Zone-project: towards a better news feed using semantic web

Christophe Desclaux[1]

Wimmics Inria, Sophia Antipolis...,
christophe@zouig.org,
WWW home page: http://www.zone-project.org

Abstract. Nowadays we can use RSS feeds, Twitter, Google Reader, Yahoo Pipes or aggregators to keep up with news. Though those solutions do not guarantee data privacy and rather manage news by origin. The zone project proposes an innovative solution to overcome those issues using the power of Semantic Web and group related informations together. ZONE-project provides a new, innovative way to follow news. At its core, the system is aggregating news items from various RSS feeds. Using the power of semantic web we are able to efficiently tag and annotate each news. Those tags are the basis of filters. Filters allow users to see only news that are relevant. For instance users can retrieve all news containing a tag, or on the contrary never see news containing specific tags.

Keywords: linked data, data aggregation, RSS

1 Motivation

A lot of news are published every day on internet, the number of news websites has increased significantly. People and organization are now building news aggregators in order to sort all this information. This systems are really important in order to clean all the amount of information.

Solutions exists, you can for example use *Google news* [1] like a trusted provider of information, but in this web application your are only a consumer and don't have a lot of interaction with the system.

The second solution is to make news forecasting using *Twitter*[2], it's a solution in which you have the hand on your sources. You can clearly choose the users you want to follow. But you can't have a selection on sources and on topics and you have a lot of noise around goods informations.

The last revelant solution that you can use is *Yahoo Pipes* [3], this tool allows the mixing of popular data feeds to create data filtering via a visual editor. It ue pipes as workflows which will help users to sort feeds.

[1] http://news.google.fr
[2] http://www.twitter.com
[3] http://pipes.yahoo.com/pipes/

With the study of this three solutions I have extract some main challenges that good aggregators need to work on.

- **filtering capacities** he need to sort all information according to [critres fins]
- **lot of informations** the tool need to have access to all news present on internet
- **privacy** users need to use the solution independently of any source provider.

Technicals solutions exists in order to solve this challenges, Google solve this problem using [article de Larry Page sur l'aggregation par mots cls] but this solution is not much efficient because it work on words instead of working on meaning. The solution to use is to use aggregation based on semantic web.

We will first present in this article how work our solution called Zone with explications about the annotation workflow, the ontologie used and the use of datamining solutions. The we will present a demonstration of the application and finally conclude and talk about future work.

2 Application: Zone

2.1 the workflow

In order to have a selection more efficient of news according to their semantic relevance, we have create two sort of workflows, according to the following figure 1 which show the general architecture: a semantic annotation workflow of news and a filtering workflow. The distinction between the two workflows is [very very] important in order to work [de manre asynchrone].

The semantic annotation workflow

The filtering workflow

2.2 The ontologie

base sur RSS on a ajout un schma RDFs spcialis on se base aussi sur l'ontologie de wikimeta/insee

2.3 Using datamining

section d'Ameni

3 Demonstration

On va sur l'application, on slctionne les entites nommes qui nous interessent dans la liste des articles prsents, ca gnre ce qu'on a envie de voir... Comment se droule la dmo, captures

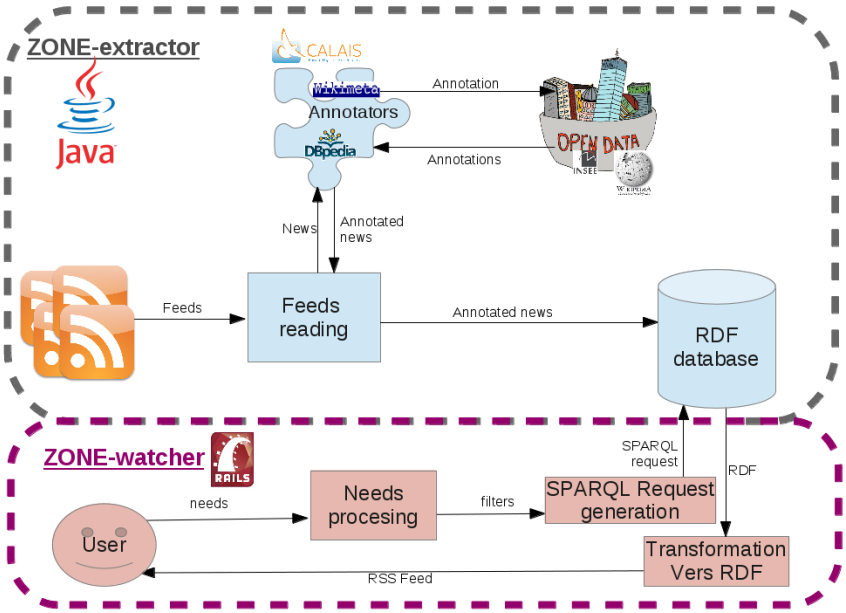

Fig. 1. Workflows d'annotation et de filtrage smantique de nouvelles

4 Conclusion and futur work

Challenges qui restent rsoudre : *tech : faire des liens vers openData et du raisoning *com : trouver moyen de prenniser le projet

5 Acknowledgments

Description du contexte de travail=¿ se passe l'inria BYC... quipe wimmics qui bosse sur semantic web

References

1. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. 78, 315–333 (1982)
2. Clarke, F., Ekeland, I.: Solutions périodiques, du période donnée, des équations hamiltoniennes. Note CRAS Paris 287, 1013–1015 (1978)
3. Michalek, R., Tarantello, G.: Subharmonic solutions with prescribed minimal period for nonautonomous Hamiltonian systems. J. Diff. Eq. 72, 28–55 (1988)
4. Tarantello, G.: Subharmonic solutions for Hamiltonian systems via a $\mathbb{Z}_p$ pseudoindex theory. Annali di Matematica Pura (to appear)

5. Rabinowitz, P.: On subharmonic solutions of a Hamiltonian system. Comm. Pure Appl. Math. 33, 609–633 (1980)