

Parliamentary proceedings in Italian Senate

Current management & perspectives

Manuela Ruisi

Head of office

Parliamentary information, legislative archive and publications

Library collection development and catalogue

UIP

Parliamentary information, institutional archives and publications Office

- Collects data from parliamentary documents and feeds institution databases (MPs register, draft bills, Official journal, Non-legislative procedures and Documents)
- Collects all parliamentary documentation in electronic format from other depts & from external Printing service and feeds the Institutional digital repository
- Publishes the pdf format of documents & proceedings on the Senate website, making them available on mobile devices
- Manages other publications edited by the Senate (books, ebooks etc.)

Point of view (1)



These activities are at the very end of parliamentary documentation workflow & deeply affected by:

- Technologies used in previous phases (quality & information richness of documents, e.g. metadata)
- Technology available for indexing
- Technology available for publishing

Point of view (2)



- As a curator I'm supposed **to be pragmatic** & strongly involved in everyday management of documents

Point of view (3)

- As a curator (& not being an ICT technician) I'm allowed to prospect new goals and **to imagine** new solutions asking the specialists to translate theories into a technological tool



This strong technology demand is aimed to achieve these goals:

- Avoid duplication in parliamentary documentation processing activities
- Saving human resources
- Saving money by fostering paperless documents circulation
- Saving money by doing most things inhouse

Goals to achieve: management

This strong technology demand is aimed to achieve these goals:

- transparency and intelligibility of documents (contents & layout)
- Searchability (access keys; contents display)
- Syndetic structure corpora allowing rich web surfing experience
- Information self service

Transparency doesn't mean simply to put everything on the web, but attains to how you offer these contents to citizens.

Goals to achieve: citizens

- Digital repository
- Parliamentary activity databases
 - Draft bills
 - TESEO
- Normattiva: portal of Italian legislation
- Heritage: Parliamentary documents of Italian Kingdom Senate (1848-1946)

Focusing on ...

- Managed by Alfresco, open source CMS, customized
- Ingestion of PDF documents produced inhouse (e.g. minutes of plenary debates) or outhouse (“pre-Printing” Service, e.g. amendments)
- Record creation with indexes based on data on cover page
- Widespread publication on website (all pages of interest)

Critical issues

- Staff **duplicates** data (copy & paste of text; selection of values from menu etc.) introducing **errors**
- **Round coverage needed** to provide immediate indexing & publishing
<http://senato.it/static/bgt/UltimiAtti/index.html>

Future development

- Automatic processing of PDF through metadata embedded in the document cover page, making the digital file self-describing. Metadata could come from automatic textual analysis of cover page powered as pdf is pushed to repository.

Digital repository

The screenshot shows a web application interface for legislative documents. On the left is a document cover page for a bill (Disegno di Legge) from the XVII Legislature, N. 2602, concerning asbestos. On the right is a metadata form titled 'Disegno di Legge' with various fields. Red arrows indicate the flow of data from the document cover page to the form fields:

- From 'Senato della Repubblica' to 'Periodo' (Repubblica).
- From 'XVII LEGISLATURA' to 'Legislatura' (17).
- From 'N. 2602' to 'N.° atto/fase' (2602) and 'N.° testo' (2602).
- From 'DISEGNO DI LEGGE' to 'Tipo' (Presentato).
- From 'd'iniziativa dei senatori FABBRI, AIELLO, BAROZZINO, BERGER, BILARDI, BORIOLI, CARDIELLO, COLLINA, D'ADDA, FASIOLO, FAVERO, LANGELLA, MUNERATO, PARENTE, PELINO, PICCINELLI, ROMANO, SERAFINI e SILVESTRO' to 'Iniziativa'.
- From 'Testo unico della normativa in materia di amianto' to 'Oggetto'.

The metadata form includes fields for: Stampato, Riservato Intranet, Validato, Link S.I., pdf corrotto, in calendario, ODG, Annullato, Periodo, N.° atto/fase, N.° testo, Subit., Id testo, id atto, Oggetto, Data ddl, Titolo, and Iniziativa.

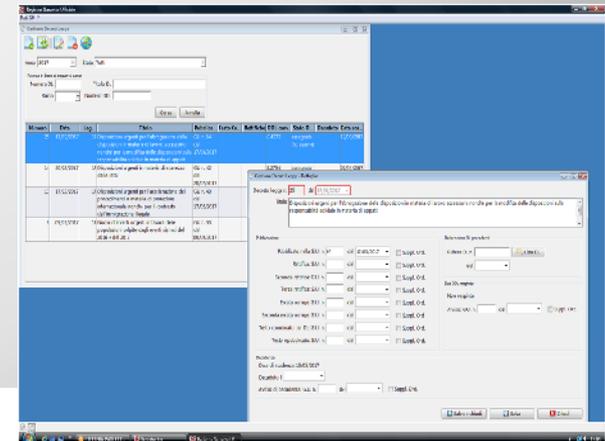
Data are manually copied in index fields (duplication, errors, delay)
Solution requested: Automatic ingestion of metadata from document cover page

Digital repository

- Managed by Java based applications developed inhouse
- Draft bills database covers both the Senate & the Chamber of Deputies
- All databases are interconnected (Composition of the house, Draft bills, Official journal, Non-legislative procedures and documents)
- Updated from 0 to 24 hours after the session of the plenary/committee

Critical issues

- Databases feeded mostly through human reading and analysis of reports (Plenary minutes and Commitees proceedings)
- Limited use of metadata for automatic indexing



Parliamentary databases

Registro Iter

Elenco Resoconti

Seduta (Antimeridiana) n.763 del 21/03/2017

- Missioni.
- Preavviso di votazioni elettroniche (ore 11.04)
- Sequ岸ito della discussione della proposta di legge: Turco: Modifiche alle disposizioni
- PRESIDENTE
- (Esame degli articoli – A.C. 1063-A)
- (Esame dell'articolo 1 – A.C. 1063-A)
- (Esame dell'articolo 2 – A.C. 1063-A)
- (Esame degli ordini del giorno – A.C. 1063-A)
- (Dichiarazioni di voto finale – A.C. 1063-A)
- (Coordinamento formale – A.C. 1063-A)
- (Votazione finale ed approvazione – A.C. 1063-A)

- Speeches loaded automatically from xml proceeding
- Explanation of vote implemented manually

Lista interventi

Politico	Qualifica
Sereni Marina	Come presidente dell'assemblea
Turco Tancredi	Contrario (a nome della componente)
Chiarelli Gianfranco Giovanni	Astenuto (a nome della componente)
Nastri Gaetano	Contrario a nome del gruppo
Farina Daniele	Contrario a nome del gruppo

- Future development**
- All information intended to be indexed should be tagged properly from the origin
 - Automatic textual analysis of declarations of vote to get contrary, abstained, favorable statement

Parliamentary databases

- In the Draft legislation database each bill is indexed by **subject matter** using TESEO (SEnate ThEsaurus for the Organization of parliamentary records) at least twice

- 1) on the original text submitted to one of the Houses
- 2) on the amended text forwarded to the other house for a second reading
- 3) and so on until the final text is approved

TESEO

**The myth of
Theseus and the
Minotaur**



ARTICOLO 7.

(Ulteriori strumenti e obiettivi per l'attuazione di iniziative congiunte).

1. Nell'ambito degli accordi di cui all'articolo 3 e dei patti di cui all'articolo 5, possono essere individuati specifici obiettivi per l'incremento dei servizi di controllo del territorio e per la sua valorizzazione. Alla realizzazione degli obiettivi di cui al primo periodo possono concorrere, sotto il profilo del sostegno strumentale, finanziario e logistico, ai sensi dell'articolo 6-bis del decreto-legge 14 agosto 2013, n. 93, convertito, con modificazioni, dalla legge 15 ottobre 2013, n. 119, enti pubblici, anche non economici, e soggetti privati.

2. Nei casi di cui al comma 1, si applicano le disposizioni di cui all'articolo 1, comma 439, della legge 27 dicembre 2006, n. 296.

In this example, the first reading of Government bill n. 4310, on urgent security provisions for cities, takes place in the Chamber of Deputies.

- **The first step** of the indexing process is the search for precedents of bills on the same subject.
- **The second step** is the text analysis (by human reading). If it contains references to laws or other legal texts, it is necessary to search for the text of the referenced provision, in order to reach its textual meaning.

- If the text being analyzed is in html format, it is possible to employ a tool that parsers the text and shows automatically such references as links to the relevant electronic archive

LEGGI D'ITALIA P.A.

Ricerca ▶ L. 27/12/2006, n. 296 **sommario** Disposizioni per la ... ▶ L. 27/12/2006, n. 296 Articolo 1, comma 439 - Misure ...

Torna ai risultati | ☰ Sfoglia Risultati | ◀ 📄 Sfoglia Documento ▶ | 📄 Vedi intero documento

▼ ⒶⒶⒶ 📄 Salva 🖨️ Stampa

L. 27/12/2006, n. 296
Disposizioni per la formazione del bilancio annuale e pluriennale dello Stato (legge finanziaria 2007).
Publicata nella Gazz. Uff. 27 dicembre 2006, n. 299, S.O.

1.439. Per la realizzazione di programmi straordinari di incremento dei servizi di polizia, di soccorso tecnico urgente e per la sicurezza dei cittadini, il Ministro dell'interno e, per sua delega, i prefetti, possono stipulare convenzioni con le regioni e gli enti locali che prevedano la contribuzione logistica, strumentale o finanziaria delle stesse regioni e degli enti locali. Per le contribuzioni del presente comma non si applica l' **articolo 1, comma 46, della legge 23 dicembre 2005, n. 266.**

TESEO: references to laws

- Next step is the actual description of the contents of the bill, using the Teseo descriptors.
- By browsing and selecting the appropriate terms on the lists we link those terms to the text fact sheet
- The bills indexing is performed on two levels: a limited number of more generic terms is used to describe the bill as a whole (level DDL Teseo); the second level describes the contents of each article (level ARTICOLI Teseo).



TESEO

Registro Progetti di Legge Camera

File Visualizza Nuovo DDL Modifica impianto Gestione DDL Sedute Finestra Aiuto

Registro Progetti di Legge

Ordinamento: per numero Ascendente

Vai a numero: 4310

Classificazione TESEO

Generalità del DDL

Numero: 4310 Conversione in legge, con modificazioni, del decreto-legge 20 febbraio 2017, n. 17 recante disposizioni urgenti in materia di sicurezza delle città

Livello DDL: CENTRI URBANI ORDINE PUBBLICO

Livello articoli

Sposta Classificazione provvisoria

Descrittore	Nota articoli
PROGRAMMI E PIANI	2, 3, 5
MINISTERO DELL' INTERNO	2, 3, 5
REGIONI	2, 3, 5
PROVINCE AUTONOME	2, 3, 5
PROGRAMMI E PIANI	2, 3, 5

Conferma Annulla

Carlo (Gentiloni Silveri-I) (17-02-2017) Question

TESEO

The image shows the terms that have been selected for our bill. Two of them (in this case Urban centres and Public Order) have been moved to the top of the screen to describe the subject of the entire bill. In the bottom half are the terms that describe the content of each article, with the number of the articles involved and sometimes, for more complex texts, even of single paragraphs.

- If the bill is forwarded to the Senate again, the "old" descriptors are automatically uploaded to the new reading of the bill **but the text analysis is repeated, taking into account the amendments approved**
- The result of indexing in the bill fact sheet:
<http://www.senato.it/leg/17/BGT/Schede/Ddliter/47841.htm>

- **Critical issues:**

- This activity is extremely valuable, but highly costly in terms of human resources
- considering the amount of documentation to be treated it can easily become unaffordable

- **Future development:**

- Automatic textual analysis of draft legislation by a tool that
 - reads & comprehends the special language of legislation (is it natural?)
 - Learns from existing knowledge base, from human teaching and from machine learning
- convergence towards a more shared Thesaurus like EuroVoc (mapping TESEO-EuroVoc)

TESEO



CAMERA VOC



Codice	
2509	politica migratoria
2517	politica sociale
2831	tutela dei soci
3320	migrante
3462	accordo bilaterale
3465	negoziato internazionale
3472	convenzione europea
3526	cittadino della Comunita'
4050	sicurezza sociale
4562	lavoratore migrante

Pertinenza
Sufficiente
Ottima
Sufficiente
Mediocre
Buona
Sufficiente
Sufficiente
Mediocre
Ottima
Sufficiente

Atto Camera
Mozione 1-00445
presentato da
FITZGERALD NISSOLI Fucsla
testo di
Giovedì 19 marzo 2015, seduta n. 395

Interpellations,
questions and
motions are
closer to natural
language

Automatic assignation of semantic descriptors to questions, motions & interrogations (moved from TESEO to EuroVoc)

Cogito at the Chamber of Deputies

- **Critical issues**

- descriptors are relevant but not always referring to the very central topics
- human supervision & revision needed
- Strong human need for training the system
- 70% of documentation on provisional classification with disclaimer

- **Future development**

- Mapping EuroVoc-TESEO to uniform the database (TESEO up to the early 2000)
- train the system with decades of earlier classification

Cogito at the Chamber of Deputies

- In a few months all legislation from 1861 on
- Each law preserved in its original text and in all following versions



Critical issues:

- No subject matter research available

Future developments:

- TESEO descriptors in last phase of parliamentary discussion of draft bill applied to the published law (available from 1991)
- Automatic semantic indexing of pre-1991 laws through a system trained by TESEO Knowledgebase

Normattiva.

The portal to Italian legislation

- 130.000 pages minutes of debates
- 180.000 pages draft bills & documents

SENATO DEL REGNO — SESSIONE DEL 1848

Il nostro concorso sarà sempre spontaneo e caloroso per conservare alla monarchia, in ogni qualunque evento, questa importante e nobilissima sua provincia.

13. La Liguria, che scende in campo colla storica sua valentia e col generoso slancio de'suoi magnati per la causa italiana, strinse da prima la destra ai confratelli suoi politici; e plaudita e gratulante nello svolgersi propizio delle nostre venture, dopo aver posto già in comune i molti interessi che a noi l'univano, mette in comune gli affetti, le simpatie, le fraterne sorti inseparabili.

14. Il Senato è lieto della concorde volontà che a noi unisce le potenze governate da istituzioni alle nostre uni-

resche fazioni che furono sempre gloria immanchevole dell'illustre vostra prosapia.

22. Ritorni a voi il glorioso padre vostro. Ritorni a noi il sovrano amato, il legislatore saggio, l'intrepido guerriero, padre pure a noi tutti. Ritorni col trionfo, con le acclamazioni dell'intera patria, con l'ammirazione dell'Europa, con la devozione e la gratitudine degli antichi e dei novelli suoi fedeli, colla rivendicata indipendenza italiana. (*Gazz. Piem.*)

IL PRESIDENTE propone la discussione in generale sul progetto di risposta al discorso della Corona. (*Conc.*)

DE CARDENAS. Mentre sono disposto ad approvare ed appoggiare in massima questo indirizzo che la benemerita

**Historical heritage
Italian Kingdom Senate
parliamentary records (1848-1946)**

Critical issues

- Amount of pages vs human resources available
- OCR full of errors due to ancient typeface
- Detailed marking requested for better enhancement of historical content

Future solutions

- Trained optical character recognition
- Linguistic analysis & automatic revision of OCR
- Automatic entity recognition & extraction (names, dates etc.) & indexing based on controlled lists.

Historical heritage

Thank you

Questions to

manuela.ruisi@senato.it

www.senato.it

www.senato.it/sitostorico

<http://dati.senato.it>

[https://github.com/SenatoDellaRepubblica/AkomaNtosoBulk
Data](https://github.com/SenatoDellaRepubblica/AkomaNtosoBulkData)

www.akomantoso.org