

refer: a Linked Data based Text Annotation and Recommender System for Wordpress

Tabea Tietz¹, Jörg Waitelonis¹, Joscha Jäger², and Harald Sack¹

¹ Hasso-Plattner-Institute, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
`firstname.lastname@hpi.de`,

² yovisto GmbH, August-Bebel-Str. 25-53, 14482 Potsdam, Germany
`firstname@yovisto.com`

1 Introduction

When searching for an arbitrary subject in weblogs or archives, users often don't get the information they are really looking for, because they are overwhelmed with an overflow of information while sometimes the presented information is too scarce to make any use of it. Without further knowledge about the context or background of the intended subject users are easily frustrated because they either cannot handle the amount of information or they might give up because they cannot make sense of the topic at all. Furthermore, authors of online-platforms often deal with the issue to provide useful recommendations of other articles and to motivate the readers to explore more of the available but often hidden content of their blog or archive.

In the demo presentation, we present *refer*, a semantic annotation and visualization system integrated into the Wordpress platform. *refer* enables content creators to (semi-)automatically annotate their texts with DBpedia resources as part of the original writing process and visualize them automatically. With *refer* users are encouraged to take an active part in discovering a platform's information content interactively. They can discover background information as well as relationships among persons, places, events, and anything related to the subject in current focus and are inspired to navigate the previously hidden information on a platform.

Related systems include Pundit [2] and WYSIWYM [1]. The original concept of the presented user interface and a first prototype have already been presented in [4].

2 *refer* Components

refer consists of the following tools and visualizations, integrated into Wordpress (cf. Fig. 1). An **Annotator** is implemented as an extension of the Wordpress text editing interface to create semantic text annotations based on DBpedia. To create annotations, the author selects a text fraction in the Wordpress editing environment. The author can further choose between manual and automated annotation by means of automated Named Entity Linking (NEL). For automated

annotation, *refer* deploys KEA-NEL [6], which implements entity linking with DBpedia entities [5]. Nevertheless, there are still errors that have to be revised manually. The user has to be supported in selecting the correct entity from the large knowledge-base which is a challenge especially for highly ambiguous textual mentions. Further, utilities are applied to rank and organize the candidate lists according to e.g. string similarity with the entity mention, or popularity of the entity [3] to increase the users' annotation performance. The *refer* Annotator provides two configurable user interface modes: (1) The **Modal Annotator** is triggered via a button in the TinyMCE control panel, whereby the suggested entities are displayed in a modal dialog window (cf. Fig. 1.1) in a table-based layout. Entities are divided into the four categories Person (green), Place (blue), Event (yellow) and Thing (purple). The text abstract and entity IRI is displayed on mouseover. The selected annotation is encoded in RDFa markup, which is added to the according text fragment. (2) The **Inline Annotator** (cf. Fig. 1.2) is triggered automatically via highlighting a text fragment and opens a circular category menu which allows the user to instantly show suggestions from the respective category (Person, Place, Event, Thing). By selecting a category, the suggested entities are displayed. In order to provide more context within the relatively small space, these entities are divided into dynamically retrieved sub-categories. While the *Inline Annotator* provides fast and simple means of semantic text annotation without losing the context of the written text, the *Modal Annotator* provides a larger overview and enables a better comparison of entities.

Infoboxes are used to visualize annotations in the article view, and annotated entities are indicated directly in the article text. On mouseover, an infobox as in Fig. 1.3 is shown right below the annotated text fragment, which contains basic information about the entity, e.g. a label, a thumbnail, and additional data in a table-layout. The infobox allows the user to gather basic facts about an entity and relations to other entities. While some basic information can be derived just from the webpage's RDFa microdata and is displayed instantly, additional content is asynchronously loaded from the web service. When the text fragment or any of the infobox entities are clicked, the *Relation Browser* slides down from the top of the page with the selected entity in focus.

The **Relation Browser and Recommender** visualizes relationships between annotations as well as suggestions for further reading. It allows users to navigate and explore relations among entities. It can be opened by the user either via click on the *refer* icon bar on top of the page or by selecting an entity of interest in the article text, (e.g. *Jules Verne*). The entity *Jules Verne* then becomes the focus-entity. Its DBpedia abstract and image are displayed on the bottom of the Relation Browser and its background color depends on its category (Person, Place, Event, Thing). Based on the focus-entity, related entities (derived from DBpedia and all annotations available on the platform) are displayed in a four column grid. The pagination bars on the right hand side of each column allow the users to browse through all entities in a category. When hovering one of the displayed entities (e.g. *Jacques Cousteau* in Fig. 1.4), relations to the focus-entity

and to further entities in the grid-view (e.g. *Oceanography*) are visualized by line connectors. A label (property) indicates the direction and type of connection. If there are more entity-relations than displayed in the first overview, connections to hidden entities are indicated by dotted lines, which can be activated via hovering a small 'plus' icon inside the entity box. A click on an entity in the grid-view replaces the focus-entity with the selected item and refreshes the related entities in all categories. A ranked list of recommended blogposts for the entity in focus is displayed on the bottom-left. The recommendations comprise blogposts that cover the focus-entity and entities related to the focus entity. The more entities are related with the entity in focus, the higher is the rank of the recommended article in the list.

3 Summary

We have presented *refer* as a tool to help blog authors to enhance their content with the help of Linked Open Data and to engage the readers to further explore the available content as well as to provide background information from DBpedia and Wikipedia. A proof-of-concept is provided as being integrated into a daily weblog³. The Wordpress-plugin of *refer* is publicly available for download⁴, and a screencast⁵ demonstrates the annotator interfaces.

Acknowledgement: This work has been funded by the German Government, Federal Ministry of Education and Research under project number 03WKCJ4D.

References

1. A. Khalili and S. Auer. WYSIWYM – integrated visualization, exploration and authoring of semantically enriched unstructured content. *Semantic Web Journal*, 1, 2014.
2. C. Morbidoni and A. Piccioli. *The Semantic Web: ESWC 2015 Satellite Events*, chapter Curating a Document Collection via Crowdsourcing with Pundit 2.0, pages 102–106. Springer, 2015.
3. J. Osterhoff, J. Waitelonis, and H. Sack. Widen the Peepholes! Entity-Based Auto-Suggestion as a rich and yet immediate Starting Point for Exploratory Search. In *Proc. of 2nd Workshop Interaction and Visualization in the Web of Data (IVDW 2012)*, 2012.
4. T. Tietz, J. Waitelonis, J. Jäger, and H. Sack. Smart media navigator: Visualizing recommendations based on linked data. In *13th Int. Semantic Web Conf.*, 2014.
5. R. Usbeck et al. Gerbil: General entity annotator benchmarking framework. In *Proc. of the 24th Int. Conf. on World Wide Web*, pages 1133–1143. ACM, 2015.
6. J. Waitelonis and H. Sack. Named entity linking in #tweets with kea. In *6th workshop on 'Making Sense of Microposts', Named Entity Recognition and Linking (NEEL) Challenge, at the 25th Int. World Wide Web Conf.*, 2016.

³ <http://blog.yovisto.com>

⁴ <http://refer.cx/>

⁵ http://s16a.org/sites/default/files/refer/refer_Screencast.mov

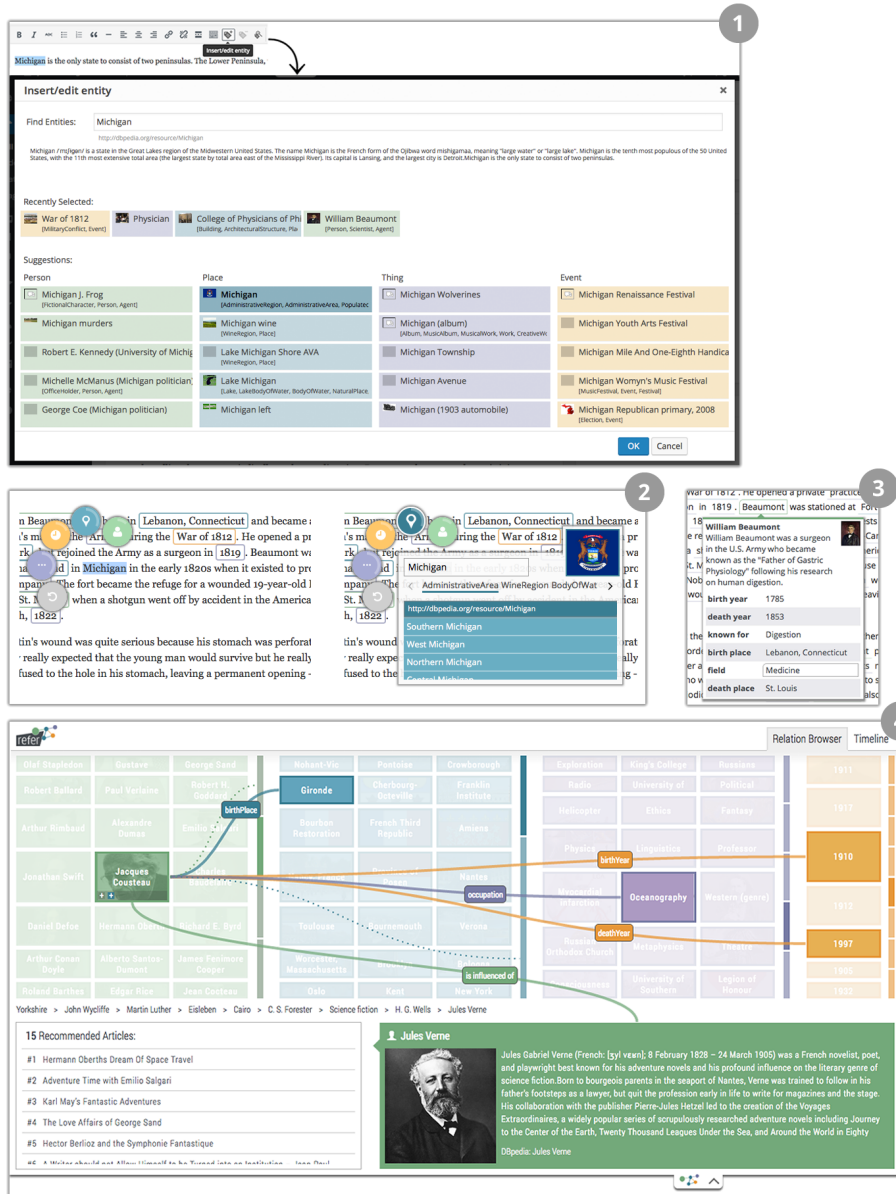


Fig. 1. refer components: Modal Annotator (1), Inline Annotator (2), Infobox (3), Relation Browser and Recommender (4)