

## 平成27年度 新学術領域研究（研究領域提案型）中間評価結果（所見）

### 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

岡田 真人（東京大学・大学院新領域創成科学研究科・教授）

### 研究領域の概要

大量の高次元計測データに隠された規則性を抽出するデータ解析の系統的技術の開発は、来るべき「データ科学時代」における全ての科学分野に共通する喫緊の課題である。本領域では、多くの自然科学分野の高次元計測データに普遍的にスパース性が存在することを基本原理としたスパースモデリングに注目し、生命分子からブラックホールに至る、幅広い自然科学分野の実験・計測研究者と情報科学者の連携により、この課題を解決する。これにより、スパースモデリングの数理的基盤を深化させるとともに、高次元データ駆動科学ともいべき新学術領域を創成し、これから本格化するデータ科学時代に向けて我が国の学術水準の圧倒的優位を確立する。

### 領域代表者からの報告

#### 1. 研究領域の目的及び意義

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。

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。

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。

#### 2. 研究の進展状況及び成果の概要

本領域では、データ駆動科学の実践、モデリング原理の確立、数理基盤の形成をそれぞれ担当する実験・

計測グループ(A01, A02), モデリンググループ(B01), 情報科学グループ(C01)の研究項目をおいた。総括班を含めた各計画研究の進捗は、2年間の活動で、4年9ヶ月で設定した半分以上の研究課題が進展しており、研究は順調に進んでいる。

本領域は先端的データ解析手法の活用による科学的方法論の革新を目的としたものであり、各計画研究においては、データ解析による限界打破を目指さない実験・観測研究、あるいは、実験・観測研究の目的を忘れた手法開発に陥らないように領域を運営した。その結果、David Marrが指摘した三つのレベルが、データ駆動科学を創成するためにも重要な着眼点となることを確信した。データ駆動科学においては、データ解析の目的である計算理論のレベルと、データ解析手法である表現・アルゴリズムのレベルの間に、モデリングのレベルが存在することに気づいた。これらの三つのレベルを新たに“データ駆動科学の三つのレベル”と名づけ、データ駆動科学の学理の原点に位置付けた。

領域活動では、この三つのレベルとレベル間の関連性を明確化するように徹底した。その結果、本質的にイノベーティブな、分野・階層を超えた共同研究が多数生まれ、それらが学術論文として結実しはじめている。これは本領域から生まれた新たな学理の圧倒的な力強さの表れである。

## 審査部会における所見

A (研究領域の設定目的に照らして、期待どおりの進展が認められる)

### 1. 総合所見

本研究領域は、大量の高次元データから科学的知識を抽出するための系統的方法論の構築を目指した、意欲的な提案である。様々な分野において実験・計測に関わる研究者と、数理的な分析技術に関わる研究者の密な連携を図るため、組織構成に工夫した点には独自性があり、領域代表者のリーダーシップにより活発な活動がなされている。発展の著しい分野であるが、当初の目標に沿った一定の成果が出ており、研究は順調に進んでいる。特に、若手研究者の育成に関しては、優れた実績が得られている。今後の研究では、本研究領域の大きな目標である、スパースモデリングの一般的な指導原理の確立を期待したい。

### 2. 評価の着目点ごとの所見

#### (1)研究の進展状況

当初の計画に基づき順調に進展している。モデリンググループにより、実験・データ収集と情報解析をつなぐアイデアはうまく機能している。個々の研究分野において、3つの研究項目を縦につなぐ成果が出ており、本研究領域の基本的な目標は十分に達成されつつある。また、審査意見に対する対応も、誠意をもってなされている。今後は、大きな目標であるデータ駆動科学の創成のための、領域全体を俯瞰する考え方を提示・共有することが期待される。

#### (2)研究成果

個々の研究分野において、十分な成果が出ている。高レベルのデータ特徴抽出や、高速なデータハンドリングが可能になったことにより、見いだされた新しい知見の例が多数報告されている。論文発表、特許申請なども良好である。いろいろな研究分野を対象としてセミナーを開催しており、普及活動として評価できる。若手研究者における研究成果も多く、優れた成果が得られている。

(3)研究組織

研究組織については特に問題はない。実験・計測グループ（研究項目 A）と情報科学グループ（研究項目 C）を媒介するモデリンググループ（研究項目 B）を介在させたことは、研究領域全体のコヒーレンスを維持する良い仕組みであり、今後も緊密な連携を期待したい。

(4)研究費の使用

特に問題点はなかった。

(5)今後の研究領域の推進方策

これまで研究は概ね順調に推移しているため、今後も基本的には現在の方針で遂行することが望まれる。ただし、国内外において競合するプロジェクトもあるため、連携や組織の強化には、なお一層の取組みが必要である。

(6)各計画研究の継続に係る経費の適切性

特に問題点はなかった。