

ナレッジグラフ若手の会  
**ESWC2022サーベイ会**

2022/08/24 オンライン

# 担当者一覧

発表順	担当者	担当セッション
1	産総研 江上	Embeddings
2	富士通 鵜飼	ML & Knowledge Graph
3	DBCLS 山本	Linked Data Analysis
4	DBCLS 山本	Reasoning
5	東芝 藤原	Reasoning Systems
6	富士通 小柳	Software Development
7	東芝 砂川	Dealing With Multiple Sources
8	静岡大 福田研	Domain Specific Knowledge Graph
9	青学大 森田研	Knowledge Graph for Tasks
10	大阪電通大 古崎研	Multiple Modalities
11	東芝 長野	Industry Track 1
12	東芝 長野	Industry Track 2

# Embeddings

# Embeddings

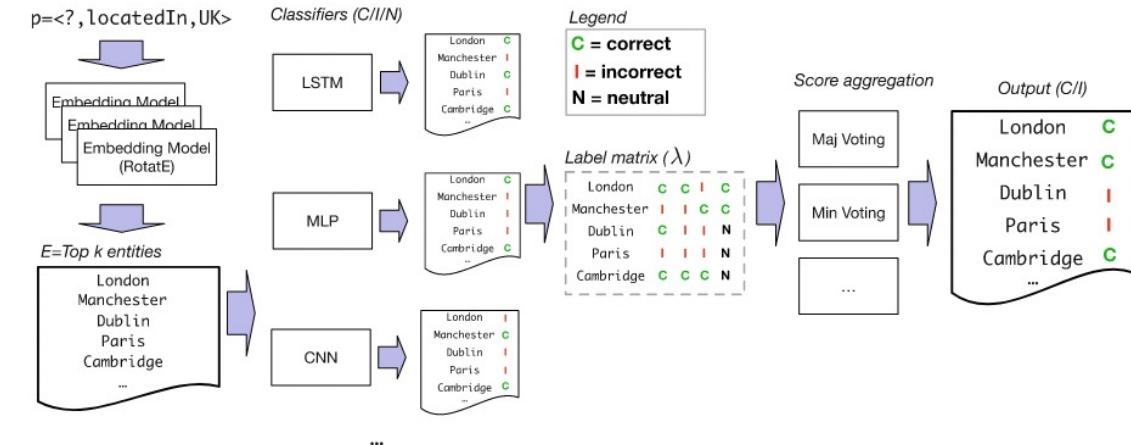
ESWC2022サーベイ会

江上周作/産業技術総合研究所

# Ensemble-Based Fact Classification with Knowledge Graph Embeddings

Unmesh Joshi and Jacopo Urbani/Department of Computer Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

- 一言でいうと  
KG埋め込みとアンサンブル学習により、リンク予測結果の真偽を高精度に分類する手法を提案
- 動機  
KG埋め込み (KGE) モデルを用いてリンク予測したものの中、本当に正しいものはごく一部であるためファクト分類が必要
- 手法
  - ・ KGEとアンサンブル学習を用いてリンク予測結果のファクト分類を行うモデル「DuEL」を提案：LSTM, MLP, CNNを利用
  - ・ CWAで学習することで再現性が損なわれる問題があるため、2つの教師なし分類器を追加：サブグラフ埋め込み、共有パスを利用
  - ・ <https://github.com/karmaresearch/duel>
- 背景  
リンク予測結果のファクト分類を行う研究はあまりされていない
- 結果  
既存手法RankClassifyと比べて大幅に（使用するKGEモデルによっては72%）F1スコアが向上。FB15k237で0.60, DBpedia50で0.51.
- 考察  
KGEはComplExが最も精度が良く、TransEは最も悪い通説通りの結果となつたため、一般化できる結果だと結論づけた。その他Ablation Studyやハイパーパラメータチューニング結果を考察
- 課題  
オントロジーの制約に基づく分類。バイアスを減らす。異なるKGEで生成されたランキングを組み合わせる。コーパス知識の取り込み。



**Fig. 1.** Schematic overview of DuEL

# Dihedron Algebraic Embeddings for Spatio-temporal Knowledge Graph Completion

Mojtaba Nayyeri, Sahar Vahdati, Md Tansen Khan, Mirza Mohtashim Alam, Lisa Wenige, Andres Behrend, and Jens Lehmann/

University of Bonn, Bonn, Germany, Institute for Applied Informatics (IfAI), Dresden, Germany. Institute for Telecommunications (INT), TH Köln, Cologne, Germany. Fraunhofer IAIS, Dresden, Germany. University of Stuttgart, Stuttgart, Germany.

- 一言でいうと

空間的・時間的側面を学習するためにDihedron Algebraを用いたナレッジグラフの埋め込みモデルを提案

- 動機

KG埋め込み (KGE) モデルではしばしば時間と空間の側面の同時考慮が無視されている。

- 手法

- Dihedron (複素数を4次元空間に拡張した超複素数系) 代数に基づくST-NewDEを提案
- (主語, 述語, 目的語, 場所, 時間) のうち1つを欠損させた不完全5つ組 (クインタップル) とその答えに分割する
- 4つの値を固定した4次元空間の中ではDihedron空間は四元数よりも広い幾何学的形状をカバーするため表現力が高い

- 背景

既存の時間的KGEモデルは代数的側面を利用したものではなく、空間的側面を考慮するものではない。

- 結果

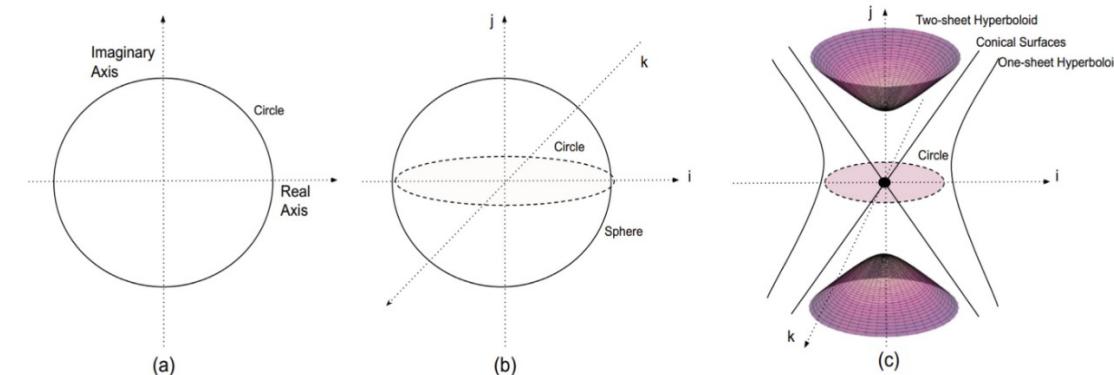
YAGO3K, DBpedia34K, WikiData53KのクインタップルのCompletionで最高精度を達成。その他クラスタリングやAblation studyで分析。

- 考察

時間・位置情報を効率的に利用する上では、平行移動・回転などの変換を組み合わせるだけでなく、時空間幾何学的表現も大きな影響を与える。

- 課題

未知エンティティの予測



**Fig. 2.** Illustration of (a) Complex plane; (b) Quaternion space; (c) Dihedrons space.

# DCWEB-SOBA: Deep Contextual Word Embeddings-Based Semi-Automatic Ontology Building for Aspect-Based Sentiment Classification

Mroos van Lookeren Campagne, David van Ommen, Mark Rademaker, Tom Teurlings, and Flavius Frasincar/Erasmus University Rotterdam, Rotterdam, The Netherlands

- 一言でいうと  
深い文脈に基づく単語埋め込みを用いて、領域感情オントロジーを半自動的に構築する
- 動機
  - ・ 感情分類のための最先端のモデルではドメインオントロジーを使用する。
  - ・ ドメイン固有の感情コーパスからオントロジーを半自動的に構築することで、感情分類の性能を向上させたい。
- 手法  
単語埋め込みを用いた領域感情オントロジーの半自動的構築手法の既存研究を拡張：(1) 多義性を解消するためにBERTから得られる深い文脈に基づく単語埋め込みを使用, (2) 副詞が感情を伝えると考え副詞を考慮, (3)感情に対応したBERT単語埋め込みの追加セットを使用, (4)新たな閾値関数を使用
- 背景
  - ・ 製品等のアスペクトを特定し、特定されたアスペクトに対する感情を測定するアスペクトベースの感情分析(ABSA)アルゴリズム
  - ・ ドメインオントロジーで推論を行い、推論ができないときはNeural Attention Modelでバックアップするハイブリッドモデルが最先端
- 結果  
提案手法DCWEB-SOBAは既存手法と比較して、カバー率の高いドメインオントロジーを構築できた。ユーザ時間は20分多くかかるが、BERTモデルの構築時間は180分短い。まとめると、かかる時間とカバレッジのバランスをうまく取れている。
- 考察  
ABSAのハイブリッド手法において、深い文脈を持つ単語埋め込みを用いると、文脈を持たない単語埋め込みに比べて性能が向上する
- 課題  
RoBERTaの利用。ファインチューニングされたモデルにおいて、pos/negがアスペクトと文脈に依存するType-3感情語を分離して学習。

# ML & Knowledge Graph

ESWC2022サーベイ会 論文紹介資料

# Section: ML & Knowledge Graphs

担当：鵜飼孝典（富士通株式会社）

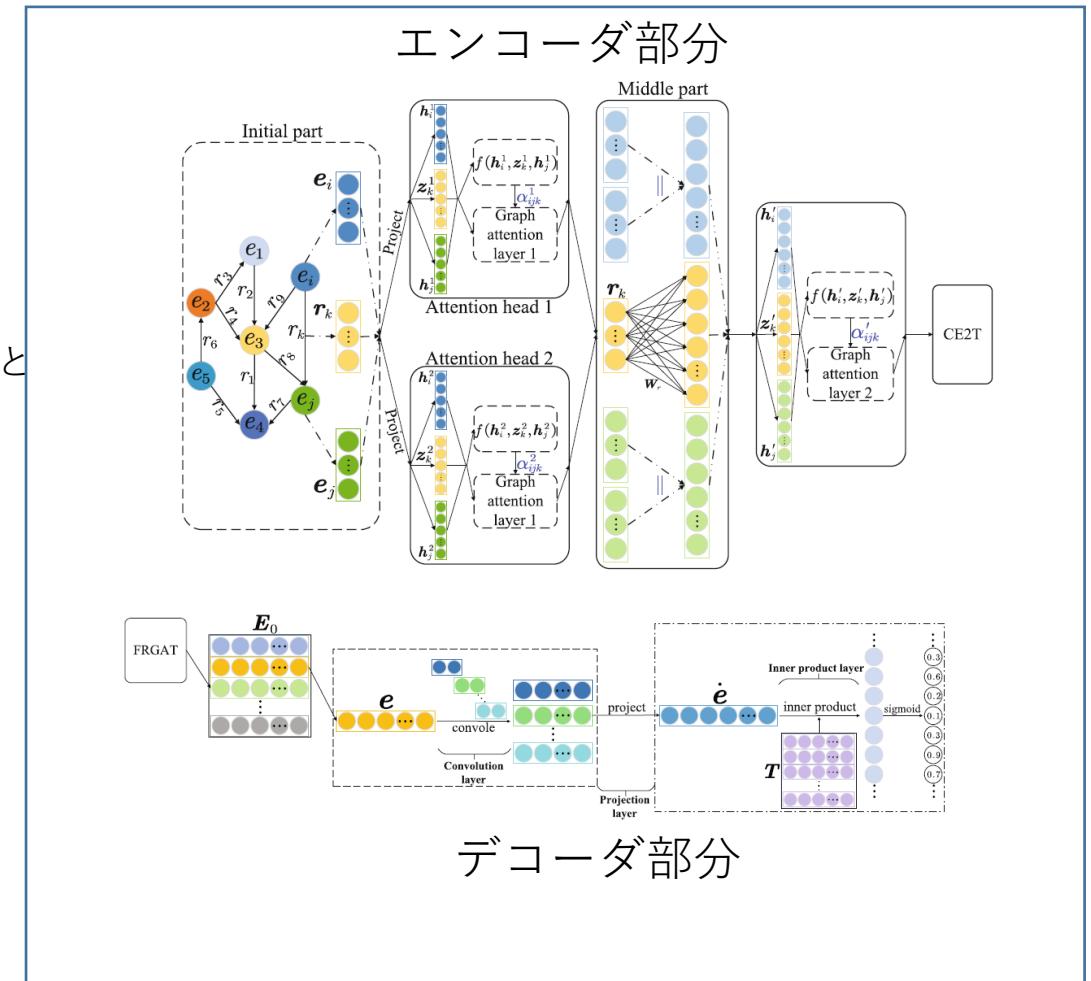
# Knowledge Graph Entity Type Prediction with Relational Aggregation Graph Attention Network

Changlong Zou, Jingmin An, Guanyu Li (Dalian Maritime University)

- 一言でいうと  
エンティティのタイプを予測するモデルを提案
- 動機  
多くの知識グラフはエンティティタイプが欠落している
- 手法  
Relational Aggregation Graph Attention Network を利用したエンコーダと  
デコーダによる予測
- 背景  
エンティティ間のリレーションが有効に利用されないが、  
リレーションには豊富な意味情報が含まれていることが多い。
- 結果  
既存の11の手法と比較してもっともよい精度
- 考察  
ノードにつながっているリレーションの数は、10くらいで精度がほぼ  
最高になり、それ以上つながっていても精度は高くならない
- 課題  
ノードの類似性を示すために分散表現に提案モデルを適用したい

コード <https://github.com/GentlebreezeZ/RACE2T>

動画 [http://videolectures.net/eswc2022\\_zou\\_knowledge\\_graph/](http://videolectures.net/eswc2022_zou_knowledge_graph/)



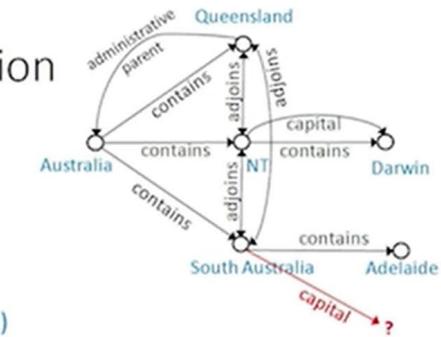
# Supervised Knowledge Aggregation for Knowledge Graph Completion

Patrick Betz, Christian Meilicke, Heiner Stuckenschmidt (University of Mannheim, Germany)

- 一言でいうと  
リンク予測を行うルールを機械学習で作り出す
- 動機  
ナレッジグラフは知識処理可能なので、ルールベースにすれば、明確に説明可能
- 手法  
(s,p,o)をアグリゲーションすることで、可能な限り一般的なルールを作る
- 結果  
生物学のナレッジグラフの予測において、既存の手法より良い精度を得た
- 考察  
副産物的にスパースなルールをアグリゲートする手法を作ることができた
- 課題  
この論文では、説明性について議論できていない

## Knowledge Graph Completion

- Knowledge Graph Embeddings (KGE)
- Rule-based approaches (e.g. AnyBURL\*)
  - $\text{hypernym}(X,Y) \leftarrow \text{hyponym}(Y,X)$
  - $\text{contains}(\text{Australia},Y) \leftarrow \text{contains}(\text{Victoria},Y)$
  - $\text{speaks}(X,\text{English}) \leftarrow \text{livesIn}(X,\text{London})$

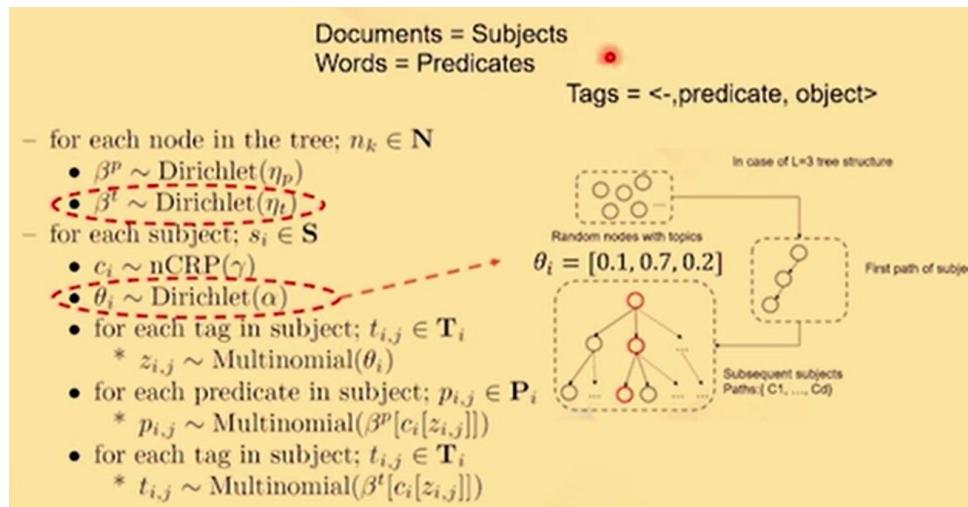


Approach	FB15k-237			WNRR			Codex-M		
	h@1	h@10	MRR	h@1	h@10	MRR	h@1	h@10	MRR
TransE	0.221	0.497	0.312	0.053	0.520	0.228	0.223	0.454	0.303
ComplEx	0.253	0.536	0.347	0.438	0.547	0.475	0.262	0.476	0.337
ConvE	0.248	0.521	0.338	0.411	0.505	0.442	0.239	0.464	0.318
RESCAL	0.263	0.541	0.355	0.439	0.517	0.467	0.244	0.456	0.317
Hitter <sub>nc</sub>	0.268	0.549	0.361	0.437	0.531	0.469	0.262	0.486	0.339
RotatE	0.240	0.522	0.333	0.439	0.553	0.478	—	—	—
$M^2GNN$	0.275	0.565	0.362	0.444	0.572	0.485	—	—	—
pLogicNet	0.237	0.524	0.332	0.398	0.537	0.441	—	—	—
MINERVA	0.217	0.456	0.293	0.413	0.513	0.448	—	—	—
DRUM [35]	0.255	0.516	0.343	0.425	0.586	0.486	—	—	—
Neural-LP [45]	—	0.362	0.240	0.371	0.566	0.435	—	—	—
Max	0.246	0.506	0.331	0.457	0.572	0.497	0.247	0.450	0.316
Noisy-or	0.247	0.494	0.329	0.391	0.559	0.446	0.218	0.427	0.289
Dense	0.245	0.510	0.335	0.466	0.587	0.507	0.261	0.465	0.331
Sparse	0.266	0.526	0.352	0.459	0.574	0.499	0.266	0.467	0.335
D+S	0.267	0.527	0.354	0.469	0.593	0.511	0.273	0.476	0.342

# Hierarchical Topic Modelling for Knowledge Graphs

Yujia, Marcin Pietrasik, Wenjie, Marek Reformat (University of Alberta)

- 一言でいうと  
トピックモデリングに用いられる確率的手法を利用した知識グラフの  
ノンパラメトリック階層的生成モデルの提案
- 動機  
知識グラフの潜在的な構造を明らかにするためのモデリングが盛んに行われている
- 手法  
Subjectを文書, predicateをwordと考えてトピックモデルの手法を適用
- 結果  
既存の階層型クラスタリング手法と同等



Method	FB15k-237		YAGO3-10		DBpedia	
	ARI	NMI	ARI	NMI	ARI	NMI
RDF2VEC						
K-means	.308 ± .012	.567 ± .007	.070 ± .019	.199 ± .017	.223 ± .005	.416 ± .005
OPTICS	.087 ± .000	.283 ± .000	.009 ± .000	.172 ± .000	.001 ± .000	.311 ± .000
Agglom.	.455 ± .000	.601 ± .000	.038 ± .000	.174 ± .000	.236 ± .000	.414 ± .000
Spectral	.539 ± .000	.678 ± .000	.071 ± .000	.218 ± .000	.218 ± .000	.410 ± .000
TransE						
K-means	.405 ± .049	.632 ± .009	.263 ± .009	.367 ± .003	.247 ± .029	.389 ± .024
OPTICS	.031 ± .000	.253 ± .000	.049 ± .000	.150 ± .000	.001 ± .000	.198 ± .000
Agglom.	.491 ± .000	.599 ± .000	.226 ± .000	.337 ± .000	.198 ± .000	.383 ± .000
Spectral	.658 ± .000	.684 ± .000	.270 ± .000	.345 ± .000	.057 ± .000	.321 ± .000
DistMult						
K-means	.269 ± .011	.559 ± .013	.174 ± .012	.326 ± .015	.400 ± .008	.587 ± .010
OPTICS	.016 ± .000	.189 ± .000	.029 ± .000	.175 ± .000	.002 ± .000	.184 ± .000
Agglom.	.379 ± .000	.621 ± .000	.202 ± .000	.382 ± .000	.389 ± .000	.594 ± .000
Spectral	.505 ± .000	.600 ± .000	.035 ± .000	.124 ± .000	.150 ± .000	.478 ± .000
ComplEx						
K-means	.271 ± .020	.562 ± .016	.137 ± .012	.342 ± .009	.462 ± .013	.630 ± .015
OPTICS	.019 ± .000	.202 ± .000	.017 ± .000	.152 ± .000	.002 ± .000	.235 ± .000
Agglom.	.385 ± .000	.630 ± .000	.181 ± .000	.299 ± .000	.442 ± .000	.628 ± .000
Spectral	.563 ± .000	.613 ± .000	.016 ± .000	.204 ± .000	.203 ± .000	.550 ± .000
ConvE						
K-means	.332 ± .031	.619 ± .013	.004 ± .003	.004 ± .001	.474 ± .019	.612 ± .013
OPTICS	.040 ± .000	.254 ± .000	.012 ± .000	.088 ± .000	.002 ± .000	.238 ± .000
Agglom.	.384 ± .000	.630 ± .000	.003 ± .000	.005 ± .000	.458 ± .000	.614 ± .000
Spectral	.556 ± .000	.703 ± .000	.002 ± .000	.006 ± .000	.439 ± .000	.639 ± .000
ExCut	.343 ± .011	.651 ± .002	.130 ± .007	.322 ± .011	.380 ± .016	.595 ± .005
Our Method	.656 ± .005	.669 ± .021	.044 ± .006	.218 ± .002	.406 ± .042	.582 ± .022

動画 : [http://videolectures.net/eswc2022\\_zhang\\_topic\\_modelling/](http://videolectures.net/eswc2022_zhang_topic_modelling/)

# Linked Data Analysis

# An Analysis of Links in Wikidata

Haller, A., Polleres, A., Dobriy, D., Ferranti, N., Rodríguez Méndez, S.J.

- 一言でいうと  
Wikidataとその他のデータセットの間で結ばれているリンクについて調査
- 動機  
様々な一次データベースとのリンクが増えているので、実態を調査したい
- 手法  
幾つかのリンクに対する仮説を立てて、それを実際に解析した
- 背景  
Linked DataのハブとしてWikidataが機能していると想定されるので
- 結果  
一次データソースからWikidataへのリンクは多数あるが、逆が少ない
- 考察  
現状、WikidataはLinked Dataのハブとまでは言えない。
- 課題  
Linked Dataのハブとして機能するために、Wikidataから一次データソースへのリンクをもっと充実すべきではないか。また、より多くのリンクを加えやすいUIも検討すべき。

# The Problem with XSD Binary Floating Point Datatypes in RDF

Keil, J.M., Gänßinger, M.

- 一言でいうと  
RDFデータで数値を表すXSD名前空間の型は広く使われているが、特に実数値については注意して利用しないと意図しない結果になりかねない
- 動機  
実数値の型を示すxsd:decimal、xsd:float、xsd:doubleについては広く使われているが、これらの違いと、生じうる問題について調査が必要である
- 手法  
195億程度のリテラル値を含む2020 Web Data Commonsデータセットを対象として上記データ型の利用動向を調査
- 背景  
xsd:floatはメモリ空間や処理速度を考慮した表現方法で、誤差が伴うが、RDFでは正確な実数値を表す型として定められており、より正確に表現できるxsd:decimalの利用が比較的少ない
- 結果  
xsd:floatが広く使われている
- 考察  
複数のxsd:float型の数値をかけるなど処理すると丸め誤差が大きくなる
- 課題  
RDF(Turtle)では小数点を含む数値表現を、形無しで記載すると、xsd:floatとして扱われるが、これは望ましくないとして今後コミュニティで対応を促す

# Reasoning

# Union and Intersection of all Justifications

Chen, J., Ma, Y., Peñaloza, R., Yang, H.

- **一言でいうと**  
対象オントロジーにおいて、すべての弁明の和と積を高速に計算する手法を提案
- **動機**  
データもオントロジーも大きくなっている中で、高速に弁明を計算する手法が必要
- **手法**  
既に計算済みの情報を使うなどしつつ、全体ではなく一部のみを処理して高速に結果を得る
- **背景**  
これまで、一つもしくは全ての弁明を効率的に取得する手法について研究されてきたが、全ての弁明を列挙することなく、和と積を効率的に取得する手法については研究されてこなかった
- **結果**  
ベンチマーク用のデータを用いて既存の手法と比較し、高速化を確認
- **考察**  
弁明を得るのに必要な推論エンジンは別途必要であるが、和と積を取得する処理の効率化は確認できた
- **課題**  
より大きなオントロジーに対しても評価を行う

# Optimal ABox Repair w.r.t. Static EL TBoxes: From Quantified ABoxes Back to ABoxes

Baader, F., Koopmann, P., Kriegel, F., Nuradiansyah, A.

- 一言でいうと  
DLにおいてエラーがある場合、ABoxの最適な修正の有無を指数時間内に計算する手法を提案
- 動機  
DLにおけるエラーを修正するためには、その原因となっている論理的帰結を取り除く必要があるが、削除する帰結をなるべく少なくしたい
- 手法  
必要最小限の帰結を削除する最適なABoxの修正法は常にあるとは限らないので、その有無を指数時間に提示すると共に当該ABoxを指数時間で取得する
- 背景  
エラーを修正する従来手法は、必要以上に論理的帰結を取り除いてしまう
- 結果  
論理的な証明と、同じ著者らによる以前の研究で利用したデータを用いて、提案手法の妥当性を確認
- 考察  
提案手法の問題点である、全てのABoxの修正が必要になる事例は実際のデータを調査することで、多くは生じないことが確認できた
- 課題  
ABoxの修正とは異なる修正セマンティクスを採用することを検討する

# Learning Concept Lengths Accelerates Concept Learning in ALC

Kouagou, N.J., Heindorf, S., Demir, C., Ngomo, AC.N.

- 一言でいうと  
DL (ALC) における概念学習の時間を削減する手法を提案
- 動機  
現状のDLの概念学習法はデータが大きくなると遅くて実用性が下がるのを何とかしたい
- 手法  
概念長をNNで予想することで探索空間を狭め、概念学習にかかる時間を短縮
- 背景  
これまで概念長を予想して探索空間を狭める手法は提案されていなかった
- 結果  
ベンチマークデータで、ランダムモデルや既存手法に対する有効性を確認
- 考察  
概念長が11までなら提案手法が高効率だが、それより長い場合は不明なほか、一般的な埋込手法では局所解に陥る可能性もある
- 課題  
与えられた知識グラフと、長い概念長の双方を合わせて学習する手法に取り組む

# Reasoning Systems

# CONSTRUCT queries performance on a Spark-based Big RDF triplestore

Adam Sanchez-Ayte, Fabrice Jouanot, and Marie-Christine Rousset / Université Grenoble Alpes, France

- **一言でいうと**

大規模ナレッジグラフに対してCONSTRUCTクエリをサポートするアーキテクチャ TESSの設計と実装

- **Resource Type**

software tools/services

- **動機**

大規模トリプルストアにおけるCONSTRUCTクエリのボトルネックの解消

- **リソースの設計方針**

並列処理、Sparkをベースとしたアーキテクチャ

- **新規性**

データセットの縮小やマージを可能とするが、  
あまり着目されていなかったCONSTRUCTクエリの効率化

- **リソースの再現性／活用実績**

実験のdocker環境が公開されている(可用性参照)

- **品質**

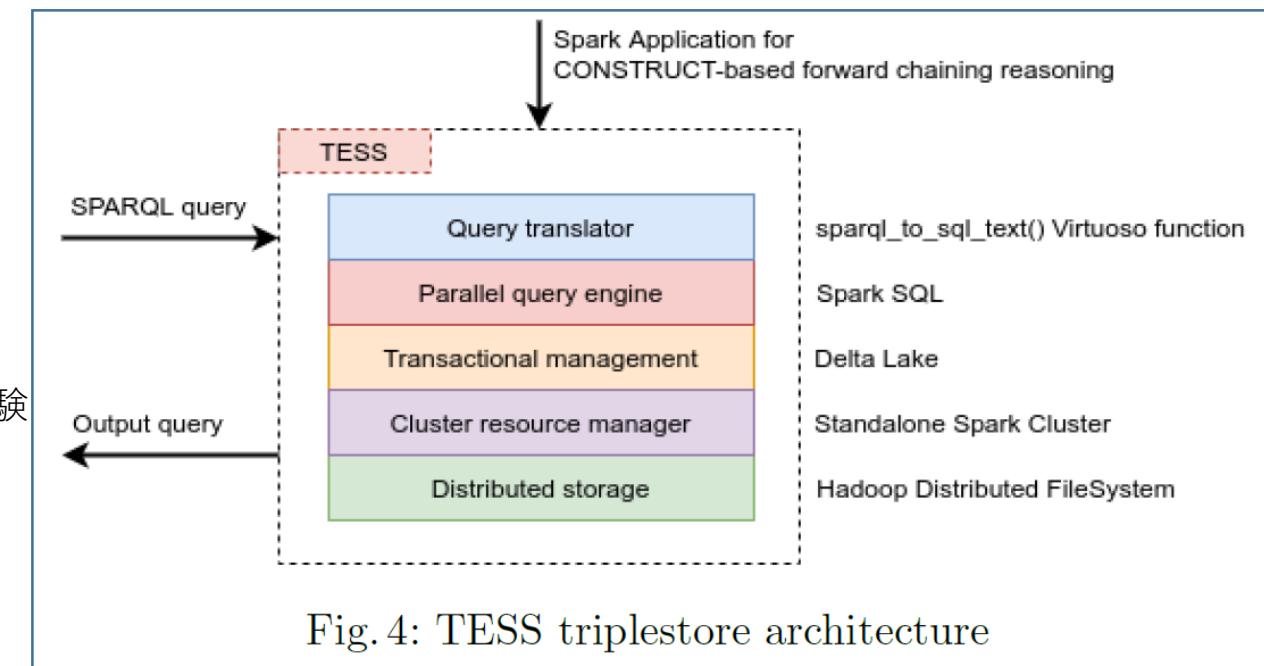
既存のトリプルストア(Virtuoso、GraphDB)を使った評価実験

- **リソースの可用性**

<https://github.com/asanchez75/ontosides-bpe>

- **次にすべきこと**

CPU/GPUベースの比較、パフォーマンス調整



# Stream Reasoning Playground

Patrik Schneider<sup>1;2;\*</sup>, Daniel Alvarez-Coello<sup>3;4</sup>, Anh Le-Tuan<sup>5</sup>, Manh Nguyen-Duc<sup>5</sup>, and Danh Le-Phuoc<sup>5</sup>

1 Vienna University of Technology; Vienna, Austria 2 Siemens AG Österreich; Vienna, Austria 3 University of Oldenburg; Oldenburg, Germany  
4 BMW Technologies E/E Architecture, Wire Harness; Garching, Germany 5 Technical University of Berlin; Berlin, Germany

- **一言でいうと**

ストリーム推論(SR: Stream Reasoning)のアプローチを比較、評価するためのPlaygroundを提供

- **Resource Type**

software tools/services

- **動機**

アプローチごとのデータモデルと推論タスクの制限を取り扱う

- **リソースの設計方針**

6つの要件(Stream reasoning、Consumer agnosticism、Extensibility、Availability and agility、Base scenarios、Multiple tasks,)

- **新規性**

既存の手法は複雑なイベント検出ベースアプローチに単純適合できない

- **リソースの再現性／活用実績**

実験のdocker環境が公開されている(可用性参照)

- **品質**

SR Hackathon 2021 に基づくケーススタディで評価

- **リソースの可用性**

<https://github.com/patrik999/stream-reasoning-challenge>

- **次にすべきこと**

ベンチマーク機能、確率を含むストリームと推論への拡張、シナリオとデータセットの拡張、新しいフォーマッターの追加

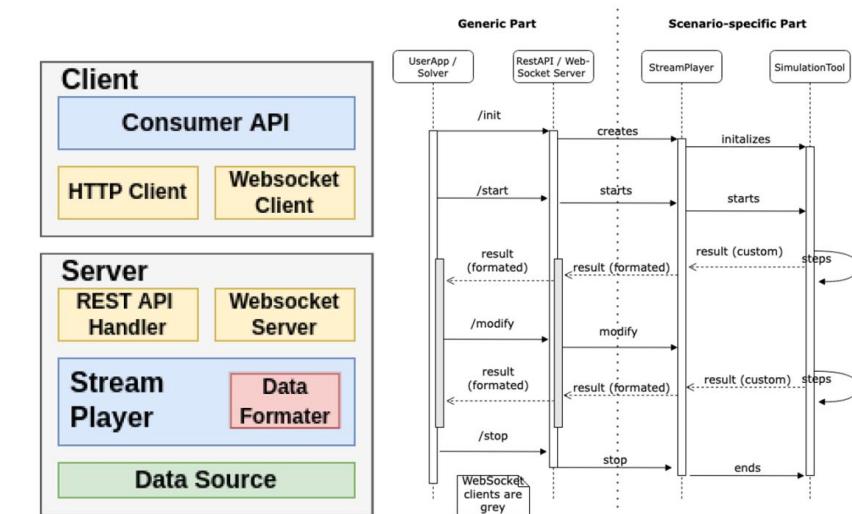


Fig. 1: (a) Overview of architecture and (b) interaction between the components, where the data source is a traffic simulation tool.

# Do Arduinos dream of efficient reasoners?

Alexandre Bento<sup>1</sup>, Lionel M'edini<sup>1</sup>, Kamal Singh<sup>2</sup>, and Frédéric Laforest<sup>1</sup>  
 1 Univ Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, F-69621 Villeurbanne, France  
 2 LaHC UMR CNRS 5516, Université Jean Monnet, Saint-Étienne, Franc

- **一言でいうと**  
Semantic Web Of Things(SWoT)実現のための軽量なインクリメンタル推論LiRotを提案

- **Resource Type**  
software tools/services

- **動機**  
リソースの制約が大きいエッジで推論を行う

- **リソースの設計方針**  
既存のRETEアルゴリズムをメモリに焦点をあてて最適化

- **新規性**  
RETEのメモリ使用量の大幅な改善、既存システムに対する優位性

- **リソースの再現性／活用実績**  
<https://gitlab.com/coswot/lirot-experiments-eswc-2022>

- **品質**  
既存の推論器(RDFox, Jena)と比較し、factの数が200～3000の場合に高速で、メモリ使用量は非常に少ない

- **リソースの可用性**  
<https://gitlab.com/coswot/lirot>

- **次にすべきこと**  
他のルールセットでの実験、様々な最適化との比較、分散型および協調型の推論アルゴリズムの調査

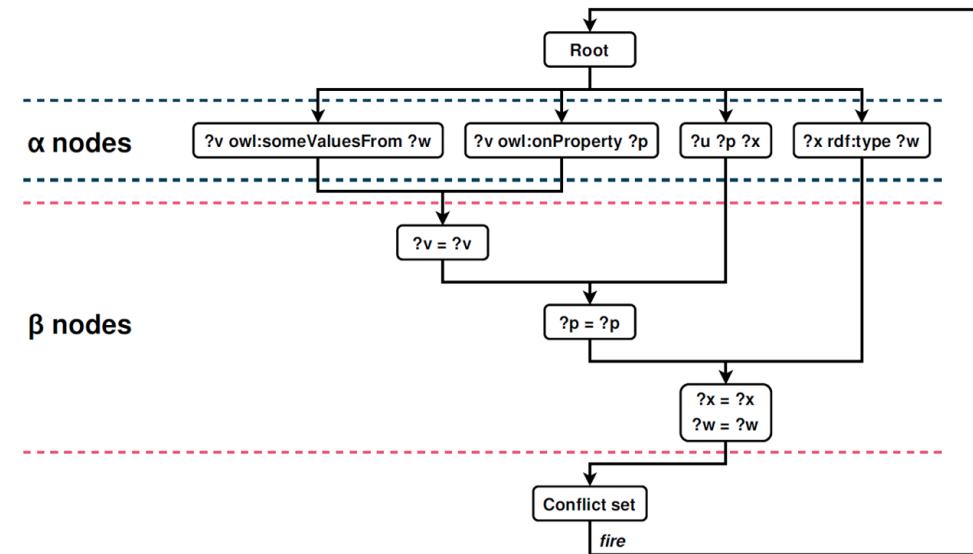


Fig. 1: Example of a RETE network for rule `cls-svf1` from the OWL 2 RL profile:  
 $(?v \text{ owl:someValuesFrom } ?w) \wedge (?v \text{ owl:onProperty } ?p) \wedge (?u \text{ ?p } ?x) \wedge (?x \text{ rdf:type } ?w) \rightarrow (?u \text{ rdf:type } ?v)$

# Software Development

# ESWC2022 サーベイ

富士通株式会社 小柳佑介

※資料内の図は原典からの引用です。

# Software Development:

- **Best Research Paper** Never Mind the Semantic Gap: Modular, Lazy and Safe Loading of RDF Data (Research)  
Eduard Kamburjan, Vidar Norstein Klungre, and Martin Giese
- A programming interface for creating data according to the SPAR Ontologies and the OpenCitations Data Model (Resources)  
Simone Persiani, Marilena Daquino, and Silvio Peroni
- Chowlk: from UML-based ontologyconceptualizations to OWL (Resources)  
Serge Chávez-Feria, Raúl García-Castro, and María Poveda-Villalón

# Never Mind the Semantic Gap: Modular, Lazy and Safe Loading of RDF Data

Eduard Kamburjan, Vidar Norstein Klungre, and Martin Giese/Department of Informatics, University of Oslo, Norway

[video](#), [GitHub](#)

- 一言でいうと  
KG(RDF)からデータをオブジェクト指向言語にロードするための新しい機構を提案 **※Best Research Paper**

- 動機  
RDFをオブジェクト指向言語のクラスにマッピングする際に生じる、インピーダンス・ミスマッチとセマンティックギャップの存在
- 手法  
RDF知識グラフからデータをオブジェクト指向言語にロードするための新しい機構を提案

- (1) クラスのインスタンスを取得するためにSPARQLクエリーを使用してクラス宣言に注釈を付与、(2) ネストされたクラス構造にfutureを使用  
オブジェクト指向とRDFの接続のためのLiskov原理を形式的に定義
- 背景  
OWLの概念をオブジェクト指向言語のクラスに直接関連付ける既存のアプローチは存在するが、インピーダンス・ミスマッチの課題は残る。

- 結果  
クラス数毎の、runtime overhead/メモリ使用量の評価
- 考察  
遅延評価が複雑なデータのロードに対して著しい性能向上をもたらすことを示した。
- 課題  
表現力を高めるために、主流の言語のために一般化。

## example

```

1 List<Nodes> it =
2 query("SELECT * WHERE { ?o :id ?id; :stamp ?stamp; :back ?w1; :front ?w2.
3   ?w1 :wheelId ?wId1; :stamp ?last1. ?w2 :wheelId ?wId2; :stamp ?last2.
4   FILTER(?wId1 != ?wId2).");
5 Int i = it.next().get("id"); //dynamic cast to Int
6 Bike bike = new Bike(i, ...);

```

Fig. 1. Dynamic data access with SPARQL.

```

1 class Wheel anchor ?o
2   (Int wheelId, Int last) //id is the id of the wheel, not the IRI
3 end retrieve SELECT ?wheelId ?last
4   WHERE{ ?o :wheelId ?wheelId; :stamp ?last. }
5 ... List<Wheel> it = load Wheel();

```

```

1 class Bike anchor ?o(
2   Int id, Int last, //id is the id of the bike, not the IRI
3   link(?o :front ?front) QFut<Wheel> front,
4   link(?o :back ?back) QFut<Wheel> back)
5 end retrieve SELECT ?id ?last WHERE { ?o :id ?id; :stamp ?last. }
6 ...
7 List<Bike> it = load Bike(); Bike bike = it[0];
8 List<Wheel> it = load bike.front;
9 Wheel w = it[0];

```

# A programming interface for creating data according to the SPAR Ontologies and the OpenCitations Data Model

[video](#), [GitHub](#)

Simone Persiani<sup>1</sup>, Marilena Daquino<sup>2,3</sup>, and Silvio Peroni<sup>2,3</sup>/<sup>1</sup> Department of Computer Science and Engineering, University of Bologna, Bologna, Italy <sup>2</sup> Research Centre for Open Scholarly Metadata, Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy <sup>3</sup> Digital Humanities Advanced Research Centre (/DH.arc), Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

- 一言でいうと  
oc\_ocdm: セマンティックWeb技術に精通していなくてもOCDM 準拠RDFデータを作成するためのPythonライブラリ
- 動機  
OpenCitationsにおけるデータモデルOCDMは、準拠データを扱うアプリケーションの作成を単純化するライブラリが必要
- リソースの設計方針
  - 要件1: 開発時にエラーを見つけやすくする開発方法論を採用
  - 要件2: 他の OpenCitations のアプリケーションで使用されているプログラミング言語に準拠したプログラミング言語を使用
  - 要件3: 定義されたエンティティについて、どのような変更が適用されるかを理解するためのメタニズムの設計(運用上のもの)
- 新規性  
既存のpublishingドメインにおける オントロジーに従った体系的なデータ作成ソリューションは不足。RDA語彙などに基づくLinked Dataへの変換ツールはあるが、学術会議などで作成されたデータは対象外。
- リソースの再現性／活用実績  
WikidataのWikipedia Citationsなど、4つのアプリケーションと共同プロジェクトで活用。  
SPARオントロジーアダプタのコミュニティは広く、再利用可能。
- 品質  
本ライブラリはOCDM準拠データをサポートしているが、従来の形式に関係なく書誌データや引用データを簡単に作成可能。
- リソースの可用性  
[GitHub](#)
- 次にすべきこと  
並列実行への対応。PyShExの開発が原則しているため、pySHACL対応を検討中。OCDM準拠でなく、任意のRDFデータに適用できる部分は、rdflibプラグイン化。

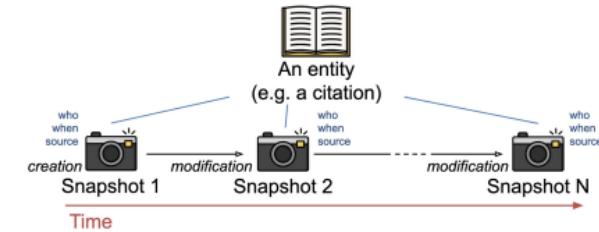


Fig. 2. The high-level description of the provenance layer of the OCDM to keep track of an entity's changes.

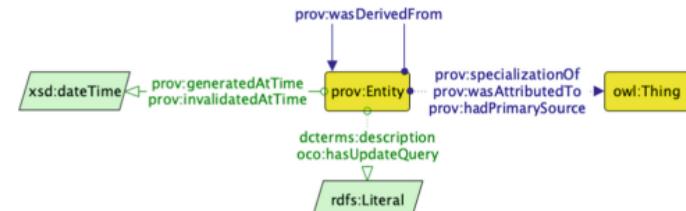


Fig. 3. The Graffoo diagram describing snapshots (prov:Entity) of an entity (linked via prov:specializationOf) and the related provenance information.

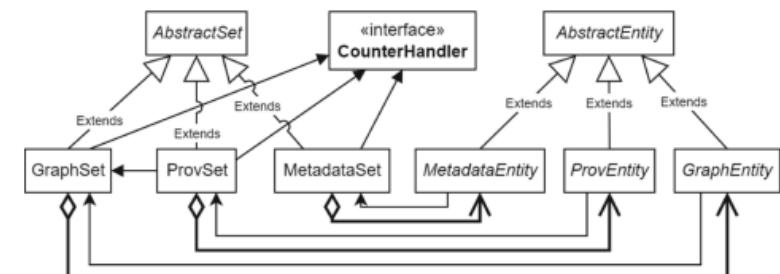


Fig. 6. A UML diagram showing the main relationships between classes in *oc\_ocdm*.

# Chowlk: from UML-based ontologyconceptualizations to OWL

[video](#), [GitHub](#)

Serge Chávez-Feria, Raúl García-Castro, and María Poveda-Villalón/Ontology Engineering Group, Universidad Politécnica de Madrid, Spain

- 一言でいうと  
Chowlk: 機械処理可能なデジタル オントロジー概念図を OWL に変換するためのフレームワーク
- 動機  
Ontology conceptualization (オントロジーの概念化) で具体化された概念は、オントロジーの実装において、手動で行われ、エラーが生じやすい
- リソースの設計方針  
Chowlk フレームワーク  
1) UMLベースの視覚表記 2) diagrams.netテンプレート 3) diagrams.net XML ダイアグラムからOWLへの変換  
グラフィカルなオントロジー エディターを構築せず、概念化の出力をオントロジー開発の一次成果物として取得し、概念化の作業から実装の最初のバージョンへのスムーズな移行
- 新規性  
1) 投資できる OWL ファイルの最初のバージョンを生成する時間を短縮  
2) 両方に使用できる概念図に焦点を当てる  
3) オントロジー ユーザーと開発者の間のコミュニケーションを改善  
4) オントロジーのドキュメンテーション段階で再利用されます。
- リソースの再現性／活用実績  
研究チーム内外でも、複数ドメインでも、W3Cなどのコミュニティでも、実績あり
- 品質  
実行サンプルが示されている
- リソースの可用性  
[GitHub](#)  
オントロジーエンジニアリングツールスイートのロードマップの一部として計画。
- 次にすべきこと  
他のビジュアル表記のサポートを検討中。SVGなどの他の標準形式のサポートを次バージョンで検討予定。

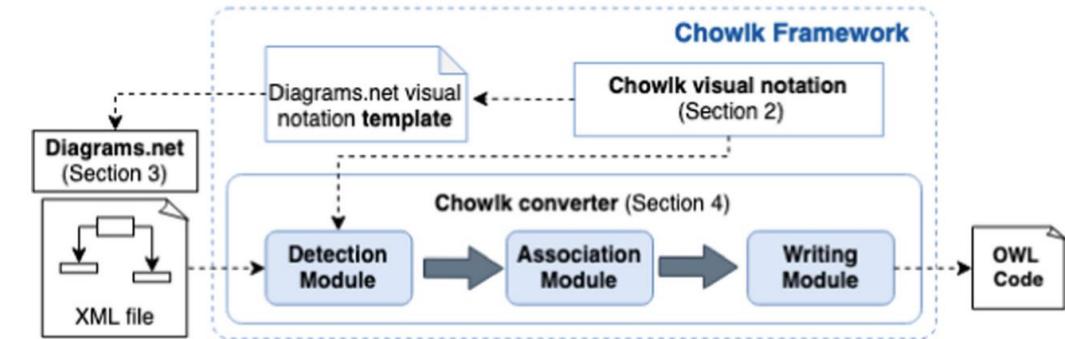


Fig. 1. Chowlk framework.

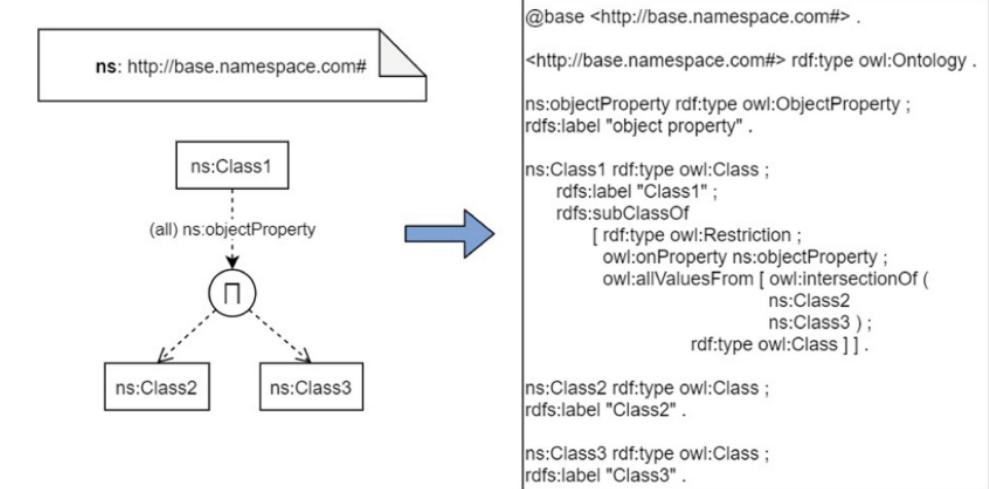


Fig. 6. Test conversion example.

# Dealing with Multiple Sources

# Impact of the Characteristics of Multi-Source Entity Matching Tasks on the Performance of Active Learning Methods

Anna Primpeli and Christian Bizer / Data and Web Science Group, University of Mannheim, Mannheim, Germany

- 一言でいうと

マルチソースでのエンティティマッチングにおけるタスクを生成する  
ALMSERgenを提案

- 動機

マルチソースで有効なデータプロファイリングを確立したい

- 手法

データソースを3つの側面（エンティティの重複度、値の不均一性、  
値のパターンの重複度）で評価し、学習に与える影響を評価

- 背景

2つのデータソースのマッチングは研究されているが、  
マルチになった場合、2者間で有効な特徴をそのまま展開できない

- 結果

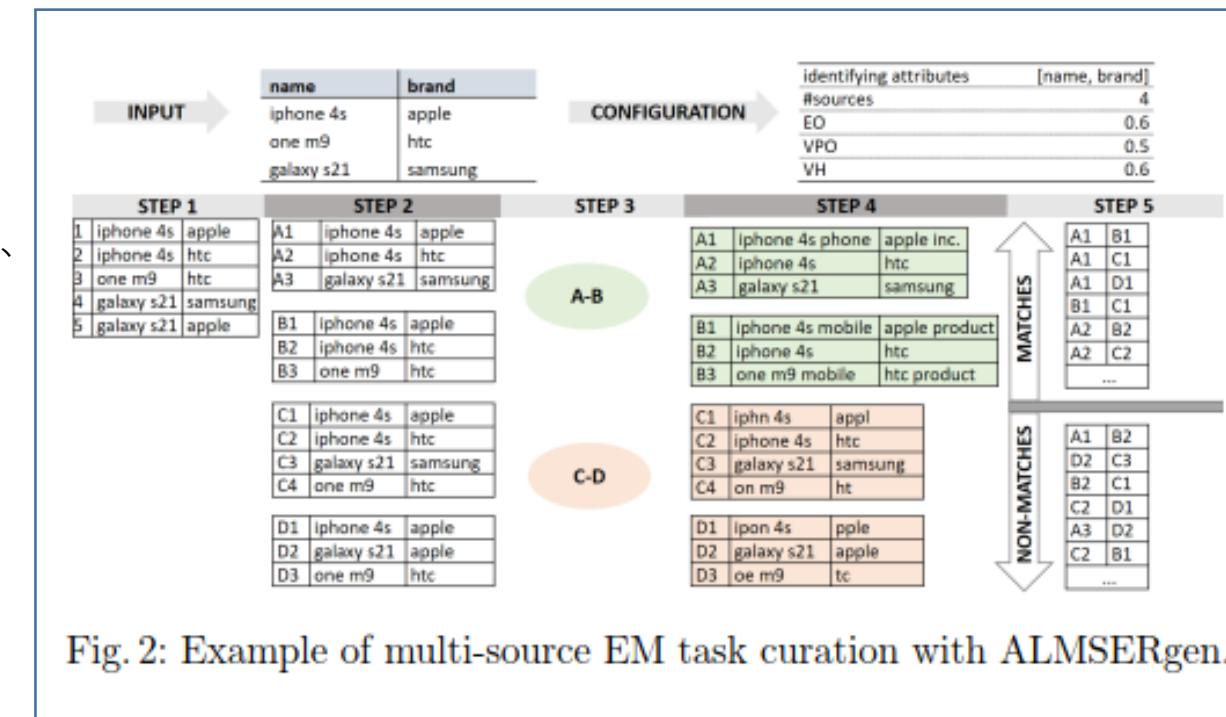
データプロファイルと、マッチングタスク生成の既存手法3つ  
(HeALER、ALMSER、ALMSERgroup)の関係を実験で比較。

- 考察

同じエンティティで値が不均一なマッチングを大量に含む場合は  
グラフ構造を用いる手法が有効、少ない場合はグルーピングを用いる  
手法が有効、といった知見が得られた。

- 課題

?



# Matching Multiple Ontologies to Build a Knowledge Graph for Personalized Medicine

Marta Contreiras Silva, Daniel Faria, and Catia Pesquita / LASIGE, Faculdade de Ciências da Universidade de Lisboa, Portugal

- 一言でいうと  
複数のオントロジーをクラスタリングしてからアライメント
- 動機  
個別医療向けKG構築のため、関連する28のオントロジー  
アライメントを行いたい
- 手法  
意味の重複をもとにオントロジーをクラスタリングし、  
クラスタ内でのマッチングをもとにオントロジーアライメント
- 背景  
従来手法は1ペアでの照合がベース。スケーラビリティや  
マルチドメイン特有の問題がある。
- 結果  
ペアワイズ分類とインクリメンタルな分類を実装。後者は前者の  
2割の時間でアライメントを行えた。
- 利用実績  
EUプロジェクト「KATY」が関連（実績は記述なし？）
- 学んだ教訓やベストプラクティス  
アライメントの評価、オントロジーの品質に対するアドホックな対策、  
ロード時間の圧縮が重要

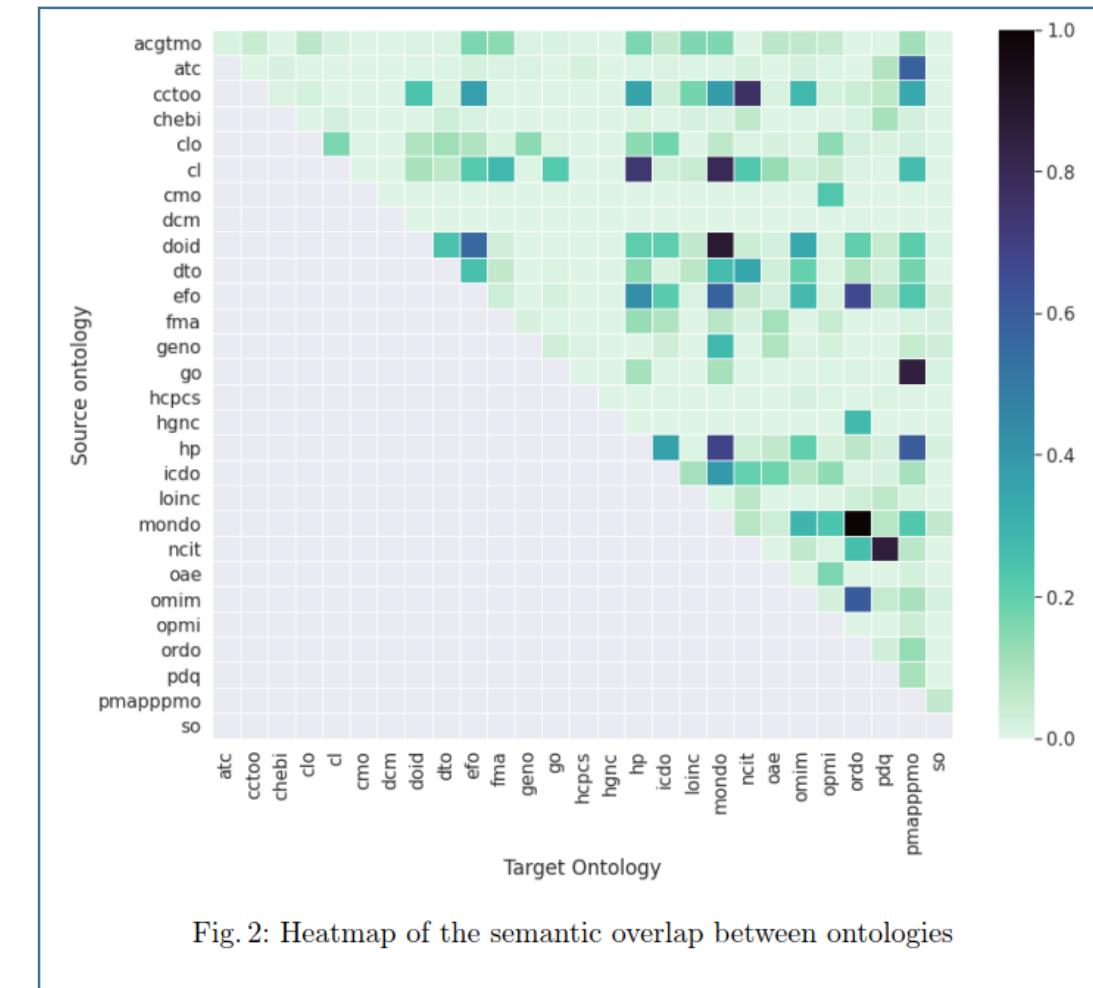
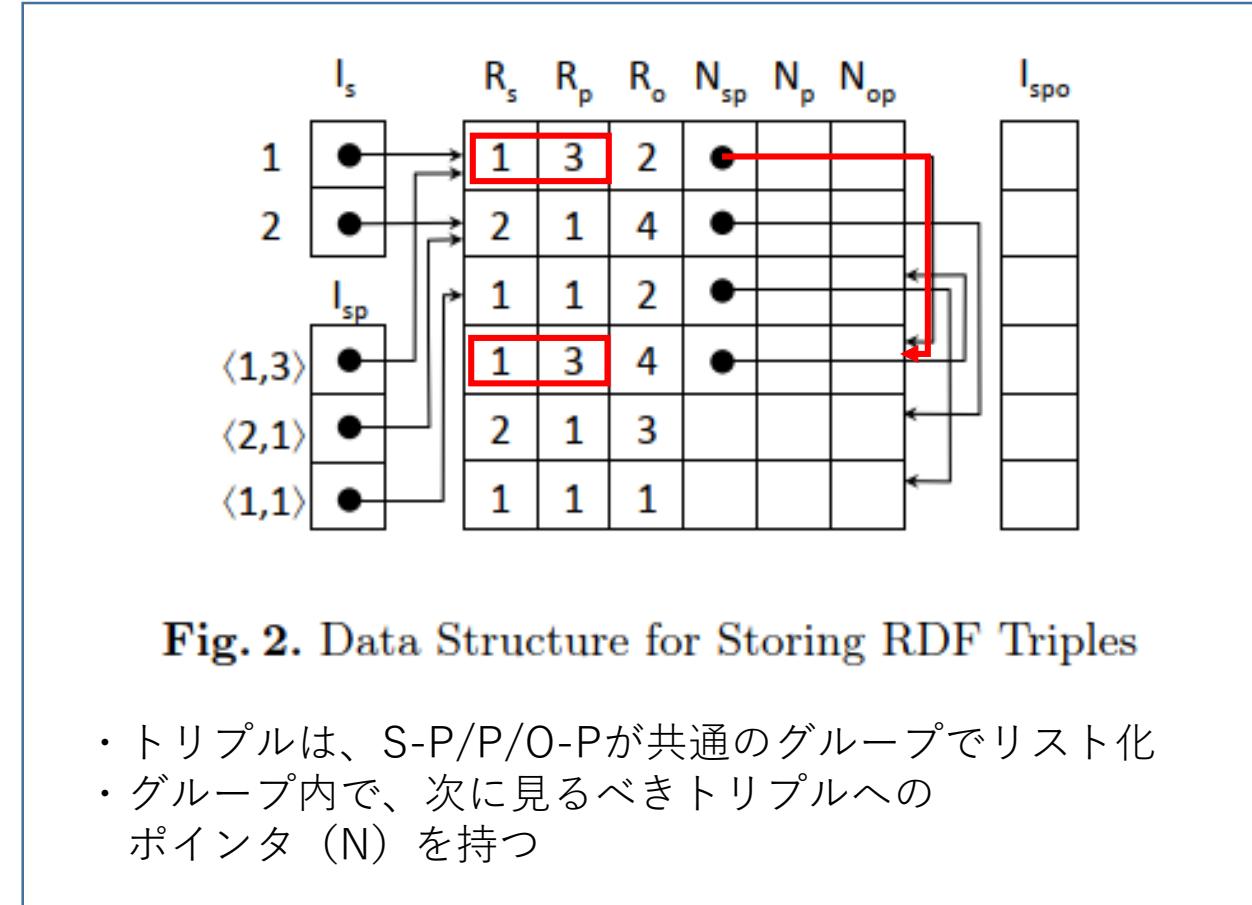


Fig. 2: Heatmap of the semantic overlap between ontologies

# The Dow Jones Knowledge Graph

Ian Horrocks<sup>1</sup>, Jordi Olivares<sup>2</sup>, Valerio Cocchi<sup>3</sup>, Boris Motik<sup>1</sup>, and Dylan Roy<sup>2</sup> / <sup>1</sup>University of Oxford, <sup>2</sup>Dow Jones, <sup>3</sup>Oxford Semantic Technologies

- 一言でいうと  
ダウ・ジョーンズ社向けに、RDFoxを使い、異なるソースからKGを作りクエリ処理する
- 動機  
企業DB、ニュースフィード、雑誌記事など様々なソースから意思決定に役立つ有益な情報を提供する
- 手法  
大規模KGの構築を、ロックフリー、処理結果の重複予防などにより並列化。
- 背景  
大規模KGの並列高速処理（構築、探索）
- 結果  
26億トリプルのKGを26分で構築
- 利用実績  
ニュース記事からのシグナル抽出や競合関係データの整備など一部業務プロセスで利用
- 学んだ教訓やベストプラクティス  
月に一度KGをリロードしているが、データソース連動でダイレクトに変更が反映される仕組みがさらに必要



# Domain Specific Knowledge Graph

# FindSampo: A Linked Data Based Portal and Data Service for Analyzing and Disseminating Archaeological Object Finds

Heikki Rantala<sup>1</sup>, Esko Ikkala<sup>1</sup>, Ville Rohiola<sup>3</sup>, Mikko Koho<sup>1,2</sup>, Jouni Tuominen<sup>1,2</sup>, Eljas Oksanen<sup>2</sup>, Anna Wessman<sup>4</sup>, and Eero Hyvönen<sup>1,2</sup>

<sup>1</sup>Semantic Computing Research Group (SeCo), Aalto University, Finland, <sup>2</sup>HELDIG - Helsinki Centre for Digital Humanities, University of Helsinki, Finland,

<sup>3</sup>Finnish Heritage Agency, Finland, <sup>4</sup>University Museum of Bergen, Norway

## ・ 一言でいうと

- Linked Open Data (LOD) に基づいたpublic考古学オブジェクトの発見を分析し広めるためのシステム「FindSampo」

## ・ 動機

- 遺産管理の課題を解決し、考古学データの国を超えた研究やデジタル人文科学への活用を目指したい

## ・ 手法

- フィンランドの Ontology for museum domain (MAO) と国際的に使われる同種のオントロジーとのマッピングを準備し、フィンランドの検索データを同様の国際的なデータと簡単に比較できるように

## ・ 背景

- 過去10年間の娯楽としての金属探知の人気の高まりによって多数の新しい考古学的データが生み出されている
- public考古学（一般の人々を関与させる考古学？\*）を国を超えて行いたい → LODを活用した相互運用性の実現

\* <https://www.saa.org/education-outreach/public-outreach/what-is-public-archaeology>

## ・ 結果

- データサービス\*\*と、検索機能・分析ツールを備えたセマンティックポータルで構成されるFindSampoデモンストレータ\*\*\*を実現
- フィンランド遺産庁によって2015-2020年の間にカタログ化された一般の人々による約3000の考古学的オブジェクト発見情報で構成

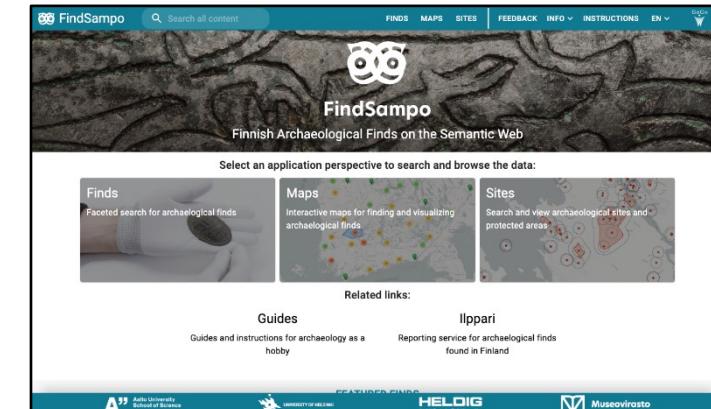
\*\* <https://www.ldf.fi/dataset/findsampo>

## ・ 利用実績

- デモンストレータ\*\*\*は2021年5月から一般公開されており、最初の6か月間で3000を超えるユーザが利用した

## ・ 学んだ教訓やベストプラクティス

- データの表現に用いられたファセットオントロジーの作成ではスプレッドシートで分類階層を作成後PythonでRDFに変換  
→ ファセットオントロジーを作成する考古学の専門家は非常に複雑なオントロジーエディターの使用法を学ぶ必要がなくなった
- しかし、階層が実際にどのように動作するかをより適切に示すことや、スペルミスを回避できることなどのオントロジーエディターによる利点が享受できないなどのオントロジーが使われて初めて明らかになる問題もあった



# LD Connect: A Linked Data Portal for IOS Press Scientometrics

Zilong Liu<sup>1</sup>, Meilin Shi<sup>1</sup>, Krzysztof Janowicz<sup>1</sup>, Blake Regalia<sup>1</sup>, Stephanie Delbecque<sup>2</sup>, Gengchen Mai<sup>1,3</sup>, Rui Zhu<sup>1</sup>, and Pascal Hitzler<sup>4</sup> / <sup>1</sup>STKO Lab, UC Santa Barbara, Santa Barbara, California, USA, <sup>2</sup>IOS Press, Amsterdam, NL, <sup>3</sup>Department of Computer Science, Stanford University, Stanford, California, USA, <sup>4</sup>Data Semantics Lab, Kansas State University, Manhattan, Kansas, USA

- 一言でいうと  
IOS Pressの書誌データを提供, 取得, 可視化, 分析する Linked Data ポータル 「LD Connect」
- Resource Type  
Knowledge graph
- 動機  
既存の学術的ナレッジグラフの問題点を克服し, 一般の人が学術的ナレッジグラフに基づいた研究やアプリケーション開発が容易になるようなリソースの開発
- リソースの設計方針  
LD Connect  
LD Connectの設計方法の説明→ IOS Press社の持つ書誌データから生成
- 新規性  
LD Connectは, 事前トレーニング済みのドキュメントとナレッジグラフのembeddingの両方を有する最初のもの
- リソースの再現性／活用実績  
LD ConnectサービスをセマンティックWebコミュニティだけでなく, 大規模な学術ナレッジグラフに基づく研究やアプリケーション開発を行う一般の人々に使用される可能性がある.  
また, IOS Pressの書誌データに限らず, 他の書誌データへの応用の可能性があり, より広範な人々に使用される可能性がある.
- 品質  
[http://stko-roy.geog.ucsb.edu:7200/iospress\\_scientometrics](http://stko-roy.geog.ucsb.edu:7200/iospress_scientometrics) にVisualizerあり (Webからブラウズ可能)
- リソースの可用性
  - Is the resource (and related results) published at a persistent URI (PURL, DOI, w3id)?  
<https://github.com/stko-lab/LD-Connect>
  - Does the resource provide a license specification? (See creativecommons.org, opensource.org for more information)  
CC BY-NC 4.0 (<http://ld.iospress.nl/about/licence/>)
- 次にすべきこと  
Space2Vecなどの手法を使用し, embeddingを使った類似性検索機能の開発

# QuoteKG:A Multilingual Knowledge Graph of Quotes

Tin Kuculo/L3S Research Center, Leibniz Universituät Hannover, Hannover, Germany ,  
 Simon Gottschalk /L3S Research Center, Leibniz Universituat Hannover, Hannover, Germany ,  
 Elena DemidovaData Science & Intelligent Systems (DSIS), Universituat Bonn, Bonn, Germany

- 一言でいうと  
Quoteを用いた多言語知識グラフの提案
- Resource Type  
Knowledge Graph (Datasets produced by novel algorithms or Community-shared software frameworks that can be extended or adapted to support scientific study and experimentation or Ontologies, vocabularies and ontology design patterns, with a focus on describing the modelling process underlying their creation)
- 動機  
有名人の「名言」をKGにしよう！(その文脈も考慮して多言語で)
- リソースの設計方針  
Wikiquoteから生成してEmbedding類似度から多言語に対応付け
- 新規性  
「名言」とそれが発現された文脈を結びつけた多言語KGは初めて？
- リソースの再現性／活用実績  
Githubに再現コードあり(dumperと多言語対応付けコード)
- 品質  
55言語で 69,000人以上の公人(の一部)分をデータ化
- リソースの可用性  
Publicly available (SPARQL endpoint)
- 次にすべきこと  
リソースの応用法についての議論

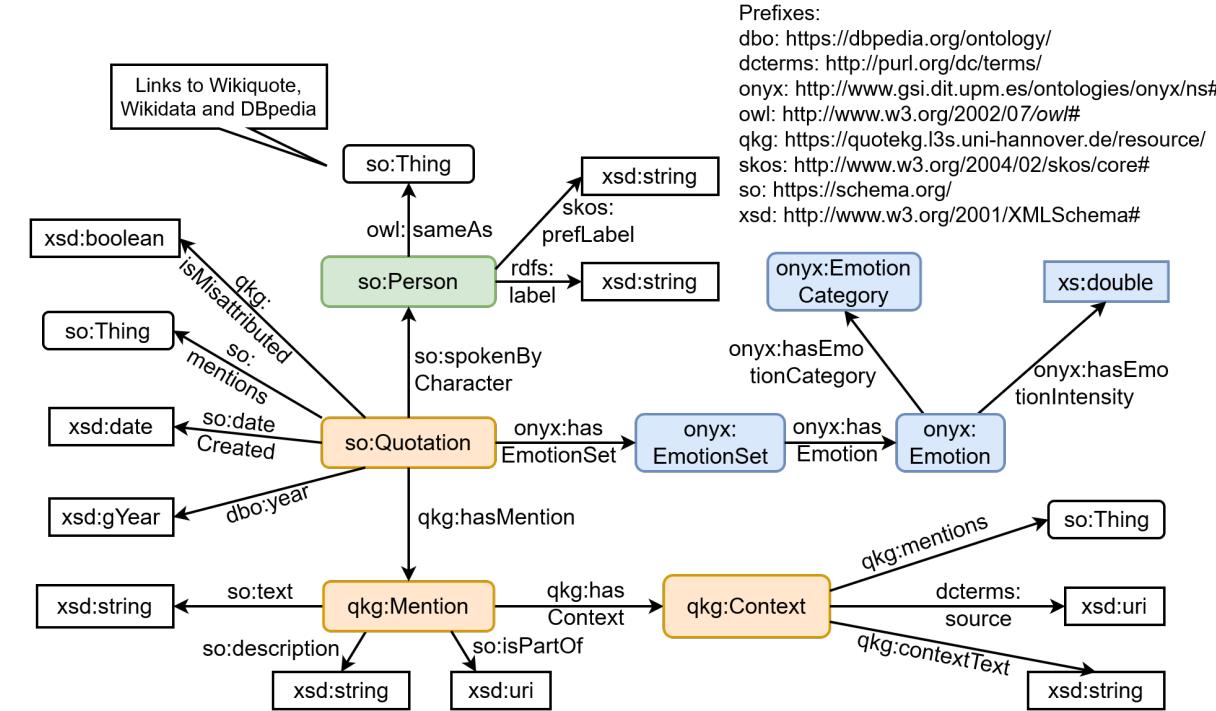


Fig. 1 (from元論文)

# Knowledge Graph for Tasks

# Improving Question Answering Quality Through Language Feature-based SPARQL Query Candidate Validation

Aleksandr Gashkov<sup>1</sup>, Aleksandr Perevalov<sup>2,3</sup>, Maria Eltsova<sup>1</sup>, and Andreas Both<sup>3,4</sup>

• <sup>1</sup> Perm National Research Polytechnic University, Perm, Russia

• <sup>3</sup> Leipzig University of Applied Sciences, Leipzig, Germany

• <sup>2</sup> Anhalt University of Applied Sciences, Kothen, Germany

• <sup>4</sup> DATEV eG, Nuremberg, Germany

- 一言でいうと

知識グラフを対象とした質問応答システムにおいて、SPARQL候補クエリにバリデーションを行うことで不要なSPARQL候補クエリを削除する手法の提案

- 動機

知識グラフを対象とした質問応答システムの精度向上

- 手法

ユーザーが入力した自然言語の質問文と、SPARQL候補クエリを自然言語に変換したものとの比較によってバリデーションを行う

- 結果

既存の質問応答システムQAnswerに対し本手法を用いてバリエーションを行い、大幅な精度向上を達成

- 考察

正解がないような質問（ナレッジグラフ上に存在しないエンティティが対象となる質問）を考慮した手法に検討の余地あり

- 課題

英語以外の言語に対するアプローチ  
異なるシステムやモデルを用いた本手法の検証

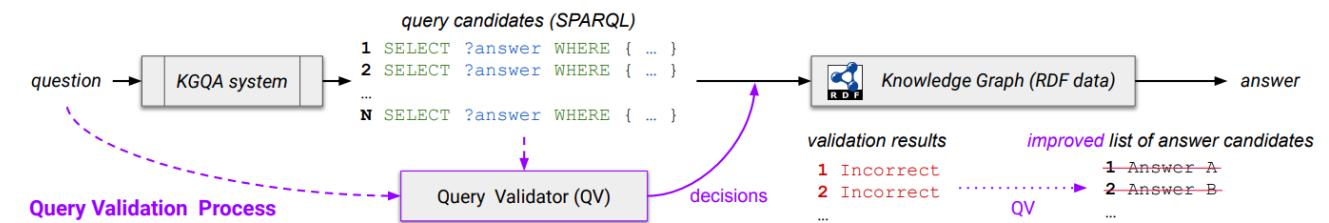
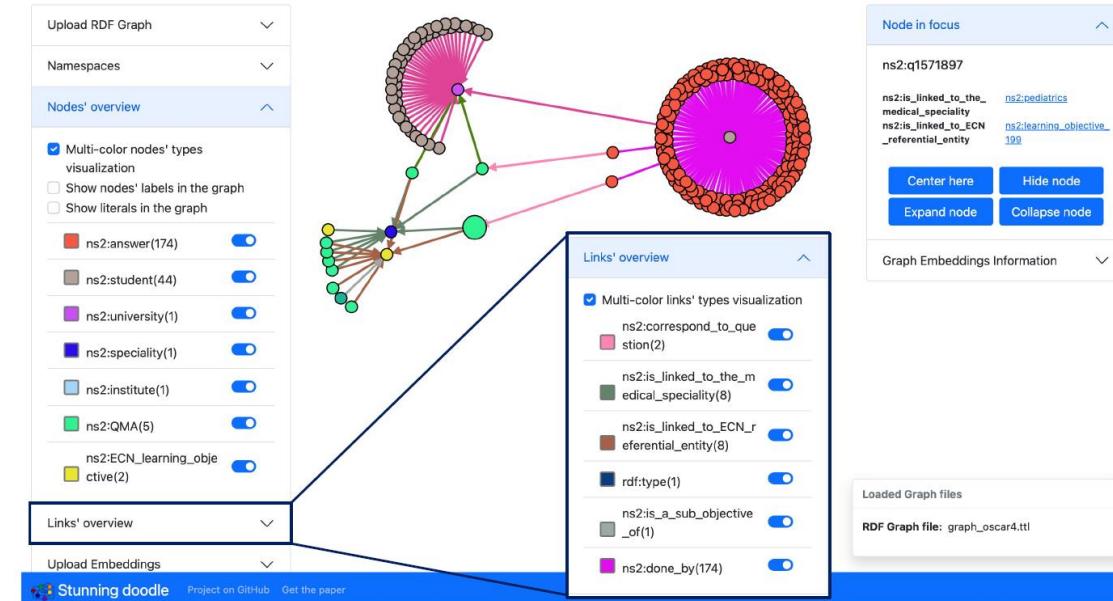


Fig. 1: General overview of the Query Validation process. The core component is the *Query Validator* intended to filter incorrect query candidates.

# Stunning Doodle: a Tool for Joint Visualization and Analysis of Knowledge Graphs and Graph Embeddings

Antonia Ettorre B, Anna Bobasheva, Franck Michel, and Catherine Faron / Université Côte d'Azur, CNRS, Inria, I3S, Sophia Antipolis, France

- 一言でいうと  
KGEが持つ情報の解析と解読を支援する可視化ツールであるStunning Doodleの利用方法の紹介と説明
- Resource Type  
Software
- 動機  
KGの内容や構造の理解、GEで得られた情報の分析・解釈のため
- リソースの設計方針  
専門家ではないユーザによるKGの探索と理解の支援と簡素化
- 新規性  
シンプルで分かりやすいKGEの視覚的分析
- リソースの再現性／活用実績  
大規模KG OntoSIDES の内容と構造の理解
- 品質  
PythonとJavascriptを使用した軽量なウェブアプリケーション
- リソースの可用性  
オープンソースソフトウェアとして提供、DOIにより識別
- 次にすべきこと  
SPARQLクエリの結果の可視化、有名なSPARQLエンドポイントへの直接アクセスなど



**Fig. 5.** Screenshot of *Stunning Doodle* showing the basic entities and relations in the OntoSIDES graph.

# Enhancing Sequential Recommendation via Decoupled Knowledge Graphs

Bingchao Wu<sup>1,3</sup>, Chenglong Deng<sup>1,3</sup>, Bei Guan<sup>1,3</sup>, Yongji Wang<sup>1,2,3</sup>, and Yuxuan Kangyang<sup>2,3</sup>/

<sup>1</sup> Collaborative Innovation Center, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup> State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China

- 一言でいうと

エンティティ間の関係のタイプに応じて、知識グラフを2つのサブグラフに分離します。さらに、2つのサブグラフに基づいてCrbiaNetを提案

- 動機

シーケンシャルレコメンドは、ユーザーインタラクションシーケンスに基づいてユーザーの動的な関心パターンをキャプチャできます。最近、知識グラフ (KG) をシーケンシャルレコメンドに統合することに关心が集まっている

- 手法

推薦ドメイン内の元のKGを、クロスユーザー動作ベースと固有の属性ベースのグラフの2つのサブグラフに分離します。次に、GNNによって独立して異なるセマンティクスを持つエンティティの高次の特徴を取得するために、2つの知識サブエクストラクタを提案。

最後に、高次のセマンティック機能が組み合わされ、順次推薦モデルに入力されて、ユーザー設定の表現が強化されます。

- 結果

提案されたCrbiaNetが、最先端の知識強化シーケンシャル推薦モデルを一貫して大幅に上回っていることを示した

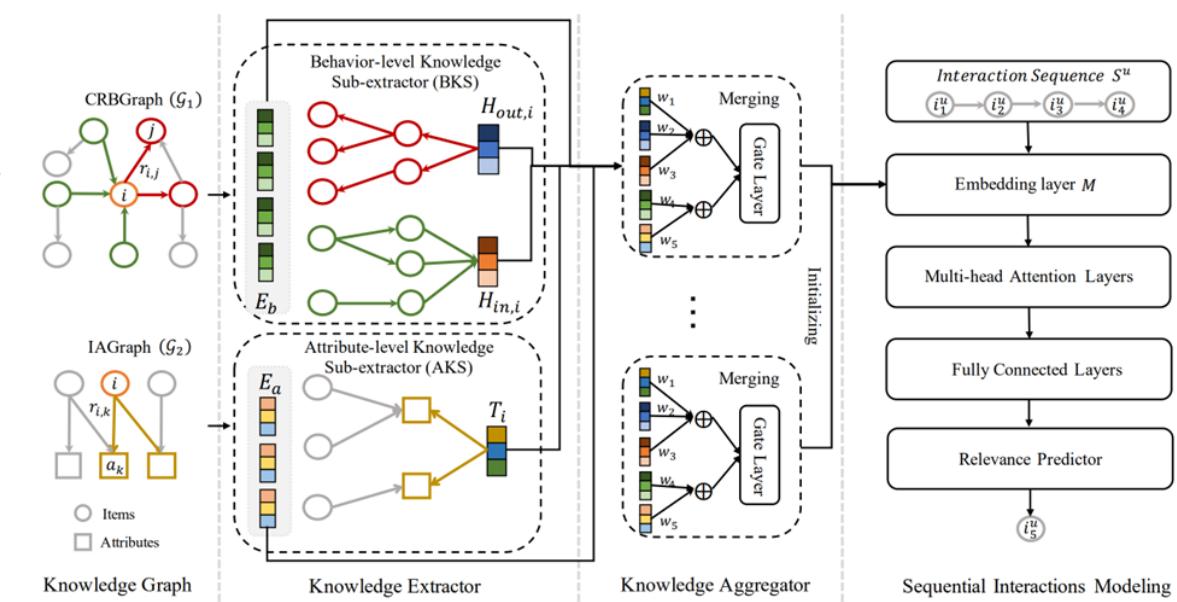


Fig. 2: The overall framework of our proposed model.

# Multiple Modalities

2022/08/24

ナレッジグラフ若手の会

ESWC2022サーベイ会

# Multiple Modalities

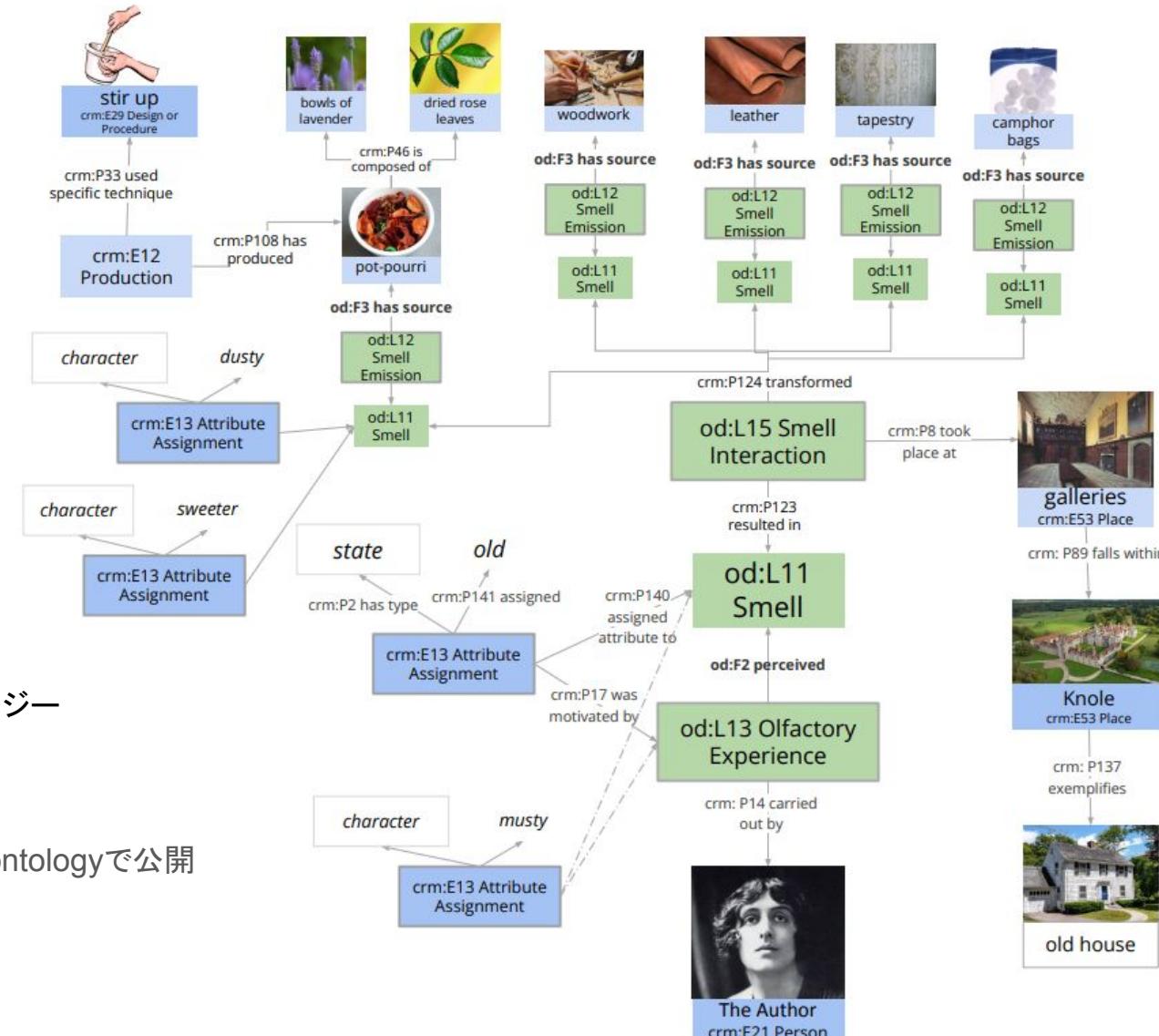
発表者：大阪電気通信大学 古崎研究室

志賀拓馬・吉岡舞香

# Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information

1 Sophia Antipolis/ EURECOM / 2 Ljubljana/Jozef Stefan Institute 3 Amsterdam/ KNAW Humanities Cluster/ 4 Cambridge/ Anglia Ruskin University  
 5 Cambridge/ University of Cambridge

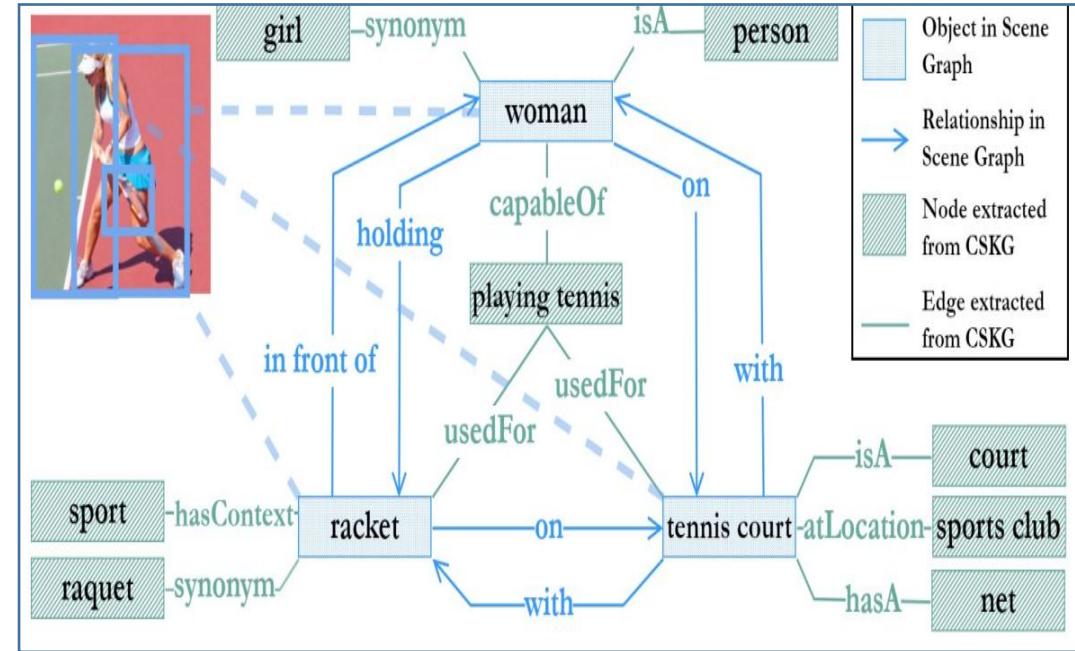
- 一言でいうと  
 嗅覚情報のデータモデル「Odeuropa」の紹介
- Resource Type  
 Ontologies and vocabularies
- 動機  
 嗅覚情報学は分子の化学構造をもとに分子がどのような匂いを発するかを計算することに焦点を当てられていたが、コンピュータサイエンス技術を用いて嗅覚体験を歴史的・文化的な文脈で捉える
- リソースの設計方針
  - ・CIDOCRMのような確立されたオントロジーを再利用・拡張し、関連する情報を相互作用の集合として表現する
  - ・CRMSciのクラスとプロパティを拡張する
- 新規性  
 歴史、文学、芸術、文化遺産などの分野で感覚体験の記述を可能にするオントロジー
- リソースの可用性
  - ・RDFを用いてHTTP APIから利用することが出来る
    - ・Oderopaデータモデルは、CC-BYライセンスの下、<http://data.odeuropa.eu/ontology>で公開
- 次にすべきこと  
 香りの遺産と香水産業との間のギャップを埋める



# Expressive Scene Graph Generation using Commonsense Knowledge Infusion for Visual Understanding and Reasoning

Muhammad Jaleed Khan, John G. Breslin, Edward Curry

- 一言でいうと  
常識を注入し表現豊かなシーングラフの生成を提案
- 動機  
オブジェクト検出や視覚的関係検出などの視覚的理解タスクにおけるいくつかの困難な問題の解決
- 手法  
異なる7つの知識ベースの統合したものを用いてシーングラフ生成。  
既存の最先端技術と比較して評価する。
- 背景  
シーングラフ生成分野の研究者は最先端技術を用いてアプローチしてきたが 表現力、精度、頑健性にはまだ大きな改善の余地がある。そのため、より高いレベルにするにはシーンとその構成要素に関する背景情報と人間の認知能力であるコモンセンス推論が必要である。
- 結果  
既存の状態よりも高い再現率を導き出せた
- 考察  
コモンセンス知識は、シーングラフ生成の性能とシーングラフ表現の表現力を向上させた
- 課題  
既存のシーングラフ生成アプローチでは、高いレベルの視覚的理解と推論に対するパフォーマンスと表現力が制限されていること



# Industry Track 1

# Enhancing Knowledge Graph Generation with Ontology Reshaping – Bosch Case

Bosch Center for AI, University of Oslo, Oslo Metropolitan University, Nanjing University, University of London

- 一言でいうと

専門家が構築した領域オントロジを、設備データの分析に適したコンパクトな知識データ（KGスキーマ）に変換

- 動機

多様な産業データ（rawデータ）を知識データ（領域オントロジや知識スキーマ）に対応付ける処理は以前として複雑でコストがかかる

- 課題

トレードオフ問題：knowledge-oriented vs data-oriented

- 手法

Ontology Reshaping[27]：領域オントロジをコンパクトなKGスキーマへ変換。  
領域オントロジから選択したノードとエッジをつないでグラフを生成。  
直感的には、rawデータに登場しないクラスとプロパティを除外して再構成。

- 評価

自動車製造ラインの溶接品質モニタリングデータで評価。  
データレコード件数400万、溶接回数1,000、自動車数100。  
KG生成処理時間7~8倍、エンティティ数1/2~1/6、ストレージ領域2/3、  
ブランクノード数0、クエリ実行時間の高速化、などを確認

- 考察

特になし

- 知見

Ontology Reshapingの理論確立、リアルタイム環境での評価、  
他のユースケースによる評価、など

Fig. 1. An overview of our approach of ontology reshaping enhanced KG generation.

Figure 1: Why do we need ontology reshaping? Domain ontology (partially shown in a) reflects the knowledge; the KG schema (partially shown in b) needs to reflect raw data specificities and usability. Blue boxes: classes that can be mapped to attributes in the raw data; black boxes: classes that cannot be found in the raw data.

[27] Zhou, et. al, Towards Ontology Reshaping for KG Generation with User-in-the-Loop: Applied to Bosch Welding, IJCKG2021

# Semantic Description of Equipment and its Controls in Building Automation Systems

Siemens

- 一言でいうと  
ビル設備の監視制御に関するデータの相互運用性を可能にする3層構造のオントロジ・アーキテクチャを提案
- 動機  
ベンダーが異なる多様なビル設備の監視制御におけるデータの相互運用性
- 課題  
異なるベンダー/ツール/情報モデルを横断した水平統合と垂直統合  
制御プログラムの役割が機械可読な形式で与えらず、相互接続性が無い  
複数の既存オントロジを手作業でつないで構築しており、保守性が低い  
管理・計画・制御で必要とする知識の粒度が異なる
- 手法  
3層構造のオントロジを構築（構造層、機能層、振舞層）。構造層と機能層は既存オントロジを取り込み共通化。設備と制御に関する専門知識は機能層と振舞層で扱い、ノウハウは振舞層に集約。設備情報をWoT Thing Descriptionで記述し、用語を各オントロジへリンク
- 評価  
実際のビル5棟に対し、3層のオントロジを利用してデータを統合した知識グラフを構築  
エンジニアリングや異常検知に関するユースケースで、外部ツールを利用することなくデータへアクセスできたことを確認
- 考察  
人間と機械の両方に対して可読性を有した、データ相互運用性の実現方式である
- 知見  
特になし



Fig. 1. Layered ontology for BA.

# Extracting Subontologies from SNOMED CT

University of Manchester, SNOMED International

- 一言でいうと  
SNOMED CTから関心事に関連する部分オントロジを抽出
- 動機  
医療オントロジの1つであるSNOMED CTは35万以上の概念から構成される  
検索や推論の処理に時間がかかるため、関心事に応じて抜粋して利用したい
- 課題  
専門家が抽出した概念リスト（Refset）はセマンティクスが元のオントロジに依存  
モジュール化は関心事と関連しない不要な概念が含まれる
- 手法  
与えられた概念に対してセマンティクスの完全性が保証された部分オントロジ<sup>\*1</sup> の  
抽出手法を開発。OWL APIとELK推論器を利用し実装  
<sup>\*1</sup>: a)抽出された概念が部分オントロジと元のオントロジとで同等に扱われること、かつ、  
b)部分オントロジで推移閉包が成立するならば元のオントロジでも推移閉包が成立すること
- 評価  
実際に運用されているRefsetを題材に、抽出された公理と概念の件数をSTARモジュール化と比較し有用性を確認
- 考察  
専門家からのフィードバックでRefsetの概念不足を指摘されるなど、モデリングやキュレーションの洗練化にも有用
- 知見  
特になし

**Table 1.** A comparison between the sizes of the extracted subontologies and locality-based (STAR) modules for a collection of refsets.

Refset Name	Refset Size (Concepts)	Axioms	Subontology Size Concepts	STAR Module Size Axioms	STAR Module Size Concepts
ERA-EDTA	184	485	475	3076	3086
Dentistry	226	455	642	1449	1478
Nursing	1337	2616	2616	5579	5708
Orphanet	5681	9209	9189	27595	27625
IPS	8182	12793	12745	53736	53708
GPS	26159	33970	33907	86167	86374

#### Refsets Key:

Dentistry	General Dentistry
ERA-EDTA	European Renal Association / Dialysis and Transplant Association
GPS	International Global Patient Set
IPS	International Patient Set
Nursing	Nursing Activities and Nursing Health Issues (combined)
Orphanet	Rare diseases, orphan drugs

# Industry Track 2

# "Semantify" business and content to meet demands for expert solutions in professional markets

Walters Kluwer

- 一言でいうと  
文書と業務を“semantify”し、顧客課題に対しタイムリーに解決策を提案
- 動機  
財務、法務等の専門分野向けソリューションにおいて、顧客に対し  
タイムリーかつ実行可能な知見の提供が求められる
- 課題  
文書等のコンテンツにメタデータを付与し検索可能とするだけでは解決  
できない
  - ・税務担当者は中小企業を経営する顧客向けに2022年の税率を知りたい
  - ・医療従事者は臨床研究におけるリサーチクエスチョンの先行研究を知りたい
  - ・弁護士は事件における主張と事実が一致した関連する判例を知りたい
- 手法  
[3]で開発の上位オントロジ（詳細不明）に対する領域拡張を作成・管理  
するプロセスを開発。ビジネスアナリストとナレッジエンジニアが領域  
専門家にヒアリングし、領域拡張や上位オントロジの洗練化を判断
- 評価  
特になし
- 考察  
特になし
- 知見  
領域やコンテンツの知識の共有を促進することで顧客業務の理解を深め、顧客課題に対してタイムリーな解決策の提案が可能に  
変革プロセスに求められる課題：スピード、コミュニケーション、オーナーシップ、コントロール、専門知識など（詳細不明）

## Semantifying content

Semantifying content means making **knowledge** in and about content **explicit, normalized and machine readable**. It also means that domain knowledge is realized as **knowledge objects** in their own right, finally completely independent of the content source they come from.

This semantification of content has an impact on all stakeholders, e.g.

- From Content Worker to Knowledge Worker
- From Software Developer to Expert Solution Developer
- From Product Manager to Knowledge Product Owner

- Textual content remains one important information channel from WK to customers
- 

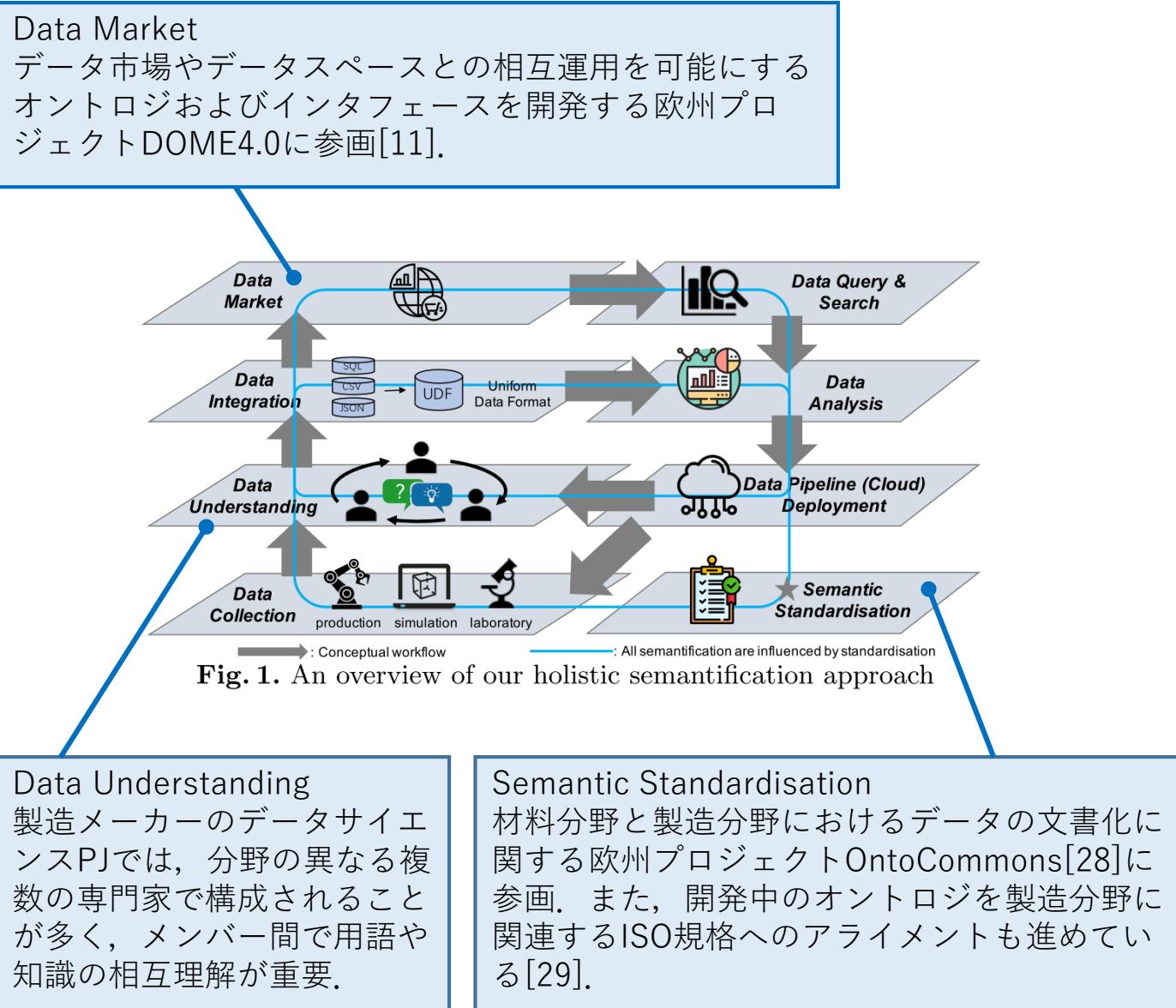
プレゼン資料より抜粋

[http://videolectures.net/eswc2022\\_dirschl\\_professional\\_markets/](http://videolectures.net/eswc2022_dirschl_professional_markets/)

# The Data Value Quest: A Holistic Semantic Approach at Bosch

University of Oslo, Bosch Center for AI, Oslo Metropolitan University, Nanjing University, Norwegian University of Science and Technology, University of London, SINTEF

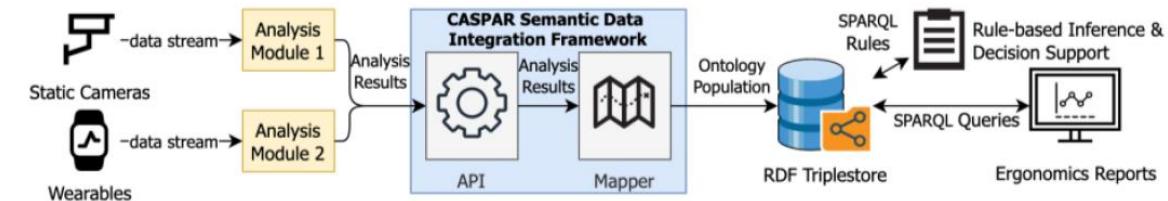
- 一言でいうと  
Boschが開発中のデータ活用プロセスを紹介
- 動機  
多種多量データからの価値の抽出・活用が産業分野で十分に進んでいない
- 課題  
以前として多くの技術課題があり、多くの解決手段が検討されている
- 手法  
データに意味を与える"semantification"に関して、Boschが開発を進めている包括的プロセスを紹介
- 評価  
プロセスの各段階で開発・採用した技術の適用実績や評価は文献を参照。本資料では、Understanding, Market, Standardisationについて紹介。
- 考察  
特になし
- 知見  
特になし



# Semantic Data Integration for Monitoring Operators' Ergonomics in an Automotive Manufacturing Setting

Catalink Limited, University of Patras, Centro Ricerche FIAT SCpA, Athena Research Center

- 一言でいうと  
カメラ画像とウェアラブルセンサから自動車組立作業者の姿勢を推定
- 動機  
自動車組立ラインにおける作業者とロボットの協調作業において、  
作業者を怪我等の身体的負担を軽減する
- 課題  
作業者の健康・安全上のリスクを回避するために、作業者の姿勢や  
体の使い方を高精度で認識する
- 手法  
組立作業を撮影したカメラ映像と作業者が身につけたウェアラブルセンサから収集した知識グラフを生成  
ルールに基づくクエリを利用し、知識グラフから作業者の姿勢や体の使い方に関する情報を抽出
- 評価  
自動車シャーシへのフロントガラス取り付けを題材に、Unityを利用して構築した仮想環境で評価実験を実施
- 考察  
a)組立知識の高い表現力、b)結果に対する説明性、c)データから直接観察されない隠れ状態の把握可能性、を有する
- 知見  
ロボットのセンサやアクチュエータの情報を加味した、作業者の姿勢推定



**Fig. 1.** Workflow overview.

ご協力いただき  
ありがとうございました！