

Title	二次的生成データに対するメタデータ自動生成に関する研究
Author(s)	荒川, 彰太郎
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/18297">http://hdl.handle.net/10119/18297</a>
Rights	
Description	Supervisor: 丹 康雄, 先端科学技術研究科, 修士 (情報科学)

## 二次的生成データに対するメタデータ自動生成に関する研究

2110006 荒川 彰太郎

近年、政府が主導する Society5.0 におけるデータ利活用に代表されるように、国、地方公共団体、民間企業などで分散して保有するデータを連携して、活用することで生産活動の高度化・効率化、サービスの利便性の向上等が実現すると期待されている。

しかし、産官学のさまざまな分野のデータ連携において、データカタログに記載される情報の表記が統一されていないなど、相互運用の観点で課題があった。こうした状況を受け、センシングデータやデバイスデータを活用するために、データカタログ作成のガイドラインがデータ社会推進協議会によって推進されている。これにより、メタデータを説明する際の記載方法の標準化が可能になり、メタデータを用いたデータの検索や分野間のデータ連携を行うことが今後可能になると考えられる。

データ利活用の場面において IoT デバイスなどから直接収集した生データである一次データのみならず、複数の一次データ等を加工して生成された二次データの利用が増加傾向にある。元データと加工処理の組み合わせの数で種類が増える二次データに対し、人手によるメタデータ付与が困難になることが考えられる。また、メタデータをどのように付与していくかが問題になりつつある。また、複数回の加工処理を行なった複雑な構造の場合、最後に生成されたデータから、元となったデータの情報や加工処理の情報を得ることができず、その二次データがどのようにして生成されたのかわからなくなることが挙げられる。

本研究では、これらの問題を解決するシステムを提案した。提案システムとして、組み合わせの数で増加する二次データの問題に対しては、一次データを加工処理し二次データを生成する際に、加工者がメタデータ生成のプログラムを作成する。これにより、種類が増える前の段階でメタデータ付与が行えるため、組み合わせで増加し続ける二次データの問題を解決することができる。生成された二次データの素性に対する問題に対しては、二次データのメタデータに履歴として、加工処理に使用した一次データのメタデータを付与することで、生成された二次データの加工経路などの説明をすることができる。これにより、二次データをデータ流通に登録することができる。と考える。

提案したシステムの活用することで、二次データを提供することができるようになることで、ユーザーがより良いサービスが実現できると考えられる。また、これらのデータを更に利用できるようになり、データ流通市場が活性化することも期待できると考える。