

# Semantic Web Technologien: FIFA's Fußball-Weltmeisterschaft 2014 in Brasilien bekommt eine semantische Unterstützung.

Manuel Dudda, Robert Brylka, and Frank Reichwein

Hochschule RheinMain Informatik Master of Science  
Fachbereich Design Informatik Medien  
Campus Unter den Eichen 5 65195 Wiesbaden , Deutschland  
{manuel.b.dudda, robert.c.brylka, frank.b.reichwein,}@student.hs-rm.de  
<http://www.semanticwc.herokuapp.com>

**Abstract.** Große Sportereignisse haben schon immer große Menschenmaßen begeistert. So auch die Events die die Disziplin Fußball betreffen. Dieses Paper beschreibt ein Projekt, das sich als Ziel die Vorstellung der Ereignisse der diesjährigen Fußballweltmeisterschaft gesetzt hat. Der Fokus dieser Anwendung liegt auf der Aktualität der präsentierten Daten sowie einer auf die Mobilien-Endgeräte angepassten Repräsentation. Für die Vielfalt der präsentierten Daten sowie deren Richtigkeit sorgte eine Informationssynthese aus mehreren Quellen. Um die Informationen weitestgehend für automatische Weiterverarbeitung vorzubereiten wurden alle Daten um semantische Bedeutung aus gängigen Sport-Ontologien ergänzt.

## 1 Introduction

In den vielen Informationsseiten, die die diesjährige Fußballweltmeisterchat in Brasilien präsentieren, such man oft vergeblich nach gut strukturierten und leichtgewichtigen Applikationen, die schnell mit mobilen Endgeräten zu erreichen sind. Zu dem werden viele Seiten nicht aktuell gehalten, was gerade bei so ereignisreicher Sportdisziplin wie Fußball unerlässlich ist. Beim unseren Projekt haben diese Aspekte die höchste Priorität erhalten. Um einem aktuellen Trend in der Webentwicklung zu folgen wurde unsere Applikation mit Konzepten des semantisches Web entwickelt. Dabei werden alle Informationen mit einer für die Maschine verständlicher Syntax ergänzt. Dazu wurden die Informationen der FIFA Homepage extrahiert und mit der Ontologie von BBC-Sport und der Ontologie Soccer-Voc des Hasso PLatners Instituts versehen. Für die Darstellung dieser Daten wurde eine Ruby on Rails Web Application, welche zur Präsentation JQuery Mobile einsetzt, entwickelt. Des weiteren werden weiter führende Informationen, durch die einbeziehung von DBpedia Einträgen, eingeblendet.

## 2 Related Work

Bei der Recherche in dem Bereich der Fußballereignisse mit semantischen Hintergrund hat sich unsere Aufmerksamkeit auf drei Quellen gerichtet, die Sport-Ontologie<sup>1</sup> der BBC, die Soccer Voc Ontologie<sup>2</sup> des Hasso-Plattner-Institut sowie der World Cup<sup>3</sup> JSON API.

Die Future Media and Technologie Abteilung der BBC setzt semantische Technologien um die intern relevanten Inhalte automatisch verwalten zu können. Der erste große Einsatz der semantischen Technologien war die Gestaltung der Internetseite<sup>4</sup> mit Informationen zu dem FIFA World Cup 2010. Die dabei eingesetzte Ontologie ist seither frei zugänglich.

Die Arbeitsgruppe "Semantic Technologies"<sup>5</sup> an dem Hasso-Plattner-Institut im Potsdam veröffentlichte 2012 ein Paper zu Linked Soccer Data[1]. Im Rahmen dieses Papers wurde ein Datensatz mit Fußballinformationen vorgestellt, der mit der LOD Cloud<sup>6</sup> verknüpft wurde. Die dabei verwendete Ontologie kann frei verwendet werden und stellt eine gute Ergänzung der Sport-Ontologie der BBC dar.

Bei der World Cup API handelt sich um ein von Eric Stiens<sup>7</sup> betreutes Projekt. Dessen Ziel ist es die aktuellen Fußballereignisse für jeden frei verfügbar zu machen. Neben der frei verfügbaren Informationen, die als JSON Dateien vorliegen, ist auch die API selbst für Weiterentwicklung als Projekt<sup>8</sup> auf Github freigegeben.

## 3 Bestimmung der Informationsquellen

Die Fußballweltmeisterschaft ist allgemein ein belebtes Thema in der digitalen Welt. Um geeignete Quellen für die gesuchte Informationen zu finden wurde eine tiefgründige Recherche durchgeführt. Die unten aufgeführte Tabelle zeigt ein Ausschnitt aus der Zusammenstellung in der die geeigneten Quellen aufgeführt wurden. Wie man an der Zusammenstellung erkennen kann verfügen nicht alle Quellen über die gleiche Informationsfülle. Dabei ist die Open Football Datenbank aufgrund der nicht aktuellen Informationen für weitere Verwendung uninteressant. Die Football Data sowie die World Cup Datenbank zeigen beide eine sehr gute Fülle an brauchbaren Informationen. Bei einer weiteren Recherche zu der Word Cup Datenbank wurde die offengelegte API gefunden die einen direkten Zugriff auf die offizielle Seite der FIFA ermöglicht. Wie die letzte Spalte der Zusammenstellung zu entnehmen ist, ist die offizielle Seite des Internationalen

<sup>1</sup> <http://www.bbc.co.uk/ontologies/sport>

<sup>2</sup> <http://purl.org/hpi/soccer-voc/>

<sup>3</sup> <http://www.http://worldcup.sfg.io/>

<sup>4</sup> [http://news.bbc.co.uk/sport2/hi/football/world\\_cup\\_2010/default.stm](http://news.bbc.co.uk/sport2/hi/football/world_cup_2010/default.stm)

<sup>5</sup> [http://hpi.de/meinel/knowledge\\_tech/semantics.html](http://hpi.de/meinel/knowledge_tech/semantics.html)

<sup>6</sup> <http://linkeddata.org/>

<sup>7</sup> <http://softwareforgood.com/soccer-good/>

<sup>8</sup> [https://github.com/estiens/world\\_cup\\_json/](https://github.com/estiens/world_cup_json/)

Fußballverbandes die beste Informationsquelle. Daraus resultierend wurde ein Fork der Soccer for good API erstellt und dieser unseren Bedürfnissen angepasst. Durch die Anpassungen wurde die API um Informationen zu den Spielern, Trainern und Schiedsrichtern erweitert. Es wurden damit alle für die Anwendung relevanten Informationen in die API integriert. Dieser Fork Cup bietet bi-etet entsprechende REST<sup>9</sup> Endpunkte an. Die Rückgabe dieses Services ist eine JSON<sup>10</sup> Datei mit allen relevanten Informationen.

	Football Data	Open Football	Footbal DB	World Cup	FIFA
References	11	12	13	14	15
Teams	+	+	+	+	+
Gruppen	-	+	-	+	+
Spiele	+	+	-	+	+
Stadien	+	+	-	+	+
Uhrzeiten	+	+	-	+	+
Spieler	+	+	-	-	+
Spielergebnisse	+	+	+	+	+
Torschützen/min	+/+	+/+	-	+	+/+
Schiedsrichter	-	-	-	-	+
Trainer	-	-	-	-	+
Spieltag	+	+	+	+	+
Aktuell	-	-	+	+	+
Besuchierzahl	+	-	-	-	+

Eine wichtige Rolle bei der Wahl der API hat letztendlich auch der Finanzaspekt gespielt. Es gibt eine ganze Reihe an gut organisierten und stets aktuellen Datenbanken, die jedoch eine monatliche Gebühr mit sich tragen. Für dieses Projekt wurde jedoch eine Vereinbarung getroffen, rein Quelloffene Lösungen zu verwenden. Dadurch sind Finanziellabhängige Projekte direkt bei der Recherche ausgeschlossen worden.

<sup>9</sup> **R**epresentational **S**tate **T**ransfer

<sup>10</sup> **J**avaScript **O**bject **N**otation

<sup>11</sup> <https://github.com/footballdata/fifadata>

<sup>12</sup> <https://github.com/openfootball/world-cup>

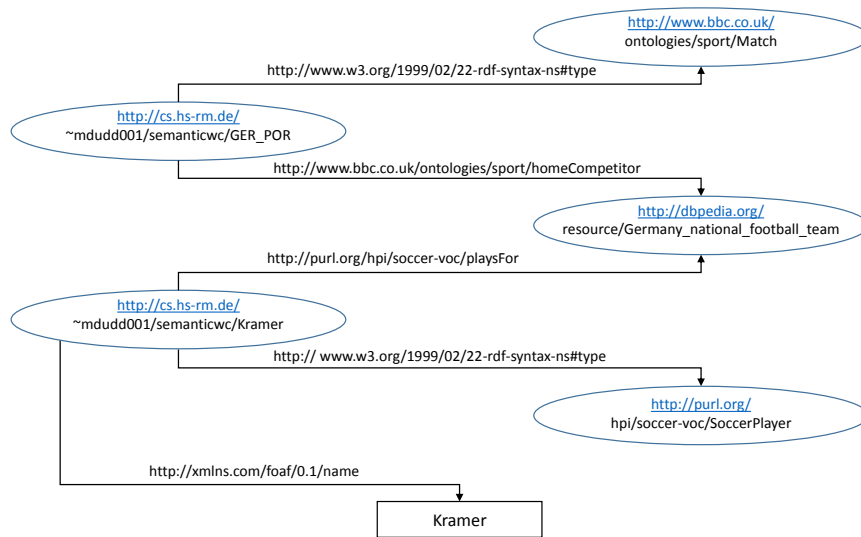
<sup>13</sup> <http://footballdb.herokuapp.com/api/v1/>

<sup>14</sup> <http://www.worldcup.sfg.io/>

<sup>15</sup> <http://www.fifa.com/>

## 4 Ontologien

Um die Daten semantisch korrekt darzustellen wurden verschiedene Ontologien verwendet. Für die Darstellung des FIFA World-Cups wurde die BBC-Sport-Ontologie verwendet. Mit dieser wurden die verschiedenen Gruppenphasen, Spieltypen, Spiele sowie die zugehörigen Stadien abgebildet. Für die Fußball spezifischen Daten wie Spielevents (Tore, Ein- und Auswechselungen sowie Karten) wurde die Soccer-voc-Ontologie des Hasso-Platner Instituts verwendet. Diese erweitert die Friend of a Friend Ontologie. Außerdem wurde die Event-Ontologie verwendet, für die zeitliche Einordnung der Spielevents.



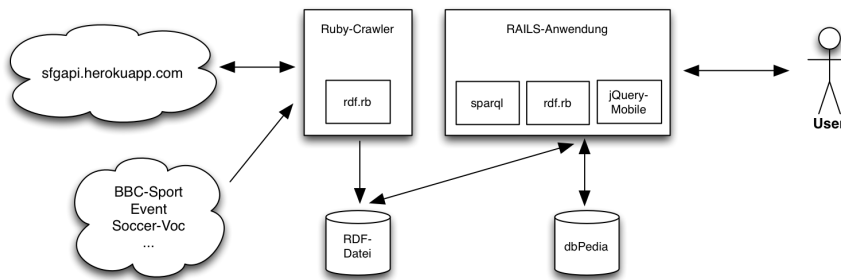
**Fig. 1.** Exemplarische Ausschnitt aus der RDF.

Die Abbildung 2 zeigt einem Graphen, der für einen exemplarischen Ausschnitt aus der mit den Ontologien versehenen Daten, aufgebaut wurde. An diesem Beispiel ist die Verflechtung der Ontologien sehr gut sichtbar. Aus dem Graph ist ersichtlich, dass für jedes Spiel ein eigener Knoten mit den Länderkürzeln der Spielteilnehmer erstellt wurde *GER\_POR*. Dieser Knoten ist vom Typ *bbc-sport:Match* und hat als *bbc-sport:homeCompetitor* Attribut den Knoten der deutschen Nationalmannschaft. Des Weiteren gibt es einen Knoten für jeden Spieler, welche vom Typ *soccer-voc:SoccerPlayer* sind. Diese sind mit dem Attribut *soccer-voc:playsFor* mit den Mannschaften verbunden. Das zuletzt genannte Attribut und der zuletzt genannte Typ stammen von der Soccer-voc Ontologie. Da *SoccerPlayer* eine Unterklasse von *foaf:Agent* ist, kann ein SoccerPlayer mit einem Attribut für einen Namen versehen werden.

## 5 Application // Implemetierung

Bei der Wahl des Framework für diese Anwendung fiel die Entscheidung zu Gunsten von Ruby on Rails <sup>16</sup>. Dieses quelloffene Web Application Framework bietet sehr gute Unterstützung bei Entwicklung von Semantik Web Applikationen um die RDF Datenbank oder den SPARQL Client zu nennen. Für Gestaltung der Seiten sowie als Unterstützung für die Mobilen-Endgeräte, z.B. für den Swipe-Effect, wurde die jQuery mobile<sup>17</sup> Bibliothek eingesetzt.

Die Abbildung 2 zeigt eine Übersicht über den Technick-stack. Für die immer aktuellsten und vollständigen Informationen aus der FIFA Seite wurde ein REST Server aufgestellt, der die bereits angesprochene und an dieses Projekt angepasste Word Cup API ausführt. Dieser Server ist auch separat über `sfgapi.herokuapp.com` erreichbar. Die so vorbereiteten Daten werden mit den zuvor beschriebenen Ontologien aufbereitet. Da die Anwendung sich zunächst auf ein Sportliches Ereignis beschränkt, in diesem Fall die Fußballweltmeisterschaft, wurde hierfür keine zusätzliche Datenbank angelegt. Stattdessen werden alle Datensätze als RDF-Trippel in einer Datei gespeichert. Für die Serialisierung der Datenbestände wurde hierbei Turtle Syntax verwendet.



**Fig. 2.** Technik-stack.

Als Abfragesprache für die so vorbereiteten Informationen wurde SPARQL ausgewählt. Neben der RDF-Datei, die als primäre Endpoint fungiert, wird für zusätzliche Informationen die DBpedia als sekundäre Endpoint eingesetzt.

<sup>16</sup> <http://www.rubyonrails.org/>

<sup>17</sup> <http://www.jquerymobile.com/>

In der Abbildung 3 sind mehrere Screenshots, aus unterschiedlichen Ansichten der Anwendung, zu sehen. Oben links ist die Startseite zu sehen, auf der immer der aktuelle Spieltag angezeigt wird. Links unten ist ein Ausschnitt aus der Ansicht die Informationen zu dem Stadion präsentiert. Die Rechtheälfte des Screenshots zeigt wiederum eine Ansicht die die Informationen zu einem Trainer einer Mannschaft präsentiert. Da die Anwendung vorwiegend auf Mobile-Endgeräte ausgelegt ist, wurde bei der Gestaltung der einzelnen Ansichten ein großer Wert auf die Übersichtlichkeit gelegt. Für eine schnelle Navigation wurden zudem zwei Buttons eingepflegt. Der 'Home' Button, der in jeder Ansicht in der oberen linken Ecke zu sehen ist, führt immer zu der Seite mit den aktuellen Spieltag. Mit dem 'Bar' Button, in oberen rechten Ecke der Spiele Ansicht, kann eine Seitenansicht mit mehreren Filteroptionen eingeblendet werden. Dabei kann der Inhalt nach dem Team, der Gruppe, dem austragenden Stadion oder dem Spieltag sortiert werden.

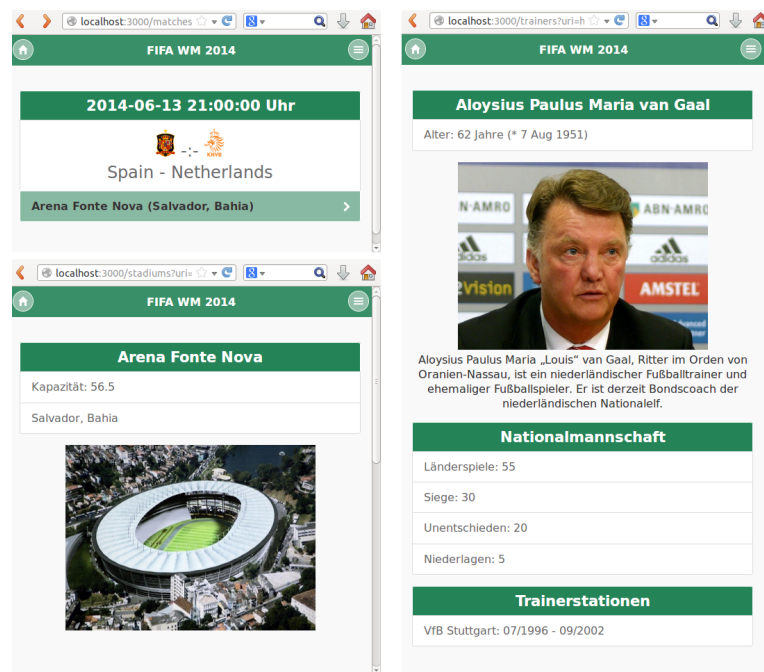


Fig. 3. Screenshots der Anwendung.

Aus der Hauptansicht, in der immer der aktuelle Spieltag angezeigt wird, kann durch eine Wischgeste nach links oder rechts zu den bereits stattgefundenen respektive in der Zukunft liegenden Spielen gewechselt werden. Die gleiche Funktionalität liegt auch bei Ansichten, die durch anwenden der Filteroperation; *nach Team*, nach Gruppe sowie *nach Tag* gewählt werden können.

## 6 Zusammenfassung und Ausblick

Im Rahmen dieses Projekts wurde eine Anwendung entwickelt, die alle wichtigen Informationen, zu der 2014 in Brasilien stattfindenden Fußball Weltmeisterschaft, bereithält. Als Informationsquelle dient hierbei die offizielle FIFA Webseite, die mit Daten aus der DPpedia ergänzt wurde. Die semantische Aufbereitung wurde mit Hilfe der BBC Sport sowie der Socker Voc Ontologie gestaltet.

Unsere Anwendung bietet eine solide Grundlage für Weiterentwicklungen die nicht nur den Bereich Fußball betreffen. Bezogen auf die Fußballereignisse kann eine leichte Portierung auf die nächste Weltmeisterschaft oder Europameisterschaft stattfinden. Denkbar ist aber auch eine Abwandlung zur Informationspräsentation der Spiele der Bundesliga Mannschaften.

## References

1. Linked Soccer Data, Tanja Bergmann, Stefan Bunk, Johannes Eschrig, Christian Hentschel, Magnus Knuth, Harald Sack, and Ricarda Schüler  
Hasso Plattner Institute for Software Systems Engineering, Potsdam, Germany  
Letzter Zugriff: 22. Juni 2014  
<http://ceur-ws.org/Vol-1026/paper6.pdf>
2. BBC ONTOLOGIES, A simple ontology for representing competitive sports events.  
Letzter Zugriff: 26. Juni 2014  
<http://www.bbc.co.uk/ontologies>
3. jQuery mobile, Letzter Zugriff: 26. Mai 2014  
<https://demos.jquerymobile.com/1.4.2/>
4. Semantic Media Mining,  
Letzter Zugriff: 22. Juli 2014  
<https://smm2013.blogspot.de/2012/11/soccer-voc.html>
5. DBpedia, Letzter Zugriff: 22. Juli 2014  
<http://http://dbpedia.org>
6. Ruby. Der beste Freund eines Programmierers, Letzter Zugriff: 22. Juli 2014  
<https://www.ruby-lang.org/de/>