

JADH2012

9月15日から17日にかけて、東京大学工学部2号館にて、日本デジタル・ヒューマニティーズ学会主催による国際大会 JADH2012 が開催された。メディア・コンテンツ総合研究機構のデジタル・ヒューマニティーズ・コアは同大会に共催として関わり、石田英敬（副委員長）、吉見俊哉、永崎研宣（事務局）、滝浪佑紀、中路武士が組織委員として参加した。また、吉見俊哉と石田英敬はそれぞれ、開会式と閉会式において、基調講演を行った。また、メディア・コンテンツ総合研究機構が招聘した Harold Short 教授も、閉会式において基調講演を行った。

以下は、メディア・コンテンツ総合研究機構が招聘した Elena Pierazzo 博士によるプレカンファレンス・ワークショップ（永崎研宣「TEI の現在・未来 JADH2012：TEI ワークショップを振り返って」）と同大会（Espen S. Ore「JADH 2012 報告」日本語訳：滝浪佑紀・永崎研宣）の報告である。各々の初出は、『人文情報学月報』No. 19（2013年2月、<http://www.dhii.jp/DHM/dhm19-2>）および『人文情報学月報』No. 15（2012年10月、<http://www.dhii.jp/DHM/dhm15>）であり、以下はそこからの転載である。

「TEI の現在・未来 JADH2012 : TEI ワークショップを振り返って」

: <http://www.jadh.org/jadh2012>

(永崎研宣：人文情報学研究所)

やや旧聞に属する話で恐縮だが、デジタル・ヒューマニティーズと人文学の将来を考える上で重要なワークショップが開催されたにも関わらず十分な報告をする場がなかったため、この場を借りてご報告させていただきたい。

Japanese Association for Digital Humanities (日本デジタル・ヒューマニティーズ学会) による年次学術大会「JADH2012」が東京大学工学部 2 号館にて開催されたのは 2012 年 9 月半ばのことであった。世界各地で開催されるデジタル・ヒューマニティーズ関連学会の例にならい、「JADH2012」においても、シンポジウムに先立ってワークショップが開催された。例年はテキスト解析とテキストアーカイビングの二つのワークショップが開催されてきていたが、「JADH2012」では、テキストアーカイビングのみのハンズオン・ワークショップが開催された。

折しも、Text Encoding Initiative (以下、TEI) 協会 (*1) では、文学・言語学研究者を中心としつつデジタル化に取り組む様々な人文学研究者を巻き込みながら 1987 年より作成・改良を続けてきた人文学資料の符号化 (マークアップ) 手法、TEI ガイドラインの最新版、P5 ver. 2.0 を前年末にリリースしたところであった。この版での改訂とは、それまでの TEI ガイドラインの前提であった「テキストは媒体に左右されない論理的な構造を持っておりそれをマークアップして共有することを基本とする」という考え方に加え、「媒体上に見られる物理的な構造をもマークアップして共有する」という考え方を導入した、画期的なものであった。

TEI 協会の手稿分科会 (Manuscript SIG) の長年に渡る成果であり、「JADH2012」では、この手稿分科会を議長として率いつつキングズカレッジ・ロンドンでデジタル・ヒューマニティーズ学科の講師を務める Elena Pierazzo 博士が招聘され、この新しい TEI ガイドラインに関するハンズオン・ワークショップが開催されたのであった。

とは言え、この TEI ガイドラインの最新版の意義を把握し、ハンズオン・ワークショップで例示される操作を理解しながら講義に追随していくためには、その前提となる TEI の基本的な知識と、XML エディタを用いた基本的な操作方法の習得が必須であった。もちろん、TEI を実際に利用する際には、たとえば、MediaWiki のボタンを押すと裏では必要なタグをつけた上に表ではスタイルシートで対応する見た目を表示してくれる、というような形で、特に XML を意識することなく利用できるような仕組みを用意しておくのが最近は一般的に

なりつつある。しかしながら、TEI ガイドラインにおいて提示されている XML の構造とは、欧米の人文学研究者達が中心になって人文学に関わるテキスト資料の構造について長年の議論を重ねてきた結果としての一つの見解であり、それを直接に知ることは様々な面で大きな意義がある。

すでに、TEI ガイドラインのワークショップとしては、前年の JADH2011 ではトリニティカレッジ・ダブリンの英文学科にてデジタル・ヒューマニティーズ・コースの教鞭をとる Susan Schreibman 博士を招聘しての Versioning Machine に対応した複数写本のパラレル・コーパスの作成を取り扱っており、そのコースを習得した人に関しては特に問題なく「JADH2012」のワークショップにも対応できる見込みであった。そこで、「JADH2011」のワークショップに参加せず「JADH2012」に参加希望をした人々に対しては、準備のためのプレワークショップが開催されることとなった。

結果として、東京大学及び近辺にて 3 回、大阪大学にて 2 回にわたり、無料参加のプレワークショップが開催され、TEI の基本的な知識と XML エディタ oXygen を用いた具体的な操作方法についてのレクチャーが行われた。oXygen は JAVA で書かれた有償の XML エディタソフトであり、TEI 協会等のアカデミックなコミュニティと関係を築きながら開発が続けられてきたものである。XML タグを直接記述したいという人にとってはかゆいところに手が届くたくさんの機能が用意されており、一度 XML に馴染むことができれば大変使いやすいものである。

そして、アカデミック価格が用意されているため、研究者や学生は比較的安価に利用でき、また、このようなワークショップなどの際には期間限定の無償ライセンスキーが配布されるというサービスもある。当然、この一連のプレワークショップ及びワークショップでもこの無償ライセンスキーの恩恵にあずかったのは言うまでもない。このワークショップでも oXygen の基本的な使い方にある程度の時間をかけることとなったが、参加者のほとんどはすぐに習熟していた。なお、本ワークショップには参加できないもののプレワークショップのみ参加して TEI に関する知識を得ようという人も少なからずおり、近年の TEI への関心の高まりを感じさせた。

そのような準備期間を経た後、2012 年 9 月 15 日に開催された Elena Peirazzo 博士によるワークショップでは、TEI についての一通りの説明が行われた後、最新版の TEI ガイドラインに従い、手稿に書かれた様々な情報を XML マークアップによって構造的に記述するというので、ダーウィンの手稿 (*2) をはじめ、いくつかの手稿のデジタル画像と翻刻テキスト等が用意された。英文学やエジプト学から日本史に至るまで様々な分野の研究者が集い、それぞれに TEI の新しい考え方に基づくマークアップの手法に取り組んでいた。

ここまで触れてきたように、TEI の新しい考え方というのは、たとえば、縦横無尽に書かれた手稿の縦横無尽さを、デジタル画像上の座標情報を用いてきちんとすべて記述しているというものである。その考え方自体は以前から存在しており、たとえば、画像上のある箇所を切り出してその座標情報を取得する Web アプリケーションといったものもすでにデジタル・ヒューマニティーズの分野では開発済みである。（<http://mith.umd.edu/tile/>等）。ここで言う新しさというのは、こうした流れを受けて、それ自体を「論理構造を持つテキスト」と同格に扱おうとするという点である。TEI の歴史のなかでは大きな転換であり、今後、何らかの形でこれに関わらざるを得なくなる研究者がますます増えていくことだろうことが予想される。

しかし一方で、このワークショップの参加者から聞かれ、筆者も感じた課題として、記述するための操作がきわめて煩雑になってしまっているという点があった。このワークショップでは画像上の座標情報を取得するために Inkscape というフリーソフトウェアの XML エディタ機能を用い、そこで得た情報を oXygen に貼り付けるという仕方をとっていた。ほんの少しのデータであればそれでもなんとかかなるかもしれないが、研究者が日常的に取り組む場合、あるいはプロジェクトとして大規模にデータ構築しようという場合には、このようなやり方ではとても十分とは言えない。もちろん、このワークショップは TEI そのものについての理解を深めるためのワークショップであったため、敢えてやや複雑な、しかし、裏側がどうなっているかよくわかるような手法をとったのであり、たとえば上述の Web アプリケーションなどをうまく組み合わせることで、大規模な Web コラボレーションを効率的に遂行することはそれほど難しくはないだろう。

ただ、現状では、個人が容易に扱えるクライアントアプリケーションという観点からすると、この新しい考え方に基づくマークアップはまだかなり敷居が高い状態にある。さらにまた、座標情報に基づいてデジタル画像上に散在することになるテキストをいかにして利用者にとって利便性の高い形で提示するか、ということに関してもまだ大いに改良の余地がある。これは、「論理構造を持ったテキスト」を前提としたこれまでの TEI 符号化テキストにおける多様なアプリケーション群と比べると、なお明らかである。現状では、まだ、ほんのいくつかの表示するためのアプリケーションが試作段階であるに過ぎない。すなわち、この手法はまだ始まったばかりであり、特に良質なアプリケーションが様々に開発されることが国際的にも急務となっていると言える。すでにさらなる開発への取り組みは欧米各所で始まっているものの、日本の研究者・開発者が TEI の将来に貢献し得るところはこういったアプリケーションの面でさえもまだ十分に残っている。

さらに言うなら、ここでは、単に技術的な事柄だけが必要なのではなく、むしろ、人文学研究者が研究者として情報をどう扱おうとしているのかという見識をいかに適切に技術に反映し、いかにして有益な情報交換や知識の蓄積をおこなえるようにするか、という点が重要なのであり、人文学研究者が貢献し得る領域はきわめて大きい。そのような試みを通じてデジタル媒体上で人文学が一定の確かさをもって展開していくことができたなら、人文学の将来にもまた違った角度からの光が射してくることだろう。20年を過ぎてなお発展を続ける TEI はそのようなところにもつながり得る拠り所の一つであると感じた次第である。

(*1) <http://www.tei-c.org/>

(*2) <http://darwin-online.org.uk/manuscripts.html>

「JADH2012 報告」 : <http://www.jadh.org/jadh2012>

(Espen S. Ore : University of Oslo)

(日本語訳 : 滝浪佑紀・東京大学大学院情報学環、永崎研宣・人文情報学研究所)

第二回 JADH シンポジウムは、9月15日のプレカンファレンス・ワークショップと共に、9月16日と17日に、東京大学において開催された。会議に関しては、同時進行的に Twitter でコメント・報告され、収集されたつぶやきは以下の Topsy において見ることができる。

<http://topsy.com/s?order=date&q=%23jadh2012&window=a>

また、少なくとも一人の参加者 (Ryan Hunt) によって、報告書が公開されている。

<http://ivrytwr.com/2012/09/19/links-of-the-week-jadh-2012-edition/>

大阪大学における今年の会議は、新しく設立された JADH 主催の最初の会議であった。JADH とは、新しく国際組織 ADHO (Alliance of Digital Humanities Organizations - <http://www.adho.org/>) の傘下に入った、日本デジタル・ヒューマニティーズ学会 (Japanese Association for Digital Humanities) のことである。

3年前、東京大学と大阪大学は、欧州デジタル・ヒューマニティーズ学会 ALLC (The European Association for Digital Humanities) と共同で、テキスト分析とテキストエンコーディングをはじめとする、デジタル・ヒューマニティーズのさまざまな側面に関するワークショップを開催した。同年および翌年のワークショップでは、基本的なテキストエンコーディングと TEI に関する講義があった。これはさらに、大阪での JADH (OSDH) 2011 におけるプレカンファレンス・ワークショップで発展させられた。

今年の東京では、キングス・カレッジ・ロンドン (TEI 役員会の議長でもある) の Elena Pierazzo が終日にわたって、「TEI における文書に基づく生成的編集入門」——とはいえ、初心者のためのコースでは全くなく、テキストエンコーディングと TEI に関して確固たる基礎知識が必要だったが——と題されたプレカンファレンス・ワークショップを開いた。Pierazzo はグラフィック情報をエンコードするために、どのようにページにゾーンを設け、テキストをこれらのゾーンにリンクさせ、SVG (Scalable Vector Graphics - XML でベクタグラフィックスを保存するフォーマット) を使用するのかを示した。手書き原稿 (あるいはテキスト一般) をこのようにエンコードするにあたって必要とされるものの一つは、ユーザーがページ上の画像をクリックすると自動的に、位置座標と当該のテキストをリンク

してくれる GUI であり、できることなら oXygen のようなツールに組み込まれていることが望ましいものである。

土曜日のワークショップの後、イリノイ大学アーバナ・シャンペーン校の J. Stephen Downie が「HathiTrust リサーチセンター——大規模テキスト分析の最前線」と題して、プレカンファレンス・プレナリーレクチャーを行った。この講演でとりわけ興味深かった点は、このような方法で組織されたアーカイブは、いかにして著作権問題を切り抜けつつ、研究者が大規模なテキスト・コレクションで研究することを可能にするのかということにあった。議論は主として、英語で書かれた文書と英国および米国の著作権法に向けられていたが、この問題は原理的に、他の言語や地域に対しても応用可能である。

公式のシンポジウムは 9 月 16 日、日曜日の午前、開会式でのいくつかの講演とともに幕を開け、その後、文学研究と学術的共同作業に関する並行セッション、さらにそれに続くセッションへとわかれていった。極めて充実した投稿を集めたカンファレンスではしばしば起こるように、どのような関心でどういったセッションを辿っていくかを定めることは難しかった。こうした状況において、アブストラクト集はカンファレンスの重要な成果のひとつだと言える。アブストラクト集は PDF ファイルの形式で、オンライン上から入手可能である。

<http://www.jadh.org/JADH2012-Abstracts-Online.pdf>

様々な発表を通して、繰り返し現れてきたように思われる論点の一つは、テキストがある版やテキストアーカイブに収められるにあたって、それは単なるエンコードされたテキストを超える必要があるというものである。すなわち、コレクションあるいは版が、単純で小さなコレクション以上のものであるためには、何かしらのデータベース・システムが必要だということである。

○基調講演

カンファレンスにおける基調講演は、個々のフィールドや学術的トピックから、デジタル・ヒューマニティーズの全体的な議論へと、私たちの関心を広げてくれるものであった。HathiTrust 研究センターに関するプレカンファレンス・レクチャーの後、公式となるシンポジウムの初日は、Susan Schreibman の講演で締め括られた。それは、『デジタル・ヒューマニティーズ入門』と現在計画中の新しい版の構成をベースとした、デジタル・ヒューマニティーズの様々な側面についてのものだった。

2日目午前の Elena Pierazzo による基調講演「DH——馬鹿げているか、必要か？」は、様々な DH コースやカリキュラムの盛衰の例を示し、DH の学生のための前提条件について議論をしたという点において示唆に富んでいた。彼女は指摘したことの一つは、伝統的に DH コースは、人文科学における伝統的な分野での学術的なバックグラウンドを持っており、そこにデジタルツールを追加した人たちによって発展させられてきたということだった。しかし、学生——とりわけ学部生——には、それは期待できないのであり、これが、例えばキングズ・カレッジ・ロンドンでもはや学部レベルで DH のコースがないことの原因だろう。

バージニア大学の Bethany Nowviskie は、「失敗するには小さすぎる——バージニア大学図書館におけるスカラーズ・ラボ」と題し、スカラーズ・ラボの機能と組織について発表した。彼女の説明するところのラボは、多くの仕方で、フラットで非階層的な構造を持っており、開かれた共同作業に基づいている。聴衆の多くは、Bethany Nowviskie を上司として、そこで働きたいと思っただろう。実際、そこは幸せな職場であるばかりでなく、多くのものが集められているという意味で、生産的な場でもあるのだ。

キングズ・カレッジの Harold Short は、閉会式で基調講演を行い、関連する様々な状況のなかでの DH 関連組織の国際的な構造についてのスナップショットを示した。

○並行セッション

シンポジウムの大部分で2つのセッションが同時並行していたことを考えると、私が多くの興味深い発表を聞きのがしたことは明らかである。以下は、いくつかのセッションの簡単なレポートであり、言及されていないからといって、それが重要でない、あるいは興味深くないというわけではない。

○学術的共同作業

学術的共同作業に関する日曜日午前のセッションは、ウィキブックス http://en.wikibooks.org/wiki/The_Devonshire_Manuscript で公開されている『Devonshire 写本』（BL Add 17, 492）を例として使用した、ビクトリア大学の Ray Siemens の発表から始まった。Siemens は、これは〔この事例にとって〕非常に適した写本であると説明した。というのも、この写本自体が男性と女性による共同作業の成果だからである。採り上げられたプラクティカルな問題の一つとしては、若い研究者がいかにして共同作業から必要な学術的メリットを得るかという問いがあった。人文学では、医学や科学一般の場合に比べて、この伝統が弱いというのである。

Devonshire 写本の版は、ウィキブック版に収められる前に、研究者が個々にふさわしい出版物で自らの成果を発表するという考えに基づいて進められている。一部の雑誌は執筆者に対して自らの論文をオープンアクセスのウェブ・サイトに載せて公開することを許可していないという問題があるかもしれないが、しかし、いずれ解決されるだろう。

英語ルネッサンスの作品から敦煌チベット語のテキストに移り、東京外国語大学の松田訓典と星泉は、ウィキのシステム (PukiWiki) から始まった仕事を発表した。しかし、より高度に整形されたデータの必要性が感じられたため、現在では、データは PostgreSQL リレーショナル・データベースに保存されているとのことである。テキストは <http://otdo.aa.tufs.ac.jp/> で利用可能であり、KWIC ツールを用いるのと同様に、テキストのリストおよび書誌データの両方を通してエントリーできる。ウェブベースの出版物に加えて、テキストはモノグラフ・シリーズとしても出版されている。Devonshire 写本の版と同様に、モノグラフ・シリーズとして出版される一つの理由は、版を出版した学者が必要な学術的メリットを得られることを確実にすることにある。

花園大学の後藤真は、正倉院文書のウェブ版について発表した。このコレクションでは、物理的な順序 (巻物や巻物の字面) は、必ずしもテキストの内容を直線的に表してはいない。以前に出版された、これらのテキストの紙ベースの選集版は、それぞれが独自の連続性を確立させていた。テキストは XML でエンコードされたテキストデータとして、データベース (SOMODA-正倉院文書データベース) に保存されている。これによって、[ある特定の仕方での] 選集、前の版に倣ったテキスト、見るにあたっての物理的順序 (巻物) など様々な仕方、テキストを抽出することができる。

セッションは、デジタル版とは保存された——望むらくはエンコードされた——テキストとそれらを様々な視野のもとでフィルタリングするツールの組み合わせであると主張した、京都大学の Christian Wittern の発表によって締め括られた。これはすでに、ArtSpace Tokyo guide (<http://read.artspacetokyo.com/>) などの商業的製品に使われている。そして、私たちはオスロでイブセンの電子版で同じようなことをしており、このことに驚いた。私たちは以下のページで XML 文書、PDF 版および電子ブック版を公開している。

(http://www.ibsen.uio.no/DRVIT_Du%7CDuht.xhtml)

○時間と空間

午後のセッションの最初の二つの論文は、歴史・考古学的分析における統計ツールと GIS の利用についてのものだった。両プロジェクトとも、フリーで使えるソフトウェアを使用

している。同志社大学の松森智彦は、遺跡の地理的クラスター分析のためのドロネー分割の使用法について発表した。ドロネー分割を使用する理由については異論があり、発表者はドロネー分割によって、すべての点はそれにもっとも近い点と結ばれると主張したが、さらなる議論の余地があるだろう。同ソフトは GNU のライセンス下で、<http://tmats.net/pine3/> において利用可能である。京都大学の清野陽一は、GIS システムのオープンソース GRASS (<http://grass.fbk.eu/>) とティーセン多角形を使用した。このシステムは、いわゆる「律令国家」の行政の中心地を分析するために使用され、「郡」という行政単位内でもっとも移動に便利であるという論理的だと思われる場所に対して、実際に中心地はどこに位置していたかを見ることを可能にする。清野は続いて、実際の中心地の場所は必ずしもティーセン多角形に合っていないことを示し、中心地の場所を説明するためには、純粋に地理要因以外のファクターも考慮に入れなければならないと主張した。

梶山女学園大学の杉藤重信と共に、セッションは地理的考察から、時間と家族関係に関する議論へと移行した。アライアンス・プロジェクトは、関係性や家系図データを蓄積し、表示するためのデータベースおよびモデルを開発している。杉戸の研究はオーストラリアのノーザンテリトリー州アーネムランドにおける調査に基づいているが、アライアンス・プロジェクトとそのデータベース・モデルはより一般的に適用可能だろう。というのも、このモデルは家族関係を構成するものに対し、いかなる予め想定された形式も押し付けないからである。ソフトウェアは、カンファレンスのブログ <http://study.hs.sugiyama-u.ac.jp/e/> で利用可能である。

○データベースの使用法

月曜午前の Elena Pierazzo による基調講演（上記参照）の後、私が出席した最初のセッションは、ニュージーランドの災害の考証、ネパールの写本コレクションから、オーストラリア先住民の言語の目録化までを扱った。

ニュージーランド・カンタベリー大学の James Smithies は、2010 年と 2011 年に同国南島の最大の都市であるクライストチャーチを襲った地震を記録するために開設されたデジタル・アーカイブについて発表した。このアーカイブは、単一の大きなデータベースというよりも、複数のアーカイブとデータベースを組み合わせたものである。この多層的な構造により、ブログをはじめとする私的な情報源からも記録を集めやすいうになっている。そして、システムを中心部で、個々に独立して保存された情報の断片をリンクすることができる。Smithies によれば、オープンソースのソフトウェアを使用したこのモデルは、他の似たようなアーカイブの構築にとっても、利用可能なパッケージとして考えることができるということである。

ニュージーランドからネパール・ドイツに移り、ドイツ・ハンブルグ大学の Kengo Harimoto は、ネパールとドイツの写本目録プロジェクトにおける写本カタログについて発表した。このプロジェクトは、長い歴史を持ち、1970 年から 2002 年にかけて遂行されたネパール・ドイツの写本保存プロジェクトを引き継いでいる。この長い歴史とマイクロフィルムやワープロのテキスト文書に基づくカタログの初期の版とともに、現在の仕事の重点は、カタログにとって有意義なウェブ・インターフェイスを提示できるメタ構造を構築することに置かれている。もしデータがリレーショナル・データベースに保存されたとしたら、多対多の関係とは實際上どのようなものか、という問題をはじめとする、克服されるべき問題のいくつかは解決されるだろう。とはいえ、現在のところ、主な作業は、既存のデジタルソースをウェブ上で利用可能とすることに集中しているとのことである。

オーストラリア・アボリジニ・トレス海峡諸島研究所 (AITSIS) の Kazuko Obata は、オーストラリアの先住民族の言語データベース (AUSTLANG) について発表した。ある意味において、ここには、系統図データベースに関する前日の杉藤重信の発表と類似点があるだろう。オーストラリアの言語にとって、言語とは何であるのか、それが何と呼ばれているかについては必ずしも明確ではない。政治的な理由のために、あるグループの人たちの言語はその隣のグループ (隣の島など) の言語と異なっていると主張したほうが便利であるだろう。そして、植民地時代と現代のオーストラリアの入植を通じて、入植者が出会った先住民族の言語は、多くの異なった名を与えられている。AUSTLANG データベースは、さまざまな言語とそれらを記述するために使用された名前に関するユニークな識別子を保持していると思われる。データベースは、Google マップとリンクし、地図上にクエリの結果を表示することができるようになっている。

○データベース構築

9 月 17 日の月曜日、私が最後に参加した並行セッションは、1675 年あたりにおける英国の学校ドラマのコレクションを含んだ、手稿のオンライン版に関する発表から始まった。作業はフィンランドのオウル大学で行われ、Lisa Lena Opas-Hänninen を中心とする学者のグループによって発表された。この版は一方において、学術校訂版である。他方、そのコンテンツである戯曲は、おそらく舞台制作者にとっても読むことが可能な形式でダウンロードできなければならない。いくつかの点で、これは前日の Christian Wittern による発表に私たちを連れ戻してくれる (上記参照)。私が重要であると信じている一つの瞬間は、Ilkka Juuso (オウル大) が「ツールの再利用は話題になっているが、皆が各々独自に発明している」と言った時である。ヘンリック・イプセン作品のオンライン版が最終段階

にあり、多くの人と意見交換をしていると、私はこの言明に多くの真実があることを認めざるをえない。

京都大学の原正一郎と株式会社ナレッジ・シナジーの内藤求は、我々を、王政復古期の英国学校ドラマから漫画という現代日本（およびインターナショナル）文学へと連れて行った。漫画に関するマルチメディア・データベースをめぐっては、多くのアプローチを同時にしなければならないというのは理に適っているだろう。システムは、書誌メタデータ（ダブリン・コアを使用して）、フレーム、絵、テキスト間の物語上のつながりを定義する構造に関するメタデータ、映画やテレビの場合と同じように定義されるシーンに関するメタデータを持っている。このシステムはトピックマップを使って組織され、tolog をクエリ言語として、トピックマップの検索が可能となっている。またこのシステムでは、異なる言語で同じシーン——あるいはフレーム——を保存することが可能となっている。しかし、トピックマップの構築、あるいはシーンやフレームなどのユニットのエンコーディングは手作業で行われているようなので、大量のものを記録することはどれほど実現可能だろうかという疑問も感じた。とはいえ、骨格の登録は自動的に行うことができるだろうし、おそらくその後、部分的に選択された箇所を深く分析することもできるだろう。

鶴見大学の矢野一志はこのセッションを、記述言語学を共有するためのコーパスに関する計画の発表で締め括った。矢野は配布資料としてフルペーパー版を配布し、シンポジウムのセッションでは、この中のいくつかの項目を議論した。いくつかの問題は、同日の早い時間に Kazuko Obata によって導入された議論に関連していたが（上記参照）、ここでの問題は、AUSTLANG の場合とは異なり、メタデータであるというよりも、言語データそのものであった。矢野は、テキストデータの異なったシステム、モデル、表現（エンコーディング）に由来するコーパス共有に関する問題を列挙した。理想を言えば、記述言語学者は共通システムの利用に一本化すべきだろう。

○他の発表セッション

上記のように、他の並行セッションに出たため、出席することのできなかった発表が少なからずあった。しかし、私は発表者の何人かとは話をしたし、アブストラクトも読んだので、聴き逃したものについてもいくらかは知っている。「文学の分析」と題された1日目の最初のセッションでは、文学作品の分析についての発表ばかりでなく、著者とその関係性のネットワーク解析についての発表もなされた。このアプローチは、日本語以外の言語地域においても興味深い結果を提示してくれるだろう。例えば、比較的最近のノルウェー人作家 Ragnhild Joelsen の伝記に関して言えば、彼女は通常主張されているよりもずっと、当時のノルウェーの文学コミュニティに深く関わっていたと示すことができるとい

うように。JADH2012のこの発表では、そのようなネットワークを記述し、分析するためのツールに関して議論された。このセッションでは、DHツールの聖書への適用を論じた発表がさらに二つあり、最後の発表はより最近の文学に戻って、テキストマイニングを論じた。

昼食後、OCRに必要な、ドキュメント・ページの解析についての発表があった。これは同じ研究者たちによるポスターセッションでの発表にも関係していた。セッションは、日本の歴史的なテキストのn-gram解析の発表、さらには日本の歴史的なテキストのための写本システムの発表と続いた。

2日目に行われたセッションの一つは、人間のインタラクションを扱った。DHに度々現れる問題の一つは、グループのメンバーが異なったタイムゾーンに広がり、異なった文化的背景を持っているという国際的な仕事に関わっている。また、別の発表は、ある文化的背景に基づきながら、TED (Technology Entertainment and Design) をDH内で使い得るプレゼンの例として使用したものだった。また、ストーリーや文化を紹介する別の方法は、稲葉光行によって発表された(彼はポスター発表で、ヴァーチャル世界のデモも行った)。彼は日本の伝統文化の例を例に、セカンドライフの世界を構築した。しかし、現代日本文化のいくつかの側面は、この枠組みではちょっと問題があった。というのは、アニメ映画に登場するトトロ(スタジオジブリ)のような外見を持つアバターは、神社でのお浄めのような行為をするのには、ちょっと体型がふさわしくなかったからである。

昼食後、ハミルトン大学の学者のグループによる教育に関するパネルがあった。ある意味において、これは、同日午前のElena Pierazzoによる基調講演と並んで、DHと学部学生に関するもう一つの見解だと見なすことができるだろう。ハミルトン大学では、学部生も巻き込んだ学際的プロジェクトができるようになっている。

○ポスター

ポスターセッションは、文学的・言語学的分析から、古代エジプトのヒエログリフ文字への古文書学的アプローチやゲーテンベルク聖書(慶應ゲーテンベルク聖書)のインタラクティブディスプレイまで、幅広いテーマにまたがっていた。文体分析の事例は、村上春樹の小説における言語の通時的研究(とりわけ)から、多くのフォロワーを持つ投稿者のツイートが投稿され、リツイートされる度合いを基とした、つぶやきの文体分析(Twitterのメッセージ)まで多岐に渡っていた。発表されたプロジェクトの中には、コーパス構築に関するものもあった。テキストコーパスは重要であり、しばしば長期にわたって永続的

な価値を持っている。今回は、テキスト選択や日本語の通時的コーパスのためのプロトタイプが示された。

○企画者への謝辞

最後に、この重要かつ興味深いシンポジウムにお招きいただき、企画者の皆様に感謝をしたいと思います！それから、月曜日の昼休みに素晴らしい琴の演奏を聞かせてくれたミュージシャンにも感謝します。