

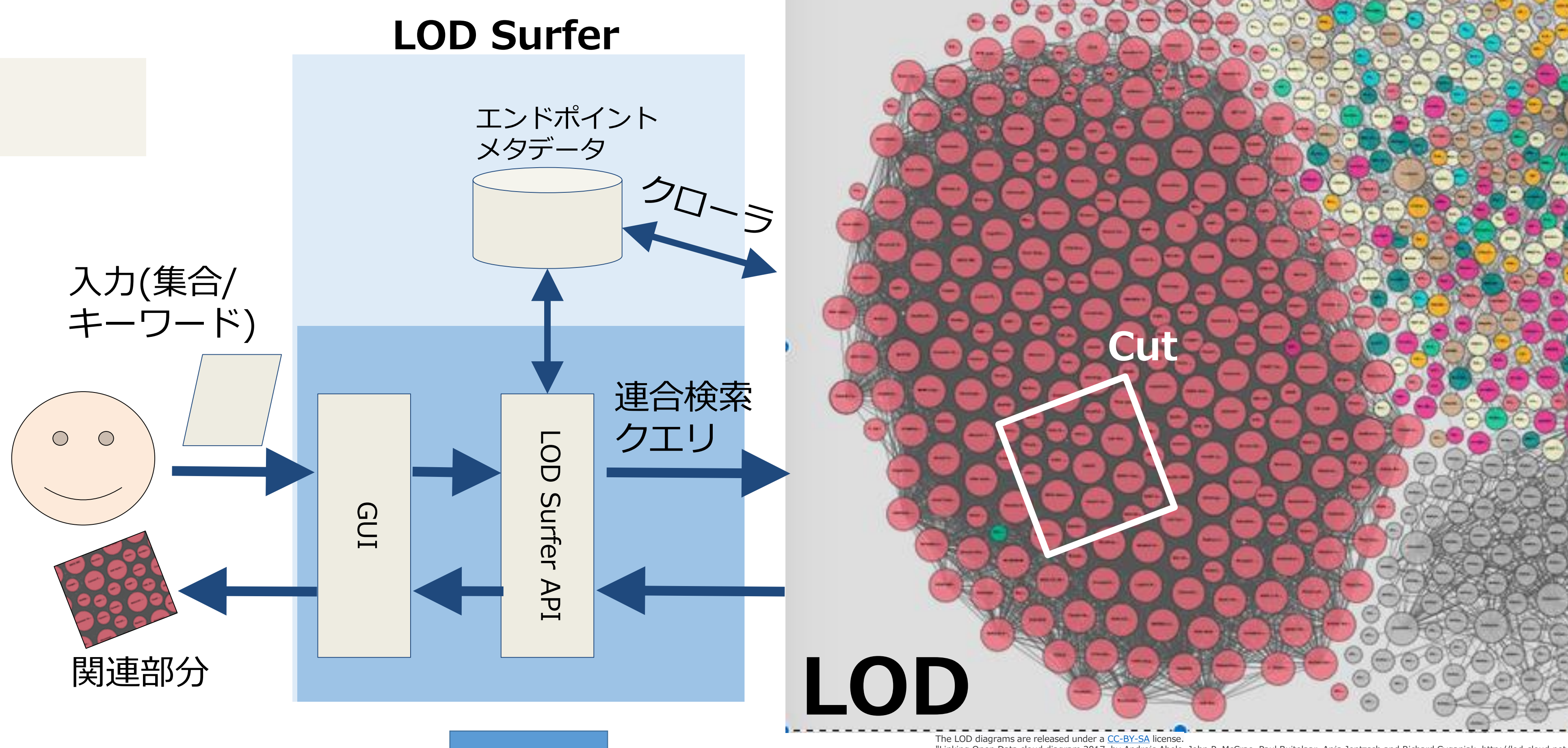
Linked Open Data (LOD)上の情報を自在に切り取るために

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、今年度は、ドメインを絞り、タンパク質アラインメントビューアという具体的なアプリケーション上で利用することで、LOD Surfer-APIの実用上の課題の洗い出しと、それらの課題解決による総合的な改良をめざした。その結果、(1) SPARQLエンドポイントを跨ぐ場合、クラス間パスの数が爆発的に増えるため、ユーザにアプリ上でクラス間パスを動的に指定させるのは現実的ではない、(2) 入力数(タンパク質数)が大きくなると分散したエンドポイントへの連合検索をアプリケーション上でユーザが負担に感じない時間で動かすのは困難である、という二点が課題となった。その解決のため、課題(1)に関しては、アプリケーション開発者が推奨するパスを事前を選んで蓄積する手法を、課題(2)に関しては、事前に必要な情報を取得して蓄積する手法を採用した。本発表では、これらの開発した技術を紹介しつつ、今後のLOD Surferシステム全体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。

入力されたデータやキーワードは、システム内で、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。



PSurfer: タンパク質特化型 LOD Surfer

パス選択ツール

SPARQLエンドポイントを跨ぐ場合、クラス間関係の数が爆発的に増えるため、ユーザにアプリ上でクラス間パスを動的に指定させるのは現実的ではない。そのため、事前にアプリケーション開発者にパスを選ばせ、そのパスを蓄積する。

クラス情報

エンドユーザに人間可読な情報を提供するために、rdfs:label 以外のプロパティで、有用な情報を提供するものをマニュアルで蓄積したもの。

PSurfer

ASHViewerからアノテーション種類(=パス)の指定を受け取り、パスとクラス情報を利用して、人間可読な情報を返す連合検索クエリを作成し、検索を行う。一度行った検索の結果は蓄積する。

パスの蓄積

選択されたパスは以下の形式で保存される：

```
@base <http://sparqlbuilder.org/lodsulfer/core/> .
@prefix ldp: <http://sparqlbuilder.org/lodsulfer/path/> .
```

```
ldp: {id} rdfs:label "path name" ;
  rdf:type <Path> ;
  <description> "description for this path" ;
  <inputClass> <classURI of input class> ;
  <outputClass> <classURI of output class> ;
  <classList> ldp: {id}/class/0 ;
  <propList> ldp: {id}/prop/0 .
```

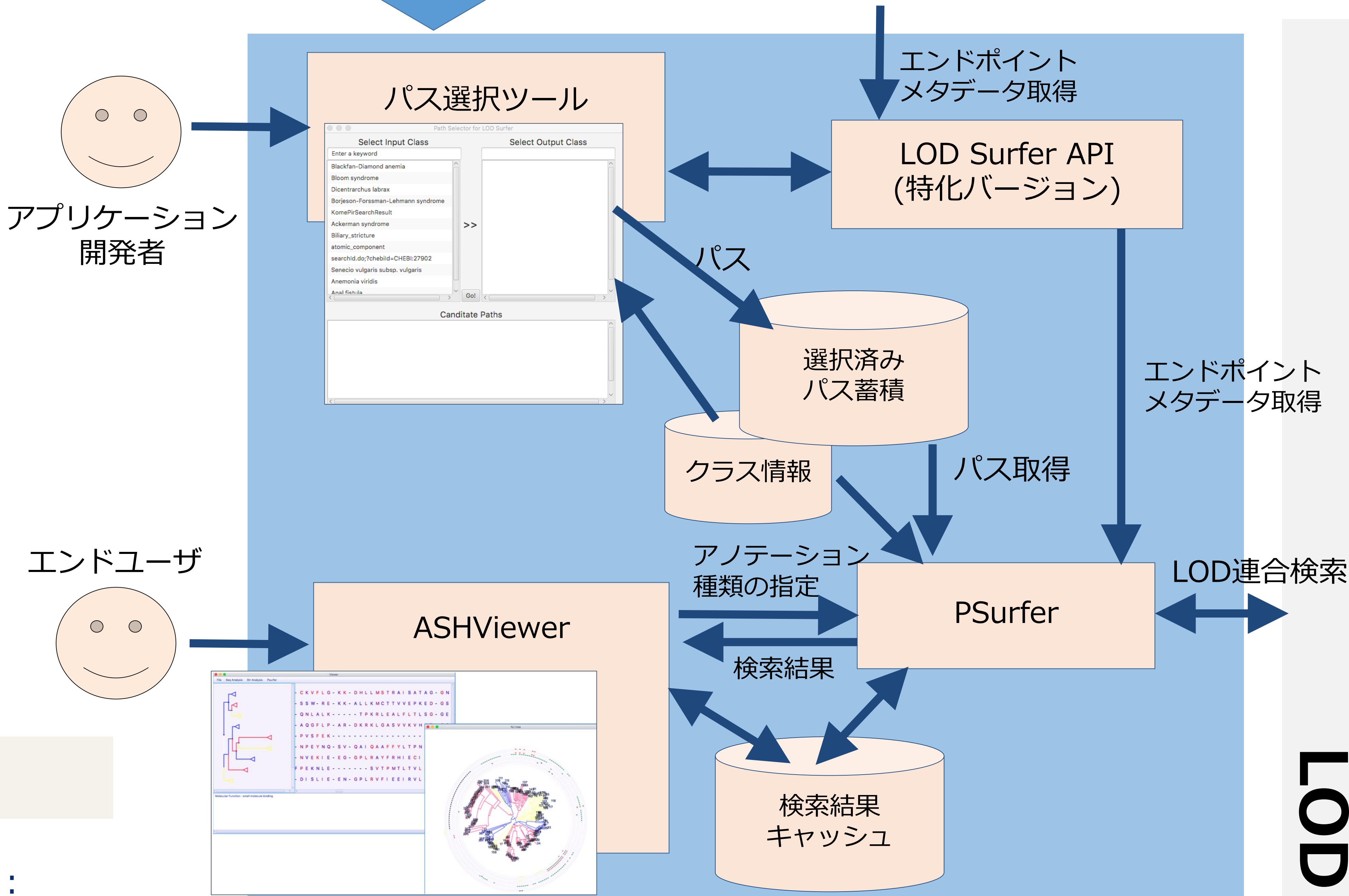
```
ldp: {id}/class/0 rdf:type rdf:List;...
```

```
ldp: {id}/prop/0 rdf:type rdf:List;...
```

※ASHViewerについては、ポスター9で説明される

クラス情報の例

http://purl.uniprot.org/core/Taxon に対し、
 http://purl.uniprot.org/core/scientificName
 http://purl.uniprot.org/core/commonName
 が表示すべき情報として検索エンジンに渡される。



LOD