



テレビ番組 ナショナルアーカイブに向けた メタデータモデル構築 「番組」の階層構造に着目して

第1回DAフォーラム

2021年12月19日

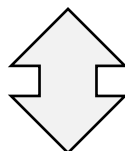
慶應義塾大学大学院文学研究科博士課程

関根禎嘉

sekine.sadayoshi@keio.jp

研究の背景

- 映像配信サービスの普及による放送されたテレビ番組のアーカイブ利用の拡大
- ジャパンサーチでのテレビ番組利用
 - 放送番組センター「放送ライブラリー公開番組データベース」
 - NHKアーカイブス



- 一方で、日本国内におけるテレビ番組の包括的なアーカイブは存在しない
 - アーカイブ各機関はメタデータ作成指針やモデルを基本的に公開していない
 - 研究目的
 - **テレビ番組のアーカイブ利用を前提とするメタデータモデル構築を行うこと**
 - ナショナルアーカイブ構築の基礎となることを目指す
-

研究方法

テレビ放送番組アーカイブから番組を記述する要素を抽出し、
それに適合するメタデータモデルを設計

公開されているデータベースの
出力などから、採用されている
モデルや記述規則を遡及的に推測

放送番組を対象とした
メタデータモデルの事例を
モデル構築のために参照

※相互運用性に鑑み、RDFで構築

※既存の番組アーカイブのメタデータモデルの推定は、本研究の目的に沿った、既存事例に関する調査として行う

「番組」とは

- 基本的だが明確な範囲の存在しない概念
- 本来は「演芸・映画・相撲などの演目・組み合わせや、その順序、またはそれを書いたもの」
- 「番組」に書かれた対象が「番組」と呼ばれるようになり、もとの「番組」は「番組表」と呼ばれるようになるという意味の転倒が起こった

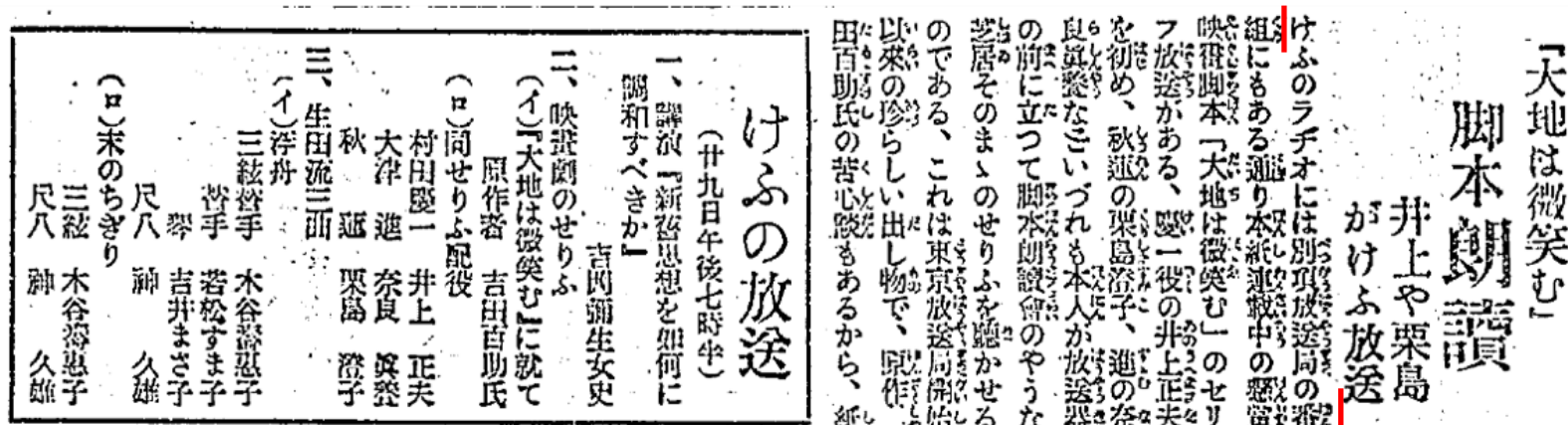


図1 ラジオ放送開始当初の新聞紙面¹⁾

1) “「大地は微笑む」脚本朗読 井上や栗島がけふ放送”. 朝日新聞. 1925年4月25日, 朝刊, 東京版, 11面

「番組」とは

- 「番組」と呼ばれるもの
 - 1回の放送
 - 一連のまとまり(例:放送初回から最終回まで)
 - テレビやラジオなど放送メディアによって流通するコンテンツ全般
 - ネット配信番組で視聴できる映像コンテンツ
 - テレビ番組にとっての時間編成
 - 「番組」として編成されることになった
 - 「連続ドラマ」という形式の成立
 - シリーズの組成
 - メタデータモデルはこうした特性を反映する必要
-

メタデータモデル事例

- BBC Programmes Ontology (BBC PO) ²⁾
 - 放送の実体を4つのドメインに整理
 - Content
 - Medium
 - Publishing
 - Temporal annotations
 - Programmeクラス
 - Contentドメインの最上位クラス
 - Brand, Series, Episodeの各クラスをサブクラスに持っている
 - Programmeの階層関係を表現

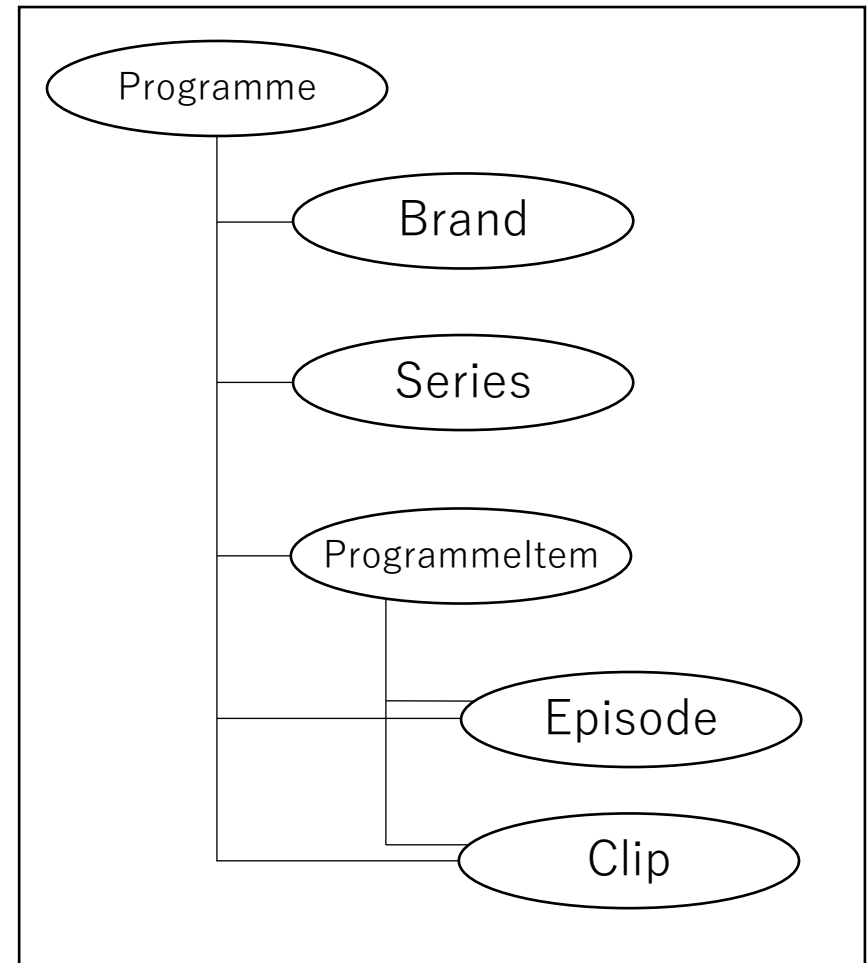


図2 Contentドメインの階層構造

2) 英国放送協会. “Programmes ontology”. Ontologies. 2009-02-20.
<https://www.bbc.co.uk/ontologies/po>, (accessed 2021-11-03).

メタデータモデル事例

- EIDR³⁾
 - Entertainment ID Registry
 - 映像コンテンツにユニークなIDを付番する体系

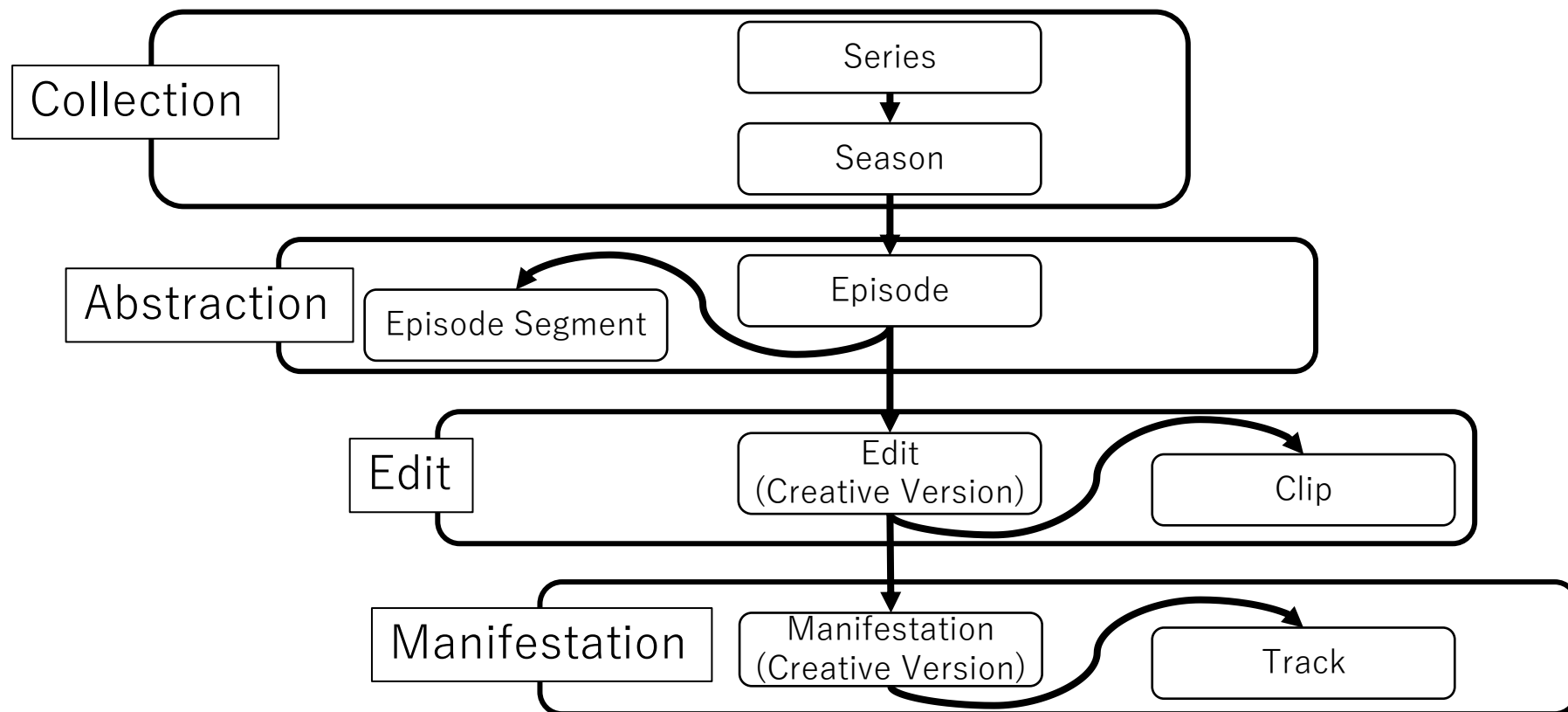


図3 EIDRにおける階層構造

3) Entertainment ID Registry Association. Introduction to the EIDR Data Model.

<https://www.eidr.org/documents/Introduction%20to%20the%20EIDR%20Data%20Model.pdf>, (accessed 2021-11-03) .

メタデータモデルの構築

- 放送ライブラリー⁴⁾公開番組データベースを例に
 - 国内唯一の放送番組専門のアーカイブ施設
 - NHK, 民間放送局のテレビ, ラジオ番組およびCMを保管
 - 所蔵番組の検索サービスをウェブ上で提供
- DB検索によって出力されるデータ項目

番組名	種別	番組ID	放送日	分数
ジャンル	放送局	製作者	制作社	出演者
スタッフ	概要	受賞歴		

- 各項目の記述内容の定義は説明されていない
- 番組名と種別…項目名として明記されていないため便宜的に設定
- 出演者名…人名の羅列
- スタッフ…番組に関与した人物名が役割とともに列挙
- 概要…番組の内容説明, 放送期間など雑多な内容

4) 放送番組センター. 放送ライブラリー.<https://www.bpcj.or.jp/>, (参照 2021-11-03)

放送ライブラリー公開番組データベース

- 右図…「土曜ドラマスペシャル」をクエリとして検索した例(一部)
- 検索結果に表示された各レコードの相互関係は不明
 - 相互の関係は概要のテキストを参照するか、あるいは当該データベース外の知識を参照することによって把握することになる。
- 放送された1話に対して番組IDを付番し、視聴の単位としている。
 - 本発表では、この放送された1話のことを「エピソード」と呼ぶことにする。
 - 複数話が放送されて一連のシリーズやシーズンをなしていることが明らかな番組であっても、シリーズやシーズンに相当する単位があるわけではない。

テレビ番組	
番組ID▼▲	タイトル▼▲
008473	土曜ドラマスペシャル 堀の中の懲りない面々 希望の出所
004521	土曜ドラマスペシャル 東京恋物語
004599	土曜ドラマスペシャル 恋子の毎日2
008476	土曜ドラマスペシャル ポルノ女優小夜子の遺産
005251	土曜ドラマスペシャル 四万十川 あつよしの夏
005247	土曜ドラマスペシャル 風よ、鈴鹿へ
007016	土曜ドラマスペシャル ラストバラードは君に
008482	土曜ドラマスペシャル あなたが大好き
008486	土曜ドラマスペシャル カラダ記念日
007467	土曜ドラマスペシャル 純愛せんせい取材日記 覗き部屋の女
007017	土曜ドラマスペシャル 新・東京物語
006996	土曜ドラマスペシャル まくりの勝っちゃん一発逆転
008490	土曜ドラマスペシャル スティル・ライフ 霧子とマリエの犯罪の同棲生活
008506	土曜ドラマスペシャル ポルノ女優小夜子の最後の冒険 ヴェルーシュカって何！？
209123	土曜ドラマスペシャル 蝶々さん ～最後の武士の娘～ 前編 誇りの代償
209124	土曜ドラマスペシャル 蝶々さん ～最後の武士の娘～ 後編 遠いアメリカ
209128	土曜ドラマスペシャル とんび 前編
209129	土曜ドラマスペシャル とんび 後編
209131	土曜ドラマスペシャル それからの海
210090	土曜ドラマスペシャル 永遠の泉
1-20 / 22件中	

図4 放送ライブラリー公開番組DBの検索結果（一部）

放送ライブラリー公開番組データベース

- 右図：放送ライブラリーDBの記述内容に対して、BBC PO（接頭辞「po:」）の階層構造を適用しRDFグラフ化
- 前ページの検索結果のうち、前後編の2話から構成されている「土曜ドラマスペシャル とんび」を例に採用
- Programmeクラス
 - 「とんび（NHK版）」
 - 操作的な実体
 - 同じ原作を持つ同名のドラマが他局に存在するために区別した名称
- Brand, Series, Episodeの各クラス
 - Programmeのサブクラス

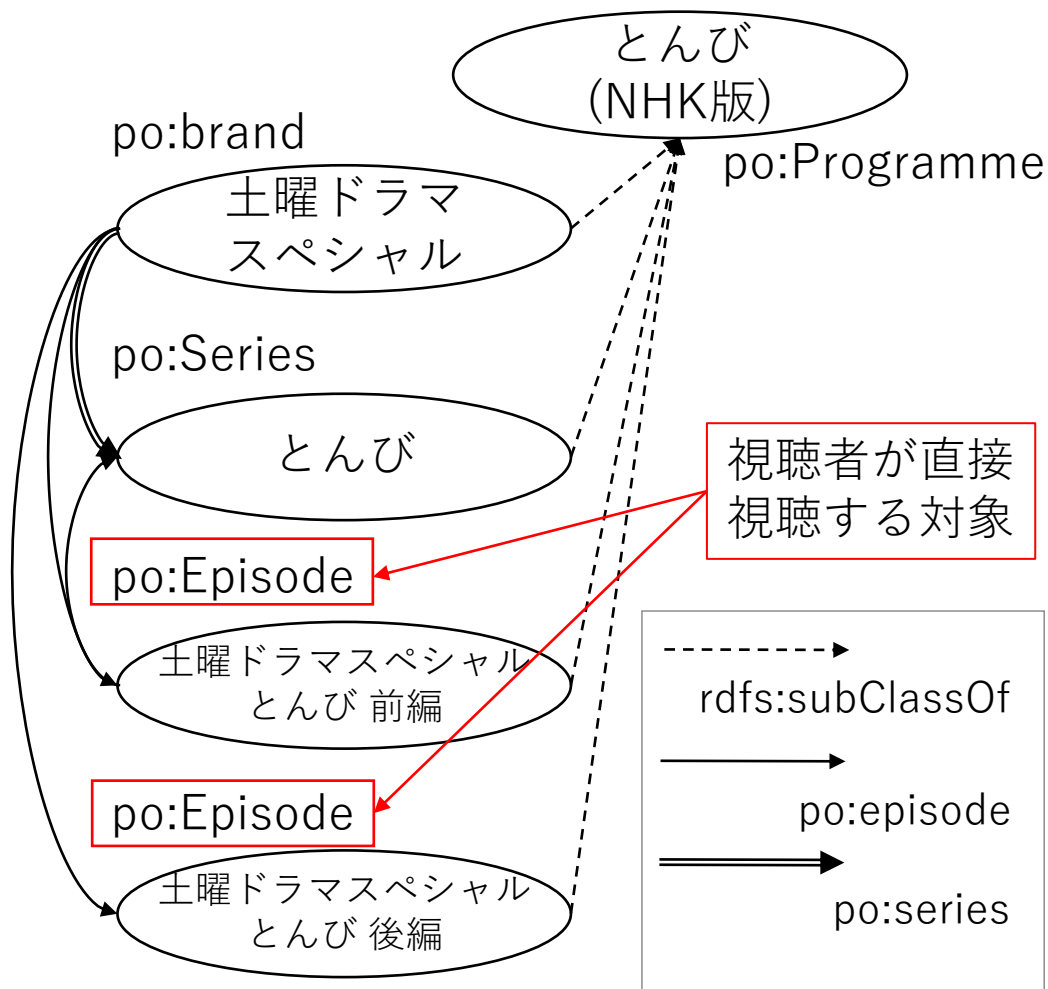


図5 放送ライブラリーDBの「土曜ドラマスペシャル とんび」のRDFグラフ

構築したメタデータモデル

- 右図：構築したメタデータモデル
- 放送ライブラリーDBのデータの特性を考慮し、必要に応じて独自語彙を設定して構築
 - 関与者情報の厚い記述
 - 利用者にとって番組へのアクセスポイントとして機能する、放送局や放送日時
- コンテンツ (Contents), エージェント (Agent), 放送イベント (Event) の3ドメインを設ける
- コンテンツドメイン
 - 「番組」クラス…基礎的な概念
 - 「シリーズ」「シーズン」…とともに複数のエピソードの一連
 - 「シーズン」は、「シリーズ内のエピソードの一連」として「シリーズ」と区別する

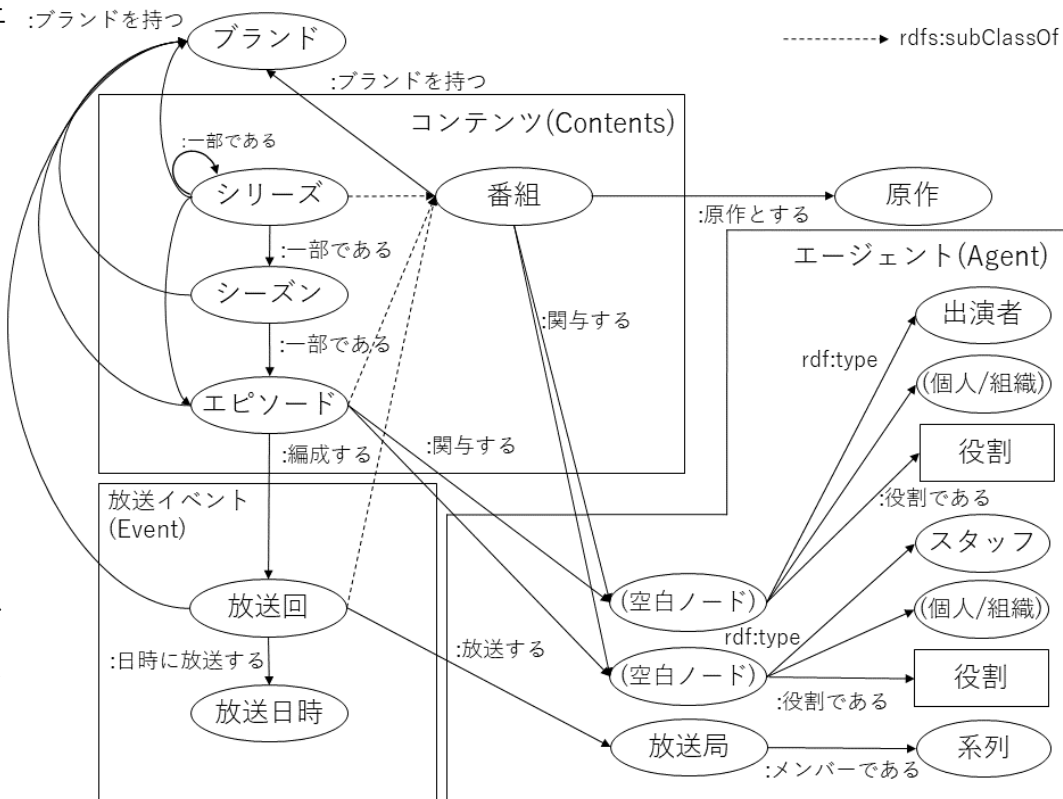


図6 放送ライブラリーDBデータのメタデータモデル

放送ライブラリー公開番組データベース

- 放送イベントドメイン
 - 特定のエピソードが特定の日時および放送局で放送されるその一回をひとつの出来事 (Event) と捉える
- エージェントドメイン
 - 関与者を表す
 - 個人 (自然人) である出演者やスタッフを含める
 - 番組クラスと出演者クラスとはプロパティ「出演する」(「関与する」のサブプロパティ), スタッフクラスとはプロパティ「関与する」, 放送局クラスとはプロパティ「放送する」(「関与する」のサブプロパティ) によって関連づける。
 - 出演者およびスタッフ
 - 役割記述はリテラルで表現

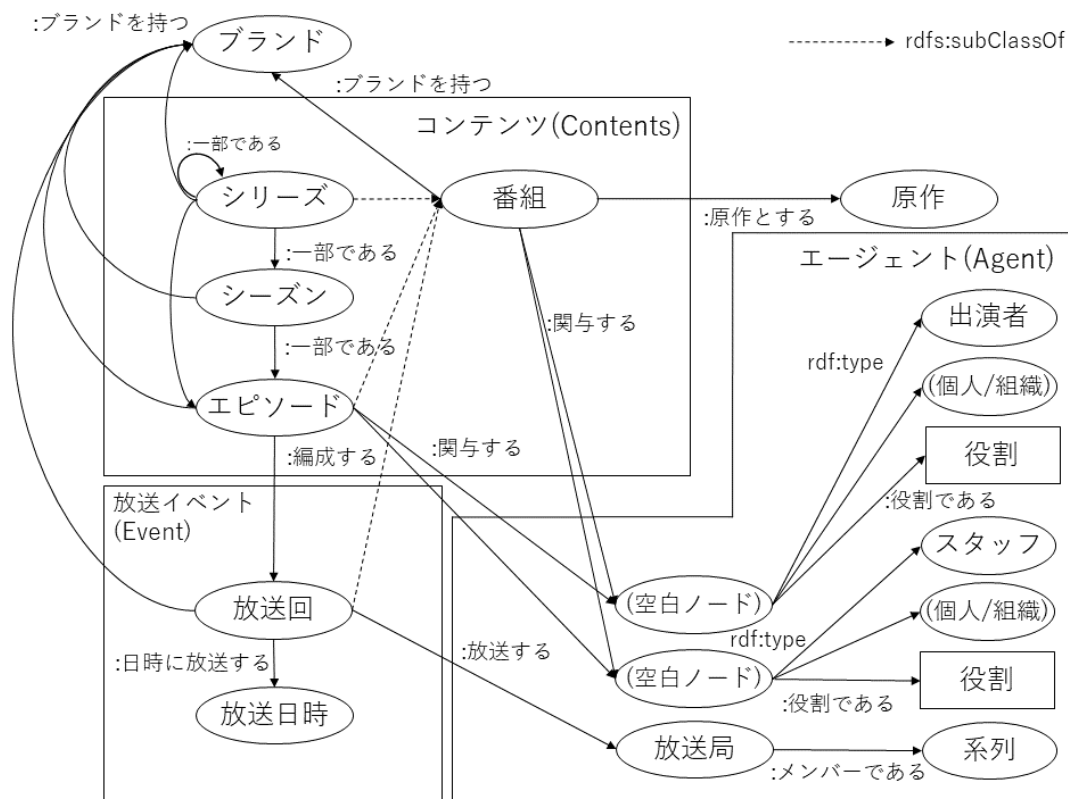


図6 放送ライブラリーDBデータのメタデータモデル (再掲)

まとめ・今後の課題

- 本研究で構築したメタデータモデルは「番組」クラスを中心的な実体として設定
 - このクラスのインスタンスが最も基本的な探索の対象となることを想定
 - 放送ライブラリーDBではすべてのエピソードが同一階層に存在するように見えるが、「番組」とそのサブクラスに構造化することにより、利用者が目的とするコンテンツの識別を容易にすることを旨とする
 - 今後の課題
 - 提案したモデルのさらなる精緻化
 - 提案したモデルの利用者タスクに鑑みた評価
 - 国内外の事例探索
 - 例えば、放送局局内事例
-



テレビ番組 ナショナルアーカイブに向けた メタデータモデル構築 「番組」の階層構造に着目して

第1回DAフォーラム

2021年12月19日

慶應義塾大学大学院文学研究科博士課程

関根禎嘉

sekine.sadayoshi@keio.jp
