Hackathon project
«Reference data for Anime and Manga: Semantic Linking and Publishing of Diverse Data - Sets»

Team lead: Victor Agroskin, vic5784@gmail.com

Ontology Summit 2014
A problem and an answer

- Ontology-based integration of engineering data should be transparent for engineers and to logicians. What does it mean? To understand simultaneously plant models, equipment, P&ID diagrams, and classes, properties and inverse properties, inheritance.
  - Few people know both. Even fewer among them use the same terms for objects they probably share.
  - Engineering datasets come from industrial proprietary CAD systems, ontology tools are open-source university apps.
  - Engineering data is guarded behind corporate firewall, ontology research in public.

- We need non-industrial example of engineering ontology work to test and publicize approaches, method and tools. What can it be?
Anime and manga production as engineering

• Studio and distribution as **Enabling system** and “anime world” as **System-of-interest** - complexly related.

• Derivative works as Product lines and Custom-tuned models: what is our typical system?

• Complex data with high variety:
  • Texts, video, picture datasets – and relational production and fan datasets.
  • Blurred border between structured and unstructured data (lessons from IBM Watson: ontologies help to answer questions but aren’t a primary source of information)
The Datasets
The Plans

A) The Reference Data
- Ontology commitments - ISO 15926 4D extensional ontology
- Reference data reuse - to be a part of federated Reference Data Libraries (RDL).
- Develop domain ontology (classes and templates – n-ary relations) rich enough to map a sensible selection of anime-manga online resources.

B) The Semantization of Datasets
- Develop patterns for anime-manga semantic data structures (bases on classes and templates of domain RDL).
- Program adapters to access selected datasets and map native data model to developed patterns.
- Run adapters and create semantic data from anime-manga sources in ISO 15926-8 RDF.

C) Publish Linked Data
The Event

• On the 29th of March 2014 12:00 – 20:00 MSK
  • Debug and corrections – 3 more days
• Overall 9 participants and some guests
• Working language - Russian, with occasional use of English
• Environment - .15926 Editor (specially tailored version downloadable from http://techinvestlab.ru/files/15alpha/dot15926Editor15alpha.rar)
• Open project repository - http://github.com/ailev/anird (initial data sets, reference data, patterns, adapter code and project results)
The Process

- Data sets to access (APIs and XML exports):
  - http://www.animenewsnetwork.com/
  - http://anidb.net

- Modeling
  - Alternative identifications – names and internal DB IDs
  - Work division – connecting title, team and roles in an activity

- URI schema

- Reference data development

- Set of patterns to describe native DB models

```json
{ 'name': 'ANNTitle',
  'signature': { 'Title': '', 'TitleId': '', 'Year': '' },
  'options': [
    { 'name': 'main',
      'parts': [
        { 'AMTitle': 'self',
          'uri': 'http://anird.techinvestlab.ru/rdl/id83475f51-fcb0-48cf-ab80' },
        { 'ANNMainTitle': 'self',
          'uri': 'http://anird.techinvestlab.ru/rdl/id602498b4-5737-4ee4' },
        { 'ANNID': 'self',
          'uri': 'http://anird.techinvestlab.ru/rdl/idbc8794fd-3091-4ca9-8af' },
        { 'Anime': 'self',
          'TitleId': 'annDataBaseID',
          'type': 'http://rds.posccaesar.org/2008/02/OWL/ISO-15926-2_2003#ClassOfInformationObject',
          'Year': 'annYearAired',
          'Title': 'label',
          { 'type': p7tpl.Classification,
            'Anime': 'hasClassified',
            'AMTitle': 'hasClassifier' },
          { 'type': p7tpl.ClassifiedIdentification,
            'Anime': 'hasObject',
            'Title': 'valIdentifier',
            'ANNMainTitle': 'hasContext' },
          { 'type': p7tpl.ClassifiedIdentification,
            'Anime': 'hasObject',
            'TitleId': 'valIdentifier',
            'ANNID': 'hasContext' },
      ]
  ]
}
```
The Results (1)
The Results (2)
Further plans?

• Not enough time to complete Linked Data presentation of the results
  • Look for Linked Data engine in the new release of .5926 Editor

• Integration of two data sources not complete

• Shortcomings and bugs in .15926 Editor pattern mapping engine discovered
  • Look for enhances mapping engine in the new release

• All are welcome to test our environment - changing, augmenting or completing the project
Thank you!

Team lead: Victor Agroskin, vic5784@gmail.com
TechInvestLab.ru

.15926 Editor freeware: http://techinvestlab.ru/dot15926Editor
(documentation: http://techinvestlab.ru/15926EditorDocumentation)