

ParliamentRAG: An Authority-Aware Multi-View RAG System for Italian Parliamentary Proceedings

Mirko Tritella¹, Riccardo Pozzi¹ and Matteo Palmonari¹

¹University of Milano-Bicocca, Milan, Italy

Abstract

Large Language Models applied to parliamentary text risk over-representing the most vocal speakers, ignoring topic-dependent expertise, and hallucinating citations. ParliamentRAG is a Retrieval-Augmented Generation (RAG) system for the Italian Chamber of Deputies that addresses these challenges and enables citizens, journalists, and researchers to explore parliamentary debates through five interactive functionalities. Given a topic query, users can: (1) obtain balanced, multi-view summaries of parliamentary group positions with authority-weighted speaker selection and verbatim quotations from official transcripts; (2) search over parliamentary proceedings using filters for MP, group, date, and document type; (3) discover the most authoritative MPs on a given topic through a transparent, component-based authority ranking; (4) visualize how parliamentary groups position themselves relative to each other via an interactive ideological compass; and (5) browse a chronological timeline of proceedings with AI-generated summaries. The system operates over a knowledge graph of the current XIX Legislature and supports multilingual access (Italian, English, French, German, Spanish, Portuguese). ParliamentRAG is publicly accessible at <https://www.parliamentrag.it/>.

Keywords

Retrieval-Augmented Generation, Parliamentary NLP, Knowledge Graphs, Expert Finding, Multi-Perspective Summarization

1. ParliamentRAG: an Authority-Aware Multi-View RAG System

Parliamentary debates constitute a dense documentary record of democratic deliberation, leading to several initiatives to publish parliamentary data as open data¹ and Linked Open Data in RDF, including the European Parliament Open Data Portal² [1] and other initiatives [2, 3]. One such example is the Open Data Portal of the Italian Chamber of Deputies³. Even when these data are openly available, obtaining valuable insights into the different political perspectives expressed by different actors (parliamentary groups, parties, and politicians) on different topics remains challenging.

A first obstacle to supporting this multi-perspective access is the nature of the data itself: most of the content is textual, while the corpus is both large and highly fragmented. The ongoing XIX Legislature of the Italian Chamber of Deputies has produced 694 plenary sessions, 6,921 debates, and 45,666 speeches. Manually navigating such a corpus is excessively costly, and keyword-based search portals return isolated documents rather than synthesized, multi-perspective answers. Large Language Models (LLMs) are well suited for synthesizing textual content. However, even when grounded on an authoritative data layer—which is possible but not trivial—their application to political text risks over-representing the most vocal actors [4], treating all speakers as equally authoritative regardless of their expertise on the queried topic, and hallucinating or misattributing citations [5, 6].

This demonstration presents *ParliamentRAG*, a RAG system for the Italian Chamber of Deputies, which has been recently accepted in the ISWC In-Use track [7]. It shows how the system generates **multi-view** summaries of parliamentary debates, while ensuring balanced representation across all ten parliamentary groups, dynamically modeling speaker expertise for **authority-aware** retrieval, and providing **grounded quotations** that correspond to verbatim passages from official transcripts. This

ISWC 2026 Companion Volume, October 25–29, 2026, Bari, Italy

✉ m.tritella@campus.unimib.it (M. Tritella); riccardo.pozzi@unimib.it (R. Pozzi); matteo.palmonari@unimib.it (M. Palmonari)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹<https://developer.parliament.uk/>.

²<https://data.europarl.europa.eu/>

³<https://dati.camera.it/>

demonstration focuses on the user-facing functionalities of ParliamentRAG, while the technical details of the system architecture and underlying methods are described in the In-Use paper [7].

ParliamentRAG operates over a knowledge graph integrating parliamentary proceedings, metadata, legislative activity, and voting records (6,589,092 nodes and 13,318,266 relationships, of which 6.3 million are individual voting records; 266,824 nodes describe proceedings, acts, and political actors).⁴ The graph is built from the Italian Chamber of Deputies’ open data infrastructure [8], which provides RDF-based semantically structured records following the *Ontology of the Chamber of Deputies* [9]. It includes debate transcripts, legislative acts, deputy profiles, and temporally valid institutional affiliations. Both the In-Use paper [7] and this demonstration focus on the current legislature (XIX), while future extensions will cover previous legislatures. In the technical paper [7], we evaluate ParliamentRAG using automated methods and a user study, discussing its advantages when compared with a strong baseline based on NotebookLM [10], which relies on prompting rather than system-level enforcement of such requirements. On 15 policy topics, with blind A/B judgments by six domain experts, ParliamentRAG reaches perfect quotation faithfulness (1.00 vs. 0.95) and higher coverage across parliamentary groups (0.97 vs. 0.95); evaluators prefer it on source-related dimensions (coverage, authority, and relevance of the cited sources), while NotebookLM scores higher on prose quality and overall satisfaction is comparable.

Building on the system described in the In-Use paper [7], we introduce three improvements. First, ParliamentRAG now supports **multilingual access**: every page includes a dropdown menu to select the preferred language (currently Italian, English, French, German, Spanish, and Portuguese). Journalists, researchers, and citizens beyond Italy can thus analyze Italian parliamentary debates in their own language, a prerequisite for comparative studies of political debates across countries. Second, we have added a **parliamentary timeline with AI-generated summaries**, a chronological view of proceedings with automatically generated recaps of debates and legislative activity. Third, quotation selection has been strengthened by an index-time LLM classification of chunk citability, which pre-filters procedural and rhetorical content and pre-extracts the most quotable verbatim sentence of each chunk.

The source code is available at <https://github.com/Emeierkeio/ParliamentRAG> under the Apache-2.0 license, with a live deployment available at <https://www.parliamentrag.it>.

2. Related Work and Novelty

To the best of our knowledge, no existing application combines interactive multi-view reports of parliamentary debates with query-dependent authority modeling and verbatim quotation grounding. The closest related systems explore hybrid retrieval approaches combining dense retrieval and knowledge graphs over parliamentary data [11] or AI-supported enrichment and visualization [12], but do not jointly address these requirements.

Recent work on **parliamentary NLP**, including corpora such as ParlaMint [13] and IPSA [14], as well as LLM- and RAG-based systems [12, 11], has improved access to parliamentary debates. Compared to these approaches, ParliamentRAG explicitly supports authority-aware, multi-perspective analysis. **Parliamentary knowledge graphs** such as LinkedEP [1], ParliamentSampo [2], DemocraSci [3], and the OCD ontology [9] provide structured access to parliamentary data. ParliamentRAG instead exploits a parliamentary KG as the foundation for query-dependent retrieval and analysis. **Expert finding and authority** ranking methods [15, 16, 17], as well as source-aware RAG approaches [18], estimate authority or source reliability. ParliamentRAG introduces a query-dependent authority model tailored to parliamentary debates with multiple legitimate viewpoints. **Diversification and fairness**-aware retrieval [19, 20], together with GraphRAG [21], improve coverage or global retrieval. ParliamentRAG explicitly guarantees balanced representation across parliamentary groups while combining authority-aware retrieval with multi-view generation. **Quotation Faithfulness**. Attribution and verification

⁴Compared to the version used for the experiments in [7], the knowledge graph has been rebuilt from scratch to improve data quality: it now integrates 6.3 million individual voting records, covers all proceedings to date, removes duplicate speeches, and attributes Mixed Group speeches to their political sub-components.

methods [22, 23, 24] improve citation reliability. In contrast, ParliamentRAG guarantees verbatim quotations by deterministically grounding every citation in the original parliamentary transcripts. For more details, we refer to the technical paper [7].

3. Main Features and Demonstration

ParliamentRAG supports two main families of processes: the knowledge graph construction processes, discussed in [7], and the live analytical processes. Our demonstration covers the latter, explored through ParliamentRAG’s frontend web interface.

We present five scenarios, each centered on one functionality of the system and on an example user need it addresses.

Citations are automatically translated from Italian. Hover to see the original.
Don't show again

What is the position of parliamentary groups on the topic: Justice Reform?

Share

Chamber of Deputies — 19th Legislature · 14:19

Query analysis

Committees

Authoritative sources

Speeches

Statistics

Ideological Compass

Generation

Saving

Introduction

The parliamentary discussion concerned the communications of the Minister of Justice on the administration of justice, with particular attention to the revision of previous reforms, the separation of careers, the balance between the powers of the State, and constitutional guarantees. Ninety-one interventions by [60 deputies](#) were analyzed in the period from 25 October 2022 to 28 May 2026.

Government Position

For the Government, justice is placed among the priorities of the agenda, with awareness of the need to overcome the critical issues of previous reforms and to promote a thorough revision of the current regulations. [Carlo Nordio](#) states that [«for the most part I agree with you that a profound revision of the reform that was made is necessary»](#) %. The executive declares itself in favor of legislative initiatives that modernize the judicial system, aiming at greater efficiency and protection of rights.

Majority Positions

For Fratelli d'Italia, the justice reform represents a historic priority and a mandate received from the voters, with the goal of overcoming decades of failed attempts and affirming the political legitimacy of the intervention on the judicial system. [Ciro Maschio](#) maintains that [«there is a Government, there is a Parliament, democratically elected, with a program and a clear mandate from the voters, with a cohesive and broad enough majority to be able to carry out, over the course of an entire legislature, the reforms that have been attempted for over 30 years but have never been achieved»](#) %. The group positions itself as a promoter of a structural reform, rejecting criticisms and aiming at a more efficient and transparent judicial system.

For Lega - Salvini Premier, the justice reform requires structural and ambitious interventions, in line with the needs of the penitentiary police and local communities, recognizing the innovative scope of the government program. [Jacopo Morrone](#) observes that [«we, Mr. Minister, share your assessment that this program is very broad, I would add ambitious and revolutionary, and we believe you when you say that, at the end of this legislature, in five years, it will largely be completed»](#) %. Lega supports a deep reform, conditioning its support on the concrete and timely realization of the changes, with attention to the balance between powers.

Opposition Positions

For Partito Democratico - Italia Democratica e Progressista, the justice reform proposed by the majority is strongly criticized because it does not address structural problems and appears motivated by punitive political intents. [Federico Fornaro](#) denounces: [«I believe it must be said forcefully, because at the origin, essentially, there is a punitive intent towards the judiciary, a signal is wanted in this direction»](#) %. The group opposes changes that risk compromising the autonomy of the judiciary, supporting the need for shared and profound interventions.

For Movimento 5 Stelle, the Government's justice reform does not respond to the real needs of the system and risks weakening the autonomy of the judiciary. [Enrica Alifano](#) laments that [«we with this amendment are on the suppressive because we do not like this reform, I believe we have said it in every possible and imaginable way»](#) %. The group strongly opposes the proposal, demanding solutions that strengthen the judicial system without compromising its independence.

Mixed Group

For the +Europa - Stati Uniti d'Europa component of the Mixed Group, the justice reform is strongly supported, claiming the liberal and guarantist tradition as the foundation of the proposal. [Benedetto Della Vedova](#) assures that [«we vote "yes," because we do not want to leave this principle of liberal order and this flag in the wrong hands»](#) %. The component sees in the reform an opportunity to strengthen liberal and guarantist principles, considering it consistent with a modern vision respectful of fundamental rights.

Figure 1: ParliamentRAG’s authority-aware, multi-view summary in response to the query “What is the position of parliamentary groups on the topic: Justice Reform?”, shown in the English interface: citations are automatically translated from Italian, and hovering reveals the original wording. Exact citations are delimited by « and ».

Topic Search: Authority-aware Multi-view RAG. This feature answers the user need to understand the positions of all parliamentary groups on a specific topic, with authoritative sources and

verbatim quotations. For example, when asking “*What is the position of the parliamentary groups on the topic: Justice Reform?*”, the system generates a structured report (see Figure 1) showing: the Government position through Minister of Justice Nordio («for the most part I agree with you that a profound revision of the reform that was made is necessary»); the majority positions—Brothers of Italy (Fratelli d’Italia) claiming a voter mandate for «the reforms that have been attempted for over 30 years but have never been achieved» (Ciro Maschio), and the League endorsing the ambition of the government programme (Jacopo Morrone: «at the end of this legislature, in five years, it will largely be completed»); the opposition positions—the Democratic Party denouncing «a punitive intent towards the judiciary» (Federico Fornaro) and the Five Star Movement firmly opposed (Enrica Alifano: «we do not like this reform, I believe we have said it in every possible and imaginable way»); and the Mixed Group rendered per political component, with the +Europa component supporting the reform in the name of the liberal and guarantist tradition (Benedetto Della Vedova). The frontend displays:

- A clean, readable layout with dedicated sections for each parliamentary group
- Speaker names with authority scores and institutional roles (available by scrolling down)
- Verbatim quotations highlighted with links to original transcripts
- Positions of the Mixed Group presented per political sub-component (resolved at speech date from the KG), since the group spans opposing orientations
- Language selection on every page (Italian default, English, French, German, Spanish, Portuguese)

Acts Search. This feature answers the user need to find who has spoken about a specific topic, or to locate specific parliamentary speeches and legislative acts. For example, searching for “*Justice Reform*” retrieves the matching floor speeches and legislative acts, which can then be narrowed down by author, document type, and time period. The frontend provides:

- A free-text search bar over parliamentary acts and floor speeches
- Filter panels for MP, parliamentary group, date range, and document type
- Search results displayed as cards with speaker names, dates, and excerpts
- Direct links from search results to original parliamentary transcripts

Authority Analysis. This feature answers the user need to identify which MPs are the most authoritative experts on a given topic, and to understand why they are considered experts. For example, entering the topic “*Justice Reform*”, the system ranks the 391 deputies represented in the knowledge graph by expertise (see Figure 2). The top-ranked MPs include Federico Cafiero De Raho and Carla Giuliano (both Movimento 5 Stelle), members of the Justice Committee. The score of a speaker s for a query q is a weighted sum of six components, each normalized to $[0, 1]$:

$$authority(s, q) = \sum_i w_i c_i(s, q), \quad (1)$$

where each component c_i measures one signal from the knowledge graph. *Profession* and *education* measure the embedding similarity between the query and the speaker’s declared profession and education titles. *Committee memberships* and *institutional roles* score membership and offices (e.g., committee president) in bodies whose remit is relevant to the query. *Signed acts* and *speech interventions* aggregate legislative activity: each act or speech contributes its embedding similarity to the query, discounted by an exponential time decay that favors recent activity and, for acts, weighted by the speaker’s role (first signatory vs. co-signatory). The score considers only mandates and affiliations valid at the reference date, so past coalition switches do not bias present-day estimates. The weights w_i are set empirically rather than learned from data, keeping the function interpretable: $w_{\text{profession}} = 0.15$, $w_{\text{education}} = 0.10$, $w_{\text{committee}} = 0.25$, $w_{\text{roles}} = 0.05$, $w_{\text{acts}} = 0.20$, $w_{\text{interventions}} = 0.25$. They are documented in the open-source configuration, and users can adjust them in the application settings (see [7] for the full formulation). The value is an interpretable proxy for topic expertise, not a claim of objective authority: the per-component breakdown lets users inspect, and contest, every ranking. The frontend displays:

- A ranked list of MPs with their party and authority scores for the given topic

- Breakdown of the six authority components for each MP
- Visual indicators showing which components contribute most to each MP’s authority
- Links to the MP’s profile and their relevant speeches and legislative acts

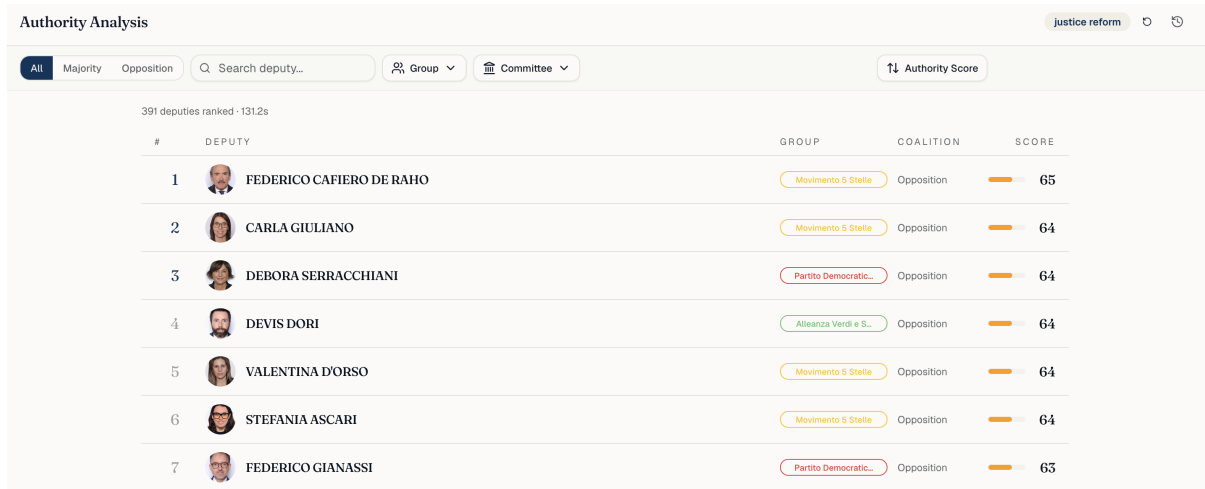


Figure 2: Authority ranking of MPs on the topic “Justice Reform”, showing the interpretable components contributing to each speaker’s authority score.

Ideological Compass. This feature answers the user need to understand how parliamentary groups position themselves relative to each other on a topic, and to identify the main dimensions of political disagreement. Given the query, an LLM generates two axes of disagreement, each defined by a pair of opposite poles, so the axes are readable before any point is placed on the map. A second LLM pass classifies the stance that each retrieved fragment argues on each axis, scoring it in $[-1, 1]$ or discarding it when the fragment takes no position. Classifying stance rather than projecting embeddings protects rebuttals: a speech attacking a position quotes its vocabulary, and a lexical projection would place it on the position it opposes. Group positions aggregate the fragment scores, and if an LLM step fails the system falls back to embedding projection onto the pole vectors. For example, when exploring Justice Reform, the compass derives a first axis opposing procedural efficiency to the protection of defense guarantees, and a second axis opposing centralization to decentralization of the justice system. The resulting map shows that, on this topic, distances between groups do not simply mirror the government–opposition divide: on the centralization axis the League sits closer to the opposition groups than to its own coalition partners, surfacing a dimension of disagreement that a coalition-based summary alone would not reveal. The frontend presents:

- An interactive 2D map with parliamentary groups positioned based on their debate positions
- Automatically derived axes representing the main dimensions of political disagreement
- Interactive party markers revealing each group’s position and placement along the axes
- Axis labels translated into the interface language, while the underlying poles stay anchored in Italian so that positions are identical across languages

Chamber Proceedings (new w.r.t. [7]). This feature answers the user need to stay informed about recent parliamentary activity on a specific topic, with AI-generated summaries for quick overviews. For example, a user who wants to know what happened in Parliament during the last week on justice-related debates can browse the timeline of recent sittings, skim the AI-generated recap of each session, and use full-text search to jump directly to the relevant debates and interventions. The frontend provides:

- A chronological timeline of parliamentary sittings with debates and speeches
- AI-generated summaries of individual sessions, providing quick overviews of key discussions
- Drill-down capability from high-level summaries to specific speeches and interventions

- Filtering by date and full-text search to explore parliamentary activity over time
- Per-speaker AI summaries of each deputy’s interventions within a debate

MCP Access for LLM Agents. Beyond the web interface, ParliamentRAG exposes its knowledge graph through a public Model Context Protocol (MCP) server,⁵ mentioned in [7] and described here in its demonstrable form: LLM-based assistants query parliamentary data as tools instead of scraping the browser interface. The server offers tools to search proceedings, retrieve debates and sessions, and inspect roll-call votes, including per-vote hemicycle charts served as stable public images. Every tool result carries the link to the official record on the Chamber’s website, so agents inherit the same traceability guarantees as the web application.

4. Conclusions and future work

This paper presents ParliamentRAG, a RAG system that supports the exploration of debates in the Italian Chamber of Deputies through multi-perspective summaries and interactive analyses, combining a knowledge graph built from authoritative RDF data with LLMs. The demonstration extends the work presented in [7], in particular through multilingual support, which opens the system to international users such as journalists and political scientists. Current limitations concern data coverage and quality. The knowledge graph covers the Chamber of Deputies only, and its transcripts only plenary sittings, leaving out Senate proceedings and committee debates. It spans a single time period, the ongoing XIX Legislature, with earlier legislatures planned as extensions. Some fields in the official source records are misfiled (e.g., education titles appearing as professions), and these errors can propagate to the authority components built on them. Planned work includes automated quality checks at ingestion time, precomputed authority rankings for recurring topics to reduce response latency, and a structured usability evaluation of the web interface, complementing the user study of the generation pipeline in [7]. Although ParliamentRAG is built on data from the Italian Chamber of Deputies, its architecture is largely data-agnostic and can be adapted to other national legislatures with comparable open parliamentary data.

Acknowledgments

This work was supported by the EU Horizon Europe programme through grants No. 101189771 (DataPACT).

Declaration on Generative AI

During the preparation of this work, the authors used Claude (Anthropic) and ChatGPT (OpenAI) in order to: grammar and spelling check, paraphrase and reword, improve writing style. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication’s content.

References

- [1] A. van Aggelen, L. Hollink, M. Kemman, M. Kleppe, H. Beunders, The debates of the european parliament as linked open data, *Semant. Web* 8 (2017) 271–281. URL: <https://doi.org/10.3233/SW-160227>. doi:10.3233/SW-160227.
- [2] E. Hyvönen, L. Sinikallio, P. Leskinen, S. Drobac, R. Leal, M. La Mela, J. Tuominen, H. Poikkimäki, H. Rantala, Publishing and using parliamentary linked data on the semantic web: Parliamentsampo system for parliament of finland, *Semantic Web* 16 (2025) SW–243683. doi:10.3233/SW-243683.

⁵<https://mcp.parliamentrag.it/>

- [3] L. Salamanca, L. Brandenberger, L. Gasser, S. Schlosser, M. Balode, V. Jung, F. Perez-Cruz, F. Schweitzer, Processing large-scale archival records: The case of the swiss parliamentary records, *Swiss Political Science Review* 30 (2024) 140–153. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/spsr.12590>. doi:<https://doi.org/10.1111/spsr.12590>. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/spsr.12590>.
- [4] S. Santurkar, E. Durmus, F. Ladhak, C. Lee, P. Liang, T. Hashimoto, Whose opinions do language models reflect?, in: *Proceedings of the 40th International Conference on Machine Learning, ICML'23*, JMLR.org, 2023. URL: <https://openreview.net/pdf?id=7IRybndMLU>.
- [5] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, P. Fung, Survey of hallucination in natural language generation, *ACM Computing Surveys* 55 (2023) 1–38. URL: <http://dx.doi.org/10.1145/3571730>. doi:10.1145/3571730.
- [6] J. Maynez, S. Narayan, B. Bohnet, R. McDonald, On faithfulness and factuality in abstractive summarization, in: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Online, 2020, pp. 1906–1919. URL: <https://aclanthology.org/2020.acl-main.173/>. doi:10.18653/v1/2020.acl-main.173.
- [7] M. Tritella, R. Pozzi, M. Palmonari, Who speaks matters: Authority-aware multi-view RAG over Italian parliamentary proceedings, in: *Proceedings of the 25th International Semantic Web Conference (ISWC 2026)*, 2026. In-Use Track, forthcoming. Accepted version: <https://emeierkeio.github.io/papers/who-speaks-matters-iswc2026.pdf>. Preprint: <https://arxiv.org/abs/2608.13410>.
- [8] Camera dei deputati, Portale dei dati aperti della camera dei deputati, <https://dati.camera.it>, 2026. Accessed: 2026.
- [9] Camera dei deputati, Ontology of the chamber of deputies: RDF classes, <https://dati.camera.it/ocd/classi.rdf>, 2026. Accessed: 2026.
- [10] Google, Google notebooklm, <https://notebooklm.google.com/>, 2026.
- [11] S. Mosbach, J. Lai, K. Rustagi, D. N. Tran, M. Kraft, E. Bindereif, M. Azzam, Parliamentary debates in the world avatar: A hybrid retrieval-augmented generation system, 2025. URL: <https://como.ceb.cam.ac.uk/media/preprints/c4e-preprint-338.pdf>, preprint.
- [12] J. Azzopardi, LLMs and knowledge discovery in low-resource language parliamentary corpora: The pq dashboard case study, in: *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025) - Volume 1: KDIR*, SciTePress, 2025, pp. 159–170. URL: <https://www.scitepress.org/Papers/2025/138351/138351.pdf>. doi:10.5220/0013835100004000.
- [13] T. Erjavec, M. Kopp, N. Ljubešić, T. Kuzman, P. Rayson, P. Osenova, M. Ogrodniczuk, Ç. Çöltekin, D. Koržinek, K. Meden, J. Skubic, P. Rupnik, T. Agnoloni, J. Aires, S. Barkarson, R. Bartolini, N. Bel, M. Calzada Pérez, R. Dargis, S. Diwersy, M. Gavriilidou, R. van Heusden, M. Irukieta, N. Kahusk, A. Kryvenko, N. Ligeti-Nagy, C. Magariños, M. Mölder, C. Navarretta, K. Simov, L. M. Tungland, J. Tuominen, J. Vidler, A. I. Vladu, T. Wissik, V. Yrjänäinen, D. Fišer, Parlamint ii: Advancing comparable parliamentary corpora across europe, *Language Resources and Evaluation* 59 (2025) 2071–2102. URL: <https://doi.org/10.1007/s10579-024-09798-w>. doi:10.1007/s10579-024-09798-w.
- [14] V. Frasnelli, A. Palmero Aprosio, There's something new about the Italian parliament: The IPSA corpus, in: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, ELRA and ICCL, Torino, Italia, 2024, pp. 16037–16046. URL: <https://aclanthology.org/2024.lrec-main.1394/>.
- [15] K. Balog, Y. Fang, M. Rijke, P. Serdyukov, L. Si, Expertise retrieval, *Foundations and Trends in Information Retrieval* 6 (2012) 127–256. doi:10.1561/15000000024.
- [16] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank Citation Ranking: Bringing Order to the Web, Technical Report, Stanford Digital Library Technologies Project, 1998. URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.1768>.
- [17] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, *J. ACM* 46 (1999) 604–632. URL: <https://doi.org/10.1145/324133.324140>. doi:10.1145/324133.324140.
- [18] J. Hwang, J. Park, H. Park, D. Kim, S. Park, J. Ok, Retrieval-augmented generation with estimation of source reliability, in: *Proceedings of the 2025 Conference on Empirical Methods in*

- Natural Language Processing, Association for Computational Linguistics, Suzhou, China, 2025, pp. 34279–34303. URL: <https://aclanthology.org/2025.emnlp-main.1738/>. doi:10.18653/v1/2025.emnlp-main.1738.
- [19] M. Zehlike, F. Bonchi, C. Castillo, S. Hajian, M. Megahed, R. Baeza-Yates, Fa*ir: A fair top-k ranking algorithm, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17, ACM, 2017, p. 1569–1578. URL: <http://dx.doi.org/10.1145/3132847.3132938>. doi:10.1145/3132847.3132938.
 - [20] J. Carbonell, J. Goldstein, The use of mmr, diversity-based reranking for reordering documents and producing summaries, in: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '98, Association for Computing Machinery, New York, NY, USA, 1998, p. 335–336. URL: <https://doi.org/10.1145/290941.291025>. doi:10.1145/290941.291025.
 - [21] D. Edge, H. Trinh, N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, D. Metropolitansky, R. O. Ness, J. Larson, From local to global: A graph rag approach to query-focused summarization, 2025. URL: <https://arxiv.org/abs/2404.16130>. arXiv:2404.16130.
 - [22] H. Rashkin, V. Nikolaev, M. Lamm, L. Aroyo, M. Collins, D. Das, S. Petrov, G. S. Tomar, I. Turc, D. Reitter, Measuring attribution in natural language generation models, Computational Linguistics 49 (2023) 777–840. URL: <https://aclanthology.org/2023.cl-4.2/>. doi:10.1162/coli_a_00486.
 - [23] L. Gao, Z. Dai, P. Pasupat, A. Chen, A. T. Chaganty, Y. Fan, V. Zhao, N. Lao, H. Lee, D.-C. Juan, K. Guu, RARR: Researching and revising what language models say, using language models, in: A. Rogers, J. Boyd-Graber, N. Okazaki (Eds.), Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Toronto, Canada, 2023, pp. 16477–16508. URL: <https://aclanthology.org/2023.acl-long.910/>. doi:10.18653/v1/2023.acl-long.910.
 - [24] N. Liu, T. Zhang, P. Liang, Evaluating verifiability in generative search engines, in: Findings of the Association for Computational Linguistics: EMNLP 2023, Association for Computational Linguistics, Singapore, 2023, pp. 7001–7025. URL: <https://aclanthology.org/2023.findings-emnlp.467/>. doi:10.18653/v1/2023.findings-emnlp.467.