

LOD Surfer API ver.2: クラス間関係に基づく連合検索を利用したLOD探索のためのウェブAPI

ライフサイエンス統合データベースセンター
山口 敦子

理化学研究所 情報システム本部
小林 紀郎
理化学研究所 バイオリソースセンター
榎屋 啓志

ライフサイエンス統合データベースセンター
山本 泰智
大阪電気通信大学 情報通信工学部
古崎 晃司

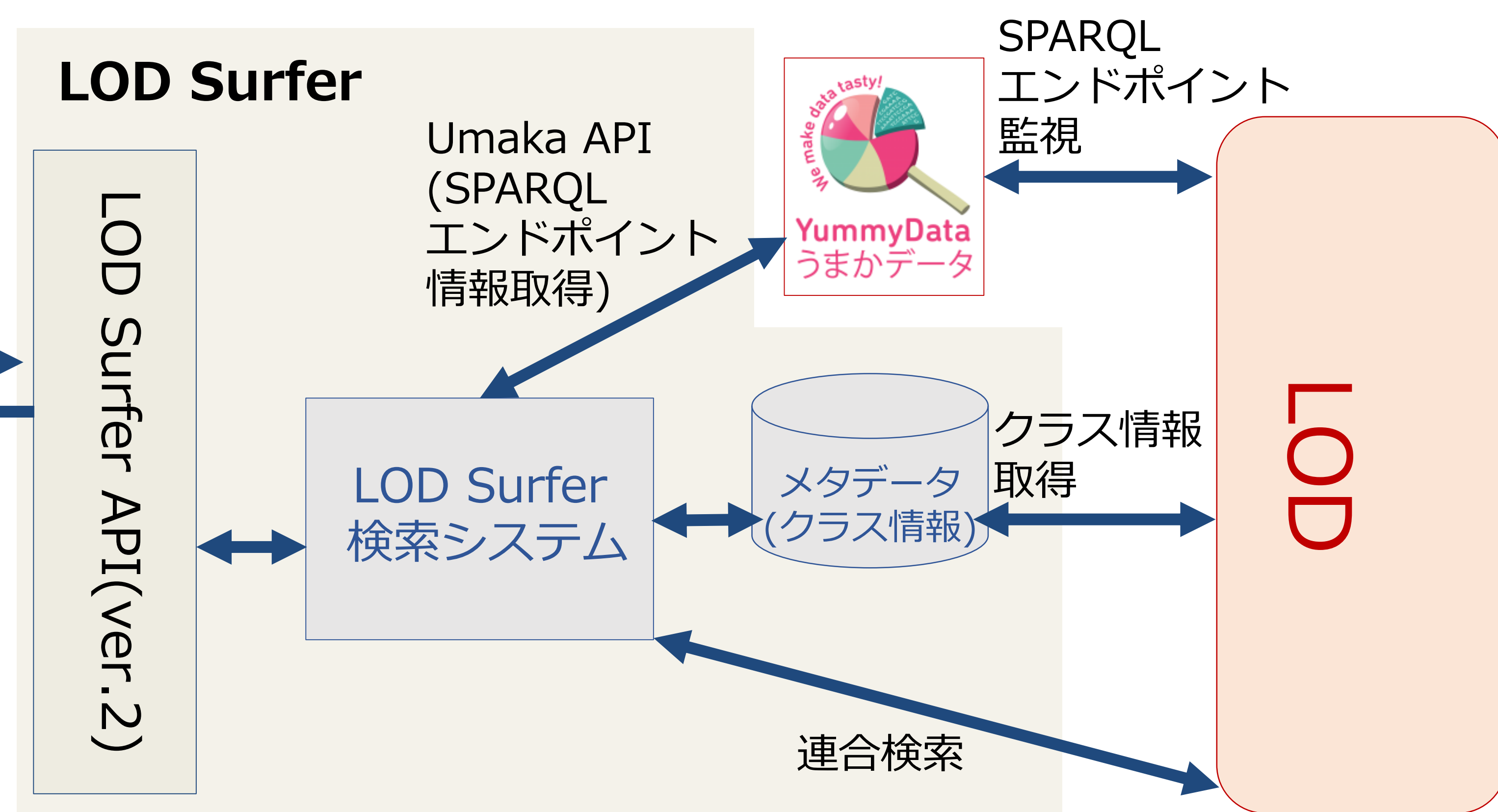
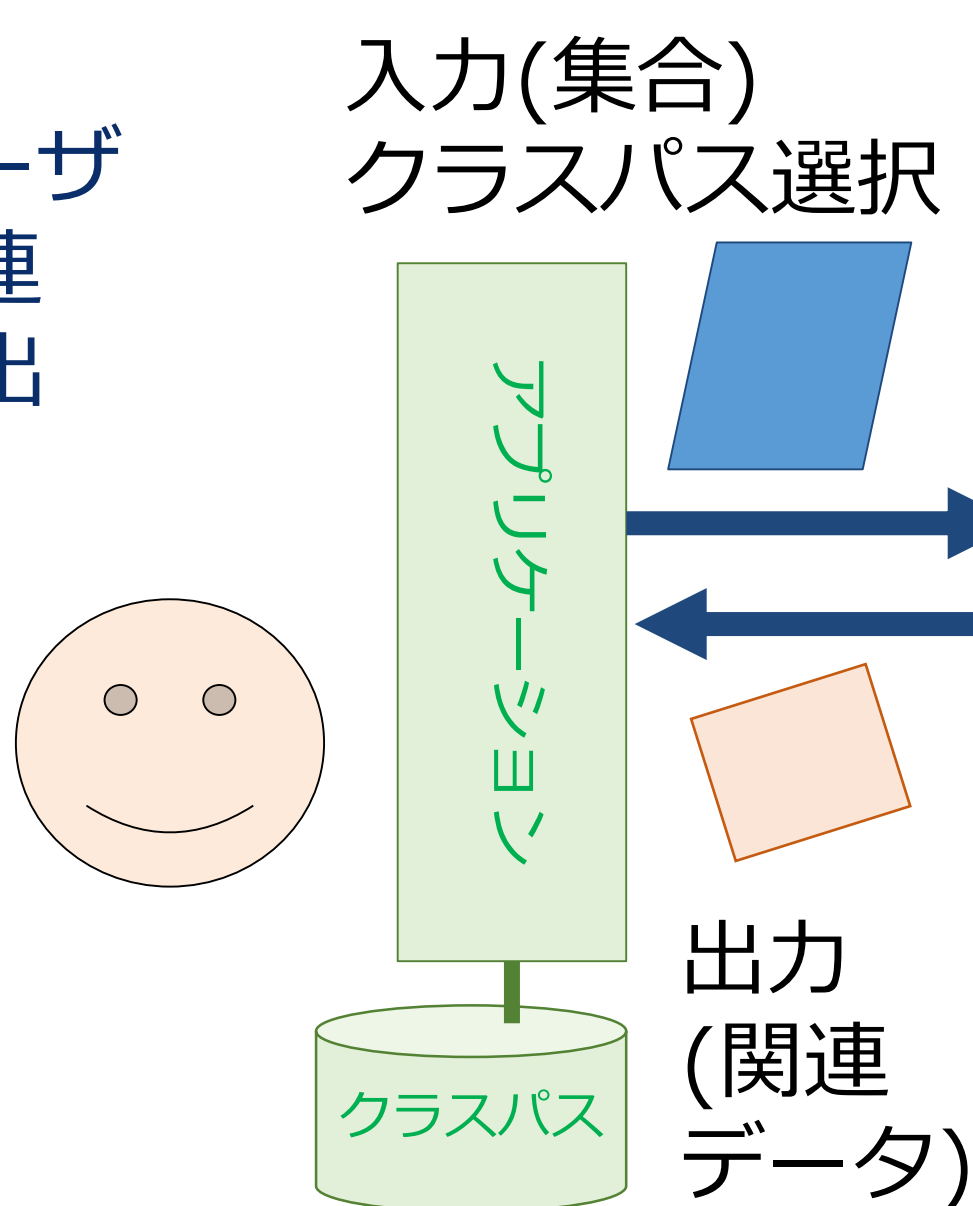
Linked Open Data (LOD)上の情報を連合検索を用いて柔軟に取得

発表者らは、連合検索を利用して、LOD上の情報をクラス間関係で切り取るシステム LOD Surfer の実現へ向けた開発と研究をすすめてきた。まず、LOD Surfer上で必要となる機能を検討し、それらの機能をLOD Surfer APIとよぶウェブAPIとしてプロトタイプ実装した。さらに、タンパク質アラインメントビューアという、ドメインを絞った具体的なアプリケーション上で利用することにより、実用上の課題の洗い出しを行った。本発表では、実用上の課題の洗い出しによって必要となった機能のうち、特にクラス間のパス選択に関するものに着目し、LOD Surfer API に追加実装した LOD Surfer API ver.2 について紹介する。このウェブAPIを基盤として、LOD上に分散するデータの統合利用を可能とするために、必要となる技術を議論しつつ、今後の展望について述べる。

LOD Surfer

LOD Surfer とは、対話的なGUI を介して、ユーザが興味をもつデータやキーワードに対して、関連するLOD上の情報をクラス間関係で切り取り、出力するシステムである。

入力されたデータは、アプリケーションを通して、LOD上のクラスに紐づけられ、さらに、そのクラスを始点としたクラス間関係探索により、入力と関連する部分をLODから切り取り提示することができる。

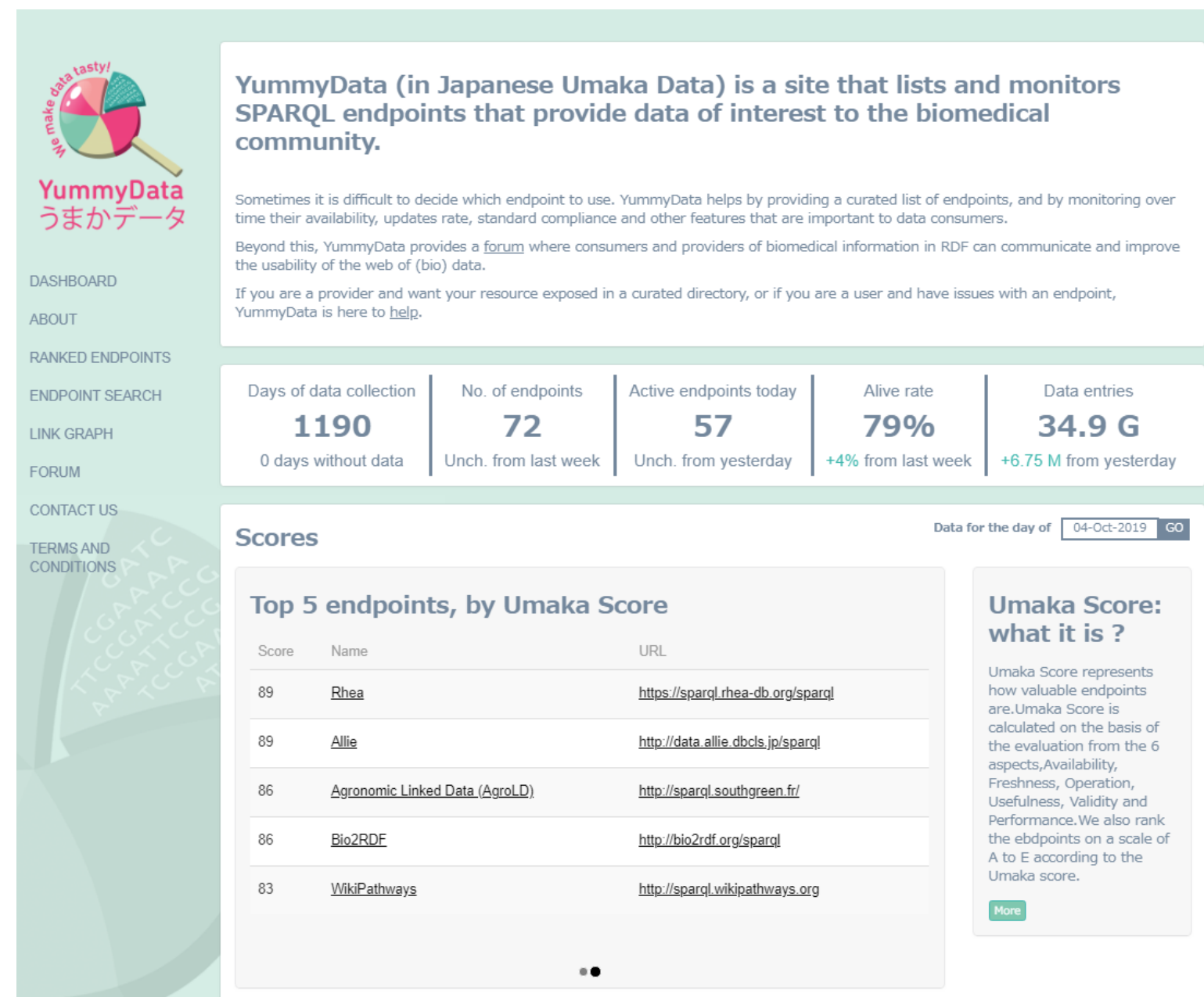


Umaka APIの利用

Umaka-YummyData は主にライフサイエンス分野のSPARQLエンドポイントの監視サービスである。

このサービスのAPI (Umaka API)を用いて、
・落ちていない、
・応答速度が速い、
SPARQLエンドポイントを選ぶことができる。

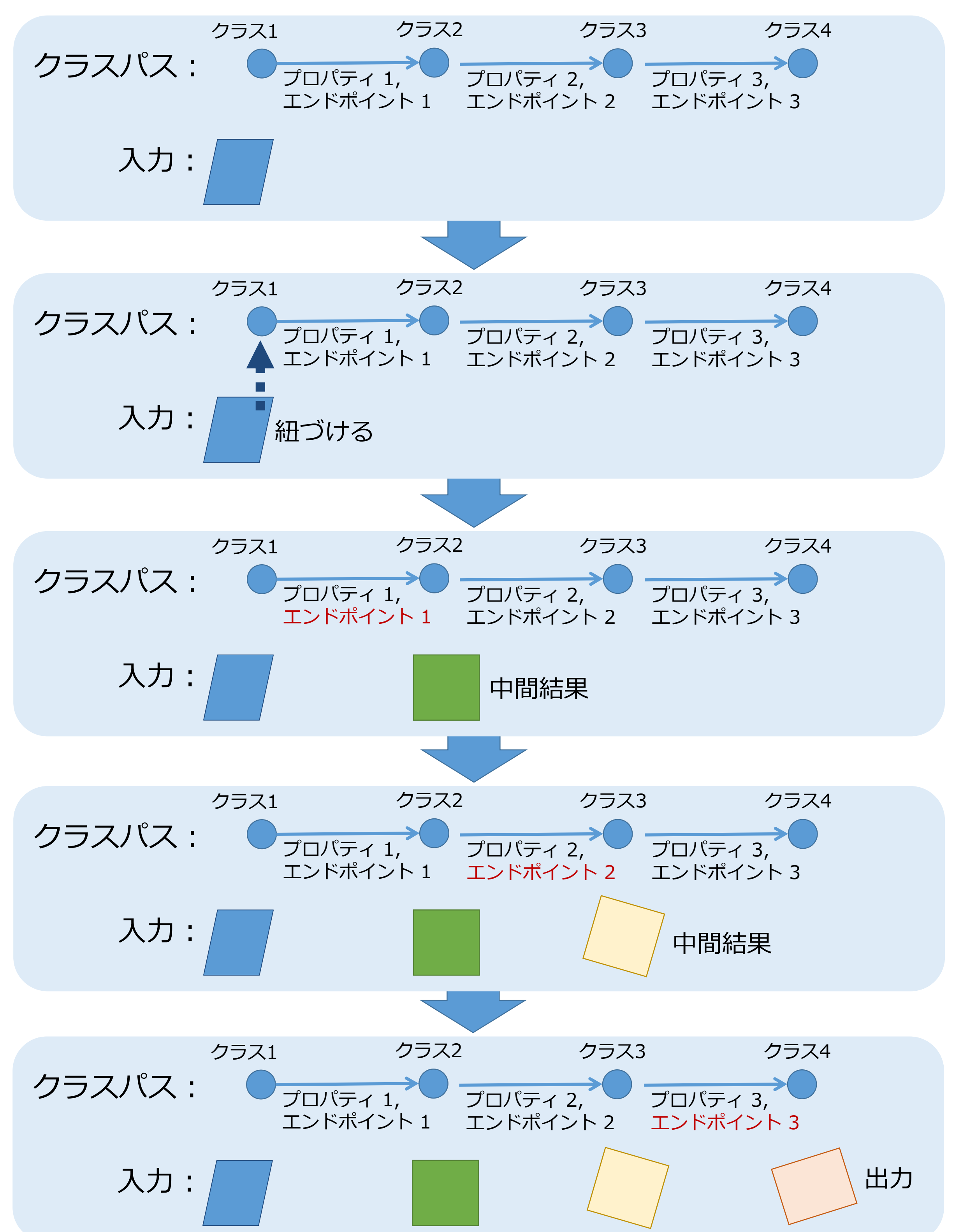
選んだSPARQLエンドポイントのみを利用することで、安定した連合検索を提供できる。



<http://yummydata.org/>

クラスパスを用いた連合検索

入力をパスの先頭のクラスのインスタンスとみなして紐づけ、パスを辿りつつ検索を進めることで、最終的な出力を求めることができる。

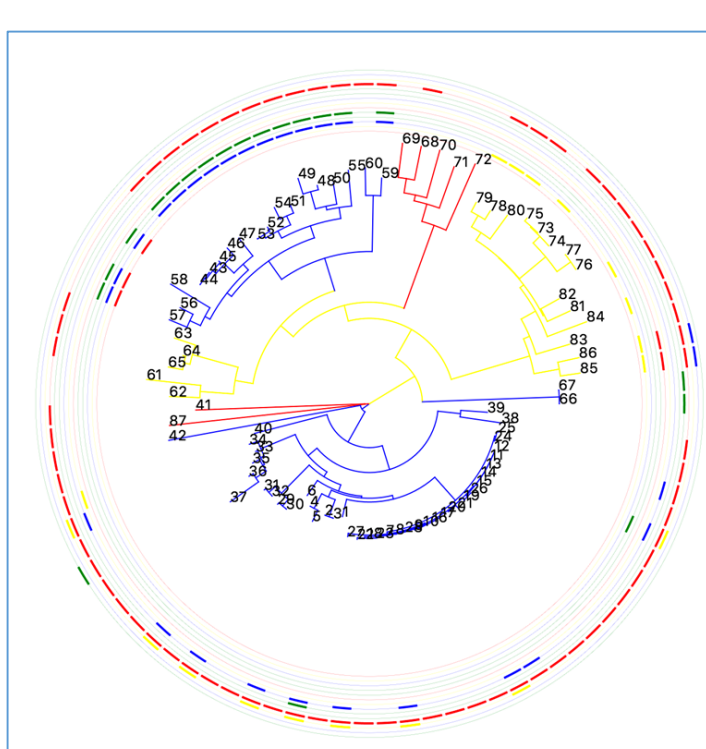
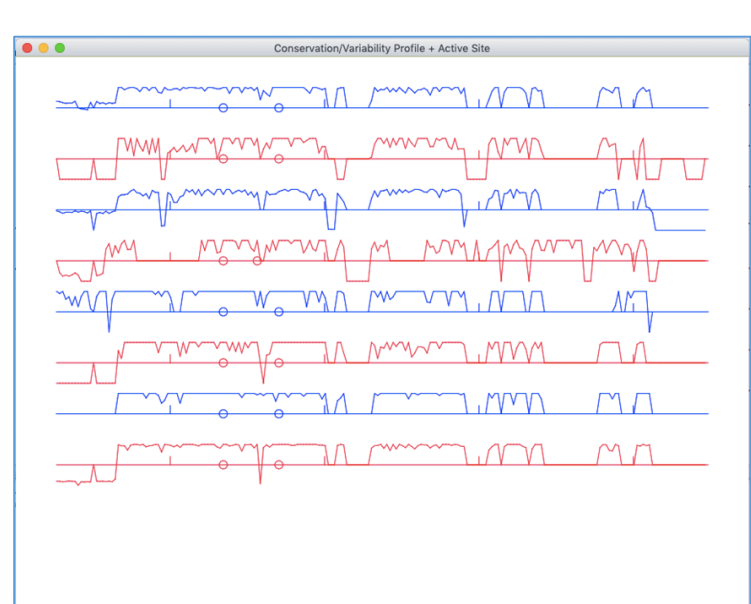


LOD Surfer API ver.2 とその利用例

LOD Surfer API ver.2

- SPARQL エンドポイントのリスト: /api/eplist
- クラスの全リスト: /api/clist
- キーワードをラベルかクラスURLに含むクラスのリスト: /api/clist?keyword=[keyword]
- class1からパスでたどれるクラスのリスト: /api/clist?class1=[classURI]
- class1からclass2へのパスリスト: /api/plist?class1=[classURI]&class2=[classURI]
- pathに対応した連合SPARQLクエリ: /api/sparql?path=[path]
- inputを入力としたときのpathに対応した出力: /api/result?path=[path]&input=[input]

利用例: ASHViewer



タンパク質マルチプルアラインメントビューワのアノテーションに利用

今後の課題

- 利用例の増加
- 入力とクラスの紐づけの自動化
- Umaka search(URL/リテラル→エンドポイント) との連携