

データ分析支援を目的としたオープンデータの可視化ツールの開発と検証

唐澤拓実

インターネットの普及に伴って全世界にあふれる情報に触れることが増えてきたことで、データを正確に読み解く力“データリテラシー”を身につけることが不可欠となっている。データの集合から意味を見出し、正しく解釈し、活用するためには、データを可視化することが有効である。データの全体像を眺めることから始まり、注目点を探し、可視化することで、知見を得ることができる。このような視覚的な情報探索のプロセスを繰り返すことで、情報を様々な視点から読み解くことができるようになる。

そこで本研究では、可視化によるデータ分析を支援することを目的として、利用者に最適な項目やグラフを選択する力、能動的にデータ分析を行うスキルを身につけることができるシステムを目指した。オープンデータを利用することで、身近なデータから興味を引き出し、可視化されたグラフをインタラクティブに操作することによって、他のデータとの関連性を探ることができる仕掛けを実現した。

開発したシステムは、利用者の端末環境に影響されないよう web ブラウザ上で動かすことができるアプリケーションとした。利用者が登録したオープンデータを自動的に可視化するにあたり、まずオープンデータに登場する項目名の一覧を表示させた。その次に、単一のデータから他のデータとの関連性を眺めることができるようにグラフの提示順を棒グラフ、鶏頭図、マップチャート、パラレルコディネート、対散布図の順に提示した。このようにシステムを利用する手順を設定することで、利用者が探索行動において迷わない仕掛けを持たせた。実際にシステムを利用する際には、手順に沿って可視化されたデータからフィルタなどのユーザ操作によって特徴を発見し、ユーザが選択する別のデータとの関連性を探るという工程を繰り返し行う。その繰り返し過程で得られた知見を還元し、さらなる発見へと結びつけることを可能にした。

システム利用によるデータ分析支援の効果を見るために、筑波大学の大学生を対象に実験を行った。47 都道府県別のオープンデータを用いて、データとグラフを見比べながらシステムを利用してもらった結果、視覚的な情報探索のプロセス習得に効果が見られた。また、多くの人が未知のグラフに興味を持ち、その表現方法から新たな知見獲得につながった。一方で、視覚的な情報探索のプロセスを行ってもらうために、ある一定のタスクを課すことが効果的であると考えた。また、今後の発展性として、データ解析の分野と絡めることで、さらに他者との情報共有により、協調的なデータ分析へとつながり、より幅広くデータ分析支援を達成することができると思われる。

(指導教員 時井真紀)