

領域番号	4503	領域略称名	疎性モデリング
研究領域名	スパースモデリングの深化と高次元データ駆動科学の創成		
研究期間	平成25年度～平成29年度		
領域代表者名 (所属等)	岡田 真人（東京大学・大学院新領域創成科学研究科・教授）		
領域代表者 からの報告	<p><u>(1) 研究領域の目的及び意義</u></p> <p>より深く自然を知りたいという飽くなき探究心が、とどまることを知らない計測技術の向上をもたらし、大量の高次元観測データを日々生み続けている。</p> <p>これを好機と捉え、科学技術の水準を革新的に向上・強化させるために、情報科学と自然科学が緊密に融合した革新的な自然探究の方法論である高次元データ駆動科学を構築する。我々は、そのためのキーテクノロジーが、スパースモデリング(SpM)であると考え。SpMは、高次元データに普遍的に内在するスパース(疎)性を利用することで、データから最大限の情報を効率よく抽出できる技術の総称である。これまでもSpMは個別分野において萌芽的成果を生み出しており、それらの背後にある共通原理を明確化し、自然科学全体に革新的展開をもたらしている。</p> <p>本領域では、SpMや高次元データ解析で顕著な実績をあげている情報科学者と、生命分子からブラックホールに至る、幅広い自然科学の実験・計測研究者がSpMというキーテクノロジーを軸として緊密に連携することで、大量の高次元データを効率的に科学的な知へとつなげる高次元データ駆動科学を創成する。これにより、これまで個々の分野ごとに探求されていた課題に対して、共通原理に基づく革新的な科学的方法論を確立する。このようにして、あらゆる科学分野の研究に大きな波及効果をもたらし、来たるべきデータ科学の時代に向けて、我が国の学術水準の圧倒的優位性を確固たるものにする。</p>		
	<p><u>(2) 研究成果の概要</u></p> <p>自然科学の発展は、円滑な仮説・検証ループに依拠する。仮説・検証ループを回す上で、計測データにフィットするモデルを探すことに留まらず、得られたモデルの妥当性の評価、および、妥当と判断されたモデルを前提知識とした新たなデータの解釈が重要である。この視点に立脚して、シンプルだが過度に単純化しすぎない普遍的なデータ解析手法の開発を推進し、「仮説・検証ループを回すための統計的手法の深化と創成」を図ることが本領域の基本学理として見出された。また、情報科学と自然科学の協同においては、仮説・検証ループの共通認識が重要で、その協同体制に関する暗黙知を形式知に定型化した指導原理として「データ駆動科学の三つのレベル」が生まれた。これらの基本学理と指導原理に基づき、自然理解のための普遍的・系統的な数理・情報科学的手法に基づく「モデリング基盤」の構築が実現し、「高次元データ駆動科学」が確立された。</p> <p>データ駆動科学により、自然科学の仮説・検証ループの爆発的な加速が実現され、横串型連携の確立により、分野・階層を超えた1000件以上の研究成果が挙げられた。これは bioinformatics の流れを汲む米国 Materials Genome Initiative に代表される、単一の自然科学分野と情報科学分野の連携である「X-informatics」型のプロジェクトには不可能であり、我が国の学術水準の圧倒的優位性が確立された。また、本領域に参画した若手研究者から常勤研究職に計52名も就職し、次世代のデータ駆動</p>		

	<p>科学を担う若手人材育成の骨格が形成された。本領域の活動は、JST CREST・さきがけの情報計測領域の戦略目標で引用されたほか、三つのレベルの協同体制は、領域代表がマテリアルズインフォマティクスアドバイザーに任命されるなど(研)物質材料機構の統合型材料開発・情報基盤部門に影響を与えている。</p>
--	--

<p>科学研究費補助金審査委員会における所見</p>	<p>A (研究領域の設定目的に照らして、期待どおりの成果があった)</p>
	<p>本研究領域は、スパースモデリング技術を核として、情報科学と自然科学の研究者が緊密に連携することにより、従来の統計科学、情報科学の細分化した研究領域を超える高次元データ駆動科学の創成の実現を目指している。</p> <p>情報科学と自然科学の複数の分野にまたがる 985 人の研究者が参画し、1,000 件を超える論文発表等、非常に優れた研究成果を上げている。データ駆動科学の学理である様々な統計手法を自然科学の仮説検証ループに取り込むための普遍的知見を得るなど、スパースモデリングのモデリング原理を堅固にした。公開シンポジウムの参加者は延べ 2,000 人に上るなど、研究成果の公表・普及に非常に積極的であった。特に次世代のデータ駆動科学を担う人材を多く輩出した点は高く評価できる。</p> <p>総合的に、本研究領域の三つの重点目標、データ駆動科学の実践、モデリング原理の確立、数理基盤の形成は十分に達成され、研究領域の設定目的に照らして、期待どおりの成果があったと評価できる。今後も引き続き、国内外のシンポジウム開催等を通じて関連研究者との連携強化・拡張に努め、データ駆動科学を持続的に発展させていくことを期待する。</p>