

4<sup>th</sup> September 2023

# Public perceptions of foundation models

---

Report prepared for the Centre for Data Ethics  
and Innovation (CDEI)

## 1. Our approach

**We conducted 6x 2-hour online focus groups with members of the public, between Wednesday 31<sup>st</sup> May and Monday 5<sup>th</sup> June 2023, with c. 7 participants per group.**

*Sample specifications required that participants were:*

Group	Audience
1	General public, 18-39, ABC1
2	General public, 18-39, C2DE
3	General public, 40+, ABC1
4	General public, 40+, C2DE
5	Early adopters / lay experts – all to say they have used LLMs and feel confident in explaining how they work
6	Public sector workers – including healthcare professionals and education staff

Within each group there was a mix of gender, ethnicity and UK regions represented.

## 2. Key findings

### 1. Participants are open towards the use of foundation models within the public sector.

- They feel that the benefits of speed and the ability to analyse complex data would add value to existing ways of working.

### 2. Participants see foundation models having the biggest positive impact in use cases which will:

- **Help the general public, rather than 'just' helping public sector workers.**
  - For example, amongst the use cases tested, foundation models are felt to be most beneficial in healthcare research and development where they could possibly help to cure disease, therefore benefitting all of society.
  - On the other hand, participants feel that internal government use cases (e.g. drafting ministerial documents) would have a smaller positive direct impact on the public. This is because the benefits of a more efficient public sector (particularly in non-public facing roles) are not felt to be relevant to them.

- **Have a tangible, rather than abstract, benefit.**
  - “Efficiency” is recognised as a benefit of foundation models, however, it can feel abstract and difficult to grasp. Instead participants want to know what efficiency means in practice.
  - Use cases which are able to communicate what can be achieved with increased efficiency are more compelling e.g. developing medication more quickly.

**3. Accuracy is the biggest concern. Participants will only support the use of foundation models if they reliably produce accurate outputs.**

- They feel that the benefits of faster decision making or increased productivity will not be realised if the models are unreliable and outputs require greater review from humans.

**4. Participants are most comfortable with use cases where they feel potential inaccuracy poses the least risk. These are use cases where:**

- The outputs or decision making will have a smaller direct impact on individuals. For example, participants are more comfortable using foundation models for assisting with policymaking (where inaccurate decisions only indirectly affect members of the public), rather than assisting doctors (where inaccurate decisions would directly cause detriment to specific individuals).
- There is human oversight over the outputs to review and catch inaccurate outputs before further decisions are made or outputs are created.

**5. Participants see the lack of emotional intelligence as a limitation of foundation models, which means their use is only felt to be suitable in certain situations.**

- This includes ‘cold’, analytical tasks such as data analysis where many see technology already being used.
- Or, if use of foundation models corresponds with human interpretation of the situation. For example, participants feel that in some cases such as education and healthcare, the ‘correct’ decision can only be made with an emotional judgement too e.g. taking into account someone’s body language as well as their symptoms.

**6. Overall, participants want there to be human accountability over decisions and outputs taken from foundation models. They see models as assisting and augmenting human capability, rather than replacing it.**

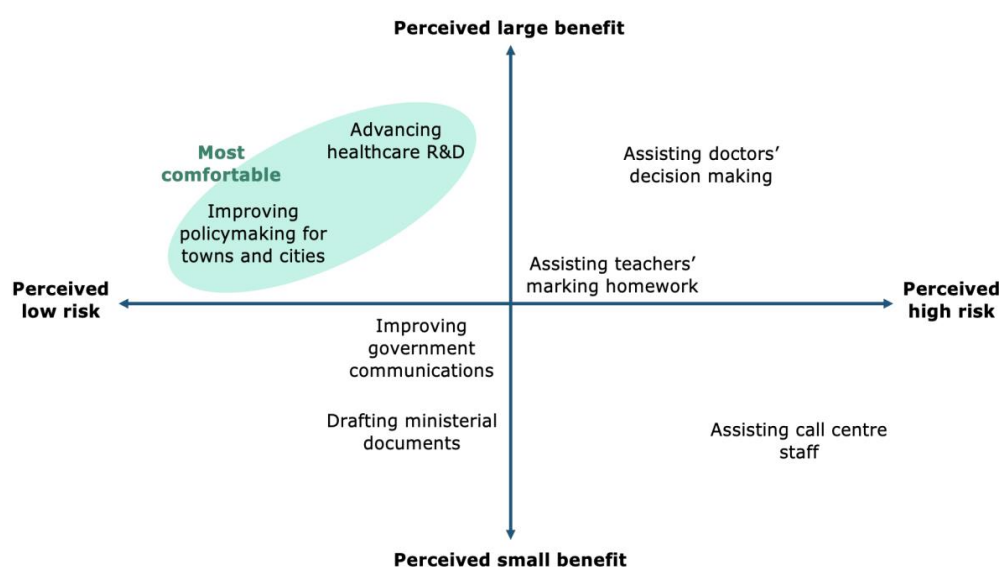
- They do not believe a foundation model can be accountable for its outputs, as it cannot explain why it produced them.
- Therefore, participants feel it is important that humans are accountable for decision making, so they can explain them if they are questioned. This is particularly true in the context where inaccuracy poses a risk.

## 7. Participants are concerned about the potential impact of foundation models on the labour market.

- This includes, to a greater extent, concern that humans risk becoming less skilled, and to a lesser extent, fear that certain jobs or industries will become redundant altogether.
- However, public sector workers feel less concerned about these risks, due to:
  - a perception that there is a shortage of public sector workers, particularly in healthcare and education, meaning their jobs feel necessary and less 'at risk';
  - feeling considerable pressure in their job, meaning they are more receptive to the benefits of time saving and less concerned by the widening scope of technology.

## 8. Advancing healthcare research and development and improving policymaking for towns and cities are the use cases which participants are most comfortable with. (See diagram below)

- This is due to a trade-off between having most potential for public good and being perceived to involve the lowest risk.



### 3. Recommendations based on this research

**The findings of the research point to several considerations for using and communicating about foundation models in the public sector:**

#### **Use and governance of foundation models**

- **Start with perceived 'low risk' use cases, which feel furthest away from directly impacting negatively on individuals.**
  - Participants feel more comfortable with use cases where the potential risk and impact of harm is perceived to be smaller, e.g., assisting policymaking for local areas and healthcare research and development.
  - For those worried about privacy, the 'population-wide' or non-personal data input in these cases also carries less of a privacy concern.
  
- **Focus on use cases where the accuracy of foundation models' outputs can be guaranteed (or at least, where inaccuracy can be mitigated).**
  - Inaccuracy is felt to be the risk of using foundation models which could be most damaging in particular use cases. In general, participants also feel that the benefits (e.g., quickly producing summaries) can be quickly undone if models are inaccurate and outputs require greater review from humans.
  
- **Ensure clear channels of human accountability for foundation model supported decisions, particularly in cases which directly affect members of the public.**
  - This will allay concerns about foundation models' accuracy (mentioned above), as well as aiding with the problem of models' lack of explainability, and 'computers taking over' type fears.

#### **Communicating about foundation models**

- **Focus on the capability of foundation models to enhance human capability as a rationale for their introduction into the public sector.**
  - Participants feel most comfortable with this way of using foundation models as it implies that humans will be overseeing and taking accountability for decision-making.
  - In addition, it can help allay concerns that foundation models are being introduced to cut jobs.

- **Tied to the above, emphasise there will be human accountability for decision making.**
  - Participants feel uncomfortable with the idea of a machine making decisions which will affect their lives. Emphasise that the ultimate decision making will rest with humans.
- **Don't say that there are no risks to using foundation models.**
  - This will not feel credible. In particular, participants feel that the risk of inaccurate or low quality outputs presents a real risk which needs to be mitigated.

#### **4. Starting awareness and perceptions of foundation models**

**The majority of participants are not aware of the term 'foundation models', however awareness of specific models (e.g. ChatGPT) is much higher.**

- With the exception of a handful in the lay experts group, there was no awareness of the term 'foundation models'.
- Participants across the groups are aware of specific foundation models, mostly ChatGPT and to a lesser extent DALL-E. Awareness is much higher amongst ABC1 participants, public sector workers and (unsurprisingly) the lay experts.
- Most had heard about specific models via the media, social media (particularly TikTok) and word of mouth.
- Where participants had used foundation models it was most often for 'a bit of fun', rather than anything more advanced e.g. to support their working life.
  - The exception to this are the lay experts who have used ChatGPT for more sophisticated purposes (e.g. supporting with coding at work; research to support university study).

*"I've heard of ChatGPT but not heard it being referred to as a 'foundation model'." 18-39, ABC1*

*"I've heard of ChatGPT because it's all over the media at the moment." Public sector worker (Education)*

*"I heard about it a few months ago; I saw it on TikTok for the first time." Early adopters / lay experts*

*"At the moment [ChatGPT] is like a new toy." 40+, C2DE*

**There is an openness to foundation models after hearing an explanation and seeing how they work in practice.**

- Participants are readily able to identify benefits with foundation models, unprompted. Most notably the speed of producing outputs and the ability to handle complex tasks (e.g. summarising complicated concepts) much better than a human can. In this context, there is an openness to using models in UK society.
- Furthermore, openness to foundation models – and artificial intelligence in general – is also built on a sense that advancement in technology is the ‘way the world is going’ and this is the next step.

*"Speed, if it is basic tasks, if an AI writes the minutes for something. That would be good. It would free someone up to do something more important."* 18-39, ABC1

*"It's in its early stages. When it progresses and is more perfected, it will eliminate human error."* Early adopters / lay experts

**That said, participants do see limitations to foundation models – particularly around accuracy and a lack of ‘human touch’ – which mean they feel their use should be restricted to certain situations.**

- Concerns about foundation models centre on two key factors:
  - **Accuracy / reliability of models:** Participants believe that models will only add value if they are accurate. Furthermore, they have the potential to do harm if they are inaccurate.
  - **Lack of emotional intelligence:** Participants feel that foundation models will be less useful in situations which require an understanding of an individual’s wider context or creativity. For example, some draw comparisons to chat bot features on websites, which fail to take into account wider contextual factors (e.g. the ‘journey’ the person has been on so far).
- As a result, participants feel that foundation models should be used as a tool or an assistant to work delivered by human beings. Importantly, humans should review and have accountability for decisions.

*"[Inaccuracy] is a really big deal. It will add to the spread of fake news and fake information."* 18-39, ABC1

*"They don't have emotional intelligence. They can't empathise. You can't expect to replace a mother and a mother's love with a language model."* Early adopters / lay experts

**Beyond this, there are concerns about the ‘slippery slope’ into over reliance on technology.**

- As well as acknowledging the benefits and limitations of foundation models, participants across the groups described them as ‘scary’.
- There is some concern about what the existence of foundation models will mean for jobs and human skills. Some participants also express concern

that foundation models could be used for malicious intent e.g. cheating on tests.

*"It might limit how much we do ourselves. If I was a student I could see opportunities to cheat very easily."* 40+, ABC1

*"Should artists be concerned? Should poets? It's interesting – but also a bit scary!"* 40+, C2DE

## 5. Contextual views on the use cases

**The public are open to using foundation models in the public sector. Overall, there is little fundamental opposition to any of the use cases tested (providing the models are reliable).**

- The public are able to identify important benefits to all use cases tested and feel they will – to varying extents – have a valuable place in the public sector.
- Models being unreliable or low quality (e.g. by producing inaccurate outputs) is the biggest concern:
  - The public feel the benefits will be lessened or even nullified completely if the models are not producing accurate outputs.
  - In some cases, the public feel that the risk of harm from inaccuracy is higher (explored further below).
- Assuming that the foundation models are accurate/reliable, other broader concerns about their scope are surmountable.
  - For example, the public would feel reassured if the use cases changed in scope (e.g. restricting their usage to certain situations) and/or were accompanied by governance (e.g. human oversight).

*"Echoing what everyone else has said, [I'd be happy] as long as it is tried and tested as much as it can be and the information it produces is accurate."* 40+, C2DE

### **Perceived benefits and where the public see most potential for use cases:**

**Consistent with views of foundation models in general, speed and handling complexity are perceived to be the key benefits of the use cases.**

- In particular, the public are receptive to use cases which are seen to increase the productivity and efficiency of completing tasks. For example, requiring fewer experiments than a human to facilitate healthcare research and development.
  - The public also perceive that increased speed and efficiency will mean workers will have more time to do other tasks which may require more emotional intelligence, e.g. spending more time with students on the homework example.



*"There are lots of positives in [the healthcare R&D use case]. Reduction in timescales, the ability to pool lots of information, fewer experiments..."* 40+, ABC1

*"Go for it! Then I could spend so much more time with patients, talking to them."* Public sector worker (Healthcare)

**Putting aside their concerns (explored below) participants see the most potential for use cases which will benefit the public, rather than public sector workers.**

- Amongst use cases tested, the public have greatest support for using foundation models for improving healthcare research and development as it has potential to cure disease – something which is of high importance for the public.
- On the other hand, the public see the least potential in using foundation models to support with drafting ministerial documents and government communications. Both of these use cases feel less relevant to the public overall and therefore the perceived impact of their benefits is reduced (i.e. only public sector workers will feel the benefits).
- Equally, in the use case of foundation models assisting call centre workers, key benefits are perceived to be saving employees time and helping them to make efficient decisions. However, participants did not find these compelling if they would also mean worse customer service or the possibility of wrong decisions for social security claimants.

**In addition, participants feel use cases with a tangible benefit are more compelling than those which are more abstract.**

- Participants agree that efficiency and saving time are benefits of foundation models, but they are more compelling when the use case highlights what is achieved with improved efficiency or time saving.
  - For example, allowing doctors to spend more time with patients or having access to new medication sooner.
- On the other hand, use cases are less compelling when the benefit of 'efficiency' is in the abstract.
  - For example, whilst participants recognise that using foundation models to draft ministerial documents will save civil servants time, it is less clear what this means in practice e.g. what will they be able to do instead?

**Perceived concerns and overall comfort with the use cases:**

**Consistent with views of foundation models in general, the potential for inaccurate outputs is seen as the main risk of the use cases.**

- The public believe that faster decision making and/or greater productivity are of no use (and no quicker) if the foundation models make bad decisions which negatively impact on individuals or produce low quality outputs which will need to be heavily reviewed and edited by a human before being used.
  - For example, the benefit is nullified if it takes an individual just as long to review and check outputs from foundation models as it would for them to do the task themselves.
- This concern is reinforced by the lack of explainability in the event of wrong AI decisions.
  - If the models are producing inaccurate outputs, the public are concerned that users will not be able to explain why that is the case and, therefore, 'fix' the problem.

*"I don't see how I feed the piece of [homework] into the model. I don't know if in the time that I have to set it up and feed it the objectives and then review afterwards, whether I could have just done the marking myself?"* Public sector worker (Education)

### **The lack of emotional intelligence is also felt to be a limitation for foundation models.**

- In some cases, the public feel the 'human touch' is necessary to make good decisions. The public feel only humans are able to understand the wider context and emotional factors at play e.g., teacher relationship with student, doctor being able to see/feel a symptom rather than just describe it, dealing with a social security claim.

*"Imagine yourself on that call. You need the personal touch for difficult conversations."* 40+, ABC1

### **Participants do share other concerns when prompted, (e.g., lack of explainability, being resource intensive, and bias) however these are less likely to be raised spontaneously when discussing the use cases.**

- The lack of explainability in foundation models' decision making is not a concern for participants when the models are accurate, rather it is felt to simply compound the problem when models get things wrong.
- Participants are concerned about bias, particularly with respect to health outcomes and inequalities. However, this concern is often not as front of mind as others (e.g., inaccuracy).
  - Most participants believe that bias can be addressed by cleaning input data, ensuring it is as representative and accurate as possible, so that the output is accurate and fair as a result.
- The fact that foundation models are resource intensive was rarely mentioned, with the exception of healthcare research and development. A small number of participants worried that large pharmaceutical companies

would be able to access (and profit from) foundation models more easily than smaller companies or charities in the health sector would.

*"There needs to be accurate input information and it needs to be well-rounded; all the information inputted will ultimately make the decision."* 40+, C2DE

*"The only problem is... pharma companies would rinse it, while smaller companies may not be able to afford the computing infrastructure."* 40+, ABC1

**The public's concerns with the use cases affect their risk appetite. In line with their perceived benefits and drawbacks, the public feel most comfortable with use cases when:**

- **There will be a smaller *direct* impact on individuals** if/when a foundation model provides inaccurate or harmful information. On the other hand, the public feel more comfortable with a degree of inaccuracy where the use case feels further away from individuals. For example, they are more comfortable using foundation models for traffic planning / R&D rather than assisting call centre staff / doctors.
  - In addition, those with more knowledge of foundation models also feel the risk is lower – and therefore they feel more comfortable – with use cases that do not require some input of personal data for example, their homework or health record.
    - They feel that use cases which do not require a personal data input pose less of a privacy risk.
- **There is human oversight of decisions** as the public feel they are able to provide a layer of quality assurance over the outputs and mitigate some of the risks regarding inaccuracy.
  - Human quality assurance is particularly important in use cases which have a direct impact on individuals

*"I'm more comfortable with [foundation models] looking at mass amounts of data rather than taking on a more human role."* 18-39, ABC1

*"There needs to be a human safety net to filter in case anything goes wrong."*  
40+, C2DE

**In addition, participants want humans to be accountable for decision-making, rather than foundation models. As a result, the most comfortable and appropriate use for foundation models is to augment human capability, rather than to replace it completely.**

- Participants do not believe a foundation model can be accountable for its outputs. Instead, they see models as a 'tool' or an 'assistant' to support humans – who are able to be accountable for their actions.
- Therefore, they feel it is important for there to be human accountability for decision making – particularly if there's potential for inaccuracy.

- Participants want decisions to be explained if needed and for someone to be held accountable if things go wrong.

*"There needs to be a line that AI isn't allowed to cross. It is an assistance tool not a decision-making tool."* 40+, C2DE

### **Other wider contextual views on foundation model use cases:**

#### **Existential risk is not front of mind.**

- Although there is some awareness of these type of concerns, participants did not hold or spontaneously raise these concerns themselves.
  - For example, some participants mentioned recent media coverage about the researchers letter and had familiarity about extinction from Sci-Fi films but did hold these opinions themselves.

#### **However, participants across all ages and socio-economic grades – although typically not public sector workers – express concerns about the potential impact of foundation models on the labour market.**

- These concerns took two forms:
  - **(To a greater extent) that humans risk becoming less skilled:** For example, in the case of drafting ministerial documents and government communications, participants feel these are important skills for civil servants to learn and know how to do.
    - There is some optimism that foundation models could take the administrative burden off professionals (e.g., doctors, teachers, civil servants) to allow them more time for complex, interesting or emotionally intelligent work. However, this is often met with a worry that if professionals lose the skills over time to *do* the more 'boring' work themselves, they will become less able to think critically about a foundation model's outputs, feeding a cycle of greater reliance on AI.
  - **(To a lesser extent) making certain jobs redundant:** For example, in the case of assisting call centre staff, some participants assumed that the foundation model would reduce the need for a human altogether. There was also a worry that certain creative industries might suffer if AI can generate content a) very quickly and b) in the style of great artists (singers, writers, poets, graphic designers etc.).

#### **Overall, concerns about the labour market are not insurmountable.**

- To an extent, participants feel that the growing usage of technology is a given in the labour market.
  - Although there was minimal suggestion that foundation models would lead to the creation of new jobs outright, participants did

acknowledge that jobs may change due to increased involvement from the models, e.g., more jobs becoming 'data analyst'-type roles.

- In response to their concerns, participants would like to see safeguards to avoid the overreliance on foundation models to ensure human beings remain adequately skilled.

*"The call centre staff would become reliant on what the computer is telling them instead of becoming more skilled at their job and developing in their role. It could stop them wanting to learn because there is no need to."* 18-39, ABC1

*"A drawback would be that [using foundation models] would reduce availability of jobs in the areas of graphic design and copywriting, but that's just a risk that comes with technology."* 18-39, C2DE

*"The public sector has suffered so much with cutbacks; they could be making mistakes because they're doing too much. So automation with human oversight would be good."* 40+, C2DE

*"They've been making music sound like someone singing it. Deepfakes are really crazy. It's really going to kill some industries."* Early adopters / lay experts

#### **Public sector workers did not tend to hold these concerns:**

- A perception that there is a shortage of public sector workers (particularly in healthcare and education) means they do not fear being replaced by foundation models.
- Feeling considerable pressure from their job means they are more receptive to the benefits of time saving and less concerned by the widening scope of technology.
  - Some have already experienced technological 'shifts' at work and recognise these shifts have saved them time, which allows them to spend a greater amount of their time with their students/patients.

*"I work for NHS 111 and... it's already in place at a basic level; the computer is deciding whether you need to see a doctor or go to A&E. I'd see it as being beneficial."* Public sector worker (Healthcare)

#### **Participants want to see transparency about the data input and how this leads to the models' outputs:**

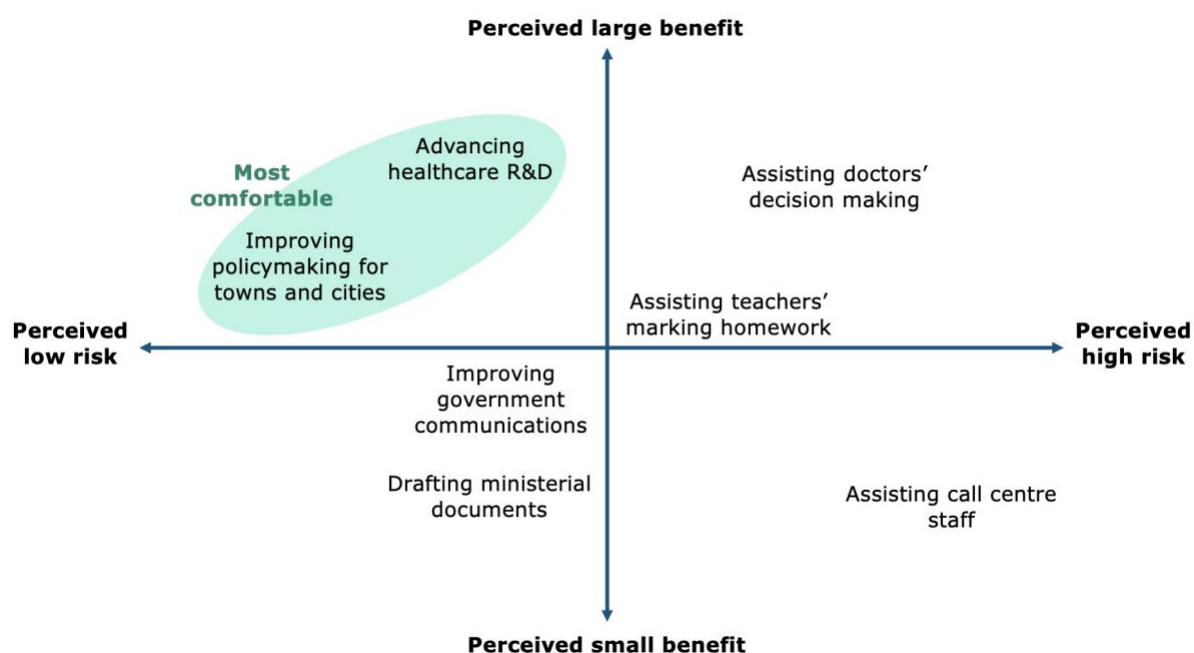
- This would help to address concerns about accuracy and reliability, as well as privacy.

*"[I want the Government to consider] transparency – we should be declaring where AI has been applied. And it's about where the information is coming from, ensuring it's as correct as it can be and mitigating bias as much as possible."*

40+, ABC1

## 6. Public perceptions of the use cases (in order of comfort)

**Diagram:** showing public perceptions on use cases according to the examples which they felt had the most potential to deliver benefits, and those which they saw as the riskiest.



Use case 5: Advancing healthcare research and development	
Perceived benefits	High
Perceived risks	Low
Relevance / personal impact	Medium

- **Participants see considerable potential for this use case:**
  - Participants see the benefit in foundation models speeding up the analysis process and finding new patterns in data which humans would not be able to spot.
  - The best case scenario (e.g. discovering the cure to a disease) is felt to have the highest rewards and biggest impact for the public out of all the use cases tested.
- **There is some latent awareness about using AI in healthcare development, which helps participants feel comfortable with this use case.**

- Using 'the latest technology' is already associated with modern science.
- One participant could recall a recent headline of AI being used to develop a medicine for one of the three main antibiotic-resistant 'superbugs', which served as a positive success story.
- **Participants are less concerned by inaccurate outputs in this use case.**
  - This is because they believe they would have a smaller direct negative impact on individuals. There is a sense that if an inaccurate output were produced it simply would not be used.
- **In addition, participants feel that the potential for foundation models to produce wrong information could be mitigated by human oversight / further testing.**
  - This also retains the need for scientists, whose role may change (e.g., to a data analyst or product testing role) but not be replaced, easing concerns of loss of jobs or skills.
- **It is less obvious to participants what this example's worst case scenario is, other than waste of resources.**

*"If it can find cures to diseases quicker, then that is something that is worth pursuing."* 18-39, ABC1

*"Medicine is science and it needs to keep moving forward, so this is probably quite positive."* 40+, C2DE

*"It would need that human element and the experts who have studied that field for years. It'd be a tool to use but I don't think it can completely take away the experts."* Public sector worker (Healthcare)

Use case 2: Improving policymaking for towns and cities	
Perceived benefits	Medium to high
Perceived risks	Low
Relevance / personal impact	Medium to low

- **Participants see this use case as enhancing human capability. The model is adding value by doing something a human couldn't (synthesising large amounts of population data).**
  - The nature of the task feels like a familiar use for technology (numerical and statistical, rather than emotional).
  - Another key benefit is speed and efficiency, in what is thought to be a bureaucratic discipline (urban planning).

- **The impact of a foundation model getting a decision wrong in this case is perceived to be less direct to the individual.**
  - Therefore, the worst case scenario is neither as obvious nor felt to be as bad as in the other use cases.
- **Participants like that policy makers retain decision making and execution, but are simply better informed.**
- **However, there are questions about the inputs to this model and the potential for privacy concerns:**
  - While deidentified geospatial data is received positively from a privacy point of view, there is some confusion about what data would be needed, what exactly “deidentified” means and how, if at all, this is different to data that is currently used.

*“I think in the last scenario you were dealing with one person calling, and the outcome of that call would have a big impact on that person. This is more generic and smaller errors are less likely to have a big impact; a small impact on traffic flow won’t make a massive difference.” 18-39, ABC1*

*“It would enable local authorities to be a bit more effective in how they create infrastructure moving forward.” 40+, C2DE*

*“This is what these models are good at... You don’t need an emotional side to it – it’s just raw data.” Early adopters / lay experts*

Use case 7: Assisting doctors’ decision making	
Perceived benefits	High
Perceived risks	Medium to high
Relevance / personal impact	High

- **Overall, participants respond positively to the idea of foundation models being used to generate a summary dashboard of patient history.**
  - There is an expectation that this could be used in both routine and emergency situations.
  - Some participants mention already seeing doctors using Google during appointments, making the idea of using a foundation model seem less new or different to the status quo.
  - Another key perceived benefit is the possibility of speeding up appointments.
- **However, participants worry about the impact of inaccurate information which could be directly harmful to an individual (e.g., if it recommended the wrong medication), and therefore feel that**



**suggesting treatments or diagnoses is a step too far for a foundation model.**

- Participants strongly feel that medical professionals must maintain accountability for this type of decision making.
- This sentiment is reinforced by participants not wanting medical practices to lose their 'human touch', for example:
  - The fact that a foundation model cannot physically see or feel someone's condition or symptoms
  - The fact that healthcare workers can draw on professional experience
  - Emotional intelligence is required to understand and adequately respond to patient needs
- **Healthcare workers tend to think in line with the wider public about this use case, seeing how it could positively affect their role, while cautious of giving too much authority to AI.**
  - One healthcare worker welcomed the idea that a dashboard could be used to draw together reports from different specialists and parts of the healthcare system, saying it would help them to perform their role more efficiently.
  - Another mentioned that doctors can make mistakes, especially since healthcare staff are often overworked, and so humans and foundation models could work together effectively to spot each other's inaccuracies and biases.
  - Some would like to see a trial period before full implementation, including an explanation of how patients' sensitive data, such as medical records, would be used by models in a secure way.

*"I want my doctor to be my doctor."* 40+, C2DE

*"The model should be trained and constantly looked after. Any kind of misdiagnosis and the risk is very high."* Early adopters / lay experts

*"I would be more confident with it summarising patient health history, rather than it making recommendations with medication."* Public sector worker (Healthcare)

*"I can see how it [would help] me in my role. I write primary discharges and I write it all up and I check it again, but if it can help pull it all together for me and then I can just check it, that would save me a whole lot of time and I can spend more time with patients."* Public sector worker (Healthcare)

Use case 6: Assisting teachers to mark homework	
Perceived benefits	Medium

Perceived risks	Medium
Relevance / personal impact	Medium

- The key perceived benefits of this use case are:
  - **Saving teachers time on administrative tasks**
    - This is perceived to be particularly useful for larger class sizes.
  - **Spotting plagiarism**
    - Anti-plagiarism software is already known to be in place in university settings, meaning participants are more comfortable with this as a use of foundation models.
    - Teachers say they are already able to recognise whether students use ChatGPT or overperform from using bots, easing others' concerns about over-reliance on the foundation models.
  - **Reducing teachers' bias in marking**
    - Participants note that a foundation model 'would not have favourites' in the same way as some teachers do.
- **However, participants feel the risks are considerable, particularly teachers over-relying on foundation models, and loss of 'human touch'.**
  - The teacher-student relationship is thought to be crucial to a child's development, and some fear that key elements of this could be lost with greater reliance on foundation models to mark work.
  - Similarly, over-reliance on AI could, over time, limit teachers' ability to mark homework effectively themselves.
  - Some feel that foundation models would be better for marking scientific subjects where there are 'right and wrong answers' in mark schemes, and fewer questions which require understanding of context or emotional intelligence.
- Some feel that using **foundation models could confuse who is accountable for students' grades**, further confusing any appeal processes.
- To mitigate these perceived risks, participants are **more comfortable when the teacher makes the final decision** as to a student's grade.
  - Participants are more open to the possibility of using foundation models to highlight key parts of homework to aid the teacher in their marking, rather than deciding grades outright. However, this use case still lends itself to concerns around loss of teacher's skills over time, and some also question how this would work in practice.

*"I like this one. I like the 'retaining the human decision' element at the end. And it must be hard to check for plagiarism in a class of 30." 40+, ABC1*

"This would be really helpful. A lot of teachers have favourites and the AI wouldn't." 40+, C2DE

"In order to be a good teacher, you need to know your children and have that connection with their parents and I don't see how AI can replace that." Public sector worker (Education)

"I don't see how I feed that piece of writing into the model. I don't know if by the time I have to set it up and feed the objectives into the model, I could have just done it myself?" Public sector worker (Education)

Use case 3: Improving Government communications	
Perceived benefits	Medium
Perceived risks	Medium
Relevance / personal impact	Medium (in case of localised information) to low

- **A key benefit is understood to be Government employees saving time:**
  - This could either be by producing multiple designs quickly for humans to then 'whittle down', or to more quickly tailor content to particular audiences, regions or demographics etc.
- **However, this case's popularity varies according to two sub-use cases.**
  - **The foundation models simply (re)formats or tailors information / content provided by a communications team:** Participants feel that this is an appropriate use of AI, speeding up the process by which pre-agreed content is shaped into a campaign.
  - **The foundation models 'generate' content themselves based on, e.g., previous campaign materials:** In this case, there is concern that a foundation model might create inaccurate or offensive campaign materials, due to a lack of understanding of context. This would have a negative knock-on effect in terms of time and resources, and there would also be potential for confusion around accountability for the materials produced.
- A minority are worried about a resulting **loss of jobs in the design industry** in both of the cases above.
- **Concerns about accuracy and/or manipulation of content in this use case can be mitigated through humans having final sign-off on materials.**
  - This would reassure participants, both in terms of accountability for materials published and quality control.

- **However, despite recognising benefits for Government employees, enthusiasm for this use case is more muted compared to others.**
  - This is because it is perceived to deliver limited public good (i.e. participants feel the benefits are only felt by employees and not the public). Namely, the benefits of a more efficient public sector are not felt to be directly beneficial to the public.

*"It's just a really quick way to get information out quickly and I don't really have any concerns about it because it is being reviewed by a human. If there's no human review then no, it could be offensive."* Public sector worker (Healthcare)

*"It's fine if a person has the final sign-off for it. It's fine for tailoring to use a robot rather than manpower. It saves time and means things can keep up to date."* 18-39, ABC1

*"If something really stereotypical came out that would be really bad."* Early adopters / lay experts

<b>Use case 4: Speeding up drafting ministerial documents</b>	
Perceived benefits	Medium to low
Perceived risks	Medium (higher when foundation models used for drafting briefings vs. lower when used for drafting standard contracts)
Relevance / personal impact	Low

- **This use case is seen to benefit only public sector workers, rather than the general public.**
  - While participants can see the appeal in having content quickly reformatted into a particular style to suit ministers' preferences, this is not felt to be as 'worthy' a use as in other use cases (e.g., medical research and development).
    - Furthermore, similar to the Government communications use case, the benefits are not felt to be relevant to the general public which lessens their perceived impact.
  - Other gains, such as saving time and resource, are perceived to trade off with concerns around skills and quality (see points below).
- **Drafting briefings is thought to require skill, leading to concerns around loss of skills if models are relied on too heavily.**
  - A briefing document may contain sensitive information which requires special consideration.
  - Equally, it may require a political stance (or at least knowledge of the political context and nuance behind the terminology used).

- Due to the skilled nature of drafting these documents, some participants fear that civil servants' ability to *review* briefing documents may decrease if they are no longer required to *write* briefings themselves.
- **Furthermore, participants believe that providing inaccurate content to ministers would have clear, negative consequences.**
- **On the other hand, the idea of models drafting standard contracts is more positively received.**
  - The fact that these contracts have a set template, are routine, and require less strategic thinking reduces fears about loss of skills.

*"Just because a minister likes a specific style, I don't think that's an adequate reason to get a computer to draft it up. You'd hope that people at the bottom of the chain were drafting those documents so they have experience doing that."*  
18-39, ABC1

*"I write briefings within local authority. Some like 1 side of A4, some like bullets, and often they want an urgent briefing. Each briefing usually has risk analysis so I can see a use for this, but only as a first draft and then a human intervenes."*  
40+, ABC1

*"I can understand a standard contract, but something that's a new or important subject should be drafted by a human being."* 40+, C2DE

Use case 1: Assisting call centre staff	
Perceived benefits	Medium to low
Perceived risks	High
Relevance / personal impact	High (for individuals in social security system)

- **Participants are concerned about AI generated advice providing a worse customer service for clients, compared to the existing model:**
  - They are accustomed to experiencing difficulties talking to chatbots in customer service settings. For example, chatbots that do not recognise the extent and nuances of their issue and redirect them to the wrong place.
  - They are aware that poor customer service is even more damaging for a vulnerable audience, such as social security claimants.
- **This use case has a direct impact on individuals, which means that a human touch and need for accurate information are felt to be doubly important:**

- Successful calls require the staff member to have empathy and emotional intelligence on top of the logical information that a foundation model could provide.
- There are severe and direct stakes for an individual claimant if a decision is incorrect, for example, being wrongly told that they cannot receive benefits that month. The risk of a foundation model producing inaccurate information is therefore heightened in this example.
- **There is concern about call centre staff over-relying on the advice from the foundation model:**
  - This leads to a concern that there will be either a loss of incentive for workers to be skilled at their job, or a slippery slope to the foundation model eventually replacing the staff member altogether.
  - There is further worry from some about the potential for a foundation model to recommend decisions which exacerbate biases from previous social security decisions.
- **However, participants are less opposed to this use case when considering the model's potential to detect fraud:**
  - This has a clear benefit and is seen as something which the corresponding human staff couldn't do.

*"I don't know if you've used a chatbot before but I find them so frustrating because they never give you the answer I need. It's really irritating when they don't understand you."* Public sector worker (Education)

*"The emotional side of things... I would worry a lot as people call because they have issues. You need that bit of emotional caring to make decisions. I would worry about the coldness of it all."* Early adopters / lay experts