



Medicines & Healthcare products
Regulatory Agency

10 South Colonnade
Canary Wharf
London
E14 4PU
United Kingdom
[gov.uk/mhra](https://www.gov.uk/mhra)

The Rt Hon the Baroness Stowell of Beeston MBE
Chair of the Select Committee on Communications and Digital
House of Lords
London
SW1A 0PW
Email: holcommunications@parliament.uk

Our reference: 19802

16 October 2023

Dear Baroness Stowell,

Communications and Digital Committee Large Language Models Inquiry

Thank you for your letter regarding the Communications and Digital Committee Inquiry into large language models. I am sorry for the delay in responding to your letter.

The MHRA responses to your list of questions are attached. I look forward to hearing the outcome of this important Inquiry.

Please do not hesitate to contact me should you require any further information.

Yours sincerely,

Dr June Raine, DBE
Chief Executive
Medicines and Healthcare products Regulatory Agency

E: Executive.Office@mhra.gov.uk



Medicines & Healthcare products Regulatory Agency

MHRA RESPONSE

1. How many staff do you employ specialising in AI-related governance (if any)?

Within Software Group at MHRA we have 1.5 full time equivalent staff (FTE) specialising in AI-related governance.

2. How many data scientists (or equivalent) do you employ (if any)?

We have no data scientists within the Software Group at MHRA. Within Digital & Technology (D&T) Group we have 2 x FTE Data Scientists, as well as some other related roles such as Data Architect and Data Engineers.

3. Do you anticipate these figures changing significantly in the next three years?

We intended to grow our AI focused capacity by 3 FTE over the next 2 years within the software group. Within the D&T Group, we also intend to increase our capacity in order to keep up with the pace of technology as well as the projects on the MHRA's Technology Roadmap, and also to reduce dependence on contingent labour. Alongside Data Scientist roles, we need critical enabling roles such as Cyber Specialists and Data Architects and Engineers. As a minimum we need 14 additional roles in the D&T Group, since as it stands, the proportion of D&T roles within MHRA is significantly lower than in other comparable organisations.

4. Do you expect any new regulatory issues of concern to arise within your remit in the next three years in relation to advances in large language models?

There are a few areas where large language models, and other forms of AI, generate concerns that must be managed to create safe and effective medical devices. For the most part, such concerns would fall into three categories:

A. Existing Issues seen in other software medical devices and mitigated accordingly

Issues within this category are broader in nature, for example ensuring that the claims and functionality of the software or AI product has a specific enough intended purpose to effectively navigate the regulations. This includes determining whether the product is a medical device and if so, what risk classification it falls under and beyond into the type and extent of supporting evidence needed to demonstrate that the benefits outweigh the risks.



Medicines & Healthcare products Regulatory Agency

As LLMs (and other foundation models) may demonstrate highly generalised functionality that can be applied to a wide range of use cases resulting in an intended purpose that is too broad to demonstrate safety against.

A second broader issue that is more common with LLMs is Software of Unknown Provenance (SOUP), whereby a legal manufacturer of a medical device bases their product on a large amount of “off the shelf” software that falls outside of their direct control.

From a design perspective this make it harder for the manufacturer to manage and design out risks in the product that relate to the SOUP component. Whether open source or paid for, currently, many LLMs and foundation models would be treated as SOUP for onward development into a medical device and may not have sufficient documentation or control to fully mitigate inherent risks.

By extension, the large and complex nature of LLMs, means that a third broad issue of transparency, whilst not exclusive to LLMs, is a concern. This relates to several sub-issues such as understanding the model’s decision-making process and explaining this to humans both for design controls and for useability and trustworthiness. If end users, who are typically regulated healthcare professionals, are unable to sufficiently trust or validate the AI product this may result in inappropriate decision making.

This not only creates an unsafe situation but currently leaves the clinical professional ultimately liable for a technical failure. A second issue relating to product transparency is the ability to identify and manage biases that may become baked into the training of the model during development. Compounded by the SOUP concern above, this can lead to medical device products that are unable to perform as intended for all use cases.

B. Novel issues for LLMs than can be managed through existing mitigations

The tendency for large language models to generate incorrect outputs, broadly termed hallucinations, is an issue for their use as a medical device. The specifics and propensity for different LLMs to hallucinate is an active area of research, as are the attempts to manage this issue through design controls (parameter modification and fine tuning) or process controls (prompt engineering and management).

However, the medical device regulations and associated governance frameworks already contain requirements to understand and mitigate the impact of these concerns. For example, beyond the design and process controls mentioned above, the requirement to conduct a structured clinical evaluation is in place to validate the performance of the LLM product.



Medicines & Healthcare products Regulatory Agency

C. Novel issues for LLMs than will require novel solutions

One area of novel regulatory issues that LLMs (and potentially other AI solutions) may generate is in the ability for them to utilise other assets. It is common for software products to access databases or other objects to achieve their functions, the management of these tasks is typically well understood and explicitly built-in during development. However, LLMs have demonstrated the ability to utilise software tools, not explicitly designed for use with an LLM, to produce a more accurate output.

For example, several studies into LLMs have demonstrated improved mathematical outputs when a software calculator program is accessible and demonstrated to the LLM. Such activities would require extensive validation and transparent understanding of how the LLM functions to meet the medical device regulations. Additionally, there is the potential for LLM based models to be provided with other software medical devices akin to the calculator in the above example, to achieve its functionality. In this scenario the products may be more complex than existing validation requirements could manage.

The pharmaceutical industry (and their software vendors) is also exploring use of LLMs across a range of applications, including summary of datasets used in regulatory submissions. Such application of LLMs would already be covered by existing requirements for robust quality management systems and subject to pharmacovigilance inspections. However, these tools may create novel challenges which will need to be addressed.

Additionally, there is a concern that data generated using LLMs could be used to disrupt existing vigilance processes, as it may become challenging to determine whether (for example) an incident report has been created by a citizen or has been generated via an LLM. Provenance of data will become increasingly important, to determine whether it is both real and accurate and should be included within regulatory processes.

5. Do you anticipate a requirement to conduct an audit of algorithmic processes used in large language models in the next three years? (For reference we refer to the criteria set out in a paper by the Digital Regulation Cooperation Forum on algorithmic audits: governance, empirical, and technical. Please provide an indication for each of these three types)

We do not anticipate a requirement to conduct an algorithmic audit by the MHRA directly. However, audits are a fundamental component of regulating medical devices, all types of audits outlined by the DRCF are conducted by different stakeholders throughout a product's lifecycle and also apply to audits of the organisation legally responsible for the



Medicines & Healthcare products Regulatory Agency

product. Where a large language model qualifies as a medical device, then the following outline of our governance involving audits may provide additional context:

Product Auditing

Prior to deployment, all products, apart from the lowest risk category (Class 1), are audited as part of the premarket assessment process. This is conducted by a third-party organisation called an approved body that operates on behalf of the MHRA as the UK regulator. The MHRA does not have the capacity to directly audit every product prior to being placed on the market. We therefore audit and oversee the approved bodies.

Once placed on the market, legal manufacturers are required to audit their development processes and products frequently. In the event of a safety concern, manufacturers are required to investigate the root cause of the issue, mitigate the problem and report this to the MHRA. Auditing here would both form an investigative tool and a validation activity to ensure the risk has been managed effectively.

Finally, whilst not mandated under MHRA regulations there is a growing focus on end users conducting audits on products proactively, especially in the AI space.

Organisational Auditing

Within the medical device sector of healthcare, there are governance mechanisms both for products and for the organisations that are legally responsible for them. One such governance mechanism is auditing of a manufacturer's staff and processes against internationally recognised standards. Auditing is performed by an approved body against a standard (typically ISO13485) and repeated on an annual basis, or more frequently if there is a concern or non-conformity issue being managed.

Organisational issues must then be resolved within a given timeframe to maintain certification to legally manufacture medical device products. Additionally, these organisational audits will review the internal audits conducted by the manufacturer (as outlined above) as an additional layer of governance.

In summary, auditing forms a vital component of the regulatory and governance systems within medical device safety. These activities are conducted by several actors in a repetitive way across the lifecycle of all medical device products which would include LLM based medical devices.



Medicines & Healthcare products Regulatory Agency

- 6. Notwithstanding your answer to question 5, do you believe your organisation is well positioned to conduct audits of large language models at the (a) development and (b) deployment stage, to ensure compliance with your existing duties and the expectations in the Government's white paper?**

In general, the MHRA would be lacking the key skills to conduct a technical audit for regulatory purposes as described by the Government's white paper. However, expertise from across the agency could potentially be convened to work collaboratively in exceptional circumstances. This hypothetical scenario would most likely need to be in response to a safety concern after deployment and not as part of assisting product development unless as an activity covered by a clinical investigation application.

Regarding governance and empirical audits, whilst outside our scope of work, current skills within regulatory technical teams would be able to conduct such activities. This would also need to occur as a result of a safety concern once deployed on the market, or via a clinical investigation application. Whilst we would then be able to respond to any safety driven concerns, this would have a knock-on impact on the capacity of the team.

Capacity provides a greater challenge and the MHRA currently oversees in excess of 3 million medical device products on the UK market.

- 7. (If applicable): Please could you provide an estimate of how many comparable audits you have undertaken in the last five years?**

Not applicable

- 8. In your view, what are your main powers and duties most applicable to delivering on the five principles set out in the white paper?**

Application of these five principles within the healthcare, and specifically the medical device space for the MHRA, will require a nuanced approach to ensure compatibility with the extensive existing regulatory infrastructure in place for health and social care. If implemented poorly there is a risk of additional regulatory burden for regulators and manufacturers of medical products.

This increased burden may result in fewer healthcare products for UK patients. We have attached our response to the white paper which outlines in greater detail how the MHRA's remit interacts with the five principles. It is also worth noting that there are many other regulators and assessors within the healthcare sector whose remits must also be considered against the principles.



Medicines & Healthcare products Regulatory Agency

9. Are there any areas where you anticipate your existing powers and duties will fall short of what is needed to deliver against the expectations in the white paper?

At this point in time the MHRA has not conducted a detailed review of implementing the existing principles via existing or planned legislative powers. At a higher level the MHRA cannot identify any clear gaps in our ability to satisfy the aspects of the principles that fall within our remit. Greater clarity on the interpretation and expectations of meeting the principles is required before a more detail legal analysis can take place.

10. How confident are you that there are sufficient co-ordination structures, independent of Government, between regulators outside the Digital Regulation Cooperation Forum in relation to regulating large language models?

There are several AI focussed projects between governance bodies within healthcare and projects such as the AI and Digital Regulation Service (AI&DRS) look specifically at assisting with regulations. Large language model discussions reside within several of these forums. However, the scale, volume and pace of developments in the LLM innovation landscape does pose a risk to the capacity of our service provision as one of several regulators.

11. a. (If applicable): please provide examples of established structures that are working well.

[AI and Digital Regulations Service for health and social care - AI regulation service - NHS \(innovation.nhs.uk\)](https://www.digitalregulations.innovation.nhs.uk/)
[\(https://www.digitalregulations.innovation.nhs.uk/\)](https://www.digitalregulations.innovation.nhs.uk/)

[Guidance to civil servants on use of generative AI - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai)
[\(https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai\)](https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai)