

学位被授与者氏名	槇 俊孝 (Toshitaka Maki)
学位の名称	博士 (工学)
学位番号	博 (一) 第 5 0 号
学位授与年月日	平成 3 0 年 3 月 2 0 日
論文題目	Linked Data の知識ベース化を指向したオープンプラットフォームの研究
論文題目 (英訳または和訳)	Study on Open Platform to Reinforce Knowledge Bases in Linked Data
論文審査委員	論文審査委員会 委員主査 : 福岡工業大学大学院知能情報システム工学専攻教授 若原 俊彦 同審査委員: 福岡工業大学大学院知能情報システム工学専攻教授 西田 茂人 同審査委員: 福岡工業大学大学院知能情報システム工学専攻教授 徳安 達士 同審査委員: 福岡工業大学大学院物質生産システム工学専攻教授 大山 和宏
論文審査機関	福岡工業大学大学院工学研究科
論文内容の要旨 (和文)	<p>Linked Data は、Resource Description Framework (RDF)に基づき、主語、述語、目的語の 3 つ組 (triple)で各リソースを整理したラベル付き有向グラフのデータである。Internationalized Resource Identifier (IRI)形式のリソースは、リソース間で横断的なリンクを形成可能であり、体系的に意味関係を参照することで知識ベースを構築できる。</p> <p>2009 年にオバマ政権がオープンガバメントを提唱して以降、オープンデータに関する諸活動が広がり、Linked Data をオープンデータとして公開した Linked Open Data (LOD)の公開件数が飛躍的に増加した。LOD はウェブ上に大規模な知識ベースを構築できることから、各国で地理や統計、辞書、文献などの様々な LOD が公開され、地域課題の解決や研究動向の分析などの活用が期待されている。しかし、現存の LOD は、述語やリソース型の定義に課題がある他、活用法が認知されていないため二次利用が限定的である。例えば、Wikipedia を LOD 化した DBpedia における約 80%のリソースは横断的リンクが可能である。これに対して LinkData.org における地理や統計、施設などの地域に関する LOD は、横断的リンクが可能でリソースは約 10%に過ぎない (2017 年 11 月 6 日現在)。つまり、地域に関する LOD は、知識ベースとなり得るリソースが比較的少数であることが分かる。また、LinkData.org に 31、560 種類もの述語が個別に定義されており、それぞれのリソースを網羅的に取り扱うことができないため知識として活用できない。本論文では、知識として活用できない Linked Data を知識ベース化する Resource Propagation Algorithm (RPA)、観光語彙基盤、動向分析システムをパッケージ化した新しいオープンプラットフォームについて述べる。RPA は、文字列型のリソースからキーワードを抽出してカテゴリを同定し、IRI リソースに変換した後に、グラフ構造に基づいて潜在的なリンクを推定する新しいアルゴリズムである。RPA は、従来の Label Propagation とは異なり、述語に基づいてリンク (エッジ) の柔軟な重み変更を実現しており、また、教師データが不要なため様々な種類の LOD を対象として知識ベースの生成が可能である。観光語彙基盤は、乱立した観光領域の述語を統一化する述語セットであり、IRI 型のリソースを基軸としている。LOD の知識ベース化によって二次利用が容易になるだけでなく、推定されたリソースの概念であるキーワードやカテゴリを用いることで高度な活用が可能となることを確認した。例えば、従来の統計データだけは計測できない地域の特色を定量的に分析できる他、ナビゲーションとして利用可能な質問応答システムを実装できることを確認した。本オープンプラットフォームは、Linked Data を知識ベース化してデータインフラの構築に貢献する非常に有用なものであり、LOD Challenge 2016 において優秀賞、並びに LOD for 地域創生賞を受賞し、また、電子情報通信学会 I-Scover 主催のコンテストにおいても最優秀賞、並びに優秀賞受賞して脚光を浴びている。</p> <p>本論文の構成は以下の通りである。</p> <p>第 1 章では序論として研究背景や関連研究、研究目的、特色と独創的な点を述べる。第 2 章ではオープンデータについて概説し、第 3 章で Linked Data の概説と</p>

	<p>ともに現状課題を議論する。第 4 章でオープンプラットフォームである RPA、共通語彙基盤、及び動向分析システムについて述べる。第 5 章で実験と考察を議論し、第 6 章では本論文の総括と今後の展望を述べる。</p>
<p>論文内容の要旨 (英文)</p>	<p>Linked Data are data of labeled directed graph that are a triple set based on the Resource Description Framework (RDF). The each resource of subject, predicate, and object on the RDF constructs the triple. The resource of Internationalized Resource Identifier (IRI) data type can constructs knowledge bases because it can set cross-links between resources and can refer semantic relationship systematically.</p> <p>Since the Obama administration advocated Open Government in 2009, various activities related to Open Data have expanded, and the number of Linked Open Data (LOD) that are published Linked Data as the Open Data has increased dramatically. The LOD have characteristics to be able to construct a large knowledge bases on the Web. In each country, various LOD such as geography, statistics, dictionaries, and literatures are published, and utilization such as regional problem resolution and technology trend analysis have been expected. However, currently released LOD have not been used much because they have problems such as unique predicates and sparse graph structures. For example, about 80% resources in DBpedia that is converted from Wikipedia can link to other resources. On the other hand, about 90% resources in LinkData.org related to region such as geography, statistics, and facilities cannot link to other resources. In the other words, users are difficult to use the LOD related to the region as the knowledge bases.</p> <p>This paper presents a new innovative open platform that is constructed by three functions: a Resource Propagation Algorithm (RPA) to reinforce the knowledge bases in the Linked Data, a tourism core vocabulary, and a trend analysis system. The RPA is a new algorithm to estimate potential links between resources based on the graph structure after extracting concepts of IRI data type. The RPA can generate the knowledge bases from various LOD because it can set flexible edge (link) weight based on the predicates and it does not require the teacher data. The tourism core vocabulary is a set of predicates that are based on IRI data type for unifying predicates related to the tourism field. The platform has been confirmed that secondary usage becomes easy and user can advanced use by reinforcing the knowledge bases of the LOD. For example, user can analysis regional features quantitatively by using the trend analysis system. In addition, we can implement the question answering system using the knowledge bases that is generated by the platform. The platform contributes for constructing the data infrastructure by reinforcing the knowledge bases. As a concrete achievement of the platform, the platform received the Best Award and “LOD for Region Creation Award” in LOD Challenge 2016. In addition, the platform received the Excellence Award and the Best Award in IEICE I-Scover competition, and got attention from many LOD specialist.</p> <p>This paper is organized as follows. Chapter 1 presents the study background, related works and the purpose, etc. as the introduction. Chapter 2 presents about the Open Data, and Chapter 3 discusses the Linked Data and the currently problems. Chapter 4 presents the RPA, the tourism core vocabulary, and the trend analysis system as the open platform. Chapter 5 discusses the experiments and consideration, and Chapter 6 presents the conclusion and future prospects.</p>
<p>論文審査結果</p>	<p>博士後期課程知能情報システム工学専攻 3 年の榎 俊孝氏が提出した学位論文を審査するとともに、最終審査を行ったので、その結果について報告する。</p> <p>(学位論文の概要)：近年のインターネット技術の発展とともに Web 技術も飛躍的に進化しており、Linked Data と Web 技術の融合によって次世代 Web であるセマンティック Web が発展している。セマンティック Web は、政府や自治体の主導のもとに進んでいるオープンガバメントによりオープンデータとして公開された Linked Open Data (LOD) などを利用して、主語、述語、目的語の三つ組(Triples)の関係から意味ベースを構築することが出来る。しかし、現状の公開されている三つ組には用語の表記や意味の統一化の問題の他、述語の構造やリンク構造に問題があり、コンピュ</p>

ータによる機械判読が難しく二次利用が進んでいない。本論文では、観光語彙基盤として新たに観光分野での用語や意味構造の統一化を図った。さらに、三つ組のグラフ構造においてリンクが張られていないノード間に新たに伝搬定数を定義してリンク伝搬する機構を導入し、類似した意味のノード間のカテゴリやキーワードのラベル推定機能を持たせ、しきい値以上のリンク重みを有するエッジにノード間のリンクを設定することにより、従来よりも 20%以上のリンクを設定できる新たな Resource Propagation Algorithm(RPA)を考案した。これらの効果を糟屋郡新宮町との共同研究で構築した観光情報ポータルサイト「たのしんぐう」の質問応答システムで動作確認し、その有効性も検証した。また、電子情報通信学会の文献検索システム I-Scover においても LOD を利用した時系列分析システムを作成して有効性を実証した。

(学位論文の構成)：第 1 章では、研究背景および研究の目的、論文の構成について述べている。第 2 章では、オープンデータの概要とその形式、第 3 章では Linked Data として Resource Description Framework(RDF)、LOD の課題、共通語彙基盤とオントロジーについて述べ、第 4 章では、オープンプラットフォームとして新たに構築した観光語彙基盤とその設計方針、知識ベース化のための RPA の原理、キーワード推定法およびキーワード特性、動向分析システムについて述べている。さらに、第 5 章では実験と結果の考察、第 6 章では結論を述べている。

本研究の成果は、博士後期課程において学術論文 3 編 (第 1 著者 1 編)、国際会議 9 編 (第 1 著者 3 編) に纏められている。特に、学術文献検索システム I-Scover 活用コンテストで LOD を利用した共起語グラフの可視化による支援システムを提案し最優秀賞を受賞するとともに、LOD チャレンジ Japan2016 において意味検索システムを出展して優秀賞および LOD for 地方創成賞などを受賞し、本研究の実用面からも新規性と有効性が評価された。

以上の理由により、学位論文審査委員会は本論文が学位論文に適合すると判定した。

学位論文公聴会においては、論文内容を丁寧に説明し理解を得るとともに、関連する分野の工学的及び技術的な質問に対していずれも適切に回答した。また公聴会後の最終試験においても、関連する専門分野の学識を有していることを確認し、今後さらにビジネス化をめざした研究を進めていく研究能力を十分に備えていることを確認した。以上の結果から、当学位論文審査委員会は本論文が博士 (工学) の学位に適格であると判定した。

主な研究業績

参考論文 4 2 編 1 冊

(1) 学術雑誌等

[査読有り]

1. "Beacon-based tourist information system to identify visiting trends to tourists", Journal of Robotics, Networking and Artificial Life (JRNAL), Vol.4, Issue.3 (2017)

Authors: Akihiro Yamaguchi, Masashi Hashimoto, Kiyohiro Urata, Yu Tanigawa, Tetsuya Nagaie, **Toshitaka Maki**, Toshihiko Wakahara, Akihisa Kodate, Toru Kobayashi, Noboru Sonehara

2. "Trend and Factor Analysis of Office Research in LOIS Technical Committee", IEICE Trans. Inf. & Syst., Invited Paper, Vol.E100-D, No.10, pp.2383-2390 (2017)

Authors: Toshihiko Wakahara, **Toshitaka Maki**, Noriyasu Yamamoto, Akihisa Kodate, Manabu Okamoto, Hiroyuki Nishi

3. 「I-Scover 文献メタデータを用いた時系列・技術要因分析システム」、電子情報通信学会論文誌、Vol.J99-D、No.10、pp.1002-1012 (2016)

著者：**榎 俊孝**、若原俊彦

[査読無し]

1. 「I-Scover を使用した研究、アプリケーションについて」、  
電子情報通信学会誌、Vol.98、No.12、pp.1106-1118 (2015)  
著者: 千村保文、宇野毅明、榎 俊孝、若原俊彦、Ruiyu Fang、Lu Fang、Qingllang Miao

国際会議

[査読有り]

1. "LOD Conversion System for generating Large Knowledge Base from Web Contents", 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE 2017), OS-CPS4, pp.155-158 (2017)  
Authors : Kazuki Takahashi, Toshitaka Maki, Toshihiko Wakahara, Toru Kobayashi, Akihisa Kodate, Noboru Sonehara
2. "Metadata Complement Method by Linked Open Data for Literature Search", The 20th International Conference on Network-Based Information Systems (NBIS 2017), NBIS-S9, pp.484-491 (2017)  
Authors : Toshihiko Wakahara, Toshitaka Maki, Kazuki Takahashi
3. "Resource Propagation Algorithm to Reinforce Knowledge Base in Linked Data", The 20th International Conference on Network-Based Information Systems (NBIS 2017), NBIS-S9, pp.476-483 (2017)  
Authors : Toshitaka Maki, Kazuki Takahashi, Toshihiko Wakahara, Akihisa Kodate, Noboru Sonehara
4. "Beacon-based tourist information system to identify visiting trends of tourists", The 2017 International Conference on Artificial Life and Robotics (ICAROB 2017), pp.480-483 (2017)  
Authors : Akihiro Yamaguchi, Masashi Hashimoto, Kiyohiro Urata, Yu Tanigawa, Tetsuya Nagaie, Toshitaka Maki, Toshihiko Wakahara, Akihisa Kodate, Toru Kobayashi, Noboru Sonehara
5. "Research Field Analysis Using Linked Data of I-Scover", 2016 IEEE 5th Global Conference on Consumer Electronics (GCCE 2016), OS-CNW1, pp.541-543 (2016)  
Authors : Toshihiko Wakahara, Toshitaka Maki, Kazuki Takahashi
6. "Link Prediction of LOD by Multiple Label Propagation Algorithm Considering Semantic Distance", 2016 IEEE 5th Global Conference on Consumer Electronics (GCCE 2016), OS-CPS4, pp.517-521 (2016)  
Authors : Toshitaka Maki, Kazuki Takahashi, Akihiro Yamaguchi, Toshihiko Wakahara, Toru Kobayashi, Akihisa Kodate, Noboru Sonehara
7. "A New Laboratory Assignment Support System for Undergraduate Students", The 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS 2016), FIMIS-S1, pp.361-366 (2016)  
Authors: Kazuki Takahashi, Toshitaka Maki, Toshihiko Wakahara
8. "Tourism Local Community System using LOD", 2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS

2016),ECDS-S1,pp.332-336 (2016)

Authors: Toshihiko Wakahara, Toshitaka Maki, Kazuki Takahashi, Akihiro Yamaguchi, Shinichiro Kimoto, Akinori Takagi, Yu Ichifuji, Noboru Sonehara

9. "A New Multiple Label Propagation Algorithm for Linked Open Data",The 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS 2016),IMIS-S4,pp.202-208 (2016)

Authors: Toshitaka Maki, Kazuki Takahashi, Toshihiko Wakahara, Akihiro Yamaguchi, Yu Ichifuji, Noboru Sonehara

国内学会や報告書等

[査読無し]

1. 「I-Scover SPARQL API 利活用レポート 1」  
電子情報通信学会 I-Scover ビジネスライアルチーム、Vol.1 (2017)  
著者：榎 俊孝,千村保文,若原俊彦
2. 「ふくおか IT Workouts 2016 参加報告 -新宮町でのおもてなしに向けた ICT 活用法の検討-」、福岡工業大学 FD Annual Report、Vol.7、pp.55-61 (2017)  
著者：山口明宏、若原俊彦、榎 俊孝、高橋和生、松前洋祐、橋本雅史
3. 「PBL Summit2017 参加報告」、  
福岡工業大学 FD Annual Report、Vol.7、pp.51-54 (2017)  
著者：高橋和生、山下拓弥、橋本雅史、榎 俊孝、山口明宏、若原俊彦
4. 「観光 LOD のための観光語彙基盤の実装」、  
2017 年電子情報通信学会総合大会、D-9-42 (2017)  
著者：榎 俊孝、高橋和生、山口明宏、若原俊彦、小館亮之、小林透、曾根原 登
5. 「観光情報の LOD 変換手法の検討」、  
2017 年電子情報通信学会総合大会、D-9-41 (2017)  
著者：高橋和生、榎 俊孝、若原俊彦、山下拓弥、坂井大輝、松前洋祐、小館亮之、曾根原 登
6. 「オープンデータを活用した地域の観光関連情報発信サービスにおける課題」、  
電子情報通信学会技術研究報告、Vol.116、No.488、LOIS2016-84、pp.123-128 (2017)  
著者：小館亮之、佐野未来、ティモシー ボルト、榎 俊孝、若原俊彦、田中康裕
7. 「LOD の汎用化を図るメタデータの設定手法」、  
電子情報通信学会技術研究報告、Vol.116、No.488、LOIS2016-82、pp.111-116 (2017)  
著者：榎 俊孝、高橋和生、若原俊彦、小館亮之、小林 透、曾根原 登
8. 「観光業の振興を図る情報ポータルサイト構成法の検討」、  
電子情報通信学会技術研究報告、Vol.116、No.405、LOIS2016-58、pp.57-62 (2017)  
著者：松前洋祐、坂井大輝、山下拓弥、高橋和生、榎 俊孝、山口明宏、若原俊彦
9. 「新宮町の LOD を用いた意味検索システム構成法の検討」、

電子情報通信学会技術研究報告、Vol.116、No.405、LOIS2016-57、pp.51-56 (2017)

著者：若原俊彦、榎 俊孝、山下拓弥、坂井大輝、松前洋祐、高橋和生、山口明宏、小林 透、小舘亮之、曾根原 登

10. 「観光コミュニティサイトにおける LOD 変換手法の一検討」、  
電子情報通信学会技術研究報告、Vol.116、No.405、LOIS2016-56、pp.45-50 (2017)

著者：高橋和生、坂井大輝、松前洋祐、榎 俊孝、山口明宏、若原俊彦

11. 「観光オントロジーの構築のための語彙基盤の提案」、  
電子情報通信学会技術研究報告、Vol.116、No.291、LOIS2016-37、pp.51-56 (2016)

著者：榎 俊孝、高橋和生、若原俊彦、山口明宏、小舘亮之、小林 透、曾根原 登

12. 「LOD のプロパティを考慮したマルチプルラベル伝搬アルゴリズムの検討」、  
電子情報通信学会技術研究報告、Vol.116、No.220、LOIS2016-24、pp.49-54 (2016)

著者：榎 俊孝、高橋和生、若原俊彦、山口明宏、小舘亮之、小林 透、曾根原 登

13. 「Broadic の導入によるマルチプルラベル伝搬アルゴリズムの性能改善と評価」、  
第 15 回情報科学技術フォーラム (FIT 2016)、O-024 (2016)

著者：榎 俊孝、高橋和生、若原俊彦、山口明宏、小舘亮之、曾根原 登

14. 「オフィス情報関連システムの研究動向とその分析」、  
第 15 回情報科学技術フォーラム (FIT 2016)、O-023 (2016)

著者：若原俊彦、榎 俊孝、高橋和生

15. 「ふくおか ITWorkouts2015 における新宮町でのおもてなしに向けた ICT 活用法の検討—

新宮発見隊の活動—」、福岡工業大学 FD Annual Report、Vol.6、pp.36-44 (2016)  
著者：若原俊彦、榎 俊孝、高橋和生、山口明宏、佐藤夏姫

16. 「地域の Linked Open Data の現状課題とその解決法に関する考察」、  
電子情報通信学会技術研究報告、Vol.116、No.138、LOIS2016-17、pp.51-56 (2016)

著者：榎 俊孝、高橋和生、若原俊彦、山口明宏、曾根原 登

17. 「CityData の LOD を対象としたマルチプルラベル伝搬アルゴリズムの適用手法」、  
電子情報通信学会技術研究報告、Vol.116、No.23、LOIS2016-10、pp.101-106 (2016)

著者：榎 俊孝、高橋和生、若原俊彦、山口明宏、一藤 裕、曾根原 登

18. 「I-Scover を用いた時系列・技術要因分析システムの活用法」、  
2016 年電子情報通信学会総合大会、TK-8-4 (2016)

著者：榎 俊孝、若原俊彦

19. 「地域観光 Triple データの生成とオントロジー化手法の研究」、  
電子情報通信学会技術研究報告、Vol.115、No.486、LOIS2015-73、pp.57-61

(2016)

著者：若原俊彦、榎 俊孝、山口明宏、木本紳一郎、高木昭典、一藤 裕、曾根原 登

20. 「Linked Open Data 生成のためのマルチプルラベル伝搬アルゴリズムの提案と性能評価」、  
電子情報通信学会技術研究報告、Vol.115、No.486、LOIS2015-72、pp.51-56  
(2016)  
著者：榎 俊孝、若原俊彦、山口明宏、一藤 裕、曾根原 登
21. 「Linked Open Data 生成のための自治体 CMS の構成法」、  
電子情報通信学会技術研究報告、Vol.115、No.410、LOIS2015-57、pp.53-58  
(2016)  
著者：榎 俊孝、若原俊彦、山口明宏、木本紳一郎、高木昭典、一藤 裕、曾根原 登
22. 「技術用語の類似度を用いた研究室配属支援システムの検討」、  
電子情報通信学会技術研究報告、Vol.115、No.410、LOIS2015-56、pp.47-51  
(2016)  
著者：高橋和生、榎 俊孝、若原俊彦
23. 「リンクトオープンデータを用いた地域コミュニティシステム支援手法」、  
電子情報通信学会技術研究報告、Vol.115、No.219、LOIS2015-26、pp.61-65  
(2015)  
著者：若原俊彦、榎 俊孝、吉井一希、佐藤夏姫、山口明宏、一藤 裕、曾根原 登
24. 「I-Scover を用いたメタデータの概念拡張による論文分析手法」、  
2015 年電子情報通信学会ソサイエティ大会、TK-1-2 (2015)  
著者：榎 俊孝、若原俊彦
25. 「オープンデータ自動生成を考慮した自治体 CMS の構成手法の検討」、  
第 23 回電子情報通信学会九州支部学生会講演会、D-14 (2015)  
著者：吉井一希、高橋和生、榎 俊孝、若原俊彦
26. 「講義シラバスと卒業論文データを用いた研究室配属支援システム」、  
第 23 回電子情報通信学会九州支部学生会講演会、D-13 (2015)  
著者：高橋和生、榎 俊孝、若原俊彦
27. 「徘徊高齢者検索メールを用いた人物特徴生成による検索支援システム」、  
第 23 回電子情報通信学会九州支部学生会講演会、D-12 (2015)  
著者：小山泰希、高橋和生、榎 俊孝、若原俊彦
28. 「平成 26 年度 FIZOPEN 試行事業ふくおか IT Workouts 報告一大野城市：観光・商店街の活性化を目指すコミュニティシステムの開発」、福岡工業大学 FD Annual Report、Vol.5、pp.28-32 (2015)  
著者：若原俊彦、榎 俊孝、大塚信吾、本田泰希
29. 「I-Scover および Wikipedia を用いた専門用語辞書ツール構築法」、  
電子情報通信学会技術研究報告、Vol.115、No.138、LOIS2015-11、pp.13-17  
(2015)  
著者：榎 俊孝、若原俊彦