

Jimmy Tidey, PhD

DESIGN RESEARCH PORTFOLIO

Project 1

AI Assistant for a central government department

Project 2

Google-internal team management tools

Project 3

PhD research



Hello, I'm Jimmy.

I bring technical knowledge and academic rigour to design.

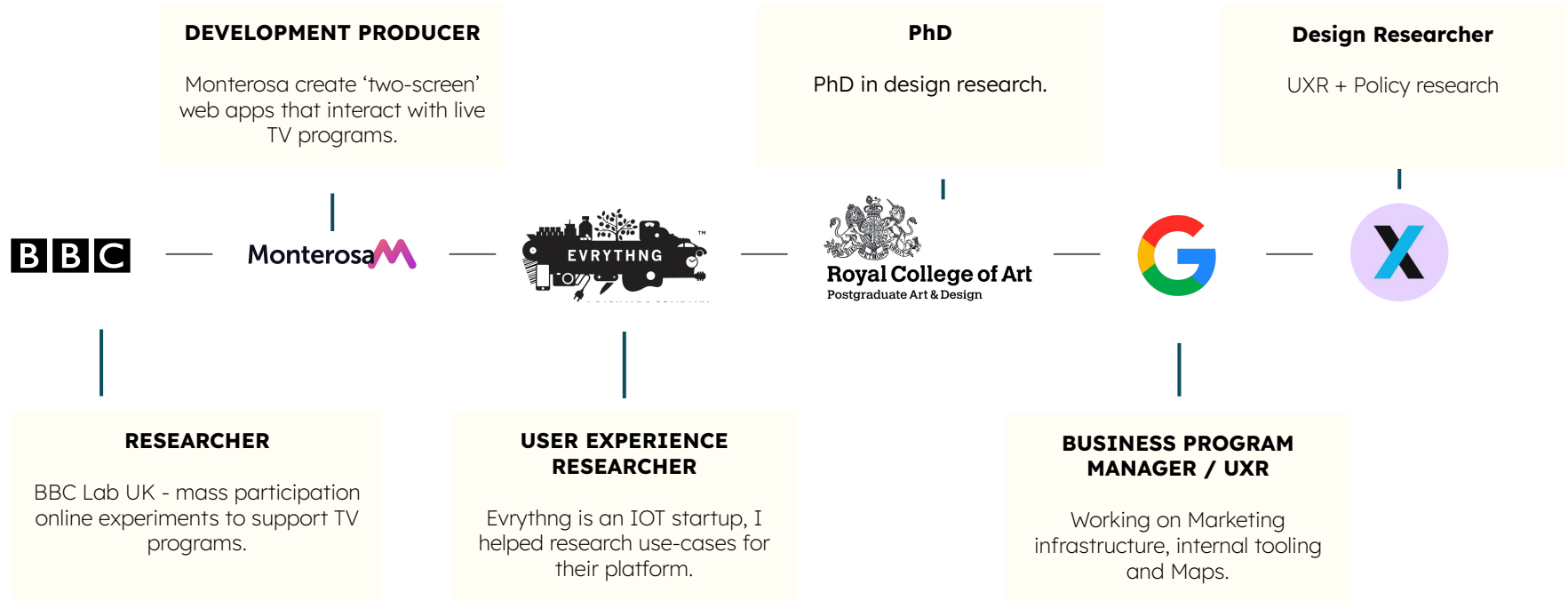
My approach draws on ten years of industry experience and a design research PhD.

My focus is on generative AI.

I love working at the intersection of cultures and specialisms; my diverse background has helped me to make unique contributions in settings ranging from academia to startups to working at Google.

Employment history

(Summary)



Portfolio summary

I've included a lot of detail in my portfolio. To make it easier to navigate, I've provided direct links to projects (underlined) and a list of the research methods demonstrated in the project.

PROJECT 1

Central Gov AI Assistant

My most recent work at TPXimpact.

Methods

- AI Assistant evaluation
- LLM prototypes
- Content review
- In-depth user interviews

PROJECT 3

Team management tool

A detailed description of the work I did as UXR responding to a feature request.

Methods

- In-depth user interviews
- Concept testing
- Data analysis

PROJECT 3

PhD Research

Building and evaluating a social media analytics tool.

Methods

- Physical prototypes
- Data / network analysis
- Focus groups
- Sociometric surveys

PROJECT 1

HR AI Assistant

(for a central government department)

My Role: Leading design research

Team: Interaction designer, tech lead, data scientist, tech lead and account manager.

Duration: 3 months (ongoing)

Outcome: Research approach identified, concept tested and validated

Context

Staff in 'digital, data and technology' roles in a UK central government department receive bonus payments to ensure their pay is competitive.

Staff are often confused by how their bonus is calculated and how it interacts with other HR policies, such as parental leave.

Problem statement: HR staff spend over 40 hours per week replying to emails about the 'digital, data and technology' bonus policy.

Proposed solution: Using a RAG-based assistant to provide first-line responses using Microsoft Azure's Language Studio and AI tooling.

Gov Department

AI HR Assistant

HR Assistant

Morning! You can use this service to ask questions about pay, performance and people policies. You can also provide feedback and make suggestions for improving our organisation. Everything you do here is anonymous.

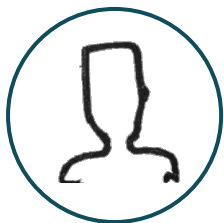
What do you want to ask?

Do not disclose any personal or confidential information as we keep a record of questions asked for the purposes of continuous improvement.

Send

Prototype UX

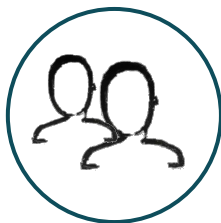
Team & stakeholders



Me



Interaction Designer



Data scientist &
Tech Lead



Account manager



*Government
department*

Research structure

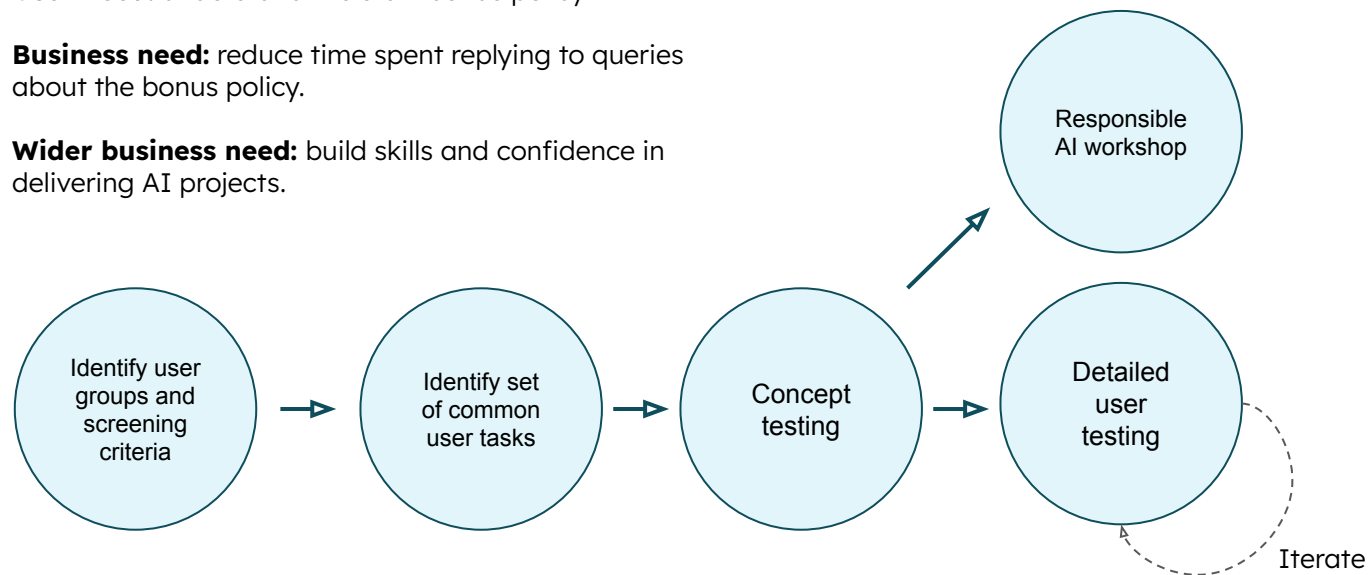
The design process was structured in four phases. It also included a Responsible AI workshop to help address government guidelines on deploying AI systems.

The goal was to ensure a proposed AI assistant (based on a RAG model) met user and business needs.

User need: understand the staff bonus policy

Business need: reduce time spent replying to queries about the bonus policy.

Wider business need: build skills and confidence in delivering AI projects.



User groups & recruitment criteria

User groups were identified through a content review of the HR bonus policy documents we anticipated ingesting into the RAG pipeline for use in the assistant. Building the RAG pipeline based on this material was a fundamental design constraint. Three user groups were identified: *individual contributor*, *line manager*, and *head of practice*.

User testing recruitment criteria - recruitment criteria were identified as an even mix of users from each user group. We identified a goal of recruiting a range of tenures at the department and tenures working for the UK government more widely. Demographic factors were considered but rejected as recruitment criteria.

Selecting common tasks

Common tasks were identified in two phases:

- 1) **Collecting** - a set of common user tasks (eg. 'find out if bonus payments continue during parental leave') were by collected by analysing email and slack support channels for the most frequently mentioned topics.
- 2) **Filtering** - asking the support team to identify which of those tasks could feasibly be addressed by the AI assistant.
- 3) **Classifying tasks** - Academic research suggests different tasks would lead to different questions entered into the AI Assistant, that may in turn need different UX approaches. Examples of question types include 'list', 'factoid', 'hypothetical'. This typology is suggested by [Mishra & Jain](#).



Concept testing

We used a '[wizard of oz](#)' approach to concept testing.

Participants were presented with a prototype that gave preset answers when it detected keywords in the user's question.

Participants were then prompted to use the interface to perform four tasks based on the common tasks identified in the previous phase. (eg. 'Find out if bonus payments continue during parental leave.')

Justice Digital

AI HR Assistant

HR Assistant

Different staff may have different leave allowances so ensure you check. Across the Ministry of Justice people work on different terms and conditions, so please ensure you refer to the relevant information. For colleagues who work reduced or compressed hours you will need to calculate their leave allowances and manually calculate bank holiday entitlement.

Use annual leave form

► [Documents used to generate this response](#)

Related questions

[How do I book annual leave?](#)

[What is my sick leave entitlement?](#)

[How do I tell my manager about expected paternity leave?](#)

What do you want to ask?

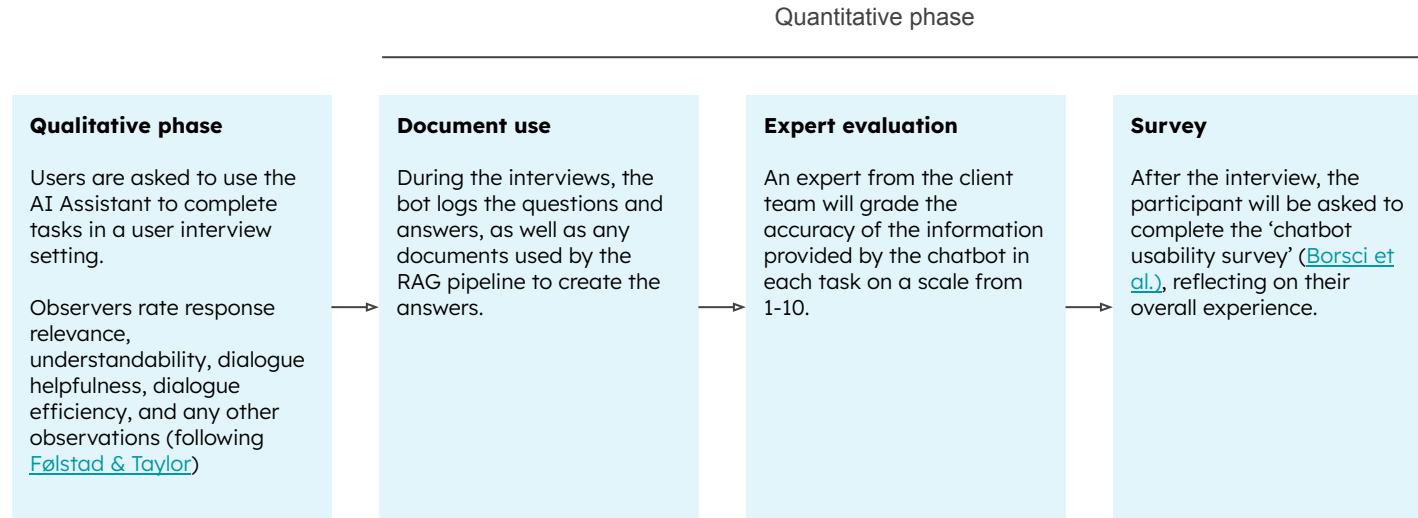
Do not disclose any personal or confidential information as we keep a record of questions asked for the purposes of continuous improvement.

Send

Outcomes

- Users were very positive about the AI Assistant, even when they had previously had unsatisfactory interactions with 'chatbots'.
- Users had diverse expectations of the level of privacy they could expect and often did not consider that chat logs may be kept.
- Some users did not understand that answers were 'generated' in response to their question, ie. they thought it functioned as a search tool and expected answers they saw verbatim from the HR documentation.
- None of the users realised they were using a 'Wizard of Oz' mockup, instead believing they were using a fully functional AI assistant.

Detailed user testing



Outcomes

- A detailed, structured approach to evaluating the AI Assistant has given the client confidence to rely on our evaluation.
- Drawing on existing academic research brings rigour to our work and saves 'reinventing the wheel'.
- The quantitative aspect of the evaluation will allow us to iterate on the AI Assistant and be confident that our work is increasing the quality of the Assistant's responses.

Responsible AI workshop

The 'Responsible AI' implications of an internal-only AI Assistant were considered to be limited.

However, the client wanted to build skills and confidence in delivering AI projects, and wanted to upskill staff in Responsible AI.

To build Responsible AI skills, and give participants a wider range of Responsible AI considerations, the workshop focused on a speculative project to create an external-facing AI assistant that helped candidates applying for jobs at the government department.

I delivered the workshop to 18 participants, including an introduction from a data science professor and two hours of group work.

The workshop adapted the [Ethical OS toolkit](#), and drew on the [Generative AI Framework](#) published by the UK Government.

Outcomes

- Three clear anxieties were observed in the workshop: 'AI will take my Job', 'Our department is behind the curve on AI', and 'AI is all hype'. These are not directly 'Responsible AI' considerations, but we felt that these anxieties need to be addressed in future workshops to get the most out of participants.
- The UK central government has very strong commitments to ensuring that job candidates receive fair treatment. Some participants worried that an AI might not give fair or impartial advice to candidates, while others saw the advantages of an auditable, automated approach.
- We observed that Responsible AI workshops rely on a base-level understanding of how AI can be applied, something that we could spend more time developing in future workshops.

PROJECT 2

Google Team Management tool

My Role: Leading the project as a User Experience Researcher

Team: Support from a UX manager, working with a UXR and engineering team.

Duration: 8 months (40% commitment)

Outcome: Internal team management tool upgraded to better support ~3000 engineers

Context

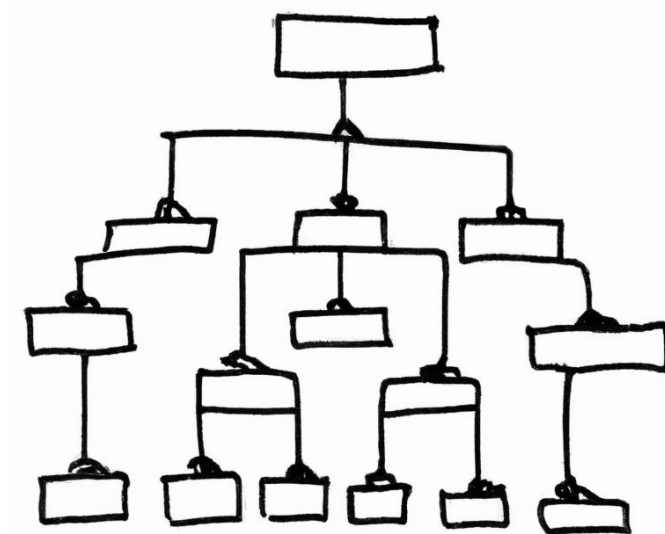
Google has an internal tool for managing teams. Teams are represented as a tree, where every team has a single 'parent' team.

A VP of engineering made a feature request for the tool to allow teams to have two parents to reflect the way teams work.

The request implied a major change to the UX and underlying data model.

Key challenge

Taking a tightly-specified but hard-to-implement feature request from a senior stakeholder, using UXR methods to find a deliverable, usable solution that meets the requester's original intentions.



Some details, including the original assets, are sensitive and cannot be included.

What I did

Working with the engineering and UXR teams, I led the project, including:

- Understanding the context of the ‘two-parent’ feature request.
- Running a scoping phase to frame the problem in terms of user groups and their user needs; building a mockup of how the ‘two-parent’ solution could work.
- Evaluating the ‘two-parent’ solution through in-depth user interviews.
- Building a case to address the ‘two parent’ feature request with a combination of existing features and incremental improvements.

Team & stakeholders



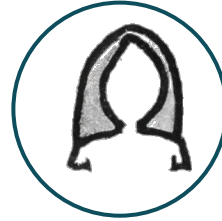
“The client” - a VP of an engineering department



Me



UX manager



UXR advising on integration into existing UX programs



Engineering team

Research structure

At the outset of the project, I ran an extensive scoping phase to define a research question based on specific user needs, deliberately moving away from the implementation details of the initial feature request. We then ran in-depth interviews to validate the user needs and concept-test the original feature request.

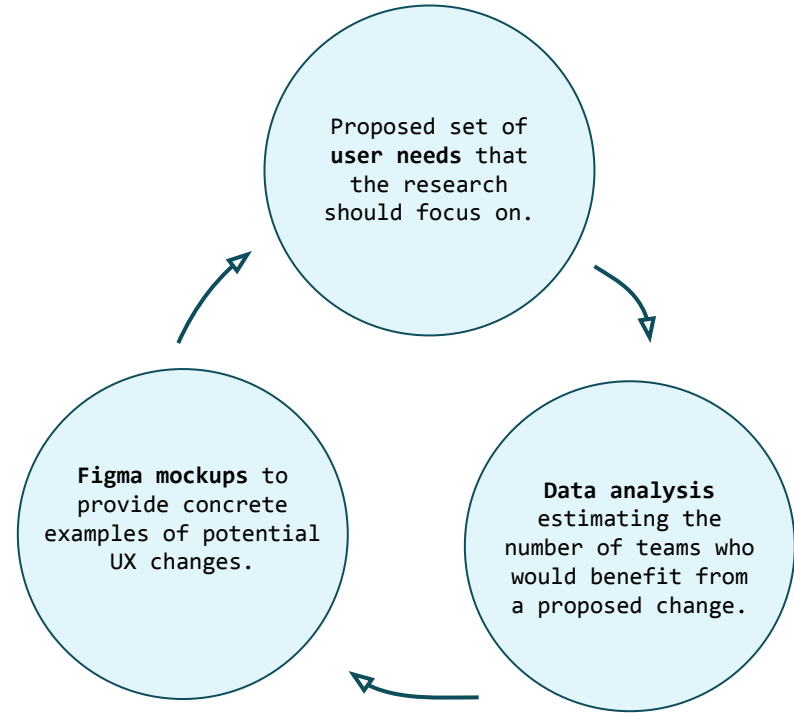


Iterative Scoping

Five iterations of scoping were required to define the user group and user needs that best reflected the goals of the feature request.

In each iteration, we discussed different criteria for teams that might benefit from having two 'parents'. My data analysis is described in more detail on the next slide.

I also used **Figma to generate mockups** to generate consensus around how the original feature request might work.

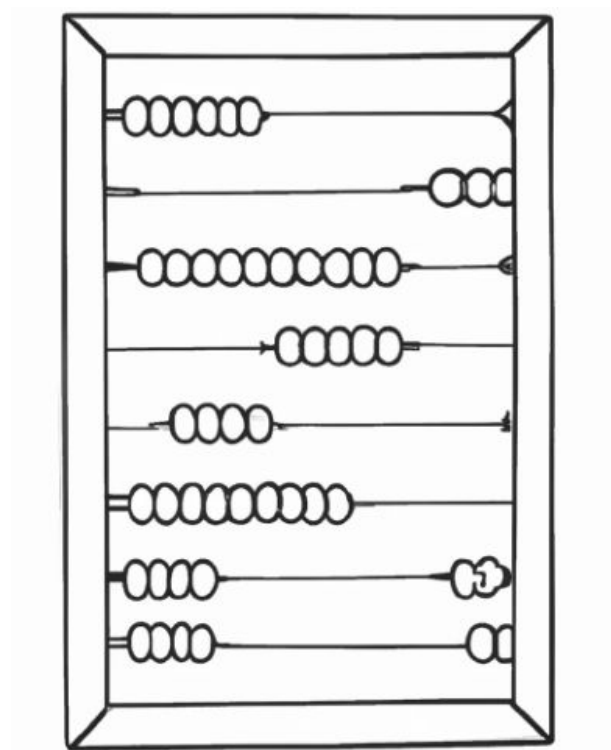


Data analysis

I used BigQuery and Jupyter Notebook scripts to estimate how many users would benefit from each iteration of the 'two parent' function. (eg. '132 teams at Google have staff from two or more departments'). The analysis ensured ensure any new features would benefit a broad user base.

The data was in the form of two 'graphs' (sets of nodes and edges); I wrote complex SQL queries to discover which teams had particular properties (e.g. 'team with at least three subteams with distinct groups of team members')

I used the Python NetworkX library to simplify the graph data and the Gephi visualisation tool to create network diagrams that articulated different team structures at Google.



Scoping outcomes

- **Four categories** of potential user benefits for senior managers were identified which we could test in the in-depth user interviews.
- The initial brief anticipated benefits primarily for **individual team members**. By systematically considering user benefits, we identified that **senior managers** were, in fact, the most important user group. This informed who we included in user testing interviews.
- I iteratively developed a mockup of the proposed feature, which we subsequently used for **concept testing**.

User interviews - recruiting

I created a recruiting survey and sent it to ~300 engineering team leads. We wanted to get a representative sample of team leaders with respect to tenure, region and team structure - categories we identified as strongly shaping users' experience of the Team Management Tool.

I ensured that all stakeholders approved our sampling method while guarding against manipulation of the sampling process.

Eight users were selected for interviews.

*Mockup of the
recruiting survey with
selected questions*

Engineering team leads UX interview invite

The Team Management Tool user experience team are looking for engineering leads to participate in a 45 minute remote interview about the way you use TMT. Please complete the form below if you are interested. As thank you gift, we'll make a donation to a charity of the participant's choice.

What is the composition of the team(s) you lead?

- ☐ One or more of my teams have a mixture of different types of engineers
- ☐ My teams are mostly or completely composed a single type of engineer
- ☐ Other: _____

What region do you work in?

- ☐ US / AMER
- ☐ EMEA
- ☐ APAC
- ☐ Other: _____

How long have you worked at Google in engineering?

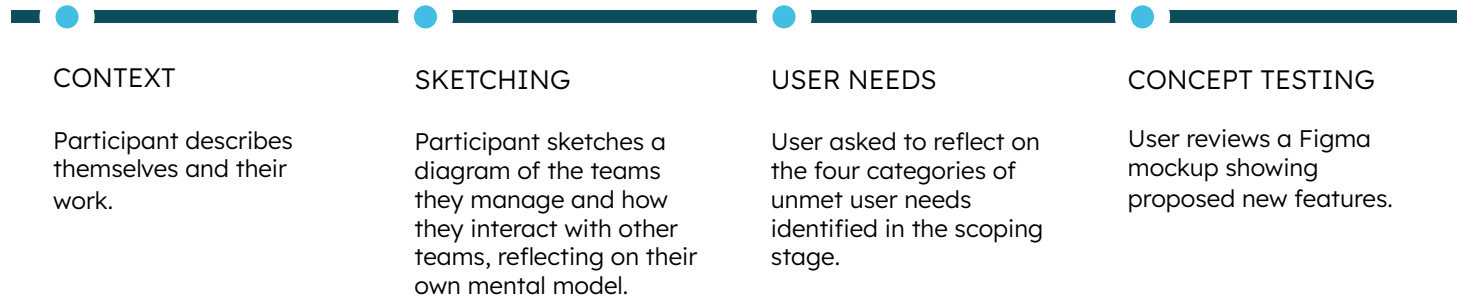
- ☐ One year or less
- ☐ Between one year and three
- ☐ Longer than three years

In-depth user interviews

I conducted eight semi-structured remote interviews, including moderating them and writing the script.

The interviews were recorded with engineering team leads from across Google. We planned to consider a second wave of interviews, however, after the first wave a clear picture had emerged.

Interview structure



Interview Findings

- The sketching phase indicated that participants' mental models of their team structure were highly diverse and complex - much more complex than would be captured by allowing teams to have two parents.
- The sketching phase also indicated that team leads would likely use the 'two parent' functionality to create extremely confusing team hierarchies.
- In the concept testing phase of the interview, almost no support was found for the prototype version of the feature request.
- Over half the users were not aware of an existing 'partnerships' feature that could be used to indicate collaborations.

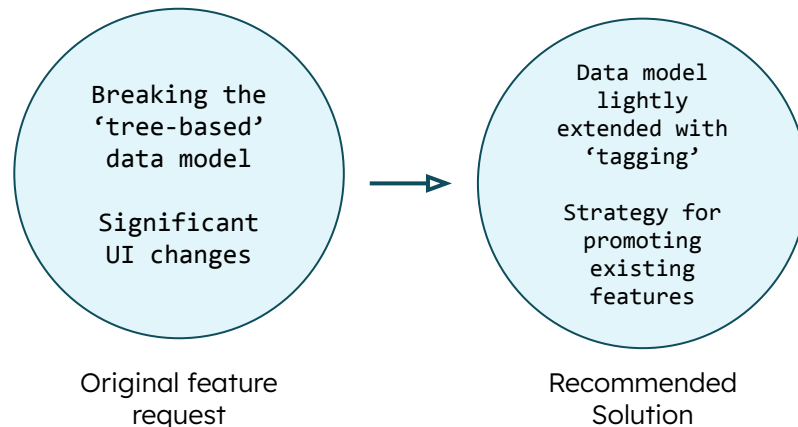
Impact

Product impact

- The original feature requester agreed that a naive implementation of their request would not meet their goals.
- A small extension to the existing data model was suggested using 'tags'.
- We identified a set of users who should be using an existing 'collaboration' feature and suggested a strategy for promoting this feature to them.
- This approach will benefit teams across Google, specifically thousands of engineers who use the system.

Further work

Producing a guide or set of templates to articulate best practices for using the team management tool was identified as a next step.



UXR secondment background

My secondment arose from an on-spec research project, demonstrating my ability to use speculative design and editorial writing to initiate and build momentum for research projects.

Speculative design

What we miss about the office, an 'on spec' research project about opportunities for collaboration during the covid lockdown, which circulated to hundreds of Googlers.

Research methods: speculative design, literature review.

Tech skills: BigQuery analysis



Research article

Article in an internal journal discussing my research.

Research methods: editorial writing

Tech skills: Network visualisations using Python Networkx library & Gephi



UXR secondment

My article led to a project with an infrastructure team to improve an internal collaboration tool used by over 2,000 engineers.

Research methods: Figma mockups and concept testing.

Tech skills: BigQuery analysis

PROJECT 3

PhD Research

My Role: Researcher / doctoral candidate

Team: Multiple collaboration, including with a regional NHS Trust and central government.

Duration: 4 years

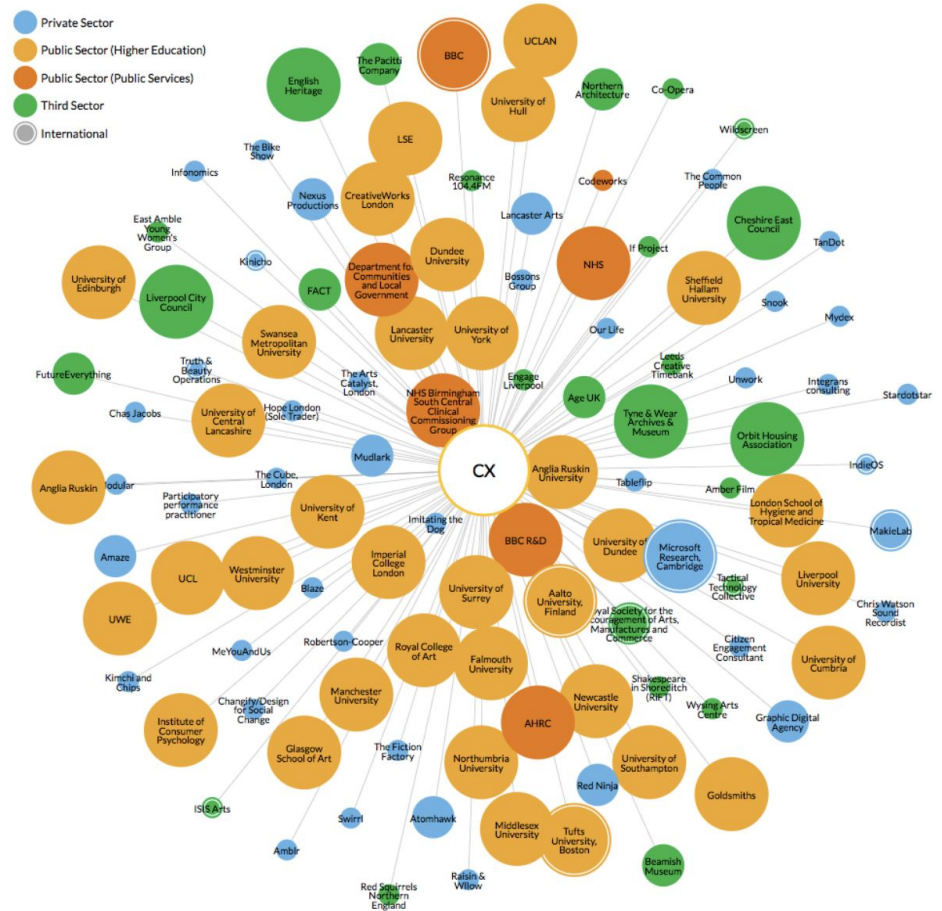
Outcome: a paid-for software product, a set of design principles, a doctorate

Context

My PhD was part of the Creative Exchange program. It was funded by the Arts and Humanities Research Council to explore 'digital public space'. There were 21 other doctoral researchers in the program.

The research program was a collaboration between the Royal College of Art (where I was based), Newcastle University and Lancaster University.

The Creative Exchange program acted as a hub between many other institutions. My collaborations were predominantly with local government institutions.



PhD Summary

Real-world problem

My PhD developed a social media analytics tool for local government. Social media hosts many conversations about topics relevant to local government - local planning issues, crime, the environment etc. However, local government rarely uses social media analysis as a means to understand the communities they serve.

In my PhD I built and tested a Twitter analytics tool (*LocalNets*) to explore social media analytics as a tool to support local government policymaking.

Theoretical framing

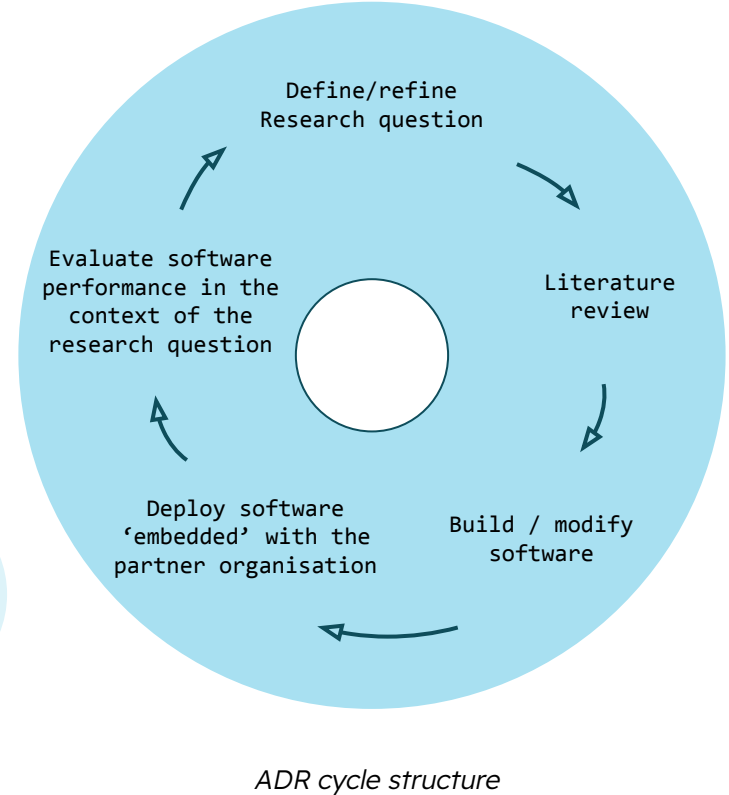
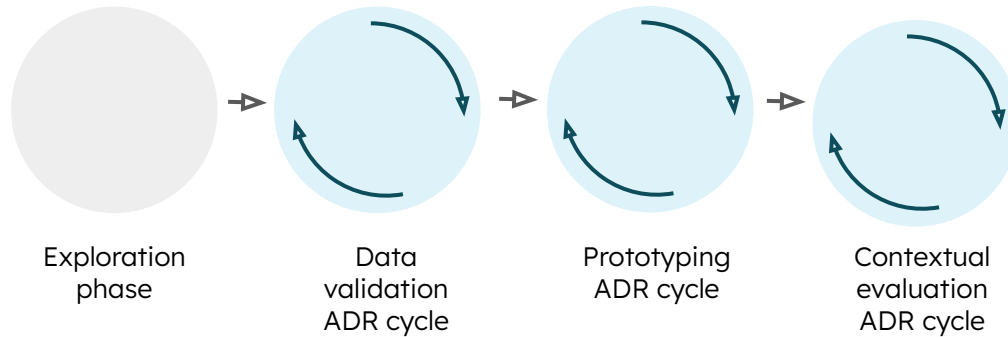
Drawing on design research methods, my PhD was highly interdisciplinary. The project was framed in relation to the concept of collective action from the discipline of political economy and used tools from network science to detect communities and process data.

Outcomes summary

1. A social media analytics product. Three organisations **paid a monthly subscription** to continue using my tool after the research completed.
2. A **set of design principles** for social media analytics for local government as a contribution to design research.

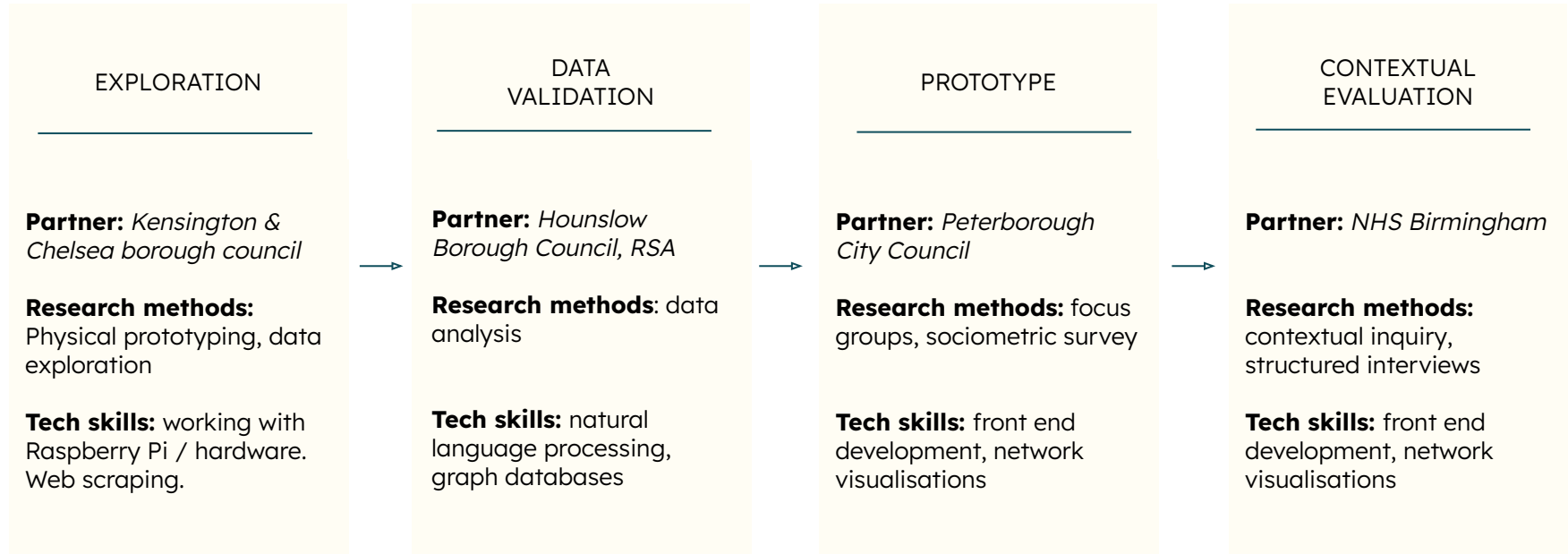
Research structure

My research used the Action Design Research framework, a cyclical approach to analysing software in context. The research consisted of three research cycles, each with the same ADR structure.



Research cycles & partner projects

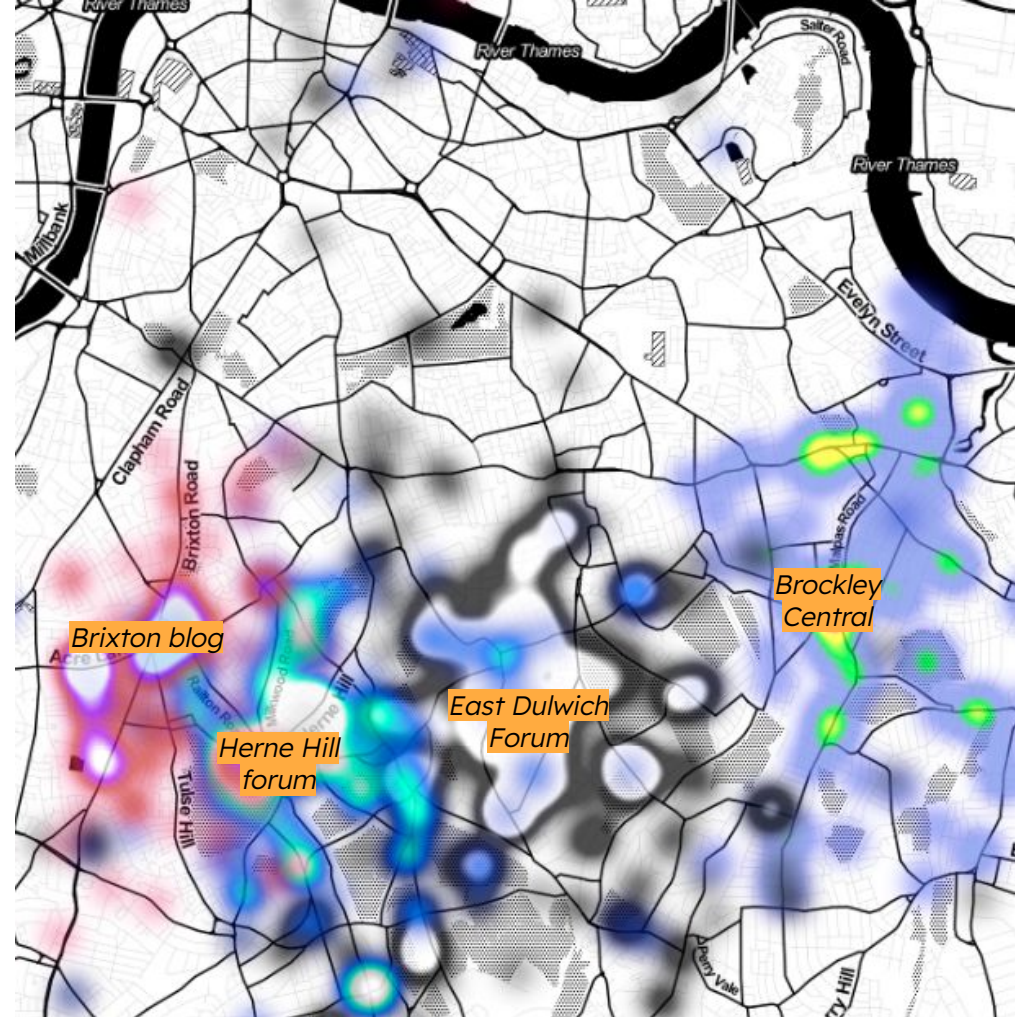
My PhD developed the LocalNets analytics software through a series of projects with partner organisations.



Exploration phase - data sources

At the outset of the project, an exploration phase investigated a wide range of options for using social media analytics in the context of Local government.

One strand of the exploratory phase considered I explored using data from 'hyperlocal' blogs and forums.



Exploration phase - tweet-based newspaper

As part of the exploration phase I created a tweet-based 'newspaper' (using the *Little Printer* receipt printer).

It was deployed in various locations, including in the staff canteen of Kensington & Chelsea borough council.

Tweets were selected manually for local relevance.

Printing an 'instant newspaper' using Twitter data. →

Printer in action. ↓



Prototyping phase selected findings

- In comparison with the Twitter API, collecting data from 'hyperlocal' forums is complex, and geographic coverage is poor. Based on the exploratory work, I decided to further investigate data from the Twitter API.
- In Kensington & Chelsea, there was a significant corpus of locally-oriented Twitter activity that was of interest to K&C staff.
- Taken together, these points supported the idea that Twitter analytics might be relevant for local government.

Quantitative data validation phase

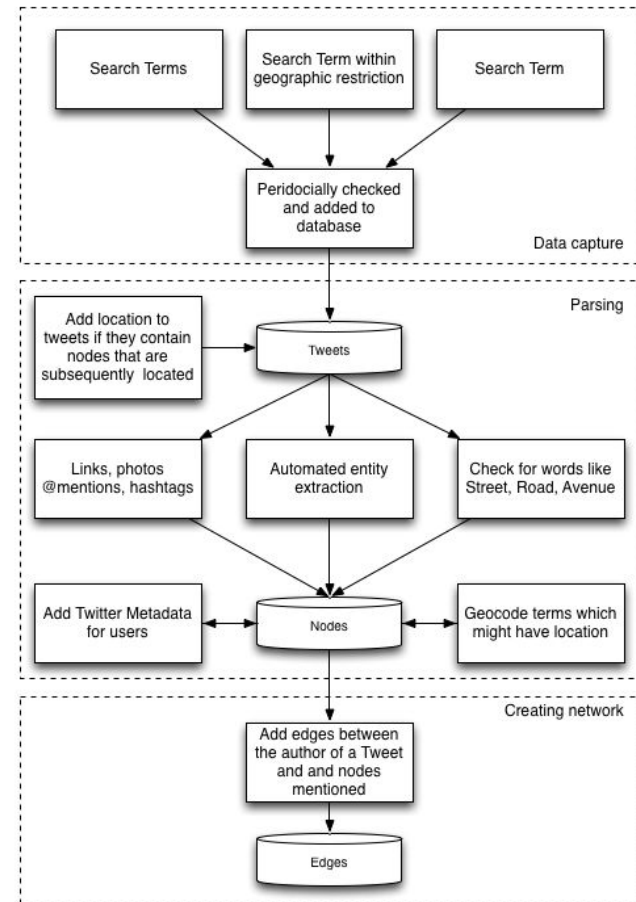
My second project was with the Royal Society of Arts (RSA) and Hounslow Borough Council.

I built a data analysis pipeline to add metadata about the locations and topics mentioned in tweets.

As part of their research, the RSA prepared a detailed asset map (a list of community centers, community events, community organisers etc.) of a single ward in Hounslow through a door-to-door survey.

In this phase of the research a comparison of the manually gathered and Twitter-based data was made, and explored visualising the data.

There was overlap between the manually mapped assets and the assets from the Twitter corpus (~10%).



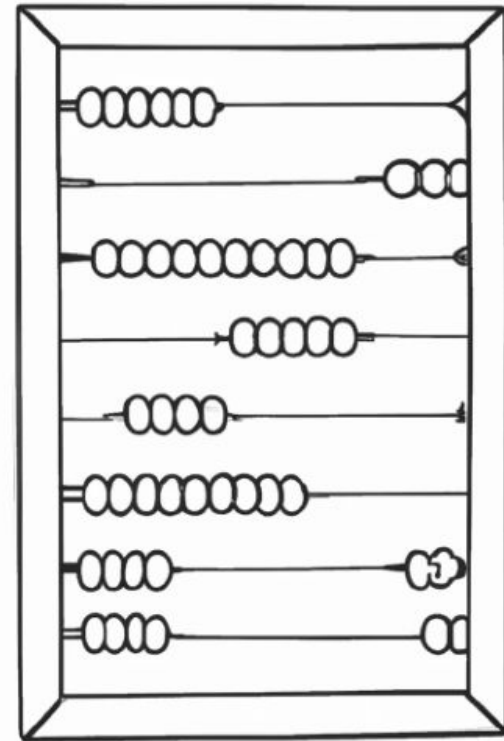
Software architecture diagram

Data Analysis

My PhD research drew on a number of data analysis techniques, including:

- **Natural language processing.** I used an NLP library to parse tweets for entities (eg. locations, people and institutions mentioned in tweets).
- **Network analysis.** Data from Twitter was stored in both a Mongo database and Neo4J graph database. Neo4J stored Twitter users and entities as nodes and mentions as edges. Using Neo4J I ran community detection algorithms (see next slide) and performed queries on properties such as betweenness and centrality.
- **Geolocation.** Localnets inferred the locations that tweets referred to by geocoding extracted location entities. It also inferred areas that specific Twitter users were connected to by aggregating their geolocated tweets. Geolocating

Community detection and geolocation were subsequently used to allow Localnets to find Twitter users to participate in local government focus groups.

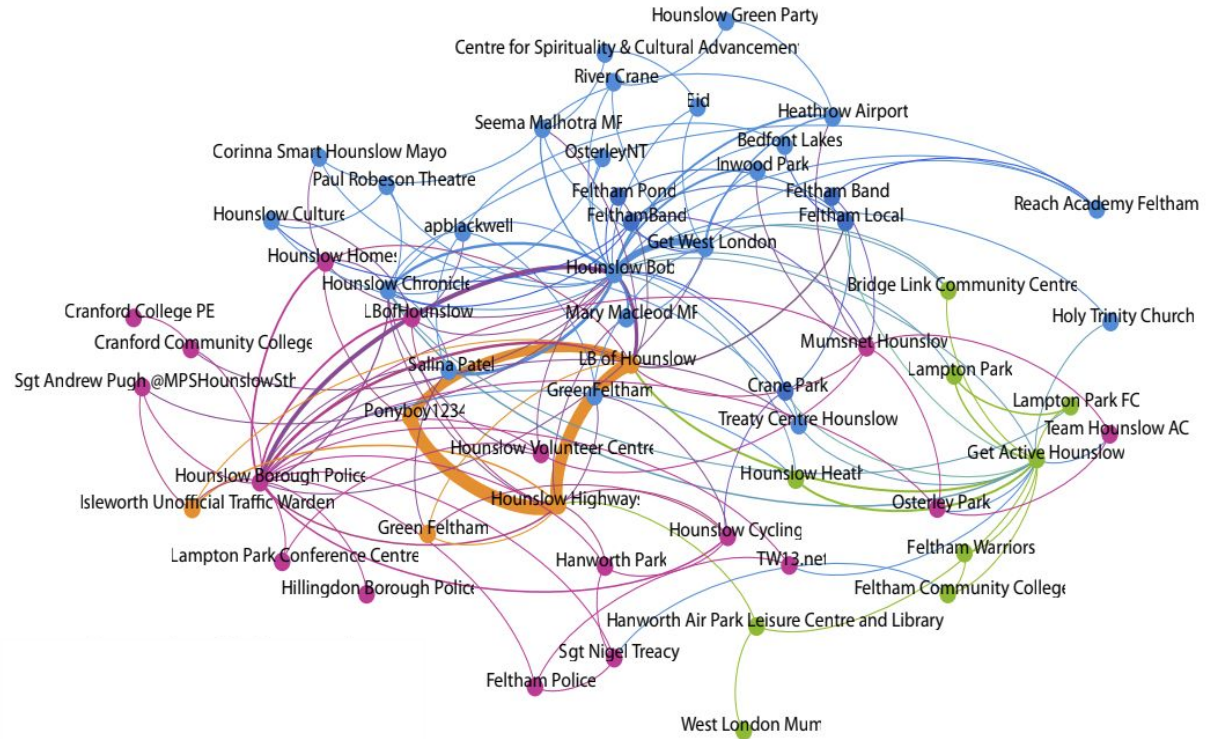


Data visualisation

To visualise the data collected by LocalNets, I created network diagrams of local communities. Within the LocalNets software, network analysis algorithms were used to discover relevant Twitter accounts. They were also used as part of the UI.

This diagram shows interactions between Twitter accounts in Hounslow, where nodes are accounts and the thickness of the edges shows interactions.

Colours show algorithmically detected communities.



Network map of community asset Twitter accounts in Hounslow.

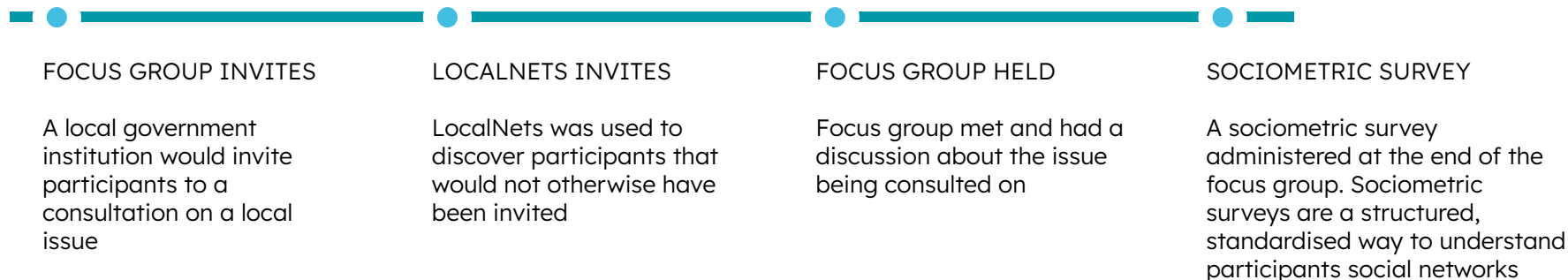
Data Validation Findings & outcomes

- Tweets were successfully tagged and geolocated to specific areas using natural language processing and network analysis.
- There was overlap between the manually mapped assets and the assets from the Twitter corpus (~10%).
- Individuals, as opposed to locations or events, were best represented in the Twitter data.
- **Outcome:** The data suggested that a potential use-case for LocalNets would be finding participants for local government consultation programs. The Twitter was able to find many individuals who were highly engaged in local topics (eg. traffic management) and lived locally in Hounslow.

Focus groups

The data validation phase indicated that LocalNets data might be effective at finding participants for local consultations. This led to the development of a method for using existing local consultation focus groups to evaluate LocalNets.

The method was designed to evaluate how LocalNets could contribute to broadening focus group participation and strengthening community connections.



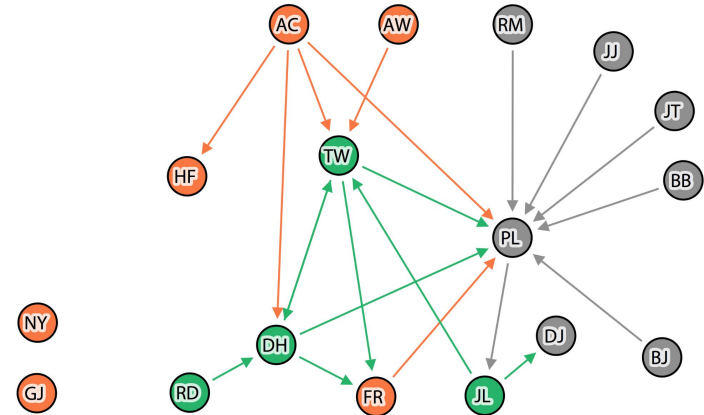
Sociometric surveys

This slide describes in detail how sociometric surveys were administered and analysed to evaluate LocalNet's impact.

Please indicate how familiar you are with today's focus group participants.

| | Anton Coles | Hugh Franklin | Andrea Warner | Tina Jones |
|---|-------------|---------------|---------------|------------|
| I know this person from Twitter | | | | |
| I know this person, other than through Twitter | | | | |
| I have previously worked with this person | | | | |
| I would like to keep in touch with this person | | | | |
| Would like to work with this person in the future | | | | |

Example of a sociometric survey question



Sociogram of a focus group in Hounslow. Nodes indicate participants, Arrows indicate who previously knew whom. Orange nodes were invited through LocalNets. Grey nodes were invited by Peterborough City Council. Green nodes were invited through both.

The diagram shows that LocalNets may have helped invite users who would not normally be part of the social group that attends local focus groups.

Findings & outcomes

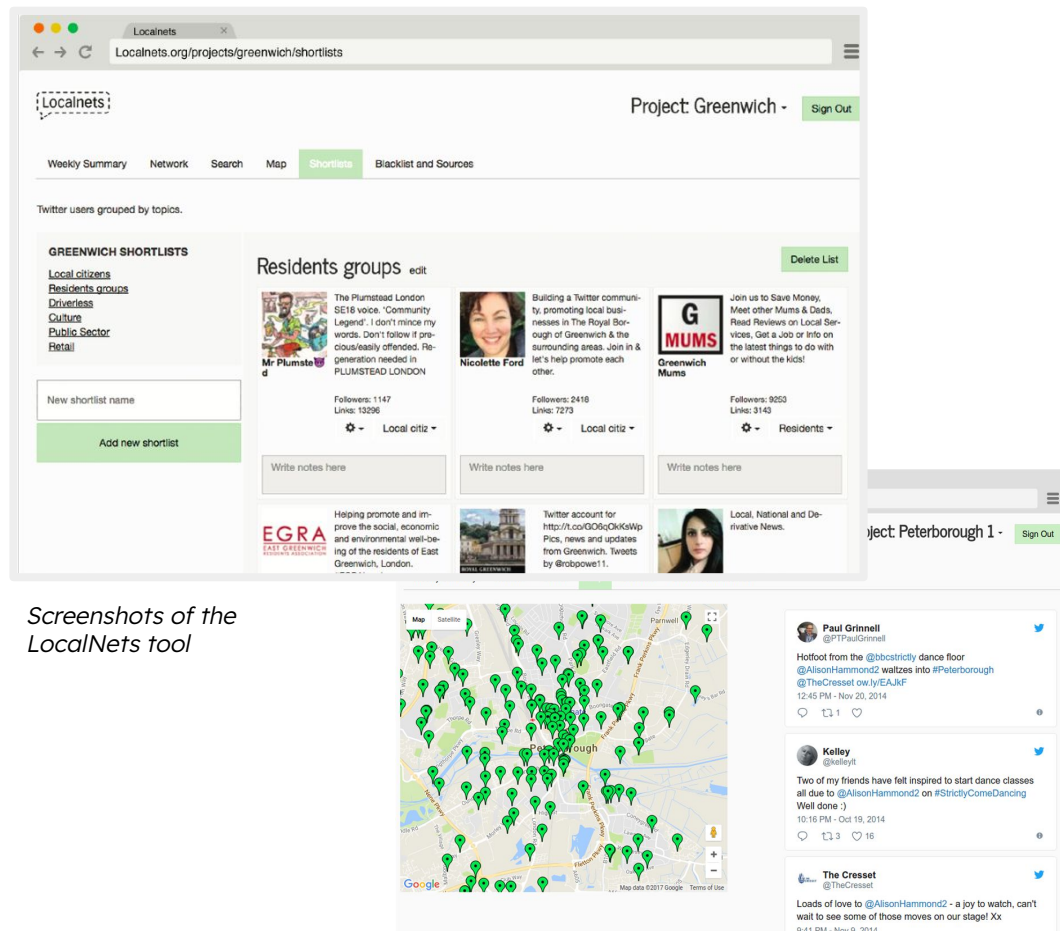
- This sociometric surveys built a picture of how LocalNets contributed new participants to the focus groups, indicating that many new participants were known to existing ones.
- Despite not delivering a wholly new set of participants, local government organisations valued the ease with which they could extend focus group participation.
- **Outcome:** The findings built confidence that LocalNets could address a genuine user need for local government organisations, and supported moving to a more detailed review of how LocalNets would be used in context by Local Government staff.

Contextual evaluation

A contextual evaluation was conducted with three partners, with NHS Birmingham testing the final iteration.

In this stage of the research, the tool was used by NHS Birmingham staff to select participants in a local consultation about an underperforming medical practice.

I used semi-structured interviews to understand their experience of using LocalNets. The interviews were in person and took place in the user's workplace.



Screenshots of the LocalNets tool

Contextual evaluation outcomes

- Network diagrams, which were presented as part of the interface, were rarely used. While they convey a 'summary' of the data LocalNets, when it came to task completion, such as selecting participants, they are too complex to convey sufficiently precise information.
- On the basis of these interviews, functionality was added to create shortlists of potential participants, responding to a need to generate and review multiple sets of potential invitees.
- I also added functionality to export shortlists to Excel, which was the most common means for sharing and discussing potential participants within the local government institutions I worked with.

Impact

My PhD research resulted in:

- LocalNets - a working web app, with three institutions signing up to pay a monthly fee to use it.
- A set of four design principles for social media analytics tools for public sector institutions.
- Avenues for further enquiry, including the use of network algorithms to detect communities in the context of local government consultations.
- A set of situated design findings based on the specific set of case studies where it was tested.
- A doctorate.



END