

# Media Meets Semantic Web

How the BBC uses DBpedia and Linked Data to Make Connections

---

Andreas Müller

11<sup>th</sup> June, 2017

Technische Universität Berlin

# TABLE OF CONTENTS

1. Problem
2. Objectives
3. Interlinking of concepts
4. Interlinking of documents
5. Content Link Tool
6. Conclusion

# Problem

---

# PROBLEM

Large amounts of BBC online content: **TEXT, AUDIO, VIDEO**

Domain specific microsites: **FOOD, GARDENING, SPORT**, etc. ...

# PROBLEM

Large amounts of BBC online content: **TEXT, AUDIO, VIDEO**

Domain specific microsites: **FOOD, GARDENING, SPORT, etc. ...**

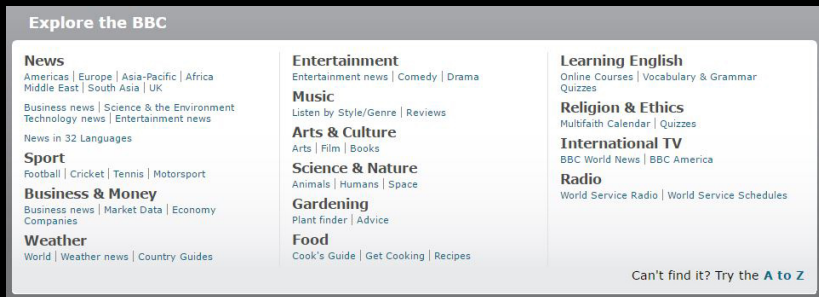


Figure 1: BBC microsites in June, 2009

# PROBLEM

Large amounts of BBC online content: **TEXT, AUDIO, VIDEO**

Domain specific microsites: **FOOD, GARDENING, SPORT**, etc. ...

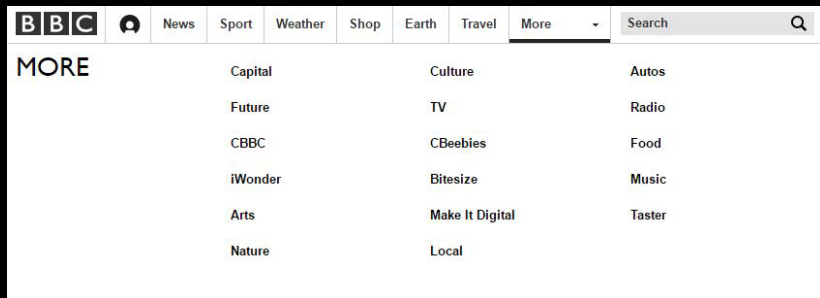


Figure 2: BBC microsites in June, 2017

# PROBLEM

## Not possible to...

Find **everything**, BBC has published to a given subject

Navigate between BBC domains following a **semantic thread**

No interlinking between microsites = not using full potential of available data

# Objectives

---



# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

1. Service to link all radio and TV programmes

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

1. Service to link all radio and TV programmes
2. Develop a new music offering

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

1. Service to link all radio and TV programmes
2. Develop a new music offering
3. Retrofit simple navigational elements

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

1. Service to link all radio and TV programmes
2. Develop a new music offering
3. Retrofit simple navigational elements
4. Provide a common set of web scale identifiers

# OBJECTIVES

Make the BBC website **more coherent** and **more useful**

Better connections and interlinking of existing systems

**Soft transition** and reducing impact on existing systems while adding new services to **maximize interlinking** of domains

1. **Service to link all radio and TV programmes**
2. Develop a new music offering
3. Retrofit simple navigational elements
4. **Provide a common set of web scale identifiers**

## Interlinking of concepts

---



Legacy auto-categorization system: CIS

## Limits:

- Difficult to cover every single entity
- No relations between terms are available
- Only internal identifiers

# INTERLINKING OF CONCEPTS — DBPEDIA

## Common Vocabulary: DBpedia

### DBpedia Label Lookup

Find most likely matches to a given term, calculate relevance with number of backlinks

### Context-based Disambiguation

Disambiguate possible matches by clustering them and finding according context in DBpedia

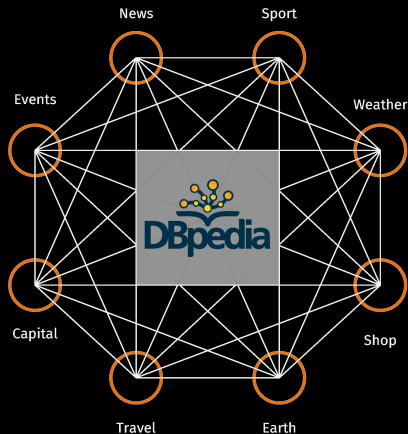


Figure 3: Linking BBC Domains

## Interlinking of documents

---

# INTERLINKING OF DOCUMENTS — MUDDY BOOTS

Identify main actors in a piece of content: **Muddy Boots**

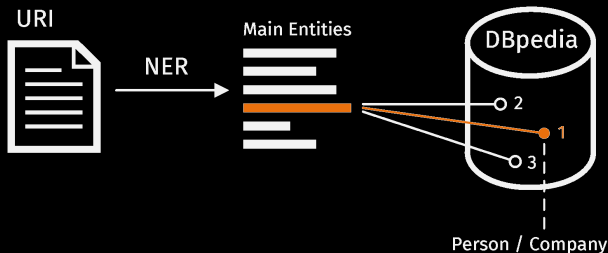


Figure 4: How Muddy Boots works

# Content Link Tool

---

Annotation tool to manually edit metadata  
High quality automated suggestions

## Conclusion

---

# CONCLUSION

User experience in the center of efforts

Smart interlinking to internal and external resources

Well integrated and hidden systems



Thanks for your attention!

Questions?



G. Kobilarov, T. Scott, Y. Raimond, S. Oliver, C. Sizemore, M. Smethurst, C. Bizer, and R. Lee.

**Media meets semantic web - How the bbc uses dbpedia and linked data to make connections.**

*Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5554 LNCS:723–737, 2009.