# ELES: Combining Entity Linking and Entity Summarization

Andreas Thalhammer and Achim Rettinger

AIFB, Karlsruhe Institute of Technology
{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 jQuery-based 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 as well as semantic browsing or recommendation. In this work, we consider the use case of providing a knowledge panel that pops while the mouse pointer hovers over entities that were previously identified and linked. The knowledge panel is used to explain the entities identified by the annotator with concise fact-based information units and also enables further browsing. 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 an entity linking tool) and LinkSUM [4] (as an entity summarization tool interfaced via the SUMMA API [5]). Both applications use DBpedia[1] as a knowledge base.

[1] DBpedia – http://dbpedia.org

**Listing 1.** Example use of `its-ta-ident-ref` in a `<span>` tag.

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

## 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 `<span>` 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).

We extended the DBpedia Spotlight jQuery plugin in order to enable ITS 2.0 output.[2] The system uses a DBpedia Spotlight deployment 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 summaClient[3] implementation [5]. The original summaClient implementation has been extended by the qSUM method. It registers mouseover events for all elements with an `its-ta-ident-ref` attribute. The respective knowledge panel is then shown at the position of the respective annotation on mouseover. The LinkSUM entity summarization system can be easily replaced by any other entity summarization system that implements the SUMMA API. A screenshot of an automatically annotated text is provided in Figure 1 and a do-it-yourself example is provided in Listing 2.

## 3 Related Work

Our work on ELES was inspired by Denny Vrandečić's qLabel[4] 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 refer[5] 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

[2] ITS 2.0 for DBpedia Spotlight – `https://github.com/dbpedia-spotlight/demo/pull/5`
[3] summaClient – `http://athalhammer.github.io/summaClient/`
[4] qLabel – `http://googleknowledge.github.io/qlabel/`
[5] refer – `http://refer.cx/`

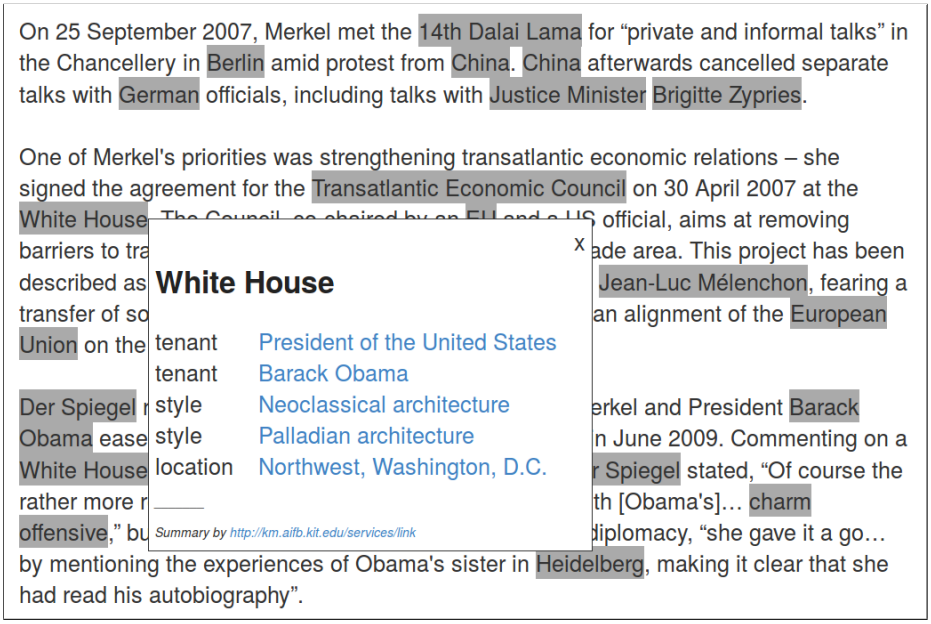


**Fig. 1.** Automatically annotated excerpt of a Wikipedia article[9]and the summaClient knowledge panel with a summary by LinkSUM.

that can be enabled at the top of each page. Other proprietary solutions include the Bing Knowledge Widget[6] and Ontotext's Now[7]. Most of the proprietary 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.

**Acknowledgement.** 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).

[6] Bing Knowledge Widget – `https://www.bing.com/widget/knowledge`
[7] Ontotext Now – `http://now.ontotext.com/`
[9] `https://en.wikipedia.org/w/index.php?title=Angela_Merkel&oldid=709980123`

**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"
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>
$(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>
```

## References


1. J. Daiber, M. Jakob, C. Hokamp, and P. N. Mendes. Improving Efficiency and Accuracy in Multilingual Entity Extraction. In *Proceedings of the 9th International Conference on Semantic Systems (I-Semantics)*, 2013.
2. D. Damljanovic and K. Bontcheva. Named Entity Disambiguation using Linked Data. In *Proceedings of the 9th Extended Semantic Web Conference*, 2012.
3. D. Filip, S. McCance, D. Lewis, C. Lieske, A. Lommel, J. Kosek, F. Sasaki, and Y. Savourel. Internationalization Tag Set (ITS) Version 2.0. W3C recommendation – `http://www.w3.org/TR/2013/REC-its20-20131029/`, W3C, Oct. 2013.
4. A. Thalhammer, N. Lasierra, and A. Rettinger. LinkSUM: Using Link Analysis to Summarize Entity Data. In *Proceedings of the 16th International Conference on Web Engineering (ICWE 2016)*. To appear, 2016.
5. A. Thalhammer and S. Stadtmüller. SUMMA: A Common API for Linked Data Entity Summaries. In *Engineering the Web in the Big Data Era*, volume 9114, pages 430–446. Springer International Publishing, 2015.
6. T. Tietz, J. Waitelonis, J. Jäger, and H. Sack. Smart Media Navigator: Visualizing Recommendations based on Linked Data. In *Industry Track at the International Semantic Web Conference 2014 (ISWC 2014)*, volume 1383. CEUR-WS, 2015.
7. D. Vrandečić and M. Krötzsch. Wikidata: A Free Collaborative Knowledgebase. *Commun. ACM*, 57:78–85, 2014.
8. L. Zhang and A. Rettinger. X-lisa: Cross-lingual semantic annotation. *Proceedings of the VLDB Endowment (PVLDB), the 40th International Conference on Very Large Data Bases (VLDB)*, 7(13):1693–1696, September 2014.