

To: "Dr. Baruch Fischhoff - Chair, National Academy Committee on Improving Intelligence" <baruch@cmu.edu>,
From: Lloyd Etheredge <lloyd.etheredge@policyscience.net>

Subject: 160. Wolfram 4: The BBC's Leading-Edge Ontology/Indexing System - World Soccer & the London Olympics

Dear Dr. Fischhoff & Colleagues:

You might be interested in the enclosed overview slides from the BBC presentation [john.odonovan@bbc.co.uk] at the recent Wolfram Data Summit re data ontologies and high performance search/display/analysis.

The BBC - with enough financial support to be a global, open source, intelligence service - monitors more than 200 global TV channels and has about 12,000 journalists and contributors in the UK and 80 global locations. They also have some smart people assigned to store, index, retrieve, display, and analyze all of their data (including 1000 hours of their own TV broadcasting/week). The data are not simply text or speech but (now) video. And for display - for example, for each player in World Cup Soccer - they are pushing the envelope on graphic design of index pages.

One of the most interesting challenges - analogous, perhaps, to some needs of the DNI system - is World Cup Soccer and (now) preparing for the upcoming London Olympics. There are huge demands for rapid search/discovery, retrieval, analysis and display and people who are very sharp and serious about accuracy and the fast performance of the system.

Traditional BBC metadata systems have become deeply inadequate - for example, because they are made for different audiences/users. The BBC has developed a system of domain ontologies [based in part on behavioral/user research] for tagging data. BBC journalists/original content providers are trained to do the tags for their input using the ontology system.

I hope that this file (3MB+) comes through as an email attachment. I will place the file on my www.policyscience.net Website where it might be downloadable if the email attachment system does not work directly.

[Re another presentation about the generic need to develop ontology coding systems for searching and finding data (perhaps by different methods - e.g., Bayesian analysis along the lines of Google): The typical "unabridged" English language dictionary contains about 500,000 words although there now appear to be 2-3 million. Google has identified about ten million distinct strings that occur more than 200 times per trillion instances.] There is too much noise to just search huge databases without coding/retrieval ontologies. These are the kinds of generic problems that probably can be tackled more quickly by shifting as much of the work as possible into the public domain.

LE

Dr. Lloyd S. Etheredge

URL: www.policyscience.net

301-365-5241 (v); lloyd.etheredge@policyscience.net

lloyd.etheredge@aya.yale.edu (email)

[The Policy Sciences Center, Inc. is a public foundation that develops and integrates knowledge and practice to advance human dignity. Its headquarters are 127 Wall St., Room 322 PO Box 208215 in New Haven, CT 06520-8215. It may be contacted at the office of its Chair, Michael Reisman (michael.reisman@yale.edu), 203-432-1993. Further information about the Policy Sciences Center and its projects, Society, and journal is available at www.policysciences.org.]

Future of BBC Content and Data

...or how metadata came to rule the world

John O'Donovan, Chief Architect, BBC News, Sport and Knowledge

Our goal...

“Moving from a page publishing model to a content publishing model where metadata and asset management is used to render the appropriate content to different target platforms for the lowest possible cost ”

The BBC has a lot of data...

The BBC provides a host of services but some of the things we are...

- Over 20,000 people
- Over 12,000 journalists and core contributors across 60 UK and 80 Global locations
- 8 UK and global TV channels, up to 16 interactive TV channels
- Over 1000 hours of TV broadcast every week
- Over 60 UK and global radio channels
- Hundreds of websites in over 40 languages
- Monitors over 200 global TV channels and thousands of sources
- Both Public Service and Commercial arms

Making content creates a lot of data...



This is one important image from one frame of video from one match in one World Cup

Generally, each item of content generates between 10 and 100 items of discrete data

And we make millions of items of content every day

How much content and data are we talking about...

Millions of news stories, items of sport data and business data in multiple data centres

100 km of shelves for physical storage

1.5M items of film and video

750,000 hours audio

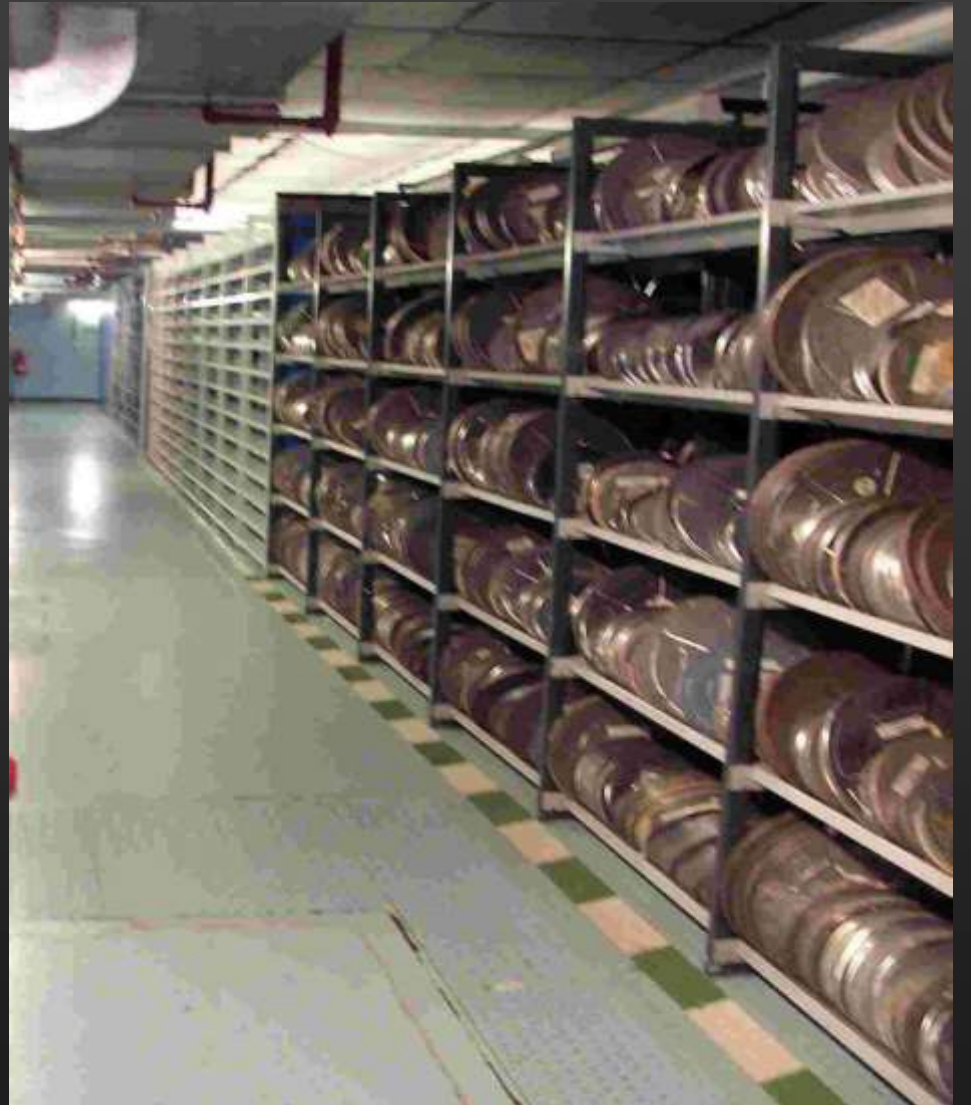
22M newspaper cuttings

3M photographs of which only $\frac{1}{4}$ online

4M items sheet music

500k phonetic “pronunciations”

1.5M titles in “grams library” – vinyl and CDs



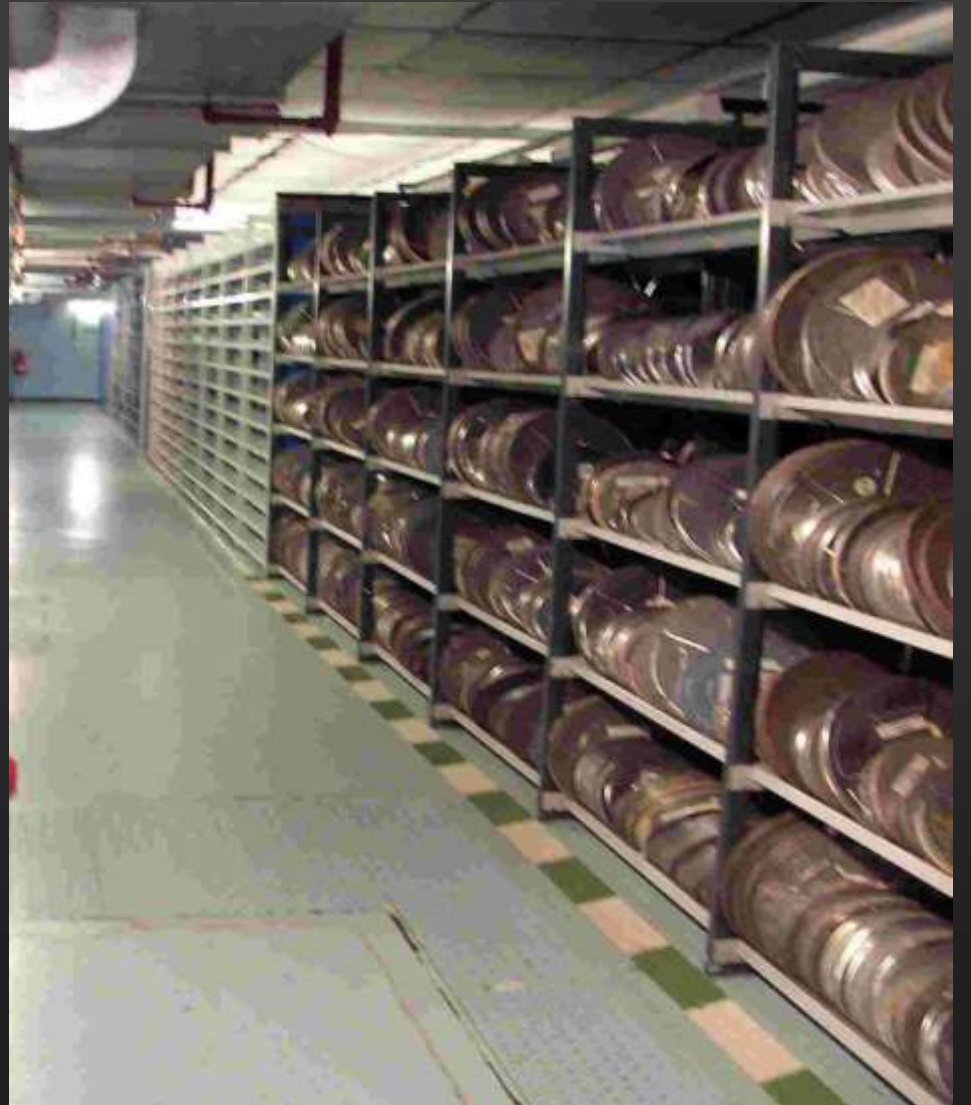
How much content and data are we talking about...

Users are mostly internal to the BBC:

Content producers	70%
News	20%
Commercial exploitation	6%
Others	4%

Clearly, this feels imbalanced given the breadth and depth of content and data...

The key question is how do you make content discoverable and how do you present it?



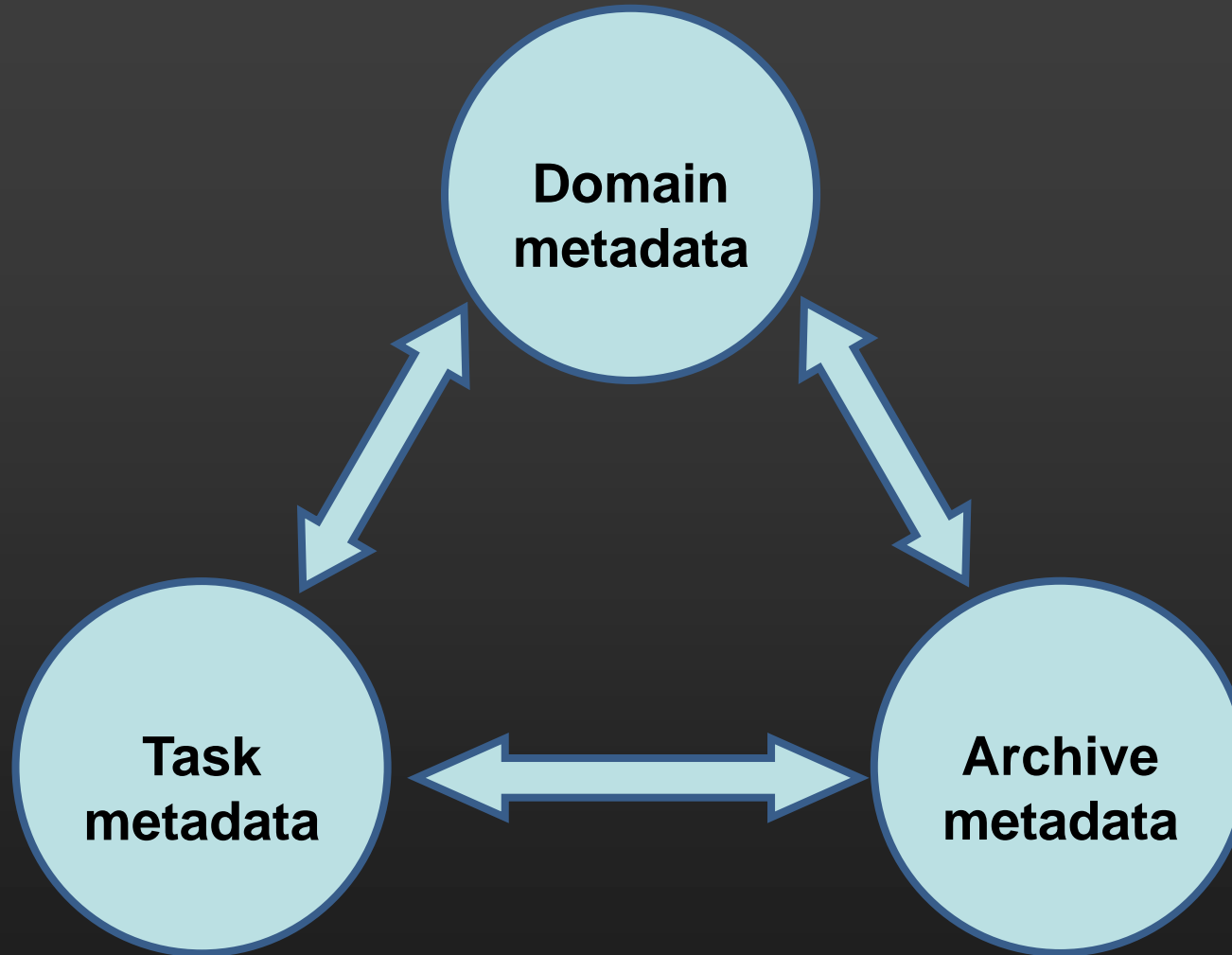
Information about content is inadequate...

Tape No. BRD1234567		Length 120 mins		Prog. No. LSF G554W/72		Date 13/06/2004		RMS Booking No. 250021239	
Series Title: THE SINGLE-PROGRAMME VTRR								M/c No: 123456	
Core Title: This Is The Main One We Will Use								Source: BRD525252	
								Tech code: 6CBDSO	
								Aspect ratio: 16F16	
Time code	PC	Tape Contents						Stereo <input checked="" type="checkbox"/> Mono <input type="checkbox"/>	
09:57:00:00		line up						A1	F/Mix
09:59:30:00		clock						A2	
10:00:00:00	72	programme						A3	M/E
10:55:00:00		clean backgrounds						A4	
: : :								Cue	
: : :								Tape tracked, Yes <input type="checkbox"/>	
Incoming programme/editing report (Tape numbers, source Facility House and notes for Reviewer) ex Bankside TV. Astons and Generic titles added. Some poor quality material ex CCTV cameras.									
Name Ace Editor									
Sound dub to Tape #				In: Date		Sound dub to tracks 1&2		In Th 5	
Name				Date		Name Dubbing Dude		Date 15/06/04	
Tape status		O <input type="checkbox"/>	M <input type="checkbox"/>	T <input checked="" type="checkbox"/>	D <input type="checkbox"/>	S <input type="checkbox"/>	U <input type="checkbox"/>		
Review Suite: rs2 Technical review remarks minor breakup on cctv material at 10:23:45-10:23:59:00. Strobing at 10:56:00:00-10:57:00:00 due to flash photography causing Gordon alarm- prod dealing.									
Please enter exact timecodes below as they will be used for ingest for TX. 'In' should be first <i>frame</i> of S&V for TX. 'Out' is last <i>frame</i> of S&V for TX (3 second hold on the last credit or until end of music is normal)									
In: 10:00:00:00				Out: 10:59:25:00			Dur: 59 Min 25 Sec		
Passed for TX <input checked="" type="checkbox"/>				Name Eagle Eye			Date 17/06/2004		
SPECIAL TX AUTHORISATION: strobing agreed by Maria West commissioning ed x67654									
Programme re-edited since review Details:									
Name					Date				

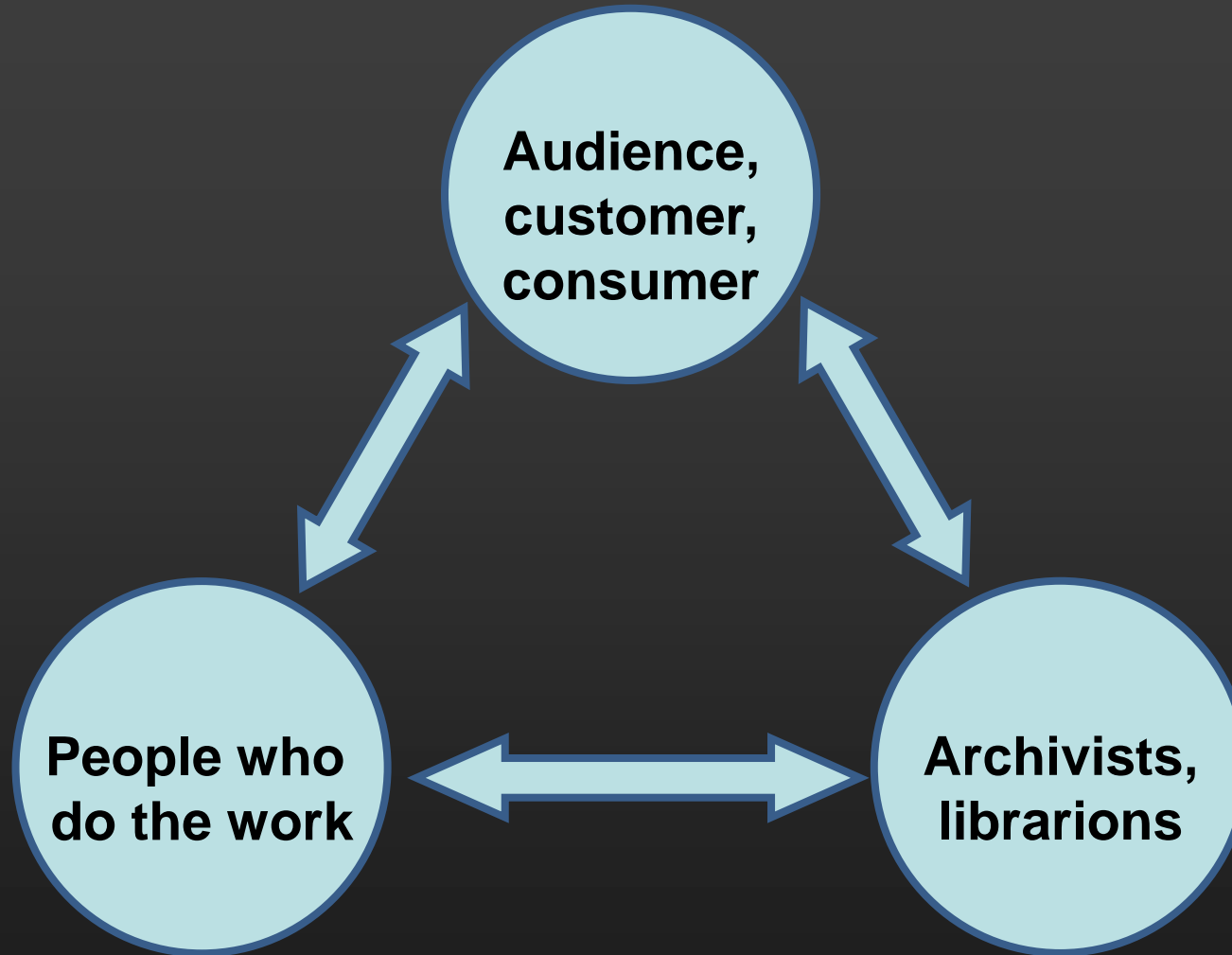
Which is why our people are not happy...



The biggest mistake...data is made for different audiences



Which means you manage your data at least three times...



Task based data... "I know what that means..."



Archive Data... "That has been correctly categorised"

Category LONPROG

Catalogue number 9674938

Item title ELECTION 2005:PUBLIC OPINION **Medium** VT

Programme/Item ITEM

Duration 0:12:07 **Time in** 0:38:52

Frank LUNTZ public opinion analyst, reports from Milton Keynes with the views of a small group of undecided swing voters & analyses what will attract them in tomorrow's election.

> Sp s vt, intvs, ex lib INTVS: Peter HYMAN strategist to the Prime Minister 1997-2003 re: it has to be about the economy (41:41-42:08) Danny FINKELSTEIN Conservative strategist: re: it's got to be action not words (45:33-46:07) Olly GRENDER former Lib Dems communications director: re focus group & Lib Dems hearing what they want to hear (48:32-48:54)> Sp s vt, intvs, ex lib DEPT ACTY: MILTON KEYNES GALLERY: Frank LUNTZ listening to swing voters (38:52-39:23) Frank LUNTZ with panel re reactions to Tony BLAIR: "insincere", "un-opinionated", "liar", "untrustworthy". (39:40-41:20) Panel reaction to Michael HOWARD "a bit of a wimp, not very confident", "evil", "he's an imposition", "don't trust him" (46:07-48:11) Panel Reaction to Charles KENNEDY: "He's sincere enough", "undecided", "a loser" (48:54-50:15)

Domain data... "How do I present this well"

Research.Gateway - Microsoft Internet Explorer provided by the BBC

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites

Address http://grain03.radio.bbc.co.uk/grain/InfaxSearchAgain.do?searchOption=InfaxQuickSearch

Gateway home

research.gateway Desktop access to research resources exclusively for the BBC. Tel: 02 02266 Email: research.gateway

HOME BOOKS CELEBRITY INFO EVENTS & ANNIVERSARIES MUSIC NEWSPAPER CUTTINGS PRONUNCIATION REFERENCE COLLECTION STILLS TV & RADIO

Top 20 Research Sites

TV & Radio Simple search

Search for archive material from BBC TV and Radio programmes.

Search options

- Simple search
- Advanced search
- TV Subject search

Programme title newsnight

Free text

Search Reset

Dates

Any date

Specific date 04/05/2005

From To

Search category selector

- Search all TV and radio categories
- Search my categories
- Search categories from the list below

Categories to be searched

- TV
- RADIO
- National TV news bulletins
- TV programmes

BBC Home News Sport Weather iPlayer TV Radio More Search

NEWS

LIVE BBC NEWS CHANNEL

Page last updated at 09:15 GMT, Friday, 14 May 2010 10:15 UK

News Front Page

BANGKOK CLASHES: Two journalists reported shot during confrontation in Thai capital

Cameron to finish naming his team

David Cameron is expected to put the finishing touches to his ministerial team on Friday, before going to Scotland for talks.

- At-a-glance: Coalition policies
- Who's who: Cameron's cabinet
- Cameron in Scotland for talks
- Hague in US for Afghan talks
- Web revolution in Whitehall

Inside Cameron's cabinet

UK COALITION: FULL COVERAGE

- Adonis condemns dissolution block
- Brown to remain as back-bench MP
- Samantha Cameron quits company post

Blog standard

Bloggers discuss the role they played in the general election

OTHER TOP STORIES

- Students could pay more for loans
- Somali pirates release UK tanker
- Report reveals police DNA failure
- Three held in New York bomb plot
- US 'will not destroy Kandahar'
- Airport parking fees 'ridiculous'
- Crackdown on rural road speeding
- Memorial to Bomber Command heroes

ALSO IN THE NEWS

- Graham Norton takes Jonathan Ross's Radio slot
- TV talent show winners beat Pope to Classical Brit

VIDEO NEWS

- Wallaby lawnmower - it's Odd Box
- UK-born astronaut on shuttle's end

FEATURES, VIEWS, ANALYSIS

- About a Boy: Teenage actor gets under Boy George's skin in new drama
- Football numbers: How the figures add up for World Cup bids
- Keep the love: How long do honeymoons tend to last?

WORLD CUP 2010 OUR COVERAGE STARTS HERE...

SPORT HEADLINES

- Beckham hands over 2010 Cup bid
- Classy England cruise into final
- Capello considering new formation

ADD YOUR LOCAL NEWS, WEATHER AND SPORT HERE

HAVE YOUR SAY

The challenge...

- **Make more content for more devices and platforms**
- **Apply precise metadata to it...**
- **...so that it can be presented well**
- **And all for less money than we used to have**

To address this first it's necessary to understand a little about describing content...

Headlines describe content...sometimes not that well

'We hate math,' say 4 in 10
— a majority of Americans

WASHINGTON
relationship w

Committee appoints committee to appoint committee

By Meghann Ackerman
STAFF WRITER

Although the final decision is nearly a year off, the Newton School Committee is starting the process of hiring a replacement for Superintendent Jeff Young.

At a meeting last week, the School Committee agreed to hire a consultant to help in the nationwide search and to form a search committee made up of school and community representatives to

Several School Committee members, including Geoff Epstein, said getting the word out about the committee would be the best way to attract people with expertise in many areas.



Volume 50 Number 27

June 28, 1996

Locally Owned

Gooding County's Community Connection

Printed On Recycled Paper

Poverty meeting attracts poor turnout

By Mary Ann Blagoe
Gooding County Leader

It is not so surprising to find nobody came.

At least that was the theme of a poverty-related meeting about

and poverty rate in the Magic Valley. The region also ranked

third highest in poverty in the state.

Southern Idaho Center for New Directions, and Beulah Nordrup

from the South-Central Community Action Agency in

summarize. "We really accomplished no

that we think we did." The event attracted about 19

Headlines can be misleading...

When Inverness Caledonian Thistle FC beat Celtic FC in the Scottish Cup 3-1 in February 2000...



The Sun reported the story with the headline "Super Caley go ballistic, Celtic are atrocious".

And Mary Poppins turned in her grave...

And without context, content is difficult to understand...

Lesbian albatrosses to raise chick

And without context, content is difficult to understand...

Lesbian albatrosses to raise chick

Two female royal albatrosses at a New Zealand breeding colony have successfully incubated a chick.

And without context, content is difficult to understand...

Telegraph.co.uk

Home News Sport Finance Lifestyle Comment Travel Culture Technology
UK World UK Politics Celebrities Obituaries Weird Earth Science Health News Edu
Earth News Copenhagen Climate Change Conference Environment Wildlife Outdoors

HOME > EARTH > WILDLIFE

Lesbian albatrosses to raise chick

Two female royal albatrosses at a New Zealand breeding colony have successfully incubated a chick.

By Ben Leach
Published: 8:02AM GMT 03 Feb 2010



There are about 140 royal albatrosses on the colony. This season 17 chicks have hatched from 17 fertile eggs, a rare 100 per cent success rate. Photo: ALAMY

The father, one of several males at the Taiaroa Head Royal Albatross Centre on the South Island's Otago Peninsula, appears to have disappeared, according to the centre managers.

"It's quite unusual in the albatross population here at Taiaroa Head to have two

Share | f | | |
291 diggs digg it
88 retweet
Email | Print
Text Size + -
Wildlife
News
UK News
How about that?
Ben Leach
Animals of the week
IN PICS

And without context, content is difficult to understand...

Great tits cope well with warming

And without context, content is difficult to understand...

The screenshot shows the BBC News website interface. At the top left is the BBC logo with 'Low graphics' and 'Help' links. A search bar and 'Explore the BBC' button are on the right. The main header features the word 'NEWS' and a 'LIVE BBC NEWS AT ONE' indicator. A navigation menu on the left lists various news categories, with 'Science & Environment' highlighted. The main content area displays a news article titled 'Great tits cope well with warming' by Richard Black, an environment correspondent. The article text states that great tits are laying eggs earlier in the spring due to climate change. A photograph of a bird's beak is included. To the right, there are sidebars for 'CLIMATE CHANGE' with an animated guide link and 'LATEST SCIENCE' with several news items.

BBC Low graphics Help

Search Explore the BBC

NEWS

LIVE BBC NEWS AT ONE

Page last updated at 18:04 GMT, Thursday, 8 May 2008 19:04 UK
[E-mail this to a friend](#) [Printable version](#)


Great tits cope well with warming

By Richard Black
Environment correspondent, BBC News website

At least one of Britain's birds appears to be coping well as climate change alters the availability of a key food.


Researchers found that great tits are laying eggs earlier in the spring than they used to, keeping step with the earlier emergence of caterpillars.

Writing in the journal Science,



Food for hungry mouths

CLIMATE CHANGE

 [Animated guide: Find out how the greenhouse effect works and more...](#)

LATEST SCIENCE

- ▶ [Pause in Arctic's melting trend](#)
- ▶ [Four degrees of warming 'likely'](#)
- ▶ [UK 'must plan' for warmer future](#)

GLOBAL POLITICS

- ▶ [UN climate talks split on treaty](#)

Pros and cons of the current approach...

Using the power of the following tools to organise content

- **Semantic Concept Extraction with Unique Identifiers**
- **Linked Data**
- **Content Tagging**
- **Ontology to organise content**

Where did Linked Data start...

CERN DD/OC

Information Management: A Proposal

Vague but exciting ...

Tim Berners-Lee, CERN/DD

March 1989

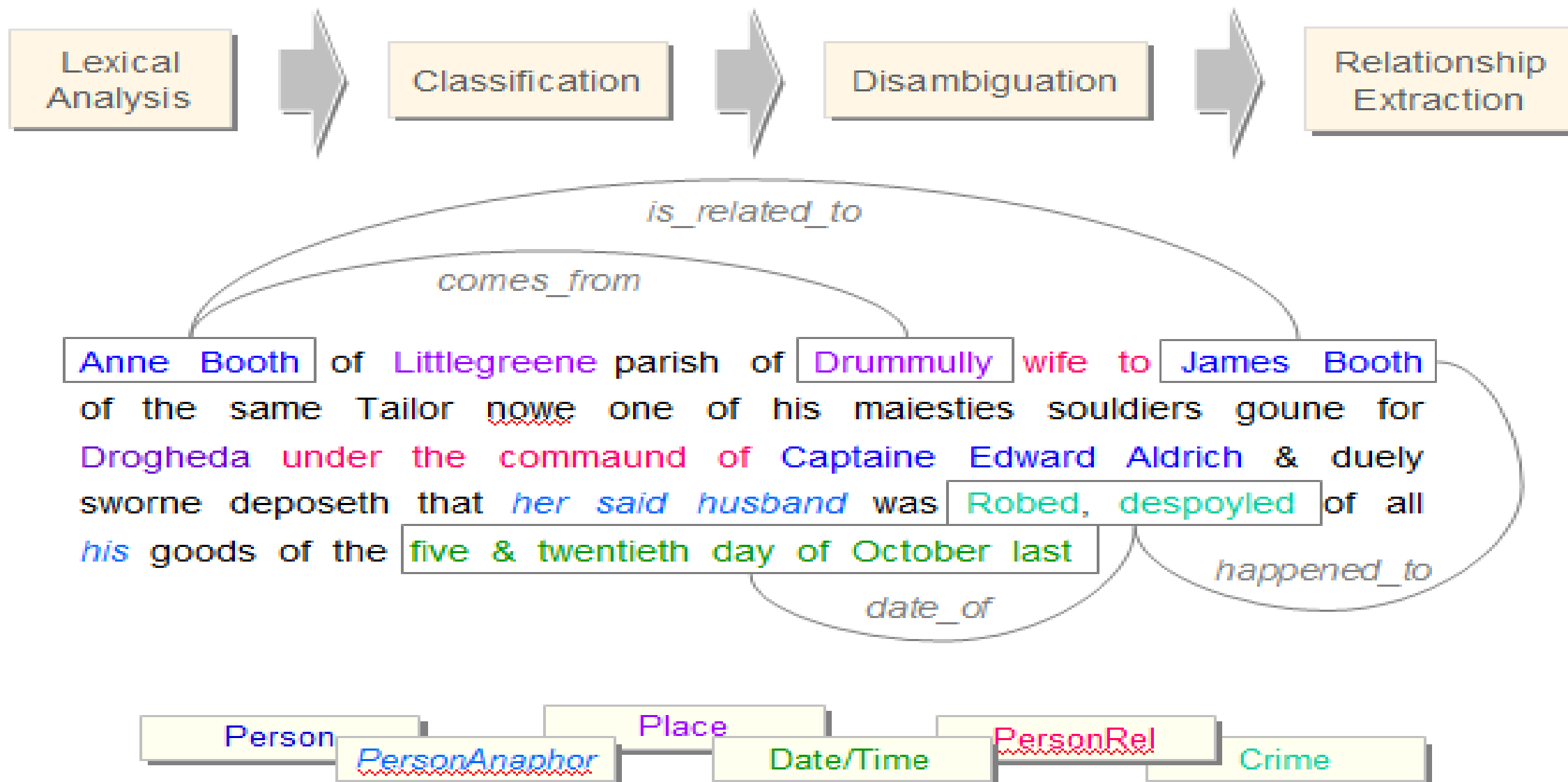
Information Management: A Proposal

Abstract

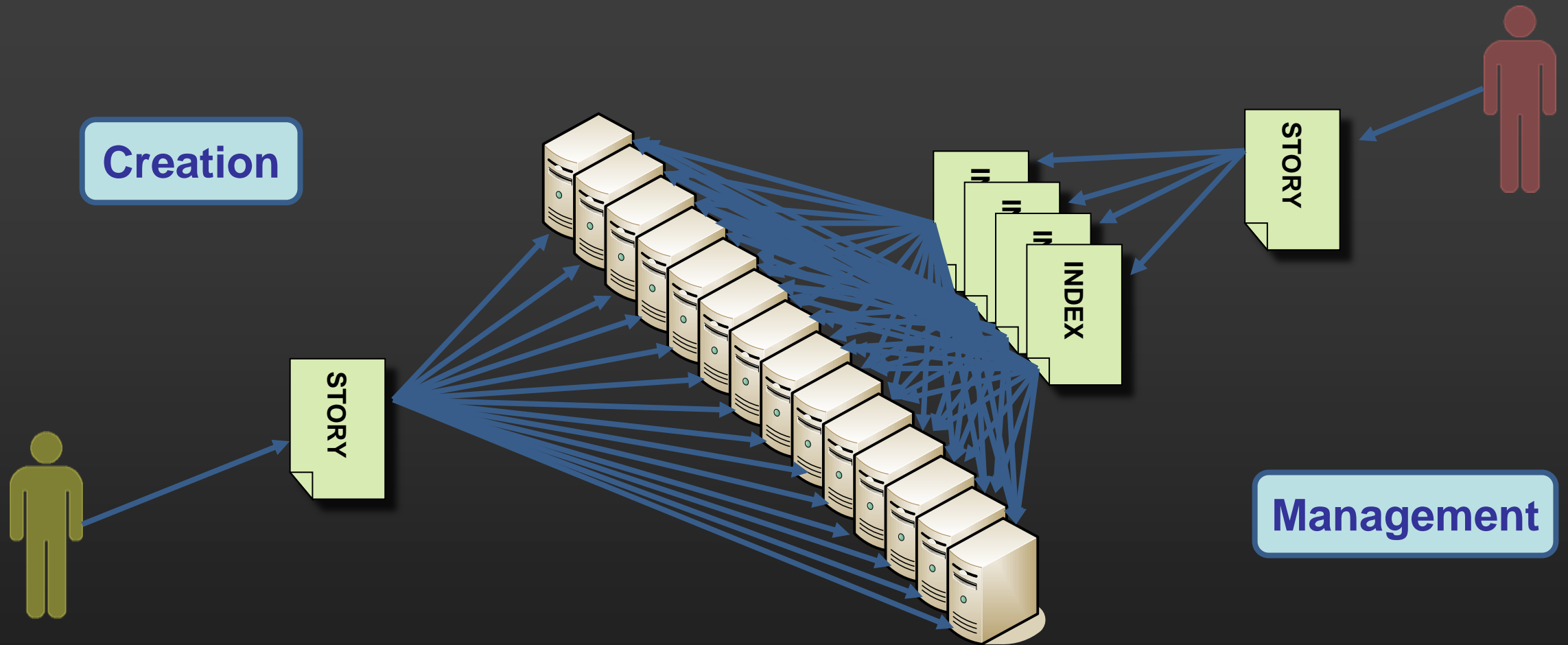
This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

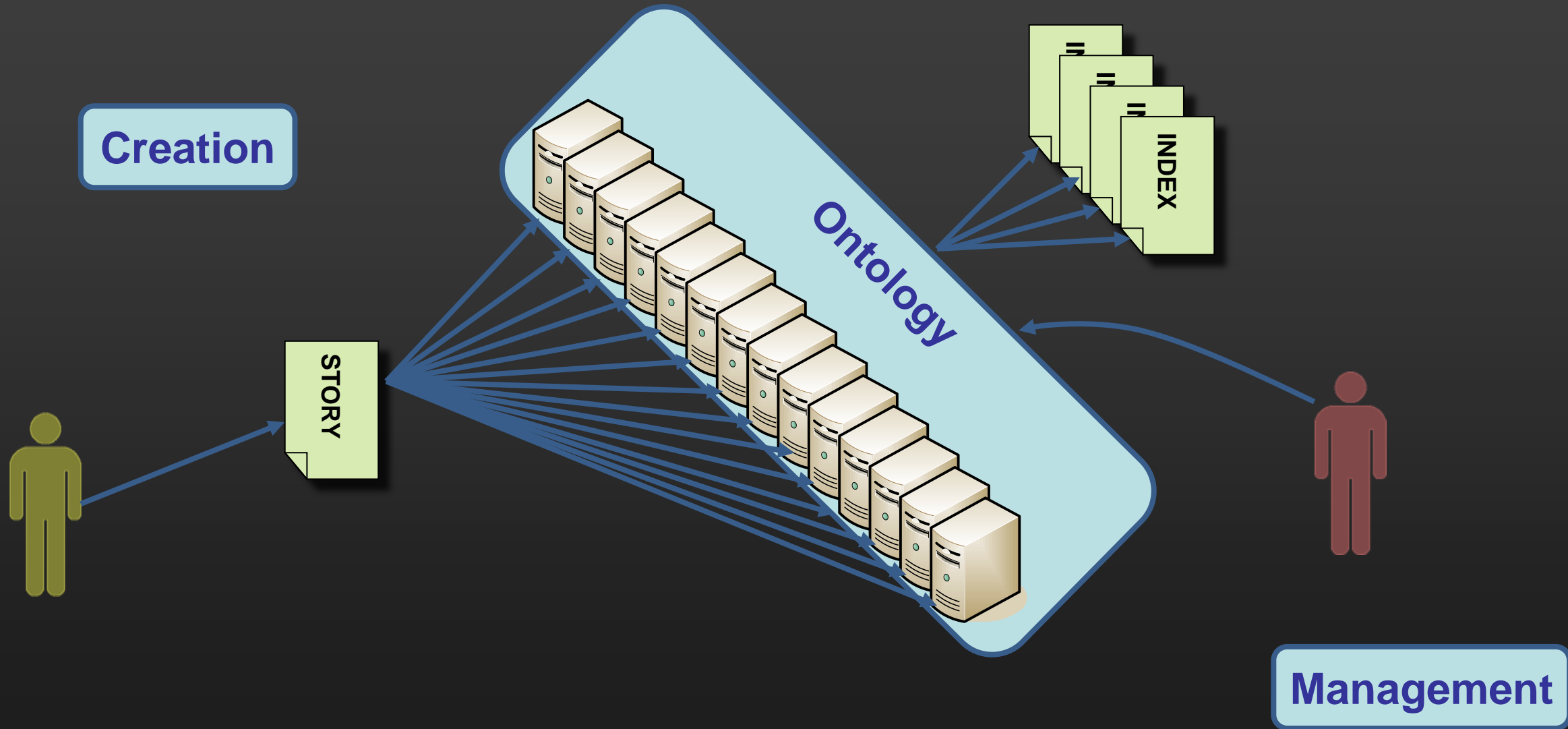
Why do you need an Ontology...what does it do



What do we do now...



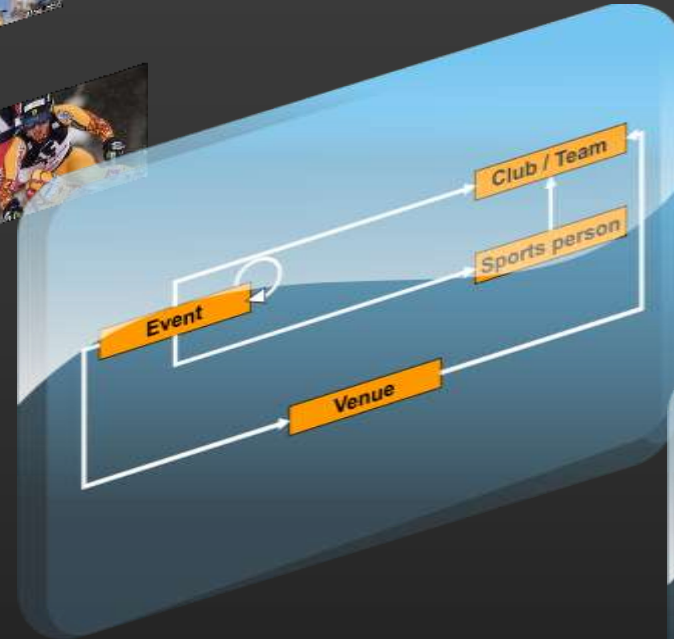
How does this help...



Fitting the pieces together...where we are heading



Content Objects

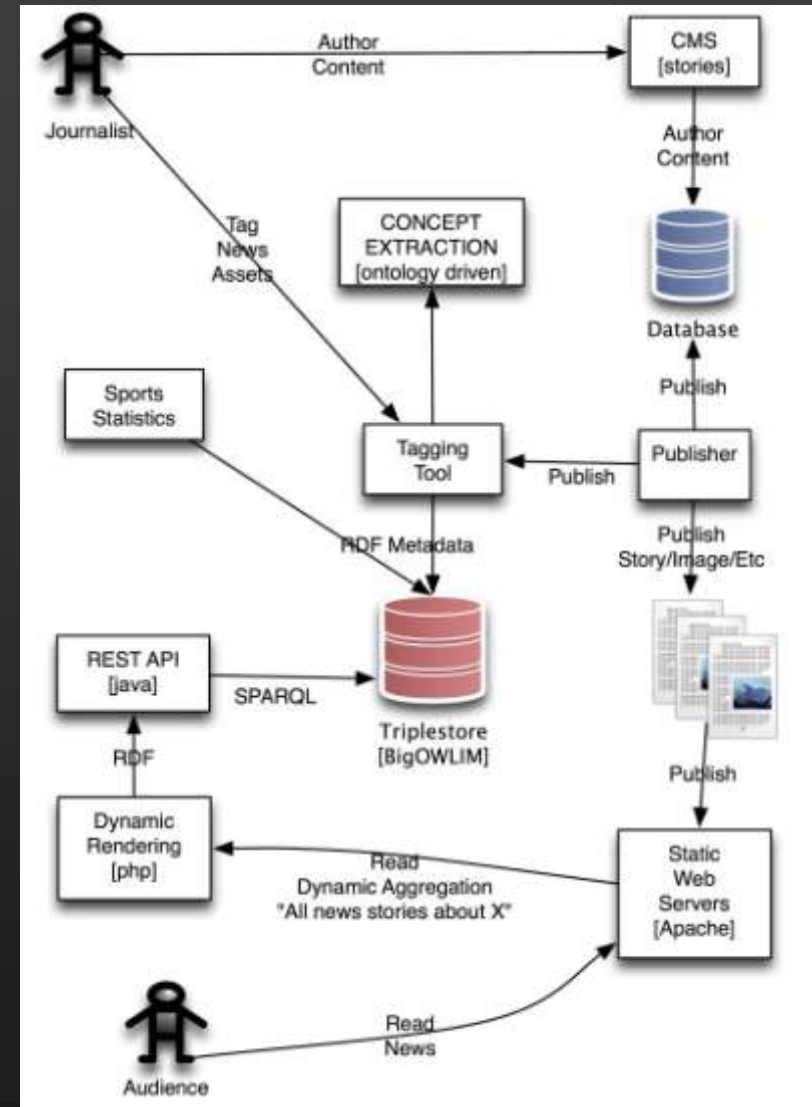
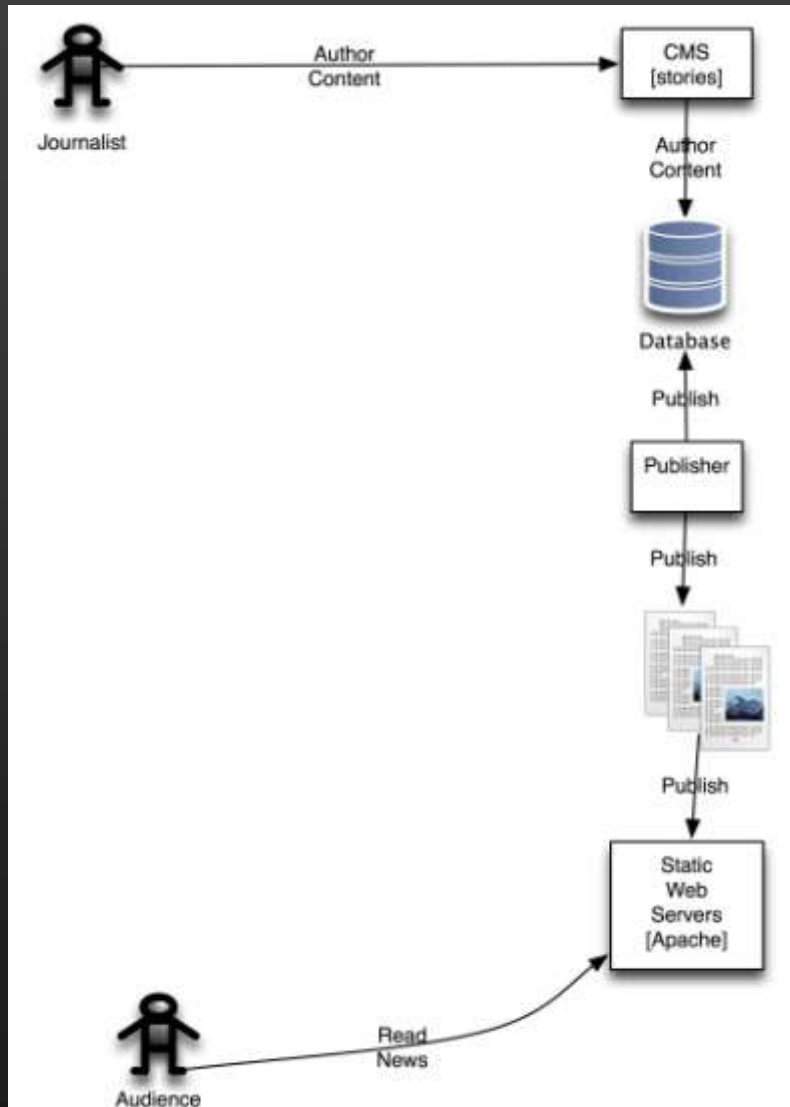


Tagging through ontology

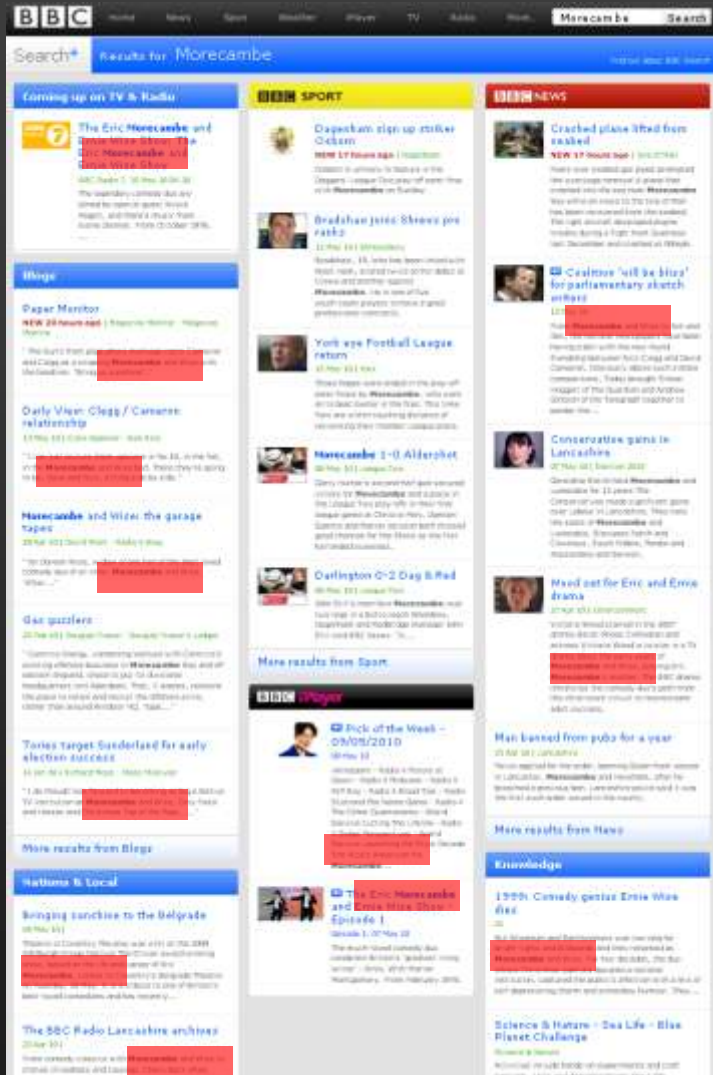


User Experience

What does this look like architecturally...



Why not just use search?



Example page for a place called Morecambe in the UK...

It's schizophrenic...

Most aggregations built this way suffer from disambiguation

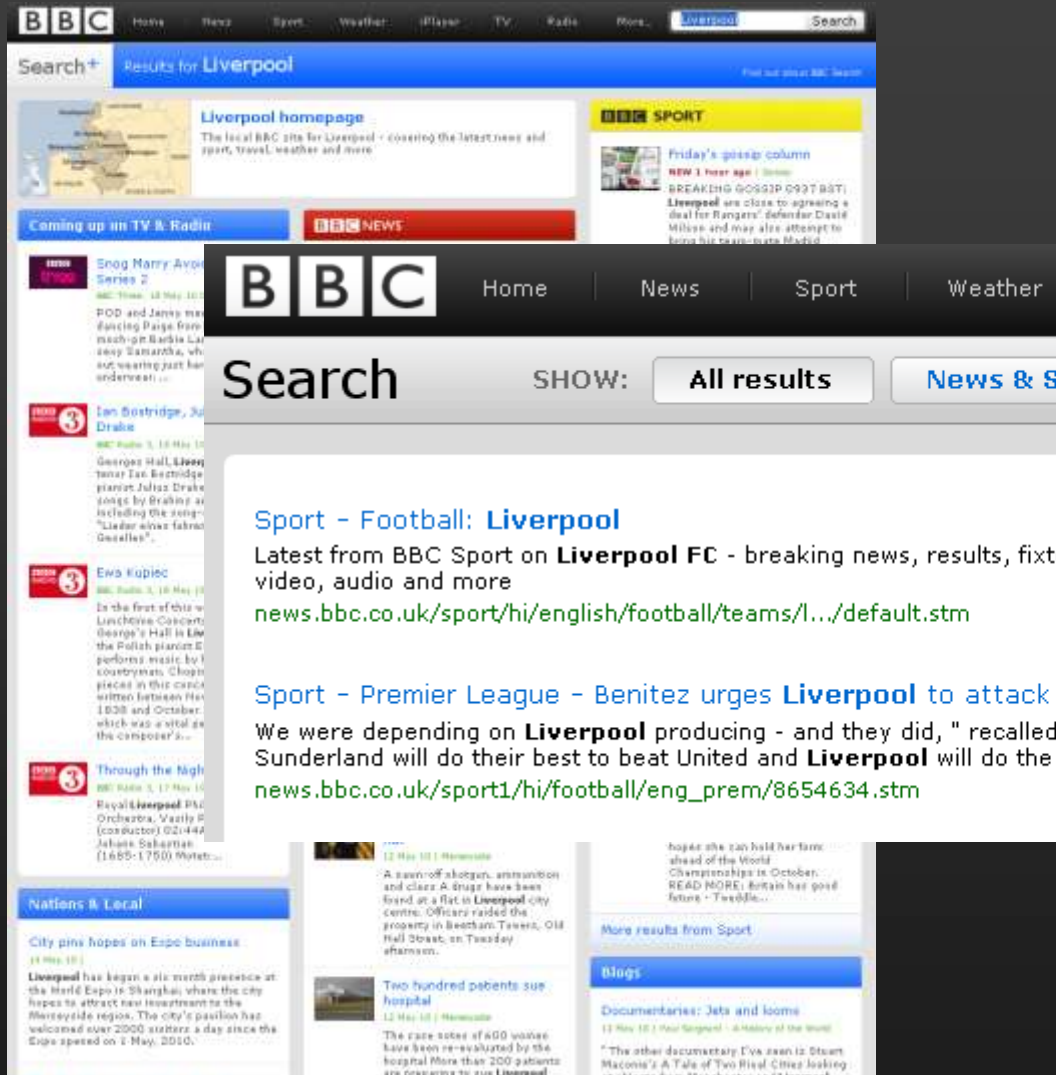
Search also does not know which area of the country Morecambe is in as it has no knowledge

This means it doesn't know to include stories near Morecambe

No idea what it is linking to and whether it is relevant – content may mention Morecambe but is it *about* Morecambe

Feedback has been clear – our audience do not tolerate errors in this type of product

Search doesn't know how things are related...



The screenshot shows the BBC search results for 'Liverpool'. The search bar at the top contains 'Liverpool' and the search button is labeled 'Search'. Below the search bar, there are several search results. The first result is 'Liverpool homepage' with a sub-heading 'The list of BBC sites for Liverpool - covering the latest news and sport, travel, weather and more'. Below this, there are more search results, including 'Friday's gossip column', 'Enag Marry Ayo Series 2', 'Ian Bostridge, Julia Drake', 'Ewa Kopiec', and 'Through the High Royal Liverpool'. The search results are displayed in a grid format with various images and text snippets.



The screenshot shows the iPlayer search results for 'Liverpool'. The search bar at the top contains 'Liverpool' and the search button is labeled 'Search'. Below the search bar, there are several search results. The first result is 'Sport' with a sub-heading 'TV HIGHLIGHTS'. Below this, there are more search results, including 'Formula 1', 'The Football League', 'Formula 1', and 'Formula 1'. The search results are displayed in a grid format with various images and text snippets.

How do you tag things... “Graffiti” as part of the process

The screenshot shows the 'Graffiti' web application interface. At the top, there are navigation buttons for 'CPS', 'Blogs', 'Admin', and 'Help', along with a 'System status: OK' indicator. The user is logged in as 'nationalisaact01'. The main content area displays a blog post titled 'Green and Barry on road to recovery' by Phil McNulty, dated Mon 14 Jun 2010. Below the post is a 'Tag story' section with a 'Select main topics' instruction and a list of suggested topics including England, United States, Algeria, Robert Green, Gareth Barry, and Joe Hart. A 'This story is mostly about:' section shows selected tags for England, Robert Green, and Gareth Barry. A 'Tagging progress' bar is visible at the bottom of the main content area. On the right side, there is a profile for Gareth Barry, featuring a photo of him on a football pitch and a table of personal and club information.

Currently viewing: Phil McNulty: Green and Barry on road to recovery (<http://www.bbc.co.uk/blogs/philmcnulty/2010>)

Green and Barry on road to recovery

Posted **Mon 14 Jun 2010, 1:32 PM** by **Phil McNulty** in **Phil McNulty**

World Cup 2010: Rustenburg Robert Green's work was not being done under the icy gaze of an expectant nation - but he was literally the focus of attention as England got back to World Cup business at their Rustenberg base....

Tag story > Select main topics

Select the topics that this story is primarily focused on. Topics could be anything from a person or organization to a broad subject such as 'women in the workplace'.

Suggested topics:

- England
- United States
- Algeria
- Robert Green
- Gareth Barry
- Joe Hart


This story is mostly about:

- England
- Robert Green
- Gareth Barry

<< Back Next >>

Tagging progress:

Gareth Barry



Personal information

Full name	Gareth Barry
Date of birth	23 February 1981 (age 29)
Place of birth	Hastings , England
Height	1.83 m (6 ft 0 in)
Playing position	Midfielder

Club information

Current club	Manchester City
--------------	---------------------------------

How do you tag things... "Graffiti" as part of the process

The screenshot displays the 'Graffiti' application interface, which is used for tagging content. The top navigation bar includes 'CPS', 'Blogs', 'Admin', 'Help', and 'System status: OK'. The user is logged in as 'national\visaact01'. The main content area shows a news article titled 'Man killed as four lorries crash' from BBC News - England, dated August 18, 2010. Below the article, there is a section for tagging locations. A search box contains the text 'milton', and a list of suggested locations is shown, including 'Milton Keynes, Borough of Milton Keynes, GB'. To the right, a map of the United Kingdom is displayed with a red pin marking the location of Milton Keynes. The interface also includes a 'Tagging progress' bar at the bottom.

Currently viewing: [BBC News - England: Man killed as four lorries crash \(http://www.bbc.co.uk/go/rss/int/news/-/n...](#)

Man killed as four lorries crash

Posted **Wed 18 Aug 2010, 11:31 AM** by **unknown author** in **BBC News - England**

A man dies in a crash between four lorries, leading to the closure of the M1 and 12-mile tailbacks.

Tag story > Select locations
Select the locations that appear in this story.

Suggested locations

Location search

milton

- Milton Keynes, Borough of Milton Keynes, GB
- Milton Ernest, GB
- Milton Regis, County of Kent, GB
- Milton Damerel, County of Devon, GB
- Milton, County of Oxfordshire, GB
- Milton of Murtle, Aberdeen City, GB

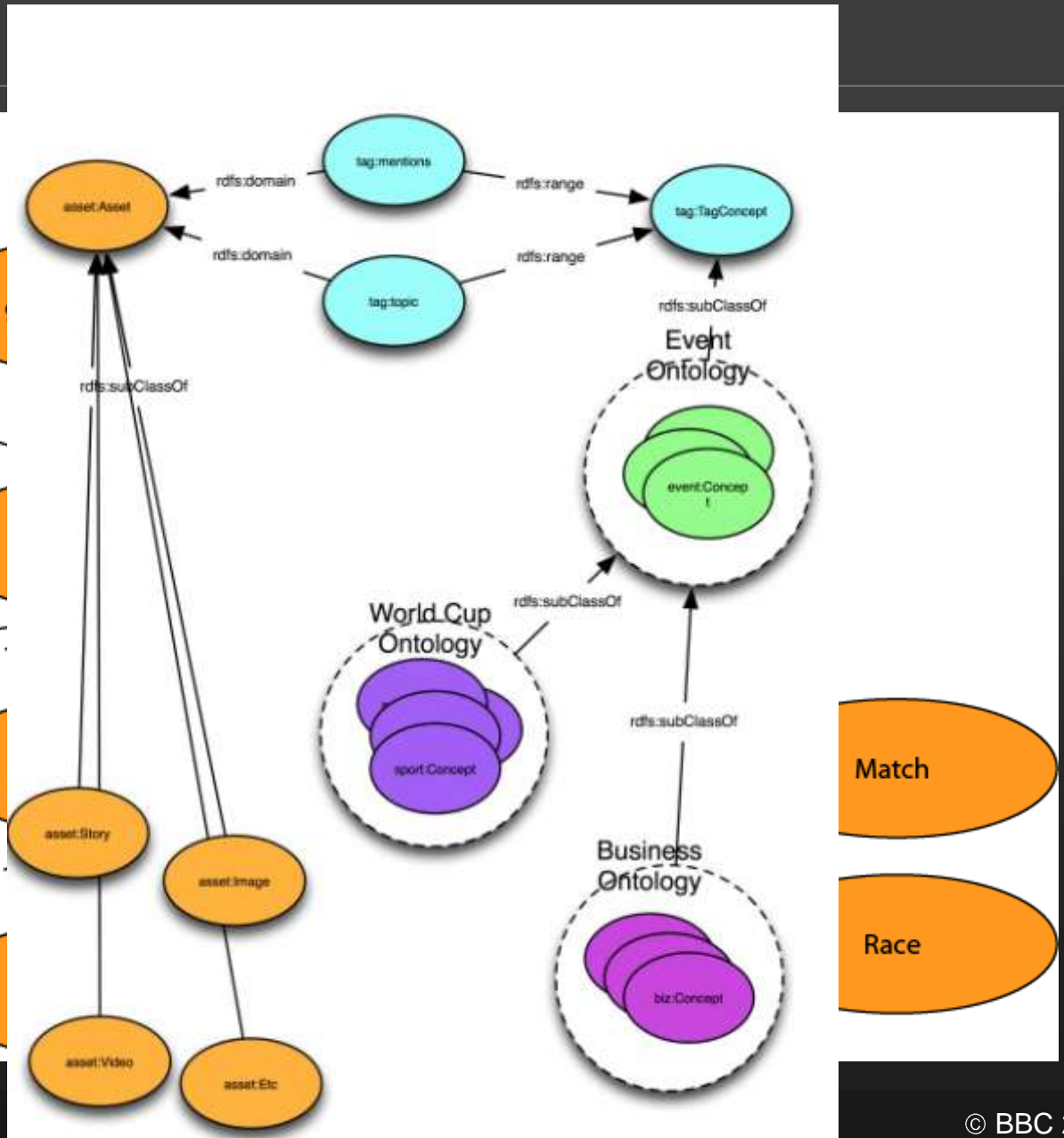
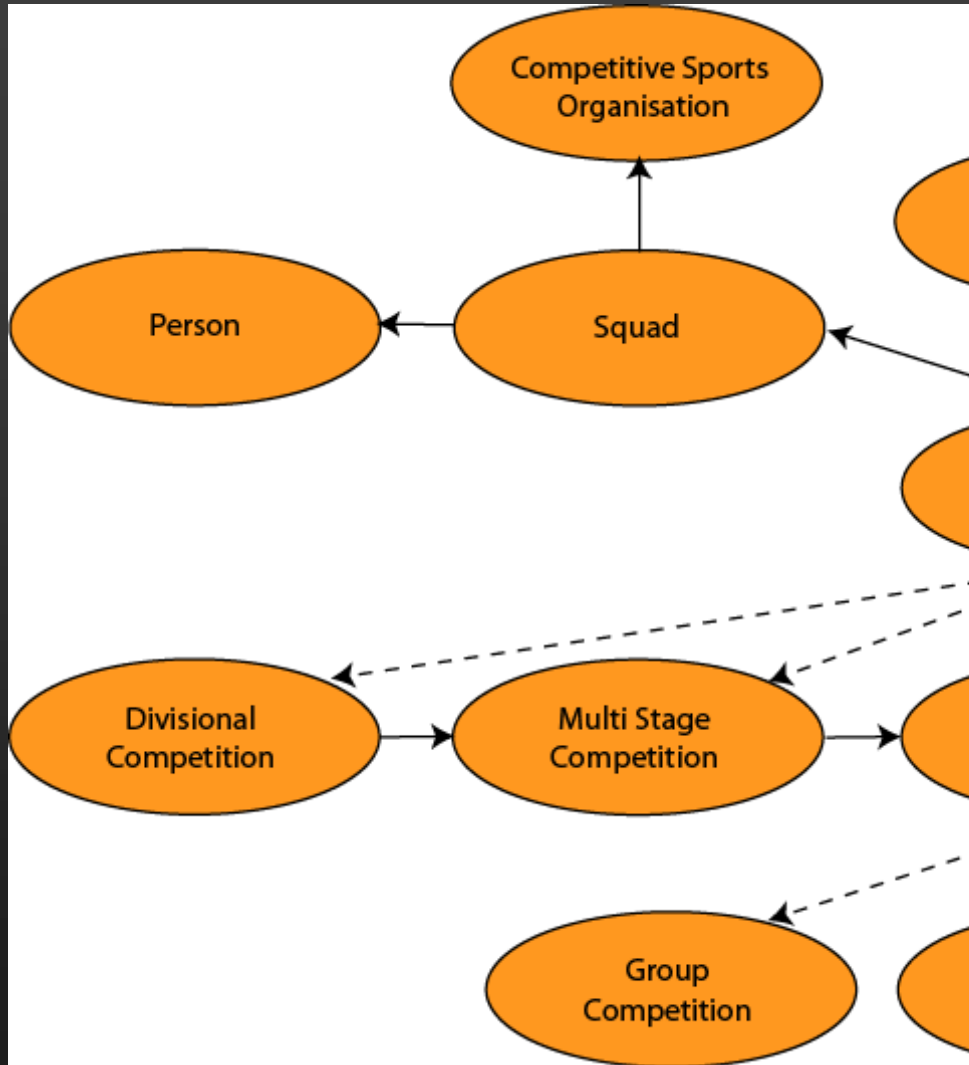
These locations appear in the story:

- Milton Keynes, Borough of Milton Keynes, GB

<< Back Next >>

Tagging progress

What is an ontology anyway...



Using an ontology...

- **Rich expandable ontological knowledge model. Not easy to model in a generic fashion using a relational model**
- **OWL reasoning, forward chaining and inference**
- **Open models/ontologies re-use and publishing**
- **Simplified query language. SPARQL vs SQL**
- **Content sharing and linked open data**

Metadata quality makes the best of what you have automatically...

The screenshot shows the BBC Sport website for the 2010 World Cup. The top navigation bar includes links for Home, News, Sport, Weather, iPlayer, TV, Radio, and More... A search bar is located on the right. Below the navigation is a large banner for 'SPORT WORLD CUP 2010' with a decorative trail of flags and the World Cup trophy. A secondary navigation bar includes links for SPORT, FOOTBALL, WORLD CUP 2010, GROUPS & TEAMS, FIXTURES & RESULTS, VIDEO, and BBC COVERAGE.

Latest matches

- NED 2-0 DEN**
Highlights & report
- JPN 1-0 CMR**
Highlights & report
- ITA 1-1 PAR**
Highlights & report

Italy

- Italy 1-1 Paraguay (Monday, 14 June) - Match report
- Italy v New Zealand (Sunday, 20 June, 15:00) - Preview
- Slovakia v Italy (Thursday, 24 June, 15:00) - Preview

Latest stories

- Buffon confident after back pain
- Paraguay boss delighted with draw
- Gattuso to retire after World Cup
- Italy's Pirlo to miss two matches
- Sunday's World Cup 2010 round-up
- Rossi left out of Italy Cup squad
- Tuesday's World Cup 2010 round-up
- Prandelli to be next Italy boss
- Lippi in post-World Cup exit hint

Group F Teams

	A	B	C	D	E	F	G	H
Group F Teams	W	D	L	GD	PTS			
Italy	0	1	0	0	1			
Paraguay	0	1	0	0	1			
New Zealand	0	0	0	0	0			
Slovakia	0	0	0	0	0			

Features

- World Cup 2010 performance analysis: All the facts and figures from the World Cup 2010 tournament
- My World Cup - Noel Gallagher
- Guide to the teams at South Africa 2010

Metadata quality makes the best of what you have automatically...

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

Latest matches

NED 2-0 DEN

[Highlights & report](#)

JPN 1-0 CMR

[Highlights & report](#)

ITA 1-1 PAR

[Highlights & report](#)

NZL v SVK

LIVE BUILD-UP

Group F

TEAMS	P	W	D	L	F	A	GD	PTS
Italy	1	0	1	0	1	1	0	1
Paraguay	1	0	1	0	1	1	0	1
New Zealand	0	0	0	0	0	0	0	0
Slovakia	0	0	0	0	0	0	0	0

Latest stories

- Buffon confident after back pain**
 - ▶ Gattuso to retire after World Cup
 - ▶ Skrtel hopeful for opening game
 - ▶ Rossi left out of Italy Cup squad
 - ▶ Skrtel makes final Slovakia squad
- Paraguay boss delighted with draw**
 - ▶ Italy's Pirlo to miss two matches
 - ▶ Sunday's World Cup 2010 round-up
 - ▶ Tuesday's World Cup 2010 round-up
 - ▶ Cameroon secure Slovakia draw

Matches

- Italy 1-1 Paraguay** Monday, 14 June [Match report](#)
- New Zealand - Slovakia** Today 12:30 **LIVE**
- Slovakia v Paraguay** Sunday, 20 June, 12:30 [Preview](#)
- Italy v New Zealand** Sunday, 20 June, 15:00 [Preview](#)
- Paraguay v New Zealand** Thursday, 24 June, 15:00 [Preview](#)
- Slovakia v Italy** Thursday, 24 June, 15:00 [Preview](#)

Features

World Cup 2010 performance analysis
All the facts and figures from the World Cup 2010 tournament

- ▶ Slovakia performance analysis
- ▶ Paraguay performance analysis
- ▶ New Zealand performance analysis
- ▶ My World Cup - Noel Gallagher

Latest audio and video

- World Cup 2010 extended highlights
- Highlights - Italy 1-1 Paraguay
- De Rossi equalises for Italy

Metadata quality makes the best of what you have automatically...

The screenshot shows the BBC Sport website interface for the 2010 FIFA World Cup. The main navigation bar includes 'SPORT', 'FOOTBALL', 'WORLD CUP 2010', 'GROUPS & TEAMS', 'FIXTURES & RESULTS', 'VIDEO', and 'BBC COVERAGE'. The central focus is the player profile for Wayne Rooney, an English striker. His statistics are presented in a grid format, comparing his performance to the total for England. The statistics include:

- Games played:** Wayne Rooney (1), Total England (1)
- Shots on target | off target:** Wayne Rooney (2 | 2), Total England (9 | 7)
- Goals:** Wayne Rooney (0), Total England (1)
- Assists:** Wayne Rooney (0), Total England (1)
- Fouls by | on:** Wayne Rooney (0 | 2), Total England (8 | 10)
- Cards yellow | red:** Wayne Rooney (0 | 0), Total England (3 | 0)

Additional sections include 'Latest matches' (NED 2-0 DEN, JPN 1-0 CMR, ITA 1-1 PAR, NZL v SVK), 'Features' (Who will win the Golden Boot?), 'Top 5 World Cup stories', and 'Top TV and radio' (BBC iPlayer). The bottom section lists 'England games played' with links to match reports and previews.

Value...where did we start...where are we going...

The screenshot shows the BBC Sport website for the Vancouver 2010 Winter Olympics. The main headline is "Swiss men claim curling bronze". Below the headline is a video player showing a curling match. To the right, there are tweets from @bbcsp0rt_0lle. Below the tweets are several news articles with thumbnails, including "Canada's curlers strike 'perfect' gold", "Swiss win dramatic curling bronze", and "Canada throw away gold on last stone in curling final". The website also features a search bar, navigation menu, and a weather forecast for Vancouver, Canada.

15 indexes automatically managed with 446 stories had 3,500 tags

Total of about 60 minutes per day to tag stories instead of 1-3 people spending 30-50% of their time managing the site

For 20, 30 or 50 indexes the tagging time stays the same.

For the World Cup we had 800 index pages (every team, player, match and venue) and for the London Olympics we will have thousands

Reach...

Web [Images](#) [Videos](#) [Maps](#) [News](#) [Shopping](#) [Mail](#) [more](#) ▼

Google [Advanced Search](#)

Search: the web pages from the UK

Web [+ Show options...](#)

Image results for **tiger** - [Report images](#)



[Tiger](#) - [Wikipedia, the free encyclopedia](#) ☆

The **tiger** (*Panthera tigris*) is a member of the Felidae family; A subspecies called the Trinil **tiger** (*Panthera tigris trinilensis*) lived about 1.2 ...

en.wikipedia.org/wiki/Tiger - [Cached](#) - [Similar](#)

[BBC - Wildlife Finder - Tiger \(video, facts and news\)](#) ☆

Tigers are the largest of all the big cats. This magnificent animal is a heavily muscled and powerful predator that employs stalk and ambush tactics to ...

www.bbc.co.uk/nature/species/Tiger - [Cached](#) - [Similar](#)

Web [Images](#) [Videos](#) [Maps](#) [News](#) [Shopping](#) [Mail](#) [more](#) ▼

Google [Advanced Search](#)

Search: the web pages from the UK

Web [+ Show options...](#)

Results 1 - 10 of ...

Image results for **lion** - [Report images](#)



[Lion](#) - [Wikipedia, the free encyclopedia](#) ☆

The **lion** (*Panthera leo*) is one of the four big cats in the genus *Panthera*, and a member of the family Felidae. With some males exceeding 250 kg (550 lb) in ...

[Etymology](#) - [Taxonomy and evolution](#) - [Physical characteristics](#)

en.wikipedia.org/wiki/Lion - [Cached](#) - [Similar](#)

[Literature Online - Marketing Site](#) ☆

ProQuest web site. ProQuest's web site will open in a new browser window. Literature · HOW TO SUBSCRIBE REQUEST A TRIAL DOWNLOAD A BROCHURE JOIN OUR MAILING ...

lion.chadwyck.co.uk/ - [Cached](#) - [Similar](#)

[BBC - Wildlife Finder - Lion \(video, facts and news\)](#) ☆

Lions are unique in that they are the only cats to live in groups (prides). Male **lions** are also the only cats that have manes, giving them a regal ...

www.bbc.co.uk/nature/species/Lion - [Cached](#) - [Similar](#)

A final example...

The screenshot displays the Elan software interface for video analysis. The main window is titled "Elan - one-o-clock-news.eaf". The interface includes a menu bar (File, Edit, Search, View, Options, Help), a video player showing a man in front of a "KENT COUNTY CONSTABULARY HEADQUARTERS" sign, and a table of sections.

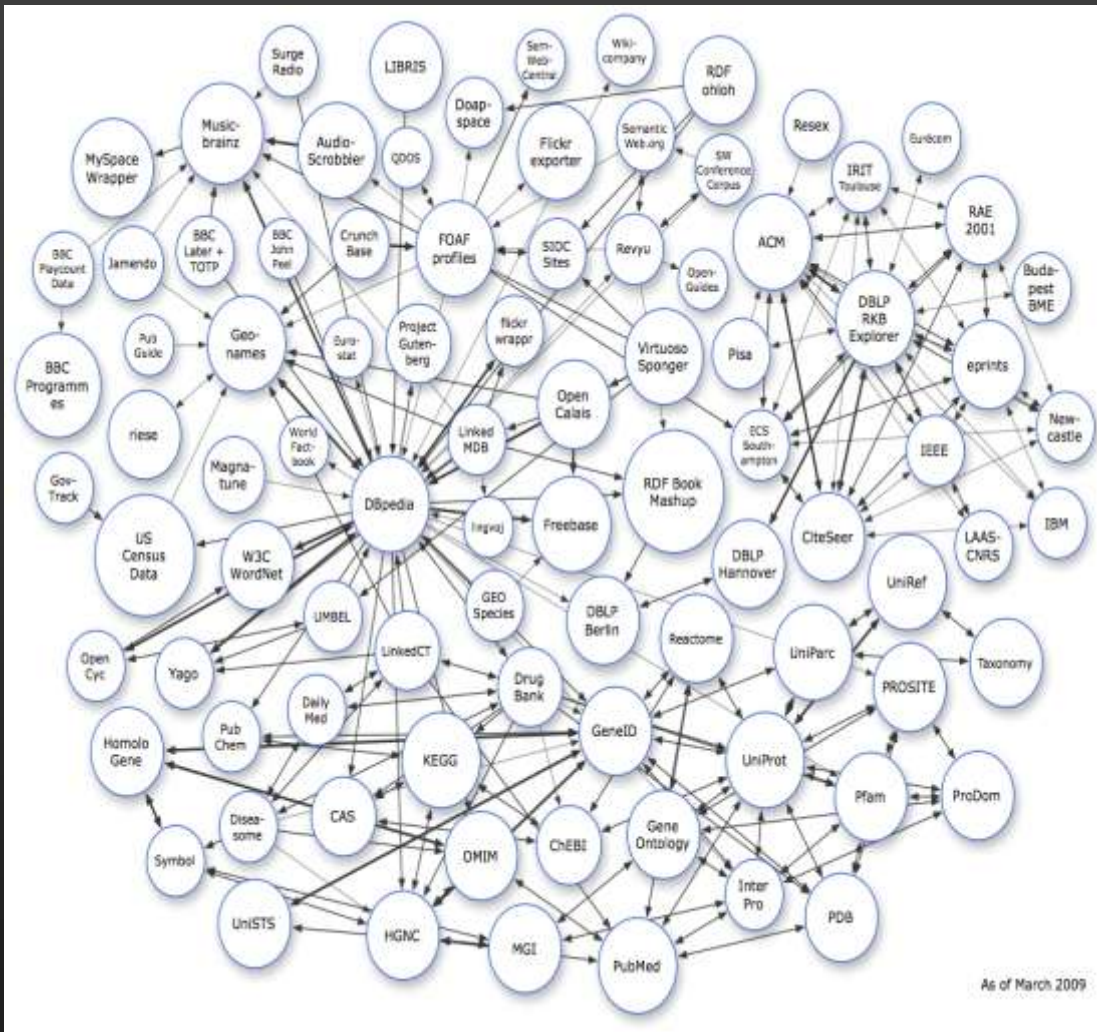
Nr	Annotation	Begin Time	End Time	Duration
2	Section=UK	00:01:51.888	00:04:01.760	00:02:09.872
3	Section=UK	00:04:01.760	00:05:31.568	00:01:29.808
4	Section=UK	00:05:31.568	00:06:52.496	00:01:20.928
5	Section=En...	00:06:52.496	00:09:43.008	00:02:50.512
6	Section=Edu...	00:09:43.008	00:12:27.312	00:02:44.304
7	Section=Mid...	00:12:27.312	00:15:06.224	00:02:38.912

The video player shows a timestamp of 00:09:44.850. Below the video player is a control bar with various playback buttons and a selection range of 00:09:43.008 - 00:12:27.312 (164304).

The bottom section of the interface shows a detailed timeline view for the selected segment. The timeline is divided into several tracks:

- Words:** he | can | initially | be | hell | for | twenty | four | hours | <SIL> | up | to | seventy | m
- Person:**
- Location:**
- Section:** Section=Education
- Organization:**
- Lonclass:** Lonclass=37
- Headline:** Headline=Splitting up A-levels
- Description:** Description=The BBC's Mike Baker also suggested the long-term solution to A-levels that has now been recom

Summary



Ontologies simplify data

Build an ontology, pour the data in

The web is going to evolve to use this

Real benefits that pay back fast

It's easier than you think but don't bite off too much. Most data sources are full of errors...you need to clean it up.

If your ontology looks too complicated, it's probably wrong...

Back to our goal...

“Moving from a page publishing model to a content publishing model where metadata and asset management is used to render the appropriate content to different target platforms for the lowest possible cost ”

Thanks for listening...

Any questions...?

john.odonovan@bbc.co.uk

Quality in breadth, depth and flexible publishing

The screenshot displays the BBC Sport website's layout for the 2010 World Cup. The top navigation bar includes 'Home', 'News', 'Sport', 'Weather', 'iPlayer', 'TV', 'Radio', and 'More'. The main header features 'SPORT WORLD CUP 2010' with a search bar and a trophy image. Below this, a secondary navigation bar lists 'SPORT', 'FOOTBALL', 'WORLD CUP 2010', 'GROUPS & TEAMS', 'FIXTURES & RESULTS', 'VIDEO', and 'BBC COVERAGE'.

The page is divided into several content areas:

- Fixtures and Results:** A grid showing match schedules for various teams.
- Blogs:** A section with articles from Dan Roan, John Motson, and David Bond, such as 'Pienaar ready for defining moment' and 'Capello's rating takes a tumble'.
- News:** A central section with a large image of Fabio Capello and headlines like 'Capello considering new formation' and 'Terry '100% fit' for FA Cup final'.
- Group Tables:** A table for Group A showing France, Mexico, South Africa, and Uruguay with their respective records (W, D, L, GF, PTS).
- Video choice:** A list of video thumbnails with titles like 'World Cup guide - England' and 'BBC World Cup coverage kit'.
- Features and analysis:** A section titled 'The story of the 1970 World Cup' with a photo of Pelé and a list of key moments.
- Group and Teams:** A detailed table showing the composition of all 32 teams across eight groups (A-H).
- Sport feeds:** A social media-style feed at the bottom right.