

# 企業コードとXBRLデータのLOD化

## Making LOD of XBRL Data and Company Codes

小出誠二<sup>1\*</sup> 加藤文彦<sup>1</sup> 小林巖生<sup>2</sup> 大向一輝<sup>3</sup> 武田英明<sup>3</sup>  
 Seiji Koide<sup>1</sup> Fumihiko Kato<sup>1</sup> Iwao Kobayashi<sup>2</sup> Itsuki Ohmukai<sup>3</sup> Hideaki Takeda<sup>3</sup>

<sup>1</sup> 情報・システム研究機構 国立情報学研究所

<sup>1</sup> Research Organization of Information and Systems, National Institute of Informatics

<sup>2</sup> 有限会社スコレックス

<sup>2</sup> Scholex Co., Ltd.

<sup>3</sup> 国立情報学研究所

<sup>3</sup> National Institute of Informatics

**Abstract:** The eXtensible Business Reporting Language (XBRL for short) is an international standardized reporting language based on XML in financial affairs. In Japan, diverse dialects of XBRL is utilized as EDINET by Japanese Financial Services Agency, as TDnet by Tokyo Stock Exchange, etc. The infusion of Linked Data into XBRL enables us not only to promote the synergy among these data in respective XBRLs, but also to link such financial data to non-financial Linked Data, i.e, industrial statistics, energy statics, economic census, etc. However, EDINET and TDnet use their own company codes, which are different from each other. Therefore, we also made Linked Data of company codes in order to link respective data mutually, using the legal personality registration codes to Ministry of Justice so that a registration code mediates between a relevant EDINET code and a TDnet code. The Linked Data technology is available to unify diverse aspects of company activities.

## 1 はじめに

XBRL (eXtensible Business Reporting Language) は国際的に普及している電子化された財務情報の報告形式であり、XML Schema や XLINK などの XML 技術をベースとし、日本では金融庁への報告 EDINET、東京証券取引所への報告 TDnet、日本銀行への報告のほか、国税庁 e-Tax、地方税 ELTax などの実現に利用されている。このように、すでに財務報告分野で実用化され、企業における財務データの生成から所管部署におけるデータハンドリングまで、一貫した処理がコンピュータ上で可能となっているにもかかわらず、さらにそのデータをリンクト・オープンデータ (LOD) 化することには価値がある。それは EDINET データと TDnet データなどの異なる XBRL データの内容を比較したり、XBRL ベースのデータを (工業統計データなど) 他分野のデータと一緒にシームレスに連携利用することができるようになるからである。

本報告では、日本の 2012 年度会計期間期末における

代表的業種 10 の売上トップ 5 社の金融庁への財務報告形式 EDINET と東京証券取引所への TDnet 報告形式のデータを LOD 化した結果について述べる。<sup>1</sup> また、2013 年 12 月以降報告の EDINET の仕様はインライン XBRL へ大きく変更されたため、2014 年末に入手可能な 2013 年度会計期間期末の同 EDINET 報告データの LOD 化も行った。<sup>2</sup> 本報告では、2012 年度会計期間期末報告の EDINET を旧 EDINET、<sup>3</sup> 2013 年度会計期間期末報告の EDINET を新 EDINET<sup>4</sup> と称する。

本研究は XBRL に代わるものとしてのリンクトデータ化を意図して行ったものではない。今日ではリレーショナル・データベースや LDAP をはじめ、CSV や XML など世の中にはすでに多くの構造化データやフォーマットがあるが、複数分野の異なる構造化データ

<sup>1</sup> EDINET および TDnet の 50 社の 2012 年度会計期間期末データにおける LOD 化は、平成 25 年度経済産業省「オープンデータ推進に関する調査研究」の一部として実施された。

<sup>2</sup> EDINET 報告のインライン XBRL 化に伴い、報告形式が変わると同時にそのファイル命名ルールやタクソノミ構成は 2012 年度と 2013 年度では大きく異なっている。今後の報告はこのいわゆる「次世代 EDINET タクソノミ」に従って行われるものと思われる。

<sup>3</sup> [http://www.fsa.go.jp/search/20130301/02\\_b.pdf](http://www.fsa.go.jp/search/20130301/02_b.pdf) を参照のこと

<sup>4</sup> <https://disclosure.edinet-fsa.go.jp/download/ESE140112.pdf> を参照のこと

\*連絡先：情報・システム研究機構、国立情報学研究所  
 〒 101-8430 東京都千代田区一ツ橋 2-1-2  
 E-mail: koide@nii.ac.jp

を互いに流通させようとする、原理的にはN種の構造化データにおいてN×Nの変換が必要となる。しかし、これらのすべてをリンクトデータ化することで、RDFへのN個の変換を用意するだけでよいことになる。XBRLデータをLOD化することで、エネルギー統計や産業統計のデータと財務報告データの関係を探索することが可能になる。

XBRLに代わることを意図したものではないが、本報告ではリンクトデータ化にあたって、極力XBRLの形式と元来の意図を尊重する態度をとった。それにより、XBRLの専門家のリンクトデータへの参入を容易にし、XBRLの特徴と利点にLODの特徴と利点を加えることができると考えたためである。これは表形式データのLOD化のためのW3C勧告RDF Data Cube Vocabulary<sup>5</sup>が、統計データに関するISO標準であるSDMX<sup>6</sup>をベースにその特徴を生かすように開発された経緯と同様の趣旨である。

EDINETで用いられる企業コードとTDnetで用いられる企業コードは異なっている。そのほか現状では複数の独自企業コードが民間および官公庁で用いられている。近い将来、企業マイナンバーとも言われる法人コードが実施される予定であるが、それが普及すれば、それをハブとしてLOD技術によって従来の多様な企業コードと関連情報を互いに連携することが容易となる。法人コード実施時の企業コード連携を目標に、今回2014年2月時点での上記50社の登記情報(crdata)を法務局から取り寄せ、会社法人等番号をハブとしてEDINETおよびTDnetの三者の企業コードをLOD化した。

## 2 LOD化の基本方針

XBRL形式データのLOD化にあたっては、以下に示す諸点について、注意を払った。

1. XBRLではデータの表示に関する情報もHTMLデータや表示用メタデータとして存在するが、LODではそのような情報は本来意味がない。表のディメンションも含め今回それらの情報はRDF化されていない。
2. XBRL報告書は複数のファイルからなるファイルセットであるが、EDINETおよびTDnetともにメタデータを用いた命名ルールがあり、ファイル名によって企業、報告種類、会計期間が分かり、自分のほしい情報が分かれば、唯一に個々のファイルを同定できるようになっている。RDF化においても、このファイル名をそのまま

各RDFファイル名(リソース名)に生かすよう変換した。

3. 旧EDINETあるいはそれに対応する旧TDnetに用いられる勘定科目は、標準タクソノミと、業種別タクソノミと、企業別タクソノミから成っているが、これらのタクソノミについてもメタデータの略号を用いたディレクトリ構造とファイル名になっている。ところが新EDINETでは府令による報告書式に依存したタクソノミ構成となり、名前空間抜きのタクソノミ名称は同一でも、タクソノミディレクトリ構造、ファイル名、名前空間も旧EDINETと共通性はなくなった。今回新旧それぞれの名前空間の違いをそのまま尊重してLOD化を行った。ただし、それらのスキーマに相当するタクソノミファイルについてのLOD化は今回未実施である。
4. 一つのXBRL報告書内では同じ勘定科目名でも複数のエントリが異なるコンテキスト(当年度/前年度/前々年度、連結/個別など)を伴って存在する。したがって、勘定科目名とコンテキストを指定してデータを同定しなければならない。さらに、LOD化において同一企業でも複数年度のXBRL報告書をRDFストアに収めてフラットに取り扱うことにすれば、データを指定するには何年度のXBRL報告書に記載のデータかも指定する必要がある。そこで、EDINETやTDnetの報告書自体にもURLを振り、RDFノードとして存在させると同時に、報告書内のコンテキストを各該当報告書ごとにLOD化し、さらに上記XBRL報告書ノードへのリンクもコンテキスト属性に持たせることで、企業ID、報告年度、勘定科目、報告書内コンテキスト等を指定して、望みのデータをSPARQLエンドポイントから得られるようにした。
5. データをRDF化する際にURLでグローバルに唯一の名前を個々のデータにつけずに、ブランクノードとすることもできるが、その場合にはそれを外部から直接リンクすることはできない。LOD化においてなるべくブランクノードを避けるというのが一つのプラクティスとなっている。今回、コンテキストをURLノードにするかブランクノードにするか、データノードをURLノードにするかブランクノードにするか、試行錯誤を繰り返したが、結局、コンテキストはURLノードにし、データノードはブランクノードにすることにした。データノードをブランクノードにした最大の理由は、今回の経験で現実のXBRL報告書においては一つの報告書においてさえ、データ部

<sup>5</sup><http://www.w3.org/TR/vocab-data-cube/>

<sup>6</sup>[http://sdmx.org/?page\\_id=10](http://sdmx.org/?page_id=10)

分に ID が降られていたりなかったりして、データ ID がない場合にはブランクノードとせざるをえず、変換プログラムが煩雑になることを恐れたからである。なお、たとえデータノードがブランクノードであってもそれにリンクされているグローバルに唯一のコンテキストノードさらにはリンクされている文書ノードを合成クエリに含めることで、問題なく所望のデータを検索可能であることを注意しておく。

### 3 旧 EDINET データの LOD 化

#### 3.1 旧 EDINET の RDF グラフ

前章で述べた基本方針にしたがって、旧 EDINET をリンクトデータにした。その RDF グラフ（勘定科目 NetIncome のみ）を図 1 に示す。

ここで、中央上部のノードは EDINET 文書ノードであり、中央のノードがコンテキストノード群である。下部のブランクノードがデータノードに相当するが、ここでは二つだけを記載している。EDINET コード+追番<sup>7</sup>と報告書提出日を指定すれば該当報告書が同定でき、EDINET コード+追番と報告対象期間の開始日あるいは終了日を指定すれば、該当のコンテキスト群が得られることが分かる。コンテキスト名称は以下のような情報の組み合わせであり、読めばその意味は理解される。

名称 := 相対年+連結情報+時間情報  
 相対年 := {Prior2Year/Prior1Year/CurrentYear}  
 連結情報 := {Consolidated/NonConsolidated}  
 時間情報 := {Instant/Duration}

図 2 に経産省の DataMETI SPARQL エンドポイント (<http://datameti.go.jp/sparql/>) に対し、EDINET コード+追番と報告書提出日を指定して、該当報告書のコンテキストを得る SPARQL クエリの例を示す。

また図 1 の RDF グラフから分かるように、勘定科目と一つのコンテキストのノード（これは各 EDINET 報告書ごとに存在する）を指定すればそれに該当するデータが得られる。例として、E00048-000 の ID を持つコンテキストと、勘定科目 t-cte-owl:NetIncome を指定して、該当ノードの値を列挙する SPARQL クエリの例を図 3 に示す。このクエリの結果、異なるコンテキストの 4 個のデータが得られる。

<sup>7</sup>報告書インスタンスが複数の場合、その順序を示す。



図 2: 旧 EDINET の SPARQL クエリ例。

#### 3.2 旧 EDINET データの名前空間

旧 EDINET の XML 表記においては、業種別タクソノミや企業別タクソノミに図 4 に示すような名前空間が用いられている。

これらのアノテーションを RDF に表現したときには、そのプレフィックスと名前空間を図 5 に示すように変えた。なお図 5 に示すもの以外にも、新しい語彙について edinet や xbrl-owl など新しく導入された名前空間もある。

一方、旧 EDINET の XBRL ファイル名は以下のような命名規則になっている。

jpfr-{ 報告種類 }-{ EDINET コード }-{ 追番 }-{ 報告対象期間期末日 }-{ 提出回数 }-{ 提出日 }

XBRL ファイルの名前空間は定義されていないが、XBRL ファイル提出時には、提出者ごとの企業タクソノミファイルも同時に提出される。したがって、XBRL ファイルの名前空間も当然図 4 にある提出者別企業タクソノミファイルの名前空間と同様な名前空間になるべきだと考え、このときは URL パス部の最後に単純に XBRL ファイル名を追加して次のような http アドレスを想定した。

<http://info.edinet-fsa.go.jp/jp/fr/gaap/> { EDINET コード } - { 追番 } / { 報告種類 } / { 報告対象期間期末日 } / { 提出回数 } / { 提出日 } / { ファイル名 } -xbrl

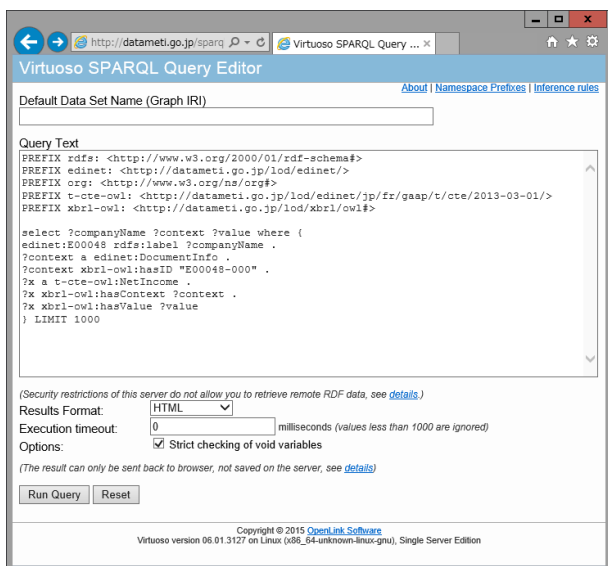


図 3: 旧 EDINET の SPARQL クエリ例.

この「<http://info.edinet-fsa.go.jp/jp/fr/gaap/>」を「<http://datameti.go.jp/lod/edinet/jp/fr/gaap/>」に置き換えたものが図 1 に示した RDF ファイルの URL である。しかし、これは RDF ストアを前提とすれば URL パス部とファイル名部に重複する情報を持ち冗長であった。そのようにした理由は、旧タクソミを一括取得するとわかることであるが、ファイル名はディレクトリ構成を反映した名前になっており、冗長ではあるがタクソミファイル名のみから、タクソミのメタデータ（層カテゴリや業種カテゴリなど）やディレクトリ構成が分かるようになっている。それと同様に、XBRL ファイルや RDF ファイルについても一括取得の場合を想定してファイル名にディレクトリ情報を持たせることに何か意味があるように思えたからである。

## 4 新 EDINET データの LOD 化

### 4.1 新 EDINET の RDF グラフ

EDINET インライン XBRL では報告書提出者が作成したインライン XBRL ファイル (iXBRL) を EDINET に登録した時点で、自動的に XBRL インスタンスファイルが生成される。旧 EDINET の XBRL ファイルと比較されるべきは、こちらの自動生成された XBRL インスタンスファイルである。

先に述べたように EDINET の仕様は新旧で大きく異なっているが、リンクトデータ化の方針はそのまま新 EDINET にも適応された。その結果の RDF グラフを図 6 に示す。

薄墨部分が新旧で異なる部分である。グラフの構造としては変化はないが、新 EDINET では URL パス部における情報の冗長性を除いた。その理由は、旧 EDINET のディレクトリ構成が業種別になっていて、オントロジー的に何か意味があるように見えたが、新 EDINET では報告書の書式を反映していて、オントロジー的にはディレクトリ構成を保存することに意味がないと思われたからである。新タクソミの名前空間に合わせて RDF データの名前空間も変えたことも異なる。旧 edinet:DocumentInfo はより適切な名前の edinet:Context に変更した。また、日本円の定義が報告書ごとに変わることはないため、簡単に edinet:JPY とした。

### 4.2 新 EDINET データの名前空間

勘定科目名は名前空間は別として新旧で変わることはないが、コンテキスト名は完全に新旧で不一致である。ただし、対応の推察は可能である。図 7 に主要な新 EDINET における名前空間を、図 8 にそのアノテーションを RDF にした場合の名前空間を示す。

なお新 EDINET におけるファイルの命名規則は以下のようになっている。<sup>8</sup>

```
jp{ 府令略号 }{ 様式番号 }-{ 報告種類 }-{ 報告書連番 }-{ EDINET コード }-{ 追番 }-{ 報告対象期間末日 }-{ 提出回数 }-{ 提出日 }
```

## 5 TDnet データの LOD 化

旧 TDnet ファイル内容と、旧 EDINET ファイル内容は、東証コードと EDINET コードの違いや企業別タクソミの名前空間の違いを除けば、酷似している。従って、それらの違いに注意すれば、旧 EDINET と同様な方法でリンクトデータにすることができる。今回一部のデータのみについてリンクトデータ化を行い、EDINET と TDnet とで同一の勘定科目データであることを確かめた。

また、新 EDINET ファイルがインライン XBRL に変わったのと軌を一にして TDnet もインライン XBRL に変わったようであるが、新 EDINET では無料で提供される XBRL インスタンスファイルに相当するファイルが、東証からのダウンロードでは得られない。新 TDnet の iXBRL ファイルから XBRL ファイルを生成することが必要となるが、今回それは省略した。

TDnet の HP を見ると、オープンデータとするにはためらわれるような記載がある。EDINET はオープン

<sup>8</sup>報告書連番および追番の詳細は、<https://disclosure.edinet-fsa.go.jp/download/ESE140111.zip> 提出者別タクソミ作成ガイドライン 添付 4 各種命名規約の略号、連番及び追番一覧を参照されたい。

データと言えるが、LOD の利点を最大限活用するために、TDnet のオープン化が期待される。

## 6 企業データの Linked Data 化

旧 EDINET データを取得した 50 社について、2014 年 2 月時点での登記情報提供サービスの「商業・法人登記情報」（以下、企業登記情報と呼ぶ）を取得し、この法人番号と EDINET コードおよび TDnet の東証コードについて、下記のような RDF データを作成した。

```

@prefix tdnet: <http://datameti.go.jp/iod/tdnet/> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix edinet: <http://datameti.go.jp/iod/edinet/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix cri: <http://datameti.go.jp/iod/cri/> .
@prefix corporate: <http://datameti.go.jp/iod/corporate/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix sicc: <http://datameti.go.jp/scheme/sicc/2003/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix crv: <http://datameti.go.jp/scheme/crc/2013/> .
@prefix vCard: <http://www.w3.org/2006/vcard/ns#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix roc: <http://datameti.go.jp/scheme/roc/2013/> .
@prefix time: <http://www.w3.org/2006/time#> .
...
corporate:C120001077450 a org:Organization ;
  rdfs:label "住友電気工業"@ja ;
  skos:prefLabel "住友電気工業"@ja ;
  org:identifier "120001077450" ;
  owl:sameAs edinet:E01333 , tdnet:T58020 ,
               cri:R120001077450 .
...
edinet:E01333 a org:Organization ;
  rdfs:label "住友電気工業"@ja ;
  skos:prefLabel "住友電気工業"@ja ;
  org:identifier "01333" .
...
tdnet:T58020 a org:Organization ;
  rdfs:label "住友電気工業"@ja ;
  skos:prefLabel "住友電気工業"@ja ;
  org:identifier "58020" ;
  org:classification sicc:C3500 .
...
cri:R120001077450 a org:Organization ;
  rdfs:label "住友電気工業株式会社"@ja ;
  dct:issued "1920-12-10"^^xsd:date ;
  org:purpose "1. 次の各種製品、複合製品の開発 ... "@ja ;
  skos:prefLabel "住友電気工業株式会社"@ja ;

  crv:publicNotice [
    a crv:RegistrationInformation ;
    dct:description "当会社の公告方法は、電子公告とする。
... "@ja
  ] ;
  org:hasPrimarySite [
    a org:Site ;
    rdfs:label "本店"@ja ;
    org:siteAddress [
      a vCard:Address ;
      vCard:locality "大阪府"@ja ;
      vCard:street-address "大阪府中央区北浜四丁目 5 番 3 3
号"@ja
    ]
  ] .
...

```

SPQRQL エンドポイントからデータを検索するのではなく、RDF ノードの URL が分かっている URL を直接 http アドレスとしてアクセスして情報を得ることもできる。住友電気工業株式会社に関する LOD データを表示した例を図 9 に示す。



図 9: 参照解決可能な RDF データの例。

なお、sicc は証券コード協議会による業種別分類項目を LOD 化したものである。

そのほかにも、工業統計商品分類表、事業所分類、エネルギー統計分類、経済センサスデータ、工業地区コード、日本標準産業分類、工業統計データ、標準地域コード都道府県、標準地域コード市区町村、大都市圏構成市町村名一覧、会社役員スキームなどを LOD 化したのが、その詳細は別報告に譲る。

## 7 関連研究と考察

先行研究に鈴木らの研究 [1, 2] がある。旧 EDINET データについて RDF 化しているが、その動機や向かうところは我々とは微妙に異なっている。我々は XBRL をリンクトデータ技術の応用の場としてとらえ、EDINET、TDnet、企業情報などの LOD 化を行って、それらの



データを公開したが、鈴木らは会計ドメインにおけるオントロジー構築を意識している。また LOD の応用として企業情報を検索閲覧するためのアプリケーションを開発している。我々の研究を水平指向とすれば、鈴木らの研究は垂直指向と言ってよい。また、鈴木らのモデルではブランクノードがなく、データノードは（固有の名前空間において）コンテキストと全く同じ名前を持っている。また、プロパティ語彙も我々のものとは異なっている。

XBRL の LOD 化を推進するためには、最終的に世界規模でのモデルと語彙の統一が必要であることは言うまでもないが、そのような研究が他国においてそれほど多く行われているわけではない。Dave Raggett によって一時期 W3C Interest Group on financial data が計画されたようであるが、<sup>9</sup>その後 W3C での進展は見られない。

一方、XBRL International では OMG MOF に似た抽象オブジェクトモデル<sup>10</sup>を作ることにより、その意味論を明確にしようとしている。会計ドメインオントロジーを考慮するにあたっては、XBRL International による XBRL 抽象モデルを踏まえて LOD スキーマを考える必要がある。

過去の研究には、XML データから OWL データを自動生成するもの [3] や、XML Schema から自動的に RDF モデルを生成する研究 [4] や、ReDeFer の XSD2OWL と XML2RDF<sup>11</sup> を用いて XBRL の構造から自動的に RDF モデルを作る試み [5] もあるが、本来 XML Schema には意味論がなく、XML から自動的に RDF モデルを抽出する一般的方法を見つけることは至難の業である。ドメインに特化したマッピング情報を追加することで XBRL から RDF モデルを生成することは可能とは考えられるが、XML2RDF の一般的なツール開発を目的とするのではなく、XBRL2LOD が目的であれば、ReDeFer を用いるまでもなく、そのような変換プログラムを開発すればよい。実際我々も RDF モデルを確定したのちは、旧 EDINET および旧 TDnet では XSLT により、新 EDINET ではプログラムによって自動変換を行った。

本研究では、東証 33 業種のうち、建設業、パルプ・紙、石油・石炭製品、ゴム製品、鉄鋼、非鉄金属、情報・通信業、卸売業、小売業、銀行業の 10 業種を選び、それぞれの売上トップ 5 社を選んで、LOD 化を行った。一方、工業統計は標準産業分類に基づいて行われており、東証業種分類と標準産業分類との関係についても調査したが、それらの調査結果を今回の研究に生かすこと

はできなかった。しかし、売上トップ 5 社のみならず、ある産業分類における主要な企業の EDINET データがすべて LOD 化されれば、工業統計データと EDINET データを結び付けることも可能になると思われる。

LOD を利用したのではないが、本多ら [6] は 2007 年 12 月時点での EDINET 有価証券報告書を取得し、各会社役員の日経 225 企業内における兼任状況と、過去に日経 225 企業で働いていた経験の有無を調べ、企業間の人材交流のネットワークを分析した。同様に、もし新 EDINET の「大株主の状況」と「大量保有報告書」の情報を LOD 化すれば、株式持ち合いから見える企業間のネットワークも EDINET データから見えてくるであろう。一部上場企業 1,869 社、さらには二部上場その他含めて 3,472 社のほぼすべての EDINET データは入手可能と思われる。それらのオープンデータをリンクトデータとすることで、日本経済分析や政策立案に資する環境がさらに整備されるであろう。

## 8 おわりに

RDF ストアと SPARQL の利用を前提に、EDINET と TDnet の LOD 化を実施した。また、法務局への法人登記情報を LOD 化し、相異なる企業コード、EDINET コード、TDnet コード、登記法人コードをリンクした。すでに EDINET は政府標準利用規約（第 1.0 版）に準拠しているの、ほぼオープンデータと言ってよいし、政府の所管するデータは次々とオープン化されている。企業マイナンバーと言われる法人コードが実施されれば、さらに LOD に有利な社会環境が準備される。これらのオープンデータをリンクトデータとすることで、オープンデータの有効利用が促進される。

本研究の次のステップとして、1) 主要な日本企業の EDINET データの継続的な LOD 化の実施、2) XBRL International が進める XBRL 抽象モデルに整合する、OWL スキーマの構築と標準化、が考えられる。いずれも一朝一夕に済む話ではない。多くの専門家の参加と多くの組織の協力を得て、着実に日本の企業データの LOD 化を進めていきたい。

## 謝辞

本研究の一部は、日立コンサルティング主導の下、NPO 法人リンクト・オープン・データ・イニシアティブの名前で、平成 25 年度経済産業省「オープンデータ推進に関する調査研究」の一部として実施されたものである。経済産業省と日立コンサルティング、および協力いただいた日立中央研究所に感謝する。

<sup>9</sup><http://www.w3.org/2009/Talks/0623-dsr-xbrl-semweb.pdf>

<sup>10</sup><http://www.xbrl.org/Specification/abstractmodel-primary/PWD-2012-06-06/abstractmodel-primary-pwd-2012-06-06.html>

<sup>11</sup><http://rhizomik.net/html/redefer/>

## 参考文献

- [1] 鈴木 健太, 山口 高平: 会計ドメインにおける RDF モデルの構築と Linked Data との連携, 第 25 回人工知能学会全国大会, 3E3-OS20-5 (2011)
- [2] 鈴木 健太, 玉川 奨, 山口 高平: 大規模会計 Linked Data のためのシステムアーキテクチャ, 第 25 回人工知能学会全国大会, 3C2-OS-13b-7 (2012)
- [3] Yahia, N., Mokhtar, A. S., Ahmed, A.: Automatic Generation of OWL Ontology from XML Data Source, *Int. J. Computer Science Issues*, Vol 9, Issue 2, (2012)
- [4] Bedini, I, Matheus, C., Patel-Schneider, F, P., Boran, A., Nguyen, B.: Transforming XML Schema to OWL Using Patterns, *Semantic Computing (ICSC)*, Fifth IEEE International Conference, pp.102–109 (2011)
- [5] García, R., Rosa, G.: Triplifying and Linking XBRL Financial Data, *Proc. 6th Int. Conf. Semantic Systems (I-SEMANTICS '10)*, pp.3:1–3:8, (2010)
- [6] 本多 元, 大向 一輝, 武田 英明: 役員の職歴を用いた企業間ネットワークの定量分析, 情報処理学会第 70 回全国大会, pp.“5 – 243”–“5-244” (2008)

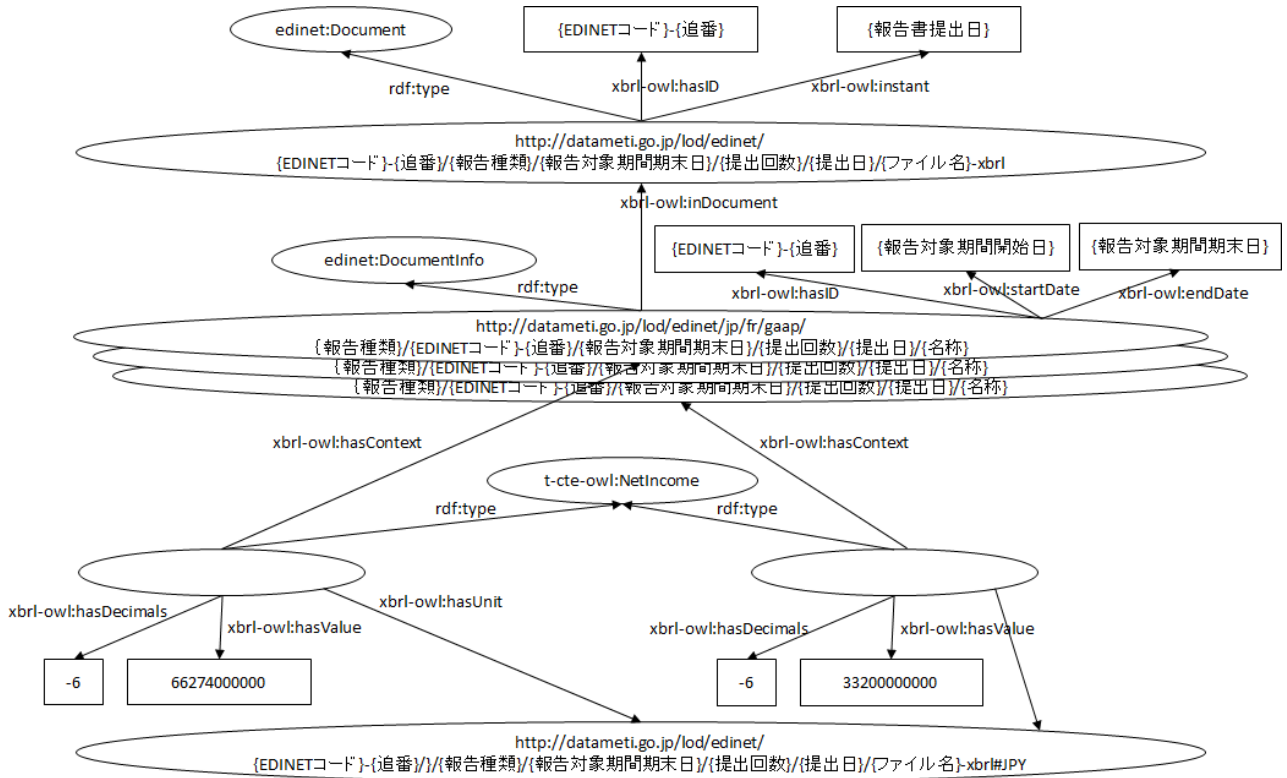


図 1: 旧 EDINET の RDF グラフ例.

タクソノミ種類	プレフィックス	名前空間
業種別タクソノミ	jpfr-〔層〕-〔業種略号〕	http://info.edinet-fsa.go.jp/jp/fr/gaap/〔層〕/〔業種略号〕/〔発行年〕
提供者別タクソノミ	jpfr-〔報告種類〕-〔EDINETコード〕-〔追番〕	http://info.edinet-fsa.go.jp/jp/fr/gaap/〔EDINETコード〕-〔追番〕/〔報告種類〕/〔報告対象期間期末日〕/〔提出回数〕/〔提出日〕
文書情報タクソノミ	jpfr-di	http://info.edinet-fsa.go.jp/jp/fr/gaap/o/di/〔発行年〕

図 4: 旧 EDINET の名前空間.

タクソノミ種類	プレフィックス	名前空間
業種別タクソノミ	〔層〕-〔業種略号〕-owl	http://datameti.go.jp/od/edinet/jp/fr/gaap/〔層〕/〔業種略号〕/〔発行年〕/
提供者別タクソノミ	〔報告種類〕-〔EDINETコード〕-〔追番〕-owl	http://datameti.go.jp/od/edinet/jp/fr/gaap/〔EDINETコード〕-〔追番〕/〔報告種類〕/〔報告対象期間期末日〕/〔提出回数〕/〔提出日〕/
文書情報タクソノミ	di-owl	http://datameti.go.jp/od/edinet/jp/fr/gaap/o/di/〔発行年〕/

図 5: 旧 EDINET の RDF 名前空間.



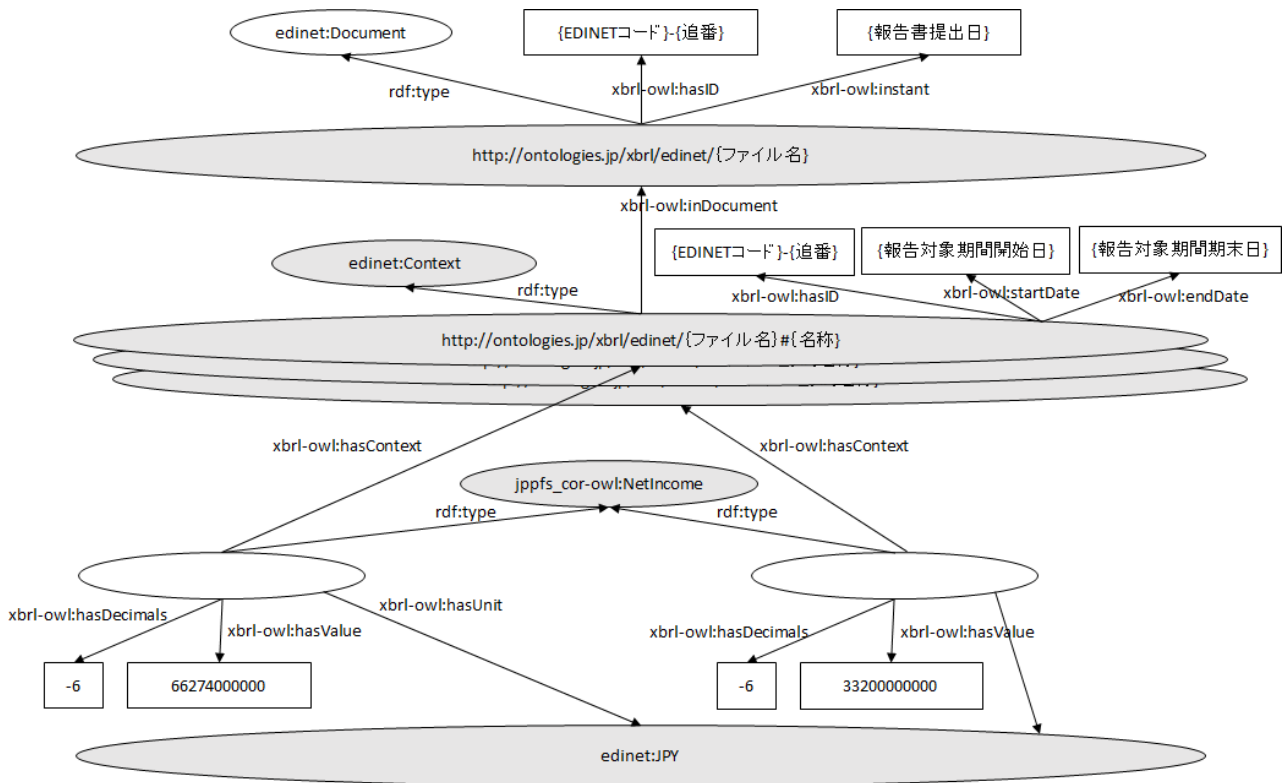


図 6: 新 EDINET の RDF グラフ例.

タクソノミ種類	プレフィックス	名前空間
府令別語彙	jp(府令略号){-(報告種類)}.cor	http://disclosure.edinet-fsa.go.jp/taxonomy/jp(府令略号){-(報告種類)}/{(発行年)}/jp(府令略号){-(報告種類)}.cor
財務諸表本表語彙	jppfs.cor	http://disclosure.edinet-fsa.go.jp/taxonomy/jppfs/{(発行年)}/jppfs.cor
DEI語彙	jpedi.clr	http://disclosure.edinet-fsa.go.jp/taxonomy/jpedi/{(発行年)}/jpedi.cor

図 7: 新 EDINET の名前空間.

タクソノミ種類	プレフィックス	名前空間
府令別語彙	jp(府令略号){-(報告種類)}.cor-owl	http://ontologies.jp/xbrl/edinet/jp(府令略号){-(報告種類)}/{(発行年)}/jp(府令略号){-(報告種類)}.cor
財務諸表本表語彙	jppfs.cor-owl	http://ontologies.jp/xbrl/edinet/{(発行年)}/jppfs.cor
DEI語彙	jpedi.clr-owl	http://ontologies.jp/xbrl/edinet/jpedi/{(発行年)}/jpedi.cor

図 8: 新 EDINET の RDF 名前空間.