

論文審査の要旨

| | | | |
|---|------------|-------|-------|
| 報告番号 | 理工研 第 473号 | 氏名 | 久永 忠範 |
| 審査委員 | 主査 | 淵田 孝康 | |
| | 副査 | 小野 智司 | 重井 徳貴 |
| 学位論文題目 オープンデータの連携に関する研究 (Study on the cooperation of open data) | | | |
| 審査要旨 提出された学位論文及び論文目録等を基に学位論文審査を実施した。本論文は、地方行政団体の保有するオープンデータの連携に関して、データ間の連携度の高いものを数値化して利活用を推進することについて述べたものである。以下の全文5章より構成されている。 | | | |
| <p>第1章は序論で、オープンデータの背景と状況の説明を、政府の施策や取組の状況と共に説明している。またデータ連携を行うためには、機械判読しやすいデータ形式でなければならないことと、あまり高度な知識が必要なデータ形式は、地方公共団体職員がそのデータを作成することは容易ではないことを述べている。そのために政府は、地方公共団体の職員向けにデータ作成のための研修やe-learningサイトも開設して、オープンデータの推進を行っていることも述べている。</p> <p>第2章では、オープンデータに関する関連研究について述べた。政府の取り組むデータ連携のための共通語彙基盤の運用がどのように行われているか説明している。またオープンデータを活用または推進するために、自治体職員の負担減やオープンデータ運用のための費用対効果を考えた受容性向上を目的とした研究や、防災情報を速やかに配信するためのデータ連携の共有についても述べている。そして機械判読によるデータの連携を行うためには、共通のキーとなる語彙が必要であることを述べている。</p> <p>第3章では、本研究に必要なオープンデータ間の連携度を測るための述語ベクトル法について述べた。実験に必要なオープンデータの収集方法やそれらのデータの項目名と列データの関係、そして述語ベクトルを生成するための項目判定関数を用いた列間類似度について述べている。またオープンデータ間の連携度を計算する手法やデータの偏りをなくするための重み付けの手法、連携度の計算式、重み付けのない平均、正規分布、減衰関数、閾値を用いた4パターンのうち、どの計算式の連携度の正確性があるのかの考察をおこなっている。</p> <p>第4章では、述語ベクトル法を用いて、全国の地方公共団体が公開しているオープンデータから抽出した300のCSVファイルの連携度を列データ間類似度と項目名間類似度による2つのパターンで比較して正確性の実験を行っている。列データ間類似度によるオープンデータ間の上位の連携度を考察すると項目名が違っても列データが似通って入れば高い連携度を示していることを示している。また列データ間の連携度と項目名間の連携度の比較を行ない、列データ間の連携度の優位性を実験で明らかにしている。</p> <p>第5章では、本研究のまとめと今後の課題について述べている。本研究は、地方自治体の保有するオープンデータの連携度を測り、オープンデータを機械判読によって活用することを目的としている。今後は、地方公共団体のオープンデータの連携内容の詳細を調査し、連携度の計算時間を短縮し、連携度の高いデータ同士を活用するアプリケーションの作成について述べ、総括としている。</p> | | | |
| <p>以上、本論文はオープンデータの連携に関する研究で、オープンデータ間の連携について検討を行い、データ間の連携度を数値化することで連携をしやすいデータを抽出できることを明らかにした。これは今後の地方自治体のオープンデータの利活用に大きく寄与する。</p> <p>よって、審査委員会は博士(工学)の学位論文として合格と判定する。</p> | | | |