The use of Artificial Intelligence in Planning: Frequently asked questions

Mark Green[1], Alex Lord[1], Alex Singleton[1], Terry de Sousa[2], Samantha Johnston[2].

1 Department of Geography and Planning, University of Liverpool, UK

2 Greater Cambridge Shared Planning, Cambridge, UK

Introduction

This document provides a discussion around the use of Large Language Models (LLMs) and Artificial Intelligence (AI) to summarise representations made about planning documents. It details the methods underpinning the technology, outcomes associated with the evaluation of its use, and describes some concerns commonly raised about the work – along with some responses to those concerns. The creation of this document was recommended following engagement with public and stakeholder groups on the use of these technologies to help build trust and reassurance. It has been created in the spirit of having an open discussion around the tool that has been created. The document has been structured in relation to the key themes of questions.

Methods

How does it work?

The tool is designed to replicate the current process of planning officers reviewing and summarising public representations to consultations. It creates two reports: (i) a short summary of each individual submission describing what the representation said, and (ii) a report which provides an executive summary of the main issues raised, a more detailed list of individual points raised both in support and objection to the document (including citations to link points to individual submissions so you can see who said what), and a geographic analysis of who submitted evidence (e.g., which areas had more or less submissions, how submissions varied by level of deprivation).

These two reports are created through using a generative AI. Rather than designing a bespoke LLM developed explicitly only for planning services, we utilise an existing LLM that we repurposed for the specific tasks outlined above. This has the benefit of using a 'state of the art' product that has been trained on billions of pieces of information and validated elsewhere. These LLMs are easily adaptable to be applied to different tasks, saving on costs of development and environmental impact. We currently embed the system within Microsoft's Azure service and use LLMs developed by OpenAI.

Can the AI tool 'make up' representations, miss out key points, or provide misleading summaries?

The tool has been designed so this is not possible. It only considers all submissions stored within the database and can only draw information from these sources. We avoid any misleading summaries through introducing an 'agent' (i.e., a second AI resource)

within the tool that reviews the summary produced by the initial LLM to check it is a fair account of what was written helping to moderate the generated summary. Where the summary diverges from the original submission, the tool will revise the summary and this is then further checked – the content is stored as an output only when there is agreement (i.e., this is an interactive loop). We have extensively trialled this to make sure it performs well at doing just this and have found that it captures all points effectively. We also ask the AI tool to provide citations so where it reports key points raised, they can always be traced back to a specific submission.

What does the tool do when it encounters offensive language? Will this be included in generated summaries or reports?

We include content filtering within all summarisation and generated text. Validated classification models check whether there is any harmful content (e.g., discriminatory language, swear words, violent comments, mention of self-harm) and removes them from consideration in our outputs.

Can the AI discern the nuances of human language (e.g., notice sarcasm)?

The LLM we use has been trained on billions of records of text and human conversation which gives it some understanding of the nuances of English. In our tool, the LLM will report the implied meaning of statements rather than take things like sarcasm on face value so that comments are still captured. To some extent, this mimics a human planning officer who may or may not understand all nuances.

Could the AI tool be designed to ask leading questions or produce a report that gives a more favourable outcome based on a planning authority's needs?

We have designed the tool so it generates a fair and clear summarisation of the information presented. Its goal is to repeat back the main points raised in representations and not offer any judgement about them. Our tool and systems are designed with high security so it would be almost impossible for someone with malicious intent to reprogram the code.

Who created the tool?

The tool was created by the University of Liverpool (Professors Alex Lord, Alex Singleton and Mark Green) with input and advice from Greater Cambridge Shared Planning service (Terry de Sousa and Samantha Johnston).

Ethics and data protection

What are the potential biases of the tool?

The underlying LLM is trained on a large body of information across the Internet, books and other text sources. These tend to reflect the dominant perspectives of the authors of these sources, meaning that it was trained on information over-represented by mainstream Western, English-speaking and often male voices. To minimise these potential biases, we have constrained the tool so it focuses on just summarising what public representations include (i.e., providing an objective description of what others say). This means that it approaches summarisation tasks with a 'statistical view' of what was said, rather than reflecting on any lived experience of subjective or cultural nuances. We also acknowledge that the tool may repeat back biases within public submissions when summarising viewpoints. This is not any different from the traditional approach where planning officers need to make a judgement on the final decisions as to how these issues affect their knowledge.

Our overall summary report presents an analysis of the types of areas and peoples who submit evidence. This is not something currently done by Greater Cambridge Shared Planning service and allows them to examine the extent of inequalities in who is submitting representations. This new information should help assess the extent of bias across public submissions and potentially allow them to target under-represented communities during future consultations.

Is this tool General Data Protection Regulation (GDPR) compliant?

The tool is GDPR compliant. It redacts any personal information from representations before summarising them. All information is processed in the cloud using servers secured to industry standard GDPR compliant servers that are based in the UK. No information, personal data or text from submissions is stored by the cloud-based services used in the tool.

How accountable is the AI tool? What happens if a wrong decision is made based on the AI tool?

Planning officers make all decisions relating to the reports. No decisions are made by the AI tool and it does not offer suggested actions – it just provides a summary of public representations. Greater Cambridge Shared Planning service will continue to review the original representations when preparing the future stages of the Local Plan or other planning documents and the final decision on whether to consult, amend or adopt a

document will lie with the relevant Councils through existing Committee or Cabinet meetings. The final say is always with humans.

Will people who submit representations be allowed to check their AI generated summaries?

Currently, Greater Cambridge Shared Planning service emails all summaries to people who submit representations, and this will continue with the tool. People / groups who submit can challenge or ask for revisions to be made to any summary. We are currently exploring the potential to offer the AI generated summary in real-time when submissions are made to speed up this process.

Evaluation of the tool

How has the AI tool been trained?

The development of the tool has been based on trialling the code on 100,385 representations submitted across 164 documents (2012-2024) in the Greater Cambridge area. We generated summaries of submissions and overall reports and then compared them to the published human created outputs to see how well our tool performed. This provided qualitative insights into how to refine the tool further to ensure it gives an accurate report.

How reliable and valid is the tool in producing accurate summaries / reports?

As part of the evaluation of the tool, we applied the tool 'live' to three public consultations in January 2025. At the same time, human planning officers continued to manually summarise each submission which allowed us to compare the quality of both methods. Greater Cambridge Shared Planning service then conducted an independent comparison of human and AI generated summaries. They found that there were no noticeable differences in quality of the AI summaries when compared to the human generated ones. Where text differed, this was either due to small divergences in discussion that were not meaningful (e.g., same comments described differently) or occasions where the AI tool added parts that human planners did not mention.

Greater Cambridge Shared Planning service have been active in reviewing and feeding back on outputs throughout product development which has helped contribute to ensuring the accuracy of reports.

Is the AI tool faster than humans?

During the live evaluation of three consultations in January 2025, it took human planning officers a total of 65 hours to process, summarise and respond to each representation, of which 18.5 hours were just on summarising. In contrast, it took the AI tool 16 minutes to summarise each submission and create an additional summary report detailing the main points raised.

Are there opportunities for the tool to learn based on feedback?

The tool is designed to not store any knowledge or information relating to submissions so that it remains GDPR compliant. We will revise the tool considering any feedback and tailor the code so that it continues to offer the best service possible.

What is the environmental impact of using this tool?

Each document the AI tool summarises generates approximately 4.32 grams of CO2 emissions. This is equivalent to sending two emails, growing three strawberries, streaming a 10 second video, or boiling half a cup of water in a kettle.

There is significant environmental footprint of training, running and maintaining large language models both in electricity generation and water consumption. For example, it was estimated that OpenAI's ChatGPT-3 model consumed enough energy during training to power 1,000 households for one year. We use Microsoft's Azure which has outlined plans to achieve 100% renewable energy usage by 2025. The service is carbon neutral and is constructing more sustainable data centres as well.

Implementation within planning services

When should planning officers use this tool?

We recommend that the tool is only used for the summarisation of representations linked to planning documents.

Does the AI tool make planning decisions?

The AI tool creates a summary report of all public submissions only. It is then for planning officers to read and interrogate the report and decide based on the findings from the report (and other consultations / evidence). The final say is always with humans.

Was this tool created to replace human planners and jobs?

We designed the AI tool to always keep the 'human in the loop' with the idea of having tools that support rather than replace humans. This can then 'free up' the time of planning officers, minimising laborious and repetitive tasks and allowing them to spend more time on more complex and key planning issues where their experience and expertise is better directed.

It is recommended that planning officers review a sample of the summaries generated by the LLM to ensure they are accurate and true reflections of the original comments, and that the full, original representation should always be reviewed when officers are reviewing and responding to public submissions and not solely relying on the LLM outputs when doing this task.

Can this tool be used by other planning authorities?

Currently, it is designed to only work with the systems that Greater Cambridge Shared Planning service use. However, with some minor adjustments we can easily deploy this tool elsewhere. Anyone interested in finding out more should contact Professor Alex Lord via alexlord@liverpool.ac.uk.

Contact

If you have any questions about this report, please get in touch with:

Professor Mark Green

Department of Geography and Planning

University of Liverpool

Liverpool

L69 7ZT

Email: mark.green@liverpool.ac.uk

Acknowledgements

The project was funded by the Ministry for Housing, Communities and Local Government's (MHCLG) 'PropTech' fund. This work was supported by the Economic and Social Research Council [grant number ES/L011840/1 and ES/Z504464/1].