

# CDE-NIBRS DISCOVERY

## KEY INSIGHTS & RECOMMENDATIONS

January 30, 2018

## AGENDA

# **EVOLVING THE CDE**

## KEY INSIGHTS & RECOMMENDATIONS

OVERVIEW & KEY INSIGHTS

1:00-1:30pm

DEVELOPING A STRATEGY

1:30-2:00pm

BREAK/DISCUSSION

2:00-2:15pm

HIGHLIGHTING NIBRS

2:15-3:00pm

CONCEPT DESIGNS

3:00-4:00pm

Q&A; WRAP-UP

4:00-4:30pm

## EXECUTIVE SUMMARY

This discovery effort was aimed at helping CJIS set a trajectory for the future of the CDE as it evolves from a minimum viable product (MVP) and into the “preferred platform for publishing UCR data.” Within this context 18F was not focused on specifying features to be delivered, or defining long-term development plans. Rather, our intent was to surface key opportunity areas and considerations to guide CJIS in its future decision-making around the product.

Another important aspect of our work involved preparing CJIS for the “cost of ownership of the CDE,” by identifying the internal processes and capabilities that are necessary for supporting the product in a sustainable way. Part of this will require new ways of thinking and doing, both in terms of how CJIS delivers software and how it chooses to publish UCR data in a digital way.

### GO SLOW TO GO FAST

We recommend a phased approach that starts with building up internal processes and capabilities at CJIS, such as defining automated data validation processes, to ensure a strong foundation for maturing the CDE. Next, we recommend a focus on enabling “responsible access to the data,” by adding capabilities that allow for more dynamic footnotes and notations and validating the core product offering once the [FBI.gov](#) URL is obtained. With these foundational pieces in place CJIS can then build on what it has learned from users to pursue “growth-oriented” enhancements, such as potential changes to the interface, that help further the mission of the CDE.

### THE CDE AS A PLATFORM FOR DEMONSTRATING THE VALUE OF NIBRS

Given the push to achieve an all NIBRS reporting standard by 2021, we recognize the pressure to expand the presence of NIBRS in the CDE. However, the NIBRS program still only covers 30-40% of the US population, which suggests that there are higher value opportunities for CJIS to pursue in the short-term than redesigning the web interface around NIBRS. Doing so would not only be costly, but result in an inconsistent user experiences as many states are likely to report only SRS data for the foreseeable future. Moreover, our research to date suggests that our target users of the CDE (media and other “advanced consumers of data”) prefer to access UCR data in its rawest form, as downloads and the API are easier to work with when performing deep analysis, than static charts or visualizations in the interface.

### INVEST IN SPECIALIZED DOWNLOADS

Even though the volume of NIBRS data is limited, the CDE can still help demonstrate its value while the program matures and is more widely adopted nationally. On this front we recommend investing in “specialized” downloads, or datasets built around specific topics, to both showcase the value of NIBRS and learn more about which datasets are of interest to users and how they wish to interact with them. In this sense, specialized downloads provide a low cost way of expanding the presence of NIBRS in targeted directions, while also setting the stage for more robust enhancements in the future. CJIS may also consider developing “data stories”, or brief narratives around specialized datasets, to further highlight the power of NIBRS and appeal to a wider range of users.

### JUST START

NIBRS is complex and so there are many potential views of the data that could be offered. Our research suggests that there is no clear consensus among users as to which NIBRS perspectives they would like to see next, but there are a few logical places to start. As such, we recommend starting with a short list of topical areas, such as women against violence and violent crime in large cities, and adding new downloads iteratively based on what is learned from user feedback and website analytics.

# **OVERVIEW**

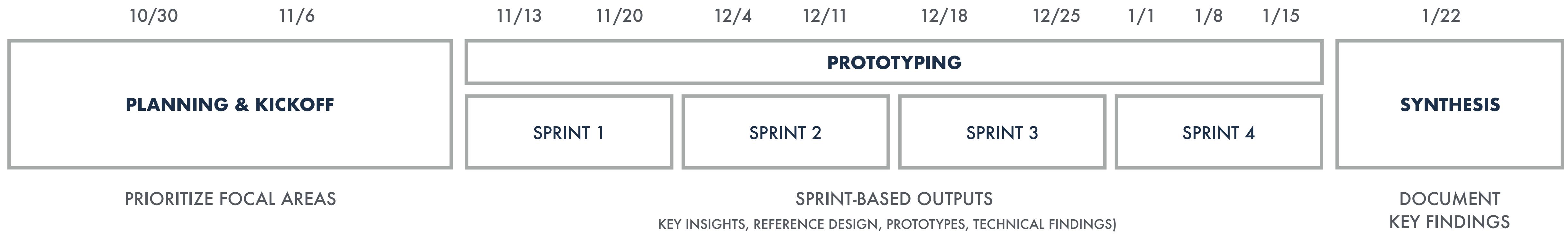
## THE GOAL FOR THIS EFFORT IS TO DEVELOP A **STRATEGY** FOR EVOLVING THE CDE.

18F & the FBI have been working together for the past year to build a website called the Crime Data Explorer (CDE). The CDE makes it easier for law enforcement and the general public to access data collected from nearly 18,000 law enforcement agencies as part of the Uniform Crime Reporting (UCR) program. The team delivered an initial release, or minimum viable product (MVP), of the CDE in the summer of 2017 and is now preparing the FBI's Criminal Justice Information Service (CJIS) to own and evolve the product in an iterative and human-centered way.

Since its initial release, the vision for the CDE has sharpened. It is now being positioned as the "primary platform for publishing UCR data" and the "digital front door for the UCR program", by offering broader access and more dynamic views of available crime data. This focus raises new questions, such as how can crime data be best presented via a digital medium, and how can broader access to UCR data lead to better, more timely data?

Given these goals CJIS engaged 18F in a 12-week discovery effort that was aimed at setting a trajectory for the CDE. Unlike previous development efforts, 18F was not asked to deliver production-ready code; rather it was tasked with answering key questions that would inform the future of the product, such as how to make NIBRS data easier to work with and how to expand its presence in the CDE.

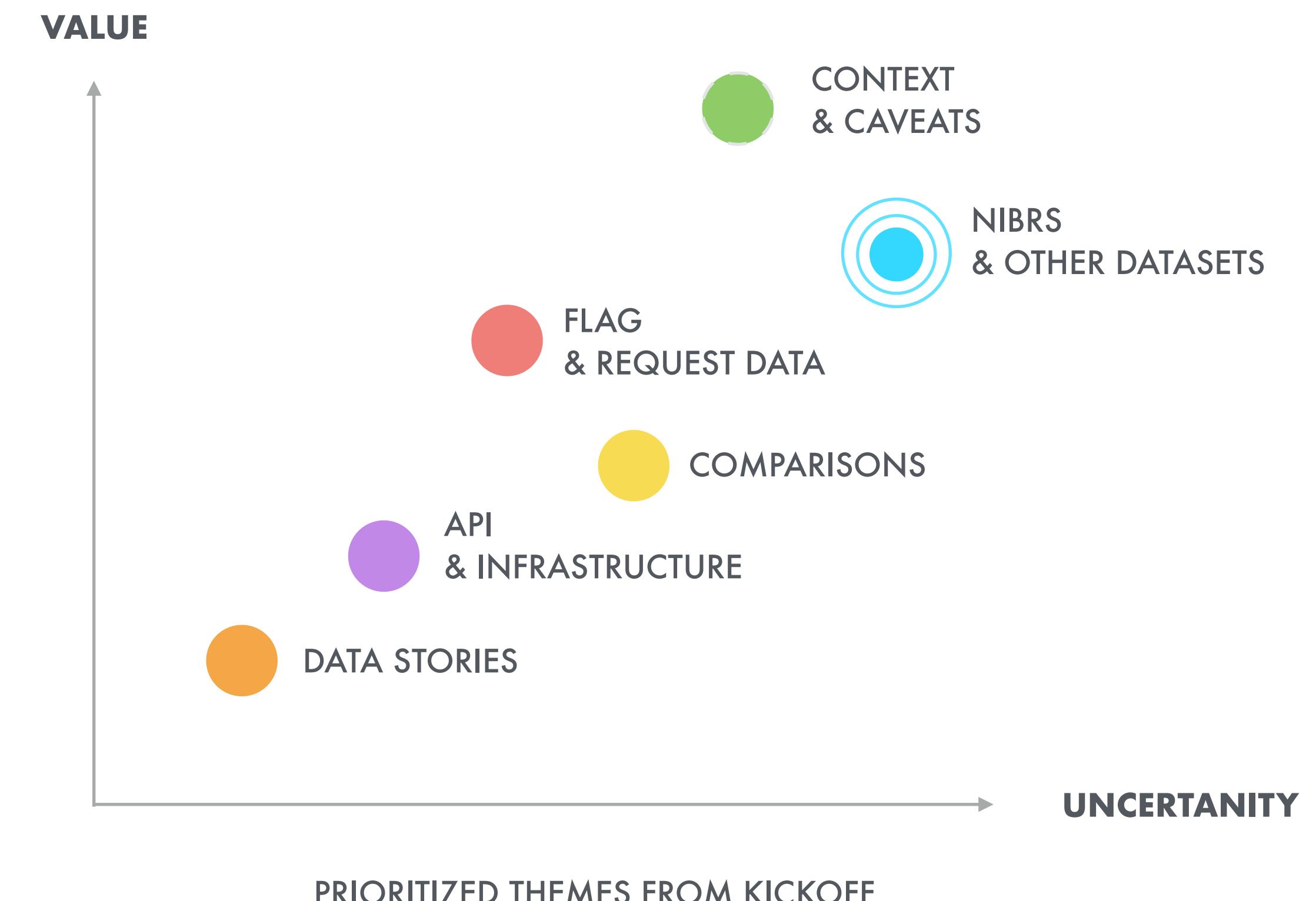
## TIMELINE



## PRIORITIES FROM KICKOFF

During the project kickoff we identified several themes to explore. We asked CJIS to prioritize these themes based on their perceived value to the CDE program and what we thought we could learn through discovery.

While all of these areas deserve future consideration for the CDE product roadmap, we set “expanding the presence of NIBRS” and improving “context & caveats” as our focus for the 12 weeks.



## GUIDING QUESTIONS

### WHY SHOULD NIBRS BE AN IMMEDIATE FOCUS FOR THE CDE?

How does “expanding the presence of NIBRS in the CDE” compare with other opportunities to broaden & improve access to UCR data?

What’s the value of NIBRS today considering that it only covers 30% of the U.S. population? How is it being used to facilitate analysis and drive decision-making?

What types of stories do people want to tell with NIBRS data and what perspectives can the data reliably support?

### HOW ACCURATE & EFFECTIVE ARE OUR CURRENT OFFERINGS?

What questions and concerns do users have about the data? How can we improve trust and confidence in what is published?

How do we provide users with broad and flexible use of the data while still maintaining ease of use?

How can can represent NIBRS data accurately and responsibly and empower others to do the same?

### HOW CAN THE CDE SERVE AS A CATALYST FOR CHANGE?

How can the CDE demonstrate the value of NIBRS, while also help promote improved awareness of the UCR program?

How might the unique capabilities of the CDE influence how CJIS thinks about how to publish UCR data going forward?

How can working in a more open and responsive way help CJIS better respond to reporting inaccuracies and criticism about the UCR program?

# **WHAT WE LEARNED**

WHO WE SPOKE WITH

12  
INTERVIEWS

6 DATA JOURNALISTS  
4 UCR STAFF/LAW ENFORCEMENT  
2 OPEN DATA ADVOCATES

## KEY TAKEAWAYS FROM USER RESEARCH

### BULK DOWNLOADS ARE VIABLE PATHWAYS FOR ADVANCED USERS

**MOST OF OUR TARGET USERS ARE COMFORTABLE WORKING WITH LARGE & COMPLEX DATASETS.** We recently changed the format of the bulk NIBRS downloads to provide a more “unfiltered” or “normalized” view of the data. This helps represent the data more accurately, but also results in larger, more complex files to work with. In our testing we wanted to know how usable these new formats were. We found that advanced users well equipped for working with normalized datasets, assuming that the data was well-documented. However, we also learned that the current single-year view of the data is an impediment to deep analysis and an opportunity area to consider in the future.

### CONTEXT IS KEY & CONTEXT IS DATA

**OPPORTUNITIES EXIST FOR IMPROVING THE CONTEXT BEHIND THE DATA.** In our testing we found that users didn’t always understand what the data represents, or how they should be working with it. This not only results in a poor user experience, but increases the likelihood that the data could be misrepresented and negatively affect the UCR brand. As such, more dynamic use of footnotes, methodologies, and other forms of context would help build confidence in the data, while also mitigating the potential for controversy. However, while we explored ways of presenting additional context in the user interface, CJIS does not yet have a programmatic way of supplying this data to the CDE; the current process for compiling this information is largely manual and the data is spread out over many different spreadsheets with little to no version control. For this reason and others we suggest treating context like data and creating internal content management tools and workflows so that it can be more easily added and maintained going forward.

### TRUTH FROM A CERTAIN POINT OF VIEW

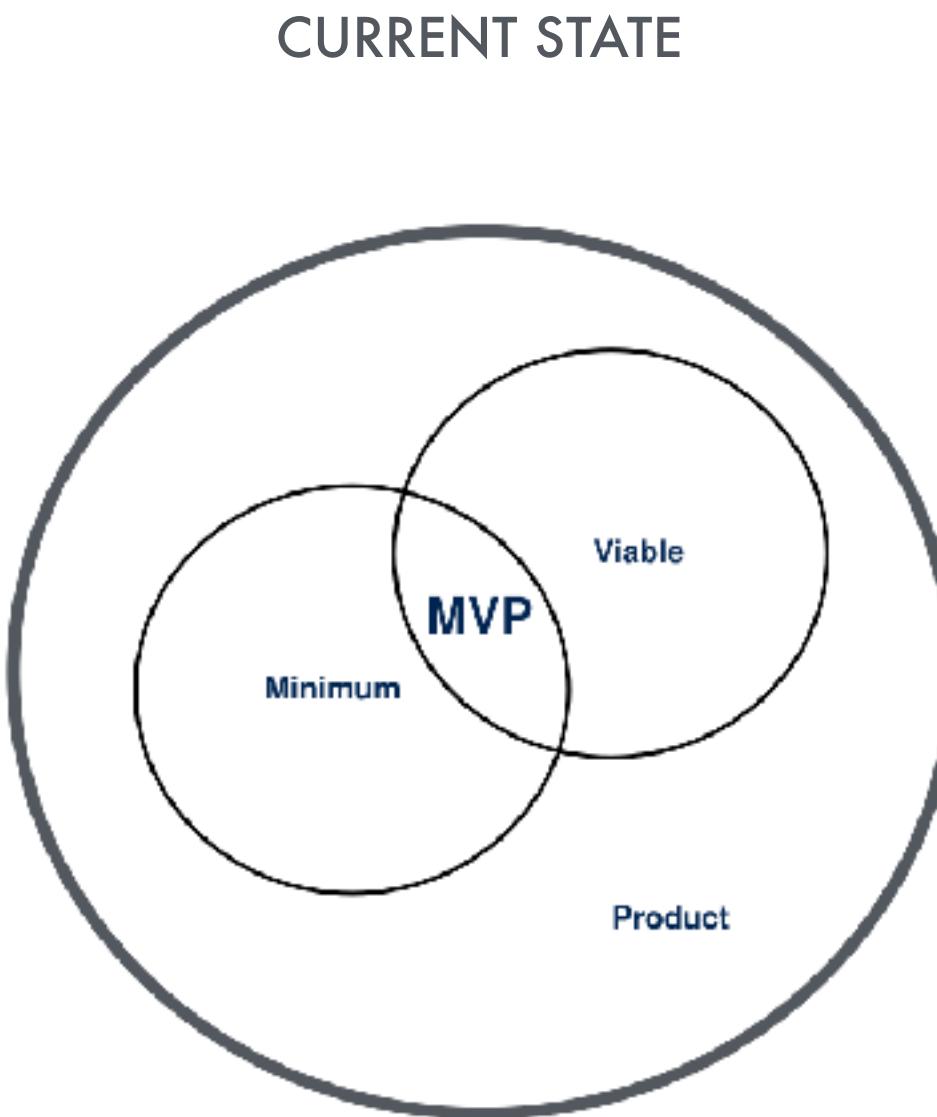
**THE DYNAMISM OF NIBRS CAN LEAD TO DIFFERENT PERSPECTIVES ON CRIME AND VARYING ACCOUNTING PRACTICES.** One of the advantages of NIBRS is that it allows people to explore granular dimensions of crime, such as how specific victim demographics are affected by gang violence. However, how a user arrives at an accounting for gang violence (e.g. total number of victims) depends on how they define that scenario and how they decide to work with its related data elements. While there are official rules and standards for working with NIBRS, the interpretation of crime data can still be a subjective process that leads to different representations of crime throughout the UCR community. As such, CJIS should focus less on the “accuracy” of the data and more on ensuring that the methodologies behind the data it reports are clear to the user community and consistently applied.

### NO CONSENSUS ON WHICH ASPECTS OF NIBRS TO HIGHLIGHT

**THERE IS NO CONSENSUS AMONG USERS AS TO WHICH NIBRS PERSPECTIVES TO ADD NEXT, BUT THERE ARE A FEW LOGICAL PLACES TO START.** When we asked participants what aspects of NIBRS they would like to see highlighted in the CDE next, we received a range of suggestions, but no clear answers. Similarly, CJIS has a sense as to which issues, or types of crime, people are interested in based on the data requests they receive, but these requests aren’t necessarily representative of the general public. For these reasons we recommend starting with a few topical areas, such as violence against women, and creating feedback loops to measure demand and user engagement so that data can more readily inform product decisions.

# **DEVELOPING A STRATEGY**

## WHERE DO WE WANT TO GO?



## FUTURE STATE

**THE CDE IS THE “DIGITAL FRONT DOOR” FOR THE UCR PROGRAM.**  
**IT PROMOTES PROGRAM MODERNIZATION THROUGH IMPROVED ACCESS TO UCR DATA.**

**THIS IS WHAT WE ASPIRE THE CDE TO BE.** Use this vision statement to align stakeholders and to guide decision making. It is your North Star.

View national, state, and local trends for major SRS & NIBRS offenses

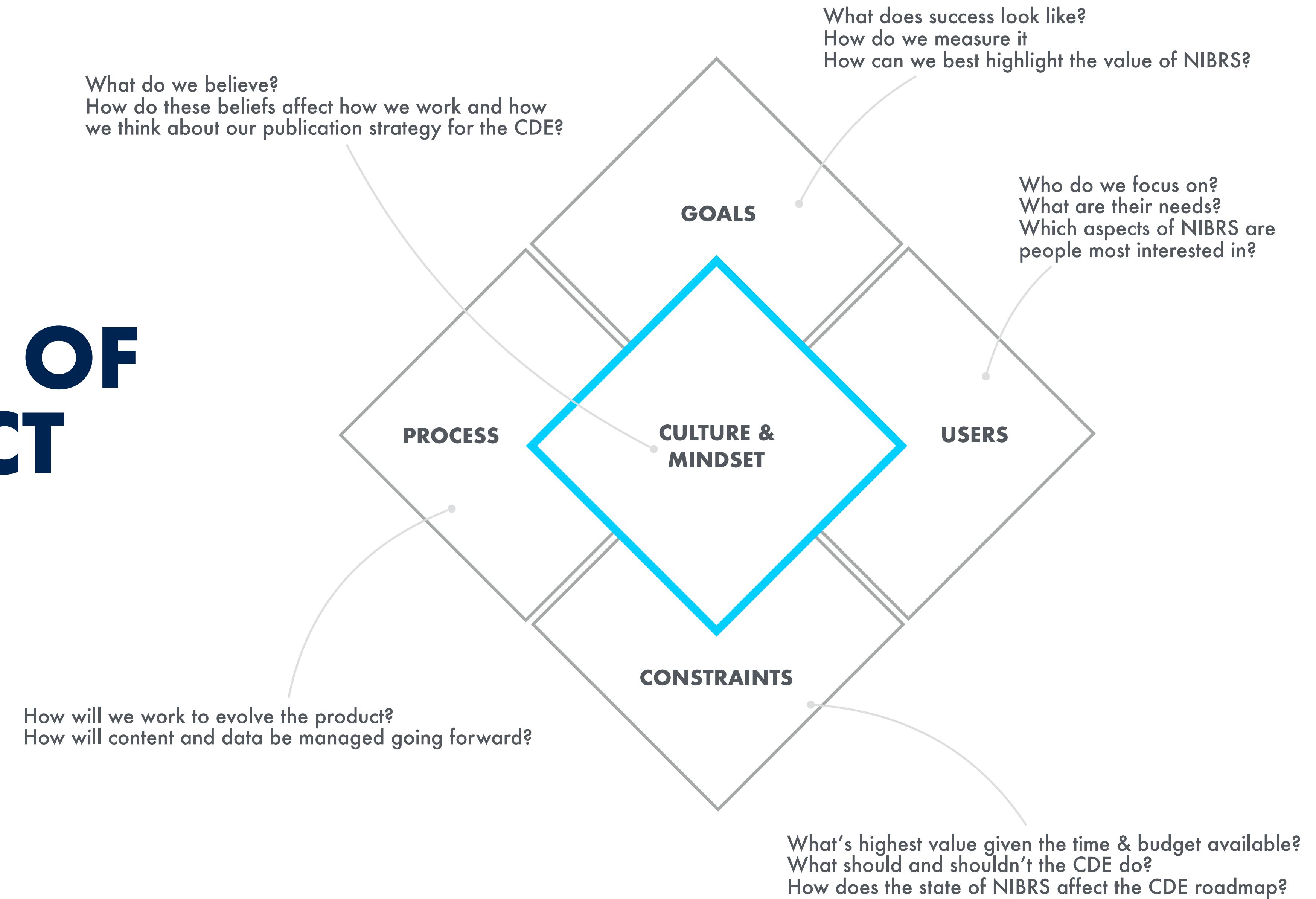
Download bulk incident data by state and year, as well as other selected datasets, like hate crime, as a CSV

View definitions and related links that provide context for the data

Access UCR data via an open API

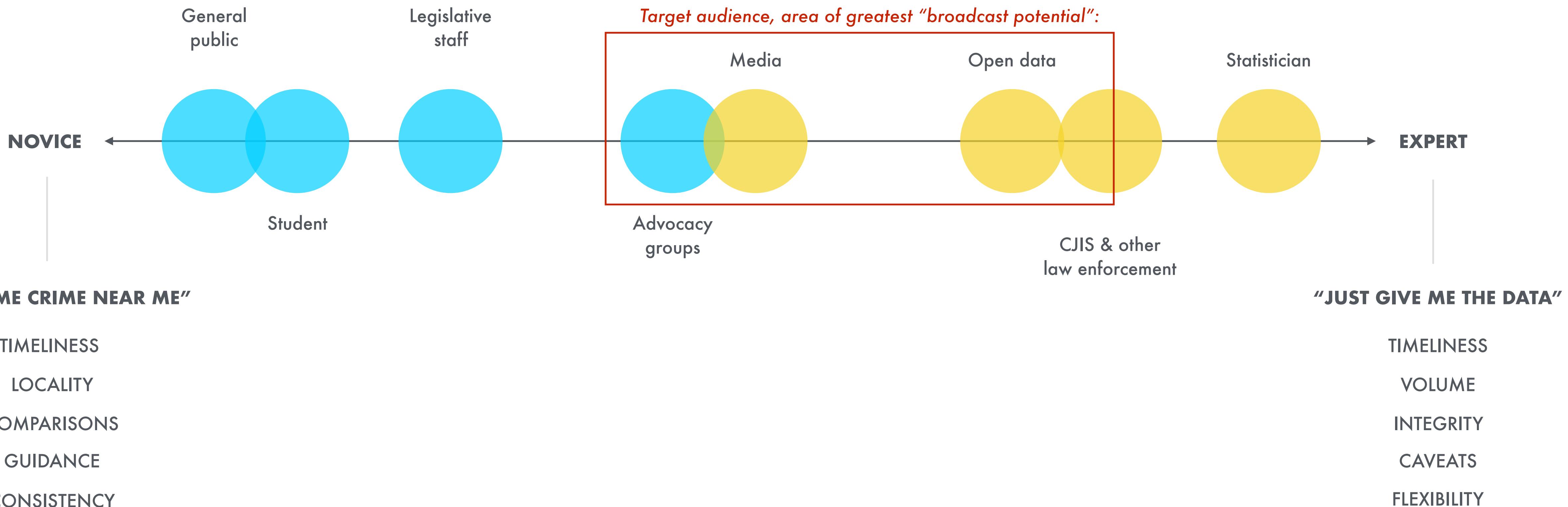
HOW DO WE GET THERE?

# THE 5 ELEMENTS OF A PRODUCT STRATEGY



## THE USER COMMUNITY

Research suggests that the typical CDE users is familiar with the UCR program and or working with open data. Similar to the “novice”, they value context, but prefer working with the data in it’s primary form and favor downloads & the API over the CDE interface.

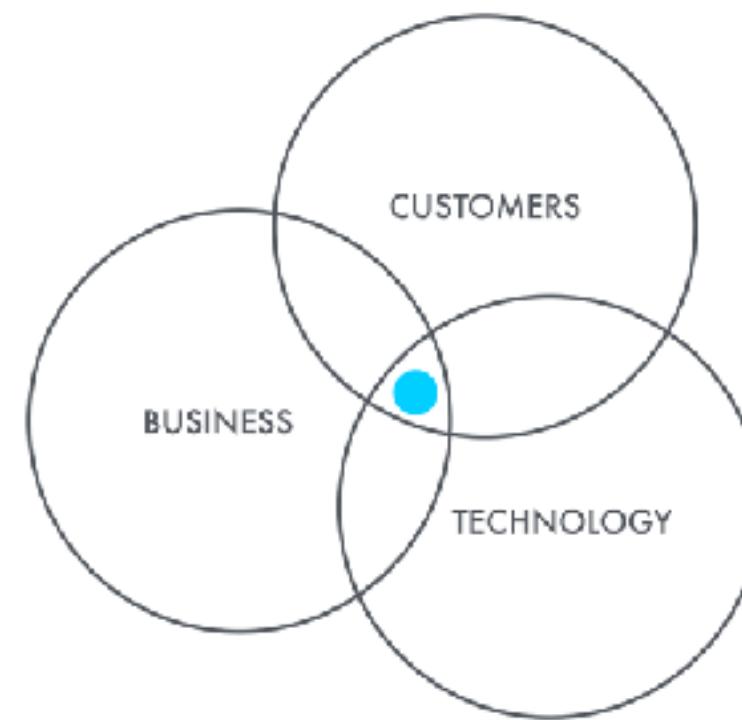


# THE CDE IS A PRODUCT, NOT A PROJECT

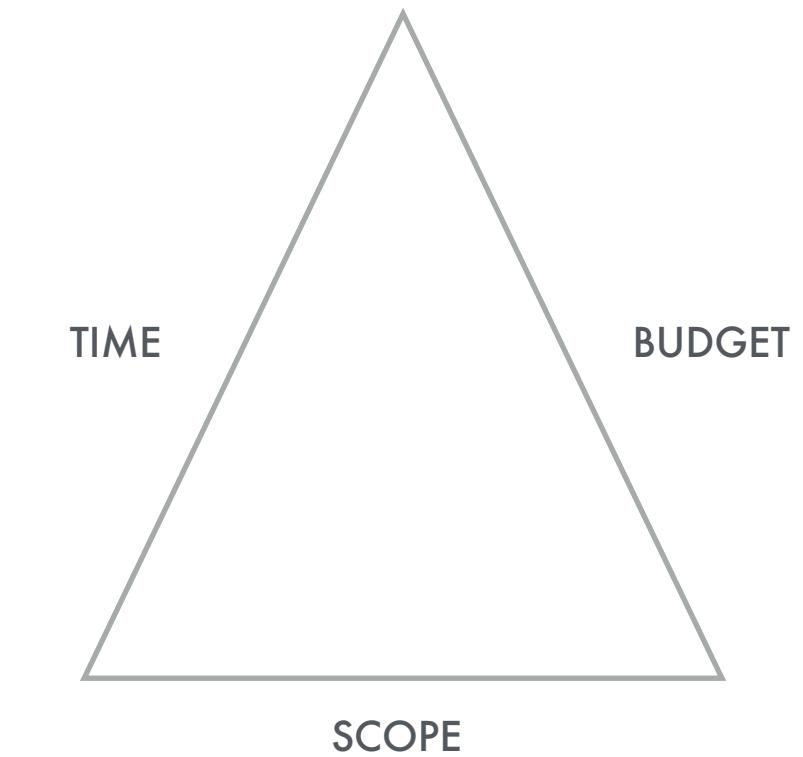
A product is designed to meet a specific user need as part of a larger business strategy.

A product mindset is focused on people, problems, and outcomes and requires a fundamentally different approach than traditional IT project management.

## PRODUCTS DELIVER VALUE



## PROJECTS ARE MANAGED



Learning & adapting

Design

Iterative development

Hypothesis-based

Purposeful

Value generation

Predictability & efficiency

Execute

Waterfall development

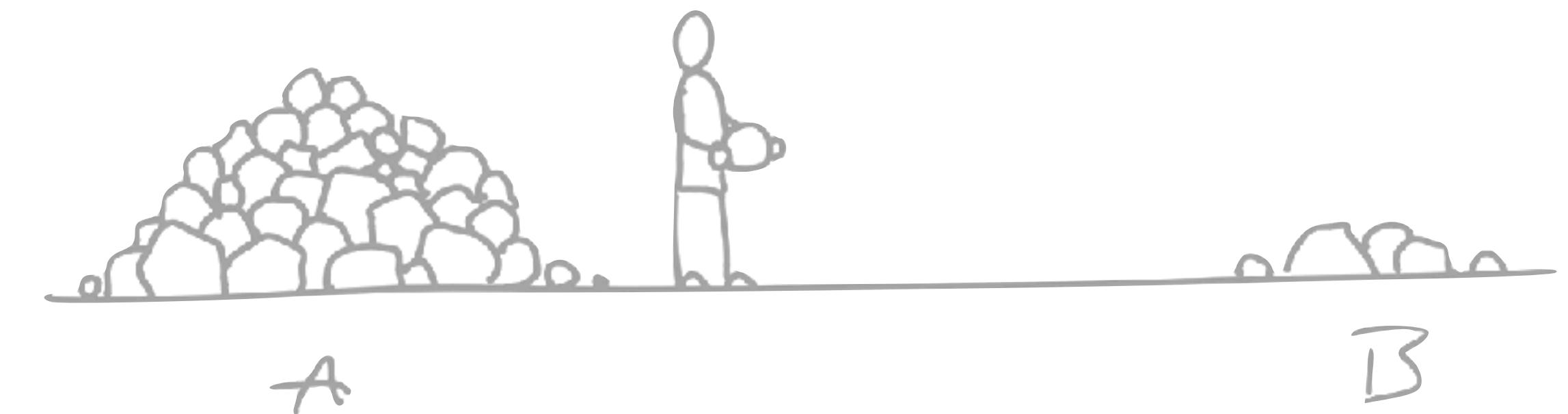
Requirements-based

Reactive

Conformance to a plan

# PRODUCT DEVELOPMENT ISN'T A LINEAR PROCESS - IT'S ABOUT LEARNING

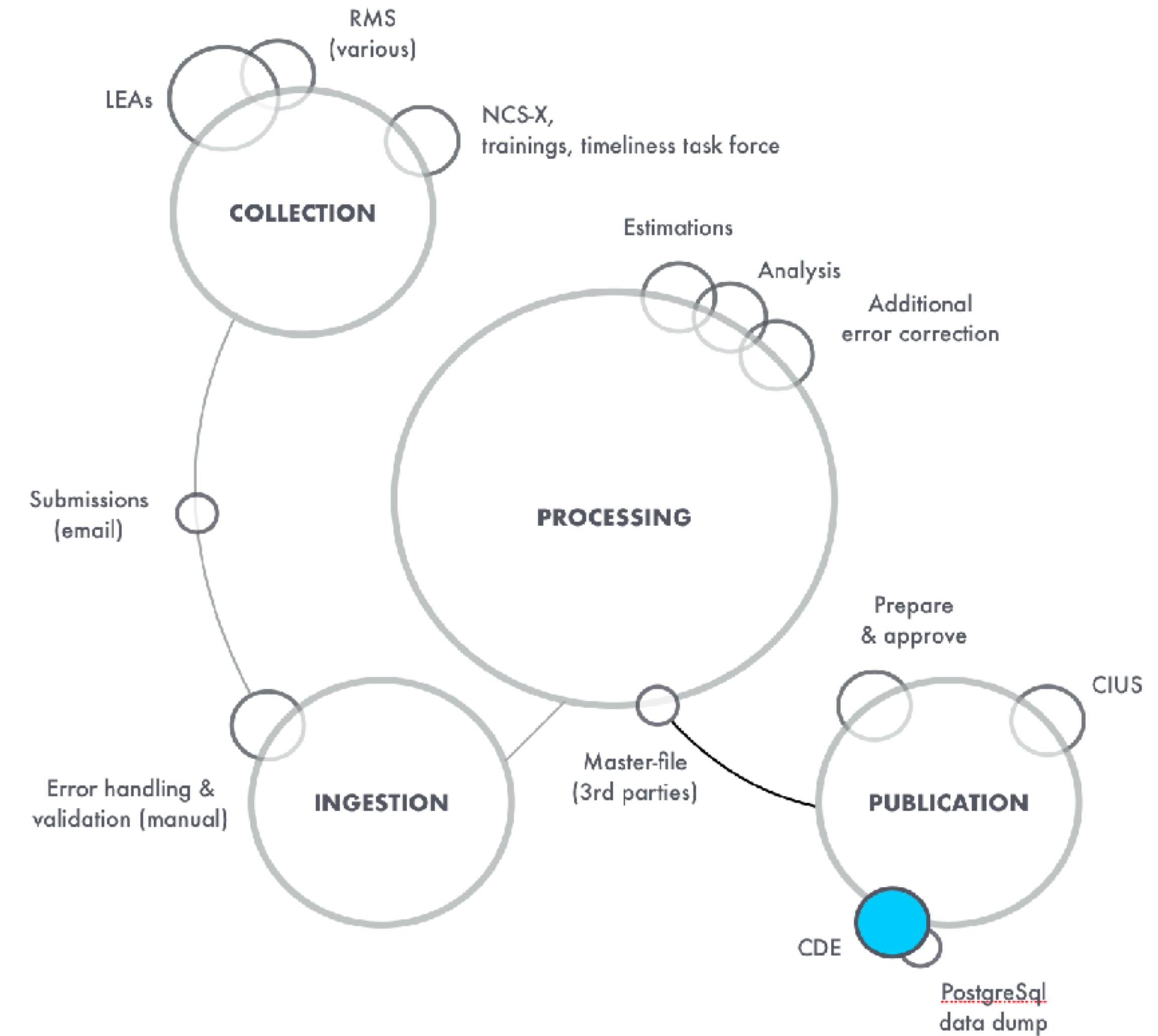
Traditional “command and control” approaches to project management, with its focus on predictability, can distract from what truly matters - people, problems, and outcomes - and increase the risk of building the wrong things.



Ryan Singer, *Why Agile Isn't Working and What We Can Do Differently*

# THE CDE IS ONE PART OF A DIGITAL ECOSYSTEM

Do all of these pieces work in isolation, or are they part of a cohesive whole? What goals do they help CJIS work towards and how can we measure them? How do we coordinate teams and efforts across these various work-streams?



## PHILOSOPHY FOR EVOLVING THE CDE

### PRODUCT PHASES

Enhancements that extend the product and further the mission. The growth layer is only worth investing in when there is a solid foundation in place and the core product has been validated.

Core functionality that enables “responsible access to the data”, meet user needs, and validates product direction. Requires close contact with users and an iterative, experiment-driven approach.

Processes & people are in place to support the success & sustainability of the product.

### GROWTH

### CORE VALUE PROPOSITION

### FOUNDATION

### OPPORTUNITY AREAS

Expand the API

Add detail to the interface as NIBRS participation improves

Consider new visualizations based on frequently requested data and what is learned from downloads

Start treating context like data

Meet users where they are by investing in downloads

Use specialized views to highlight the value of NIBRS & as a platform for informing future perspectives

Define qualitative & quantitative success metrics

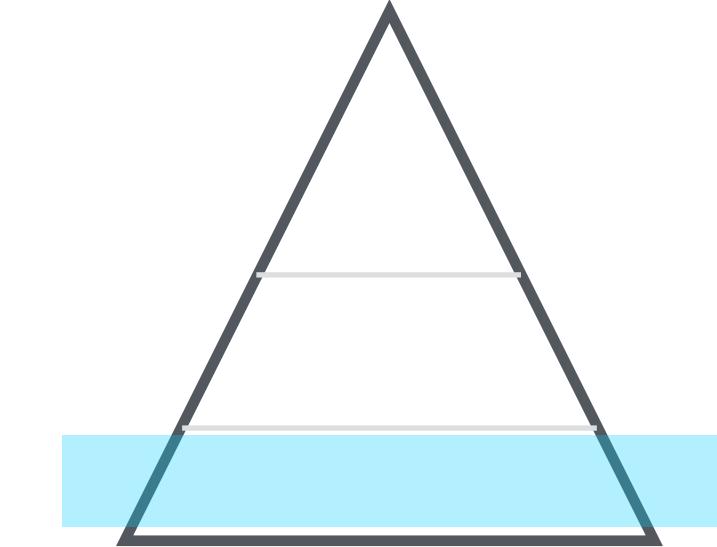
Develop processes for validating & maintaining the data

Adopt iterative development practices

*John Vars, "Product Hierarchy of Needs"*

## SETTING THE FOUNDATION

- Define qualitative & quantitative success metrics
- Develop processes for validating & maintaining the data
- Adopt iterative development practices



# DEVELOP PROCESSES & CAPABILITIES FOR SUSTAINING THE PRODUCT

Establish a team-wide, shared understanding of what success looks like based on a basket of qualitative and quantitative success metrics and regular measurement intervals.

Automate as much data ingestion and production as possible, with an emphasis on the ability to audit the process. Ideally, the audits are automated in some fashion too with indicators that are visible team wide.

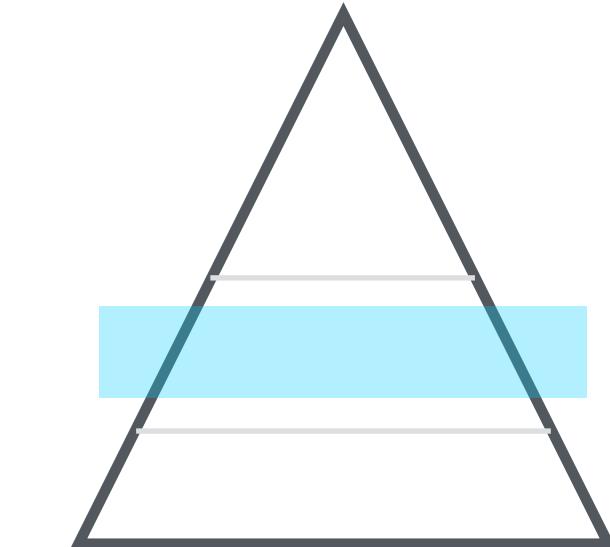
Validate ideas early and cheaply through prototypes and user testing with a trusted user community. Aim to ship something at the end of every sprint cycle, and use the end of a sprint cycle to reflect and improve on the internal CJIS processes, not just the product.

## CORE VALUE

Start treating context like data

Meet users where they are by investing in downloads

Use specialized views to highlight the value of NIBRS  
& as a platform for informing future perspectives



# ADD CORE FUNCTIONALITY THAT VALIDATES PRODUCT DIRECTION

The core value proposition of the CDE is to provide accurate and easy to use UCR datasets.

The largest risk to our current core value proposition is to provide confusing or, worse, inaccurate data. The lack of caveats, footnotes, and methodology are holes in the current value proposition that should be remedied.

Build a supporting system that helps CJIS manage contextual data, such as caveats, footnotes, and methodology with the same rigor as the rest of the data.

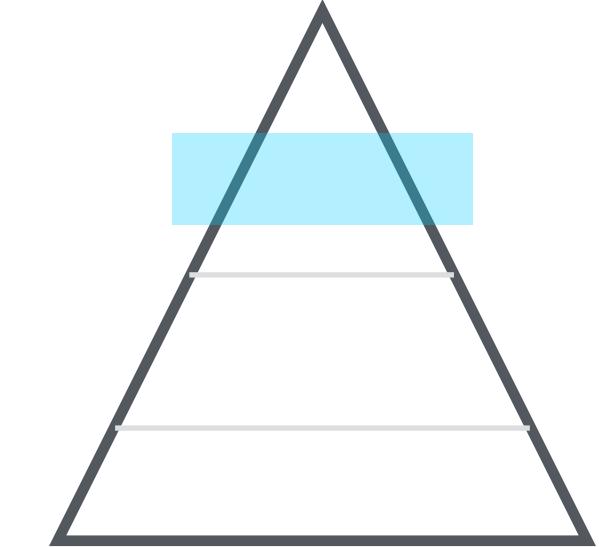
Validating core functionality requires close contact with users and an iterative, experiment-driven approach.

## GROWTH

Expand the API

Add detail to the interface as NIBRS participation improves

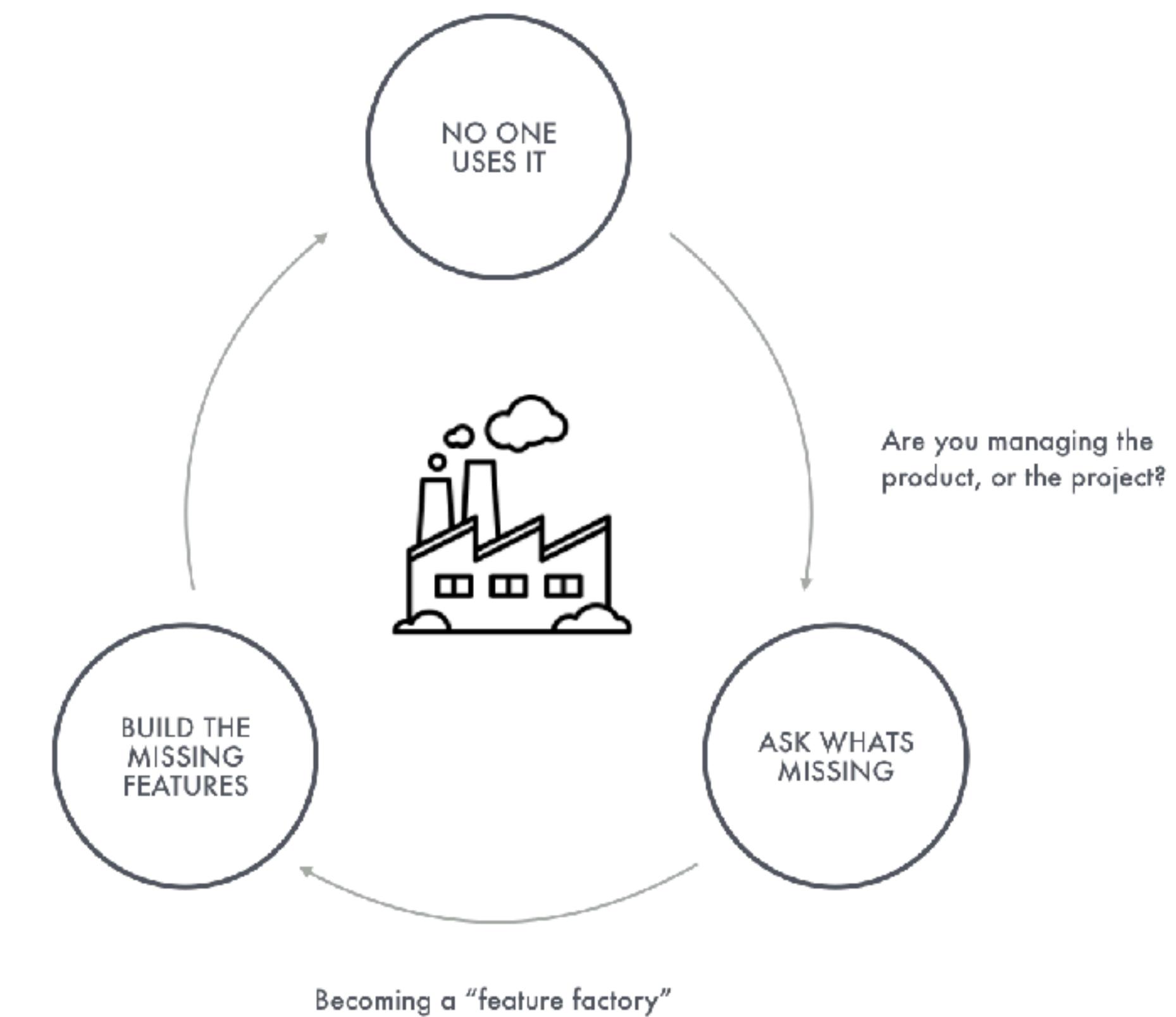
Consider new visualizations based on frequently requested data



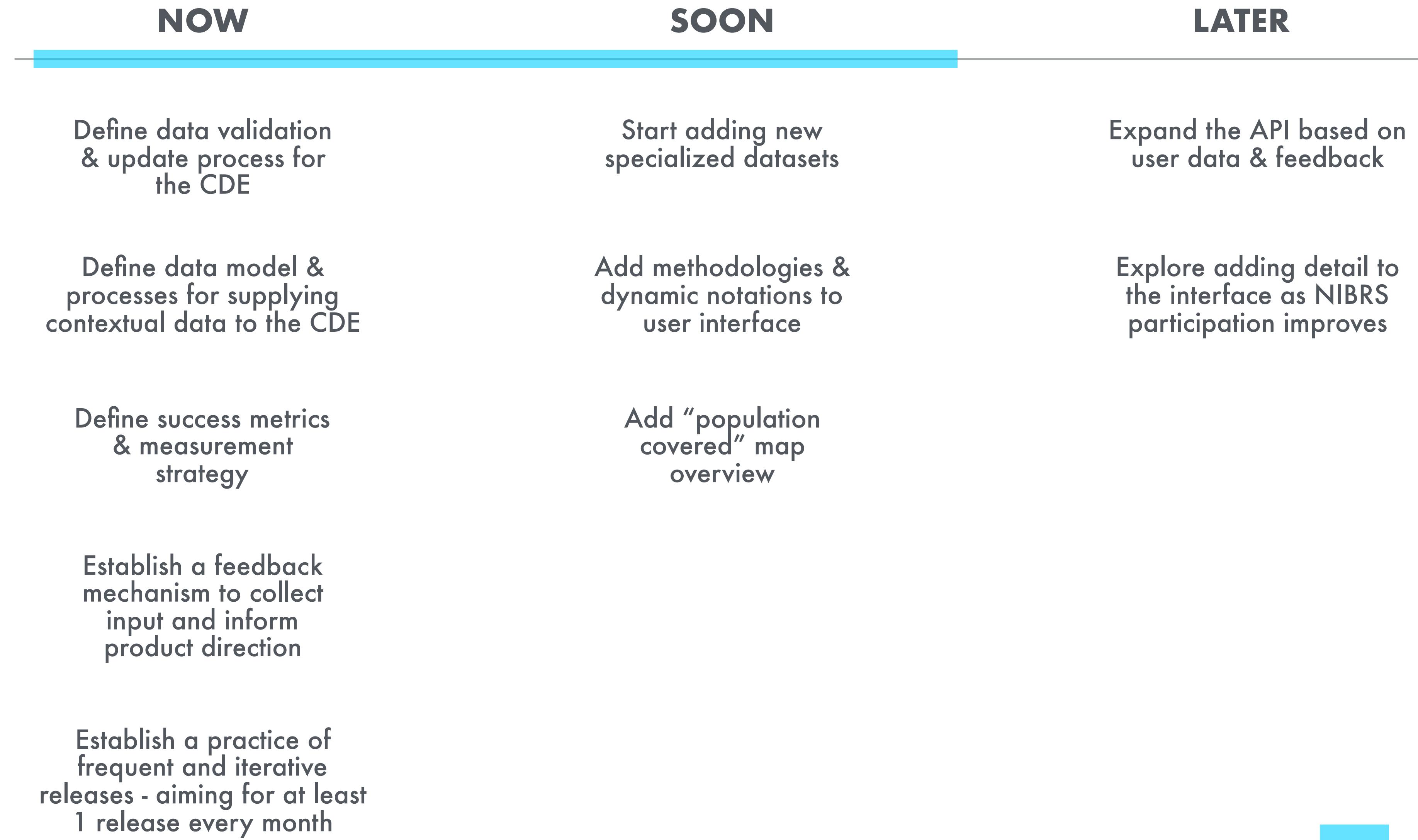
# PURSUE ENHANCEMENTS THAT FURTHER THE MISSION

Growth features are built on a strong foundation, and a clear core value proposition. Without these two prerequisites, the risk of becoming a "feature factory" is increased.

The cost of being a feature factory is not only using limited resources to build features that users don't want, but also the cost of maintaining those (unused) features over the lifetime of the product.

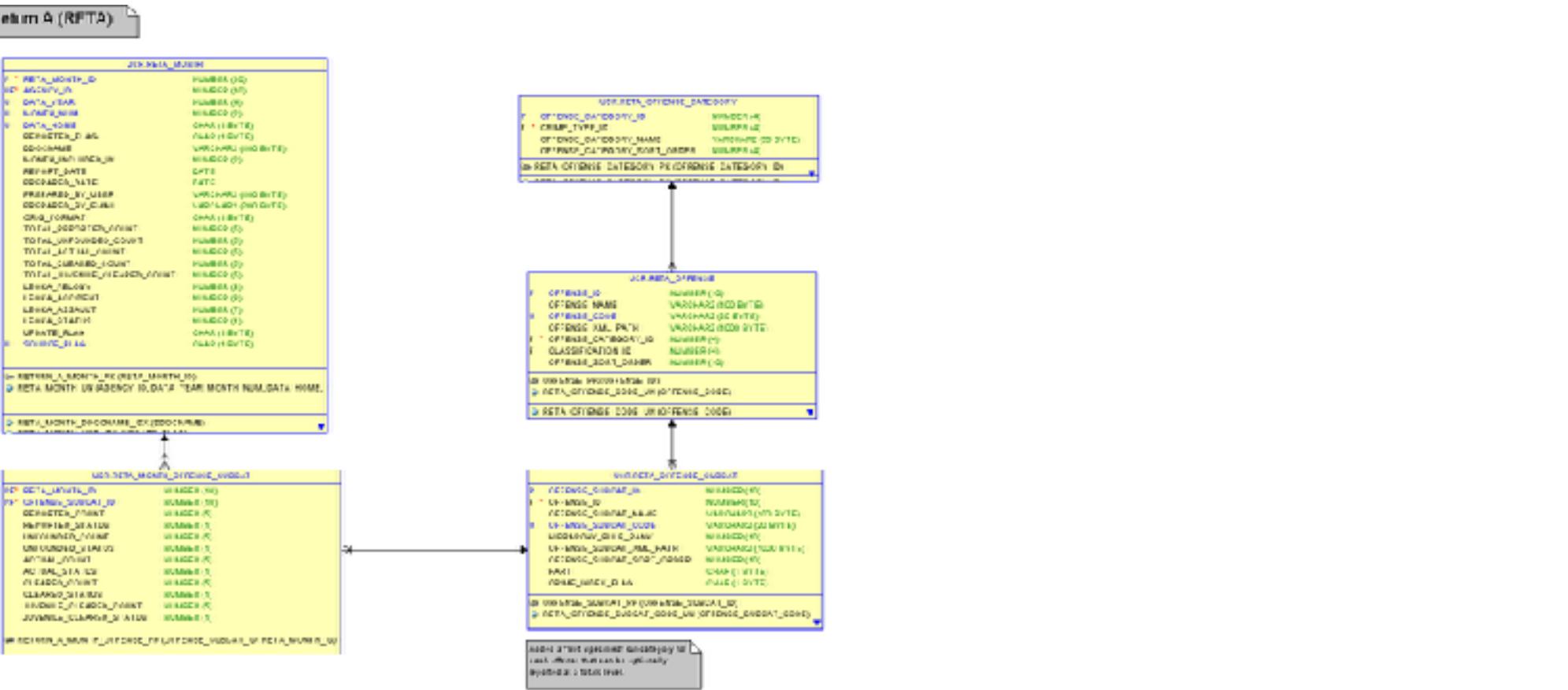


## HIGH-LEVEL ROADMAP



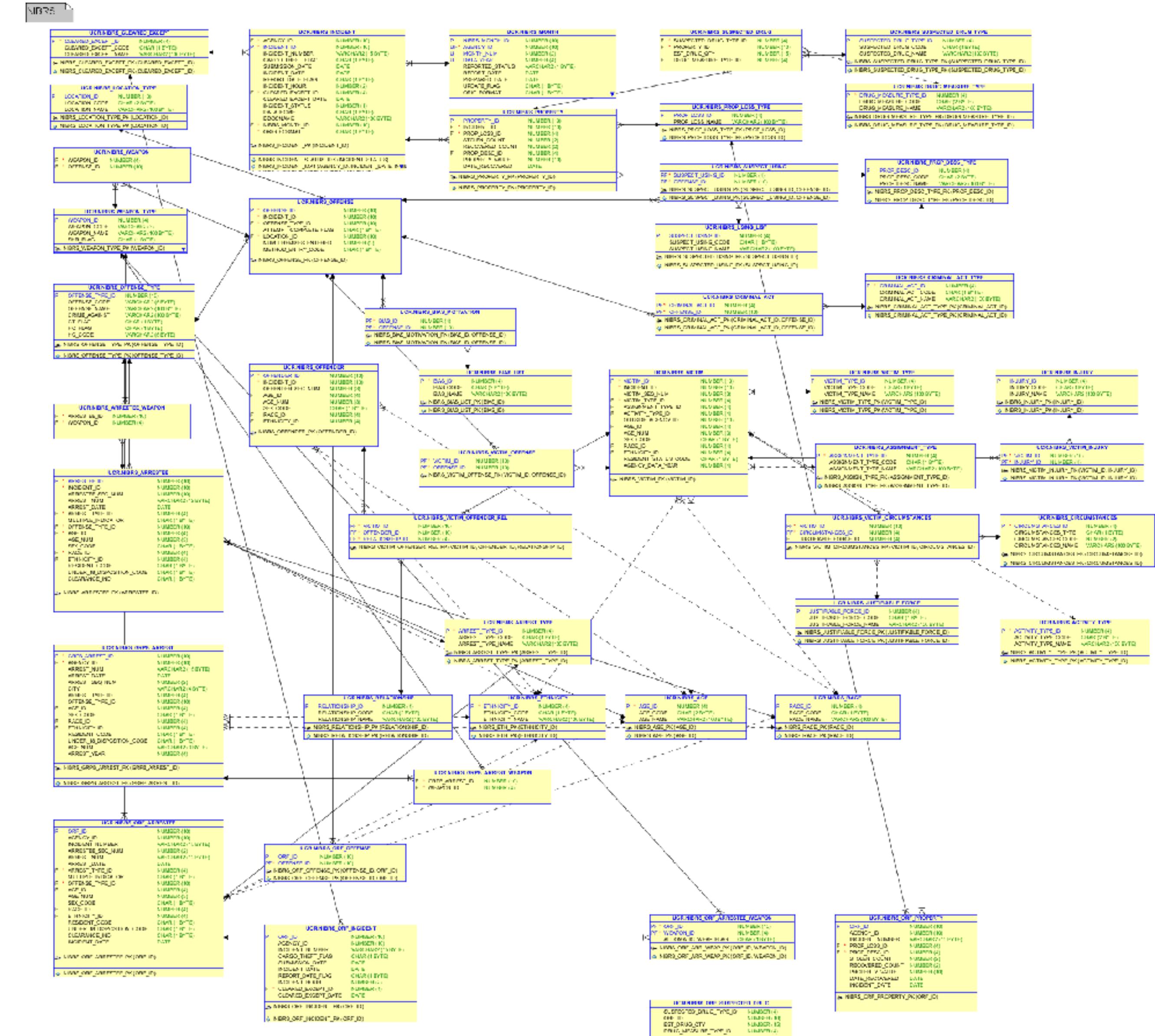
# **HIGHLIGHTING NIBRS**

## DEFINING THE NIBRS-CDE DATA MODEL



# WHERE WE STARTED

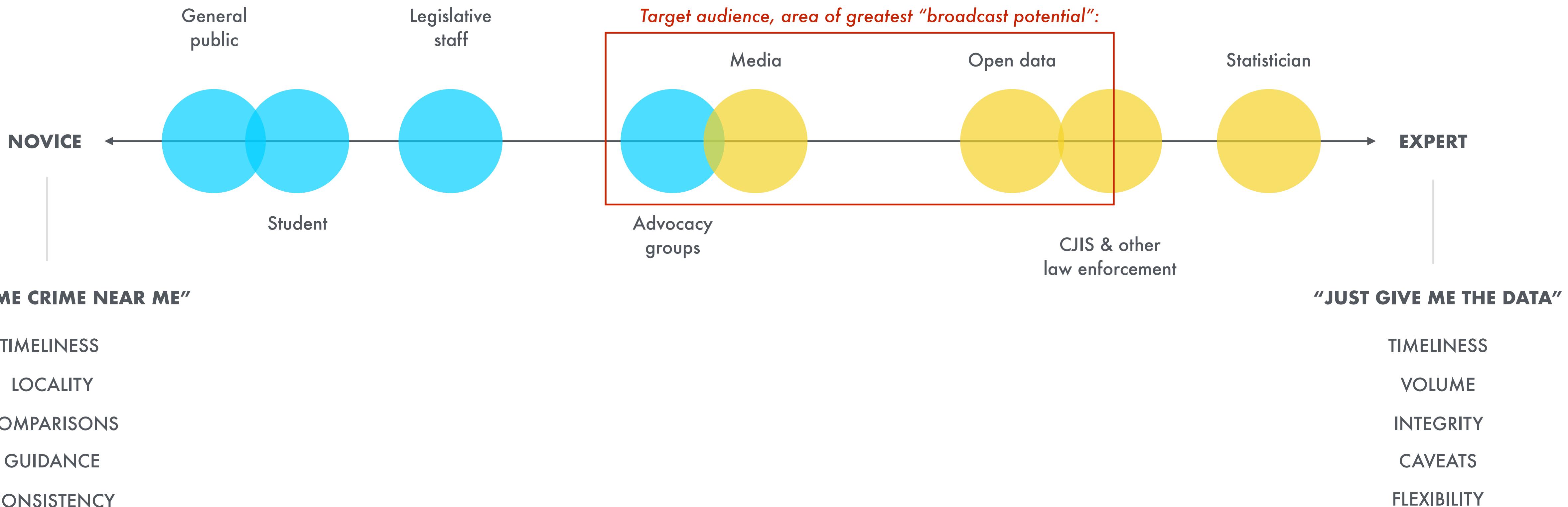
# DEFINING THE NIBRS-CDE DATA MODE



# **NIBRS IS RICHER, BUT ALSO MORE COMPLICATED**

## THE USER COMMUNITY

Research suggests that the typical CDE users is familiar with the UCR program and or working with open data. Similar to the “novice”, they value context, but prefer working with the data in it’s primary form and favor downloads & the API over the CDE interface.



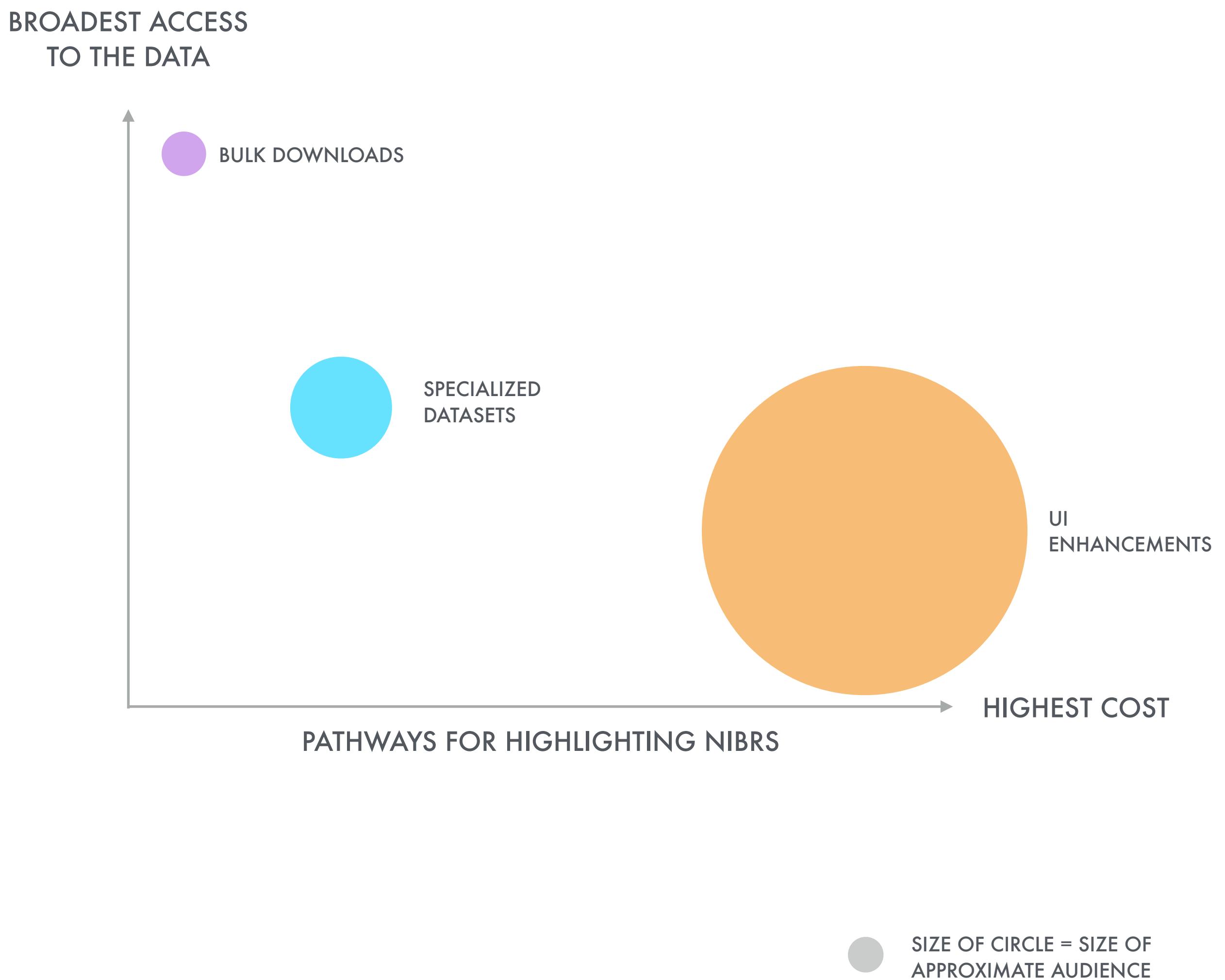
## MEETING USERS WHERE THEY ARE

"Meeting users where they are", or providing multiple pathways to the data based on user needs and preferences, has been a key principle of our work to date. CJIS will need to revisit how it makes the data available and which channels (i.e. downloads, API, or user interface) it wants to invest in as part of its broader strategy for evolving the CDE.

The channel that provides the broadest access to the data today are the NIBRS bulk downloads, which are available by state and year. These files allow users to load their own local copies of the NIBRS database, which provides incredible flexibility for exploring the data, but likely reaches only a small audience of "power users."

Our current discovery work has explored the value of more specialized datasets - single-table downloads that can be loaded into a spreadsheet instead of a database and are thus more focused in their subject matters. At the same time, specialized datasets are built in the CDE database and provided as downloads, making them cheaper to build than potentially costly changes to the user interface.

Looking ahead we recommend that the FBI prioritize building out more specialized datasets. This approach would bring two main benefits: first, it would allow CJIS to score a couple of quick wins since these datasets can be added at relatively low cost, second, the new datasets could serve as a "proving ground" that helps validate those datasets that are of most interest to users before representing them in the interface.



# HOW DOES THE STATE OF THE NIBRS PROGRAM AFFECT ITS PRESENCE IN THE CDE?

NIBRS is a huge leap forward over prior crime-reporting datasets, but it is not without drawbacks that can affect user perceptions

---

## LIMITED PARTICIPATION REDUCES USEFULNESS

- Currently only 30% of police agencies participate in NIBRS.
- Only 21% of 84 city agencies that cover more than a million people participate in NIBRS
- Participation will improve as 2021 deadline nears, but for some states there won't be five years of NIBRS data until 2026.

## MORE DATA, BUT MISSING KEY DETAILS

- Users in testing sometimes expected NIBRS to be an aggregated police blotter
- Geocoding was most common thing users expected to be in there.
- Detailed only up to a point (e.g., can get listings of gun thefts but not types or models or counts of guns stolen)

## COMPLEXITY CAN BE OVERWHELMING

- Users often struggled to understand what crime data was "only in NIBRS"
- Lack of a hierarchy provides freedom to explore crime from any direction but it makes it harder to get started

## DATA COLLECTION PROCESS AFFECTS QUALITY

- NIBRS submission rules may affect what data is recorded for different types of offenses
- Some fields are completely optional or mandatory only after a certain year
- Agencies and states have wide discretion in classifying notable attributes like hate crimes or gang violence

## BULK DOWNLOADS

NOVICE←



Statistician  
Data Journalist

Direct dumps of the NIBRS database tables for a single state and year, for skilled database analysts to load into local databases for their own analysis.

### OPPORTUNITIES

With minimal work to create individual data files, this is the most “honest” view of the data.

Ideal format for database experts who just want to work with NIBRS data in local databases and can build complicated SQL queries without much help.

Users are on their own. CJIS does not need to perform additional validation or support specific tables here.

### CHALLENGES

Downloads are very large. A single zip for a single year of data from a medium-sized state is 35MB compressed

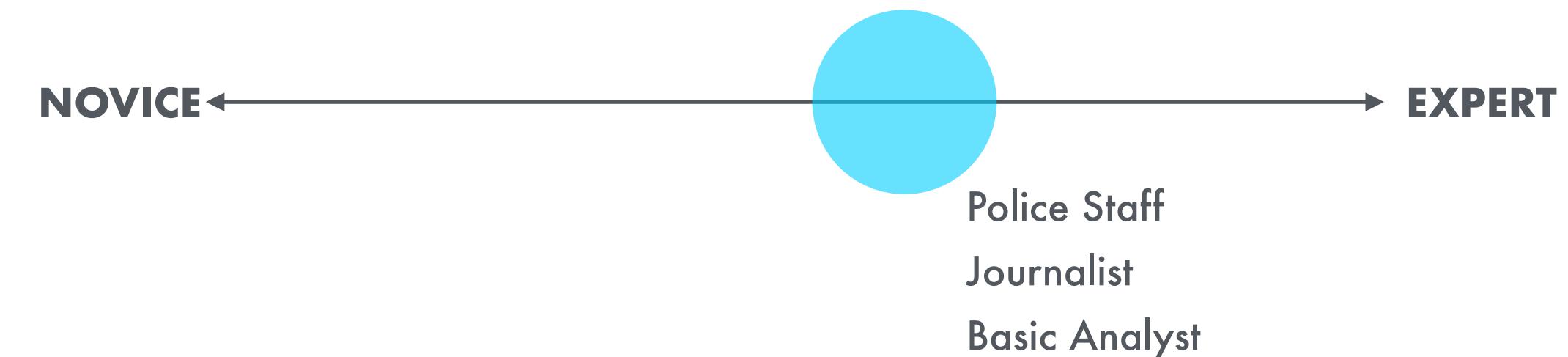
The audience of users who can work directly with this data is relatively small, although they might produce visualizations that would be viewed by larger audiences

State and police agencies are looking to CJIS to guide how NIBRS data is responsibly used in reporting

Some users will make mistakes when they build their own reports from this data. How much should CJIS care?

Not easily apparent to user how much NIBRS coverage is in a particular state or the quality of the data

## SPECIALIZED DATASETS



NIBRS data distilled into single-file downloads for people comfortable with using a spreadsheet. This can include both information on NIBRS incidents or higher-level summary aggregations.

## OPPORTUNITIES

By building simpler files that can be loaded in spreadsheets, CJIS can share NIBRS data with a much larger user base.

CJIS can lead by example, establishing proper techniques and methodologies for making datasets from NIBRS

Specialized datasets are still large enough to support many potential types of additional analysis

Specialized datasets and the intermediate tables that would be created to help build them could be shared in API methods and other public-facing tools

## CHALLENGES

Building specialized datasets involves making many design decisions that require careful planning and consistency

There will need to be an automated build system for generating these files and tracking dependencies between all the tables used to make them

There is a risk that decisions on what data to include, what datasets to prioritize, and if any datasets are discontinued would be seen as politically-motivated

Downloads are often still large. A single zip for a single year of data is 20MB compressed, making it hard to provide data for all years

## UI ENHANCEMENTS

There are multiple areas in which NIBRS data could enhance what is currently on display in the Crime Data Explorer. All of these would reach larger audiences but at a significantly greater cost.

## OPPORTUNITIES

Expand the NIBRS-derived incident details on states and agencies to add histograms specific to that offense type – e.g., *building types for arson, drug type for drug possession charges, or property types for burglary* – or to remove panels that don’t apply – e.g., *victims in crimes against society*.

Add new offense type categories to the Explorer to include new offense types like drug possession or fraud offenses that currently are not included in any FBI reporting.

Add “data stories” as interactive views of certain specialized datasets, allowing users to drill into specific aspects of the data



## CHALLENGES

New detail panels would require significant work to both build the underlying data tables and select which incident details are displayed for which offense types. Many interesting details would only apply to offenses that aren’t part of the Crime Data Explorer yet.

These NIBRS-only crimes would lack any corresponding SRS trend data, so CJIS would have to validate the proper methodology for deriving state-level trends from NIBRS data and compute the data.

Data stories would often be oriented around victims or other entities that are quite different than the incident-level orientation of the CDE

## STEPS FOR BUILDING A SPECIALIZED DATASET

Violence Against Women, an example specialized dataset is a 73-column CSV. Each row represents a single victim of one or more major offenses in a criminal incident, and it provides basic demographic information about the victim, weapons used, injuries received and other circumstances of the crime. In addition, each row contains a count of offenders that were family of the victim, known to the victim or otherwise unknown and detailed information about the 3 most familiar offenders to the victim.

### BASIC STRUCTURE

**1**

- Pick focus: incidents, offenses, **victims**, offenders or property
- Determine parameters to select the correct records
- Denormalize code values into the final table

### SUPPORTING TABLES

**2**

- Build many “flat” tables that reorient rows like weapon type or criminal acts to columns
- Make a complex flat table that maps 3 most familiar offenders for each victim

### FINAL PRODUCTION

**3**

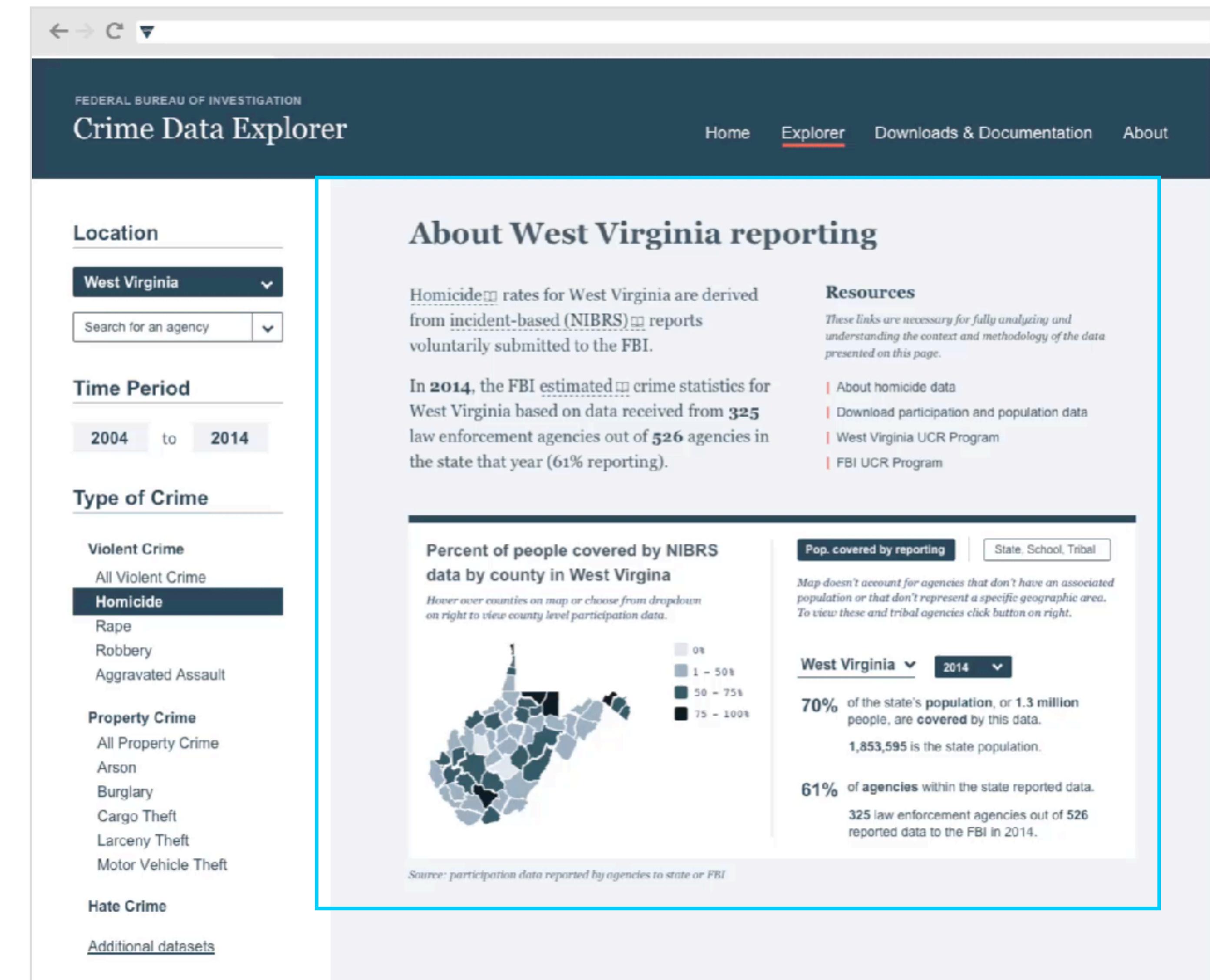
- Create a final table that combines supporting tables to make each row a single record
- Determine how to split into one or more files to add to downloads page

# CONCEPT DESIGNS

# MAPPING POPULATION COVERED

Many of the users we spoke with explained that they take steps to assess the reliability of a dataset before they use it. Determining reliability can be a subjective process, but typically involves an examination of the source/publisher of the data and its reputation, a high-level review of the volume and integrity of the data, and a look at the quality of its supporting documentation. Data from the CDE rates highly in each of these areas, but since participation in the UCR program is voluntary, the volume of available data can be uneven. Within this context, we found that users were especially interested in the participation rate for the area they were viewing data for, with high participation suggesting that the data was more likely to be trusted.

This concept explores how percent of the population covered (a measure of participation that takes into account population density) may be mapped over a geographical areas to build trust in the data, or to warn users when participation is low. Similarly, the concept provides a county-based view of participation within a state, to provide an at-a-glance view of the maturity of the UCR program in that area. The UCR staff we spoke with suggested this might be a helpful feature for promoting broader visibility and accountability within the program.



*Click image to play*

## KEY DESIGN DETAILS

- 1** **Intro paragraph** stays almost the same, but we suggest adding a percent indication at the end of the second paragraph.
  - 2** **Resource list** is moved from below the intro paragraph to the upper right of the page, where the map used to be. A sentence is added describing the importance of these resources. We learned from testing that users can sometimes miss useful information, so we want to call this sections out more and give better direction/context to users.
  - 3** **Interactive map** (main feature of this concept) of percent of population covered by county, and a toggle button for State, School, and Tribal coverage.
- 3A** **Choropleth map** of the state, with county distinctions and colored based on percent reporting (0%, 1-50%, 50-75%, 75-100%). When user hovers over map, the county information will appear on the right. When user first lands on the page however, the overview for the whole state is pre-populated.
- 3B** **Dynamic data section**, similar to the legend section of the trend charts currently implemented on the CDE. This section displays the name of the state or county selected and accompanying data. County can be selected from the map or the dropdown, as well as year.

# About West Virginia reporting

**1**

Homicide rates for West Virginia are derived from [incident-based \(NIBRS\)](#) reports voluntarily submitted to the FBI.

In [2014](#), the FBI estimated crime statistics for West Virginia based on data received from **325** law enforcement agencies out of **526** agencies in the state that year (61% reporting).

**2**

## Resources

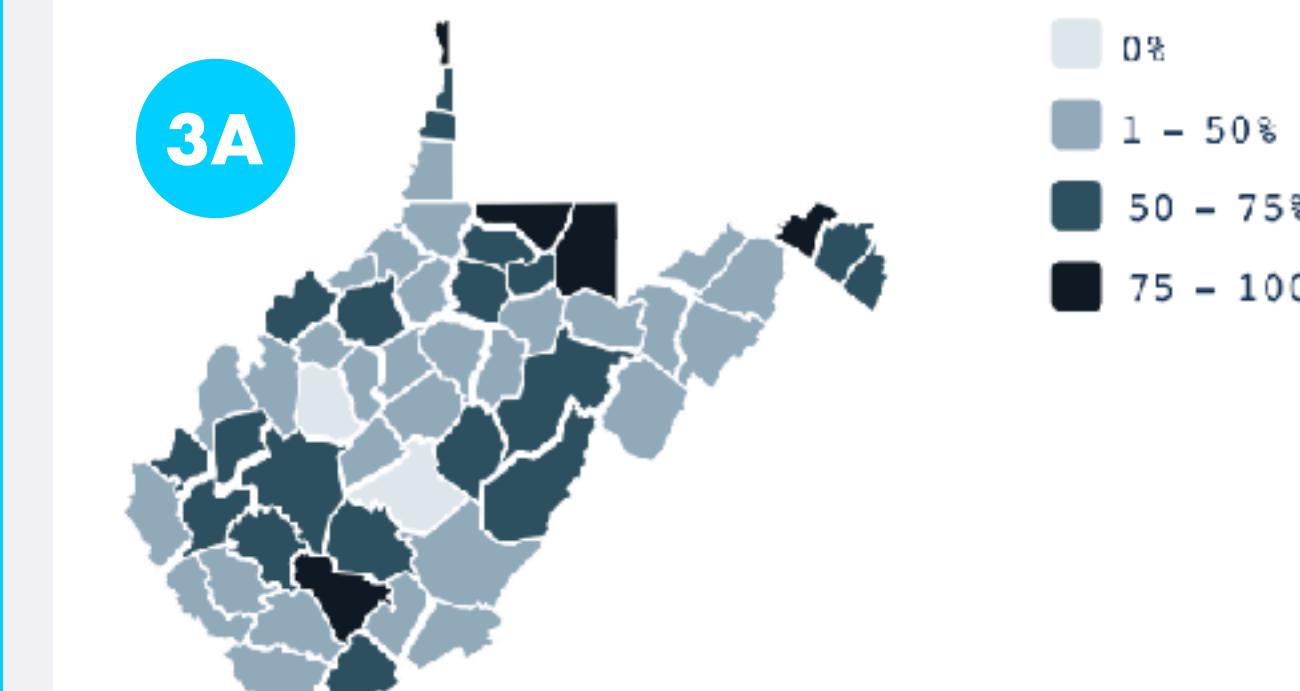
*These links are necessary for fully analyzing and understanding the context and methodology of the data presented on this page.*

- [About homicide data](#)
- [Download participation and population data](#)
- [West Virginia UCR Program](#)
- [FBI UCR Program](#)

**3**

## Percent of people covered by NIBRS data by county in West Virginia

*Hover over counties on map or choose from dropdown on right to view county level participation data.*

**Pop. covered by reporting****State, School, Tribal**

*Map doesn't account for agencies that don't have an associated population or that don't represent a specific geographic area. To view these and tribal agencies click button on right.*

**West Virginia****2014****3B**

**70%** of the state's **population**, or **1.3 million** people, are **covered** by this data.

**1,853,595** is the state population.

**61%** of **agencies** within the state reported data.

**325** law enforcement agencies out of **526** reported data to the FBI in **2014**.

## CONCEPT DESIGNS: BUILDING CONFIDENCE IN THE DATA

### KEY DESIGN DETAILS

**4** **County selection** interaction: this shows what a selected county looks like (outlined in red) after the user has hovered over the map or used the dropdown on the right. To stay on the county selected, the user can click the map instead of hover, or use the dropdown. The dynamic information on the right has updated to the county's data, instead of the default state view.

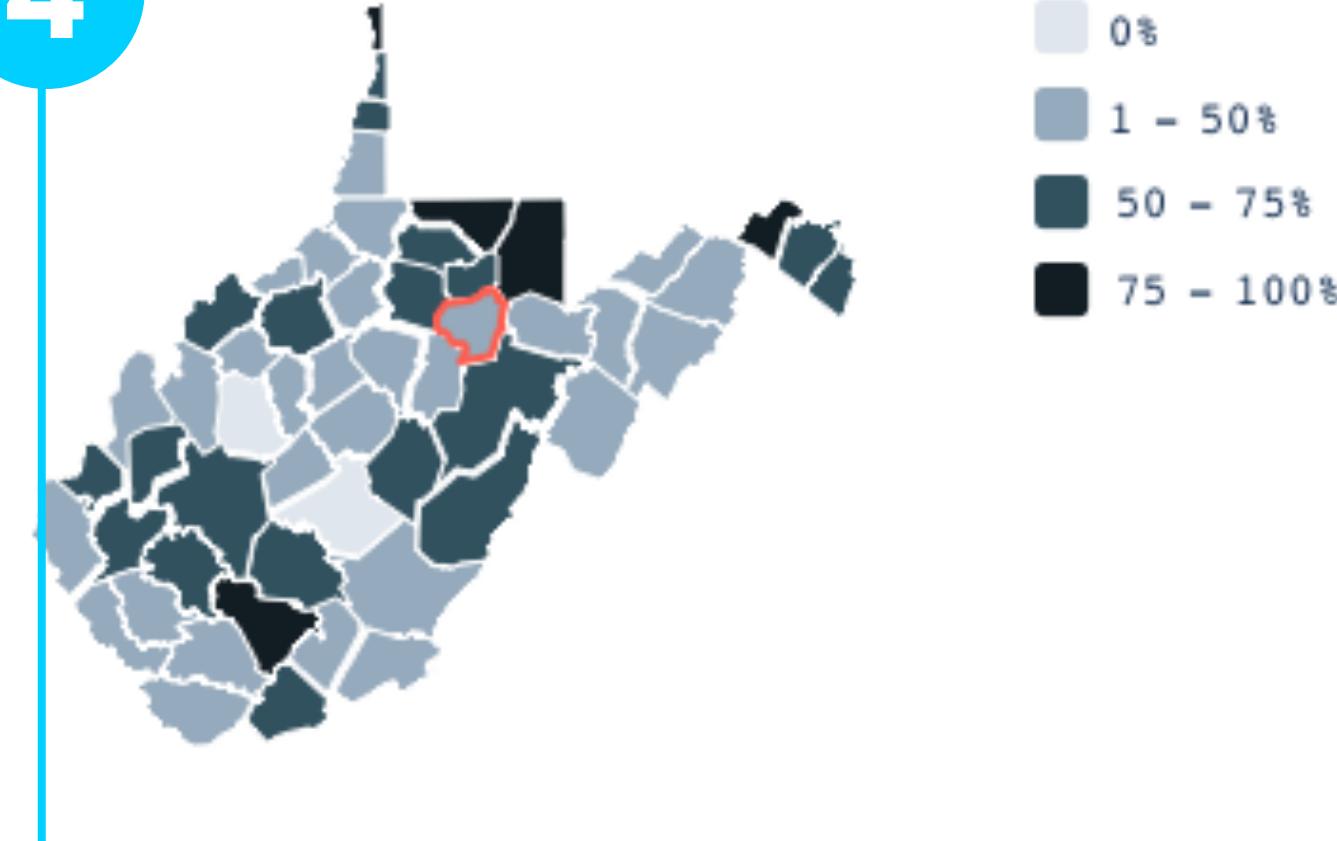
**4A** **County selector** drop down.

**5** **Button to switch** from population covered to State, school, tribal agency participation data. Since these are agencies that don't have a chartable geography or reliable population count, we have to treat this information differently, but within the same feature.

### Percent of people covered by NIBRS data by county in West Virginia

*Hover over counties on map or choose from dropdown on right to view county level participation data.*

**4**



**5**

**Pop. covered by reporting**

**State, School, Tribal**

*Map doesn't account for agencies that don't have an associated population or that don't represent a specific geographic area. To view these and tribal agencies click button on right.*

**Barbour County** ▾

**2014** ▾

**70%** of the county's **population**, or **11,693** people, are **covered** by this data.

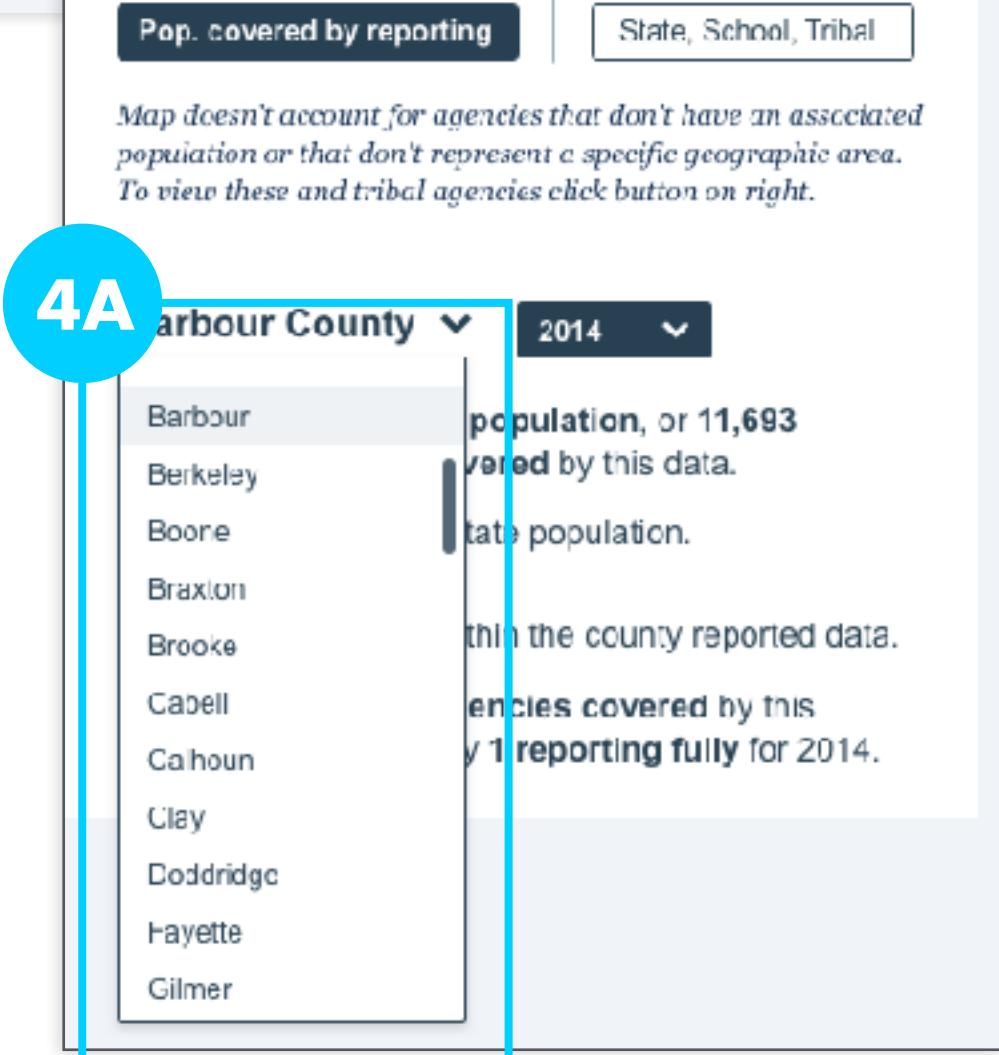
**16,704** is the state population.

**61%** of **agencies** within the county reported data.

There are **4 agencies covered** by this county, but only **1 reporting** fully for 2014.

*Source: participation data reported by agencies to state or FBI*

**4A**



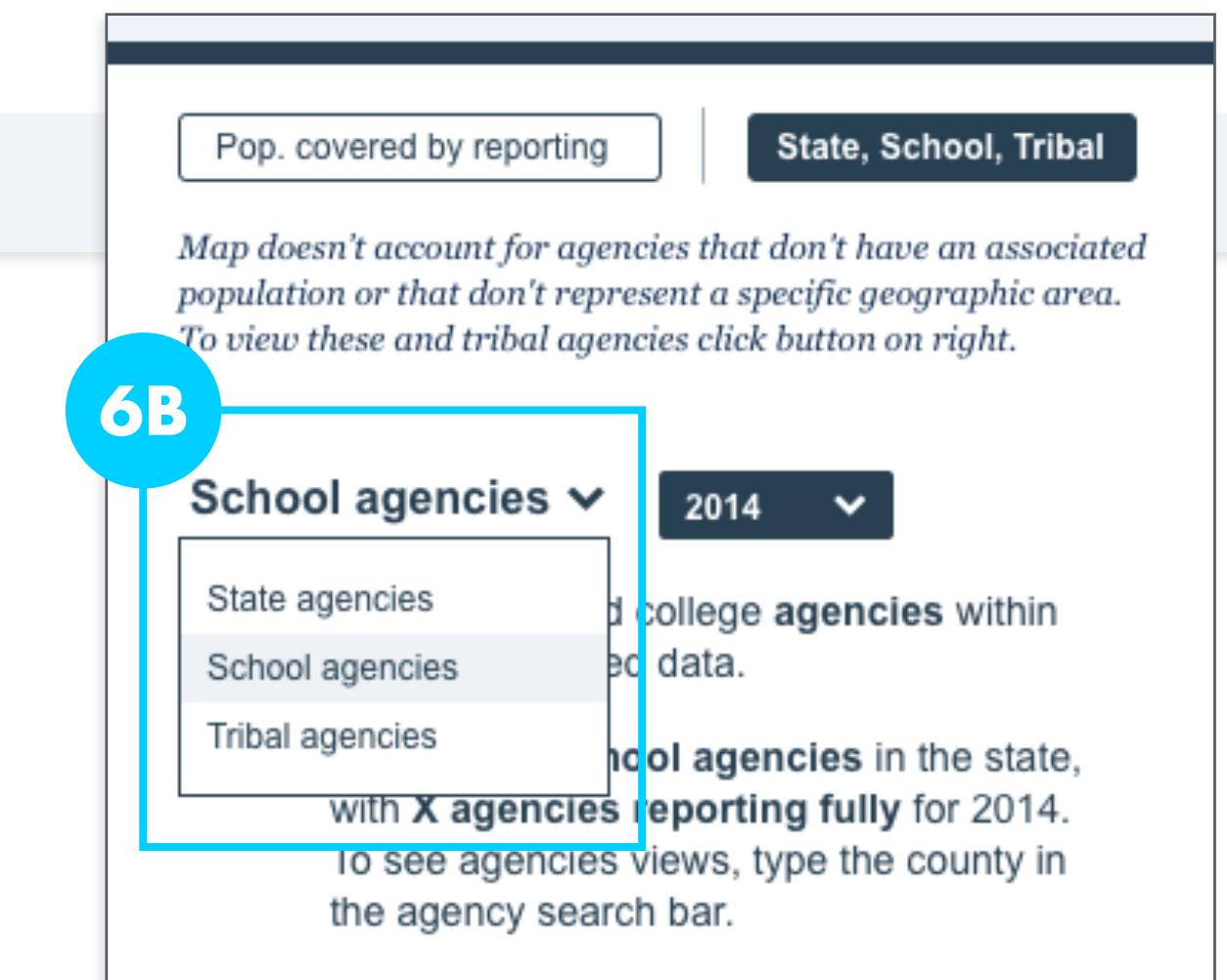
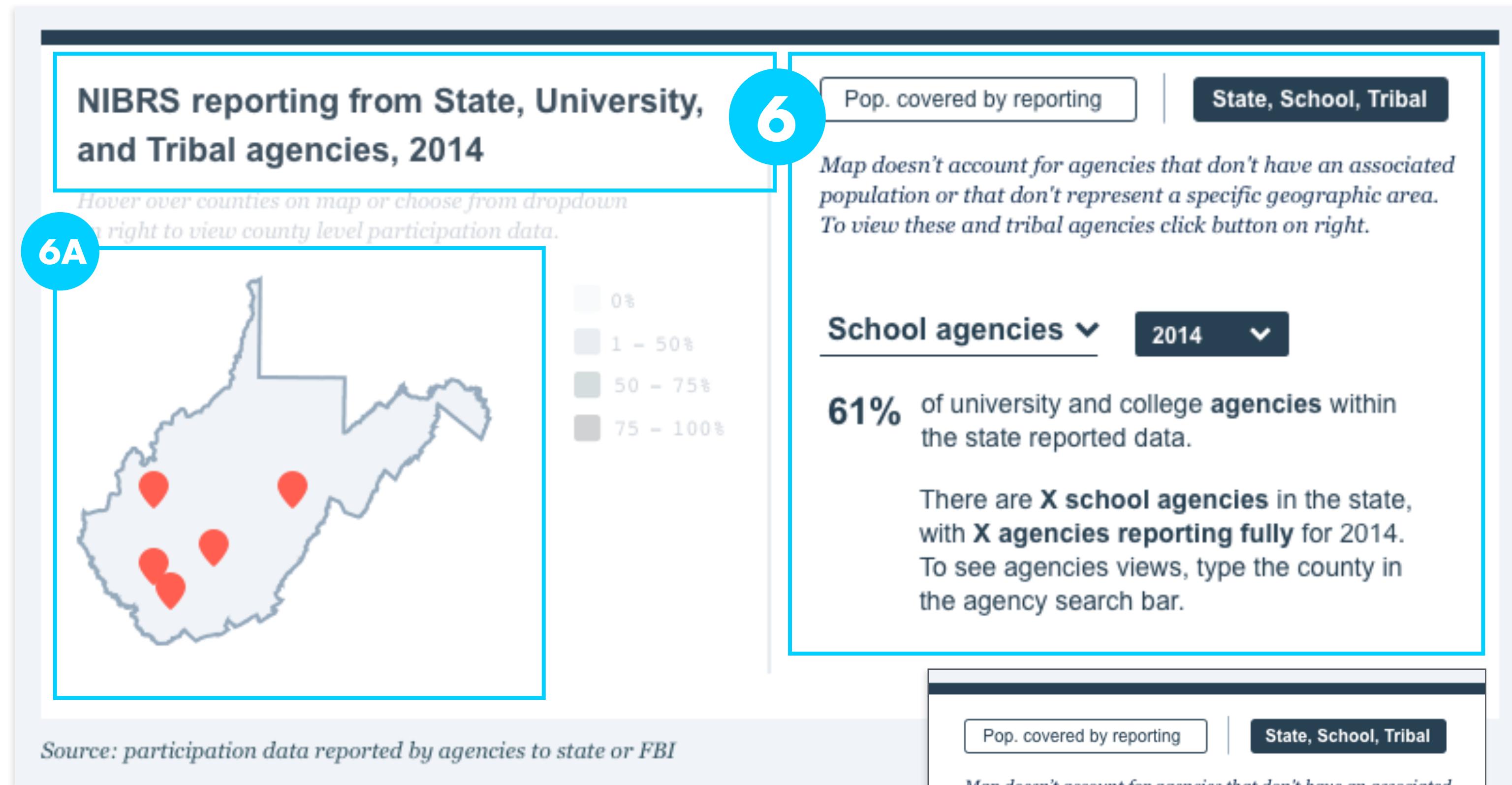
## CONCEPT DESIGNS: BUILDING CONFIDENCE IN THE DATA

### KEY DESIGN DETAILS

**6** View of State, School, Tribal option in interface. The button replaces the headline, removes the county lines from the map and replaces county name dropdown with the other options (State, School, or Tribal agencies).

**6A** The map here could be used to show where University or college agencies are located for example, but may not serve a function for state or Tribal agencies (Tribal could perhaps show general geographic location).

**6B** Dropdown selector example.



## CONCEPT DESIGNS: BUILDING CONFIDENCE IN THE DATA

### KEY DESIGN DETAILS

**7**

**Participation layout for Downloads & Documentation page.** This feature can also be utilized in the new preview design for the Downloads & Documentation page. Since many users, especially data professionals, go straight for the download section and don't spend much time on the Explorer pages, they might miss this useful participation data, which is key to fully understanding the downloads themselves. Since there is more space in the layout of the Downloads & Documentation pages (no side navigation like in the Explorer), the design for this feature, while still having the same basic functionality, is slightly different.

**7A**

This section stays as an overview of NIBRS for the state.

**7B**

This section is dynamic for countries and state, school, and tribal agencies.

**8**

**Participation data for specialized dataset (violence against women):** in this case, the left side would stay as a national view overall, and the right would be state overviews. This is helpful for users because they can see whether or not a state they are interested in reports NIBRS.

**7**

### West Virginia | 2016 state reporting overview

**7A**

**73%** of agencies participated in the NIBRS program.

**83%**

of this state's population was covered by the data.

**1,831,102**

is the state's population.

In 2016, the FBI received NIBRS data from 437 law enforcement agencies out of 528 agencies in West Virginia.

**7B**

#### Percent of people covered by NIBRS data by county in West Virginia

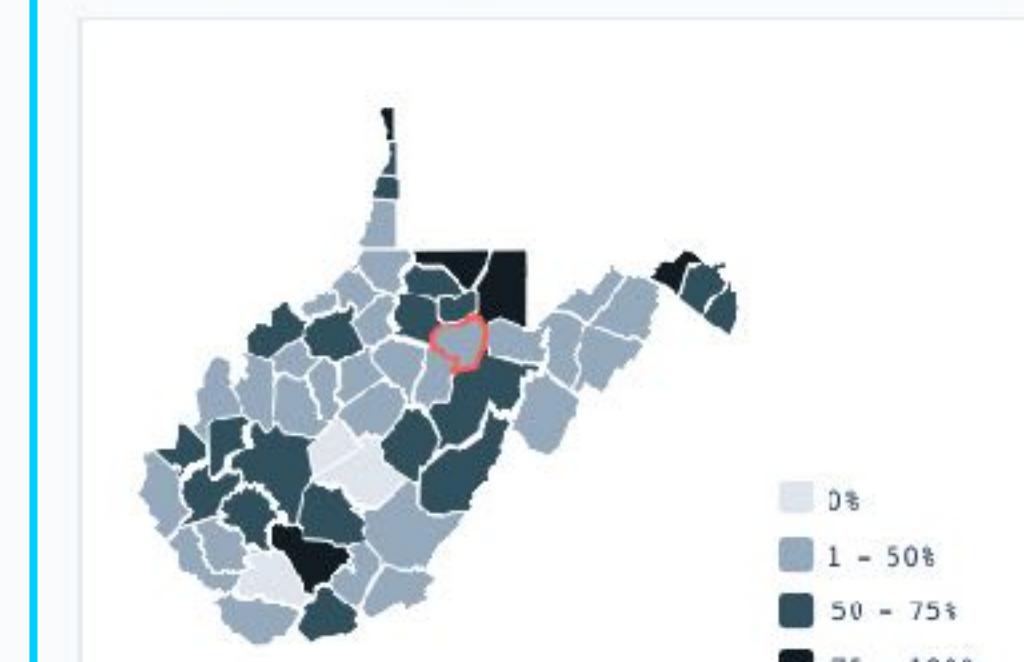


Chart does not account for State and University departments that do not have a population or chartable geography. To view these and tribal, click option on below.

**Pop. covered by reporting**

**State, School, Tribal**

**8**

### Violence against women | 2016 U.S. reporting overview

**73%**  
of agencies participated in the NIBRS program.

**83%**  
of this country's population was covered by the data.

**323.1 million** is the national population.

In 2016, the FBI received NIBRS data from X law enforcement agencies out of X agencies nationwide.

Percent of population covered by NIBRS reporting by state



● Territories including American Samoa, Puerto Rico, Guam, and the U.S. Virgin Islands submit voluntary data to the FBI. Alaska and Hawaii submit data directly to the FBI.

**Pop. covered by reporting** | **State, School, Tribal**

West Virginia

**70%**  
of people covered by reporting in state.

**1.9 million**  
is this state's population.

In 2014, the FBI estimated crime statistics for West Virginia based on data received from 125 law enforcement agencies out of 526 agencies in the state that year.

**Barbour County**

**70%**  
of people covered by reporting in county.

**16,704**  
is this county's population.

There are **4 agencies** covered by Barbour County, but only **1 reporting fully** for 2016. To see agency views, type the county in the agency search bar.

## KEY QUESTIONS AND CONSIDERATIONS

This concept requires an accounting of population covered by the UCR program by county and state. The groundwork for visualizing this data is already in place; although some API endpoints may need to be changed or created in order to fully support the view. Similarly, the view requires a revised and expanded participation download, which includes participation data at the county level. Currently the Explorer only provides access to a roll-up of participation rates by year for the state.

The design also requires a clear methodology for how coverage rates are derived. This methodology should be written in plain language in order to be accessible to a wide range of users and explain how population is assigned by the UCR program and how the CDE accounts for agencies that span county boundaries, or which have no counties assigned to them at all (i.e., schools and tribal agencies.)

While this concept tested favorably, we were not able to assess its potential impact on the mobile experience.

## SAMPLE USER STORIES

As a user I can view a map showing the percentage of population covered by available UCR data by state and county.

As a user I can download participation data for the state that includes a breakdown of participation by county and years.

As a user I can view and download a list of agencies in a state that aren't assigned population or counties, including state agencies, tribal agencies and schools.

## RELATED RESOURCES

### INVISION PROTOTYPE

<https://gsa.invisionapp.com/share/3RFHY8JKG8F>

### CDE STYLEGUIDE

<https://github.com/18F/crime-data-explorer/wiki/Style-Guide>

## CONCEPT DESIGNS: IMPROVING CONTEXT

# DYNAMIC FOOTNOTES & NOTATIONS

In our testing we learned that users didn't always understand what the data they were viewing meant and how it was derived. This confusion was exacerbated when there were gaps or unusual changes in an agency's reporting from year to year. Lacking context, a user might assume a surge in Orlando's homicide rate in 2016 was due to other factors beyond the inclusion of the Pulse nightclub shooting in the city's homicide totals.

This concept explores how the user interface could be extended to dynamically include specific annotations of CDE data. Currently the CDE includes only limited information about methodology, buried in static footnotes or disclaimers that redirect the user offsite. Adding functionality that allowed for these details to be more easily accessible and updates would not only improve the user experience, but help mitigate risks that users would misinterpret UCR data.

### Offender Demographics

**GENERAL METHODOLOGY**

The NIBRS database captures details about specific crime incidents, such as the gender or race an offender. This data feature provides context and insight into criminal activity, such as where a crime occurred and how that compares to where other similar crimes have occurred.

For the selected date range, the Crime Data Explorer groups crimes based on their details with other incidents in the NIBRS database and then counts them to present a summary view.

**OFFENDER METHODOLOGY**

We calculate offender demographics by a simple mechanism of group-and-count.

First we group offenders by specific metrics, which include:

- State
- Year
- Type of offense
- Agency that reported the offense

To make sure we account for every offender involved in a specific incident, we map offenses to offenders.

After grouping by metric we count the total number of occurrences for each value. In addition, we count the total number of offenders for that specific grouping (i.e., state-year-offense\_code) to include offenders for which demographic information is not known.

**NOTES:**

1. Offenders are counted once for every unique type of offense in an incident. An offender who is suspected of multiple counts of the same offense will only be counted as a single offender. When an offender is suspected of multiple crimes in an incident, that offender may show up in the counts for other offenses. In some cases, specific demographic information may not be known for an offender. Those offenders will not be included in the demographic charts, but would be counted in the number of total offenders.
2. Ethnicity is not a required reporting field, which may affect the amount of data we are able to provide.
3. To understand how race and ethnicity are reported and why, please see the definition provided by the Office of Management and Budget.

[View more information about Homicide data below.](#)

For more information about working with NIBRS data, view [system user manual](#) from CJIS.

### Victim Demographics

**Sex of Victim**

Sex	Count
Female	25
Male	75
Unknown	0

There were 100 reported victims.

**Age of Victim**

Age Group	Count
0-10	10
10-20	20
20-29	30
30-40	25
40-50	15
50-60	10
60-70	5
70-80	2
80-90	1
90-100	1

VICTIM AGE

There were 1,700 incidents involving victims with a reported age of 20-29.

[Download data](#)

Sources: FBI, [FBI Uniform Crime Report data for Tennessee, 2004–2014](#).

### Rape offenses reported by Chicago Police Department, 2004–2014

**RAPE COUNT METHODOLOGY**

We calculate rape count by a simple mechanism of group-and-count.

- State
- Year
- Agency that reported the offense

To measure rape count for every offender involved in a specific incident, we map offenses to offenders.

After grouping by metric we count the total number of occurrences for each value. In addition, we count the total number of offenders for that specific grouping (i.e., state-year-offense\_code) to include offenders for which demographic information is not known.

**NOTES:**

YEAR FOOTNOTE

2010 The data collection methodology for the offense of forcible rape used in Chicago, Illinois, does not comply with national Uniform Crime Reporting Program guidelines. Consequently, figures for forcible rape and violent crime (of which forcible rape is a part) are not published in this table.

2013 The FBI determined that the agency's data were underreported. Consequently, those data are not included in this table. The data collection methodology for the offense of rape used by Chicago, Illinois does not comply with national UCR Program guidelines. Consequently, its figures for rape and violent crime (of which rape is a part) are not published in this table.

[Download data, methodology & notes](#)

View more information about rape data below.

For more information about working with NIBRS data, view [system user manual](#) from CJIS.

## CONCEPT DESIGNS: IMPROVING CONTEXT

# KEY DESIGN DETAILS

**1** **Methodology buttons** are added to all charts reading, "How these numbers are calculated." Clicking any of these will make the methodology view slide in to the frame of the chart, covering the charts.

**2** **Methodology and notes view:** When this feature is open, the light blue background should cover the charts at a 95% opacity. The user can exit this view and see the charts again by clicked the "Close X" in the top right corner.

The general content template for the has space for these sections:

1. General methodology
2. [Specific attribute e.g. "Offender"] methodology
3. Notes (footnotes and other important information)

**2A** Additional links to the About section at the bottom of each explorer page, and a link to the NIBRS manual are available here, if users are interested in learning more. This methodology feature, in combination with the About section below, covers what a user needs to understand to work well with the data we are showing.

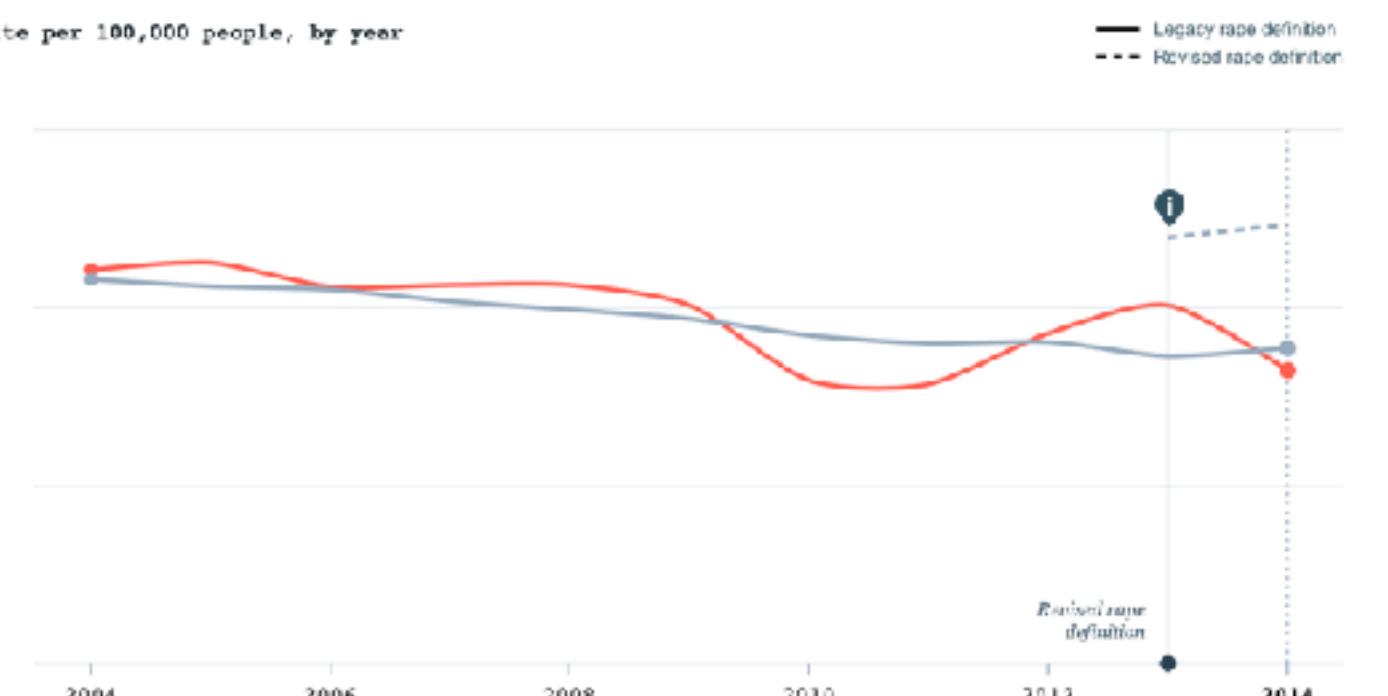
## Crimes reported

### Rape rate in Illinois, 2004–2014

How these numbers are calculated

In 2014, the rate at which rape was reported using the legacy definition was 24.7 per 100,000. Rape was reported using the revised definition at a rate of 41.5 per 100,000 people.

2014	Rate	Total	Population
Illinois	24.7	3,182	12,882,189
United States	25.6	84,864	318,907,401



Source: FBI, Estimated data for Tennessee, 2004–2014.

### Rape incidents details reported by Illinois

There were 1,250 individual rape incidents reported to the FBI in Illinois between 2004 and 2014 by 2 law enforcement agencies reporting data.

#### Offender Demographics

How these numbers are calculated

##### Sex of Offender

View charts by  %

Female	11%
Male	83%
Unknown	6%

There were X reported offenders.

#### Victim Demographics

How these numbers are calculated

##### Sex of Victim

View charts by  %

Female	25%
Male	75%
Unknown	1%

There were X reported victims.

2

## Offender Demographics

### GENERAL METHODOLOGY

Close X

The NIBRS database captures details about specific crime incidents, such as the gender or race an offender. This data feature provides context and insight into criminal activity, such as where a crime occurred and how that compares to where other similar crimes have occurred.

For the selected date range, the Crime Data Explorer groups crimes based on their details with other incidents in the NIBRS database and then counts them to present a summary view.

### OFFENDER METHODOLOGY

We calculate offender demographics by a simple mechanism of group-and-count.

First we group offenders by specific metrics, which include:

- State
- Year
- Type of offense
- Agency that reported the offense

To make sure we account for every offender involved in a specific incident, we map offenses to offenders.

After grouping by metric we count the total number of occurrences for each value. In addition, we count the total number of offenders for that specific grouping (ie, state-year-offense\_code) to include offenders for which demographic information is not known.

**NOTES:**  
593 Black or African American  
<13 Native Hawaiian or Other

1. Offenders are counted once for every unique type of offense in an incident. An offender who is suspected of multiple counts of the same offense will only be counted as a single offender. When an offender is suspected of multiple crimes in an incident, that offender may show up in the counts for other offenses. In some cases, specific demographic information may not be known for an offender. Those offenders will not be included in the demographic charts but would be counted in the number of total offenders.
2. Ethnicity is not a required reporting field, which may affect the amount of data we are able to provide.

3. To understand how race and ethnicity are reported and why, please see the definition provided by the Office of Management and Budget.

2A

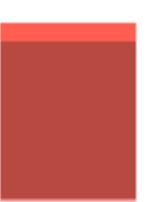
View more information about Homicide data below.

For more information about working with NIBRS data, View system user manual from CJIS.

## Victim Demo

How these numbers

### Sex of Victim



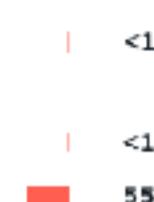
There were X reported

### Age of Victim



There were 1,700 in a reported age of 20

### Race of Victim



Black or African American

Native Hawaiian or Other

Asian

Hispanic

Middle Eastern

American Indian

White

Asian Indian

Arab

Other

Two or more races

Two or more ethnicities

Asian Pacific Islander

Asian Indian

Arab

Other

Asian Pacific Islander

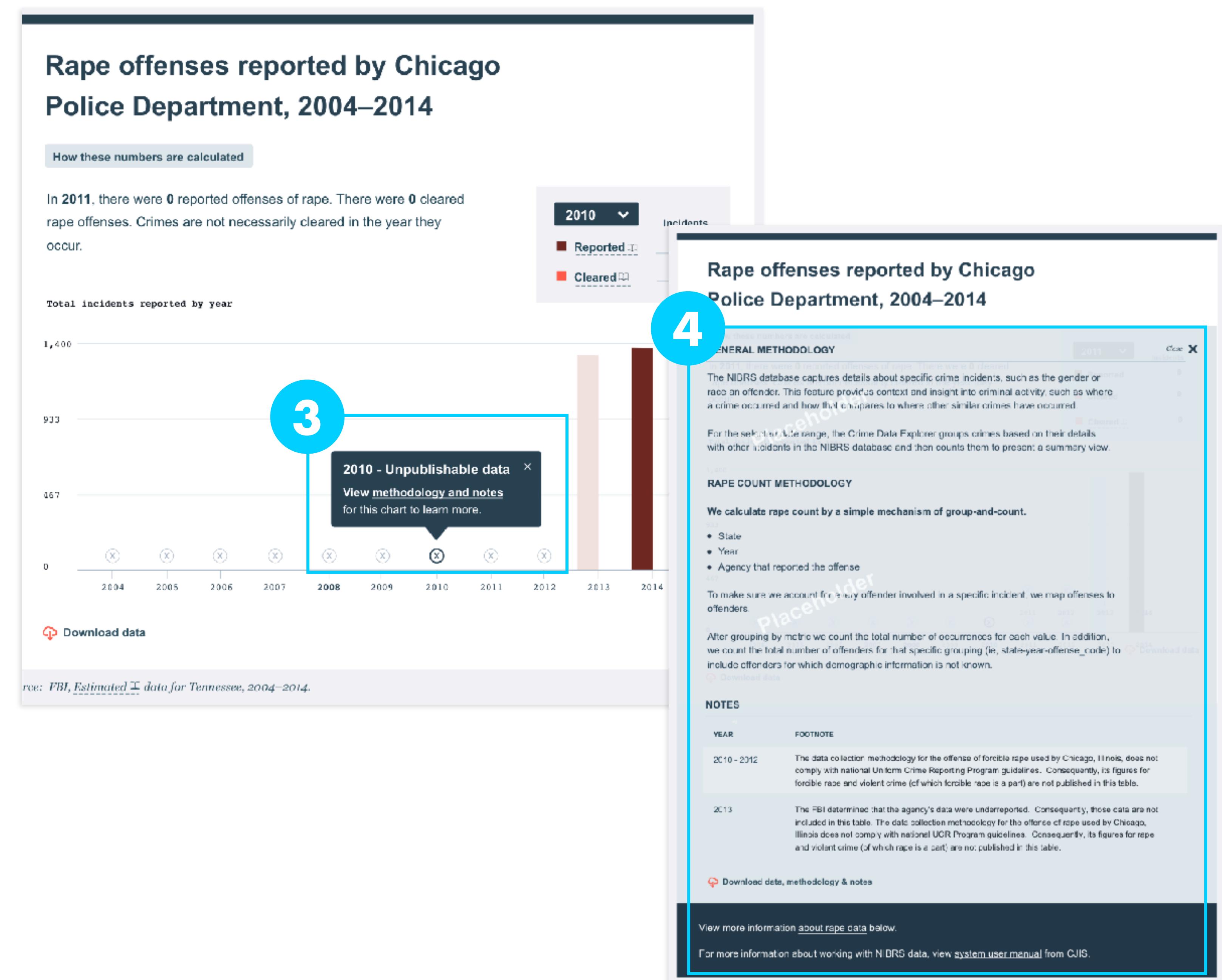
## CONCEPT DESIGNS: IMPROVING CONTEXT

### KEY DESIGN DETAILS

**3** **Interactive tooltips** for missing data or other important information to note about a specific year: this example shows several missing years of data for Chicago PD. Currently on the CDE, there is nothing to help a user fully understand why this data isn't being displayed. With the addition of this tooltip feature, a user can hover over the year and get at-a-glance information, as well as click the link to view the full methodology and footnotes for the chart.

**4** **Methodology view for line or bar charts** in the Explorer: this has the same content template as the NIBRS charts, but in this example we also show a table of the footnotes associated with the chart. There is a link to download the data with the methodology/footnotes as well.

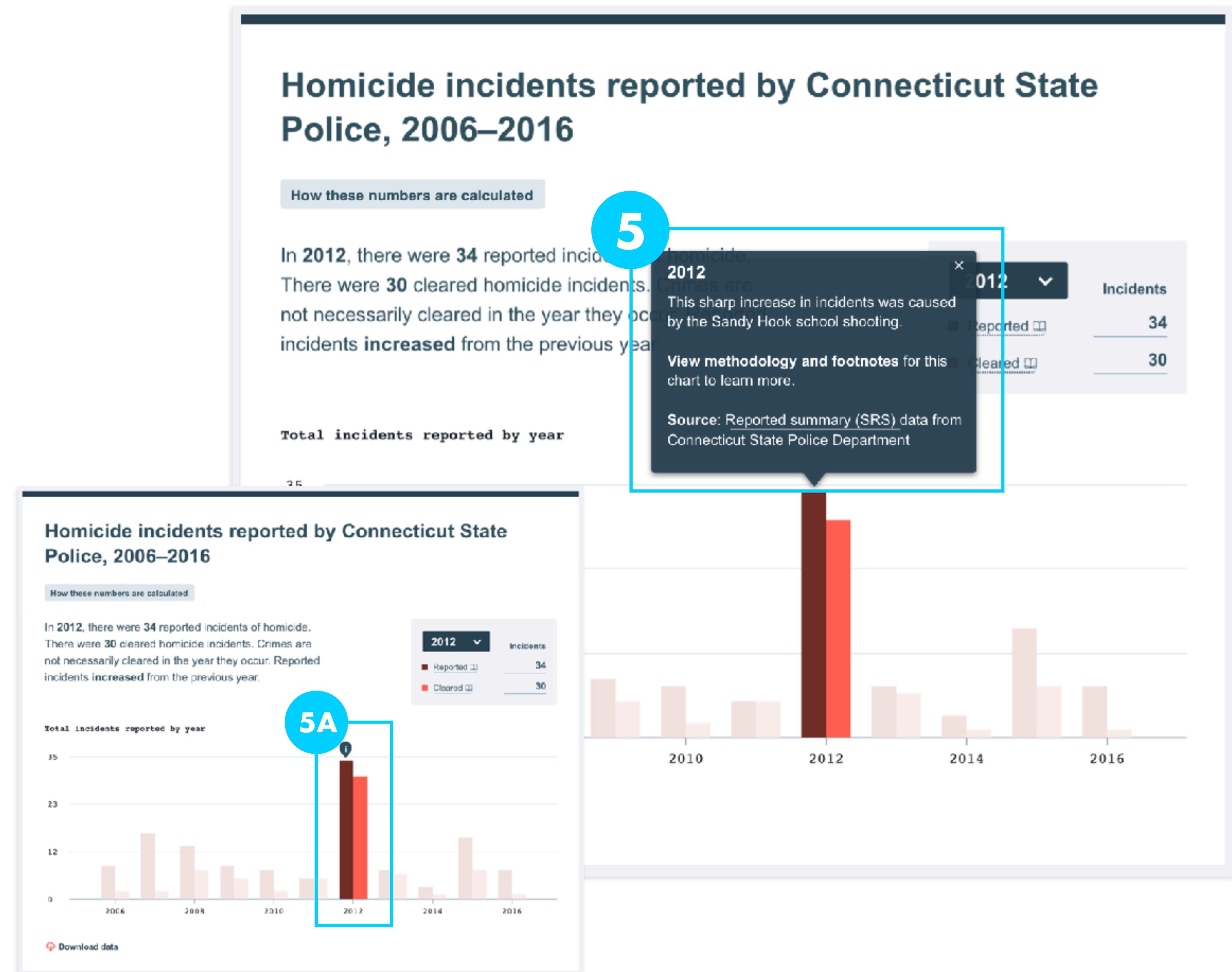
For implementation, this view would get cut off wherever the chart naturally ends. User can scroll within the frame to view the whole segment (there is a "Scroll for more" direction and icon included, as well as a scroll bar included in the design).



## CONCEPT DESIGNS: IMPROVING CONTEXT

### KEY DESIGN DETAILS

- 5** This example shows how the tooltip function can be implemented for a specific event or point of interest in the data, like a spike in crime from a large violent attack. The functionality is the same as the tooltip for other charts, and has the same link to view the full chart methodology and footnotes.
- 4A** Icon notifying user of a clickable point of interest that would open the tooltip.



## KEY QUESTIONS AND CONSIDERATIONS

The key question for this concept is how contextual data - footnotes, annotations, etc. - will be supplied to the CDE. Currently it is manually collected and stored in multiple spreadsheets with little to no version control. This not only makes it difficult to aggregate and display this information in meaningful ways, but also raises the cost of maintaining it.

CJIS should begin to treat the contextual data as a full-fledged part of the CDE database, ideally also including it in the new UCR, so that annotations can be handled programmatically. CJIS should also define a data model for supplying this information to the CDE and consider content management tools and workflows that would allow for this information to be more easily maintained going forward.

## SAMPLE USER STORIES

As a user I want to view the associated methodology for the chart or table I am viewing.

As a user I want to view available annotations for a single year in the time series I am viewing data for.

As a user I want to view and download all footnotes for a time series I am viewing data for.

## RELATED RESOURCES

### INVISION PROTOTYPE

[https://gsa.invisionapp.com/share/J9FI477QBGW#/275373308\\_Methodology-Detail\\_1](https://gsa.invisionapp.com/share/J9FI477QBGW#/275373308_Methodology-Detail_1)

### CDE STYLEGUIDE

<https://github.com/18F/crime-data-explorer/wiki/Style-Guide>

## CONCEPT DESIGNS: DOWNLOAD PREVIEW

# EXPANDING THE DOWNLOAD EXPERIENCE

Similar to the “population covered” concept, we learned that users value the context behind the datasets they’re downloading because it helps them roughly assess its reliability and develop a mental model for how to work with it. We also saw examples of how other government and open data sites accomplish this, by providing a brief narrative and “preview” of the data that is available for download.

While this concept incorporates some of these ideas for previewing the data, it also accounts for the broader product strategy outlined in this report – that CJIS should prioritize adding new downloads to the CDE because they are the preferred way of accessing the data for our target users and they may serve as a proving ground for future user interface changes. Under this approach, the current design for the “downloads and documentation” view will need to be extended in order to account for longer file listings.

FEDERAL BUREAU OF INVESTIGATION  
Crime Data Explorer

Home Explorer Downloads & Documentation About

## Downloads & Documentation

Download incident-based (NIBRS) data by year and location. Estimated data and other crime-related datasets are also available for download. Data is provided as CSV files and can be accessed via the [Crime Data Explorer API](#).

**Resources**

- | Readme
- | Data dictionary

### Choose dataset to preview and download

See dataset descriptions below

Dataset Location Year View data preview

#### Summary and Incident datasets

Available for download by state and year

Incident-based (NIBRS) Data Expand for descriptions

Summary (SRS) Data with Estimates

#### Specialized datasets

Available for download by year

Violence Against Women

Hate Crime

Assaults on Law Enforcement Officers

Police Employee Data

Uniform Crime Reporting Program Participation Data

Cargo Theft

Human Trafficking

US Territory Data

#### West Virginia | 2016 state reporting overview

Percent of people covered by NIBRS data by county in West Virginia

Barbour County ▾

70% of people covered by reporting in county

16,704 is this county's population.

There are 4 agencies covered by Barbour County, but only 1 reporting fully for 2016. To see agencies viewer, type the county in the agency search bar.

Prep. covered by reporting | State, School, Title

Source: preliminary data reported by agencies to the FBI

#### Download details

The NIBRS dataset offers detailed information about the victims, offenders, locations, offenses, and property types associated with reported crime incidents. It also shows the relationships between and among these data elements to enable new perspectives of reported crime data.

For more information on how to work with NIBRS please see the resources on the right.

**STATE**  
West Virginia

**DATA TYPE**  
NIBRS

**YEARS**  
2016

**LAST MODIFIED**  
January 1, 2016

**RELEASE DATE**  
January 1, 2016

**VIEWS**  
50

**DOWLOADS**  
20

**LICENSE**  
Open data

**FILE SIZE**  
5.2 MB

32 MB unzipped

**What's in the download**

FILES	DESCRIPTION
README.html	Useful documentation about the dataset
README.md	Useful documentation about the dataset
agency_participation.csv	Agency, county, and state participation for 2016
cde_agencies.csv	List of agencies and information
nibrs_activity_type.csv	Type of NIBRS activity related to areas:
nibrs_agency.csv	Agencies...
nibrs_arrest_type.csv	Arrest type...
nibrs_arrestee.csv	Arrestee...
nibrs_arrestees_weapon.csv	Weapon used by arrested
nibrs_assignment_type.csv	Assignment...
nibrs_incident.csv	Full list of crimes committed
nibrs_incident_detail.csv	Des motivation of offenses
nibrs_circumstances.csv	Circumstances of...

Scroll for more rows

Download data

## CONCEPT DESIGNS: DOWNLOAD PREVIEW

### KEY DESIGN DETAILS

**1 Redesign of downloads navigation:** to accommodate more datasets as they become available on the CDE, and to provide a more comprehensive download experience, this page has been modified. In this section, the user can select from any available dataset for download, the location (if applicable, since some are nationwide only), and year. After the fields have been chosen, the “View data preview” button becomes active and clickable.

**2 Dataset descriptions:** The datasets are split into two sections, summary and incident datasets, and specialized datasets. The user can click on the title or caret to expand a full description of each.

**2A Expanded description and select button:** User can see the full description of the dataset and choose to select that dataset, which populates their choice into the navigation above, and prompts them to finish choosing a location (if applicable) and year to move on to preview the data.

The screenshot shows the 'Downloads & Documentation' page of the Crime Data Explorer. At the top, there's a header with the FBI logo, 'Crime Data Explorer', and navigation links for 'Home', 'Explorer', 'Downloads & Documentation' (which is underlined in red), and 'About'. Below the header, the main content area is titled 'Downloads & Documentation'. It features a section for 'Resources' with links to 'Readme' and 'Data dictionary'. A large blue circle labeled '1' highlights a box containing a 'Choose dataset to preview and download' form with dropdowns for 'Dataset', 'Location', and 'Year', and a 'View data preview' button. Another blue circle labeled '2' highlights a 'Summary and Incident datasets' section. This section is divided into three categories: 'Incident-based (NIBRS) Data', 'Summary (SRS) Data with Estimates', and 'Specialized datasets'. Under 'Specialized datasets', there are links for 'Violence Against Women', 'Hate Crime', 'Assaults on Law Enforcement Officers', 'Police Employee Data', 'Uniform Crime Reporting Program Participation Data', 'Cargo Theft', 'Human Trafficking', and 'US Territory Data'. A blue circle labeled '2A' highlights a detailed description of the 'Incident-based (NIBRS) Data' category. The description explains that NIBRS was implemented to improve data quality and includes a note about its transition from SRS format. A red box surrounds the 'Select dataset for preview' button.

## CONCEPT DESIGNS: DOWNLOAD PREVIEW

### KEY DESIGN DETAILS

**3 NIBRS dataset bulk download preview:** After the user selects their dataset to preview, the page refreshes to show the relevant reporting participation data for the location (3B - explained in greater detail in previous section of presentation) and the download preview below. First is a spot for a “Downloads detail” paragraph and any other useful resources that may not be in the bulk download itself.

**3A Download metadata:** the left side of the data preview section is where useful metadata about the files the user is about to download can be found.

**3B List of files included in download and descriptions:** Since our downloads have more files in them for NIBRS now, it's even more important to provide users with a more comprehensive view of the large group of files they are about to get. Since there are many rows, this section (when applicable) should scroll. While this shows all of the files listed together, another option for this section would be splitting the names and descriptions into two sections, one being supporting or informational files like the readme and data dictionary, and the rest being the data itself.

**4 Specialized dataset preview example:** violence against women. This is to show what a download with less files, in this same design, would look like.

**3**

**Download details**

The NIBRS dataset offers detailed information about the victims, offenders, locations, offenses, and property types associated with reported crime incidents. It also shows the relationships between and among these data elements to enable new perspectives of reported crime data.

For more information on how to work with NIBRS please see the resources on the right.

**3A**

STATE West Virginia	What's in the download
DATA TYPE NIBRS	FILES DESCRIPTION
YEARS 2016	README.html Useful documentation about the dataset
LAST MODIFIED January 1, 2016	README.md Useful documentation about the dataset
RELEASE DATE January 1, 2016	agency_participation.csv Agency, county, and state participation for 2016
VIEWS 50	cde_agencies.csv List of agencies and information
DOWNLOADS 20	nibrs_activity_type.csv Type of NIBRS activity related to arrest
LICENSE Open data	nibrs_age.csv Age of ...
FILE SIZE 5.2 MB 32 MB unzipped	nibrs_arrest_type.csv Arrest type....
	nibrs_arrestee.csv Arrestee....
	nibrs_arrestee_weapon.csv Weapon used by arrestee
	nibrs_assignment_type.csv Assignment....
	nibrs_bias_list.csv Full list of biases possible
	nibrs_bias_motivation.csv Bias motivation of offense
	nibrs_circumstances.csv Circumstance of...

**3B**

**Download data**

**4**

**Download details**

This dataset provides information about all female victims of several major offenses as reported to NIBRS by participating agencies in a single year. Each row represents a single victim of one or more major offenses — homicide and voluntary manslaughter, involuntary manslaughter, kidnapping/abduction, robbery, rape, sodomy, sexual assault with an object, fondling, simple assault and aggravated assault — in a criminal incident, and it provides basic demographic information about the victim, weapons used, injuries received and other circumstances of the crime. In addition, each row contains a count of offenders that were family of the victim, known to the victim or otherwise unknown and detailed information about the 3 most familiar offenders to the victim.

**Resources**

- Readme
- Data dictionary
- NIBRS attributes
- NIBRS manual

DATA TYPE	What's in the download
Violent Crime	FILES DESCRIPTION
YEAR 2015	violence_against_women.csv Information about the victim, circumstances, and offender in a reported crime incident against women
LAST MODIFIED January 1, 2016	README.md A document defining related terms
RELEASE DATE January 1, 2016	Data dictionary Open data licensing document
VIEWS 200	DOWNLOADS 150
LICENSE Open data	FILE SIZE 9.1 MB 65 MB unzipped

**Download data**

## CONCEPT DESIGNS: DOWNLOAD PREVIEW

### KEY DESIGN DETAILS

**5** Preview page sections, to give a complete view of the layout of this concept.

**5A** Datasets selector navigation bar.

**5B** Reporting participation data and interactive map (detailed in previous section of presentation).

**5C** Download preview

**5D** List of available datasets and expandable descriptions, from Downloads & Documentation landing page, can slide beneath specific dataset preview content after selection.

**5**

**5A**

**5B**

**5C**

**5D**

**5D**

## KEY QUESTIONS AND CONSIDERATIONS

The idea of the Download Preview – which includes a brief description and other useful information about the file – has broad appeal among the users we spoke with. However, some aspects of this design may not translate to all types of potential downloads. For example, a bulk download is made up of 40+ individual files that depend heavily on each other, so a preview of the first 10 rows of any of them is not particularly helpful. Specialized datasets are usually only a single file, but since many are often over 50 columns, a preview of the first 10 rows will likely not work well for them.

More importantly, while the added context is appreciated by users, it may not be necessary as many users report receiving far less context from other sites they've used. The cost/benefit ratio of implementing this concept is thus unclear, as it doesn't necessarily solve an clear problem and it would require CJIS to tabulate and maintain preview data for all files.

Then again, as CJIS adds more specialized datasets, the Download Preview may serve as a good mechanism for explaining what each is for and when they were generated. And the Download Preview may be the basis for a tool to help users filter datasets as more downloads are added to the page.

## SAMPLE USER STORIES

As a user I want to view the associated methodology for the chart or table I am viewing.

As a user I want to view available annotations for a single year in the time series I am viewing data for.

As a user I want to view and download all footnotes for a time series I am viewing data for.

## RELATED RESOURCES

### INVISION PROTOTYPE

[https://gsa.invisionapp.com/share/U5FJNCORS9Z#/275847520\\_D-D\\_1](https://gsa.invisionapp.com/share/U5FJNCORS9Z#/275847520_D-D_1)

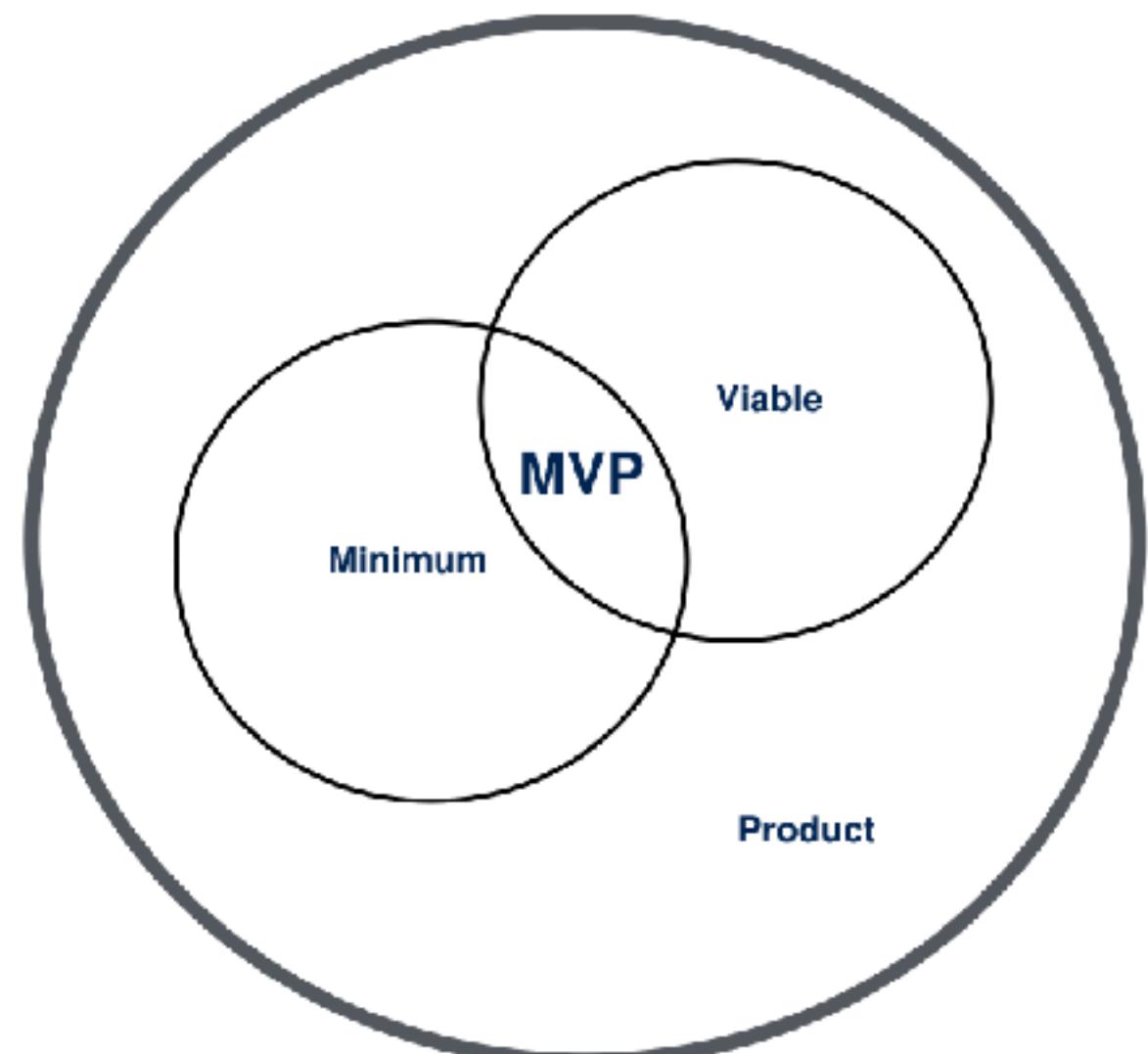
### CDE STYLEGUIDE

<https://github.com/18F/crime-data-explorer/wiki/Style-Guide>

# **APPENDIX**

## WHERE ARE WE TODAY?

**WE DELIVERED AN INITIAL OFFERING, OR MVP,  
THAT IMPROVES ACCESS TO UCR DATA.**



AN MVP IS A START, NOT AN END. IT  
IS MEANT TO VALIDATE DIRECTION  
AND INFORM NEXT STEPS.

Providing multiple pathways to the data via the CDE



National, state, local  
trends. Limited to  
major SRS & NIBRS  
offenses.

Major & minor  
incident data by state  
& year. Some  
specialized datasets,  
like arrests.

Promotes the broadest access  
to the data, but is aimed at  
more technical audiences.  
Lacks capability to query  
individual incidents.

## **DO ONE THING REALLY WELL**

**A GOOD PRODUCT HAS FOCUS.** For the CDE that means providing broad access to UCR data; it is not a dedicated tool for data analysis or visualization.

## **MEET USERS WHERE THEY ARE**

**OUR USERS HAVE A RANGE OF NEEDS & PREFERENCES.** To serve these different groups we provide multiple pathways to the data, such as a web interface, downloads, and an API, rather than attempt a “one-size fits all” approach.

## **BE INTERESTING, BUT RESPONSIBLE**

**ORIENT & GUIDE PEOPLE TO THE DATA, BUT LET THEM DRAW THEIR OWN CONCLUSIONS.** A persistent challenge for us was striking a balance between presenting information in a way that was engaging to the public, while staying true to the FBI's reporting standards.

## **PROVIDE A CONSISTENT EXPERIENCE**

**USERS DON'T UNDERSTAND THE NUANCES OF THE UCR PROGRAM, LIKE THE DIFFERENCE BETWEEN SRS & NIBRS.** We sought to minimize these differences by centering the web experience around a common set of crime types and maintaining a similar look and feel regardless of the data type being displayed.

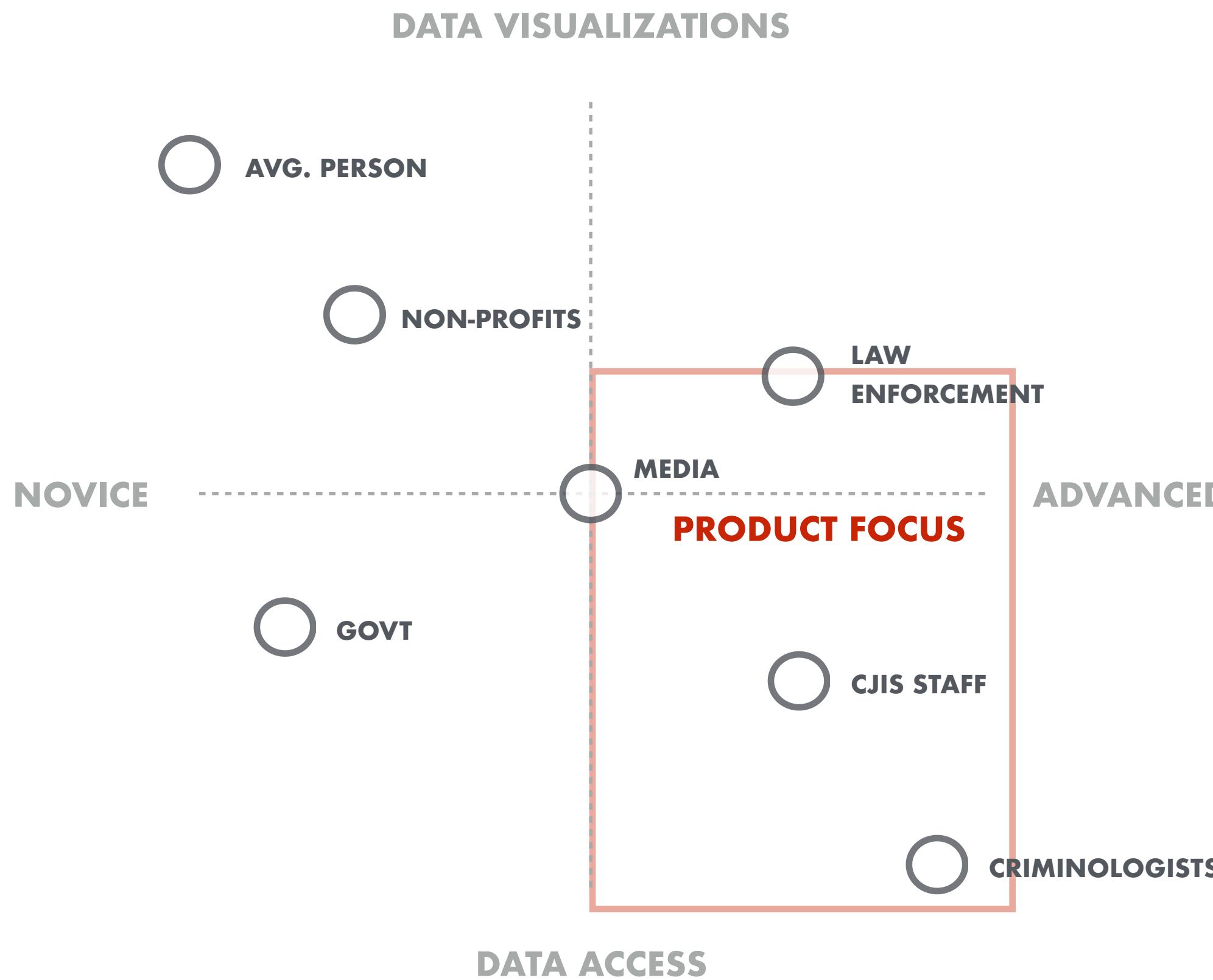
## **INSPIRE TRUST**

**THE CDE IS AN EXTENSION OF THE FBI BRAND.** We sought to build trust in the CDE by explaining how the data should be used and what its limitations are. Similarly, we prioritized usability, performance, and data integrity over expanded functionality.

## SETTING FOCUS

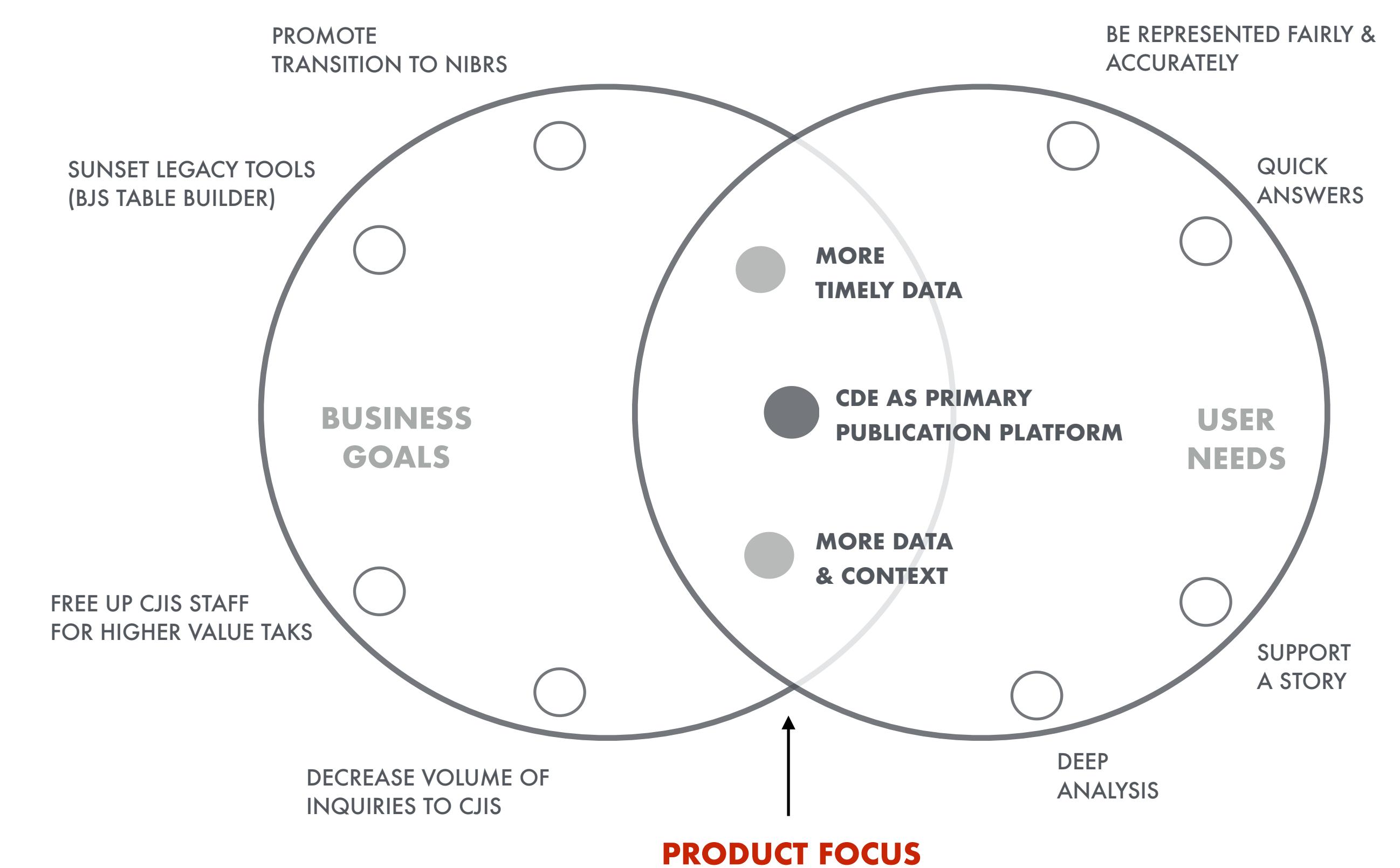
### TARGET USERS

HOW USERS DIFFER IN BACKGROUND & PREFERENCES

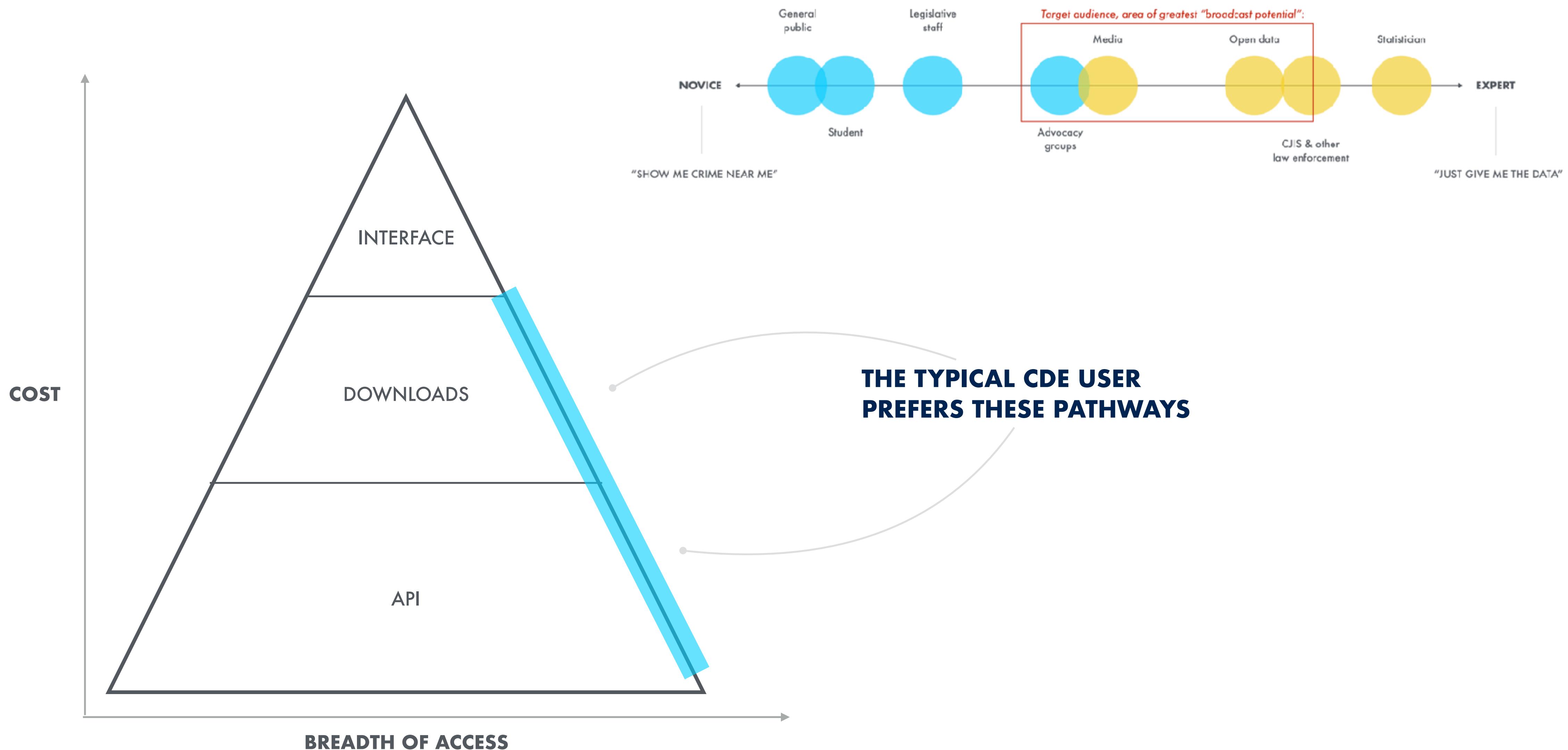


### THEMES & CAPABILITIES

HOW BUSINESS GOALS & USER NEEDS INTERSECT

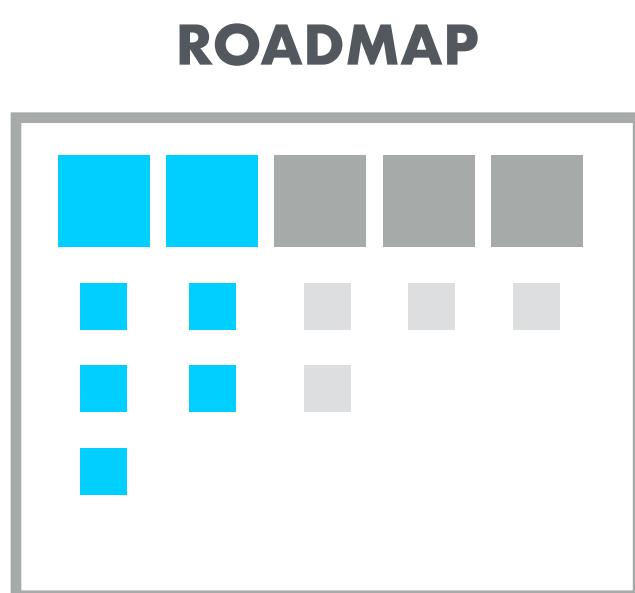


## WHY AVOID CHANGING THE INTERFACE?



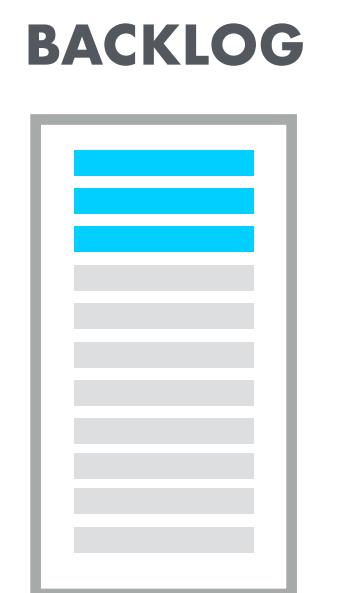
## GOAL FOR ROADMAP EXERCISE

**SET A TRAJECTORY FOR THE CDE THAT GUIDES NEXT STEPS. THE ROADMAP SHOULD BE REVISITED EVERY 3-6 MONTHS.**



High-level plan that promotes alignment around product direction

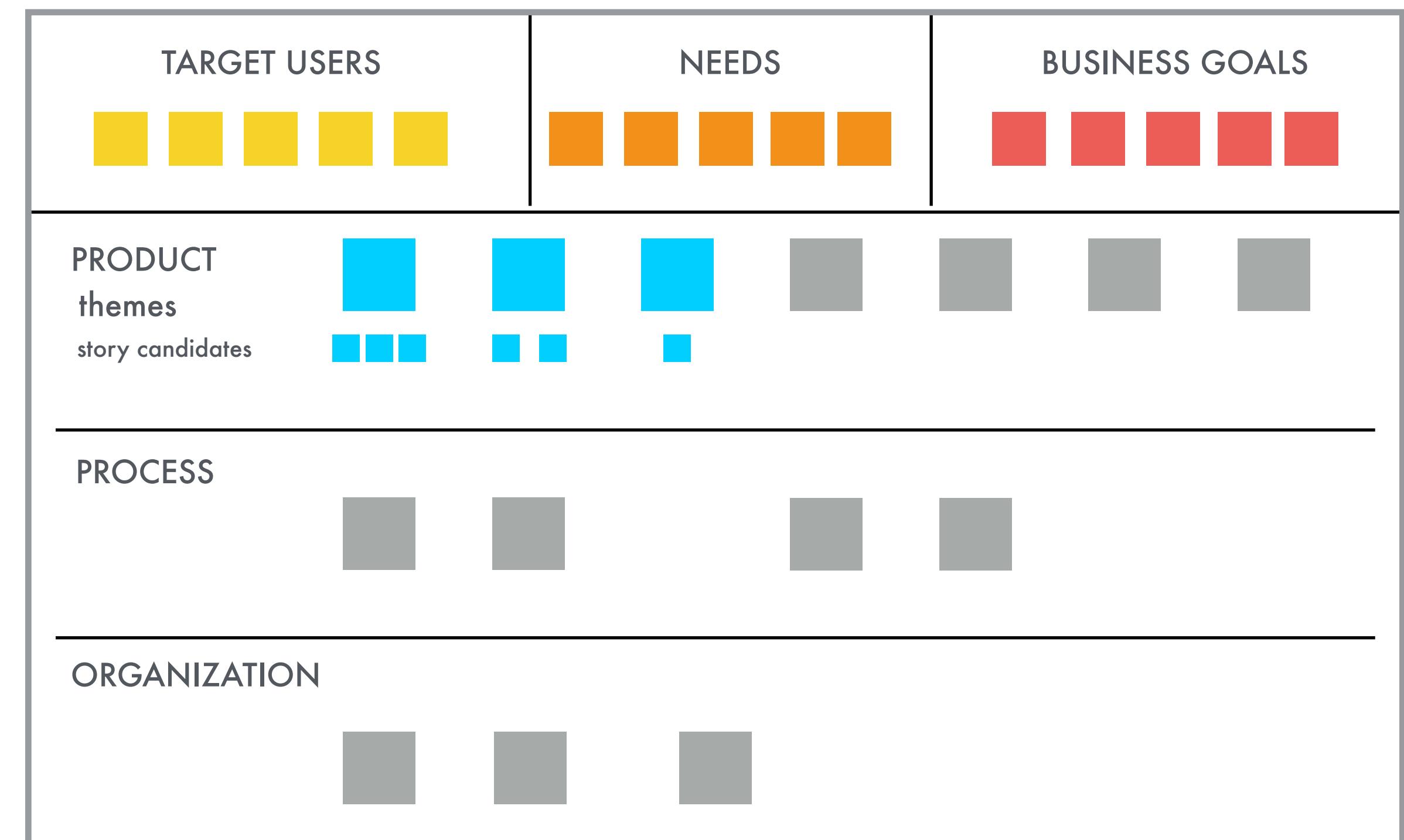
A canvas for big picture planning; focused on broad themes to pursue, not specific features



A tactical view of the work to be done based on the roadmap

The backlog is constantly changing.

**PRODUCT VISION:** the vision for the CDE explained in 1-2 sentences



## PRODUCT ROADMAP - V1

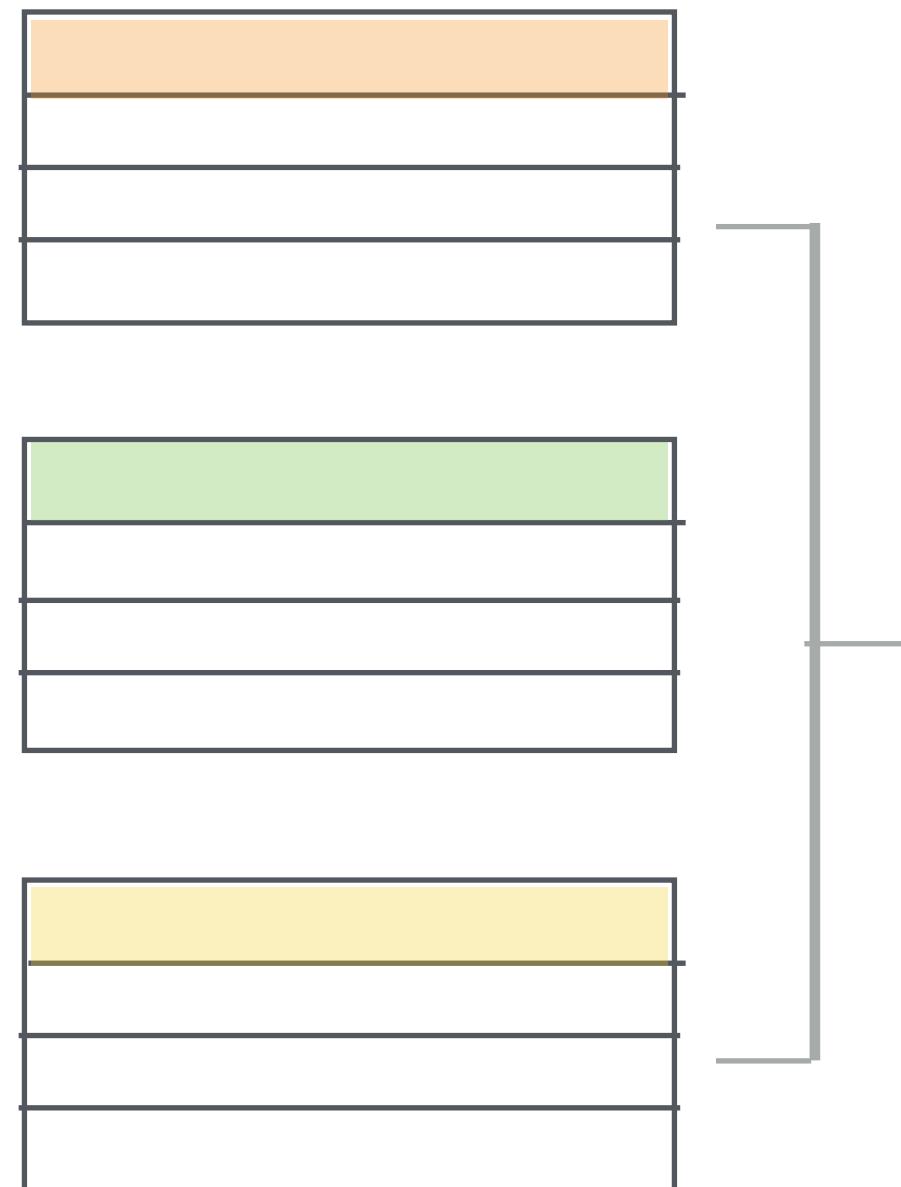
PRODUCT THEMES	NOW (2-4 MONTHS)	SOON (3-6 MONTHS)	LATER (6+ MONTHS)
	<ul style="list-style-type: none"> <li>● MORE DATA</li> <li>● ENABLE A DIGITAL ORGANIZATION</li> </ul>	<ul style="list-style-type: none"> <li>● IMPROVED CONTEXT &amp; CAVEATS</li> <li>● MORE ROBUST ANALYTICAL CAPABILITIES</li> </ul>	<ul style="list-style-type: none"> <li>● IMPROVED GRANULARITY</li> <li>● MORE TIMELY DATA</li> </ul>
STORY CANDIDATES	<ul style="list-style-type: none"> <li>● Add incident endpoint (expand API)</li> <li>● Add dimensions for existing NIBRS offenses</li> <li>● Add new NIBRS offenses to Explorer</li> <li>● Add SHR dataset for download</li> </ul>	<ul style="list-style-type: none"> <li>● Enable dynamic flagging/footnotes for the data</li> <li>● Enable trend-based comparisons between states</li> <li>● Add demographics to complement crime data</li> <li>● Add a BJS-like table builder for custom queries</li> </ul>	<ul style="list-style-type: none"> <li>● Enable new geographic perspectives of the data, such cities, counties, and regions.</li> <li>● Allow users to flag potential issues with the data for correction</li> </ul>
PROCESS	<ul style="list-style-type: none"> <li>● Collect feedback on which datasets to add</li> <li>● Consider prioritizing access to the data via</li> <li>● the API &amp; downloads over changes to the interface that require more UI/UX work</li> </ul>	<ul style="list-style-type: none"> <li>● Conduct research to determine if a custom query tool similar to the BJS table-builder is needed, or if pre-generated downloads will suffice</li> </ul>	<ul style="list-style-type: none"> <li>● Decide how to represent non-standard reporting areas, like cities, responsibly</li> <li>● Explore opportunities for tighter integration between CDE and New UCR to facilitate more regular updates</li> </ul>
ORGANIZATION	<ul style="list-style-type: none"> <li>● Establish a cross-functional digital team that is fully dedicated to the CDE. Hire UI/UX help</li> <li>● Seek out agile &amp; product owner training</li> <li>● Start building a culture that embraces more widespread use of agile &amp; open source</li> </ul>	<ul style="list-style-type: none"> <li>● Work with the publication team to develop a mechanism and process for mapping annotations with specific agencies &amp; years</li> </ul>	<ul style="list-style-type: none"> <li>● Consider collecting more GIS information from LEA's to improve the granularity of the data available.</li> <li>● Explore reforms to the UCR program that improve the volume and quality of the data available to the CDE.</li> </ul>

## NORMALIZED VS DENORMALIZED DATASETS: FINDING THE RIGHT WAY TO PRESENT THE DATA

An attempt to make NIBRS data easier to work with (denormalization) led to an oversimplification of the relationships between elements, specifically offender-victim counts.

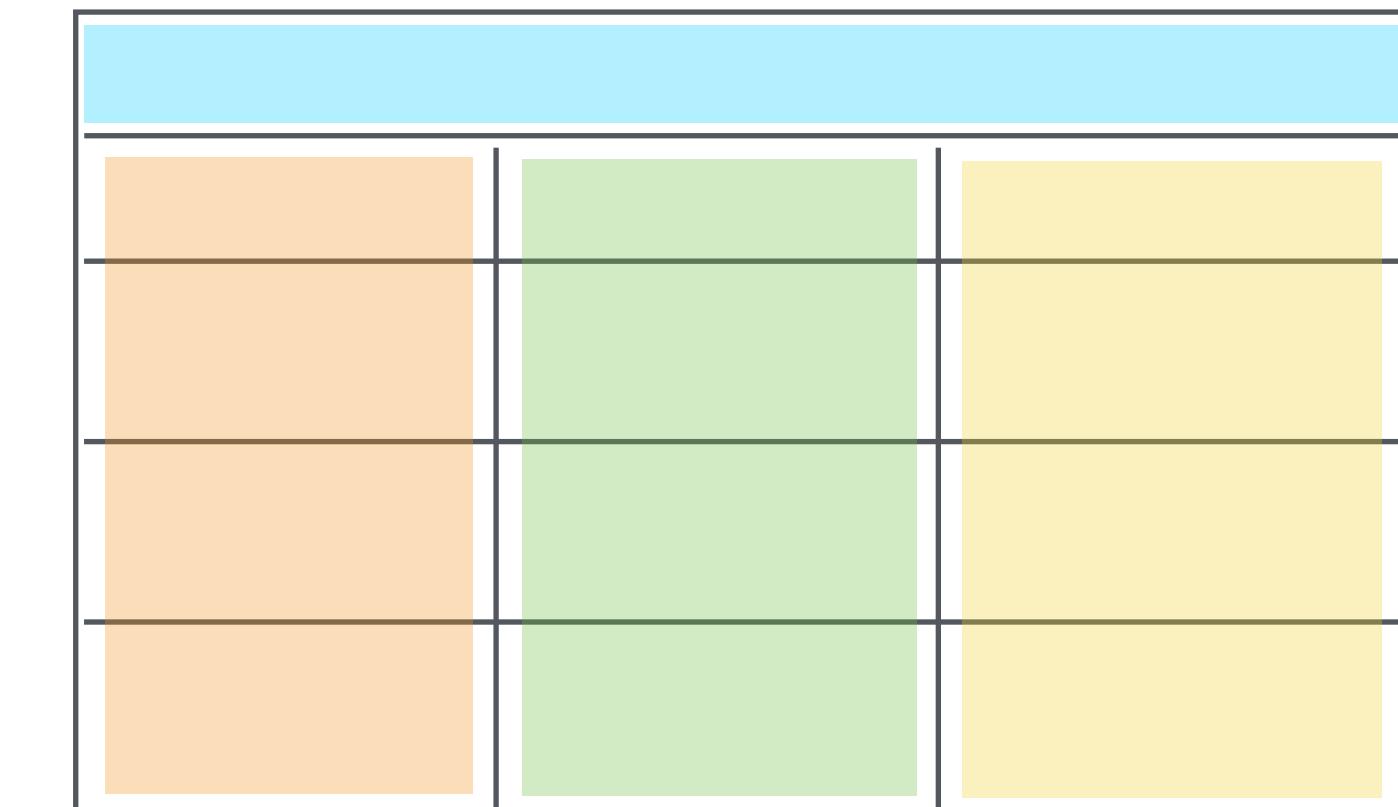
We are revising this model for improved accuracy and are considering the best denormalization strategy going forward. As we think about the best way of “packaging” NIBRS data as a download, we should consider the tradeoffs between greater flexibility (more tables) and ease of use.

**NORMALIZATION** IS A WAY OF ORGANIZING DATA THAT INVOLVES BREAKING COMPLEX TABLES INTO SMALLER ONES



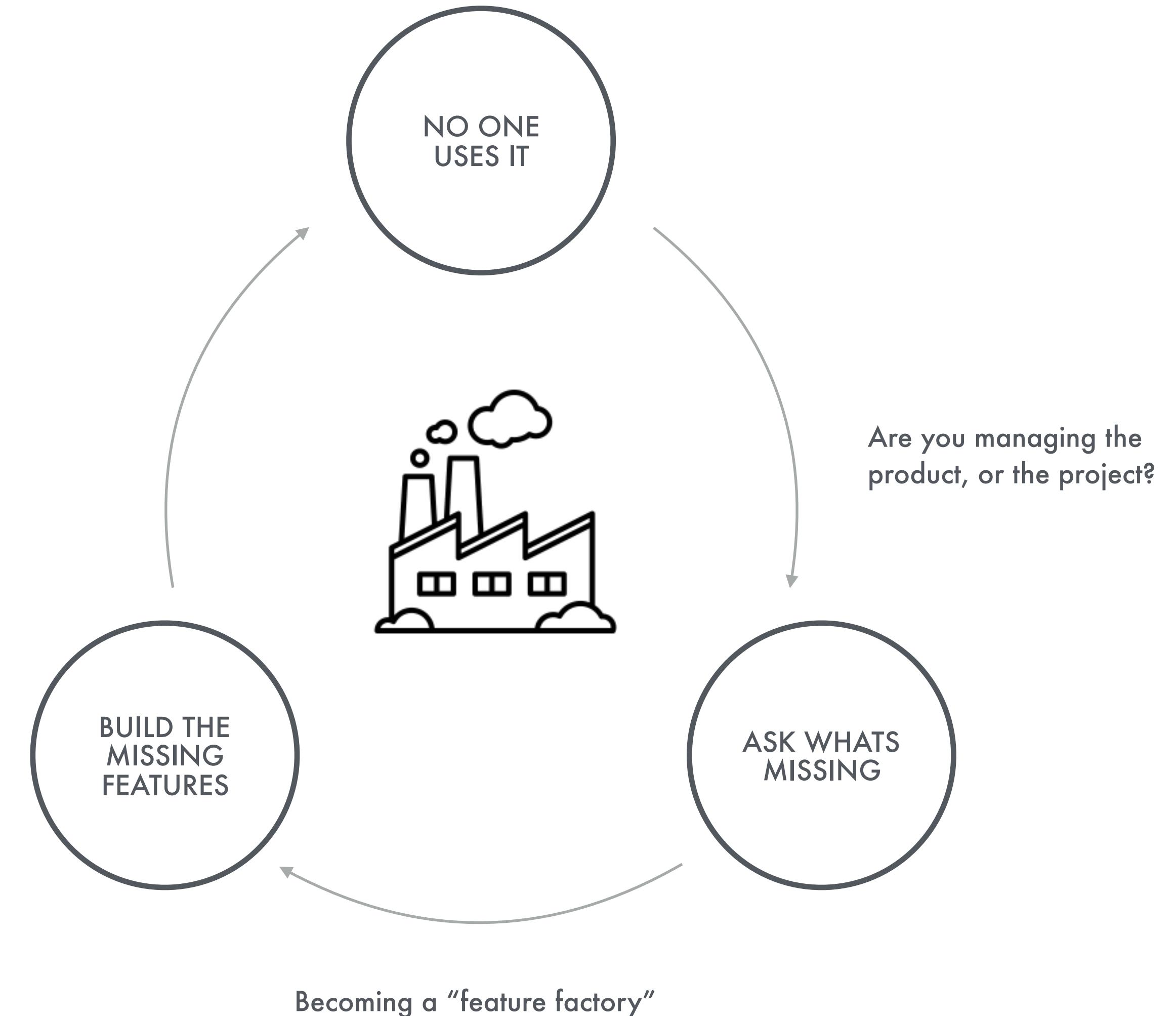
IT REDUCES THE SIZE OF A DATABASE, BUT REQUIRES MORE “TABLE JOINS” THAT RESULT IN LESS EFFICIENT QUERIES

**DENORMALIZATION** REDUCES THE “COST” OF MAKING CERTAIN QUERIES BY CREATING SPECIALIZED GROUPINGS OF DATA



IT CAN LEAD TO SOME DUPLICATION OF THE DATA, BUT RESULTS IN FEWER TABLES TO WORK WITH, WHICH IS HELPFUL AS A DOWNLOAD

# **“TACTICAL HELL” - A PLACE OF ENDLESS REACTING & REACTION**



## LESSONS LEARNED

**UCR DATA WILL NEVER BE  
PERFECT.**

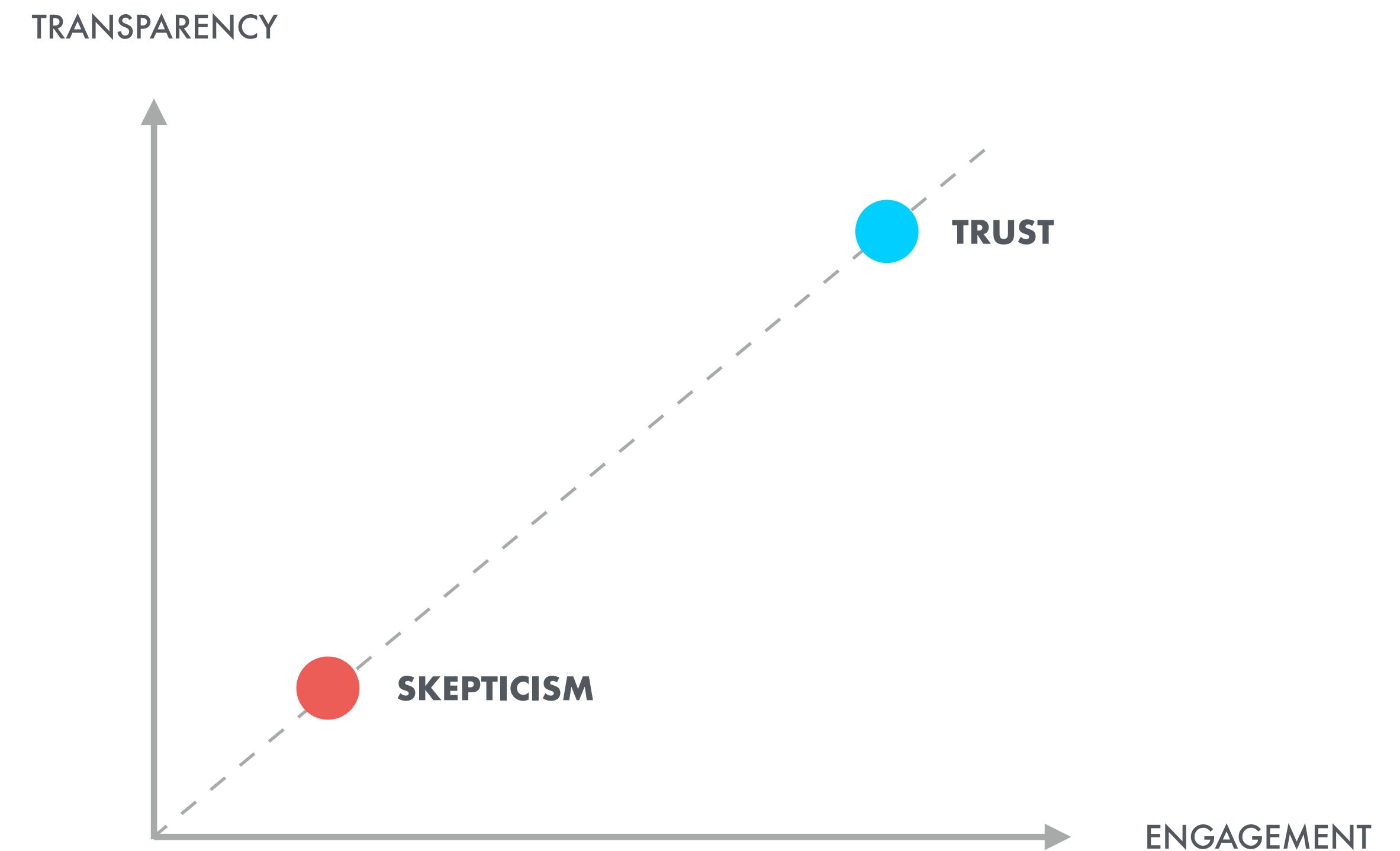
**BEING MORE TRANSPARENT  
AND WORKING IN THE  
OPEN CAN MITIGATE  
CRITICISM ABOUT WHAT IS  
PUBLISHED.**

DEFINE A PROCESS FOR VALIDATING CDE DATA & FUTURE UPDATES

EMBRACE OPEN SOURCE STANDARDS IN USE ACROSS GOVERNMENT & KEEP  
THE CDE PROJECT REPO OPEN IN ACCORDANCE WITH OMB M-16-21

CREATE & MAINTAIN FEEDBACK LOOPS WITH USERS TO FLAG POTENTIAL  
DATA QUALITY ISSUES & TO INFORM THE DIRECTION OF THE PRODUCT

ENSURE THAT THE METHODOLOGIES FOR TABULATING CDE DATA ARE EASILY  
UNDERSTOOD BY THE PUBLIC & CONSISTENTLY APPLIED.



18F