

How can chatbots help to improve government information provision on EA?

Van Ravesteijn, M.P., Muspitasari, Y. & van Eck, G.

As Dutch national government EA-experts we have been experimenting with ways that AI can assist in the way we inform other governments and the public about the EA instrument. This paper describes the results of this experiment.

Introduction

The use of Artificial Intelligence (AI) is growing rapidly. In 2024, all EU Member States recorded higher shares of enterprises using AI technologies compared with 2023 (Eurostat, 2025). Governments are also increasingly using AI. One of the ways that governments use AI is to promote and improve information provision (Van Noordt & Misuraca, 2022). Local governments and environmental impact assessment practitioners can benefit from improved information on the EA-instrument.

The main question in this paper is: how can AI enhance information provision regarding the EA instrument in the Netherlands?

Sub-questions are:

- How reliable are the AI-generated answers?
- What lessons can we learn from AI-answers to EA helpdesk questions?
- What are the barriers governments face when using AI?

Information provision in this paper refers to the way in which information on EIA and SEA in the Netherlands is provided by the national government to public organizations, focusing on information provision through the Dutch national 'IPLO' helpdesk and website.

About the experiment

In 2024 we conducted an experiment to answer helpdesk questions with three chatbots. The experiment consisted of five steps:

1. We provided 24 representative anonymized questions from the helpdesk. For example: what are the differences between the procedures of EIA and SEA? We made sure the questions couldn't be traced back to persons, organizations or projects.
2. The questions were fed to IPLO AI, DemoBot Friese Aanpak and ChatGPT in December 2024. These chatbots generated answers including references to source pages. IPLO AI was not set up to answer follow-up questions. No follow-up questions were asked to be able to compare the answers of the different chatbots.
3. We reviewed these answers and sources and rated them on a five-point scale (see below). The three criteria we used, in this order, were: correctness (is the law explained correctly), completeness (are all the elements present) and understandability of the answer.

1	Totally wrong or incomprehensible answer or no answer
2	Largely incorrect or incomprehensible or very incomplete answer
3	Largely correct answer but some basic information is missing
4	Correct answer but nuances are missing
5	Near perfect answer, only small optional corrections

4. We discussed the findings with a target group. In preparation for the discussion, members of the target group were asked to score the answers of helpdesk questions on the quality of the answer.
5. We interviewed AI-experts and government employees who have experience with using AI about our findings and about the barriers for governments wanting to use AI.

As explained before, we used three different chatbots in our pilot. Each had different characteristics. This is explained below.

About IPLO AI

IPLO AI is a chatbot with access to specialized content only. The goal of the IPLO AI was to explore whether AI can help experts find the right input to answer questions quickly and adequately. The IPLO AI works in the similar way as the general language model by the user writing questions or requests and the IPLO AI providing the answers. In contrast to open chatbots such as ChatGPT, the IPLO AI was restricted to only search answers within the approximately 7.000 IPLO webpages.

About DemoBot Friese Aanpak

The DemoBot Friese Aanpak is a trained AI chatbot developed by de 'Stichting Digitale Intelligentie' (foundation for digital intelligence) and the 'Friese Aanpak' (an cooperation initiative of local and regional governments in the province of Friesland) to help frontend employees in Friesland answer helpdesk questions about the new environmental law. The DemoBot Friese Aanpak is restricted to only use the law (Omgevingswet), related legal texts and the IPLO-pages. Therefore, it has more sources than the IPLO AI but less than ChatGPT.

About ChatGPT

ChatGPT is a generative artificial intelligence chatbot developed by OpenAI and launched in 2022. Generative Pre-trained Transformer (GPT) architecture is the foundation of ChatGPT. ChatGPT enables users to discuss with the AI by inputting prompts (Haleem, Javaid & Singh, 2023). ChatGPT searches the whole internet to find an answer.

Results

The quality of the answers of the IPLO AI and the other chatbots differed greatly. Scores 3, 4 or 5 were given for largely correct answers on which the experts can build their answer. IPLO AI gave largely reliable answers in 63% of the EA-questions compared to 58% for the DemoBot and 25% for ChatGPT.

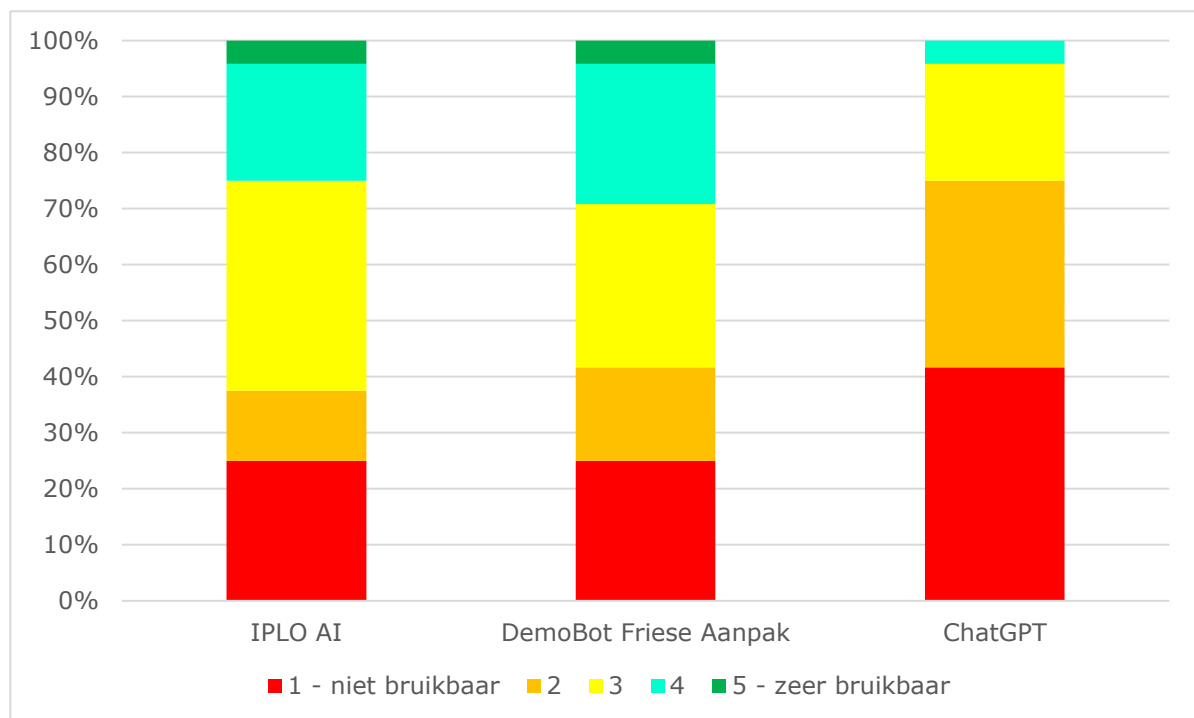


Figure 1 Results of the review of the answers on a scale from red (incorrect and not usable at all) to dark green (correct and very usable). Source: pilot with only questions EA-experts, December 2024

Reliability of chatbot answers to helpdesk questions

Our experiment showed that the reliability of the AI-answers differed greatly. This was due to three main factors: the quality of the webpage, the type and complexity of the prompt, and the type of sources that the chatbot used to find the answer. These factors are explained below.

1. IPLO AI produced well formulated answers in the cases that the answer was clearly present on one of the IPLO-webpages.

IPLO AI tried to source the input for the answers from other IPLO-page, often resulting in wrong answers. For example, sometimes a question about EIA was answered using the rules for SEA. For some answers or conclusions in the answers it is not clear how the tool reaches the conclusions. It seemed that the chatbot made up an answer (or in terms of chatbots started to 'hallucinate'). This happened more with ChatGPT than with the other chatbots.

2. The quality of the answers depends on, among other things, the type and complexity of the prompt.

The reliability of the answers can be improved with better prompts (questions) and/or asking follow-up questions. One of the interviewed experts explains: "The first answer you get from a chatbot is never correct. You always have to chat a while before you get the complete correct answer." According to the expert a good prompt meets the following criteria (interview expert DemoBot Friese Aanpak, January 2025):

- Give a clear assignment
- Add guidelines like 'Give a complete answer' or 'Give an answer at language level B1'
- Add your role; e.g. are you asking a question as an expert on this subject or as a member of the public?
- Add the context; why are you asking this question?

3. IPLO AI and the DemoBot Friese Aanpak gave more reliable answers to the helpdesk questions on environmental assessment.

ChatGPT sometimes referred to older laws, instead of current laws. This resulted in incorrect and unreliable answers. In most cases, incorrect answers were the result of outdated sources. Having access to relevant and up-to-date source material proved to be highly beneficial for the IPLO AI and the DemoBot Friese Aanpak.

Lessons AI teaches us for better information provision

This experiment with the IPLO AI and the other chatbots taught us three main lessons regarding the way we provide information in the helpdesk and on our website. These are described below.

1. Accessible language

The chatbot uses more simple words and less grammatically complex sentences. This was especially appreciated by the target group. One member of the target group perceived short, to-the-point answers as more reliable than extensive complicated answers. However, others missed relevant information in the short AI-answers (Target Group Discussion, February 2025). The IPLO website aims to provide information on language level B1/B2 (intermediate level). The challenge that experts often face is translating the technical and judiciary to accessible language. The 'language models' of the chatbots help to formulate answers in a simple language that is easy to understand.

2. Well structured answers

AI answers break down questions into several sub-questions and address them accordingly and logically. However, the answers to these sub-questions must be consistent and not contradictory. Experts handling helpdesk inquiries can learn from this approach to structuring complex questions.

3. Identifying knowledge gaps in webpages

Sometimes the AI generated incorrect answers due to outdated or incomplete information on IPLO-webpages. In this sense, AI acts as both a filter and processor of information, offering insights into the completeness and accuracy of the website's content. This helps us improve the pages that explain the EA instrument.

Risks and barriers governments face when using AI

This experiment has looked at the technical possibility of using AI. However, we also wish to address the question whether governments should use it and if this involves any risks and barriers. There are two main concerns: confidentiality of information versus the use of public (commercial) tools, and the energy consumption associated with intensive AI use.

Currently, the available public AI tools are provided by commercial parties. The use of these tools is therefore often regarded as risky due to uncertainty about where the submitted information is stored or transmitted. Thus, this hinders many government employees to use AI since they often work with information that is still classified before publication. However, the Dutch government is currently experimenting with ways to use generative AI safely and to understand the risks without missing the opportunities provided by this new technology.

The IPLO AI pilot is one of these experiments. One barrier in the IPLO AI pilot was that the actual helpdesk answers could not be used as a source of information for the IPLO AI. This was due to privacy restrictions regarding the personal data of the people who asked the helpdesk questions and the specific project information which is provided in order to get a tailored answer.

For government officials working in the field of environmental impact, the energy consumption of AI can also be a barrier to using AI. AI models can spur huge energy-consuming datasets and big data centers (Nishant, Kennedy & Corbett, 2020).

Conclusions

The experiment shows that while experts are still better in formulating professionally correct answers, chatbots are sometimes better in formulating accessible answers. The easier language used by the chatbots allowed us to improve our IPLO answers and webpages.

We also learned under which conditions the chatbots are able to provide good answers. Chatbots can provide good initial answers when the appropriate input is available in their databases. However, when the information is missing or incomplete, expert input is required to answer the questions accurately. Furthermore, the experiment confirmed that chatbots are good at answering specific assignments instead of questions. Lastly, chatbots perform better when multiple consecutive prompts are given with clear instructions.

The experiment confirmed that specially trained chatbots (DemoBot Friese Aanpak), or those with access to specialised content only (IPLO AI), generally provide more accurate answers than general-purpose chatbots. We discern three features for chatbots that are crucial for getting good answers.

Firstly, restrict your sources to the sources that are needed to answer most of the questions. In the experiment we noticed that this was an important factor why the IPLO AI and DemoBot Friese Aanpak performed better. However, none of the chatbots was able to answer questions about the interpretation of EIA-categories because this information was not in the datasets that they relied upon. We recommend to provide access to that information. It would also be valuable to include previous helpdesk responses in the dataset. For this, barriers regarding the privacy have to be overcome.

Secondly, it is important to train the chatbot on the difference between current and former (outdated) legislation. Many incorrect answers were due to the chatbot using outdated laws or regulations.

Thirdly, we believe that the ability to have a conversation by asking follow-up questions with the chatbot is needed for the chatbot to provide a complete and correct answer.

Coming back to the main question: we believe that specially trained chatbots help to provide better information on EA. It helps experts to formulate more accessible helpdesk answers and webpages and to discover knowledge gaps and improvements.

Sources

Eurostat, 2025, Use of artificial intelligence in enterprises - Statistics Explained (accessed February 19, 2025)

Halleem, Javaid & Singh, 2023, An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges

Hao, 2019, Training a single AI model can emit as much carbon as five cars in their lifetimes

Interviews and target group discussion, 2025

IPLO, 2024, Presentation in September 2024 on the Experiment AI for helpdesk questions (in Dutch: Experiment AI vraagbaak)

Nishant, Kennedy & Corbett, 2020, Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda

Van Noordt & Misuraca (2022), Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union