

## Linked Dataの企業での活用について

西野 文人 桑 照宣<sup>†1</sup>

知識集約型組織では、経営戦略上の各種の意思決定に有用な知識や洞察を生み出すために、企業内外の事実に基づく情報を組織的にかつ系統的に蓄積・分析し、それを活用することが求められている。我々は一つの例として研究所がどんな研究をしているのかを見える化し、営業・SEが研究所の技術を探せる/発見があるようにするために、社内外の文書、展示会、人物、組織、技術、研究、ソリューション等に関するデータを Linked Data 化するとともに、この Linked Data を活用するツールを作成した。ツール作成にあたっては、探索の効率化のためのネットワーク表示機能、知識を増殖したり維持管理するための集合知活用などを取り入れた。

### Utilization of Linked Data in Enterprises

NISHINO, FUMIHITO and KUME, TERUNOBU

To produce useful knowledge and insight for making various decisions on corporate strategies, it's important for knowledge-intensive organizations to exploit intelligence generated by systematically accumulating and analyzing information based on facts inside and outside of enterprises. We have produced Linked Data from data sources including documents, exhibitions, persons, organizations, technologies, researches, solutions, etc., inside and outside our company, and developed a browser for navigating Linked Data. When we were developing the browser, important considerations were to display network diagrams for exploring linked data efficiently and to utilize collective intelligence for accumulating and managing knowledge.

<sup>†1</sup> 富士通研究所  
Fujitsu Laboratories

## 1. はじめに

Linked Data<sup>1)</sup> はデータの新しい共有方法として認知され、欧米を中心に実践が進んでいる。特に、広く公開データを作成しようという Linking Open Data(LOD) 活動<sup>2)</sup> は、日本でも公的機関を中心に広まりつつある<sup>3)</sup>。一方、企業内においては、技術文書の語彙を Linked Data 化して利用しようという取り組み<sup>4)</sup> などがあるものの、企業内の様々な知的活動においてデータを Linked Data 化した後、具体的にどんな効果があるのかと問われると、まだ他人を十分に説得できるだけの答えを持っていない。また、現在ある企業内のデータを Linked Data 化しようとする、何か問題になるのかといったことも整理しておく必要がある。我々は、Linked Data 技術を「大量にあるデータの中から、データを戦略的に選択・抽出・結合することで情報を生み出し、これらの情報を戦略的に活用することで知的活動を変革することを可能とする技術」と捉えている。ここで「戦略的」とは、目的に関係しそうなデータや情報の発見、抽出、構造化、分析、推論などをどう行うかを決定するものであり、そのために、情報抽出技術、スクリーンスクレイピング技術、分類技術、情報の価値判定技術などを利用することになる。本稿では、まず企業における知識経営とはどうあるべきかを示し、データの重要性を示す。続いて多くの企業でありがちなデータの取り扱いに関する現状を示すことで、あるべき姿とのギャップから多くの企業が抱えているデータの取り扱いに関連する問題を考察する。その上で、企業内のデータを Linked Data 化することによって何が解決できそうなのかを示し、我々が何をどう考えて自社内データの Linked Data 化と、Linked Data を活用するためのツールを開発しているのかを示す。そして実際に自社内のデータを Linked Data 化して何ができるようになったのかを示す。

## 2. 企業における課題

本節では企業で求められている知識経営について述べ、データの重要性を示す。これに対して多くの企業で起こっているデータの取り扱い関連の問題を示し、どのようにあるべきかを示す。そして、企業内のデータを Linked Data 化することのメリットと、Linked Data 化する上での問題点について考察する。

### 2.1 知識経営

企業活動を進めていく上では、判断・行動をするためのインテリジェンスが必要である。そこで多くの企業では、自社の方針、戦略意思決定、運営に影響する経営環境、市場、顧客、技術、競合などの外部情報を組織的に収集・分析し、意思決定や戦略に活かすことを狙

いとした活動を行っており、そのために内部情報・外部情報の収集・分析を支える ICT 環境の整備と、ICT 環境を活かす組織 IQ<sup>5)</sup> の向上（知識循環型組織の構築）につとめている。特に研究部門では、技術革新や、研究成果の市場に投入するまでのリードタイムの短縮が求められており、そのためには、研究のやり方・質を高めること、研究の現状分析や計画予測をできるようにすること、研究パフォーマンス・マネジメントを可能にすること、情報ニーズに機敏に対応できるようにすること、研究活動を見える化し活用されることを目指すこと、研究部門外の人が必要な技術や人を見つけ出せるようにすること、研究者間あるいは研究者と開発・営業との間のコミュニケーションを強化することなどが求められている。そのためには以下のような対応が必要になる。

- 研究戦略・目標の設定
  - － 背景・内容・妥当性などの説明や戦略上での抜け等の明確化
- 外部動向調査
  - － 特許・論文調査結果（レビュー）の共有
  - － 調査知識（ノウハウ）・ブックマーク等の共有
  - － 市場変化、顧客変化、技術変化情報の獲得
- 知財マネジメント
  - － 特許の分野、課題、手法、発明者、利用情報の整備
  - － 発明部門、研究所推進部門、特許部、製品での発明の紐付け
  - － 知財（技術力など）、人材（人脈など）の評価
  - － 知財権が適切に利用されているかの確認
- 技術コラボレーション
  - － ツール・データ類のネットワーク共有
- 研究成果広報
  - － 研究成果の広報・検索環境
  - － ニーズ調査
- 研究成果分析
  - － 研究価値の見える化（自社売上貢献度、学会活動貢献、特許活動貢献、他社インパクトの分析）
  - － 失敗プロジェクト分析

これらに対応するには、散在する社内外のデータを戦略的に収集・構造化して「情報」とし、経営層、研究員、営業などがそれぞれの立場で情報を自由に結合・分析して「インテリ

ジェンス」を作り出す環境を整えることが重要になる。また、それと同時に企業活動の中で自然にデータを収集できるようにすること、情報を提供すべき人が積極的に情報を提供するような意識を持たせることが重要になる。

## 2.2 企業の ICT 活用の現状

今では情報の共有や Web システム化が進んでいる企業も多いであろう。しかし全社的なものではなく、小さなグループでの情報共有では未だにメールでの情報のやり取り、ファイル置き場にファイルを置くだけの情報共有などが行われている。また、世の中では色々な Web サービスがあるのだが、セキュリティ等の問題から利用が制限されていることも多い。また、制度やシステムが単目的に作られていることも多いのではないだろうか。すなわち、情報は既に提供しているはずなのに、それぞれの目的に応じて申請が必要になっているなど、社内のシステム・制度がサイロになっていることも多いのではないだろうか。社内 Web 横断検索システムのようなものもあるだろうが、ここで検索される情報は構造化されておらず、またメンテナンスされていない Web ページの存在などにより、知りたい情報が見つからなかったり、見つかったとしても情報が古かったり、自分の部署には関係のない情報だったりしないだろうか。より信頼度の高い情報を得ることができるようになるためには、信頼のできる社内の公式情報を収集し、一定の規則に従ってメタデータを作成し、そのメタデータを検索するシステムが求められている。

また社外システムとの連携はどうであろうか。複数のウィンドウを立ち上げて、コピー＆ペーストを繰り返しながら作業を繰り返していたりはしないだろうか。もちろんマッシュアップ技術を使えば作業効率化できる。我々も個人レベルで簡単に Web サービスの画面を別の Web ページにマッシュアップさせるツールを開発している<sup>7)</sup>。しかし、このようなことを自分でする人は非常に少ない。情報系の人間からすると驚くほど世の中の人は保守的であり、例えば自分の姓名すらかな漢字変換の単語辞書に登録せずに毎回苦労して自分の姓名を入力している人さえ珍しくない。基本的に ICT の作業環境を良くしようとする意欲に欠けている（あるいは、どうしたらよいかわからない）人も少なくない。

情報はあるだけでは価値がない。しかしながら、当初の目的が何であったのかが不明になり、今では目的を見失ったにもかかわらず、習慣で情報を収集していることも多い。そして、利用目的のないままに見える化して満足していることも多い。情報はつなげて分析をし、行動に結びつけてはじめて意味のあるものとなり、そのように活用されてこそ原データを収集し共有することに意味がある。

社内でのオンライン化、情報共有は進んだが、課題も生まれてきた。昔は情報処理関係者

だけが計算機を使っていたので、自分らの開発したシステムを使い、フォームに記入するものも構造化文書で作成して適切なフォーマットに変換して印刷するというようなことが行われていた。しかしながらスプレッドシートの利用が進み、誰でも簡単にフォームが設計できるようになると、単に印刷形式に合わせた書式が増え、情報の意味の明確化という意味では退化した。その後、Web サービス化なども進んだが、早い段階で Web 化されたシステムはそのサービス単目的でのユーザの利便性のみを追求し、データとしての活用というようなことを考えていないために様々な問題を抱えている。また情報の Web 公開も進んでいるが、事務管理部門が作成する Web ページの HTML は酷いものがある。Web ページ作成ガイドラインのようなものもあるが、提示すべき情報や見栄えについては指定も多いが、データが多観点での活用ということまでしっかり考えてガイドラインが作成されている企業はまだ少ないのではないだろうか。

### 2.3 社内データのあるべき姿

例えば、以下のような問いに簡単に答えられるようになっていないだろうか？

- ある分野のセミナーを開催したいが、その分野に詳しい外部の先生と、その先生の社内の知り合いを探したい。
- 画面の操作性に関して参考となる情報や開発実績を知りたい。
- 顧客訪問に際して、関連情報を知りたい（研究の有無、関連資料、他社動向、当社の強み、業界全体の動きなど）
- 社内でのどの学会にどのくらいの人数が入会しているかを知りたい
- 5年間で特許出願が3件以下の研究員は？
- 最近の情報処理学会の特集にあった技術で、自社で関連する研究がない技術は？

このような要望に応えるためには、社内にある様々な情報が、個々のファイルや Web システムの中に存在するというだけでなく、情報を後で柔軟に利用できるように情報の意味が明確化されて自由に扱え、そしてそれぞれのデータが有機的に繋がっていることが必要になる。例えば、「IE」という単語は、自然言語処理研究者では情報抽出 (Information Extraction) の意味で使うが、ブラウザの名称としても使われている。あるいは研究者が勝手に作った用語もある。企業でこれらの用語に関連する技術を活用するには、様々な利用者の存在を考慮して、用語の意味範囲を定義しておくことが重要になる。

意味記述とともに重要となるのが集合知の活用である。アイデア、コツ、裏ワザ、経験則などの組織内の身近なレベルの情報も組織の情報資産である。これらをいかに吸い上げるかが重要である。企業活動ではチームで一つの取り組みを行ったり、チームは別でも互いに

関連する分野に興味を持っていることもある。そのようなチームのメンバーや別のチームのメンバーが関心を持った情報を利用したり、他人が得た知識を活用することが求められている。

また、これらの情報は様々なアプリケーションから使用できるようにすることも考える必要がある。そして企業内の様々なデータを結びつけるとともに、社外データとも結びつけることにより、企業活動、特に研究活動を活性化することが望まれている。社内と社外の知識をやソーシャル情報を収集・結合・分析・活用することで、注目の情報を浮かび上がらせたり、全体の傾向を知ることができるのである。

### 2.4 Linked Data 化への道

Linked Data の重要性は情報技術者にとっては明白であるが、Linked Data 化によるメリットと、Linked Data 化するためのコストを説明できないと、社内で広くシステムを利用するというにならないし、既存のデータの Linked Data 化や、既存の Web システムの作り直し、新しいシステムの作成にまで至らない。既に使い慣れているシステムがあると、システムを作り直すことへの抵抗、システムの使い方が変更されることへの利用者の抵抗は非常に大きい。また、情報の公開範囲についても秘匿すべき範囲や内容など気にすべき点が多い。そして、システムが持つ情報を提供したり、公開範囲を変更することなどに対して、誰が責任を持っているのかも明確でないこともあり、情報が提供されない場合もある。

そこで我々は Linked Data 化に際し、次の方針をとることにした。

- (1) 実データを Linked Data 化し、実利用でメリットを示す。
- (2) 既存のシステムをそのまま利用する。
- (3) 集合知を活用する。

まずは、何か具体的なデータを Linked Data 化して、実際の活用場面においてメリットを示す。大きな目標としては、研究のやり方・質を高める研究パフォーマンス・マネジメントとして、研究経営層が研究所の現状分析・計画予測をできるようにするとか、商談をしやすくするなどがあるが、これらは具体的なメリットを示しにくいので、データや技術の観点から、より具体的にまず実現できそうなものとして、「研究所がどんな研究をしているのか見える化し、営業・SE が研究所の技術を探せる/発見があるようにする」というのを設定した。

二つ目の方針は、既存のシステムをそのまま利用するというものである。既存のシステムのデータベースを直接利用したりデータを入手できることはほとんどないので、ブラウザ経由で利用するような形でデータを取得することが必要になる。またこれは単発の処理ではなく定期的に更新することが必要になる。我々の方針としては、はじめからたくさんの情報を

入れようとせず、情報の公開において難のあるもの（公開範囲に制限がつくものや、明らかでない関係者に何らかの許可を必要としそうなもの）は置いておき、既に社内での Web 上に公開されているシステム上のデータの Linked Data 化を進めた。その上で、以下で対応を行っている。

- (1) 既存の Web サービスからのスクレイピング
- (2) 新たな情報管理システムの提供

電話帳のようなシステムは部署内にとどまるものではなく、全社的なシステムになっているので、これを置き換えることは難しい。またこのシステムは人事情報が自動的に反映されるので、定期的にアクセスし、スクレイピング技術などを利用して情報を抽出し、更新されている情報を見つけてメタデータに反映することにした。一方、月報のようなものは、比較的狭い範囲での運用であったので、Semantic Media Wiki<sup>6)</sup> を利用して、ここに情報を入力させるとともに、メタデータを抽出して反映している。

三つ目の方針は、誰でも自由に情報を追加・編集できるようにすることで、情報を充実させるというものである。初期システムを如何に構築するかとともに、初期に保有していない知識をいかに補充するか、また、活動の中で生じるアイデア、論文や特許公知例等の検索結果やレビュー、市場調査などでの気づきを自然な作業の中から如何に蓄積し、如何に情報を正しい状態に保持した管理をするかが大きな問題である。そこで、システムの検索・閲覧で得られた結果に対して簡単に情報を追加・編集できるようにするとともに、コメントの付与等をサポートすることにした。また前述の月報システムのように、活動の一環のシステムの中にセマンティック機能を組み入れ、簡単な制約を設けた自動抽出によって情報を追加する。さらに、ソーシャルブックマーク等の連携、ログからの知識獲得によって知識の獲得を図っている。

### 3. データ整備

データ整備にあたっては、どこにどんな情報があるかだけでなく、誰がどういうことをしたいかが重要になる。研究所活動の見える化という点では、以下のような検索の場面を想定した。

- お客様に向けて、ある分野の技術に関するセミナーを開催したり、説明資料を作成するうえで、その分野の技術に詳しい人物を知りたい
- お客様から解決したい課題が与えられたが、それを解決できる技術はあるか？  
このようなニーズに応えるには以下のようなことが考えられる。

- 分野として技術分野がひととおり網羅されている。
- その分野に通じるような技術キーワードが関連付けられている。
- 自社の研究の技術的な位置づけ（差異化要素、特許情報、特許のポイントなど）
- 社会で話題になっている技術と自社の関係
- 最終的に担当者に連絡をとるための情報

ということで、人物、文書、組織、イベント、技術、ソリューションなどを Linked Data 化することにした。

データは各クラスに対して、そのクラスが持つプロパティを定義している。プロパティとしては、そのクラスにおけるわかりやすい名称、RDF としてのタグ名、各種の制約や暗黙値のほか、専用ブラウザにおける検索・表示・編集に対する属性（アイコン、サマリー表示の対象、検索の対象など）を定義している。

#### 3.1 人物

人物に関しては、まず誰を整備対象とするのが問題になる。我々の場合、研究所の人間のみを対象の範囲としたが、それでも退社した人、事業部に異動した人、海外の研究員をどうするかが問題になった。また、これらの情報をどうメンテナンスするかも問題である。さらに ID として何を利用するかも問題である。ID として、姓名自身を利用するのか、比較的一意性が保たれる情報を利用するのか、一連の番号を利用するのかも問題になる。

属性としては、姓、名、姓の読み、名の読み、所属、役職、メールアドレス、内線電話番号などがある。しかし、ここでも欧米人や中国人・韓国人は姓名としてどのような表記にするのか、読みとして何を登録するのかが問題になる。また、兼務や社外委員等はどうするのかが問題になる。所属・役職は人事異動があると変わるので、人事異動情報をどこからどのように反映するか、何は上書きしていいが、何は上書きできないかが問題になる。上記の情報が社内のデータベースにもあるような情報であるが、上記以外に登録すると良い情報としては、写真、所属学会、興味を持つ分野、現在従事するプロジェクトや過去に従事したプロジェクト、各種 SNS の ID などがある。また、論文、特許、社内の技術文書なども、それらの著者としてリンクされるべきものであろう。

#### 3.2 応用、ソリューション、技術情報

前述のようなニーズに応えるためには、単に現在誰がどのような研究に従事しているかだけでなく、過去に誰がどのような研究に従事していたかも重要な情報になり、現在は研究をしていないものもリストアップすべきということになる。特に、お客様向けに技術を紹介したいという意味では、社内・社外の展示会の情報も重要な情報である。ただし、展示会等

に記述されているキーワードについては、自動で抽出したとしても、その用語が既に登録されているものならばよいが、そうでない場合は、クラスの推定、適切なプロパティの付与、解説の付与が必要である。特に、研究者が独自に作成した用語では問題があった。

また、技術の商品につなげるためには、製品・サービス一覧情報の Linked Data 化、個々の技術と製品・サービスの紐付けが必要である。しかし、これは誰が紐付けられるのかという問題もあるが、誰かが紐付けてくれたとしたらその情報を通知できるようにすることは有用であろう。

社外情報としては、話題の技術（例えば情報処理学会の特集やマーケティングリサーチ会社の調査・予測など）、即時利用可能なオープンな技術、注目すべき論文などを収集することも有用である。

#### 4. Linked Data の閲覧・検索

我々のシステムは、様々なデータから知識処理技術により情報を抽出して RDF を生成することで Linked Data 化するツールと、Linked Data の格納・管理を行う RDF リポジトリ、Linked Data 化したデータを Web から検索・閲覧・編集を可能とすることで発展的活用を支援するブラウザからなっている（図 1）。以下では Linked Data を活用するためのブラウザについて述べる。

##### 4.1 提供機能

本ブラウザが提供する主な機能を以下の通りである

- **メタデータ検索** 入力されたキーワードを持つメタデータと、それに関連するメタデータを検索する。
- **関係性表示** 特定のメタデータに関するメタデータを、その関係性を含めてビジュアル化して表示する。
- **メタデータ入力** 社内・社外に散在する様々な情報をメタデータ化して収集し、保存する。
- **メタデータ編集** メタデータが持つプロパティの変更や、他のメタデータとの関係性の追加・削除を行う。
- **フィードバック/リコメンデーション** 利用者からのコメントやログを分析し、ランキング表示やおすすめ情報として活用する。

##### 4.2 リソース検索

本ブラウザの検索ではユーザは概念（基本的には単語）を入力し、その概念と関連するリソースを提示するという方針を採用した。文字列検索ではなく概念の検索なので基本的には

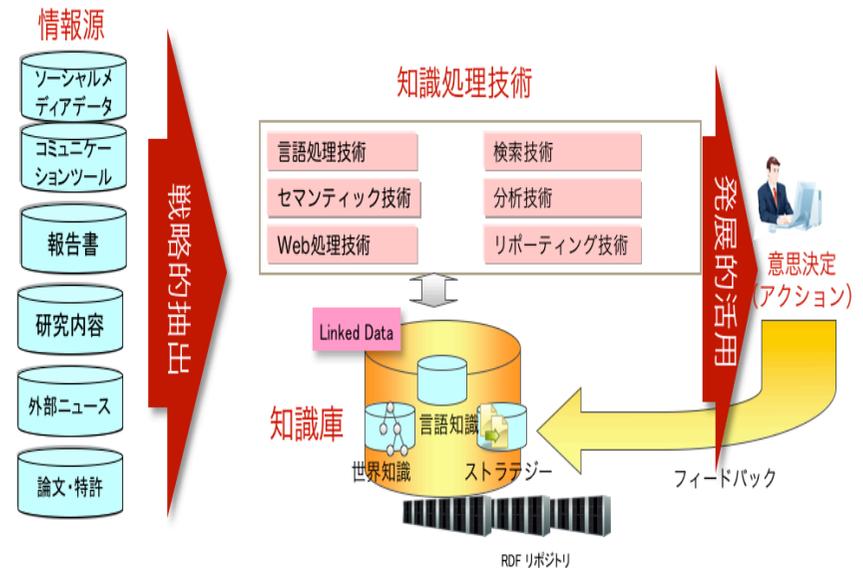


図 1 システムの概要  
Fig.1 System outline

対応するリソースの情報が検索結果として表示されることになる。リソースの情報としては詳細情報（そのリソースに付随する各プロパティの値）を表示するとともに、そのリソースの周辺にどのようなリソースが存在するのかをリソースをノードとするリソース間関係のネットワーク図として表示することにした（図 2）。ネットワーク図のノードとしてはリソースの名称程度しか表示できないのでこれらのリソースのサマリー情報も下部に表示するとともに、情報の探索を支援するという意味でリソース詳細情報、ネットワーク図のリソース名、関連リソースの要約情報から関連するリソースへアクセスできるようにした。

##### 4.3 関連するリソースがない場合

検索しようとした概念に関連するリソースがなかった場合、想定されるのは、1) 入力する文字列を誤った、2) 入力された文字列では検索できなかったが、それと近い文字列の概

念に関するリソースがシステム内に存在する、3) 入力された概念に関連するリソースがシステム内に存在しない、という場合である。1) の場合には文字列を訂正して再度検索してもらえば良い。一方、2) の場合は文字列の微妙な差や単語の表現の違いなどで関連するリソースが見つからなかったということがあるので、「もしかして」という形式で可能性のありそうな候補のリソースを提示している。そして、2) の「もしかして」で見つからなかった場合は、3) ということで、利用者がメタデータ（リソースやプロパティ値）を追加できるようにしている。

#### 4.4 メタデータの追加・編集

社内の Web サービスから情報を収集するだけでなく、利用者にメタデータの追加・編集を開放しており、検索した概念がなかった場合にデータを追加できるようにしてあるのが本システムの特徴である。新規入力ではタイトルとクラスを入力することで URI を割り当てる。URI はタイトルとクラスから適切な変換によって割り当てられる。またクラスに対して必須プロパティの値は定義ファイルで割りつけられた暗黙値が割り当てられる。

メタデータを編集しようとする、そのメタデータのクラスに応じたプロパティ表を持つ編集画面が現れ、その画面上で簡単に編集できるようになっている (図 3)。

プロパティ編集でその値としてリソースを入力するには、どのようなものがリソースとして存在するのかを適切に提示してやるが必要になる。リソース指定はリソース検索と同様に単に完全一致による概念の提示だけではなく、類似文字列によるリソース候補の提示やクラスによる絞り込みを行って検索を行い、検索結果の中から意図するリソースを選択させる。

#### 4.5 コメント・ピックアップリスト

展示会・論文だけにとどまらずすべてのリソースに対して利用者はコメントを記入したり、他人のコメントを参照することができるようにしている。さらに、リソースの集合を作成して新たにリソース化して共有するピックアップリストという機能も導入している。

## 5. 評価

社内データの Linked Data 化のメリットとしては以下のようなものがあると考えている。

- (1) 探索の効率化
- (2) 内容の確認
- (3) 統制語彙
- (4) フィードバック

### (5) 外部との連携

#### 5.1 探索の効率化

Linked Data 化によって、技術や人を探す上で効果があったのであろうか。そこで、以下のような簡単な実験を行った。実験は本 Linked Data ブラウザを利用して情報を探した場合と、本ブラウザを利用せずに情報を探した場合とで、情報を探すのに必要とした時間と探された結果の有用度を比較したものである。情報ニーズとしては、実際に営業・SE から研究所の窓口にお問い合わせがあった案件を使った。これはもちろん何を Linked Data 化したかということ、一般検索の UI などに大きく依存するのだが、おおよその目安にはなるかと考えた。その結果、探索にかかる時間としては、一般検索に対して Linked Data での探索は 40%~60%の時間であった。また、検索結果の有用度はほぼ同等であった。これは Linked Data としての整備が進めばさらに効率化すると考えられる。Linked Data での探索時間の短さは、利用者に迷いが無いことによる。リソースが見つからなければ、利用者が他にやれることが少ないので、あきらめが早くなるということである。一般の検索では、どうしても非公式の文書や古い資料が見つかってしまい、検索結果の中から良い情報を探し出すのに時間がかかる。また、検索キーワードを変えたり、できそうなことが多いため、色々試してしまうので探索時間がかかることになる。また、回答有用度は、Linked Data ではデータとして取り込んでいないものは見つからないということはあるが、企業内だと公開すべき情報はそれなりに公式のものに記述されるので、検索結果の有用度はそれほど落ちない。むしろ担当者の内線番号などは、一般検索では見つかった資料の値が古いものであるなど間違っているときにも気づかないという問題があった。ただし、Linked Data 化したデータでも、内線番号や所属などの定期的に見直しをかけるデータは良いが、Linked Data 化するときのデータしかない場合は怪しい情報もあるので、最終更新日などがわかるようにしておくことが必要であろう。いずれにしろ、Linked Data 化することによって、社内情報の探索では大いに役立つことが示された。

#### 5.2 内容の確認

研究開発では、世間で一般に知れ渡っている技術用語だけでなく、独自の用語も作り出すことが多い。このような用語に関しては説明を与えるとともに、世間に知れ渡っている用語との関係などを記述することで、用語の意味も明確になるし、また一般の用語からの探索も可能になる。また、研究では、そこでの成果の技術的ポイントなどがわかるようになることで、研究の価値がわかるようになる。この効果についてはまだ評価はできていないが、開発部門等を含め評価されている。

### 5.3 統制語彙

Linked Data 化の一つのメリットとして、これを統制語彙として利用することである。一つの使い方は、調査におけるセマンティックタグとしての利用があり、調査業務の効率化に結びつくものと考えている。別の使い方としては、文書作成において利用するもので、マニュアルをはじめとして研究計画や成果報告などでも勝手に用語を使用しないようにするために使うものである。文書の中で Linked Data で定義されている単語を抽出する、あるいは適切な言葉に置き換えることで、用語の統一を図るあるいは未定義な用語に対してその意味等の登録を促すことができる。このような使い方に対してはユーザインタフェースが重要になり、今後検討・評価を進めていく予定である。

### 5.4 フィードバック

社内で公開してコメントを付与できるようにしてはいるが、まだ研究所外と研究所内とを結ぶコミュニケーションの確立には至っていない。SNS と連携するなどの工夫を考えるとともに、社内での展示会での利用などを考えていきたい。

### 5.5 外部との連携

外部のデータを取り込んだり、リンクで結合するなどを行っているが、まだ有用な情報として活用する段階にまでは至っていない。外部アプリからのアクセスという点でもまだ活用されていない。今後の課題としたい。

## 6. 現状と今後

これまで第 1 ステップとして、もともと社内外で流通している電子データをメタデータとして取得して Linked Data 化を図ってきた。現在、第 2 ステップとしては、月報などの通常の活動のための環境を提供し、ここに情報を蓄積することで、そこからメタデータを自動的に抽出し Linked Data 化することを試みている。ここでは MediaWiki を機能拡張した Semantic MediaWiki を利用するとともに、さらにキーワードの自動抽出のための機能拡張を導入することで、自動的に情報を抽出するとともに、Linked Data のブラウザとも連携することで、月報の作成者や閲覧者に Linked Data の存在を認識させることで、自動的だけでは抽出しにくいメタデータの記述の促進を図っている。さらにこれまでは電子的に蓄積されていなかった情報、あるいはされていたとしても流通していなかった情報を取り込むことで社内データの Linked Data の拡充を図りたいと考えてる。またこれまで、本システムは、研究所の技術や人を探すという使い方をしてきたが、活動の見える化、報告書等にある用語の確認、統制語彙としての活用なども進めていきたいと考えている。これまでの活

動は、どんなデータを Linked Data 化できるか、そのデータを使って何ができるかということが中心であった。しかし、実際には、xx の分析をしたい、xx を予測したいというようなニーズがあり、そのためには LOD としてどんな情報が公開されているのか、例えば、ある人に連絡をとりたいというようなときに、欲しい情報は何か、メールや電話番号が記述されているデータはあるのか、ないならばどうするか、その人の連絡先を知っている人をどうやって探すか、などが必要であり、それらがどこにどのような情報として存在するかなどを整理することも重要と考えている。

そしてもう一つ、知識の社内共有と高度利用を実践する知識循環型組織を目指すことが必要である。ICT システムの効果をあげるためには、ツールの開発だけでなく、利用組織の組織 IQ をあげることに、特に意識をもって情報の追加・編集、コメントの付与などに関与する意識をもってもらうようにすることも重要と考えている。

## 参 考 文 献

- 1) linkeddata.org: Linked Data - Connect Distributed Data across the Web, linked-data.org (online), available from <http://linkeddata.org/> (accessed 2011-09-06).
- 2) W3C: LinkingOpenData, W3C (online), available from <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenDataer.url/> (accessed 2011-09-06).
- 3) 武田英明：日本における Linked Data の現状と普及に向けた課題, 情報処理, Vol.52, No.3, pp.326-333 (2011).
- 4) François-Paul Servant: Linked Enterprise Data, Linked Data on the Web (2008).
- 5) 鈴木勘一郎：組織 IQ—小さなチームから大企業まで 本当の実力度, 角川 one テーマ 21, 角川書店 (2001).
- 6) semanticweb.org: Semantic MediaWiki, semanticweb.org (online), available from [http://semanticweb.org/wiki/Semantic\\_MediaWiki](http://semanticweb.org/wiki/Semantic_MediaWiki) (accessed 2011-09-06).
- 7) 糸照宣, 西野文人: ソーシャルな情報再構成による情報共有システムの提案, 情報処理学会研究報告会, Vol. 2010-GN-74, No. 1, pp. 1-6, 2010



図 2 メタデータ詳細表示画面のイメージ  
Fig.2 Linked Data Browser



図 3 プロパティ編集画面のイメージ  
Fig.3 Linked Data Editor