

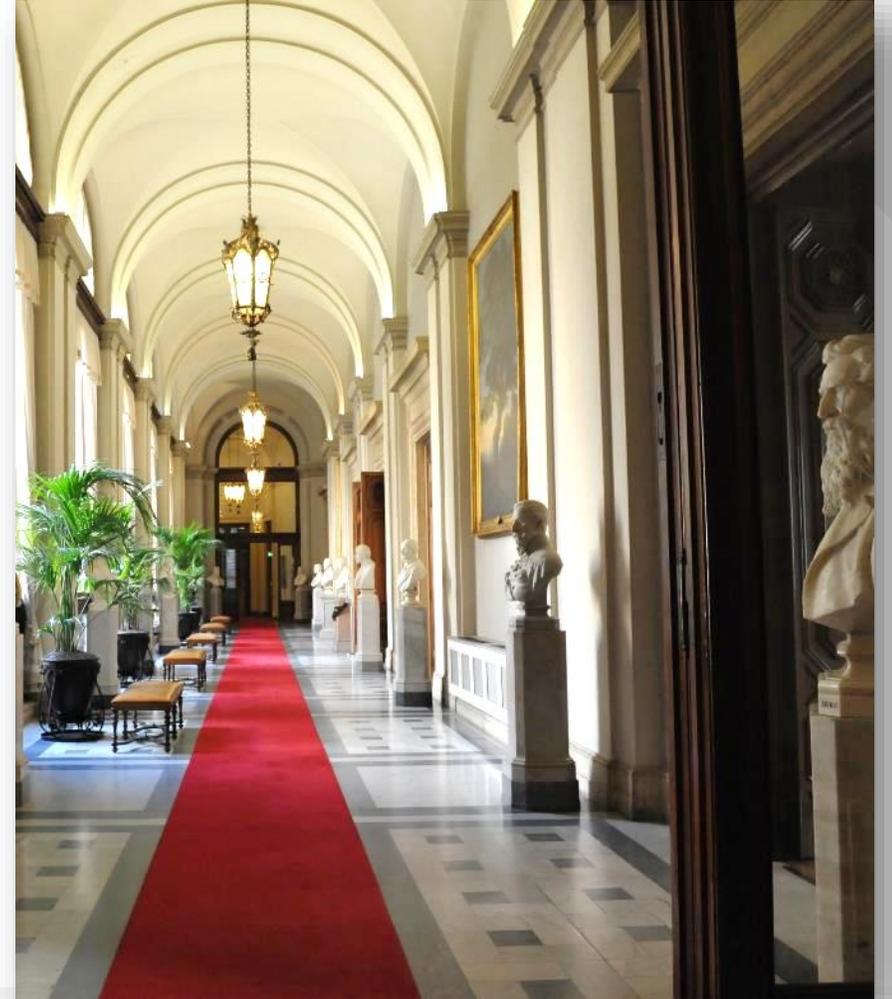
AI experiences at Italian Chamber of Deputies

ICON • S Italian Chapter – University of Trento



Agenda

- **Projects and experiments with AI and genAI**
 - consolidated solutions
 - early experiments
 - recent experiments
- **Rules of conduct**
- **Challenges**

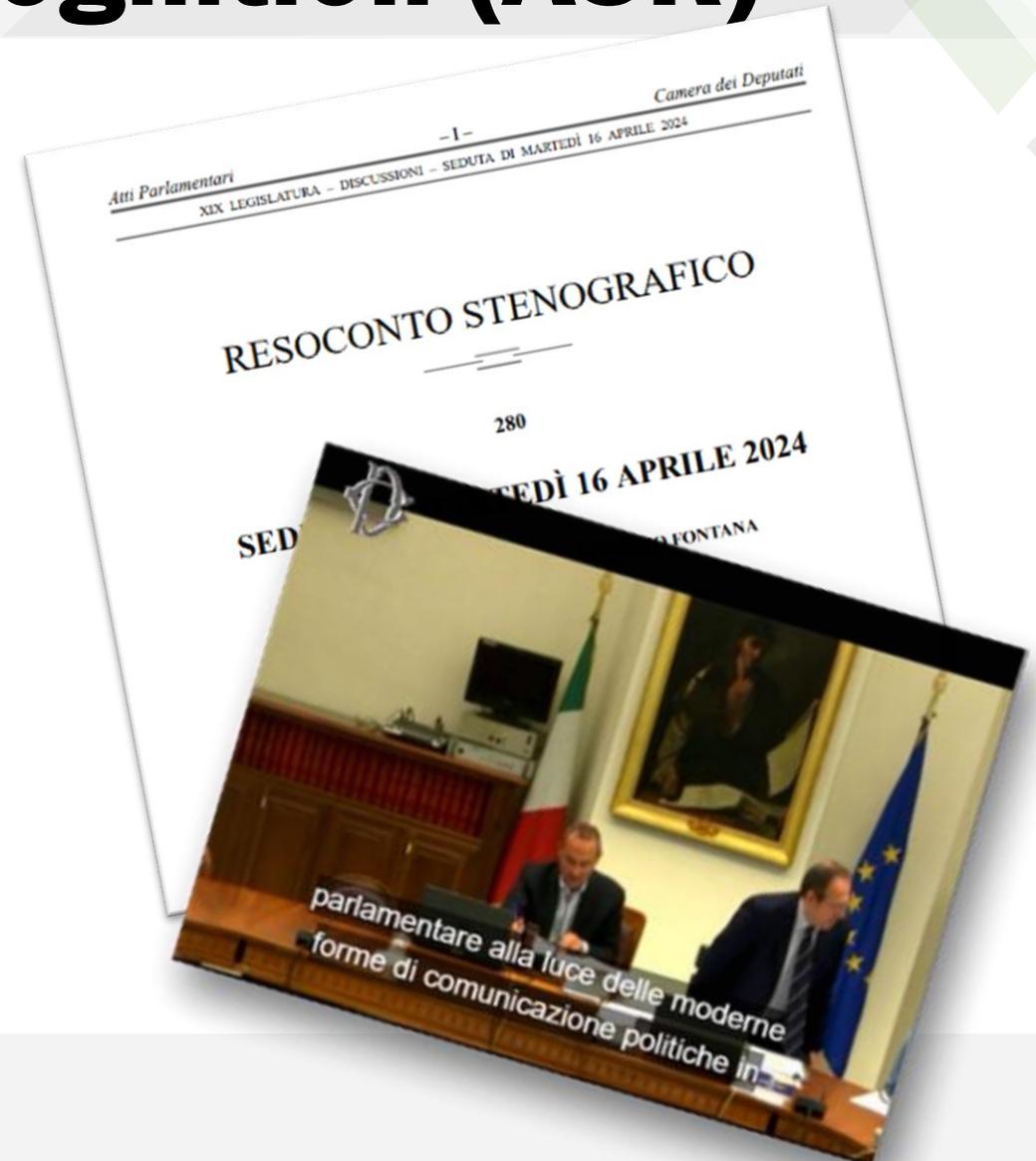


Consolidated solutions

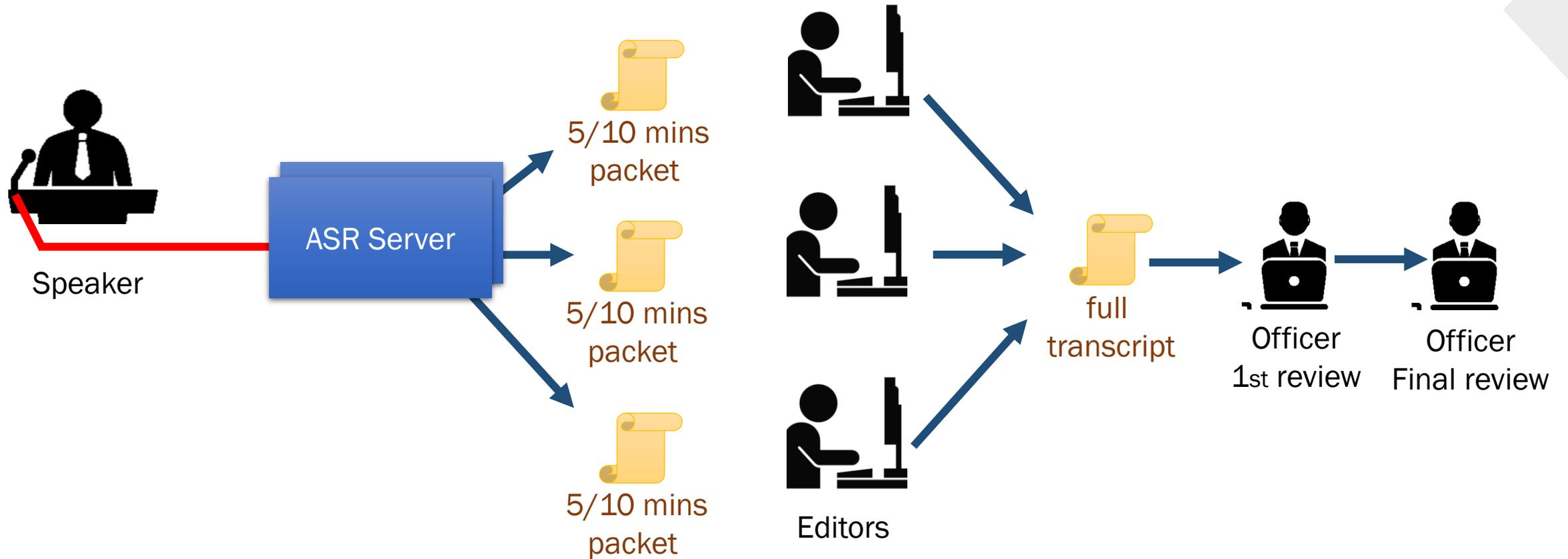


Automatic Speech Recognition (ASR)

- Automatic production of **verbatim reports** of *Assembly sessions*: transcription of speech plus human supervision ("*human in the loop*")
- Automatic production of **assembly sessions subtitles** and their insertion **inside the video** available through web TV and institutional website



ASR – human review for assembly sessions

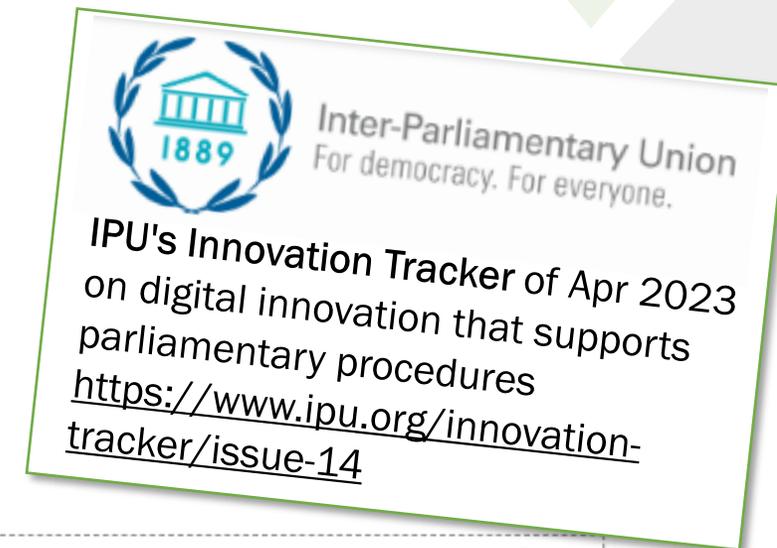


Amendments sorting and similarities

The file of amendments, referred to a particular bill, is processed by an AI algorithm for:

- **sorting the amendments** according to voting order
- computing the **similarity scores** among amendments

This solution is based on *text mining*, *NLP* and *semantic distance* computation.



Esame Ammissibilità Concluso Drafting concluso

Al comma 3, primo periodo, dopo le parole: da adottare aggiungere le seguenti: , previa intesa in Conferenza unificata ai sensi dell'articolo 9 del decreto legislativo 28 agosto 1997, n. 281,.

**1.10. (ex 1.1.)(Prima Riformulazione) Grimaldi (AVS), Zaratti
Proposte emendative identiche: 1.10, 1.11, 1.12, 1.13, 1.14, 1.15

Proposte simili: 1.15.(0,991)

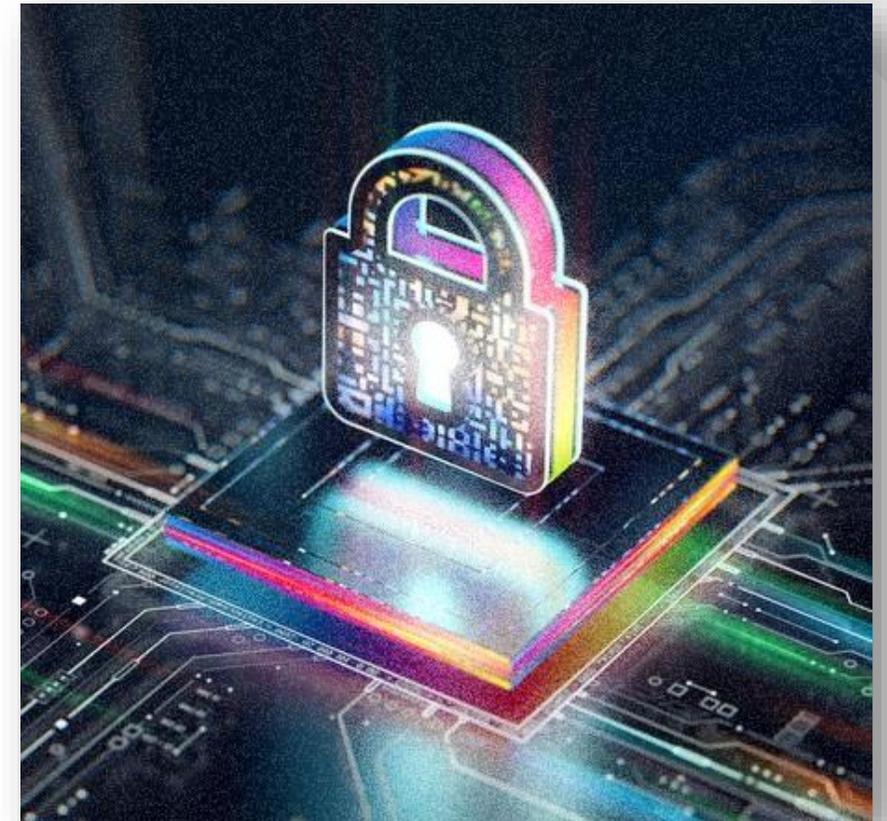
Iter in Commissione: Segnalato Derivato (28/03/2024); Riformulato (12/04/2024); Approvato (12/04/2024)



Tools used for cybersecurity

The *Security Operation Center (SOC)*, which is part of the *Logical Security Office*, is equipped with tools and technologies that use *AI techniques* for activities such as:

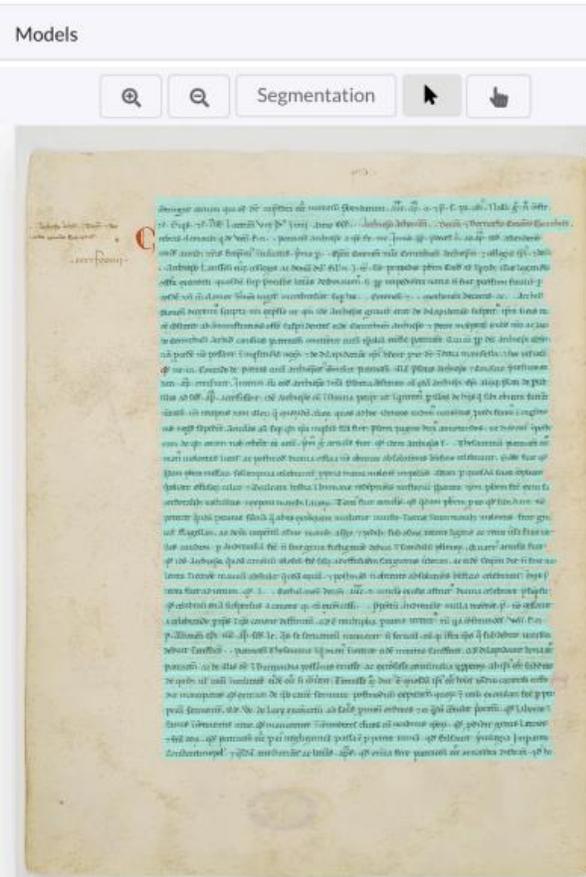
- anomaly detection
- user behavioral analysis
- correlation of security events
- passive analysis of network traffic
- detection of differences from the usual behavior of users and analysis of the dangerous behaviors
- **vulnerability verification** of software systems: AI simulates the behavior of a potential attacker to test defense systems



Early experiments



Transcription of manuscripts



Lines	PDF	Raw Text
		contingat certum qui esse dicitur in prefata ecclesie numerum prebendarum auctoritate apostolia onfirmam e presentis scripti patroinio omnium. Nulli ergo nosto confirmtionis
		etcetera Siquis etcetera Datum Laterani Vij Idus Iunij Anno Octotri Archiepiscopo Athenensis Decino et degnardo Canonico Bntbinis
		rebris lamoribus qui de Venerabilm fratrem sotu paracensis archiepiscopo a tipsoe fe maenete Juniocenis pee predecesis satr ad apostolice sedis ascenderant
		tandus auribus nostris frequentius inculcatis prstmo per Episcopum Oronensim nuricite Corinthiensis Archiepiscopum et collegis ipsius et deeniem
		Archiepiscopo Larissenses eiusque collegis ac demum dilectus filium vetu Jt tituli sancte praxedis presbiterum Cardinlm tunc inpartibus illis legationis
		constertum exercenti quasdam super premissis litteras destinavimus se peptr impedimenta varia non fuit processum finaliter per
		eosdem Unde cum clamor huiusmodi igitur increbrescet super hoc Coronensi et Mothonensi decanis ac Archidiaconm
		Mlonensis direximus scripta nostra ex presso ut quai idem Archiepiscopus graviter erat de dilapidatione suspectus ipsum si eis ita
		esse constaret ab amminustratoin offro susprudentes eidem Cornthiensis Archiepiscopo et pero malpens Subdis nostro ac lalacunes
		bos Corinthiensis Archidiaconus canonicis Patracensis committerent curam temporalium ecclisle paracensis Qui cum peper dicti Archiepiscopi absen
		tuam procedere non possent in inquisitionis negotio et de dilapidatione ipsius haberent prout dicitur inditia manifesta et hoc intercera

The research collaboration with the University of Rome3 realized an OCR system capable of **transcribing handwritten manuscripts** from documents in the *Historical Archive*.

The OCR uses AI models based on *deep learning*.

The application allows users to:

- 1) upload the jpg image of the page to be transcribed
- 2) start **segmentation**, the identification of all the text lines in the page
- 3) execute **text transcription**, the extraction of the text content from the identified lines

EuroVoc classification of documents

Text to be classified

10.7.2019 IT Gazzetta ufficiale dell'Unione europea L 184/3 ACCORDO tra l'Unione europea e il Principato del Liechtenstein ai fini dell'applicazione di talune disposizioni della decisione 2008/615/GAI del Consiglio sul potenziamento della cooperazione transfrontaliera, soprattutto nella lotta al terrorismo e alla criminalità transfrontaliera, della decisione 2008/616/GAI del Consiglio relativa all'attuazione della decisione 2008/615/GAI sul potenziamento della cooperazione transfrontaliera, soprattutto nella lotta al terrorismo e alla criminalità

Suggested labels (and relative scores)

913 - Europa occidentale

913 - Europa occidentale	86%
3489 - politica di cooperazione	85%
4040 - relazioni dell'Unione europea	84%
122 - paesi membri del Consiglio d'Europa	84%

Automatic EuroVoc classification of legislative documents, in collaboration with *CINI* (*National Inter University Consortium for Informatics*).

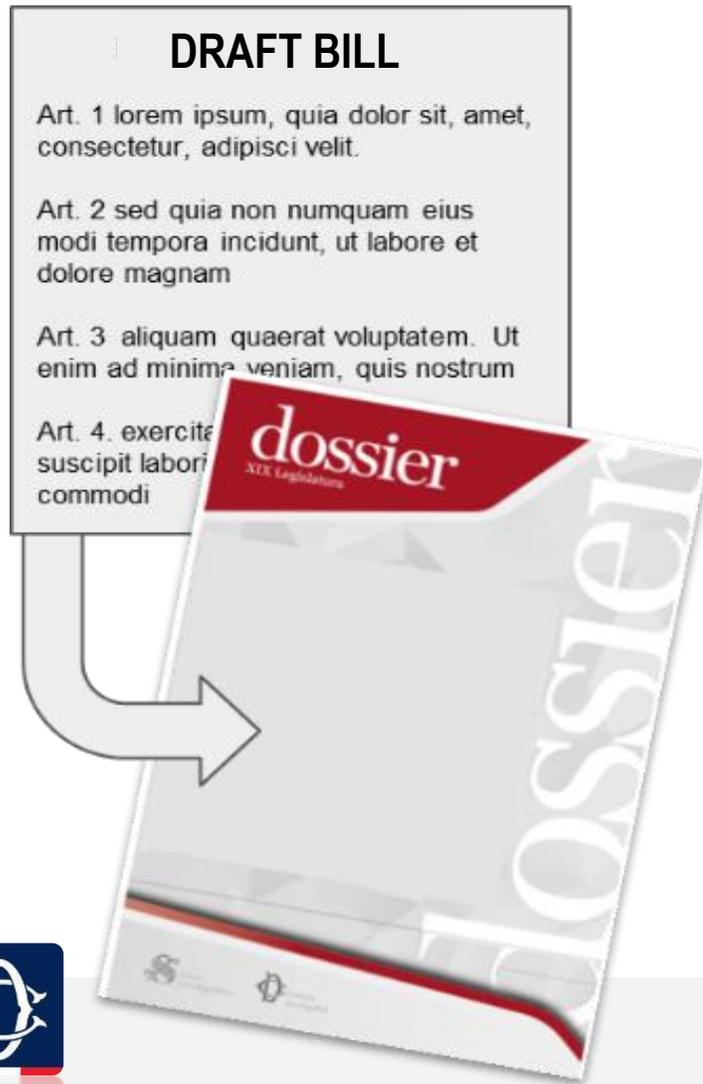
The **training data** consists of Italian **legislative documents** published in the Official Gazette from 2013 to 2023.

Experiments performed with **different models**:

- legal BERT
- Italian legal BERT
- BERT base Italian XXL (better)

Result: classifier with items **up to the level of EuroVoc descriptors** (546 classes); good compromise between expressiveness of classes and availability of data.

Production of summary notes from bills



The goal is to build a system that uses generative AI models (LLama2) to suggest an outline for writing an explanatory dossier for a draft bill.

It was trained from examples of draft bills and their explanatory dossiers actually produced.

Considerations/complications: the *dossier* is not just a summary of a bill, but is an analysis of it, enriched with:

- explication of the cited **legislative references** and relative impact
- specific knowledge
- particular contextual elements
- inclusion of tables/images



Latest experiments



Transcription / translation podium

Interactive podium with subtitling and automatic translation in real time

The speaker talks in his native language, the podium translates into another language (English, French, Spanish, ...) and shows the realtime translation on the display

Benvenuti all'incontro di oggi in cui presenteremo la nostra nuova soluzione



Welcome to today's meeting where we will present our new solution

cabolo podium



Querying Gov's DB on implementing acts

Natural language querying of Government's database related to monitoring and parliamentary control on implementing acts.

The solution combines NLP and Generative AI techniques to:

1. Understand the user intent
2. Generate executable SQL code that the database can process
3. Transform the query results into a human friendly text

Summarize the 3 main measures, based on the funding received, of DRAGHI's government

ASK

The three main measures of DRAGHI's government, according to the funding received, are:

- **Terms and methods of allocating resources from the University Housing Fund to acquire new student beds:** this measure received funding of 660 million euros.
 - **Terms and modalities for the disbursement of resources for the support of economic operators in the tourism, entertainment and dekk'automobile sector seriously affected by the Covid-19 emergency:** this measure received funding of 150 million euros.
 - **Indication of the amount, within the limit of 150 million euros, to be allocated to projects for decarbonization of the steel production cycle at the Taranto plant:** this measure received funding of 150 million euros.
- Translated with DeepL.com (free version)

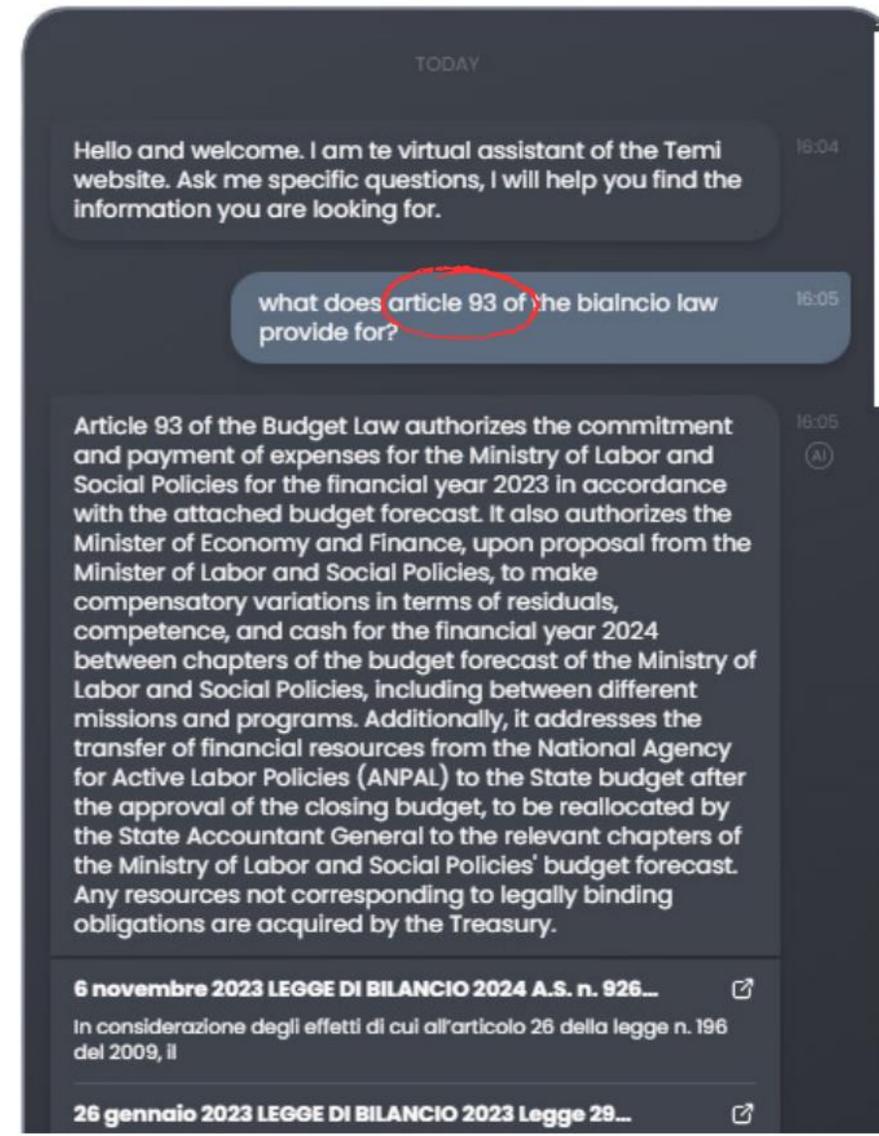
Chatbot for the documentation website

Chatbot to support consultation and search on the *parliamentary documentation website* for:

- Retrieve more quickly information in documents
- Make data more accessible through natural language conversations

How it works:

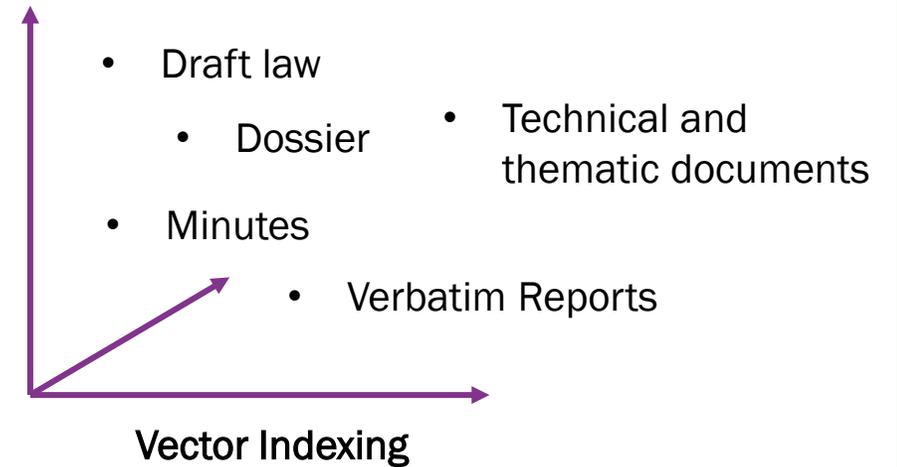
1. **Query Processing:** AI transforms the NL question into a query vector for document alignment using deep learning models
2. **Paragraph ranking and merging:** Semantic analysis combined with vector space modeling to rank and merge document paragraphs. LLMs for response generation
3. **Answer Extraction & Synthesis:** the system extracts precise data snippets directly correlating to the document and create contextually relevant summaries and responses.



GenAI to understand the process of a bill.

It is a **conversational interface** to interact in a new way with knowledge base: minutes containing discussion in Committee and in Assembly, bill text, amendments, hearings, votes, verbatim reports, dossier.

Responses are cross-camera, the system is able to **interpolate the information about the *iter*** of Camera and Senate, channeling them to obtain the answer to the question received.



Hi! How can I help you?

Can you tell me what the AC 1038 provides in terms of customs matters?

The AC 1038 provides a directive for the overhaul of customs regulations. The goals include: 1) Simplifying the documentation and procedures for customs, making increased use of digital technologies. This aims to decrease administrative burdens for taxpayers. 2) Ensuring that every taxpayer participates in the customs procedure and providing a reasonable period for taxpayers to make observations on proposed assessments. 3) Rationalizing and reorganizing the rules concerning risk analysis, in compliance with the rules on the protection of privacy and access to documents, in order to prevent harm to taxpayer guarantees. 4) Introducing specific forms of cooperation between national and foreign administrations that carry out activities to control the correct fulfillment of tax and social security obligations. The purpose of this is to minimize impacts on taxpayers and their economic activities. 5) Promoting the use of digital technologies, including the use of artificial intelligence systems. This is with the aim of achieving full interoperability between databases, making relevant information available and ensuring its timely use. In addition, AC 1038 provides for the enhancement of the collaborative compliance scheme and enables domestic companies that do not qualify for the scheme under current regulations to benefit from its application.

GenAI to understand Italian legislation

- The aim of this collaboration with *State Mint and Polygraphic Institute* is to have a system to query a set of laws in the *Normattiva database* (consolidated versions of normative acts) in natural language
- The system answer includes a summary of the relevant laws and the link to the full text of the law.



The screenshot displays the Normattiva website interface. At the top, it features the logo of the Presidency of the Council of Ministers and the text 'NORMATTIVA IL PORTALE DELLA LEGGE VIGENTE'. Navigation links include 'Home', 'Il progetto', 'Collegamenti veloci', 'Legislazione Regionale', and 'Guida all'uso'. A central banner announces a website update: 'NORMATTIVA SI RINNOVA CON IMPORTANTI E NUOVE FUNZIONALITA''. Below this, a search section titled 'Ricerca degli atti statali in multivigenza' offers 'Ricerca semplice' and 'Ricerca avanzata' options. A 'In evidenza' section highlights three key updates: 'NORMATTIVA SI RINNOVA CON IMPORTANTI E NUOVE FUNZIONALITA'' (dated 18 DICEMBRE 2023), '"MINT E MASE - MODIFICHE REGOLAMENTI DI ORGANIZZAZIONE"' (dated 18 DICEMBRE 2023), and '"MASAF - REGOLAZIONE RIORGANIZZAZIONE"' (dated 14 DICEMBRE 2023). A chat window on the right introduces 'Ciao! Sono l'assistente virtuale della presidenza del consiglio dei ministri! Posso rispondere a domande sul codice civile e penale, chiedi se hai bisogno'.



Tags and summaries of parliamentary reports

Use of generative AI to facilitate the comprehensibility of *plenary reports*.

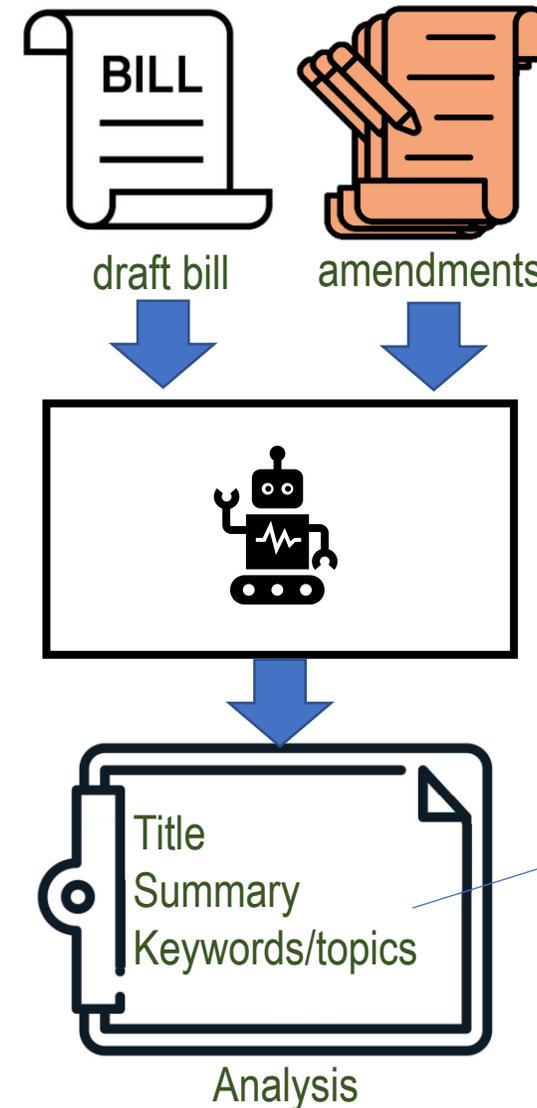
1. OCR the PDF report
2. Perform “**segmentation**”: divide it into micro-sections, based on *topic*
3. **Summarize** the micro-sections.
4. Extract the **keywords** (tags) having a semantic meaning for each section.

Section title	Section tag	Section summary	Original text
<p>Comunicazione del risultato della votazione per l'elezione di due Segretari di Presidenza, ai sensi dell'articolo 5, commi 5 e 6, del Regolamento</p> <p><i>Da pag. 2 a pag. 6</i></p>	<ul style="list-style-type: none">• Presidenza• Votazione• Deputati	<p>Durante la seduta della Camera dei Deputati del 8 novembre 2022, è stato votato per l'elezione di due Segretari di Presidenza in ottemperanza all'articolo 5 del Regolamento. Alla votazione hanno partecipato 348 deputati, con Alessandro Colucci e Benedetto Della Vedova che hanno ricevuto rispettivamente 185 e 119 voti. Sono stati registrati anche 1 voto disperso, 31 schede bianche e 12 schede nulle, con i due deputati proclamati eletti come Segretari di Presidenza.</p>	<p>Comunicazione del risultato della votazione per l'elezione di due Segretari di Presidenza, ai sensi ...</p> <p>... Lupi Maurizio Molteni Nicola Pichetto Fratin Gilberto Polidori Catia Prisco Emanuele Tajani Antonio</p>



Analysis of amendments with genAI

Usage of artificial intelligence tools to understand the normative effect of proposed amendments relying exclusively on official legal sources.



Title: Support for advanced energy storage technologies, overcoming limitations of current batteries.

Summary: Provides support for next-generation energy storage technologies, with a focus on technologies that overcome the disposal problems of electrochemical batteries - Includes support for CO2-based technologies and Energy Transition Combined Cycle (ETCC) facilities.

Topics: energy, green transition, technology innovation, industry

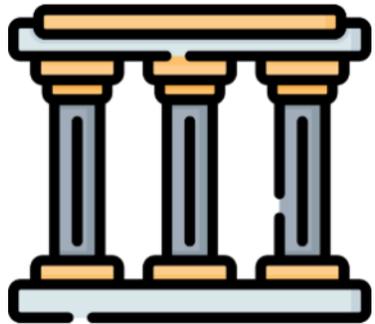


Rules of conduct

Rules of conduct

Rules of conduct for the usage of tools based on Generative Artificial Intelligence at Italian Chamber of Deputies

They were published in April 2024



Pillars:

- fundamental human rights,
- ethics,
- privacy protection



Rules of conduct

Data management:

- Ensure the **confidentiality and integrity** of data
- **Sensitive data** and data covered by **professional secrecy** shall not be used (risk assessment)

IT Service Catalogue:

- IT services using GenAI shall be **listed explicitly** within IT Service Catalogue

Education and awareness:

- Users shall be **informed on GenAI usage** through disclaimers
- Colleagues' education and training on AI
- "*Human in the loop*" approach (support and not replace humans)

Application development:

- "Security by design" approach for solution development process



Next Challenges

GenAI projects awarded in the Chamber of Deputies

(25th July 2024)

In early 2024, the Chamber of deputies promoted a call for ideas to encourage **the participation of academics and researchers** in defining and proposing scenarios for the use of generative AI to serve parliamentary documentation.

Three proposals were selected:

- the first concerns the **analysis of the text** of laws and bills - on historical series too - with respect to **readability and comprehensibility indexes**
- the second includes a set of tools for the structuring of legislative texts and for the **semantic analysis** of bills and amendments
- the third suggests a **conversational interface** for citizens to learn about the activity of Parliament and its Members.



**THE NEXT CHALLENGES → WILL BE
HOW TO PROFITABLY
GATHER THESE CONTRIBUTIONS
AND IMPLEMENT THEM
OBTAINING THE EXPECTED BENEFITS.**





Thanks for your attention.

Claudia Di Andrea

Giovanni Ciccone

Chamber of Deputies of the Italian Parliament

