Inter-Parliamentary Union
For democracy. For everyone.

THE NATIONAL ASSEMBLY OF
THE REPUBLIC OF KOREA

World e-Parliament
Conference 2014

8 - 10 May 2014 // National Assembly of the Republic of Korea // Seoul

Open data, documents and related policies @ Senato.it

Senate of Italy

Carlo Marchetti, PhD

Head of Information Systems Development Office
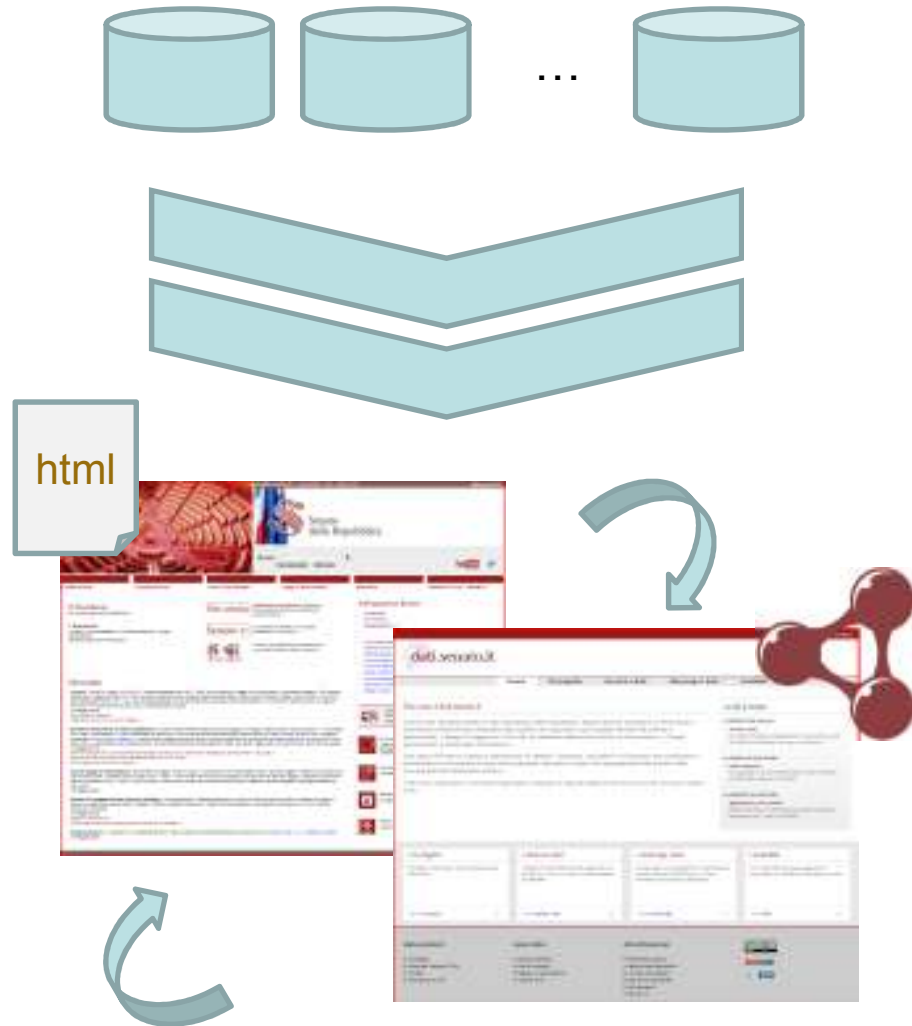
Figures on IT Dpt.

- **IT Department Staff**
  - ✓ 1 Director
  - ✓ 5 Officials
  - ✓ 4 Information specialists
  - ✓ 18 Senior IT Analysts
  - ✓ 12 IT Operators
  - ✓ 10 non-IT staff

- **Selective outsourcing:**
  - ✓ Infrastructure management
  - ✓ On-site technical support and assistance
  - ✓ Help desk
  - ✓ Application maintenance

- **Internal users**
  - ✓ 3.000 managed users
  - ✓ 2.400 PCs
  - ✓ 4.500 mailboxes

- **Infrastructure**
  - ✓ 2 Data Centers
    - ✓ DR configuration
  - ✓ 170 (Virtual) Servers
  - ✓ 40Tb Storage
  - ✓ 5000 Network sockets

- **Applications**
  - ✓ more than 500 custom apps
  - ✓ Website of 500.000 pages
  - ✓ 300.000 pdf online

Open data in Italy

- demand, activities and hype since 2010

- first portals: Q1 2011 – see dati.gov.it

- Chamber of Deputies published a comprehensive open data endpoint and website in Q1 2012 based on W3C LOD standards

- Senate followed in Q1 2013, in tight coordination

- meanwhile, development of legislative and regulatory framework (Code for Digital Administration - CAD)
  - set goals and obligations for all Agencies
  - A specific Agency "Agenzia per l'Italia Digitale" is in charge of emitting and maintaining guidelines

Open data @Senato.it

- IT Dept. started activities 1Q 2012
- Focus on W3C standards
  - Linked Open Data and Semantic Web

- Idea: "LODify" website information

- Results:
  - triple store on parliamentary activity
  - open data website
  - Akoma Ntoso XML documents

html

dati.senato.it/sparql : LOD endpoint

- contains all data on parliamentary activity since 1996:
  - Senators, Parties, Political Groups, Committees, Bills, Docs, Plenary Votes, …
  - 34MLNs triples updated every night
- strict update SLAs (same as website)
- publicly available to everyone with sufficient programming skills

Virtuoso SPARQL Query Editor

About | Namespace Prefixes | Inference rules

Default Data Set Name (Graph IRI)

Query Text

`select distinct ?Concept where {[] a ?Concept} LIMIT 100`

based on Virtuoso open source sw

(Security restrictions of this server do not allow you to retrieve remote RDF data, see details)
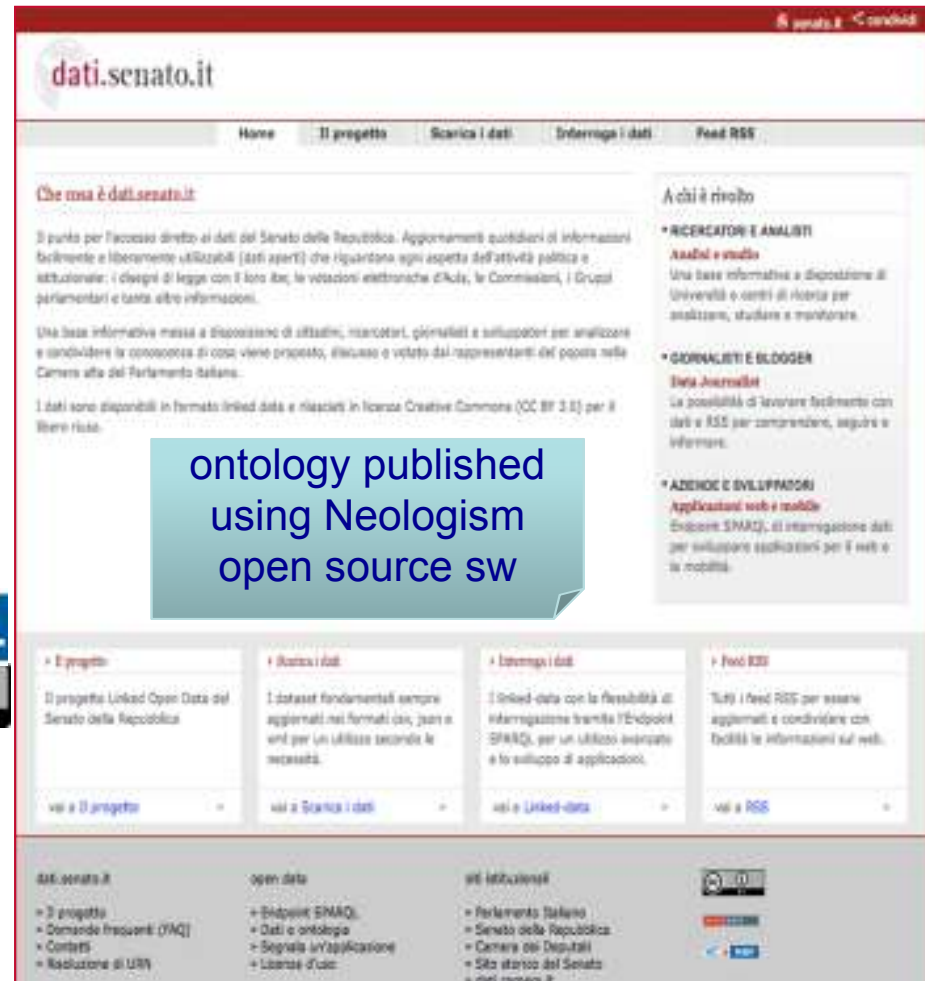
Results Format: HTML

Execution timeout: 0   milliseconds (values less than 1000 are ignored)

Options: ☑ Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see details)

Run Query   Reset

dati.senato.it website

- triple store is pretty unfriendly

- website lays on the TS and eases the burden of
  - understanding data
  - downloading data

- relevant contents:
  - OWL ontology
  - CC-BY licence
  - ready-made parametric datasets with easy interfaces

ontology published using Neologism open source sw
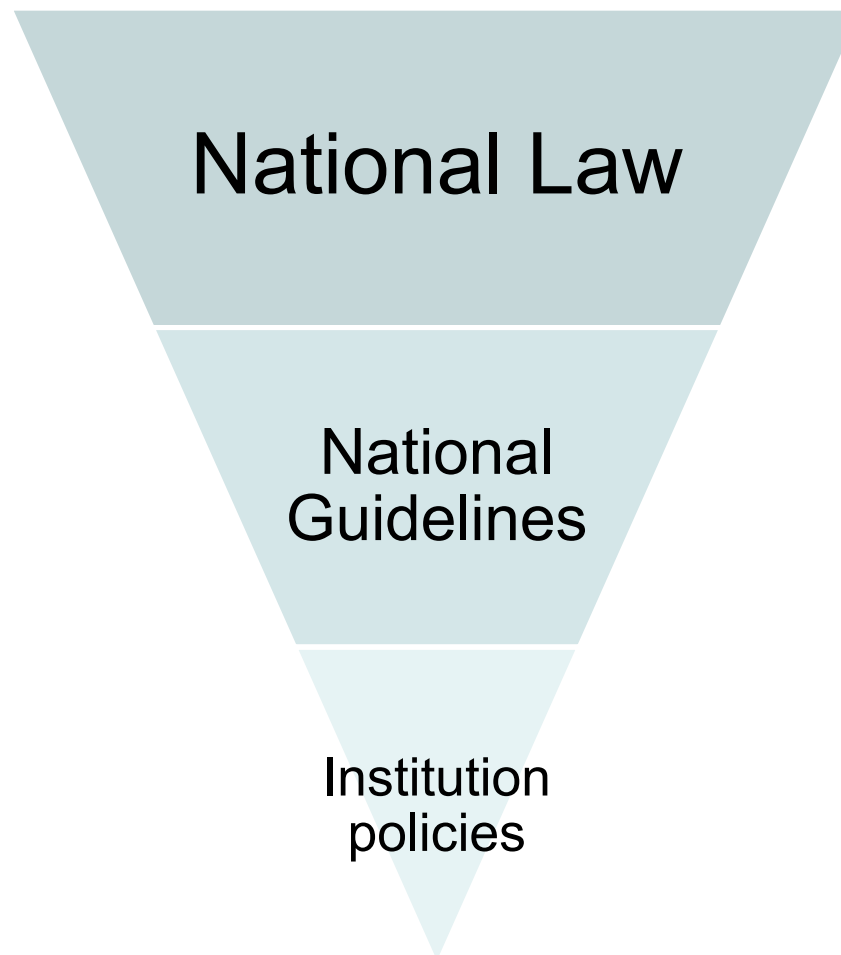
Akoma Ntoso XML docs

- most data are based on/refer to acts

- acts contain lots of data
  - MPs promoting a bill
  - references to other acts

- machine readability of documents must be taken into account along with open data

- adoption of Akoma Ntoso for bills and dossier
  - next: amendments and reports

Policies and guidelines on open data

- in force legislation with mandatory obligations for all public administration (ministries, agencies, etc.)

- a specific "Agency for Digital Italy" has to coordinate technical activities and decisions, and emits standards and guidelines

- Implementation-specific policies: licences, update SLAs, ontologies etc.

National Law

National Guidelines

Institution policies

Legislative prescriptions

- The Digital Administration Code Act (CAD in Italy) – 2005 - has been heavily integrated in 2012 for promoting open data among Agencies
- Main prescriptions (art. 52):
  - all Agencies must publish on their website a catalogue of managed DBs and access rules (except for tributary and sensitive data)
  - in each tender requiring data management activities (gathering, handling etc), opening data becomes mandatory
  - Government must yearly refer to the Parliament on the actuation of public policies on open data
  - The Agency for Digital Italy must publish and yearly review:
    - national standards, methodologies and rules for the publication of Open Data from all Agencies
    - report on advances made by agencies

Technical Guidelines

- ample set of specifications resulting from a cooperative and collaborative managed process involving all public Agencies

- focus: semantic interoperability of data published from different sources to avoid "information silos"

- consistent set of recommendations "RFC style", e.g. :
  - ✓ adopt – whenever possible – W3C standard and LOD publishing
  - ✓ publish data along with their metadata/ontologies
  - ✓ use a website to describe data, metadata and user license
  - ✓ reuse existing ontologies (e.g. foaf)
  - ✓ link data to dbpedia whenever possible

Agenzia per l'Italia Digitale
Presidenza del Consiglio dei Ministri

Senate policy

- Adhere and contribute to National Law and Guidelines
- View: open data new frontier for transparency, civic media feeders
- President Pietro Grasso statement for International Open Data Day – February the 21st, 2014

"The term ' transparency ' acquires new meanings today. Being present on the Web with a continuously updated and enriched site is just the first step, necessary but not sufficient. Technological progress poses new challenges that Institutions must cope with, by opening their archives and providing information in standard formats and according to international protocols that foster interchange, reprocessing and reuse of data. This is what the Senate is doing, with the site dati.senato.it that adds a new and important step to a range of information services already very used and valued - based on databases and search engines of the site www.senato.it that every day, for almost 20 years, answered the questions of a growing number of users."

Lessons learned

- LOD are more difficult than OD datasets, but have better ROI
  - building a triple store pays in terms of freedom of dataset extraction and user autonomy
  - PMOs and companies increasingly using (e.g., openpolis and votali#)

- Open Data is a new service: need for support and communication activities

- Ontologies and Knowledge management are central elements for communication&support

- Policies and technical guidelines are live objects: obsolescence is very fast

- Documents play a central role but are hard to be managed as "open": more work is needed in this direction

Future horizons

- Senate level
  - publish open data on leg. I-XII
  - publish amendments and reports in Akoma Ntoso

- National level
  - search engine on regional legislation based on LOD
  - Dossier on OD in Italy together with Chamber of Deputies
  - Extend guidelines on open documents

- International level
  - we're taking different ways, approaches, to represent similar data: MPs, bills, amendments
  - opportunity for inter-parliamentary cooperation: agreeing on basic datasets and ontologies

References

- Summary of the Guidelines (in English):
    - http://www.agid.gov.it/sites/default/files/documentazione_trasparenza/semanticinteroperabilitylod_en_3.pdf

- Legislation (in Italian):
    - Codice dell'amministrazione digitale
        - http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2005-03-07;82

    - Open Data by Default (in Italian)
        - http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2012-12-17;221

THANK YOU