that are informative to that interest. Generally, data are relevant when appropriate measurements of the population of interest are taken at the right time and place.

Consider a few examples. If a program is considering native plants restoration in an imperiled habitat and requires information on the status of the habitat, then five-year-old data from a different region is less relevant than two-year-old data from a parcel at or near to the restoration site. If an office is evaluating a program's efficiency, salary cost is less relevant than the combined salary, equipment, administrative, and travel costs. If IT staff are allocating service support time by application, then the number of active users is a more relevant measure than the number of application installations.

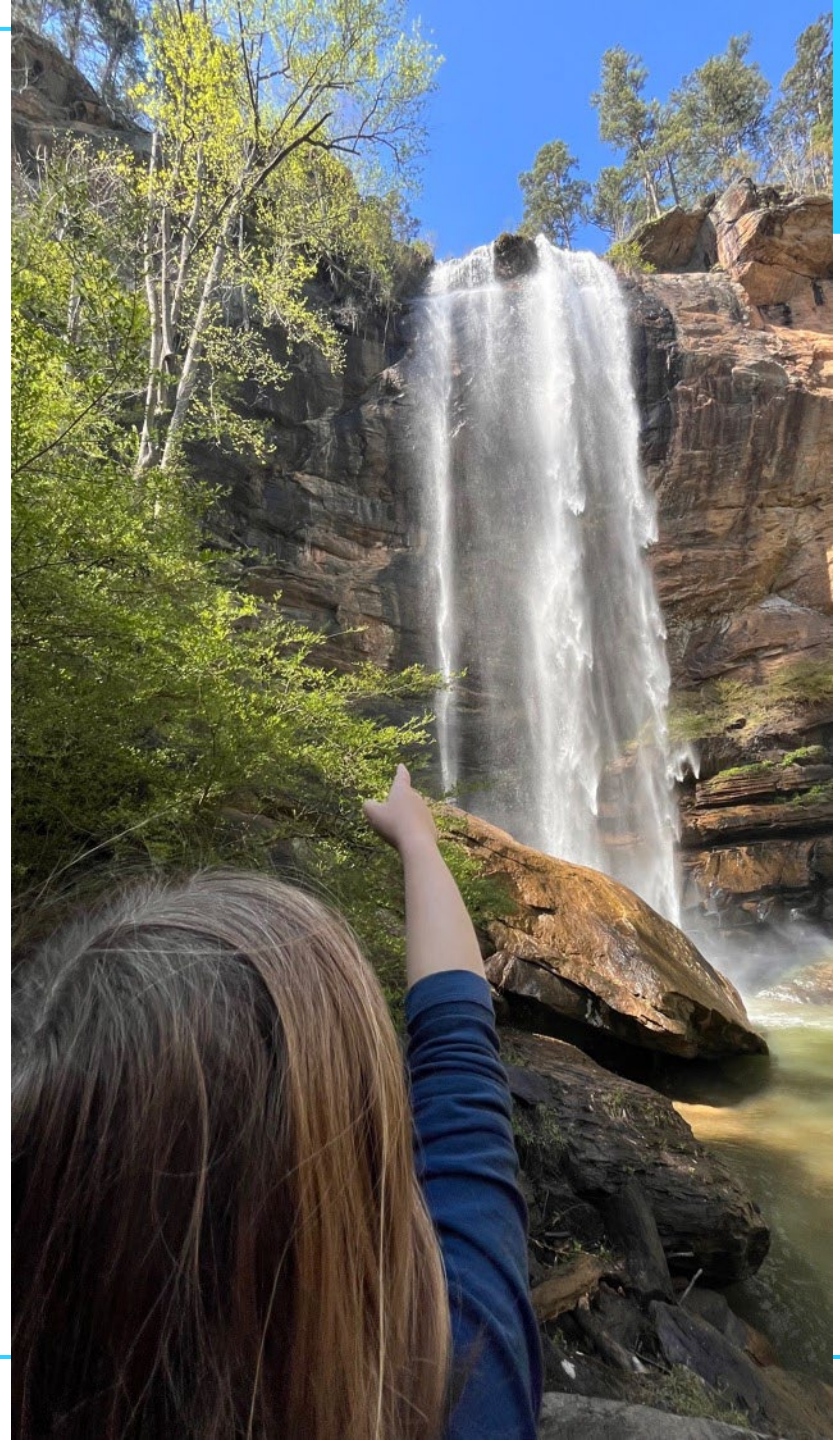
Program / Policy _____ **Assessor** _____

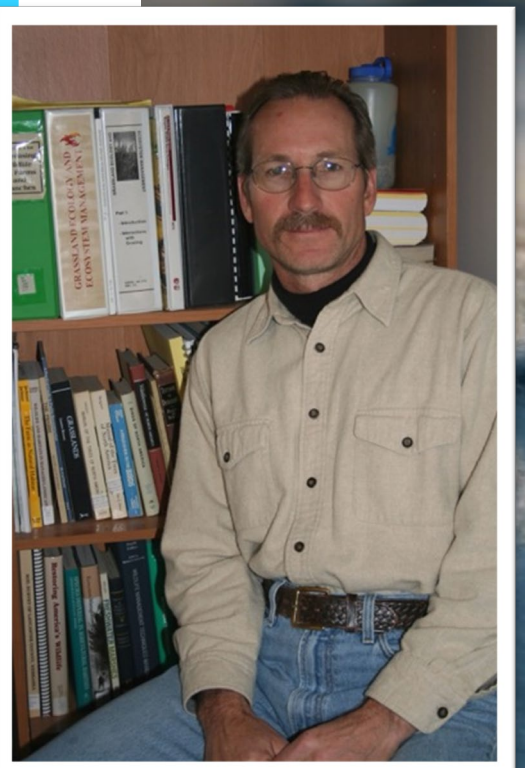
Bureau / Office _____ **Date** _____

	Dimension	Question	Addressed	Notes
Planning & Design	**Relevance**	Are the data relevant to the questions they are used to answer?	☐	
	Granularity	Do the data represent the appropriate temporal, spatial, or domain resolution to inform the questions at hand?	☐	
Collection or Acquisition	**Coherence**	Does data collection follow statistical standards and use consistent methods and definitions so data can be combined with other relevant data?	☐	
	Accuracy & Reliability	Are the data unbiased, precise estimates of the true value? Are values consistent across repeated measurement under similar conditions?	☐	
	Scientific Integrity	Do data collection and handling procedures meet principles of scientific integrity?	☐	
Analysis & Use	**Timeliness**	Are the data available for use close to the event or phenomenon they describe?	☐	
	Punctuality	Are the data available for use on or around their planned release date?	☐	
Storage & Maintenance	**Security**	Are the data stored securely with suitable backups to ensure that they are not corrupted or falsified?	☐	
	Confidentiality	Does data management and sharing meet confidentiality requirements in law, regulation, and Interior policy?	☐	
	Credibility	Do users have confidence in the data based on current or past practices for data collection, storage, and management?	☐	
Sharing & Publication	**Accessibility**	Are the data easy to discover and access in an understandable format?	☐	

Next steps and actions

- Staff capacity-building
  - Community of Practice
  - Resources for training
  - Public-private partnerships?
- Management and leadership capacity
  - Early action essential
  - Decision-making with variable quality
- What else???





CC-BY Susan Cook


U.S. DEPARTMENT OF THE INTERIOR
MARCH 3, 1849