

Accessible to whom and for what purposes? Researchers reflecting on working across 125 years of parliamentary proceedings¹

Sam Hames², Michael Haugh², Naomi Barnes³, Eve Ansell^{2,4}

Abstract

Timely digital publication of accurate parliamentary proceedings unquestionably supports democracy and democratic deliberation “as it happens”. However, because of their high quality and relatively consistent setting, these same proceedings also have significant value over centuries as records of society, culture, history, and language.

In this paper we reflect on this broader value of parliamentary proceedings, and in doing so consider the challenges and enablers of how we can help more people to engage with and access these materials in transformative ways. We draw on our experience with two research projects in the fields of linguistics and education. Both projects analysed the same 125 years of the Proceedings of Australian Federal Parliament using different combinations of computational methods and the close-reading that is fundamental to the humanities. We introduce the specific confluence of factors that made these projects possible, and in doing so we outline the limitations and challenges of current approaches to making parliamentary proceedings accessible and available.

We situate these limitations and challenges in the context of the modern information landscape and how difficult it can be for even a well informed citizen to refute misinformation. This means that accessibility and availability considerations for parliamentary proceedings should also treat those proceedings as data collections. Bulk publishing of proceedings in convenient, machine accessible formats enables broader networks of civil society organisations to close the gap between access and availability for more diverse audiences, and better access is key to responding to mis- and dis-information.

Introduction

Timely publication of what is said in parliament by elected representatives is now taken for granted in many governments around the world. This is particularly true in jurisdictions that have adopted

¹ This project was supported by the University of Queensland’s Centre for Digital Cultures and Societies, the Language Data Commons of Australia (LDaCA) and the Reusable and Accessible Public Interest Documents (RAPID) project. LDaCA and RAPID are supported through a co-investment partnership with the Australian Research Data Commons (ARDC) as part of the HASS and Indigenous Research Data Commons (DOI: 10.3565/kq2v-9g52). The ARDC is enabled by the Australian Government’s National Collaborative Research Infrastructure Strategy (NCRIS).

² The University of Queensland

³ Queensland University of Technology

⁴ QCIF Digital Research

Westminster style parliamentary systems including former British colonies such as Australia. As a practice, publication of transcripts supports democratic engagement through transparency and accountability of elected representatives. Public service teams within parliaments routinely prepare and release transcripts in very short order, allowing reading and searching within hours of words being spoken in a parliament.

The transition to transparent publication and reporting on proceedings in Westminster systems can be traced to changes in British Parliamentary practices along with the rise of newspapers during the late 1700s and 1800s. The timing of this shift meant that many then British colonies adopted the practice of producing parliamentary proceedings very early in their history. For example Australia's Federal Government has prepared and published complete proceedings since Australian federation in 1901. It is this long span of time for the production of proceedings of parliaments in many different contexts around the world that amplifies the value of these proceedings even further than their important primary purpose of supporting democratic governance.

This paper will examine what this broader value and use of parliamentary proceedings looks like in the context of two research projects in linguistics and education that worked across the complete Proceedings of Australian Federal Parliament from 1901-2025. We reflect on the critical enablers of our access to these proceedings, and how this access depended in our case on a critical combination of skills, experience, and the foundational work of the public service. We use our consideration of these enabling factors for our research projects to reconsider what access looks like in the broader context of civil society and access to parliamentary materials in the modern information landscape, and suggest that thinking about parliamentary proceedings as data collections is one promising pathway to collaboratively improving access for everyone.

The Broader Value of Parliamentary Proceedings for Research

Parliamentary proceedings have a broader value than solely as records of elected representatives in support of democratic governance. Because parliamentary proceedings cover long time periods, are prepared by dedicated public service teams with broadly comparable processes, and are prepared with a commitment to integrity and a particular kind of accuracy of presentation (typically described as “substantially verbatim”^{5,6}) parliamentary proceedings are a unique resource that only grows in value over time. There are few comparable resources that extend through multiple decades, are easily accessible and are prepared across multiple countries around the world with broadly similar characteristics. Newspaper collections are comparable in many respects, but run into issues such as the “copyright cliff of death” that prevents more recent materials being made available⁷. This

5 Stef Slembrouck, ‘The Parliamentary Hansard ‘Verbatim’ Report: The Written Construction of Spoken Discourse’. *Language and Literature: International Journal of Stylistics* 1(2) 1992, pp. 101–119. <https://doi.org/10.1177/096394709200100202>.

6 For a practice oriented view of what substantially verbatim means see: Emma Hardman, ‘Word for Word: Finding Meaning beyond Verbatim’. *Griffith Review* 31(Autumn) 2011, pp. 58–66. <https://search.informit.org/doi/10.3316/informit.371033652088423>.

7 Tim Sherratt, ‘Beyond the Copyright Cliff of Death’. *Tim Sherratt – Recent Updates and Work-in-Progress*, November 13, 2020, <https://updates.timsherratt.org/2020/11/13/beyond-the-copyright.html>.

accumulation of effort and extension into the current day means the parliamentary proceedings can be a unique resource for examining society, history, culture, and language in new ways.

These multiple perspectives mean that proceedings are used to create new knowledge in a very broad variety of research disciplines and contexts. We will not attempt a comprehensive survey here, but we will illustrate the diverse landscape of research drawing on parliamentary proceedings with a focus on the recent Australian context. Some research use cases start from detailed considerations of the impact of and responses to single notable speeches, as in the 2012 case of then Prime Minister Julia Gillard's "misogyny speech"⁸. Other research use cases slice across different dimensions of parliamentary speech and how that interfaces with different aspects of society and governance. Examples include: examining why and how particular interest groups are represented in parliamentary speech⁹; where and in what ways Indigenous languages are being represented in Australian parliaments¹⁰; or whether women in parliament are interrupted more than men¹¹.

Changes in parliament, whether reflecting broader changes in society or the specifics of parliamentary practice provide a wealth of options for research. A transition to more valedictory speeches as parliamentarians leave parliament have provided an opportunity for detailed investigation into the role of different genres of parliamentary speeches and how these speeches can support democratic transitions¹². The long term nature of proceedings mean that with appropriate care they can be used to study how language itself changes over time¹³.

We would, of course, be remiss not to also acknowledge that these proceedings are not neutral or strictly "representative" of broader society, and that the broader value we are illustrating here can only be accessed with appropriate care and attention to both the broader context of parliamentary speech (such as who is eligible to be elected), and also to the procedural operations of how parliaments conduct their business and the way the transcripts themselves are produced. Without consideration of the broader context we might make assumptions that reproduce inequities and injustices, and ignoring the procedural details may mean overlooking artificial gaps and silences in the records. This might be as simple as the common transcription policy that interjections that are not acknowledged by the

8 Ngaire Donaghue, 'Who Gets Played By 'The Gender Card'?: A Critical Discourse Analysis of Coverage of Prime Minister Julia Gillard's Sexism and Misogyny Speech in the Australian Print Media'. *Australian Feminist Studies* 30(84) 2015, 161–178. <https://doi.org/10.1080/08164649.2015.1038118>.

9 Darren R. Halpin et al., 'What Explains Interest Group Prominence in Parliamentary Speech? Policy Agenda, Partisanship, or Conflict Expansion'. *Political Studies* 73(1) 2025, pp. 196–217. <https://doi.org/10.1177/00323217241232934>.

10 Alexandra Grey, 'Celebrating Indigenous Linguistic Diversity in Australia's Parliaments'. 42(2) 2025, pp. 198–236. <https://doi.org/10.1080/07268602.2025.2503862>.

11 Lindsay Katz et al., 'Do Women Politicians Face More Interruptions? An Analysis of Interjections in the Australian Parliamentary Debates (1998-2025)'. *Australian Journal of Political Science* [online first] 2026, pp. 1–27. <https://doi.org/10.1080/10361146.2026.2642819>.

12 Peter Ferguson et al., 'Valedictory Speeches in the Australian Parliament: Democratic Rituals, Legislative Rotation and Autobiographical Self-Disclosure'. *The Journal of Legislative Studies* [online first] 2026, pp. 1–19. <https://doi.org/10.1080/13572334.2025.2610528>.

13 Haidee Kruger and Adam Smith, 'Colloquialization versus Densification in Australian English: A Multidimensional Analysis of the Australian Diachronic Hansard Corpus (ADHC)'. *Australian Journal of Linguistics* 38(3) 2018, pp. 293–328. <https://doi.org/10.1080/07268602.2018.1470452>.

member with the call or the chair are not included¹⁴, or more extreme examples where editorial decisions are made to expunge particular events from the record¹⁵.

In reflecting on these kinds of methodological issues facing research that draws on parliamentary proceedings, and the critical importance of contextualising those research findings, there are two broader considerations that also come into play. First, do the ways in which they are made available or accessible shape or even constrain the kinds of research questions that are - and can be - asked by researchers? Second, to what extent can new methodological directions push against those constraints, thereby broadening the scope of possible lines of research? In the following section, we reflect on the methodological issues facing researchers wishing to draw on parliamentary proceedings through the lens of two research case studies.

Research Case Studies with the Proceedings of Australian Federal Parliament

To motivate our discussion on access to parliamentary proceedings we briefly cover the process underlying two research projects we have undertaken. The first is situated within linguistics and examines unparliamentary language, while the second examines why and how political rhetoric around education frequently uses crisis narratives. Both of these case studies draw on the complete Proceedings of Australian Federal Parliament from 1901-2025. As well as the primary disciplinary questions, in both of these projects we were also aiming to articulate new methodological directions using computational approaches to large collections of materials in the humanities.

Throughout these projects, and especially at the beginning, we relied heavily on the Australian Federal Parliamentary Library's ParlInfo search interface¹⁶ to access and read materials. A limitation that we quickly ran into for both projects was the limited affordances of a search based interface to support our exploration before we had clearly articulated the exact direction of each project. Creating an effective search query required us to simultaneously understand what we were searching for and also the materials we were searching over. Developing both of these at the same time proved challenging and conventional search interfaces such as ParlInfo did not straightforwardly allow for iterative development of those understandings.

Consequently for both of these projects we took the decision to create a complete speech-level copy of the proceedings, treating them as a structured data collection. Because we had multiple projects covering the same topic and were interested in covering the complete Australian Federal Parliament, and we had someone with the appropriate technical skills to do so on the team, we conducted a speech level web-scraping exercise¹⁷ to reassemble a complete database of the proceedings, which then

14 D.R. Elder and P.E. Fowler, eds., *House of Representatives Practice*, 7th ed. (Department of the House of Representatives, 2018).

15 Charlie Feldman, 'You Can't Print That in Hansard: Surveying Hansard Expungements in Canada, Australia and New Zealand'. *Australasian Parliamentary Review* 38(2) 2023, pp. 97–117.
<https://doi.org/10.3316/agispt.20240105101494>.

16 Parliament of Australia, 'ParlInfo - Basic Search'. Accessed at:
<https://parlinfo.aph.gov.au/parlInfo/search/search.w3p>.

17 Apart from not recommending this as web-scraping has now become a significant source of problems for website operators in general, it turns out that this also wasn't necessary: we could have used the XML and SGML transcripts for this purpose if we had known that they were publicly available for the complete time period.

enable us to apply computational text analytics methods to a copy of the complete proceedings. Our efforts follow similar lines of work to transform proceedings into data collections for a variety of purposes. Other examples in Australia include the Open Australia Foundation's work underlying the They Vote For You project¹⁸, Tim Sherratt's Historical Hansard website¹⁹, and work by multiple teams to assemble various datasets from Federal Parliamentary materials^{20 21}.

With a complete local copy of the proceedings we were able to proceed much more quickly with detailed computational analysis for each project. This helped us address both our immediate exploratory needs across the proceedings, but also created the opportunity for us to innovate methodologically in ways that were not possible through a web-based search interface. The flexibility afforded by creating a bulk dataset cannot be understated in terms of enabling these two projects.

Unparliamentary Language

In our first case study we systematically examined how the word “unparliamentary” was used by parliamentarians across the history of Australian Federal Parliament.²² Our specific focus was to understand specifically how the word unparliamentary was used in a metapragmatic sense, that is how “unparliamentary” was used when characterising other speech acts (for example, to contest a particular characterisation of another member), but also in how unparliamentary language itself became the topic of discussion (for example, through debates about whether a particular statement was unparliamentary).

Having chosen unparliamentary language as the central concept of the study, we used a computational approach to articulate the broad sweep of how “unparliamentary” was used across the entirety of the proceedings. In very brief, we clustered the words used across all speeches based on how similar the sets of speeches they occurred in were: we developed a custom algorithm for iteratively improving which words were grouped with which. As words that are related are likely to be used together (Firth's distributional hypothesis) this computational procedure results in groups of topically or functionally similar words. We used the set of speeches containing the word “unparliamentary” as a probe into this automatic groupings to identify groups and words related to unparliamentary, a subset of which are shown in Figure 1.²³ This gave us a computational overview to explore in aggregate the context and use of “unparliamentary”.

18 ‘How Does Your MP Vote on the Issues That Matter to You?’ *They Vote For You*. Accessed at: <https://theyvoteforyou.org.au/>. [accessed 30 August 2026]

19 ‘Home : Historic Hansard’. Accessed at: <http://historichansard.net/>. [accessed 30 August 2026]

20 Lindsay Katz and Rohan Alexander, ‘Digitization of the Australian Parliamentary Debates, 1998–2022’. *Scientific Data* 10(article 567) 2023, pp.1–14. <https://doi.org/10.1038/s41597-023-02464-w>.

21 Alfie Chadwick et al., ‘Hansard DB: A Relational Database of Australian Parliamentary Speech’. SoDa Laboratories Working Paper Series 2026-01, Monash University. Accessed at: https://ideas.repec.org/p/ajr/sodwps/paper_1779665234420_559.html.

22 Sam Hames et al., ‘“How Is That Unparliamentary?": The Metapragmatics of ‘Unparliamentary’ Language in the Australian Federal Parliament’. *Lingua* 320(article 103932) 2025, pp. 1–24. <https://doi.org/10.1016/j.lingua.2025.103932>.

23 Our complete annotations are available. See Samuel Hames et al., *Metapragmatics of “Unparliamentary”: Detailed Annotation of Word Clusters That Co-Occur with “unparliamentary” in the Proceedings of Australian Federal Parliament*, March 6, 2025. Accessed at: <https://doi.org/10.5281/zenodo.14977495>.

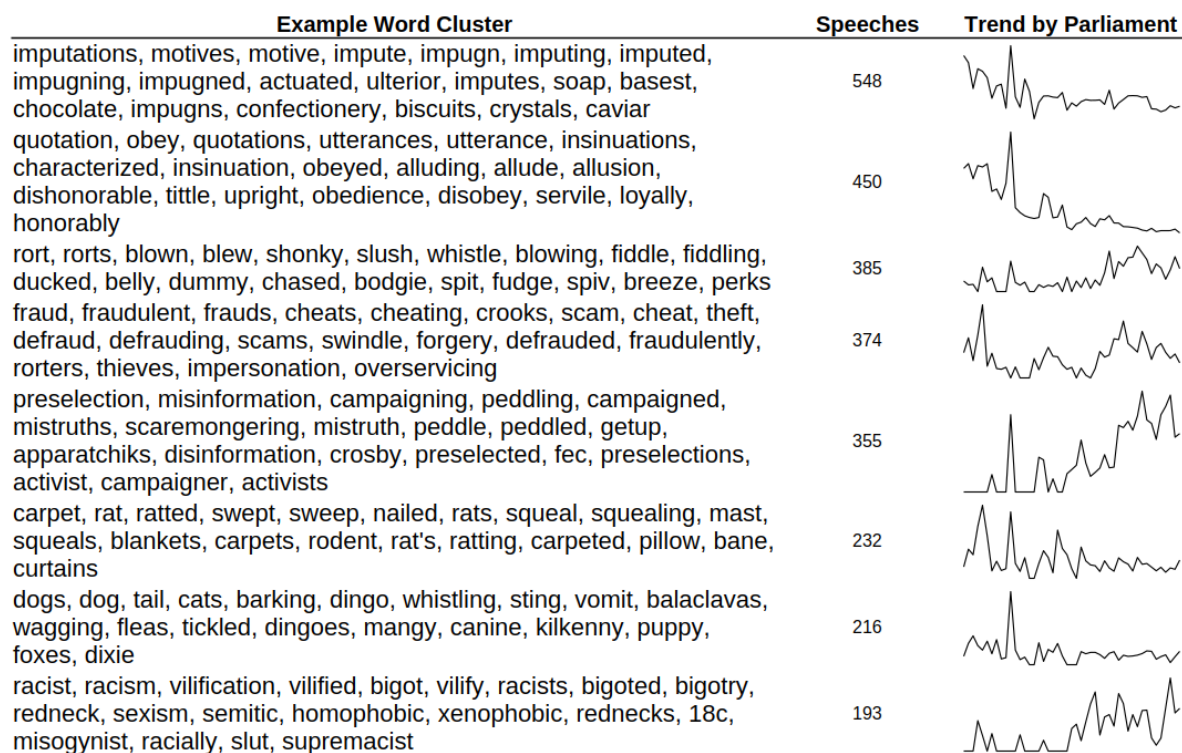


Figure 1: Words and groups of words associated with “unparliamentary” by their occurrence in speeches. The trend line is the proportion of speeches containing any of the words in the group and the word “unparliamentary”, by parliament.

Combined with careful close reading of speeches to understand the contexts of word use, we were able to use this computational overview to construct a typology of unparliamentary language. Apart from identifying some expected patterns and hotspots for the occurrence of unparliamentary language (such as accusations of lying and the procedures relating to withdrawing remarks), this annotation process also serendipitously created a number of suggestions for further analysis, in particular the identification of “racism” and “racist” as words associated with “unparliamentary” language. By further examining “racism” and “racist” using corpus linguistics approaches to see when and where they were first used and examining specific cases where they were used in unparliamentary ways we were able to study at multiple levels of detail how the use of language and particular terms can be construed as unparliamentary.

Ultimately we showed that the use of the word unparliamentary is more complex than might be expected from definitions in Standing Orders, and also that parliamentary discourses about unparliamentary language have changed over time as members contest and enforce boundaries. In an increasingly mediated parliamentary environment where an off-microphone insult that is subsequently withdrawn can receive significant media coverage²⁴ considerations of what is unparliamentary will continue to be important.

24 Blair Jackson, ‘Angus Taylor Doubles down on Albo ‘Arrogant Prick’ Jab’. *News.Com.Au*, 28 May 28 2026. Accessed at: <https://www.news.com.au/national/politics/angus-taylor-doubles-down-on-albo-arrogant-prick-jab/news-story/32e64375691eae13e23568ef9e0d36b9>.

Education Crises

Schooling is a continuing political issue in Australia. Despite public schooling being primarily a state responsibility, the Federal Government still exerts considerable influence through a variety of means, including national policy such as the National Curriculum “setting the expectations for what all young Australians should be taught, regardless of their background or where they live.”²⁵ Political narratives about schooling and education frequently use ‘crisis narratives’: a projected future shortfall of teachers might be described as a ‘system in crisis’. The aim of this research project was to examine the history of where and how crisis narratives were used in political debates about education at the national level in Australia.

Using a similar algorithmic procedure to our unparliamentary language investigations we identified a broad set of keywords relating to specifically to education (including expected words like “student”, “teacher” and “skills”), but also picking up on organisation names and education specific acronyms like “ATAR” (the Australian Tertiary Admission Rank). Rather than trying to categorise these different words we instead used these as a broad exploratory net for speeches likely to contain education related crises. We did this by searching for every speech where one of the education related words occurred within ten words of “crisis” or “crises”, identifying 2600 speeches for closer initial examination. The trend of these speeches over time are shown in Figure 2.

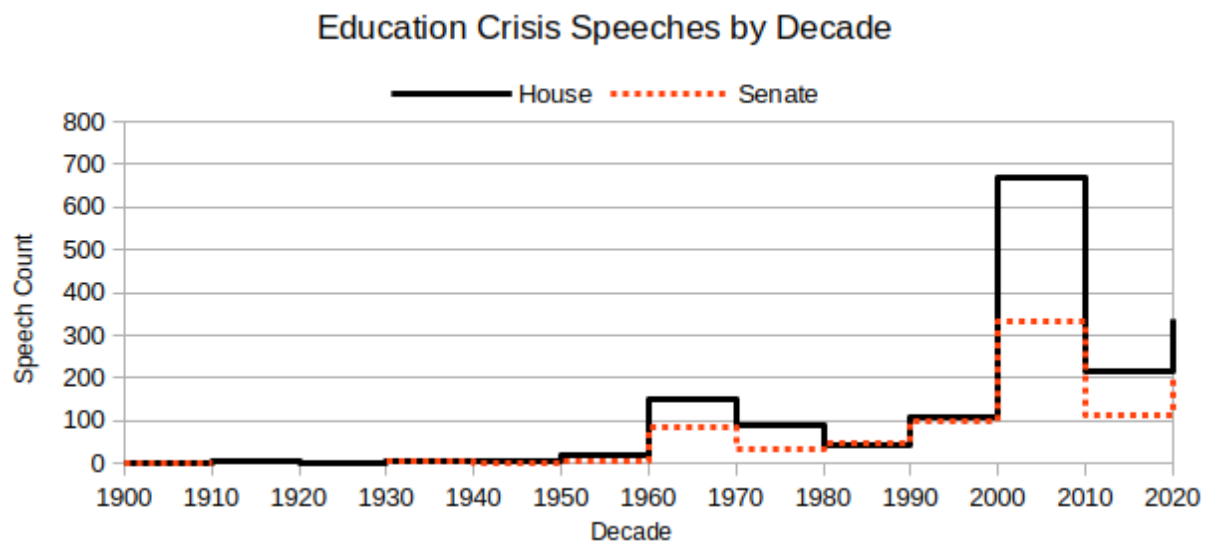


Figure 2: Counts of speeches by decade where “crisis” or “crises” occur near education related keywords.

This exploratory search approach immediately gave us a wealth of entry points into the education crisis narratives and also helped us refine our specific questions. For example, the spike in speeches in the 2000s was associated with the “skills crisis” that was said to be limiting economic growth during the mining boom because of a shortage of appropriately skilled tradespeople, while the smaller spike in the 1960s was associated primarily with the Whitlam opposition’s critique of government policy. Even the small number of speeches prior to 1950 that used such narratives were interesting because they linked education and training to national preparedness for crises like war. Based on these

²⁵ ‘Australian Curriculum V9’. Accessed at: <https://www.australiancurriculum.edu.au/>. [accessed 28 August 2026]

overview analyses enabled by access to the proceedings as a data collection we have elected to focus our further work on the changes in education ‘crisis narratives’ during the transition from the Whitlam Opposition to the Whitlam Government.

Helping Researchers Make the Most of Proceedings

Our work across the complete Federal Parliamentary Proceedings has enabled us to consider new approaches to studying complex topics in large collections. By working across an entire large collection systematically with computational methods, we were able to gain a better understanding of our topics, while also identifying high value areas for more intensive close reading and analysis. We believe elements of the methods we employed in these case studies are applicable to other research contexts and questions. However, there are two key factors that enabled us to do this work at all. Firstly the extensive and ongoing work that the Federal Parliamentary services such as the Parliamentary Library do in transcription and preparation of new proceedings and the digitisation of old proceedings. Secondly we have a team with the appropriate technical and research skills to reconceptualise what would otherwise be more constrained, close-reading humanities projects, as such methods could be significantly enhanced through large scale data processing and preparation methods.

Our analyses were not possible without the combination of these two factors, and the ability to create a complete data collection out of Federal Hansard. But most researchers, let alone broader society, do not have access to the resources and expertise to work in this way. Many people’s experience of the proceedings are completely determined by what the web interface allows (or more importantly, doesn’t allow). While we shouldn’t forget that web based access and search systems are significant improvements over the distribution of printed proceedings, this does not mean we should not continue looking for ways in which to further improve access to and use of those proceedings.

One example is the prominence of search as point of entry for proceedings. Mitchell Whitelaw has critiqued conventional approaches to digital collections where search is framed as the main entry point for access, concluding that search is inadequate and that “Keyword search is ungenerous: it demands a query, discourages exploration, and withholds more than it provides.”²⁶ If search is the only entry point for a collection, we certainly agree that it would be inadequate, as we found when trying to use search systems for initial exploratory analysis. If, however, we conceptualise search systems as powerful tools for a *specific set of purposes* (from very specific tasks like verifying a quote all the way through to search based reading as a new mode of textuality),²⁷ and also broadly consider other ways to enable access *in addition to* search, then we are likely to be able to improve how researchers access and work with proceedings. We think Gerben Zaagsma’s question, “How to digitally enable serendipity?”²⁸ in the context of digital archives is a key question to continue to explore.

26 Mitchell Whitelaw, ‘Generous Interfaces for Digital Cultural Collections’. *Digital Humanities Quarterly* 9(1) 2015, pp.1–16. <https://doi.org/10.63744/4pg836gkjaeb>.

27 Alan Bilansky, ‘Search, Reading, and the Rise of Database’. *Digital Scholarship in the Humanities* 32(3) 2017, pp. 511–527. <https://doi.org/10.1093/lc/fqw023>.

28 Gerben Zaagsma, ‘Digital History and the Politics of Digitization’. *Digital Scholarship in the Humanities* 38(2) 2023, pp.830–851. <https://doi.org/10.1093/lc/fqac050>.

One Size Doesn't Fit All

With our own experiences as a starting point, and under the banner of the Reusable and Accessible Public Interest Documents (RAPID) project we have been having conversations with a broad cross section of researchers in Australia who are interested in working with parliamentary proceedings generally and Federal Parliament specifically.

One of the most common complaints reflected on the lack of affordances for enabling more systematic engagements with the full text of all or selected parts of proceedings. For quantitatively inclined researchers like corpus linguists this might look like access to complete components of the proceedings in data standards appropriate for large scale data analysis (cite: parlamint). For qualitatively inclined researchers in fields like history or political science, systematicity might look like being able to download the full text of a set of speeches matching a search query²⁹, in a format that encourages browsing and close reading. The search query affects a content based sampling strategy that contains the speeches in the direction of the research question. While analysing materials researchers want to be able to return again and again to a particular well specified slice of materials rather than navigating repeatedly through a web interface.

Beyond this desire for systematicity we find that the needs of researchers become specialised along more disciplinary and methodological lines. Political scientists and sociolinguists might be particularly interested in different subsets of information about parliamentarians such as age, gender, party and electorate. However, while this information may be partly present in the transcripts themselves or available in places like the Federal Parliamentary Handbook³⁰, it is not necessarily consistent enough within the transcripts or cross-linked with the search and filtering interfaces of the proceedings. Researchers in disciplines that focus considerable analytical effort on a small number of items using methodologies like critical discourse analysis, still want to be able to read that speech in a broader context: a PDF of a single speech may not capture everything they need but the PDF of the entire session is too unwieldy. It is not uncommon for materials to be printed for close reading because the digital presentation is presented at an awkward level of granularity.

Web based API's are commonly presented as a solution to address some of these problems³¹, but they are unlikely in and of themselves to directly address any of the issues above. It is rare for a collection API to depart from the logic of display and navigation that a conventional website interface would. Certainly more convenient data representations for computational access are beneficial, but given that most people lack the technical skills to make use of web API, there may be more effective places to invest resources.

Access from a Broader Perspective

While it would be easy to say that the particular complaints about access aired here are specific to the needs and workflows of researchers studying parliamentary proceedings, and that parliaments have primary responsibilities to broader society rather than researchers, we believe it's important to

29 Of course, this is assuming they have muddled through the earlier exploratory phases to be able to formulate such a query.

30 'Home - Parliamentary Handbook. Accessed at: <https://handbook.aph.gov.au/>. [accessed 28 August 2026]

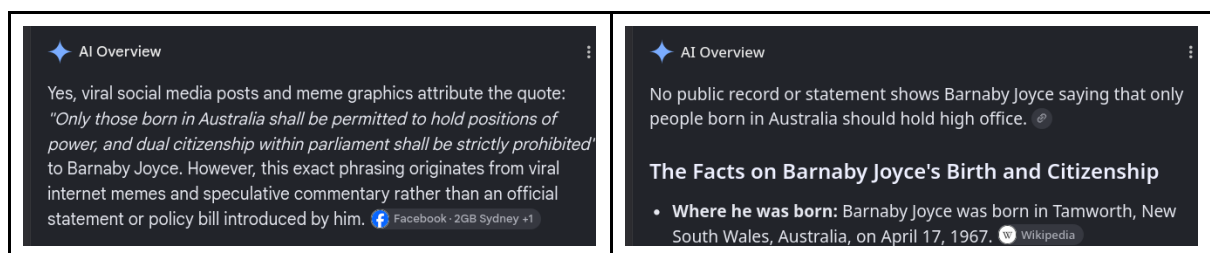
31 Robin Howlett, 'Online Publication and the Accessibility of Parliamentary Information'. *Australasian Parliamentary Review* 41(1) 2026, pp. 103–119.

consider access from a more holistic perspective. If there is no one size fits all approach for researchers, then it is hard to see how there can be one size fits all for all the stakeholders in a democratic society. It stands to reason that a civics and citizenship educator working with high school students, a corpus linguist, and a journalist are likely to benefit from different approaches to the same parliamentary proceedings.

In this broader context of access, mis- and dis-information are prominent challenges for democracy. As one example, a coordinated campaign identified by AAP FactCheck in March 2026³² was publishing politically polarising messages targeted at Australian audiences through a network of websites and social media pages. Posting polarising messages generated significant engagement for the pages involved in the campaign.

One example narrative falsely claimed that Barnaby Joyce had introduced a bill that would make dual citizens ineligible to sit in parliament. While someone with a passing memory of Australian politicians being ruled ineligible due to dual citizenship might suspect there is something off about this narrative, we have to consider the steps a person would need to take to specifically verify whether a bill like that was introduced in parliament, or whether words to those effect were even spoken at all. Verifying that something *was* said in parliament is directly supported by proceedings search interfaces: keywords, phrase search and filtering by date can be used to quickly identify material that exists. What is more difficult is demonstrating the *absence* of a statement (i.e., that a claimed statement never occurred).

In the broader information landscape the typical person skeptical of a claim and wanting to research further is more likely to turn to a search engine than directly to the proceedings. If we start with a search that mirrors the disinformation claim the same query can produce many different AI overviews as shown in Figure 3. In this particular case, while Barnaby Joyce never introduced the particular bill claimed, the AI overviews generate a range of interpretations from immediate confirmation of the false claim (first panel, but the statement is attributed to “viral internet memes”) to a variety of rejections and contextualisations. His history and statements about related issues including his own constitutional ineligibility to sit in parliament, combined with indexing of, and commentary on, disinformation material leads to significant variation in the overview.



32 ‘Vietnam Disinformation Pages Masquerade as Australian News Sites | AAP’. AAP Factcheck. Accessed at: <https://www.aap.com.au/factcheck/vietnam-disinformation-pages-masquerade-as-australian-news-sites/>. [accessed 27 August 2026]

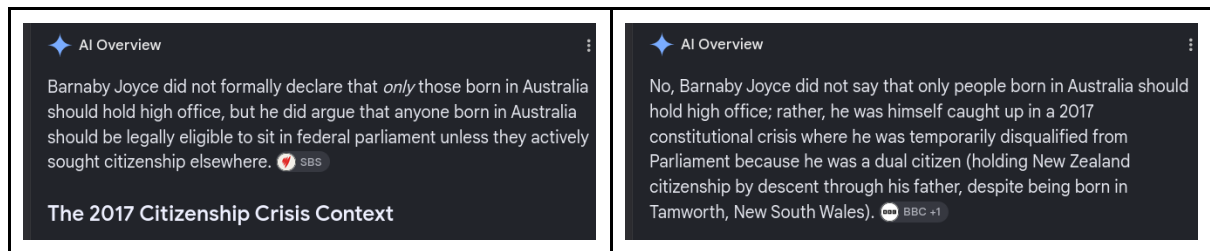


Figure 3: Four Google AI overviews produced in response to the query: *did barnaby joyce say only those born in Australia should be allowed to hold the highest offices of authority.*

We also note from conducting this exercise over the last three months that the variety of sources used to ‘support’ the claims in the overview have varied significantly. While we originally found the AAP FactCheck article on this campaign through the AI overview, we have not been able to repeat that since. Similarly while the proceedings themselves are indexed and can be surfaced through the overview, we did not observe them referenced in this particular case. This may simply be because the proceedings do not support the claim.

To us this exercise further supports the idea that access to parliamentary proceedings, whether as direct records of democracy or for their further value, requires a range of deliberate approaches and potentially different access mechanisms. If we assume that one single access point is sufficient for all use cases we are likely leaving people unprepared to critically evaluate misinformation.

Parliamentary Proceedings as Data Collections to Support Transformative Access

If access to parliamentary proceedings is best supported through a range of different approaches, how can we achieve this? As we have already discussed, web APIs might be useful but they are not fully transformative, and through our own research we have demonstrated that being able to work across an entire proceedings in bulk fundamentally enables new kinds of exploratory approaches at scale. We argue that conceptualising parliamentary proceedings as data collections is one way to support and encourage the variety of access mechanisms we collectively need.

Importantly, work towards this goal can proceed incrementally. Rather than creating a ‘perfect’ view of proceedings from a data perspective it may be enough to prepare complete collections of proceedings ‘as-is’ for others to work without needing to laboriously re-assemble proceedings from individual items. Further work (in collaboration with stakeholders) can progressively improve how proceedings are shared in bulk by identifying the most useful bulk data formats, or improving how procedural nuances are explicitly represented. The text heavy documents that make up proceedings are typically small and amenable to sharing in bulk: Federal Parliament’s 125 years of proceedings can fit in less than 10GB of data (depending on the file formats used).

Even if data volume was a concern, many governments already operate data sharing infrastructure (for example, data.gov.au³³) which directly supports bulk dissemination of data. Consistent publishing under appropriate open licences would also enable materials to be more transformatively shared, meaning that parliaments become publishers of data to a collaborative network of downstream users

33 ‘About | Data.Gov.Au’. Accessed at: <https://data.gov.au/about>. [accessed 30 August 2026]

and organisations, each with their own audience and networks that complement and support parliamentary engagement efforts. At the same time, a network or ecosystem of downstream data users allows more rapid experimentation with access methods that are not dependent on, or conflicting with the critical resources the public service needs to produce high quality proceedings.

This approach may not be without risks. As has been pointed out, there are risks of trust and credibility when parliamentary information is misused by presenting information out of context or in a misleading fashion.³⁴ While this risk is worth consideration, we must also weigh risks against benefits, as there are also considerable benefits that accrue from making proceedings as a whole easier to access and work with. At the same time, a disinformation campaign can be a pure fabrication, have a negative impact on a democracy, and parliaments may have little effective recourse to respond to such fabrication if coordinated and delivered over the web and social media. In light of the risks to society from the current information landscape we think consideration and broader sharing of parliamentary proceedings as data collections can only strengthen democratic engagement.

³⁴ Robin Howlett, 'Online Publication and the Accessibility of Parliamentary Information'. *Australasian Parliamentary Review* 41(1) 2026, pp. 103–119.