特集 「編集委員今年の抱負 2011」

統・電子化担当の今年の抱負

森田 武史

慶徳義塾大学大学院理工学研究科

2年前の特集では、「電子化担当の今年の抱負」と題して、「1. 論文査読のワークフロー化」、「2. スパムメール対策」、「3. Webサイトの半自動更新」の三つを掲げた。しかしながら、1と3については、手付かずとなってしまった。2については、根本的な解決はできていないが、スパムメールと正式ML配信依頼メールの件名やメールアドレスを分析して、10数個のフィルタをGmailで作成し、スパムフィルタと併用することで、誤って正式なML配信依頼メールがスパムメールと判定される場合は、最近では、ほとんどなくなった。

1については、JSTが構築したJ-STAGEをAI学会論文誌の査読のため導入するという話であったが、AI学会全国大会で利用しているシステムが再利用できるのであれば、という話も出され、それ以降、この話はストッピングをしていると思われる（自分の知らないところで、準備が進んでいるかもしれない）。

3については、全く手付けずで、特に「注目の話題」については、更新が滞っていることが多い、関係者の皆様はご迷惑をおかけしております、大変申し訳なく思う。

2010年11月号（Vol.25, No.6）の「編集室」でも書かせていただいたが、この問題を解決するために、Webページにメタ情報埋め込むための仕様であるRDFaを利用できるではないかと考えている。RDFaはGoogleのリッチメタデータなどで利用されており、Webページ作成者がWebページにメタデータを埋め込むことで、検索結果に表示されるスニペットのカスタマイズが可能になる。

現在、AI学会内には、19の研究会があるが、それぞれ個別にWebサイトを構築しており、特にそのアイデアが重複されているため、Webスクレイピングなどで、最新の研究会開催日や開催回数などを各研究会Webページから導入することを試みる。そこで、RDFaにより各研究会Webサイトのトップページに、最新の研究会開催日や開催回数に関する二つの属性情報を埋め込んでいただくことができれば、自動的に各研究会のトップページが更新される。それに伴い、それらの属性情報を抽出することで、AI学会HPトップページの研究会日程や注目の話題の自動更新が可能になると考えられる。さらに、研究会の開催地の情報や、特に知らせたい内容についても、別途属性情報を埋め込んでいただくことができれば、注目の話題などをまとめて自動的に掲載することも可能である。

私自身、セマンティックWebの研究に携わってきたこともあり、上記の案も含めて、何かしらAI学会WebサイトにセマンティックWeb技術を導入することを電子化担当の今年の抱負としたい。

せっかくの機会なので、研究に関する抱負についても述べさせていただきたい。私はこれまでセマンティックWebとオントロジーを研究分野として、特に、オントロジーやRDFコンテンツの半自動構築の研究に取り組んできた。これに関連して、セマンティックWebの分野では、近年、各Webサイトで公開されている政府、地理、科学、音楽、写真といったさまざまなデータベースをRDF化して連携する、Linked Open Data（LOD）と呼ばれる活動が活発に行われている。残念ながら参加することができなかったが、セマンティックWeb分野とのトッププロジェクトであるISWC2010では、LODの話題が中心であったと聞いている。


以上の背景から、私が所属する研究室でも、日本語ウィキペディアからオントロジーをRDF化を含む研究に取り組み始めており、構築したオントロジーやRDFデータを公開している[4,5]。現状では、約300万トリルと欧米のLODにはほど遠いが、少しでも日本においてLODがいても、セマンティックWebの関係が高まるように、研究を進めていきたいと思う。

そこで、今年の研究に関する抱負としては、日本語ウィキペディアから構築したオントロジーと日本語WordNet[6]を統合し、公開することとした。

◇参考文献◇