

# 「Wine Ontology」 活用の探求(4)



日付： 2024年11月27日  
インフォラボ游悠 [研究所]  
中岡 実

備考： 本内容は、オントロジー・サンプルとして提供されている“Wine Ontology”の内容を元に、Class、プロパティ等を追加し、グラフDBでの利用を研究するために作成したものです。

- <参照資料>
- <http://www.w3.org/TR/owl-guide/wine.rdf> で公開されている“Wine Ontology”
  - 利用データ： UCI 赤・白ワイン評価データ <https://archive.ics.uci.edu/ml/datasets/wine+quality>

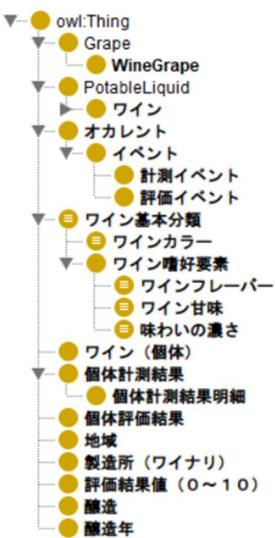


および名称は、作成者(中岡実)の登録商標です。

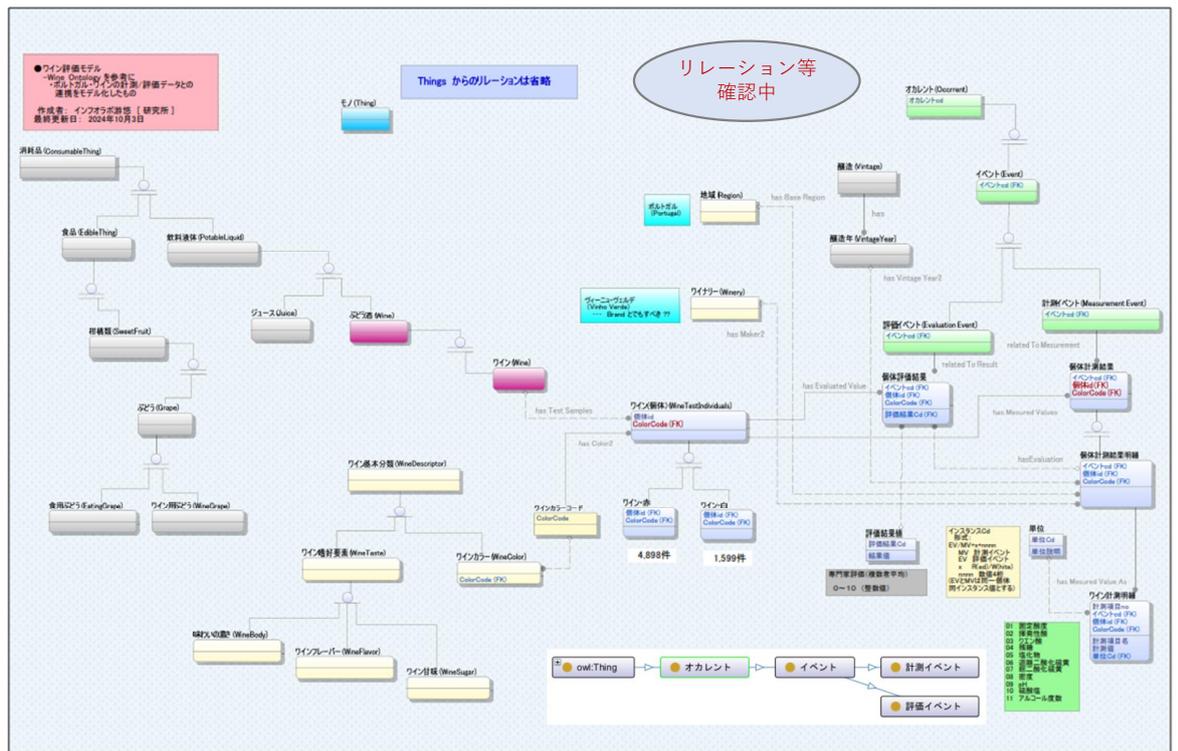
©2024 インフォラボ游悠 [研究所]

1

## ● ワイン・オントロジーを参考にした拡張モデル図 (検討中 10/5時点)



Wine-Ox13j-value7.rdf



©2024 インフォラボ游悠 [研究所]

2

● 利用データ(インスタンス)情報について (10/5時点) … 元のデータセットから一部を選択使用

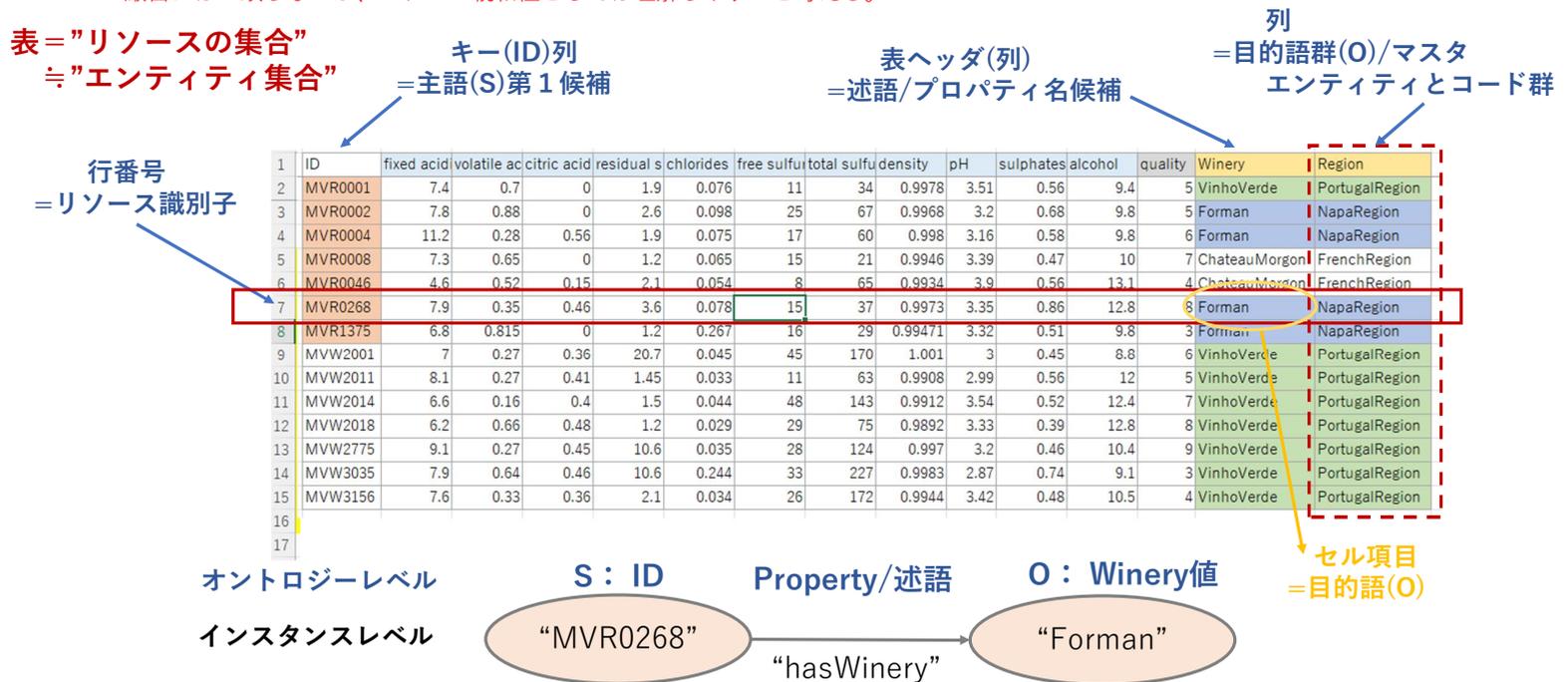
- ・ワイン赤 7 件 (ID : MVRnnnn)、ワイン白 7 件 (ID : MVWnnnn) を利用した。変数値に関する説明は第 3 回「参考 1 表」を参照。
- ・IDからqualityまでのデータは、元のテストデータ値組合せをそのまま適用表記した。
- ・Winery、Regionについては、VinhoVerde、PortugalRegion 以外は仮想的な値を当てている (Ontologyファイル内の登録値を利用)

ID	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality	Winery	Region
MVR0001	7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5	VinhoVerde	PortugalRegion
MVR0002	7.8	0.88	0	2.6	0.098	25	67	0.9968	3.2	0.68	9.8	5	Forman	NapaRegion
MVR0004	11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	6	Forman	NapaRegion
MVR0008	7.3	0.65	0	1.2	0.065	15	21	0.9946	3.39	0.47	10	7	ChateauMorgon	FrenchRegion
MVR0046	4.6	0.52	0.15	2.1	0.054	8	65	0.9934	3.9	0.56	13.1	4	ChateauMorgon	FrenchRegion
MVR0268	7.9	0.35	0.46	3.6	0.078	15	37	0.9973	3.35	0.86	12.8	8	Forman	NapaRegion
MVR1375	6.8	0.815	0	1.2	0.267	16	29	0.99471	3.32	0.51	9.8	3	Forman	NapaRegion
MVW2001	7	0.27	0.36	20.7	0.045	45	170	1.001	3	0.45	8.8	6	VinhoVerde	PortugalRegion
MVW2011	8.1	0.27	0.41	1.45	0.033	11	63	0.9908	2.99	0.56	12	5	VinhoVerde	PortugalRegion
MVW2014	6.6	0.16	0.4	1.5	0.044	48	143	0.9912	3.54	0.52	12.4	7	VinhoVerde	PortugalRegion
MVW2018	6.2	0.66	0.48	1.2	0.029	29	75	0.9892	3.33	0.39	12.8	8	VinhoVerde	PortugalRegion
MVW2775	9.1	0.27	0.45	10.6	0.035	28	124	0.997	3.2	0.46	10.4	9	VinhoVerde	PortugalRegion
MVW3035	7.9	0.64	0.46	10.6	0.244	33	227	0.9983	2.87	0.74	9.1	3	VinhoVerde	PortugalRegion
MVW3156	7.6	0.33	0.36	2.1	0.034	26	172	0.9944	3.42	0.48	10.5	4	VinhoVerde	PortugalRegion

©2024 インフオラボ游悠 [研究所]

● リンクデータ表現 (三つ組表現) とオントロジー/データモデル的考え方の親和性に関して

- ・ここでは、表形式データ表現と、リンクデータなどで用いられるトリプル(三つ組)表現、オントロジー/データモデル表現との類似性/親和性といった面の考察概要を提示する (ここでは、リレーショナルモデルで考察対象とする「正規化」等は一旦省く)。
- ・厳密には一致しないが、モデルの親和性としては理解しやすいと考える。

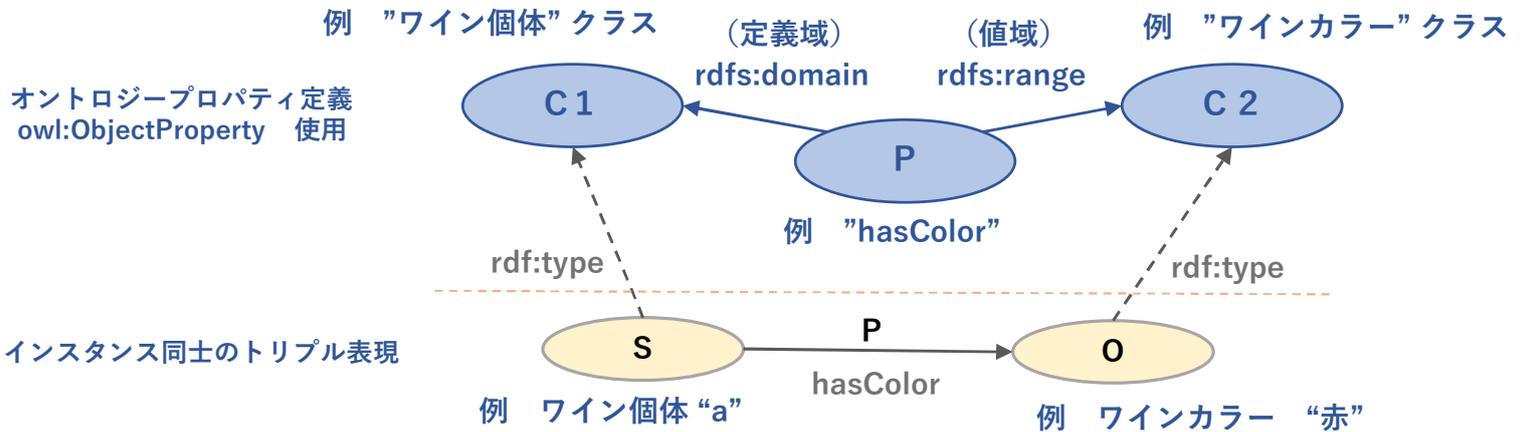


©2024 インフオラボ游悠 [研究所]

## 【補足】第三回分科会（10/23）質問への追加説明

Q. 1 Wineオントロジー利用でObject Property と Datatype Property 使用方法の違いは何か？

- A. 1 ObjectProperty は、インスタンスとインスタンスを結ぶ個体値型プロパティが属するクラス。この型のプロパティは、目的語がURI参照を持つノードもしくは空白ノードになる。一方、DatatypeProperty は、インスタンスをデータ値に結び付けるデータ値型プロパティが属するクラス。この型のプロパティは、目的語がリテラルになる。



©2024 インフオラボ游悠 [研究所]

5

## 【参考2】グラフDB基盤活用の主な種類に関する考察（By 中岡）

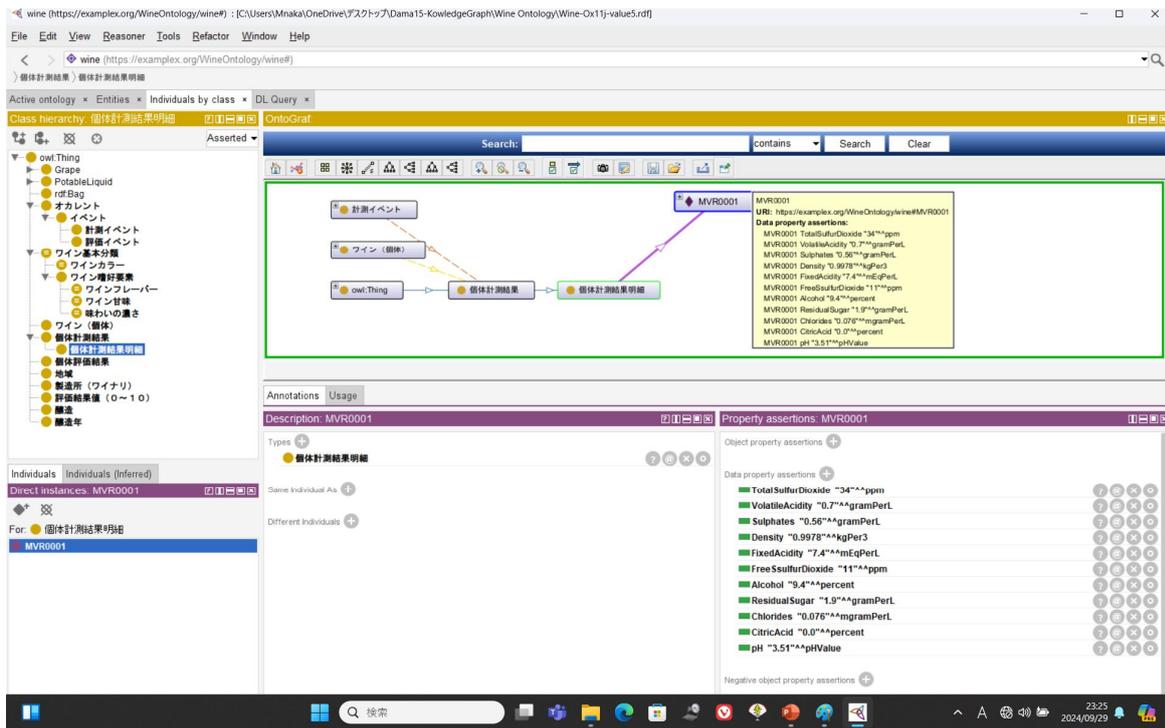
No.	利用場面による種類	概要 by 中岡
1	LOD利用	<ul style="list-style-type: none"> <li>・トリプル（三つ組み）を格納する入れ物としての利用。</li> <li>・トリプル格納バケツとしての代表格は、所謂リンクト・オープンデータ(LOD)の取込みを目的としたもの。</li> <li>・これは広範囲な全体構造を把握・管理するよりは、外部で公開されたデータとの情報の拡張・接続性を主な期待値として利用するものと考えられる。ここでは基本的にはインスタンスの集合という形態であり、データの拡張利用においては語彙の一致性が鍵になると云える。但し語彙というモノは、同じ名称を使っているにもかかわらず意味合いが異なるという「同名異義」、また異なる語彙であっても実は同じ意味を表しているという「異語同義性」といった現象を表すことが頻繁にあり、正確性の揺らぎを生じる傾向がある。その上で三つ組みレベルの集合という点が、対象データの急激な増加を生むという現象に関わる。異なった言語で表現されたトリプルを結合する場合、更にこの三つ組み集合の表す結果の正確性／品質が問題となる。この点を許容する範囲で納める結合の工夫が重要性を持つ。</li> </ul>
2	Ontology管理/利用	<ul style="list-style-type: none"> <li>・オントロジー構築中心の構造管理を主目的とした基盤としての使い方。</li> <li>・1つめの利用との関連としては、三つ組みの単なる大規模集合に意味の基礎を与えるという役割と云える（これはボトムアップ的アプローチということになり、手間も掛かるものになると考えられる）。</li> <li>・一方データで表したい意味構造を概念レベルから設計するという方向もあり、これはトップダウン的アプローチという正攻法と云える。</li> <li>・このオントロジー構築の過程で、所謂マスターデータの設計と、その構成コード体系具体化という意味でマスタに属するインスタンス群の生成という流れが生まれる。こうして構築されたオントロジー構築が意味メタデータとして位置づけられ、品質を伴ったグラフデータが生まれることになる。この手法でビッグデータの機械学習連携が精度を向上する効果を生むことが期待される。</li> </ul>
3	知識データベース活用(KDB)	<ul style="list-style-type: none"> <li>・知識データベースとしての利用を目的とする知識グラフ格納庫／構築基盤といった方向。</li> <li>・この有効利用においては、2つめのオントロジー構築に加えてグラフデータベースを元にした推論規則構築の仕方と、これを用いた探索言語の整備という点が重要性を持つ。</li> <li>・これに関しては既に様々な試みが行われている状況で、基本的には対象領域（例えば、製造、金融、物流管理など）に特化した構築方法が主体になると想定される。大域的な基本構造と、それを利用する個別プロセスの場面依存という課題があるためと考えられる。こうなると、探索された知識を反映して知識グラフにフィードバックするという方向の技術が、更に意味をもつことになると想定される。この分野においては、機械学習／AI技術の連携として、現在LLM (Large Language Model; 大規模言語モデル) や RAG (Retrieval Augmented Generation) 技術の連携利用が試みられているが、筆者の理解では、まだ実証実験段階であるように捉えられる。</li> </ul>
4	リポジトリの利用	<ul style="list-style-type: none"> <li>・カタログ／語彙管理の基盤としての利用。メタデータ側の管理が重要。具体的な方法は今後の検討課題。</li> </ul>

©2024 インフオラボ游悠 [研究所]

6

## 【補足】第三回（10/23）質問への追加説明（続き）

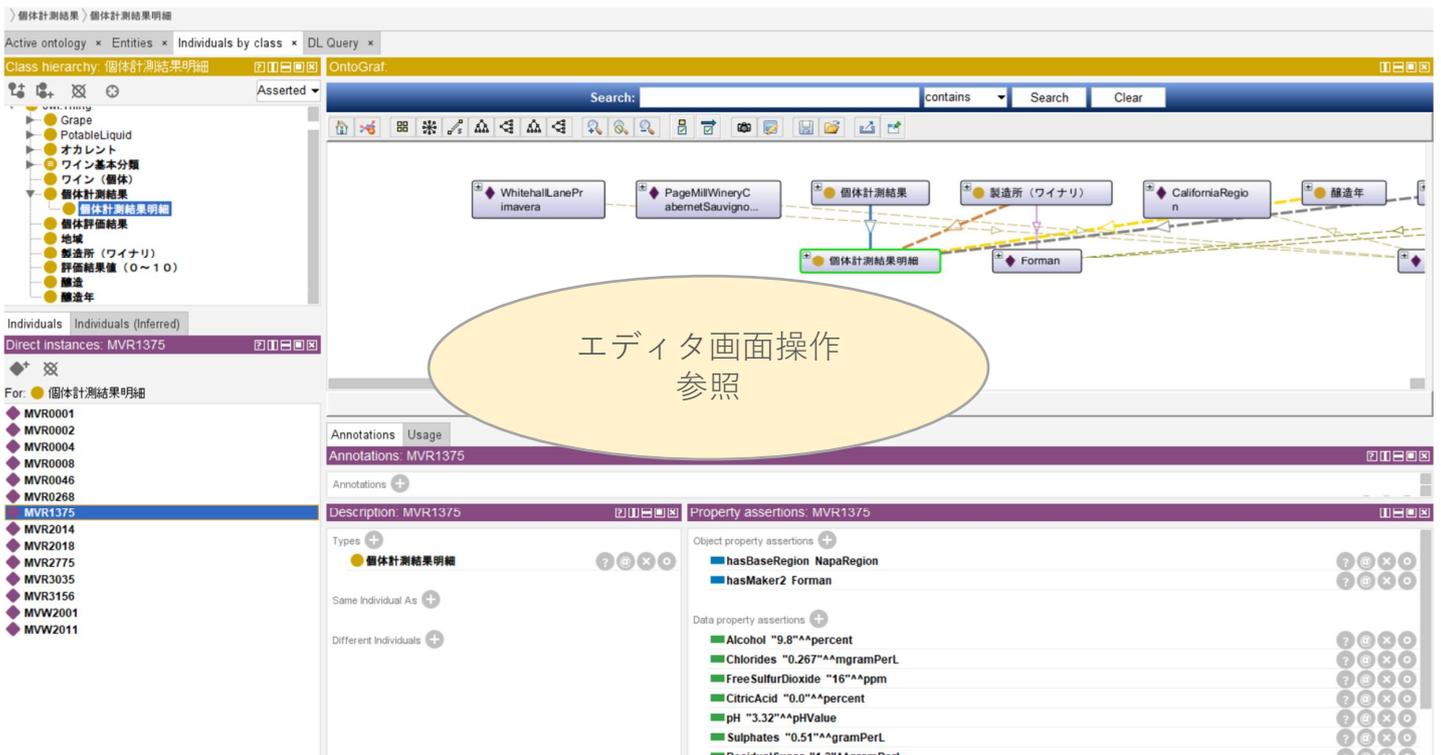
### ● DatatypeProperty の利用の様子（Protégé エディタ画面）



©2024 インフオラボ游悠 [研究所]

7

### ● RDF/OWLファイルの展開の様子（Protégé エディタ画面を参照）



©2024 インフオラボ游悠 [研究所]

8

# 【参考 1】 GraphDB Desk Top (無料ライセンスのあるグラフDBソフトウェア)

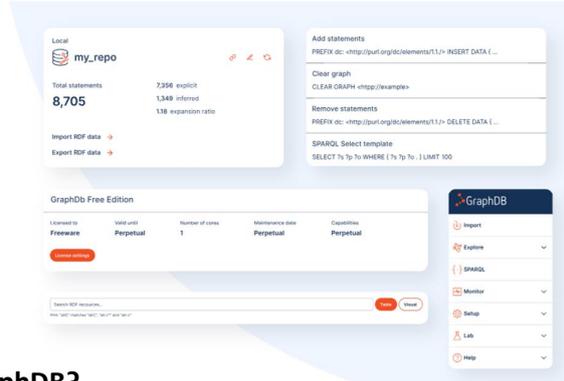


<https://www.ontotext.com/products/graphdb/>

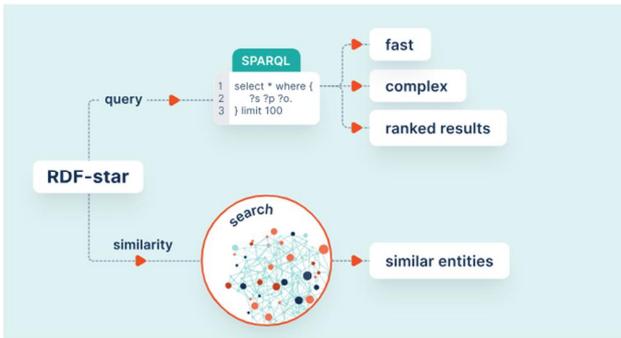
## Ontotext GraphDB

Link diverse business data, index it for semantic search, and enrich it via text analysis. Unlock insights and turn information into business knowledge.

[Download GraphDB](#)



### Why GraphDB?



### Build, consume, govern enterprise knowledge graphs with GraphDB

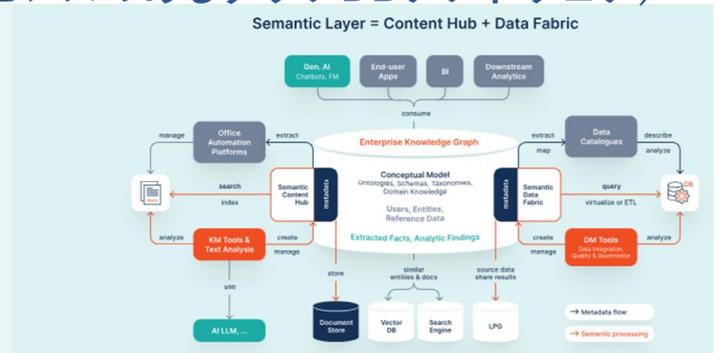
- Unify data via a shared conceptual model.
- Avoid misinterpretation via formal semantics.
- Manage data quality and trace provenance.
- Uncover new relationships with reasoning.
- Employ standards to avoid vendor lock-in.

©2024 インフォラボ游悠 [研究所]

# 【参考 1】 GraphDB Desk Top (無料ライセンスのあるグラフDBソフトウェア)

## Connect the dots across your IT systems

- Implement a data fabric to gain unified access to your structured data.
- Use conceptual models to search and recommend documents.
- Ask natural language questions across multiple data sources.



## Unlock the power of AI & LLM with GraphDB

Allow users to search for specific terms and related concepts to help them get faster insights.

- Talk to your graph: Use no-code RAG chat in GraphDB Workbench.
- Explain SPARQL queries and their results.
- NL Query your graph from LangChain.
- Simplify knowledge graph building.



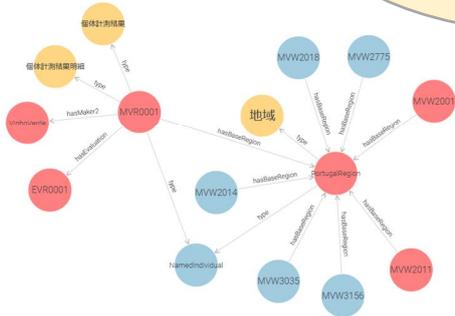
©2024 インフォラボ游悠 [研究所]

● グラフデータベースに、作成したRDF/OWLを取込んだ結果の展開の様子（GraphDB画面を参照）

The screenshot shows the GraphDB interface with a class hierarchy on the left and a dependency table on the right. The dependency table lists classes and their link counts:

Class	Links
:Region	149
:WineSugar	86
:WineFlavor	86
:WineBody	82
:Winery	61
:WineColor	20
wine.EvaluatedValue	14
wine.EvaluatedResult	11

Below the table is a circular sunburst chart showing the distribution of these classes. A central yellow oval contains the text "GraphDB画面操作 参照".

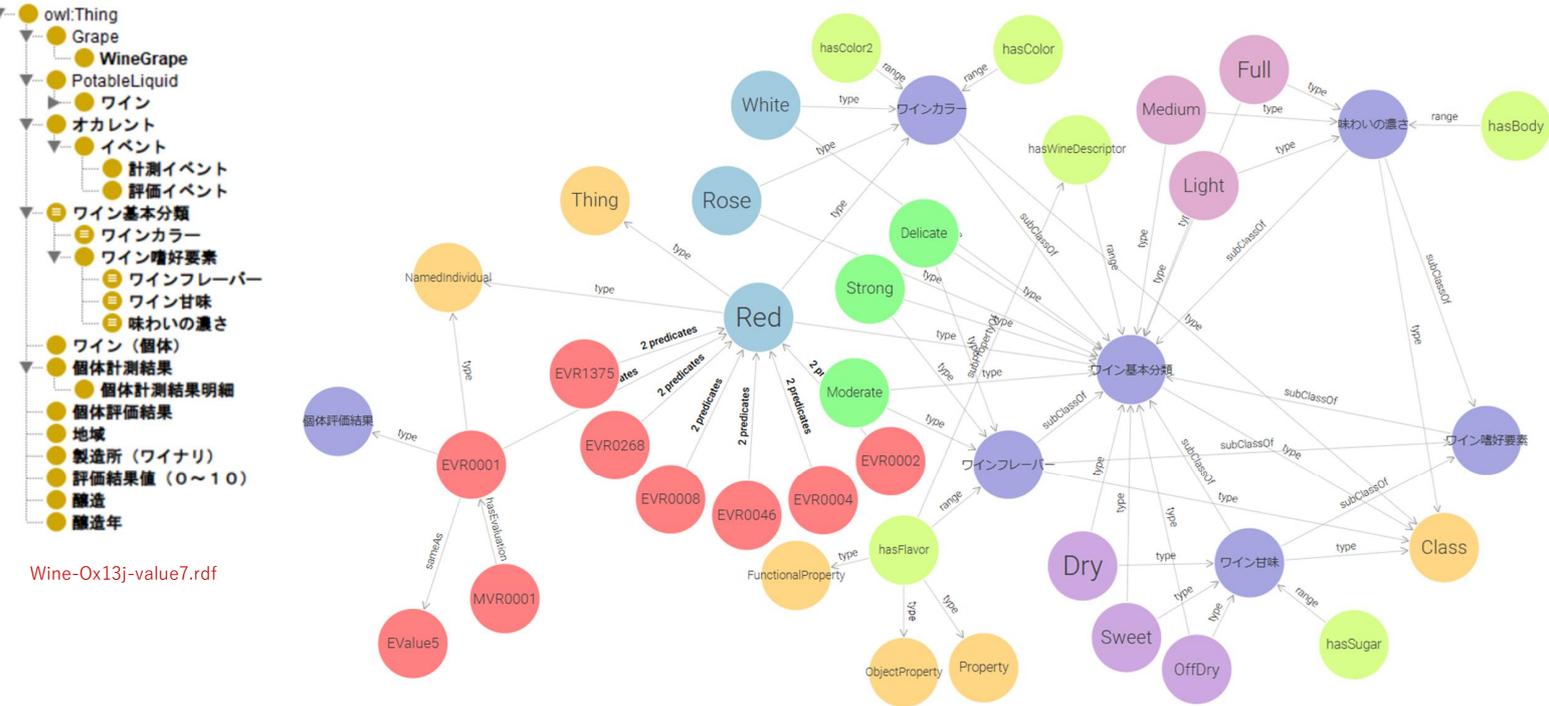


SPARQL Query & Update

```

1 prefix onto:<http://www.ontotext.com/>
2 select ?s {
3   ?s a <https://example.org/WineOntology/wine#EvaluatedValue> .
4 }
    
```

・ 14件インスタンス(赤7件/白7件) …… 赤/EVR0001からの展開。ワイン嗜好要素3分類とワインフレーバーが展開表示されている。

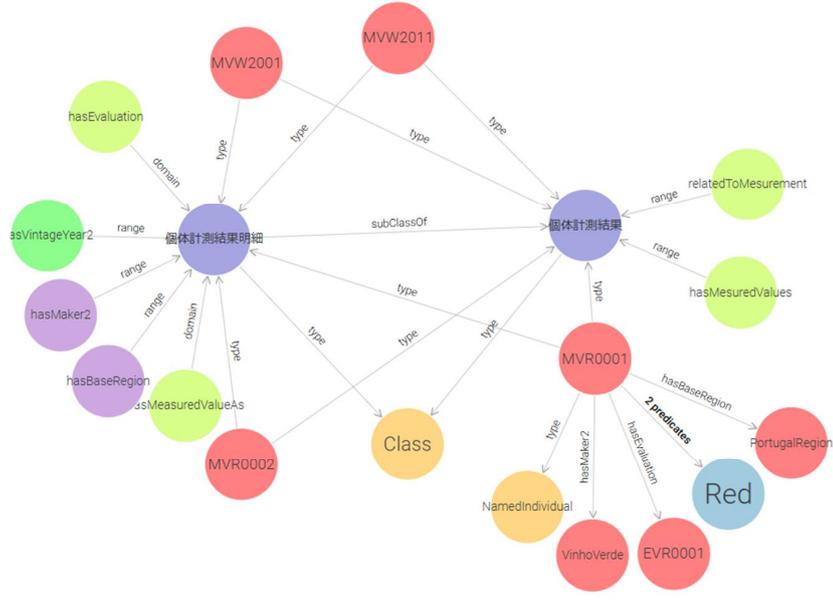


Wine-Ox13j-value7.rdf

- 個体計測結果クラスから MVR0001を選択し、明細のリレーションまでを表示し、MVR0001の計測結果の計測値詳細を表示したグラフ



Wine-Ox13j-value7.rdf



**MVR0001**

MVR0001  
ワイン個体の計測値詳細

Types:  
owl:NamedIndividual  
wineIndivMeasuredResult

RDF Rank:  
0

Q Search instance properties

measuredVal:Alcohol  
9.4 <https://example.org/datatype#percent>

measuredVal:Chlorides  
0.076 <https://example.org/datatype#mgramPerL>

measuredVal:CitricAcid  
0.0 <https://example.org/datatype#percent>

measuredVal:Density  
0.9978 <https://example.org/datatype#kgPer3>

measuredVal:FixedAcidity  
7.4 <https://example.org/datatype#mEqPerL>

measuredVal:FreeSulfurDioxide  
11 <https://example.org/datatype#ppm>

measuredVal:ResidualSugar  
1.9 <https://example.org/datatype#gramPerL>

measuredVal:Sulphates  
0.56 <https://example.org/datatype#gramPerL>

measuredVal:TotalSulfurDioxide  
3.4 <https://example.org/datatype#ppm>

measuredVal:VolatileAcidity  
0.7 <http://www.w3.org/TR/2003/PR-owl-guide-20031209/gramPerL>

measuredVal:pH  
3.51 <https://example.org/datatype#pHValue>

2024/11/24  
時点