

平成25年度 次世代IT基盤構築のための研究開発
「ビッグデータ利活用のためのシステム研究等」採択課題一覧

資料5
情報科学技術委員会
(第82回) H25.8.7

事業名	課題名	課題概要	事業代表者
データ連携技術等に 係るフィージビリティ スタディ及び予備研 究	ビッグデータ利活用 のためのデータ連 携技術に関する フィージビリティスタ ディ及び予備研究	<p>超大規模データ異種情報源の利活用のためには、データの生成からアプリケーション・利用者による利活用に渡るデータのライフサイクル全体をカバーする技術が必要となる。本事業では、以下の4課題を設定することでこれらの技術を網羅したフィージビリティスタディ及び予備研究を実施する。</p> <p>各種センシングデバイスの発展やソーシャルセンシングへの関心の高まり等、センシングデータの重要性が高まっている。【課題1 異種センシングデータ高度収集基盤】においては、実世界及びサイバースペースにおいて生成されるデータを収集するための基盤技術を対象とする。ハードセンサ、ソフトセンサを統一的に扱うための異種センシングデータ処理基盤について検討し、様々なイベントやトレンドをリアルタイムに取得、発信することを可能とする。</p> <p>センシングデータ収集基盤によって取得された各種ストリームデータを含む大規模異種データに対して、多様な情報源との連携・統合を行った上で分析処理を行うことが求められる。現状ではこれらの処理はアプリケーションレベルで個別に対応がなされているが、本格的な利用に際してはミドルウェアレベルでのサポートが必要である。このため、【課題2 異種データ連携・統合基盤】においては、異種データ連携・統合のためのストリーム処理基盤、異種データの高度連携基盤技術、データ連携におけるプライバシー等を対象とする。</p> <p>センシングデータ基盤によって収集された異種ストリームデータや、異種データ連携・統合基盤によって生成されるデータは、様々な分析目的や情報の保管・記録等のためにアーカイブする必要がある。情報の保管は、セキュリティや説明責任の観点からも多くの場面で必要とされる。このため、【課題3 超大規模データ格納管理基盤】において、高帯域なストリームにも対応可能な超大規模データ格納管理基盤や情報可視化を対象とする。</p> <p>既存のデータマイニング手法は、基本的に、特定の目的のために収集されたデータに対して、全体的な処理を行うことで新たな知見を発見することを目指していた。一方、超大規模なリアルタイムデータを扱う際には、これらの前提を置くことは困難である。このため、【課題4 超大規模データ分析・リアルタイムマイニングアルゴリズム】では、超大規模データに対する新たなデータマイニングの計算パラダイムに関する検討を行う。</p>	筑波大学 システム情報系 北川 博之 教授

事業名	課題名	課題概要	事業代表者
アカデミッククラウド 環境構築に係るシス テム研究	コミュニティで紡ぐ次 世代大学ICT 環境 としてのアカデミック クラウド	<p>本事業では、大学において ICT に関わる様々な立場の人々が集うコミュニティに依拠しながら全国の高等教育機関 783 校を対象にした調査を行い、研究・教育・管理運営等に関わるデータの量・分布を明らかにする。そして、調査結果をもとに、アカデミッククラウドのあるべき方向性と実現手段を検討し、標準仕様を策定する。</p> <p>現在、大学におけるクラウド化は、集約化・共有化によるサーバ管理コストの圧縮を主眼として大きな潮流になりつつある。大学間連携により地域でのICT基盤の共有を実現する場合も、集約化・共有化により効率化を目指す側面が強い。しかしながら、教育研究の質的向上や事務・大学経営の高度化のためには、各種サーバ群の集約化・共有化を通じて形成される巨大なデータの利活用を可能とする革新的な枠組みが必要である。これこそが、アカデミッククラウドが備えるべき要件であり、ビッグデータ時代の大学間連携の姿である。事業の実施に際しては、九州大学を中核とした事業推進グループと、研究・教育・コンテンツ・事務・大学経営・システムアーキテクチャ・セキュリティ・プライバシー・ネットワーク・認証連携の10分野の研究者で構成されるシステム研究グループを構成し、調査内容作成フェーズ・調査実施フェーズ・調査結果分析フェーズを通じてコミュニティエンゲージメントを明確にしながらか推進する。中核となるコミュニティは、各大学のベストプラクティスの共有や各種課題への共同対応を目指して設立された大学 ICT 推進協議会とする。これにより、国公立であるは研究大学・教育大学など、様々な形態・機能の大学の中で生じる多様なニーズを的確に掴(つか)み、成果の適用可能性を確保する。このようなコミュニティとの密接な連携は、事業終了後のアカデミッククラウド実現フェーズを見据えたものでもある。クラウドサービスを開発・運用していくためには、小さな成功を積み重ねながら持続的かつ柔軟に対応する必要がある。本事業を支えるコミュニティは、事業成果の波及・展開においても威力を発揮する。</p>	九州大学 附属図書館付設教 材開発センター 岡田 義広 教授

事業名	課題名	課題概要	事業代表者
	データサイエンティスト育成ネットワークの形成	ビッグデータをイノベーションに結びつける人材、いわゆるデータサイエンティストとは、1)データベース・分散処理などの情報技術、2)統計・機械学習などのデータ分析手法、それに3)ビジネスやサイエンスなどの適用分野の3領域をまたがる学際的専門的知識を有し、かつそれらの知識を現場において実践できる人材である。このような人材の育成のため、情報科学、統計・機械学習の研究教育機関、それにデータを持つ現場としての研究機関・企業を広くネットワークし、認知度の向上、人材の流動化、ベスト・プラクティスの調査と共有、教育コンテンツの開発、それに国際連携を行う。本事業終了後も、人材育成が継続的に行われるよう、人材の出口戦略に特に留意する。	統計数理研究所 樋口 知之 所長
ビッグデータ利活用によるイノベーション人材育成ネットワークの形成	スキルと実践を重視したビッグデータ・イノベーション人材育成プログラム	本教育プログラムは、第一必修科目として新規開講する「ビッグデータ技術演習」によりビッグデータスキルを体得させ、第二必修科目として新規開講する「ビッグデータ実践」により、その分析結果の関係機関による評価を受けながら成果をまとめ、実践経験を積む。また、3種類の選択必修科目群「ビッグデータ技術論」、「ビッグデータシステム論」、「ビッグデータ活用論」の各科目群から1科目以上履修させ、スキルと実践経験と関連知識を修得させながら、ビッグデータに関連する修士論文(研究)を開始させ、M1修了時の第1回研究発表会、M2途中の第2回研究発表会、M2修了時の第3回研究発表会の審査にすべて合格すれば、本教育プログラム修了認定書を取得できるものとする。 また、データ提供機関は、現在、(生命と生活分野)納豆菌ゲノムデータ(国立遺伝学研究所)、就職関連情報とライフスタイル(リクルート)、(グリーン分野)運転行動プローブ(HQL:人間生活工学研究センター)、(公共・安全・防災分野)行政データ(LOD (Linked OpenData)イニシアチブ)、企業の安全文化醸成度調査アンケート(東京地下鉄、京阪電鉄、東武鉄道、札幌市交通局、西部石油)、防災位置情報(IMESコンソーシアム)、(ビジネス)ゴルフ商品購入履歴データ(ゴルフダイジェスト・オンライン)を想定し、数値、時系列、画像、動画、日本語テキストを含む、Volume(大容量)、Variety(多様性)、Velocity(多頻度)の3Vのビッグデータである。	慶應義塾大学 理工学部 山口 高平 教授