LLMs and Knowledge Discovery in Low-Resource Language Parliamentary Corpora: The PQ Dashboard Case Study

Joel Azzopardi[a]
Department of AI, Faculty of ICT, University of Malta, Malta

Keywords: Parliamentary Questions (PQs), Large Language Models (LLMs), Artificial Intelligence (AI), Natural Language Processing (NLP), Civic Engagement, Distant Reading, Interactive Dashboard.

Abstract: Parliamentary Questions (PQs) are a critical mechanism for democratic oversight and accountability. However, their comprehensive analysis can be hindered by limitations such as single-language availability (especially when the language is a low-resource language such as Maltese) and a lack of structured thematic organisation or interlinking. This paper introduces the PQ Dashboard, a web-based platform developed to enhance the accessibility and analytical utility of Maltese Parliamentary Questions. The system employs AI and open Large Language Models (LLMs) to automate PQ collection, translate content into English, classify it according to the COFOG-99 taxonomy, extract key terms, and identify interconnections. The interactive dashboard provides users – including the public, journalists, and academic researchers – with functionalities to navigate PQs by category or keyword, visualise thematic distributions, and analyse trends in MPs' activity and ministerial responses. This enhanced data accessibility aims to facilitate deeper insights into parliamentary discourse, policy development, and governmental accountability. The PQ Dashboard demonstrates a practical application of AI-driven solutions for transforming unstructured public data into a more accessible and analysable format, thereby contributing to increased transparency and informed public engagement.

1 INTRODUCTION

Parliamentary Questions (PQs) constitute a cornerstone of democratic governance, serving as a vital mechanism through which Members of Parliament (MPs) hold ministries accountable, seek information, and scrutinise government policy. The effective functioning of this oversight process relies heavily on the accessibility and interpretability of these parliamentary records for various stakeholders, including the public, journalists, and academic researchers.

However, traditional parliamentary portals, such as the official Maltese Parliament website hosting the PQs (https://pq.gov.mt; last accessed: September 2025), typically offer basic search functionalities and categorisation based on metadata like date, MP, or ministry. While these provide fundamental access, they often lack advanced semantic search capabilities, thematic overviews, or tools for longitudinal analysis. This limitation is particularly pronounced for parliamentary data in low-resource languages, such as Maltese, where readily available Natural Language Processing (NLP) tools and pre-trained models are less common, posing a significant barrier to comprehensive analysis and public engagement (Koehn and Knowles, 2017; Ranathunga et al., 2022). Early work on Maltese parliamentary data, such as Analysing and Visualising Parliamentary Questions: A Linked Data Approach (Abela and Azzopardi, 2018), has explored methods to enhance accessibility through linked data and visualisations, but these did not incorporate advanced Artificial Intelligence (AI) for semantic enrichment or multilingual access to the extent now possible with Large Language Models (LLMs).

Specifically, the current Maltese parliamentary portal publishes all answered PQs exclusively in Maltese. While it offers basic search and categorises PQs by criteria such as category, heading, MP, ministry, and sitting, it does not readily reveal overarching topics, enable tracking of trends in topics' popularity, or facilitate analysis of MPs' activity over time. Furthermore, references within PQs to other questions are not directly linked, which limits navigability and the ability to trace interconnected legislative discourse. These deficiencies collectively restrict the ability of users to gain high-level overviews, thematic insights,

[a] https://orcid.org/0000-0001-6709-8530

Azzopardi, J.
LLMs and Knowledge Discovery in Low-Resource Language Parliamentary Corpora: The PQ Dashboard Case Study.
DOI: 10.5220/0013835100004000

In *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025) - Volume 1: KDIR*, pages 159-170

and a deeper understanding of the parliamentary process.

This paper introduces the PQ Dashboard, a novel online platform designed to address these critical limitations. By leveraging advanced AI and NLP techniques, namely open LLMs, the PQ Dashboard aims to transform the accessibility and analytical utility of Maltese Parliamentary Questions, thereby promoting greater transparency and informed civic participation. As a publicly accessible web portal available at https://pq.ir.mt, the PQ Dashboard's innovative features considerably enhance the value of the available data. It achieves this by offering content in both Maltese and English, providing categorisation into COFOG-99 categories, identifying relevant keywords, and pinpointing citations between Parliamentary Questions. This comprehensive approach facilitates the presentation of rich insights and aggregate information, including the profiling of MPs and Ministries' activities through the PQs being submitted, and enables seamless navigation between PQs that are linked together via citations.

The rest of this paper is structured as follows: Section 2 first provides an overview of similar systems. Section 3 then describes the underlying process of the system, from data acquisition via web scraping to the processing of the data utilising LLMs, and finally, the updating of the server hosting the web application. Section 4 details the functionalities of the developed online dashboard and presents some key insights that can be extracted. In Section 5, we describe the evaluation that was carried out. Finally, Section 6 presents the conclusions and outlines our plans for future work.

2 SIMILAR SYSTEMS

Modern legislatures table thousands of Parliamentary Questions (PQs) each session – Malta alone recorded almost 30,000 PQs between May 2022 and July 2025, and this legislature is not yet complete (Parliament of Malta, 2022). While most parliaments now publish PQs and responses online, practices vary widely. Basic portals that provide documents (in *docx* or *pdf* formats) remain common, while only a minority offer machine-readable formats or public APIs (Inter-Parliamentary Union (IPU) and UN/IPU Global Centre for ICT in Parliament, 2024). Metadata is often inconsistent, and keyword-based search systems limit the discoverability of relevant material (Inter-Parliamentary Union (IPU) and Parliamentary Data Science Hub, Centre for Innovation in Parliament, 2024). Further hindrances to transparency include delayed responses or evasive replies such as "data not held" (TMID Editorial, 2024).

Recent advances in Artificial Intelligence (AI) and Natural Language Processing (NLP), particularly through the development of Large Language Models (LLMs), offer the potential to overcome these limitations (Zhuang et al., 2025). These tools can deliver semantic search that transcends literal keyword matching (Alvarez and Morrier, 2025), generate concise summaries of technical responses, and extract topics for building interactive dashboards. They can also support classification and trend detection, link related questions, and assess response quality—enabling deeper analysis of political discourse.

2.1 Comparative Practices and International Benchmarks

Several research initiatives and civic technology projects demonstrate best-practice applications of AI using data from different national parliaments:

- **United Kingdom:** Provides comprehensive APIs and machine-readable formats via https://explore.data.parliament.uk, which supports civic tools like *TheyWorkForYou* (Parliament, 2024).
- **Brazil:** Uses machine learning in its "Ulysses Suite" to categorise citizen input and assist legislative drafting (Inter-Parliamentary Union (IPU), 2022).
- **Italy:** Applies generative AI to cluster and assess thousands of amendments before committee review (Citino, 2024).
- **Finland:** Leverages semantic web technologies in the *ParliamentSampo* platform for concept- and speaker-based exploration (Hyvönen et al., 2022).
- **France:** Employs the *LLaMandement* LLM to summarise complex legislative amendments (Gesnouin et al., 2024).

These systems are commonly underpinned by open data standards, cross-parliamentary collaboration, and responsible AI frameworks (Inter-Parliamentary Union (IPU) and UN/IPU Global Centre for ICT in Parliament, 2024; Inter-Parliamentary Union (IPU) and Parliamentary Data Science Hub, Centre for Innovation in Parliament, 2024).

2.2 Current Capabilities in Malta

While Malta has established a foundational level of transparency through its official PQ portal (Abela and Azzopardi, 2018), the absence of a structured data

pipeline or a public API continues to limit both internal analysis and external reuse. A linked data prototype developed in 2017–2018, known as *PQViz*, demonstrated the feasibility of graph-based exploration (Abela and Azzopardi, 2018). Building on this foundation, the PQ Dashboard presented in this paper – accessible at https://pq.ir.mt – incorporates a range of LLM-based features, as detailed in Section 4.

Despite these advancements, persistent delays and incomplete responses from ministries remain a significant barrier to effective scrutiny (TMID Editorial, 2024). Table 1 provides a comparative overview of PQ access capabilities across various national parliaments.

2.3 AI Techniques Utilised in Parliamentary Analysis

2.3.1 Semantic Search and Retrieval

Conventional search in legislative portals is limited by exact term matching (Bryłkowski and Klikowski, 2025). Semantic search techniques, driven by LLMs, enable retrieval based on meaning and intent (Alvarez and Morrier, 2025). Embedding-based models transform language into vector representations, allowing for similarity matching even in the absence of direct keyword overlap (Bryłkowski and Klikowski, 2025). This significantly enhances search and retrieval through large, complex corpora.

Retrieval-Augmented Generation (RAG) architectures combine traditional retrieval with generative capabilities, enabling factual, grounded responses (Sharma, 2025). This is particularly valuable for civic applications, where hallucinated content could erode public trust. Although RAG improves accuracy, it also introduces challenges such as retrieval noise and outdated knowledge bases. Ongoing research aims to refine these systems for reliable use in governance (Sharma, 2025).

2.3.2 Summarisation and Topic Modelling

LLMs can generate high-quality summaries of PQ responses, aiding readers in quickly understanding detailed content (Alvarez and Morrier, 2025; Siino et al., 2025). France's *LLaMandement* project exemplifies the effectiveness of such summarisation in legislative settings (Gesnouin et al., 2024). Topic modelling further enables the discovery of dominant themes and emerging trends (Polat and Korpe, 2022; Hyvönen et al., 2022), supporting policy research and media scrutiny.

2.3.3 Discourse and Answer Quality Analysis

LLMs can classify PQs based on rhetorical purpose (e.g., factual, policy, accusatory) and assess the quality of responses (e.g., explanatory, evasive) (Alvarez and Morrier, 2025). This supports analysis of discourse strategies and highlights avoidance tactics (Alvarez and Morrier, 2025). Additional work has investigated potential algorithmic bias and the need for domain-aware evaluation frameworks (Cunningham et al., 2025; Rozado, 2024).

2.4 AI in Broader Public Sector Applications

AI is increasingly embedded in public-sector workflows to optimise service delivery, automate approvals, and assist decision-making (Zhao et al., 2025). These applications have been shown to enhance efficiency and public trust by enabling responsive and transparent government operations.

In the United States, over 1,700 AI use cases have been documentedi in December 2024, including fraud detection at the Veterans Administration and decision support at the Social Security Administration (Martorana, 2025). Malta has also seen innovation through AI-based analysis of legal judgments using cross-lingual information retrieval and rhetorical role labelling (Azzopardi, 2024).

2.5 Ethical and Regulatory Considerations

The deployment of AI in public-sector contexts raises important concerns around transparency, bias, and privacy. The *Artificial Intelligence Act* adopted by the European Union (European Parliament and Council of the European Union, 2024) establishes a regulatory framework for trustworthy AI, including classification of high-risk applications such as those used in governance and public services. The Inter-Parliamentary Union (IPU) has complemented this with its *Guidelines for AI in Parliaments*, which outline principles for responsible AI use within legislative contexts, emphasising human oversight, fairness, and accountability (Inter-Parliamentary Union (IPU) and Parliamentary Data Science Hub, Centre for Innovation in Parliament, 2024). These frameworks encourage parliaments and related organisations to adopt ethical practices that safeguard public trust while enabling innovation.

Table 1: Comparative Analysis of PQ Access Features.

Feature	Malta (Current State)	UK (Best Practice)	Canada (Best Practice)	Brazil/Italy/Finland (Advanced AI Examples)
Online Portal Access	Yes, via https://parlament.mt	Yes, via https://questions-statements.parliament.uk and data API	Yes, Open Parliament portal and LEGISinfo	Integrated into main websites
Search Functionality	Basic keyword search	Advanced filtering and API support	Advanced portal search and filtering	Semantic/NL search, AI dashboards
Data Format Availability	PDF and DOC only	Open formats (XML, JSON, CSV)	CSV/XML via API	Structured data and Linked Open Data
Public API	Not available	Fully documented API	API access supported	APIs power civic and internal tools
Visualisation / Analysis	PQViz (legacy)	External tools (e.g., TheyWorkForYou)	Tools like OpenParliament.ca	AI-driven dashboards and semantic portals
AI Integration Status	Minimal / exploratory	Emerging	Early-stage interest	Fully operational (Brazil, Italy), custom LLMs (France)
Answer Timeliness / Quality	Significant delays and evasive answers	Some evasions but procedural accountability	Formal processes	AI being used to assess quality

3 METHODOLOGY

The PQ Dashboard is an innovative online platform developed using entirely open-source technologies, demonstrating a commitment to transparency, reusability, and data sovereignty. This section details the system architecture, operational workflow, and the specific Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques employed.

3.1 System Architecture and Technologies

The system is built upon a robust, Linux-based infrastructure, specifically utilising Ubuntu 24.04 for its operating environment. The core components of the web application are developed using Python Flask, served by Gunicorn, with the front-end interface constructed using standard HTML, CSS, and JavaScript. Data persistence is managed by MongoDB, a NoSQL database, chosen for its schema-less design which is optimal for handling document-oriented data, typically in Binary JSON (BSON) format. MongoDB also supports efficient data retrieval through indexing, crucial for the interactive dashboard.

The PQ Dashboard operates across a two-server architecture:

- **Processing Server:** This in-house server, located within the University of Malta, is dedicated to data acquisition and intensive processing tasks. Its specifications include an Nvidia RTX 4090 GPU with 24GB of dedicated memory, 32GB of RAM, and an Intel i7-13700K CPU (providing 24 virtual processors). This hardware configuration is specifically chosen to support the computational demands of Large Language Models (LLMs) and other NLP operations, enabling the hosting of LLMs up to approximately 12 billion parameters.
- **Publicly Accessible Virtual Private Server (VPS):** This server hosts the web application, making the PQ Dashboard publicly available. It is configured with 8GB of RAM and an Intel(R) Xeon(R) CPU E5-2680 (2 virtual processors). The VPS focuses solely on serving user queries via the web interface, with all heavy data processing offloaded to the dedicated processing server.

3.2 Operational Workflow and Data Acquisition

The system operates on a daily automated cycle to ensure the PQ Dashboard remains up-to-date with the latest parliamentary information. Each night, the data processing server performs the following sequence of operations:

1. It connects to the official Maltese Parliament portal (https://pq.gov.mt) to check for newly available Parliamentary Questions.

2. The system maintains a record of previously processed parliamentary sittings, identifying any new sittings for which data has been submitted.
3. New PQs from these sittings are then scraped from the portal. The scraping process is executed using *Selenium*, which operates a headless Google Chrome browser to interact with the website. JavaScript is employed to identify and extract the salient parts of each PQ from the web page, with Python then retrieving this data and harvesting it into the MongoDB database.
4. The newly acquired PQ data is then processed using the LLM.
5. All processed PQs are stored in the MongoDB database on the data processing server.
6. To maintain synchronisation, the processed PQs are also pushed to the publicly accessible VPS via a custom-built API, ensuring data consistency across both servers.

Each Parliamentary Question is accompanied by several core metadata fields, including: title, question and answer texts, the Member of Parliament (MP) submitting the question, the responding minister and ministry, and supplementary information such as the legislature, PQ number, sitting, date, and classification category.

3.3 Large Language Model (LLM) Processing

The core of the data enrichment process relies on an open Large Language Model. The system currently utilises the *Gemma 2 9B Instruction Tuned* model, released by Google and available via Hugging Face (Google, 2024). This model was selected following comparative testing with other open LLMs available at the time of development, including Meta Llama 3 and Mistral. Gemma demonstrated superior performance in processing Maltese language texts. The choice of LLM size was constrained by the available 24GB GPU memory on the processing server, which efficiently supports models up to approximately 12 billion parameters. Running the LLMs on CPU was tested but found to be prohibitively slow.

The LLM performs several critical tasks on the scraped PQ data:

- **Translation:** The Title, Question, Answer, and associated ministry fields of each PQ are automatically translated from Maltese into English.
- **COFOG-99 Categorisation:** The LLM identifies and assigns the relevant COFOG-99 (Classification of the Functions of Government) category to each PQ (United Nations Statistics Division, 2024). This United Nations standard provides a thematic classification of government activities.
- **Keyword Extraction:** Key terms are extracted from the PQ content to facilitate thematic browsing and analysis.
- **Inter-PQ Link Identification:** The LLM identifies references within PQs to other Parliamentary Questions, establishing outgoing links.

Given the inherent limits of LLM queries, particularly for longer PQs, the system implements a strategy where lengthy PQs are split into sections, processed separately by the LLM, and then the results are merged to ensure comprehensive analysis.

3.4 Additional Data Processing and Synchronisation

Beyond the LLM-driven tasks, additional processing is performed to enhance data navigability. Once outgoing links from a PQ to other PQs are identified, the system automatically establishes corresponding incoming links for the referenced PQs. This allows the user interface to display both outgoing and incoming links, improving the user's ability to trace interconnected legislative discourse, a feature not available on the official parliamentary portal. All extracted and processed data is then harvested into the MongoDB database on the processing server.

Data synchronisation between the processing server (located within the University of Malta's secure network and not publicly accessible) and the public-facing VPS is managed via a custom-built API. The processing server pushes newly processed data to the VPS over HTTPS through a custom-built API. The API, residing on the VPS, listens for incoming requests via HTTPS and is only accessible from a predefined range of trusted IP addresses, enhancing security. This API exposes calls to check which parliamentary sittings are currently stored on the server database, and which PQs are stored for that sitting. It then provides methods to store the details about a new parliamentary sitting and to add PQs to a sitting. Through this API, the databases on the data processing server and the VPS are kept automatically synchronised, ensuring up-to-date content on the public dashboard without exposing the processing infrastructure. The data processing server specifically uses this API to push the data relevant to newly processed PQs to the VPS thus ensuring that the PQ Dashboard is kept updated.

3.5 Web Portal

The online PQ Dashboard, accessible at https://pq.ir.mt, serves as the user-facing component, hosted on the VPS. It does not perform additional data processing but provides an intuitive interface for users to query and explore the processed parliamentary data. The system was developed using open-source technologies. Further details on the functionalities provided within this dashboard are described in Section 4 below.

The backend is powered by Python Gunicorn, proxied behind Apache2. The frontend is built with HTML, CSS, and JavaScript, utilising Bootstrap to ensure a responsive design across various screen sizes. Other libraries incorporated include DataTables for efficient data presentation and quick searching within tables, and Chart.js for data visualisations. MongoDB is used as the database.

4 PQ DASHBOARD: FUNCTIONALITIES AND EXTRACTED INSIGHTS

The PQ Dashboard is available publicly on https://pq.ir.mt. It currently covers Parliamentary Questions (PQs) from the current (14th) legislature, and the system is updated nightly with newly published PQs available on the official PQ portal (process described in Section 3). Figure 1 shows a screenshot of the dashboard.

Key functionalities available to users include:

- **Browse by Categories:** Users can explore PQs classified into COFOG-99 Categories, enabling thematic navigation of government functions.
- **Browse by Keywords:** The dashboard allows users to search and browse PQs using extracted keywords, facilitating the discovery of frequently discussed topics.
- **Analysing MPs' Activity:** The platform provides tools to analyse the activity patterns of Members of Parliament (MPs).
- **Overview by Ministry:** Users can view PQs organised by the ministries they were directed to, offering insights into ministerial engagement.
- **Advanced Filtering:** Users can filter by date, MPs, ministry, category and keyword. Any combination is possible, and multiple values for MPs, ministry, category and keyword can be entered. Select2 is utilised to allow users easy search through the dropdown values.
- **Bilingual Access:** The dashboard provides access to PQs in both their original Maltese and the automatically translated English versions, catering to a wider audience.
- **Official Source Linking:** For each PQ, a direct link is provided to the official version on https://pq.gov.mt. This ensures transparency and allows users to verify any information or inconsistencies with the authoritative source.
- **Inter-PQ Navigation:** Users can seamlessly browse between PQs via identified incoming and outgoing links. This feature significantly enhances research capabilities, as PQs often refer to previous questions, and the dashboard eliminates the need for manual searches for referenced documents.

The subsequent subsections provide further details and some insights obtained when utilising the different functionalities.

4.1 Browse by Categories

Users can view the distribution of PQs across the different COFOG-99 categories. They can then drill down by selecting a particular category to view the list of PQs categorised within that category. This action opens the list of PQs as a clickable table, allowing users to view individual PQs. Furthermore, horizontal bar charts are displayed, showing the distribution of ministries targeted within these PQs and the distribution of MPs who posed these questions.

For instance, the majority of PQs published from the start of the current legislature (16th May 2022) to mid-July 2025 related to General Public Services (8855 PQs). The second most popular COFOG-99 category within the same period is Economic Affairs, with 4358 PQs. Of these PQs (related to General Public Services), 866 PQs were addressed to the Minister for National Heritage, the Arts and Local Government, and 594 were addressed to the Ministry for Transport, Infrastructure and Public Works. The MP who asked the majority of PQs from this category was Jerome Caruana Cilia (929 PQs), followed by Ivan Bartolo (658 PQs).

4.2 Analysing MPs' Activity

Users can view a list of MPs who posed questions, along with statistics on the number of PQs posed as per the filtering selected. For instance, Jerome Caruana Cilia submitted the most PQs (2195) in this legislature, followed by Graziella Galea (2006 PQs). However, if only PQs related to Environmental Protec-

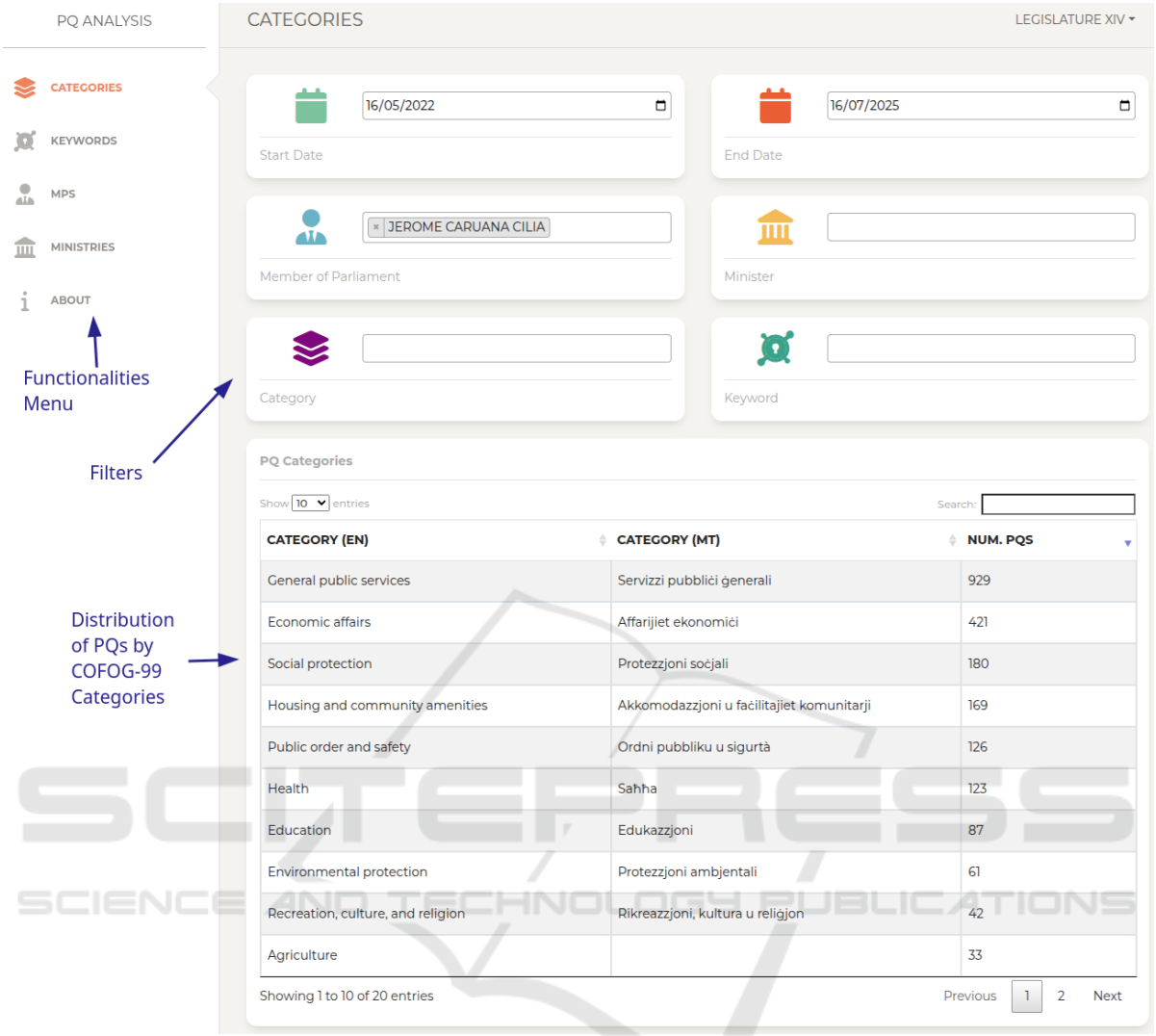

Figure 1: The PQ Dashboard landing page showing the distribution of PQs by Jerome Caruana Cilia according to the different COFOG-99 categories.

tion are considered, Rebekah Borg submitted the most questions (117), followed by Graziella Galea (95).

Users can then select a particular MP to drill down into that MP's activity. This opens a clickable table listing PQs, along with horizontal bar charts showing the distribution of ministries targeted by that MP, the distribution of categories to which those PQs belong, and the distribution of keywords associated with those PQs.

For instance, when considering the PQs related to Environmental Protection by Rebekah Borg, the majority of them were submitted to the Ministry for the Environment, Energy and the Regeneration of the Grand Harbour (54 PQs), the Ministry for the Environment, Energy and Public Health (32 PQs) and the Ministry for the Environment, Energy and Enterprise (9 PQs). It should be noted that when ministerial portfolios are modified by the government, the affected ministries appear as 'new' ministries. Most of the PQs submitted by Rebekah Borg related to Environmental Protection concerned the Environment and Resources Authority. This was clearly highlighted by the visualisation of keyword distribution.

4.3 Overview by Ministry

This section displays a list of ministries along with the number of PQs addressed to each of them. Users can select a ministry to view a list of PQs addressed to that ministry (subject to any additional filtering provided by the users), a horizontal bar chart showing the distribution of MPs who posed these questions, a hori-

zontal bar chart showing the distribution of categories to which those PQs belong, and a horizontal bar chart showing the distribution of keywords associated with those PQs.

For example, the Ministry for Education, Sport, Youth, Recreation and Innovation is the most commonly addressed ministry (2911 PQs). Over 20% of these PQs (632) were posed by Justin Schembri, with Graziella Galea submitting the second most (249 PQs). The great majority of these PQs were categorised within the obvious COFOG-99 category 'Education'. Others were categorised within 'General Public Services', and 'Recreation, Culture and Religion', amongst others. Some of the most popular keywords include 'Schools', 'Students' and 'Sport'.

Figure 2 provides additional insights, showing the breakdown of PQs directed to the Ministry for Gozo and Planning between May and July 2025, classified by the Member of Parliament (MP) posing the question and the corresponding COFOG-99 categories.

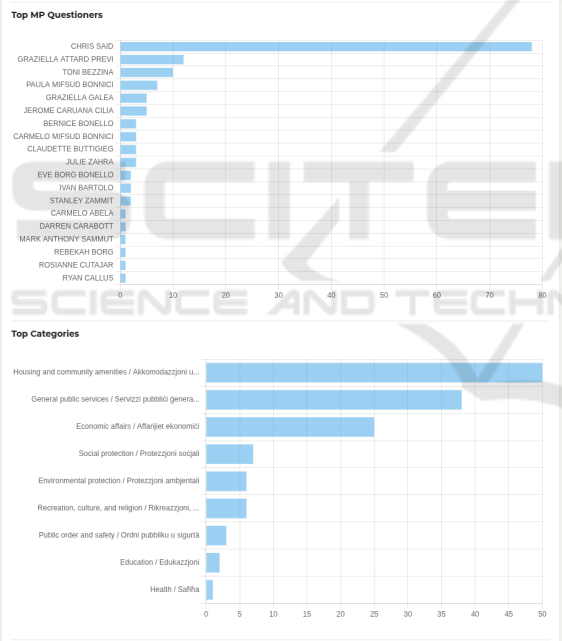

Figure 2: Insights into Parliamentary Questions (PQs) directed to the Ministry for Gozo and Planning from May to July 2025, highlighting the activity of posing MPs and the distribution across COFOG-99 categories.

4.4 Viewing a PQ

When a user is presented with PQs in a datatable, these are clickable to view the details of that PQ. The presented details for a PQ include:

- Details available in the original official PQ, namely:
  - PQ number
  - Sitting
  - MP posing the question
  - PQ Type (Oral vs. Written)
  - The Ministry to which the PQ is addressed (in Maltese)
  - The Minister answering the PQ
  - PQ Title (in Maltese)
  - Question Text (in Maltese)
  - Answer Text (in Maltese)
- Translated version of the applicable fields, more specifically:
  - Ministry name
  - Title
  - Question Text
  - Answer Text
- Other details retrieved from the LLM processing:
  - COFOG-99 Category (in both Maltese and English)
  - Extracted Keywords (in both Maltese and English)
  - Links to PQs cited within this PQ (a.k.a. outgoing links)
  - Links to PQs citing this PQ (a.k.a. incoming links)
- Link to the original PQ from the official Parliament of Malta portal (https://pq.gov.mt).

Figure 3 shows the PQ Dashboard's detailed view of a specific PQ.[1] It should be noted that the original PQ may contain links to documents laid on the Clerk's table. These 'documents laid' are to date not processed or displayed in the PQ Dashboard.

5 EVALUATION

The PQ Dashboard was informally evaluated through demonstrations to various key stakeholders, including officials from the Parliament of Malta, civil service officials within the Government of Malta, and former Members of Parliament. Qualitative feedback was recorded during these sessions. While acknowledging that a more rigorous and systematic evaluation would have been desirable, the insights gathered from this qualitative feedback still proved valuable in assessing the system's utility and identifying areas for future enhancement.

[2]The original PQ is accessible at https://pq.gov.mt/pqweb.nsf/06d013e9f9ab0283c12568f50054014f/c1257d2e0046dfa1c1258c8100253c15 (last accessed: September 2025).

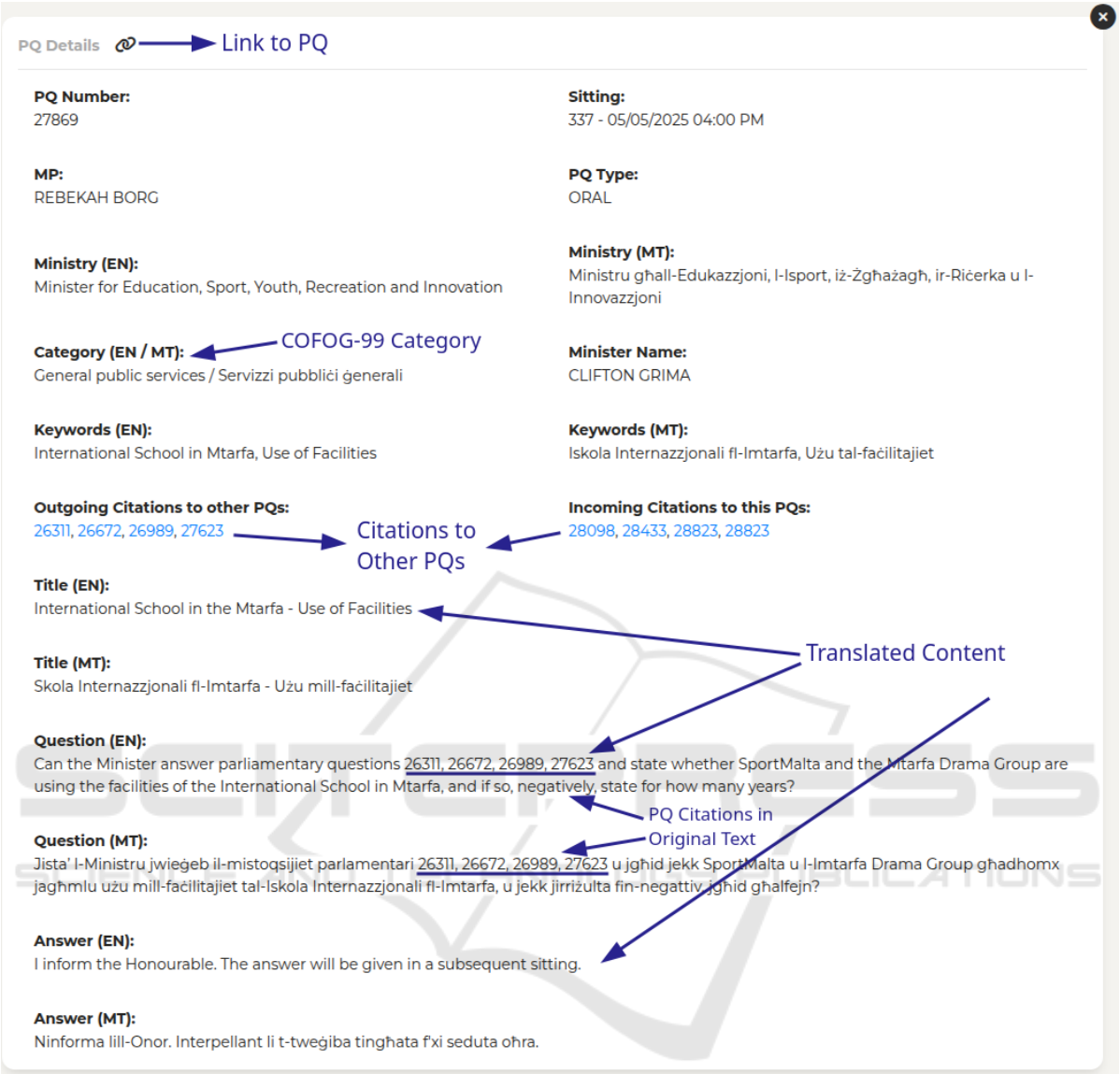

Figure 3: A detailed view of a Parliamentary Question (PQ) within the dashboard, displaying all relevant information.[2]

Positive Feedback

The overall feedback received was overwhelmingly positive, particularly highlighting several key strengths of the system:

- **Multilingual Search Capability:** A significant advantage highlighted was the ability to search Parliamentary Questions (PQs) using English, which greatly enhances accessibility for non-Maltese speakers and facilitates broader research.
- **Simplified Navigation through Inter-PQ Links:** Users appreciated the simplification of PQ searches, especially when different PQs refer to each other. The system's ability to identify and link related questions streamlines the research process, eliminating the need for users to initiate new searches for referenced PQs, a limitation of the official parliamentary portal.
- **Thematic Overviews:** The dashboard provides clear overviews of the typical subjects of questions, allowing users to quickly grasp prevailing topics and trends in parliamentary discourse.

Beyond general research into PQs, specific useful applications of such a system were highlighted:

- **Tracking Civil Service Projects:** The system was identified as a valuable tool for keeping track of current projects related to the civil service. It was noted that, until recently, PQs were often

the only means to acquire certain governmental data (e.g., data related to governmental authorities) when not shared through other official channels, making the dashboard a key tool for data acquisition in such scenarios.

- **Ministerial Consistency Checks:** Ministerial officials found the system useful for preparing answers to new PQs, enabling them to ensure consistency with information previously provided in past PQ answers.

Identified Shortcomings and Areas for Improvement

Despite the positive reception, a number of shortcomings and areas for improvement were also noted, providing crucial guidance for future development:

- **Translation Nuances:** While the automatic translations were generally considered to be of good quality, it was observed that they occasionally failed to capture the exact meaning implied in the original Maltese text. Such issues were typically present in "canned" parts of the text—frequently recurring phrases in PQs. For instance, the phrase "Ninforma lill-Onor. Interpellant illi..." ("I inform the Honourable Member that...") was sometimes translated as "I inform the Honourable...", omitting the word 'Member'. A suggested workaround involves implementing rules to apply pre-prepared, accurate translations for such common phrases, thereby correcting the AI-generated output.
- **Clear Source Attribution:** It was recommended that the system explicitly mark which parts of the displayed text were obtained directly from official sources (e.g., original Maltese question and answer texts) and which were generated or processed by AI. This would enhance transparency and user trust.
- **System Independence Disclaimer:** To avoid misconceptions due to potential inaccuracies in AI-generated data, it was suggested that the system explicitly state its separation from the official parliamentary version. This clarifies that the PQ Dashboard is a supplementary tool and defers to https://pq.gov.mt as the authoritative source.
- **Limited Historical Coverage:** The current system is limited to PQs from the 14th legislature. Stakeholders expressed a strong desire for the system to be extended to cover all PQs available online, from the 9th legislature onwards, as provided by the official PQ portal. This expansion would unlock a wealth of historical data for more comprehensive longitudinal analysis.
- **User Language Preference:** The current interface displays both Maltese and English versions of PQs side-by-side. Feedback suggested implementing a user-selectable language flag, allowing users to choose their preferred language for all subsequent interactions within the dashboard, rather than a dual display.

The qualitative feedback, despite its informal nature, has been instrumental in validating the core utility of the PQ Dashboard and in clearly delineating a roadmap for its future development, ensuring that subsequent enhancements directly address user needs and improve the system's accuracy and usability.

6 CONCLUSIONS AND FUTURE WORK

The PQ Dashboard stands as a compelling demonstration of how accessible public data can be achieved through the strategic utilisation of relatively low-cost hardware and the integration of Machine Translation. By offering access to Parliamentary Questions (PQs) in both their original Maltese and automatically translated English versions, the system broadens substantially the potential audience for this vital public information. Furthermore, the extraction of key details from each document, such as COFOG-99 categories, keywords, and inter-document citations, serves to enhance accessibility further. These extracted details are then leveraged to generate valuable aggregations, enabling users to 'profile' Members of Parliament (MPs) and Ministries, and identify their primary interests, thereby offering deeper insights into parliamentary activity.

A fundamental principle underpinning the system's development was the exclusive reliance on free and publicly available open-source technologies, purposefully eschewing commercial services. This approach not only ensures inherent cost-effectiveness but also serves as a robust showcase for the development of secure and privacy-preserving Artificial Intelligence (AI) solutions within the public sector. This is particularly relevant even when dealing with publicly available information where immediate privacy concerns might appear less pressing, as it underscores a commitment to data sovereignty and ethical AI practices through on-premise Large Language Model (LLM) deployment.

Future work will expand the capabilities and scope of the PQ Dashboard. The first priority is to ad-

dress the limitations identified during evaluation (see Section 5), with emphasis on extending coverage to earlier legislatures—starting from the 9th—whose records are already publicly available. Secondly, the system will be updated to incorporate more recent or advanced Large Language Models, ensuring continued state-of-the-art performance. In addition, discourse analysis is planned to capture different question types (e.g., factual, policy-oriented, or accountability questions) and to categorise response types (e.g., factual, explanatory, or evasive answers), offering a more nuanced understanding of parliamentary dialogue. Furthermore, a more formal and structured user study is planned to systematically assess the system's strengths and weaknesses.

ACKNOWLEDGEMENTS

The authors acknowledge the use of Artificial Intelligence (AI) language models in the drafting and refinement of this manuscript. Specifically, the Gemini family of Large Language Models (developed by Google) and OpenAI's ChatGPT (version GPT-4, July 2025) were utilised to support summarisation, language refinement and structural organisation. All AI-assisted content was thoroughly reviewed, edited, and validated by the authors to ensure accuracy, originality, and compliance with academic standards.

REFERENCES

Abela, C. and Azzopardi, J. (2018). Analysing and visualising parliamentary questions: A linked data approach. In Szymański, J. and Velegrakis, Y., editors, *Semantic Keyword-Based Search on Structured Data Sources*, pages 32–43, Cham. Springer International Publishing.

Alvarez, R. M. and Morrier, J. (2025). Measuring the quality of answers in political q&as with large language models. *Political Analysis*, page 1–18.

Azzopardi, J. (2024). Translating justice: A cross-lingual information retrieval system for maltese case law documents. In Goharian, N., Tonellotto, N., He, Y., Lipani, A., McDonald, G., Macdonald, C., and Ounis, I., editors, *Advances in Information Retrieval*, pages 236–240. Springer Nature Switzerland, Cham.

Bryłkowski, A. and Klikowski, J. (2025). Large language models in legislative content analysis: A dataset from the polish parliament. Available at https://arxiv.org/abs/2503.12100.

Citino, Y. M. (2024). Leveraging automated technologies for law-making in italy: Generative ai and constitutional challenges. *Parliamentary Affairs*, 78(3):625–647.

Cunningham, E., Cross, J., and Greene, D. (2025). Identifying algorithmic and domain-specific bias in parliamentary debate summarisation. Available at https://arxiv.org/abs/2507.14221.

European Parliament and Council of the European Union (2024). Artificial Intelligence Act (regulation (eu) 2024/1689). Entered into force 1 August 2024; full text available via EUR-Lex. https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

Gesnouin, J., Tannier, Y., Silva, C. G. D., Tapory, H., Brier, C., Simon, H., Rozenberg, R., Woehrel, H., Yakaabi, M. E., Binder, T., Marie, G., Caron, E., Nogueira, M., Fontas, T., Puydebois, L., Theophile, M., Morandi, S., Petit, M., Creissac, D., Ennouchy, P., Valetoux, E., Visade, C., Balloux, S., Cortes, E., Devineau, P.-E., Tan, U., Namara, E. M., and Yang, S. (2024). Llamandement: Large language models for summarization of french legislative proposals.

Google (2024). Gemma 2 9b instruction tuned model. Hugging Face. https://huggingface.co/google/gemma-2-9b-it.

Hyvönen, E., Leskinen, P., Sinikallio, L., Mela, M. L., Tuominen, J., Elo, K., Drobac, S., Koho, M., Ikkala, E., Tamper, M., Leal, R., and Kesäniemi, J. (2022). Finnish parliament on the semantic web: Using parliamentsampo data service and semantic portal for studying political culture and language. In *Digital Parliamentary Data in Action (DiPaDa 2022) Workshop, CEUR Workshop Proceedings, Vol. 3133*, pages 69–82, Uppsala, Sweden. CEUR-WS.org. Presented 15 March 2022; licensed under CC BY 4.0.

Inter-Parliamentary Union (IPU) and Parliamentary Data Science Hub, Centre for Innovation in Parliament (2024). Use Cases for AI in Parliaments. Published in partnership with the Parliamentary Data Science Hub, following IPU Assembly resolution on AI (October 2024). Available at: https://www.ipu.org/resources/publications/reference/2024-12/use-cases-ai-in-parliaments.

Inter-Parliamentary Union (IPU) (2022). Brazil: A digitally mature parliament. Case study; published 1 June 2022 on the IPU News & Case Studies page https://www.ipu.org/news/case-studies/2022-06/brazil-digitally-mature-parliament.

Inter-Parliamentary Union (IPU) and Parliamentary Data Science Hub, Centre for Innovation in Parliament (2024). Guidelines for ai in parliaments. Published December 2024; launched at event on 3 December 2024; framework for responsible AI use in parliamentary contexts https://www.ipu.org/resources/publications/reference/2024-12/guidelines-ai-in-parliaments.

Inter-Parliamentary Union (IPU) and UN/IPU Global Centre for ICT in Parliament (2024). World e-parliament report 2024. Based on survey of 115 parliaments in 86 countries; licensed under CC BY-NC-SA 4.0. https://www.ipu.org/resources/publications/reports/2024-10/world-e-parliament-report-2024.

Koehn, P. and Knowles, R. (2017). Six challenges for neural machine translation. In Luong, T., Birch, A., Neubig, G., and Finch, A., editors, *Proceedings of the First*

Workshop on Neural Machine Translation, pages 28–39, Vancouver. Association for Computational Linguistics.

Martorana, C. (2025). Ai in action: 5 essential findings from the 2024 federal ai use case inventory. Accessed July 2025; official .gov publication https://www.cio.gov/ai-in-action/.

Parliament, U. (2024). Explore data from the uk parliament. Online data portal. https://explore.data.parliament.uk/ Accessed: 2024-10-09.

Parliament of Malta (2022). Parliament of malta — parliamentary questions. https://pq.gov.mt. Accessed: 2025-07-23.

Polat, H. and Korpe, M. (2022). Estimation of demographic traits of the deputies through parliamentary debates using machine learning. *Electronics*, 11(15).

Ranathunga, S., Prifti Skenduli, M., Shekhar, R., Alam, M., and Kaur, R. (2022). Neural machine translation for low-resource languages: A survey. *ACM Computing Surveys*, 55.

Rozado, D. (2024). The political preferences of llms. *PLOS ONE*, 19:1–15.

Sharma, C. (2025). Retrieval-augmented generation: A comprehensive survey of architectures, enhancements, and robustness frontiers. Available at https://arxiv.org/abs/2506.00054.

Siino, M., Falco, M., Croce, D., and Rosso, P. (2025). Exploring llms applications in law: A literature review on current legal nlp approaches. *IEEE Access*, 13:18253–18276.

TMID Editorial (2024). Pqs should be answered. *The Malta Independent*. Editorial ("TMID Editorial: PQs should be answered").

United Nations Statistics Division (2024). Revision of the classification of the functions of government (cofog). Webpage, UN Statistics Division. https://unstats.un.org/unsd/classifications/cofog/revision.

Zhao, X., Huo, Y., Abedin, M. Z., Shang, Y., and Alofaysan, H. (2025). Intelligent government: The impact and mechanism of government transparency driven by ai. *Public Money & Management*, 0(0):1–12.

Zhuang, Z., Chen, J., Xu, H., Jiang, Y., and Lin, J. (2025). Large language models for automated scholarly paper review: A survey. *Information Fusion*, 124:103332.