

メディアとLOD

2010/09/24

株式会社MetaMoji 乙守信行

2. 1 苦境に立たされる出版メディア業界

イギリスで開催された News Linked Data Summit

■概要

-2010年1月25日英国に開催される
-The Guardian, the Media Standards Trust, Talis、the BBC
が参加

■開催のモチベーション[1]

- LODに英国出版関係が入っていないことの危機感
- 紙ベースの新聞業界にLOD技術をどう導入するか

チャタム・ハウス・ルールのため発言者・発言内容は不明である
が、関係者のブログから英国での新聞業界の動向が垣間み
える

日本でも同様の問題意識、危機意識があるに違いないはずで
ある。

報道機関がLODを利用すべき10の理由

先行事例であるBBC, NYTなどから新聞業界としてLODを利用すべき理由をMartin Moorがまとめている。

1. Linked DataによってSEOが向上する
“LION”のキーワードでBBCのページが上位10位に出てくる。これは以前のBBCではなかったことである。
2. Linked data によって他者があなたのサイトに簡単にリンクを張ることができる
Searchエンジンのアルゴリズムは賢いが以下のような意味的に異なるものを区別することはできない。
Martin Moore (me), [Martin Moore](#)(kitchens), and [Daniel Martin Moore](#) (the Kentucky singer songwriter)
3. 自社のコンテンツをベースにしたサービスを構築するのに助かる
ある記事だけで有料にするのは難しくても、各地域の学校の成績比較サービスなどは、個々のデータが機械可読になっているれば開発可能。
4. 他の人があなたのコンテンツベースでサービスを開発することが出来る-これによってあなたも利益を得ることができる
5. 情報源を直接にリンクすることができる
報道機関であれば、情報の信頼性は重要なものとなる。犯罪データなど直接<http://data.gov.uk> にリンクさせれば信頼性は高まる。
6. ジャーナリストの仕事を助ける
関連情報の入手や、記事に対する参照情報を提示するなど記事執筆に役に立つ。
7. 有料化コンテンツに対しての導入
有料化された時点で、コンテンツは見えなくなるが、情報がリンクされていれば、そこから購入すべきかどうかの判断ができる。
8. コンテンツに対して動的にデータを関連付けられる
データが動的にリンクされれば、2006年の試合結果のままで止まっているページはなくなる。
9. ニュースを参照の基準点 ‘canonical reference point’ (CRP)とする
音楽業界でのハブとなる<http://musicbrainz.org>、政府データのハブとなる<http://dbpedia.org> などLinked Dataにつなげることで業界のハブとなる
10. 知的なプラットフォームをつくるために
LODはより知的なWEBであり、セマンティックWebへ進むために必要である

2. 2 出版メディア業界の挑戦(ビジネスモデル変革に向けた各社の取り組み)

2. 3 Linked Dataの活用(データの実例、および活用例)

先行事例：The New York Times

ニューヨーク・タイムズは2011年から従量課金による有料化を開始する。
将来は紙ベースの新聞はストップせざるを得ないとも発表[8]
LOD技術の導入を進めているのも、有料化の中での魅力的なサービス構築のひとつとして位置づけているのかもしれない。

2009/01/26	■ニューヨーク・タイムズシソーラスをLinked Data Cloudへ[3] 件名、人名、組織名、地名と創作物（著書、映画、演劇など）の5つの統制語彙をベースにした百万以上の用語のシソーラス
2009/10/29	■最初の5000タグをLinked Data Cloudへリリース[4] 人名に関する5000データをLODとしてFreeBase,DBpediaへリリースする。これらのデータはRDFとして第三者が利用可能となる。 記事検索APIも提供
2010/01/13	■新規50000のタグをLODへリリース[5] 人名に追加して、組織名、会社名、地名など新しい50000のタグをGeoNamesにもリリース
2010/03/30	■あなただけのLinked Data を使ったアプリケーションを開発しよう[6] 同級生検索サービス「 Who Went Where 」を例にLODを利用したWebアプリの開発方法をソースコードを含めて紹介
2010/06/24	■より多くの件名をLODへ[7] 単語レベルのデータだけでなく、上位下位関係を含めたデータをLODに提供する。これにより「誰(who)」 「どこ(where)」だけでなく「何what」での検索も可能となる。こうした関係記述詞(subject descriptor)は現状498であるが将来的には約3500を提供予定

先行事例:BBC放送

■前回報告分

BBC Program, BBC Wildlife Finderでこれまでの独立したサイトのコンテンツを
LODとオントロジーでリンク



■セマンティック技術を用いた動的なコンテンツ生成によるWORLD CUP2010年のBBCサイト = Dynamic Semantic Publishing

Text only | Help

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 2-1 BGR
Highlights & report

URU 1-1 GHA
Highlights & report

ARG 0-1 GER
Highlights & report

PAR 0-1 ESP

England

England 1-1 United States
Saturday, 12 June

England 0-0 Algeria
Friday, 18 June

Slovenia 0-1 England
Wednesday, 23 June

Germany 4-1 England
Sunday, 27 June

Match report

Match report

Match report

Match report

Latest stories

Gerrard commits future to England NEW

England sponsorship likely to end

Capello to remain England manager

Mueller blames England imbalance

Capello receives Gartsdale backing

Pressure got to Rooney - Ferguson

FA unfit for purpose says Caborn

England's fear of crossing borders

England duo bypass London event

Barwick baffled by dismal England

Group C Teams

	A	B	C	D	E	F	G	H
				W	D	L	GD	PTS
USA				1	2	0	1	5
England				1	2	0	1	5
Slovenia				1	1	1	0	4
Algeria				0	1	2	-2	1

Features

German lessons
Jurgen Klinsmann on how to revolutionise England

- A German view on English football
- Redknapp backs England to shine
- BBC pundits on England
- Roy Hodgson Q&A
- World Cup goals analysis

Around the web

- BBC Search+ country page
- England Fifa Profile

Dynamic Semantic Publication

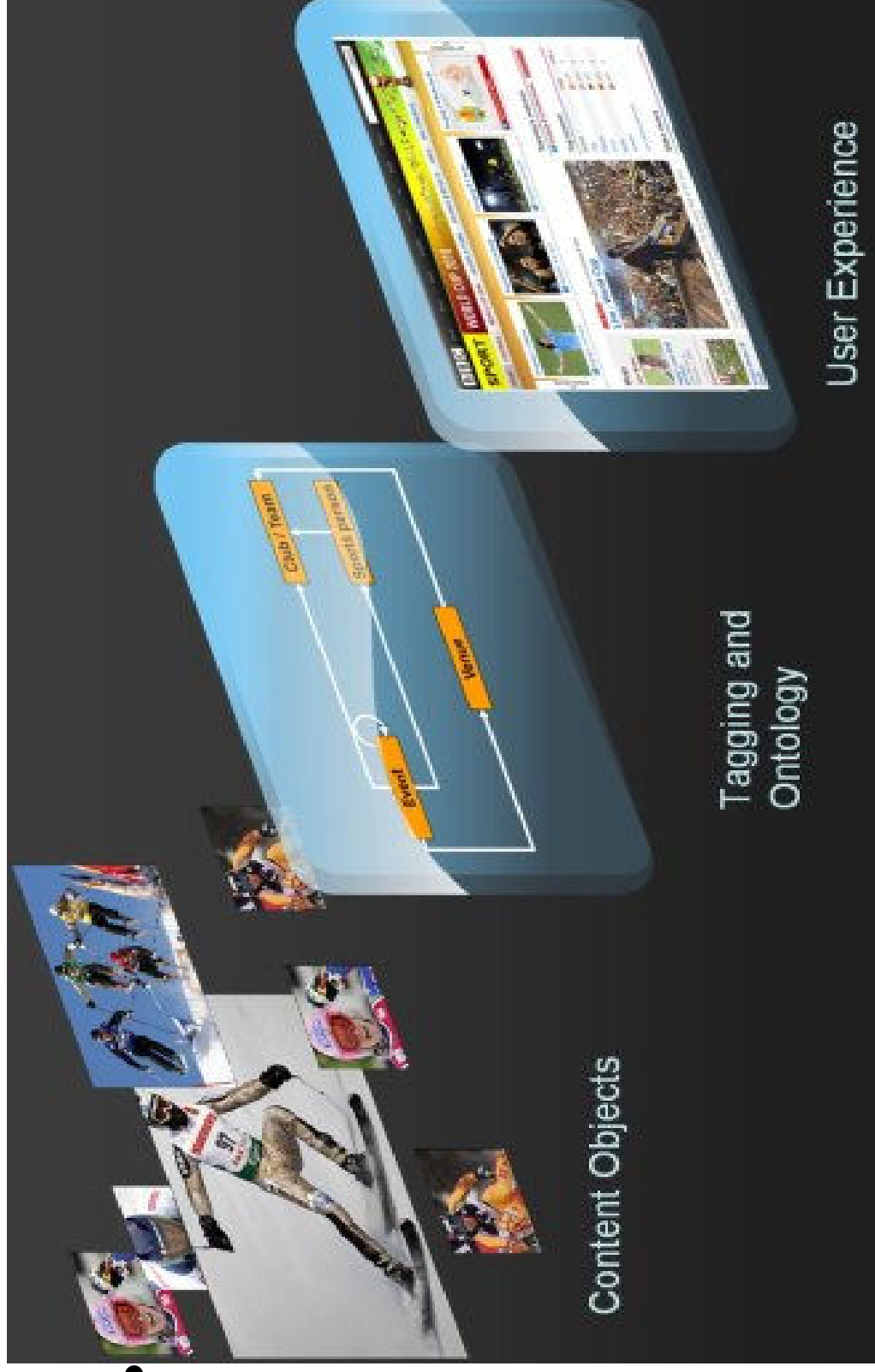
1. ワールドカップのサイトはグループやトーナメント表、各チームの情報など700ページ以上に及ぶ
2. こうしたindexページに対して、試合経過、談話などの最新情報を編集していかないといけない
3. そこで個々のコンテンツに対して自動的に関連情報のリンクを行ないWEBページを作成する
4. 単純な検索では、同一名の選手がリンクされてしまうなど正確性にかけてWEBページが生成してしまう
人名や地名など(World Cupサイトでは選手名、チーム、試合名など)の概念タグを各コンテンツに付与して正確にページをさせる

BBCの目指すところ

自社コンテンツだけでなく、外部のサイトやコンテンツにリンクしてより魅力的なサービスをユーザーに提供する

=>[10]ではLODレベルからSW技術に踏み込んだ技術説明がされているが、今回論文の趣旨としては簡単に触れる程度でいい？

Dynamic Semantic Publishingのイメージ図 [9]より



2. 4 出版メディアは生き残れるか(期待される効果、将来ビジョン)

2. 5 今後の課題

[検討の方向性1]

Martin Moorの提案も参考に国内事情を含めて今後検討する。

1. Linked DataによってSEOが向上する
2. Linked data によって他者があなたのサイトに簡単にリンクを張ることができる
3. 自社のコンテンツをベースにしたサービスを構築するのに助かる
4. 他の人があなたのコンテンツベースでサービスを開発することが出来る-これによってあなたも利益を得ることが出来る
5. 情報源を直接にリンクすることができる
6. ジャーナリストの仕事を助ける
7. 有料化コンテンツに対しての導入
8. コンテンツに対して動的にデータを関連付けられる
9. ニュースを参照の基準点 'canonical reference point' (CRP)とする
10. 知的なプラットフォーム

[検討の方向性2]

根本さん提案(勝手に拝借しました m_m)

多分、出版のみならず、コンテンツ(この場合定義が難しいですが。)をWeb上で配信する「媒体(メディア)」業界すべての問題だと思います
1. まず、アナログ(紙、アナログ放送)で配信するか? 否か? (Web 0.0)
2. 次に、Web上で配信するか? 否か? (Web 1.0 ~ 2.0)
3. そして、現在のWeb上で、LOC や Semantic Publishing を使うか? 否か? (Web 2.0から、セマンティックであるWeb3.0へ)
ちよつと乱暴で単純化しすぎているかもしれませんが、上記のようではないでしょうか?
(あくまで仮説です。)
海外事例として、N.Y. TimesやBBCは、2→3への遷移の問題ではないでしょうか?
日本の出版業界は、1→2への遷移でそもそもつまづいているとは、考えられないでしょうか? 仮説の上での考察(具体事例をマッピング)ですが。。。日本のテレビ業界はどうでしょうね? ちよつと考えます。

參考資料

- [1] What is the value of Linked Data to the news industry?' Martin Belam 2010/1/25
<http://www.guardian.co.uk/help/insideguardian/2010/jan/25/news-linked-data-summit>
- [2] Why news organisations should try publishing in linked data
<http://www.mediastandardstrust.org/medianews/blogs/blogdetails.aspx?sid=49891>
Martin Moore 2010/03/17
- [3] [NYT to Release Thesaurus and Enter Linked Data Cloud](#)
<http://open.blogs.nytimes.com/2009/06/26/nyt-to-release-thesaurus-and-enter-linked-data-cloud/>
- [4] [First 5,000 Tags Released to the Linked Data Cloud](#)
<http://open.blogs.nytimes.com/2009/10/29/first-5000-tags-released-to-the-linked-data-cloud/>
- [5] More Tags Released to the Linked Data Cloud
<http://open.blogs.nytimes.com/2010/01/13/more-tags-released-to-the-linked-data-cloud/>
- [6] Build Your Own NYT Linked Data Application
<http://open.blogs.nytimes.com/2010/03/30/build-your-own-nyt-linked-data-application/>
- [7] More Subject Headings Published as Linked Open Data
<http://open.blogs.nytimes.com/2010/06/24/more-subject-headings-published-as-linked-open-data/>
- [8] Arthur Sulzberger on charging online: to succeed, we need to take risks
http://www.editorsweblog.org/newspaper/2010/09/arthur_sulzberger_on_charging_online_to.php
- [9] The World Cup and a call to action around Linked Data 2010/07/09
http://www.bbc.co.uk/blogs/bbcinternet/2010/07/the_world_cup_and_a_call_to_ac.html
- [10] BBC World Cup 2010 dynamic semantic publishing 2010/7/12
http://www.bbc.co.uk/blogs/bbcinternet/2010/07/bbc_world_cup_2010_dynamic_sem.html