

科学的可視化による因果関係の探索

小山田 耕二 (京都大学 学術情報メディアセンター 教授)

世界中の公的・科学由来のデータがインターネット経由で開放されるオープンデータ時代に突入しつつある。そして、市民がこれらオープンデータにアクセスして、科学に参加できるオープン科学時代が到来しようとしている。このようなオープン科学時代には、市民は、科学的手法について、傍観者ではなく、積極的に関与する必要がある。この科学的方法では、対象の観察、問題設定、仮説構築、仮説検証、社会実装という流れを辿るが、これらのほぼすべてにデータ可視化技術が関与するようになった。特に最後の社会実装においては、科学的方法で重要な反証性を担保するために可視化は重要な役割を果たす。

科学的方法において、仮説は、因果関係で構成されることが多い。因果関係とは、XX ならば YY という形式で記述されることが多く、仮説を操作化したうえで、これらにデータを対応させることができる。統計的手法によりデータ間の相関は、計算できるが、因果関係になると順序性の決定が含まれるために、人間の判断が重要となり、ここにデータ可視化の役割の重要性が確認される。

本講演では、計測結果や計算科学シミュレーション結果データから因果関係の発見を支援する可視化技術について講演する。因果関係を発見するには、まず、相関関係を統計的手法で明らかにし、そのうえで、順序性を明確化する（図を参照）。順序性の発見については、時間的なずれを導入した相関関係の分析や観測変数に影響を及ぼす潜在変数の対話的導入により、研究者による発見の支援を行う。さらに、議論や思索中において、想起した仮説について、対応するオープンデータが存在するかどうかを効率よく判定する技術についても論点を提供したい。

図：線虫胚発生の核分裂動態を解明するために可視化された大規模因果関係ネットワーク

(データ提供：理化学研究所生命システム研究センター チームリーダー 大浪修一氏)

