Evolving a Data Enterprise to Support Relevant, Timely, and Equitable Statistical Products—*We Need Your Help!*

Erika Becker-Medina,

Senior Advisor to the Chief Scientist and Associate Director

Research and Methodology Directorate

U.S. Census Bureau

May 16, 2023

The views expressed in this perspective are those of the presenter and not the Census Bureau.

Forces Driving Innovation Today

" *As we advance into the 21st century, we are experiencing* ***increased demand for our data***, *struggling with* ***challenges to traditional data collection*** *methods, and exploring* ***rich new data sources and tools*** *that can revolutionize what we do and how we do it. Our success critically depends on our ability to seize the opportunities in front of us to deliver* ***statistical products that address the increasingly complex and diverse needs of our users.***"

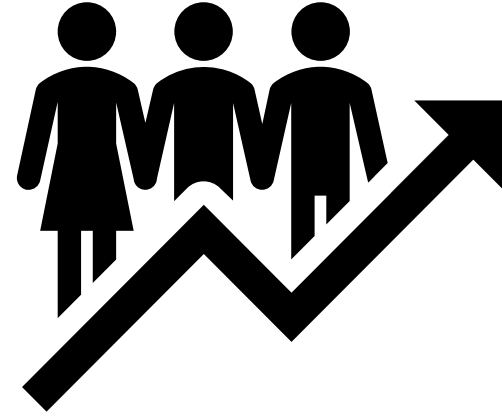**Ron S Jarmin, PhD.,** Deputy Director
U.S. Census Bureau

Flipping the Focus

Determine what information stakeholders need to reach their objectives

From there, shape the statistical products to be developed

Statistical Product First Approach
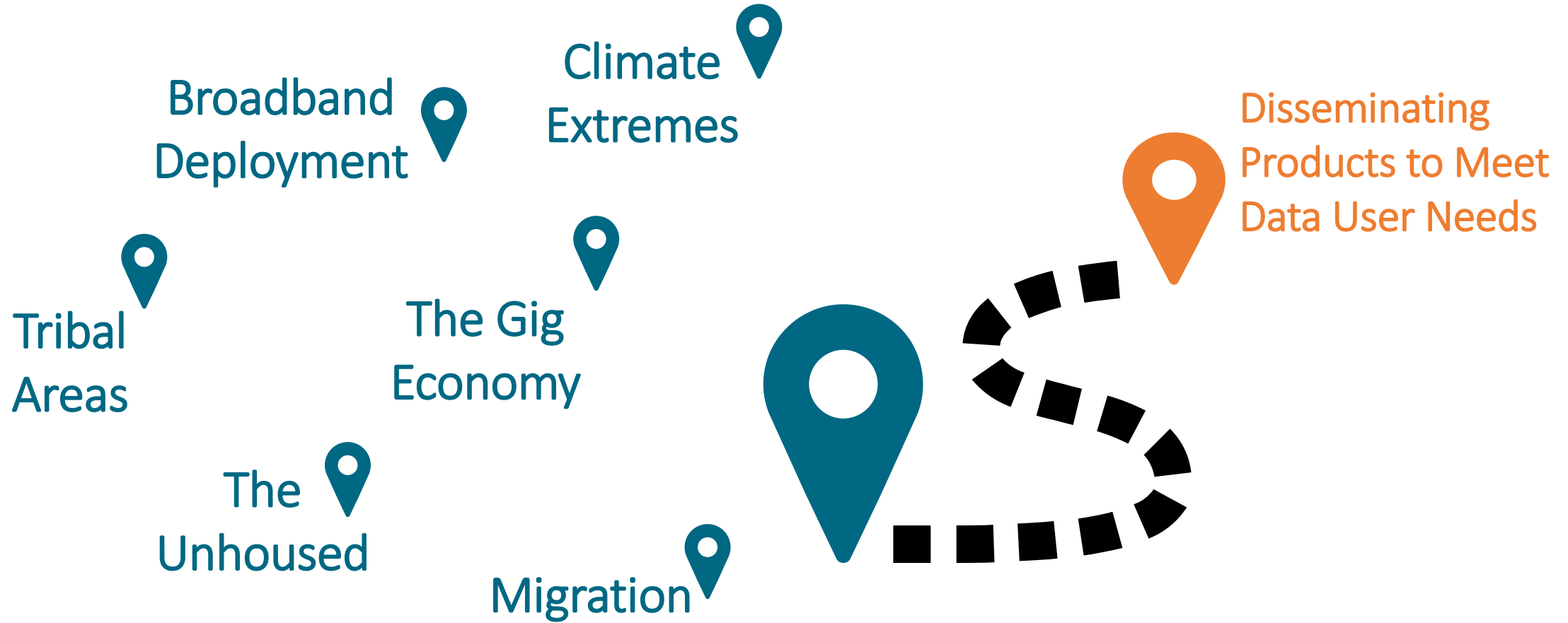
Statistical Product First Approach

Ensures Data Support Purposes and Uses

Disseminating Products to Meet Data User Needs

Purposes and Uses
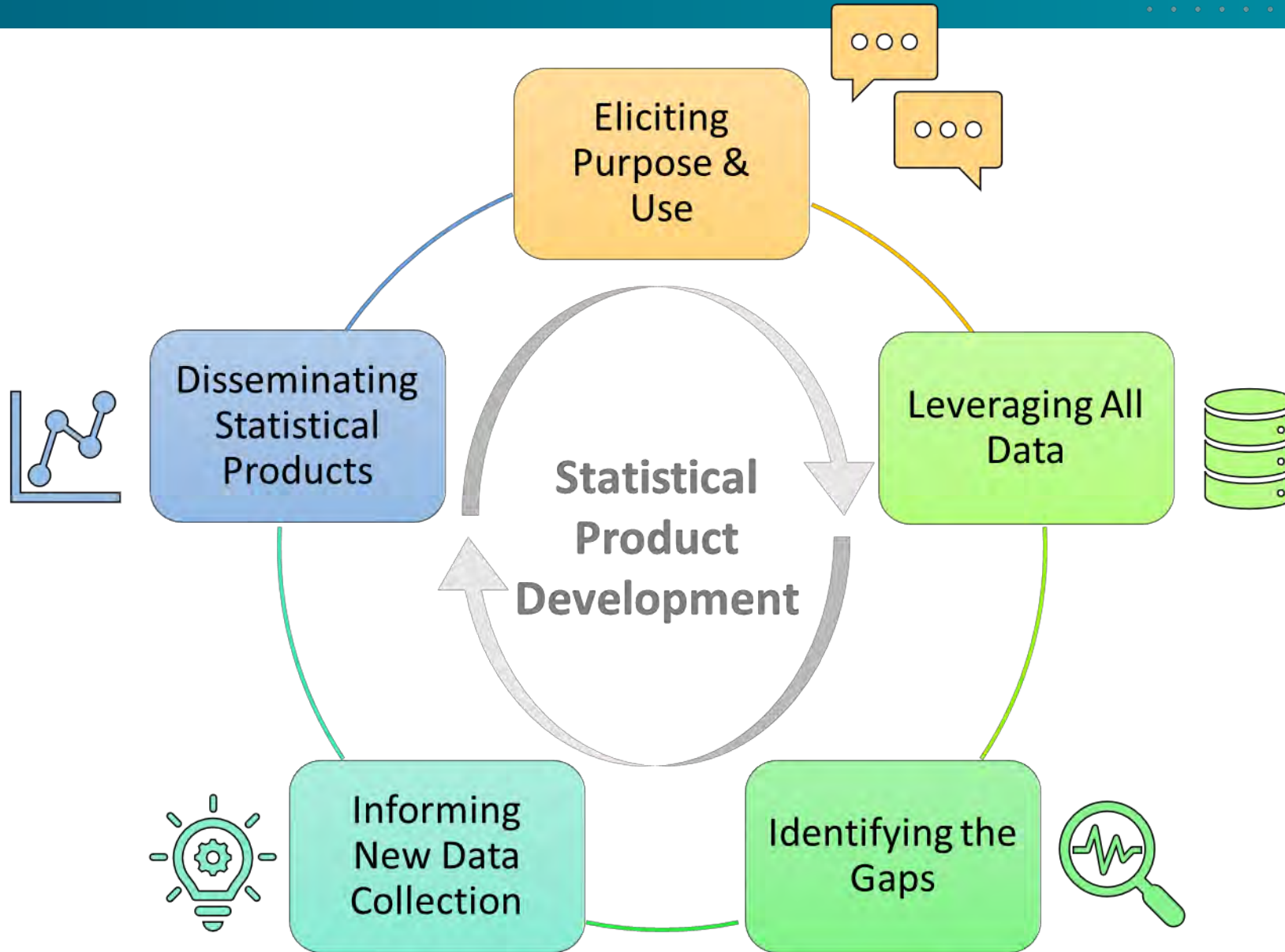
What are some purposes and uses?

Climate Extremes

Broadband Deployment

Disseminating Products to Meet Data User Needs

Tribal Areas

The Gig Economy

The Unhoused

Migration

What are some purposes and uses?

Differential Undercount

Commuting Patterns

Disseminating Products to Meet Data User Needs

Counting Children

Local Businesses

Equity and Inclusion

Grant Applications

United States® Census Bureau

Exemplar: Measurement of Mental Health

Eliciting Purpose & Use

There are societal questions on the impact of the COVID-19 pandemic on mental health care.

Disseminating Statistical Products

Leveraging All Data

Statistical Product Development

Informing New Data Collection

Identifying the Gaps

Exemplar: Measurement of Mental Health

Eliciting Purpose & Use

Disseminating Statistical Products

Statistical Product Development

Leveraging All Data

Questions on the receipt of prescription medication, counseling or therapy, and unmet mental health needs were included on the National Health Interview Survey (NHIS) in 2019.

Informing New Data Collection

Identifying the Gaps

Exemplar: Measurement of Mental Health

The NHIS asks these questions with a 12-month reference period. There is a need for more timely information.
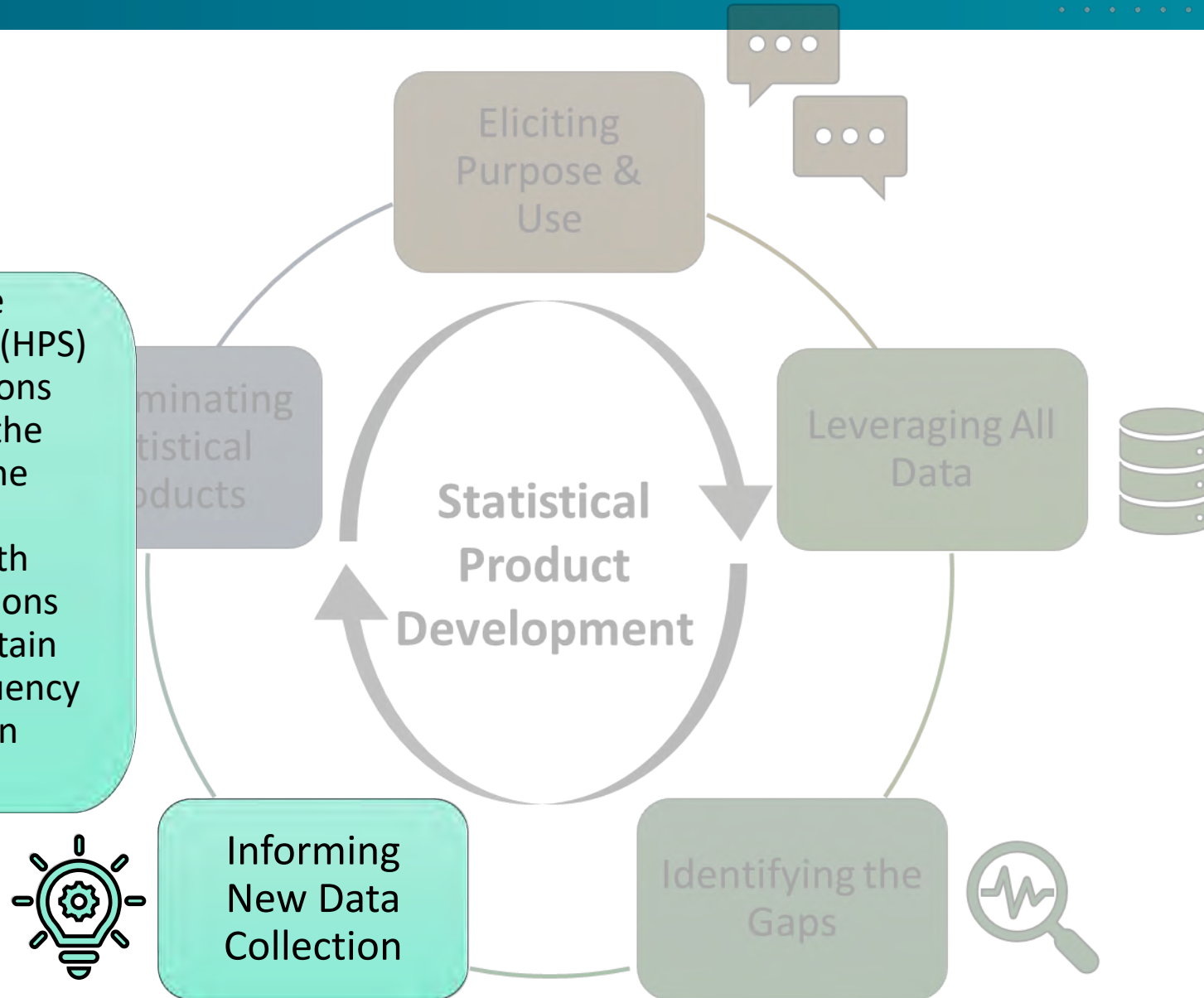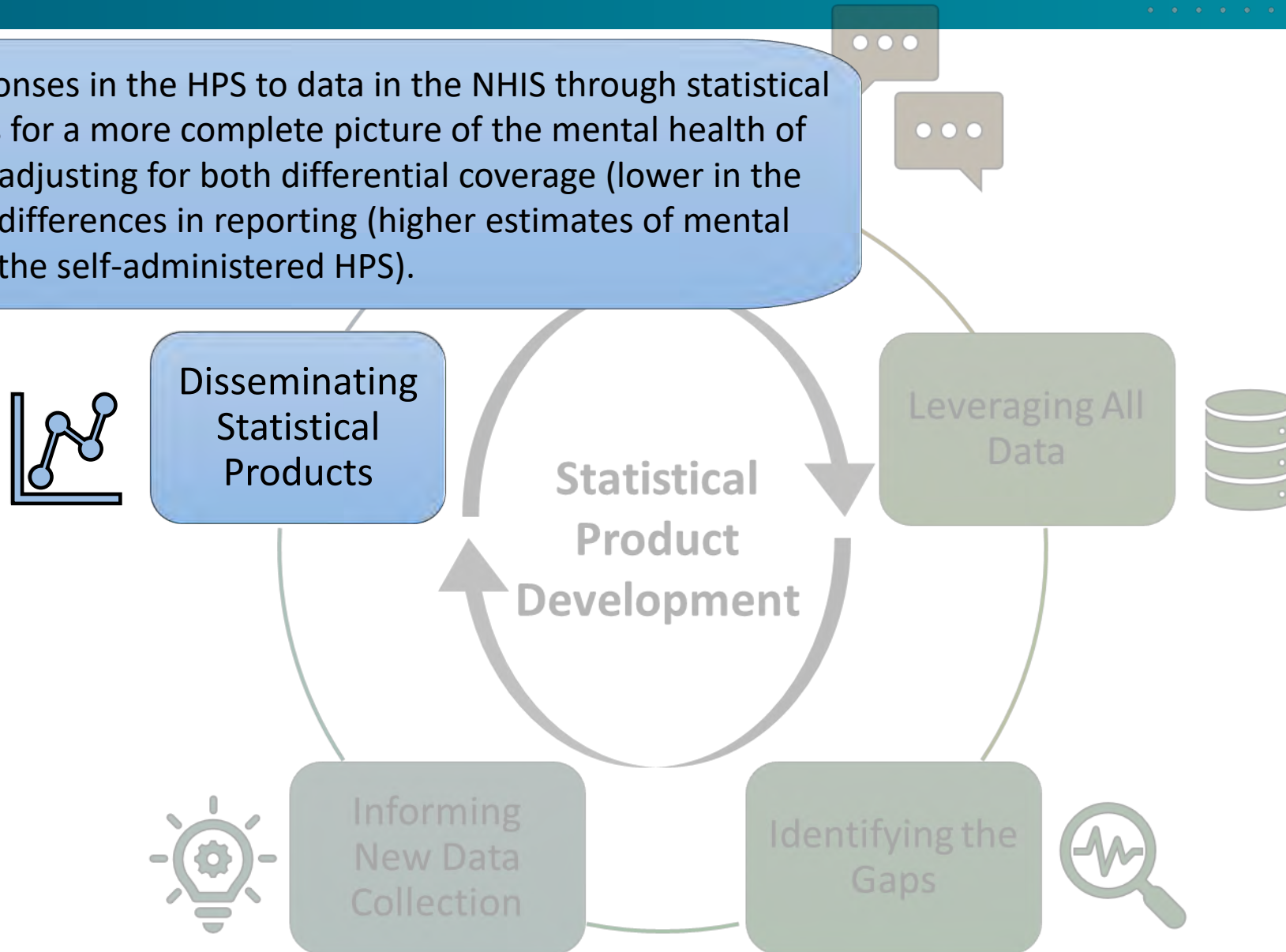
Exemplar: Measurement of Mental Health

Rapid deployment of the Household Pulse Survey (HPS) (July, 2020) asked questions about receipt of care in the past 4 weeks. Through the COVID-19 pandemic, the National Center for Health Statistics included questions based on the NHIS to obtain information on the frequency of anxiety and depression symptoms.
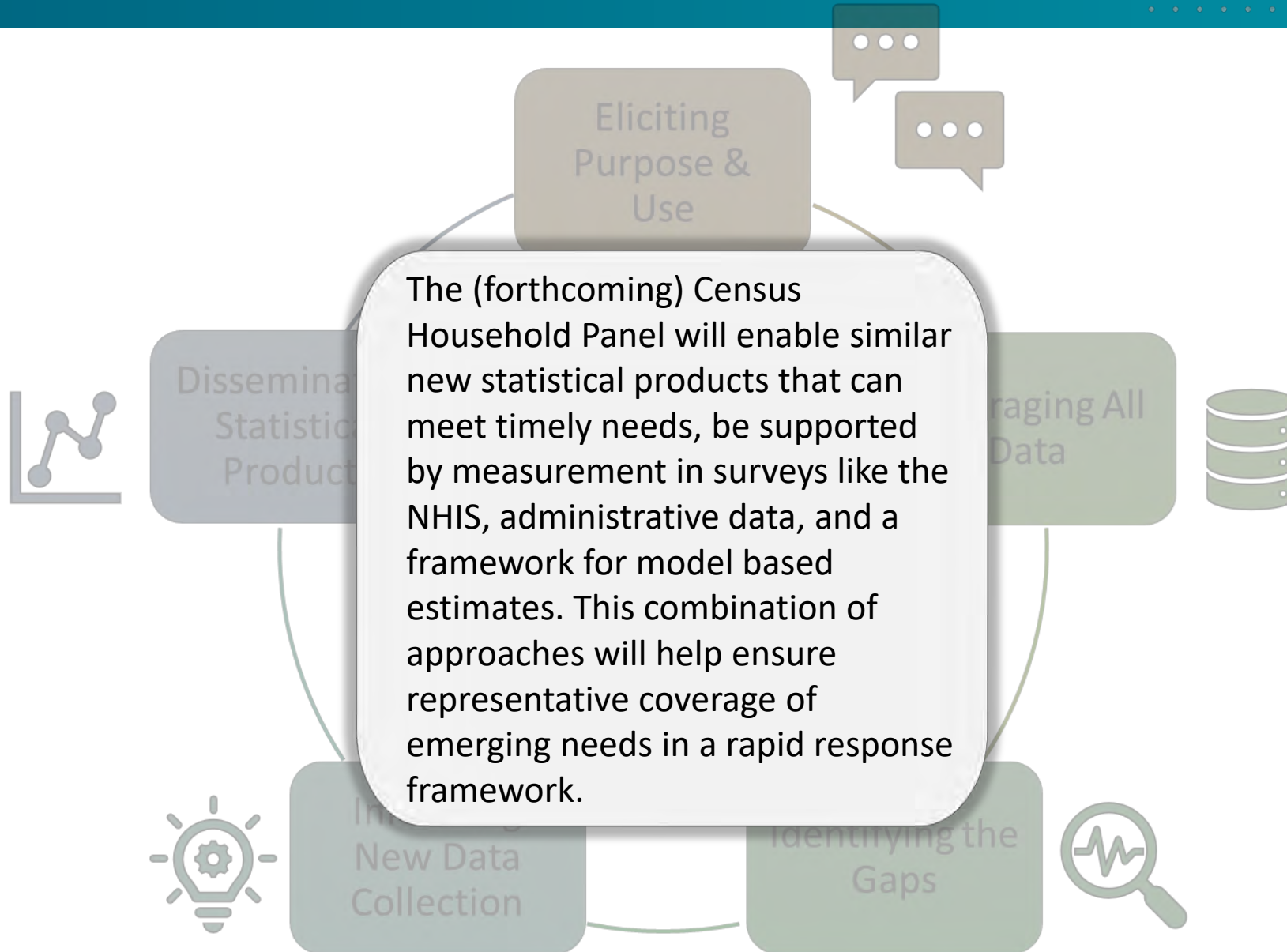
Eliciting Purpose & Use

Disseminating Statistical Products

Statistical Product Development

Leveraging All Data

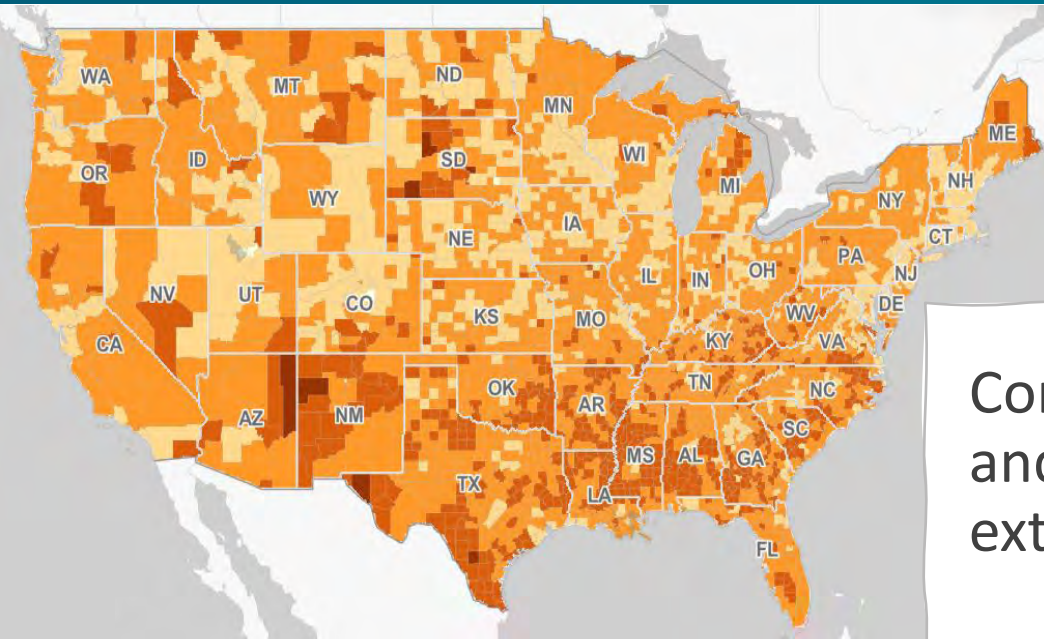Informing New Data Collection

Identifying the Gaps

Calibrating responses in the HPS to data in the NHIS through statistical modeling allows for a more complete picture of the mental health of the population, adjusting for both differential coverage (lower in the HPS) and mode differences in reporting (higher estimates of mental health issues in the self-administered HPS).

Disseminating Statistical Products

Leveraging All Data

Statistical Product Development

Informing New Data Collection

Identifying the Gaps

Eliciting Purpose & Use

Disseminating Statistical Products

...raging All Data

Incorporating New Data Collection

Identifying the Gaps

The (forthcoming) Census Household Panel will enable similar new statistical products that can meet timely needs, be supported by measurement in surveys like the NHIS, administrative data, and a framework for model based estimates. This combination of approaches will help ensure representative coverage of emerging needs in a rapid response framework.

Exemplar: Community Resilience Estimates (CRE) for Heat

Eliciting Purpose & Use

Community resilience is the capacity of individuals and households within a community to absorb the external stresses of a disaster.

The CRE was published as an experimental data product to provide information about the COVID-19 pandemic but also to garner feedback from data users on the quality and usefulness of the new product.
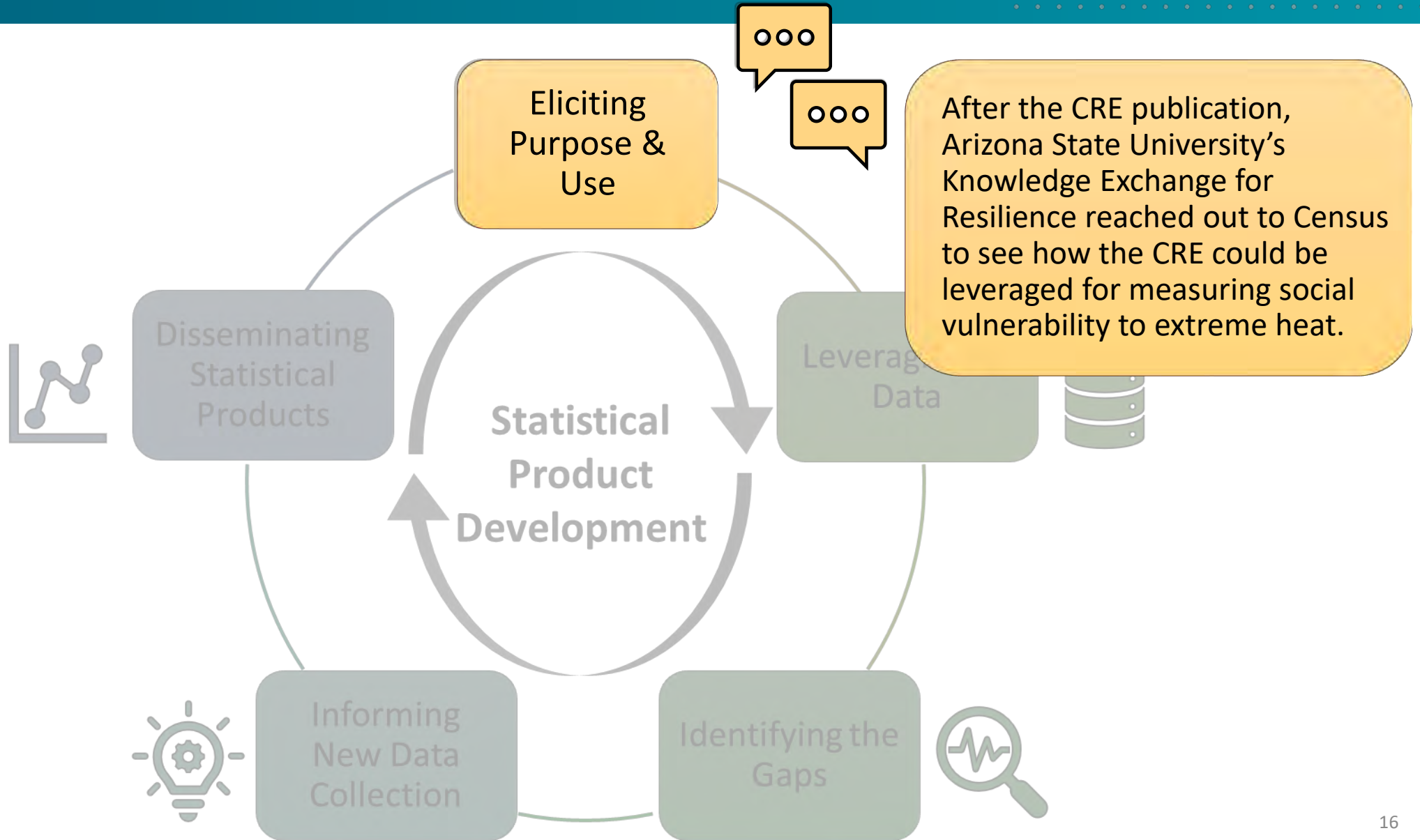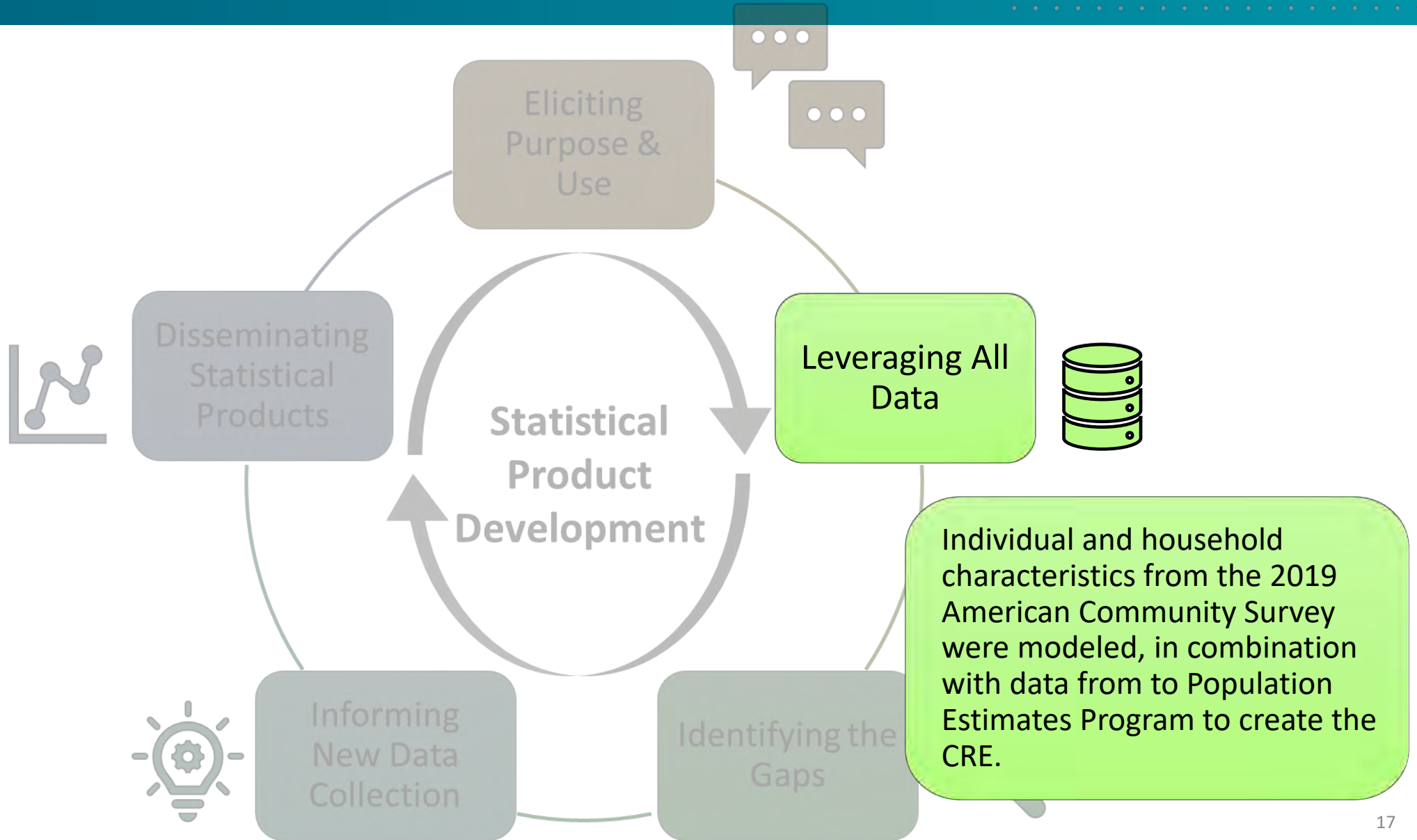
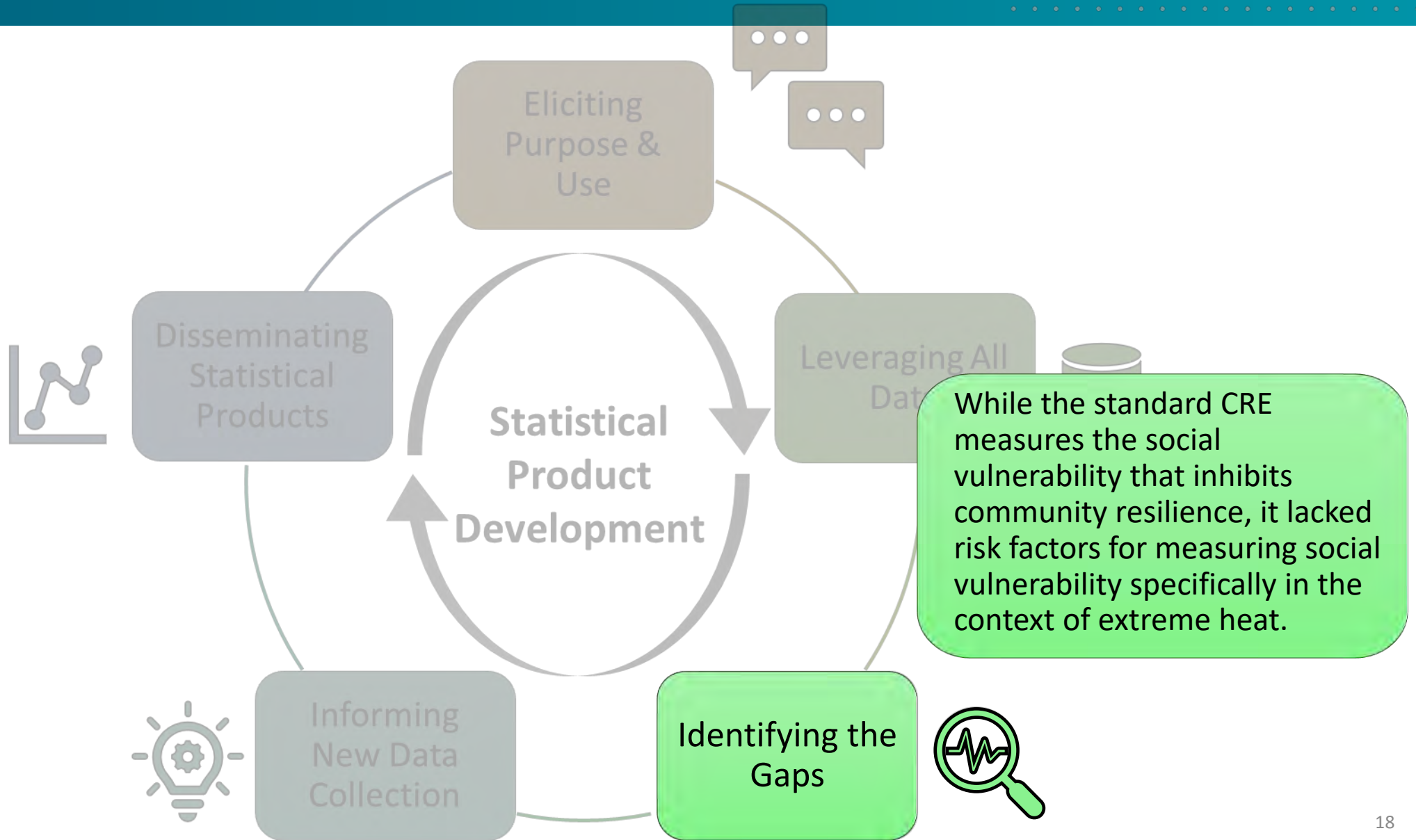Informing New Data Collection

Identifying the Gaps

Eliciting Purpose & Use

After the CRE publication, Arizona State University's Knowledge Exchange for Resilience reached out to Census to see how the CRE could be leveraged for measuring social vulnerability to extreme heat.

Disseminating Statistical Products

Leveraging Data

Statistical Product Development

Informing New Data Collection

Identifying the Gaps

Exemplar: CRE for Heat

Eliciting Purpose & Use

Disseminating Statistical Products

Statistical Product Development

Leveraging All Data

Informing New Data Collection

Identifying the Gaps

Individual and household characteristics from the 2019 American Community Survey were modeled, in combination with data from to Population Estimates Program to create the CRE.

Eliciting Purpose & Use

Disseminating Statistical Products

Leveraging All Data

Statistical Product Development

Informing New Data Collection

Identifying the Gaps

While the standard CRE measures the social vulnerability that inhibits community resilience, it lacked risk factors for measuring social vulnerability specifically in the context of extreme heat.

With consultation from Arizona State University's Knowledge Exchange for Resilience, the CRE was updated for vulnerability to extreme heat.

Three of the ten standard CRE risk indicators were modified, taking into account:

- housing structure type,
- commute type, and
- whether the household's housing costs were greater than 50%.
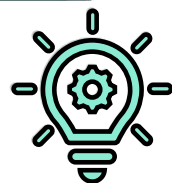
Eliciting Purpose & Use

Leveraging All Data

Statistical Product Development

Identifying the Gaps

Informing New Data Collection

Disseminating Statistical Products
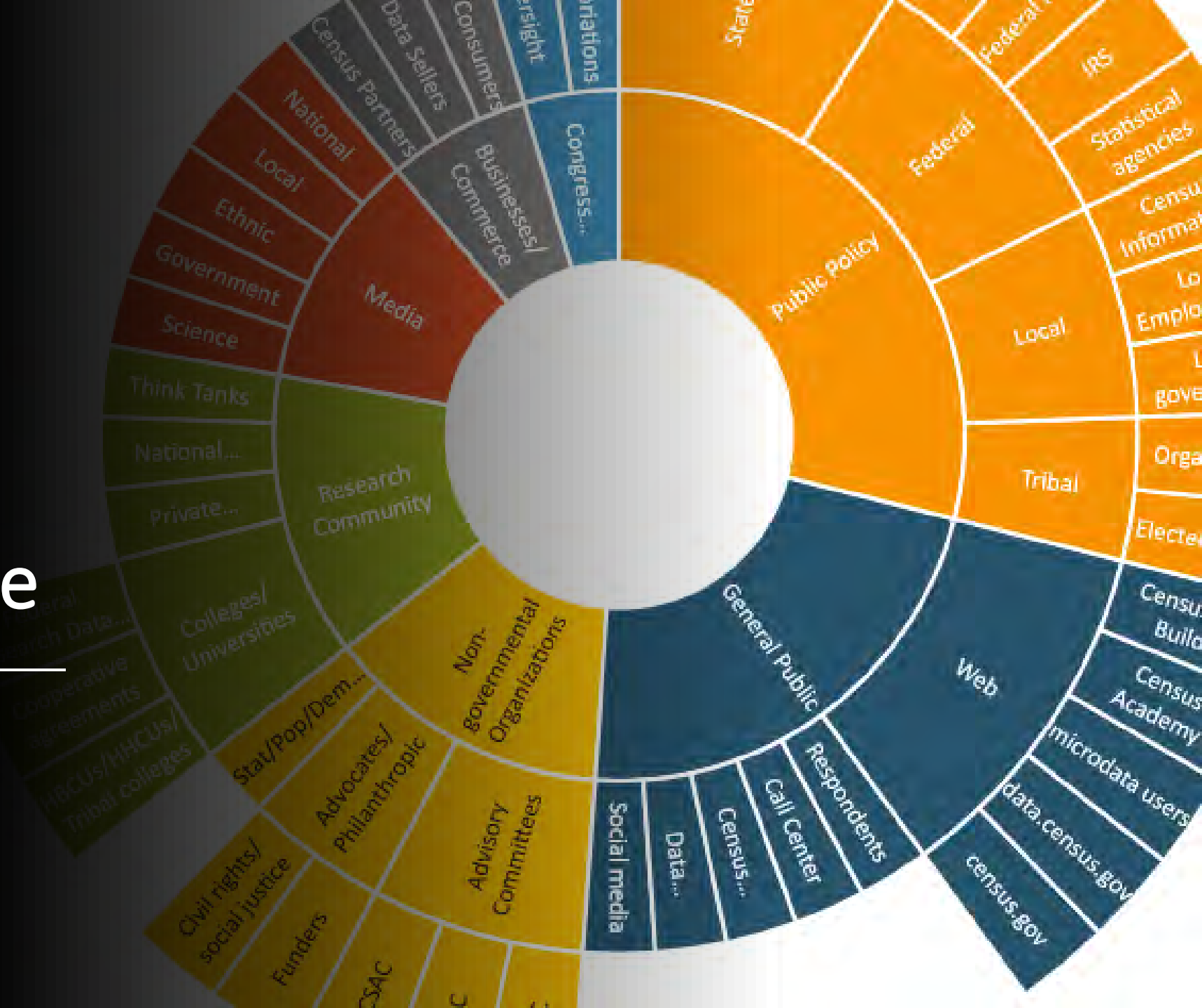
United States® Census Bureau

Exemplar: CRE for Heat

The CRE for Heat identifies the specific areas across the country most socially vulnerable to the effects of extreme heat. Local planners, policy makers, public health officials, and community stakeholders can use the estimates as one tool to help assess the potential resiliency of communities and plan mitigation and recovery strategies.
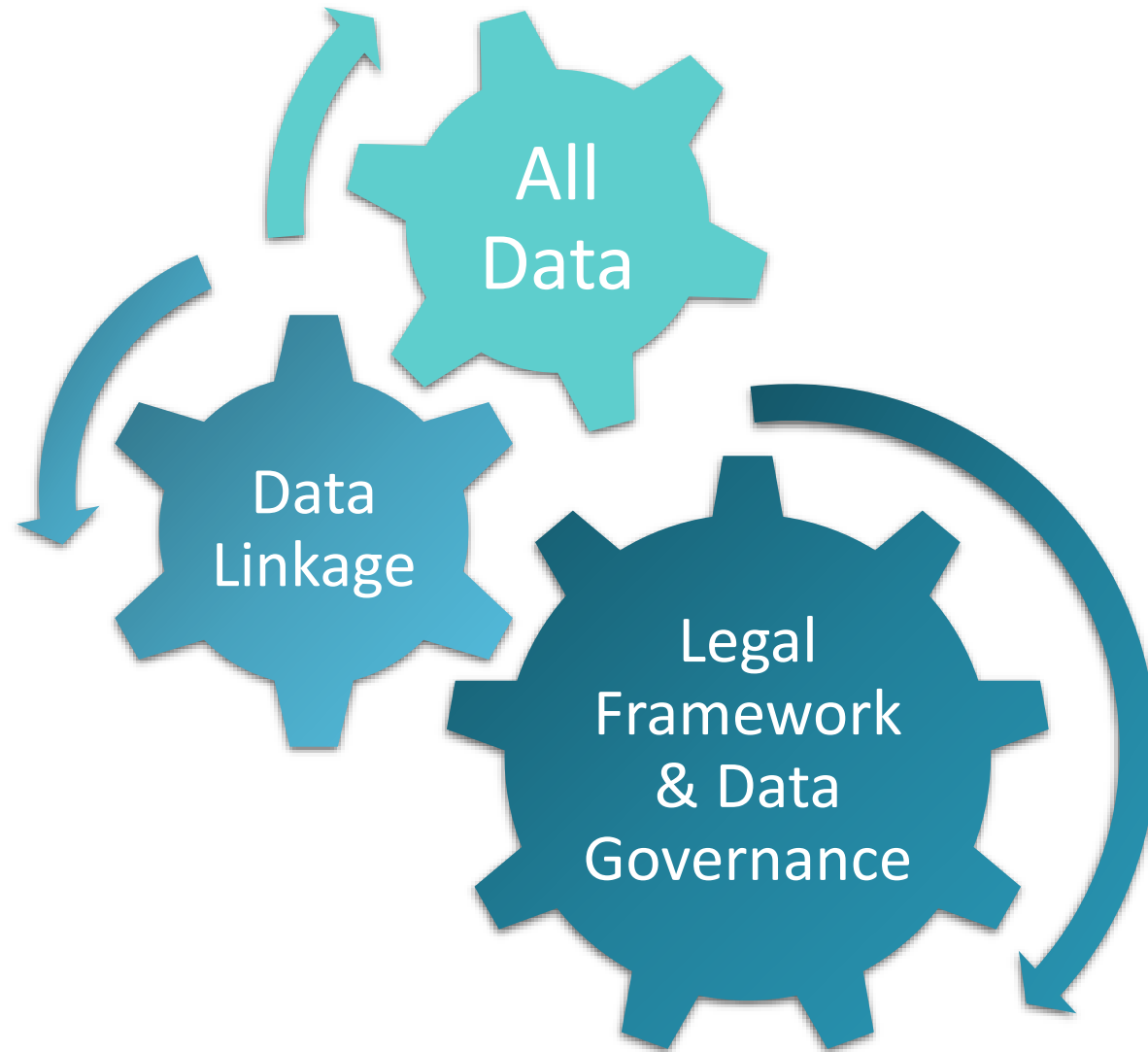
Disseminating Statistical Products

Leveraging All Data

Statistical Product Development

Informing New Data Collection

Identifying the Gaps

United States® Census Bureau

The CRE will continue to inspire new statistical products, including methods for the development of custom geographies as well as the integration of climate extremes and disaster data.

Through the use of small area modeling techniques, the CRE remains flexible and can be easily modified for a broad range of other uses, like economic recovery.

Illuminating the Ecosystem of Data Users to Elicit Purpose and Use

Leverage Data Infrastructure

All Data

Data Linkage

Legal Framework & Data Governance

United States® Census Bureau

Use *ALL* Data Assets

Going beyond the survey data we collect

Designed Data	Administrative Data	Opportunity Data	Procedural Data

Exploit Data Linkage

We collect a variety of linkage fields and tailor linkage methodology to each linkage type – people, places, jobs, and organizations.

Adhere to Legal Framework and Data Governance

Title 13 directs us to acquire and use external data records for statistical purposes.

Foundations for Evidence-Based Policymaking Act promotes and encourages data sharing.

Confidential Information Protection and Statistical Efficiency Act sets forth functional separation of statistical versus administrative uses of data.

Ethical Data Stewardship

- ✓ Ensure only statistical uses
- ✓ Disclosure review
- ✓ Punishing of wrongful disclosure
- ✓ Secure computing environments

United States® Census Bureau

26
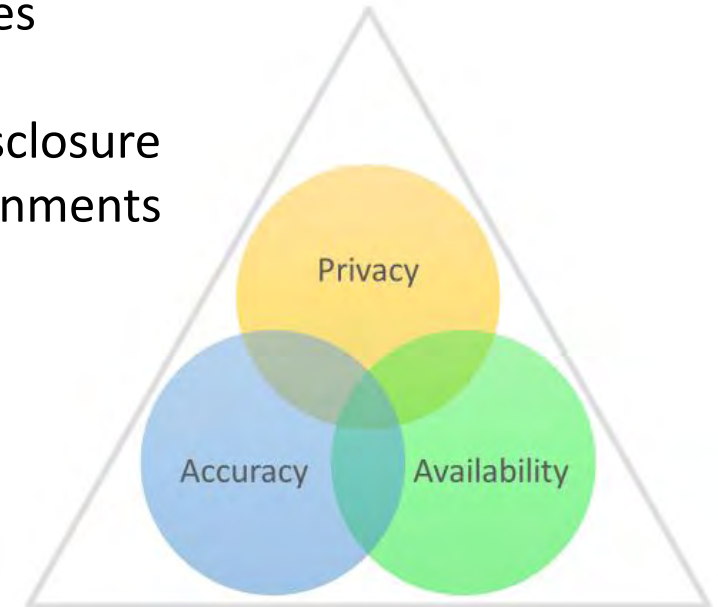
Statistical Product Spectrum for *ALL* Users

Protected
microdata
(in FSRDCs)

An opportunity
to identify and fill
the space in between
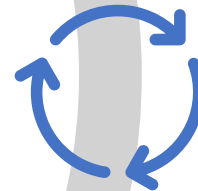
Simple, public
use products

Summing it up: A Curated Data Enterprise

Enabling Technologies for the Enterprise

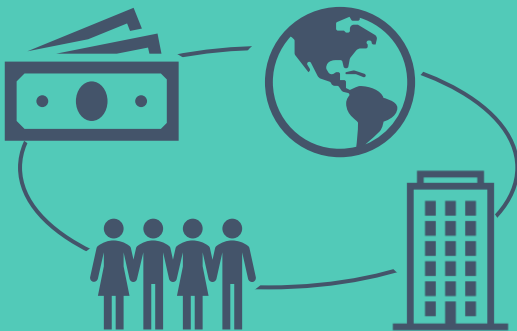**Data Ingest and Collection for the Enterprise (*all* data assets)**

Enterprise Data Lake
Data processing, computing, and management

Enterprise Linked Frames
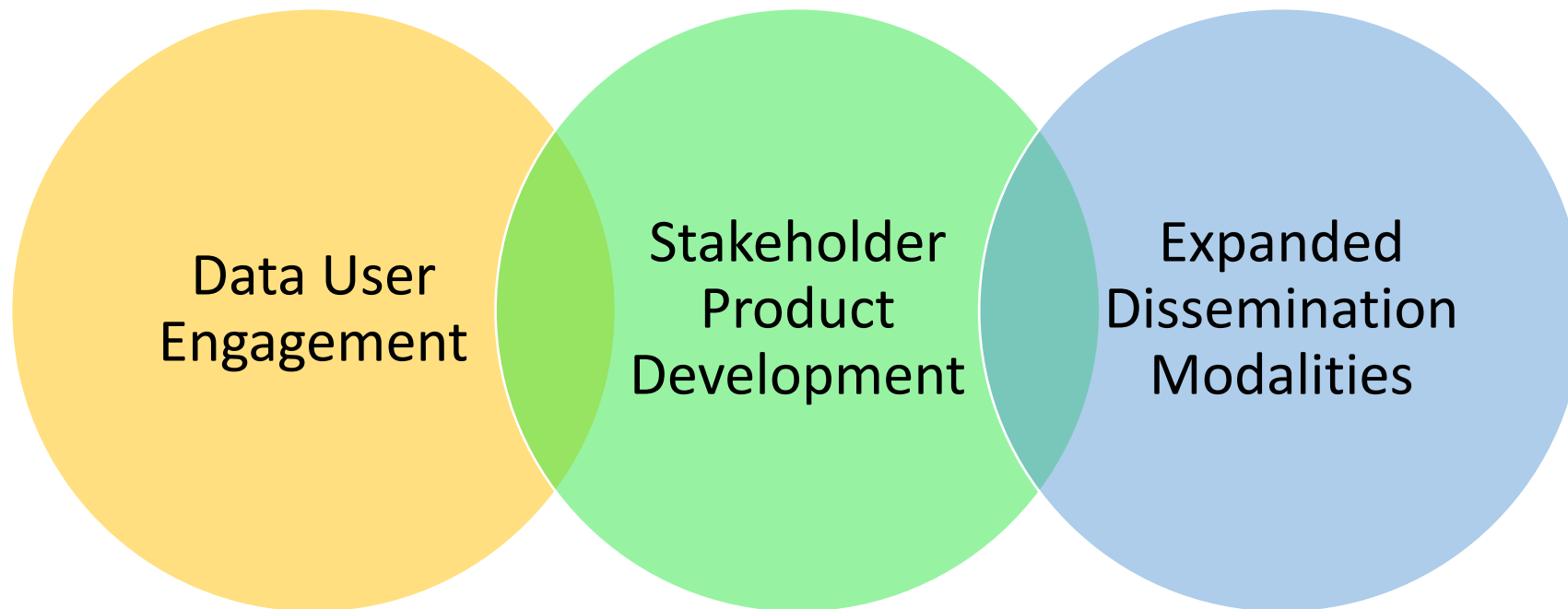4 seamless and linkable frames

Enterprise Dissemination Services

Moving this Endeavor Forward

Our current research efforts for the Statistical Product First approach are focused around three core elements:

Data User Engagement

Stakeholder Product Development

Expanded Dissemination Modalities

Statistical Product First Team Leads

Erika Becker-Medina,
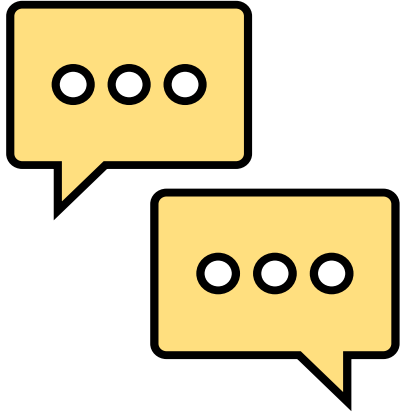**Data User Engagement**

Nate Ramsey,
Stakeholder Product Development

Emily Molfino,
Expanded Dissemination Modalities

Data User Engagement Update

- Conducted multiple external engagement sessions and have heard similar feedback from different stakeholder groups: *Census does a lot of engagement already—it's the feedback loop that needs work!*

- Planning an internal data user engagement workshop for all Census data user outreach program representatives to map current processes, identify gaps, streamline disjointed efforts, and engage more effectively.

- Reviewing state data center sites to identify existing user-friendly data tools and take inventory of what state offices are doing to avoid redundancies and identify statistical product gaps.

United States®
Census
Bureau

# Stakeholder Product Development Update

- Designing a demonstration pilot focused on a singular use case: identifying authoritative statistics to satisfy grant application requirements among state and local government and non-profit applicants.

- The demonstration pilot will encompass the three core elements: data user engagement, statistical product development, and expanded dissemination modalities.

- Similar strategies and techniques developed for this pilot could be applied toward a broader problem set (such as statistical needs among private sector organizations seeking government contracts.)

United States®
Census
Bureau

Expanded Dissemination Modalities Update
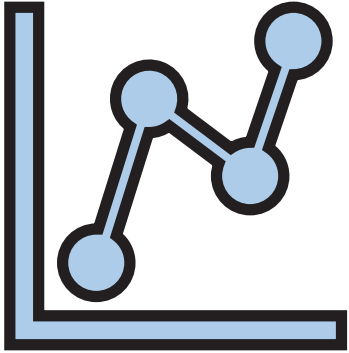
- Working with OpenAI and the Department of Commerce's Chief Data Officer to determine how to better deliver our data for use in AI platforms.

- Improving and expanding the contextual information about the data, putting it in a format that will be captured AI systems.
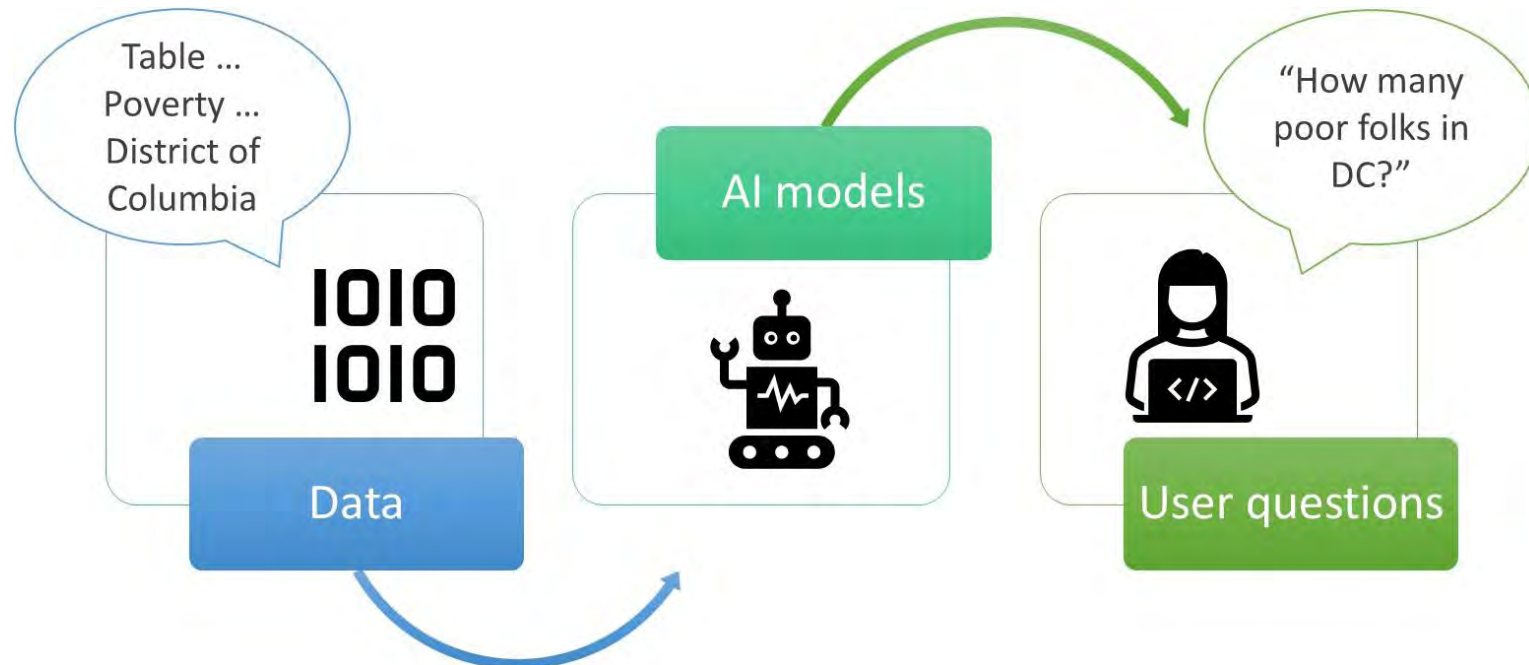
Imagine the Art of the Possible

We welcome your ideas!

➢ What are **major purposes** the statistical products need to support?

➢ What would **add value to your** statistical products?

➢ What **modalities of statistical product dissemination** support data user needs?

➢ **How do we elicit** that information?

➢ **Who should participate** in development of products?

➢ How do we ensure the **solicitation and development** of statistical products **happens equitably**?

Thank You

Erika Becker-Medina,
Senior Advisor to the Chief Scientist and Associate Director
Research and Methodology Directorate
U.S. Census Bureau

erika.h.becker.medina@census.gov