

CREST・さきがけ
「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」
複合領域事後評価報告書

総合所見

本研究領域は、“情報科学・数理科学とビッグデータの利活用による社会的インパクトを生むアプリケーション分野との協働”と“アプリケーション分野での課題解決を通じたビッグデータからの新たな知見の発見”を目標に取り組みました。採択にあたっては4つの技術分野に法制度を加えた5つを重点的な取り組み分野として設定し、高い独創性・新規性と提案者の明確な目的意識を重視して、ビッグデータ解析の基礎となる数理基礎技術、言語や画像のメディア処理技術、セキュリティやシステムアーキテクチャ、さらには法制度の分野の研究課題が選考された。さらに、さきがけではビッグデータの取得や提供を目指す「データさきがけ」の設定により、複合領域全体のビッグデータ統合解析の研究推進に貢献できたことは特筆に値する。

研究マネジメントでは、国内のみならず欧米の研究者を海外アドバイザーとして迎え、NSFの関心を集めながら、EUを含めた国際連携の活発な活動を行った。また、“アカデミア・企業等が様々な分野のビッグデータを統合分析できる共通基盤を利活用して、研究開発や実用化を推進する環境作り”の実現に向け、大学や国立研究機関に加えて企業およびアプリケーション分野の研究者を含め、ビッグデータの社会利用とその先のイノベーションへの発展の可能性を示した。総合的に“分野を超えたビッグデータ利活用により新たな知識や洞察を得るための革新的な情報技術及びそれらを支える数理的手法の創出・高度化・体系化”を大変強力で推進した。

研究成果は、CREST・さきがけ合計で、発表論文1,762報*（うち国際論文1,616報*）、特許出願88件（うち国際出願48件）が発表されている。特出すべきは、「ERAでAランク」と「QualisでA1ランク」を同時に満たすトップカンファレンスで計198件の論文発表があり、さらにビッグデータの重要な国際会議であるIEEE Big Dataでも計23報が発表されている。また、社会的・経済的観点については緑内障診断、企業の製品改善、自治体サービス、遺伝子情報解析による予防医療支援など、実社会の問題解決への貢献をめざした実証実験につなげている。

以上により、本複合領域は総合的に特に優れていたと評価できる。

*本数値は領域事後評価委員会（2021年2月9日）時点の件数に新型コロナウイルスの影響による2研究課題の延長期間の成果件数を合算したものである。

1. 複合領域としての成果について

1.1 CREST

(1) 研究領域としての研究マネジメントの状況

ビッグデータの統合解析のための次世代基盤技術の創出・体系化に向けて、重点的に取り組むコア分野の設定、研究用データの整備、分野を超えた連携、および、国際連携等、領域のねらいを明確に定め、技術に加えて、コア分野の1つとして技術の社会における利活用に不可欠な「ビッグデータ法制度」を設定し、明示的に組み込んだことは社会的な意義があり、評価できる。

研究課題は、上記5つのコア分野において高い独創性・新規性に加え、提案者が明確な目的意識を有することを重視して選考が行われた。結果的に、ビッグデータ解析の基礎技術である数理基礎技術、自然言語処理や画像処理、そしてセキュリティとシステムアーキテクチャ等、我が国の技術を一層発展させる重要な分野の研究課題を広く採択した点は評価できる。

領域アドバイザーについては、高い実績を有する多様な専門分野の研究者を選定し、戦略目標の達成ビジョンにある「イノベーションによる新産業・新市場の創出」を強化するため、企業経験のある研究者から約半数を配置した。また、国際競争力強化と国際連携推進のために、米国およびフランスの研究者を海外アドバイザーとして選定している点も評価できる。

研究マネジメントでは、国際連携において米国、ヨーロッパ、アジア、日本の動向を踏まえ、国際的な観点から本研究領域の重要性を強く反映した運営、推進が行われ、米国 NSF や EU との研究プロジェクト連携でワークショップや会合を積極的に行い、NSF と情報通信研究機構 (NICT) の共同公募事業の採択につなげた。また、イノベーション創出につながる実用化が期待できる項目への重点的な予算配分により、促進された実証実験やオープンソース化など具体的な成果が確認できる。さらに、大学や国立研究機関に加えて企業およびアプリケーション分野の研究者を含めることで、ビッグデータの社会利用とその先のイノベーションへの発展の可能性を示した点は特筆される。

以上により、研究領域としての研究マネジメントは特に優れていたと評価できる。

(2) 研究領域としての戦略目標の達成状況

① 研究成果の科学的・技術的な観点からの貢献

CREST 全体では、発表論文 1,644 報* (領域中間評価時には 911 報。括弧内は以下同様)、そのうち国際論文 1,527 報* (850 報)、口頭発表が 2,504 件* (1,152 件)、また招待講演は 935 件 (570 件) としての研究成果を発表している。このうち、「ERA で A ランク」と「Qualis で A1 ランク」を同時に満たすトップカンファレンスの発表論文が 179 報、またビッグデータの重要な国際会議である IEEE Big Data で 22 報の発表を行っており、非常に高く評価できる。また、書籍は 68 冊* と多数あり、これらの中には技術と法制度分野に関する書籍も出版され、本研究領域の独創性を示している。加えて、松岡チームの記事掲載・報道 305 件

も社会からの注目度の高さを如実に示している。

以上により、研究成果の科学的・技術的な観点からの貢献については、特に高い水準にあると評価できる。

② 研究成果の社会的・経済的な観点からの貢献

主要な対象としている“アプリケーション分野との協働”、“アプリケーション分野での課題解決”や本研究領域の目標である“アカデミア・企業等が様々な分野のビッグデータを統合分析できる共通基盤を利活用して、研究開発や実用化を推進する環境作り”に関して積極的に取組み、将来の活用に向けた社会実装への取り組みが活発に行われた。具体的には、今後、深刻化が予想され失明リスクのある緑内障について従来と比較し被検査者の負担が軽い緑内障進行予測技術、大都市の複雑な状況を粗視化し現実的な時間で問題を解けるようにした大規模道路ネットワークの避難計画策定技術、実社会の散在情報を集約し有効に活用する(製品の不満など意見を製品改善に反映、自治体ホームページ内サービス問い合わせに回答する対話エージェントなど)自然言語処理技術による社会活動支援技術、教師データが少ない場合でも、高い分類性能を持ちモデルを学習できるクラス間学習手法によるデータ理解・要約・検索技術、ゲノム情報と医療情報の収集・統合・統計解析の一連の作業を暗号化したまま秘密裏に計算する準同型暗号を用いた生活習慣病の予防医療技術、単独では抽出困難な情報分析について秘密分散された情報(医療施設や学校などの公共データ)の統合情報分析技術、ビッグデータ処理性能に優れるスーパーコンピュータ設計技術、特定クラウド事業者に依存せず、資源割当の自動化等による環境設定の省力化、迅速性、再現性を実現するオーバーレイクラウド技術などが創出され、実利用・実運用に向けた取り組みが始まった。これら研究成果は、社会的に大きなインパクトを与える実用性、応用可能性を示しており、今後の波及効果も期待できる。

また、自然言語・画像基盤技術、セキュリティ基盤技術、システムアーキテクチャ基盤技術の各分野では、多数のソフトウェアが公開されており、知的財産権としては特許出願 79 件、うち国際出願 47 件と多数出願されている。

以上により、研究成果の社会的・経済的な観点からの貢献については、高い水準にあると評価できる。

1.2 さきがけ

(1) 研究領域としての研究マネジメントの状況

「数理基礎技術分野」、「自然言語・画像基盤技術分野」、「システムアーキテクチャ基盤技術分野」、「ビッグデータ法制度分野」の4つの重点分野に加えて、アプリケーション分野のビッグデータの新規取得・準備や他研究者へのデータ提供を目指す「データさきがけ」を創設し、研究用データ整備により複合領域全体のビッグデータ統合解析の研究推進に貢献した点が特筆に値する。

領域アドバイザーは、CREST・さきがけ兼任の他にデータベースとデータマイニングを専門とする幅広い研究分野に対応できる、さきがけ専任アドバイザー2名を配置した。

研究マネジメントについては、研究者の自由な発想を活かしつつ、トップダウン施策やアドバイザーとのバランスを重視して研究実施状況を適宜確認し、必要に応じて合同領域会議、サイトビジット等の機会できめ細かなフィードバックができる体制で臨んだ。実社会の情勢が大きく変化する中で、研究目標や成果の出し方も適宜修正が必要だが、その場合にも、きめ細かな対応ができる柔軟な体制で推進したことが評価できる。実際、他分野連携が積極的に行われ、個人研究からチーム型研究への発展につながるケースもみられ、特に「データさきがけ」での成果を基にCREST研究課題・AIP加速課題に採択されるなど新たな研究課題の創出、および人材育成についても成果が認められる。

以上により、本研究領域の研究マネジメントは特に優れていたと評価できる。

(2) 研究領域としての戦略目標の達成状況

研究成果は、論文118報(領域中間評価時には93報。括弧内は以下同様)で、そのうち国際論文89報(68報)、口頭発表443報(307報)、招待講演179報(118報)として発表されている。特に、このうち「ERAでAランク」と「QualisでA1ランク」を同時に満たすトップカンファレンスの発表論文数が19報(11報)、またビッグデータに特化した重要な国際会議であるIEEE Big Dataには1報(1報)の発表を行っており、研究成果の国際的水準は非常に高いと評価できる。特許出願は9件(5件)で、そのうち国際出願は1件(0件)であった。

各々の研究成果では、(i)数理基盤技術分野では、基盤技術の研究で優れた成果を上げ、その結果を応用に結びつけることで科学・技術の発展に貢献、(ii)自然言語・画像基盤技術分野では、主要なメディアのビッグデータを統合的に活用するための基盤技術に関する研究で発展に大きく貢献し、医療画像を用いた種々の研究、また、宇宙観測に動画像を用いる研究で成果を上げ、(iii)システムアーキテクチャ基盤技術分野では、未来のビッグデータ処理基盤を見据えた研究で、(iv)ビッグデータ法制度分野においては、個人情報保護やオープンデータに関連した法制度に関する研究で成果を上げた。(v)データさきがけでは、これまで利用可能な形で存在しなかったデータ(大豆の栽培、脳活動、腸内環境の3種)を収集・整理した。このように教育・医療という社会の根幹を支える活動を主な対象にして、実社会の実データを用いた実証実験を成功させ、社会に大きく貢献できる潜在力を持っていることを示した。

以上により、戦略目標の達成状況については、特に高い水準であると評価できる。

以上