SDM Ontology: Software Defined Mediaのメタデータ管理のためのOntology

菰原 裕[1]　塚田 学[1]　江崎 浩[1]　曽根 卓朗[2]　池田 雅弘[2]　高坂 茂樹[3]　新 麗[4]　新 善文[5]

概要：インターネットが広く普及し，全天球カメラのような新しい記録デバイスも登場している現代において，インターネットを前提としたオーディオビジュアルシステムの需要が高まっている．そのような状況において，Software Defined Media(SDM) コンソーシアムでは，オブジェクトベースのメディアとインターネットを前提とした環境における視聴空間の制御を目的として研究を進めている．本研究では，SDM が収録を行なったクラシックコンサートのデータを Linked Open Data(LOD) として保存・公開する為に”SDM Ontology”を作成した．その際，SDM のデータを適切に表現出来るようにクラスの構造やプロパティの種類を議論し，構築を行なった．

SDM Ontology: Ontology for managing metadata of Software Defined Media

Yu Komohara[1]　Manabu Tsukada[1]　Hiroshi Esaki[1]　Takuro Sone[2]　Masahiro Ikeda[2]　Shigeki Takasaka[3]
Rei Atarashi[4]　Yoshifumi Atarashi[5]

1. はじめに

1.1 背景

近年，スマートフォンやタブレットといったデバイスが普及し，また全天球カメラのような新しい記録デバイスも登場している．そのような状況において，対象の三次元情報も含めて記録する，立体音響のような手法が普及しつつある．記録対象をソフトウェア制御すること，三次元情報とメタデータを記録しバーチャル空間上で再現すること，またバーチャル空間上で可能な演出を組み合わせて新しい表現を創造することを目的に，現在 SDM(Software Defined Media) コンソーシアムが活動している．[*1] [1] [2]

SDM コンソーシアムでは，データを記録・再現するためのシステムを構築するだけでなく，実際にライブやコンサートのデータを記録し公開する活動も行なっている．しかし記録したデータは，対象の位置情報や音といったデータから，録音プロセスの詳細や空間のメタデータまでを含み，整理が困難である．このデータを，アプリケーションの作成に有用な形式で公開して新しい表現を創出するために，適切なデータ構造を考案する必要がある．

本論文では，まず第 2 章で SDM の詳細な説明と保有するメタデータについての解説を行う．第 3 章では，データ管理の要素技術である RDF/LOD と，その関連研究について述べる．続く第 4 章で，今回構築した SDM Ontology の詳細について説明し，第 5 章で結論と今後の展開を述べる．

2. Software Defined Media

2.1 SDM の概要

スマートフォンのように収録機器を内蔵したモバイル端末が普及し，またインターネットも一般的に使用されている現代では，記録した映像・音声をインターネット上で共有するのが当然の時代となっている．さらには全天球カメラや HMD(Head Mount Display) のような新しい記録デバイス・再生デバイスも登場している．そのため，単一視点

[1] 東京大学 The University of Tokyo
[2] ヤマハ株式会社 YAMAHA Corporation
[3] エスイーディー 株式会社 Sophisticated Engineers Department (SED)
[4] 株式会社 IIJ イノベーションインスティテュート IIJ Innovation Institute
[5] アラクサラ ネットワークス株式会社 ALAXALA Networks Corporation
[*1] http://sdm.wide.ad.jp

や 2 チャンネルのステレオ録音といった従来の記録方式ではない，対象の三次元情報も含めて記録する手法が登場しつつある．すなわち，対象の位置情報や向き，音楽イベントであればそれぞれの演者が発する音などのデータも記録する，対象メディアをオブジェクトとしてとらえオブジェクトベースでデータを管理する手法が注目を集めている．空間内における音の聴こえ方を再現した立体音響などはその例といえる．

オブジェクトベースの情報記録は，単一の対象であれば容易である．しかし，コンサートやライブ会場のように複数の記録対象が存在する空間そのものを一度に記録することは，同期や遅延の観点から難しい．また，複数のデバイスで別々に記録したとしても，収録したデータの集約コストが高いため現実的なシステムの構築は困難であった．しかし，IP ネットワークの発達により，サブシステム間をネットワークで繋ぎ，全体を 1 つのシステムとして制御する技術が登場してきている．また，スマートフォンのようにメタデータを自動的に記録出来るデバイスも普及している．このような状況によって，空間を複数のデバイスで一度に記録することが可能になりつつある．

以上のような，オブジェクト指向のディジタルメディア，ネットワーク環境をデザインに組み込んだシステムの構築により，以下のようなことが可能となる．

- 記録空間のソフトウェア制御
- 三次元情報とメタデータを記録しバーチャル空間上で再現すること
- バーチャル空間上で可能な演出を組み合わせて新しい表現を創造すること

これらを目的として，2014 年 1 月より SDM(Software Defined Media) コンソーシアムが活動している．

2.2 データセットとその活用

SDM では，システムの設計・構築以外に，実際にイベントのデータを記録する作業も行なっている．2016 年 01 月 10 日，慶應義塾大学日吉キャンパス内に建つ，509 席の席数を有する藤原記念ホールにて開催された，慶應義塾大学 コレギウム ムジクム古楽アカデミーのコンサートを収録した．収録対象は，17 世紀のドイツ宮廷音楽であるヨハン・フリードリヒ・ファッシュの「管弦楽組曲ト長調 FaWV K:G2」など，最大 24 名の演奏家による全てアコースティックの演奏である．楽器には，テオルボ，チェンバロ，バロック・ヴァイオリン，バロック・オーボエ，ヴィオラ・ダ・ガンバなどの，現代では珍しい楽器が含まれており，楽器ごとの演奏と各楽器の合わさった音を分離して収録した [3]．カメラとマイクは，Fig.1 のように配置し，記録した．

こうして記録したデータには，楽器そのものの音だけではなく周囲の環境音や残響音が含まれている．SDM では，このデータから楽器単体の音を抽出してバーチャル空間上に配置することで，会場における音の聴こえ方を再現することを試みている．

実際に抽出した楽器の音を使ったアプリケーションとして，SDM では”360Square”というアプリケーションを開発した．[4] これは，収録会場の 3D モデルを作成し，位置情報に基づいてパートごとの音源や映像を配置することで，自由聴点と複数視点のインタラクティブな切り替えを可能にしたアプリケーションである．実際の動作画面は Fig.2 のようになる．ユーザは，観客席の複数の箇所からの視点を切り替えて観ることができ，また舞台上の音源は自由に有効化・無効化することが可能である．

SDM では他にも以下のようなユースケースを考えている．

- スポーツ観戦やコンサート鑑賞において，ユーザの要求に応じて特定の選手や演奏者にズームアップしたり，スタンド・客席の自由な視点から鑑賞できるアプリ．
- バーチャル空間上で演出を付加し，現実空間にフィードバックすることによる拡張演出
- 複数コンテンツの自動ミキシング

しかし上記のようなアプリケーションは，現在は容易に作ることが出来ない．実際に 360Square の作成にあたっては，デバイスと楽器の位置情報を元に音声・映像データを手動で配置したため，実装の工数が多くなってしまった．データを構造化し，データベースとやり取りする際の標準フォーマットを決定することで，作業が自動化され効率的なアプリケーションの作成が可能になり，また外部のアプリケーションやデータとの連携も容易になると考えられる．

2.3 三次元試聴オブジェクトのメタデータ

SDM では，対象の映像や音だけでなく，位置情報や向き，演奏された曲目や会場の情報，録音プロセスといった情報も記録するため，相互に関連するが粒度がバラバラなデータを扱うことになる．そこで，SDM が想定しているユースケースを実現するために，データを構造化し相互にも外部にも連携可能とする必要がある．

多数のデバイスを用いて記録した視聴空間のデータを管理する，SDM のデータ形式の要求事項として以下の 3 つが挙げられる．

(1) 3 次元映像音声メディアを管理出来ること
(2) アプリケーションに非依存な形式で保存されていること
(3) ユーザが求める表現のレベルに応じてデータが取得できること

1 つ目は，3 次元映像音声メディアを管理するのに特化した形式で保存するということである．SDM で記録するデータはデバイスの位置情報や対象などにとどまらず，録音プロセスの情報なども含んでいる．このような SDM に



図 1　オーケストラ収録のカメラとマイクの配置



図 2　360Square のスクリーンショット

特有の情報を記述できる形式で保存する必要がある．

2 つ目は，様々なアプリ開発に利用可能な形式で保存するということである．SDM データを用いたアプリには，ユーザとインタラクティブな挙動を示すアプリや，パッケージメディアのミキシング，半自動ミキシングなどが考えられる．どのようなアプリでもデータを利用できるよう，特定の用途にスペシフィックではなく汎用的な形式で保存することが必要となる．

3 つ目は，目的に応じて柔軟にデータの取得が可能ということである．SDM が保有するデータには「イベント名」のような単純な情報から「録音に使用した 5 番目のマイクの位置座標」のような細かい情報まで含まれている．データベースは，ユースケースごとに異なるユーザのニーズに応えられるよう，全ての情報に対して柔軟にアクセス可能である必要がある．

上記の条件を満たすデータ形式として，本研究では RDF・LOD によるデータの管理を行う．その際，SDM の LOD を作成するために必要な語彙として，SDM Ontology を作成し，データの定義を行う．

次章では，要素技術である RDF/LOD の解説と，関連研究を紹介する．

3.　要素技術と関連研究

3.1　RDF・LOD

インターネットが普及した現代において，web 上に存在する無数のデータが検索によって容易に手に入るようになった．また，自分の持つデータを世界中の人々とシェアするために，オープンデータとしてインターネット上に公開することも可能となった．こうした状況において保有するデータをオープンデータとして広く利用してもらうには，誰もが容易に利用できる形式，または機械が自動的に処理可能な形式で公開することが望ましい．

データの公開方式の評価方法として，World Wide Web の考案者である Timothy Berners-Lee は 5 Star Open Data[*2]というものを提案している．これは，公開するデータの形式を 5 段階で評価したものであり，1 段階目から 3

[*2] http://5stardata.info/en/

段階目までは次のような形式での公開を表している．

(1) pdf や jpg での公開：文章データを jpg で公開するといった，自動処理が困難である形式での公開

(2) xls や doc での公開：特定のアプリケーション以外では前処理なしに扱うことのできない形式での公開

(3) csv や tsv での公開：データ解析において一般的に用いられている形式での公開

この指標の 4 段階目では RDF(Resource Description Framework)[*3]という形式を推奨している．これによって記述されたデータは URI によって参照可能であり，機械による自動処理に適している．この RDF が外部と連携したものは LOD(Linked Open Data)[*4]と呼ばれ，5 Star Open Data の最高レベルである 5 段階目となっている．LOD の形式でデータを公開することで web 上のデータ同士がリンクしあい，オープンデータの利用が推進される．

3.2 関連研究

RDF・LOD を用いたデータ管理の研究には，RDF・LOD を用いたアプリケーションの作成に関する研究や，データベースを構築するための語彙の研究などがある．

RDF・LOD で管理されているデータを用いたアプリケーションの研究例として，T. Di Noia ら [9] による映画のリコメンドシステムの研究を紹介する．この研究では，ユーザの嗜好を表す user profile という値と，映画間の類似度を表す similarities という値を組み合わせて学習させている．

ユーザの嗜好を計算するのに Noia らが用いたデータセットは，MovieLens[*5]というものである．これはリコメンデーションシステム開発やベンチマークのために作成された，映画のレビューについてのデータセットである．Noia らが用いたバージョンには，6,040 のユーザによる，3,952 の映画に対する，1,000,029 のレビューが含まれている．このデータセットから，各ユーザの嗜好を計算する．

また，映画間の類似度を計算するのに用いたデータは，DBpedia[10], LinkedMDB[11], Freebase[12] という LOD から抽出された，MovieLens で扱っている映画に関係する RDF のグラフである．類似度は，ある映画とある映画が「どれだけ同じ情報を共有しているか」によって計算される．例えば「主演が同じである」「ジャンルが同じである」といった内容が RDF で記述されている場合，高い類似度となる．

こうして RDF・LOD から計算された映画間の類似度と，ユーザの嗜好のデータを組み合わせることで，高精度なリコメンドシステムを構築することに成功している．

これにより，あるユーザがある映画に対してどのような評価を下すかを推定し，高評価を下すと推定された映画がリコメンドされることとなる．

また，語彙の定義に関する研究として，Y.Raimond ら [5] による The Music Ontology の研究がある．The Music Ontology は，音楽に関するメタデータを詳細に記述することを目的として構築された RDF Vocabulary である．The Music Ontology は，3 つの ontology を元に構築されている．1 つ目の ontology は the Timeline Ontology[6] である．これは時刻 (time instant)・時間間隔 (time interval)・時間の流れ (timeline) を表現することを目的として構築されたオントロジーであり“this performance happened the 9th of March, 1984”といったデータを RDF で記述することが出来る．

2 つ目は the Event Ontology[7] である．“this musical event was performed in Suntory Hall in Tokyo”のようなイベントの情報を記述するための ontology であり，場所や公式ホームページといったイベント関係の情報を記述することが出来る．

3 つ目は Friend-of-a-friend ontology[8] である．これは“the artist of this classic song is Johann Sebastian Bach”といった，人物に関係する情報や人間関係を記述するための ontology であり，様々な LOD において広く用いられている．

以上のオントロジーを元に，音楽に関係する情報を記述するための語彙を追加・再定義したものが The Music Ontology である．具体的には“performed_in”，“conductor”，“recorded_in”，“remix_of”などといった，演奏や楽曲の情報を詳しく表すための語彙が定義されている．

しかし，SDM で扱うデータには録音プロセスや録音デバイスの情報などが含まれており，これは Music Ontology ではカバーしていない範囲となる．SDM のデータを表す語彙は現状存在しないため，LOD 作成のためには新しくオントロジーを作成する必要がある．

4. SDM Ontology

4.1 オントロジー構築のアプローチ

本研究のアプローチとして，Fig.3 のような流れを想定している．まず，SDM が保有しているデータを LOD に変換するために，SDM 用の RDF Vocabulary である”SDM Ontology”を定義する．SDM Ontology の構造を議論し，構造を決定したのちに，これを実際に構築する．その際は構築支援ツールである protege[*6]というものを用いる．

SDM Ontology を用いてデータを LOD として記述したのちに，それを外部アプリケーションから参照できるよう公開する．公開方法としては，LOD 用の問い合わせ言語である SPARQL を投げられる SPARQL Endpoint というものを構築し，そこに LOD をアップロードすることで，

[*3] https://www.w3.org/RDF/
[*4] https://www.w3.org/DesignIssues/LinkedData.html
[*5] https://movielens.org/
[*6] http://protege.stanford.edu/

外部から問い合わせが可能となる．SPARQL Endpoint には，fuseki[*7]，Virtuoso[*8]，AllegroGraph[*9]といった様々な種類があるが，RDF 表記形式のサポート数などの観点から比較検討した結果，今回は StarDog[*10]という Endpoint を用いる予定である．

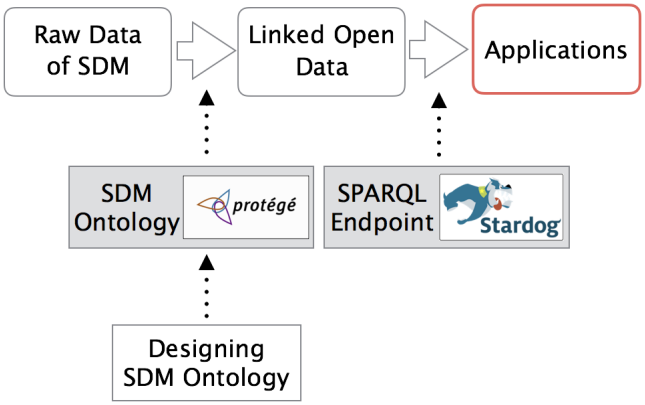


図 3 Approach for LOD of SDM

データの制作の利便性や，加工，利用における機能性や応用の自由度を高めるために，SDM Ontology は記録対象の構成や作業フローの実態に即したデータ構造を持つ必要がある．そこで，まず記録対象の空間を構成する要素を以下の 4 つに分解することを考える．

- Event: イベント全体の情報 (施設の詳細やプログラムの内容など)
- Target: 収録対象の情報 (対象の位置情報や対象の種類など)
- Recording: 収録行為の情報 (デバイスの位置情報やデバイスの種類など)
- Signal: 生成されるシグナル (シグナルの種類など)

以上の 4 要素によって空間は構成されていると考え，これらをカバーできるように SDM Ontology の構築を行なった．

4.2 オントロジーの詳細

今回考案した SDM Ontology の構造は Fig.4 の通りである．図中の色分けは，データベースを作成する際に便宜上取り決めたグループを表している．黄，青，緑，紫はそれぞれ Event, Target, Recording, Signal を表現する為のクラスにあたる．

まず，RecordingContext クラスがイベント全体を表現している．現在はコンサートのデータを表現するための構造となっており，コンサート (Concert) クラスがそこに繋がっている．また，そこから会場 (Facility) とそのコンサートの演目 (Program) にも関連付けてある．演目には楽曲 (Song) を関係づける．

次に，演目を演じる側のオブジェクトとして RecordingTarget クラスが定義されている．そのサブクラスとして，演奏を表すパート (Part)，楽器 (Instrument)，演奏者 (Player)，オーケストラ (Orchestara)，観客 (Audience)，スタッフ (Staff) などがある．また，対象の位置情報を記録するためのクラス (SDM Geometry) も定義している．

一方で三次元視聴オブジェクトを記録する行為及び記録物を定義するためのオブジェクトとして，RecordingEvent クラスと Signal クラスを定義した．RecordingEvent は記録機器 (Recorder) に紐づけられており，Recorder は収録対象 (RecordingTarget) を示すことが出来る．また，RecordingTarget と同様に Recorder は位置情報 (SDM Geometry) を持つ．Recorder には，コンフィグ情報を保存するクラス (RecordingSystem) と，記録の種類によって Recorder の種類ごとにクラスが用意されている．音楽を収録する場合は AudioRecorder クラス，映像を記録する場合は VideoCamera クラス，それ以外のセンサーデータ等を記録する場合は DataRecorder クラスを用いることを想定している．また，AudioRecorder には Mic クラスをつなげ，録音データの内容を表現する手がかりとしている．InputChannel クラスは Audio Mixer 等の音声加工装置 (機能) を抽象化したクラスであり，必要に応じて Mic からの音声を加工して AudioRecorder に出力する．

Recorder クラスが音声，映像，それ以外のデータの記録用のクラスを持つため，書き出される Signal クラスにも同様にそれぞれに対応したクラスが存在する．音楽に関する Signal は AudioSignal クラスに，映像に関するものは VideoSignal に，それ以外は DataSignal へ格納される．また，Signal クラスは録音データの格納場所のリンクを情報として持ち，このリンクを辿ることで録音データを取り出すことを想定している．

こうしてデザインを行なったオントロジーをファイルとして記述して web 上に置くことで，SDM Ontology の語彙の定義を外部から参照することができ，SDM Ontology を用いた LOD の作成が可能となる．そのためのファイルの作成は，前述の protege を用いて行う．

5. 結論と今後の展開

本論文では，SDM コンソーシアムが保有するデータを LOD として表現するための SDM Ontology の構築について説明した．まず SDM の活動の目的と，実際に行なっている活動について述べた．データの記録作業について，慶應で行なった収録の説明を行ったのちに，そこで得たデータの説明を述べた．つづいて，そのデータを用いたアプリケーションとして”360Square”を紹介し，さらにデータを格納・公開して外部から利用してもらうためのデータベースの要求事項を整理し，実装として RDF/LOD を用いるこ

*7 http://jena.apache.org/documentation/serving_data/index.html
*8 https://virtuoso.openlinksw.com
*9 http://franz.com/agraph/allegrograph/
*10 http://stardog.com/

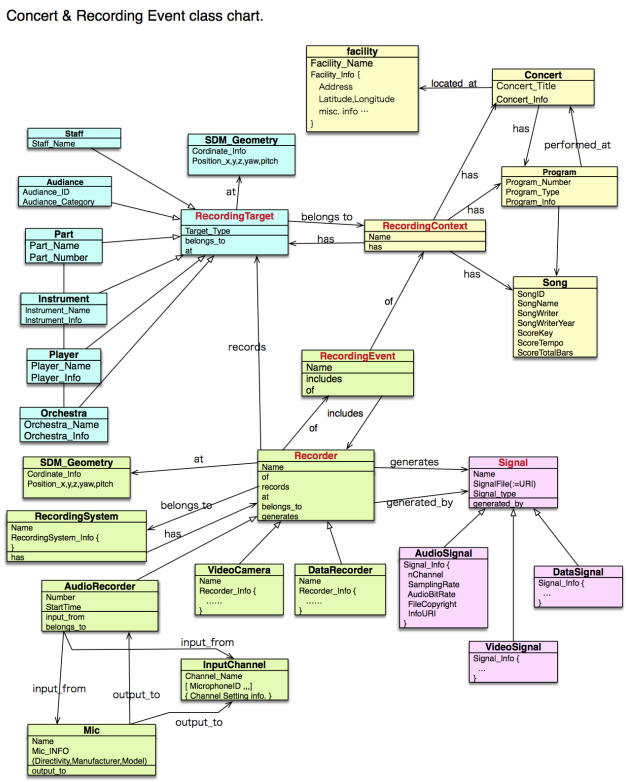


図 4　SDM Ontology のデザイン

とを説明した．さらに，実際に SDM のデータを表現する枠組みとして，記録対象の空間を Event, Target, Recording, Signal の 4 要素に分割して考え，それぞれを表現する為のクラスとプロパティについて議論し，SDM Ontology の構築を行なった．SDM Ontology の構築により，SDM が保有するデータを LOD として公開することが可能になり，データの利用が促進される．

今後の展開として，まだ議論が不十分といえる SDM Ontology の詳細を議論していく．RecordingEvent や Signal の部分は，映像やその他のデータを保存するための拡張も含めて仕様が固まりつつあるが，RecordingContext と RecordingTarget に関してはまだ詳細が詰められていない．例えば，RecordingContext は現在コンサートのデータを保存するのに特化したクラス構造となっている．SDM は記録対象として，コンサートに限らずライブであったりそれ以外のパフォーマンスであったりと多様なものを想定している．それらを包括的に管理できるよう，SDM Ontology はコンサートに特化せず汎用的なクラス構造を持つ必要がある．また，RecordingTarget については，クラス間の関係にあたるプロパティの詳細などが決定していない．また，これは人を表すクラスにあたるため，SDM Ontology で定義せずとも既に存在するオントロジーで表現可能な部分が存在する．RDF の思想として，データを公開する人が各々の語彙を勝手に定義するとデータの共有が難しくなるため，出来る限り既存の語彙を使用することが推奨されている．SDM Ontology の構築に際しても，そのことに注意を払いながら構築を進める必要がある．

デザインしたオントロジーを LOD から参照するためにファイルとして書き出す protege の使用方法の調査や，LOD をアップロードし公開するための SPARQL Endpoint のサーベイは終了しているため，今後は SDM Ontology が出来上がり次第，データを LOD で記述し、公開を進めていく．SDM ではデータの使用方法として複数のユースケースを想定しているため，LOD での公開が完了したのちに，360Square 以外の様々なアプリケーションからも実際に使用できるかの検証を進めていく予定である．

参考文献

[1] 塚田学, 小川景子, 池田雅弘, 曽根卓朗, 丹羽健太, 齊藤翔一郎, 粕谷貴司, 砂原秀樹, 江崎浩 (2017)「Software Defined Media: 視聴空間サービスのソフトウェア制御」『日本ソフトウェア科学会学会誌『コンピュータソフトウェア』「ネットワーク技術」特集』

[2] Manabu Tsukada, Keiko Ogawa, Masahiro Ikeda, Takuro Sone, Kenta Niwa, Shoichiro Saito, Takashi Kasuya, Hideki Sunahara, and Hiroshi Esaki ”Software Defined Media: Virtualization of Audio-Visual Services”, IEEE International Conference on Communications (ICC2017), May 2017

[3] Masahiro Ikeda, Takuro Sone, Kenta Niwa, Shoichiro Saito, Manabu Tsukada, and Hiroshi Esaki. New recording application for software defined media. In Audio Engineering Society Convention Paper, 141st AES Convention, Los Angeles, USA, September 2016.

[4] 塚田学, 菰原裕, 新居英明, 粕谷貴司, 高坂茂樹, 小川景子, 江崎浩 (2017)「SDM360^2:音楽イベントのための自由視聴点映像音声のインタラクティブ再生」『マルチメディア，分散，協調とモバイル (DICOMO2017) シンポジウム』

[5] Y. Raimond, S. Abdallah, M. Sandler, F. Giasson, ”The music ontology”, In: Proceedings of the International Conference on Music Information Retrieval, pp. 417-422, September 2007

[6] Y. Raimond and S. A. Abdallah, “The timeline ontology”, OWL-DL ontology, 2006. [Online]. Available: http://purl.org/NET/c4dm/timeline.owl

[7] Y. Raimond and S. A. Abdallah, “The event ontology,” OWL-DL ontology, 2006. [Online]. Available: http://purl.org/NET/c4dm/event.owl

[8] D. Brickley and L. Miller, “Foaf vocabulary specification,” Working draft, 2005. [Online]. Available: http://xmlns.com/foaf/0.1/

[9] T. Di Noia, R.Mirizzi, V. Claudio Otuni, D.Romito, M.Zanker, “Linked Open Data to support Content-based Recommender Systems”, I-SEMANTICS, Sept. 2012

[10] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. Dbpedia - a crystallization point for the web of data. Web Semant., 7:154-165, September 2009.

[11] O. Hassanzadeh and M. P. Consens. Linked Movie Data Base. In Proceedings of the WWW2009 Workshop on Linked Data on the Web, April 2009.

[12] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. In SIGMOD ’08, pages 1247-1250. ACM, 2008.