



June 2011 Brochure

**Web Mining is Chasing Jobs
in the UK**

**UK Government Web Archive
goes Semantic**

**Semantic Biomedical Tagger
Scores in Mexico**

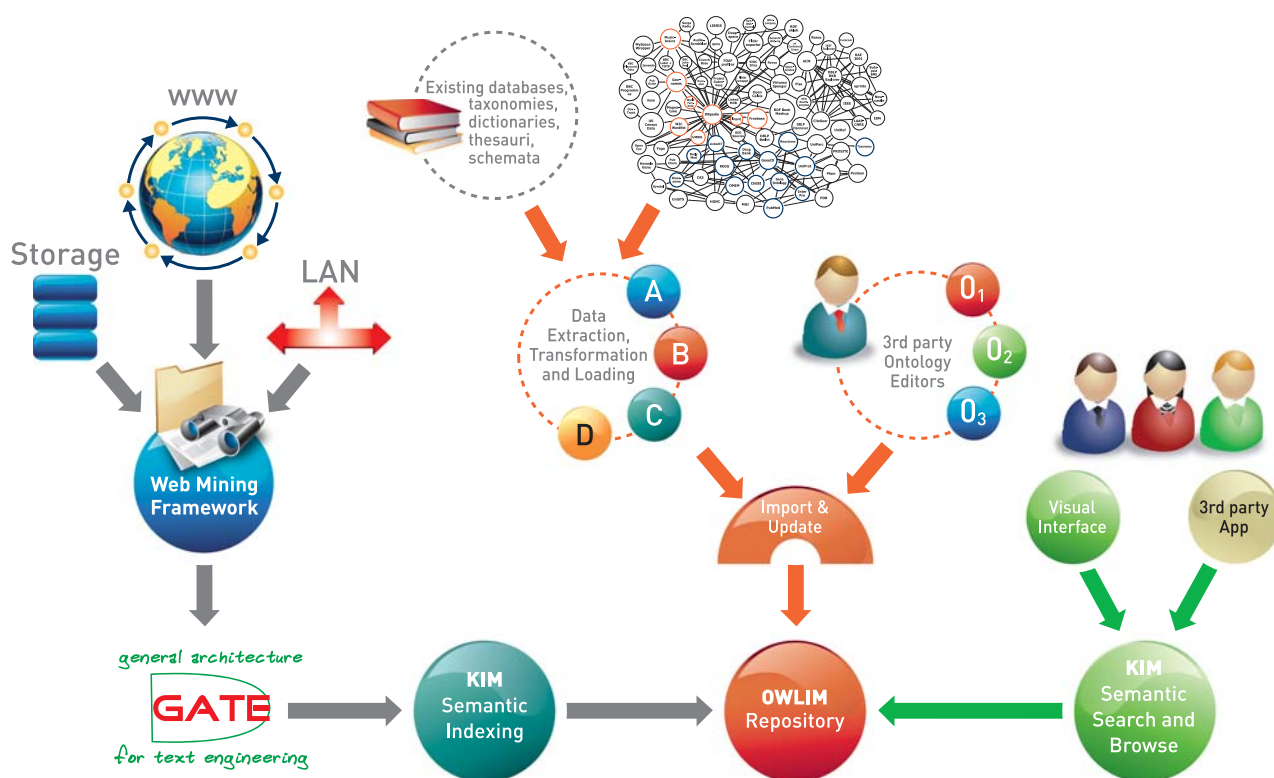
**BBC SPORT
Wins with Ontotext**

Linked Data for the Enterprise

Every organization struggles to get the right information to the right people when they need it – whether these people are internal analysts, executive decision-makers, or customers. But often the right data is in multiple formats, various locations and being accessed by different systems.

Enterprises need a better approach to find information where it lives, link it based on meaning, search it efficiently, and manage changes more easily.

This approach is *Linked Data for the Enterprise*, powered by semantics, built by Ontotext.



The Ontotext Value Proposition



Ontotext's unique interdisciplinary expertise provides a single access point into all technologies that turn the promise of semantics and linked data into reality.

Our technology unveils implicit relations and finds hidden links, matching facts scattered across huge volumes of diverse information. It delivers more answers in less time and with less effort.

Ontotext focuses on technology leadership that is sustainable, practical, and readily available now.

THE BUSINESS CASE:

A Shorter Adoption Curve for Long-Term Benefits

- Flexible data schemas adapt to changing information more easily
- Search across multiple sources without increasing complexity
- Text mining saves manual indexing & data entry time, without loss of depth and quality
- Our solutions can be applied in cooperation with existing data, storage, search, and publishing systems

OUR PROMISE: Sustainable Value

- We deliver semantic technologies at enterprise scale and reliability
- Ontotext has expertise in the full range of semantic technologies
- The up-front investment is lower than you might expect
- Open standards protect your investment
- We are committed to training, knowledge transfer, and support



Products & Client Profiles

OWLIM Repository

RDF Data Management

The OWLIM family of RDF semantic database engines offers high performance and reliability at practical enterprise scale. OWLIM is efficient and flexible, enabling query responses based on inference and integrated full-text search.

Compliant with a wide range of standards, including RDF(S), OWL 2 RL and QL, and SPARQL 1.1, OWLIM comes in three varieties:

OWLIM Lite: A free light-weight database. The ideal entry point for researchers and those new to RDF applications.

OWLIM SE: A single-server RDF database platform with transparent reasoning support throughout the entire life cycle of the data and many advanced features, such as native support for geospatial data. Suitable for group-level information needs.

OWLIM Enterprise: All the features of OWLIM SE, in a replication cluster configuration and reliability features to take RDF data management to true enterprise grade.

<http://www.ontotext.com/owlim>

BBC, London

Dynamic Web Publishing

BBC's 2010 World Cup web site was the first project running on a Dynamic Semantic Publishing (DSP) architecture. DSP provides improved navigation, content re-use and re-purposing, and supports semantic advertising.

OWLIM serves as an RDF store in the middle of the architecture facilitating metadata-based management and retrieval of content. Later on, Ontotext also implemented another important part of the DSP architecture - the Concept Extraction service.

"We are happy to have chosen OWLIM, after evaluation of several RDF stores, and SPARQL in our approach to build a high-performance semantic stack for the World Cup 2010 site"

"The Concept Extraction Service developed by Ontotext becomes the main content enrichment facility for BBC SPORT"

Jem Rayfield,
Senior Technical Architect, BBC



Web Mining Framework

Grab structured data from the deep web

Web Mining Framework (WMF) is a suite of data access and acquisition technologies, including crawlers, connectors, scrapers and text miners, enabling efficient extraction and consolidation of structured information from web sites, blogs, aggregators, and social media.

WMF is modular and can be used as a flexible, manageable platform for projects involving OWLIM and KIM, or other data storage and search systems.

The WMF system is in use in several commercial services, including UK-based Innovantage.

<http://www.ontotext.com/wmf>

Innovantage, Cardiff

Online Recruitment Intelligence

Innovantage is a leading UK provider of labour market intelligence. By tracking millions of live online job postings each month, advertised across 200 job boards as well as tens of thousands of employer websites, the software provides users with the most comprehensive analysis of recruitment demand.

"Ontotext's solution allows Innovantage to maintain the most comprehensive database of job adverts in the UK. Semantic technology allows us to be way more efficient and flexible than our competitors"

Matthew Dewstowe
Managing Director of Innovantage

Products & Client Profiles



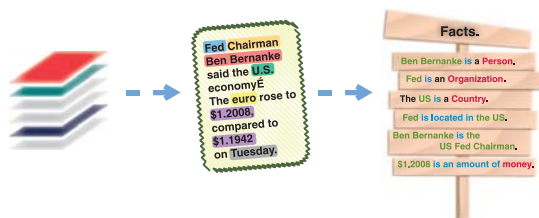
KIM Platform

Semantic enrichment, linking and search

KIM empowers you to link, enrich and search your information space to find data, content and facts.

It creates semantic links between your documents, data and ontologies by finding entities and facts.

As a platform, it is highly adaptable, scales as your needs grow, and smoothly integrates into your existing software environment.



<http://www.ontotext.com/kim>

The National Archives, Richmond, UK

Semantic KB of the Government Web Archive

The Semantic KB project applies text mining to the UK Government Web Archive to interlink it with DATA.UK.GOV and other LOD datasets. Here KIM delivers semantic search at ultimate scale: 150 million unique documents annotated against 5 billion triples.

The tender was granted to a consortium led by Ontotext and comprising of the GATE team from the University of Sheffield and System Simulation Ltd.

"The Semantic KB implementation delivered by Ontotext solves two problems: the collection contains a large amount of near-duplicate material which makes conventional search tools not very effective; another major issue is how users continue to find topical information when government structures change over time"



Amanda Spencer
Head of Web Continuity at TNA

Life Sciences

Linking Biomedical Knowledge

Ontotext maintains unique expertise and technology to solve information management problems in the biotechnology, pharmaceutical, healthcare, and electronic patient records industries.

Our applications efficiently process billions of biomedical facts and link them to semi- or unstructured enterprise documents using semantically rich ontologies.

For an example of linked bio-medical facts, see Linked Life Data on page 6.

<http://www.ontotext.com/life-sciences>

Preventive Health Technologies, Miami

Semantic Health Management

Preventive Health Technologies is working with Mexican healthcare company HDO to create a smart answering system for preventive health. The system will match participating patient records with basic treatment information and a database of common questions and answers accessible via SMS messaging. It would allow users to not only receive reminders, but to be given answers to common questions based on their needs. And, as the system grows, its ability to match questions with known answers will increase thanks to the semantic data matching provided by Ontotext.

"We are very excited about this project. We've chosen Ontotext because they have great semantic technologies, understand the health field, and are committed to open standards, which are critical to our success."

Gabriela Sabate,
CEO of Preventive Health Technologies

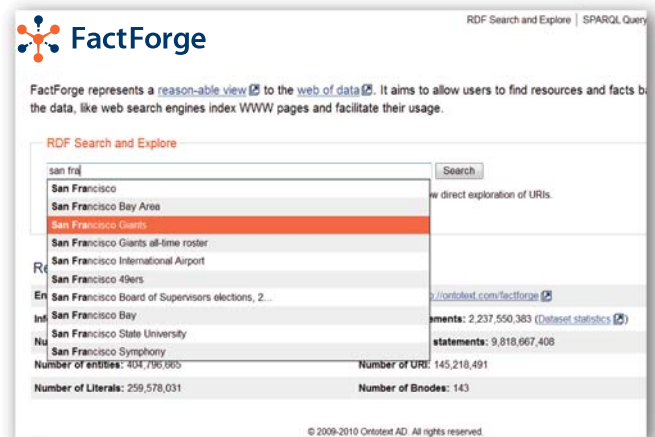
Public Data Services

FactForge

Free linked data portal, offering fast track to the center of the data web. One can query and explore about 2 billion triples, integrating 10 of the most popular Linking Open Data (LOD) data sources, including DBpedia, GeoNames, Freebase, Wordnet, and MusicBrainz. Powered by OWLIM, the portal provides a unique experience of using this data, considering its semantics.

FactForge offers a quick start for using these datasets into enterprise search systems.

<http://www.factforge.net>

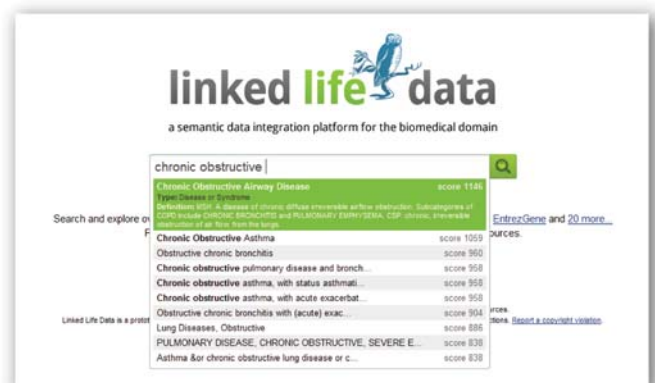


Linked Life Data

Linked Life Data service is a public showcase of Ontotext's semantic warehousing platform. The service semantically integrates and aligns more than 26 diverse data sources by identifying similar concepts. The warehouse aggregates more than 1 billion biomedical objects described by more than 5 billion RDF statements generated using reasoning and text analysis methods.

Exploit Interactive Relationship Discovery in billions of biomedical entities aggregated from a multitude of heterogeneous data sets.

<http://www.linkedlifedata.com>



Lupedia

Lupedia is a text enrichment service that can look up LOD entities in plain text. These entity references can be used as metadata for further semantic operations. The service is multilingual and currently uses DBpedia and LinkedMDB as lookup datasets.

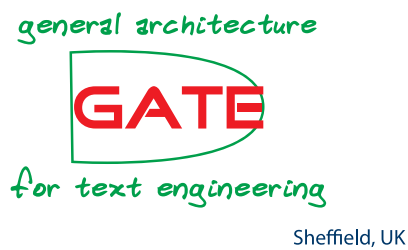
Lupedia is used in the NoTube project to add LOD metadata to TV content descriptions. The metadata is then used as a basis for recommending other programmes of potential interest to the viewer.

A recent evaluation of Lupedia against similar existing services (OpenCalais and AlchemyAPI) showed that Lupedia had a much greater coverage and comparable precision.

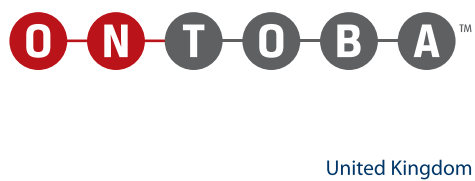
<http://lupedia.ontotext.com/>



TECHNOLOGY PARTNERS



REGIONAL PARTNERS



RESEARCH PROJECTS

Ontotext participates in the following projects funded under the Seventh Framework Program (FP7) of the European Commission:

CUBIST - combines essential features of Semantic Technologies and Business Intelligence
<http://www.cubist-project.eu/>

INSEMTIVES - a platform for user involvement in semantic content enrichment through social and economical incentives.
<http://www.insemtives.eu/>

Khresmoi - to build a multi-lingual multi-modal search and access system for biomedical information and documents
<http://www.khresmoi.eu/>

LarKC - The Large Knowledge Collider: a platform for large-scale integrated reasoning and Web-search
<http://www.larkc.eu/>

MOLTO - cross-lingual search and ontology-language interoperability. Asking questions in a natural way.
<http://www2.molto-project.eu/>

NoTube - semantic technologies for personalised creation, distribution and consumption of TV content
<http://www.notube.tv/>

RENDER - developing methods, techniques, software and datasets that will leverage diversity as a crucial source of innovation and creativity
<http://www.render-project.eu/>

SOA4ALL - combines a service delivery platform based on Web 2.0, Semantics and Context Management to enable a Web of billions of services
<http://www.soa4all.eu/>

MEMBERSHIPS



STI · INTERNATIONAL

Proud Sponsor of:



**2011 Semantic
Technology Conference**
JUNE 5-9, 2011 SAN FRANCISCO
semanticweb.com



Ontotext is part of Sirma Group
Copyright © 2011 Ontotext AD. All rights reserved.
Specifications are subject to change without notice.
All product names cited are trade names or trademarks of their respective owners.

Ontotext USA Inc.

2490 Black Rock Turnpike #331
Fairfield
Connecticut 06825-2400, USA

Contact: Mr. Matthew Petrillo
Tel. +1 (718) 785 9692
Toll-free for N. America:
+1 (866) 972 6686

Ontotext AD

135 Tzarigradsko Chaussee, Floor 3
1784 Sofia, Bulgaria

Contact: Ms. Kamelia Atanasova
Tel: (+359 2) 974 61 60
Fax: (+359 2) 975 32 26

info@ontotext.com
www.ontotext.com