

## 情報フロンティア研究会（第3回） 議事要旨

1. 日時 平成17年4月12日（火） 16時～18時
2. 場所 総務省 1101会議室
3. 出席者（敬称略）
  - (1) 構成員  
國領二郎（座長）、木村忠正（座長代理）、岡田仁志、栗原聡、齋藤義男、津田宏、西村毅、藤沢久美、柳沼裕忠
  - (2) 総務省  
鈴木政策統括官、松井審議官、武田情報通信政策課長、村手地方情報化推進室長、内藤情報通信政策課課長補佐
4. 議事概要
  - (1) 開会
  - (2) 議題
    - ICTの利活用におけるフロンティア
      - ・ネットワークを流通する膨大なデータ情報の活用

※資料3に基づく事務局説明、資料4に基づく津田構成員からの発表、資料5に基づく栗原構成員からの発表の後、質疑応答・意見交換が行われた。

  - ・情報フロンティアの世界では、個人がモバイルで1テラバイトのWebサーバを持つことも可能であり、個人の持つ1テラの情報がWeb上から見えるような仕組みを通じて、人と情報が張りついて、Webの世界とリアルな世界がつながった社会が生まれるかもしれない。その場合、どういう情報交換の仕方があり、新しいアーキテクチャはどのようなものが問題となる。また、情報家電やパソコンは人格を失いつつあり、ディスプレイ、ハードディスク、チューナー、タグといったものがセンサーネットワークと連携して、テレビになったり、パソコンになったりするといったWeb連携、デバイス連携が進むと思われる。そういう意味で、機器のユビキタスというよりは機能のユビキタスであり、これはフロンティア的な考え方だと思う。
  - ・すべての情報家電がモジュール化されれば、新しいシステムをつくる際にはそのモジュールを組み合わせるだけで済み、重要な課題は組み合わせるための共通プロトコルになる。さらにWebとどのように連携させていくのかも問題の一つ。
  - ・UWBになれば、プリント基板上のものがアンバンドルしても遜色ない世界になるのではないか。そういう意味では、アンバンドルがキーワードになるのではないか。
  - ・セマンティックWebや電子タグも同様だが、これは要するに方法論の問題であり、何を

ユニットとして考えれば一番うまくいくかという問題に過ぎない。

- ・電子タグについては、セマンティックWebのRDFを使おうという動きがある。情報家電についても、セマンティックなものを取り入れようという動きはあるが、総じて下位の規格が乱立している状況。情報家電の場合は、キラーアプリ、キラーシーンを明確に打ち出さなければ普及しない。
- ・情報家電をモジュール化すれば、消費者は無駄なものを重複して買わなくて済むというメリットがある代わりに、知恵を使わなければならない。メーカーがそれをある程度まとめてくれると、消費者は重複して買わなければならないかもしれないが、代わりに手軽さを手に入れることができる。おそらく日本の消費者の9割近くは、多少重複があっても手軽さを求めるであろうし、その方がメーカー、消費者の双方にとって案外コストが安い場合もある。また、セマンティックWebについては、誰がメタデータを書くかが重要で、そのためのインセンティブモデルを行政などが用意しないとうまくいかないと思う。その際、Google や goo などの検索エンジンをうまく取り込むべき。
- ・インターネットはボランタリーな世界であり、インセンティブについても、金や地位・名誉とは違う切り口からやりたくなるような仕組み・環境づくりをすべきではないか。
- ・(オントロジーの構築に対する消極的な意見に対し) 要はレベルの問題であり、例えばオントロジー技術による満足度の度合いが非常に高いものであるということに皆が納得できれば、自発的なオントロジーの構築が促進されるだろう。納得させる部分を、Google などがオントロジーを積極的に採用し、押し上げてくれるというストーリーは現実味が高い。現状では方法論が先行しており、出来た技術を何に使うかをその後で考え始めるというように、目的が二の次となっている傾向も見受けられる。
- ・セマンティックWebはカテゴリーに対する考え方が認知主義的であり、明確な定義づけが可能だが、我々のカテゴリー形成の仕方は必ずしも実世界と1対1に対応できるわけではない。それでも、人間の思考を広げていくことには学術的な意味があり、セマンティックWebのような考え方を進めることは重要。Webを基盤にして様々な技術開発が行われているが、実際に使える範囲は非常に限定されている。情報フロンティアを考える上での重要な戦略的な分岐点として、ある特定の目的のためだけに技術開発するという考え方だけでなく、目的にとらわれずWebに対してチャレンジしていくことにより新たなものを生み出すというモデルもありうる。
- ・セマンティックWebに対する過剰な期待は確かに存在する。ただ、セマンティックWebで記述できる意味は、何の知識もないソフトウェアが曖昧なく処理できるレベルに過ぎない。ただし、そのレベルであっても、特定のアプリケーション、特定のコミュニティーの中で、現行のWeb検索ではできない有用性があり、徐々に普及していくのではないかと。

- ・データマイニング技術が発達すると、セマンティックWebの考え方は必要なくなるかもしれない。また、セマンティックを考えるということは、共通の言語体系を持つグループを前提として、グループ間のインターオペラビリティを考えることになるのではないか。
- ・自然言語処理などAI的な技術は、新聞などの明瞭な文章に対してはかなり自動化できるが、Webのような雑多な文章に対しては難しいのが現実。セマンティックWebは閉じたコミュニティの中で言葉の概念を共有するもので、考え方としては単純だが、そうしないと情報のやりとりで不正確性が生じてしまう。異なったコミュニティ間では、そうした概念の変換を行えばよいという発想だが、現実のビジネスでは、案外そういう考え方でもうまくいくことも多いと思う。
- ・セマンティックWebは一過性では終わらないと思う。セマンティックWebは基本的には文献の中身のある程度把握した上で検索するものであり、マイニングを行う際に、ベースとしてセマンティックWebがある場合とない場合では、結果が大きく異なってくる。
- ・通常、マイニングするには前処理として生データをマイニングにかけられる状態にする必要がある。その際、WebにおいてセマンティックWebのタグを使うことが出来れば、前処理を含むマイニングの実行効率が向上するだろう。そもそもWebのコンテンツは人によって作られているのであり、様々なWebのコンテンツ自体を解析しようとしても限界がある。メタなタグや、また、Webの特徴であるリンク関係に着目すれば、コンテンツに踏み込まずに効果的に情報抽出が行えるのではないか。
- ・我々の実社会には、単純にディフィニションで収集して検索結果を一覧表で表示するのは違う世界があると思うが、そうした側面はネットワーク科学と関係するのか、あるいは、セマンティックWebはそこまでカバーしようとしているのか。
- ・現状のようにブログや掲示板を検索してその内容を読んでといったことをやっている間は時間がかかる。ブログのトラックバックを利用して特定の商品やサービスの善し悪しがリアルタイムに判定できるようになると面白いと思う。リンクの書き方を拡張し、ブログのコメントに対する善し悪しなどをタグに入れられれば、すぐにでも実現可能。あるいは、自然言語処理で文面から、良さそうなことを言っている、悪そうなことを言っているというのを判別することもある程度可能。
- ・ネットワークロボットに関して、セマンティックWebによる意味づけによって、人間がやらなくてもいい領域をロボットにやらせたり、ロボット間で情報交換をさせたりすることが可能になるのではないかという議論があった。高齢化社会の進展により人以外が行うことが増えていくという意味では、この領域を残しておくのは重要ではないか。
- ・特に情報家電と絡めると、日本はそうした分野に強みがある。介護ロボットなどに取り入

れることも可能。

- ・オントロジーやセマンティックWebで鍵になるのが、「どのように構築するか」ということ。様々な場面において、アドホックに効果的な構築方法を見つけていくという現状が見受けられるが、そのような過程を繰り返し多くの事例を蓄積していくというやり方が実は現実的なものかもしれない。
- ・セマンティックWebやオントロジーは、仮に枠組みができたとしても、誰がやってくれるのが問題となり、結局はボランティアというのが答えになるという考え方がある。調べたいことがGoogleでうまく見つからない場合、オントロジーを共有しているコアな人を見つけて、その人のホームページを見た方が実は早いことがある。それはブログでも同様で、オントロジーを共有している人のブログを見つける方が効率的であり、ブログの幅もオントロジーを共有するコミュニティーに限られていくのではないか。
- ・セマンティックWebやマイニングは、今まで不可能だったことを可能にしてくれる技術ではなくて、人間がやらなければならない低レベルのことをサポートしてくれるという技術に過ぎないのに、期待し過ぎている分だけ評価が下がってしまうことを懸念している。現実には、ネットを検索するよりも気のきいた友達に聞いた方が圧倒的に良い結果が得られることも多く、不可能を可能にしてくれるという幻想を抱かさないようなプロモーションが必要なかもしれない。
- ・今後は様々なものが複合してくると考えられる。大きくて薄い口コミサイトは広告をとりやすいものの、実際に興廃をトリガーしているのはブログのような小さくて濃いコミュニティー。こうした小さいコミュニティーは広告モデル、ビジネスモデルがないのが弱みだったが、最近では収益モデルが確立されつつある。大きな口コミサイト業者は、いかにブログの現象とつき合い、ビジネスモデルに取り込むかを真剣に考えており、Small World論でいくと、レギュラーな世界とリンクをいかにハイブリッドで組み合わせていくかという問題が生じつつあるのではないか。
- ・今後は様々な技術が複合化するため、セマンティックWebとマイニングだけでなく、他の問題も複合して考える必要がある。例えばセマンティックWebが進展した際に、WebサービスのUDDIはどうなるのかというように全てが絡んでくるので、一通り色々なテーマで話した後、全体を踏まえて考える必要がある。
- ・今までのビジネスプロセスでは、特定の市場やコミュニティーに絞って何が売れるのかを考えてきたが、インターネットが広がり様々な情報が混沌と存在している中で、そのコミュニティーを再度つくるには共通言語が必要であり、そのためにはメタデータが普及するように定義づけを進めなければならない。何らかの共通言語に基づくコミュニティーが出来上がれば、規則性・法則性は見出しやすいので、手段が何もない状況下よりは精度は上がり、結果としてWebの世界で新しい市場の創出につながる。また、メタデータに関し

ても、誰がどう付けるのかといった成功モデル、インセンティブモデルの議論は重要。

- ・メタデータの主導権は検索エンジンがとる。検索エンジンの主導権は日本にはない。そうすると、メタデータの主導権は日本にはないことになる。
- ・5、6年前は日本にもサーチエンジンが多数存在したが、ここ5年のうちに全て Google に置き換わってしまい、今では Google ベースでないものはほとんどない状況。Google、Yahoo、MSNの3強状態を変えるためのツールの1つがメタデータ。メタデータが付くことにより詳細な検索が可能となり、一般のサーチエンジンで満足できない人を取り込むことが可能となる。
- ・テレビのリモコンのボタンをどれくらい押したか、DVDに何を録画したかといった情報は検索エンジンには引っかけられないが、パーソナライズする際に価値のある重要な情報となる場合もある。そういう意味で、情報家電の分野では Google 一人勝ちというわけではなく、要はやり方ではないか。
- ・現在のところ Google が優勢であるのは、検索エンジン自体が優れているのと同時に、検索エンジンとしてのシステムの拡張性や耐故障性が非常に高いからである。既に Google が膨大な検索データを保有しているのは事実であり、例えば Google がある画期的なメタデータ規格よりも劣るメタデータ規格を採用したとしても、膨大なWebデータに後者の規格にてタグが付けられてしまえば、前者の規格が優れていても実際問題として使われることはない。必ずしもよい製品が世界を席巻するわけではないのが悩ましいところだが、放っておく訳にもいかない。
- ・我々はITリテラシーの比較的高い人間であり、Google の存在の大きさを認識できるが、一般的には検索エンジンに対する認識はそれほど高くない。したがって、決して Google が一人勝ちしている状況ではないと思う。
- ・ITリテラシーの低い人は、ネットで検索するのではなく誰かに聞く。個々のニーズはキーワードをいくつか入れる程度で済むものではなく、より感覚的なものである。そういう意味では、Google はインターネット発展の過程におけるプリミティブな部分に過ぎず、今後は多くの背景情報も含めて、共感する誰かが教えてくれるといったことが必要となってくるのではないか。
- ・将来に向けてどのような戦略が立てられるかを考えていきたい。これまで様々な議論を重ねてきたが、全体を貫通するテーマを導き出す必要があると思う。

以上