Case Study

—

Intelligent tagging with metadata

Eidosmedia S.p.A.
Mac 7, Via C. Imbonati 18
20159 Milano
Italy
Tel. +39 02 36732000
Fax +39 02 36732999
info@eidosmedia.com
http://www.eidosmedia.com

# Table of contents

Background

The daily is one of Europe's leading sources of financial and business information and debate. Since 2009 it has used a multi-channel Eidosmedia platform to produce its print edition, Web portals and e-paper and manage part of its editorial archives.

Challenge

For a business publication whose news content has significant reference value (even after publication), the retrievability of items through search and reference mechanisms is an important part of the service it offers its paying subscribers.

The ease and speed with which an article can be located by search mechanisms is significantly increased by tagging the item with the right metadata. This classification is normally done by the author of the article, typically by choosing terms from a list and entering names. The step is often skipped or done badly by busy journalists.

What was needed was a process that would create richer and more comprehensive semantic metadata without over-burdening the authoring staff.
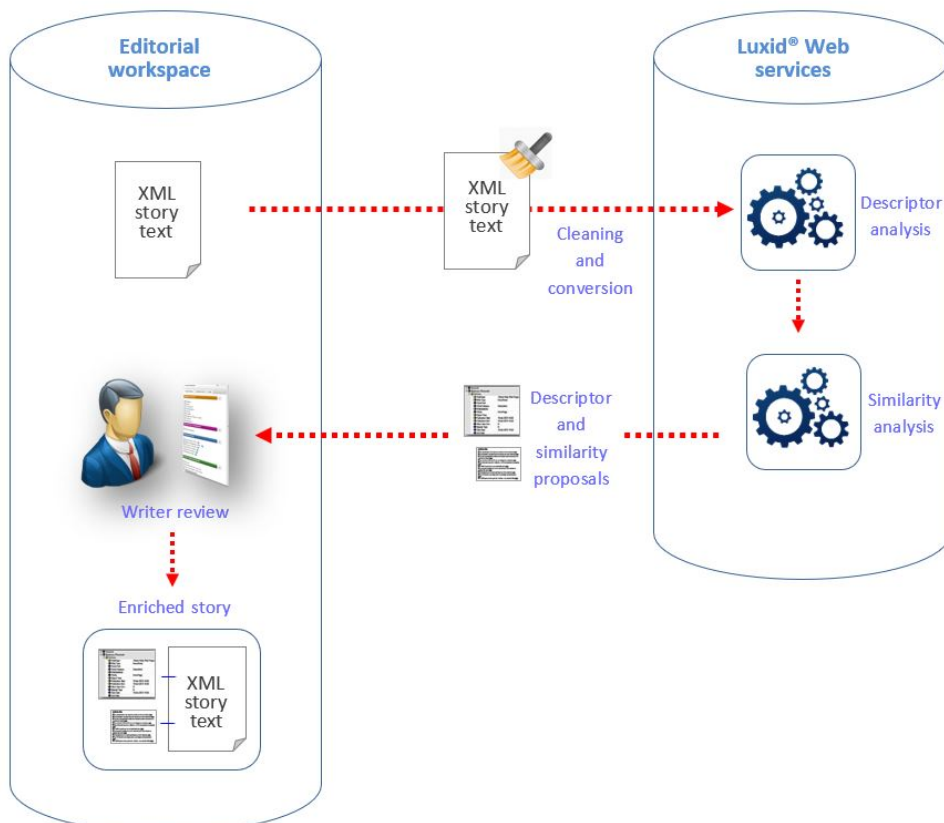
Solution

The solution adopted is based on the Luxid expert system for the semantic analysis of unstructured data developed by French company Temis.

When an author releases a story for publication, the XML story file is first 'cleaned up' and converted into a format that Luxid's analytical tools can read. It is then sent to the Luxid remote platform where two kinds of web-based analyses are undertaken.

The first analyses the content to generate 'descriptors' (people, locations, categories etc.) which will be used to enrich the metadata, including tagging with IPTC topic newscodes.

The second identifies other documents within the editorial corpus, kept in an external archive, which exhibit 'similarity' with the analyzed text.

The results are then returned to the editorial workflow and the author is invited to review the tagging, removing the checks from boxes next to terms that do not apply.

The author may also add locations to the geodata using an auto-completion mechanism fed by a link to an external database.

At the end of the operation, the corrected metadata is appended to the news item and it is archived.

Outcomes

The auto-tagging solution significantly improved the quality and relevance of the metadata associated with the published articles, while cutting the time needed to complete the task.

Authors were far more willing to review and confirm metadata than to originate it.

There were immediate benefits in terms of the accessibility of the paper's archived content and the solution provided a sound basis for the advanced search and retrieval functions planned for the future.