

CIVIC DIGITAL FELLOWSHIP

Streamlining the SGRD Frame Maintenance Process

Yong Li Dich | Harvard University, Computer Science and Folklore & Mythology

Supervised by Elizabeth Willhide and Michael Flaherty

User research with Millicent Grant and Vicki Kuppala

Economic Reimbursable Surveys Division

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau.

Shape
your future
START HERE >

Survey of State Government Research and Development

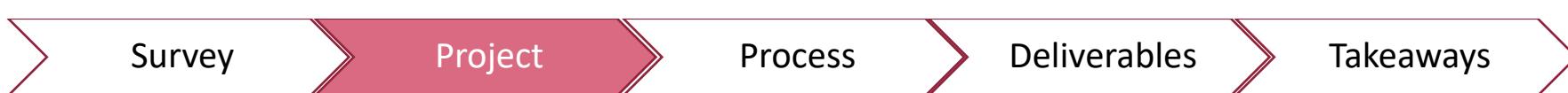
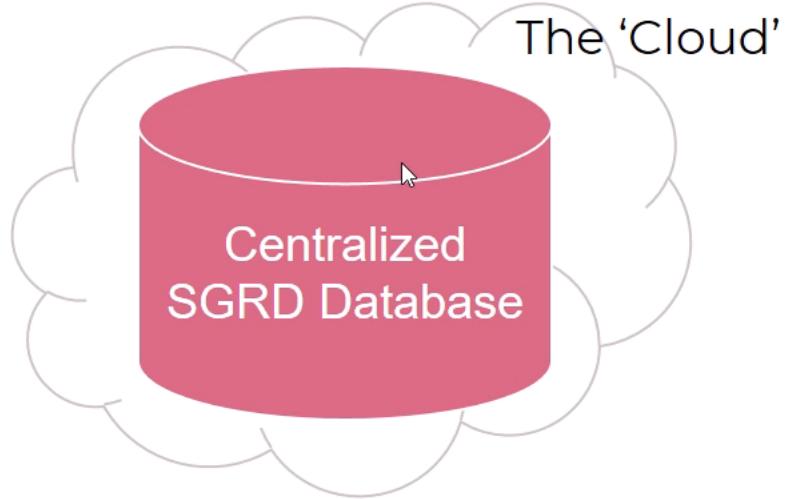


- Conducted jointly with National Science Foundation's National Center for Science and Engineering Statistics
- Measures R&D performed or funded by departments and agencies in the 50 states, DC, and Puerto Rico
- ~750 Agencies
- Example result: Health R&D = 44% of State R&D with \$1.1 billion in expenditures

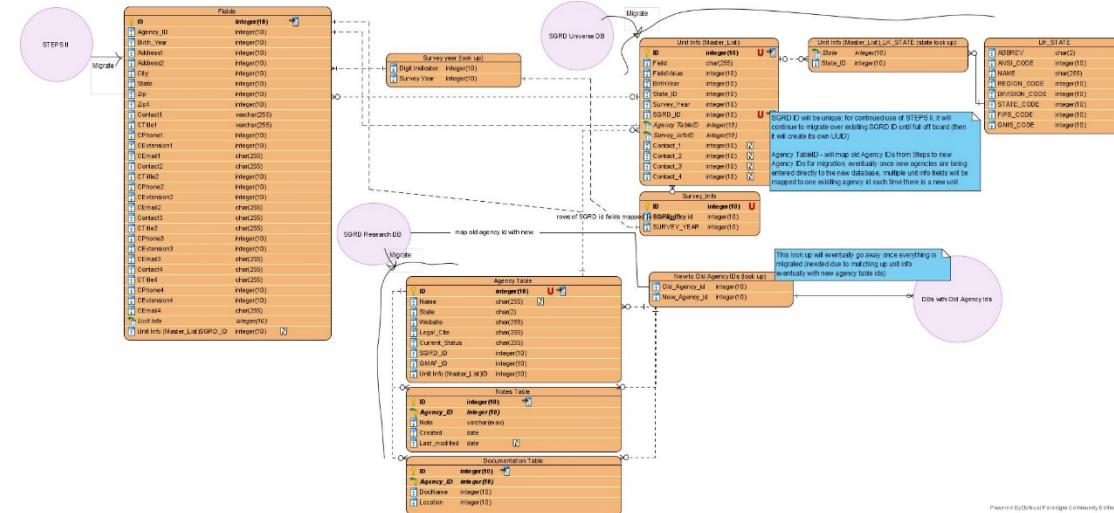
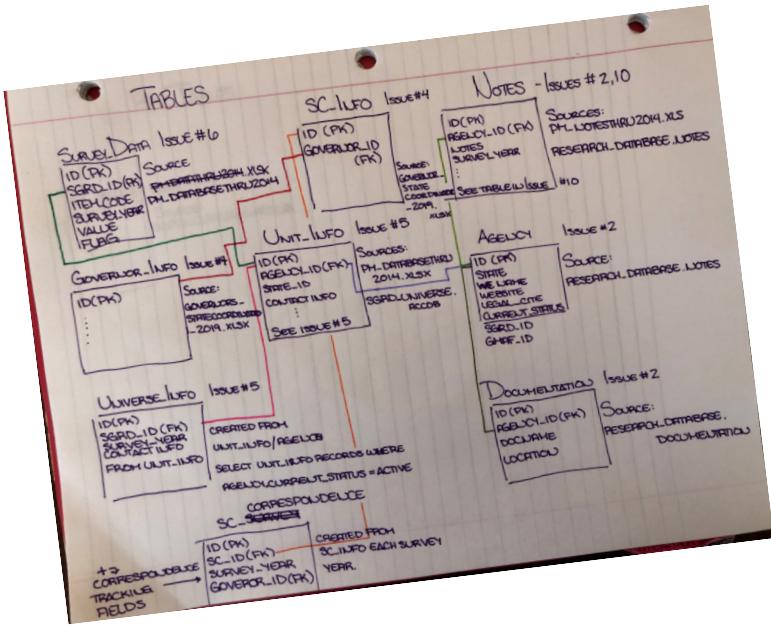
Current Pain Points

- > Key survey information resides on two systems
- > Communication between systems is clunky and manual
- > Updates between analysts, thus, *can take up to 2 weeks (i.e. reconciling information)*



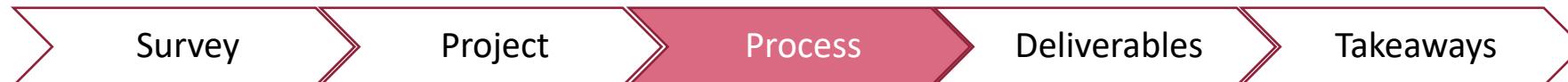


DATA MIGRATIONS INTO CENTRALIZED DATABASE

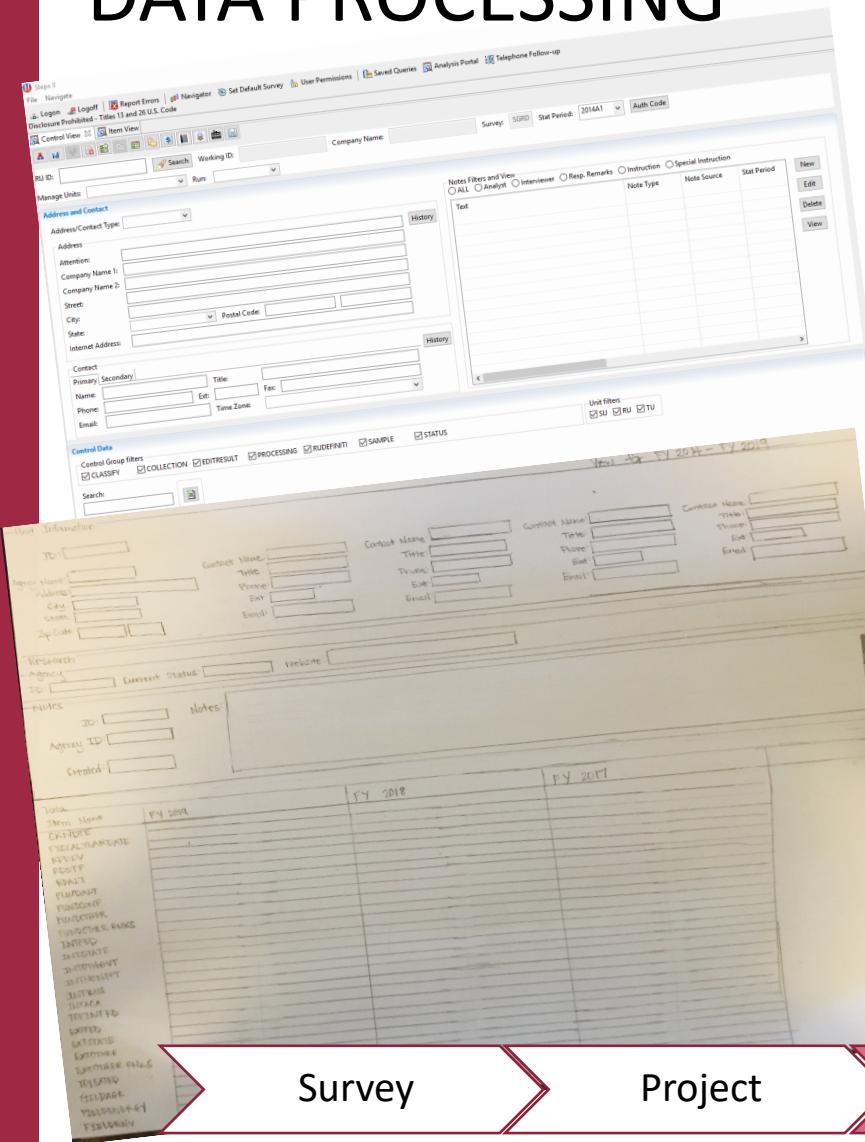


Challenges:

- Same types of data stored (and thus needed to be updated) in multiple locations
- One system carried data about agencies that didn't exist in the other
- Defining new model schemas that have values from different sources



DATA PROCESSING



Survey Data Processing

Check in date: Start week → End week

Birth year: Start year → End year

12347 Department of Transportation
Out of Scope, State: MN, Last Check in Date: 6-10-2019

12348 Department of Transportation
Pending NSF Review, State: CA, Last Check in Date: 5-10-2019

12349 Department of Transportation
Active, State: MA, Last Check in Date: 4-10-2019

12350 Department of Transportation
Dead, State: MA, Last Check in Date: 4-10-2014

Survey

Project

Process

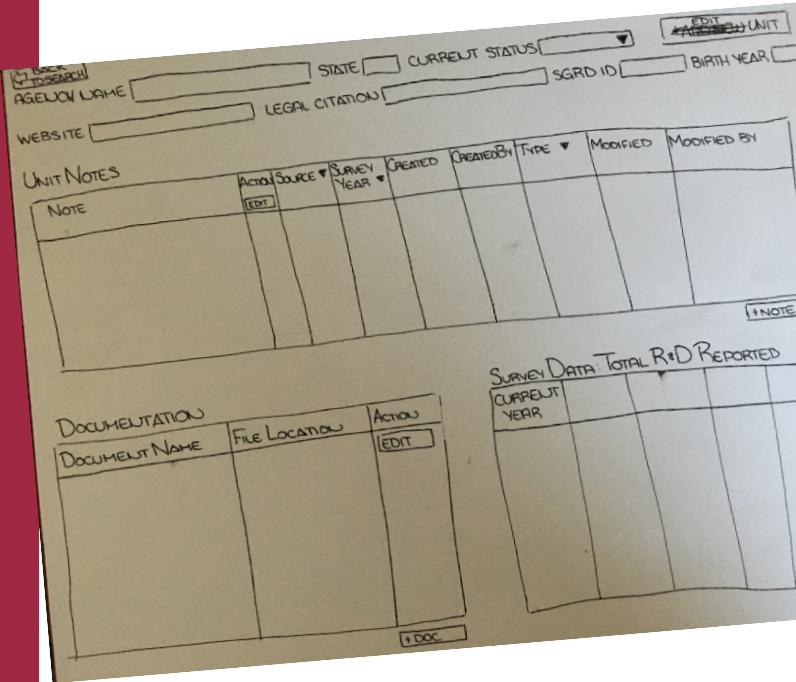
Deliverables

Takeaways

Shape
your future
START HERE >

United States®
Census
2020

RESEARCH PROCESSING



Screenshot of a Survey Data Processing application. The interface includes a search bar, filters for 'Check in date' (Start week, End week) and 'Birth year' (Start year, End year), and a main list area. The list displays five records for the 'Department of Transportation':

- 12347 Department of Transportation: Out of Scope, State: MN, Last Check in Date: 6-10-2019
- 12348 Department of Transportation: Pending NSF Review, State: CA, Last Check in Date: 5-10-2019
- 12349 Department of Transportation: Active, State: MA, Last Check in Date: 4-10-2019
- 12349 Department of Transportation: Dead, State: MA, Last Check in Date: 4-10-2014
- 12350 Department of Transportation: Inactive, State: OH, Last Check in Date: 4-10-2015

SEARCH VIEW

The screenshot shows a search interface for "Survey Data Processing". The search bar contains "input here". The results list five entries for "Department of Transportation" with the following details:

- 12346 Department of Transportation: Active, State: OK, Last Check in Date: 7-14-2019
- 12346 Department of Transportation: Dead, State: OK, Last Check in Date: 7-14-2014
- 12347 Department of Transportation: Out of Scope, State: MN, Last Check in Date: 6-10-2019
- 12348 Department of Transportation: Pending NSF Review, State: CA, Last Check in Date: 5-10-2019
- 12349 Department of Transportation: Active, State: MA, Last Check in Date: 4-10-2019

The left sidebar includes filters for "SGRD ID", "Agency Name", "Status" (Active, Inactive, Dead, Out of scope, Pending NSF Review), and date ranges for "Check in date" and "Birth year".

Survey

Project

Process

Deliverables

Takeaways

DELIVERABLES

Deliverable	Technology Used
Documentation	GitLab (wiki, READMEs, etc) [gitlab.com]
Data migration	Django (ORM) [djangoproject.com]
Centralized database (db)	Prototype db: SQLite [sqlite.org] Research/Production db: PostgreSQL [postgresql.org]
Functional frontend interface connected with database	React [reactjs.org] Django REST framework [django-rest-framework.org]
Research deployment	Backend (i.e. database): AWS GovCloud (internal) Frontend: internal web app hosting



WHAT'S NEXT?

Next steps:

- Training team members
- Iterating on frontend interface and API
- Prepare project for research level production

What this pilot project could benefit:

- *Short term:* SGRD survey data integrity, new interface platform provides shared language between analysts
- *Long term:* Scaling future surveys, more robust systems, encourage collaborative workflows, focus energy and efforts on survey quality

