

オープンデータに基づくサイエンス基盤およびサービス Science Infrastructures and Services based on Open Data

小島 功^{1*}
KOJIMA, Isao^{1*}

¹ 産業技術総合研究所情報技術研究部門
¹ Information Technology Research Institute, AIST

計算機やインターネットを使った科学研究の比重が高まるにつれ、シミュレーション結果や実験データなど計算機上にある「データ」を中心とした研究手法が重要性を増している。この場合、新たな科学的知見の実体は新たなデータであり、論文はそれを説明するものであるため、データ中心の科学研究は自然な方向性である。観測データやその解析が本重要な地球物理の分野に限らず、研究成果をデータの形で発表、公開することは普及しつつあり、論文や論文誌と同様にデータに対する引用やデータジャーナルといった様々な試みも進められている。

さて、デジタルデータはその柔軟性から、引用のためのデータの単位だけでも様々な要求があり、また更新の可能性もあり、扱いには新しいポリシーや方法が必要になっている。また、データは計算機上において共有や加工が容易なために、効果的な共有や再利用を進めると科学研究を大きく促進できる一方、不適切な方法では逆に研究を阻害することも考えられる。さらに、データは論文と異なり維持管理を行わなければならない、特に公開データでは維持管理のコストは無視できない。

一方、データが膨大になるにつれそれを効果的に管理するニーズや新たな知見発見のために高度に解析するニーズが明らかになっており、データだけでなくその周辺の情報（研究者であるとか関連論文であるとか）を効果的に連携させてデータの利活用を促進することもイノベーションや成果普及に有効であり、ビジネスとして成立しうるものとして様々なサービスの提供がなされつつある。

そこで本発表では、このようなデータ中心科学における現状をまとめると共に、その利点や課題についても明らかにする。特に、クラウドコンピューティングの普及に伴って、データがクラウド上に置かれ、他のクラウドのサービスと連携されることで研究そのものがクラウド上で遂行されて基盤となっている状況に注目して詳説する。

キーワード: オープンデータ, オープンサイエンス, サービス, doi, リンクトデータ, クラウド
Keywords: Open Data, Open Science, Science Service, doi, Linked Data, Cloud computing