

US Census Bureau Final Presentation

April 22, 2016

Prepared by Luciano Arango, Nidhi Badaya, Aaron Capizzi, Rebecca Scharfstein, and Peter Willis



client



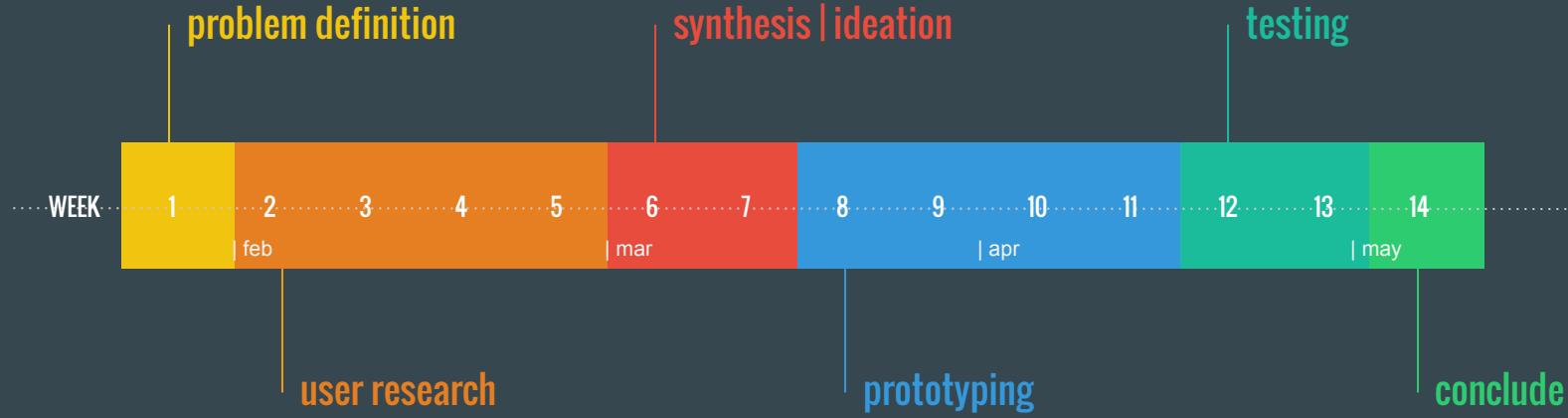
The Census Bureau, part of the Department of Commerce, serves as the leading source of quality data about the nation's people and economy. We honor privacy, protect confidentiality, share our expertise globally, and conduct our work openly.



problem statement

How might we improve **user experience** - search, retrieval and incorporation - of US Census data, particularly **income distribution** data?

timeline



USER RESEARCH



methods

Our research aimed to understand how users find, retrieve, and use data.

We engaged in a three step process in order to identify user personas:

- (1) identify and interview user groups;
- (2) synthesize findings into key insights; and
- (3) articulate user personas that cut across user groups.

user group focus

Our preliminary research focused on three user groups with high leverage and practicality.



includes 2 site visits



key insights

We generated 8 key insights from our user interviews:

Data

1. Users need data to do their jobs properly.
2. For users, data is personal and local.
3. Users go to extraordinary lengths to get the right data.
4. Manipulating data takes work, users don't expect shortcuts

Census

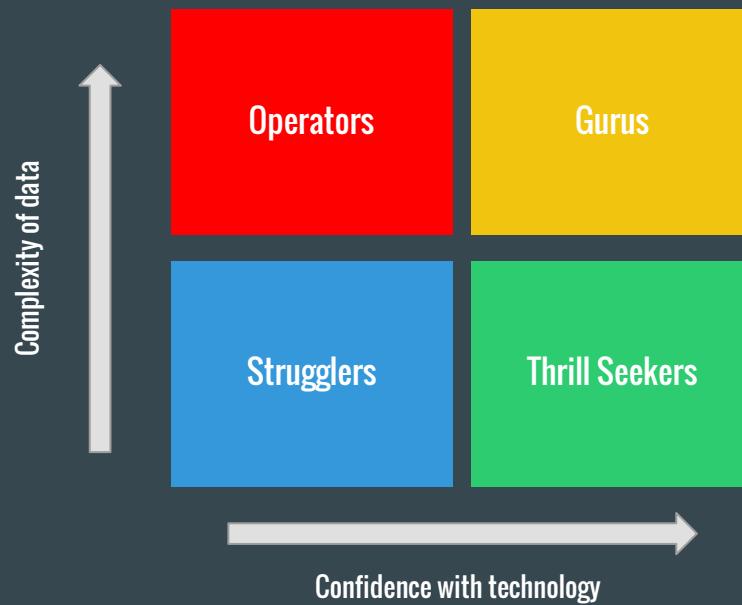
5. Census data can overwhelm and deter users.
6. In the minds of users, Census data exists in an ecosystem.

Trust and support

7. Users want assurance they are using data properly.
8. Users value personal support.

personas

To create personas we segmented users according to their confidence with technology and the complexity of the data they use.



PROTOTYPING



converging on a single problem

We used four steps to narrow our focus from a wide range of users, insights, and possible solutions down into three prototype options

1

Brainstorm
solutions

2

Group by
insight(s)

3

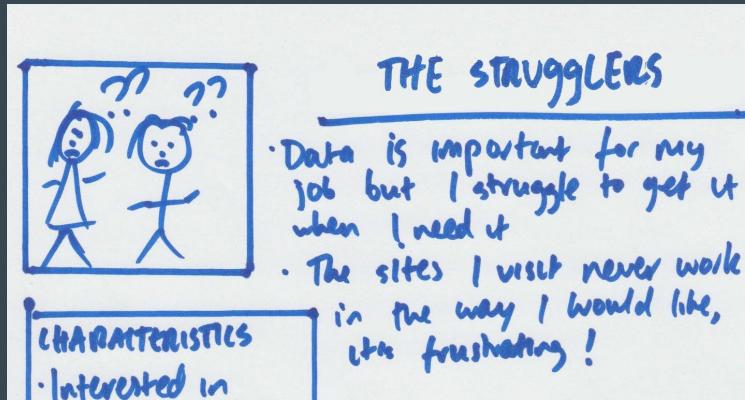
Rank solutions
by practicality

4

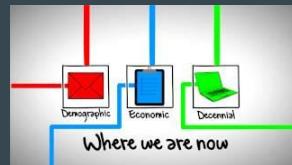
Aggregate into
prototype forms

strategy behind the prototype we chose

Improve the experience for users



Support Census modernization efforts



CEDSCI



Customer relationship mgmt.



Content migration

reimagined income inequality page

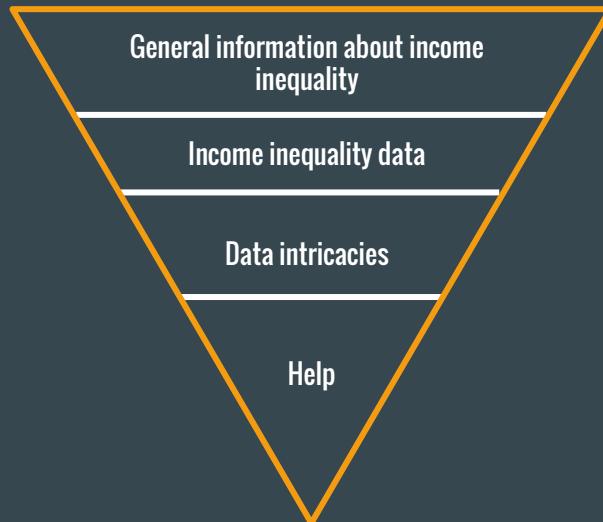
j.mp/inequality-page



tracing user insights to page features

“Daunting” / “I don’t know what they have! ”

- Single page layout
- Organized by increasing complexity



The screenshot shows a complex single-page website with the following sections:

- Header:** Income Inequality in America
- Top navigation:** Learn, Get the Data, Contact Us
- Learn section:** What is Income Inequality?, Getting Started, Income Data
- Get the Data section:** Income Inequality Report, Census Explorer, Maps of Inequality
- Community Data section:** What Our Community Has Built, New York Times, The Economic Policy Institute, Bloomberg
- Support section:** General Support, Economic Justice, Census Public Information, Create New Media & Content
- Footer:** Get the Data, Footer links (FactFinder, CitySlate, The Census Data), and a footer note: "Where Census Data Comes From".



insights to page features cont.

“I would prefer an easy source but it’s not always possible”

Census data exist in an ecosystem



Learn

What Is Income Inequality?

Income inequality refers to how evenly or unevenly income is distributed in a society. When inequality is high, a small group of people earn most of the income.

In an equal society, everyone has the same income. A uniformly rich country is just as equal as one in which everyone is poor.

Gathering Income Data

The Census Bureau surveys, analyzes and reports on income inequality in the United States every year.

- Income inequality can be measured by:
 - Comparing the share of wealth held by households with different incomes
 - Calculating single figure statistical measures like the Gini coefficient

What Census Has Built

Below are a number of projects built by the Census using income inequality data.

Income and Poverty in the United States, 2014

Download and read the 2014 report on income inequality and poverty in the U.S.

Learn more | Get the Data

Income Inequality Report

Explore a range of American Community Survey demographic, economic and geographic data on household income for states, counties and census tracts.

Learn more | Get the Data

Census Explorer

Interactive visualization median household income by county

Learn more | Get the Data

Islands of Income

Interactive visualization median household income by county

Learn more | Get the Data

What the Community Has Built

Below are a sample of projects the community has built using Census income inequality data.

New York Times

Income inequality article and interactive visualization

Learn more

The Economic Policy Institute

Interactive tools and videos that tell a story about inequality.

Learn more

Bloomberg

An income inequality and poverty dashboard for ranking Congressional districts.

Learn more

insights to page features cont.

"Why 5 mapping tools for the same data?"

Manipulating data takes work, users don't expect shortcuts

"They keep throwing datasets at you and don't explain what they do"
"All data is dirty"



Get the Data
Download the latest data

American Fact Finder
Find popular facts (population, income, etc.) and frequently requested data about your community.
[Explore](#)

CitySOK
You can get started using income data quickly with this JavaScript toolkit that makes the data accessible, easy, and collaborative friendly.
[Learn more](#)

The Census Data
Download all the data from the 2014 report on income inequality and poverty in an open, machine readable format.
[Download](#)

Where Census Data Comes From

ACS
American Community Survey
The ACS is an ongoing survey producing detailed information on a variety of topics about our nation and its people. Data helps determine how resources are allocated and how state funds are distributed each year.
[Learn More](#)

CPS
Current Population Survey
The CPS is a statistical survey conducted by the United States Census Bureau for the Bureau of Labor Statistics (BLS). The BLS uses the CPS to measure employment and unemployment in the United States.
[Learn More](#)

SUP
Survey of Income and Program Participation
SIPP collects basic and program participation data related to various types of income, labor force participation, social programs participation, and other personal and demographic characteristics to measure the effectiveness of existing federal, state, and local programs.
[Learn More](#)

How Income Inequality is Calculated
Income inequality is calculated in several ways:

The Gini Coefficient
This is the most widely used measure when discussing income inequality. The Gini Coefficient ranges from 0 to 1. If everyone in a group has the same income, the Gini Coefficient is 0. A value near 0 indicates perfect inequality - all income goes to a single person. The closer the number is to 1, the more unequal the group is. The Gini Coefficient does not take family size into account when calculating income.
[Learn More](#)

Alternate Methods
Because the Gini Coefficient has several limitations, other statistical methods have been devised in combination or as alternatives:
• The Theil Index is similar to the Gini index. It is a single statistic that summarizes the disparity between people across the entire income distribution.
• The Atkinson Measure is used to determine the effect of the income distribution on income inequality.
• The Lorenz Curve measures the gap between median and average income.
• The Gini Coefficient of income taxes only considers the number of people living in the household and how these people share resources and take advantage of economies of scale.

Income by Quintiles
The Gini coefficient reports income by dividing the US population into 5 groups, or quintiles, based on their income. The size of the population is equally divided among the 5 groups.
For example, if the top quintile would receive 20% (or 10%) of the nation's income, in a perfectly unequal society (with a Gini Coefficient of 1), the top quintile would earn 20% of the nation's income.
[Learn More](#)

insights to page features cont.

Users value personal support -

“I’m very good at telling people what I need”

Users want assurance they are using data properly
Users need data to do their jobs properly.

A screenshot of a web page titled "Get Support". The page is divided into sections for "General Support", "Income Surveys Branch", "Census Public Information Office", and "Center for New Media & Promotion".

- General Support:** Includes a brief description, contact information (301-763-4636, 800-923-8282, <http://ask.census.gov>), and social media links for Facebook, Twitter, and LinkedIn, along with a "Join Us in Slack" button.
- Income Surveys Branch:** Describes the branch's purpose, contact information (301-763-3243, SEHSD-ISB.LIST@CENSUS.GOV), and a link to the GitHub repository (<http://lucianoiscool.github.io/income-inequality/>).
- Census Public Information Office:** Describes the office's purpose, contact information (301-763-3030, pib@census.gov), and a link to the GitHub repository (<http://lucianoiscool.github.io/cpi/>).
- Center for New Media & Promotion:** Describes the center's purpose, contact information (301-763-7339, cnmp.developers.list@census.gov), and a link to the GitHub repository (<http://lucianoiscool.github.io/cnmp/>).

At the bottom of the page, there is a footer with the text "Income Inequality in America" and a link to the GitHub repository (<http://lucianoiscool.github.io/income-inequality/>). There is also a "Visit on GitHub" button.

quick design iterations based on testing

- EE The three main labels should be action words



The image shows three shield-shaped buttons arranged horizontally. The first button is blue with a graduation cap icon and the word 'Learn'. The second button is red with a grid icon and the text 'Get the Data'. The third button is gray with a question mark icon and the text 'Contact Us'. Each button has a red rectangular border around its text area.

- EE Local support is more helpful, can you add a link to connect with the State Data Centers?



The image shows a support page with the title 'Get Support'. It features two main sections: 'General Support' and 'Get Local Support'. The 'General Support' section includes a paragraph of text, social media links (Facebook, Twitter, YouTube), and a 'Join Us in Slack' button. The 'Get Local Support' section includes a paragraph of text and a dropdown menu for selecting a state. A large red rectangle highlights the 'Get Local Support' section and the state dropdown.

RECOMMENDATIONS



recommendations

We generated five key recommendations for the US Census:

1. Initiate front-end redesign of Census.gov
2. Place users - not stakeholders - at the center of existing technology projects to modernize the Census
3. Increase focus on two-way conversation in data dissemination strategy
4. Simplify personal support pathways for different user groups
5. Continue to add talent - especially user experience and front-end designers - to support effort

1

recommendations - redesign of Census.gov

Short term

Initiate project to redesign homepage

Tools – explain why user would use one tool vs. another

Census surveys – on topic pages, explain which survey is best for which purpose

Medium term

Simplify homepage – fewer menu options, get to destination quicker

Scale prototype – simplify topic page layout in-line with prototype

For all products that are based on data, upload machine-readable version of the data

Longer term

Transition to comprehensive use of web design standard

2 recommendations - users at the center

Short term

Publish sanitized specifications for current modernization projects online

Publish completed *stakeholder* feedback online (e.g. with State Data Centers)

Medium term

Transition all modernization projects to testing and iteration cycle with *external users*

Publish external user research online as completed

Initiate regulate Comms. Directorate and Census Leadership review of user feedback

Longer term

Open source code for modernization projects (CEDSCI, CRM, CMS)

Regularly release list of bugs / features / issues for all modernization projects

3 recommendations - two-way conversation

Short term

Ask for suggestions on improvement of source data, additional reports

Make web analytics accessible and share across the whole US Census organization (incl. leadership)

Medium term

For every report launch, conduct datajam / hackathon / event with journalists

Partner with innovative data orgs. (e.g. Census Reporter, DataUSA) and feature prominently on website

Use API statistics to determine most requested data, publish results across the US Census organization

Longer term

Re-evaluate wholesale vs. retail focus – to extent possible within US Census mandate

4

recommendations - streamlined personal support

Short term

Map support journeys for different user groups, based on actual requests received

Create single web page outlining the different types of support – and why users would choose them

Feature user support page more clearly on the homepage

Medium term

Clarify support mandate of Census Regional Offices, Census Information Centers, State Data Centers, Data Dissemination Specialists

Increase coordination between central training team and Data Dissemination Specialists

Nominate primary customer support liaison staff member in each topic branch in Census

Longer term

Investigate providing guaranteed response times for different support requests – publish standards

Continue to develop CRM using agile, user-centered approach

5 recommendations - UX and design talent

Short term

Create internal email group for Census staff interested in user experience and design

Advertise for new positions – more professionals with user experience background in CNMP team

Medium term

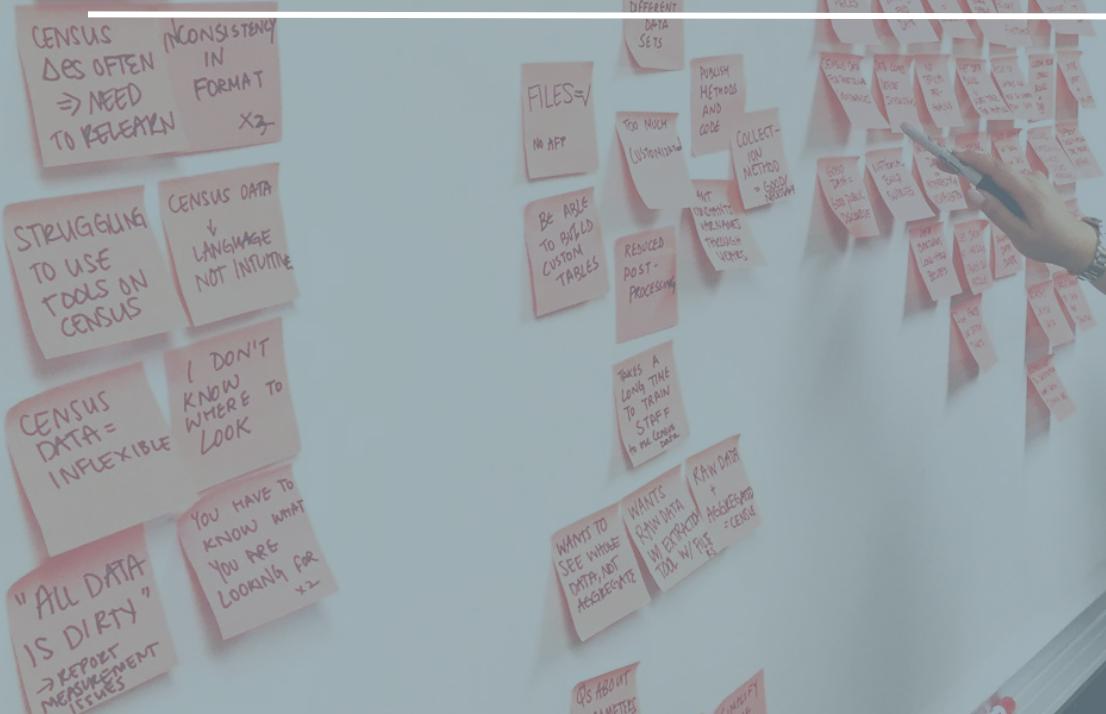
Hire extra staff to do user experience and design work on modernization projects

Establish links to user experience and design professionals in other Federal agencies, 18F, USDS

Longer term

Build user experience and design community of practice within Census – events, meetings etc.

APPENDIX



1 key insight – data

Users need data to do their jobs properly.

→ Data allow users to perform the core responsibilities of their jobs, including

- Target spending better
- Write more engaging research papers
- Overturn popular assumptions
- Improve the quality of public debate
- Prove there is a need for their service or policy

→ Users use data for a professional purpose. We did not interview users who explore data as a hobby.

“ [Data] makes the story more concrete ”

“ There is evidence in the data that your hypothesis is right or wrong ”

2 key insight – data

For users, data is personal and local.

- People want data that is important to *them*.
- Users look for data specific to their needs. Some are addressing a particular issue for the entire nation. Others want to collect all available data in a particular geographic area.
- Many users think in regional terms. They must align Census tract data with local neighborhood information to inform their activities within the community.



But I just want the
data for Medford
[City]...



If you can hover over one
region and get all the data for
that area, that would be huge



3 key insight – data

Users go to extraordinary lengths to get the right data.

- Many users spend hours searching for and extracting data.
- Other users sign up for webinars and develop personal relationships with data dissemination specialists who are employed by data providers
- Some users pay 3rd party providers for curated and difficult-to-find data that is provided for free on the Census website.



To get one data point from census.gov the first time took me 12 hours. Now it only takes a couple



If we can't get the data, we work around it



4

key insight – data

Manipulating data takes work, users don't expect shortcuts

- Some users like to play with the raw data. They want to understand the links between data sets and relish the challenge of overcoming the difficulties of the search process.
- Other users see data manipulation as a burden. When they want curated data they ask a colleague with analytical skills, such as a data journalist or Census employee, to manipulate the data for them.
- Despite their different attitudes towards the process, all users expect that they will have to do some work to get the right data. No-one expects it will be easy or straight-forward.



I would prefer an easy source but it's not always possible



All data is dirty



5 key insight – Census

Census data can overwhelm and deter users.

- The scope of Census collection can be daunting. Users may be discouraged from exploring Census data and finding if they don't know where to start.
- Too many tools complicates the issue by providing an excess of choice. Users unsure which of many choices is the right one, end up not choosing anything.

I wouldn't use Census as
my first source because I
don't know what they have!

“ ‘‘ daunting ’ ’

Why 5 mapping tools for the
same data? I just want ONE
place I can find what I need

“

6 key insight – Census

In the minds of users, Census data exists in an ecosystem.

- People use multiple data sources for their work.
- Going to multiple sources is not a cost itself; however, if it's consistently difficult to find data it becomes a very time consuming process.
- People keep downloads of data on their local hard drives and build their own databases. Their relationship with data sources is dynamic and customized for their own use.



My first step is to ask another journalist [if looking for data]



I never go straight to the Census website...instead I search through old articles to get context



7

key insight – trust and support

Users want assurance they are using data properly.

- Demographic data is complex and intricate. Professionals want to understand the methods, context, and quality of Census surveys to ensure the data supports their conclusions.
- Some users don't trust data until they inspect and exercise it. For them, accessing the data in raw form is best.
- Many users want to ensure data is searched and gathered comprehensively; missed data or geographic areas could undermine their purpose



If there's too many areas
[where] I could get it
wrong, I'll walk away



They keep throwing
datasets at you and don't
explain what they do



8

key insight – trust and support

Users value personal support.

- Data users seek out responsive, supportive providers. Even the most experienced data analysts want to establish personal relationships with data sources.
- Many users, when confronted with a data problem, reach out to a person for help by default. This can turn a potentially frustrating event into a positive experience.



Alexandra is great!!!

”

I'm very good at
telling people
what I need

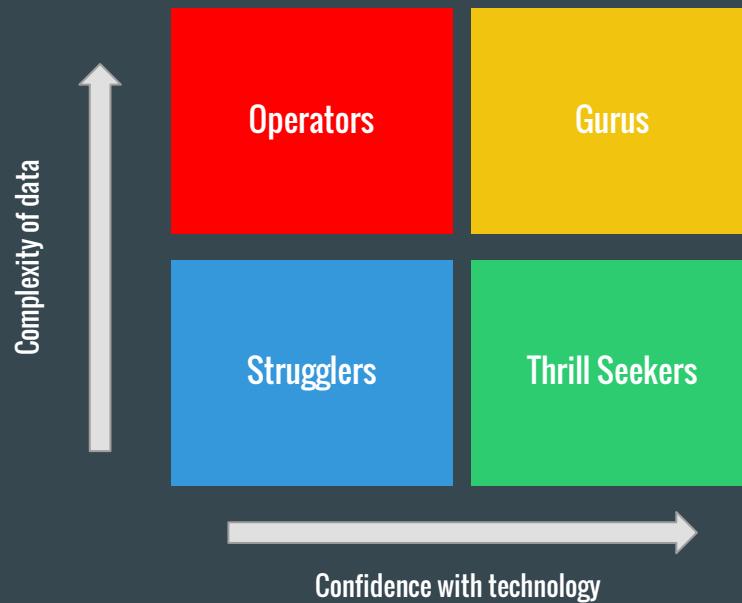
”

I'm dependent on the
people at the Census to
crunch the numbers for me

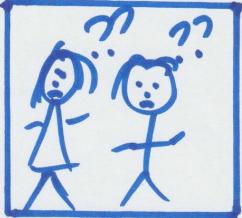
”

personas

To create personas we segmented users according to their confidence with technology and the complexity of the data they use.



persona #1: strugglers



THE STRUGGLERS

- Data is important for my job but I struggle to get it when I need it
- The sites I visit never work in the way I would like, it's frustrating!

CHARACTERISTICS

- Interested in aggregate data, pulled relatively infrequently
- Intimidated by jargon and taxonomies
- Overwhelmed with tools and tables
- Relies heavily on support from the Census employees

ACTION PROMPT

"I want to be able to ask questions to someone"

QUOTE

"I would never have known how to do this."

LOW CONFIDENCE
LOW COMPLEXITY



persona #2: operators



OPERATORS

CHARACTERISTICS

- Others find and process the data, Operators understand it.
- Works with 3rd parties (will pay) and internal specialist team members
- Technical enough to know what is hard.
- Have a "higher" purpose for the data → story, research...

- Use old data but not the latest technology.
- Work in teams with people who know how to use data better than we do → bring other things to the table

ACTION PROMPT

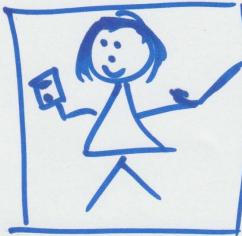
"I want my data - I don't care who gets it to me."

QUOTE

"I'm very good at telling people what I need"

LOW CONFIDENCE
HIGH COMPLEXITY

persona #3: thrillseekers



THRILL SEEKERS

CHARACTERISTICS

- Deals with large data sets
- Likes raw data, trusts it more than someone else's processed data.
- Loves finding unusual, quirky data
- Enjoys the challenge of using legacy interface websites

ACTION PROMPT

"I want direct access to all the raw data"

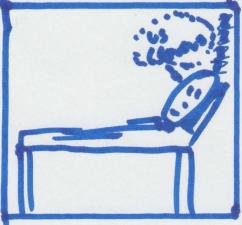
QUOTE

"Why 5 mapping tools?
I just want one place where I can find what I need."

HIGH CONFIDENCE
LOW COMPLEXITY



persona #4: gurus



GURUS

- I'm an experienced user, a specialist.
- I answer everyone's questions internally at my workplace
- I'm deep in specific areas, sources.

CHARACTERISTICS

- Uses advanced tools, coding skills
- Interested in metadata, the way classifications are made
- Does not use packaged tools, quick fact finders
- Maintains their own data bases, matches data sets

ACTION PROMPT

"I want to engage with the people collecting and curating the data."

QUOTE

"I want control... you lose freedom if you can't be completely in control of your [research] question"

HIGH CONFIDENCE
HIGH COMPLEXITY

