given topic.

In 2009, the *New York Times* started to convert its entire subject headings index to Semantic Web format. Not only is the entire headings list being published in RDF format, but the terms are also being linked to hubs on the Linked Data Web such as DBPedia, Freebase, and GeoNames. By 2010, this process was finished for about 30 percent of all subject headings.

The *New York Times* has reported that the linking of their subject headings with the Linked Data Web helps them to provide geographic information (through the link with GeoNames), helps them to align articles with other data sources that are often used by the newspaper, such as the Library of Congress, and makes it possible to rapidly build nonstandard mash-ups of their archive material, such as finding all items about the alumni of a given University.

6.6 Sig.ma and Sindice

Whereas applications such as GoodRelations help traditional search engines to improve their results without ever confronting the end-user with the underlying RDF graphs, the search engine Sindice[17] and its front-end Sig.ma[18] are different: Sindice operates directly on the RDF graph that makes up the global Semantic Web, and Sig.ma presents to the user the binary relations that are discovered in this way.

Sindice has a classical search engine architecture: a crawler, a very large indexed store, and a retrieval interface. But the difference from a classical search engine is that Sindice does not retrieve and index words and phrases, but instead retrieves and indexes RDF graphs. At the time of writing, Sindice indexes some 400 million RDF "documents," resulting in 12 billion RDF statements. These statements are indexed and available for querying through a SPARQL endpoint. This produces a database-like

[17] http://sindice.com.
[18] http://sigm.ma.