NIST-wide initiatives for improving access to scientific research data

Chandler A. Becker, PhD [1]
Robert J. Hanisch, PhD [1]
Katherine E. Sharpless, PhD [2]

[1] Office of Data and Informatics, Material Measurement Laboratory, NIST
[2] Special Programs Office, Laboratory Programs, NIST

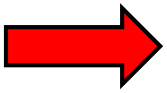
The philosophy of open data: Making data easy to find, access, and use can fuel entrepreneurship, innovation, and scientific discovery that improves Americans' lives and contributes significantly to job creation.

Government data should be open and machine-readable by default.

The practicalities of open data: It takes time and money to put our data into a form that someone else (internal or external) can understand and use.

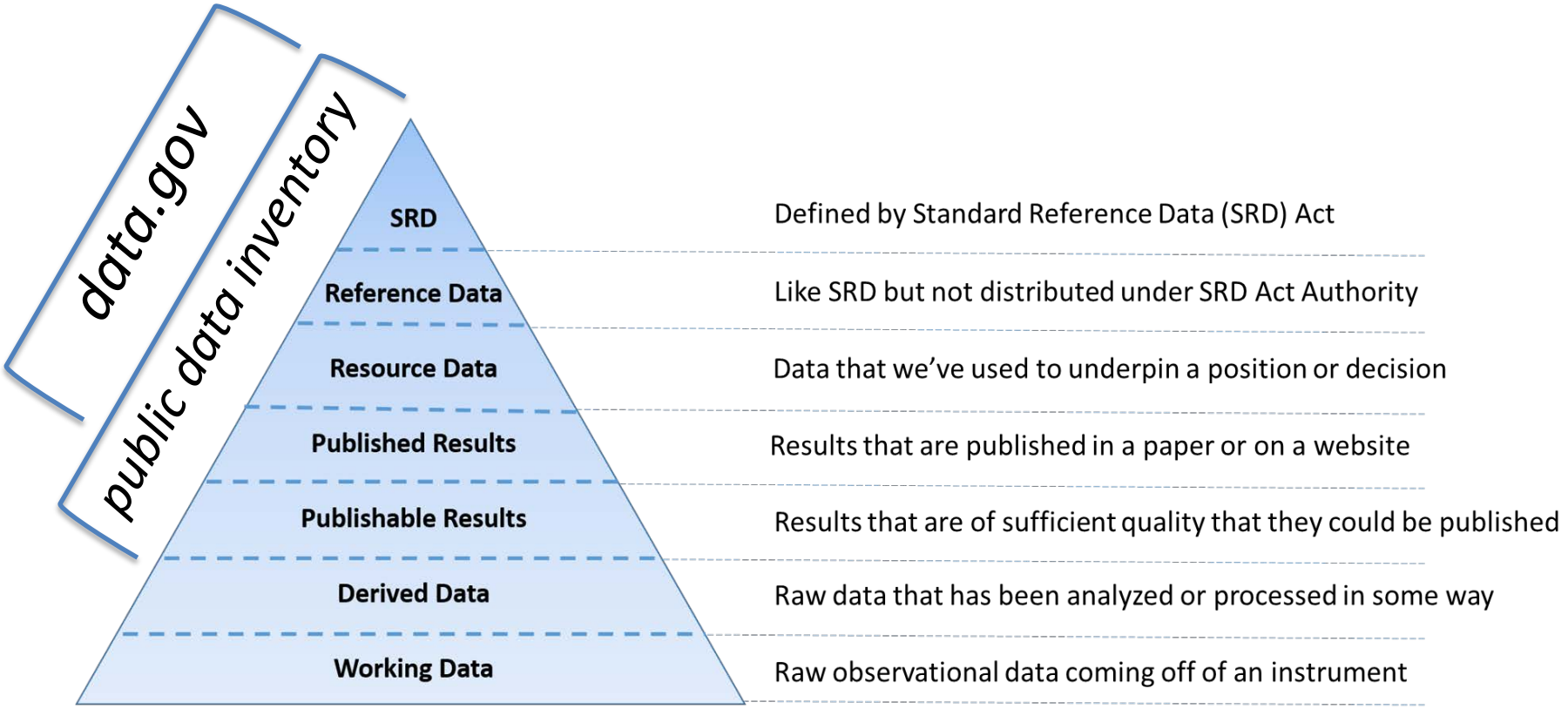
But that time and money is well spent if we can

- Use our own data to work smarter,
- Make better decisions by using our data, and
- Publish data that's frequently requested so we don't have to make it available to individuals.



Balance release of high-impact datasets and availability of resources

Data categories at NIST



NIST policy is to make our scientific research data and publications freely* available in publicly accessible databases and repositories.

What We Will Make Public:

- Standard Reference Data#
- Reference data#
- Resource data#
- Published results#

What We Might^ Make Public:

- Publishable results#
- Derived data
- Working data

*'Free' doesn't apply to SRD that we sell to recover certain costs.
^ It's up to the NIST Laboratory to decide whether their staff publish these data types.
#These data types must be listed in NIST's data inventory.

What will not be shared?

- Laboratory notebooks, results of preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as laboratory specimens;

- Trade secrets, commercial information, or other materials necessary to be held confidential by a researcher until they are published, or similar information that is protected under law; and

- Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

http://www.nist.gov/open/upload/NIST-Plan-for-Public-Access.pdf

Efforts to improve access to NIST data

- Listings of datasets available at data.gov
- Peer-reviewed publications in PubMed Central (NIH) and technical reports in the Federal Digital System (GPO)
- Infrastructural improvements to support better data management and sharing
  - Open Access to Research initiative includes a data management planning tool, data catalog support, data repository and discovery portal, faster networks, and more tools approved for researcher use
- Standard Reference Data (SRD) modernization to improve consistency across data products
- Development of data catalogs and registry systems

Data sharing at NIST

- NIST is working to make data available as resources and policy allow.

- Various efforts are supporting this goal, including:
  - Open Access to Research developing infrastructure and resources
  - SRD modernization to improve consistency across products
  - Data catalogs and repositories (stay tuned for other presentations)

- Thanks for your attention.
- Questions?