

特定領域研究 「情報爆発 (Info-plosion)」 への新展開

喜連川優 (東京大学生産技術研究所)

文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」(略称:「情報爆発IT基盤」, 領域代表者:喜連川優)に関し, 特定領域研究「情報学」との関係, 研究構想, プロジェクト体制の概略などを紹介する。

情報学特定から情報爆発特定へ

情報爆発特定は, 2004年に申請, 2005年に採択, そして2006年に計画研究ならびに公募研究を含むプロジェクト全体が始動した。本年度は, 5年間の研究期間の2年目にあたる。その生い立ちを明らかにすべく, 安西慶應義塾長が領域代表を務められた情報学特定とのつながりについて振り返る。情報学特定は, 文部科学省の科学技術・学術審議会の情報学部会における議論を踏まえ, 情報分野全般の研究強化, 学問分野としての情報学の確立を目指したトップダウン的施策として発足したものであり, 2001年度から2005年度まで実施された。その後期, 2004年になると, 情報学特定の中で, 次の研究展開に関して種々の議論がなされるようになり, 情報学特定に適用された特定研究のスキームはすでに廃止されていたが, 5年間にわたる活動の成果を基に次の段階の課題を具体的に設定するとともに, 情報分野の実質的な飛躍とプレゼンスの向上を目指し, 研究をさらに発展させたいという議論がなされた。情報学特定の執行部では, 多面的な検討を行い, 情報関連研究者がこれから取り組むべき課題として「情報爆発」が最重要であるとの認識に至り, 当該課題を軸に次の特定領域研究を発足させる活動に結実した。2004年初夏以降, 松山(京大), 松岡(東工大), 須藤(東大), 安達(NII), 喜連川(東大)を中心に, さらに多くの先生方の協力を頂戴しながら, 詳細に立案が進められ, 秋に申請書を提出した後も, 継続的に議論が進められた。幸い, 情報学特定の最終年度と重なる2005年に情報爆発特定が採択され, この年には情報爆発特定の総括班が並行して活動し, 2006年

からの実質的研究活動の立ち上げに向けて種々の具体的調整を進めるとともに, 研究公募を行った。このようにして, 情報学特定から情報爆発特定へと円滑な研究体制の移行が実現された。

現象としての情報爆発

最近の情報量の爆発的増大に関する最初のレポートは, 米国カリフォルニア大学バークレイ校により実施されたHow much informationプロジェクトによるものであり, 人類が創出する情報量について詳細な調査がなされた^{1), 2)}。加えて, 最近, 別の調査として, 2007年3月, IDCにより2006年の情報創出量ならびに2010年の予測に関するレポートが提出された³⁾。当該調査によると2010年の情報生産量は約1ゼットバイト(10^{21})に達すると予想されている。とりわけ21世紀を迎えるころより, 人類の創出する情報量は爆発的に増大しており, この最近の現象を21世紀の情報研究者の大きな課題と捉え, 当該課題への挑戦を本特定のテーマとした。大量情報の生産には少なからずIT関連研究・開発者がかかわってきたことは事実であり, であるならば, 情報関連研究者は, 一般の人々が大量情報に翻弄されることなくそれを活用し豊かな生活を送ることができるよう, 新しい技術開発を目指すべきと考えた次第である。本特定は, 現象としての情報爆発を受け止め, 新たに発生しつつある多様な課題を明確化し, 情報関連の種々の専門分野の研究者が集い, 各々のエクスペリエンスを融合し挑戦するという研究形態を採用した。このような機動的な研究体制こそが情報爆発時代においてはきわめて重要であると考えた次第である。

研究組織体制

本特定領域研究の構成と体制の概観を図-1と, 図-2に示す。全体は3つの技術系の研究柱(A01~03)と1つの社会系研究柱(B01)から構成されるが, 第1の柱A01は, 情報管理・融合・活用基盤の研究を行い, 田中(京大), 安達(NII), 下條(阪大), 黒橋(京大), 喜連川が計画班として活動する。本特定を模索していた2004年当時を振り返ると, YouTubeやSecond Lifeの現在のような隆盛は予想だにしていなかったものの, Webに代表されるインターネットコンテンツ, 電子メールやデスクトップに代表される個のコンテンツが大きく情報分野を変えていくドライビングフォースとなってゆくことは強く予見された。情報爆発時代においては, サーチの高度化は地道な研究ではあるが避けて通れないきわめて重要な研究課題であり, また, ランキング操作が社会

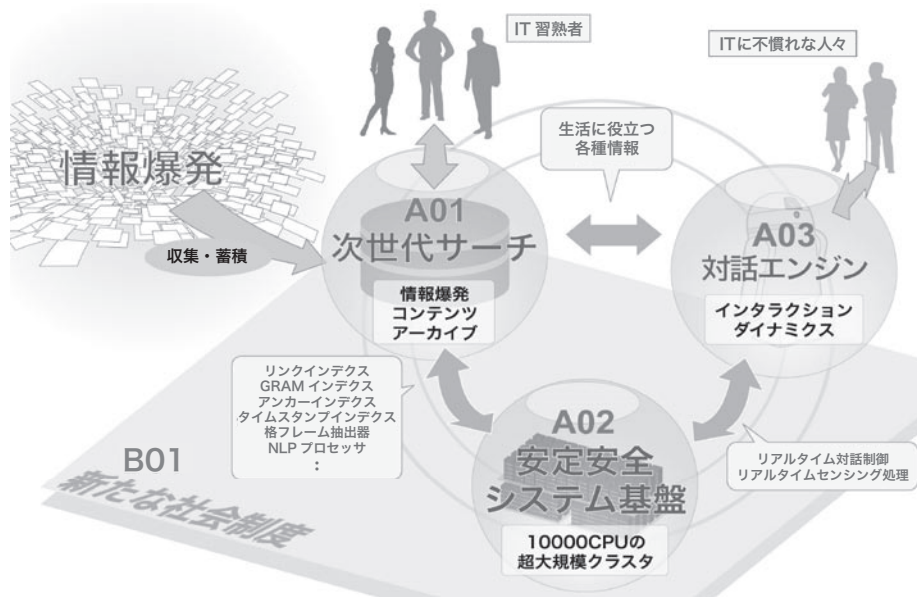


図-1 情報爆発プロジェクト概観

的に問題視される中で、多様なランク付けを可能とする基盤技術を国として研究することこそが本質であると考へA01を企画した。Webのみならずe-Scienceにおいても、情報爆発が生じており、膨大なシミュレーション結果情報や多様な実験機器からのデータに押し潰されることなく、膨大な情報空間から有意な現象、知見を見出すデータドリブンサイエンスの流れは、インターネットサーチと共通する部分も少なくなく、しっかりとした情報管理・検索・操作基盤の構築は科学者にとっての生命線となると考えた次第である。加えて、センサ技術の著しい進歩により、爆発する膨大なセンサ情報を利用したソリューション、ストリームデータ管理もその範疇である。第2の柱A02は安全・安心ITシステム基盤の研究であり、松岡(東工大)、柴山(東工大)、近山(東大)、中島(早稲田)が計画班を務める。CERNにおける高エネルギー物理実験のための計算機環境のCPU数は4万、EGEEなる欧州グリッド全体では10万である。また、Googleを代表とする新しいインターネットサービスのためのインフラは、数十万台ともいわれる巨大サーバを抱え、2004年時点ですでに、従来とは異なる意味でのスーパーコンピュータと見なせる存在となっており、きわめてエキサイティングな研究対象となりつつあった。近年の大学におけるコンピュータサイエンスのシステム研究は、携帯環境やゲーム機などのパーソナル環境の高度化に目が向けられがちであったが、超大規模システムの構築とそのリジリアンシー向上技術に関する研究が大学の重要な研究対象領域に戻ってきたといえる。安定運用には、ソフトウェア挙動モニタリングによる問題箇所同定が有効な策となると想定されるが、上述の物理的なセンサ同様、ソフトウェアセンサが生み出す爆発的な情報とそこからの異常状態の検出が肝となる。エス

トニアに見られた国家レベルサイバー攻撃が現実のものとなる今日、大量情報監視によるセキュアな大規模システム環境の実現についての研究も不可欠である。第3の柱A03はヒューマンコミュニケーション基盤に関する研究であり、松山(京大)、西田(京大)、国吉(東大)が計画班を務める。爆発する情報空間から真に必要な情報を獲得する作業は、いまだ強いスキルが必要とされており、現行のツールは一般の人々にとって自在に利用できるものとはいえず、人間に優しいインタフェースの構築においても、また、今後必須となるロボットとのコミュニケーション局面を想定しても、ヒューマンコミュニケーションの基礎研究は重要であり、本柱が当該研究を司る。人が行う多様なコミュニケーションに関する精密なセンシングから発生する爆発的な情報の分析を通じて、ノンバーバル、サブコンシャスなコミュニケーション機構の本質を見極めたい。第4の柱B01は須藤(東大)が計画班を担当し、社会制度面、情報ガバナンスの観点から研究を進め、A01～03の研究が技術的成果にとどまることなくそれらの社会への積極的な展開を調整する。とりわけ、研究者は法制度には疎い場合が多く、実社会における実証実験における種々の制約に関する社会科学研究者との連携は不可欠といえよう。

最近の斬新な情報サービスの多くは単一の新技術によって構築されるのではなく、むしろ、多様な先端的情報分野の研究の融合により創出されている点が大きな特徴であり、上述の4つの柱を機動的に連携可能な研究体制を組むことが情報爆発時代の研究においてきわめて重要であると、本特定を立案する時点で種々議論を重ねた次第である。情報爆発なる研究課題は計画研究だけでは到底カバーできず、本特定には多数の公募研究の参画がある。2006年度単年度には62件、引き続き

A01 情報爆発時代における情報管理・融合・活用基盤

- 「情報爆発時代におけるサイバー空間情報定量評価基盤の構築」喜連川優（東大）
- 「情報爆発時代の情報検索基盤技術」安達 淳（NII）
- 「情報爆発に対応するコンテンツ融合と操作環境融合に関する研究」田中克己（京大）
- 「コピキタネットワークコンテンツに対する管理・統合基盤に関する研究」下條真司（阪大）
- 「構造的言語処理による情報検索基盤技術の構築」黒橋禎夫（京大）

A02 情報爆発時代における安全・安心ITシステム基盤

- 「情報爆発に対する高度にスケーラブルな高性能自律構成実行基盤」松岡 聡（東工大）
- 「情報爆発に対応する高度にスケーラブルでセキュアなソフトウェア構成・更新方式」柴山悦哉（東工大）
- 「情報爆発に対応する高度にスケーラブルなソフトウェア構成基盤」近山 隆（東大）
- 「情報爆発に対応する高度にスケーラブルなモニタリングアーキテクチャ」中島達夫（早稲田大）

A03 情報爆発時代におけるヒューマンコミュニケーション基盤

- 「相互適応可能な実世界インタラクションのための計算モデル・システムの構築」松山隆司（京大）
- 「実世界インタラクションの分析・支援・コンテンツ化」西田豊明（京大）
- 「セマンティック・ヒューマノイドインタラクション」國吉康夫（東大）

B01 情報爆発時代における知識社会形成ガバナンス

- 「知識社会経済システムの競争的発展とそのガバナンスに関する研究」須藤 修（東大）

図-2 情報爆発プロジェクト研究体制構成

2007, 2008 年度の 2 年間の研究には 74 件の公募が採択され、全体予算の 40% 以上を占める。また、2007 年度の 74 件の中で、2006 年度からの継続研究は約半数であり、流動性はきわめて高い。公募研究の競争率はきわめて高く、採択に漏れた興味深い研究がきわめて多数あることから、すべての集会は基本的にオープンとし、誰でも、ほぼ全分野のカテイングエッジなトップカンファレンス情報は情報爆発集会で入手可能であるように工夫をしている。

わくわく共有研究プラットフォーム

日本の情報分野の研究は国際的に見て元気がないと他分野の先生方からご指摘を受けることが少なからずある。筆者は、情報分野の中でもほんの限られた分野でしか仕事をしてきていないための外れであるやもしれないが、国際会議のプログラム委員を数多く経験して、米国の論文の強さの根源は美しいシナリオと、迫力のある実験結果にあると感ずることが多々あった。コンピュータサイエンスは、ビッグサイエンスとなりつつあり、有無を言わせない力強い実験結果を出せる環境の整備がきわめて重要である。トイアプリによる評価の貧弱さや、小規模な実験環境から派生する説得力の弱さが、日本が主導する研究の潮流を作る上での障害と感ずることがある。もちろん、斬新なアイデアが最も重要であることは論をまたないが、研究基盤環境も大変重要である。今回の情報爆発特定では、各研究者に予算を細分化し配分するのではなく、一人では購入もできないし、維持もできないような環境に大きく予算を配分し、それを皆で共有する方式を採用している。共有プラットフォームは若手の研究者が率先して引っ張っていってくれるような「わくわく」したものでなくてはならない。黒橋、新里（京大）らの深い自然言語処理による精度の飛躍的改善を

狙ったオープンソース・オープンデータサーチエンジン Tsubaki, 田浦（東大）らの分散ソフトウェア実験環境 InTrigger（現在数機関で約 500CPU 規模であるが、最終的には全国十数機関・数千 CPU 規模を予定）、角、坊濃、西田（京大）らの多人数コラボレーション解明のための高度センサ実験室 IMADE, 中村, 大島, 田中（京大）らの自分のアイデアを簡単に実現可能なサーチエンジンツールキット SlothLib, 原（阪大）らのセンサネット実験基盤等いずれも、最先端の「わくわく実験環境」である。本年に入り初年度の機器導入が完了し、ようやく少しずつ環境が整いつつある。特定領域に属する研究者に閉じることなく、これらを基盤として力強い論文が数多く創出されることを強く期待している次第である。当該基盤は特定研究に設けられた支援班なる枠組みを利用して実現されており、その詳細は <http://itkaken.ex.nii.ac.jp/i-explosion/> をご覧いただきたい。

謝辞 本特定を立ち上げ、そして体制を維持・管理するにあたってはきわめて多くの方々にご支援を賜った。ここに深く御礼を申し上げたい。国立情報学研究所に事務局として特定の運営を支えていただいている。また、省を越えた NICT との連携、東工大 TSUBAME の利用などに関しても、多くのお力添えを頂戴しており、感謝申し上げる次第である。

参考文献

- 1) <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>
- 2) <http://www.horison.com>
- 3) http://www.emc.com/about/destination/digital_universe/
(平成 19 年 7 月 18 日受付)

● 喜連川優（正会員）

kitsure@tkl.iis.u-tokyo.ac.jp

1983 年東京大学工学系研究科情報工学専攻博士課程修了、工学博士。同年、同大生産技術研究所講師、現在、同教授、戦略情報融合国際研究センター長。データベース工学、Web マイニングの研究に従事。2008 年本会全国大会プログラム委員長、本会フェロー。