

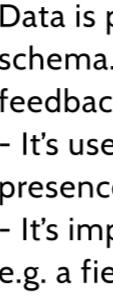
Department for Transport (DfT)

Discovery/Alpha into adding accessibility data to the national database of public transport stops.

The challenge

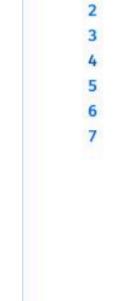
Context

 **NaPTAN stores where public transport stops are located**
A national database of bus stops, train stations, ferry terminals, etc. allowing transport operators to plan multi-operator and multi-modal journeys.

 **It's an XML file**
NaPTAN has existed since the late 90's. It's a big XML file where most local authorities (LA's) submit all their public transport stops, in a schema version (2.1) from the early 00's.

 **Data 'producers' and 'consumers'**
Our data 'producers' are local authorities (LA's), who currently upload this NaPTAN data via a spreadsheet upload. Our data 'consumers' download it and use it in their services, e.g. journey planning services like Citymapper.

The Ask

 **Add accessibility data to NaPTAN**
Discovery and Alpha (in line with GDS standards) into incorporating accessibility data for every public transport stop into NaPTAN.

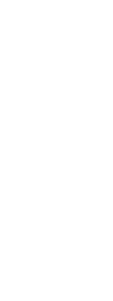
 **Understand the data**
Learn what data local authorities (LAs) have and can provide to NaPTAN. Learn what data consumers would need to show accessible journeys to passengers

 **Develop a proof of concept in 10 weeks**
Show that we can serve accessibility data alongside NaPTAN data to data consumers. We had only 10 weeks to deliver this.

What we built + how research influenced the solution

 **Ingest data**
Consume spreadsheets

 **Through UR we learned:**
- The data's almost always going to come as a .csv file.
- LA's don't have the resources to transform the data for us, so we have to accept the spreadsheets as they are.
- They will often not be directly linked to NaPTAN stop identifier (ATCO Code). It may take some work to link.

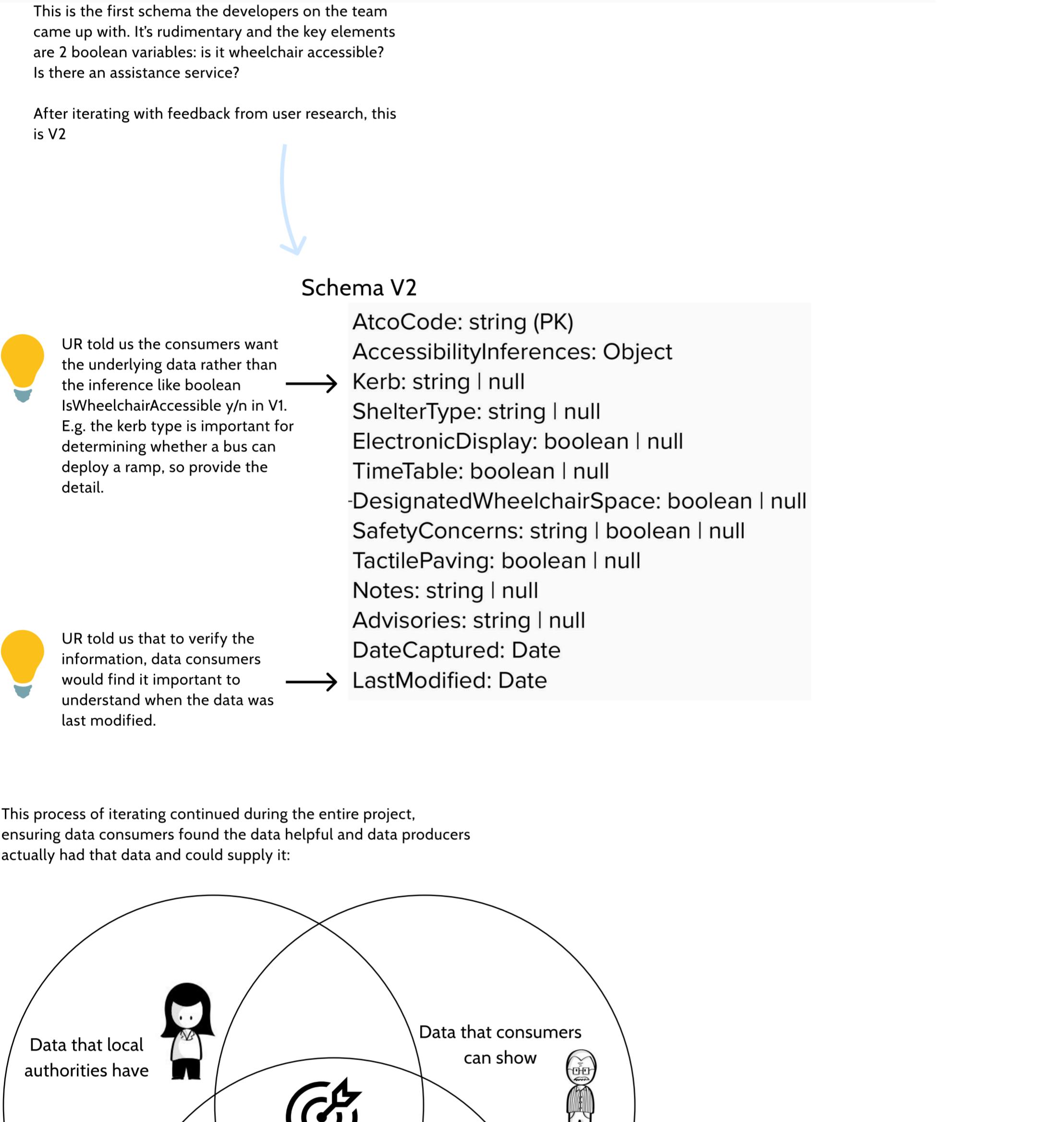
 **Process data**
Put it into BigQuery

 **Data is processed from spreadsheets into a unified schema.** That schema was iterated on using feedback from UR, for example we learned:
- It's useful to include kerb height-type, the presence of tactile paving.
- It's important to be able to verify the information e.g. a field to show when it was last updated.

 **Serve data**
Give it to consumers

 **UR told us:**
- We should have an option for consumers to bulk download the data (REST API), as well as an option to download data for individual stops to see the data they need (GraphQL).

An example of 2 queries made in the GraphQL interface to our data, to get the accessibility data between 2 dates. The second one only shows stops that are wheelchair accessible.



My approach

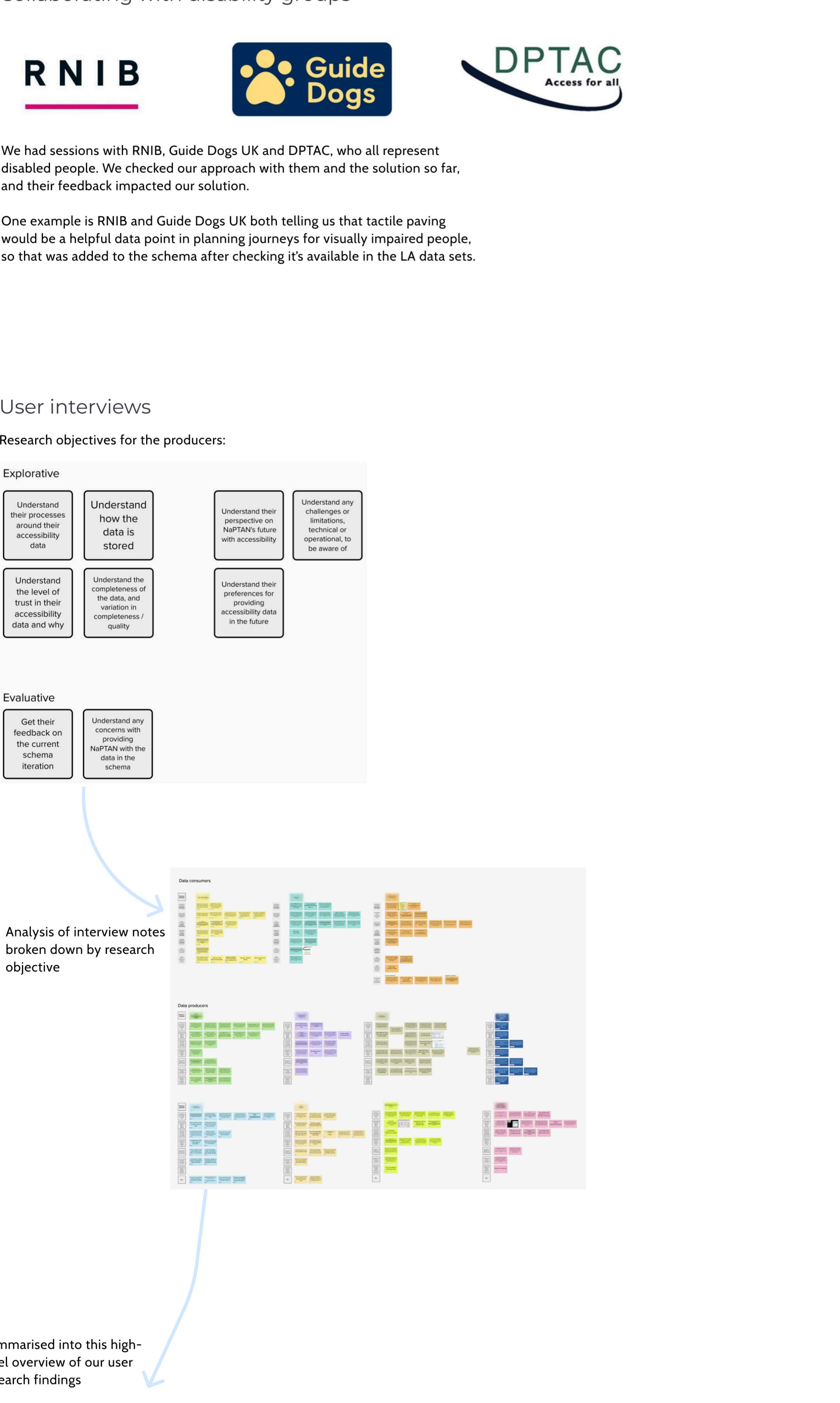
The focus

 I focused on data producers and data consumers
NaPTAN is essentially the platform that provides the data. It doesn't have control over the experience for passengers. Our primary users are data producers and consumers.

Travel planning apps have the interface to passengers. To create and test an end-to-end solution for passengers requires close collaboration with them, which I recommended as a focus for the Private Beta phase, after this discovery/alpha.

For this service to work, it needs:
1. To be able to understand the data from many different local authorities across the UK (data producers)
2. To be consumed and valuable to journey planners that provide the interface to passengers (data consumers)
3. To be valuable to the end users: passengers with accessibility needs, and provide them with agency and autonomy in making public transport journeys, but they're not the primary users of the service.

The process



Iterating on the schema with UR findings

Schema V1

AtcoCode: string (PK),
StopCommonName: string
IsWheelchairAccessible: boolean | null,
HasAssistanceService: boolean | null,
DateCaptured: Date > Date when we last updated the raw dataset (base table)

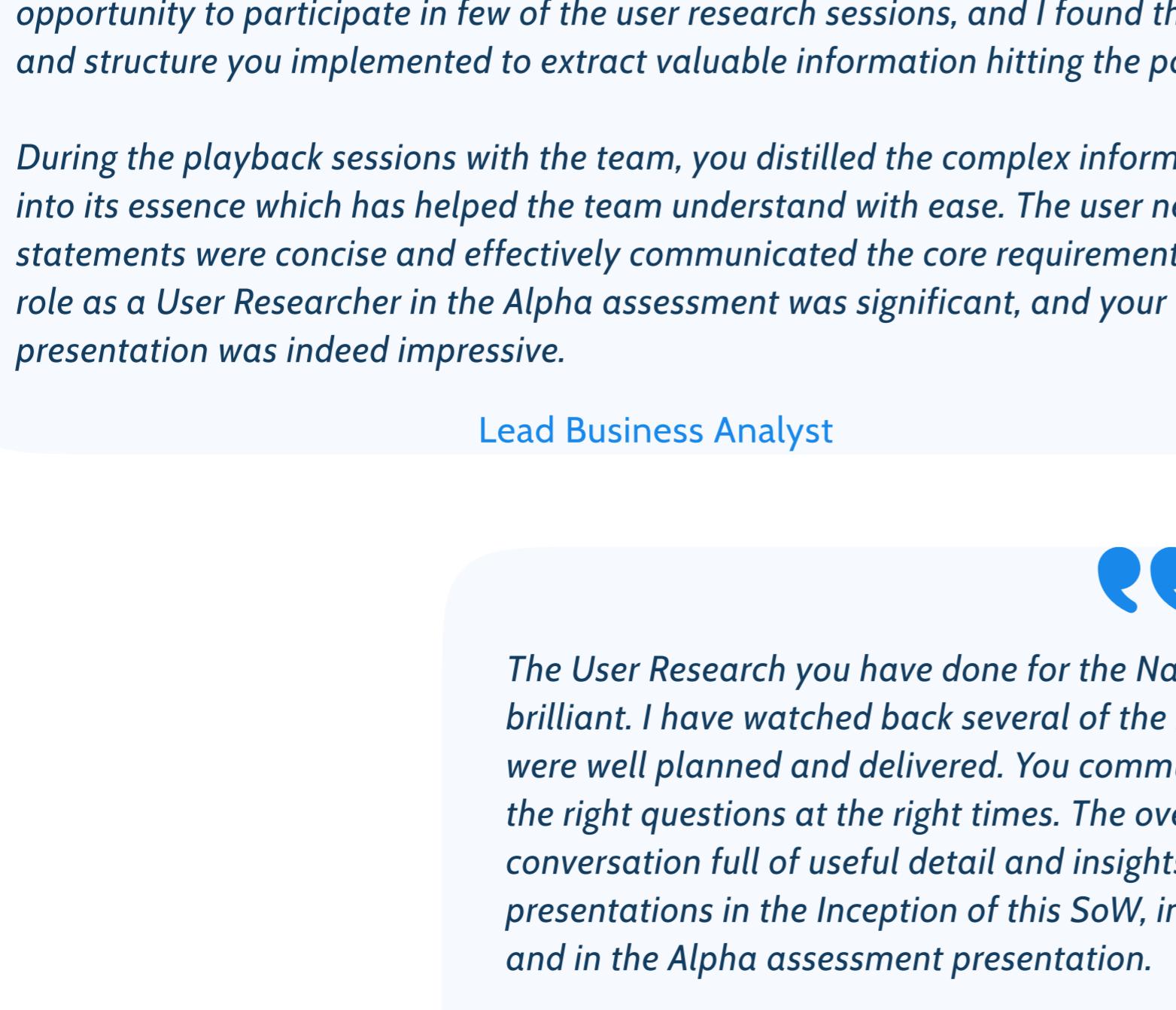
This is the first schema the developers on the team came up with. It's rudimentary and the key elements are 2 boolean variables: is it wheelchair accessible? Is there an assistance service?

After iterating with feedback from user research, this is V2

Schema V2

AtcoCode: string (PK)
AccessibilityInferences: Object
Kerb: string | null
ShelterType: string | null
ElectronicDisplay: boolean | null
TimeTable: boolean | null
DesignatedWheelchairSpace: boolean | null
SafetyConcerns: string | boolean | null
TactilePaving: boolean | null
Notes: string | null
Advisories: string | null
DateCaptured: Date
LastModified: Date

This process of iterating continued during the entire project, ensuring data consumers found the data helpful and data producers actually had that data and could supply it:



User research

The task for this user: plan and execute an accessible journey to Greenwich Observatory with public transport.

The route involved multiple modes: a bus, a London Underground ride and a ferry (Uber Boat).

The purpose was to understand first-hand the experience of a powered wheelchair user in planning and executing a journey using public transport.

We observed many issues that we had read about previously, and this was an incredibly valuable experience in developing empathy and first-hand understanding for the end users of service we are enabling.

We observed many aspects of the journey that could have been improved by information provided e.g. in a journey planning app: some examples:

- Is there assistance at the station/platforms involved in the journey?
- Is the bus stop involved accessible for my wheelchair? How steep are the ramps I'll be taking?
- Are there accessible toilets along the way?
- Is the wheelchair space occupied free?

Some of the things observed were outside NaPTAN's remit, and others were within its remit.

Collaborating with disability groups



We had sessions with RNIB, Guide Dogs UK and DPTAC, who all represent disabled people. We checked our approach with them and the solution so far, and their feedback impacted our solution.

One example is RNIB and Guide Dogs UK both telling us that tactile paving would be a helpful data point in planning journeys for visually impaired people, so that was added to the schema after checking it's available in the LA data sets.

I did the same for data consumers:

 Understand their processes for publishing accessibility data
Understand how the data is stored
Understand their perspective on NaPTAN future with accessibility
Understand any challenges or requirements for publishing accessibility data in the future
Understand their preferences for publishing accessibility data in the future

 Get their feedback on the proposed schema iteration
Understand any challenges or requirements for publishing accessibility data in the future

Analysis of interview notes broken down by research objective

Both explorative research to learn and go wide, as well as evaluative research to get feedback on our solution.

Summarised into this high-level overview of our user research findings

Stages

High level research findings

User needs statements

Key insights

Pain points / constraints

To explore

Get data

Store data

Export and send data to NaPTAN

Pre-launch

Post-launch

Feedback

Lead Business Analyst

Evaluative

Exploratory

Analysis of interview notes broken down by research objective

Both explorative research to learn and go wide, as well as evaluative research to get feedback on our solution.

Summarised into this high-level overview of our user research findings

Stages

High level research findings

User needs statements

Key insights

Pain points / constraints

To explore

Get data

Store data

Export and send data to NaPTAN

Pre-launch

Post-launch

Feedback

Lead Business Analyst

Evaluative

Exploratory

Analysis of interview notes broken down by research objective

Both explorative research to learn and go wide, as well as evaluative research to get feedback on our solution.

Summarised into this high-level overview of our user research findings

Stages

High level research findings

User needs statements

Key insights

Pain points / constraints

To explore

Get data

Store data

Export and send data to NaPTAN

Pre-launch

Post-launch

Feedback

Lead Business Analyst

Evaluative

Exploratory

Analysis of interview notes broken down by research objective

Both explorative research to learn and go wide, as well as evaluative research to get feedback on our solution.

Summarised into this high-level overview of our user research findings

Stages

High level research findings

User needs statements

Key insights

Pain points / constraints

To explore

Get data

Store data

Export and send data to NaPTAN

Pre-launch

Post-launch

Feedback

Lead Business Analyst

Evaluative

Exploratory

Analysis of interview notes broken down by research objective

Both explorative research to learn and go wide, as well as evaluative research to get feedback on our solution.

Summarised into this high-level overview of our user research findings

Stages

High level research findings

User needs statements

Key insights