Incubator for AI

Case study using ThemeFinder for analysis of responses to DSIT's Digital Inclusion Action Plan Call for Evidence

01/09/25

Executive summary

ThemeFinder is a topic modelling Python package designed for analysing one-to-many question-answer data such as survey responses and public consultations. It is the AI capability which underpins Consult – a tool which alleviates the burden of running consultations by facilitating quicker analysis for a fraction of the cost.

For a Call for Evidence accompanying the Digital Inclusion Action Plan DSIT used ThemeFinder. This case study summarises the experience.

This is not a formal evaluation, but provides supplementary evidence alongside a wider programme of evaluation and user research.

Headlines:

- ➔ **ThemeFinder enabled policy teams to conduct more detailed and meaningful analysis.** Its mapping functionality allowed users to navigate raw data effectively, resulting in insights that are more rigorous and representative of the full dataset.
- ➔ **The process of generating and editing themes was well-received**, with policy colleagues expressing strong confidence in the relevance and appropriateness of the themes produced, often requiring minimal revisions.
- ➔ **ThemeFinder's mapping was considered accurate and sensible**, bolstering trust in the analysis output.
- ➔ **ThemeFinder sped up identification of initial results**, providing rapid insights to inform ministerial engagement. However, overall report writing remained time-consuming due to the requirement to review all responses.
- ➔ **Those reviewing responses also felt ThemeFinder helped to mitigate against bias**, making it more possible to conduct a neutral and balanced overview of responses, which enhances the credibility of the analysis.

Contents

Introduction

Background

What is the incubator for Artificial Intelligence (i.AI)?

i.AI rapidly designs, tests and delivers AI products for government. We take pride in being a highly technical team, composed of experts with deep knowledge and experience from both the public and private sector. This allows us to tackle complex challenges and innovate continuously towards our overall mission: to harness the opportunity of AI for public good.

What is ThemeFinder?

ThemeFinder is a topic modelling Python AI-powered package developed by i.AI to analyse one-to-many question-answer data such as survey responses and public consultations. It is the AI capability which underpins Consult.

ThemeFinder's pipeline consists of five distinct stages, each utilising a specialised Large Language model (LLM) prompt: Theme generation, Theme condensation, Theme refinement, Theme target alignment and Theme mapping.

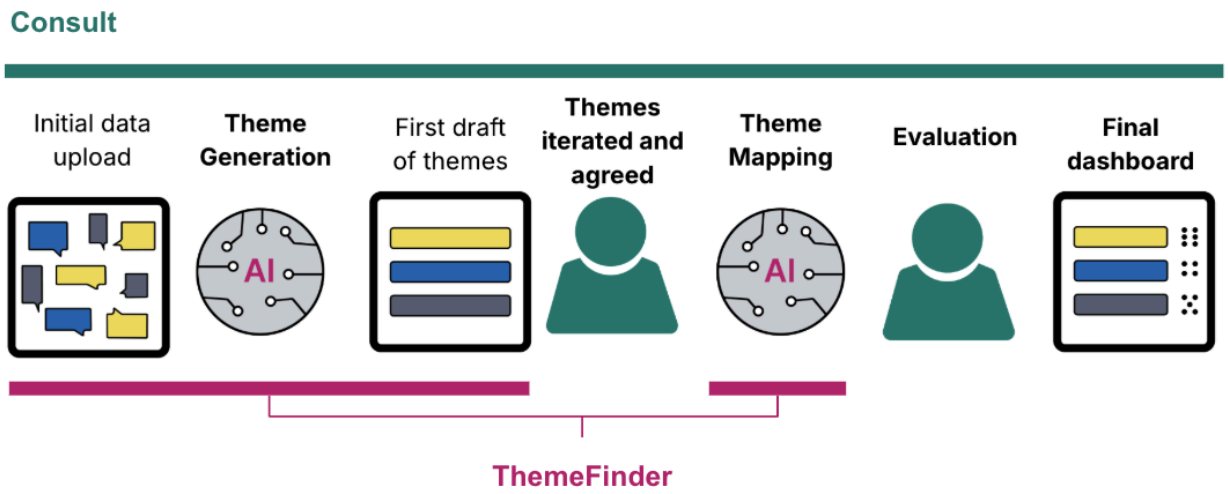
What is Consult?

Consult, aims to facilitate quicker analysis of the over 600 consultations run by the government every year, with comparable quality, for a fraction of the cost. Also developed by i.AI, it uses ThemeFinder to generate themes for each question and map responses to those themes and has additional safeguards as it's been designed to be used for consultations which need to be analysed in accordance with the Gunning Principles. So there is a human in the loop to ensure appropriate oversight and a dashboard designed to specifically support consultation analysis.

It does this using a multi stage process with a human in the loop:

1. **Theme Generation**: Identify common themes in the responses to generate a list of themes for each question.
2. **Theme Sign off**: The team running the analysis reviews a long list of themes and decides which to use for the analysis, with the ability to add, edit or remove themes.

3. **Theme Mapping**: Classify all responses using one or more of these agreed themes (or none, where appropriate).



Consult also qualifies the quality of the performance of the theme mapping before giving teams access to a dashboard to understand the relationships between responses.

Why have we conducted this case study?

Consult is still under development and the team developing it are learning from teams currently doing consultations and calls for evidence to understand their experience and how Consult and ThemeFinder can add value to the process.

Earlier this year (2025) Department for Science, Innovation and Technology (DSIT) conducted a Call for Evidence to accompany the Digital Inclusion Action Plan. DSIT decided to experiment with AI analysis alongside reading all responses.

The consultation analysis approach

The call for evidence received 800 responses; 756 via the online form hosted on GOV.UK and 44 via the email address provided. A thematic analysis of responses to the 17 open-ended questions was conducted. While all responses were read in full, ThemeFinder, in combination with extensive manual review, was used to identify key themes.

Steps taken

All responses for open (free text) questions were sent to a Large Language Model for Theme generation, Theme condensation, Theme refinement, Theme target alignment.

1. Data was gathered into one source and cleaned
2. Data was run through ThemeFinder to generate themes
    a. A long list of themes (theme name and theme description) for each question with estimated frequency
    b. Examples of responses for top 20 themes
3. Policy team reviewed all themes and examples to decide final list of themes to be used for the mapping
4. Themes were mapped by ThemeFinder
5. Policy teams used these mappings to inform their analysis while reading all responses to write reports.

Evaluation approach

The team chose not to conduct a formal evaluation of the theme mapping because they were using ThemeFinder as an aid to their existing process of reading all responses, rather than a replacement. The following findings are anecdotal. This case study documents their experience of using ThemeFinder as captured in verbal feedback to i.AI's Consult team.

Findings

- **Depth of analysis**: The team achieved more detailed and meaningful analysis more efficiently, with ThemeFinder's mapping enabling policy teams to navigate the raw data effectively and produce more rigorous, representative insights.

    "The output is more rigorous and more representative of the data… Policy teams were able to get closer to the raw data. The mapping allowed them to navigate the data "

- **Value in theme identification:** The team found the theme identification process of ThemeFinder and the fact that they could edit them positive.

"conversations with policy colleagues were really positive; they didn't have many comments for revisions at all"

- **Accuracy in theme mapping:** The team found the mapping accurate.

  "pretty accurate with sensible categorisations"

- **Speed:** The team found that ThemeFinder sped up getting first results but didn't save time overall in writing the whole report because of the need to read all responses.

  "Analysis ran so quickly we could give the Minister results quickly to feed into their engagement at the time"

- **Objectivity:** The team felt that ThemeFinder was unbiased and able to categorise responses in a way that reduced the bias of the reviewer.

  "it was neutral, not biased by who was most eloquent or landed arguments best or which points appealed to the policy official and the level of detail it goes into is very useful"

- **Qualitative analysis**: The team were using the themes to identify the breadth and diversity of responses, not to quantify them and were positive about ThemeFinder's ability to do this.

  "this does it better than a human researcher given the timescales we were working to"

- **Quality of responses**: Some responses were not addressing the points the team were looking to improve, this made them hard to draw useful insight from and categorise.

  "people want to share what they want to share"

Takeaways and implications for Consult

- In order to save teams time the Consult team needs to understand how many responses need to be read.
- The Consult team should continue to understand how teams conduct the analysis to design the dashboard to meet those needs.