

科学的知見の創出に資する可視化研究の推進

① 計画の概要

本研究計画マスタープランでは、2017年に出版した日本学術会議提言「科学的知見の創出に資する可視化にむけて」（以降、提言）を参照して、さまざまな科学分野において得られるデータを活用して科学的知見を促す対話型データ分析・可視化技術の開発とその評価を実現するための可視化研究コミュニティの構築を目指す。

本研究計画では、提言で述べたように可視化研究を以下の3つに分類する。

1. デジタル化：学術的に興味深い現象や対象からデータを生成すること
2. データ可視化：デジタル化により得られたデータから画像を生成すること
3. 可視化評価：データ可視化により生成された画像が人間の認知に及ぼす影響を定量化すること

デジタル化研究について、デジタル化が未着手の現象や対象を取り上げ、データ化を実現するための技術を開発する。本研究では、古文書を含む文書のデジタル化に注目する。文書によっては、文字データとして表現されている場合も多いが、文書間の関係性の評価が容易となるようにベクトルデータ化する。古文書の場合は、文字データにもなっていない場合が多く、前処理としてのデジタル化技術を開発する。

データ可視化研究については、データを多次元時系列データとして、これを拡張された流れデータとしてみなす。流れの可視化で開発されてきた技術を基盤として、このデータの裏側に潜む知見の抽出を促進する技術を開発する。たとえば、データ全体を効率よく俯瞰したり、また、データからシステム状態特定やその遷移の抽出を支援するための対話的可視化技術を開発する。

可視化評価研究では、究極的には、可視化のわかりやすさの定量化が目標となる。これらの研究開発を通じて、デジタル化・データ可視化における研究成果が科学的知見の創出に資するものなのかどうかを可視化評価における研究成果を用いて評価することが本研究計画の概要である。

② 学術的な意義

本研究において開発するデジタル化技術により、研究データ保存システムなどに保存された大規模文書データ分析が促進される。例えば、日本学術会議提言文書を自然言語処理技術によりベクトルに変換し、提言文書分類とその可視化を実施することができる。この結果、領域間の共通性が発見され融合領域創生の可能性が高くなる。また、三次元CT装置を使った文献デジタル化により、書籍などから、物理的な解体を行わずに、デジタル化が可能になるものと期待される。

本研究により開発されるデータ可視化技術により、多次元時系列データから状態遷移図を導出することが可能となり、様々な分野でパラダイムシフトに資する研究成果が生み出される可能性が高くなる。具体的には、生命科学分野研究者との協働により、細胞内で計測されたカルシウム濃度に関する多次元時系列データからシステム状態を同定し、その遷移を可視化することができれば、生命のメカニズム解明につながる。本研究により実施される可視化評価は、可視化の学理を究めるものである。認知科学者や脳科学者とともに、どのような画像化が気付きや閃きを促進するのかを明らかにすることができる。画像や画像生成に関して行うインタラクションとともに脳機能計測などを行い、その計測結果と画像からの気づき・閃きを紐づけ、科学的知見を創出する可視化の効能の説明モデルを作成することができる。

③ 国内外の動向と当該研究計画の位置づけ

可視化技術の発展は1986年から始まり、多くの技術が開発されたが、適用分野は限定的であった。困難な問題に対する可視化技術の適用が大きく注目されたのは、米国でテロ攻撃が起こった2001年以降である。テロに関する多くのデータがデータベースに格納されていたにもかかわらずテロ攻撃を許してしまった経験から、2004年、国土安全保障省の配下に、National Visualization and Analytics Center (NVAC) が世界に先駆けて組織され、スタンフォード大学をはじめとする6ヶ所の地域研究センターとともに全米規模で研究・開発が進められてきた。一方、欧州でも、2009年からVisMasterプロジェクトが開始された。さらに、中国において始まったChinaVis会議はいまや1000名ほどの登録を誇る勢いを示している。我が国における可視化技術の研究は、欧米中の後塵を拝しているが、流れの可視化技術の研究者のコミュニティである可視化情報学会が誕生したのは1981年である。当該計画は、可視化技術の開発とその適用を手掛ける我が国初の可視化研究コミュニティ構築に関するものである。

④ 実施機関と実施体制

本研究計画において、科学的知見については、科学的方法による研究活動で得られる知見であると考え。このうち、科学的方法において、重要な役割を果たす仮説は、因果関係として構成されることが多い。(1)目的と実施内容 での記述内容に対応して、3つのタイプの研究拠点（いずれも研究グループ単位）を提案する。

<デジタル化研究拠点>

さまざまな科学分野において、対象となる現象からデータを生成し、これを俯瞰可視化するための環境を整備し、問題設定が効果的に行える環境を構築する。立命館大学は、文化財デジタルアーカイブに関する研究分野で突出した研究成果があり、人文・社会科学に関するデータ生成を担当する。海洋研究開発機構・東京大学は、海洋科学や流体科学に関する研究分野で突

出した研究成果があり、自然科学に関するデジタル化研究を担当する。

<データ可視化研究拠点>

多次元時系列データから次元圧縮技術を使って、二次元平面上で散布図を作成し、専門家による状態の特定を支援するための対話的可視化技術を開発する。慶應大学・京都大学・大阪大学は、対話的可視化技術に関する研究分野で突出した研究成果があり、データを用いた因果探索・状態遷移抽出支援技術の整備・開発を担当する。

<可視化評価研究拠点>

本拠点では、科学的知見の創出において重要な、対話的可視化技術における効能を測定するための環境を構築する。測定対象は、可視化技術利用者の客観的生体情報（脳波、脈波、瞬目、視線）と主観的情報（閃き、気づき）などを考え、これらと可視化効能を関連付けて、可視化効能説明モデルを構築する。東北大学・芝浦工業大学・明治大学・京都大学は、可視化効能や感性の測定に関して研究成果があり、生体データを用いた可視化効能評価技術の研究開発を担当する。

⑤ 所要経費

初年次 480,000,000 円 以降毎年 165,000,000 円

<デジタル化研究拠点>

設備費：三次元 CT 装置 100,000,000 円/装置（維持費 10%）

人件費：ポスドク研究員 5 名 50,000,000 円

<データ可視化研究拠点>

設備費：高解像度表示装置 50,000,000 円/装置（維持費 10%）

人件費：ポスドク研究員 5 名 50,000,000 円

<可視化評価研究拠点>

設備費：可視化効能測定装置 200,000,000 円/装置（維持費 10%）

人件費：ポスドク研究員 3 名 30,000,000 円

⑥ 年次計画

【2019-23 年度】～基盤技術開発～

本研究では、古文書を含む文書のデジタル化に注目する。また、研究データ保存システムに格納される大量の文書データをベクトル化し、これを俯瞰的に可視化するためのシステムを開発し、科学的知見の創出が効果的にできるかどうかの評価を行う。さらに、多次元時系列データからこれを次元圧縮により可視化した散布図において、特徴あるクラスターを状態として特定するための対話環境を開発する。

【2024-28 年度】～要素技術開発～

(a) デジタル化研究拠点において、歴史的価値が高く、解体が困難な古文書に対して、本デジタル化手法を適用し、その記載内容を可視化する。大量の科学技術文書を俯瞰可視化し、新たな文書作成の際にどの程度役立つかを明らかにする。大量文書を多次元時系列データとして表現し、これらから状態を対話的に特定し、その遷移を導出することにより、ある分野における歴史的系統図を作成する。

(b) データ可視化拠点において、潜在因子推定のための対話操作環境に関する研究開発を行う。研究活動において、未観測のデータが関係する潜在因子を推定することが重要である。

(c) 可視化評価研究拠点において、データの視覚的表現の適切さに関して、本研究計画で開発を行う可視化効能測定環境において評価する。

【2029-33 年度】～技術適用検証～

要素技術開発の段階で開発された技術について、政策策定や教育の現場で活用できるように社会システムを改革する。

⑦ 社会的価値

本研究計画による成果を使って、世の中の文書において SDGs に強く関係するものを効果的に効率よく探索することが可能となる。本研究では、さまざまな分野における知の枠組みを共有する活動が重要な柱になっていることより、さまざまな分野において、その知的価値が高く評価されるものと予測される。単なる共有だけでなく、他分野での知的財産を問題解決に活用したり、それがきっかけとなって新たな融合分野が生まれ出される可能性も高いと推測される。

本研究では、革新的な因果推論・状態遷移図作成支援技術が開発されることが予想され、そのうちの多くは、ビッグデータ解析ソフトウェアシステムとしての製品化が見込まれることより経済的・産業的価値は高い。

⑧ 本計画に関する連絡先

小山田 耕二（京都大学学術情報メディアセンター）

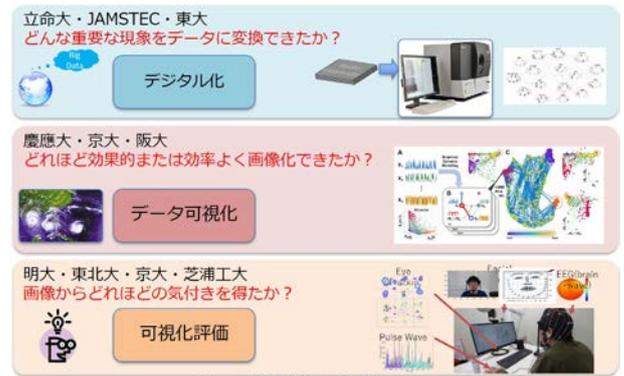


図1 可視化研究の分類と拠点