

ナレッジグラフ若手の会  
**ISWC2022サーベイ会**

2022/12/07 オンライン

# 担当者一覧

順	セッション	担当者
1	1A: KG CONSTRUCTION & USE	eiichi_sunagawa
2	1B: KG COMPLETENESS	yu_asano
3	2A: EMBEDDINGS (I)	ugai
4	2B: ONTOLOGIES & KGS FOR MANUFACTURING	kgyanagi
5	3A: BIOMEDICAL ONTOLOGIES & KGS	Yasunori
6	3B: LINKED DATA	森田研究室
7	4A: NLP (I)	KozakiKouji
8	4B: ONTOLOGIES (I)	福田研究室
9	5A: EMBEDDINGS (II)	ugai
10	5B: ONTOLOGIES (II)	S. Egami

順	セッション	担当者
11	6A: NLP (II)	KozakiKouji
12	6B: QUERYING	koji6_fujiwara
13	7A: KG REFINEMENT & SCALE	Yasunori
14	7B: TEMPORAL KGS	S. Egami
15	8A: LEARNING REPRESENTATIONS	kgyanagi
16	8B: GRAPH VALIDATION	福田研究室
17	INDUSTRY TRACK 1	Nagano
18	INDUSTRY TRACK 2	Nagano

# **1A: KG CONSTRUCTION & USE**

# Leveraging Knowledge Graph Technologies to Assess Journals and Conferences at Springer Nature

Simone Angioni 1 , Angelo Salatino 2 , Francesco Osborne 2,3 , Aliaksandr Birukou 4 , Diego Reforgiato Recupero 1 , and Enrico Motta 2  
 / 1 University of Cagliari, Italy 2 The Open University, United Kingdom 3 University of Milano Bicocca, Italy 4 Springer-Verlag GmbH, Germany

- 一言でいうと

AIDA Dashboard 2.0というコンピュータサイエンスの学術会議や論文誌の検索・比較・分析ツールをリリース

- 動機

データアナリストやエディタは、大量の文献を見つづ  
会議・論文誌について調査や評価を行わなければならず大変

- 手法

Computer Science Ontology、DBpediaなど用いて、トピック抽出、  
筆者所属、工学的な分類を行い、AIDA KG(RDF)を生成。  
学会や論文誌を所定のメトリックで評価。

- 背景

AIDA 1.0と比較し、(1)学術誌も分析可能、(2)ランキング機能、  
(3)新しい研究分野の分類、(4)SpringerNatureとのシステム連携

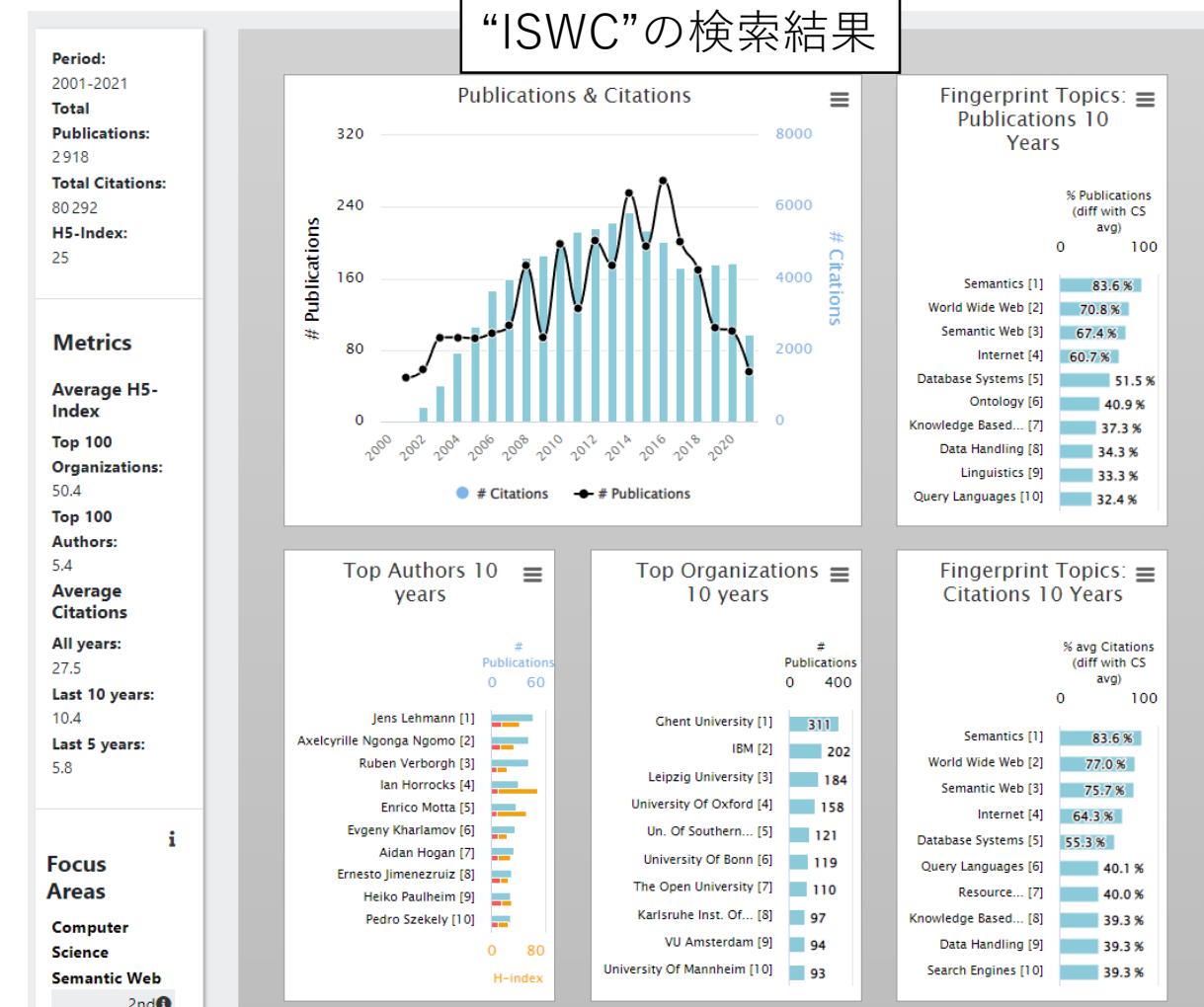
- 結果

2500万件の学術文献から15億以上のRDF三重構造。  
利用実績

一般公開中。事前に、SEMANTICS、EKAW、DL4KG、  
Text2KG、Sci-K の開催委員にリリース

- 学んだ教訓やベストプラクティス

編集者や研究者で評価したところ、関連会議のタブが特に好評。

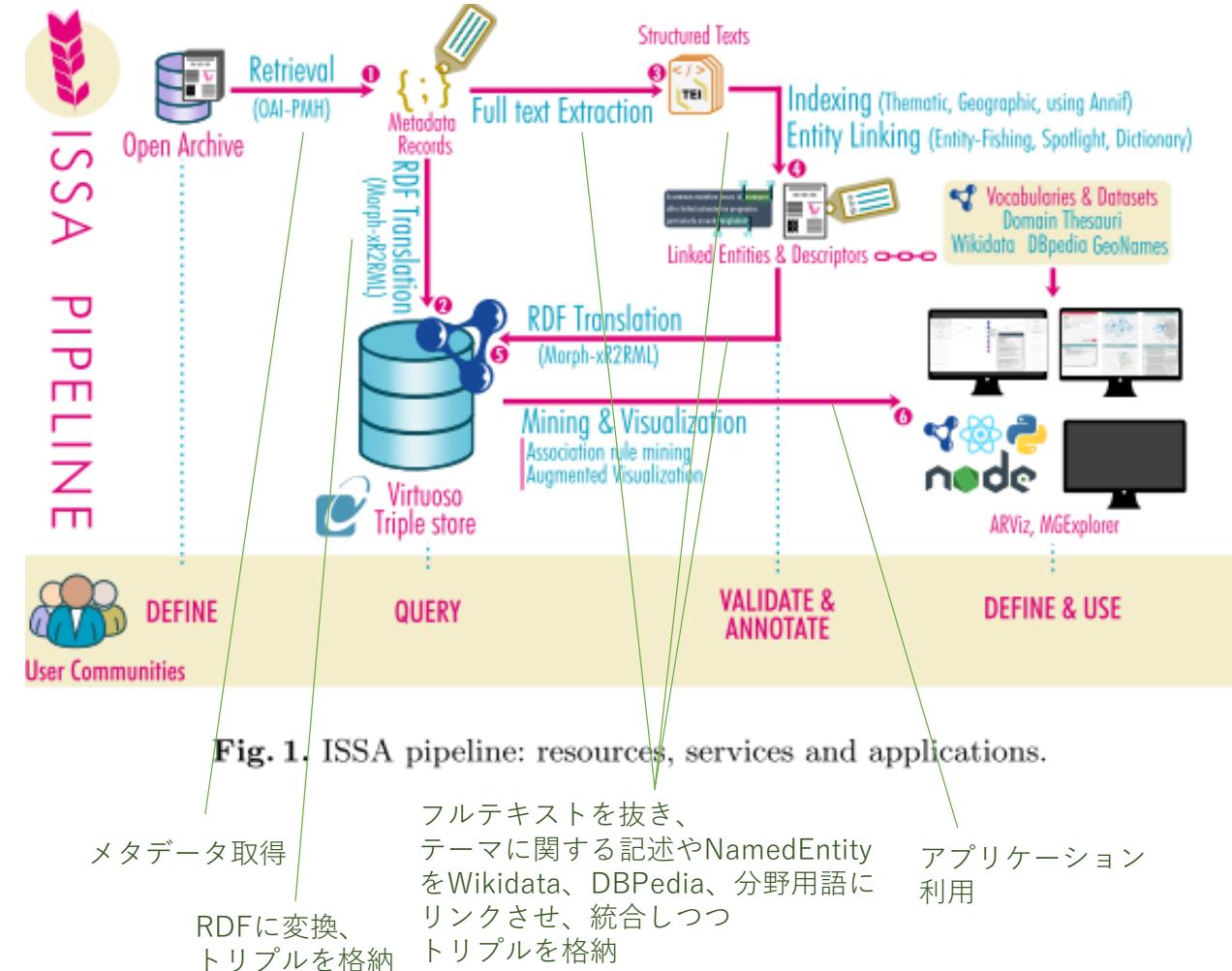


AIDAリンク↓  
<https://aida.kmi.open.ac.uk/dashboard>

# ISSA: Generic Pipeline, Knowledge Model and Visualization tools to Help Scientists Search and Make Sense of a Scientific Archive

Anne Toulet, Franck Michel, Anna Bobasheva, Aline Menin, Sébastien Dupré, Marie-Claude Deboin, Marco Winckler, and Andon Tchetchmedjiev  
 / CIRAD, University Côte d'Azur, Inria, CNRS, Euromov Digital Health in Motion, Univ Montpellier, IMT Mines Alès, Alès, France

- 一言でいうと  
文献から生成したKGを使い共著関係や共起トピックで検索
- Resource Type  
不明 (KG?)
- 動機  
キーワード検索では論文間のつながりを把握できない
- リソースの設計方針  
Wikidata, Dbpedia, GeoNames, AGROVOCなど外部知識に紐づける
- 新規性  
テーマの相関や共著関係やトピックの共起などで論文検索可能
- リソースの再現性／活用実績  
CIRADの11万件以上のリソース（本、記事、論文）に適用
- 品質  
今後の課題。特にテキスト分類。
- リソースの可用性  
データセットおよびパイプラインのソースは公開済み
- 次にすべきこと  
データの品質評価、ツールの使い勝手の評価



# CS-KG: A Large-Scale Knowledge Graph of Research Entities and Claims in Computer Science

Danilo Dessim1, Francesco Osborne2,3, Diego Reforgiato Recupero1, Davide Buscaldi4, and Enrico Motta2

/ 1 University of Cagliari, Italy 2 The Open University, UK 3 University of Milano Bicocca, Italy 4 LIPN, CNRS, Université Sorbonne Paris Nord, France

- 一言でいうと  
タスク、手法、メトリクスなども含むComputer Scienceの大規模KG

- Resource Type  
Knowledge Graph

- 動機  
既存のCS-KGは、文献情報の要約が人依存／分野限定されるため、論文数が少ない

- リソースの設計方針  
タスク、手法、メトリクスなども包括的に含む形で、670万件の論文から3億5000万のトリプルを自動生成。

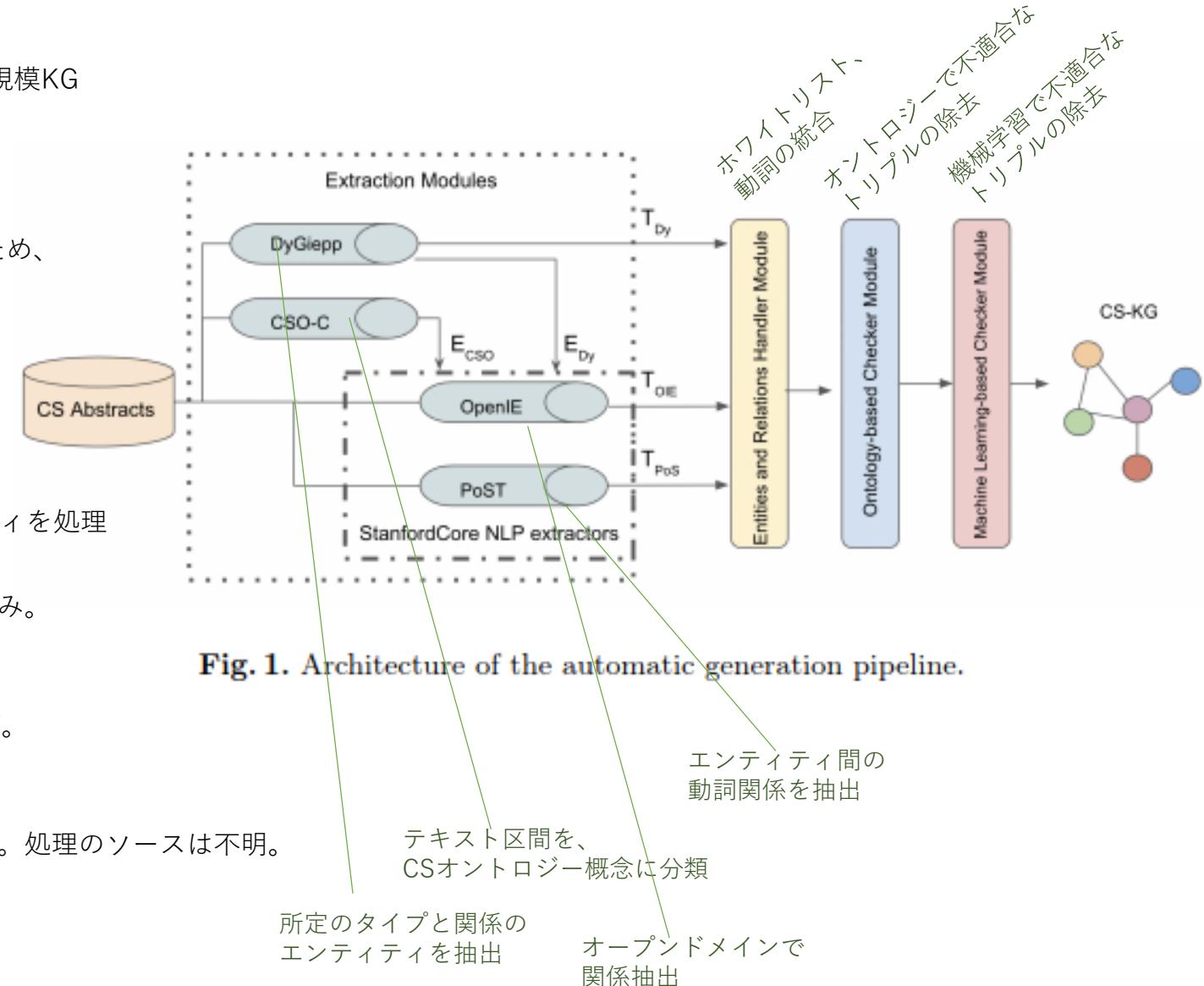
- 新規性  
規模として過去にない。  
2020年度のAI-KG(33万件の論文)で課題だった同一エンティティを処理

- リソースの再現性／活用実績  
MS Academic Graph、OpenAlexを利用して対象論文を絞り込み。  
実績は不明。

- 品質  
研究者により、生成したトリプルに正誤のラベルを付け、評価。  
CS-KGパイプラインによるトリプル生成のF値は0.76。

- リソースの可用性  
リソースはWeb (<https://scholkg.kmi.open.ac.uk/>) にて公開。処理のソースは不明。

- 次にすべきこと  
ユーザが修正したり、フィードバックを残す枠組みの開発。



# **1B: KG COMPLETENESS**

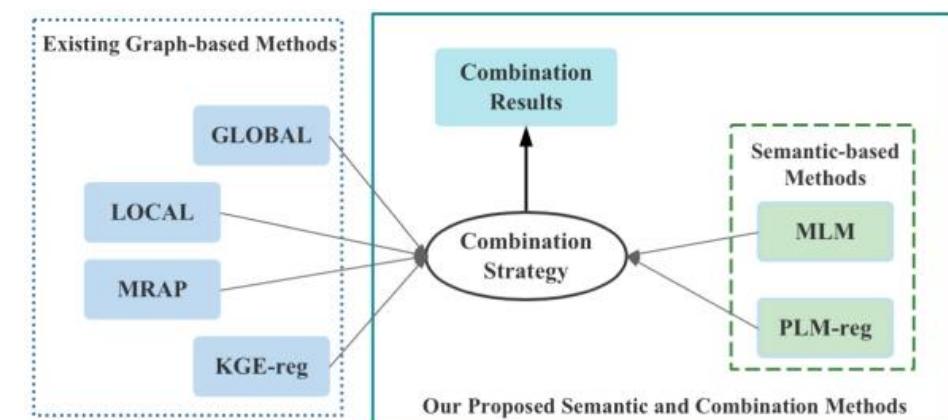
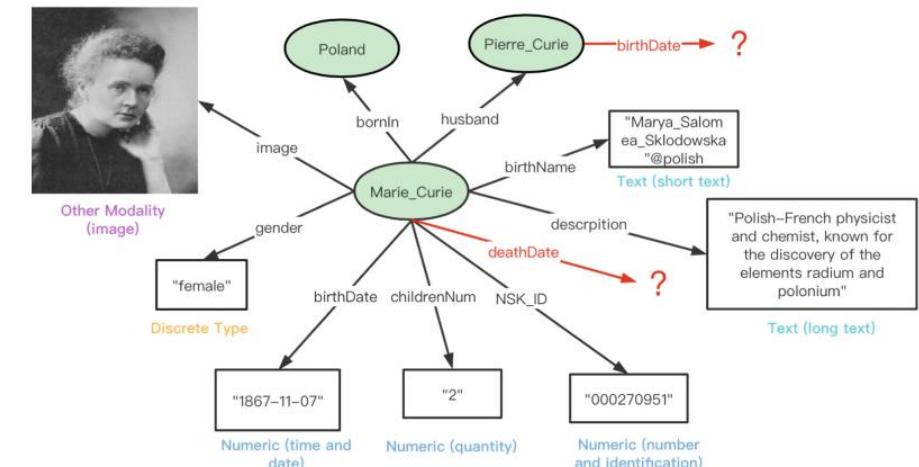
# Introducing Semantic Information for Numerical Attribute Prediction over Knowledge Graphs

Bingcong Xue<sup>1</sup>, Yanzeng Li<sup>1,2</sup>, and Lei Zou<sup>1,2</sup>

<sup>1</sup> Peking University, Beijing, China

<sup>2</sup> Beijing Institute for General Artificial Intelligence (BIGAI), Beijing, China

- 一言でいうと  
言語モデルを用いたリテラルを取る関係を予測する手法を提案し、有効性を示した。
- 動機
  - ・ 知識グラフ補完の研究は、欠けている関係を推論するリンクの予測タスクとして行われてきたが、リテラルは注目されてこなかった
- 手法
  - ・ 言語モデルを利用した手法を提案する。また、明示的な構造的な情報 (GLOBAL、LOCAL、MRAP、KGE-reg) と暗黙的な意味的な情報 (MLM、PML-reg) の両方を活用するための効果的な組み合わせ方も提案する。
- 背景
  - ・ 身長、年齢、誕生日などの数値属性は、他のリテラルと異なり、計算や推定が可能であるため、予測される可能性が高い
- 結果
  - ・ FB15KとYAGO15Kで評価し、属性によって適切なモデルは異なる。全体的に MRAP、PML-regが効果的。また、組み合わせが有効であることを示した。
- 考察
  - ・ 不明
- 課題
  - ・ KGをテキストへの変換、他タスクへの適用、数値属性についての検討



# An Analysis of Content Gaps versus User Needs in the Wikidata Knowledge Graph

David Abian<sup>1</sup>, Albert Merono-Penuela<sup>1</sup>, and Elena Simperl<sup>1</sup>

<sup>1</sup> King's College London, London, UK

- 一言でいうと
  - ・Wikidataのコンテンツギャップ分析により、人間開発指数の低い国はリクエストに従って必要以上に作業をしないことが分かった
- 動機
  - ・知識グラフのコンテンツギャップは、下流タスクのアプリケーションに影響を与える
- 手法
  - ・コンテンツギャップの原因を明らかにする仕組みを提案
    - (i)ピアプロダクションシステムに関する文献調査をもとに、コンテンツギャップのコアタイプを把握する
    - (ii)そのようなギャップのある領域において、知識グラフ上の作業のアンバランスとユーザーの情報ニーズのアンバランスを定量的に比較する
- 背景
  - ・これまでにデータの品質やオントロジーの評価に関する研究が行われてきた
  - ・ユーザニーズの文脈でギャップを定義し、効果的に対処する方法がない
- 結果
  - ・Wikidataにおける性別、更新頻度、地理的・社会経済的なギャップに適用し、2018～2021年の編集指標とWikipediaページビューを比較した
    - ・性別や更新頻度のギャップを見いだせなかったが、例外的に人間開発指数の低い国はリクエストにしたがって必要以上に作業をしないことが分かった
- 考察、課題
  - ・Wikidataの全インスタンスを対象としたコンテンツギャップの分析の継続、下流アプリケーションにおける分析結果の活用

# GraphGNNQ: A Neuro-Symbolic Approach for Query Answering over Incomplete Knowledge Graphs

Maximilian Pflueger<sup>1</sup>, David J. Tena Cucala<sup>1</sup>, and Egor V. Kostylev<sup>2</sup>

<sup>1</sup> University of Oxford, UK, <sup>2</sup> University of Oslo, Oslo

- 一言でいうと  
クエリに基づいて不完全なKGを補強し、クエリの答えを生成する手法を提案し、性能向上することを示した
  - 動機  
symbolicな問合せエンジンは、実世界のKGが不完全であるために、含まれていない答えを生成することはできない。
  - 手法  
不完全なKGに対する問合せのためのneuro-symbolicな帰納的なアプローチを提案  
Step1. クエリの一部にマッチするKGを表す事実で、入力するKGを記号的に補強する。  
Step2. 補強されたナレッジグラフに、Relational Graph Convolutional Networkを一般化した手法により、クエリの答えを生成する
  - 背景  
含まれていない答えを得るために、KGとクエリを低次元のベクトルに埋め込むMLベースの手法があるが、定数などのKGエンティティが学習時と識別時で異なるような帰納的な場合には適用できない。
  - 結果  
ベンチマーク (WATDIV、FB15k237) で評価し、提案手法が性能向上に寄与することを示した。
  - 考察  
不明
  - 課題  
より表現力豊かな問合せをサポートするための手法の拡張
- 例. almaMater (母校) を補完
- $q(x)$ :

$\mathcal{K}$  and  $\mathcal{K}^*$ :
- 出典：<https://link.springer.com/content/pdf/10.1007/978-3-031-19433-7.pdf>

# 2A: EMBEDDINGS (I)

# EaT-PIM: Substituting Entities in Procedural Instructions using Flow Graphs and Embeddings

Sola Shirai/Rensselaer Polytechnic Institute and Hyeongsik Kim/Robert Bosch LLC

- 一言でいうと

料理レシピのような特定分野の指示書から情報を抽出し、フローグラフに変換する手法(EaT-PIM)を提案する。フローグラフに出現するエンティティやアクションの埋め込みを学習し、その埋め込みを利用して、説明書に加えられる可能性のある変更を特定するアプローチを示す。

- 動機

料理で言えば、レシピの材料を置き換えるようなものだ。置き換え可能な材料は数多く存在し、その結果、元の料理に「近い」ものがで  
きる。しかし、どのような方法で作り直せばよいかを判断するのは難しい

- 手法

- (1) ドメインオントロジーと依存関係解析ツールを活用し、命令文からルールベースのフローグラフを生成する方法を提案
- (2) フローグラフのための新しいグラフ埋め込み方法を提案

- 背景

- 結果

Flow Graph Generation: precision : 0.638, recall : 0.566, F1 score : 0.600

**Table 2.** Examples of top ranked substitutions in two recipes.

- 考察

Recipe	Target ingredient	Top 3 substitutes
Pork marinade	Pork	Boneless pork, Rib, Pork loin roast
Mashed potatoes	Red potato	Dried thyme, all purpose flour, Chicken

- 課題

埋め込みにおけるより複雑な戦略の検討

EaT-PIMの異なるドメインの命令への適用

**Table 1.** Results for ranking missing ingredients in recipe flow graphs.

Model	MRR	HITS@3	HITS@5	HITS@10
COOC	0.132	0.138	0.189	0.281
DistMult	0.012	0.012	0.015	0.021
ComplEx	0.017	0.018	0.023	0.032
RotatE	0.118	0.120	0.163	0.242
TransE	0.151	0.158	0.211	0.301
GNN (Nathani et al.)	0.068	0.068	0.88	0.124
TransE (flow graph path)	0.172	0.177	0.206	0.254
EaT-PIM (ours)	<b>0.286</b>	<b>0.355</b>	<b>0.437</b>	<b>0.520</b>

# Heterogeneous Graph Neural Network with Hypernetworks for Knowledge Graph Embedding

Xiyang Liu, Tong Zhu, Huobin Tan, and Richong Zhang/Beihang University

- 一言でいうと

関係間の一般的な意味をモデル化して必要なパラメータを生成する、ハイパーネットワーク (HGNN) を組み込んだ新しいヘテロジニアスグラフニューラルネットワークを提案する。

- 動機

HGNNを用いて知識グラフを処理しようとすると、型に関するパラメータが爆発的に増加するため、HGNNはエッジの型が少ないグラフにしか適用できない

- 手法

- 背景

- 結果

- 考察

- 提案するハイパーネットワークはパラメータ数を大幅に削減することができ GPUのメモリ使用量を削減できる

- 課題

- ハイパーネットワークのバリエーションを増やす
- ハイパーネットワークをより高度なヘテロジニアスアーキテクチャと組み合わせ、KGE学習を使えるものにする

Model	FB15k-237						WN18RR					
	MRR	MR	Hits@N				MRR	MR	Hits@N			MRR
			1	3	10	1			3	10		
TransE [△]	0.313	-	0.221	0.347	0.497	0.228	-	0.053	0.368	0.520	-	-
DistMult [△]	0.343	-	0.250	0.378	0.531	0.452	-	0.413	0.466	0.530	-	-
DualE	0.330	-	0.237	0.363	0.518	0.482	-	0.440	0.500	0.561	-	-
ConvE [△]	0.339	-	0.248	0.369	0.521	0.442	-	0.411	0.451	0.504	-	-
ConvR	0.350	-	0.261	0.385	0.528	0.475	-	0.443	0.489	0.537	-	-
HypER	0.341	250	0.252	0.376	0.520	0.465	5798	0.436	0.477	0.522	-	-
CoPER [†]	0.320	390	0.234	0.351	0.491	0.442	5315	0.418	0.450	0.487	-	-
R-GCN	0.248	-	0.153	0.258	0.414	-	-	-	-	-	-	-
SACN	0.350	-	0.260	0.390	0.540	0.470	-	0.430	0.480	0.540	-	-
VR-GCN	0.248	-	0.159	0.272	0.432	-	-	-	-	-	-	-
A2N	0.317	-	0.232	0.348	0.486	0.450	-	0.420	0.460	0.510	-	-
KBGAT [◊]	0.157	270	-	-	0.331	0.412	1921	-	-	-	0.554	-
CompGCN	0.355	197	0.264	0.390	0.535	0.479	3533	0.443	0.494	0.546	-	-
HRAN	0.355	156	0.263	0.390	0.541	0.479	2113	0.450	0.494	0.542	-	-
<b>HKGN</b>	<b>0.365</b>	171	<b>0.272</b>	<b>0.402</b>	<b>0.552</b>	<b>0.487</b>	2468	0.448	<b>0.505</b>	<b>0.561</b>	-	-

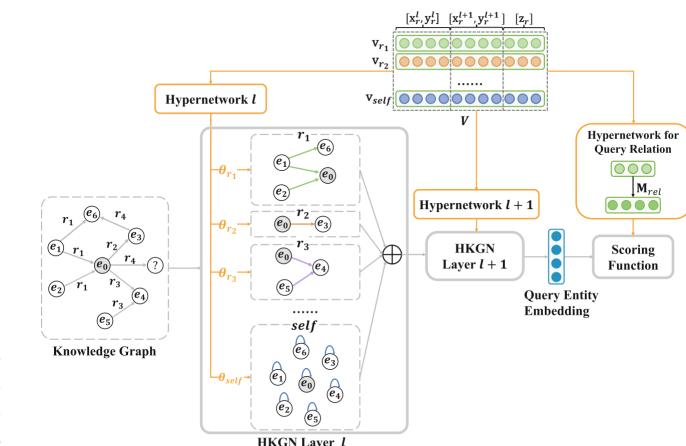


Fig. 1. An overview of HKGN. HKGN consists of two modules: the hypernetworks module used to obtain weights and the primary HGNN module utilized to process the multi-relational knowledge graph.

# $\mu$ KG: A Library for Multi-source Knowledge Graph Embeddings and Applications

Xindi Luo, Zequn Sun , and Wei Hu/Nanjing University

- 一言でいうと  
知識グラフを用いた表現学習のためのオープンソースのPythonライブラリ
- Resource Type
  - Python library
- 動機
 

近年、エンティティアライメントを用いた複数ソースのKGからの学習が注目されている。  
このようなKGの埋め込みは、エンティティのタイピングや複数ソースのKGの補完など、様々な下流タスクに有用であることが示されている。
- リソースの設計方針
 

複数ソースの知識グラフに対する共同表現学習、複数の深層学習ライブラリ、複数の埋め込みタスク、複数の並列計算モードをサポート
- 新規性
- リソースの再現性／活用実績
- 品質
- リソースの可用性
- 次にすべきこと
 

より多くのKG埋め込みモデルやマルチソースKGタスクの統合  
実装の継続的な改善

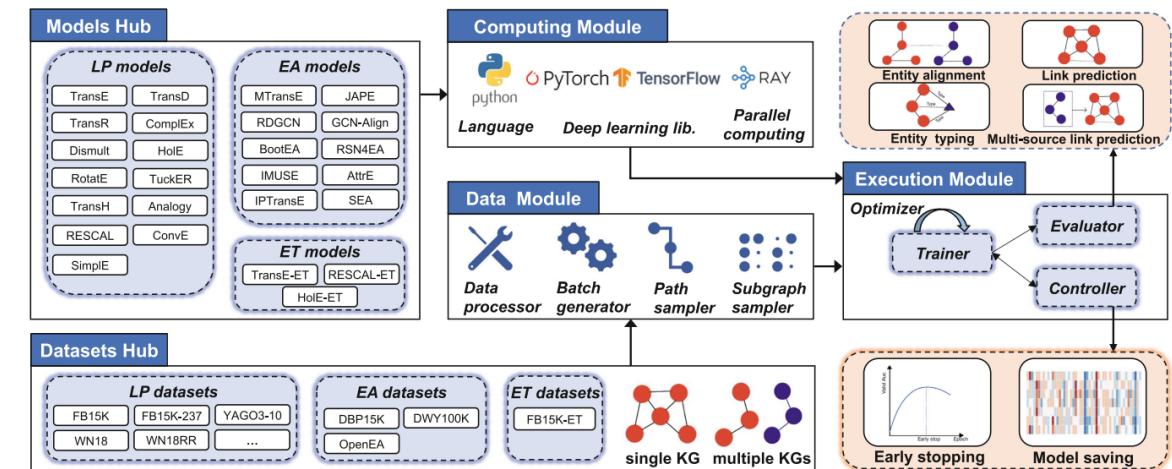


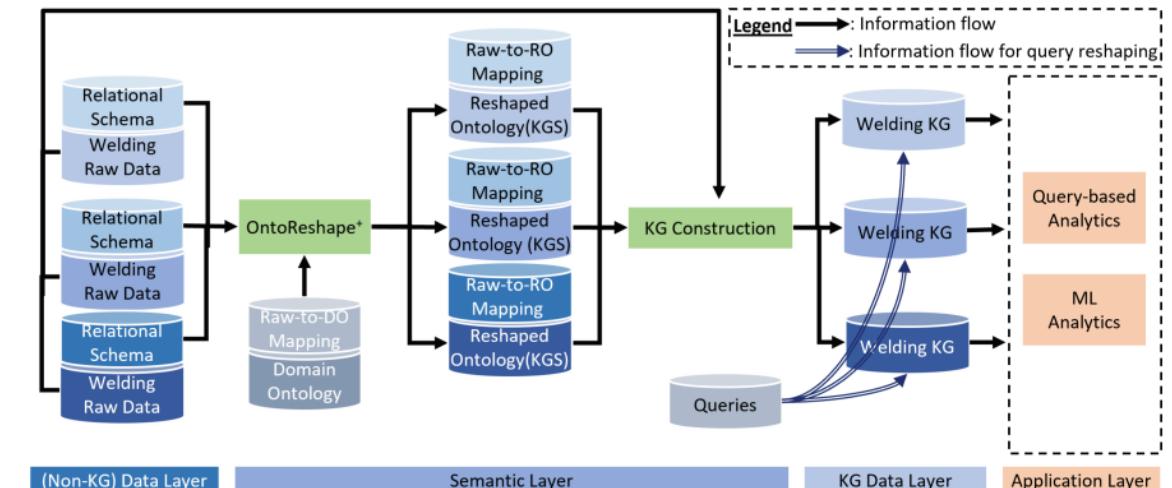
Fig. 1. Framework overview of  $\mu$ KG.

# **2B: ONTOLOGIES & KGS FOR MANUFACTURING**

# Ontology Reshaping System for Knowledge Graph Construction: Applied on Bosch Welding Case

Dongzhuoran Zhou<sup>1,2</sup>, Baifan Zhou<sup>2</sup>, Zhuoxun Zheng<sup>1,3</sup>, Ahmet Soylu<sup>3</sup>, Gong Cheng<sup>4</sup>, Ernesto Jimenez-Ruiz<sup>2,5</sup>, Egor V. Kostylev<sup>2</sup>, and Evgeny Kharlamov<sup>1,2</sup>  
 / <sup>1</sup> Bosch Center for Artificial Intelligence, Renningen, Germany <sup>2</sup> Department of Informatics, University of Oslo, Oslo, Norway <sup>3</sup> Department of Computer Science, Oslo Metropolitan University, Oslo, Norway <sup>4</sup> State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China <sup>5</sup> Department of Computer Science, City University of London, London, UK

- 一言でいうと  
オントロジー再形成システムとアルゴリズム OntoReshape+ を提案
- 動機  
自動KG構築のために、データを特定のKGスキーマにマップし、エンティティとプロパティを構築  
KGスキーマとして、ドメインオントロジーを活用することが多いが、クラスがデータにマッピングされていないために、空白ノードが作成されたり、ディープな構造が作成されたりする問題がある
- 手法  
オントロジー再形成システムとアルゴリズム OntoReshape+ を提案
  - ドメイン知識も充分にカバーしながら、データを完全にカバーするKGスキーマを生成
- 背景  
既存研究は requirements ( R1 Data Coverage, R2 Knowledge Coverage, R3 User-friendliness, R4 System Efficiency) に対して不充分
- 結果  
ユーザ調査： アンケート結果において、R1, R3, R4は高い評価、 R2(Knowledge Coverage)は比較的良好だが改善の余地あり  
4つのベースラインに対する評価： ベースラインよりもバランスの良い結果を示し、バランスの良いOntoReshapeをR2で上回る結果
- 利用実績  
現在、ボッシュの評価環境に導入
- 学んだ教訓やベストプラクティス  
(不明)



**Fig. 4.** An architectural overview of our KG solution. KG: knowledge graph. KGS: KG schema.

# SeLoC-ML: Semantic Low-Code Engineering for Machine Learning Applications in Industrial IoT

Haoyu Ren<sup>1,3</sup>, Kirill Dorofeev<sup>1</sup>, Darko Anicic<sup>1</sup>, Youssef Hammad<sup>1,3</sup>, Roland Eckl<sup>2</sup>, and Thomas A. Runkler<sup>1,3</sup> /

<sup>1</sup> Siemens AG, Otto-Hahn-Ring 6, 81739 Munich, Germany <sup>2</sup> Siemens AG, Siemenspromenade 1, 91058 Erlangen, Germany

<sup>3</sup> Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany

- 一言でいうと  
IIoT での ML アプリケーションの迅速な開発をサポートするローコード プラットフォーム上に構築されたフレームワーク (SeLoC-ML) の提案

## • 動機

MLを産業用デバイスに導入

インダストリアル IoT (IIoT) で ML アプリケーションを開発すると、ハードウェアの不均一性、ML モデルの標準化されていない表現、デバイスと ML モデルの互換性の問題、遅いアプリケーション開発など、さまざまな課題が生じる

## • 手法

SeLoC-ML: セマンティック Web テクノロジを活用して IIoT での ML アプリケーションの迅速な開発をサポートするローコード プラットフォーム上に構築されたフレームワーク

## • 背景

Thing Description(TD) ontology: IoTデバイスのメタデータとインターフェースを指定  
(W3C WoTワーキンググループによって開発)

iotschema.org : IoT ドメインのセマンティックモデル、接続されたモノの間でデータを enrichするために使用

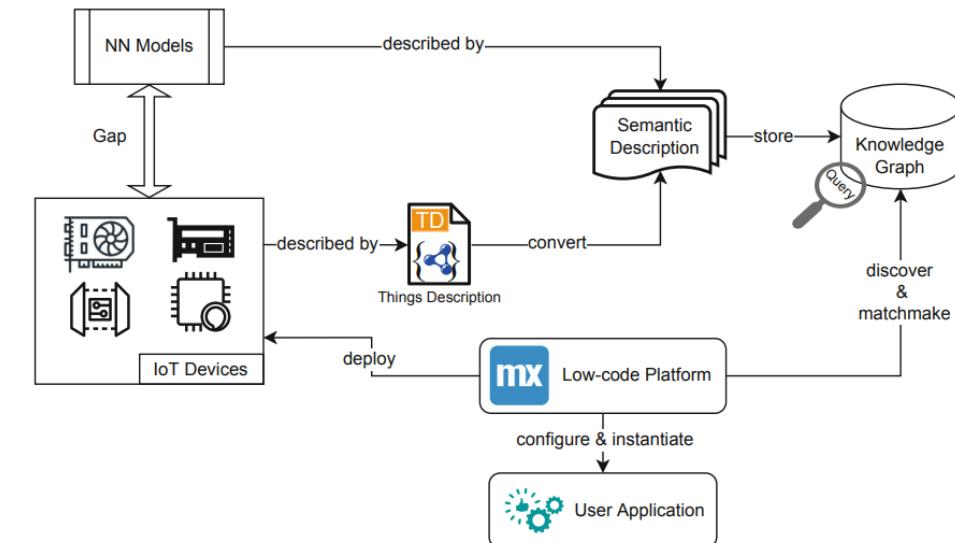
## • 結果

産業用 ML 分類のケース スタディで従来のアプローチと比較して、  
エンジニアリング作業を少なくとも 3 分の 1 に削減

## • 利用実績

産業用 ML 分類のケース スタディを紹介  
コードを共有(<https://github.com/Haoyu-R/SeLoC-ML>)

## • 学んだ教訓やベストプラクティス (不明)



**Fig. 1.** Framework architecture of SeLoC-ML.

# Executable Knowledge Graph for Machine Learning: A Bosch Case for Welding Monitoring

Zhuoxun Zheng<sup>1,2</sup>, Baifan Zhou<sup>3</sup>, Dongzhuoran Zhou<sup>1,3</sup>, Xianda Zheng<sup>4</sup>, Gong Cheng<sup>5</sup>, Ahmet Soylu<sup>2</sup>, and Evgeny Kharlamov<sup>1,3</sup> / <sup>1</sup> Bosch Center for Artificial Intelligence, Renningen, Germany <sup>2</sup> Department of Computer Science, Oslo Metropolitan University, Oslo, Norway <sup>3</sup> SIRIUS Centre, University of Oslo, Oslo, Norway <sup>4</sup> State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China <sup>5</sup> School of Computer Science and Engineering, Southeast University, Nanjing, China

- 一言でいうと

機械学習のための、Executable KGの概念とシステムを提案

- 動機

ML を含むデータ分析は、現代の産業で生産データから洞察を引き出すために不可欠  
産業用 ML は次の影響を受ける

- 非 ML 専門家に対する ML の透明性の低さ
- 見直しや理解のための ML プラクティスの不十分で統一されていない説明
- 特定のアプリケーションに合わせて調整された ML ソリューションのアドホックな方法（再利用性に影響）

- 手法

Executable KGの概念とシステムを提案

- MLの知識とソリューションをKGにエンコード
- 再利用可能かつモジュール化された形式で実行可能なスクリプトに変換

- 背景

プログラミング言語でコードとコード間の関係をKGで表現する取り組みはあるが、  
コード間のデータの情報を流れをエッジとして扱い、意味的制約を定義するなどは議論なし。  
業界での機械学習の実践を説明することに専念した研究はほとんどない。

- 結果

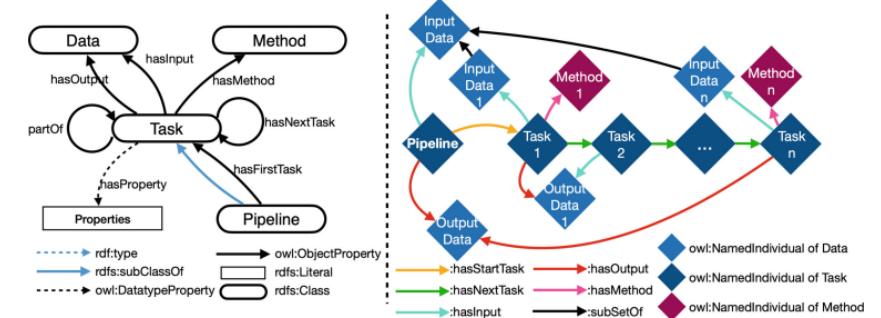
ユーザー調査、ワークショップ、スケーラビリティ評価など、ボッシュでの影響力のある  
産業用ユースケースを使用して、システムを広範囲に評価  
ML の専門家でなくとも、ML メソッドについて議論、カスタマイズ、および再利用のため  
のユーザーフレンドリーな方法を提供することを示す。

- 利用実績

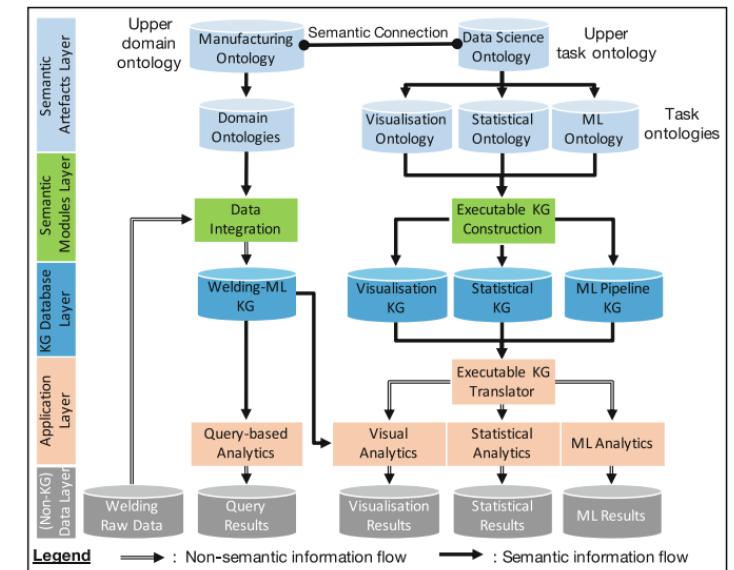
Boschの環境(Use Case: Bosch Welding Monitoring)

- 学んだ教訓やベストプラクティス

多くのユーザーが機械学習関連の知識とソリューションに非常に关心を持っているが、その知識を学び機械学習を実践する良い出発点が無かった。  
ユーザは提案システムを高く評価し、特に GUI と、ハイパーパラメータの調整について多くの改善点を提案。



**Fig. 2.** Framework of executable knowledge graph



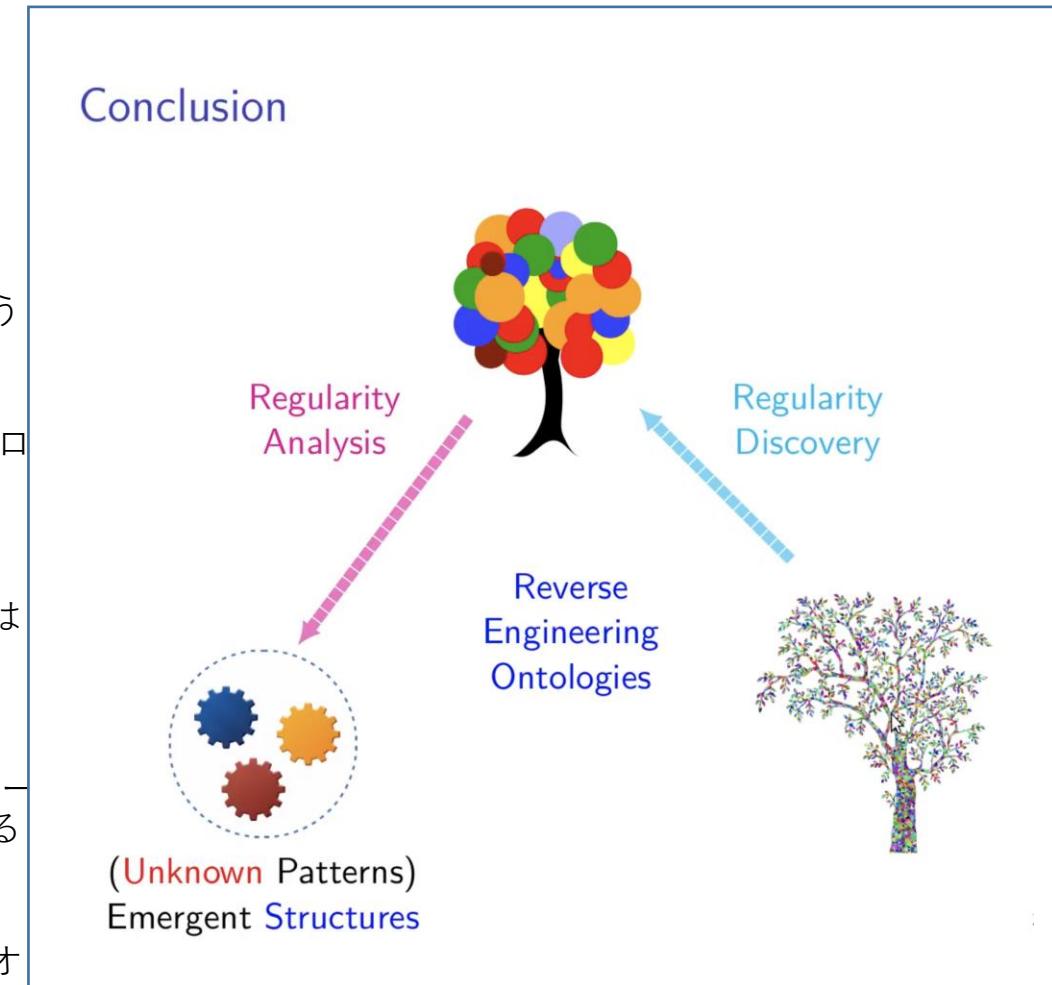
**Fig. 3.** An architectural overview of our knowledge graph solution

# **3A: BIOMEDICAL ONTOLOGIES & KGS**

# A Survey of Syntactic Modelling Structures in Biomedical Ontologies

Christian Kindermann and Martin Skjæveland / University of Oslo

- 一言でいうと  
これまでに発表されている多種多様なオントロジーの間に共通する統語構造の有無やその複雑さを調査した
- 動機  
オントロジーは次々に発表されているが、そこに共通する統語構造があれば、オントロジーを構築する際のベストプラクティスとして利用可能だろう
- 手法  
多くのオントロジーを擁するレポジトリであるBioPortalが収載するオントロジー群を対象としてサイズや複雑性、深さなどの観点から統語構造を調査
- 背景  
オントロジーは対象分野が多岐にわたるが、そこにある基本的な統語構造はオントロジー横断的に標準的な方法で調査可能
- 結果  
入れ子のないSubClassOf関係と少数の公理から成る単純構造のオントロジーが横断的に多い一方で、とても複雑な構造をもつオントロジーが散見される
- 考察  
オントロジーを調査する研究は多くあるが、その構造を、公開されているオントロジーを対象としたリバースエンジニアリングで調査した試みはない。
- 課題  
得られた統語構造の根拠を調査する。



# Machine Learning-Friendly Biomedical Datasets for Equivalence and Subsumption Ontology Matching

Yuan He, Jiaoyan Chen, Hang Dong, Ernesto Jimenez-Ruiz, Ali Hadian and Ian Horrocks / University of Oxford

- 一言でいうと

オントロジーマッチング技術評価のためのデータセットを構築し、Ontology Alignment Evaluation Initiative (OAEI)のBio-ML TRACKデータとして公開した。

- Resource Type

Ontology Matching Dataset

- 動機

オントロジーマッチング技術はたくさん発表されているが、包摂関係のマッピングや最適でないマッピング、機械学習を用いたシステムの評価が不十分。

- リソースの設計方針

上記の動機に基いて、生物医学系のオントロジーマッピングデータを5つ用意した。  
元となっているデータはMondoとUMLS。

- 新規性

同値関係と包摂関係のマッピングデータを用意し、対象オントロジーが取扱う分野に関する部分のみを取出す枝狩手法を提案、そして機械学習システム向けのデータを用意した。

- リソースの再現性／活用実績

OAEIのデータセットとしてオープンライセンスの下で公開した。

- 品質

人手によりキュレーションされたマッピングデータを公開している他、説明も充実。  
[https://krr-oxford.github.io/DeepOnto/#/om\\_resources](https://krr-oxford.github.io/DeepOnto/#/om_resources)

- リソースの可用性

DOI: <https://doi.org/10.5281/zenodo.6510086>

- 次にすべきこと

単純なクラスマッチだけではない、制限が含まれたものも含む関係を対象にするなど

**Table 1.** Information of the source ontologies used for creating the OM resources.

Mapping source	Ontology	Ontology source & Version	#Classes
Mondo	OMIM	Mondo <sup>a</sup>	44,729
	ORDO	BioPortal, V3.2	14,886
	NCIT	BioPortal, V18.05d	140,144
	DOID	BioPortal, 2017-11-28	12,498
UMLS	SNOMED	UMLS, US.2021.09.01 <sup>b</sup>	358,222
	FMA	BioPortal, V4.14.0	104,523
	NCIT	BioPortal, V21.02d	163,842

<sup>a</sup>Created from OMIM texts by Mondo's pipeline tool available at: <https://github.com/monarch-initiative/omim>.

<sup>b</sup>Created by the official snomed-owl-toolkit available at: <https://github.com/IHTSDO/snomed-owl-toolkit>, which keeps 350K classes of all the 490K classes in the original SNOMED CT.

## Ontology Matching

- Example of matching T-Box named concepts through equivalence and subsumption relationships.

### FMA Axioms

```
... VestibulocochlearNerve ⊑ CanialNerve
VestibulocochlearNerve ⊑ branch.CochlearNerve
VestibulocochlearNerve ⊑ branch.VestibularNerve
CanialNerve ⊑ NeuralTreeOrgan
```

The diagram illustrates the FMA axioms for the VestibulocochlearNerve and CanialNerve. It shows nodes for NTO, CN, Eb.CN, Eb.VN, VN, NS, CNS, PNHN, and VNS. Solid lines represent subsumption relationships: NTO is a superconcept of CN, which is a superconcept of both Eb.CN and Eb.VN. Both Eb.CN and Eb.VN are subconcepts of VN. The node VN is also labeled as a subconcept of CNS. There are dashed lines representing equivalence relationships between CNS and NS, and between PNHN and VNS.

### SNOMED Axioms

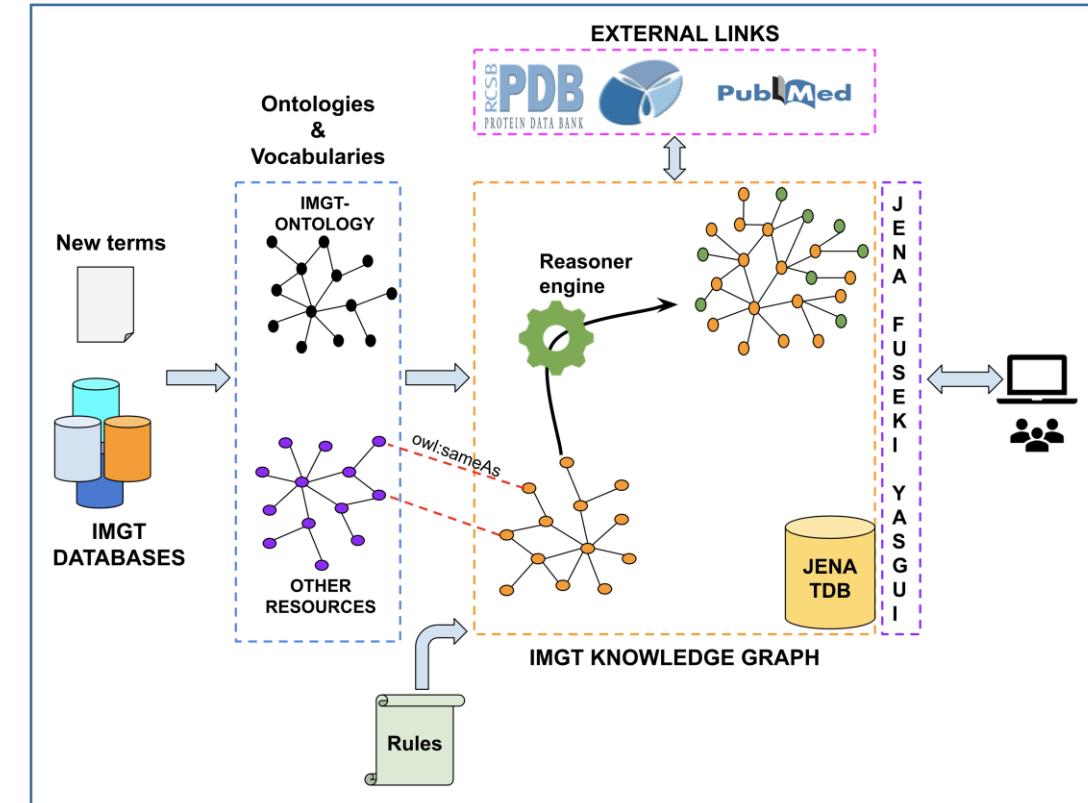
```
... VestibulocochlearNerveStructure ⊑ CanialNerveStructure
VestibulocochlearNerveStructure ⊑ PeripheralNerveOfHeadAndNeck
CanialNerveStructure ⊑ NerveStructure
```

The diagram illustrates the SNOMED axioms for the VestibulocochlearNerveStructure and CanialNerveStructure. It shows nodes for NTO, CN, Eb.CN, Eb.VN, VN, NS, CNS, PNHN, and VNS. Solid lines represent subsumption relationships: NTO is a superconcept of CN, which is a superconcept of both CNS and PNHN. Both CNS and PNHN are subconcepts of VNS. There are dashed lines representing equivalence relationships between CNS and NS, and between PNHN and VNS.

# IMGT-KG: a knowledge graph for immunogenetics

Gaoussou Sanou, Véronique Giudicelli, Nika Abdollahi, Sofia Kossida, Konstantin Todorov and Patrice Duroux / University of Montpellier

- 一言でいうと  
遺伝免疫学に関する知見を知識グラフ化した。
- Resource Type  
RDF dataset
- 動機  
遺伝免疫学に関するデータベースが幾つかあり、研究者はそれらを横断的に検索したい。
- リソースの設計方針  
関連既存オントロジーと、遺伝免疫学に関する新たに構築したオントロジーを組み合わせて、RDFデータを生成し、推論も可能とした。
- 新規性  
遺伝免疫学における統合された知識グラフを始めて構築した。
- リソースの再現性／活用実績  
SPARQLエンドポイントを公開するとともに、データセットをオープンライセンスの下で公開。[\(https://www.imgt.org/imgt-kg/void.ttl\)](https://www.imgt.org/imgt-kg/void.ttl)
- 品質  
利用しやすいように文書を公開するとともに、実用的なクエリ例や、標準的な記載方法であるVOIDに従ったメタデータを提供。
- リソースの可用性  
<https://doi.org/10.5281/zenodo.6511279>
- 次にすべきこと  
抗体に関する情報なども効率的に取得できるように、繋げる関連データベースを増やす。(個人的にはLinked Data化して欲しい)



# 3B: LINKED DATA

# LODChain: Strengthen the Connectivity of your RDF Dataset to the Rest LOD Cloud

Michalis Mountantonakis<sup>1,2</sup>, Yannis Tzitzikas<sup>1,2/1</sup> Institute of Computer Science - FORTH-ICS, Heraklion, Greece, <sup>2</sup> Computer Science Department, University of Crete, Heraklion, Greece

- 一言でいうと  
プロバイダが自分のRDFデータセットと他のデータセットとの間の接続を強化するのを助けることができるサービス**LODChain**の提案
- Resource Type  
Software Service
- 動機  
他のデータセットとの接続は、一般に発見性、閲覧性、問合せ性にとって重要だが困難
- リソースの設計方針  
入力データセットと何百ものLODデータセット間の等価推論を計算
- 新規性  
メタデータベースではなく、データベースの発見サービスである点
- リソースの再現性／活用実績  
エラー検出と修正/接続性分析の提供/高度なクエリ機能/データセット発見
- 品質  
実データセットの接続性を450%から1100%に向上
- リソースの可用性  
<https://doi.org/10.5281/zenodo.6467419>
- 次にすべきこと  
a) LODChainの並列化の検討, b) GUIの改善とデータセット所有者とのユーザビリティ評価, c) 完全に切断されたRDFデータセットに対する接続検索のためのエンティティマッチングのサポート, d) このような接続サービスの評価ベンチマークの作成

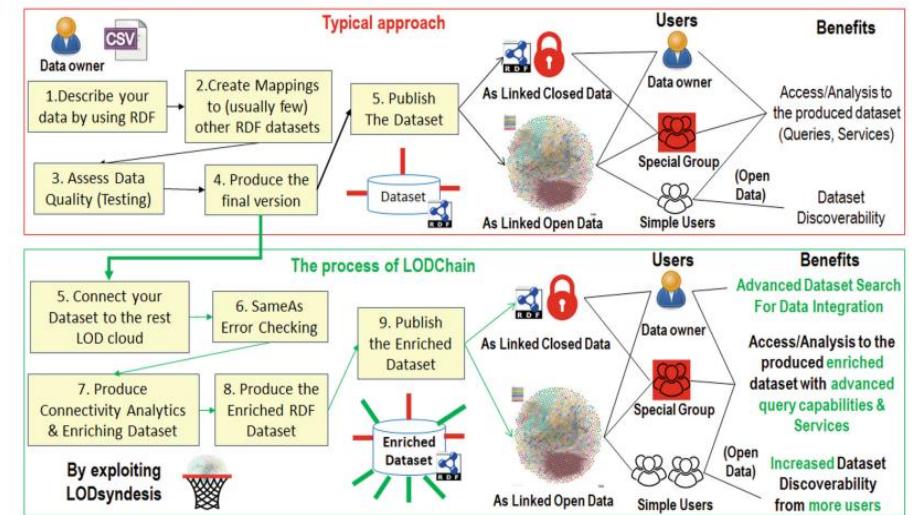


Fig. 1. Typical approach versus the **LODChain** approach for publishing an RDF dataset

# Reproducibility Crisis in the LOD Cloud : Studying the Impact of Ontology Accessibility and Archiving as a Counter Measure

Johannes Frey/Leipzig University, Denis Streitmatter/Leipzig University, Natanael Arndt/eccenca GmbH, Leipzig, Johannes Frey/eccenca GmbH, Leipzig

- 一言でいうと  
LODにおける再現性の危機（科学実験の結果を他の研究者が再現できないこと）対策としてアーカイブの有効性の調査
- 動機  
LODのアクセス性（データへのアクセスのしやすさ、安定性）の問題と、その問題に対してアーカイブは貢献できるか調査が必要
- 手法
  - DBpedia ArchivoとLinked Open Vocabulariesでアーカイブされているデータを分析
    - ・ DBpedia Archivo : web上でowlオントロジーを検出し、自動的に評価及びアーカイブを行う
    - ・ Linked Open Vocabularies : LODの用語を収集、公開している
- 背景  
LODの研究において、再現性はアクセス性に大きく依存する
- 結果  
8か月の調査で、既存のオントロジーの46%で一度以上のアクセス障害を確認
- 考察  
安定してアクセス可能であるアーカイブされたLODは再現性に大きな貢献が可能
- 課題  
DBpedia Archivoの自動収集アルゴリズムに改善の余地

# RMLStreamer-SISO: An RDF Stream Generator from Streaming Heterogeneous Data

Sitt Min Oo<sup>1</sup>, Gerald Haesendonck<sup>1</sup>, Ben De Meester<sup>1</sup> and Anastasia Dimou<sup>2,3/</sup>

<sup>1</sup> IDLab, Department of Electronics and Information Systems, Ghent University – imec, Ghent, Belgium

<sup>2</sup> Department of Computer Science, KULeuven, Leuven, Belgium

<sup>3</sup> AI – Flanders Make, Lommel, Belgium

- 一言でいうと

任意の形式 (JSON、CSV、XML など) の異種データストリームから RDF ストリームを生成する、並列で垂直方向および水平方向にスケーラブルなストリーム処理エンジンである RMLStreamer-SISO を提案

- 動機

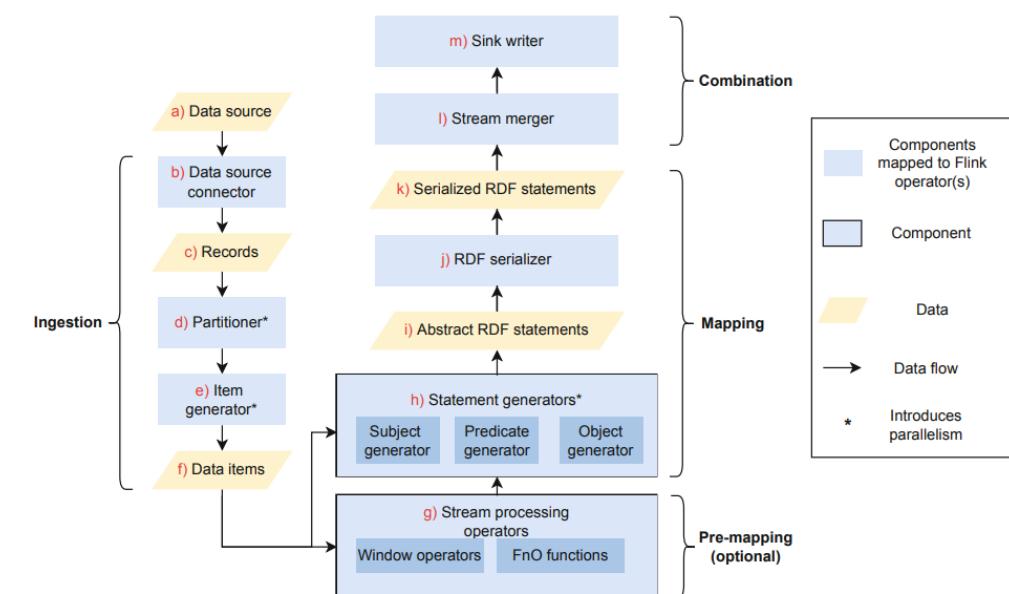
ストリーム処理フレームワークは、より多くの RDF ストリームに対するこれらの成熟した RSP エンジンの需要にもかかわらず、ストリーミングデータソースからの RDF グラフ生成の領域では実質的に使用されていない

- 手法

永続的なビッグデータソースから RDF を生成するための RMLStreamer のアーキテクチャを拡張し、レイテンシを低く保ちながら、データの速度とボリュームが高い異種データストリームから RDF ストリームも生成

- 結果

高速データストリームを処理する際に最先端のツールよりも優れたパフォーマンスを発揮し、低レイテンシを維持しながら処理できるスループットを向上させる。



**Fig. 1.** Workflow of RMLStreamer. Data flows from the *Data Source* at the top through all the components pipeline to the *Sink writer* at the bottom.

## 4A: NLP (I)

# Enhancing Document-level Relation Extraction by Entity Knowledge Injection

Xinyi Wang, Zitao Wang, Weijian Sun and Wei Hu/Nanjing University, Huawei Technologies Co., Ltd.

- 一言でいうと  
KIRE (an entity knowledge injection framework) というドキュメントからの関係抽出 (Fig.1) の手法を提案。
- 動機  
ドキュメントからの関係抽出は様々な推論を行う上で重要なタスク。特に、共参照関係(coreference)に注目しておこなう。
- 手法  
共参照関係(coreference)に注目した自然言語処理の手法と、ドキュメントをリンクした知識グラフ (KG) を用いた手法を組みあわせる (Fig.2)。
- 背景  
関係抽出の必要性。近年、様々なKGが利用可能となった点。
- 結果  
KGとしてWikidataを利用し、2つのベンチマークデータセットで評価。ベールラインの結果からエンコーディングのモデルを改良し、最終結果はF1で60-80程度。
- 考察  
既存の手法と比較して良い性能が認められた。
- 課題  
Wikidata以外の知識源との統合

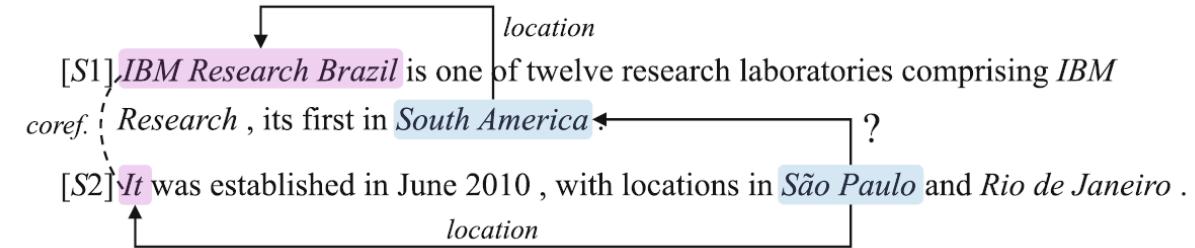


Fig. 1. An example of document-level RE excerpted from [32]

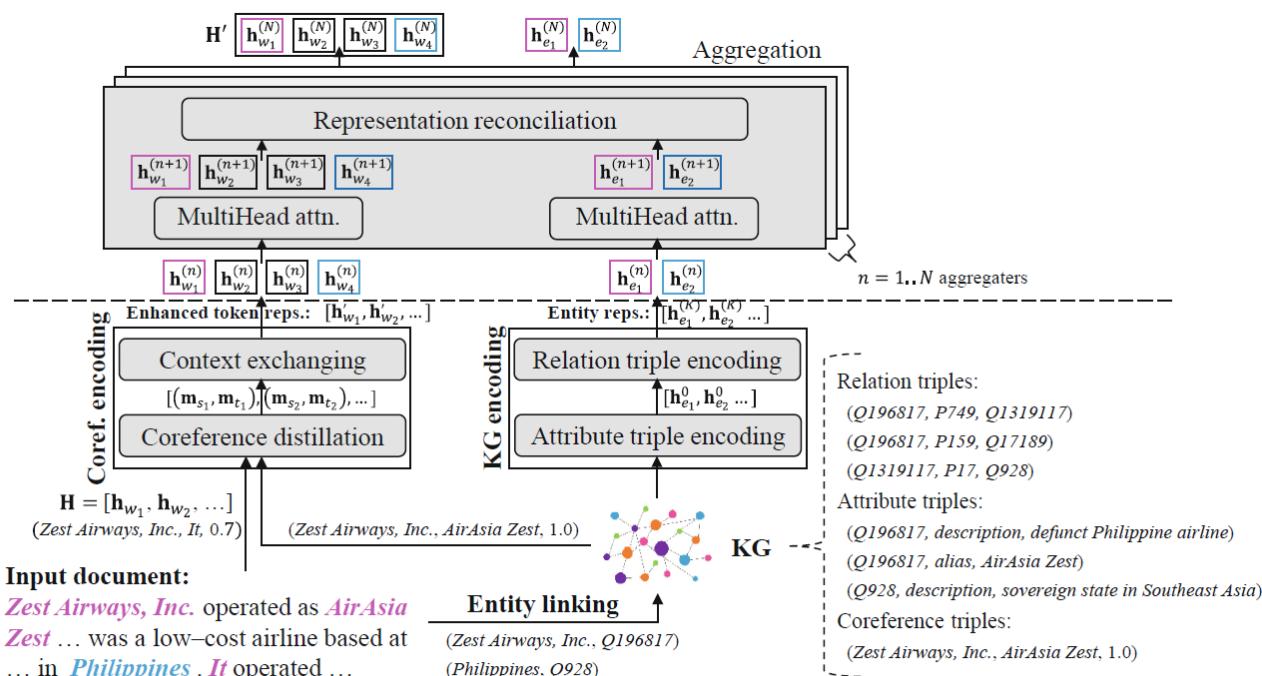
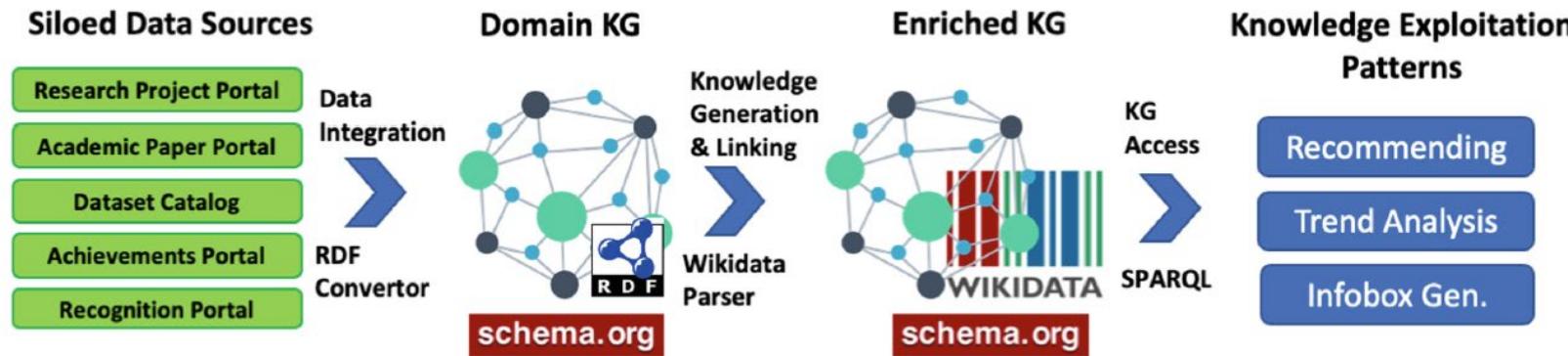


Fig. 2. Architecture of the knowledge injection layer

# Knowledge Graph Induction enabling Recommending and Trend Analysis: A Corporate Research Community Use Case

Nandana Mihindukulasooriya, Mike Sava, Gaetano Rossiello, Md. Faisal Mahbub Chowdhury, Irene Yachbes, Aditya Gidh, Jillian Duckwitz, Kovit Nisar, Michael Santos and Alfio/IBM Research AI,

- 一言でいうと  
アプリケーション開発者が知識グラフ(KG)をより利用しやすくなるようにKGを利用するための共通パターンをAPI化.
- 動機  
研究コミュニティにおいては論文、研究プロジェクト、データセット…など様々なデータを扱って分析する必要がある。  
KGの利用はその解決につながる可能性がある
- 手法  
推薦(Recommending), トレンド分析, Infobox生成のKnowledge Exploitation Patterns (KEP)を開発.
- 背景  
すでに様々なツールを有しているが改良の余地がある。推薦(Recommending)とトレンド分析に注目して行う。
- 結果  
それぞれのKEPについて評価を実験を実施。推薦は30人にユーザ評価, トレンド分析はISWCの論文データを使用, など
- 利用実績  
約6000人の利用者を持つIBM research communityにて使用。
- 学んだ教訓やベストプラクティス  
開発したKEPを通して、KGを研究コミュニティで利用することの有用性が示せた。  
今後、さまざまなタスクにおいて利用することが期待できる。



**Fig. 2.** The knowledge graph induction framework

# RT-KGD: Relation Transition Aware Knowledge-Grounded Dialogue Generation

Kexin Wang, Zhixu Li, Jiaan Wang, Jianfeng Qu, Ying He, An Liu and Lei Zhao/Soochow University, Fudan University, IFLYTEK Research,

- 一言でいうと

関係性の遷移を考慮した、知識に基づく対話生成モデル  
**RT-KGD:**の提案.

- 動機

既存の対話システムは対話に出現したEntityのみを考慮して知識グラフ (KG) 上の関係の重要性を考慮していない.

- 手法

対話における関係性の遷移をMulti-turn Heterogeneous Knowledge Transition(MHKT) Pathとして管理し利用.  
MHKT-pathによる構造化知識と、対話の流れを統合して利用する.

- 背景

対話生成におけるEntityの関係の重要性.

- 結果

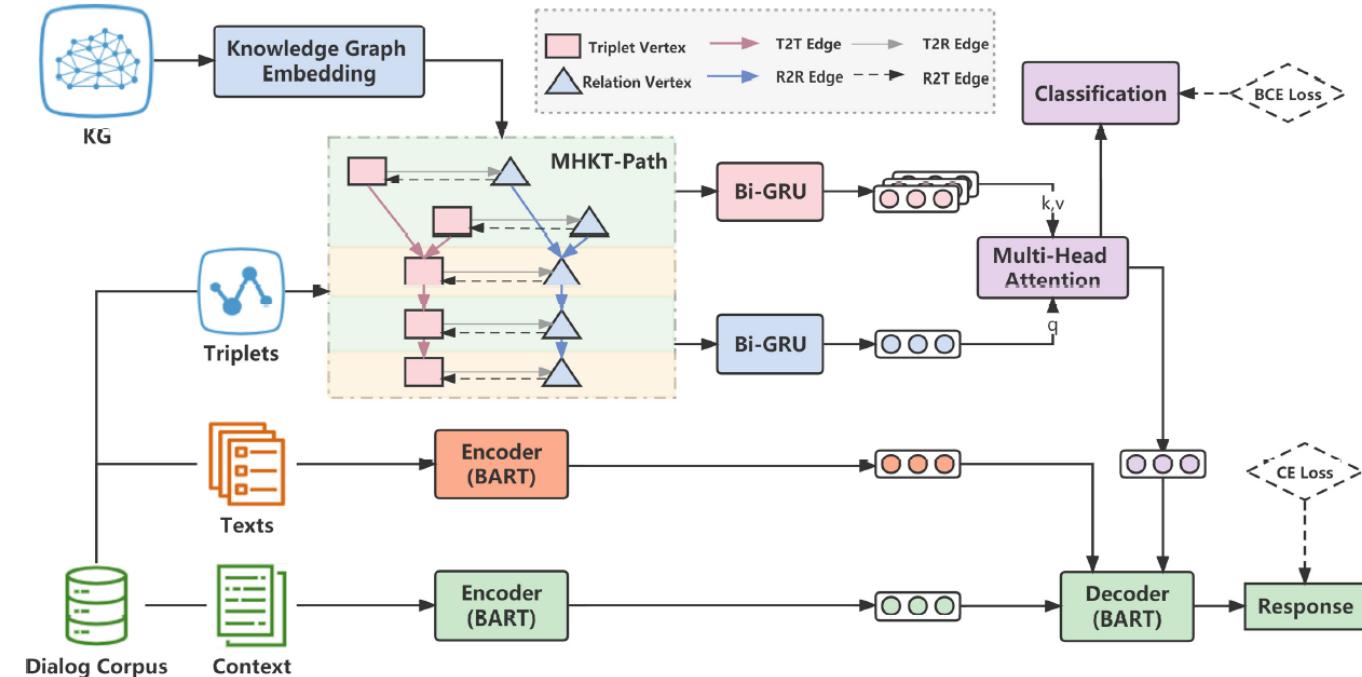
対話データセットを用いた評価+ユーザ評価を実施.

- 考察

いずれの評価実験でもベースラインモデルに勝る評価を得られた.

- 課題

記載なし.



**Fig. 2.** The architecture of the proposed RT-KGD model.

# 4B: ONTOLOGIES (I)

# How to Agree to Disagree Managing Ontological Perspectives using Standpoint Logic

Luc'ia Go'mez A'larez, Sebastian Rudolph and Hannes Strass

Computational Logic Group, Faculty of Computer Science, TU Dresden, Dresden, Germany

- 一言でいうと  
StandPoint Logicを用いた知識の相互運用性の課題の解決のための提案
- 動機  
人間の世界観の多様性や自然言語の意味の不明瞭性などによって生じる知識の相互運用性の課題を解決したい
- 手法  
知識の表現と相対する可能性のあるドメインの特性を推論することに適した立場の概念に基づく論理形式の使用
- 背景  
図では命題論理をベースとした Stand Point のフレームワークを導入しているが、このフレームワークは現代のオントロジー言語を用いた知識統合のシナリオに対して十分な表現力を備えていない
- 結果  
first-order logic を表現力のある基礎言語として用いて、様々なエージェントが各々のStand Point を確立し、それらのエージェントのStand Pointを考慮し、それらの間の組み合わせを調整できる multi-modan framework を提案
- 考察  
本研究で提案する framework は、既製の高度に最適化された推論器の上に構築された推論システムを使用して、オントロジーの調整、や概念交渉、知識の集約に適用が可能である
- 課題  
モダリティのは範囲内で自由変数の存在を許容するFOSL fragments の複雑さの研究  
Stand Point の組み合わせによる知識の曖昧さの解消や同調のための橋渡し的ルールの確立

$$* \preceq (LC \cup LU) \wedge \Diamond_{LC}[T \sqsubseteq T] \wedge \Diamond_{LU}[T \sqsubseteq T],$$

図：BOFによるモデリング（ずっと見ていると顔文字のように見えてくるモデル...）

# Efficient Dependency Analysis for Rule-Based Ontologies

Larry Gonz'alez , Alex Ivliev, Markus Krotzsch and Stephan Mennicke  
Knowledge-Based Systems Group, TU Dresden, Dresden, Germany

- 一言でいうと

正の信頼関係と拘束関係という2種類のルール依存関係に注目し、それらを効率的に実装するための最適化アルゴリズムを設計・実装する

- 動機

存在規則オントロジーの静的解析のために、いくつかのタイプの依存関係が提案されている。しかし、これらの依存関係はほとんど実装されておらず、依存関係の潜在的な可能性はほとんど示されていない。

- 手法

正の信頼関係と拘束関係の計算のためのアルゴリズムを設計（関連するチェックの数を減らすグローバルな最適化と特定のチェックを実行するローカルな最適化を行うアルゴリズムを設計）  
様々なサイズの200以上の実世界のオントロジーを用いて実験

- 背景

最適化の有効性を考慮すると、グローバルとローカルの2種類のアルゴリズムを設計することで、全体的なパフォーマンスに重要な貢献をし、様々な実用的な用途が可能であることを提示できると考えた

- 結果

本研究で提案する最適化アルゴリズムによって既存のルールで発生する複雑な形の依存関係であっても、効率的に実装できることを示し、実用的かつ理論的に興味深い多くの用途に適用することが可能であることを示した。

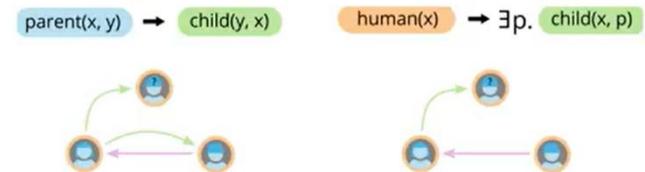
- 考察

本研究は、信頼性解析の実用化に向けた道を切り開くものである。

- 課題

優れたオントロジーデザインは、冗長性を確実に回避する仕様にすべきかについての検討  
依存関係の知識が推論を高速化するためにどのように使われるかについての研究

**Reliances** describe interactions between rules



図：依存関係の図 (発表スライドを参照)

# Controlled Query Evaluation in OWL 2 QL : A “Longest Honeymoon” Approach

Piero Bonatti<sup>1</sup>, Gianluca Cima<sup>2</sup>, Domenico Lembo<sup>3</sup>, Lorenzo Marconi<sup>3</sup>, Riccardo Rosati<sup>3</sup>, Luigi Sauro<sup>1</sup> and Domenico Fabio Savo<sup>4</sup>

<sup>1</sup>Università di Napoli Federico II, Naples, Italy, <sup>2</sup>University of Bordeaux, CNRS, Bordeaux INP, LaBRI, Bordeaux, France, <sup>3</sup>Sapienza Università di Roma, Rome, Italy, <sup>4</sup>Università degli Studi di Bergamo, Bergamo, Italy

- 一言でいうと

過去のQueryの評価に基づき、現在のQueryに対する応答を変更する動的CQE(Controlled Query Evaluation)の提案

- 動機

静的CQEでは、ユーザのプライバシー保護のため、Queryに対してどの答えを可視化し、どの答えを隠すかをあらかじめ選択していた。そのためにいくつかのケースでは、ユーザの興味を考慮しない完全に協調的ではないCQEが提案されていた。

- 手法

特定のQueryの流れに基づいて真か偽かを徐々に決定する動的CQEであるdynCQE法を用いる。

この方法は一連のQueryが与えられたとき、dynCQEは偽を返す前に可能な限り長い正直な答えのシーケンスを返すという Longest Honeymoon 特性を満たしている。

- 背景

動的CQE法は、各ステップにおいて、出来るだけ多く検閲することにより以前の回答と一貫したものを選ぶ。そうすることで、現在の検閲者のプールからQueryに正直に回答できるものを選択することで、次のQueryに正直に回答する可能性を最大化することが可能になる。

- 結果

OWLおよび記述論理オントロジーにおける制御されたQuery評価のための最も協調的なアプローチを示した。本アプローチは、従来の協調性の低いCQEアプローチと比較して計算量的に困難でないことを示した。

Query適合性を解決するための新しいQuery書き換えアルゴリズムを定義した。

- 考察

本研究では動的CQEに対する問い合わせ書き換えアプローチの可能性を示したが、実用的なQuery応答手法を定義することとそれをnon-Boolean UCQsに拡張するためには、より多くの作業が必要である。

- 課題

non-Boolean UCQsへのフレームワークの拡張

より表現力のあるデータ保護ポリシーを包含するためのポリシー言語の拡張

OWL2などの他のオントロジー言語における含意問題の研究

**Example (cont'd)**

The query  $q_3 = \exists x \text{Contain}(x, \text{phenytoin})$  can be rewritten as:

$$\begin{aligned} q_r = & \neg \text{Buy}(\text{ann}, d_1) \wedge \neg \text{DrugA}(d_1) \wedge \exists x \text{Contain}(x, \text{phenytoin}) \vee \\ & \neg (\text{Buy}(\text{ann}, d_1) \wedge \text{DrugA}(d_1) \wedge \neg(T)) \wedge \exists x (\text{Buy}(\text{ann}, d_1) \wedge \\ & \text{Contain}(x, \text{phenytoin}) \wedge \neg(d_1 = x)) \vee \\ & \neg \text{Buy}(\text{ann}, d_1) \wedge (\exists x (\text{DrugA}(d_1) \wedge \text{Contain}(x, \text{phenytoin}))) \vee \\ & \exists x (\text{Buy}(\text{ann}, d_1) \wedge \text{DrugA}(d_1) \wedge \text{Contain}(x, \text{phenytoin})) \wedge \neg(T \vee d_1 = x) \end{aligned}$$

Due to the second disjunct, we have that  $q_r$  evaluates to true in  $\mathcal{A}$  (indeed we had that  $\mathcal{S} \models q_3$ )

図：First-order rewriting の例  
(発表スライドを参照)

# 5A: EMBEDDINGS (II)

# Radar Station: Using KG Embeddings for Semantic Table Interpretation and Entity Disambiguation

Jixiong Liu, Viet-Phi Huynh, Yoan Chabot/Orange and Raphaël Troncy/EURECOM

- 一言でいうと

事前に識別された表のセル-個体のアノテーションの後に、意味曖昧性解消のステップを追加することを目的としたハイブリッドシステムを提案

- 動機

参照知識グラフを用いた表セルの値の曖昧性排除の手法のうち、ヒューリスティックに基づくものは、しばしば列の種類と、投票戦略によって集約された列間関係に依存し、最高の性能を達成するものであることが実証されている。しかし、他の列の意味的類似性を無視することが多く、エラーが伝播しやすい。

- 手法

t	Methods	Limaye			T2D			2T_v2			ShortTables		
		AP	PA	GP	AP	PA	GP	AP	PA	GP	AP	PA	GP
1	DAGOBAH-SL	0.647	0.168	0.853	0.308	0.053	0.785	0.067	0.023	0.870	0.672	0.194	0.654
	RS + TransE		0.630	0.870		0.294	0.813		0.041	0.871		0.355	0.673
	RS + RotateE		0.636	0.870		0.289	0.812		0.044	0.871		0.363	0.673
	RS + DistMult		0.391	0.861		0.163	0.798		0.034	0.870		0.229	0.658
	RS + ComplEx		0.57	0.869		0.171	0.798		0.036	0.870		0.235	0.659
0.95	DAGOBAH-SL	0.614	0.296	0.853	0.332	0.180	0.785	0.327	0.208	0.870	0.671	0.302	0.654
	RS + TransE		0.528	0.872		0.312	0.815		0.230	0.872		0.414	0.673
	RS + RotateE		0.542	0.873		0.312	0.815		0.235	0.872		0.418	0.674
	RS + DistMult		0.377	0.860		0.230	0.797		0.213	0.870		0.328	0.659
	RS + ComplEx		0.435	0.864		0.233	0.798		0.219	0.870		0.334	0.660
0.9	DAGOBAH-SL	0.653	0.432	0.853	0.336	0.241	0.785	0.500	0.300	0.870	0.714	0.414	0.654
	RS + TransE		0.570	0.872		0.323	0.815		0.313	0.872		0.532	0.684
	RS + RotateE		0.578	0.873		0.322	0.814		0.318	0.872		0.536	0.684
	RS + DistMult		0.475	0.860		0.274	0.797		0.303	0.870		0.466	0.668
	RS + ComplEx		0.494	0.862		0.275	0.798		0.306	0.870		0.471	0.669

- 結果

- 考察

- 課題

- エンティティテーブルやマトリクステーブルなど、より多くの種類のテーブルを処理することで、システムの網羅性と堅牢性を高める
- 言語モデルを活用して、より多くの文脈情報を考慮する
- 他の競合システムの上にRadar Stationをプラグインする

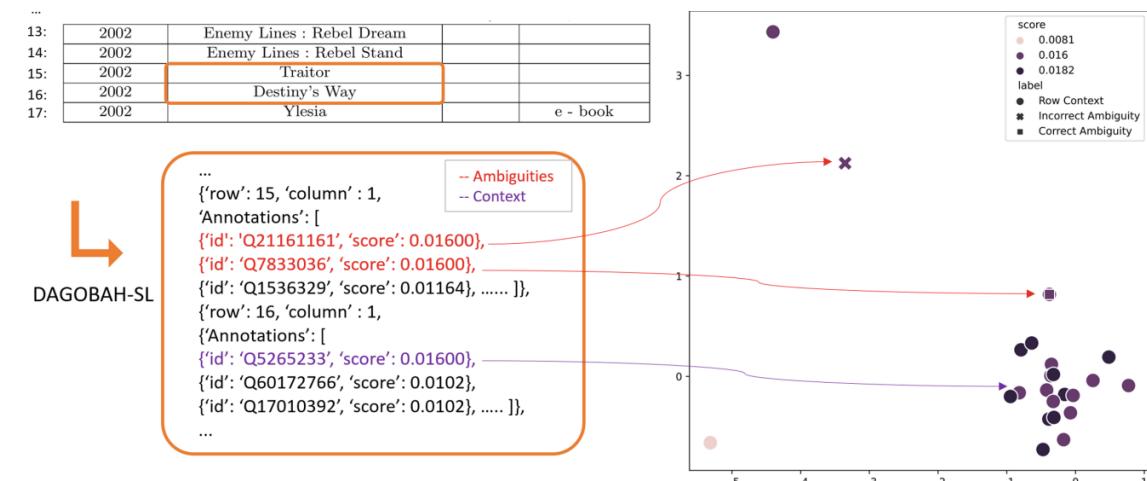
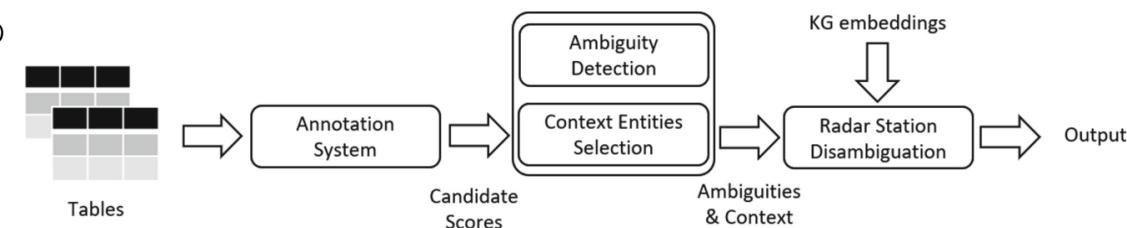


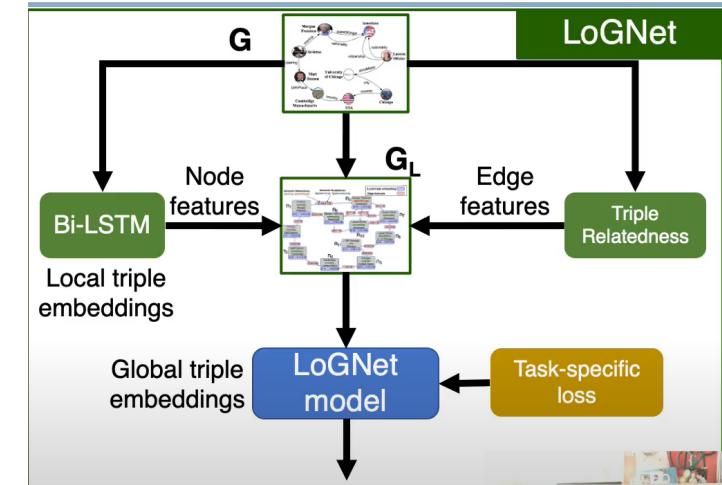
Fig. 1. Illustration of Radar Station with DAGOBAH-SL results. The plot is generated with RotatE embeddings after dimension reduction by T-SNE.



# LoGNet: Local and Global Triple Embedding Network

Giuseppe Pirr' o/ Sapienza University of Rome

- 一言でいうと  
述語の異常検知を目的として、グラフニューラルネットワークをベースに、局所的および大域的なトリプル埋め込みを組み合わせたモデルを提案する。
- 動機  
「3つ組」の分類や類似の計算を行いたい
- 手法
- 背景
- 結果
- 考察
- 課題
  - 学習モデルの改良
  - パスベースの他のアプリケーションの検討



KG	% Corr.	ConvE	RotatE	KGTtm	KBAT	KGist	T2Vec	LNet_E	LNet_G
DBP	0.05%	0.532	0.542	0.674	0.553	0.532	0.678	0.679	<b>0.684</b>
	1%	0.526	0.534	0.671	0.550	0.534	0.675	0.680	<b>0.691</b>
	2%	0.524	0.525	0.670	0.546	0.542	0.672	0.672	<b>0.687</b>
	3%	0.513	0.521	0.667	0.538	0.560	0.668	0.668	<b>0.671</b>
	5%	0.501	0.513	0.662	0.535	0.561	0.670	0.675	<b>0.680</b>
DBP1M	0.05%	0.543	0.567	0.687	0.561	0.547	0.6912	0.696	<b>0.702</b>
	1%	0.536	0.560	0.671	0.560	0.541	0.687	0.690	<b>0.696</b>
	2%	0.531	0.556	0.676	0.559	0.550	<b>0.678</b>	0.669	0.671
	3%	0.530	0.542	0.678	0.551	0.546	0.6751	0.677	0.679
	5%	0.526	0.534	0.674	0.541	0.576	0.671	0.673	0.681
NELL	0.05%	0.531	0.534	0.614	0.529	0.54	0.6214	0.621	<b>0.631</b>
	1%	0.5301	0.532	0.623	0.546	0.543	0.6245	0.625	<b>0.632</b>
	2%	0.523	0.521	0.638	0.543	0.542	0.641	0.631	<b>0.647</b>
	3%	0.513	0.516	0.622	0.526	0.552	0.623	0.624	<b>0.631</b>
	5%	0.502	0.512	0.637	0.518	0.553	0.6401	0.641	<b>0.655</b>

## Setting

- LoGNet<sub>E</sub>: local triple embeddings as concatenation of the embeddings of the elements
- LoGNet<sub>G</sub>: local triple embeddings via BiGRU
- Evaluation measure: AUC

# The DLCC Node Classification Benchmark for Analyzing Knowledge Graph Embeddings

Jan Portisch/SAP SE and Heiko Paulheim/University of Mannheim

- 一言でいうと  
埋め込みアプローチを、どのような種類のクラスを表現できるかという観点から分析するためのリソース
- Resource Type
  - DLCC (Description Logic Class Constructors) ゴールドスタンダード作成のためのフレームワーク
  - 実グラフに基づくゴールドスタンダードと合成知識グラフに基づくゴールドスタンダード
  - 埋め込みのクラス分離能力を容易に評価・比較するための評価フレームワーク
  - 様々な埋め込み手法に対する予備的な分析
- 動機  
知識グラフの埋め込みは、元のベンチマークタスクに依存せず、一般に複数のタスクに利用できる汎用性がある。しかし、どの埋め込み手法がどのようなクラス分類を学習できるかは、系統的に評価されていない。
- リソースの設計方針
- 新規性
- リソースの再現性／活用実績
- 品質
- リソースの可用性
- 次にすべきこと

TC	RDF2vec	RDF2vec <sub>oa</sub>	TransE-L1	TransE-L2	TransR	ComplEx
tc01	<b>0.882</b>	0.867	0.767	0.752	0.712	0.789
tc02	<b>0.742</b>	0.737	0.677	0.677	<b>0.531</b>	<b>0.549</b>
tc03	0.797	<b>0.812</b>	<b>0.531</b>	0.581	<b>0.554</b>	<b>0.536</b>
tc04	<b>1.000</b>	0.998	0.790	0.898	0.685	0.553
tc05	<b>0.892</b>	0.819	0.691	0.774	0.631	0.726
tc06	0.978	0.963	0.898	0.978	0.888	<b>1.000</b>
tc07	0.583	0.583	<b>0.540</b>	0.615	<b>0.673</b>	0.518
tc08	0.563	0.585	0.585	<b>0.613</b>	<b>0.540</b>	0.523
tc09	0.610	<b>0.628</b>	0.588	<b>0.543</b>	0.525	<b>0.545</b>
tc10	<b>0.638</b>	0.623	0.588	0.573	<b>0.518</b>	<b>0.510</b>
tc11	<b>0.633</b>	0.580	0.583	0.590	0.573	0.590
tc12	<b>0.644</b>	0.614	0.618	<b>0.550</b>	<b>0.513</b>	<b>0.540</b>

より最近発表されたRDF2vecのフレーバーを含む、より多くの埋め込みアプローチに系統的な評価を拡張する

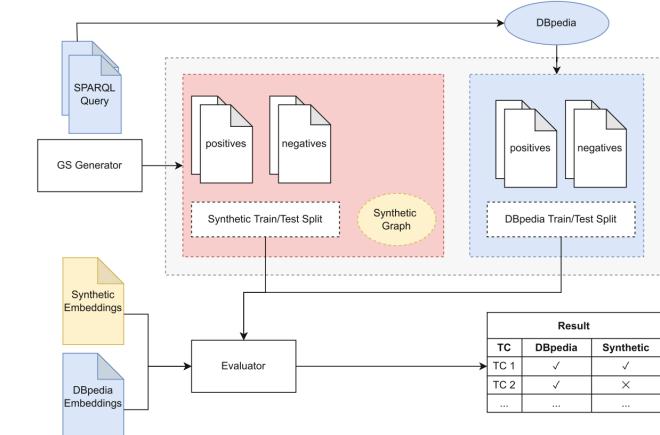


Fig. 2. Overview of the approach

# 5B: ONTOLOGIES (II)

# Semantic Knowledge Graphs for Distributed Data Spaces: The Public Procurement Pilot Experience

Cecile Guasch<sup>1</sup>, Giorgia Lodi<sup>2</sup> and Sander Van Dooren<sup>1</sup>/ (1)European Commission, DG DIGIT, Brussels, Belgium, (2)Institute of Cognitive Sciences and Technologies of the Italian National Research Council (ISTC-CNR), Rome, Italy

- 一言でいうと  
EUにおける公共調達のナレッジグラフを検証し活用するパイロットプロジェクトとその教訓について
- 動機
  - EUの250,000以上の公的機関は、GDPの約14%を公共調達に費やしており、EU経済において重要性が知られている
  - 公共調達を書類ベースからデジタルに移行するにつれて相互運用性の課題に直面している
- 手法
  - パイロットプロジェクトの目的やガイドラインに沿った分散型アーキテクチャをBig Data Test Infrastructure (BDTI)上に設計
  - eProcurement OntologyとRMLによるデータ変換のワークフローを、Apache Airflowと提案したTransformation Ontologyで管理
- 背景
  - 調達分野ではThe Buy for You Platform、公共調達ではLOTED2、PPROCなどあるがいくつかの要素に絞っている
- 利用実績
  - 欧州委員会ISA2プログラムの資金援助を受けた欧州のパイロットプロジェクト
- 学んだ教訓やベストプラクティス
  - 宣言的データ変換アプローチアプローチは人手による介入を大幅に削減
  - DCAT-AP, PROV-Oはデータソースと変換結果のトレーサビリティを高める
  - パイロット参加者の多くはKG管理のインフラを所有しないのでBDTIが必要
  - ...

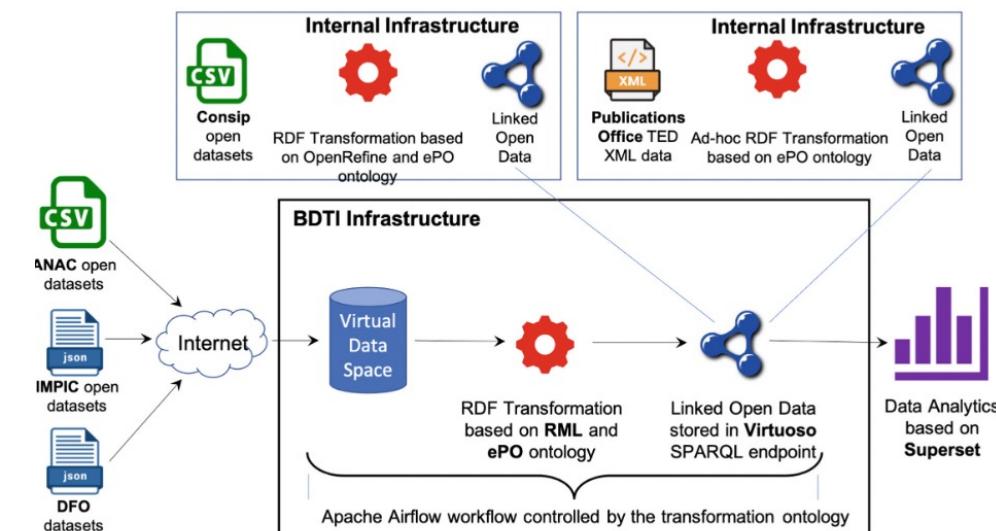


Fig. 1. Pilot architectural scenario

# Faithful Embeddings for $\mathcal{EL}^{++}$ Knowledge Bases

Bo Xiong<sup>1</sup>, Nico Potyka<sup>2</sup>, Trung-Kien Tran<sup>3</sup>, Mojtaba Nayyeri<sup>4</sup> and Steffen Staab<sup>4,5/</sup>

(1) University of Stuttgart, Stuttgart, Germany, (2) Imperial College London, London, UK, (3) Bosch Center for Artificial Intelligence, Renningen, Germany, (4) University of Stuttgart, Stuttgart, Germany, (5) University of Southampton, Southampton, UK

- 一言でいうと

記述論理  $\mathcal{EL}^{++}$  における論理構造を適切に捉える幾何学的知識ベース埋め込み手法 BoxEL の提案

- 動機

現在の埋め込み手法はデータレベル (ABox) しか埋め込む事ができないか、概念レベル (TBox) を扱う際に固有の制限を受ける

- 手法

KB内の概念をボックス（軸合わせされた超直方体）としてモデル化し、実体をボックス内にあるべき点としてエンコードし、リレーションをボックスと点との間のアフィン変換として考える

- 背景

存在量化と論理積をサポートする  $\mathcal{EL}^{++}$  の埋め込み手法 Onto2Vec, OPA2Vecなどが提案されているが、アノテーションデータを必要とし、論理構造を明示的にモデル化するものではない

- 結果

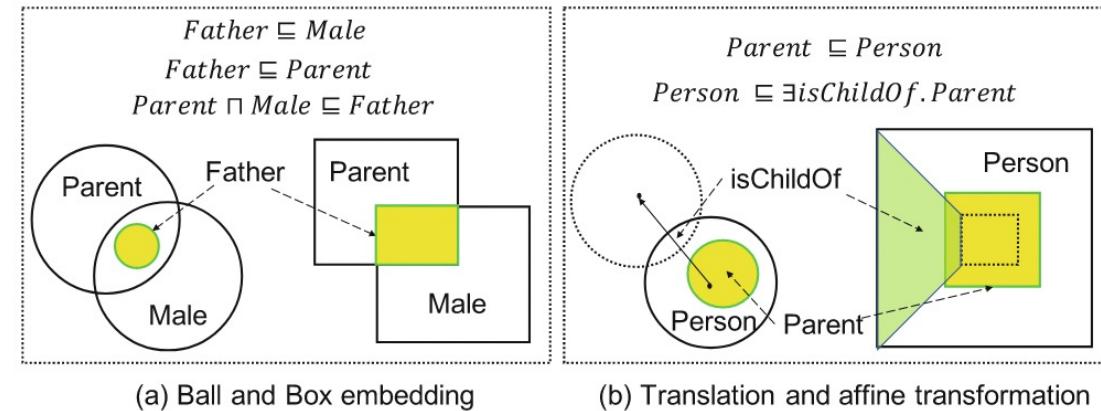
バイオ系のオントロジー (Gene Ontology, GALEN, Anatomy) を使用したタンパク質相互作用の予測実験で既存手法を上回る精度

- 考察

Translation系にくらべて Transformation系（アフィン変換）が精度を向上させる。実体を点として識別することで、実体と概念を区別する方ことに利点がある。

- 課題

確率的記述論理 StatisticalELS に拡張する



# POSO: A Generic Positioning System Ontology

Maxim Van de Wynckel and Beat Signer / Vrije Universiteit Brussel, Brussels, Belgium

- 一言でいうと  
SSNとSOSA上に構築された新たな汎用測位システムオントロジーPOSOの提案
- 動機  
様々な測位システムの相互運用性のために汎用性のある語彙が必要
- 手法  
Semantic Sensor Network (SSN) OntologyとSensor, Observation, Sample and Actuator (SOSA) ontologyをとトップレベルオントロジーとして採用
- 背景  
Basic WGS84 vocabulary, Location Ontology, GeoSPARQL, LinkedGeoData ontologyなどあるが、測位システムは常に地理的境界を必要とするわけではなく、より文脈的な情報を依拠する
- 結果  
学生の屋内・屋外追跡のためのキャンパス測位システムに応用。GPSを用いた屋外測位システム、Wi-Fiフィンガープリントを用いた屋内測位システム、ハイブリッド測位システムを含むシナリオでの応用方法の説明。  
測位システムの入出力RDFを使用することで、異なる測位システムの技術と出力を分類することができる。
- 課題  
測位技術やアルゴリズムの手続きを追加し、各手続きが提供する入力と出力をさらに記述することに重点を置く

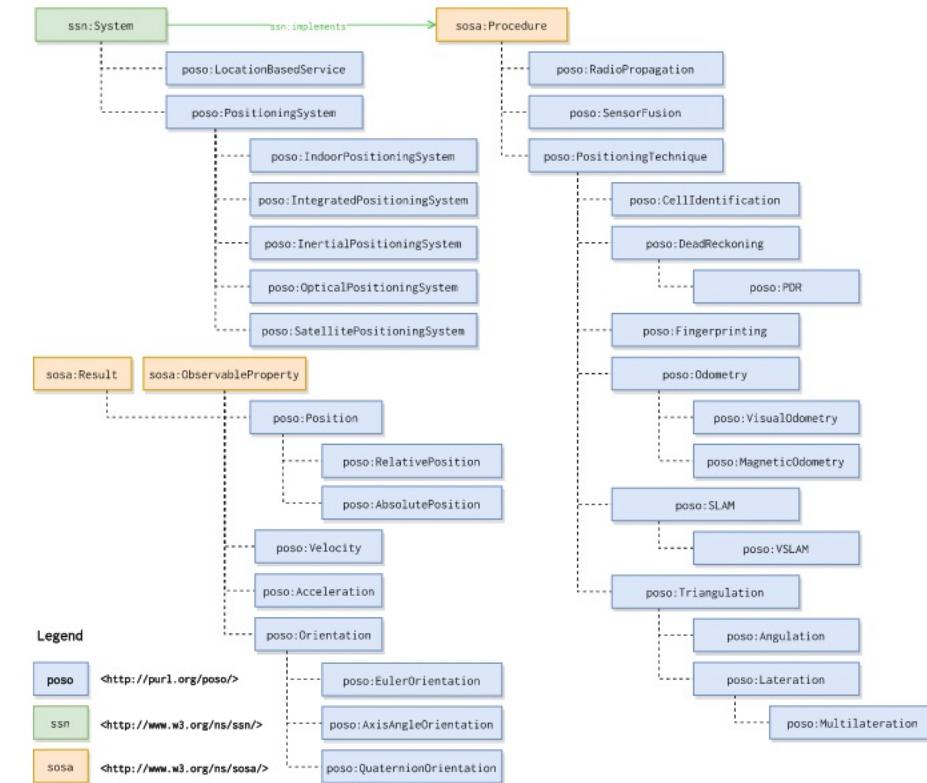


Fig. 2. Positioning systems and techniques in the POSO ontology

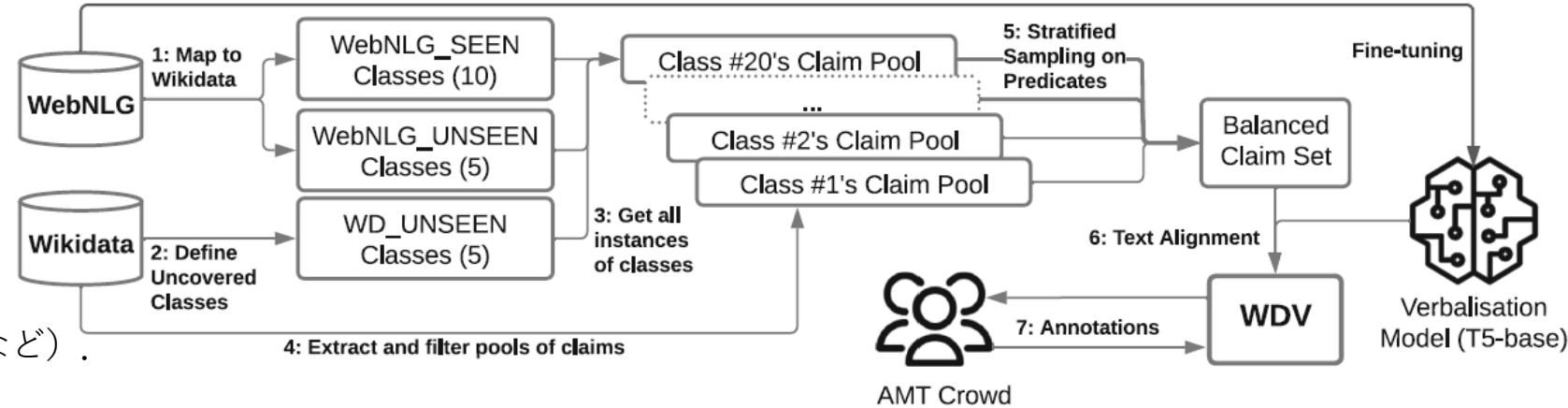
休憩

# 6A: NLP (II)

# WDV: A Broad Data Verbalisation Dataset Built from Wikidata

Gabriel Amaral, Odinaldo Rodrigues and Elena Simperl /King's College London

- 一言でいうと  
Wikidataを言語化 (Verbalisation) したデータセット.
- Resource Type  
言語リソース (+アノテーション)
- 動機  
データの言語化は重要であるが、既存のデータセットは不十分な点がある (クラス定義, 述語のバリエーション, など).
- リソースの設計方針  
Wikidataから抽出したトリプルを言語化し, クラウドソーシングでアノテーション.
- 新規性  
既存のデータセットは存在するが, KGとの対応などで不十分.
- リソースの再現性／活用実績 ツールが公開されており再利用可.
- 品質  
構築したデータセットのユーザ評価が行われている.
- リソースの可用性  
GitHubにてデータ生成用プログラムと共に公開されている.
- 次にすべきこと  
英語以外の言語への対応. (リソースの活用方法の検討?)



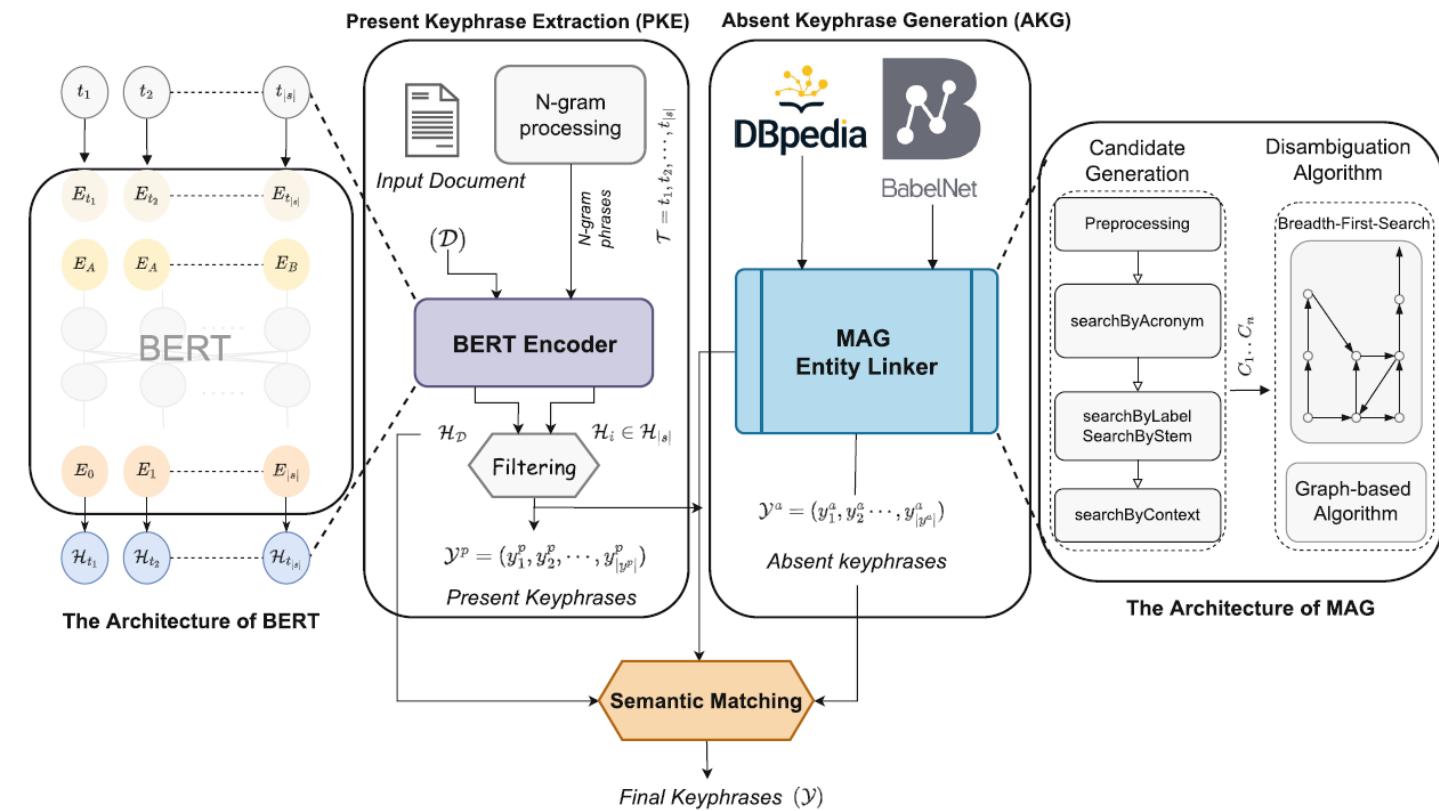
**Table 4.** Examples of claims verbalised by the model. From left to right, we see: the claims, their components' labels, the verbalisations, and their aggregated metrics of fluency (median) and adequacy (adequacy percentage).

Subject	Predicate	Object	Verbalisation	Flu.	Ade.
Eve Towards the Light	Height	268 centimetre	Eve Towards the Light is 268 centimetres high	2.0	1.0
Bydgoszcz	Flag	Flag of Bydgoszcz	The flag of Bydgoszcz is the flag of Bydgoszcz	0.0	1.0
Marawi	Köppen climate classification	Tropical climate	City of the Philippines and capital of the province of Lanao del Sur	4.0	0.4
Decius	Child	Hostilian	Decius is a child of Hostilian	4.0	0.4
UCAC2 28288086	Surface gravity	9 centimetre per square second	UCAC2 28288086 has a surface gravity of 9 centimetres per square second	4.0	1.0
Casa Batlló	Fabrication method	Trencadís	Casa Batlló is made using the Trencadís method	5.0	0.8

# MultPAX: Keyphrase Extraction using Language Models and Knowledge Graphs

Hamada Zahera, Daniel Vollmers, Mohamed Sherif and Axel-Cyrille Ngonga/Paderborn University

- 一言でいうと  
言語モデルと知識グラフを用いたキーフレーズを抽出するフレームワークMultPAXの開発
- 動機  
キーフレーズの抽出は自然言語処理において重要。
- 手法  
言語モデル(BERT)によるPresent keyphrase extraction (PKE)と,知識グラフからのAbsent keyphrase generation (AKG)の結果をSemantic Matchingする。
- 背景  
既存のアプローチはテキストを言語モデルで処理するのが主流。KGを活用することを検討したい。
- 結果  
4つのベンチマークデータセットを用いて評価。
- 考察  
既存システムよりも評価となった。KG利用部分は短いテキストの際に、より有効であった。
- 課題  
他のKGの利用。関連する単語を抽出するためにDBpediaのアブストラクトを再帰的に利用。



**Fig. 1.** The architecture of MULTPAX framework with three components: *present keyphrase extraction, absent keyphrase generation and semantic matching*

# REBench: Microbenchmarking Framework for Relation Extraction Systems

Manzoor Ali, Muhammad Saleem and Axel-Cyrille Ngonga Ngomo / Paderborn University, University of Leipzig

- 一言でいうと  
目的に応じた関係抽出の（マイクロ）ベンチマークを生成するフレームワーク
- Resource Type  
ソフトウェア (+ データセット)
- 動機  
既存のベンチマークは、データやタスクによって同じものでも結果が異なる。タスクや目的に応じたマイクロベンチマークが必要
- リソースの設計方針  
ベンチマークに必要な特徴を入力し、RELD-RDF DatasetからSPARQLを用いてデータを抽出。関係の選択、ベクトル化、クラスタリングをしたのち、最終的なものを選択してベンチマークとして出力する (Fig.2)。
- 新規性 マイクロベンチマークを生成する既存のフレームワークは存在しない。
- リソースの再現性／活用実績 オープンソースとして公開されており活用可能。
- 品質 評価実験の結果が提示されている。生成されたデータの多様性で評価している。
- リソースの可用性 オープンソースで公開。データも拡張可能。
- 次にすべきこと フレームワークを自然言語処理タスクに拡張する計画あり。

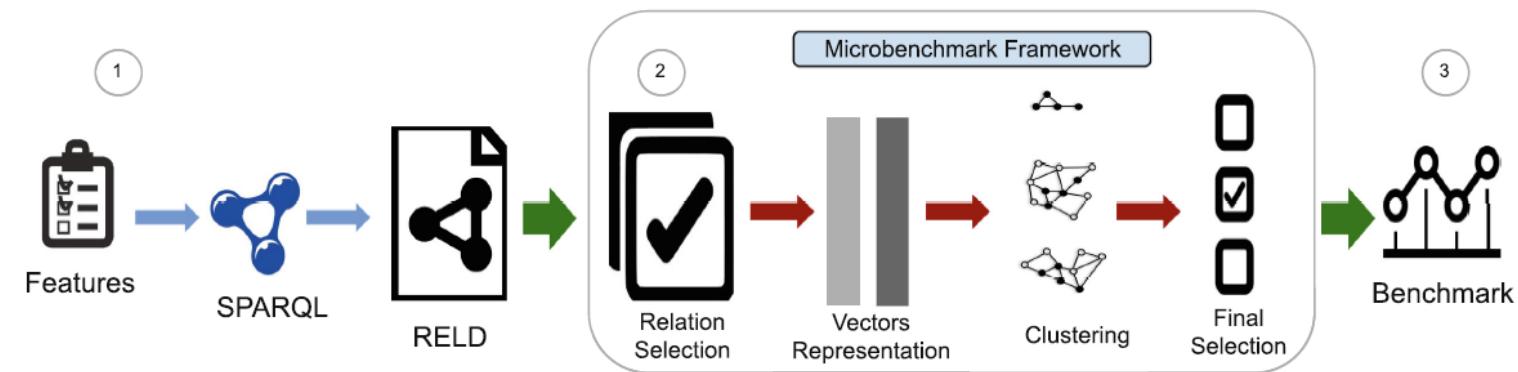


Fig. 2. REBench sampling process from input to output.

# 6B: QUERYING

# Hashing the Hypertrie: Space- and Time-Efficient Indexing for SPARQL in Tensors

Alexander Bigerl(1), Lixi Conrads(1), Charlotte Behning(2), Muhammad Saleem(3), Axel-Cyrille Ngonga Ngomo(3)

(1)DICE Group, Department of Computer Science, Paderborn University, Paderborn, Germany

(2)Department of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn

(3) CS Department, Leipzig University, Leipzig, Germany

- 一言でいうと  
Hypertrieの改良  
Hypertrie:テンソルベースRDFストアのインデックス構造
- 動機  
Hypertrieのメモリ使用量とクエリ処理効率の改善
- 手法
  - h: ハッシュ化による重複ノードの排除
  - s: 非分岐パスの圧縮
  - i: 単一エントリーの親ノードへの格納
- 背景  
Hypertrieは多くの重複ノードを含み、メモリ使用量が多い
- 結果  
4種のデータセットで実験の結果、最大でメモリ使用量を70%削減、QoSを39%向上、QMPhを740%向上
- 考察  
3つのアプローチをすべて適用したものが最良
- 課題  
インデックス構築時間が長い

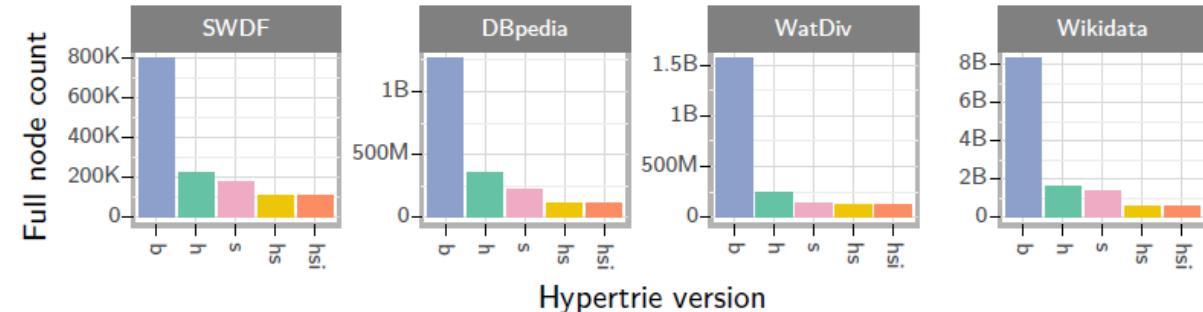


Fig. 1: Full node counts of different hypertrie versions on four datasets. The hypertrie versions are identified by their features: baseline (b), single-entry node (s), hash identifiers (h), and in-place storage of height-1 single-entry nodes (i).

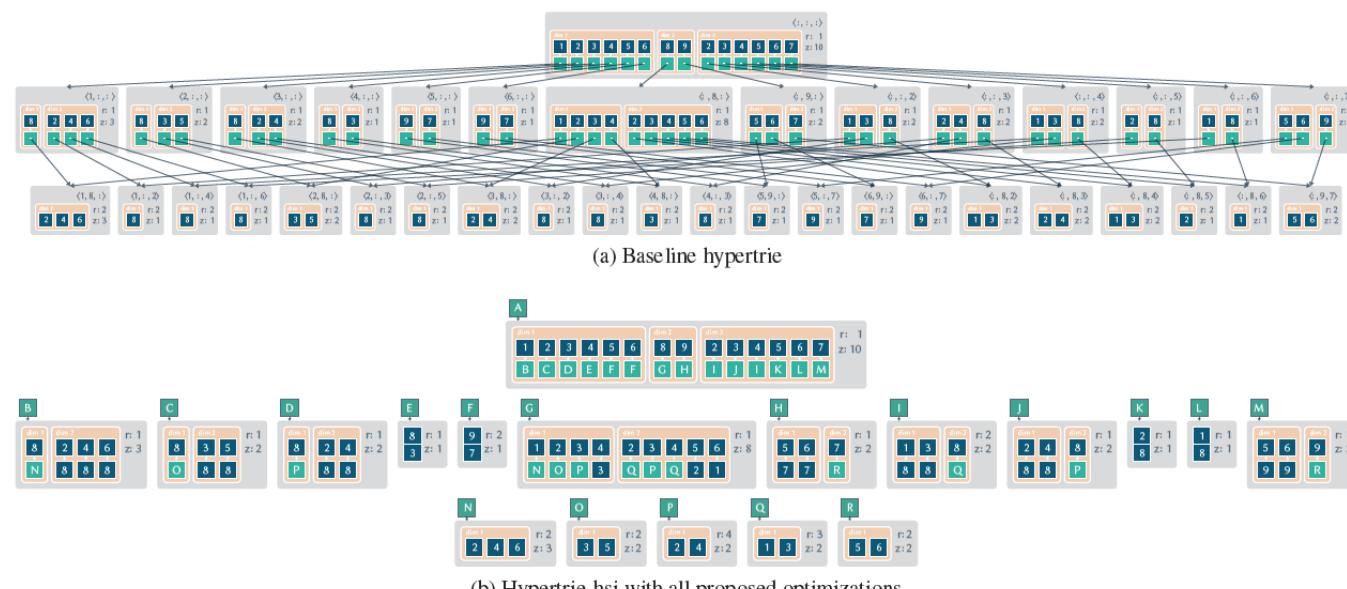


Fig. 3: Both hypertries encode the RDF graph from Figure 2. RDF resources are encoded by their ID (see also Figure 2).

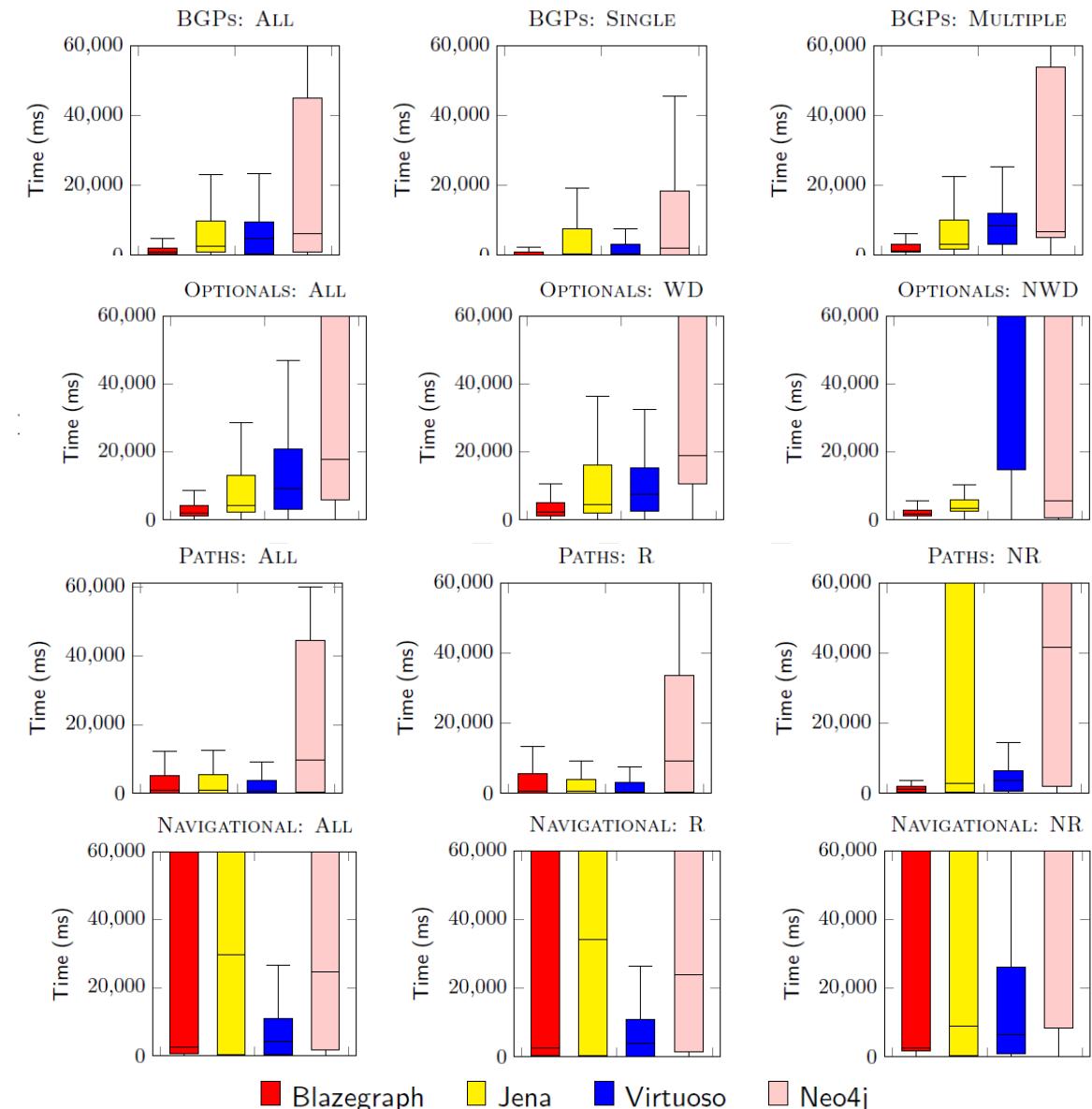
# WDBench: A Wikidata Graph Query Benchmark

ARenzo Angles(1,2), Carlos Buil Aranda(1,3), Aidan Hogan(1,4), Carlos Rojas(1), and Domagoj Vrgoc(1, 5)

(1) IMFD Chile (2) DCC, Universidad de Talca (3) Universidad Tecnica Federico Santa Maria

(4) DCC, Universidad de Chile (5) PUC Chile

- 一言でいうと  
Wikidataに基づく知識グラフのベンチマークWDBenchの作成
- 動機  
実世界のデータを使ったベンチマークが少ない
- 手法  
SPARQLやグラフDB共通の4種のクエリにフォーカス  
Basic(961) : トリプルパターンが複数結合したもの  
Optional(498) : OPTIONALを含むもの  
Path(660) : 2つのノード間のパスが正しいかどうかを検証する  
Navigational(539) : プロパティパスとトリプルパターンの結合
- 背景  
既存の実世界ベンチマークはログ分析の結果に依存したものがほとんどである
- 結果  
Wikidataログからクエリを抽出し、4つのストアでベンチマーク
- 考察  
BlazegraphとVirtuosoが最速で、Neo4jが最遅
- 課題  
提案手法に新しい機能を含めることはできるが、どんどん複雑になる



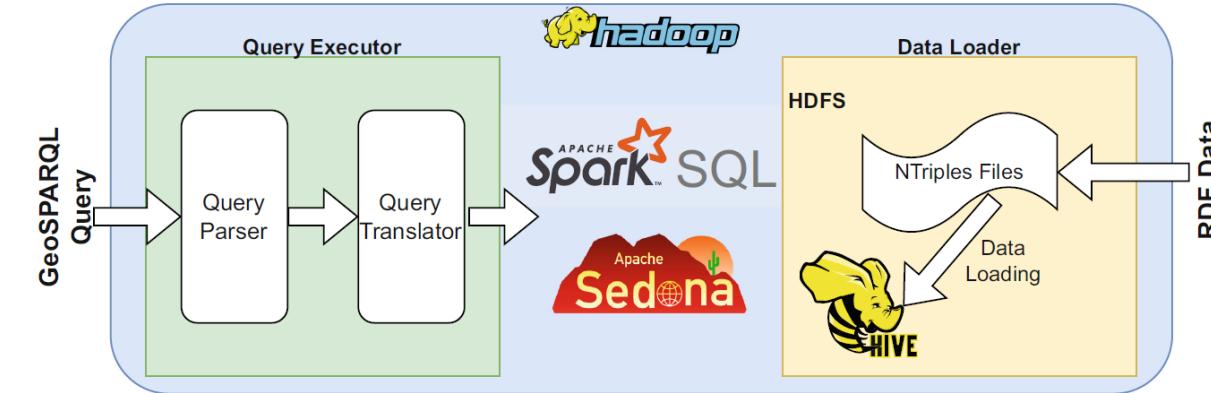
# Strabo 2: Distributed Management of Massive Geospatial RDF Datasets

ADimitris Bilidas(1), Theofilos Ioannidis(1), Nikos Mamoulis(2), Manolis Koubarakis(1)

(1) National and Kapodistrian University of Athens, Athens, Greece

(2) University of Ioannina, Ioannina, Greece

- 一言でいうと  
分散地理空間RDFストアStrabo2を提案
- 動機  
大規模RDFデータに対し、GeoSPARQLクエリを実行
- 手法
  - SEDONA：地理空間ライブラリ
  - Spark、Hadoop：分散処理フレームワーク
  - hive：永続化
- 背景  
大規模なRDFデータに対する分散RDFエンジンは  
地理空間クエリをサポートしていない
- 結果  
1.16TBのデータを生成し、72個のクエリを実行した結果、  
平均107秒で処理可能
- 考察  
データサイズが数GB以上になると、Sparkの分散処理によって処理が高速化される
- 課題  
RCC8やEgenhoferのトポロジ関係とクエリ書き換えへの対応



**Fig. 1.** Architectural overview of STRABO 2

# 7A: KG REFINEMENT & SCALE

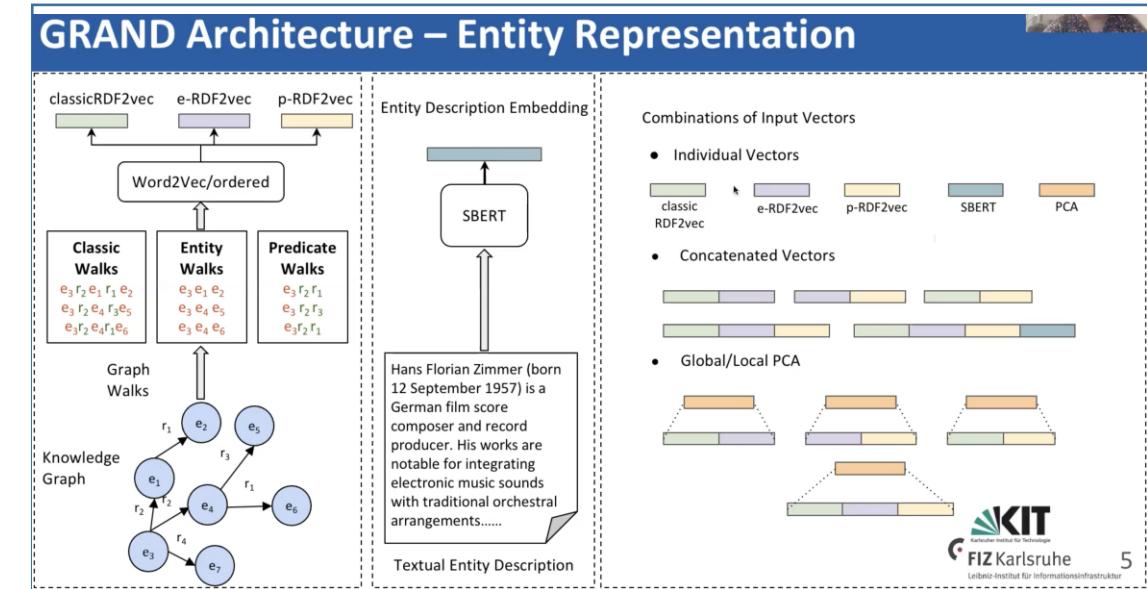
# MAIN TRACK 7A: KG REFINEMENT & SCALE

## Christian Kindermann

# Entity Type Prediction Leveraging Graph Walks and Entity Descriptions

Russa Biswas, Jan Portisch, Heiko Paulheim, Harald Sack and Mehwish Alam / Leibniz Institute for Information Infrastructure 他

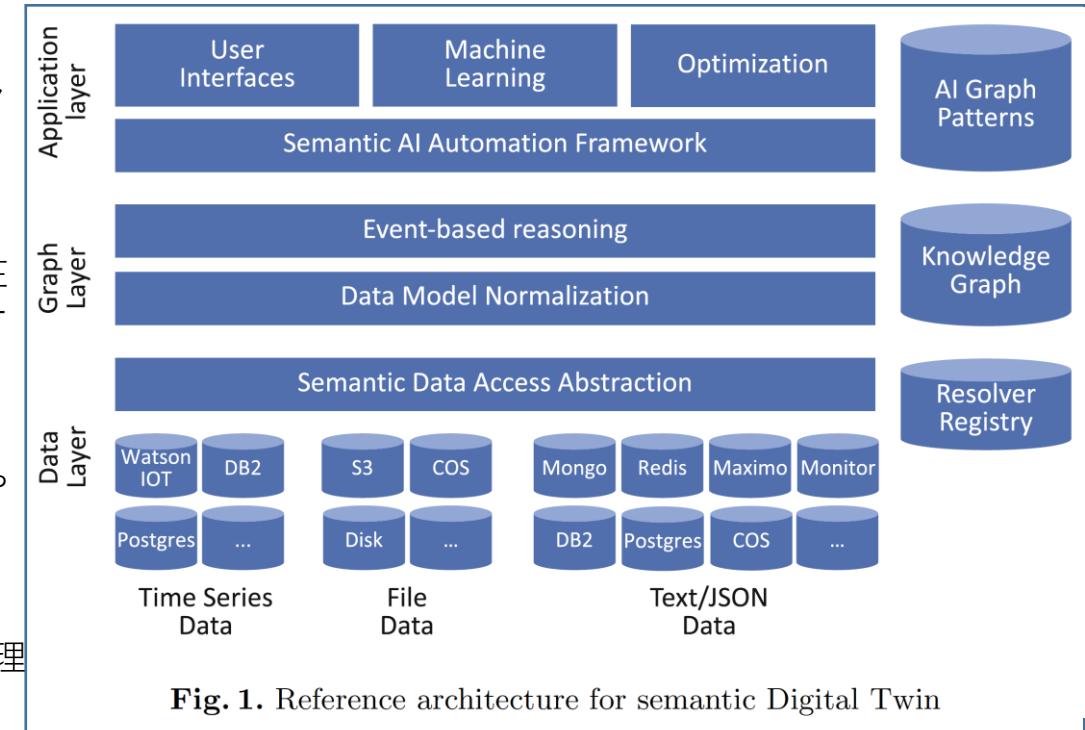
- 一言でいうと  
新しいエンティティ型予測手法GRANDを提案。
- 動機  
これまでよりも予測性能の高い手法を開発したい。
- 手法  
テキスト埋め込み情報とRDF2vecを利用した手法。
- 背景  
DBpediaやFreebaseのような知識グラフの型に関する情報は不完全。  
NLPでこれらの知識グラフを利用する際には型情報が重要。
- 結果  
これまでの提案手法よりも良い予測性能を示した。
- 考察  
ベンチマークデータセット（DBpedia630kとFIGER）に対して、ベースラインモデルと比較して良い成績を収めたことを確認。
- 課題  
さらに細かい粒度の型が定義されているYAGOやCaLiGraphに対しても試してみる。



# Scaling Knowledge Graphs for Automating AI of Digital Twins

Joern Ploennigs, Konstantinos Semertzidis, Fabio Lorenzi and Nandana Mihindukulasooriya / IBM Research Europe

- 一言でいうと  
デジタルツインで知識グラフを大規模に利活用するためのベンチマークの提案と複数のデータベースを評価。
- 動機  
デジタルツインにおいて知識グラフを利活用することは有益だが、現在利用可能なデータベースは実データにスケールしにくいのでベンチマークを提案して利活用を促進したい。そしてIBM製データベースの販促。
- 手法  
知識グラフ層とデータ層に分けるなど、階層化することでスケールしやすくした。デジタルツインデータ処理に向くデータベースを開発。
- 背景  
さまざまな産業分野でIoTが利活用され、そこから得られるデータの管理には知識グラフ技術が有益だがデータ規模が大きいという課題がある。
- 結果  
デジタルツイン向けのベンチマークデータDTBMを構築し、一般的なトリプルストア向けのベンチマークBSBMとLUBMと合わせて、既存データベース（VirtuosoやGraphDB、Neo4jなど）とIBM製データベースKITTを比較。KITTが全てのベンチマークで高い性能を発揮。
- 利用実績  
著者の所属するIBMで様々な顧客に対してサービスを提供している。
- 学んだ教訓やベストプラクティス  
RDFとプロパティグラフの利点を併せ持ち、包摂と推移律を簡単に利用可能なデータベースがデジタルツイン分野では求められる。



# Facing Changes: Continual Entity Alignment for Growing Knowledge Graphs

Yuxin Wang, Yuanning Cui, Wenqiang Liu, Zequn Sun, Yiqiao Jiang, Kexin Han and Wei Hu / Nanjing University

- 一言でいうと

知識グラフを対象としたエンティティ・アライメント (AE) について、対象グラフが変化しても全体を処理し直さずに更新できる手法を提案した。

- 動機

これまで発表してきたAE手法は、対象グラフを静的なものと想定しているが、DBpediaなど時とともに変化するものなので、効率よく更新したい。

- 手法

ContEAという新たな手法を提案した。

- 背景

従来のAEでは、DBpediaやWikidataのような更新頻度の高い知識グラフを想定していなかった。更新情報をうまく利用すれば精度を上げられる。

- 結果

ベースライン手法と比較して高精度に、そして短時間でAEを実行できた。適切なベンチマークデータが無いので、はDBP15Kを元にして新たに構築。

- 考察

ContEA手法は、ベースライン手法に対して、効率的かつ効果的であることがベンチマークデータに対する評価で明らかになった。ベースライン手法は、埋め込みに基づくEA手法を更新のたびに再学習させるものと、演繹的知識グラフ埋め込みを利用するもの。

- 課題

新たな関係が加わったり、既知のアライメントを新たに追加したり、逆にエンティティやトリプルが失われる事例にも対応する。

The diagram illustrates the Continual Entity Alignment (ContEA) process for growing knowledge graphs (KGs). It consists of two main parts: a bipartite graph alignment and a 'growing' process.

**Bipartite Graph Alignment:** This part shows entities from two knowledge graphs, KG1 (left) and KG2 (right), connected by various relations. Entities are represented by colored circles: blue for KG1 and orange for KG2. Relations are shown as arrows. Dashed lines with arrows pointing from KG1 to KG2 are labeled "align", indicating existing alignments. Some edges are labeled with French terms like "proie sur", "sous-classe", and "animal".

**Growing Process:** This part shows the evolution of the alignment over time.

- Time t:** Shows two sets of entities, \$e\_1^1\$ through \$e\_1^4\$ and \$e\_2^1\$ through \$e\_2^4\$, connected by edges \$e\_1^1 \rightarrow e\_2^1\$ and \$e\_1^2 \rightarrow e\_2^2\$. A red 'X' marks a failed alignment attempt between \$e\_1^3\$ and \$e\_2^1\$.
- Time \$t+1\$:** Shows the same entities after growth. New edges have been added (e.g., \$e\_1^3 \rightarrow e\_2^2\$), new entities have been introduced (e.g., \$e\_1^4\$ and \$e\_2^4\$), and new alignments have been established (e.g., between \$e\_1^3\$ and \$e\_2^4\$).

**Legend:**

- seed alignment
- - - new edge
- new entity
- - - new alignment

# 7B: TEMPORAL KGS

# H<sup>2</sup>TNE: Temporal Heterogeneous Information Network Embedding in Hyperbolic Spaces

Qijie Bai, Jiawen Guo, Haiwei Zhang, Changli Nie, Lin Zhang and Yuan Xiaojie / College of Computer Science, Nankai University, Tianjin, China

- 一言でいうと

時間的異種情報ネットワークのための双極空間埋め込み手法H<sup>2</sup>TNEの提案

- 動機

実世界のネットワークの多くは階層構造やべき乗則の分布を示し、ユークリッド空間と等尺性でない。近年、双曲空間における表現学習が、階層構造やべき乗則を持つデータに対して有効であることが証明されており、この性質に触発された

- 手法

時間的・ヘテロに制約された二重ランダムウォーク戦略を利用して構造的・意味的情報を取得し、実世界のネットワークのべき乗則分布を利用するためには双極空間を利用する。

- 背景

実世界のネットワークが非ユークリッド構造であることから双曲空間の表現学習が行われている。双曲線GCN、ポアンカレ球における階層的特徴の学習など。

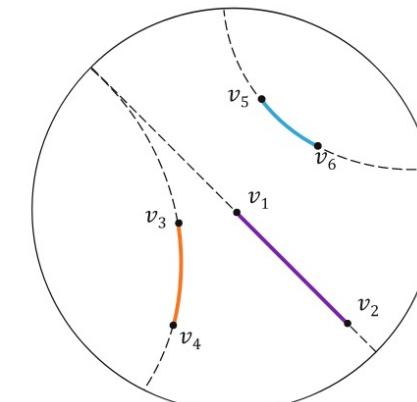
- 結果

いくつかのSOTA手法より優れた性能を達成

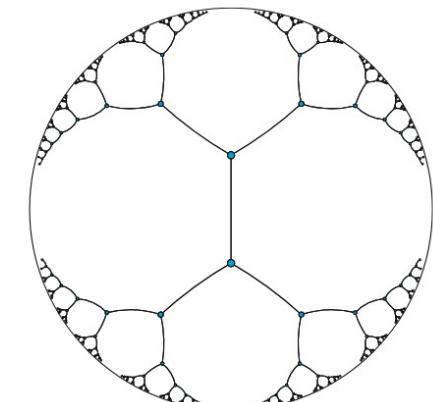
- 考察

ウォークパラメータの影響分析。複数の種類のノードを持つデータセットは $\alpha$ の影響を受けやすく、時間粒度の細かいデータセットは $\beta$ の影響を受けやすいことがわかり、ランダムウォークの制御に適合している。

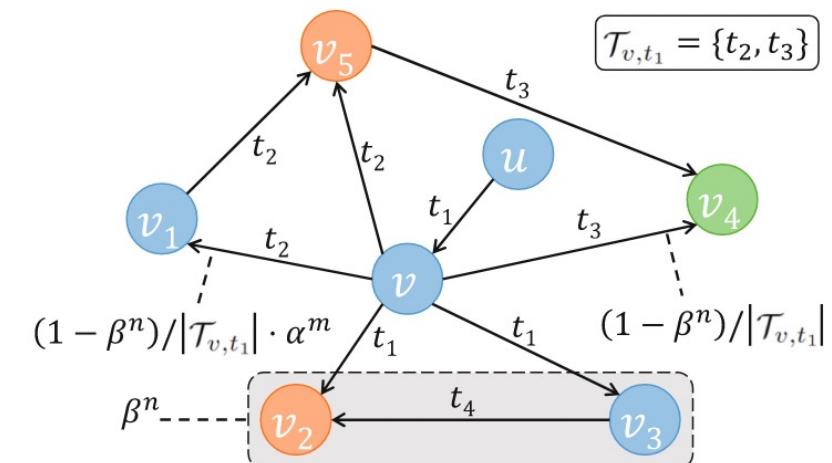
- 課題 ?



(a) Distance in the Poincaré ball



(b) Embedding a tree in  $\mathbb{D}$



# CRNet: Modeling Concurrent Events over Temporal Knowledge Graphs

Shichao Wang, Xiangrui Cai, Ying Zhang and Xiaojie Yuan / (1)College of Cyber Science, Nankai University, Tianjin, China (2) Tianjin Key Laboratory of Network and Data Security Technology, Tianjin, China

- 一言でいうと

過去と未来のタイムスタンプの両方におけるイベントの同時性を利用した時間的KG推論手法CRNetの提案

- 動機

時間的KG(TKG)推論はイベント予測などに不可欠。未来の同時発生事象の間には互いに相関し、影響を及ぼす可能性のある複雑なつながりが存在する。

- 手法

過去の同時発生イベントに対してはノードやエッジではなく完全なイベントメッセージを近傍に渡して適応的に集約するGAT(EventRGAT)を開発する。未来の同時発生イベント間の相互作用をモデル化するために2段階のフレームワークを提案。

- 背景

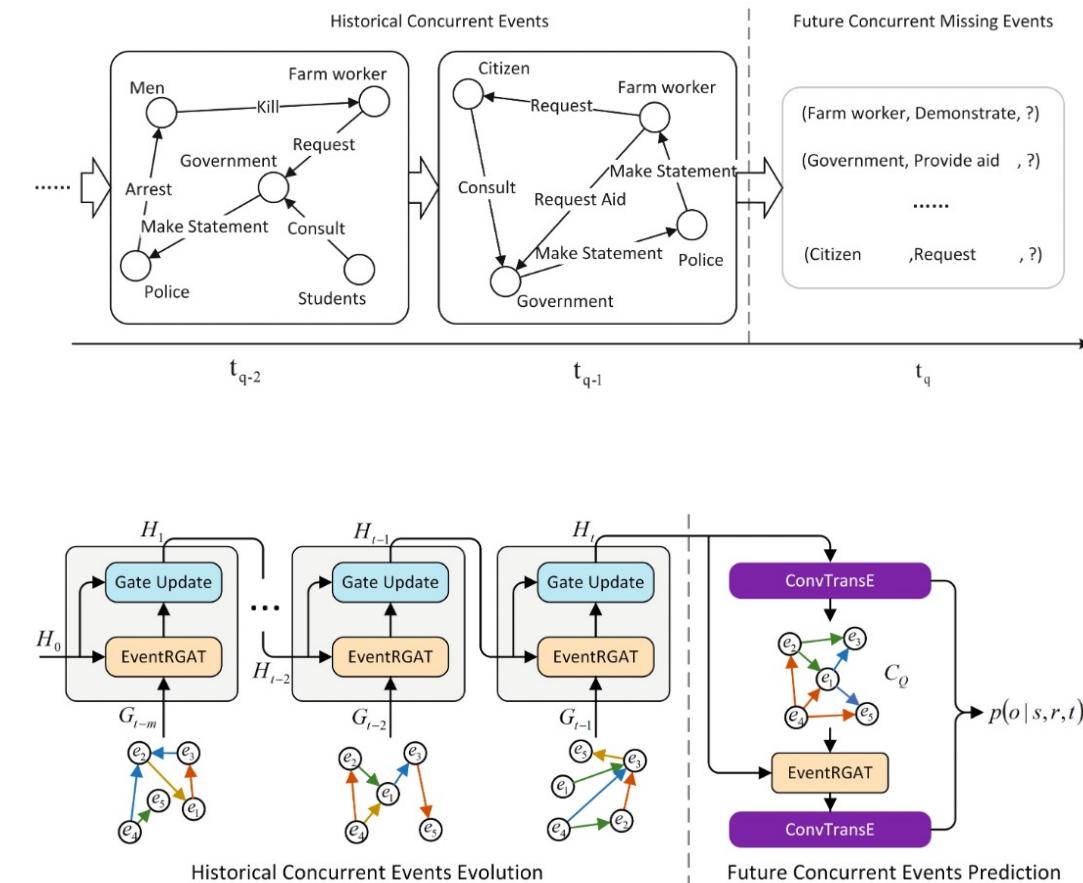
様々なTKG埋め込み・推論手法があるが、同時発生の可能性を無視し、独立に欠損データを予測している。

- 結果

ICEWS18, ICEWS14, GDELTデータセットでSOTA

- 考察

同時発生的な欠損ファクト情報の影響を調査するために、様々なケースを調査した結果、未来のタイムスタンプにおける同時発生のコンテキストが、欠落イベントの予測に有益であることが示された



# Each Snapshot to Each Space: Space Adaptation for Temporal Knowledge Graph Completion

Yancong Li, Xiaoming Zhang, Bo Zhang and Haiying Ren / (1) School of Cyber Science and Technology, Beihang University, Beijing, China (2) State Key Laboratory of Software Development Environment, Beihang University, Beijing, China

- 一言でいうと  
異なる潜在空間を異なるタイムスタンプのスナップショットに対応させる  
TKG推論モデルSANEを提案

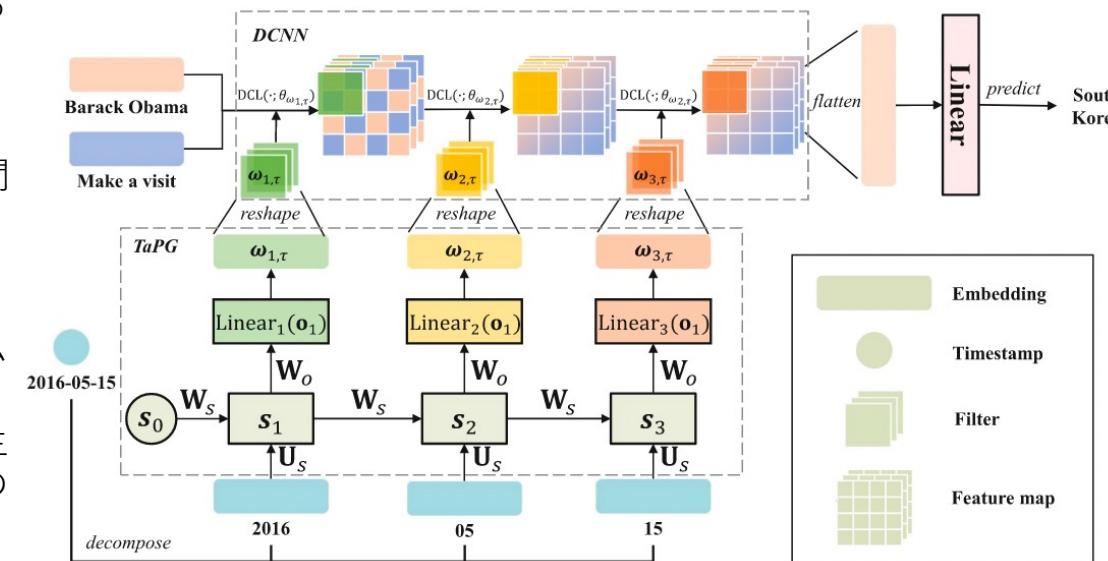
- 動機  
TKGにおける欠損補完は重要な研究テーマ。従来モデルは意味空間が時間とともに変化するにも関わらず、すべての事実を同一の潜在空間で扱う

- 手法  
スナップショットごとに時間固有のネットワークを生成し、異なるタイムスタンプの事実を異なる空間にエンコードする。  
タイムスタンプの距離に応じて潜在空間の重なりを制限するパラメータ生成器を設計し、隣接するスナップショットが異なるが類似した潜在空間の共有を可能にする。

- 背景  
様々なTKG埋め込み・欠損補完手法があるが、従来モデルはすべての事実を同一の潜在空間で扱う。

- 結果  
ICEWS14, ICEWS05-15, YAGO11k, Wikidata12kデータセットでSOTA
- 考察

時間情報を用いない場合は最悪の結果を示し、SANEにとっての時間情報の重要性を反映している。パラメータ生成器はモデル性能に大きな影響を与える。タイムスタンプを異なる時間粒度に分解することは、モデル性能の向上に有効である。タイムスタンプを異なる粒度に分解しない場合でも、提案モデルは従来のTKGC手法よりも良好な結果を得る。



# 8A: LEARNING REPRESENTATIONS

# Towards Neural Network Interpretability using Commonsense Knowledge Graphs

Youmna Ismaeil<sup>1,2</sup>, Daria Stepanova<sup>1</sup>, Trung-Kien Tran<sup>1</sup>, Piyapat Saranrittichai<sup>1</sup>, Csaba Domokos<sup>1</sup>, and Hendrik Blockeel<sup>2</sup> /

<sup>1</sup> Bosch Center for Artificial Intelligence, Renningen, Germany <sup>2</sup> KU Leuven, Leuven, Belgium

- 一言でいうと

外部の常識知識グラフによって定義された属性に整列させるためのフレームワークを提案

- 動機

最近の研究では、CNNの潜在的な表現がセマンティックな概念と一致するが、大量のラベル付きデータに依存し、取得が困難

- 手法

CNNからの隠れユニットを、外部のCommonsense Knowledgeから抽出されたクラスのセマンティック属性にマッピングするためのフレームワーク

- 背景

個々のニューロンに深い説明を与える研究はあるが、大量のデータが必要（提案手法では外部のKGを活用）

個々のニューロンをKGのエンティティとalignmentさせる提案はなし

- 結果

copy-pasteの敵対的画像分類と、一般化されたzero-shot学習タスクに対するフレームワークの有効性を経験的に示す

- 考察

結果から、CVタスクにおける常識KGの重要性と有用性が示唆される。

- 課題

様々な常識KGでの実験と考察

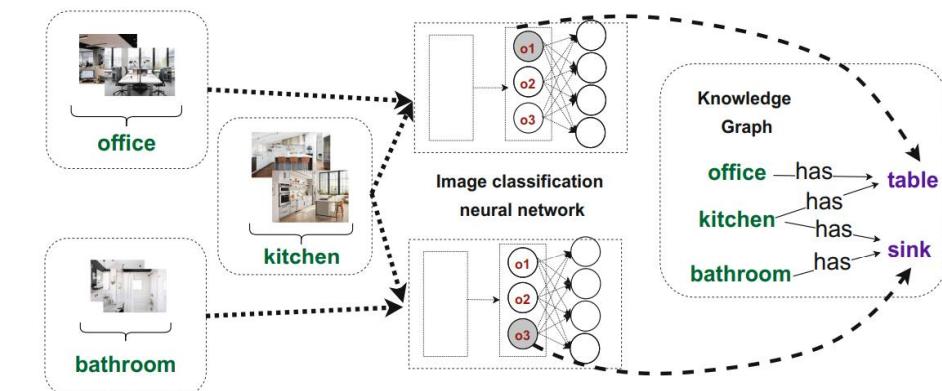


Fig. 1. The goal of our work is to assign meaning to neurons using semantic properties of classes from a knowledge graph.

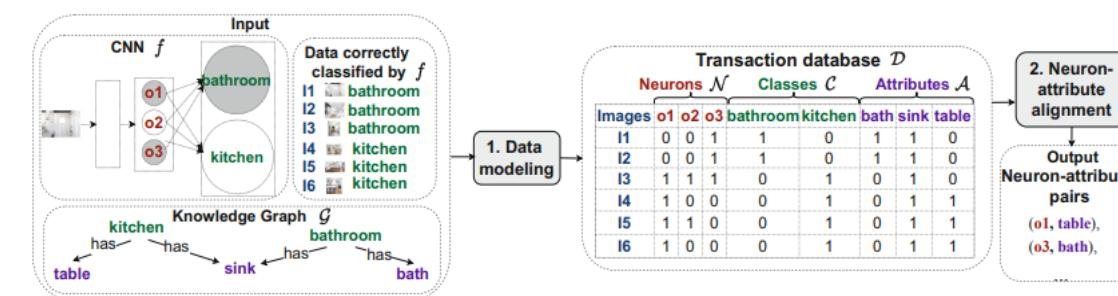


Fig. 2. Workflow of the proposed framework.

# HCL: Improving Graph Representation with Hierarchical Contrastive Learning

Jun Wang<sup>1</sup>, Weixun Li<sup>1</sup>, Changyu Hou<sup>1</sup>, Xin Tang<sup>2</sup>, Yixuan Qiao<sup>1</sup>, Rui Fang<sup>2</sup>, Pengyong Li<sup>3</sup>, Peng Gao<sup>1</sup>, and Guotong Xie<sup>1,4,5</sup> /

<sup>1</sup> Ping An Healthcare Technology, Beijing, China <sup>2</sup> Ping An Property and Casualty Insurance Company, Shenzhen, China <sup>3</sup> School of Computer Science and Technology, Xidian University, Xian, China <sup>4</sup> Ping An Health Cloud Company Limited, Shenzhen, China <sup>5</sup> Ping An International Smart City Technology Company Limited, Shenzhen, China

- 一言でいうと

階層的な方法でグラフ表現を明示的に学習する新しい階層的対照学習 (HCL) フレームワークを提案

- 動機

対照学習 (Contrastive learning) は、グラフ表現学習の強力なツールだが、

固定された粗視化スケールでグラフの特徴を学習するため、ローカル情報またはグローバル情報を過小評価する可能性あり

- 手法

階層的な方法でグラフ表現を明示的に学習する新しい階層的対照学習 (HCL) フレームワークを提案

- より包括的な対比目的のためのより合理的なマルチスケールグラフトポロジを構築するための適応学習pooling手法(L2Pool)
- 各スケール内の相互情報のより表現力豊かな学習をさらに可能にする、新しいマルチチャネル疑似シャムネットワーク

- 背景

教師なしグラフ学習： PHDは、グラフ分類では良い結果だがノードレベルのタスクは良くない

マルチスケールグラフpooling：

- graph-coarsening pooling methods (DiffPool, StructPool)は、計算が複雑で大きなグラフに適用不可
- node-selection pooling methods (gPool, SAGPoolなど)は、元のグラフ構造を失う

- 結果

- node分類タスク: Coauthor CS以外は、提案手法が最高accuracyで、最高
- node clustering タスク: NMIもARIも最高 (CiteseerのARIのみは、MVGRLが最高だがそれに近い)
- graph分類タスク: MUTAG以外は最高 (MUTAGはMVGRLが最高だが、それに近い)

- 考察

学習した表現を視覚化により、HCL がグラフの意味のある特性をうまく捉えていることを確認

- 課題

将来的には、より多くのGNNモデルと効果的に統合され、より多くのグラフ学習タスクに適用。

固有の有益なパターンをキャプチャするためのより豊富な機能の相互作用を探索可能に。

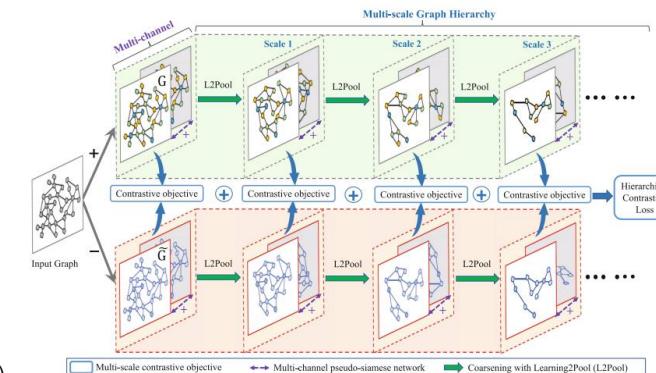


Fig. 1. Framework of the proposed Hierarchical Contrastive Learning (HCL) for graph representation.

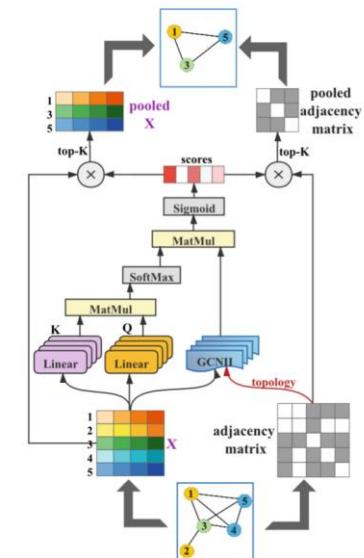


Fig. 2. An illustration of the proposed L2Pool using Transformer-style self-attention and topology information to select representative nodes and to coarsen into a graph hierarchy for cross-scale contrastive learning.

# Context-Driven Visual Object Recognition Based on Knowledge Graphs

Sebastian Monka<sup>1,2</sup>, Lavdim Halilaj<sup>1</sup>, and Achim Rettinger<sup>2</sup> /

<sup>1</sup> Bosch Research, Renningen, Germany <sup>2</sup> Trier University, Trier, Germany

- 一言でいうと

KGに基づくコンテキスト駆動型視覚オブジェクト認識のフレームワークの提案

- 動機

オブジェクト認識のための既存の深層学習方法はデータ駆動型であり、多数の訓練サンプルが必要。  
小さな偏差が発生する新しい環境に直面すると失敗する傾向あり。

- 手法

KGにエンコードされた外部のコンテキスト知識を使用してDeep Learningを強化するアプローチ  
一般的なKGからさまざまなコンテキスト ビューを抽出し、ベクトル空間に変換してDNNに注入。

- 背景

Implicit-Contextual Visual Model: 画像自体で発生する視覚的特徴間の関係をコンテキスト化。

Explicit-Contextual Visual Model: Scene graphs, Label graphs, Semantic scene graphsを活用。  
Context-aware zero-shot learningは、視覚情報 (old, cute, など) をオブジェクトに追加。

Contextual Knowledge Graph Embeddings: CoKEは、動的で柔軟なエンティティと関係の埋め込みを学習し、  
KG上の文脈を考慮。

- 結果

Experiments on Cifar10: すべてのクラスで、他のモデルよりも優れているコンテキスト モデルが異なる。

Experiments on Mini-ImageNet: ベースライン(SupSSL)と、コンテキスト ビューが異なる他モデルよりも大幅に優。

- 考察

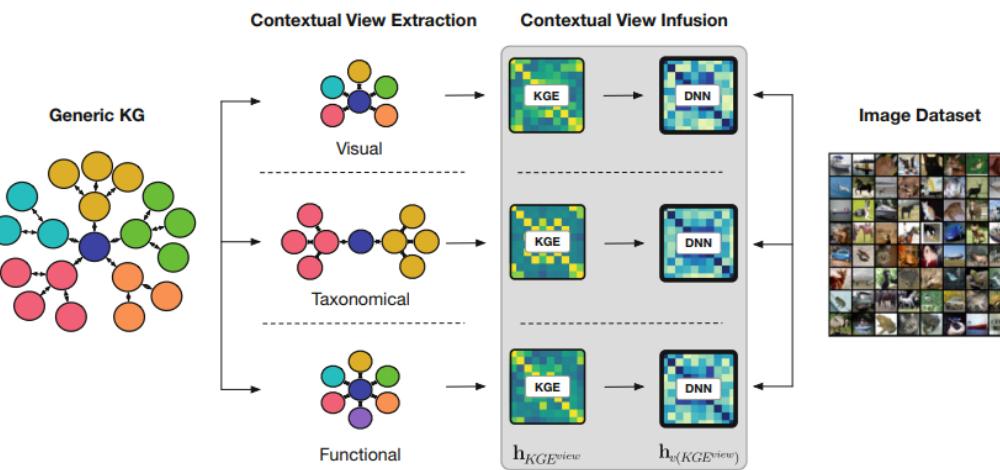
ハイパーパラメーターの選択、重みの初期化、データセットの依存性などの機械学習の問題が、  
学習した表現に強く影響するため、知識の注入は簡単ではない。

KGの形式で提供されるコンテキストが、どのような影響を与えるかを考察。

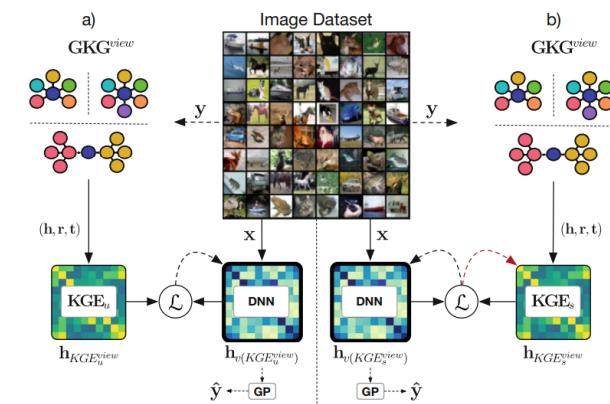
コンテキスト モデルが最終的な精度に与える影響は小さいものの、個々のクラスや画像の表現方法や予測方法には大きな影響

- 課題

Infusion methodの実装、画像データへの強い依存性（初期のデータセットのバイアス）



**Fig. 2.** Our approach to learn contextual image representations consists of two main parts: 1) the *contextual view extraction*; and 2) the *contextual view infusion*.



**Fig. 4. Contextual view infusion.** The contextual object recognition model (DNN) is trained in two different ways: a) using the KG as a trainer, where  $KGE_u$  uses no supervision of the image data; or b) using the KG as a peer, where  $KGE_s$  uses supervision of the image data. Images  $x$  are fed into the DNN, producing  $h_{v(KGEview)}$  which is compared with  $h_{KGEview}$  using the KG-based contrastive loss. In a second step, a gaussian process (GP) or linear layer is trained to predict the class labels  $y$  of  $x$  based on the trained  $h_{v(KGEview)}$ .

# **8B: GRAPH VALIDATION**

# Repairing SHACL Constraint Violations Using Answer Set Programming

Shqiponja Ahmetaj<sup>1,2</sup>, Robert David<sup>2,4</sup>, Axel Polleres<sup>2,3</sup> & Mantas Šimkus<sup>1,5</sup>

<sup>1</sup>TU Wien, Vienna, Austria, <sup>2</sup>WU Wien, Vienna, Austria, <sup>3</sup>Complexity Science Hub Vienna, Vienna, Austria, <sup>4</sup>Semantic Web Company, Vienna, Austria, <sup>5</sup>Umeå University, Umeå, Sweden

## • 一言でいうと

データ統合時のデータグラフの不整合を  
非再帰 Shapes Constraint Language (SHACL) 制約に準拠するように修復する計算アルゴリズムの提案および実装をした

## • 動機

一連の SHACL 制約に準拠するようにデータグラフを修復したい

## • 手法

Answer Set Programming (ASP) を使用して、  
説明問題をロジックプログラムにエンコードし、  
データを特定の制約と一致させる原因となる  
事実の追加と削除の集合（修復）を計算する

## • 背景

データ統合時のデータ品質確保を目的とした  
一連の形状制約に対する RDF グラフの検証を可能にする  
SHACL が W3C によって最近提案された  
しかし、制約違反を解決するための手順の説明がほとんどない  
→ 整合性を実現するためのデータグラフに適用できる修復を自動的に識別する必要

## • 結果

制約違反を解決するための最小限の変更のみが含まれる手法の提案／修復可能な最大数のターゲットを修復する最大修復の概念を導入  
Clingo ASP システムを使用してこれらのエンコーディングを実装・テスト → 実用的な RDF グラフの品質管理と品質向上に有望

## • 考察

修復プログラムと ASP の実装により、修復を実用的なシナリオに持ち込むための基盤が築かれ、実際の RDF グラフの品質が向上した

## • 課題

修復を適用できるユースケースを選択かつ実現可能性を評価し、修復の品質とスケーラビリティを調査すること  
SHACL property path 修復、recursive SHACL 制約をサポートし、class-based/property-based ターゲットに拡張すること

$P_{\text{Annotation}}$  consists of rules that collect existing atoms or atoms that are proposed to be in the repaired data graph.

$P_{\text{Repair}}$  consists of rules that repair the constraints by proposing additions and deletions of atoms.

$P_{\text{Interpretation}}$  consists of rules that collect all the atoms that will be in the repaired data graph.

$P_{\text{Constraints}}$  consists of rules that filter out models that do not provide repairs.

修復問題  $\Psi = (G, C, T)$  において、形状グラフの中で形状によって特定された  
全ての制約に対する修復プログラム  $P_\Psi$  を構成する四つのルール

# HybridFC: A Hybrid Fact-Checking Approach for Knowledge Graphs

Umair Qudus<sup>1</sup>, Michael Röder<sup>1</sup>, Muhammad Saleem<sup>1</sup> & Axel-Cyrille Ngonga Ngomo<sup>1</sup>

<sup>1</sup>DICE Group, Department of Computer Science, Universitat Paderborn, Paderborn, Germany

- **一言でいうと**

複数のfact-checkingアプローチを組み合わせKnowledge Graphs(KG)のassertionの信頼性をより良く予測する

- **動機**

KGにおけるassertionの信頼性をより良く予測したい

- **手法**

ensemble learning設定内で、複数のfact-checkingアプローチの既存のカテゴリの多様性を活用する

- fact-checking : KGにおけるassertionがtrueである尤度を測るタスク

- **背景**

五つの主要なfact-checkingアプローチは部分的に重複する制限を受けている

- text-basedアプローチ : 手動特徴量エンジニアリング

- path-based and rule-basedアプローチ : 背景知識としてのKGの排他的な使用

- embedding-basedアプローチ : 現在のfact-checkingタスクにおける低い精度スコア

- **結果**

Area Under the Receiver Operating Characteristic curveで最新研究を0.14~0.27上回った

- FactBenchデータセットにおいて

- **考察**

活用したアプローチ群が持つパフォーマンスの多様性をうまく利用できている

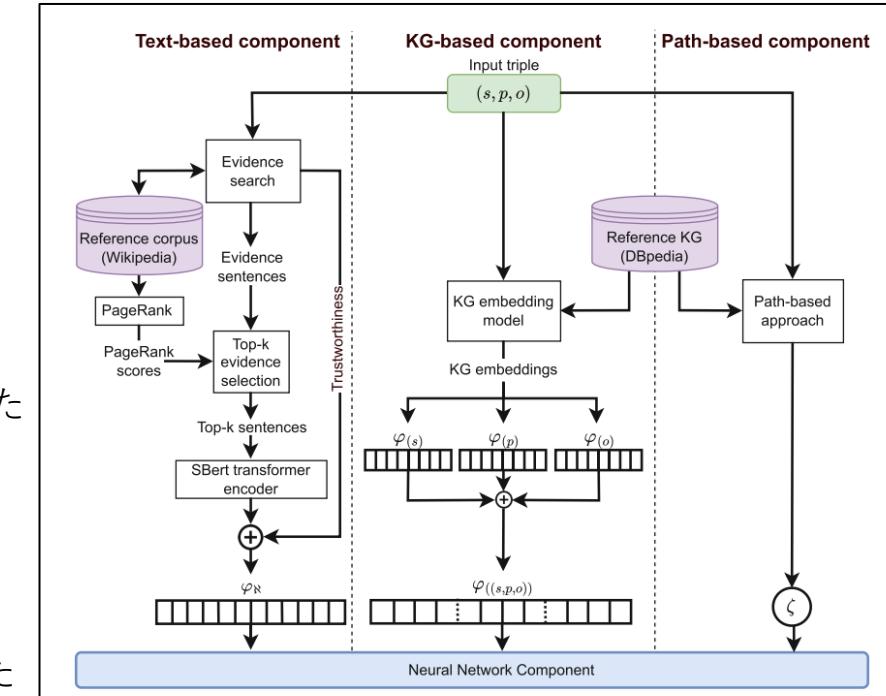
- ほとんどのプロパティ上のCOPAALによる良い性能に頼りつつ、

COPAALがうまく働かないプロパティに対しては他アルゴリズムの予測で補完できた

- **課題**

rule-basedアプローチを統合することにより提案手法のmodularityを活用すること

最良のいくつかの証拠文を選択するための他の可能性の調査



**Fig. 1.** Architecture of HybridFC. The purple color represents reference knowledge. The green color marks the input assertion. KG stands for knowledge graph. (Color figure online)

# Mapping Relational Database Constraints to SHACL

Ratan Bahadur Thapa and Martin Giese

Department of Informatics, University of Oslo, Oslo, Norway

- **一言でいうと**

RDFへのマッピングにおいて、RDBのsqlにおける整合性制約をSHACL制約に書き換えるマッピングを導入する

- **動機**

R2RMLの登場以来、RDF制約言語への関心が高まり、  
SHACLが標準化されたが、マッピングで生成されたデータセットに対して、どのSHACL制約が有効であるか保証できないため

- **手法**

リレーションナルデータの整合性制約を整形された制約形式にマッピングする semantics preserving constraint rewriting を用いて R2R 変換されたデータの意味特性を復元する

- **背景**

表現力の高いオントロジーに基づく関係データのマッピングにおける関係制約の大きな視点に、Epistemic DL axioms や OWL DL axioms に移行する試みとして、最近問題が研究されている

- **結果**

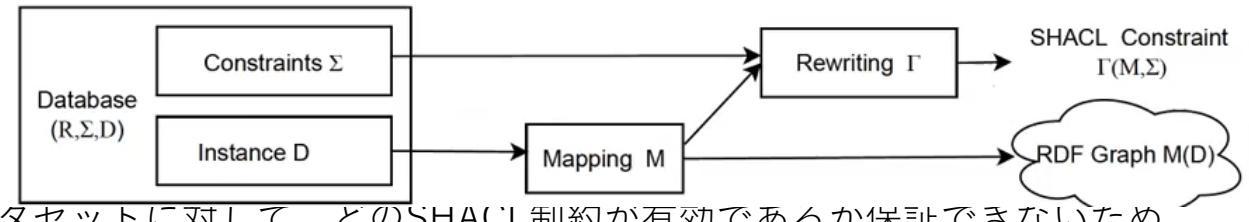
リレーションナルデータから RDF データへの変換における制約書き換えの問題を、最大意味保存の中心的性質に基づき、標準的な SQL データベース制約を RDF グラフのための SHACL 制約言語の形に変換した

- **考察**

提案した制約の書き換えは、SHACL制約を持つ品質保証されたRDFグラフの構築と維持というR2Rマッピングツールの中核を構成すると考えられる。生成されたRDFグラフのSHACL記述は、データ交換、相互運用性、クエリ最適化のためのデータ品質保証を提供する

- **課題**

オントロジーベースのデータアクセスプラットフォームにおけるリレーションナルからRDFへのデータ変換とクエリ最適化のための書き換えの実装と実用的評価



# INDUSTRY TRACK 1

# Knowledge Graph - Our Secret Recipe for Entity Disambiguation

Dmytro Dolgopolov<sup>1</sup>, Elena Romanova<sup>1</sup>, Lev Yarmovsky<sup>1</sup> and Chen Zhang<sup>1</sup>

<sup>1</sup>FINRA (the Financial Industry Regulatory Authority), Rockville, MD, USA

- 一言でいうと

多様な文書から抽出し構築したエンティティ（組織、個人）のKGに対して  
クラスタリング手法を適用し、エンティティの曖昧性を解消する

- 動機

金融取引業規制機構FINRAは、証券会社や投資家が提出する報告書を監査している。  
複数の報告書から組織と個人を特定し、規制に関わるパターンの特定作業に時間が  
かかっている

- 課題

組織名、個人名の曖昧性が多様でNLPだけでは難しい  
(郵便番号や電話番号しか手がかりがないことも)

- 手法

KGに対してコミュニティ分析手法を適用して曖昧性を解消

- “Seed”アルゴリズム：信頼できるエッジを介して繋がったノードを  
種にしてコミュニティを抽出
- CLICKアルゴリズム：重みが最小のエッジでグラフを分割

- 評価

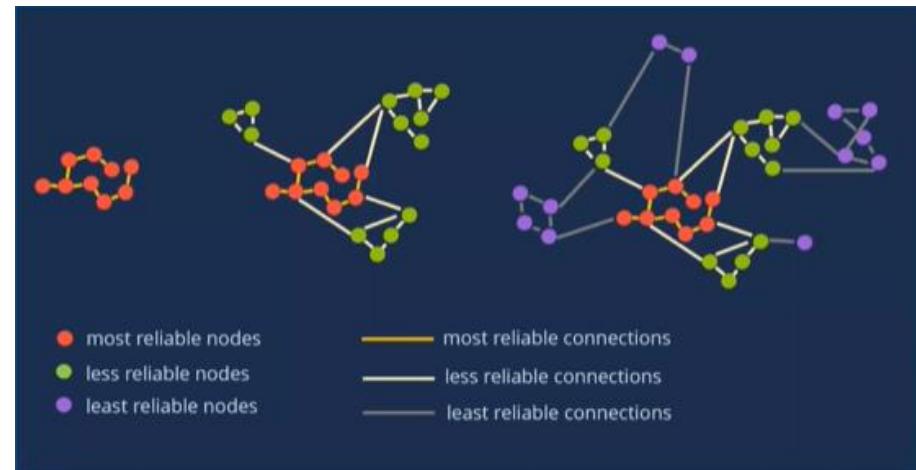
実際の報告書13M件には組織36M件、個人4.7M件（重複あり）が  
含まれていたが、ユニークな組織1M件、個人700K件を抽出できた

- 考察

多様な文書に含まれるエンティティ（組織、人）から構築したKGを分析  
することで、多様な曖昧性を効率よく解消できた（処理時間は記載なし）

- 知見

組織よりも個人の曖昧性解消がより難しい



“Seed”アルゴリズムの適用例



CLICKアルゴリズムの適用例

# Knowledge-Based News Event Analysis Toolkit

Oktie Hassanzadeh<sup>1</sup>, Parul Awasthy<sup>1</sup>, Ken Barker<sup>1</sup>, Onkar Bhardwaj<sup>1</sup>,

Debarun Bhattacharjya<sup>1</sup>, Mark Feblowitz<sup>1</sup>, Aamod Khatriwada<sup>1,3,†</sup>, Lee Martie<sup>1</sup>,

Steve Fonin Mbouadeu<sup>1,4,†</sup>, Jian Ni<sup>1</sup>, Anik Saha<sup>1,2,†</sup>, Sola Shirai<sup>1,2,†</sup>, Kavitha Srinivas<sup>1</sup>

and Lucy Yip<sup>1</sup>

<sup>1</sup>IBM Research, Yorktown Heights, NY, USA

<sup>2</sup>Rensselaer Polytechnic Institute, Troy, NY, United States

<sup>3</sup>Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA

<sup>4</sup>St. John's University, NY, USA

- 一言でいうと

報道で取り上げられるイベント情報を解析する Toolkitを開発

- 動機

過去のイベントおよび因果関係を学習し、新しいイベントを予測する

- 課題

イベントと因果関係の抽出および予測の自動化。具体的な記載なし

- 手法

Wikidataを知識として利用し、ニューラルベースの教師あり学習によるイベントおよび因果関係の抽出機能を組み込んだ Toolkitを開発。主要な機能は右図を参照（説明無し）

- 評価

記載なし

- 考察

記載なし

- 知見

主要な研究課題の提示のみ（詳細は文献[1]を参照）

- Wikidataを知識として利用した場合の網羅性、およびWikipediaからの知識獲得
- 見出しから知識へのマッピング自動化
- 知識グラフを拡充するルールベース手法と知識グラフ埋込み手法の性能比較
- イベント順序関係の抽出

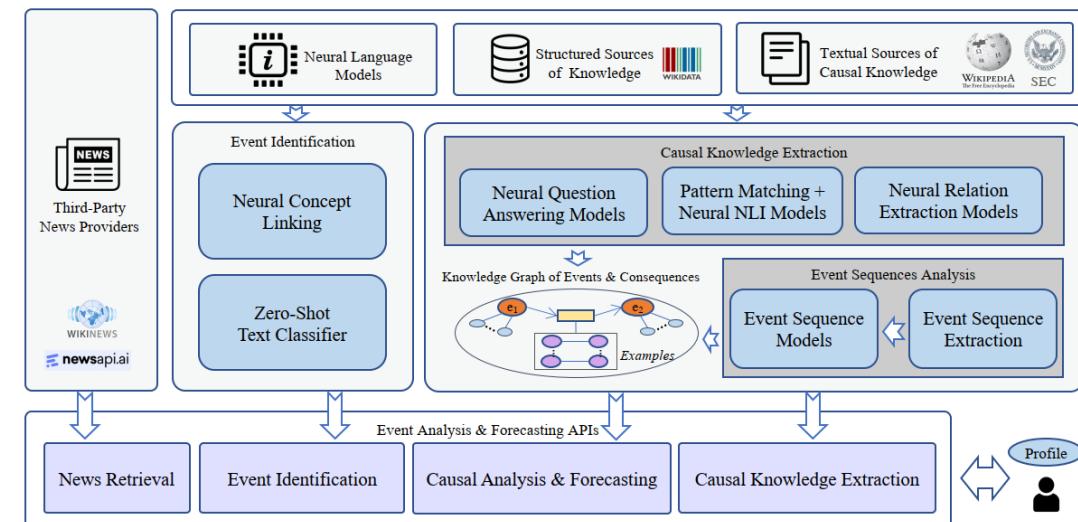


Figure 1: Toolkit Components and APIs

Toolkitのシステム構成

# RDF2TigerGraph: Towards Supporting RDF in TigerGraph Property Graph Database

Lu Zhou, Jay Yu

Innovation and Development Center, Tigergraph, Inc., 3636 Nobel Dr. Suite 100 San Diego, CA 92122, USA

- 一言でいうと  
Property Graph (PG) データベース上でRDFデータを扱う汎用的な方式を提案
- 動機  
RDFの表現力とPGのスケーラビリティの間にトレードオフがあり併用できない
- 課題
  - 既存実装の1つであるNeo4jのneosemanticsプラグインは機能が限定的
- 手法  
Tigergraph社のPGデータベースであるTigerGraphへRDFデータを取り込むためのスキーマおよびマッピングルールを定義。詳細な記載なし。マッピングルールおよびマッピング例は右図を参照
- 評価
  - RDFデータベースのLDBC SPBベンチマーク(\*1)を利用
  - SPARQLクエリをGSQL(\*2)クエリへ手作業で変換(右図を参照), 両クエリの結果が同じであることを確認
  - パフォーマンスはRDFデータベースに匹敵(比較対象の記載なし)
- \*1: <https://ldbcouncil.org/benchmarks/spb/> \*2: TigerGraphのグラフクエリ言語
- 考察  
blank node, named graph, RDFS/OWL reasoning, SPARQLのconstructとupdateなどに未対応
- 知見  
記載なし

## Example Graphs in Diagram



Figure 1. Example triples in RDF Graph

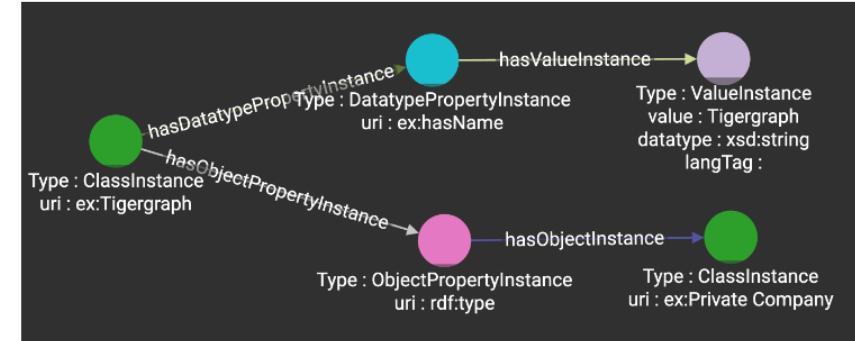


Figure 2. Example triples in Tigergraph

## SPARQL to GSQL Translation

Mappings and Examples:

Mapping Rules	SPARQL	GSQL	SPARQL Examples	GSQL Examples
Rule 1	SELECT	SELECT	SELECT ?cwork	SELECT s.uri as cwork INTO T
Rule 2	WHERE	FROM & WHERE	WHERE {?cwork cwork:dateModified ?dateModified. }	FROM ClassInstance:s - (hasDatatypePropertyInstance>:e) - DatatypePropertyInstance:t - (hasValueInstance>:e1) - ValueInstance:t1 WHERE t.uri == "cwork:dateModified"

[https://github.com/luzhoutg/ldbc\\_spb\\_tigergraph](https://github.com/luzhoutg/ldbc_spb_tigergraph)

# The Knowledge Graph Lifecycle in NTT DATA

Javier Flores<sup>1,\*</sup>, Emmanuel Jamin<sup>2</sup>, Sergi Nadal<sup>1</sup> and Oscar Romero<sup>1</sup>

<sup>1</sup>Universitat Politècnica de Catalunya, Barcelona, Spain

<sup>2</sup>NTT Data, Barcelona, Spain

- 一言でいうと  
　　歐州助成プログラムによる研究プロジェクトの成果情報を、持続可能な知識グラフ（CORDIS KG）として構築・公開する
- 動機
  - CORDISポータルサイトが持つ欧州プロジェクトの研究成果データへのアクセスと再利用を容易にする
  - 他のデータプロジェクトにとって参照となるKGライフサイクルを設計する
- 課題
  - 手作業によるタスクの自動化、および持続可能なライフサイクルの確立
- 手法
 

下図のKGライフサイクルを構築

  - Collection: EURIOオントロジを利用、RMLマッピングでRDFデータ化
  - Clean: 独自の品質ルール+SHACL+手作業
  - Enrich: CRF-NERを利用して非構造化データから組織名や人名を抽出
  - Link: アライメントツールを利用してKGを拡張。精度が出ないものは、手作業やgraph data profileを併用
  - Refactor: PAVオントロジからヒントを得て、KGのバージョンニング機構を考案
  - Publish: FOOPS! を利用して、FAIR原則に対する語彙やオントロジの適合性を評価
- 評価
 

記載なし。EUが2022年12月にCORDIS KGのSPARQLエンドポイントを公開予定
- 考察
 

記載なし
- 知見
 

記載なし

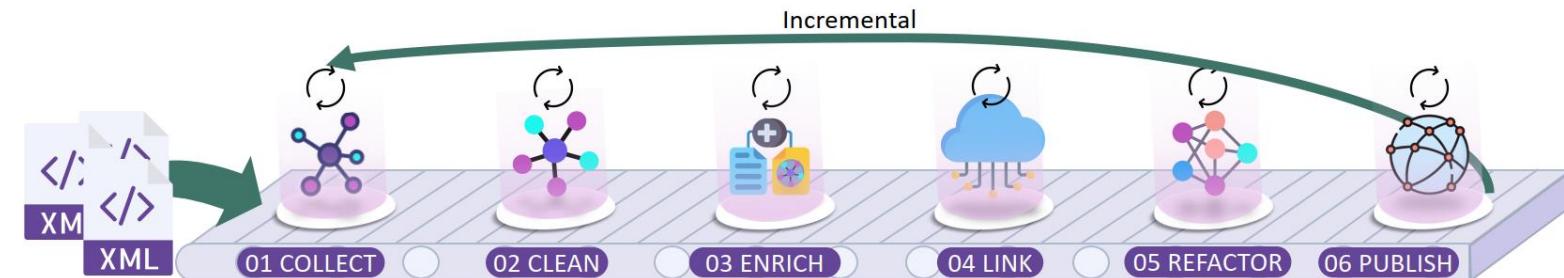


Figure 1: Overview of the implemented KG lifecycle in CORDIS.

# INDUSTRY TRACK 2

# Versioned Objects

Dag Hovland, Fredrik Chrislock

Bouvet Norway, Bergen, Norway

- 一言でいうと

石油ガス開発において、膨大かつ複雑な契約や業務に関する情報交換をRDFデータを利用して効率化する

- 動機

石油ガス開発において請負業者と作業当事者の間で生じる、契約、設備、業務等に関する膨大かつ複雑な情報交換を、完全な文書情報の単位から、小さな情報の単位へと移行し、業務のデジタル化／効率化を進める

- 課題

RDFは有力な候補であるが、単一の巨大なRDFデータセットを組織間で共有し編集履歴を管理することは解とはならない

- 請負業者も作業当事者も完全グラフの共有を望まない
- 各組織が異なる選択を検討できるように複数バージョンの維持が必要

- 手法

- 永続的に扱うオブジェクトと、オブジェクトが持つバージョン化された情報とを区別して扱う。
- IRIの形式：{永続ベース IRI}/バージョン/{オブジェクトのハッシュのバイト数}/{任意のバージョン番号付けシステム}
- 組織間では、変更のあったオブジェクト単位でRDFデータを送信する
- <https://github.com/equinor/versioned-object>

- 評価

記載なし

- 考察

記載なし

- 知見

記載なし

# Rapid Explainability for Skill Description Learning

Caglar Demir<sup>1</sup>, Anna Himmelhuber<sup>2,3</sup>, Yushan Liu<sup>2,4</sup>, Alexander Bigerl<sup>1</sup>,  
Diego Moussallem<sup>1</sup> and Axel-Cyrille Ngonga Ngomo<sup>1</sup>

<sup>1</sup>Paderborn University, Paderborn, Germany

<sup>2</sup>Siemens AG, Munich, Germany

<sup>3</sup>Technical University of Munich, Munich, Germany

<sup>4</sup>Ludwig Maximilian University of Munich, Munich, Germany

- 一言でいうと

- 工場の自動化に向け、機械が持つスキルの記述を深層強化学習により獲得する
- ドイツ連邦経済エネルギー省 (BMWi) ファンドによるRAKIプロジェクトの研究成果

- 動機

- スキルマッチングを利用して、工場における機械の自律化を実現する
- スキルマッチングは機械に対し製造工程における操作の割当てを行う

- 課題

- 機械が持つスキルの記述、および操作に対するスキル要件の記述の作成には、専門家による高度な知識と作業時間をする

- 手法

- クラス表現学習 (CEL) (文献[2]) として定式化し、RDF知識ベースと入力例からILPと深層学習を利用して記述論理の概念を獲得する
- 深層強化学習の利用による学習を高速化 (文献[5,6,7])
- 言語化モジュールの利用による解釈可能な結果を提示

- 評価

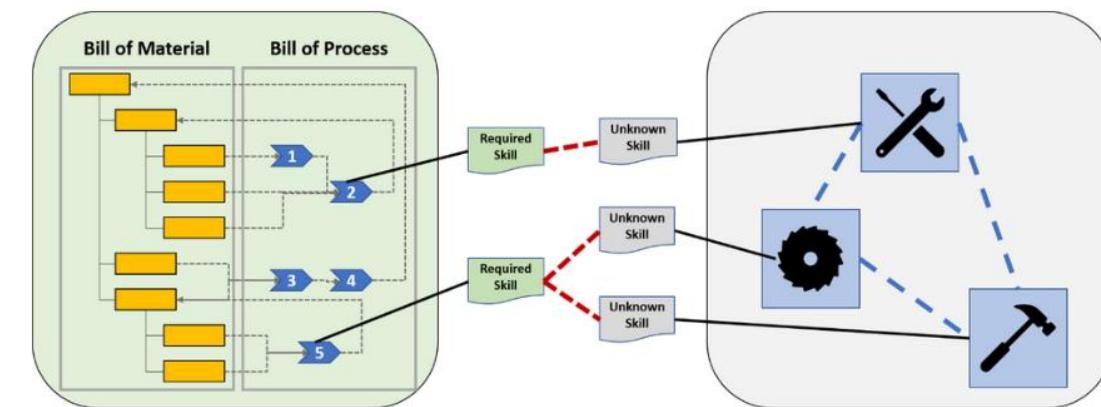
- 既存手法と比較して、高精度で概念を獲得でき、処理時間を大幅に短縮 (比較対象、および具体的な実験結果データの記載なし)
- 解釈性の評価なし (プラント運転員による理解の容易化を期待)

- 考察

記載なし

- 知見

記載なし



<https://raki-projekt.de/>

# Neuro-Symbolic AI at Bosch: Data Foundation, Insights, and Deployment

Baifan Zhou<sup>1,\*</sup>, Zhipeng Tan<sup>2,3</sup>, Zhuoxun Zheng<sup>2,4</sup>, Dongzhuoran Zhou<sup>2,1</sup>,  
 Yunjie He<sup>2,5</sup>, Yuqicheng Zhu<sup>2,5</sup>, Muhammad Yahya<sup>6</sup>, Trung-Kien Tran<sup>2</sup>,  
 Daria Stepanova<sup>2</sup>, Mohamed H. Gad-Elrab<sup>2</sup> and Evgeny Kharlamov<sup>2,1</sup>

<sup>1</sup>SIRIUS Centre, University of Oslo, Norway

<sup>2</sup>Bosch Center for Artificial Intelligence, Germany

<sup>3</sup>RWTH Aachen University, Germany

<sup>4</sup>Department of Computer Science, Oslo Metropolitan University, Norway

<sup>5</sup>University of Stuttgart, Germany

<sup>6</sup>University of Galway, Ireland

- 一言でいうと

独ボッシュによるニューロシンボリックAI(\*1)の取組みを紹介

\*1:ニューラルネットワークを用いた深層学習と、記号的表現に基づくシンボリックAIの融合技術

- 動機

多種多様なデータの利用にもとづく産業のスマート化

- 課題

ニューロシンボリックAIに関する実践事例の不足

- 手法

- ドメインオントロジの整備に基づくアセットデータの統合、及び ontology reshaping (文献[4]) により生成したKGスキーマによるデータ処理の自動化
- ExecutableKG (文献[5]) の利用による、機械学習パイプラインの標準化
- Answer set programmingを利用した解空間の探索
- KG埋込みを利用したリンク予測によるニューロシンボリック学習の開発

- 評価

具体的な記載は無いが、実証事例3件を紹介：  
プロセス診断、品質モニタリング、従業員スケジューリング

- 考察

記載なし

- 知見

記載なし

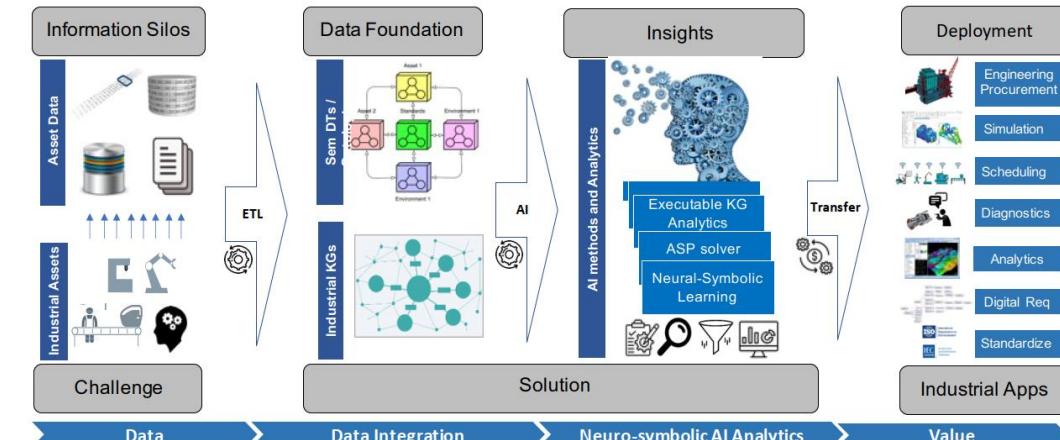


Figure 1: Overview: Semantic ETL creates the data foundation of AI-powered analytics, which extracts insights that can be transferred to value in industrial deployment.

## ニューロシンボリックAIの概要

# MetaverseKG: Knowledge Graph for Engineering and Design Application in Industrial Metaverse

Utkarshani Jaimini<sup>1,\*</sup>, Tongtao Zhang<sup>2</sup> and Georgia Olympia Brikis<sup>2</sup>

<sup>1</sup>Artificial Intelligence, University of South Carolina, Columbia, SC, USA

<sup>2</sup>Siemens Corporation, Technology, Princeton NJ, USA

- 一言でいうと  
産業用メタバース上で製品の設計製造のシミュレーションを可能にする知識グラフの提案

- 動機

- 産業用メタバースでは仮想空間に工場を再現しシミュレーションを行い、製品の設計情報や製造情報に関するメタデータが共通に利用される
- エンティティ（オブジェクトやプロセス）を仮想空間上で表現するには、多様なデータソースへのアクセス、及び異なるデータモデルとセマンティクスの協調が必要となる

- 課題

- 産業用メタバースのプラットフォーム（PF）に求められる相互運用性：  
a) セマンティクスの交換, b) リアルタイムな相互作用, c) 情報の同期, d) エンティティの再利用

- 手法

MetaverseKGの提案：a) 実世界と仮想世界をシームレスにつなぐセマンティックネットワークを構築, b) 異なるPF間の相互接続を促進, c) PF横断による開発を促進にする標準手段の提供

- 評価

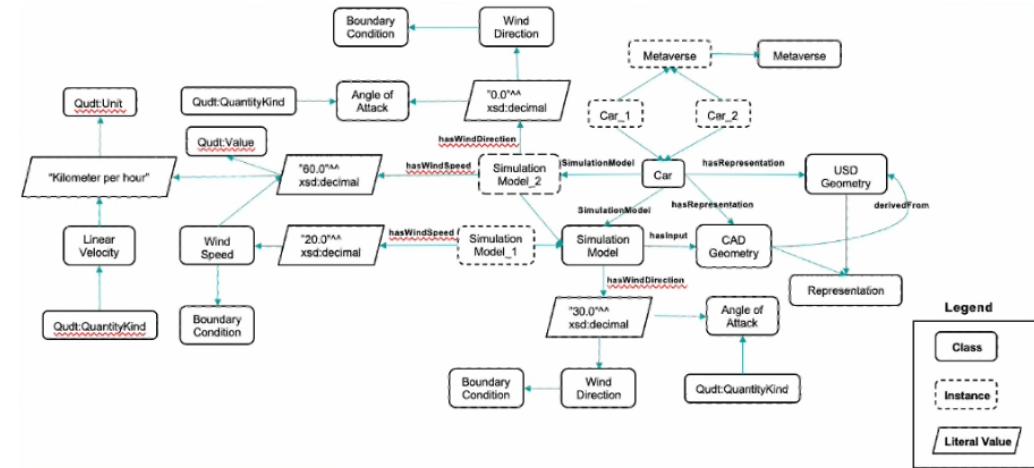
記載なし

- 考察

記載なし

- 知見

記載なし



MetaverseKGの一部

ご協力いただき  
ありがとうございました！