Customer Success Story

REEEP

REEEP is an international NGO that supports and funds renewable energy projects worldwide. In 2011, the service-oriented organisation started to develop a semantic product that is known as Climate Tagger. Organizations providing climate related information can implement the Climate Tagger as a highly precise content classification service and enrich their digital assets with consistent metadata. Affiliate platforms can also exchange data through the Climate Tagger and mesh it with their own content.

The challenge

The renewable energy and climate change sector consists of heterogeneous stakeholder groups that have a common body of knowledge. Synergy effects and the exchange of existing information is essential in this sector. However, the used vocabulary strongly varies among the political and industrial players. The inconsistent information landscape also prohibits profound decision-making. The automatized reuse of data across organisational boundaries is limited as well. In the coming years, political and economic decisions will take climate information increasingly into account. Therefore, the demand for a 360 degree view on climate issues is strongly increasing.
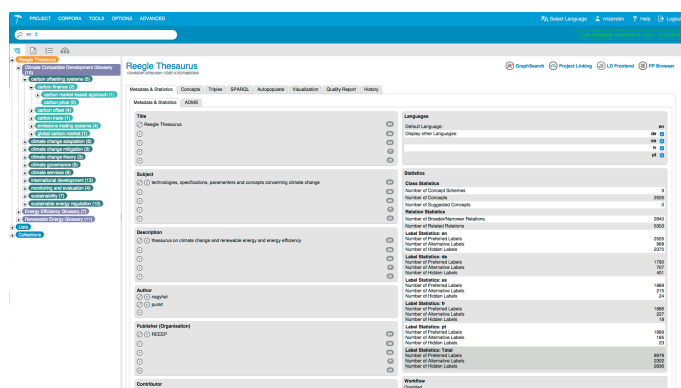
The solution

The Climate Tagger is an application that is built on top of a comprehensive taxonomy. This knowledge graph was developed in a collaborative process with over 30 international experts specializing in renewable energy policies, climate change adaptation and mitigation finance. The taxonomy structures the knowledge domain and takes synonyms and other related information into account. The common vocabulary is available in five languages. Organisations can analyze their content through the Climate Tagger and annotate their digital assets with metadata through the underlying taxonomy and text-mining capabilities. For example, when a user searches for solar plants, he/she can also get results for photovoltaic, which is synonymous with solar plant. Out-of-the box integrations are available for CKAN, Drupal and Wordpress.

The results

In 2015, already 200 interest groups were using the Climate Tagger to annotate their content according to an industry-standard vocabulary and more than 700k documents were analyzed and tagged. The wide-spread industry usage of the Climate Tagger ensures a consistent content annotation that enables a more convenient content reuse and linking. Organizations can decide to make their website externally accessible. Then the possible consolidation of knowledge bears significant innovation potential. The content gets displayed on other platforms as well, when search queries fit the annotated content. Organizations don't need to reinvent the wheel, but can focus on their area of expertise and bundle their knowledge with complementary resources.

Screen-Shot: PoolParty Taxonomy on Renewable Energy

Screen-Shot: CTCN News & Media , Content Classification by the Climate Tagger

Project insights

How to develop a common vocabulary with text-mining

STEP 1: EXTENSIVE KNOWLEDGE POOLING

First of all, the REEEP chief taxonomist had to develop a fundamental understanding of the domain. She asked experts to send her representative documents, which got analyzed by the PoolParty text-mining module. She generated a key word list and built a hierarchical taxonomy structure out of that. This taxonomy draft was the starting point for expert discussions.

STEP 2: REFINING THE TAXONOMY

The taxonomist conducted qualitative expert interviews. Then she started to extend the knowledge graph on a vertical level. The growth of the knowledge model led to learnings and guidelines for the taxonomy development:

✓ When building the taxonomy, it is essential to stick to a top-down approach and extend its depth gradually.
✓ The prioritization of the further taxonomy development must consider frequent usage, relevance and complexity of the specific knowledge sub-domains.
✓ In the role of an external taxonomist, it became obvious that many stakeholders are influenced by their internal vocabulary. To derive an industry-relevant knowledge graph, the knowledge engineer has to function as a catalyst and streamline

cross-organisational discussion results into a generic vocabulary.

STEP 3: ONGOING TAXONOMY MANAGEMENT

A taxonomy is not an end in itself, but a foundation for knowedge-intensive applications as search portals, expert finders or recommendation engines. Highly topical subject areas as renewable energy are under rapid change and also influenced by the application usage. Taxonomies are part of a learning system. The taxonomist is regularly checking and fine-tuning the knowledge graph to ensure metadata accuracy and relevance:

✓ Via text-mining, a regular gap analysis is executed. New knowledge areas get uncovered and can be included in the taxonomy.
✓ The taxonomy is enriched through Linked Data sources. Descriptive definitions for the vocabulary are not generated manually, but are automatically included from existing resources. As the Linked Data ecosystem is growing rapidly, the taxonomist follows up on its development and includes new data.
✓ Also the maintenance of the taxonomy in five languages is supported automatically via PoolParty`s Linked Data module.

TAKE A LOOK

climatetagger

http://www.climatetagger.net

REEEP

http://www.reeep.org

poolparty

http://www.poolparty.biz

REACH OUT TO THE PROJECT TEAM

Do you want to know more? We are happy to connect you with our customer and consultant.

FLORIAN BAUER
COO & Director of Open Knowledge
Mail: florian.bauer@reeep.org

DENISE RECHEIS
Thesaurus & Knowledge Specialist
Mail: denise.recheis@reeep.org

MARTIN KALTENBÖCK
Project Management
Mail: m.kaltenboeck@semantic-web.at