

論文の内容の要旨

論文題目	都市問題の分析に向けたナレッジグラフの構築と時間的・空間的・課題横断的な拡充手法の開発
学位申請者	江上周作

国内ではゴミのポイ捨て、放置自転車、ホームレス、犯罪など、都市の地理空間的な特性や人の行動に関係の深い様々な社会課題（都市問題）が発生している。近年、こうした都市問題のデータ利活用による課題解決が注目されており、政府を始め各自治体でも、統計情報や会議録など様々なデータのオープンデータ化の取り組みが進んでいる。しかし、都市問題そのものの詳細なデータに関しては、公開だけでなくデータの蓄積もそれ程進んでおらず、直接的なデータ分析を行うには十分でない。例えば放置自転車に関して言えば、自治体によっては年単位かつ広域な統計データは存在するものの、このような粒度の粗いデータをもとに、具体的な原因特定や解決施策の実施に向けた分析を行うことは困難である。また、ゴミのポイ捨て問題に至っては自治体がデータを収集することができておらず、一部自治体が企業に調査委託した結果を保持する程度であり、オープンデータとしては公開されていない。

また、都市問題は様々な原因が複雑に絡み合っている。そのため、解決施策実行に先立って自治体の抱える各種都市問題の原因や影響、都市問題同士の関係性を俯瞰して把握することや、解決に向けた費用対効果を把握することが望ましい。しかしながら、このような都市問題間の関係が整理されたデータセットは存在しておらず、自治体によっては担当地域において生じている都市問題を把握できていないこともある。

本研究では前述の背景を受けて、都市問題解決に向けた時間的・空間的・課題横断的なデータ分析活用基盤を構築し、活用例を示すことで問題解決の一助となることを目的とする。本研究では都市問題に関するデータを知識グラフ（ナレッジグラフ）として蓄積し、これをLinked Open Data (LOD) の形式で公開した。

初めに、都市問題ごとにデータを収集して時間的・空間的に詳細なナレッジグラフを構築した。まず、ナレッジグラフを構築する上で重要となるデータ構造（スキーマ）を設計し、次にソーシャルセンサを活用してデータを収集し、

設計したスキーマに基づいてナレッジグラフを構築した。しかし、ソーシャルセンサにより集められるデータには時間的または空間的な欠損が生じる。そこで、本研究では都市問題に関するナレッジグラフの時間的な欠損値について、ベイジアンネットワークを用いて推定した。空間的な欠損値について、数値流体力学による都市の気流シミュレーション結果と地理空間情報を基に、Graph Embedding手法とニューラルネットワークを用いて推定する手法を提案した。さらに、構築したナレッジグラフを活用して可視化アプリケーションを開発し、課題解決に向けて東京都との議論を通じて、青少年・治安対策本部のホームページで紹介した。

次に、複数の都市問題間の因果関係と関連予算情報に関するデータを収集し、課題横断的なナレッジグラフを構築した。これは、各都市問題の時空間的に詳細なナレッジグラフの上位概念的な位置づけである。まず、スキーマを設計し、次にWebから収集した記事、SNS情報、自治体オープンデータ、学術情報などを基に、自然言語処理とクラウドソーシングを用いて都市問題の因果関係を抽出する手法を提案した。設計したスキーマに基づいて、抽出した都市問題の因果関係や、関連する自治体事業予算情報を記述するナレッジグラフを構築した。構築したナレッジグラフを用いて、都市問題の因果関係と関連予算情報の検索例を示した。また、異なる文脈で出現したエンティティ同士の因果的な欠損について、セマンティックWebルール言語により推論することで補完した。さらに、都市問題の因果関係からなる悪循環および悪循環を引き起こす根源的課題をSPARQL Protocol and RDF Query Language (SPARQL)で発見する実験を行い、大阪市市民局協力のもとで、都市問題の専門家による評価を行った。

本研究で示したナレッジグラフの構築と拡充手法により、データ分析活用基盤の実現を通して都市問題解決の一助となることが期待できる。今後、長期的なデータ拡充を続け、自治体やNPOでの議論、市民協働イベント等での活用を通して社会実装していくことが課題である。

論文審査の結果の要旨

学位申請者氏名 江上 周作

審査委員主査 大須賀 昭彦

委員 川村 隆浩

委員 田中 健次

委員 西野 哲朗

委員 柏原 昭博

委員

委員

本論文は、『都市問題の分析に向けたナレッジグラフの構築と時間的・空間的・課題横断的な拡充手法の開発』と題して、以下に示す全6章から構成されている。

第1章『序論』では、本研究の背景、目的と意義が記述されている。背景では社会問題および都市問題におけるデータ利活用による課題解決の現状を述べた後、データ不足と因果関係の未整理により分析活用が困難という問題に対して、都市問題ごとのナレッジグラフと課題横断的なナレッジグラフを構築し、その中で生じる欠損を補完することや、その分析活用例を示し、自治体と連携することで解決の一助となることを目的とすることが明確に記述されている。

第2章『関連研究』では、ナレッジグラフとLinked Dataおよびオントロジーの概要、社会課題および都市分析に関するナレッジグラフの既存研究、ナレッジグラフの構築に関する既存研究や、ナレッジグラフの欠損補完と機械学習への適用に関して記述している。ここでは、本研究と既存研究との関係や、従来の手法では適用できずに新たな手法の開発が必要であることが記述されている。

第3章『都市問題ごとのナレッジグラフの設計と構築』では、特定の都市問題の時間的・空間的に柔軟な分析が可能なナレッジグラフを構築するため、必要な要件を抽出した上でデータスキーマを設計する手法について提案している。また、放置自転車とゴミのポイ捨て問題に関して実データを収集し、設計したスキーマに基づいてナレッジグラフを構築する手法について述べている。本手法により、都市問題ごとのナレッジグラフとして放置自転車とポイ捨てのナレッジグラフが構築されたこと、スキーマが適切であることを確認した。

第4章『都市問題ごとのナレッジグラフの時空間的な拡充と分析』では、データがソーシャルに収集されるために生じるナレッジグラフの欠損について、欠損を推定・補完することにより、都市問題ごとのナレッジグラフを半自動的に拡充し続ける手法について提案している。まず、ナレッジグラフの時間的な欠損値の推定について、ベイジアンネットワークを用いて放置自転車の台数を推定し、

ナレッジグラフに追加する手法について提案している。評価実験の結果約70%の精度で推定が可能になったことを確認している。この結果を受けて、ナレッジグラフのスキーマに含む情報が充分であることを確認し、精度向上について考察し課題を明らかにした。

次に、放置自転車の空間的な欠損値の推定について、流体シミュレーションを使用して流体のよどみ点を検出するように、欠損値を推定する手法について提案している。評価実験の結果、約30%の精度を記録しベースラインより大幅に精度が向上したことや、流体のよどみ点と放置自転車の観測地点に相関があったことを確認している。さらに、ポイ捨て問題のように問題の発生地点が一定の範囲内で空間的に密な問題に適用できるように、流体シミュレーション、Graph Embedding手法、ニューラルネットワークを使用して空間的欠損値を推定する手法を提案している。評価実験の結果、平均二乗誤差 (RMSE) として0.24を記録し、高精度に推定できることを確認している。さらに、欠損を補完した放置自転車のナレッジグラフを活用し、放置自転車の時系列変化をアニメーション表示することが可能なアプリケーションを開発し、構築したナレッジグラフの有用性を確認した。さらに、放置自転車問題解決に向けて東京都青少年・治安対策本部と連携して問題の意識向上に取り組んだ点は高く評価できる。

第5章『都市問題の課題横断的なナレッジグラフの構築と分析』では、都市問題ごとのナレッジグラフの上位概念的な位置づけとして、複数の都市問題間の複雑な因果関係を記すナレッジグラフの設計、構築、欠損補完および分析活用手法について提案している。まず、都市問題の因果関係と自治体の関連予算を記述するナレッジグラフのスキーマを設計している。次に、Web上の文書から自然言語処理技術とクラウドソーシングを使用して因果関係を半自動的に抽出し、設計したスキーマに基づいてナレッジグラフを構築している。因果関係の抽出におけるクラウドソーシングは弱から中程度の一致度が得られており、本研究においては適切であったことを確認した。次に、異なる文脈で抽出される因果単語同士の関係の欠損について、既知の因果関係を基に関連事象同士の因果関係として考えられるパターンを11個定義し、Semantic Web Rule Language (SWRL) を使用して推論補完する手法について提案している。さらに、クエリ言語を使用して構築したナレッジグラフから、都市問題の因果関係から自治体の関連予算まで検索することで、問題の解決施策検討に役立つ例や、都市問題の悪循環や根源的課題について発見する手法を提案している。悪循環について、大阪市民局協力のもとで都市問題の専門家のコメントに基づく評価を行ったことや、本研究が専門家同士の議論活性化ツールとして有用であるとの意見が専門家から得られた点は高く評価できる。

第6章『結論』では、本研究で得られた結果を整理し、本研究を一般化した際にどこまで適用可能かどうかについて、今後の研究の課題と展望について明確に記述している。

このように、本論文では都市問題のデータ利活用に向けた課題解決に向けた、ナレッジグラフの設計、構築、欠損補完、分析活用に関する一連の手法からなる枠組みを提案し、都市問題解決の一助としての有用性が確認できたことを示しており、優れた研究であると評価できる。

以上により、本論文は博士（工学）の学位論文として十分な価値を有するものと認める。