

# **Link server aggregation with BEACON**

*Jakob Voß<sup>1</sup>, Mathias Schindler<sup>2</sup>, Christian Thiele*

<sup>1</sup>Verbundzentrale des GBV (VZG), Platz der Göttinger Sieben 1, Göttingen

<sup>2</sup>Wikimedia Germany, Berlin

## **Abstract**

We demonstrate BEACON, the simplest possible file format to exchange links. It is successfully used by more than 50 institutions to provide links to be aggregated by Wikipedia and other projects.

## **Link servers and aggregation**

Although links are its most powerful feature, the WWW is primarily build of documents. The limitation to embedded, outgoing hyperlinks, instead of links as first class objects, has been criticized regularly. Dedicated link servers allow links to be created, managed, and reused in context (Signer et. al. 2009; Michaelides 2001). But link server standards such as the Fundamental Open Hypermedia Model have not been widely adopted, and link servers such as SFX (Robertson et al 2009) remain closed data silos. With SeeAlso we have developed a simple link server protocol based on OpenSearch Suggestions and unAPI (Voß 2008). SeeAlso is used for catalog enrichment in German libraries, and implemented as open source. However, the protocol is not designed for aggregating large amounts of links from many different data providers. Neither the OAI harvesting protocol seems to provide an adequate way to exchange simple sets of links, especially for small data providers with limited technical skills. A possible solution for aggregating links is providing them as Linked Open Data in RDF. However, with multiple serialization formats and ontologies, RDF adds more complexity than needed. For this reason we created the BEACON file format as the simplest thing that could possibly work. The main design goal was ease of creation: files can be generated manually or with simple tools, such as spreadsheets, provided by any transport mechanism (HTTP, OAI...). Later they can be converted to RDF.<sup>1</sup>

---

<sup>1</sup>See detailed description of BEACON at <http://meta.wikimedia.org/wiki/BEACON> and its applications at <http://de.wikipedia.org/wiki/Wikipedia:PND/BEACON> (German).

## BEACON format and its usage

The UTF-8 encoded format starts with a set of meta fields, followed by an unordered list of links, one per line. Each link is build of an identifier as source, a target URI, and an optional label and/or description. The following excerpt contains links from German name authority records (PND) to lists of appropriate holdings of the Bavarian State library, with the number of hits:

```
#FORMAT: PND-BEACON
#PREFIX: http://d-nb.info/gnd/
#TARGET: http://opacplus.bsb-muenchen.de/search?pnd={ID}
118584596|5819
11850553X|4151
```

Its condensed form (the first link is from <http://d-nb.info/gnd/118584596> to <http://opacplus.bsb-muenchen.de/search?pnd=118584596>, annotated with ‘5819’ hits) allows storing one million such links uncompressed with a size of 12.5MB (3.35MB compressed). Links in BEACON format are already provided by more than 50 institutions and harvested BEACON files are used in projects such as Wikipedia-Personensuche,<sup>2</sup> Bayerisches Musiker-Lexikon Online,<sup>3</sup> and ‘Linked History’ at University of Leipzig.<sup>4</sup>

## References

- Michaelides, D., Millard, D., Weal, M., Roure, D. (2001) Auld Leaky: A Contextual Open Hypermedia Link Server. Proc. of the ACM Conference on Hypertext and Hypermedia, Springer, LNCS 2266, p. 59–70.
- Robertson, W., Soderahl, P. (2004) Everything you always wanted to know about SFX but were afraid to ask. The Serials Librarian 47:1/2, p. 129–138.
- Signer, B., de Spindler, A., Norrie, M. (2009) A Framework for Link Sharing in Cooperative Cross-Media Information Spaces. Proc. of CDVE ’09, Springer, LNCS 5738, p. 185–192.
- Voß, J. (2008). SeeAlso: A Simple Linkserver Protocol. Ariadne Issue 57, October 30<sup>th</sup>, <http://www.ariadne.ac.uk/issue57/voss/>.

<sup>2</sup> See Wikipedia People Search at <http://toolserver.org/~upper/pd/> (German).

<sup>3</sup> See <http://www.bmlo.lmu.de/> (German).

<sup>4</sup> See <http://aksw.org/Projects/LinkedHistory/pnd/> (not limited to BEACON format).