ELES: Combining Entity Linking and Entity Summarization

Andreas Thalhammer^{$(\boxtimes)}$ </sup> and Achim Rettinger

AIFB, Karlsruhe Institute of Technology, Karlsruhe, Germany {andreas.thalhammer,achim.rettinger}@kit.edu

Abstract. The automatic annotation of textual content with entities from a knowledge base is a well established field. Applications, such as DBpedia Spotlight and GATE enable to identify and disambiguate entities of text at high levels of accuracy. The output of such systems can be used in many different ways. One way is to show knowledge panels which provide a fact-based summary of an entity and provides further information as well as browsing options. Such fact-based summaries are produced by entity summarization systems.

This paper presents ELES, a lightweight combination of DBpedia Spotlight and the SUMMA entity summarization interface. DBpedia Spotlight analyzes text and links fragments to entities of the DBpedia knowledge base. The LinkSUM summarizer (interfaced via the SUMMA API definition) produces fact-based summaries of DBpedia entities. The two applications are combined on the client side through the "Internationalization Tag Set 2.0" W3C recommendation and lightweight jQuerybased interfaces.

1 Introduction

The field of linking fragments from text to entities of a knowledge base is currently at an advanced stage: words and compounds can be identified and disambiguated at high levels of accuracy [1,2,8]. Entity linking usually enables to provide further information on the entities, browsing, or recommendation. In this work, we consider the use case of providing a knowledge panel that pops up on mouse over entities that were identified through annotation. The knowledge panel is used to explain the entities identified by the annotator with concise fact-based information units and also enables further browsing (i.e. navigating through related entities). Entity summarization enables to filter all information that is available about an entity (often more than 1000 facts) and to select a small fraction that will be presented to the user.

With this demo, we propose loose coupling between automatic entity linking and entity summarization systems with the "Internationalization Tag Set 2.0" (ITS 2.0) W3C recommendation [3]. We exemplify the feasibility of the lightweight integration with the applications DBpedia Spotlight [1] (as a entity linking tool) and LinkSUM [4] (as an entity summarization tool interfaced via the SUMMA API [5]). Both applications use DBpedia¹ as a knowledge base.

¹ DBpedia – http://dbpedia.org.

[©] Springer International Publishing Switzerland 2016

A. Bozzon et al. (Eds.): ICWE 2016, LNCS 9671, pp. 547–550, 2016. DOI: 10.1007/978-3-319-38791-8_45

2 Implementation

We make use of ITS 2.0 by using the its-ta-ident-ref attribute as the combining element between entity linking and entity summaries. This attribute can be used with HTML elements such as tags (for an example see Listing 1). The system is implemented in accordance to the following workflow:

- 1. Automatically identify and annotate entities of a text.
- 2. Register a knowledge panel for each identified entity (on mouseover).

Listing 1. Example use of its-ta-ident-ref in a tag.

```
She spent her summers in <span its-ta-ident-ref="http://dbpedia.org/resource/Dublin,_Ohio">Dublin</span>.
```

We had to extend the DBpedia Spotlight jQuery plugin in order to enable ITS 2.0 output.² The system uses a Web service in order to annotate one or more text paragraphs with entities from the DBpedia knowledge base. The entity summaries are produced by LinkSUM [4]. The produced summaries are displayed as knowledge panels via the summaClient³ implementation [5]. The original implementation has been extended by the qSUM method. It registers mouseover events

On 25 September 2007, Merkel met the 14th Dalai Lama for "private and informal talks" in the Chancellery in Berlin amid protest from China. China afterwards cancelled separate		
talks with German officials, including talks with Justice Minister Brigitte Zypries.		
One of Merkel's priorities was strengthening transatlantic economic relations – she signed the agreement for the Transatlantic Economic Council on 30 April 2007 at the		
White House The Council on chaired by an Ell and all'S official, aims at removing		
barriers to tra	2	ade area. This project has been
described as White House		Jean-Luc Mélenchon, fearing a
transfer of so		an alignment of the European
Union on the tenant	President of the United States	
tenant	Barack Obama	
Der Spiegel r style	Neoclassical architecture	erkel and President Barack
Obama ease style	Palladian architecture	n June 2009. Commenting on a
White House location	Northwest, Washington, D.C.	r Spiegel stated, "Of course the
rather more r		th [Obama's] charm
offensive," bu Summary by http://km.aifb.kit.edu/services/link		liplomacy, "she gave it a go…
by mentioning the experiences of Obama's sister in Heidelberg, making it clear that she		
had read his autobiography".		

Fig. 1. Automatically annotated excerpt of a Wikipedia article (https://en.wikipedia. org/w/index.php?title=Angela_Merkel\&oldid=709980123.) and the summaClient knowledge panel with a summary by LinkSUM.

² ITS 2.0 for DBpedia Spotlight – https://github.com/dbpedia-spotlight/demo/pull/5.
 ³ summaClient – http://athalhammer.github.io/summaClient/.

for all elements with an its-ta-ident-ref attribute. The respective knowledge panel is then shown at the position of the respective annotation. The LinkSUM entity summarization system can be easily exchanged by another entity summarization system that implements the SUMMA API.

A screenshot of an automatically annotated text is provided in Fig. 1 and a do-it-yourself example is provided in Listing 2.

Listing 2. Full HTML example with the jQuery (UI), DBpedia Spotlight, and the SUMMA client libraries included.

```
<!DOCTYPE html><html><head><title>Example</title>
<style>span {background-color:#AAAAAA}</style>
<link rel="stylesheet" type="text/css"</pre>
href="http://athalhammer.github.io/summaClient/css/summaClient.css" />
<script src="http://code.jquery.com/jquery-2.2.1.min.js"></script>
<script src="http://code.jquery.com/ui/1.11.4/jquery-ui.js"></script>
<script src="http://dbpedia-spotlight.github.io/demo/dbpedia-spotlight-0.3.js"></script>
<script src="http://athalhammer.github.io/summaClient/js/summaClient.js"></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></scri
<script>
$(document).ready(function() {
   // selector on HTML element(s)
    var select = ".annotate";
    // as soon as the annotations are ready, start registering mouseover events
    // parameters: topK, language, fixed properties, service
    $(select).bind("DOMSubtreeModified", function() {
         qSUM(5, "en", null, "http://km.aifb.kit.edu/services/link/sum");
   }):
   // DBpedia Spotlight configuration and annotation
    var settings = { "endpoint" : "http://spotlight.sztaki.hu:2222/rest", "its" : "yes",
         "spotter" : "Default" };
    $(select).annotate(settings): $(select).annotate("best"):
});
</script></head><body><div class="annotate">Angela Merkel is TIME Person of the Year 2015.
</div></body></html>
```

3 Related Work

Our work on ELES was inspired by Denny Vrandečić's qLabel⁴ project. qLabel is a jQuery-based application that uses text fragments, annotated with its-ta-ident-ref references to Wikidata [7], for translation between different languages. For this, qLabel leverages the wealth of multi-lingual labels available in Wikidata. A Wikidata-based summarization system that uses the SUMMA API can be easily combined with qLabel via qSUM.

There are also a number of proprietary solutions that combine annotation and knowledge panels. The refer⁵ application, partly described in [6], supports automatic annotation in combination with knowledge panels. The application is well integrated with additional browsing features and a complete graph panel that can be enabled at the top of each page. Other proprietary solutions include the Bing Knowledge Widget⁶ and Ontotext's Now⁷. Most of the proprietary

⁴ qLabel – http://googleknowledge.github.io/qlabel/.

⁵ refer – http://refer.cx/.

⁶ Bing Knowledge Widget – https://www.bing.com/widget/knowledge.

⁷ Ontotext Now – http://now.ontotext.com/.

solutions are highly customized and the annotation and knowledge panel parts are often strongly connected.

4 Summary

With ELES, we propose loose coupling between automatic entity linking and entity summarization systems via ITS 2.0. We exemplify the lightweight integration approach with the applications DBpedia Spotlight and the qSUM method of the SUMMA entity summarization interface.

Acknowledgements. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 611346 and by the German Federal Ministry of Education and Research (BMBF) within the Software Campus project "SumOn" (grant no. 01IS12051).

References

- 1. Daiber, J., Jakob, M., Hokamp, C., Mendes, P.N.: Improving efficiency and accuracy in multilingual entity extraction. In: Proceedings of the 9th International Conference on Semantic Systems (I-Semantics) (2013)
- Damljanovic, D., Bontcheva, K.: Named entity disambiguation using linked data. In: Proceedings of the 9th Extended Semantic Web Conference (2012)
- Filip, D., McCance, S., Lewis, D., Lieske, C., Lommel, A., Kosek, J., Sasaki, F., Savourel, Y.: Internationalization Tag Set (ITS) Version 2.0. W3C recommendation. W3C, October 2013. http://www.w3.org/TR/2013/REC-its20-20131029/
- Thalhammer, A., Lasierra, N., Rettinger, A.: LinkSUM: using link analysis to summarize entity data. In: Proceedings of the 16th International Conference on Web Engineering (ICWE 2016) (2016, to appear)
- Thalhammer, A., Stadtmüller, S.: SUMMA: a common API for linked data entity summaries. In: Cimiano, P., Frasincar, F., Houben, G.-J., Schwabe, D. (eds.) ICWE 2015. LNCS, vol. 9114, pp. 430–446. Springer, Heidelberg (2015)
- Tietz, T., Waitelonis, J., Jäger, J., Sack, H., Navigator, S.M.: Visualizing recommendations based on linked data. In: Industry Track at the International Semantic Web Conference (ISWC 2014), vol. 1383. CEUR-WS (2015)
- Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. Commun. ACM 57, 78–85 (2014)
- Zhang, L., Rettinger, A.: X-LiSA: cross-lingual semantic annotation. In: Proceedings of the VLDB Endowment (PVLDB), the 40th International Conference on Very Large Data Bases (VLDB), vol. 7, no. 13, pp. 1693–1696, September 2014