

## メタ情報「タグ」を用いた企業内情報共有

### Knowledge Sharing using Tag in the Enterprise

佐藤 一 広

**要 約** 企業内の情報共有, 特に組織をまたがるプロジェクトなど, 利用者の視点や利用シーンが多岐に渡る際の情報共有には, 柔軟な情報の分類が行えるメタ情報を用いた情報管理が適している. メタ情報を用いることで, 情報の適切な分類と効率のよい発見ができ, さらにメタ情報のひとつであるタグ(利用者が任意につける語句)を用いると, 旧来の情報提供側によるお仕着せの分類とは違う, 利用者視点での新しい分類が可能になる.

筆者は, タグを用いた情報共有ツール「tacoPot™」を作成し, 企業内情報共有の実証実験を行っている. この実証実験での経験と結果を基に, メタ情報による情報管理手法, そしてタグを用いた新たな情報分類・発見手法を報告する.

**Abstract** Management that uses meta-information and enables a flexible classification of information is suitable for the knowledge sharing where user's viewpoints and use scenes cover a broad range of topics. Use of meta-information enables an appropriate classification of information and an efficient discovery, and furthermore, use of tags (words and phrases that a user arbitrarily applies), which is one of meta-information, allows a new classification by the user viewpoints different from off-the-shelf classification provided by a conventional information supplier.

The author developed the knowledge sharing tool "tacoPot" using tags, and conducted the proving test of the knowledge sharing of the enterprise information. This paper reports on the new information classification and the discovery techniques using tags, and the information management techniques using meta-information, based on the results and the experiences in this proving test.

#### 1. はじめに

企業における従業員の活動において, 既存の情報を利用せず担当者一人だけで作業を遂行することは稀である. 特に知的生産活動, すなわち知識や知恵を結集し新たな物を創り出すという活動においては, 複数の従業員と情報交換を行うとともに, 作業に関連する企業内外の文書やweb上にある多数の情報を参照し利用する必要がある.

これらの情報はイントラネット・インターネットの普及につれ日々増大しており, 情報の蓄積方法や蓄積された情報の発見方法は煩雑・複雑化する一方である. このような状況において, 情報共有を円滑に行うには, 表1に示すような様々な課題がある. 特に, 活動に有益な情報を最小限の手間で的確に取得するために, 情報の分類方法と発見方法を検討し, 利用者が効率よく情報を利活用できる仕組みを提供しなければならない. また, 企業においては, 組織内でのみ有効な情報もしくは組織間で共有されるべき情報といった区別, 業務のワークフローに則った管理, あるいは職制などを踏まえた参照権限管理が必須であり, 情報の管理・運用をいかに簡便かつ柔軟に制御できるか, という点も重要な課題である.

表1 情報共有における主な課題

情報共有のステップ	主な課題
収集	最小限の手間でいかに多くの情報を集めるか。
蓄積	情報格納のための容量や、提供速度などの負荷をいかに軽減するか。
分類	情報を発見・共有・活用する上で最適な分類を行うには。
発見	必要な情報を高い精度で、かつ高速に発見できるか。
共有	発見した情報を複数人でわかりやすく、手間なく利用できるか。
活用	情報をいかに再利用できるか。

本稿では、企業内における情報の分類と発見に関する課題、特に利用者が効率よく情報を活用できる仕組みの提供という課題に対して、メタ情報<sup>\*1</sup>を用いた対策とメタ情報のひとつである「タグ（利用者が任意につける語句）」を用いた対策について実証実験を基に報告する。続く2章でメタ情報の概要を説明して共有フォルダ等の階層型管理とメタ情報による管理を比較し、3章でメタ情報のひとつであるタグを用いた情報の管理手法について企業内情報共有ツールによる試行を基に報告する。この試行から得られたタグの課題とその対策について4章で述べ、5章ではメタ情報とタグを用いた企業内情報共有についてまとめる。

## 2. メタ情報による情報の管理

### 2.1 メタ情報とは

メタ情報とは、対象となる情報に関係する情報、たとえば文書における「著者名」や「作成日時」などの情報である。webコンテンツの所在を示す際にはURLが用いられるが、このような「情報の所在」もメタ情報のひとつとなる。その他「コメント」や「感想」など主観的な情報もメタ情報となる。

企業内でも利用が増えている画像や動画、音声といったマルチメディア情報の分類・発見には、メタ情報が重要な役割を持つ。昨今では画像の情報発見にも全文検索を利用するが、これらは画像そのものを検索しているのではなく、画像に付加されたメタ情報の文字列を基に検索している。メタ情報を用いない類似データ検索などの手法も研究されているが、現時点ではマルチメディア情報をメタ情報なしに発見することは困難である。

### 2.2 メタ情報による情報の管理

情報の管理において、共有フォルダ等の階層型による旧来の情報分類方法に対し、メタ情報による分類方法がある。ここでは両方の分類方法を比較し、その管理方法の違いを述べる。

#### 2.2.1 階層型の情報分類とその課題

企業内の情報、特に文書の蓄積・分類には階層型の共有フォルダなどを用いるのが主流である。このような階層型管理では、情報の蓄積の際に、想定する利用シーンに応じてどの階層（フォルダ）にどのような情報を蓄積するかをあらかじめ決めておく必要がある。つまり共有フォルダの場合、蓄積場所を定めることがすなわち分類することとなり、一義的に分類が決められてしまう。このため複数の観点から情報を発見することが難しい（複製やショートカットを作成して回避する方法があるが管理が煩雑になる）。また、階層が深くなると、共有のための情報（すなわちフォルダ階層パス）が複雑になり、発見にも手間がかかる。

なお、全文検索を併用した階層を意識しない情報発見の方法が考えられるが、検索のためのインデックス作成に時間がかかることから即時性に欠け、また全文検索では不必要な情報が数多く得られてしまうので検索効率が下がることも多い。

### 2.2.2 メタ情報による分類と発見

メタ情報は図1に示すように、対象情報に対し複数個付加することができ、メタ情報の内容変更は対象情報に影響を与えない。たとえば、「著者」や「作成日」など一意にかつ固定的な情報は必須のメタ情報として付加し、「分類情報」「共同執筆者」「要約」「コメント」など多様かつ流動的な情報は必要なものを必要な数だけ付加するといった柔軟な構成が可能であり、これらのメタ情報を利用して対象情報の分類や発見ができる。

また、「情報の管理組織」や「版」など、時間や状況により変更されるメタ情報を対象情報とあわせて管理し有効活用することで、作成・更新日や有効期限といった対象情報の鮮度管理や、閲覧制限管理も可能になる。

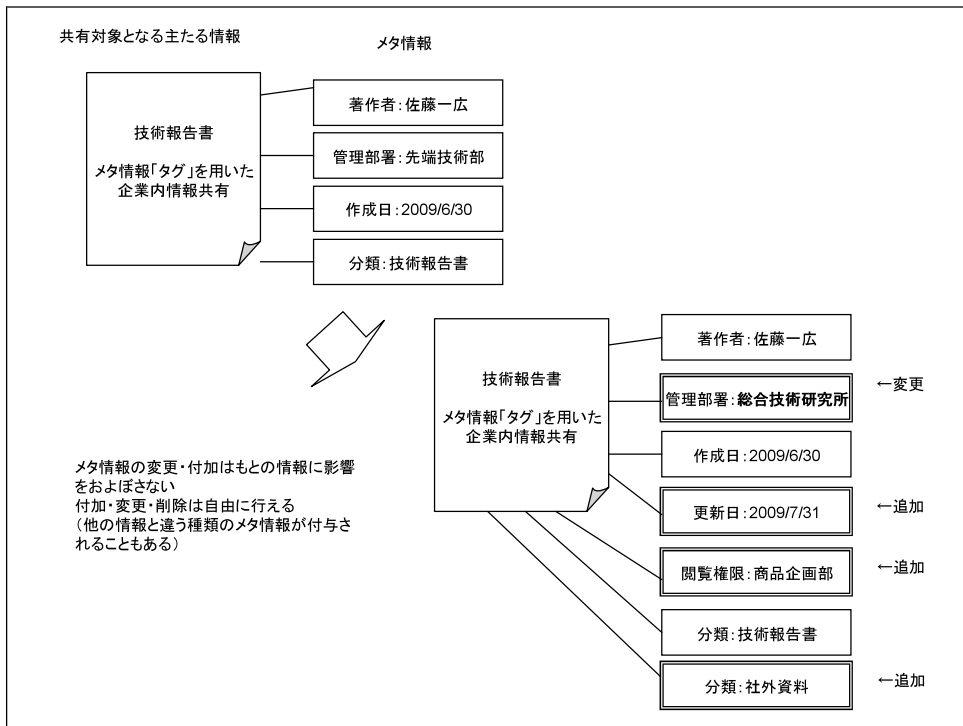


図1 メタ情報の例とその特徴

このようにメタ情報を用いた場合、付加された複数のメタ情報を基に複数の観点での分類ができる。また、蓄積場所もメタ情報で管理され、これらのメタ情報を基に情報が発見できるため、情報の実体はどこにあってもよい。加えて、閲覧権限などをメタ情報として管理すると、その情報自体が柔軟に変更できるため、組織をまたがる情報共有や命令系統を意識した閲覧制限の管理も柔軟かつ容易になる。

### 2.2.3 メタ情報のさらなる活用

前項の権限等の他に「承認状況」や「次工程の文書（文書へのリンク）」などワークフローに関する情報をメタ情報として持たせることで、承認済みの情報だけの絞り込みや次工程の文書などをあわせて発見するといったことが可能である。また、情報に対する「感想」や「分類名」など、情報参照時の条件や状況により内容が異なるような情報もメタ情報として扱い、これも分類・発見の参考情報として利用し、情報発見の効率を上げることが可能である。

なお、情報は状況変化や時間の経過により、本来持っていた意味とは違う意味に変化してしまうことがある。企業においては、提案書・契約書等の文書やそのテンプレートあるいは規則を記した情報などが多数蓄積・共有されているが、それらは利用される時期や組織によって有効性が変化し、時と場合によっては情報として無効なこともある。有効期限の過ぎた情報は、過去の経緯という観点以外にはほぼ無用な存在であり、多くの場合共有されること自体も不適切である。よって文書の「作成者」「作成者の組織」「作成時刻」「有効日時」といったメタ情報が情報本体とあわせて管理され、情報発見の際の見極めのための参考情報として有効に働くことが望ましい。

これらの例のように情報共有においては、情報を作成・分類・参照する際のコンテキスト情報<sup>\*2</sup>もメタ情報として扱い、そのコンテキストにあわせた情報を発見することが望ましい。

## 2.3 メタ情報自体の管理

メタ情報は、任意のタイミングで情報に付加される上、様々な要素で構成されるため XML のような柔軟な構造で管理される。たとえば先の例で挙げたコンテキスト情報の表現には、セマンティック web<sup>\*3</sup> に用いられる RDF (Resource Description Framework<sup>\*4</sup>) や OWL (Web Ontology Language<sup>\*5</sup>) を用いるのが主流である。これらはデータだけではなく意味の疎通にも適したメタ情報表現仕様であり、XML で記述され、語彙の意味や表現方法の標準化がなされている。

これらのことを考慮すると、メタ情報の管理には XML データベースやオブジェクト・データベースが適切であると考えられる。特に、プロトタイピングなどでアプリケーションのデータ構造が頻繁に変わる場合には、このようなデータベースの利用が有効である。

しかし、実システムの構築・運用、特に検索速度を優先させる場合は、これまでに培われた技術やノウハウの蓄積などもあり、リレーショナル・データベースの利用が最適である。実装に際してはメタ情報においてもリレーショナル・データベースのスキーマによって表現・管理することが多く、この表現・管理方法が検索性能にも大きく影響する。

なお、どちらの形態であっても、メタ情報に含まれる語句のインデクシングを行うことや n-gram<sup>\*6</sup> の検索エンジンを組み合わせることで、メタ情報の検索性能が向上する。

メタ情報をキーとバリューであらわされるシンプルな構造に単純化すると、キー・バリュー型データストアが利用できる。このようなデータストアは分散化が容易で大量データを処理することに向いていると考えられており、今後キー・バリュー型データ管理の技術と基盤が整備されるとメタ情報の管理として主流になると考えられる。

## 3. メタ情報「タグ」による情報の管理

### 3.1 タグとは

タグはメタ情報の一種であり、情報に対し任意に付加する名札＝語句のことである。ここで

述べるタグの特徴は、自由な語句であること、情報の登録者や管理者だけではなく利用者が付加することができること、タグの数や付加タイミングに制限がないことなどが挙げられる。

タグによる情報の分類とタグを検索キーにした情報の発見手法は、国内においては「はてなブックマーク\*7」などのコミュニティサイトで積極的に利用され、昨今ではニュースサイトなどでも用いられている。また、ハーバード・ビジネス・スクールのマカフィー教授は、「SLATES」すなわち Search (検索), Link (リンク), Authoring (編集), Tag (タグ), Extension (推奨), Signal (通知) といった Web 2.0 的技術が、企業内でも Enterprise 2.0 として活用されるだろう、と述べている<sup>[1]</sup>。

筆者はこれらの動向や概念に着目し、利用者が付加する自由な語句による情報の分類という手法が企業内においても有効に働き、情報提供者や管理者側がお仕着せで分類するという従来の手法よりも利便性が高く効率的な情報共有につながると考えた。そこで、図2に示すような、ファイルサーバに利用者がタグを付加できる機能、タグによる検索、タグによるリコメンドなどの機能を盛り込み、タグの利用を前面に押し出したツール「tacoPot™」を作成し、企業内情報共有の利便性および効率向上に関する実証実験を行っている。この取り組み結果を基にしたタグによる管理およびタグの活用について述べる。

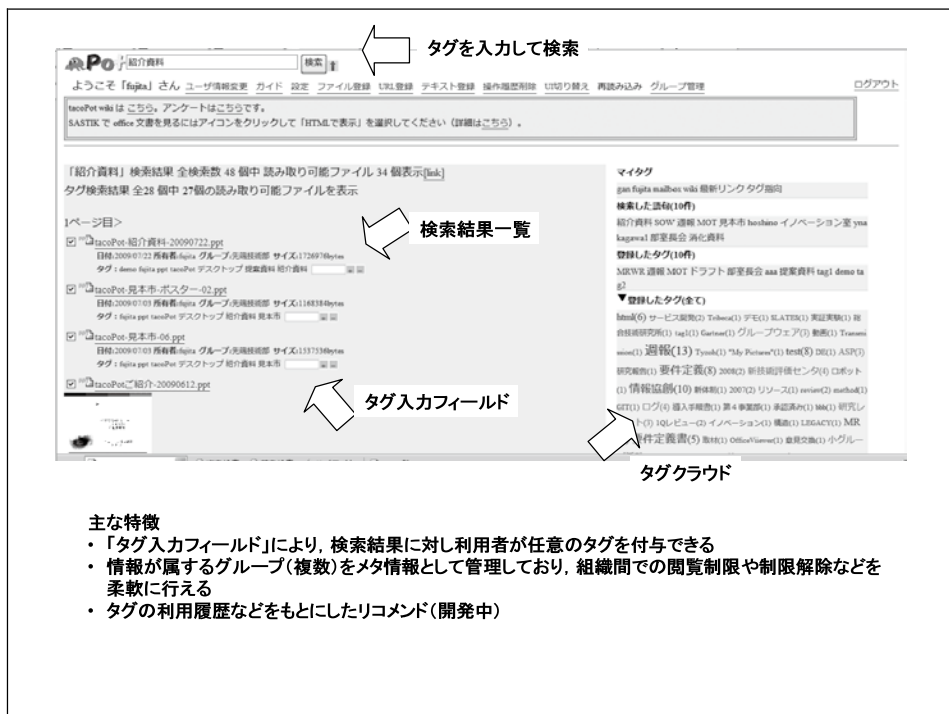


図2 タグを利用した企業内情報ツール「tacoPot™」

なお、一般的なタグは前述のように構造を持たない単なる語句の並びであるが、このような単純なデータでは蓄積されたタグ情報を活用しづらい。そこで内部的には図3で示すように語句を付加した人、付加した時刻等のメタ情報をあわせてタグとして管理している。

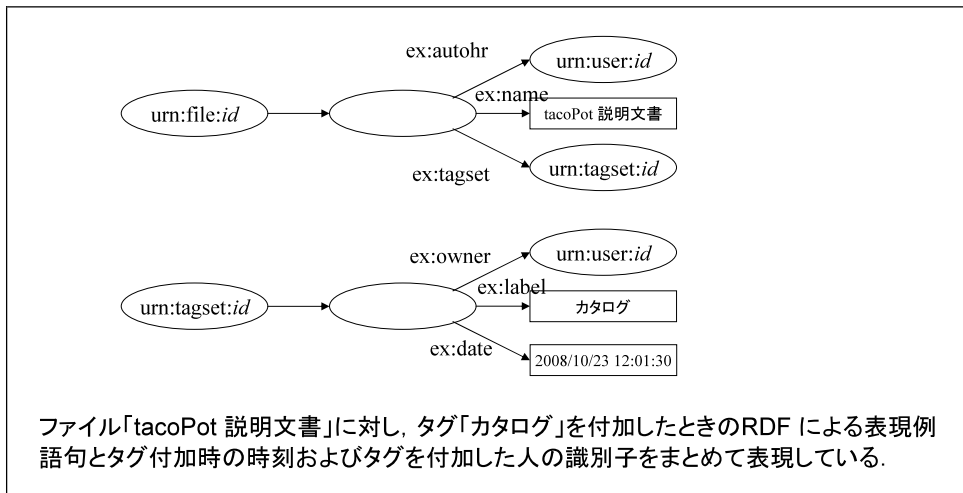


図3 RDFによるタグの構造とその表現例

### 3.2 タグによる管理「folksonomyアプローチ」

タグは任意の語句であるという性質上、必ずしも文書の「作成者」や「作成日時」のような一意に定まる語句ではなく、その表現方法にも厳密な規則はない。たとえば、タグとして“設計書”や“提案書”というようなカテゴリ情報を持たせることもあれば“A社向け提案書”といった固有名詞、あるいは“興味深い”“読みづらい”等の読み手の主観に基づく感想などの自由文の場合もある。なお“提案書”というあたかも一意の意味があるような語句であっても、実際に情報のカテゴリが提案書である必要はなく、たとえば利用者が設計書を提案にも使えろと思えば“提案書”というタグをつけてもよい。このように分類や種別を持たないという点がメタ情報の中でもタグの特徴的な点である。

タグの付加は、情報に多義的な意味（見方）を持たせ、分類・発見を容易にするために行う。複数の利用者が情報にタグを付加し分類を行う方法は folksonomy<sup>\*\*</sup> と呼ばれ、web では頻繁に利用されている。

#### 3.2.1 トップダウンとボトムアップ

多くの企業では、上長や運用管理者の指示により組織が定めた規則や取り決めに応じて、担当者（多くの場合情報の利用者でもある）が情報の収集と分類を行ってきた。この方法はトップダウン・アプローチだが、情報の分類が利用者の視点や利用シーンによっては最適ではないことがあり、情報収集・分類のモチベーションが維持できない上、担当者が分類規則を無視して情報登録を行うなど、取り決めと異なる運用が行われることがあった。

これに対しタグによる folksonomy アプローチは、文書を作成・利用する実運用者がタグをつけ分類を自ら行うというボトムアップ・アプローチである。「情報提供と自発的分类という活動により自身の業務活動がより便利になる」という実感が得やすく、明示的なインセンティブがなくとも利用の促進が進むという効果がある。

#### 3.2.2 ストック型とフロー型

前述のようにメタ情報であるタグを用いる場合にはフォルダ階層をまったく無視した管理が

できる。しかし実証を経て得た経験から、筆者はタグによる管理と階層型による管理は併用を推奨する。タグによる管理は、folksonomy アプローチにより流動的かつ柔軟に管理できるという点が利点であるが、企業内情報には、たとえば社内規定や契約書雛形など厳密かつ固定的な情報も数多く存在し、このような情報の分類・管理は、やはりこれまでの階層型管理が適していることも多いからである。

タグ - folksonomy アプローチはフロー型の情報に、階層型 - taxonomy アプローチはストック型の情報にというように管理する対象により使い分けるとというのが企業内情報管理には望ましい形であると考える。

なお、タグは前述のように厳密な規則はないが運用ルールによる制限は設けることも可能であるため、運用ルールを用いてストック型の情報管理を行うという手段もある。たとえば特定の組織やプロジェクトでは、語彙を制限する、語句の意味を一意に定義する、収集した情報に、あらかじめ定めたルールに応じてタグを付加する、などのアプローチを行い、その上で利用者の主観に基いた分類のための情報を付加することで、より適切な分類・発見が可能になる。

### 3.3 タグの活用

Google や Amazon に代表されるコンシューマ向けの web サービスでは、集められた情報を基にマーケティングを行う等の活用がなされているが、企業においてこのような取り組みを行うことは稀である。

しかし、タグを活用し、その利用履歴を蓄積することで同様の機能を提供することが可能である。タグおよびタグをつける行為に着目すると、任意につけた語句（ラベル）、語句を付ける対象（モノ）、語句を付加した人（ヒト）をタグの主要構成要素と捉えることができる。これらの要素を基にタグの構造を定めると、ラベルとモノ、ラベルとヒト、さらにはラベルとモノとラベルというような複合的関係を得ることができ、図4のような関係を基にした情報の検索ができるようになる。また、タグは利用者の主観による情報であるが、その統計をとることで利用者の総意としての分類を行うことも可能である。

- **タグから得られる情報の例**
  - ラベル(語句)を付加したモノ(情報)の発見 (ラベル-モノ)
  - ラベルをつけたヒト(付加者)の発見 (ラベル-ヒト)
  - 同じモノにつけられた別のタグの発見 (語句-モノ-語句)
  - あるラベルの出現頻度
  - あるラベルの分散状況
  - 類似ラベルの集合
  - ラベルを多く付加したヒト
  - ラベルを多く付加されたモノ
  - あるヒトが付加したラベルの集合と、別のヒトが付加したラベルの集合の重なり具合(同じようなラベルをつけたヒトの発見)
  - あるラベルを付加したモノにつく他のラベルおよびそのラベルが付くモノ(関連する情報の発見)
- タグから得られる情報を基に各種分析や、リコメンドができる。
- この他、システムによりタグの検索頻度や、タグを検索した人などを基にさらなる分析やリコメンドができる。

図4 タグから得られる情報例

#### 4. タグの課題と対策

メタ情報やタグを活用することで企業内情報共有はより効率的になったが、1年半の実運用を通じ「folksonomy アプローチの場合のタグ母集団の確保とその偏りの低減」「タグの意味のずれ（誤解）による分類・発見効率の低下対策」がタグの利用に関する今後の課題であるという認識を得た。

##### 4.1 Folksonomy アプローチのためのユーザ確保

folksonomy アプローチが有効に働くためには、タグを付ける人の数とタグの総数が重要である。これらが不足すると、タグの偏りが顕著になりタグ分類の信頼性が低下する。企業の従業員数は web サービスの利用者数として圧倒的に少なく、特にタグ利用が浸透していない初期の段階では、このような課題が顕在化しやすい。この課題に対しては、タグ利用推進者の存在、推進者による方向付けなどのタグ管理のノウハウの蓄積が重要である。また、タグの利用促進には情報共有システムのユーザ・インタフェースも重要であり、いかに簡便にタグが付加でき、いかにわかりやすくタグを利用した情報分類・発見ができるかを、利用シーンや利用集団のスキルに応じて検討・提供しなければならない。

##### 4.2 タグの意味のずれの解消

情報の分類のために同じ語句のタグを用いていても、その語句の意味を取り違え、誤った分類や意図しない検索結果が得られてしまうことがある。folksonomy アプローチにおいては「誤った分類」というのはいずれ淘汰されるか、あるいは「違った見方」という観点であればそもそも誤った分類ではないとも言えるが、運用管理の面からはイリーガルな情報の削除や、意図を一意に特定できる語句に集約させたいという状況があり得る。この場合においては、分類や検索の際に類義語辞書やグルーピングされたタグを提示するなどの手法が有効である。また現時点ではオントロジー<sup>\*9</sup>を利用する方法が主流となると考えられ、日本ユニシスをはじめいくつかの企業、研究機関で研究が進められている。

#### 5. おわりに

情報の分類・発見にメタ情報を活用すると、柔軟かつ詳細な分類と複数の視点による情報発見が可能である。特にタグを活用し利用者主体の folksonomy アプローチを用いた場合は、利用シーンに応じた分類や利用者の感性に応じた分類・発見が容易になる。

企業内情報は管理対象となる情報の形態や運用方法に応じて適切な管理が必要で、ストック型の情報など階層型管理が適している場合もある。よって、企業内情報共有では、従来の階層型管理等とメタ情報やタグによる管理をバランスよく併用することが望ましい。このような情報共有基盤の提供と運用がなされれば、情報分類・発見の効率や利便性が向上し、企業内情報共有の活性化も図られるであろう。

筆者は、将来的にはメタ情報やタグ情報、コンテキスト情報や利用履歴などの情報から、利用者が情報の分類を意識することなく、得たい情報を半自動で導き出すことができるようになって考えている。また、タグから得られるモノ・ヒト・ラベルの関係から、タグの傾向が似ている情報（モノ）、自分と同じタグをつけている人（ヒト）、その人（ヒト）がつけた別のタグ（ラベル）というような間接的な情報を提供することで、単なるキーワード検索からは思いも



よらなかった情報が得られ、利用者の知識や発想を広げることが可能であると考えている。このように、タグとその管理技術は、単なる情報「共有」から皆の叡智を結集する情報「協創」へとつながる技術と考えており、企業内情報共有ツール「tacoPot™」の開発を通じて、今後も新しい情報共有のあり方を研究していく所存である。

なお、本稿の執筆とタグにまつわる技術開発・研究には、日本ユニシス(株)総合技術研究所原氏から多大なるご助言ならびにご指導をいただいた。改めて感謝の意を表したい。

- 
- \* 1 情報に関する付加情報。たとえば文書における著作権や作成日時など。
  - \* 2 本来は「文脈」という意味だが、コンピュータ用語としては「状況」といった意味で用いられる。
  - \* 3 W3C のティム・バーナーズ＝リーが提唱した、従来のデータ交換だけではなく意味交換を行える web の仕組み、構想。
  - \* 4 web 上にある資源を記述するための統一された枠組み。XML で表現される。
  - \* 5 オントロジー（情報の概念化のための構造/仕様）によりデータ交換をするための言語。
  - \* 6 文字列を構成する一定数の文字の並びを用いて、その集合の頻度や統計情報から検索を行う仕組み。
  - \* 7 <http://b.hatena.ne.jp/> タグや人力検索など新しい形態のソーシャルブックマークサービス。
  - \* 8 「民衆 (folks)」と「分類法 (taxonomy)」をあわせた造語。人々による分類を指す。
  - \* 9 情報の概念化のための構造/仕様。

- 参考文献** [1] Andrew P. McAfee, Enterprise 2.0: The Dawn of Emergent Collaboration, MIT SLOAN MANAGEMENT REVIEW, April 2006
- [2] 野中郁次郎, 「知識創造の方法論」, 東洋経済新報社, 2003/04
- [3] Peter Morville, 浅野紀予, 「アンビエント・ファインダビリティ」, オライリージャパン, 2006/04
- [4] 吉川日出行, 「サーチアーキテクチャ「さがす」の情報科学」, ソフトバンククリエイティブ, 2007/09

**執筆者紹介** 佐藤 一 広 (Kazuhiro Sato)

1990年日本ユニシス(株)入社。オブジェクト指向開発環境、分散オブジェクト・プラットフォーム、ASP基盤などのプロダクト主管にて開発に従事。2006年よりR&D部門において新技術の研究開発に取り組む。

