

GIPO AG meeting #4: AGENDA

Date: 23 December 2015

Time: 15:00 UTC

Location: Skype

Invited: Andrea Calderaro (AC), Avri Doria (AD), Amanda Lawrence (AL), Bill Drake (BD), Daniel Dardailler (DD), Ginger Paque (GP), Lasantha De Alwis (LDA), Tom van Engers (TvE);

Present:

Secretariat / Chair: Jamal Shahin (JS)

Agenda

1. Approval of new AG members
2. AG presence in GIPO
3. GIPO tool roadmap
4. Ongoing issues:
 - a. Taxonomy (<http://bit.ly/gipotaxo>)
 - b. Multilingualism
5. Date of next meeting
 - a. TBD in early January?

Annex: GIPO Tool Roadmap 2016

Progress over the next year

By Summer 2016, the GIPO tool will have the following facilities:

Back office

- be able to analyse PDF, DOC and other non-textual feeds;
- categorisation and ontology for data linking and extraction
- resolving issues of multilinguality (see below)

Front office

- users will be able to create their own dashboard (select sources of data, themes, etc.)
- users will be able to create their own data visualisations
- the tool will provide a REST API, an OPML file and RSS output

By the end of 2016, the final tool should be launched. This version will see the addition of an 'assistant' to help users with the discovery of new sources.

Annex: Multilinguality

Note from Luis:

We are currently exploring an approach that relies on language substitution using Wikidata labels. And what is "Wikidata labels"? As you may know, Wikidata is the free knowledge base that anyone can edit. The Wikidata repository consists mainly of items, each outlined by a label, a description and likely one or more aliases.

In Wikidata, items are used to represent all the things in human knowledge, including topics, concepts, and objects. For items and properties, the labels, descriptions and aliases are all multilingual, meaning they can be displayed to users, and can be entered by users, in every language supported by the software (more than 300).

In Wikidata the label is the most common name that the item would be known by. Note that although an item will have multiple labels in different languages, there aren't labels for all items in all languages yet. But users can improve the content in Wikidata and make more knowledge accessible to

everyone. Labels, aliases, and descriptions in multiple languages can be easily added to items and properties.

So in GIPO tool we plan to avoid the need for intelligent translation algorithms in order to translate the metadata (or any other text) in the tool: it is enough to look up (using the MediaWiki API) the label for the specific items in the language desired and display them in place.

I think we can debate about this approach to see what the AG has to say...

Follow up note from Luis:

We will provide the user interface only in English and Spanish.

All the contents gathered by the tool from the sources are shown in their original language. We will not make any translation (nor human nor machine-driven).

Our proposal for language substitution (the "Wikidata approach") is only for the translation of the metadata (basically tags/keywords and terms of the taxonomies, that are associated to every content item). As these metadata are generated by the tool in English, all these entities (aspects, regions, tags, etc.) are the only ones that the tool will try to translate to at least all the rest of official languages of the United Nations as requested by the EC (Arabic, Chinese, French, Russian, Spanish). If any given entity does not have a corresponding label in Wikidata for any of the languages mentioned we will not translate it, but any user can update Wikidata/Wikipedia and add the missing translation for the term.

GIPO AG meeting #4: ACTION POINTS

Date: 23 December 2015

Time: 15:00 UTC

Location: Skype

Present: Andrea Calderaro (AC) (met with JS on 21/12), Daniel Dardailler (DD), Avri Doria (AD), Bill Drake (BD), Ginger Paque (GP).

Absent: Lasantha De Alwis (LDA), Amanda Lawrence (AL), Kasia Jakimowicz (KJ), Tom van Engers (TvE)

Secretariat / Chair: Jamal Shahin (JS)

Action Points

1. Approval of the four new members of the AG
JS to inform new members
2. AG expressed interest in physically meeting in June 2016 (EuroDIG)
JS to enquire with European Commission as to possibilities
3. AG agreed to continue discussion with technical team on the roadmap for GIPO tool, this will be carried out by email and in subsequent webinars.