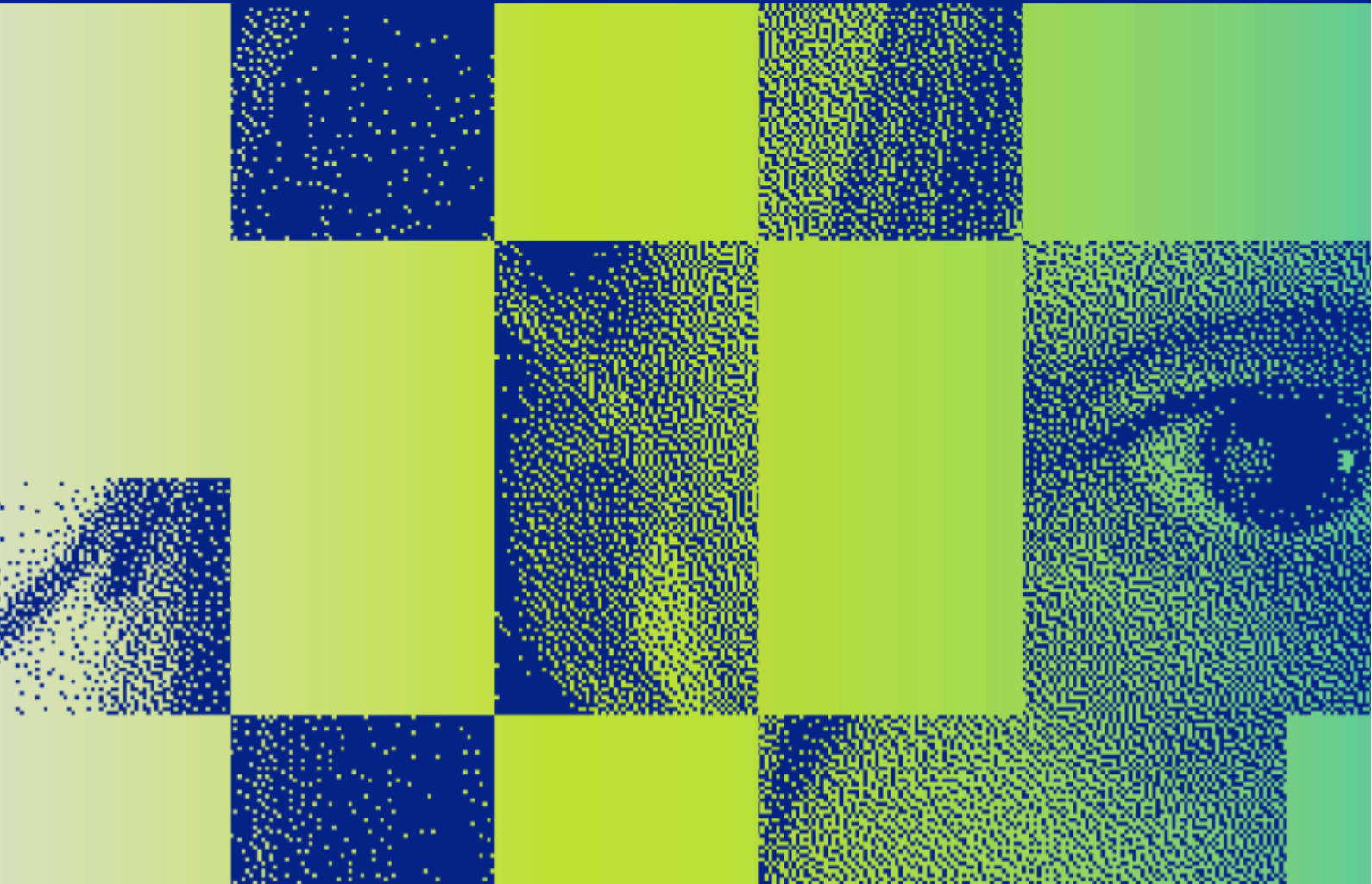




Generative AI tools for data discovery and use: Reflections from a data prompting workshop

August 2024

ODI Research
ADVANCING TRUST IN DATA



Contents

[Introduction](#)

[What did people think about the generative AI tools and how did they use them?](#)

[Data discovery](#)

[Data analysis, visualisation, and engineering](#)

[What did we learn about using these generative AI tools in the workplace?](#)

[Generative AI tools will change the way we work](#)

[Prompt Engineering is a new skill that people need to learn](#)

[Foundational literacy is useful when using generative AI models in day-to-day work](#)

[Prompting guidance for 'blindfolded assistants'](#)

[Appendix: Designing the workshop](#)

About

This report was researched and produced by the Open Data Institute, as part of our partnership with Microsoft, and published in August 2024. It was written by Joe Massey, Neil Majithia, and Elena Simperl. If you want to share feedback by email or want to get in touch, contact the ODI Research team at research@theodi.org.

We would like to thank our co-facilitators of the workshop, Johanna Walker (KCL), Georgia Panagiotidou (KCL) and Advait Sarkar (Microsoft Research) for their support in organising and running the workshop. We'd also like to thank the participants for their energy and enthusiasm during the day.

Introduction

Ensuring that people from all walks of life are able to access and use data effectively is critical for empowering people to make better decisions and creating equitable outcomes for society. However, [access to data is not always equitable](#), meaning that society cannot always realise the true value of data.

The recent emergence of generative AI models could have the potential to enable more people to access and use data. Generative AI models have the potential to facilitate easier access to data, as well as helping with data comprehension, analysis, and decision-making for individuals who may lack technical expertise. But the practical application and characteristics of these tools are not widely documented or understood.

We set out to explore how generative AI tools can be used in practice to better understand how people could use them in their day to day work. We organised a one day workshop to test how free, browser-based generative AI chatbots could be used for data discovery, analysis, engineering and visualisation, inspired by [a prior workshop on data discovery](#). The participants came from charities, civil society and the civil service, and work with data in their day to day tasks. The main goal for the workshop participants was to identify the potential of generative AI tools to help them where their technical confidence was lacking, or alternatively, to aid their colleagues' data literacy in order to increase the technical capacity of their organisation.

During interactive sessions, participants discussed their hopes and fears for the advent of generative AI tools in the workplace. They spent time exploring how some of these tools can support their everyday data activities through tackling a set of tasks. This workshop utilised the free versions of Google's [Gemini](#) and Microsoft's [Copilot](#), such that the participants could use the tools and their learnings from the workshop afterwards, regardless of their propensity to pay for a subscription version. The participants of the workshop described these tools as 'blindfolded' assistants, because they were not grounded with the datasets they were working with, in other words, they could not 'see' the data.

Overall, the participants found the tools used in the workshop to be useful for a subset of the tasks, and for certain users. They recognised the need to have a 'human in the loop' when working with these tools, to ensure the validity and

reliability of the outputs. The participants agreed that it may not be suitable to use these tools without the relevant expertise or a critical eye. There was a desire to work closely with these tools in the future, while acknowledging that their ways of working would need to change to maximise the benefits of these tools. In particular, prompting was a key skill that those using these tools will need to master.

This report details the learnings from each of the activities from the day-long workshop before exploring some more general reflections from the participants and facilitator team about how these tools could be used in the workplace. The report closes with some guidance on prompting when working with 'blindfolded' assistants. Further information on the design of the workshop is available in the [appendix](#).

What did people think about the generative AI tools and how did they use them?

Participants were excited to see how AI could be used in their workflows, to take away technical difficulty from their work and allow them to focus on the tasks they were more comfortable doing.

During the initial open discussion session, the participants demonstrated a wide-ranging interest in generative AI, in particular with regard to its potential use at the organisations they worked in. The participants were curious about the differences in the types of tools available, and their unique architectures. There were in depth discussions about the differences between generative AI models which are connected to the internet, as well as those grounded on specific datasets via Retrieval Augmented Generation. There were also concerns raised regarding accuracy, provenance, bias, and safety of chatbot responses, particularly when thinking about using these tools in the workplace.

After the session, participants noted that they enjoyed learning about and working with the generative AI tools but nonetheless found the experience “sometimes frustrating”. Some noted the tools were harder to use than they expected, with one finding it difficult to get “the kind of useful and reliable information I was hoping for”. Participants felt there were limitations in some contexts, finding them more useful as “a learning tool” and appreciating that they might have performed better as knowledge partners for the data analysis tasks had the participants had greater understanding of the subject.

Data discovery

This session aimed to help participants use a variety of conversational generative AI tools to discover and access open datasets. The participants were split into three groups, and each participant aimed to find and explore a dataset which they might find useful in their work, with facilitators providing prompting guidance halfway through the session. Participants generally utilised the generative AI models as you might use a search engine, but taking a more conversational tone, as well as asking follow up questions and clarifications.

Participants generally had a neutral attitude towards using the generative AI tools for data discovery. A subset of the participants found that the AI tools were particularly useful for finding datasets in usually hard-to-navigate data portals. Although they noted that chatbot responses could quickly get them to the web pages of useful datasets, participants felt a need to critique results and check links more so than with traditional search. A participant found “hallucinations read like fact”, putting them at risk of being misinformed.

Data analysis, visualisation, and engineering

For this session, participants were split into three groups, rotating through a set of exercises for each of the three activities. For each activity the participants worked with an adapted version of the [Kaggle International Football results dataset](#), which included 45,337 results of international football matches between 1900 and 2024. Using this dataset, the participants worked through a set of exercises which ranged from straightforward tasks, such as finding the average score of all matches, to more complex, open-ended tasks such as exploring the impact of home advantage on results.

While some generative AI tools allow users to upload datasets to analyse, such as Excel Copilot or ChatGPT 4, the tools used in the workshop did not have that functionality. Some of the workshop participants described Gemini and Copilot as ‘blindfolded’, due to the limited capabilities of the free models meaning that they could not read the data that the participants were working with. Therefore users conducted the tasks using tools they were already familiar with - usually the standard version of Excel (without Copilot), but in some cases different coding languages, using the generative AI tools as assistants providing step by step guidance through the analysis. The participants tended to use both tools in split screen, with the generative AI tool on one side of the screen and the tool of choice with the dataset open to carry out the task on the other. There were various approaches to using this side by side method, including asking for generic advice for how to carry out specific tasks, as well as trying to explain the shape of the dataset to the generative AI tool first. In the browser window, participants would take assistance from an AI tool, often lifting Excel formulas from the chatbots’ responses and pasting them directly into their own spreadsheet.

Participants on the whole had a mixed response to using the AI tools for data analysis, visualisation, and engineering. Across the activities, they found the tools to be useful as assistants to complete tasks that the participant knew were possible but didn’t know the method for, particularly appreciating that chatbots could explain the components of formulae or visualisations that were

provided in their responses. However, formulae were not always correct, meaning that participants with moderate technical knowledge felt they had to think critically and evaluate chatbot responses to really get the best of the interactions. The participants found the step-by-step nature of chatbot responses helpful and satisfying for all of the tasks. Finally, for those who could code, the ability of the tools to generate code was seen as useful, although some felt they could do it just as fast themselves (given the chatbot's response code would often require a small amount of debugging or refactoring).

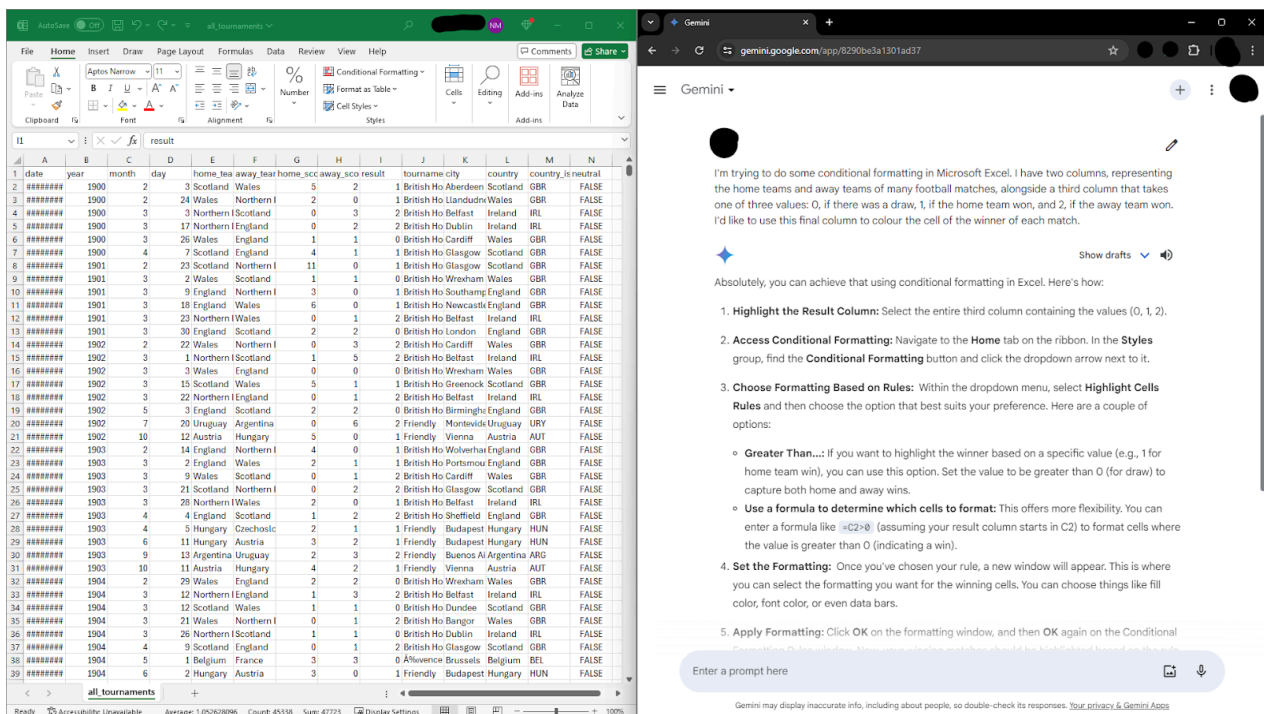


Figure 1. Screenshot of how participants used Excel and Gemini side by side

Using Excel Copilot

During the workshop several participants were able to test Microsoft Excel Copilot for data analysis and visualisation tasks. Excel Copilot functions differently to the other tools tested during the workshop, as it has direct access to the data that the user is working with in Excel. Users interact with Copilot via a chat box on the side of the window, which supports users to carry out the standard Excel functions such as creating formulas, graphs and more.

The participants who tested Excel Copilot found the integration of Copilot to be superior to using the 'blindfolded' models previously used in the workshop. They appreciated that it 'understood' which data they were working with, and could therefore suggest formulas and tables which immediately provided the right answers. Users appreciated that it suggested formulas, rather than providing the answers themselves, limiting the chances of getting the wrong output, as it was possible to check back on the workings of the formula.

However, users still found difficulty in performing some tasks with Excel Copilot, like iteratively and conversationally refining a formula (because Copilot is designed only to add column formulas, not overwrite existing ones), and reliable repetition of the same task (due to the probabilistic nature of the model). They also found that it was slower at producing answers than the 'blindfolded' models, and felt that it had a worse 'memory' regarding how it remembered previous prompts. These challenges highlight the importance of skills training for generative AI productivity tools.

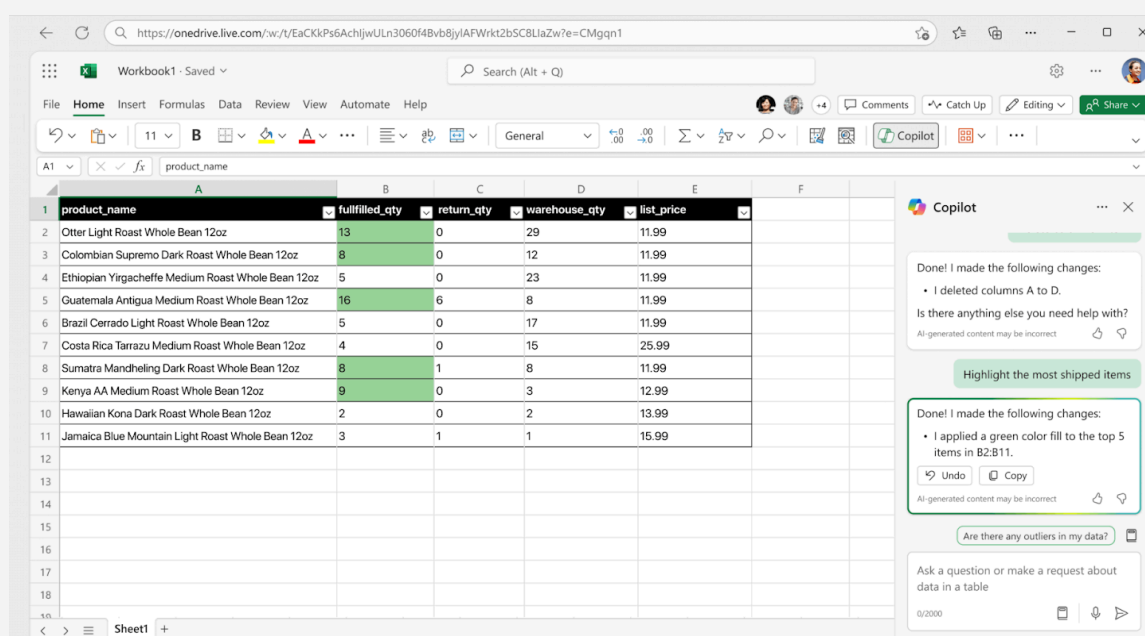


Figure 2. Example of Microsoft Excel Copilot from [Microsoft](https://www.microsoft.com/en-gb/microsoft/365/excel-copilot).

What did we learn about using these generative AI tools in the workplace?

Beyond learnings from the specific tasks, there were also some overarching takeaways which emerged from the workshop. The following sections detail some of these takeaways, and presents some ideas about how participants could see themselves using these tools in their work, and how they could impact ways of working.

Generative AI tools will change the way we work

Something which became apparent over the course of the workshop was the way in which participants could foresee how their workflows might change in a future where working with AI tools was the norm.

The participants deemed best practice to involve having a “human-in-the-loop” when working with generative AI, such that the benefits of these tools are maximised while some of the risks are mitigated. Participants could see the potential benefits to efficiency, problem solving, and idea generation through using these tools. However, they recognised there needed to be a level of human expertise involved in the process, whether that was to verify links shared as part of data discovery, or to test out the formula proposed by the AI tools. In this regard, a critical eye and a level of familiarity with the subject matter, be that data analysis or the specific dataset you are looking for, was deemed important when working with these tools.

The group noted that the requirement for user oversight when using the tools in this specific use case reduced their trust in the answers that the tools provided. This meant searching for datasets, or data analysis techniques that the participants didn't know much about was difficult, because they felt they couldn't provide the oversight required and could therefore be at risk of being misinformed. Participants noted that chatbot responses providing sources was useful for the sake of transparency and trust.

Some participants also began to work with the chatbots in new ways, beyond asking for specific prompts to support answering the task sheet. For example, asking about what a certain concept in a task was, such as a ‘what is a correlation coefficient?’. Using the tools in this way sparked some

conversations about other uses for these tools, for example to learn a specific skill. Participants noted that using AI chatbots was entirely different to the way they usually learnt technical skills (participants typically used YouTube videos, Reddit and other forums, or websites they found via internet searches). They found chatbots more accommodating and accessible given they could ask a question and get an answer immediately, with both questions and answers personalised to the complexity they needed. Similarly, the ability to go back and forth in a conversational manner matched the way that many tackled problem solving using a trial and error methodology. For some of the participants, this was a real potential benefit, and they looked forward to exploring these tools to support new learning and problem solving in their day to day work.

Prompt Engineering is a new skill that people need to learn

Overall, it took the participants some time to figure out how to best use the chatbots to maximise their utility. Utilising the prompting strategies which were shared in the prompting guide helped, but it took time for participants to use the models and get more familiar with them, building their mental model of best practices for interaction and the capabilities of the chatbots in response.

Most participants grew more comfortable with using the chatbots over the course of the day, particularly enjoying the conversational element of them and finding their experience to be “more efficient human interaction” compared to the usual tools in their workflow. Discussion suggested that a component of the chatbot’s strength here is the fact that prompts are conversational in tone, letting users feel comfortable speaking in full sentences, in turn allowing them to pose their requests more clearly in their prompts. Better prompts led to better responses from the chatbots, with clearer requests being easier to parse and use for the tools.

Participants understood the importance of prompt engineering and recognised that better prompts leads to more success, but noted that it is not always intuitive and requires practice to improve. Thinking logically or mathematically through prompts helped participants, as did reading prompts back to themselves to see if they made sense and conveyed the right meanings. Asking “Are you sure? or Did you mean that?” helped in some situations.

Participants were on the whole keen to use the workshop as a starting point to develop their skills with generative AI, especially with regards to prompting. Participants noted optimising prompts was a skill “vital for use” of the chatbots and therefore governed adoption. It was clear that the chatbots could help them in their day-to-day work or upskilling, but they wanted consistency and ease of use, both of which come from good, repeated prompting. A quote from

one of the team at Microsoft summed up this sentiment around learning prompting as a new skill perfectly: “we all need to train our prompting muscles”.

Foundational literacy is useful when using generative AI models in day-to-day work

Participants reflected that they didn’t expect to have to be so hands-on when using the chatbots given their perceptions of the potential of generative AI tools. Over the course of the day, participants were working hard to prompt correctly, verify chatbot responses, and implement the answers to their queries. Those with more limited technical literacy found all three of those components difficult, resulting in little success in dataset exploration and analysis.

Participants who were unsure about how to tackle particular tasks were also unsure about the outcome they were looking for and what to prompt, which led to worse outcomes than those with greater understanding of the tasks. One participant described this by saying “you don’t know what you don’t know”. The use of the generative AI tools did not necessarily lead to increased self-confidence in ability to answer the questions, instead simply presenting new tools that users might have the “same old problems” with. One group of participants co-created the following statement: “(A lack of) confidence in yourself leads to (a lack of) confidence in AI responses”.

Participants felt that the tools used in the workshop were best placed to enhance the skills of those using them, rather than being a direct substitute for them in their current forms. Participants found that the requirement for a “human-in-the-loop” was something of a best practice for working with generative AI tools, to realise the benefits of the tools while mitigating potential issues. Participants felt that if they were to adopt the use of these tools in the workplace, they would have to be responsible and provide both subject matter expertise and a critical eye over how it was used.

Prompting guidance for 'blindfolded assistants'

As part of the workshop, we developed a prompting guide which brought together insights from Microsoft and beyond about techniques for prompting. Many of these guides are often based on prompting in general, and we'd like to use the findings from this workshop to propose a guide on how to prompt a 'blindfolded' assistant to achieve the best results.

When using a 'blindfolded' assistant, a user should keep its functionalities at the forefront of their mind. Understanding what the chatbot is able to do (and what it cannot do) enables the user to build successful interactions in which the requests formulated into prompts are realistic and responses are therefore more accurate and useful. To build this understanding, a user doesn't necessarily need to know the actual architecture of the AI, although it can help. Instead, users can find the strengths and weaknesses of the tools by simply using them and gaining familiarity.

To remember these capabilities, a user can create a mental model or analogy that allows them to conceptualise how to communicate with an AI chatbot, much like participants in the workshop. For example, treating chat agents like ChatGPT as an "Excel expert in the other room" provides the user a framing tool regarding what information they need to provide in their prompts. Just as the Excel expert in another room would not be able to see the data in a spreadsheet, neither can the chat agent. Therefore, users need to spell out exactly what they're working with, whether it be column names, data types, or even the version of Excel they are using. Providing context is one of the most important parts of successful prompting, and to maximise the utility of these generalist chat agents like Gemini for data tasks, providing strong contextual information is critical. Setting the context and providing enough information to the AI tool is hugely important when the tool cannot 'see' the data you are asking it to analyse.

We also found that it is useful to reflect on the skillset or expertise that a user has before selecting a strategy, because while prompting strategies remain the same, the way in which users approach these tools may differ. For example, if a user has higher data literacy they can be more abstract or generalised in their prompting, asking about specific Excel formulae functions, such as pivot tables, vlookups, or if statements, and then applying the learnings to their work. Whereas someone with less confidence in their technical skills may need to provide more contextual information about the data they are working with, and the question they are seeking to answer to ensure that the outputs from the chat agents provide the responses they need, while limiting the

potential for erroneous responses.

As part of the workshop, we developed the following data prompting guide, which can be used in conjunction with the above guidance on 'blindfolded' assistants, as well as other chat agents.

1. Be specific in what you need, and include the right prompt ingredients
 - a. Task. What kind of response do you want from the generative AI model, what are you asking it to do?
 - b. Context. Why do you need it and who is involved? ‘
 - c. Expectations. How should the chatbot respond to best meet your expectations? e.g ‘Please respond in simple language’, ‘Please provide me with instructions for Excel’, ‘Please provide your sources’.
2. Ask the generative AI tool to ask you questions so that it gets all the information it needs
 - a. You can ask ‘What more do you need from me?’, to ensure you’ve provided all the important information.
3. Iterate - keep the conversation going. Following up on your prompts help you to gain more useful, tailored responses.
 - a. The different types of requests can be employed iteratively, in order to build a better picture of the challenge.
 - b. It is common for users to request generative AI chatbots to supply a draft of a text, such as a letter or report, which they then edit and improve. Similarly, the results of data activities can be used as a basis for further search, by continued prompting or additional search via traditional means.
 - c. The generative AI chatbot responds to the prompts of the users but is fundamentally passive. Building towards the goal should include explicitly reminding the generative AI chatbot of the task and previous interactions in the conversation - eg, ‘earlier, you suggested creating a graph...’.
4. Helpful hints
 - a. Know the tools limitations. These tools cannot do everything, effective prompting is important, so it’s better to give lots of details.
 - b. Communicate clearly. Pay attention to punctuation, grammar, and capitalization. Use natural language and prompt the

chatbot as you would speak to a human.

- c. Exploit the multi-modal abilities of generative AI chatbots. For instance, instead of asking for a visualisation, ask for code to create a visualisation. An incorrect visualisation with no explanation of how it was created is harder to debug than incorrect code.
- d. Use personas. It can be effective to ask the generative AI tool to take on a specific role, which can help to answer a question from different angles.
- e. Try to re-ask the question, generative AI tools are non-deterministic meaning that they can produce different outputs even when given the same input, so re-asking a question will provide a different answer with each ask (they also sometimes show multiple drafts).
- f. Don't be afraid to play and experiment!

For additional resources, please feel free to explore:

- [Microsoft 365 Copilot: The art and science of prompting](#)
- [Microsoft Copilot Prompts](#)
- [Prompt Engineering Guide](#)
- [Learn Prompting: Prompt Engineering Guide](#)
- [DataCamp: What is Prompt Engineering? A Detailed Guide For 2024](#)

Appendix: Designing the workshop

On the 29th of February 2024, we hosted a one day, in person workshop in central London, in collaboration with Microsoft Research and King's College London. Three weeks prior to the workshop we posted an [open call on the ODI website](#). The open call asked for expressions of interest from those working with data in their day to day role (regardless of whether they had a background in data analytics and/or data science), and had an interest in discovering how generative AI could help in their professional capacity. The participants were asked to fill out a survey before and after the workshop, detailing their prior experience working with data and with generative AI tools, and their perceptions of the tools after using them for multiple activities.

The workshop hosted 15 participants from a variety of organisations including charities, not for profits, and departments within the civil service. Within their organisations, their roles varied, with some working as analysts and others as managers. Half of the participants had tested one or more generative AI tools before the workshop, with two using them occasionally and three regularly. Only one participant had never tried generative AI tools before.

During the workshop the participants primarily utilised two different tools, Google's [Gemini](#) and Microsoft's [Copilot](#). These tools were deliberately selected due to them being free, enabling anyone to access and use them. As the workshop participants mostly came from the charity, civil society and government departments, where funding may not always be readily available to pay for licensed generative AI tools, the free versions of these tools were chosen to ensure that participants could take the learnings from the workshop back to their workplace. During the workshop several participants were also able to test Microsoft Excel Copilot thanks to the Microsoft team, while one of the participants also had a paid membership for ChatGPT 4. This meant that some participants on that table were able to explore how the paid tools worked in comparison to the free versions of the tools utilised for the majority of the workshop.

More than three quarters of the participants had experience with data discovery, although most identified as beginners knowing only basic methods. In their personal and professional experience, most participants used search engines to find datasets, although others had used data portals, recommendations from colleagues and Freedom of Information requests. Three participants had not previously searched for datasets before.

The majority of the participants had experience with data analysis, however many of these only knew about basic methodologies. A majority of the

participants used spreadsheets to analyse datasets, with a smaller group using programming languages and integrated data systems like PowerBI. Participants' experience with data visualisation varied but almost all of them identified themselves as beginners with basic or no knowledge. None had used bespoke data visualisation software, with most using spreadsheets, integrated data systems, or programming. For data engineering, all participants either had little to no experience or some beginner experience.