



Open Census Initiative

Emily Molfino, Ph.D.

Research and Methodology Directorate

Census Scientific Advisory Committee

September 20, 2024

“Open Science is the principle and practice of making research products and processes available to all, while respecting diverse cultures, maintaining security and privacy, and fostering collaborations, reproducibility, and equity. Science is the principle and practice of making research products”

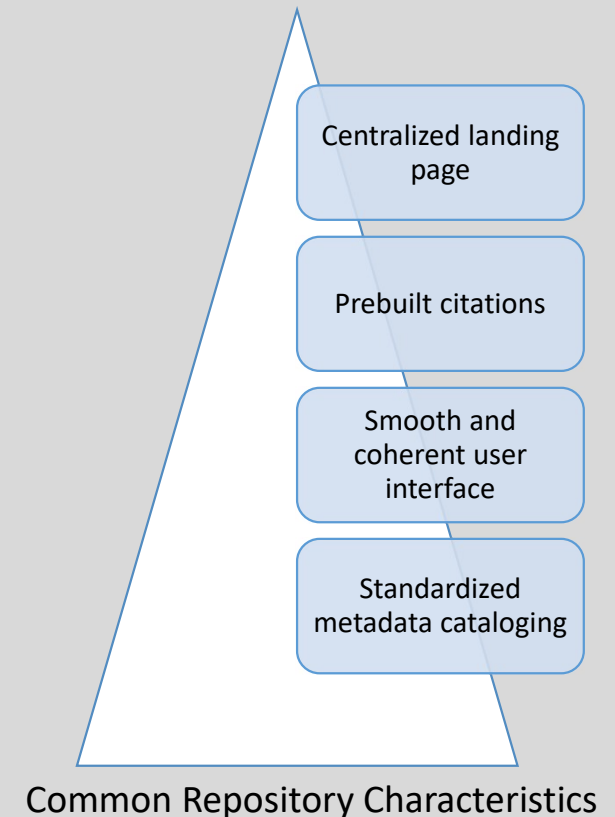
—White House Office Of Science and Technology Policy (OSTP)

YEAR OF **OPEN SCIENCE**



Learning from Federal Agencies

- Department of Energy ([DOE](#)) – Office of Scientific and Technical Information
- National Institutes of Health ([NIH](#)) – National Library of Medicine
- Department of Agriculture ([USDA](#)) – National Agricultural Library
- National Institute of Standards and Technology ([NIST](#))



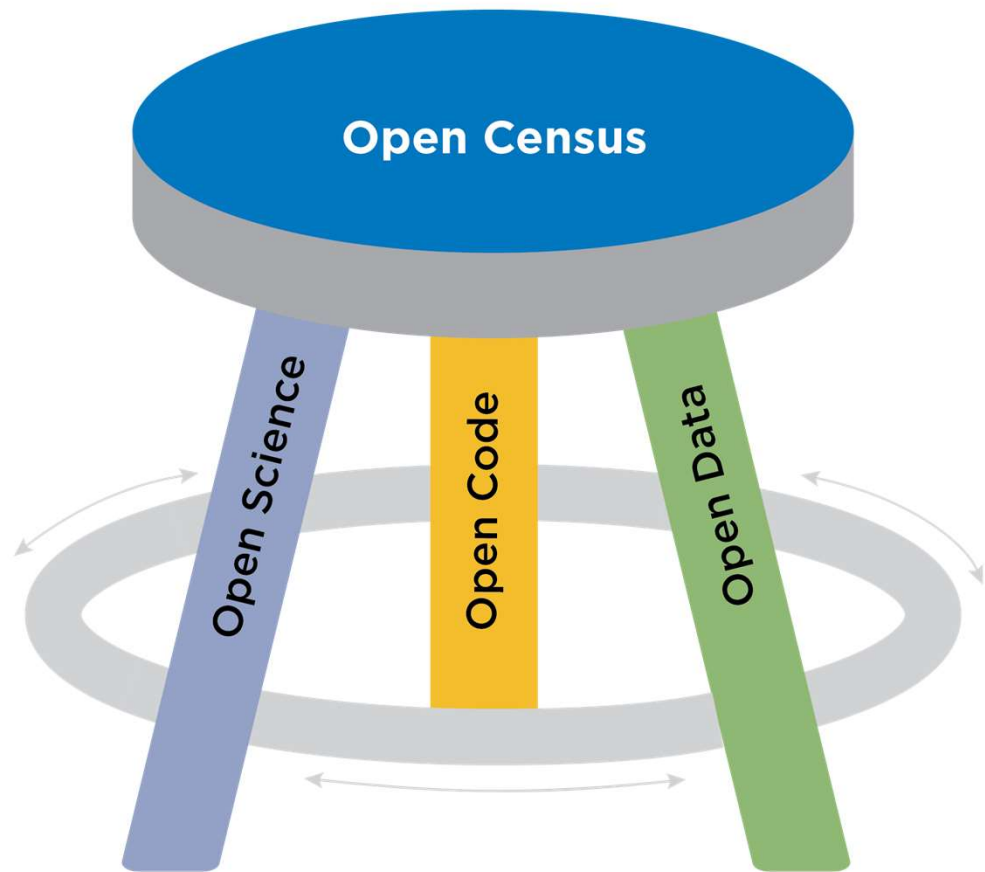
Current State of Census Bureau Dissemination

- [Census.gov](https://www.census.gov)
- [Data.Census.gov](https://data.census.gov)
- Census Application Programming Interface (API)
- Census file transfer protocol (FTP) server
- Census Library
- Census Publications
- Census Visualizations
- Tableau Public
- Census Working Papers
- Academic Journals
- GitHub
- AWS Web Service

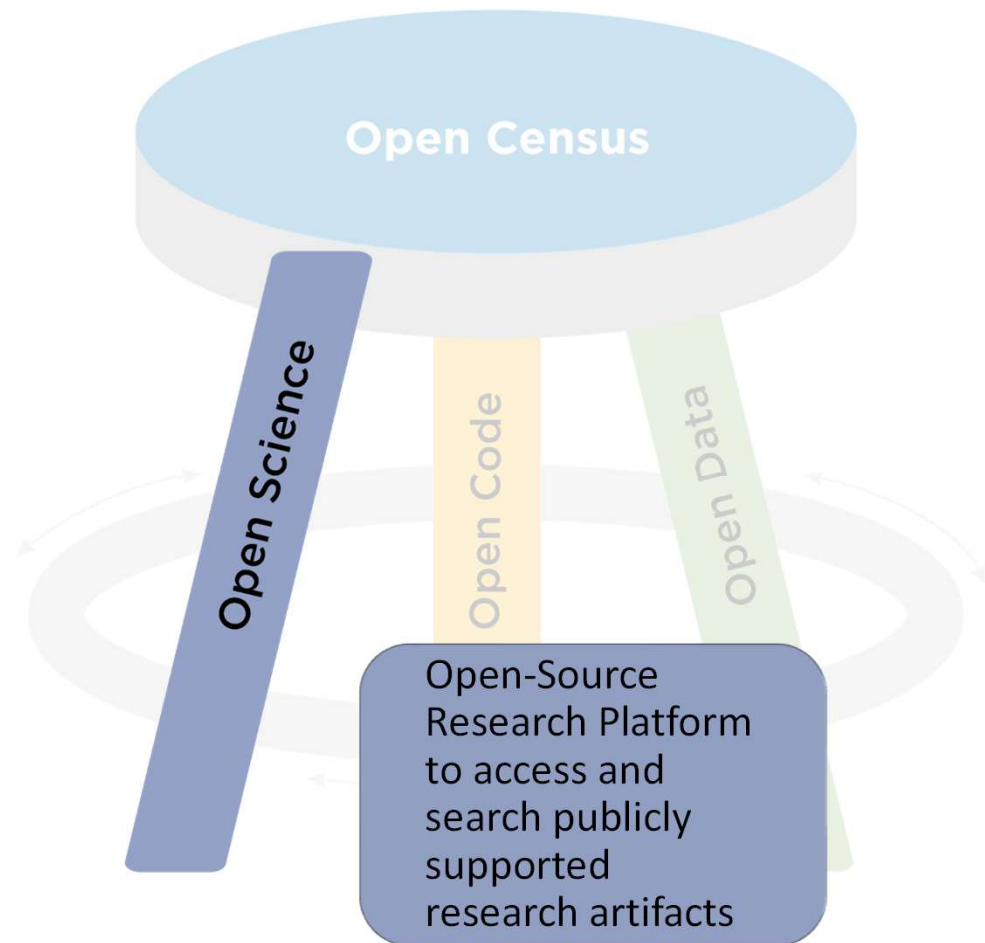
The Open Census program will make federally funded research, code, and data readily accessible and transparent, fostering a collaborative environment that promotes innovation.



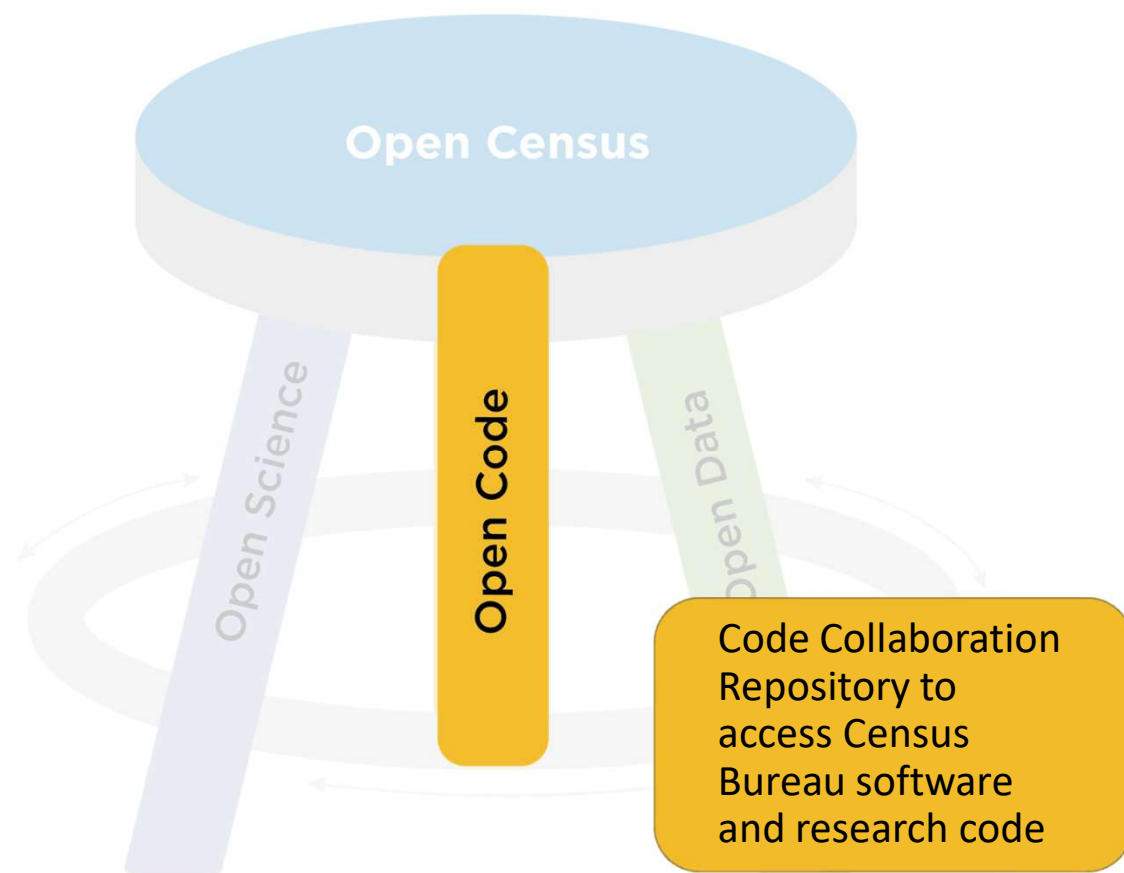
The Three Distinct But Connected Open Census Initiatives



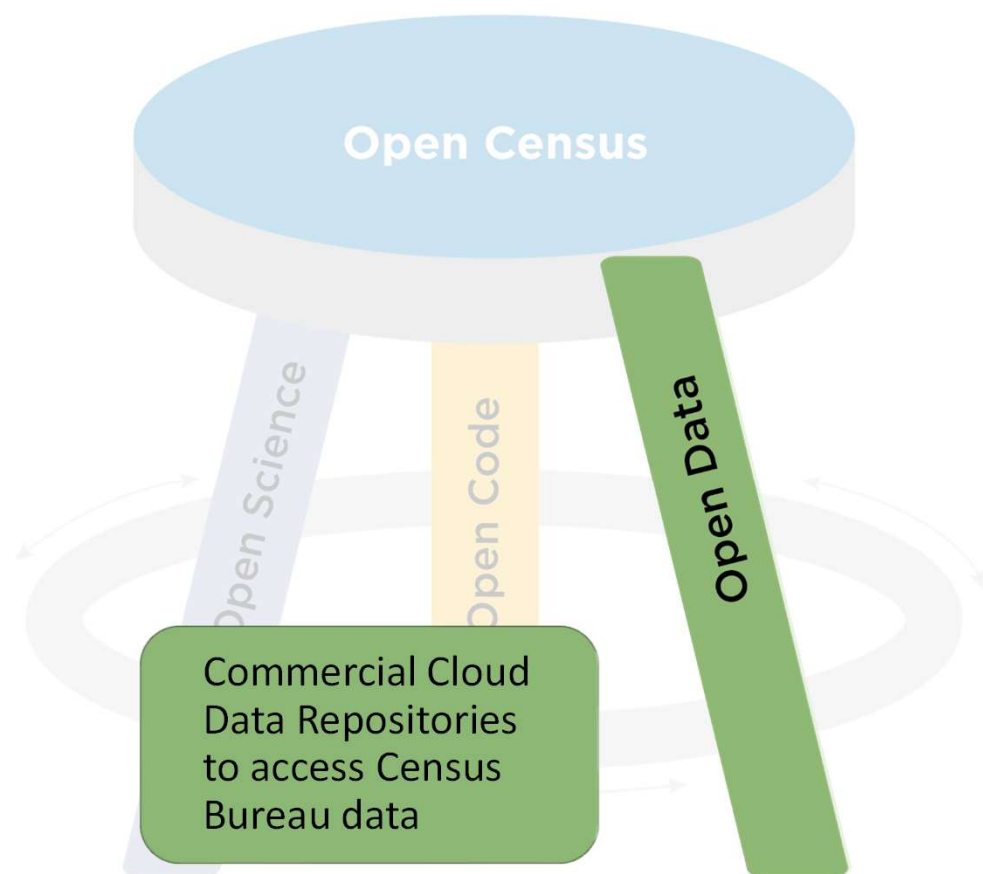
The Desired Future State of Open Census



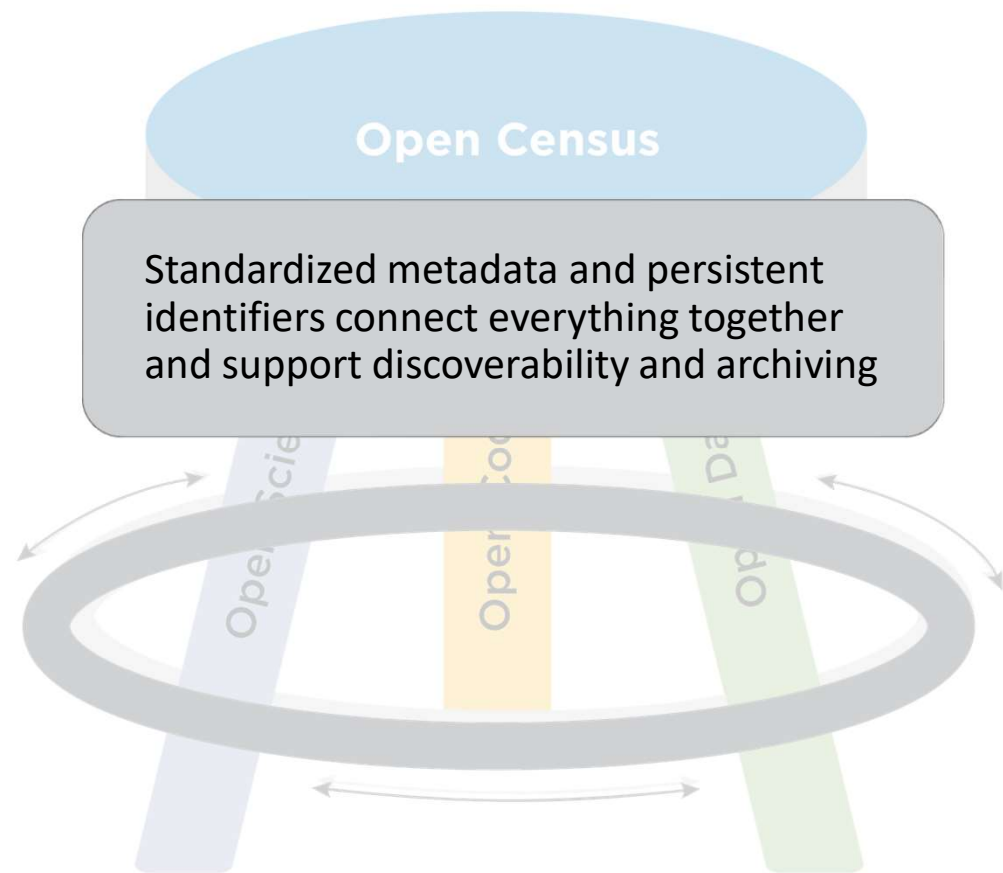
The Desired Future State of Open Census



The Desired Future State of Open Census



The Desired Future State of Open Census



Open Census | Open Science



Rationale

Provide a centralized location where research output can be released, easily discovered, and seamlessly interconnected



Strategy

Develop a Census Bureau solution for an Open Science Repository
Assign persistent identifiers
Create standardized metadata



Benefits

Increased visibility, recognition, and impact of Census Bureau research
Improved accessibility and transparency



Risks

Ability to deliver a solution that is easy and fits seamlessly into the research workflow
Migrating legacy products into a new repository

Open Census | Open Code



Rationale

Facilitate the dissemination of software products and code to increase understanding and transparency of Census Bureau research



Strategy

Leverage GitHub to host and cite code and software
Maintain internal code standards and best practices
Assign persistent identifiers



Benefits

Increases the transparency of the research process
Encourages innovation
Facilitates engagement and collaboration



Risks

Ensuring code releases adhere to established guidelines
Cultural adoption

Open Census | Open Data



Rationale

Provide platforms to release, discover, access, and cite data that support research



Strategy

Create cooperative agreements for cloud supported downloads

Assign persistent identifiers

Provide prebuilt citations

Create tools to improve data discoverability



Benefits

Increased availability

Reduced Cost

Secure and reliable

Facilitates tracking

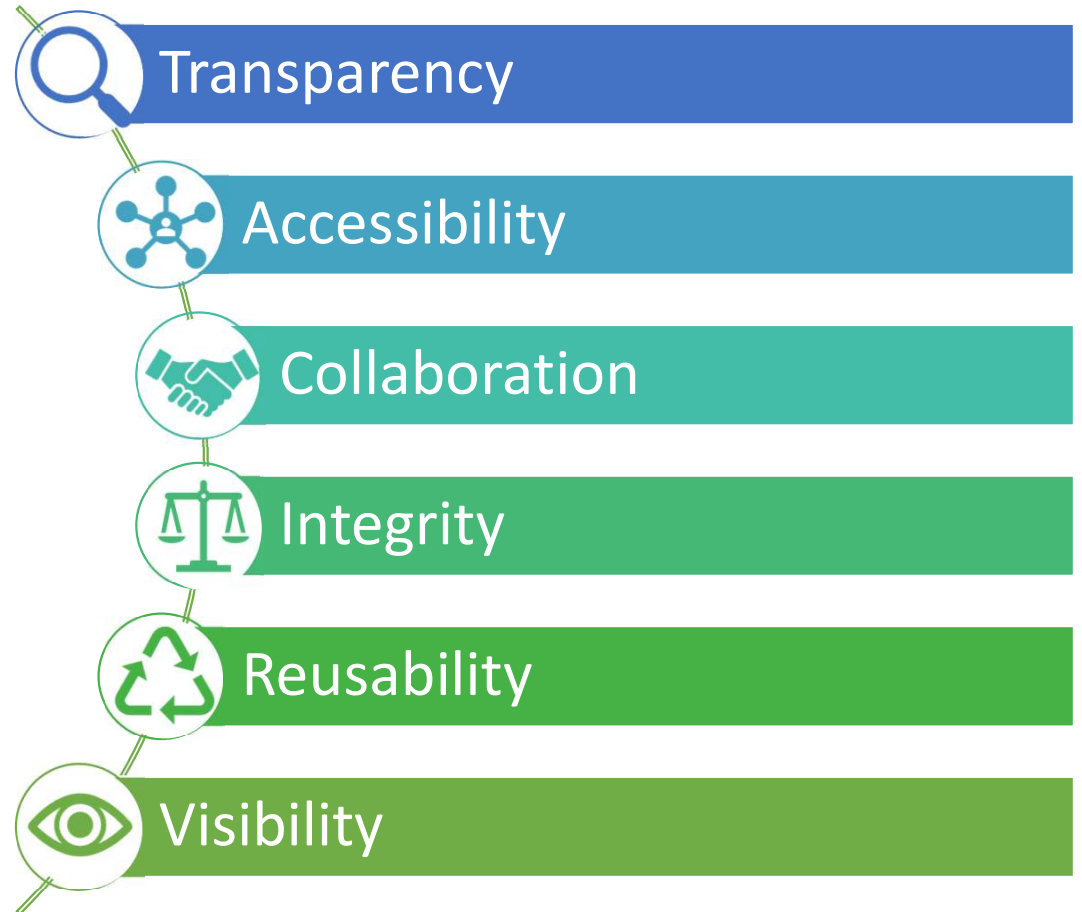


Risks

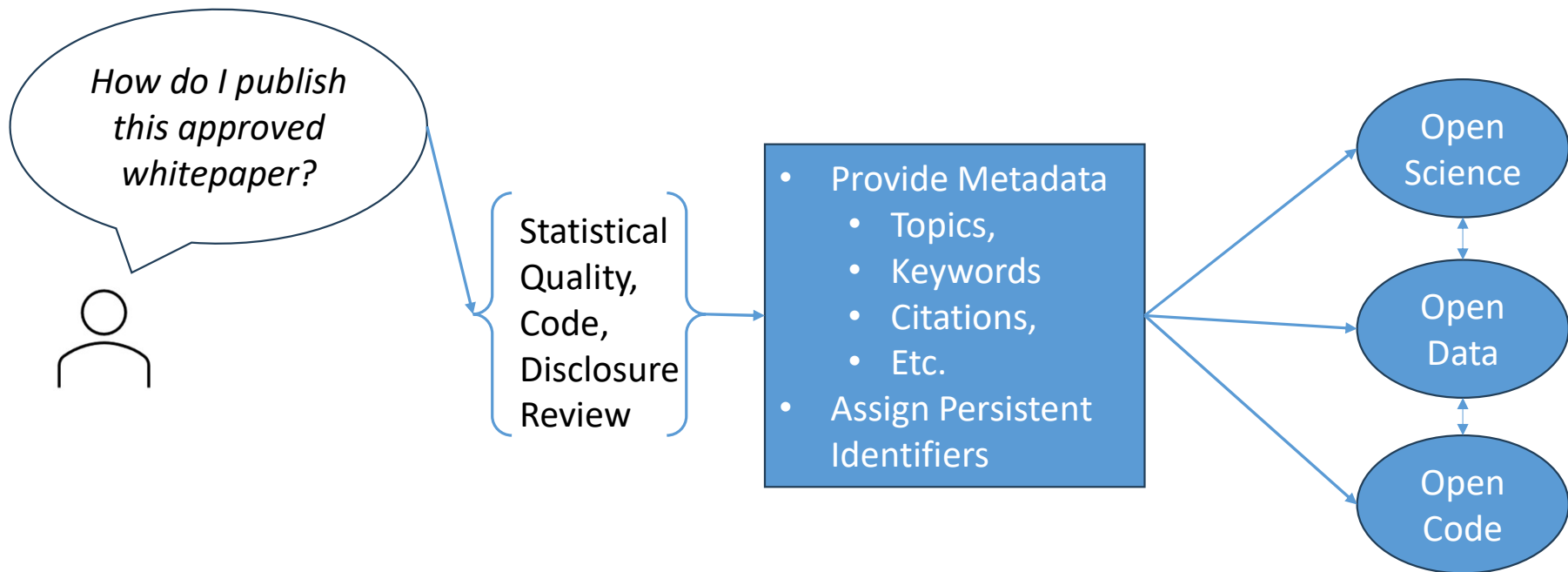
Limited agreement duration

Cultural adoption

Open Census Benefits

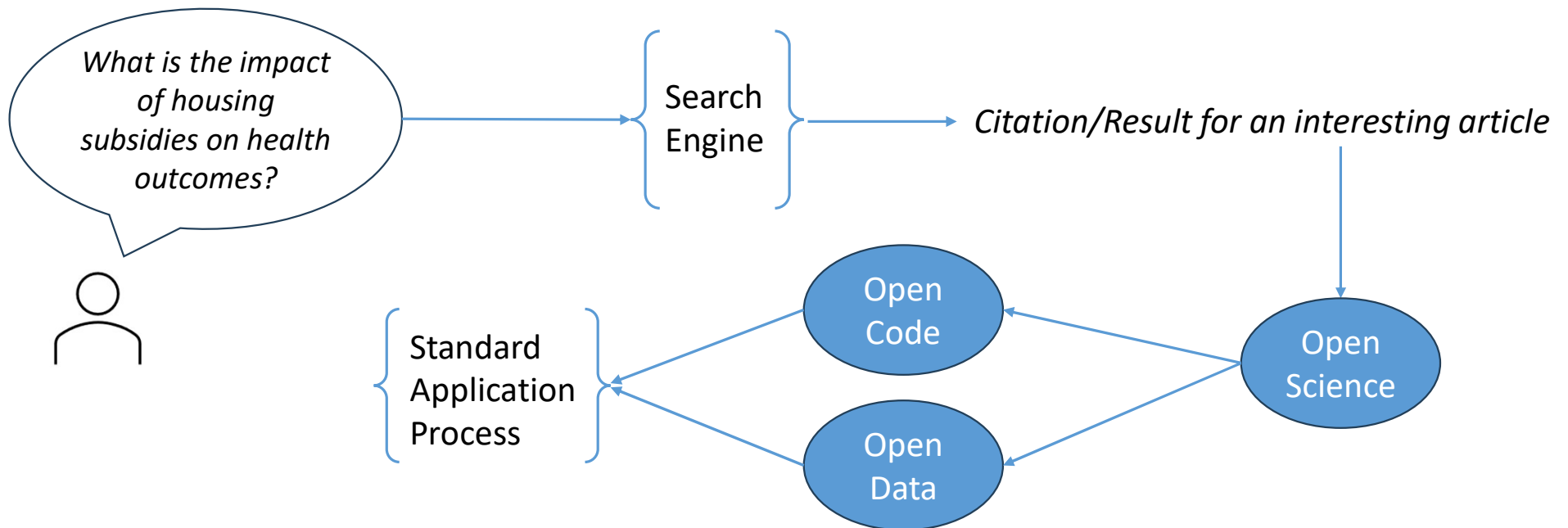


Research Workflow



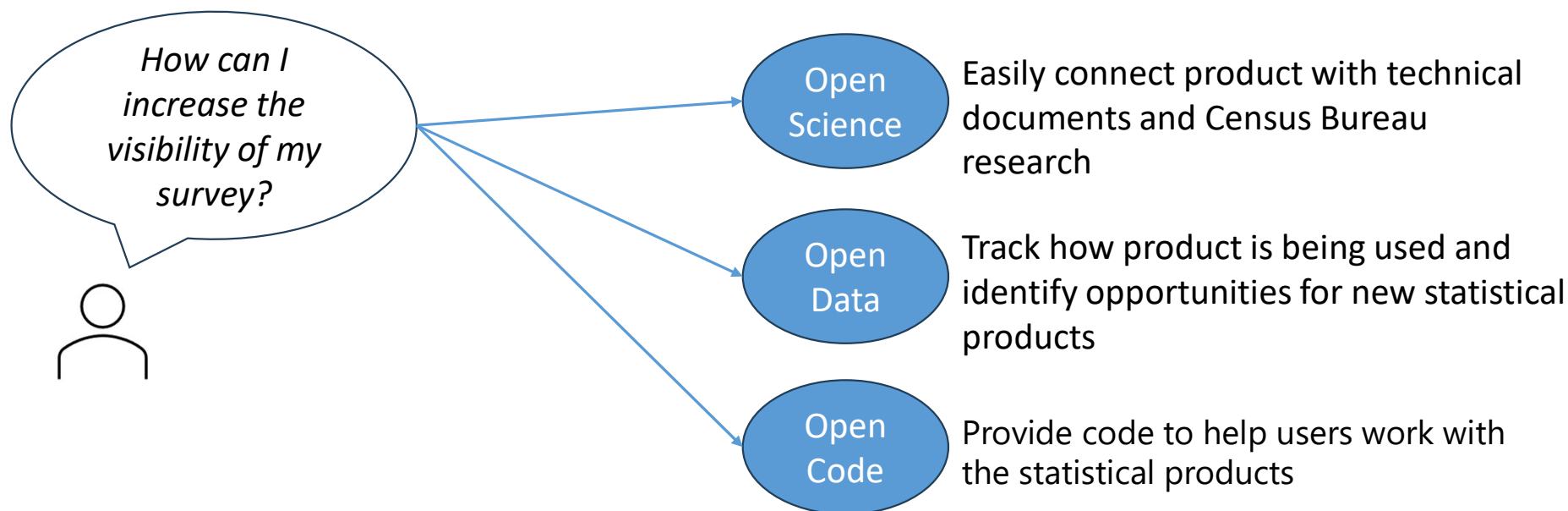
* This is a generalized example of one possible workflow.

External Researcher



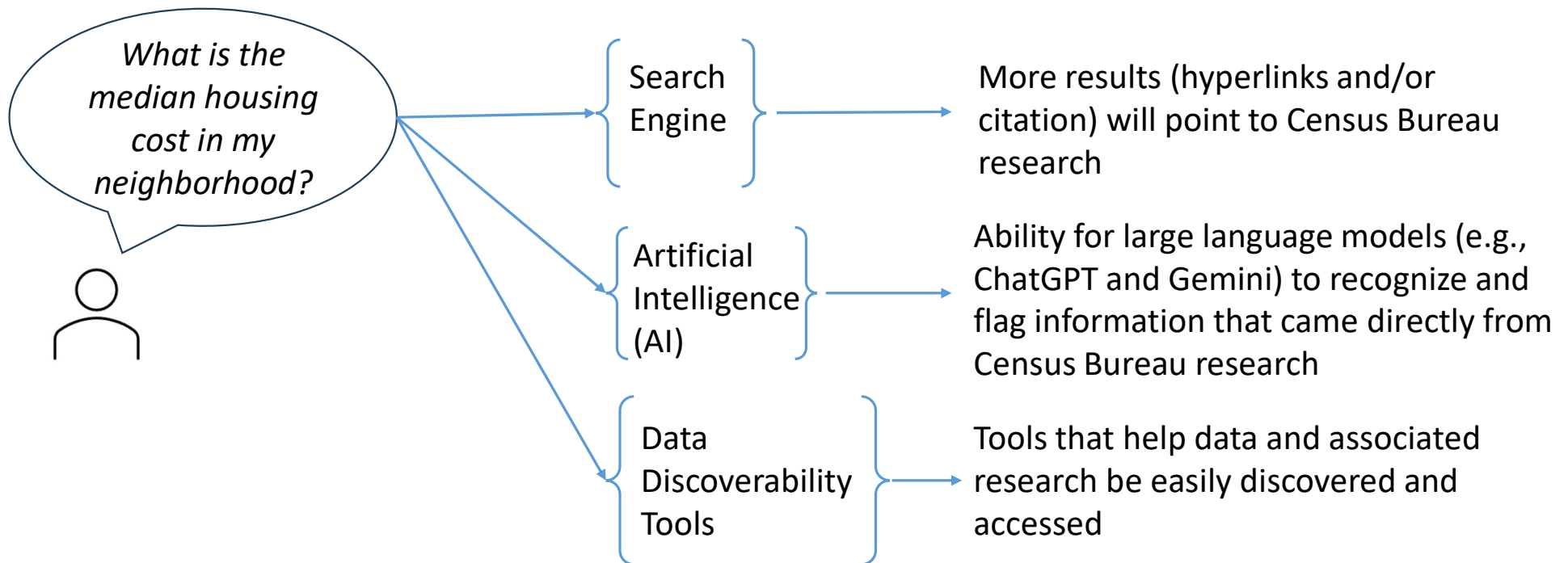
* This is a generalized example of one possible workflow.

Survey Sponsor



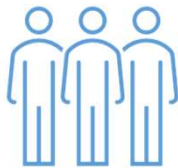
* This is a generalized example of one possible workflow.

Novice Data User



* This is a generalized example of one possible workflow.

Next Steps



Near Term

- Socialize Open Census
- Develop Concept of Operations for operation of Open Census
- Update standards guidance around Open Census (citations, persistent identifiers, code dissemination)
- Support statistical product discoverability tools



Mid Term

- Build Open Census platforms
- Formalize Census Bureau GitHub under Open Census
- Update business processes around Open Census
- Publish Research, Code, and Data on Open Census associated platforms



Long Term

- Sustainable and persistent implementation and delivery that is fully integrated into our research processes



Citation Guidance

New website with guidance on how to cite Census Bureau statistical products

- Example citations
- Focus is on components rather than a specific citation format

Next steps

- Adapt guidance for Census Bureau researchers
- Explore creating standard around citing underlying data

Metadata Discovery and Citation Tool

Data Science Training Program 2024 Cohort
Capstone Projects

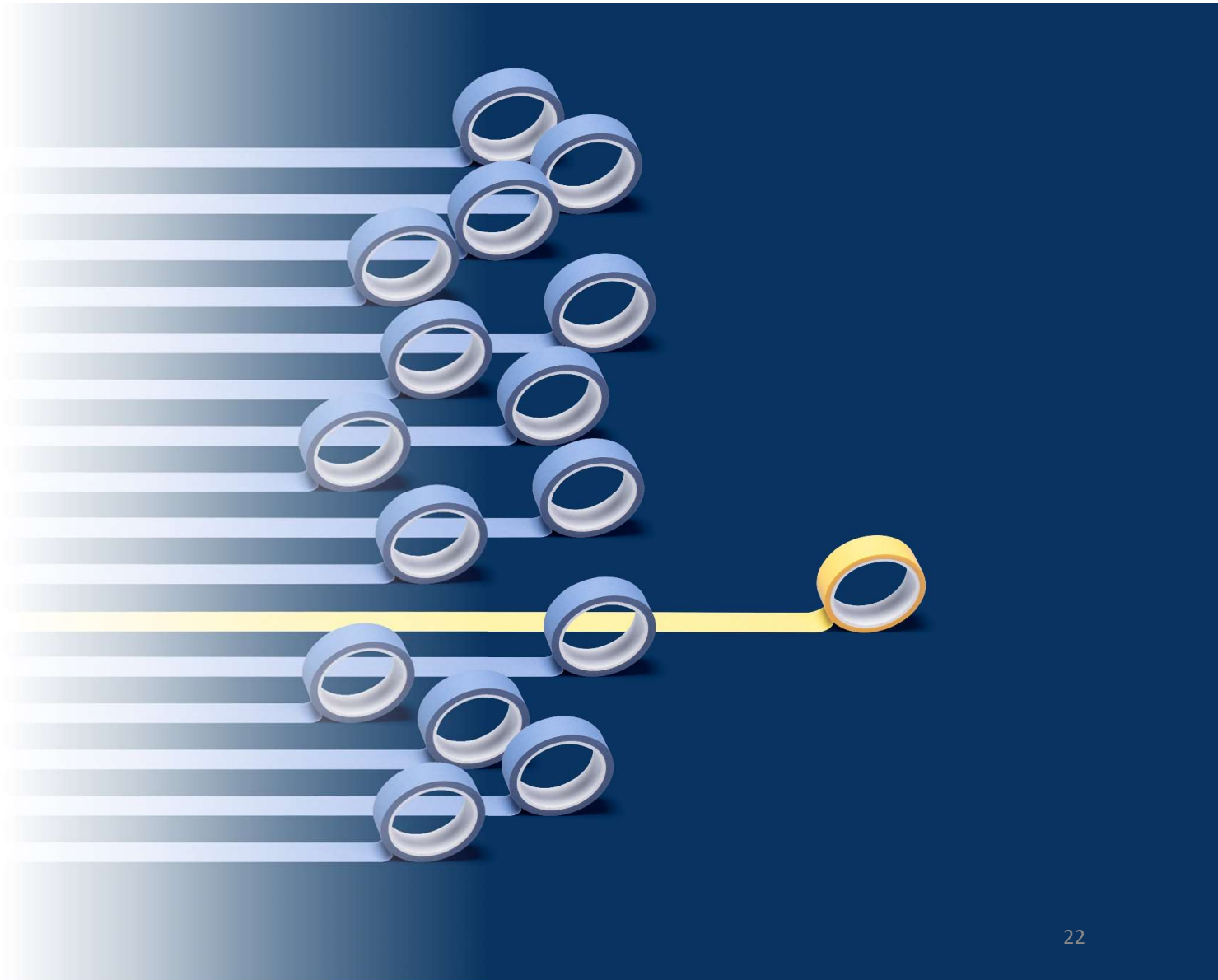
- Developed a tool to discover and display links to metadata for Census Bureau statistical products
- Expansion of Census Survey Explorer in scope, content, and functionality

Next steps

- Create internal team to build upon this effort to finalize a prototype product
- Incorporate feature to build citations

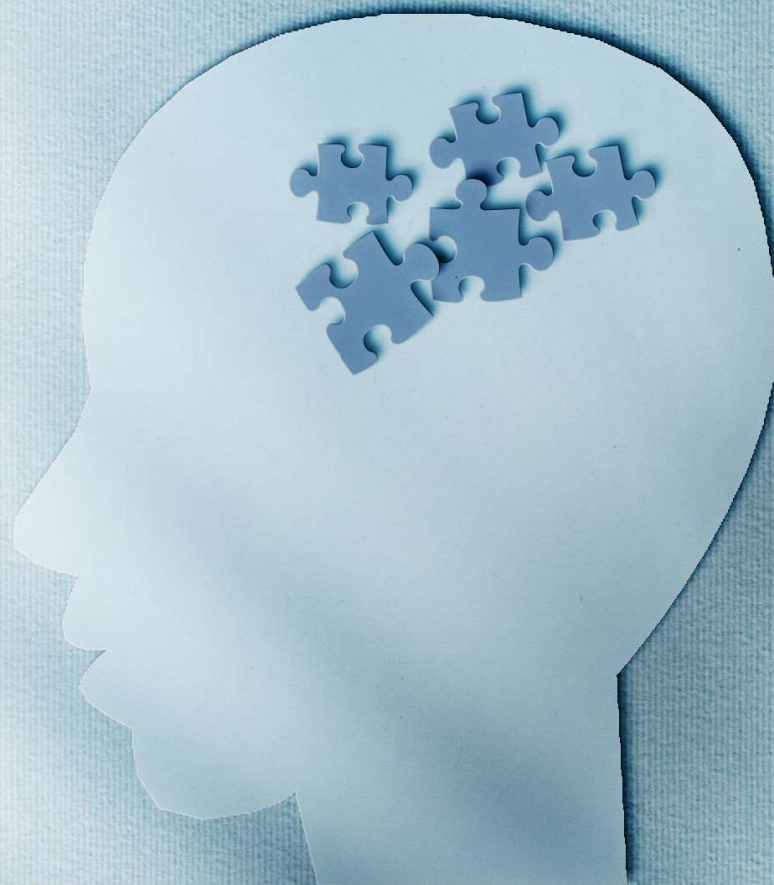


Leading the way of Open Science in the Federal Statistical Space



Questions for CSAC

- Is there anything we forgot or should keep in mind that was not addressed here?
- Does the committee have recommendations around how Open Census can increase equity in how the Census Bureau disseminates statistical products and research?
- What advice do you have as we develop standardized metadata and persistent identifiers like Digital Object Identifiers (DOI) and Open Researcher and Contributor Identifiers (ORCID)?





Thank You

Emily Molfino, Ph.D., Open Census Lead
Research and Methodology Directorate
U.S. Census Bureau

emily.molfino@census.gov