

Mining for freedom in oeil, eur-lex and a few others

Stefan Marsiske, <stefan.marsiske@gmail.com>

November 14, 2011

Background/context

- ▶ 15+ years Free software developer and advocate
- ▶ Free software is ethic-by-design.
- ▶ free software -> copyright, software patents, internet, privacy, free speech, open standards and open data
- ▶ support NGOs, deeper involvement -> realization of a few needs and opportunities.

Pippi Longstrings

- ▶ directly linkable eurlex texts
- ▶ commenting on these texts
- ▶ finding shared textblocks between documents Examples:
- ▶ EIF2 co-authored by BSA¹
- ▶ the CARIFORUM discovery (IPRED1 art 8 cypaste) recognised by Prof. Anette Kur²

¹<http://fsfe.org/projects/os/eifv2.en.html>

²<http://www.ip-watch.org/weblog/2009/12/21/acta-may-prompt-quick-restart-to-eu-harmonisation-of-criminal-enforcement-of-ip/>

Open Definition

The Open Definition sets out principles to define 'openness' in relation to content and data – that's any kind of material or data 'from sonnets to statistics, genes to geodata'. The definition can be summed up in the statement that “A piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and share-alike.”.

Parltrack

<http://parltrack.euwiki.org> is a free tool and the associated free database to track the law-making in the European Parliament.

- ▶ aggregate & publish
- ▶ notifications (email & RSS),
- ▶ analysis & commenting.
- ▶ lot's of material for datamining (e.g. deadlines for tabling amendments)

Future

- ▶ improve pippi to capture not only textual but also semantic relationships
- ▶ private commenting on parltrack,
- ▶ extending the scope to the EC and the Council

Combined information

Parltrack collects the data daily and combines the information on

- ▶ MEPs³,
- ▶ progress on dossiers from OEIL⁴,
- ▶ committee agendas⁵,
- ▶ vote results from the plenary minutes⁶.

³<http://www.europarl.europa.eu/members/public/geoSearch.do?language=en>

⁴<http://www.europarl.europa.eu/oeil/>

⁵<http://www.europarl.europa.eu/activities/committees/committeesList.do?language=en>

⁶<http://www.europarl.europa.eu/activities/plenary/pv/calendar.do?language=EN&L>

Examples. . .

- ▶ which MEP has taken responsibility for which dossiers,
- ▶ which committee is responsible or gives opinion on which dossiers,
- ▶ forecasts for next steps on dossiers in the EP *and* committees - including tabling deadlines for amendments,
- ▶ online detailed committee agendas,
- ▶ online calendar with EP *and* committee dates,
- ▶ who are the influential MEPs related to a dossier. There's much more possibilities in this data than the web-interface and your humble developers can ever present to you. Though we are willing to work for hire to dig deeper into this information-goldmine. ;)

Notifications

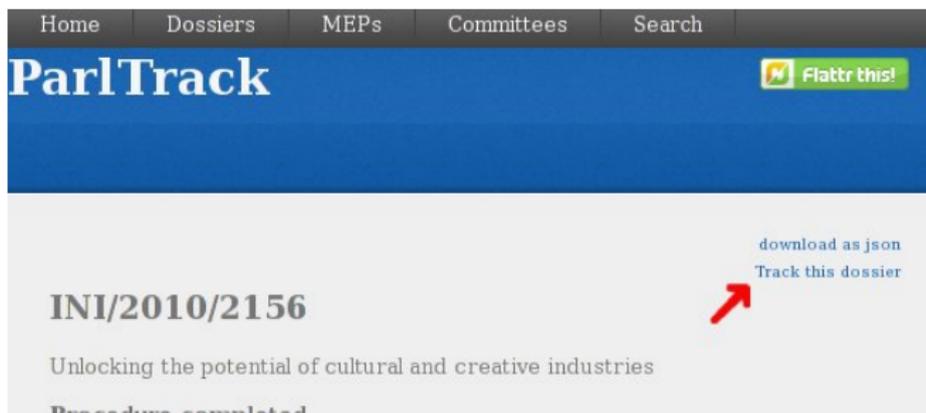


Figure: how to subscribe

Anyone can subscribe to receive email notifications on changes on dossiers. On every dossier page, there's a link "*Track this Dossier*", by providing your email here you create a new *notification group*. If you don't specify a name of the *notification group*, one will be randomly generated for you. If you specify an existing *notification group* name, the current dossier will be added to that. After you created a notification group, you can share the link to this so other

The Texts

Most of the documents created in the law-making process (initial, supplementary and final texts, opinions) are linked whenever available from OEIL. Most of these documents are also available on Eur-Lex⁷ (the official EU website publishing the texts) and thus are also automatically available in Pippi Longstrings from the *Documents* tab for dossiers. Using Pippi Longstrings you can analyse the texts⁸ and comment on them. Commenting happens by selecting the text you want to comment on. The results of pippi-analysis are also displayed as comments, from a user named Pippi Longstrings.

⁷<http://eur-lex.europa.eu/>

⁸http://www.ctrlc.hu/stef/blog/posts/pippi_longstrings.html 

The Data

All data in Parltrack is available freely in various forms for anyone to use. Most of the pages have a “*Download as JSON*” link, all new and changed dossiers are listed in an RSS feed, changes in dossiers are also available as email notifications. Furthermore a complete DB dump⁹ as JSON is also available. All this is under the ODBLv1.0¹⁰ license. Running your own instance of Parltrack is encouraged, the code¹¹ is free under the AGPLv3+¹².

⁹<http://parltrack.euwiki.org/dumps>

¹⁰<http://www.opendatacommons.org/licenses/odbl/>

¹¹<https://github.com/pudo/parltrack/>

¹²<http://www.gnu.org/licenses/agpl.html>

Searching

Currently search is possible for text in MEPs names, dossiers titles and IDs of final acts like “directive 2006/24/EC”.

- ▶ the EP should keep improving the quality of the data it makes available to the public, in particular with regards to **time and date accuracy**
- ▶ the EP should continue to ensure that data is made available in formats and through interfaces which enable free and easy access and processing of the data
- ▶ the EP should avoid investing in software and services which leads to “lock in” of citizens and risk the institution’s vendor independence, also with respect to all media
- ▶ the EP should develop strategies for making its data available for all to use and reuse as per existing best practises for opening up of public sector data

Credit

Parltrack is inspired by ehjs Tratten¹³, was conceived during the 2010 FCForum Tools for policy reformers¹⁴ workshop and is influenced by a bunch of good people, thanks!

¹³<http://euwiki.org/Tratten>

¹⁴<http://2010.fcforum.net/topics#poli>