



# NEXT GEN ARCHITECTURE MODELS: BBC WORLD CUP 2010 SITE ARCHITECTURE & PROPOSED 2012 OLYMPICS SITE APPROACH

Michael McGrath

March/2011

# BBC WORLD CUP SITE (SEMANTIC CORE)

- The BBC World Cup 2010 site is a large site with over 700 aggregation pages (called index pages)
- Uses a Semantic Web and Linked Data core
- Designed to lead users to the thousands of story pages and content which make up the whole site.
- Examples of index pages range from the Groups and Fixtures page through to detailed pages for each team or player.



**BBC Mobile** News Sport Weather Travel TV Radio More Search the BBC

**SPORT WORLD CUP 2010**

SPORT FOOTBALL WORLD CUP 2010 GROUPS & TEAMS FIXTURES & RESULTS

**Blogs**

**David Bond**  
South Africa kicks off new era  
David Bond on how the Rainbow Nation and the World Cup have changed each other forever

**Paul Fletcher**  
Tough night for Webb as Spain triumph  
English referee comes under fire from defeated Dutch

**Dan Walker**  
My World Cup memories  
Dan Walker's bus trip around South Africa nears its conclusion

**Spain return to rapturous welcome**  
Spain's victorious World Cup squad return to Madrid in a joyous celebration in the capital city.

**Top scorers**

PLAYER	TEAM	GOALS
Mueller	GER	5
Villa	ESP	5
Sneijder	NED	5
Forlan	URU	5
Higuain	ARG	4

**Argentina offer Maradona new deal**  
Ballack warns Lahm over captaincy  
Heroes' welcome for beaten Dutch

**Uruguay hails its football heroes**  
**SA to make**  
**Alonso critical of De Jong tackle**  
**England slip ranking**



**BBC** Home News Sport Weather iPlayer TV Radio More Search

**SPORT WORLD CUP 2010**

SPORT FOOTBALL WORLD CUP 2010 GROUPS & TEAMS FIXTURES & RESULTS VIDEO BBC COVERAGE

**Frank Lampard**

Position: Midfielder  
Squad number: 8  
Date of birth: 20 June, 1978 (32 years old)  
Height: 6'0" (184cm)

**Tournament totals**

Games played	Shots on target   off target	Assists	Faults by   on	Cards yellow   red
4	9   5	0	3   4	0   0
Total England	38   23	4	44   44	6   0

**Latest matches**

- NED 2-1 BRA
- URU 1-1 GHA
- ARG 0-4 GER
- PAR 0-1 ESP

**Features**

- World Cup scouting: attacking midfield
- Alan Shearer assesses James Milner's rise to prominence
- World Cup scouting: defensive midfield

**Top 5 World Cup stories**

- Blanc hints at Les Bleus shake-up
- Tuesday's World Cup round-up
- Fabregas doubtful for semi-final
- Lahm out to keep German captaincy
- Facebook role in Nigeria U-turn

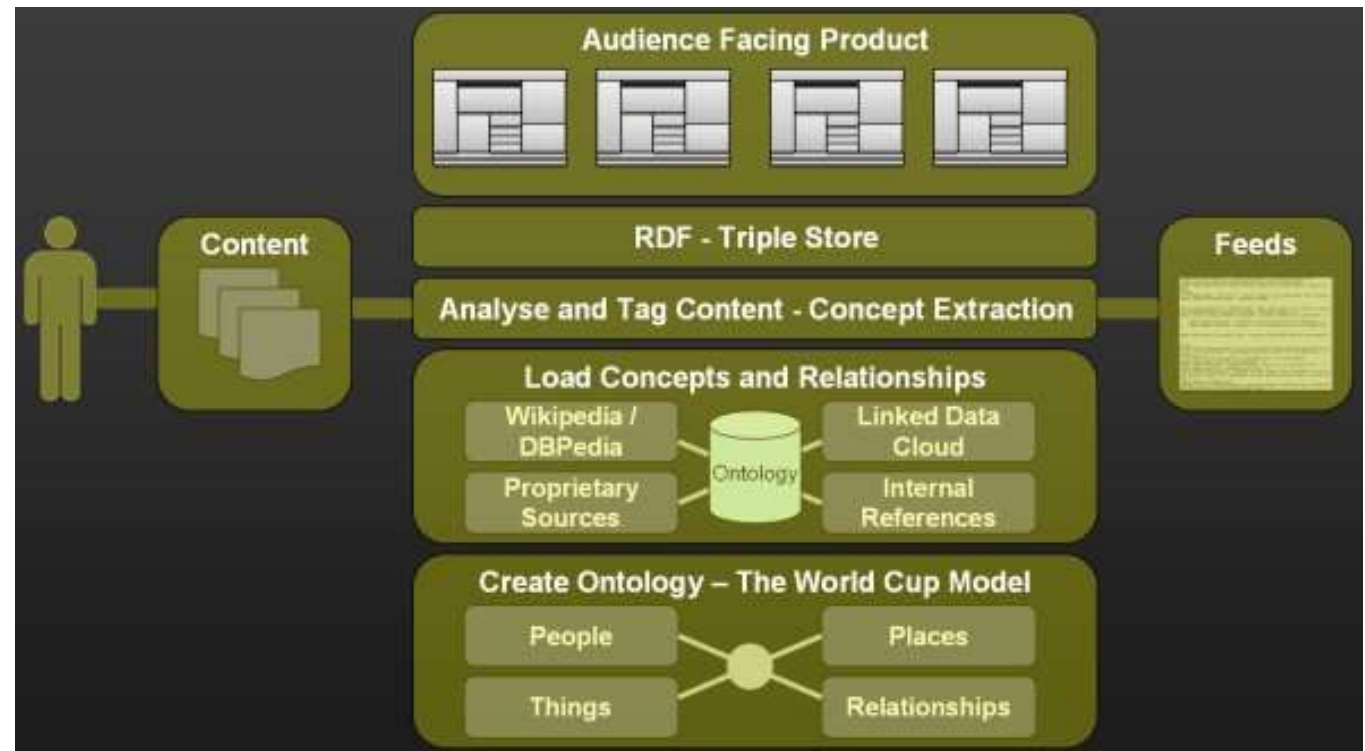
**Top TV and radio**

- Match of the Day Live: 2010 FIFA World Cup: Argentina v Germany
- Match of the Day Live: 2010 FIFA World Cup: Spain v Paraguay
- 5 live Sport: The Monday Night Club: 05/07/2010

More football on BBC iPlayer

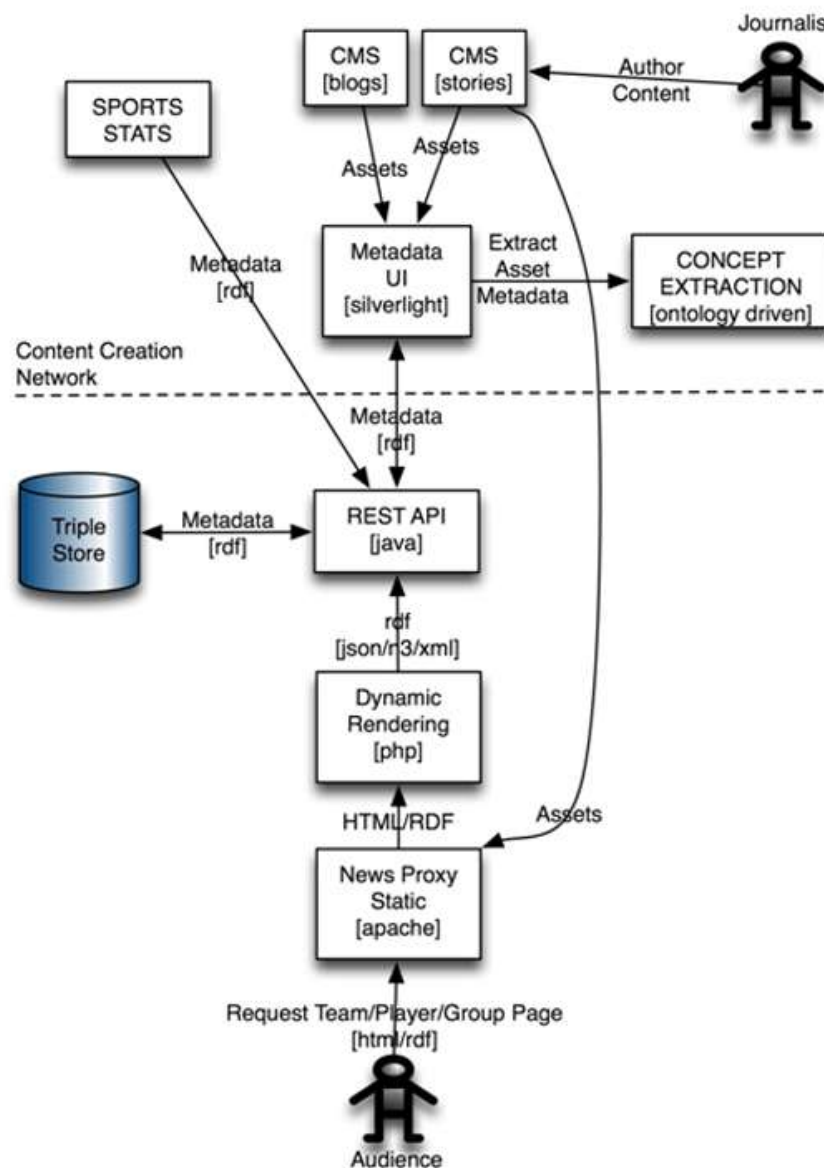
# BBC WORLDCUP ARCHITECTURE

- Using advanced methods for analysing content and deciding how to tag this content with precise metadata linked to uniquely identified concepts (a concept usually being a person, place or thing).
- In the case of the world cup the BBC are interested in players, teams, matches, etc... but the principle can be easily applied to anything. (Compare this with Telco's AFL & NRL sites, etc)
- To do this the BBC are using some technology from IBM (Language-ware) and Ontotext (BigOWLIM)



# BBC WORLD CUP: UNDER THE COVERS

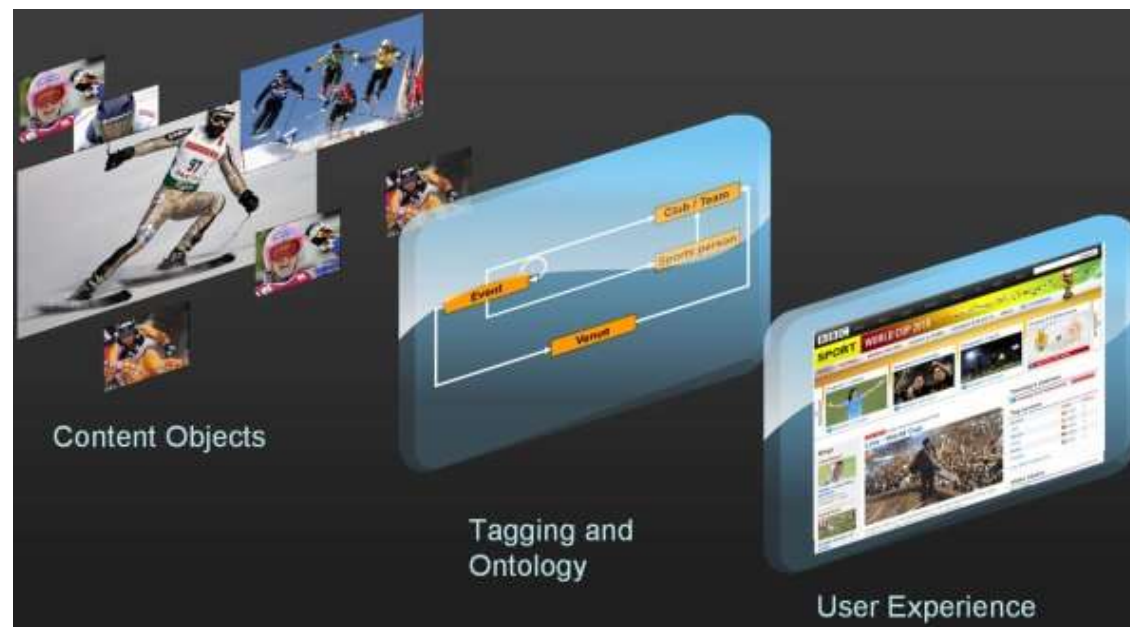
- this is a far deeper and richer use of content than can be achieved through traditional CMS-driven publishing solutions
- features 700-plus team, group and player pages, are powered by a high-performance dynamic semantic publishing framework.
- the publication of automated metadata-driven web pages that are light-touch, requiring minimal journalistic management, as they automatically aggregate and render links to relevant stories.
- The underlying publishing framework does not author content directly; rather it publishes data about the content - [metadata](#).
- The published metadata describes the world cup content, providing rich content relationships and semantic navigation. By querying this published metadata we are able to create dynamic page aggregations for teams, groups and players.
- Dynamic aggregations use a rich ontological domain model- to describes entity existence, groups and relationships between woldcup things/concepts
- For example, "Frank Lampard" is part of the "England Squad" and the "England Squad" competes in "Group C" of the "FIFA World Cup 2010".





# BBC WORLD CUP ARCHITECTURE ADVANTAGES

- The difference is in the use of Semantic Approaches and [RDF and Linked Data](#) to build and manage the site
- Another way to think about all this, is that they are not publishing pages, but publishing content as assets which are then organised by the metadata dynamically into pages, but could be re-organised into any format we want much more easily than they could before
- This means dynamic creation/translation of page content to suit multiple output channels (Telco example: Web, Mobile Phone, IP STB/TBox, etc) from the same low-level content elements (content fragments)



# BBC APPROACH – WHY IS IT APPLICABLE

- **So why is this important?**
- The principles behind this are the ones at the foundation of the next phase of the internet, sometimes called the Semantic Web, sometimes called Web 3.0.
- The goal is to be able to more easily and accurately aggregate content, find it and share it across many sources. From these simple relationships and building blocks you can dynamically build up incredibly rich sites and navigation on any platform.
- There is also a change in editorial workflow for creating content and managing the site. This changes from publishing stories and index pages, to one where you publish content and check the suggested tags are correct. The index pages are published automatically. This process is what assures the highest quality output, but still saves large amounts of time in managing the site and makes it possible for us to efficiently run so many pages for the World Cup.
- The BBC had fantastic take-up by their Sport team, engaging with new tools and workflows.
- The site approach is being extended to support the London Olympics, where there will be over 12,000 athletes and index pages to manage and so without this type of technology, the BBC will not be able to showcase and maximise all available content
- **Finally – this approach is the model for many next generation web sites and applications, and can be applied to directories, catalogues, content search, and those needing increased automation and content value enhancement**

# Semantrix: “Intelligent Information Access”



The core philosophy behind Semantrix is the delivery of “Intelligent Information Access”:

- Intelligent – as we lever new approaches to hide technologies and legacy systems, and to provide the Information the business requires, not just the data the IT department makes available.
- Intelligent Information – not just data, as we provide data context as well so the business is accessing knowledge not just data - delivering exactly what you need.
- Intelligent Information Access – To support business users by extracting meaningful information, and providing opportunities to explore correlated information that suits different users, across multiple systems, in ways not previously possible.

Leveraging Semantic Web and Artificial Intelligence technologies and approaches, Semantrix can federate and make siloed application data accessible, provide tailored interfaces adapted to *your business requirements*, and with the intrinsic flexibility of Semantic technologies add new data sources and extend business access flexibly and at low cost compared to the typical Enterprise Business Intelligence approaches.

To ensure success Semantrix have partnered with the most advanced Semantic product and services companies in the field:

- **TopQuadrant** - to exclusively sell and support their TopBraid products, services, and training in Australia- including TopBraid Maestro, Enterprise, Live & Production
- **Ontotext** – to sell and support their products and services including BigOwlIM & GATE/KIM

## • Semantrix Business Services

- Product development and supply
- Project development, integration and deployment
- Specialist consulting services
- POC and Concept Demonstrators
- Product Support and Maintenance

## • Semantrix Engagement Model

- Product Supplier (3rd party and bespoke)
  - Represent TopQuadrant & other Semantic Web Technology companies in Australia
- Project Contract
- Consortium Projects
- Specialist Consulting Engagements

## Semantrix Products and Consulting Services

Strategic Information Planning, and Enterprise, Domain, & Solution Architecture consulting	Enterprise multi-media asset management and distribution systems
Semantic Technology, “Linked Data”, and Web 3.0 Architecture consulting & development	Social network driven personalisation , VMS/DAMS recommendation, EPG systems
High Performance, high Volume Web Site Architectural consulting and development	Light-weight Semantic Content and Metadata Federation and Integration approaches
Semantic application consulting and development – including Ontology development, content management, and product construction & delivery	Enhanced user search using metadata and faceted Semantic and Commercial Search Architecture and development approaches
Knowledge-Based Systems consulting including Rules based systems, Expert Systems, and A.I technology application development	Semantic and Wireless sensor technologies to support Health Informatics
Automation approaches to apply Semantic and A.I. technologies to Search Engine Optimisation (SEO) - including Good Relation ontology applications for products and services marketing	Autonomous systems software design and development: Semantic Models, Navigation systems, Agent & rule based control & planning

Would you like more information? Please Contact:

