

UNION INTERPARLEMENTAIRE

INTER-PARLIAMENTARY UNION

ASSOCIATION DES SECRETAIRES
GENERAUX DES PARLEMENTS



ASSOCIATION OF SECRETARIES
GENERAL OF PARLIAMENTS

COMMUNICATION

by

Ms Sarah DAVIES
Clerk Assistant of the UK House of Commons

on

“Experimentation with and governance of use of AI within the UK Parliament”

Istanbul Session
April 2026

Experimentation with and governance of use of AI within the UK Parliament (and notes on updating the six-point plan for Secretaries-General on AI in Parliaments)

1. In 2023 I presented a paper at the ASGP for a six-point plan for Secretary-Generals regarding the use of AI in their Parliaments, which was subsequently updated as below.

Who is talking to who about AI in your Parliament? Is this only a “digital” discussion?

Is there a forum in your Parliament bringing together the interested groups of parliamentary staff listed in this paper? Where are your sources of expertise on data and AI? Does it span internal and external staff? How are your Libraries and data experts involved?

Has your Parliament produced staff guidance on use of Generative AI such as ChatGPT? If so, is it widely shared?

Are you assessing potential Chamber-related opportunities and risks in key areas such as questions, indexing, transcribing and legislation?

Has your Parliament made APIs available or at least have a timetable for the process of making such a decision?

Are you confident that your scrutiny function has the expertise it needs to consider these developments effectively, and is joined up to those who are assessing the implications for front-of-house and back-office operations across your Parliament?

2. This paper sets out some of the concrete uses of AI in the UK Parliament to date, with lessons learned. It first suggests two particular areas of focus for Secretary-Generals in the current climate: **governance** and **evaluation**. These suggestions came out of a series of interviews I conducted with Secretaries-General in 2025 as part of a Masters on Organisational Psychology and I’m now working as part of a policy fellowship with Imperial College London to put them into practice.
3. I’m grateful to all those I interviewed last year for their assistance in this work, as I am to colleagues who have kindly sent through material from their parliaments either directly or via Luke Hussey, Director of Digital Value in PDS.
4. All these projects are partnerships. They involve colleagues from Parliamentary Digital Service (PDS) working with the subject matter experts, or users, under time pressure, limited resources and complex operational delivery challenges. We are all striving to make more efficient use of public money, ensure that business as usual (particularly Chamber business) proceeds smoothly and respond to new and additional demands on our services.

5. The relationships that operational staff build with digital staff and contractors are critical for the success of any project – even more so for an AI-based project because of the ongoing hype/technical complexity/bias/ambiguity which can surround them.
6. **A mutually agreed, simple evaluation plan is essential for maintaining these partnerships under pressure and avoiding different narratives being told about the same project which seems to me to be a considerable but avoidable risk.**

Governance

7. **Achieving clarity is key – what is the pipeline for AI-related projects, and how does this fit into existing business case and/or digital prioritisation structures?**
8. At the moment, there is a risk in the UK Parliament that we have four concurrent pipelines:
 - AI projects in particular
 - digitisation measures forming part of the House of Commons Savings and Improvements Programme (some of which could be AI-related)
 - projects coming through our seven digital value streams
 - through citizen development, which can be a limited resource depending on the extent to which it requires testing support/cyber advice from PDS.
9. Integrating this through the existing governance mechanisms - value stream oversight boards, the AI strategy group, the Information and Digital Board – with Member input via the Commons Speaker’s Steering Group on AI in Parliaments and the various domestic committees – is a considerable leap of synthesis.

Evaluation

10. A way to unlock this knotty problem is to step up and simplify how we evaluate this work. There seems to be a view in many parliaments and indeed many institutions and companies that we need to get better at how we evaluate these projects. Projects can be technically evaluated (search accuracy, for example) and there are great examples of this. Work is done with users about their needs and journeys as part of our agreed processes. But I am struck by how regularly the outputs of these projects have different values attached to them at the end by the different groups involved and how the same project is described through multiple narratives.
11. The risk of this if not addressed and brought into the open is that first there is a breakdown of relationships between users and digital teams and secondly that lessons are not learned across different projects.
12. **This is the reason that I’ve drafted an evaluation plan in terms of questions to be considered by the different groups involved in a project (so in the UK Parliament**

case, this would be the relevant PDS team, users of the new technology and someone from the relevant part of the business with management/leadership responsibilities).

13. **We're trialling a mechanism like this, in each value stream, to see whether in practice it achieves shared understanding across technical and non-technical teams, clearly articulates the problem, surface misunderstandings early, reduces optimism bias and underlines the importance of ongoing evaluation. It will also be used to conduct a retrospective of several projects with have been or are near competition.**
14. This also clearly needs to be put into the context both of Parliament's Information and Digital Strategy and wider Government work. The UK Parliament's new Information and Digital Strategy for 2026-30 emphasises the need to be using the right tools to support modern working practices: effective evaluation of emerging technologies such as AI is crucial to that goal.
15. The UK Government's new Digital and Data Benefits Framework (published 7 April 2026) <https://www.gov.uk/government/publications/digital-and-data-benefits-framework/digital-and-data-benefits-framework> proposes a clear approach for articulating and assessing the benefits of AI investments, moving beyond assumed efficiencies to evidence productivity gains, service quality improvements and risk reduction. It deliberately sets out to align with other public sector evaluation frameworks and guidance, and for the UK Parliament, applying this framework could strengthen scrutiny and governance by clarifying the expected value of AI-enabled activities, challenging optimism bias, and supporting more informed choices at the design and delivery stages. This approach would help ensure AI is adopted where it genuinely adds value and where outcomes can be evaluated transparently over time.

Evaluation plan

Staff from PDS, potential users and a leader from within the Team concerned meet to discuss and agree:

1 What is the problem we are trying to solve?

2 Do we really need a digital or AI solution or can we change the process?

3 How much time have the staff team got to explain what they do and work with PDS/contractors? When will the team be busy or away (eg recess) and have to disengage? How can that be built into the workflows?

4 Who will lead the project from (a) PDS and (b) the team? Do we all know who is involved and what their roles and responsibilities are? What will the escalation routes be?

5 How long is it likely to take, what are the key stages and how could it get blown off track? What is the range of time it might take and what would the implications be of the longest possible time taken?

6 At what points do we lock things in and when can we change track? What are the different ways this project might end?

7 What will we say to everyone about what this project is about and how much it will cost? (this is the part to bring in technical assessments and also finalise what we think the likely benefits will be)

8 Where are our potential areas of disagreement or misunderstanding and how can we reduce the risk of this derailing the project? What is the Plan B?

Concrete uses in the UK Parliament

16. The second part of this paper sets out some of the work done so far in the UK Parliament on use cases, particularly on semantic search, automatic speech recognition, summarising/drafting and copilot.

Semantic search – definite opportunities, mixed results

17. **Table Office question editing and semantic search:** a pilot was conducted with the House of Commons Table Office which is currently experiencing very high numbers of written questions (over the last 6 months, up to over 500 a day against a historical average of around 300).
18. This took several months of developer work, initially to see whether an automatic editing tool could be imported into existing electronic systems. The work done created a new form of semantic search which has made it much easier to spot duplicate questions, saving some staff time but readouts of precisely how much vary. Less progress was made on the editing tool and in the end the team decided to reduce the number of rules applying as an emergency mechanism, rather than proceeding with use of the editing tool.
19. A recent project on e-petitions conducted by a third party **made good progress in some areas such as website accessibility but less on semantic search** – we are learning lessons from this in terms of how work done in-house and the IP involved can be shared with external partners.
20. Work has taken place with Hansard on the use of semantic search to power Hansard search, enabling users to more quickly and effectively find relevant information, compared to the existing keyword search. The completion of historic Hansard will be publicised after Easter and is an example of citizen development; semantic search has tremendous potential.
21. A proof of concept with PDS and Faculty (AI partner) was completed in December 2025 which looked at how Parliament's data can be used and accessed more easily and leveraged to enable smart search, for example with advanced Hansard search capability. This was demonstrated by creating tools that expose or search Parliamentary data, using a standardised Model Context Protocol (MCP) framework.
22. Next steps for this work are being agreed - this approach provides a template for exposing and connecting more tools and data, and background work on a target architecture needs to run alongside it. There will also be work done to explore adopting this pattern for improving access to all public Parliamentary data, for example for search.Parliament. User acceptance testing for the similar search AI feature is currently in progress by Commons Table Office users and the feature is due to be released to the Electronic Questions and Motions backend admin application before summer 2026.

Automatic Speech Recognition: has been trialled in various forms for years by Hansard, links to potential savings in subtitling, more to come

23. Commons Hansard and broadcasting are planning to use Automatic Speech Recognition instead of re-speaking for BSL subtitling in the Chambers and reinvesting the significant contract savings in subtitling for Select Committees.
24. Work has also taken place with Hansard teams in both Houses to support the use of AI in **transcribing proceedings** into a format which reflects House style and is available to edit more readily. This will be trialled **this autumn** concentrating on Select Committee transcripts in the first instance, introducing light-touch intelligent editing. At the same time, we will be testing the market for AI solutions with the re-tendering of the transcription services contract (around two thirds of Select Committee sittings) over the summer.
25. We are not aware that another Parliament has achieved savings (as opposed to expansion of services, eg Bundestag) through ASR transcription though keep a watching brief – and welcome feedback and experience from other Parliaments. We want to build elements of interoperability into our redeveloped Hansard Reporting System (HRS) but the technical work on this, particularly on production, is at a very early stage.

Summarising/drafting: ongoing trials, practical guidance being issued and iterated

26. **Consultation and evaluation:** There is ongoing work to explore the best way to identify and take forward an AI tool for the Select Committee use case. There are also trials underway in the Lords Committee Office to support the summarising of evidence using AI. Work has also taken place to use AI to support the summarisation of written evidence received by House of Lords select committee to their inquiries. This has been developed as a proof of concept by colleagues in the Lords, and has demonstrated the potential of the technology to support teams to manage and analyse the responses received.
27. Next steps include a period of testing with committees, as well considering how to convert this local innovation into a resilient tool which can be used across both Houses – and see whether it can be scaled up and appeal to a range of users.
28. **Guidance** has been now issued to all staff groups about the risks and opportunities of using Gen AI and other tools for these purposes, including information risks, alongside access to training. Teams like Research and Information (also known as the House of Commons Library) have produce specific guidance for their researchers eg on prompts and are testing new forms of semantic search and ways to assist production of papers and research without introducing bias or errors.

Copilot: relatively expensive and not the answer to everything but promising results in some areas

29. While the full version of Copilot has been trialled, problems have surfaced at a late stage with potential access to material because of how the tool uses document sharing links. Full rollout has been paused while further checks are made on historic permissions to ensure that eg parliamentary privileged information will not be surfaced.
30. The period of extended piloting and preparation will need to be used to enable better information management hygiene to be in place and for teams be clear about whether they want to use it and how to get value from it. Further decisions on this are expected later this year.
31. PDS are investigating with Microsoft a 'political content blocker' issue in M365 Copilot and Copilot Chat. When inbuilt safety filters in Copilot detect content which could plausibly be political advocacy or election influence, it blocks or deflects the request to reduce the risk of misinformation or unintended persuasion. This is an issue in the Parliamentary context and may impact the uptake of the tool – for example, an MP was blocked from writing a newsletter using Copilot.
32. Part of the UK Parliament's pilot of M365 Copilot has included experimenting with Copilot agents. Four pilots have taken place. The first was a SharePoint agent created to search and analyse the Research and Information SharePoint site. The intent was to speed up discovery and synthesis of research material. There were promising results with minimal set up.
33. The second was a HR agent to search and query the House of Commons and Joint Department Staff Handbook. The goal was to support the PACT team in responding to initial staff questions more efficiently. Outputs were human validated to reduce misinformation risk. It worked and again, was relatively straightforward to use.
34. The third was a bespoke inbox management agent built for an MP. It aimed to categorise incoming emails (casework, policy/campaign, other), automatically forward casework emails to a caseworker, and draft template responses for campaign/policy emails. The agent was not deployed live due to cost and complexity in Copilot Studio. The functionality is now being rebuilt in using Microsoft's PowerAutomate tool (at a fraction of the monthly cost of using Copilot Studio).
35. The fourth was an agent developed for the Contact Centre to help new call handlers efficiently retrieve the information they needed to respond to queries in real time. Initial attempts to use Copilot out of the box faltered (it started to inexplicably provide responses in French), but an agent has demonstrated more value.

36. Aside from the inbox management agent, all are desired to go into production. There is an outstanding question about token costs: unlike normal Copilot, there is a charge per query (£0.01p a prompt). In an agentic setting, these can add up very quickly so affordability needs to be assessed.

37. From April 2026, the UK Parliament's next stage for testing the full potential of Microsoft's Copilot tooling is one of "early adoption". This will see a campaign across teams to improve information management processes to address the emerging risks identified, as well as further explore the implications of using agentic AI in a live service setting. The intent is by April 2027 to have 1000 active M365 Copilot users. Clear evaluation criteria are being agreed, with a focus on both value for money and team productivity.