The topics and frames of parliamentary debates mentioning artificial intelligence

HHYY7

2024-04-21

1 Introduction

New technologies present both solutions and problems for society (DRCF 2022). The most powerful technologies can create a shifts in the technological paradigm (Perez 2010), which can improve the lives of many. However, they can cause hardships for others by undermining the fragile systems (society, employment, natural world) that people depend on. Governments respond by investing in and regulating technologies, aiming to enhance their benefits and reduce their disadvantages. Socio-technical landscapes can be complex and dynamic (Peters 2017) and it is difficult to determine which policies, if any, will have the desired outcomes (Flanagan and Uyarra 2016). Further, decisions and action taken depend on the definition and framing of the issue (Jasanoff (2003), p240-1), and competing interests can vie to influence political discourse (e.g. Gebru (2023), 00:17:22; Espinoza (2022)) and the policies enacted by governments. Arguably, this has a greater impact on decisions than empirically-derived evidence for policy outcomes (Dowding 2018; Kingdon 1993). This paper considers the framing of political discourse on artificial intelligence (AI) in UK parliamentary debate to understand how this has changed over time and whether influences over government can be identified.

1.1 Background

Academia has studied fluctuations in how technology is experienced and responded to (e.g. Perez (2002), Kerschner and Ehlers (2016), Edgerton (2019)). Such fluctuations determine the level of action taken, such as investment in technology and algorithmic regulation. AI experienced several "Winter" and "Spring" periods (Mitchell 2021). The "AI Winter" of the late 1990s and early 2000s was caused by earlier overconfidence and then disappointment in the technology (Mitchell 2021).

More recently, developments in technologies that produce language models able to generate convincing responses to human inputs, have caused a very rapid resurgence in focus on AI. The UK government position on AI moved rapidly from its pro-innovation stance developed over 2017 - 2022 (UKGov 2017; Office for Artificial Intelligence 2021; AI Council 2019, 2021, 2022), in which the only concerns were negative *narratives*, to a mix of excitement and existential worry in 2023 (DSIT 2023c, 2023a). It has been argued that this shift to a greater concern about risks and harms was largely due to lobbyists (such as the 2023 call for an "AI Pause"; Bengio, Russell, et al. (2023); Hogarth (2023)), who developed a narrative around AI due a range of motivations from genuine concern about risks to humanity, to a drive to create a superhuman image of their technology (Johnson 2023). Others have argued for a more nuanced approach to regulation (Ng 2023a, 2023b), or have threatened *capital strike* in response to regulation (Vincent 2023), or have continued to press for better awareness and control in areas that are concerning and affecting ordinary citizens today (Peppin 2022; Davies and Birtwistle 2023; The Ada and ATI 2023; Connected by Data 2023). It is difficult to determine why some of these narratives result in a a vigorous response from government (such as the existential risk narrative leading to the AI Safety Summit in 2023, Criddle, Murgia, and Gross (2023), DSIT (2023b)]), while other narratives appear largely unheeded (such as the many algorithmic safety reports and policy proposals developed within and outside of government over recent years, e.g. Peeters and Widlak (2018), Eubanks (2019), Ada Lovelace Institute and DataKind UK (2020), Bucknall and Dori-Hacohen (2022), DRCF (2022)).

Policy analysis is at least as much about argumentation as it is about analytical practice (Fischer and Forester (1993), p7). Understanding the interplay between narratives and policy is of interest to anyone investigating the effectiveness of the political processes, as well as those looking to influence policy. Spoken and written discourses from political actors have been used to derive their positions on, and trends in, topics of political interest. For instance, (Gabbatiss 2019) demonstrated that "climate change" is a phrase more often used by opposition parties and found its use rose steeply following protests in April 2019 (Barasi 2019). Onyimadu et al. (2014) and Abercrombie and Batista-Navarro (2018) both attempted to extract sentiments from Hansard data on UK parliamentary discourse. Other work has searched for topics within political documents, such as US presidential debates (Haddadan et al. 2023), German political manifestos and coalition contracts (Zirn and Stuckenschmidt 2014), and a database of political arguments (Ajjour et al. 2019).

Sentiment analysis, which detects opinions and/or emotions in data (Yadollahi, Shahraki, and Zaiane 2017) may be rather ambitious with political discourse, since the true positions and motivations of political actors are often guarded, resulting in sarcastic expressions and multifaceted sentences (Onyimadu et al. 2014). Ajjour et al. (2019) describe how debates are *framed* by selecting arguments to emphasis or obscure aspects of the debate topic. Identification of discourse topics and their frames may be more productive as these directly affect in which government department the policy is developed and how it is delivered and governed.

1.2 Outline

Inspired by the work of Zirn and Stuckenschmidt (2014), Ajjour et al. (2019) and Haddadan et al. (2023), this paper attempts to extract topics and their frames from UK parliamentary debates by applying text mining techniques, described in Silge and Robinson (2017) to UK Hansard data. Focussing on the topics and framing of terms related to "artificial intelligence", text search was applied to return debates in the Hansard corpus in which these terms were mentioned. Topic modelling and frame extraction were applied to the returned debates. Two contrasting search terms were identified that had different occurrences over the last century of parliamentary debate. For these two search terms, the debates had similar, but not identical, topics and some frames could be extracted.

The rest of the paper will describe the data and method used (Section 2), the results of text search, topic modelling and frame extraction are presented in Section 3 and these are discussed in Section 4 before drawing conclusions in Section 5.

2 Method

The following sections describe the data, how they were processed and the premises of these methods.

2.1 Data

TheyWorkForYou (TWFY; https://parser.theyworkforyou.com/hansard.html) maintain Hansard records in an accessible folder of xml files, which contains debates dating back to 4th February 1919. These were downloaded (on 22nd February 2024) and processed to R `dplyr` tibbles (one per xml file) to extract debate dates, headings, utterances and their speakers, and other information, before saving to RDS data files. Because of the nature of the xml structure, in which all headings (oral, major and minor) and utterances are at the same node level, headings pertaining to each utterance were derived from the preceding nodes of each type of heading. Thus there is a possibility that some utterances were mis- or unlabelled.

A simple hierarchy of utterances was devised by which a *debate* was considered to be all the utterances found in the same xml file (therefore on the same date), with the same major and minor heading. Each debate contained one or more *speech*, each of which was the combination of all the utterances in the debate by the same speaker. Whilst the data were saved as utterances, searches were performed over debates and speeches.

2.2 Data processing

Text search was performed by first combining the utterances into debates or speeches and extracting those that contained the desired *search term* (see Section 2.2.1). Initial analyses investigated the frequency of debates containing these terms over time. LDA was applied to identify *topics* in the extracted debates (Section 2.2.2). These were manually labelled and the frequencies of debates of each topic was examined. Finally, LDA was used to extract opposing *frames* of speeches within each topic group (Section 2.2.3) and, where possible, these frames were manually labelled.

2.2.1 Search terms

Search was performed over the data set using a set of terms related to AI. This set was refined (see Appendix A.1) to two main terms, referred to hereafter as: "artificial intelligence+" (which are debates mentioning "artificial intelligence", "machine learning" and/or "neural network") and "robot" (debates mentioning "robot").

2.2.2 Topic modelling

Using, Latent Semantic Analysis, Ajjour et al. (2019) (p2926) found that better results were achieved when the context of the whole debate was enabled in their analysis and so the "document" for the following analysis was the *debate*.

For each search term, the retrieved debates were tokenised into words. TF-IDF (Silge and Robinson (2017), Chapter 3) was applied to the debates (using the `bind_df_idf()` function in the R `tidytext` package), and the words that were most common across debates were removed by removing all but the 200 words with the highest TF-IDF in each debate.

Latent Dirichlet allocation (LDA) was applied to identify topic clusters (Silge and Robinson (2017), Chapter 6; using the `LDA()` function in the R `topicmodels` package). The number of topics was arrived at by performing LDA with a range of values of k and investigating the dominant words in each cluster (described in Appendix A.2). A value of $k = 10$ was found to produce a range of ostensibly coherent and distinct topics. These modelled topics were assigned text labels that best described the theme of the most words in the cluster (Appendix A.2.1). A further analysis was performed to understand how much any debate was a mix of topics (Appendix A.2.2). This found that the majority of debates had a greater than 50% probability of pertaining to a single topic and therefore each debate was assigned this majority topic as its topic label.

2.2.3 Frame extraction

Frame extraction was performed on the speeches (as "documents") within the debates assigned to a single topic. Based on Ajjour et al. (2019) and Haddadan et al. (2023), the assumption here was that, by performing LDA on speeches that were all labelled with the same topic, the differences in language would be due to topic *framing*. Further, a simple assumption was made that, within an adversarial political system, there are broadly 2 frames for any topic and thus `k`=2 clusters was specified. For each topic, the terms that most distinguished the two frames were analysed.

2.2.4 Random context inspection

Whilst working on the main analysis, a script was used to generate the sentence before, including and after, the mention of the search term. This enabled some qualitative understanding of small snippets of the data. Examples of these snippets are given in Appendix A.4.

3 Results

In total, there were 412 debates mentioning “artificial intelligence+” and 620 debates mentioning “robot”. Figure 1 shows the annual count of debates in which the search phrases were uttered. The patterns of utterance of the two search terms are somewhat contrasting. “artificial intelligence+” was first uttered in 27 November 1981, having peaks in 1986 and 2018. It was in use until the “AI winter”, which in parliamentary debate terms was between 1995 and 2012. After 2012, the use of “artificial intelligence+” rapidly increased, and possibly still is increasing. Last year (2023) there were 113 debates mentioning `a ast`. The use of “artificial intelligence+” dates back to 30 April 1926. It has been in continuous use since then, having a largely steady usage, with peaks in 1986 and 2018, both of close to 30 debates.

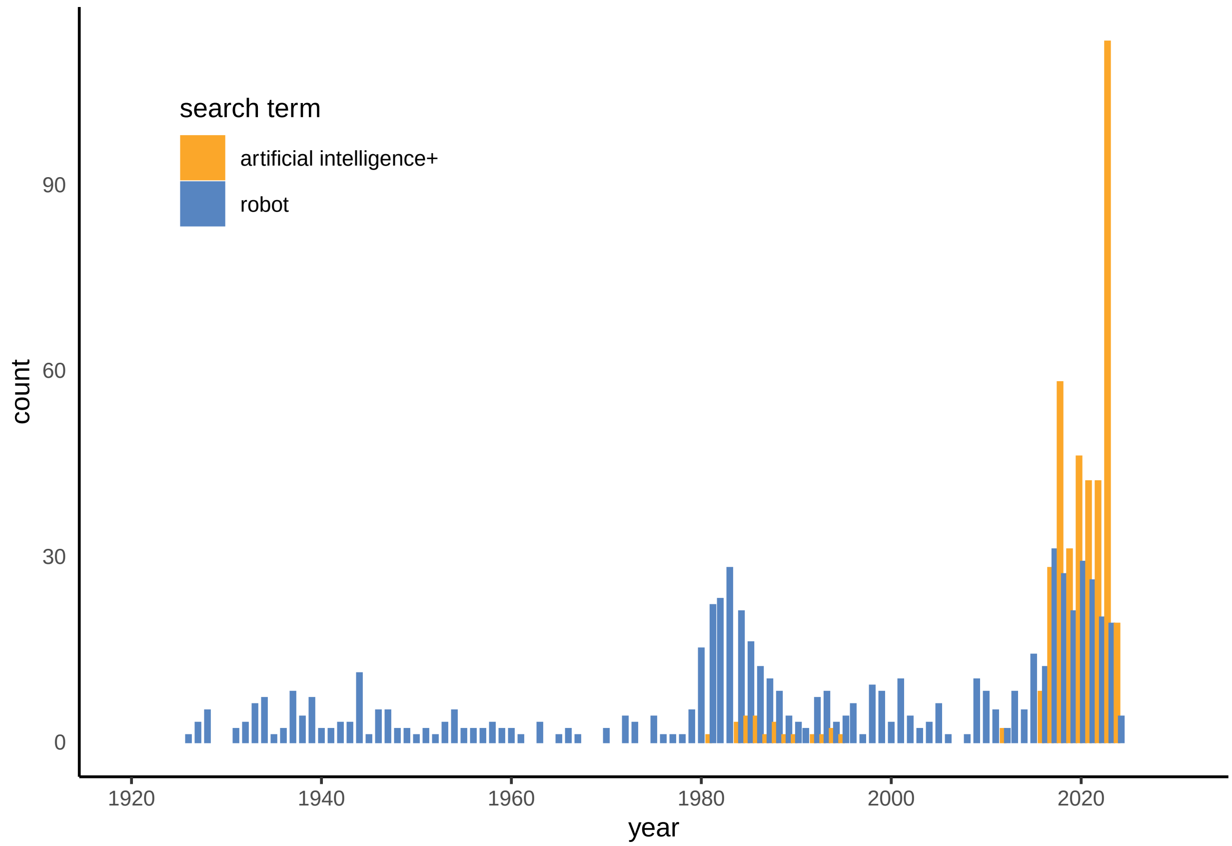


Figure 1: Count of parliamentary debates mentioning “artificial intelligence+” and “robot” by year

The sets of derived topic labels were subtly different between “artificial intelligence+” and “robot” (see Table 1). This range of debate topics for both search terms is quite broad, and covers much of the business of parliament. The main difference between the the sets of topics is that “artificial intelligence+” is mentioned in the context of “digital / online” - a set of issues that weren’t present for much of the timeline of the term “robot”. The most distinctly “robot” topics (those that didn’t emerge in the topic modelling for “artificial intelligence+”) are the very physical issues of “public services” (which often mentioned transport in the data), and “labour / housing”. It is worth noting that, although the topic did not arise in the topic model used for the final analysis, one “robot” topic model had distinctly *agricultural* words. Unfortunately, this model had other rather mixed and duplicated topics and so was not used for further analysis.

Table 1: Topic labels identified in "artificial intelligence+" and "robot" topics

artificial_intelligence_topics	robot_topics
science & technology	science & technology
state matters	governanance
int. trade / relations	int. trade
industrial strategy	int. relations
digital / online	labour / housing
police & crime	public services
education	education
health / care	health / care
economy	economy
defence	defence

Breaking down the debates into their respective topics, it is possible to see how these different contexts arise over time (Figures 2 and 3). Early (1980s) mentions of "artificial intelligence+" pertained to industrial strategy, defence, science and technology, and education. The first such utterance (see Appendix A.4.1) is in a debate with the heading: Armed Forces Technological Capabilities. The recent period over which "artificial intelligence+" has been so widely used apparently has two peak periods[1]. The early period, in late 2010s is more focused on police and crime, international trade and relations[2] and the economy. The second peak (which as previously noted, may not have ended yet), tends to be more about defence, education and digital and online matters.

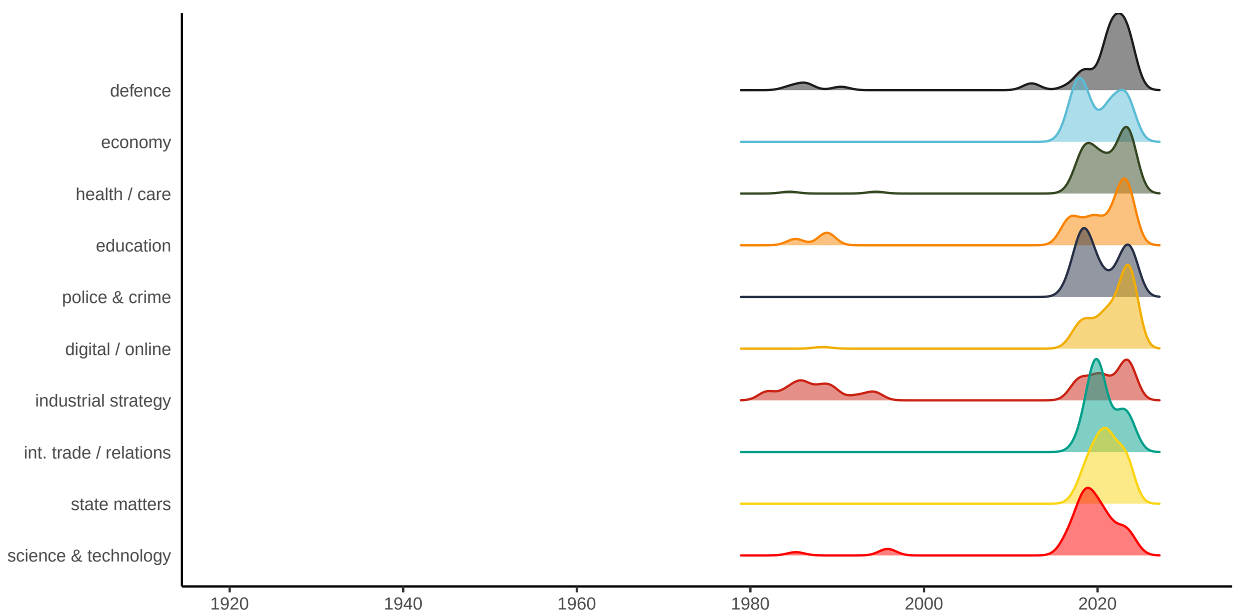


Figure 2: The different debate topics when "artificial intelligence+" is mentioned

Early mentions of "robot" were about labour and housing, international relations, and then during the years of World War II, defence became an additional "robot" topic. Defence was the dominant topic for "robot" in the 1960s. After a quiet period in the 1970s, "robot" had a renaissance in the 1980s, particularly in

[1]These peaks are a function of the smoothing filter used for the curves in Figures 2 and 3 but a "dip" around 2019 is also apparent in Figure 1

[2]Brexit was probably the biggest issue in parliament at this time

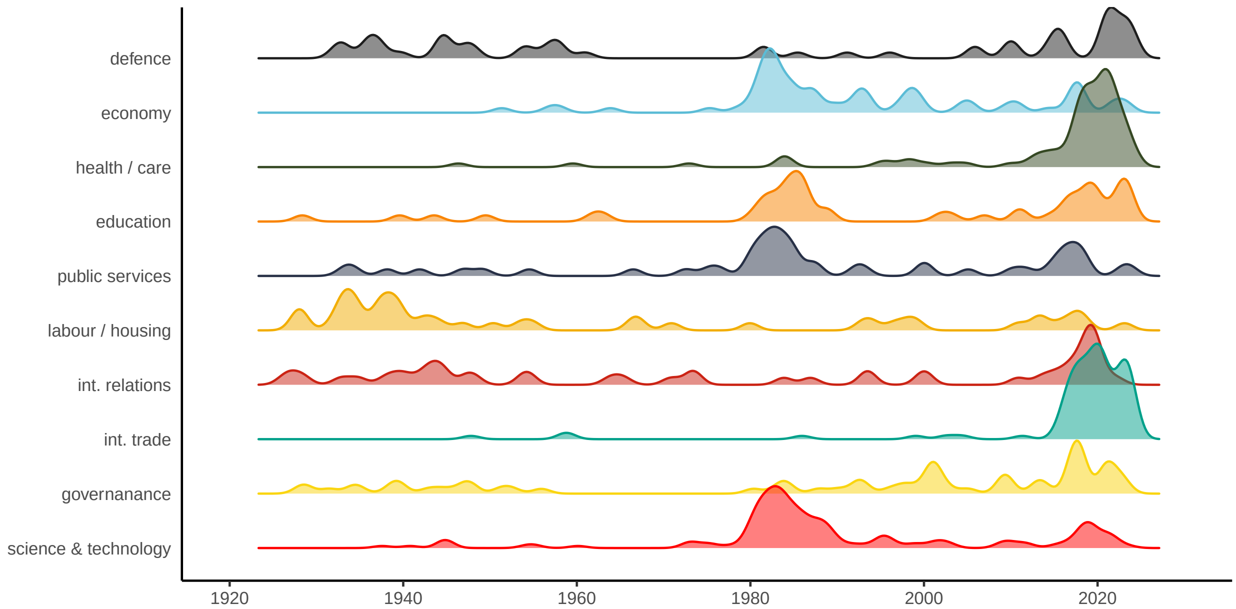


Figure 3: The different debate topics when "robot" is mentioned

relation to economy, education, public services, and science and technology. Perhaps for the same reasons that "artificial intelligence+" had a "Winter", "robot" was used less between the late 1990s and mid-2010s, before rising in use for all topic areas, but particularly health, education and international trade in the 2010s onwards.

The extracted frames are illustrated in Figures 4 and 5, for "artificial intelligence+" and "robot", respectively. Each plot in each figure represents the two frames extracted for one topic. The words on the y-axis are the words (minus stop words) that most distinguish the frames from each other, according to the LDA `beta` value. The bars to the right of the words represent these `beta` values, and the longer the line the greater the difference from the contrasting topic. For visual clarity, only 5 words for each frame are shown here (more detail can be viewed in the plots in Appendix A.3).

For many of these sets of words, it is not easy to identify a common theme that could be used to describe the framing. Tables 2 and 3 propose some possible themes from these sets of words. These suggested themes are derived from only a few words and in all cases there were words that did not match the proposed framing.

Table 2: Summary of frames suggested by more distinguishing words from "artificial intelligence" debates

topic	frame_1	frame_2
science & technology	(unclear)	(unclear)
state matters	royalty	crime
int. trade / relations	negotiation	how others see us
industrial strategy	what we offer	what we aspire to
digital / online	(unclear)	(unclear)
police & crime	investigative practice	criminality
education	support	(unclear)
health / care	technology	service
economy	(unclear)	(unclear)
defence	(unclear)	(unclear)

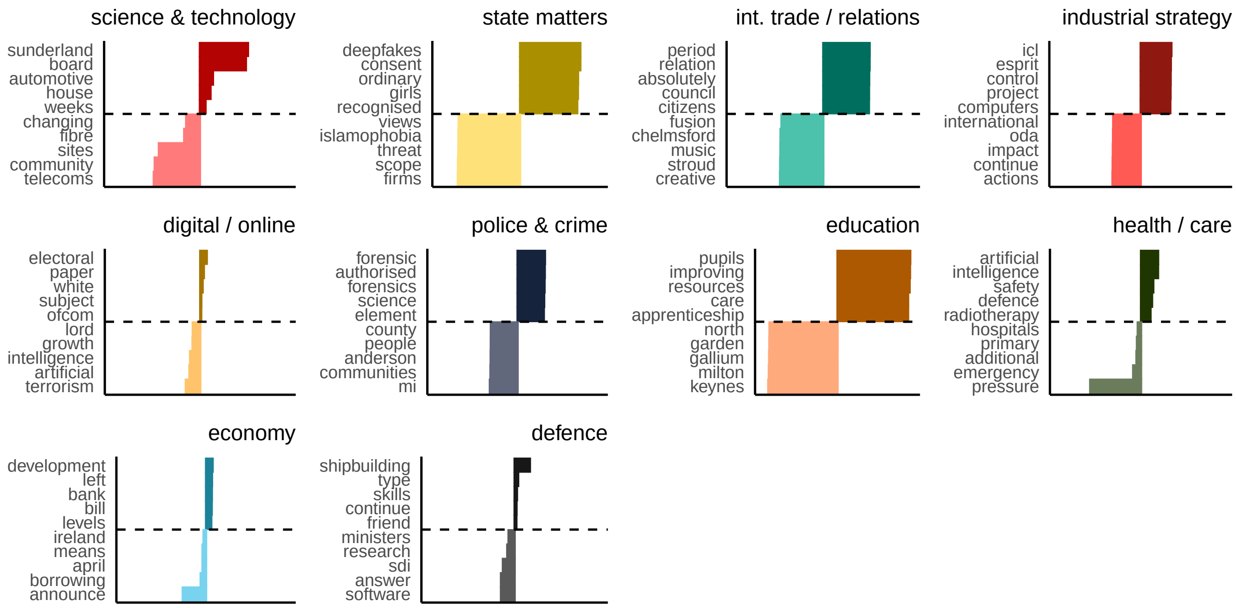


Figure 4: Pairs of frames extracted for each topic from debates mentioning "artificial intelligence"

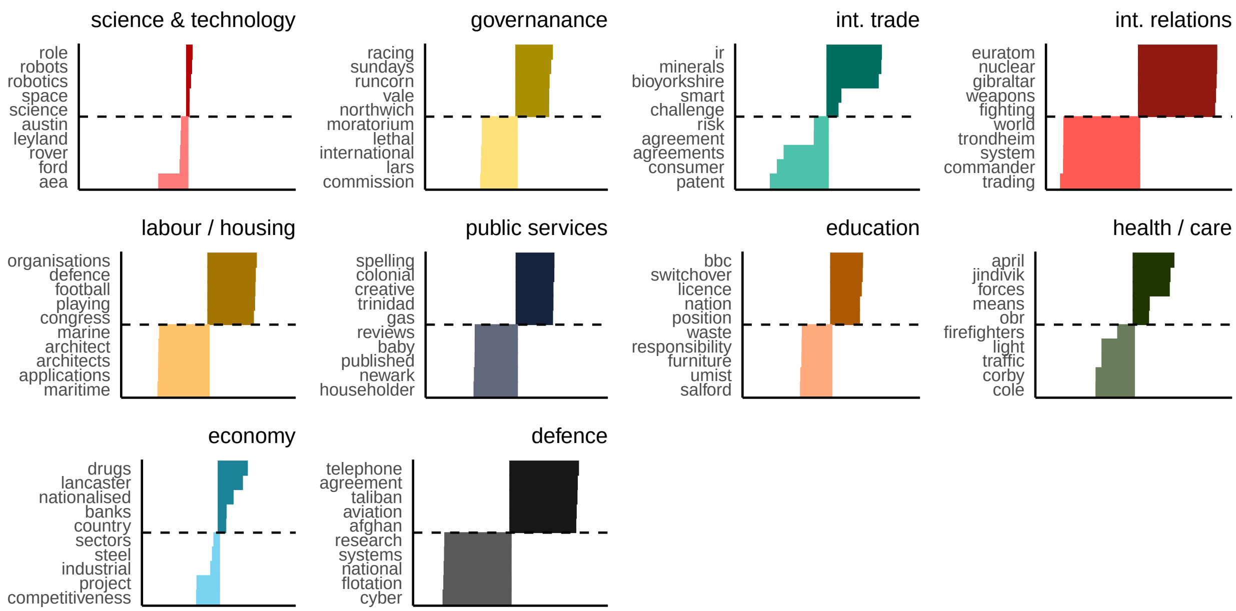


Figure 5: Pairs of frames extracted for each topic from debates mentioning "robot"

Table 3: Summary of frames suggested by more distinguising words from "robot" debates

topic	frame_1	frame_2
science & technology	(unclear)	(unclear)
governanance	(unclear)	(unclear)
int. trade	practicalities	negotiation
int. relations	dispute	diplomacy
labour / housing	(unclear)	(unclear)
public services	(unclear)	(unclear)
education	national	regional
health / care	financing	resourcing
economy	business	policy
defence	(unclear)	(unclear)

4 Discussion

The investigation described in the previous sections has produced some interesting insights into the nature of debates in UK parliament. This section considers the merits of the results presented and how they may be developed into more meaningful analysis.

The small set of search terms investigated fell into three groups. Removing those terms that were too broad, the resulting two groups provided a useful comparison of how terms are used in parliamentary debate, either steady with a small peaks or highly fluctuating. This may be related to the origins of the two terms - "robot" was a term conceived for a drama[3], whereas "artificial intelligence+" was devised to market the technology and thus generate a trend[4]. This appears to have led to hype cycles caused by expectations followed by disappointment (Mitchell 2021). Interestingly, whilst the term "artificial intelligence" wasn't used much during the "AI Winter", this was more a period of *not mentioning* "artificial intelligence" rather than *not doing* work in the field[5]. Like financial markets, technology has bubbles (Pedersen and Hendricks 2014) and there some who suggest that AI is a bubble (Yoon 2024; Naughton 2024). It would therefore be interesting to compare the patterns of these utterances to previous bubbles (such as the dot.com bubble), to identify any similarities.

The search terms applied in this work were not exhaustive and it may be valuable to extend the search to other related terms, such as "machine intelligence", "cybernetics", "network of neurons", "electronic brain", "intelligent machinery", "automaton", "android" and "intelligent machines".

The topic modelling on the debates and their occurrence over time reflects many of interests of the day. It seems that mentions of these technologies can arise within most matters of parliamentary debate. As noted in the analysis, food and agriculture appeared in the "robot" data. Similarly, manufacturing terms were evident in the words extracted for the "robot" frames. Labour and employment only appeared in a somewhat mixed topic for "robot", and apparently not for "artificial intelligence+". This is despite decades of threats that machines will be taking our jobs. One wonders if this indicates that parliament is not particularly concerned with a threat to jobs of these technologies. The differences between the topics derived for the two search terms may be because of the contrasting manifestation of the technologies: "robot" usually refers to a physical entity, whereas the technologies that make up AI are digital, as illustrated by the digital / online topic emerging from the "artificial intelligence+" debates. More work is needed to fully understand the context of the use of these terms, particularly to understand in which contexts these technologies are considered problems, or solutions.

When generating random search term contexts, it was noted that for some debates the modelled topic label didn't always to relate to its heading, even though in aggregate these appeared to align well (Appendix A.2.0.2). This doesn't mean that the modelled topics are incorrect, Onyimadu et al. (2014) note that digression from debate topics is commonplace. However, the comparison between topics and headings indicates a line of enquiry either for extracting more robust topics or for understanding the nature of debate digressions. The example in A.4.2 illustrates an apparent digression from the debate heading but also a lack of alignment between the topic label and the heading.

The approach to extracting frames naively extracts only two frames. Whilst there may be only two frames in any debate it's likely that there are many frames across debates. Also, manually identifying the frame from single words is problematic. Frames can be nuanced and full sentences are needed to understand their meaning. What was striking was that it was impossible to determine how many of them relate *at all* to the original search term!

The extracted frame words illustrate how mixed some of the topics seem to be. This may be realistic or an indication that the topic modelling needs improving. Another feature of the frames analysis was that the distance between the two frames was different for different topics, with some topics having short

[3] The term "robot" coined by Karel Čapek for a 1920's play (Love 2020)

[4] The term "artificial intelligence" coined by John McCarthy for the 1956 Dartmouth Summer Research Project on Artificial Intelligence (McCarthy et al. 1955; Bender 2023)

[5] The author developed neural network approaches for image processing during this period but would never have used the term "artificial intelligence"

bars and others very long ones. This may indicate topics for which there is a general consensus across parliament (where the bars are short) and those that are more contentious. For some topics, such as "defence", and "education" there is also a difference between the bar lengths for the two search terms ("artificial intelligence+" and "robot"), suggesting that the topic is more contentious in one data set than the other. Probably, this is less to do with the search terms than to the period in history of the debates. There is scope to extend this enquiry further, for instance by applying the approach to topics over time to identify periods of consensus and of opposition.

This initial study was not able to identify what influenced parliamentary discourse. This would require deeper investigations. For instance, it would be particularly useful to join other datasets to extract the party of the speaker and whether they were in government or opposition. Extracting terms relating to organisations or individuals who may influence parliamentary actors and policy would also provide insights into their influence. Further, it would be useful to compare the debates data with, for example, financial data and technology reports, such as those pertaining to the "bubble" and hype cycles, and data about discoveries and breakthroughs in technology. From all these it may be possible to find interactions with the peaks and troughs in parliamentary mentions. This analysis, once developed, would also be thought-provoking if applied to other terms related to human rights, health or environmental concerns (e.g. climate- and smoking-related search results are demonstrated in Appendix A.5).

LDA was used for its ease of application. However, it is not exactly suited to the analysis in this work. The underlying assumption of LDA is that a document addresses a number of different topics (Zirn and Stuckenschmidt (2014), p40). However, debates are supposed to be on single topics (barring the digressions mentioned earlier). Also, LDA assumes that topics are uncorrelated. This is not the case for most parliamentary topics. Therefore a different approach may be better, such as Correlated Topic Modelling.

5 Conclusion

Four main conclusions are drawn from this analysis:

1. The phenomenon of hype, and potentially bubbles, is apparent in parliamentary debates when counting utterances of particular search terms
2. Topics and frames extracted from debates mentioning "artificial intelligence+" and "robot" appear to have more to do with the business of parliament over time than the concept underlying the search term and as such they track changing concerns over time
3. The two-frame approach applied to each topic, whilst sometimes vague on the nature of the frame may indicate the level of consensus within debates on particular topics
4. Wider experimentation using more search terms, analytical approaches and data sets could yield greater insights into the parliamentary positions on technology and other matters

Text mining of UK Hansard records, whilst initially cumbersome, has yielded some fascinating insights into the nature of parliamentary debate. There are plenty of opportunities to extend this work, perhaps developing it towards a deeper understanding of how parliamentary decisions on contentious technologies are arrived at.

6 References

Abercrombie, Gavin, and Riza Batista-Navarro. 2018. "'Aye' or 'No'? Speech-Level Sentiment Analysis of Hansard UK Parliamentary Debate Transcripts." In *LREC 2018, Eleventh International Conference on Language Resources and Evaluation*. France: European Language Resources Association. https://research.manchester.ac.uk/en/publications/aye-or-no-speech-level-sentiment-analysis-of-hansard-uk-parliamen.

Ada Lovelace Institute and DataKind UK. 2020. "Examining the Black Box: Tools for Assessing Algorithmic Systems." Ada Lovelace Institute; DataKind UK. https://www.adalovelaceinstitute.org/report/examining-the-black-box-tools-for-assessing-algorithmic-systems/.

AI Council. 2019. "Artificial Intelligence Council Meeting 9th September 2019." Meeting minutes. Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport; Department for Business, Energy & Industrial Strategy. https://www.gov.uk/government/groups/ai-council.

———. 2021. "Artificial Intelligence Council Meeting 9th December 2021." Meeting minutes. Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport; Department for Business, Energy & Industrial Strategy. https://www.gov.uk/government/groups/ai-council.

———. 2022. "Artificial Intelligence Council Meeting 6th December 2022." Meeting minutes. Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport; Department for Business, Energy & Industrial Strategy. https://www.gov.uk/government/groups/ai-council.

Ajjour, Yamen, Milad Alshomary, Henning Wachsmuth, and Benno Stein. 2019. "Modeling Frames in Argumentation." In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processingand the 9th International Joint Conference on Natural Language Processing*, 2922–32. https://doi.org/10.18653/v1/D19-1290.

Barasi, Leo. 2019. "Extinction Rebellion's Protests Were an Unprecedented Success. Three Questions about What Comes Next." Noise of the Crowd Blog. April 28, 2019. https://www.noiseofthecrowd.com/extinction-rebellions-protests-unprecedented-success-three-questions-comes-next/.

Bender, Emily M. 2023. "Opening Remarks on 'AI in the Workplace: New Crisis or Longstanding Challenge'." October 2, 2023. https://medium.com/@emilymenonbender/opening-remarks-on-ai-in-the-workplace-new-crisis-or-longstanding-challenge-eb81d1bee9f.

Bengio, Yoshua, Stuart Russell, et al. 2023. "Pause Giant AI Experiments: An Open Letter." March 22, 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

Bucknall, Benjamin S., and Shiri Dori-Hacohen. 2022. "Current and Near-Term AI as a Potential Existential Risk Factor." In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. AIES '22. ACM. https://doi.org/10.1145/3514094.3534146.

Connected by Data. 2023. "AI Safety Summit: Open Letter to the UK Prime Minister." Connected by Data, The Trades Union Congress; Open Rights Group. October 30, 2023. https://ai-summit-open-letter.info/.

Criddle, Cristina, Madhumita Murgia, and Anna Gross. 2023. "UK to Host AI Safety Summit at Start of November." *Financial Times*, August. https://www.ft.com/content/1d3043db-b68d-496e-8fda-4654f8af15bc.

Davies, Matt, and Michael Birtwistle. 2023. "Seizing the 'AI Moment': Making a Success of the AI Safety Summit." September 7, 2023. https://www.adalovelaceinstitute.org/blog/ai-safety-summit/.

Dowding, Keith. 2018. "The Advocacy Coalition Framework." In *Handbook on Policy, Process and Governing*, 220–31. https://doi.org/10.4337/9781784714871.00020.

DRCF. 2022. "The Benefits and Harms of Algorithms: A Shared Perspective from the Four Digital Regulators." Research and Analysis. Digital Regulation Cooperation Forum. https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/the-benefits-and-harms-of-algorithms-a-shared-perspective-from-the-four-digital-regulators.

DSIT. 2023a. "AI Safety Summit: Capabilities and Risks from Frontier AI." Discussion paper. https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper.

———. 2023b. "AI Safety Summit: Introduction." September 25, 2023. https://www.gov.uk/government/publications/ai-safety-summit-introduction.

———. 2023c. "Frontier AI: Capabilities and Risks." Discussion Paper. Department for Science, Innovation; Technology. https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper.

Edgerton, David. 2019. "What Has British Science Policy Really Been?" In *Lessons from the History of UK Science Policy*, by Kieron Flanagan, S. Clarke, J. Agar, David Edgerton, and C. Craig, 31–39. https://www.thebritishacademy.ac.uk/publications/policy-histories-lessons-history-uk-science-policy/.

Espinoza, Javier. 2022. "Google in Last-Ditch Lobbying Attempt to Influence Incoming EU Tech Rules." *Financial Times*, January. https://www.ft.com/content/8c7527bc-7ab4-41cd-ba94-3145208da9c3.

Eubanks, Virginia. 2019. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* https://us.macmillan.com/books/9781250074317/automatinginequality.

Fischer, Frank, and John Forester. 1993. "Editors' Introduction." In *The Argumentative Turn: Policy Institutions and Practices.* Duke University Press. https://doi.org/10.1215/9780822381815-001.

Flanagan, Kieron, and Elvira Uyarra. 2016. "Four Dangers in Innovation Policy Studies - and How to Avoid Them." *Industry and Innovation*, March, 177–88. https://www.tandfonline.com/doi/abs/10.1080/13662716.2016.1146126?journalCode=ciai20.

Gabbatiss, Josh. 2019. "Analysis: The UK Politicians Who Talk the Most about Climate Change." Carbon Brief. September 11, 2019. https://www.carbonbrief.org/analysis-the-uk-politicians-who-talk-the-most-about-climate-change/.

Gebru, Timnit. 2023. "How to Make AI Systems More Just with Hilary Pennington and Dr. Timnit Gebru." March 21, 2023. https://www.youtube.com/watch?v=MgcUatmPnyE.

Haddadan, Shohreh, Elena Cabrio, Axel J. Soto, and Serena Villata. 2023. "Topic Modelling and Frame Identification for Political Arguments." In *AIxIA 2022 – Advances in Artificial Intelligence*, edited by Agostino Dovier, Angelo Montanari, and Andrea Orlandini, 268–81. Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-031-27181-6_19.

Hogarth, Ian. 2023. "We Must Slow down the Race to God-Like AI: I've Invested in More Than 50 Artificial Intelligence Start-Ups. What i've Seen Worries Me." *Financial Times*, April. https://www.ft.com/content/03895dc4-a3b7-481e-95cc-336a524f2ac2.

Jasanoff, Sheila. 2003. "Technologies of Humility: Citizen Participation in Governing Science." *Minerva* 41: 223–44. https://doi.org/10.1023/A:1025557512320.

Johnson, Rebecca. 2023. "X-Riskers Think Differently." August 1, 2023. https://ethicsgenai.com/x-riskers-think-differently/.

Kerschner, Christian, and Melf-Hinrich Ehlers. 2016. "A Framework of Attitudes Towards Technology in Theory and Practice." *Ecological Economics* 126: 139–51. https://doi.org/https://doi.org/10.1016/j.ecolecon.2016.02.010.

Kingdon, John W. 1993. "How Do Issues Get on Public Policy Agendas?" In *Sociology and the Public Agenda*, 8:40–53. 1. https://doi.org/10.4135/9781483325484.

Love, Damien. 2020. "Where Does the Word 'Robot' Come From?" BBC Sciene Focus. May 3, 2020. https://www.sciencefocus.com/future-technology/where-does-the-word-robot-come-from.

McCarthy, J., M. L. Minsky, N. Rochester, and C. E. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence."

Mitchell, Melanie. 2021. "Why AI Is Harder Than We Think." *arXiv:2104.12871Search...*, April. https://doi.org/10.48550/ARXIV.2104.12871.

Naughton, John. 2024. "From Boom to Burst, the AI Bubble Is Only Heading in One Direction." The Guardian. April 13, 2024. https://www.theguardian.com/commentisfree/2024/apr/13/from-boom-to-burst-the-ai-bubble-is-only-heading-in-one-direction.

Ng, Andrew. 2023a. "[Editorial]." *The Batch*, November. https://info.deeplearning.ai/cyberattack-strikes-openai-actors-reach-accord-on-ai-anthropic-goes-steady-with-google-and-amazon-1.

———. 2023b. "Written Statement of Andrew Ng Before the u.s. Senate AI Insight Forum." December 11, 2023. https://aifund.ai/insights-written-statement-of-andrew-ng-before-the-u-s-senate-ai-insight-forum/.

Office for Artificial Intelligence. 2021. "National AI Strategy." Guidance. Department for Science, Innovation; Technology, Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport,; Department for Business, Energy & Industrial Strategy. https://www.gov.uk/government/publications/national-ai-strategy.

Onyimadu, Obinna, Keiichi Nakata, Tony Wilson, David Macken, and Kecheng Liu. 2014. "Towards Sentiment Analysis on Parliamentary Debates in Hansard." In *Semantic Technology*, edited by Wooju Kim, Ying Ding, and Hong-Gee Kim, 48–50. Springer International Publishing. https://link.springer.

com/chapter/10.1007/978-3-319-06826-8_4.

Pedersen, David Budtz, and Vincent F. Hendricks. 2014. "Science Bubbles." *Philosophy & Technology* 27: 503–18. https://link.springer.com/article/10.1007/s13347-013-0142-7.

Peeters, Rik, and Arjan Widlak. 2018. "The Digital Cage: Administrative Exclusion Through Information Architecture - the Case of the Dutch Civil Registry's Master Data Management System." *Government Information Quarterly* 35 (2): 175–83. https://doi.org/10.1016/j.giq.2018.02.003.

Peppin, Aidan. 2022. "The Rule of Trust: Findings from Citizens' Juries on the Good Governance of Data in Pandemics." Ada Lovelace Institute. https://www.adalovelaceinstitute.org/report/trust-data-governance-pandemics/.

Perez, Carlota. 2002. *Technological Revolutions and Financial Capital*. Elgar. https://www.e-elgar.com/shop/gbp/technological-revolutions-and-financial-capital-9781843763314.html.

———. 2010. "Technological Revolutions and Techno-Economic Paradigms." *Cambridge Journal of Economics* 34: 185–202. https://doi.org/10.1093/cje/bep051.

Peters, B. Guy. 2017. "What Is so Wicked about Wicked Problems? A Conceptual Analysis and a Research Program." *Policy and Society* 36 (3): 385–96. https://doi.org/10.1080/14494035.2017.1361633.

Silge, Julia, and David Robinson. 2017. *Text Mining with r: A Tidy Approach*. O'Reilly Media.

The Ada and ATI. 2023. "How Do People Feel about AI? A Nationally Representative Survey of Public Attitudes to Artificial Intelligence in Britain." https://www.adalovelaceinstitute.org/report/public-attitudes-ai/.

UKGov. 2017. "Industrial Strategy Building a Britain Fit for the Future." Industrial Strategy White Paper. UK Government. https://www.gov.uk/government/publications/industrial-strategy-the-grand-challenges/industrial-strategy-the-grand-challenges.

Vincent, James. 2023. "OpenAI Says It Could 'Cease Operating' in the EU If It Can't Comply with Future Regulation." *The Verge*, May. https://www.theverge.com/2023/5/25/23737116/openai-ai-regulation-eu-ai-act-cease-operating.

Yadollahi, Ali, Ameneh Gholipour Shahraki, and Osmar R. Zaiane. 2017. "Current State of Text Sentiment Analysis from Opinion to Emotion Mining." *ACM Comput. Surv.* 50 (2). https://doi.org/10.1145/3057270.

Yoon, June. 2024. "AI Hype Has Echoes of the Telecoms Boom and Bust." Financial Times. February 14, 2024. https://www.ft.com/content/dc47c5f3-9bd4-4da0-a5cb-c795efd14c9c.

Zirn, Cäcilia, and Heiner Stuckenschmidt. 2014. "Multidimensional Topic Analysis in Political Texts." *Data & Knowledge Engineering* 90: 38–53. https://doi.org/https://doi.org/10.1016/j.datak.2013.07.003.

A Appendices

A.1 Selection of search terms

Initially, the debates were searched through for 7 search terms: “artificial intelligence”, “machine learning”, “computer science”, “deep learning”, “neural network”, “robot”, “expert system”. This set was refined by analysing the co-occurrence of each search term and our main search term “artificial intelligence” as follows:

5 debates mentioning “machine learning” but not “artificial intelligence”: 2019-01-14 - 2023-11-07

106 debates mentioning “computer science” but not “artificial intelligence”: 1965-12-07 - 2023-09-14

10 debates mentioning “deep learning” but not “artificial intelligence”: 1919-02-17 - 2012-09-17

3 debates mentioning “neural network” but not “artificial intelligence”: 1992-02-14 - 1994-06-23

540 debates mentioning “robot” but not “artificial intelligence”: 1926-04-30 - 2024-02-08

9 debates mentioning “expert system” but not “artificial intelligence”: 1981-06-17 - 2020-09-23

From these analyses, two groups of search terms were identified: “artificial intelligence+”, which incorporated results from “machine learning” and “neural network”, and “robot”. The other search terms were not used. All the decisions are described in Table 4.

Table 4: Summary of decisions made on different “artificial intelligence” search terms

search_term	reason
machine learning	very similar: combine with AI
computer science	too broad: remove
deep learning	most non-AI mentions are not relevant: remove
neural network	very similar: combine with AI
robot	relevant but different: search separately
expert system	most non-AI mentions are not relevant: remove

A.2 Determining the number and labels of topics

To determine the optimal number of topics, LDA topic modelling was applied to the debates with different values for `k`, the number of topics, and different seeds. Two approaches were taken to determine a satisfactory number of topics: "manual within-topic coherence" and "manual comparison with heading terms"

A.2.0.1 Manual within-topic coherence By inspecting the words ordered by their `beta` value for each LDA topic (the probability that a word will occur in text about that topic), it was often possible to identify a common theme in each topic's high probability words. This was achieved by extracting from the data stop words and common debate terms ("engagements", "topical", "questions", "government", "honourable"), and then plotting simple wordclouds of the words with the 20 highest `beta` values in each topic (with the word size pertaining to the word's `beta` value) (see Figures 6 and 8).

The satisficing value for `k` was chosen that demonstrated largely unmixed themes (few words that belonged better in different themes were present) and that didn't result in multiple topics with similar themes. Over several iterations, this value was identified as 10.

A.2.0.2 Manual comparison with heading terms The text from the debates' major and minor headings were *not* included in the topic modelling and so these could be used to manually verify the resulting topics. To achieve this, each debate was assigned to the topic for which it had the highest `gamma` (the probability that a debate belongs to that topic). For each of these debate topic groups, the words in the major and minor headings were extracted, and stop words, word with non-alphabetic characters and common heading terms ("engagements", "topical", "questions", "bill", "clause") were removed. A comparison between the topic words and the heading words was then performed again using simple wordclouds (see Figures 7 and 9).

This analysis is presented with a note of caution. Labelling was performed manually by one person. Also, the two datasets (topics and headings) were labelled with knowledge of the other's content. Further, there are known issues with using wordclouds. Therefore, this process is likely to be subject to bias. Future analyses should address these concerns to ensure that values for k and topic labels robustly defined (for example by starting with coherence analysis). Examples in the table are from 1 random seed and other topics emerge with different seeds and values of `k`.

A.2.1 Labelling topics

Following the above analysis, and using both sets of words (LDA topic words and debate heading words), text labels for each LDA topic were derived by selecting the a phrase that best described the theme in the highest probability (highest `beta`) topic words. These were assigned separately to "artificial intelligence+" and "robot" debate data but were then re-ordered so align similar topics between the two search term sets.

A.2.2 Probability that debates have a majority topic

A further analyis was undertaken to determine if it is reasonable to assign a single topic label to each debate. This used the LDA `gamma` values which indicate the probability that a debate has a particular topic. The topic for which each debate had the maximum `gamma` was extracted and the frequency of these values assessed. Figures 10 and 11 show this frequency for topics for "artificial intelligence+" and "robot", respectively, and illustrate that most debates have a greater than 50% probability of belonging to the single topic with the highest `gamma` value. Therefore it was considered reasonable to assign a single topic to each debate.

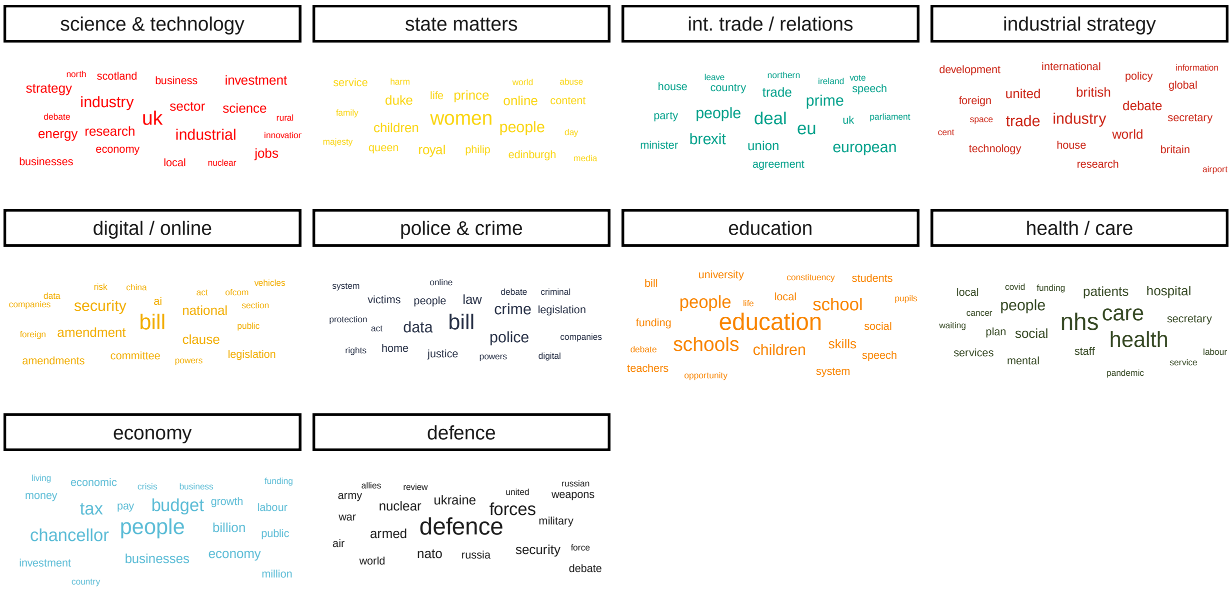


Figure 6: Wordclouds of "artifcial intelligence+" topic words for k = 10 (seed = 42)

Figure 7: Wordclouds of "artifcial intelligence+" heading words debates grouped by the LDA topic model with k = 10 (seed = 42)

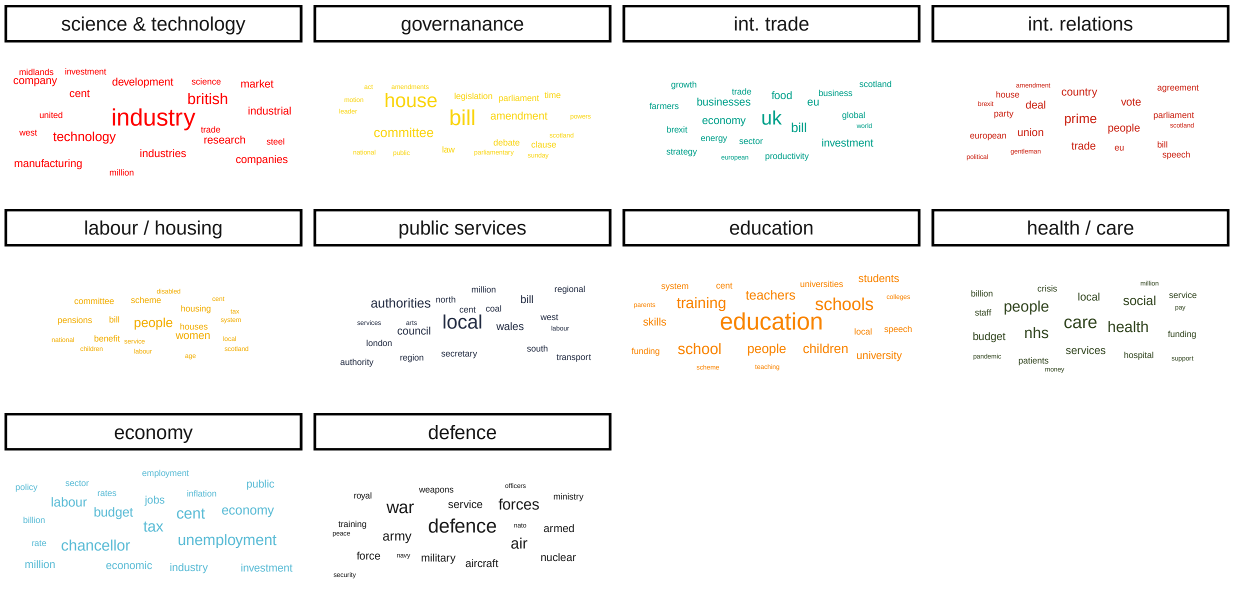


Figure 8: Wordclouds of “robot” topic words for k = 10 (seed = 1234)

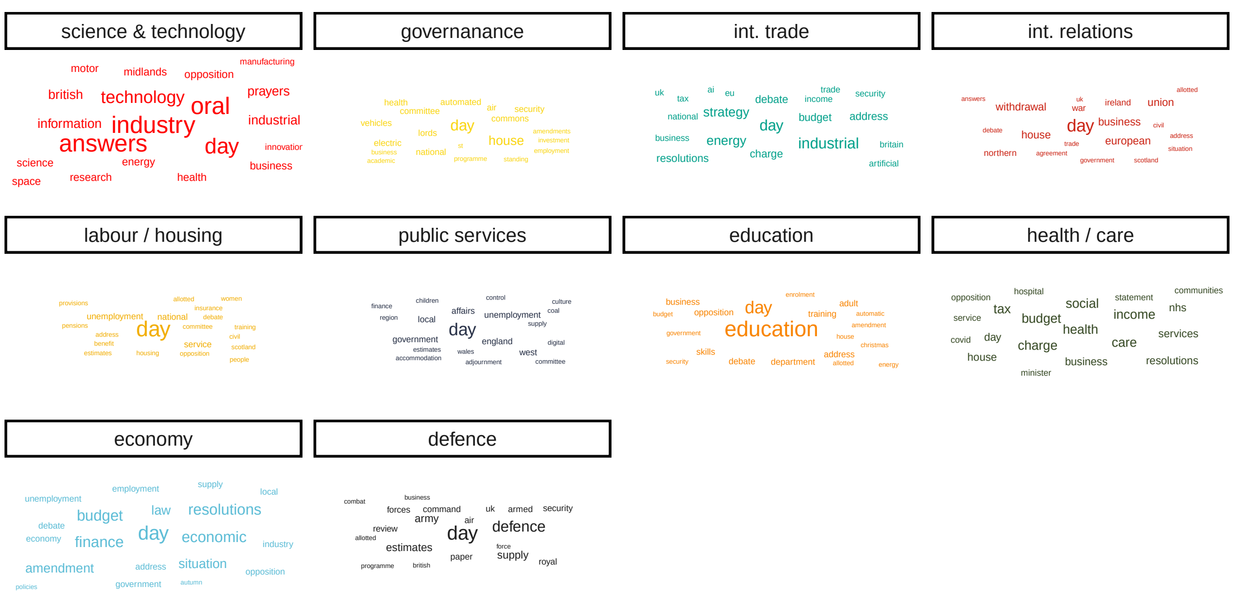


Figure 9: Wordclouds of “robot” heading words debates grouped by the LDA topic model with k = 10 (seed = 1234)

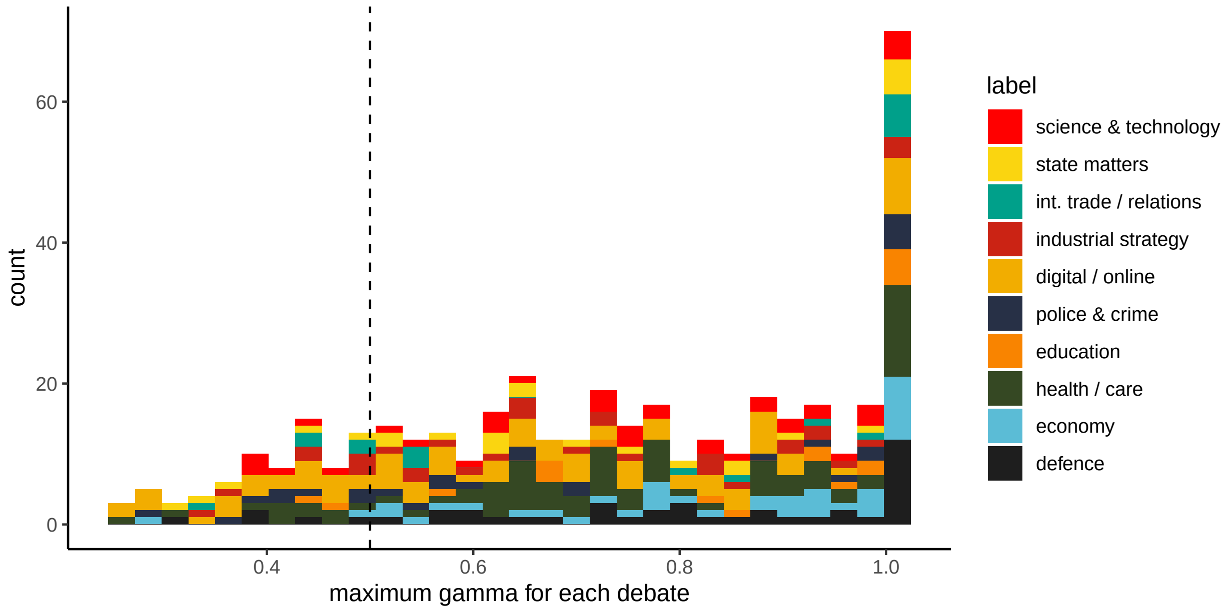


Figure 10: Frequency of probabilities that debates belong to the majority topic for "artificial intelligence" debates. x-axis is the probability (`gamma`) of the main topic, y-axis is the count of those probabilities

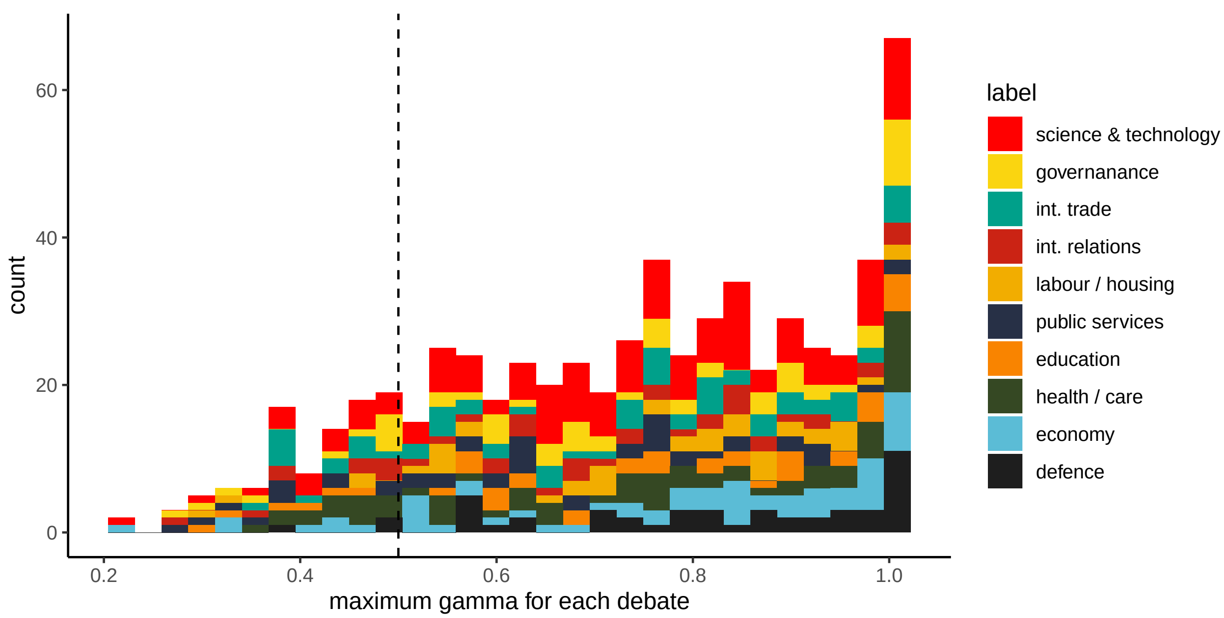
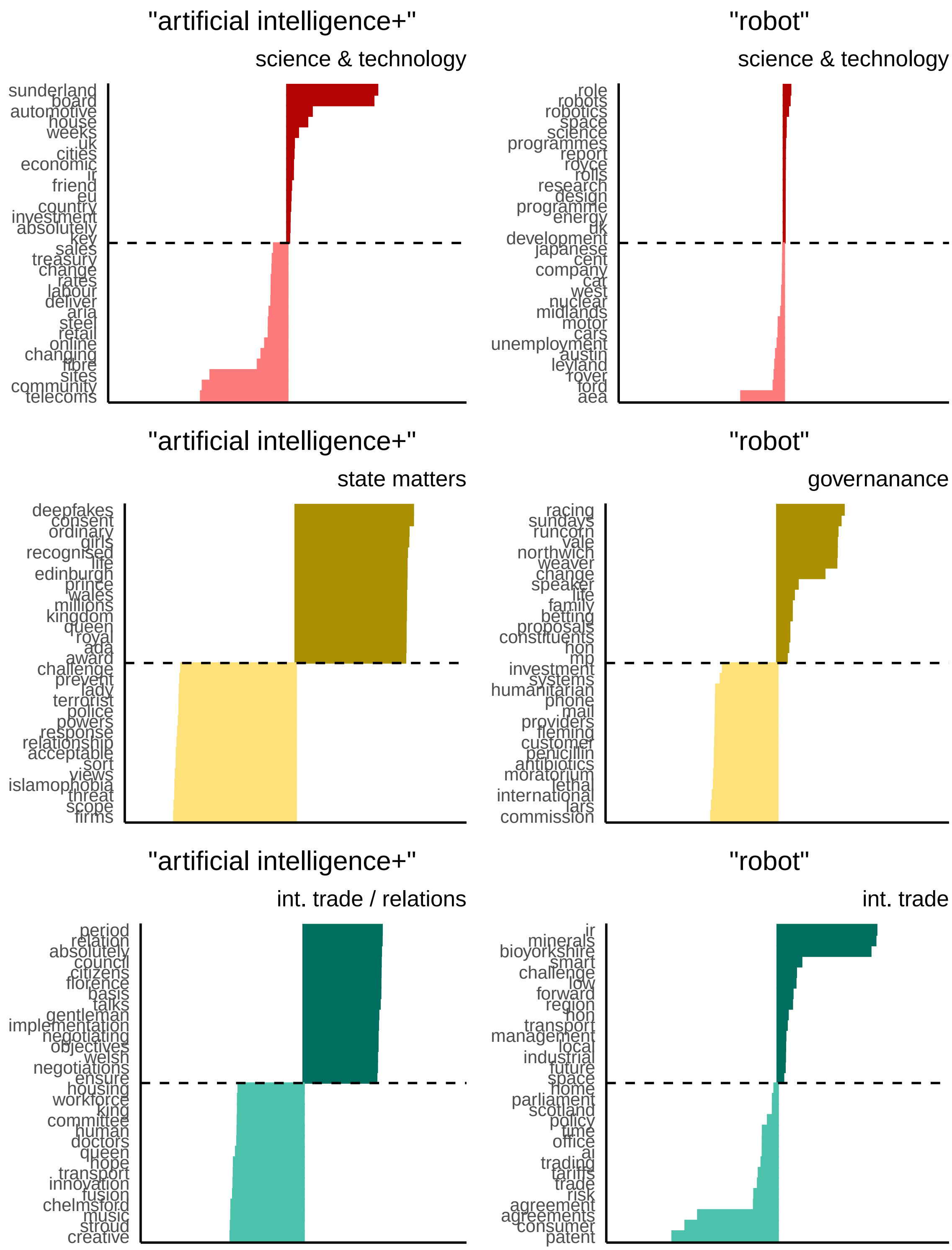
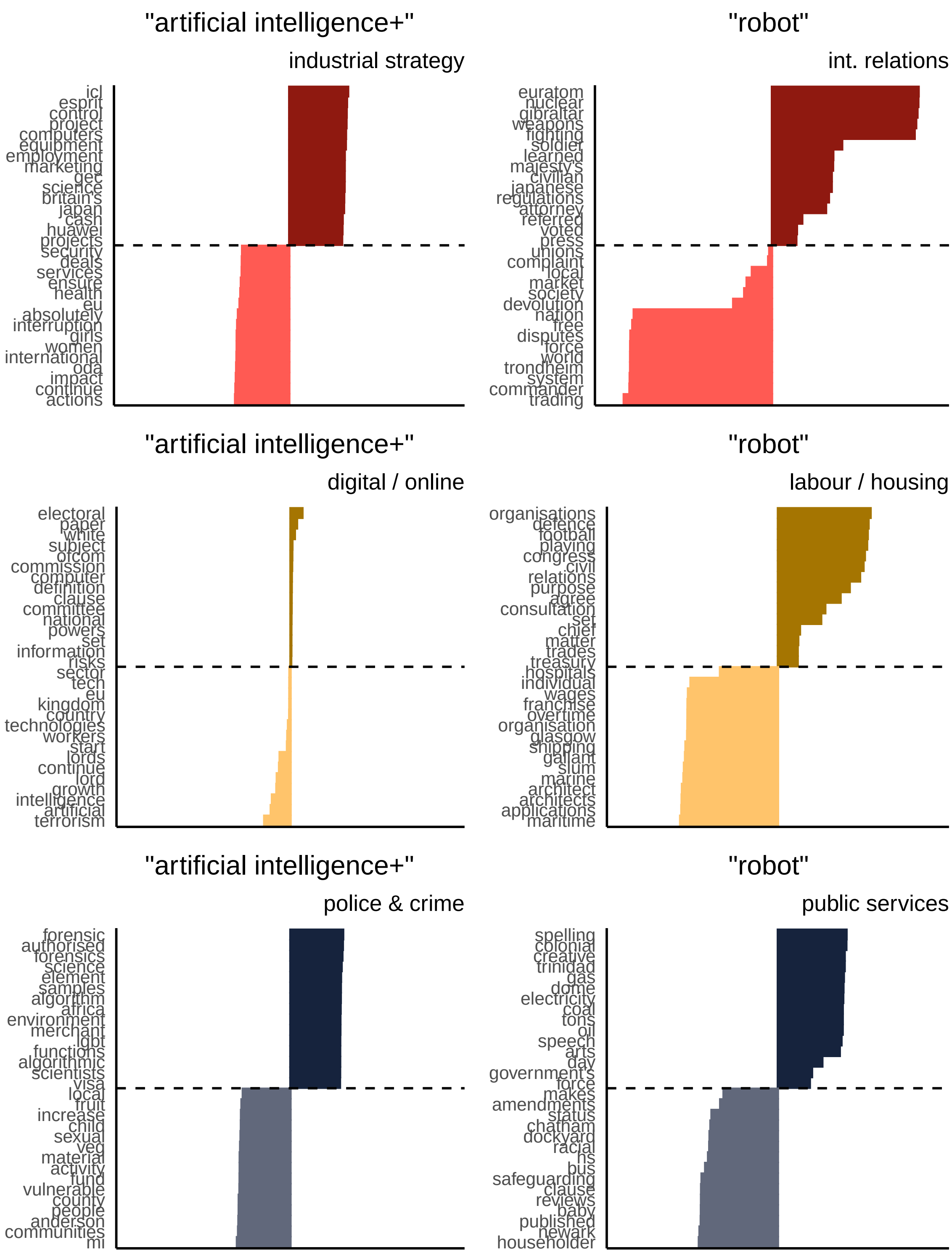


Figure 11: Frequency of probabilities that debates belong to the majority topic for "robot" debates. x-axis is the probability (`gamma`) of the main topic, y-axis is the count of those probabilities

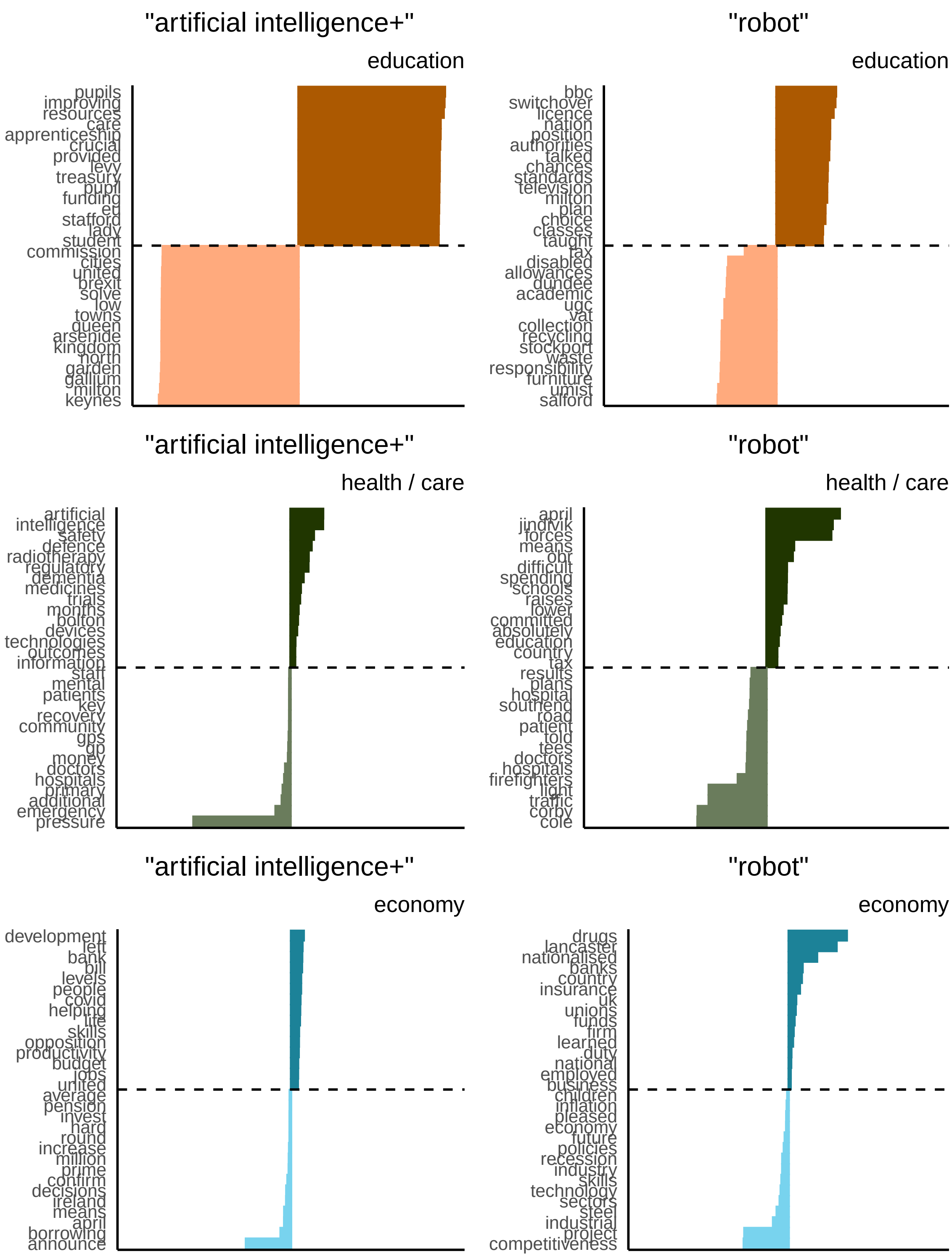
A.3 More detailed plots of the topic frames

The following plots show more detail of the two frames that each topic, for each search term, where separated into. The left column shows frames for “artificial intelligence+” debates and the right “robot” debates.

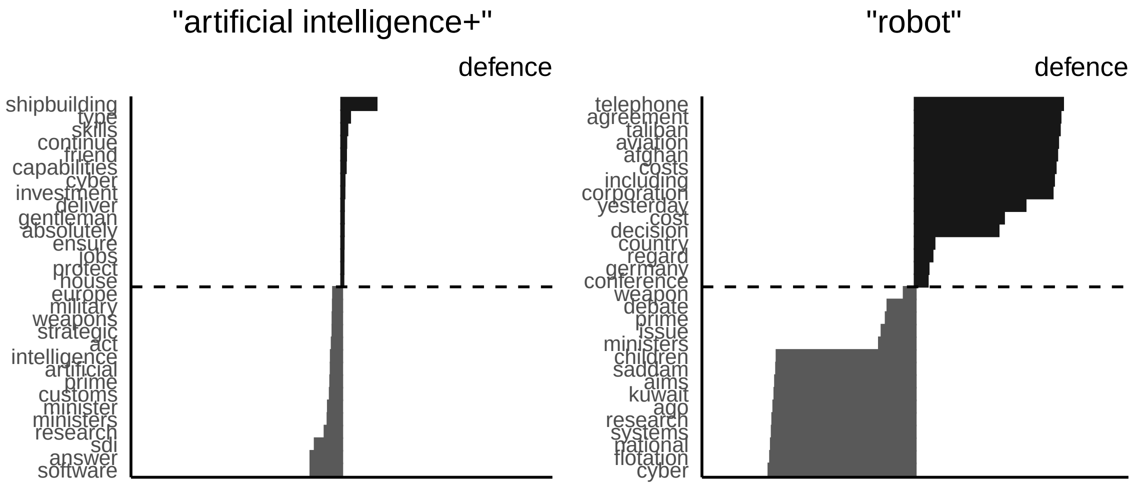




"artificial intelligence+"
industrial strategy
icl
esprit
control
project
computers
equipment
employment
marketing
gec
science
britain's
japan
cash
huawei
projects
security
deals
services
ensure
health
eu
absolutely
interruption
girls
women
international
oda
impact
continue
actions
"robot"
int. relations
euratom
nuclear
gibraltar
weapons
fighting
soldier
learned
majesty's
civilian
japanese
regulations
attorney
referred
voted
press
unions
complaint
local
market
society
devolution
nation
free
disputes
force
world
trondheim
system
commander
trading
"artificial intelligence+"
digital / online
electoral
paper
white
subject
ofcom
commission
computer
definition
clause
committee
national
powers
set
information
risks
sector
tech
eu
kingdom
country
technologies
workers
start
lords
continue
lord
growth
intelligence
artificial
terrorism
"robot"
labour / housing
organisations
defence
football
playing
congress
civil
relations
purpose
agree
consultation
set
chief
matter
trades
treasury
hospitals
individual
wages
franchise
overtime
organisation
glasgow
shipping
gallant
slum
marine
architect
architects
applications
maritime
"artificial intelligence+"
police & crime
forensic
authorised
forensics
science
element
samples
algorithm
africa
environment
merchant
lgbt
functions
algorithmic
scientists
visa
local
fruit
increase
child
sexual
veg
material
activity
fund
vulnerable
county
people
anderson
communities
mi
"robot"
public services
spelling
colonial
creative
trinidad
gas
dome
electricity
coal
tons
oil
speech
arts
day
governments
force
makes
amendments
status
chatham
dockyard
racial
hs
bus
safeguarding
clause
reviews
baby
published
newark
householder



"artificial intelligence+"
education
pupils
improving
resources
care
apprenticeship
crucial
provided
levy
treasury
pupil
funding
eu
stafford
lady
student
commission
cities
united
brexit
solve
low
towns
queen
arsenide
kingdom
north
garden
gallium
milton
keynes
"robot"
education
bbc
switchover
licence
nation
position
authorities
talked
chances
standards
television
milton
plan
choice
classes
taught
tax
disabled
allowances
dundee
academic
ugc
vat
collection
recycling
stockport
waste
responsibility
furniture
umist
salford
"artificial intelligence+"
health / care
artificial
intelligence
safety
defence
radiotherapy
regulatory
dementia
medicines
trials
months
bolton
devices
technologies
outcomes
information
staff
mental
patients
key
recovery
community
gps
gp
money
doctors
hospitals
primary
additional
emergency
pressure
"robot"
health / care
april
jindivik
forces
means
obr
difficult
spending
schools
raises
lower
committed
absolutely
education
country
tax
results
plans
hospital
southend
road
patient
told
tees
doctors
hospitals
firefighters
light
traffic
corby
cole
"artificial intelligence+"
economy
development
left
bank
bill
levels
people
covid
helping
life
skills
opposition
productivity
budget
jobs
united
average
pension
invest
hard
round
increase
million
prime
confirm
decisions
ireland
means
april
borrowing
announce
"robot"
economy
drugs
lancaster
nationalised
banks
country
insurance
uk
unions
funds
firm
learned
duty
national
employed
business
children
inflation
pleased
economy
future
policies
recession
industry
skills
technology
sectors
steel
industrial
project
competitiveness


"artificial intelligence+"
defence
shipbuilding
type
skills
continue
friend
capabilities
cyber
investment
deliver
gentleman
absolutely
ensure
jobs
protect
house
europe
military
weapons
strategic
act
intelligence
artificial
prime
customs
minister
ministers
research
sdi
answer
software
"robot"
defence
telephone
agreement
taliban
aviation
afghan
costs
including
corporation
yesterday
cost
decision
country
regard
germany
conference
weapon
debate
prime
issue
ministers
children
saddam
aims
kuwait
ago
research
systems
national
flotation
cyber

A.4 Interesting examples of utterances

A.4.1 The first parliamentary debate mention of "artificial intelligence+"

| date: 1981-11-27 | LDA topic: industrial strategy | major heading: PRAYERS | minor heading: Information Technology: "... few people in britain realise its significance or how important it is that britain should play a role in its development. i ask the minister to pay close attention to the considerable british expertise in artificial intelligence and expert systems. i was privileged to attend a conference organised by a computer company in the south of france. they left a vivid impression that we were missing a trick that could be of the greatest importance to britain. it would take far too long to endeavour to define expert systems or artificial intelligence. suffice it to say that all the knowledge that mister speaker and his clerks have of"erskine may", plus all his experience, can be codified and put into a computer if certain steps are taken. for example, the science and engineering research council supports work on distributive computing. intersting work is also being carried out in the united kingdom on artificial intelligence. my honourable friend the member for fife, east (mister henderson) waxed eloquent about the need for the house to embrace the new technologies. ..."

A.4.2 The first mention of "artificial intelligence+" after the AI Winter

| date: 2012-05-24 | LDA topic: defence | major heading: NULL | minor heading: Caravans Hull: "... it is also fair to say that he is the father of modern computing. he produced the first academic papers on artificial intelligence, which paved the way for modern computers. who knows where technology would be today without his pioneering work? ..."

A.4.3 The first mention of "artificial intelligence+" with respect to digital / online

| date: 1988-04-28 | LDA topic: digital / online | major heading: Orders of the Day Copyright Designs and Patents Bill Lords | minor heading: STATUTORY INSTRUMENTS ETC: "... it is refreshing to see this legislature getting to grips with that problem ahead of anybody else in the world. we are the first legislature to recognise the advent—not here yet, but coming—of artificial intelligence. we recognise also that a special balance has to be struck where the needs of education institutions are involved. it does not embrace all types of machine-readable information. the honourable member for dagenham (mister gould) gave the bill credit for discussing computer generation, and said that we were the first legislature to examine artificial intelligence. is the definition of computer generation in the bill clear enough? the march of technology has created much of the reason for the bill. i was impressed to see that the secretary of state for trade and industry has put into the bill a clause to deal with computer-generated products and artificial intelligence and, in committee, we shall see a closer definition of that work. i hope that, in committee, we shall touch again on the issues introduced tonight. computer software is the product of new methods of identifying and transmitting human thought. my right honourable friend the member for chertsey and walton referred to artificial intelligence when he said how difficult it was to define computer software. that is so very old. ..."

A.4.4 5G is an enabling technology for artificial intelligence

| date: 2020-11-30 | LDA topic: digital / online | major heading: Telecommunications Security Bill | minor heading: Topical Questions: "... as a chartered engineer, i want to finish by celebrating the potential of 5g, which can truly transform our businesses, our industries and our daily lives. it will not only vastly improve our connectivity and browsing experience but support new enabling technologies, from the internet of things to artificial intelligence. if the first industrial revolution was powered by engines, the fourth will be powered by data. ..."

A.4.5 Wartime mention of "robot"

| date: 1940-05-08 | LDA topic: int. relations | major heading: Orders of the Day CIVIL ESTIMATES 1940 Progress | minor heading: CONDUCT OF THE WAR: "... in the brown hours, when baffling news comes, and disappointing news, i always turn for refreshment to the reports of the german wireless. i love to read the lies they tell of all the british ships they have sunk so many times over, and to survey the fools' paradise in which they find it necessary to keep their deluded serfs and robots. the germans have claimed to have sunk or damaged 11 battleships; actually, two have been slightly damaged—neither of them withdrawn for a day from the service. ..."

A.4.6 "robot" in relation to employment

| date: 1987-04-06 | LDA topic: economy | major heading: BILL PRESENTED | minor heading: Social and Economic Policies: "... nobody is seriously suggesting that it should be committed to taking back another 10,000 people to provide them with work. although the company is taking on about 1,000 people, our problem is that, in the next era of technology, machinery will be even more robotised and even more productively efficient and will require fewer people to make an ever greater quantity of manufactured goods. that is the challenge that we face. ..."

A.4.7 Mention of "robot" in 2016 as "artificial intelligence+" would be mentioned in the 202s

| date: 2016-02-02 | LDA topic: int. trade | major heading: Enterprise Bill Lords | minor heading: Topical Questions: "... after all, it is clear that the world is now on the cusp of the fourth industrial revolution, and if we are not ready for the wave coming toward us, we will miss it. i want us to take advantage of what will be an age of rapidly advancing digitalisation, and an age of robotics and big data that is expected to transform our lives out of all recognition—and to do so much more quickly than we might expect. it will be an age that confronts us with profound questions about how to generate and share prosperity and fight for a fairer outcome for everyone in our society. ..."

A.5 Initial results for other search terms

The following are examples performing similar searches but using non AI search terms.

A.5.1 Climate-related search terms

The search terms used in this search were: “climate change”, “global warming”, “global heating”, “greenhouse effect”. Figure 12 shows the number of utterances of these terms in parliamentary debates over time.

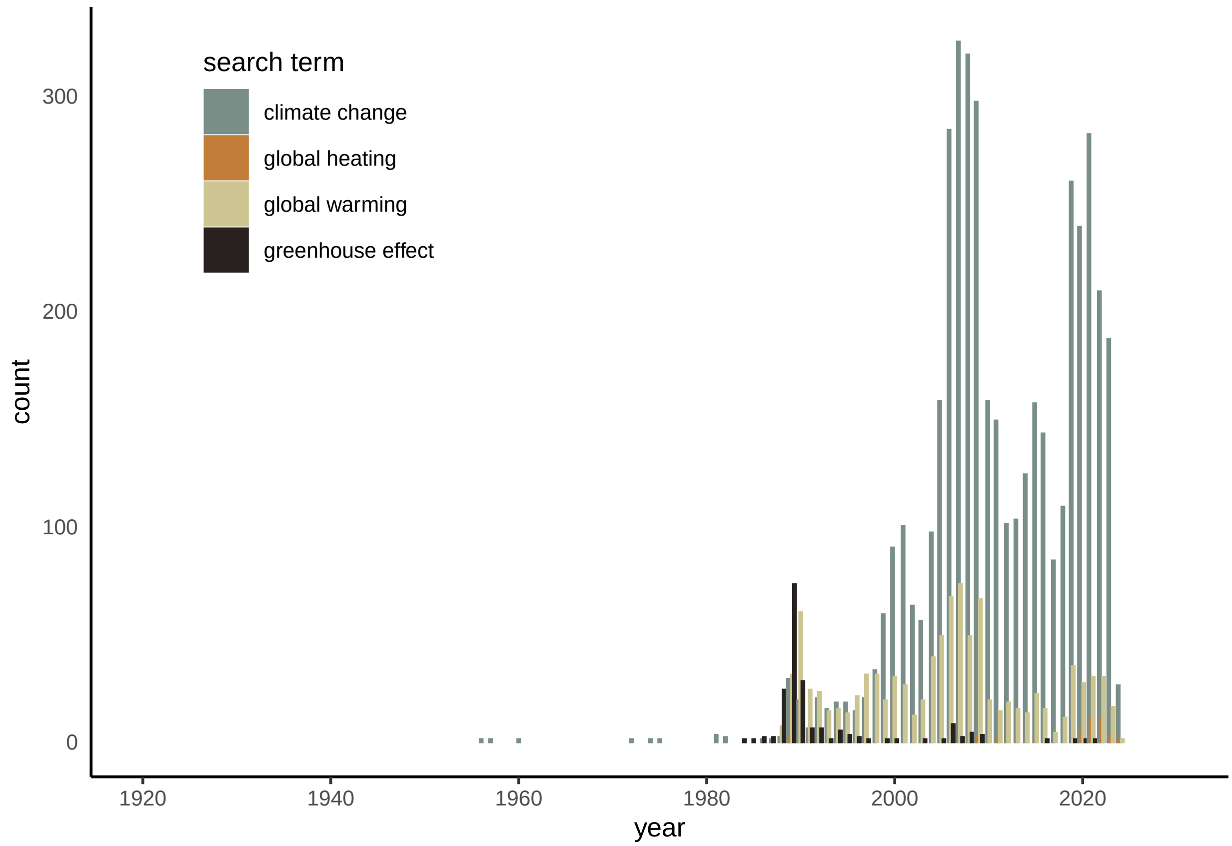


Figure 12: Count of parliamentary debates mentioning climate-related search terms by year

A.5.2 Smoking-related search terms

The search terms used in this search were: “tobacco”, “cigarette smoking”, “lung cancer”. Figure 13 shows the number of utterances of these terms in parliamentary debates over time.

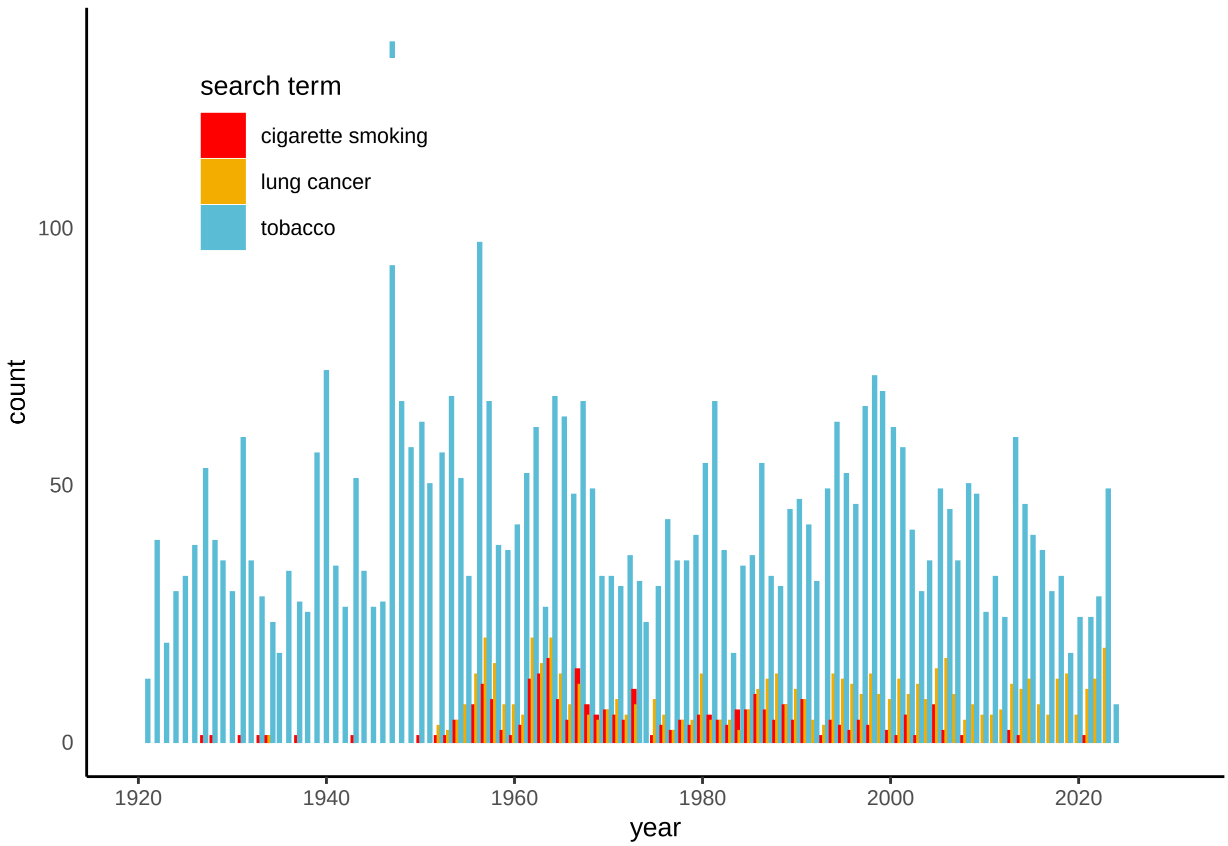


Figure 13: Count of parliamentary debates mentioning smoking-related search terms by year