

第29回

Webインテリジェンスと インタラクシオン研究会

2011年3月7日～8日
学術総合センター（東京都千代田区）

土方嘉徳 大阪大学大学院 基礎工学研究科

WI2 研究会とは

2011年3月7～8日に、第19回Webインテリジェンスとインタラクシオン研究会が、学術総合センター（東京）で開催された。上記研究会の略称は“WI2研究会”で、「ういっつ」研究会と呼んでいる（本稿でも以降は、“WI2研究会”と記す）。Web上のコンテンツからの知識発見、Webにおけるネットワーク構造からのコミュニティ抽出、Web上の人間同士のコミュニケーション支援など、Webにおける人間の知的活動支援全般を研究対象としている。Web技術の研究全般を扱う研究会としては、日本では唯一のものである。筆者は、WI2研究会の専門委員長を務めさせていただいているのであるが、本稿では、その設立の経緯、理念、主に発表される研究テーマ、これまでの取り組み、最新の研究会の様子について、紹介したいと思う。

設立の経緯と理念

WI2研究会は、2004年7月に電子情報通信学会・ヒューマンコミュニケーショングループより、第2種研究会として設立された。1989年に誕生したWebは、ネットワーク機能の充実したパソコンの普及に伴い、急速に発展してきた。既存の研究分野においても注目を集めることとなり、人工知能、データベース、自然言語処理、ヒューマンインタフェースなど、多くの分野で研究が行われるようになっ

た。筆者も1995年からWebの研究を始めたのであるが、当時日本でどのような研究が行われているのかを網羅的に知ろうとすると、これら多くの研究分野で行われている研究会や全国大会、シンポジウムなどに行く必要があった。私自身も、人工知能学会、データベースシステム研究会、言語処理学会、WISS、インタラクシオンなど、多くの学会に足を運んでいた。

それぞれの学会には個性があり、人工知能学会ではオントロジーやセマンティクスを扱ったアプローチが多く、言語処理学会ではWeb検索のためのテキスト・自然言語の処理方法に関する研究が多く、データベースの分野では、自然言語だけでなくHTMLの構造なども用いた情報検索・抽出手法などに関する研究が多かった。ヒューマンインタフェースの分野では、Web上のデータの処理方法よりも、その見せ方を工夫することにより、ユーザの情報獲得を支援するような研究が多かった。しかし、研究者らが目指しているゴールは共通しており、それはこのWebという知識の宝の山から、いかに人間が容易に目的の情報を獲得できるかという点であった。個別の分野で活躍している研究者らが、1つの問題に対して異なる見方やアプローチを示し、それらを共有できれば、そのゴールに早く近づくことができるのではないかと考えた。これが、Web技術を専門として扱う研究会であるWI2研究会ができた理由である。

WI2研究会が扱うテーマ
(A) テキスト処理関連技術
情報検索, 情報抽出, 自動要約, 質問応答, テキストマイニング
(B) 知的Web基盤技術
XML, メタデータ構築支援, リンク解析, アノテーション作成支援, セマンティックWeb, マルチメディア検索
(C) Webインタラクション技術
意思決定支援, パーソナライゼーション, 情報推薦, ユーザ行動分析, 情報可視化, ソーシャルネットワーク分析

表-1 WI2 研究会が扱うテーマ

WI2 研究会が発表を募集するときに掲げているテーマを表-1に挙げる。知識や情報の流通に必要なフレームワークから、Web 上のコンテンツを解析する技術、Web 上でのヒューマン・コンピュータインタラクション技術まで、広く包含していることが分かる。

これまでの活動

これまでに開催された研究会の発表数と参加者数を表-2に示す。研究会は、年に2～3回行っている。参加者数は、開催地により大きく変動するが、地方開催では80人程度、東京開催では200人近い参加者が集まることもある。共催はなく、すべて単独開催である。研究会の特色としては、

- シングルセッション
- ロング発表とショート発表の2種類の発表スタイル
- 招待講演やチュートリアルなどの企画
- 企業からの参加
- 懇親会
- Twitter と Ustream

などが挙げられる。1つの発表を参加者全員が聞くことになるので、情報が共有化され、その後の議論や発展につながりやすいと言える。発表は、ロング発表(20分発表, 10分質疑)とショート発表(13分発表, 5分質疑)の2つのスタイルがあり、かなり完成された研究を発表することもできるが、ポジションペーパーのようなまだアイディアすら固まっていないような研究も発表し、さまざまな人の意見

回	ロング 発表件数	ショート 発表件数	招待 講演数	参加 者数	開催地
1	9	0	3	90	東京
2	13	0	4	83	大阪
3	19	0	4	76	広島
4	9	0	2	81	京都
5	24	0	2	153	東京
6	13	13	3	71	札幌
7	15	11	2	172	東京
8	15	3	2	73	名古屋
9	13	5	0	42	湯布院
10	12	14	2	164	東京
11	10	8	6	70	北九州
12	12	12	5	85	淡路島
13	12	16	5	155	横浜
14	17	9	1	77	島根
15	7	11	4	82	広島
16	8	14	2	134	東京
17	16	15	3	120	大阪
18	8	10	1	42	佐渡島
19	11	13	2	113	東京

表-2 発表件数と参加者数の推移

を得ることもできる。いずれの発表形式においても、長めの質疑応答の時間を取っている。また、必ず招待講演やチュートリアルなども入れ、参加者のレベルアップができるよう心掛けている。企業からの参加者も多く、東京での開催では企業関係者が4割ほどを占めることもある。毎回、懇親会も企画し、交流が深められるようになっている(写真1参照)。発表を支援するサービスとしては、発表スライド用のスクリーンだけでなく、Twitter用のスクリーン(写真2,3参照)も導入し(2009年10月より導入)、議論の活性化を図っている。また、研究会に来られなかった人の意見を取り入れるべく、第19回研究会ではすべての発表のUstreamでの配信も行った(写真4参照)。いずれにしても、研究会がこれらのサービスを本格的に導入したのは、WI2研究会が初めてである。

研究会の開催以外にも、参加者が広く活躍できる



写真1 懇親会の様子(中之島センター交流サロンにて)

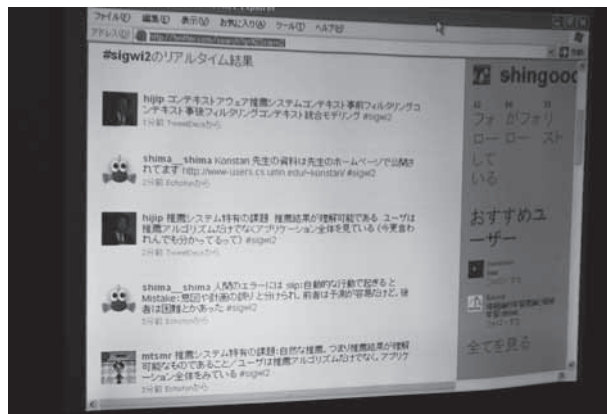


写真3 Twitterの画面の例



写真2 発表用スクリーンとTwitter用スクリーン

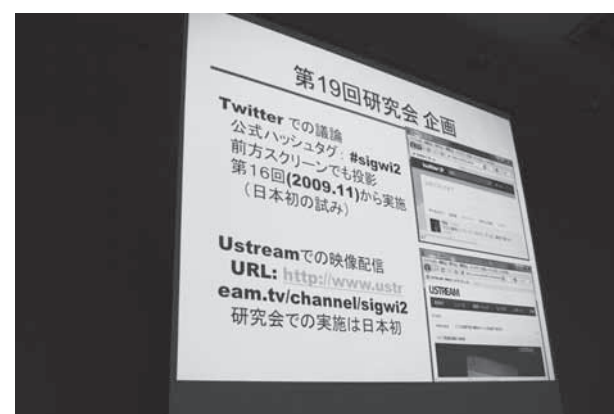


写真4 TwitterとUstreamでの配信

場を提供することも目指しており、WI2研究会の実行委員や同じ分野の研究者とともに、以下のような企画も実施してきている。

- 国際会議オーガナイズドセッション企画 (AMT'05, SMC'08)
- 国際ワークショップ企画(国際会議 DASFAA'10 内 SNSMW'10)
- 学会誌解説特集(人工知能学会誌特集「ソーシャルネットワーク時代の Web インタラクション」)
- 論文誌特集号(日本知能情報ファジィ学会, 人工知能学会, 電子情報通信学会)
- 国内シンポジウム内企画(DBWeb 特別企画)

この中でも、特に論文誌の特集号は重要な企画だと考えている。Web そのものが比較的新しい仕組みであるために、Web 技術の研究は分野横断的なものになる。そのため、Web 研究の専門家による評

価が必要となる。特に情報発信や共有のためのフレームワークや、ネットワーク・コンテンツ生成における人間の社会科学、認知科学、心理学的な影響に関する研究などを評価するには、既存の1つの研究分野の知見だけでは、正しく評価することができない。

第19回研究会

最新の研究会がどのようなものであったのかについて紹介する(写真5参照)。3月7～8日に行われた研究会では、ロング発表11件・ショート発表13件の一般発表と、1件の招待講演と特別企画があった。招待講演は、セマンティック Web 分野の Linked Data (リンクするデータ) について、ゼノン・リミテッド・パートナーズの神崎正英氏に紹



写真5 第19回WI2研究会の様子



写真6 神崎正英氏による招待講演

介いただいた（写真6参照）。セマンティック Web は、計算機のための Web、すなわち計算機による知的な自動処理を行うための Web を構築することを目指した仕組みであるが、Linked Data は、それをインスタンスの記述から取り組もうとする新しいアプローチである。セマンティクスを扱う人工知能、データからの情報を抽出するデータ工学と自然言語処理、人間によるデータ付与の仕組みを実現するヒューマンインタフェース技術を必要とする、まさに WI2 研究会に合ったテーマと言える。神崎氏は、Linked Data が誕生した経緯や理念について分かりやすく話され、感銘を受けた参加者も多かったようであった。

特別企画は、国際会議 ACM Recommender Systems 2010 の参加報告を、立命館大学の奥健太氏に講演いただいた。上記国際会議は、WI2 研究会でもよく発表される情報推薦のトップカンファレンスである。2010 年の一般発表を分類すると、ソーシャルネットワークに基づく推薦 9 件、伝統的な精度比較・アルゴリズム 6 件、推薦精度以外の評価指標 4 件、グループベース推薦 3 件、その他 3 件とのことであった。やはり、Twitter や Facebook などのニューメディアの登場により、Web 上での社会関係を用いた推薦が注目を集めているようであった。

一般発表では、情報検索・抽出、時系列データの応用、ユーザ行動分析・推薦、分類・クラスタリング、

コンテンツ閲覧と発見に関するセッションがあった。特に筆者が注目した発表としては、(1) Linked Data を利用した情報拡張手法に関する研究、(2) Web からのアドバイス文抽出に関する研究、(3) キャプチャ画像を利用した携帯専用サイト識別手法に関する研究、(4) 同姓同名人物判断に関する研究がある。

(1) では、ユーザが Web ブラウジング中に興味のあるテキストを選択することで、それに関連する情報を、Linked Data の情報を利用して提示する方法を提案している。Linked Data の情報は多種多様であり、その信頼性も多様になってくると思われる。本手法のように手軽に関連情報を閲覧できる手法が、今後必要になってくると思われる。

(2) では、Web 上の情報からアドバイス文を判別する方法を提案している。特に、ユーザの外出時の情報獲得を支援することを想定している。アドバイス文とは、「八合目より上は、傾斜がきついので、ゆっくり歩きましょう」のような、後の利用者が行動するとき役に立つような情報を含む文のことである。筆者の知る限り、このようなタイプの文を判別しようとした試みは初めてであると思われる。かなり難しい課題であると考えられ、新しい分野の開拓につながる研究であると言える。

(3) は、KDDI 研究所による研究である。外部から、ある Web サイトが携帯専用サイトかどうかを識別するには、携帯電話と PC から個別にアクセス

しないと分からない。セキュリティの厳重な携帯専用サイトは、HTTP リクエストの User-Agent だけでなく、IP アドレスのチェックも行っており、PC から User-Agent を偽装してアクセスしても、携帯用のコードを獲得することはできない。提案手法では、携帯電話の画面と PC の画面とをカメラで撮影し、同一コンテンツかどうかを判定している。

(4) では、Web 上における同姓同名の人物の判断を人手で行っている。自然言語処理の分野では、同姓同名の人物の判定は、非常にポピュラーな課題ではあるが、人間がそれをどのように識別しているのかは明らかにされていない。本研究は、このような人間の認知特性を初めて明らかにした研究であると言える。

これらはあくまで発表の一例であるが、WI2 研究会ではこのようなコンピュータサイエンスとヒューマンサイエンスの両方のアプローチから、Web 上の問題に取り組む発表が多い。また、既存の研究分野に対して一石を投じるような課題やアプローチの提案も多い。WI2 研究会が多くの研究分野から人が集まっているからこそ、このようなイノベーションが起きているのではないと思われる。

新しい潮流に向けて

これからも WI2 研究会では、Web で起こる新しいサービスや潮流に焦点を当てて、これまでにない研究テーマを追いつけることになると思われる。Web はさまざまな社会性を持つ人間が集まり活動することで、コンテンツやネットワークが形成される。純粋なコンピュータサイエンスだけでなく、社会科学、認知科学、心理学など、多くの研究者との共同が必要と考える。興味を持った方々にはぜひ参加をお勧めしたい。

WI2 研究会 Web サイト

<http://www.ieice.org/~wi2/>

(2011 年 3 月 24 日受付)

土方嘉徳（正会員） | hijikata@sys.es.osaka-u.ac.jp

1998 年大阪大学大学院基礎工学研究科修士課程修了。同年、日本アイ・ビー・エム（株）東京基礎研究所入社。現在、大阪大学大学院基礎工学研究科准教授。情報推薦、テキストマイニングなどの研究に従事。博士（工学）。

