

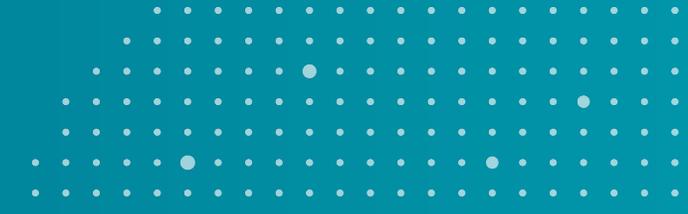
Introducing the Statistical Product First Approach

Erika Becker-Medina,

Senior Advisor to the Chief Scientist and Associate Director
Research and Methodology Directorate
U.S. Census Bureau

October 24, 2023

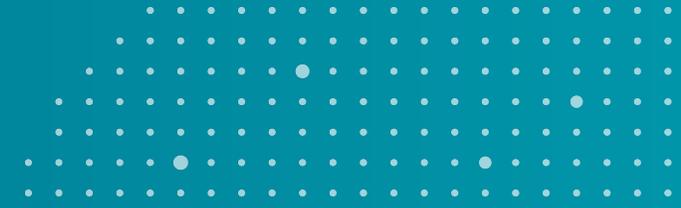
The views expressed in this perspective are those
of the presenter and not the Census Bureau.



Forces Driving Transformation Today

rapidly changing
new data sources
declining response rates
use of technology
increased demand for information
new data challenges to traditional
tools data collection
artificial intelligence methods

Forces Driving Transformation Today



*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, Ph.D., Deputy Director
U.S. Census Bureau



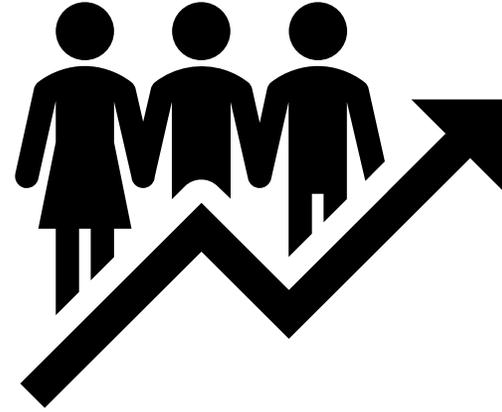
Flipping the Focus

Determine what information stakeholders need to reach their objectives

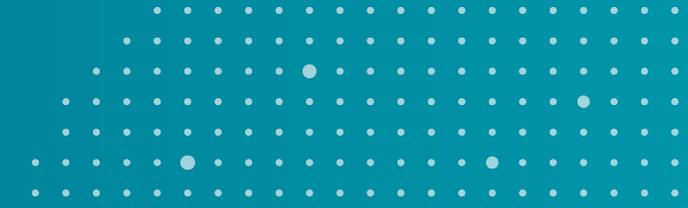
From there, shape the statistical products to be developed



Statistical
Product First
Approach



Statistical Product First Approach



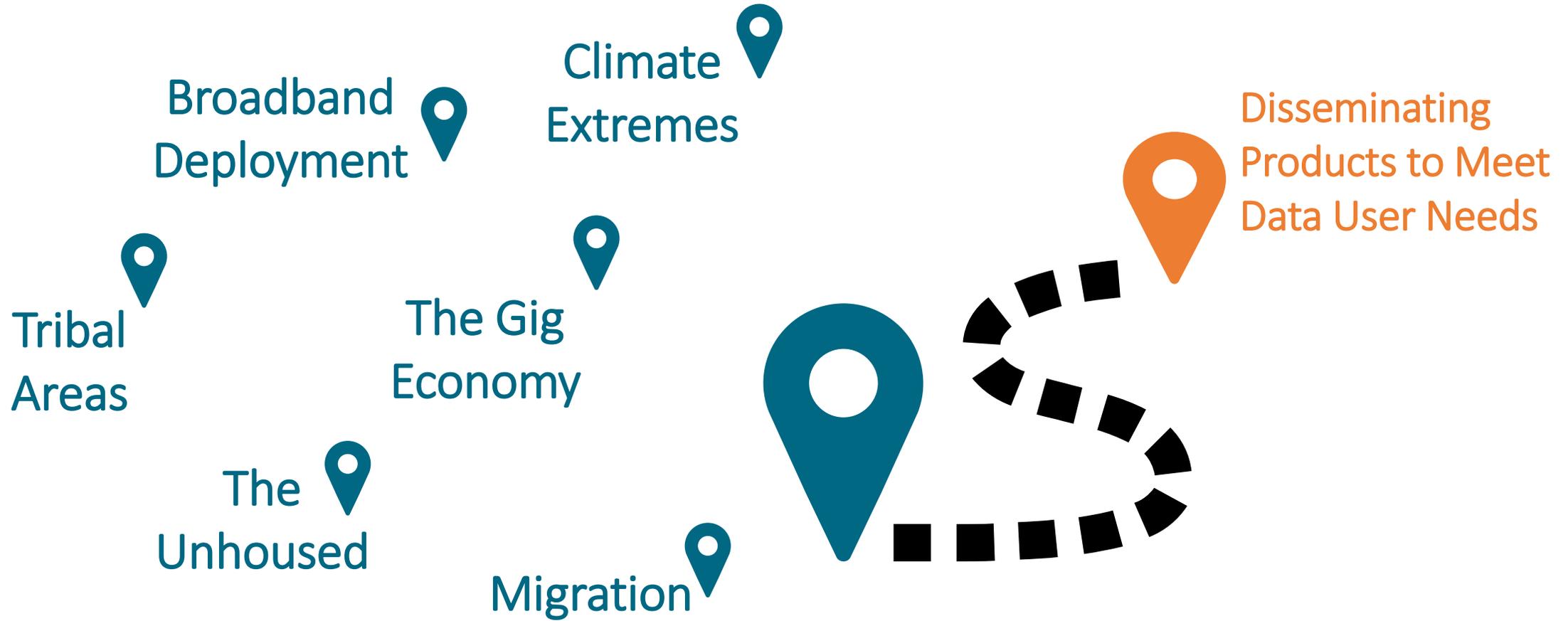
Ensures Data Support Purposes and Uses

Purposes
and Uses



Disseminating
Products to Meet
Data User Needs

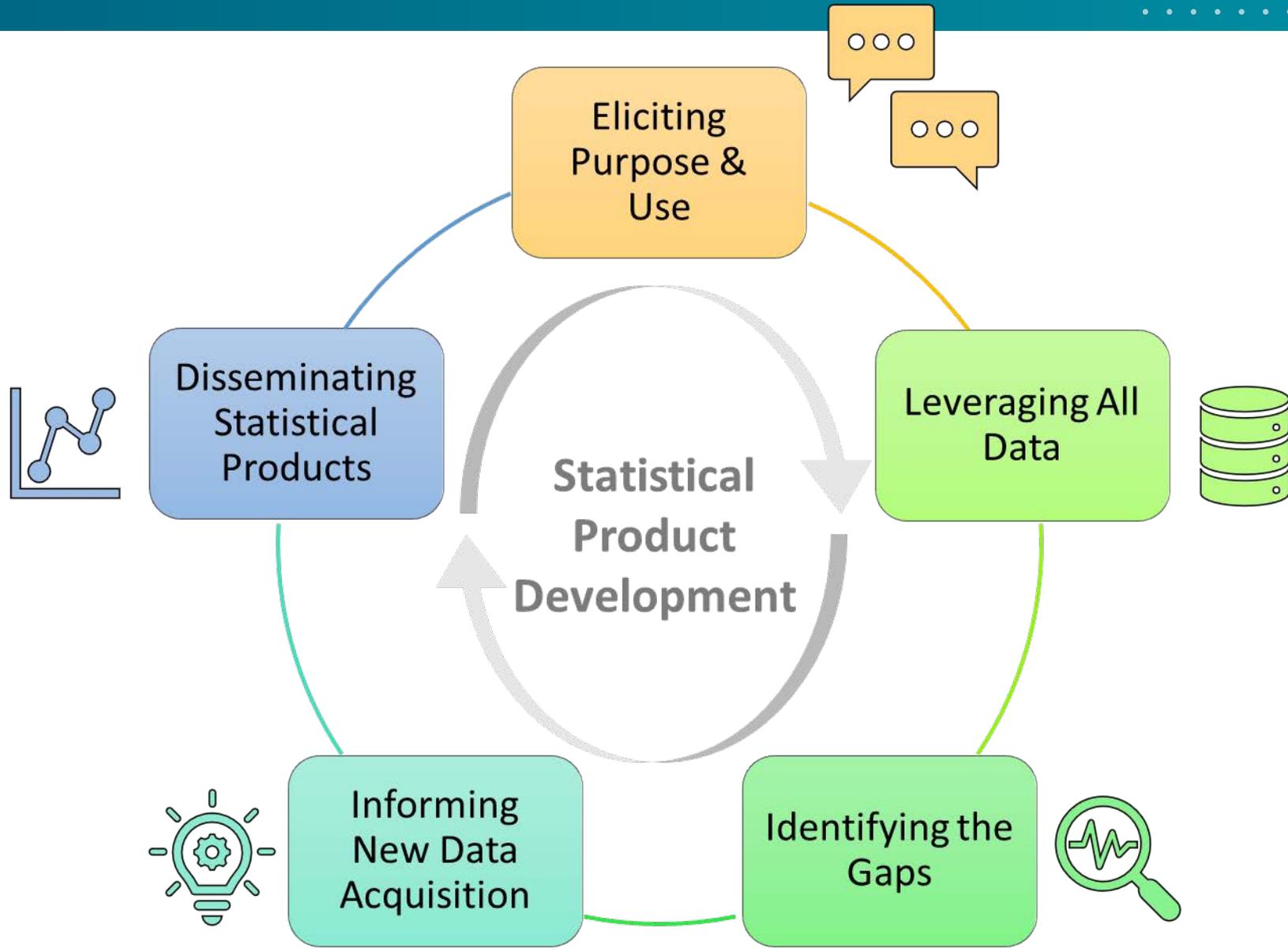
What are some purposes and uses?



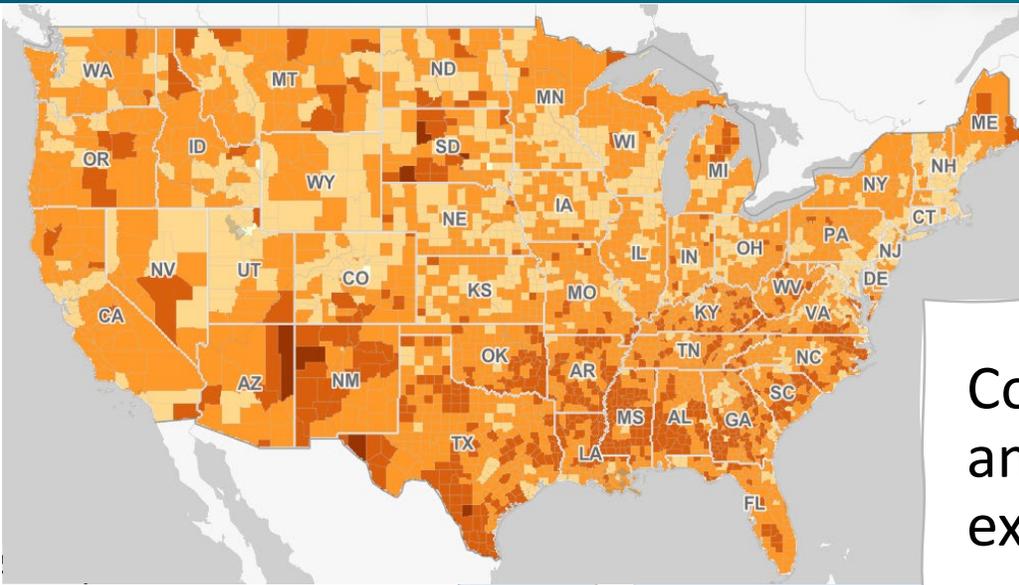
What are some purposes and uses?



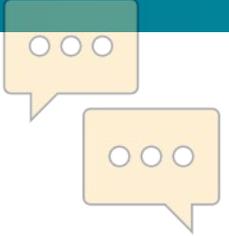
Statistical Product First Workflow



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.

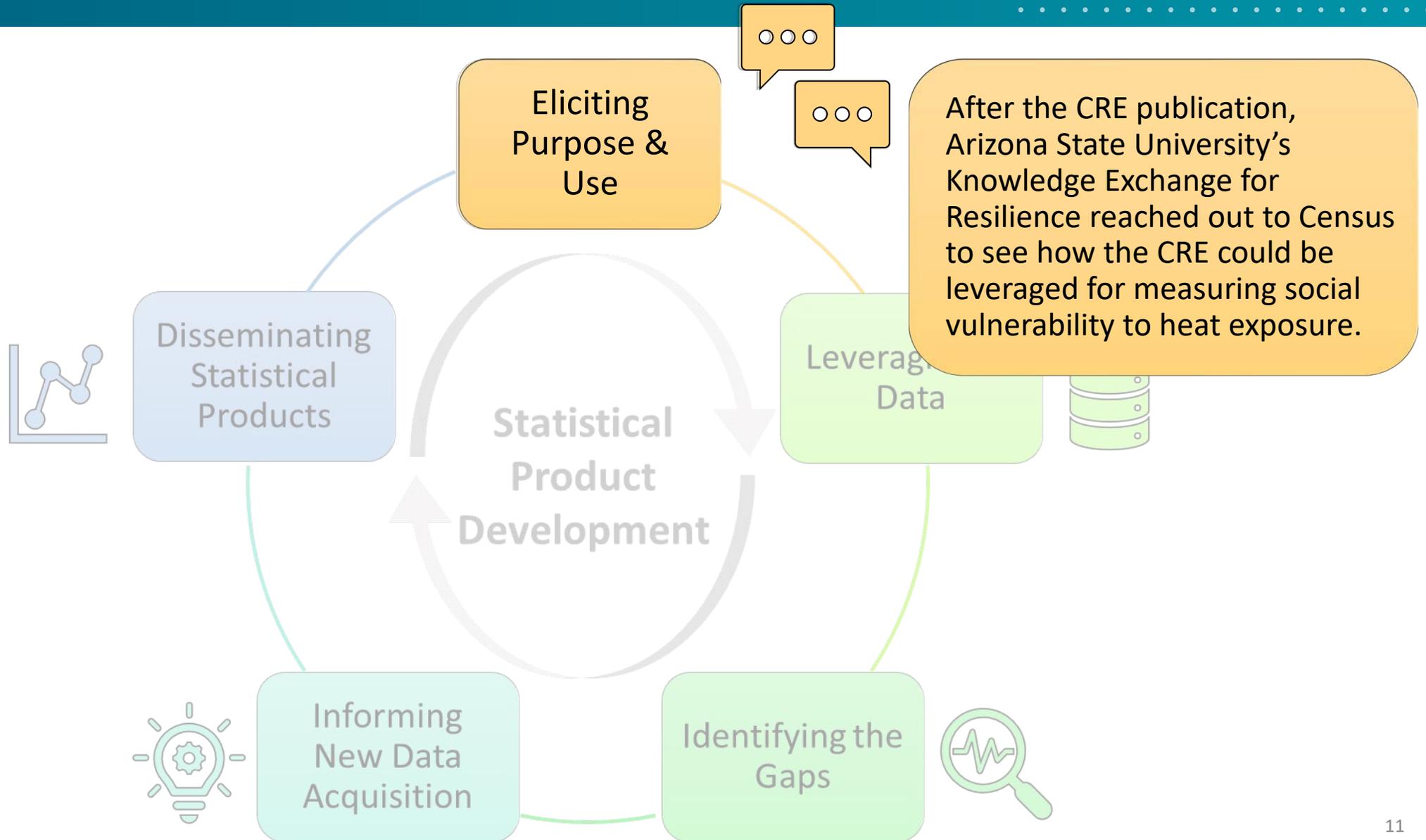


Identifying New Data Acquisition

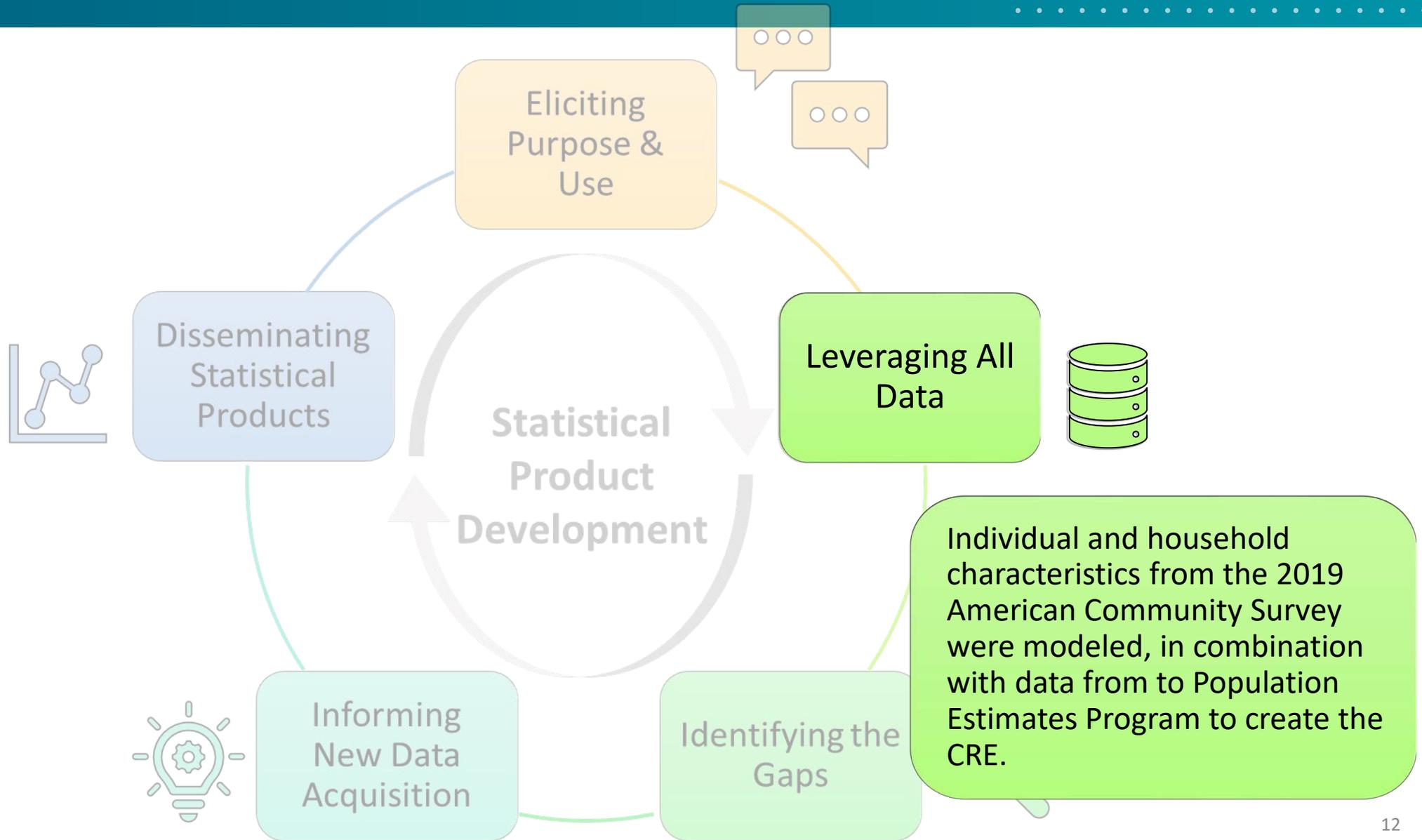
Identifying the Gaps



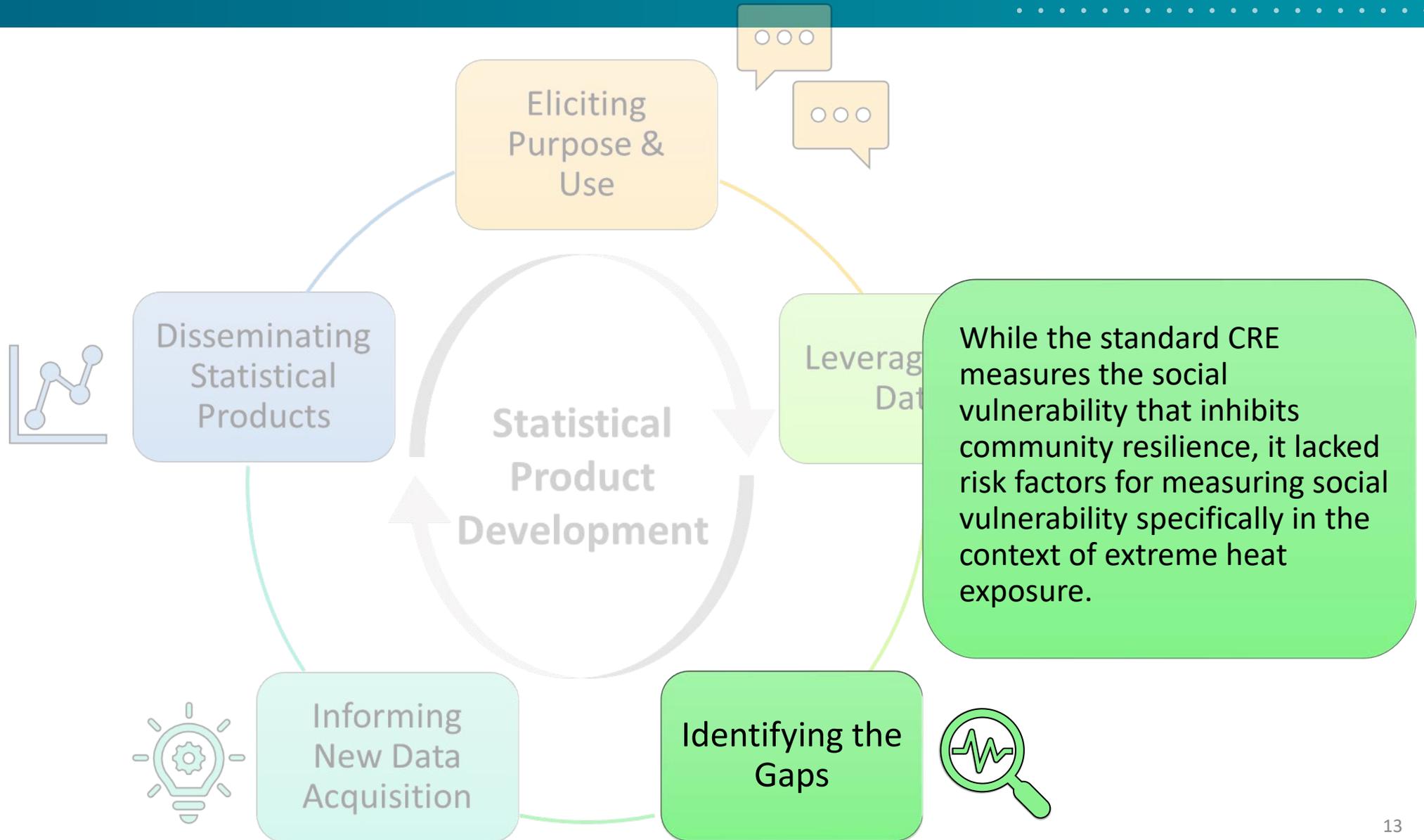
Exemplar: CRE for Heat



Exemplar: CRE for Heat



Exemplar: CRE for Heat



Exemplar: CRE for Heat

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

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%.



Informing New Data Acquisition

Eliminating Statistical Products

Statistical Product Development

Eliciting Purpose & Use



Leveraging All Data

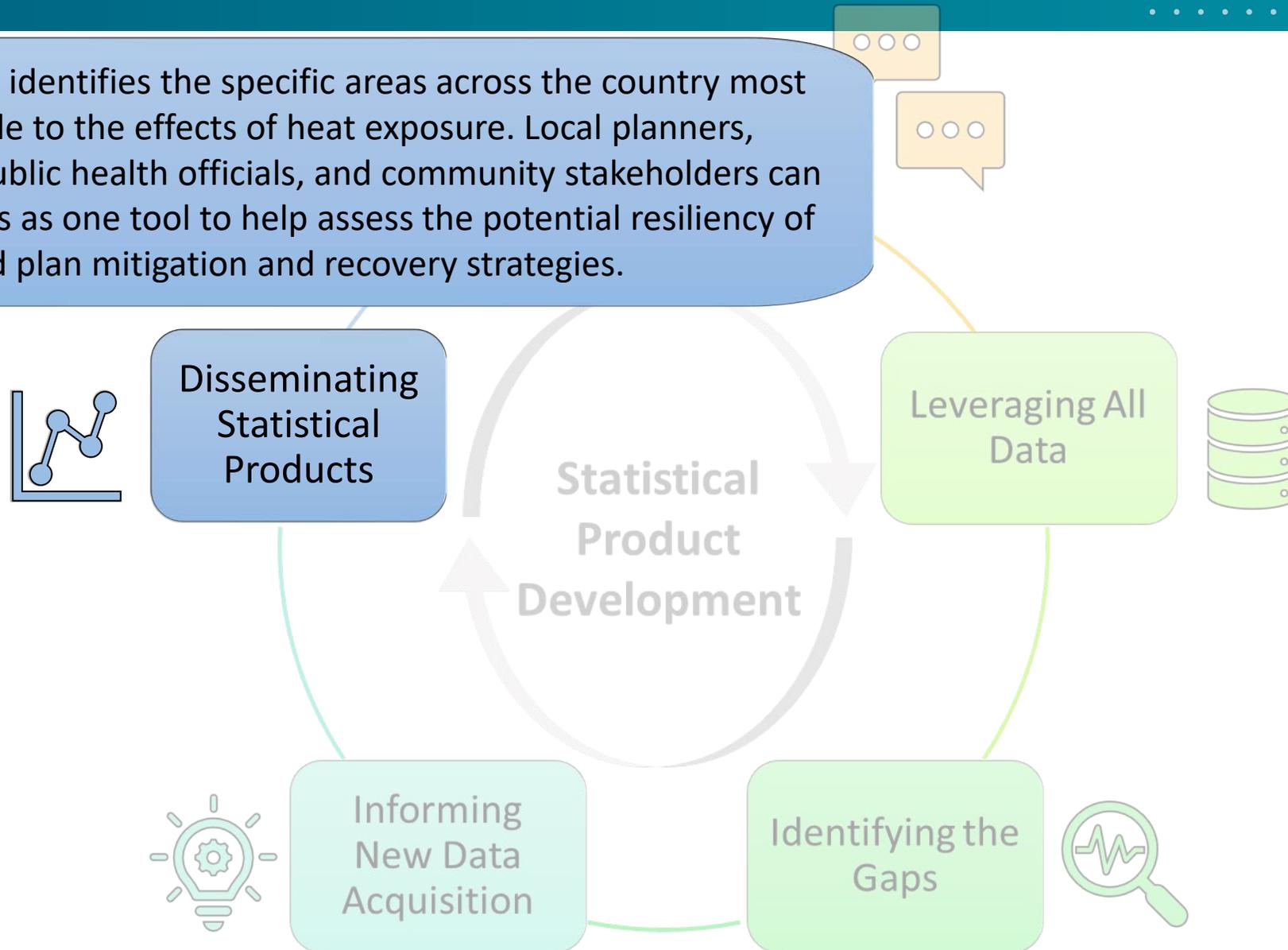


Identifying the Gaps

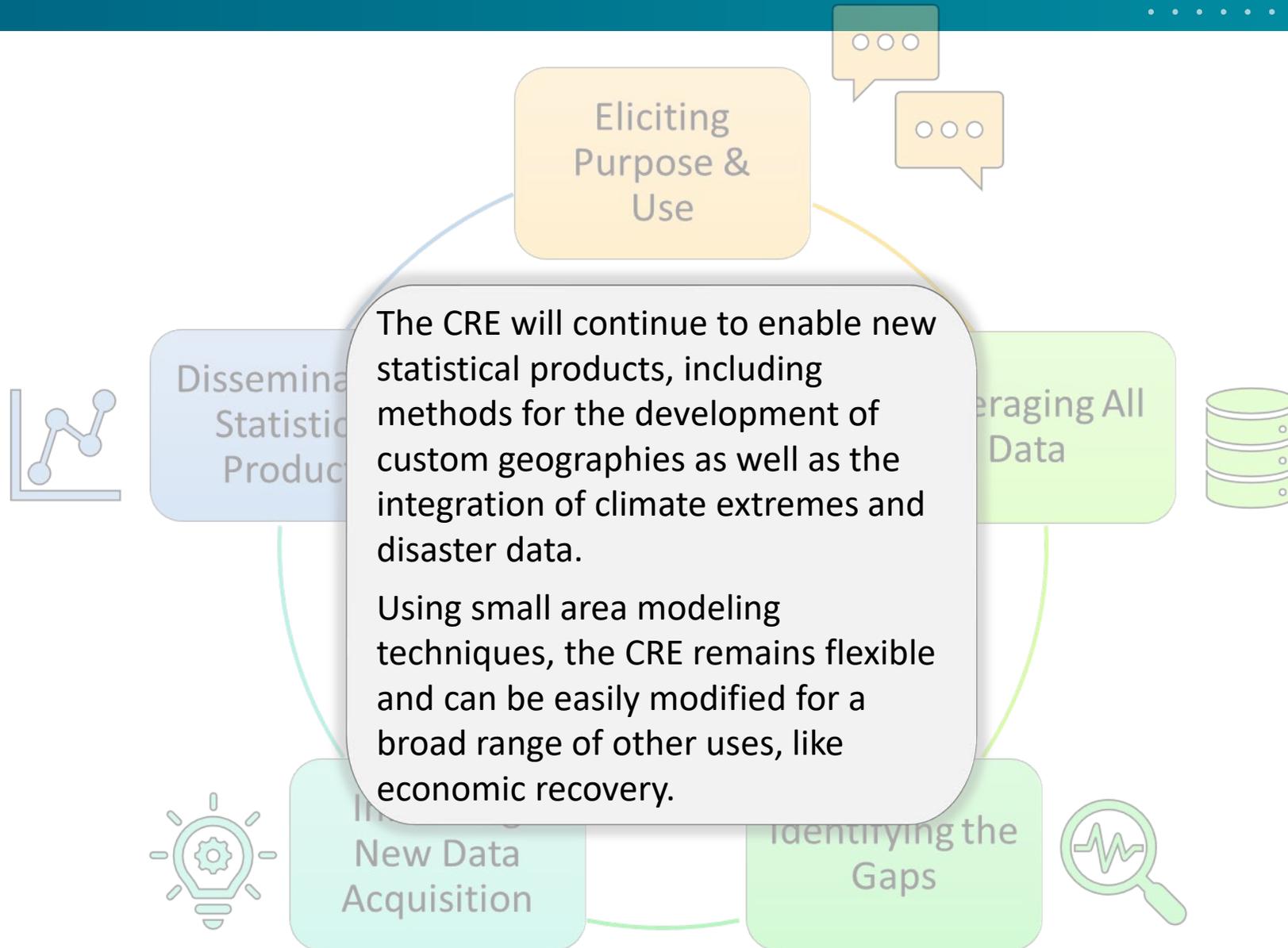


Exemplar: CRE for Heat

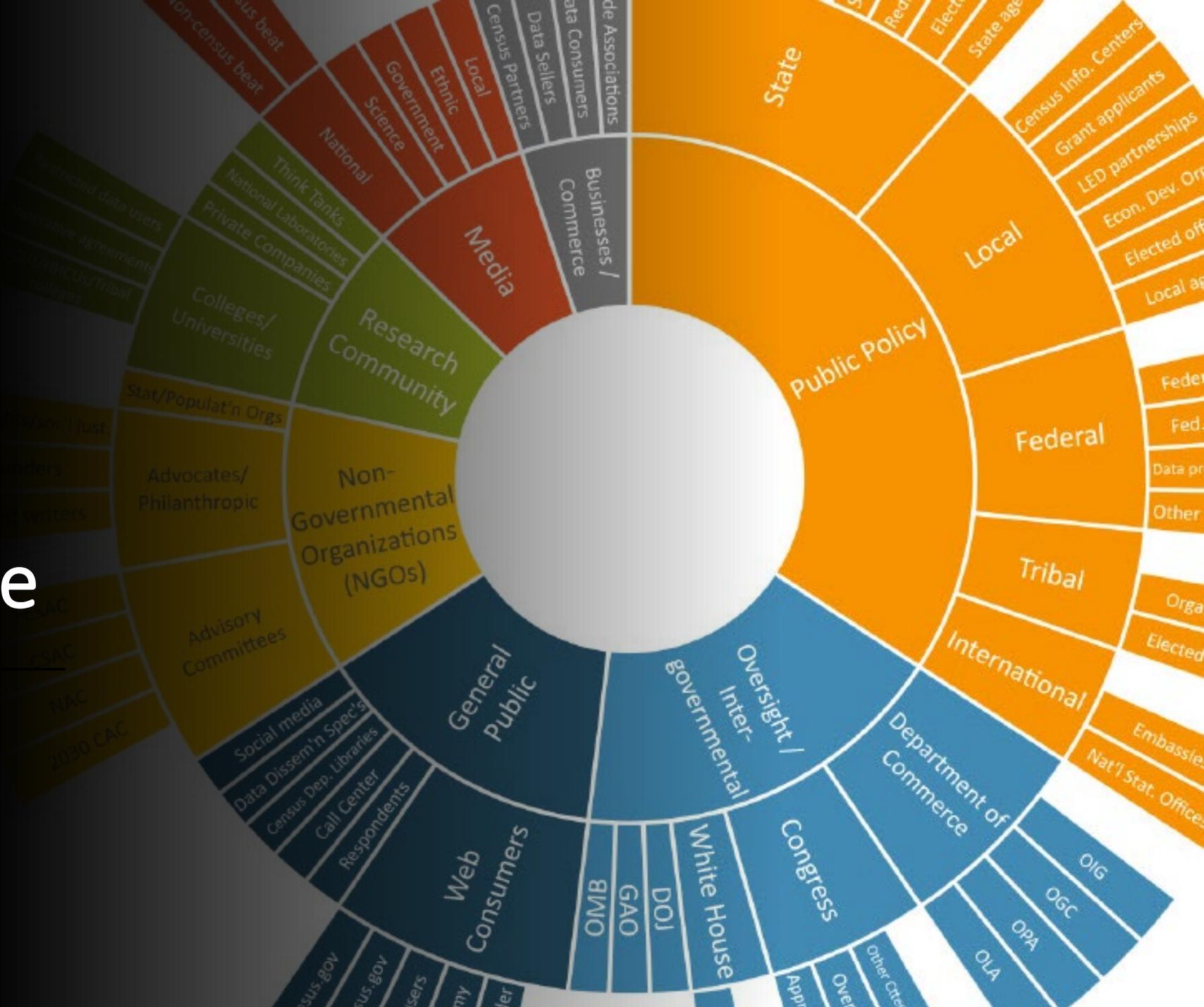
The CRE for Heat identifies the specific areas across the country most socially vulnerable to the effects of heat exposure. Local planners, policy makers, public health officials, and community stakeholders can use the estimates as one tool to help assess the potential resiliency of communities and plan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



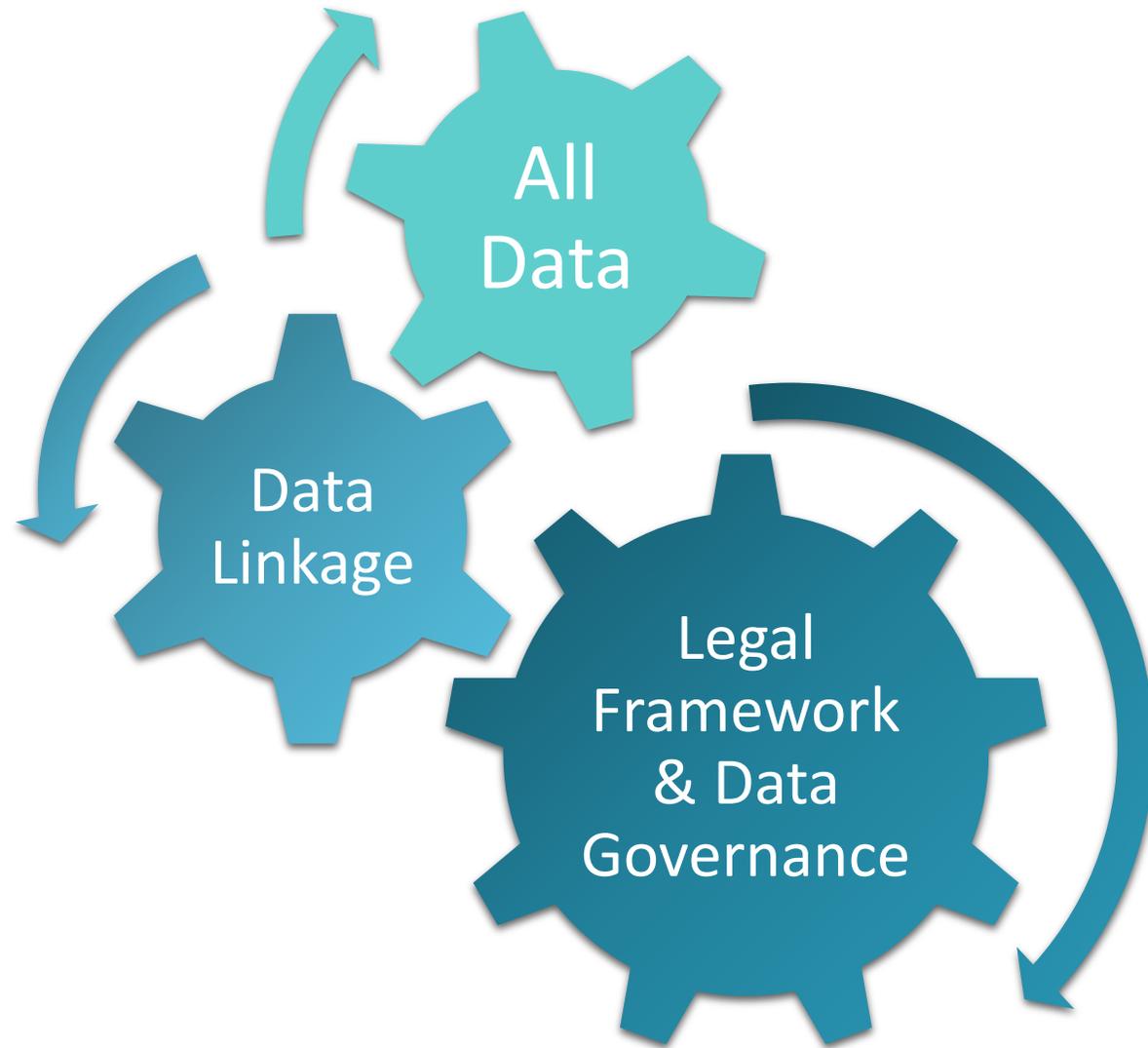
Exemplar: CRE for Heat



Illuminating the Ecosystem of Data Users to Elicit Purpose and Use



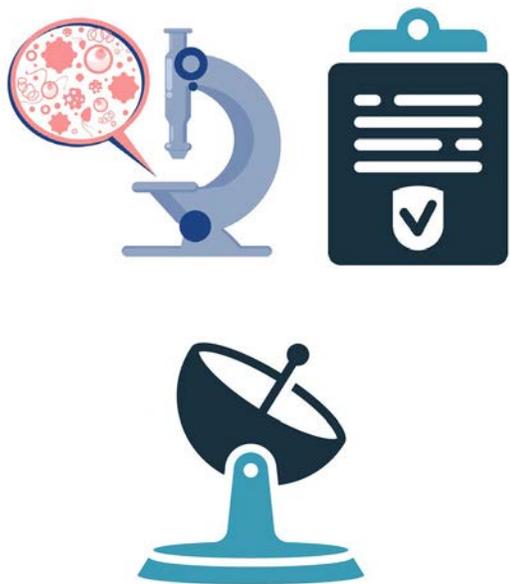
Leverage Data Infrastructure



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

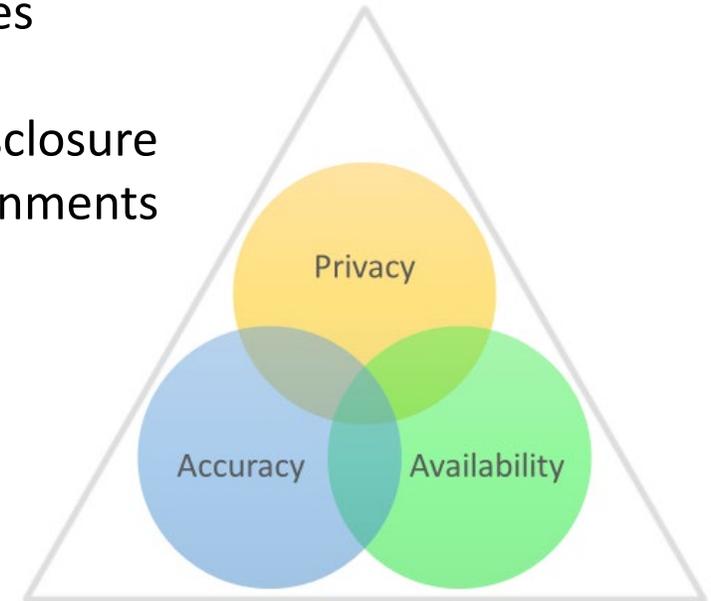
Title 13 directs us to acquire and use external data records for statistical purposes.

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

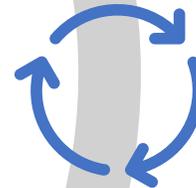


Statistical Product Spectrum for *ALL* Users

An opportunity
to identify and fill
the space in between



Protected
microdata
(in FSRDCs)



Simple, public
use products

Enabling Technologies for the Enterprise

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



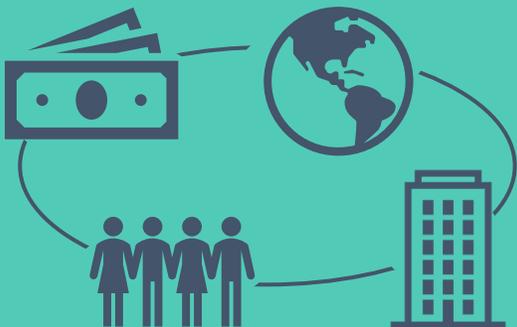
Enterprise Data Lake

Data processing, computing, and management



Enterprise Linked Frames

4 seamless and linkable frames



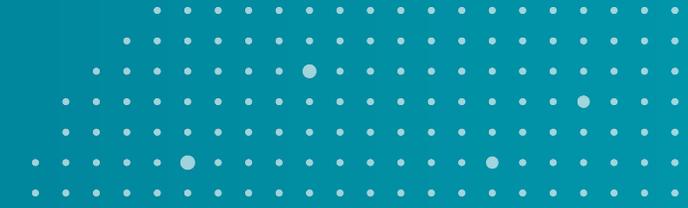
Enterprise Dissemination Services



Summing it up: A Curated Data Enterprise

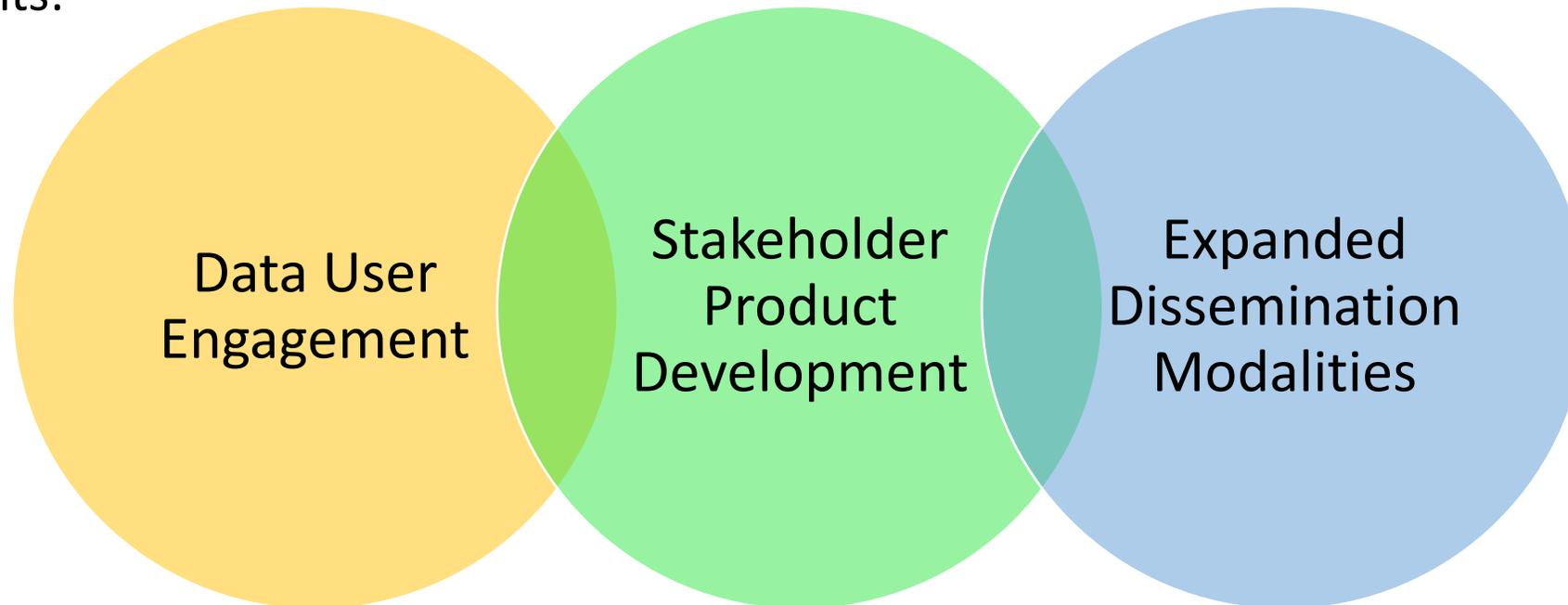


Moving this Endeavor Forward



At an agency level, we seek to determine what **types of Census Bureau data users** require what **types of statistical products** received through what **types of data dissemination mechanisms** over the coming years and decades.

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



Statistical Product First Team Leads

Erika Becker-Medina,
Data User Engagement



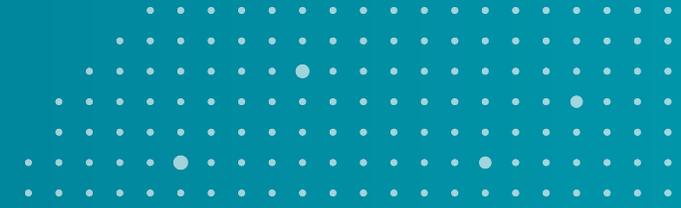
Nate Ramsey, Ph.D.
Stakeholder Product Development



Emily Molfino, Ph.D.
Expanded Dissemination Modalities



Data User Engagement



The Census Bureau has an abundance of data users and stakeholders, and we aim to serve them broadly, promptly, and equitably.

As part of this effort, we are:

- Illuminating the ecosystem of data users and stakeholders to elicit the purpose and use needs our statistics are to support.
- Segmenting data users across a multitude of factors, on an ongoing basis.
- Approaching different segments of this ecosystem in ways that will make for effective solicitation.
- Engaging data users in the statistical product development.
- Leveraging what they already do so we *add* value and are *not* duplicative.
- Implementing solutions to ensure purpose and use needs are being met.

Stakeholder Product Development

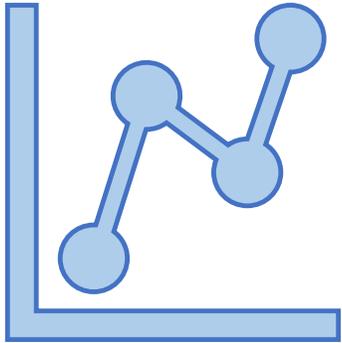


We need to address what types of statistical products data users require—from static tables to FSRDC microdata access.

As part of this effort, we are:

- Leveraging our experience in experimental product development.
- Determining the key steps for developing purpose and user-driven statistical products.
- Designing and implementing rapid demonstrations.

Expanded Dissemination Modalities



We need to determine what types of dissemination mechanisms data users require and expand dissemination modalities to serve diverse levels of data sophistication—from the novice user to the power user.

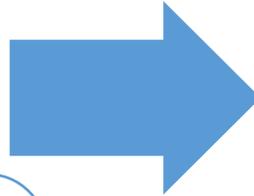
As part of this effort, we are:

- Determining what modalities of statistical product dissemination best support data user needs—including tiered access.
- Determining how to maximize data access to meet different levels of user needs—including the implications of an expanded FSRDC system.
- Improving and expanding the contextual information about the data.

What do we mean by tiered access?

Current status

- Only two options:
 - Public Products
 - Confidential Access
- Many drivers toward expansion and ease of access
 - Equity Concerns
 - Open Data
 - Evidence Building
 - Reproducibility



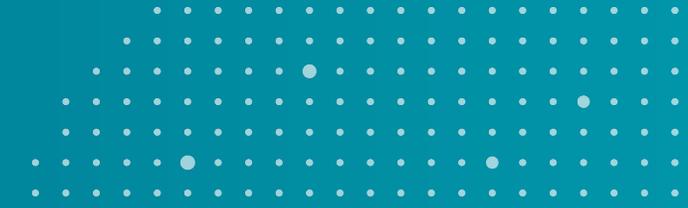
Future State

- Offer data users, providers, partners, and researchers standardized access options along a spectrum of dissemination modalities
- Recognize and accommodate varying data sensitivity assessments and risk tolerance levels for partner agencies
- Invest in solutions across the spectrum



Statistical Product First
Selected Demonstrations

Grant Applications Demonstration



We're conducting demonstration projects focused on statistics to support federal grant applications.

Teams from Census and the Georgetown Data Science Program are working with us.

Identifying	Identifying authoritative statistics to satisfy grant application requirements among state and local government and non-profit (SLGNP) applicants.
Creating	Creating statistical products based on grant application requirements that combine multiple data sources to better serve SLGNP data users.
Deploying	Exploring dissemination options to maximally benefit SLGNP grant applicants in need of authoritative statistical information.

State Demographers + State Data Centers Demonstration

We're developing tiered access demonstration based on engagements with Federal-State Cooperative on Population Estimates (FSCPE) and State Data Centers (SDC).

We anticipate this will include tabular output projects with restricted data, expanding user base to evidence builders—including state governments.

Reviewing

Reviewing SDC websites to gain insight into how SDC network *leverages data* and capturing the statistical tools they use and produce.

Looking

Looking for *early adopters* to participate in the demonstration and help characterize broad purpose and use needs.

Building

Building a *tiered access model* (FSRDC-light) to support demonstration focused on special tables and integration of local data.

AIAN Product Co-Development Demonstration

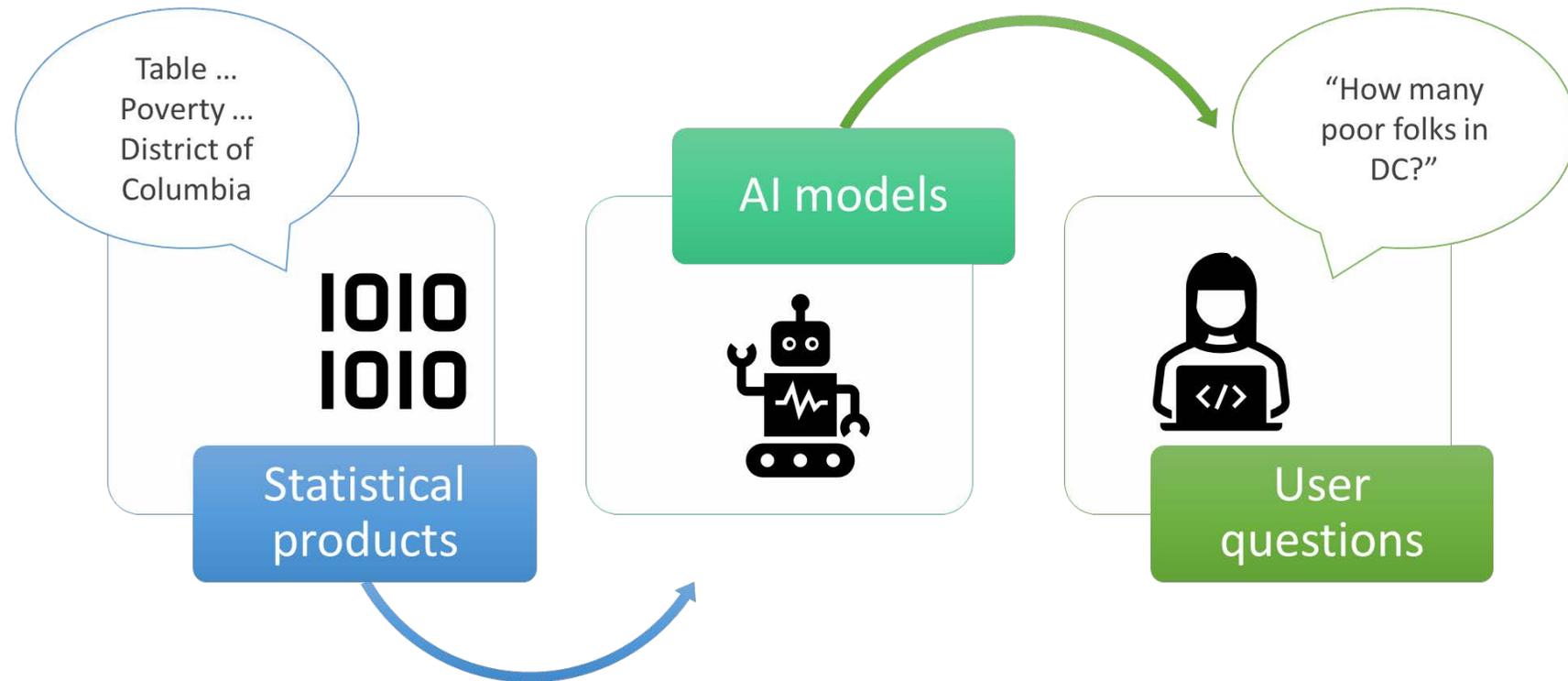
We plan to engage American Indian and Alaska Native (AIAN) stakeholders to collaborate on the design of new statistical products to meet their needs.

We'll be in Indian Country and Native Alaska this fall, beginning with the NIEA, NCAI, and BIA Providers conference to begin engagement and collaboration.

Engage	Engage with American Indian and Alaska Native (AIAN) community to seek input on what <i>questions and issues</i> they need Census statistical products to support.
Collaborate	Collaborate on the design of future statistical products designed <i>specifically</i> with the AIAN community in mind.
Leverage	Leverage existing data assets and identify new data opportunities to support the <i>co-development</i> of future statistical products.

AI Readiness Application

- Determining how to better deliver Census statistics for use in *AI platforms*.
- Enabling the *lay person* to engage with Census statistical products through large language models.
- Improving and expanding the *contextual information* about statistical products; formatting them so they will be captured by AI systems.
- *Targeting ACS* dissemination as early exemplar.





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