

# パーソナルデータストアを用いた 視聴データ利活用アプリケーションの試作

山上 悠喜† 山村 千草† 松村 欣司† 藤井 亜里紗†

NHK 放送技術研究所†

## 1 まえがき

テレビ番組の視聴が購買や旅行などの生活行動に影響したり、逆に Web の利用や移動などの行動がテレビ番組への興味に影響したりすることがある。これらの視聴および行動に関するデータを結びつけることで、放送から得た情報を視聴者にとって必要なときに便利に活用したり、新たな放送コンテンツへの接触につなげたりといった効果を期待できる(図1)。しかし、放送事業者が集めた視聴者のデータを多様なサービスに用いることは、個人情報保護に関するガイドライン[1]等を踏まえた対応が必要であり、システム設計や運用面で慎重な検討を要する。

サービス提供者がパーソナルデータを保持せずにパーソナライズドサービスを実現できる形態として、各ユーザが主体的に管理するパーソナルデータストア(PDS)を用いてパーソナルデータをサービス横断的に集約する、ユーザセントリックな管理モデルがある。欧州を中心に検討が進められており[2]、ユーザが自身のデータの提供先などを管理できる、異なるサービス間で柔軟にデータを連携できるといった特長がある。

筆者らは、ユーザセントリックな管理モデルにより、視聴者の安全・安心を担保した上で視聴データと行動データを結びつけて利活用する技術を検討している。今回、これらのデータをPDSに蓄積し利活用するためのデータ構造とインタフェースを設計し、アプリケーションを試作して実現性を確認した。

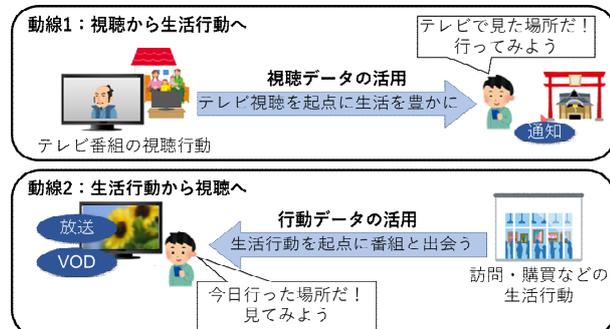


図1 視聴データと行動データの連携イメージ

Prototypes for Utilization of Viewing Data with Personal Data Stores

†Yuki Yamakami †Chigusa Yamamura †Kinji Matsumura

†Arisa Fujii

†NHK Science & Technology Research Laboratories

## 2 システムアーキテクチャの検討

視聴データを利用するにあたってのシステム要件として、以下の3点が挙げられる。

- ・ユーザが自身のデータを主体的かつセキュアに管理できる
  - ・データをどの相手に開示するか制御できる
  - ・他業種各種サービスと幅広く連携できるようなデータ構造をもつ
- PDSを実現するためのオープンソースプロジェクトのひとつであるSolid[3]は、以下の特長をもち、先に述べたシステム要件と合致する。
- ・ユーザは、自身のPDS内のデータの開示先を自分で設定でき、家族や友人などあらかじめ許可した相手にだけデータを参照させることができる
  - ・データは、W3C(World Wide Web Consortium)標準のRDF(Resource Description Framework)によって、リンクトデータとして構造化され、世の中の様々なデータと容易に連携できる。

筆者らは、視聴者主体で視聴データを蓄積・管理し、多様なサービスから視聴データを活用できるシステムを設計し、PDSの実装にはSolidを用いる構成とした。図2にシステム構成を示す。

### 2-1 データ構造

視聴データが含むべきデータとして、視聴操作(リモコン操作)の履歴と、視聴した番組情報を規定した。操作履歴の取得方法はハイブリッドキャストコネクタ(ハイコネ)の活用[4]やデータ放送の活用などが考えられる。番組情報はEPG(Electronic Program Guide)から得ることを想定した。

得られたデータはW3C標準のRDF記述手法であるTurtle形式としてPDSに格納する。語彙定義には、他のデータとの連携を考慮しSchema.org[5]を用いた。さらに、予めデータ中の全文字列を対象にキーワードを抽出・保存しておくことによって、外部アプリケーションがキーワードマッチングするたびにキーワードを抽出する手間を省き、処理を高速かつ効率的に行えるようにした。図3に視聴データの例を示す。

## 2-2 インタフェース

Solid では、PDS 内のデータを参照する際は、RDF 問い合わせ言語のひとつである SPARQL Protocol and RDF Query Language を用いることができるが、外部アプリケーションが個別の方法でデータに直接アクセスすることは効率面、セキュリティ面で適さない。そこで、外部アプリケーションからのデータの参照は、API を作成しこれを經由することとした。このうち、視聴データについて以下の関数を定義した。

- 任意の期間内の視聴データを取得
- 任意の期間内かつ任意のキーワードを含む視聴データを取得

レスポンスデータは、汎用性を考慮し JSON-LD(JavaScript Object Notation for Linked Data)形式とした。

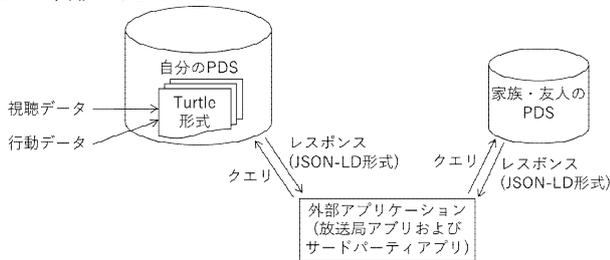


図2 システムの全体構成

```

@prefix : <#> .
@prefix c: <https://shuheitaguchi/profile/card#> .
@prefix schema: <http://schema.org> .

:[視聴行動id]
  a schema:WatchAction ;
  schema:agent c:me ;
  schema:instrument "HC" ;
  schema:startTime "2020-07-15T12:50:00+09:00" ;
  schema:object [番組id] .

:[番組id]
  a schema:TVEpisode ;
  schema:identifier "20200715147401024" ;
  schema:name "【連続テレビ小説】エール (15) 『いばらの道』";
  schema:description "ダンスホールで出会った志津(堀田真由)と";
  schema:timeRequired "900" ;
  schema:actor "窪田正孝", "中村蒼", "相島一之" ;
  schema:keywords "エール", "いばらの道", "ダンスホール" ;
  schema:genre "情報/ワイドショー" ;
  schema:publication [放送イベントid] .
    
```

図3 視聴データ(Turtle形式)の例(一部)

## 3 試作アプリケーションによる検証

PDS のデータを参照する外部アプリケーションとして、2種類の Web アプリケーションを試作し、システムの有効性を検証した。視聴データはハイコネを用いた視聴実験[4]によって取得したものを、行動データは検証のために仮に作成したものをを用いた。

### 3-1 放送局アプリ

放送事業者が開発・配布することを想定したアプリケーションを試作した。具体的には、以

下の機能を実装し、これらが 2 章で示したデータ構造およびインタフェースを用いて、放送事業者がパーソナルデータを収集せずにパーソナライズドサービスを実現できることを確認した。

- 視聴データに含まれるキーワードを基に、関連する Web コンテンツを提示する機能。
- 家族や友人の視聴データに含まれる視聴操作の履歴を取得し、追体験できる機能。
- 自分と家族や友人の視聴データを比較し、一致した番組を通知する機能。
- 視聴データや行動データと先物の EPG データをキーワードマッチングし、放送予定の番組視聴を推奨する機能。

### 3-2 サードパーティアプリ

放送事業者以外の第三者が開発・配布する想定 of アプリケーションを試作した。旅行会社が提供するサービスを想定して、視聴データに含まれる番組に関連した場所を地図上に表示する機能を実装し、事業者を問わず視聴データを活用したサービスを実現できることを確認した。

## 4 むすび

ユーザが主体的に管理するパーソナルデータストアに視聴データを蓄積し活用するためのデータ構造とインタフェースを設計した。さらに、オープンソースプロジェクトのひとつである Solid の仕様によって実装し、放送事業者がパーソナルデータを収集せずにパーソナライズドサービスを実現できることと、事業者を問わず視聴データを活用したサービスを実現できることを試作アプリケーションで確認した。

今後は、家族や友人に限らない不特定他者との安全なデータ連携の技術検討・実装を進める。

## 参考文献

- [1] “放送受信者等の個人情報保護に関するガイドライン,” 総務省告示第159号, 2017.
- [2] “BBC: Introducing the BBC Box - BBC R&D,” <https://www.bbc.co.uk/rd/blog/2019-06-bbc-box-personal-data-privacy>, 参照 Dec. 15, 2020.
- [3] “Solid Project,” <https://solidproject.org>, 参照 Dec. 15, 2020.
- [4] 山村千草ほか, “家庭環境下でのスマホ連携によるテレビ視聴関連行動データの収集と分析,” DICOMO 2020 シンポジウム論文集, pp. 155-160, 2020.
- [5] “Schema.org,” <https://schema.org>, 参照 Dec. 15, 2020.