

# LODQA

## Question-Answering over Linked Open Data

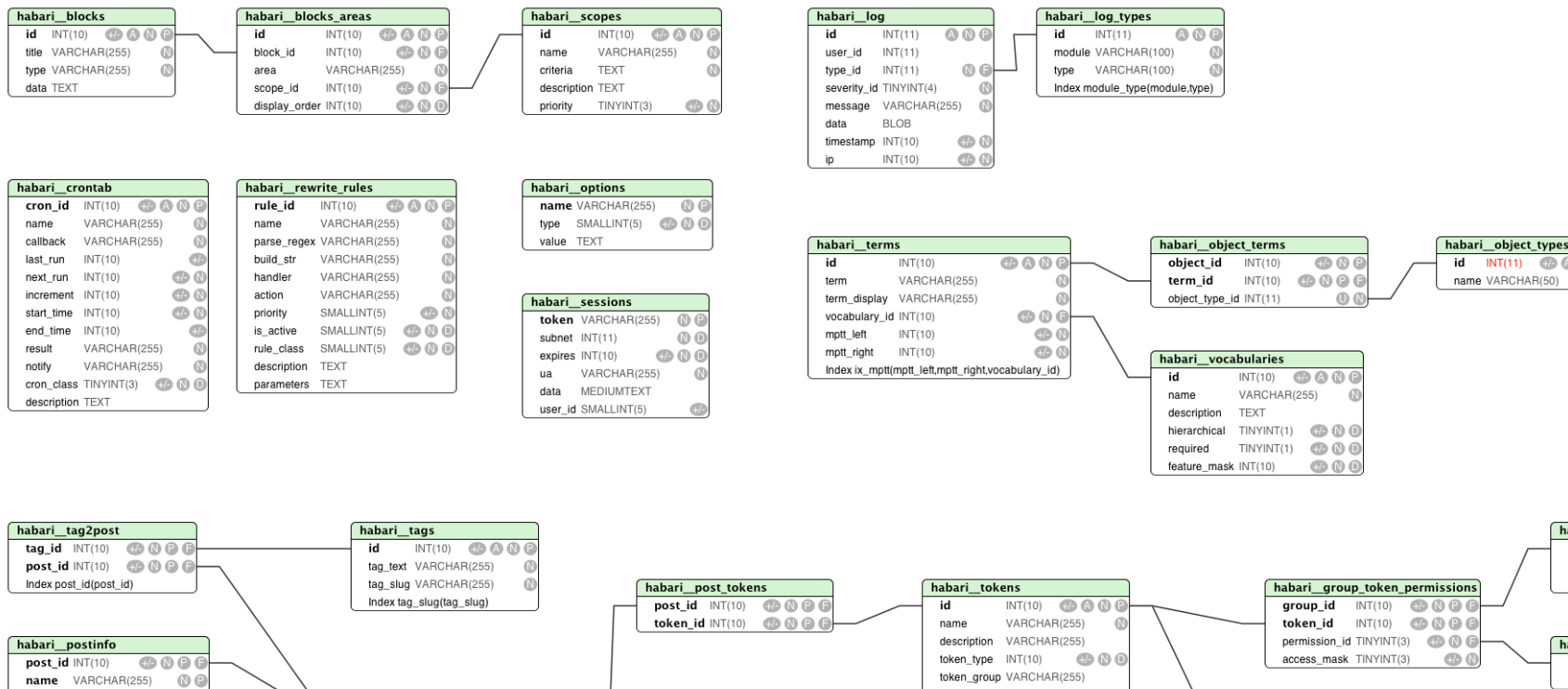
<http://lodqa.org>  
(demo system)

Jin-Dong Kim  
Database Center for Life Science



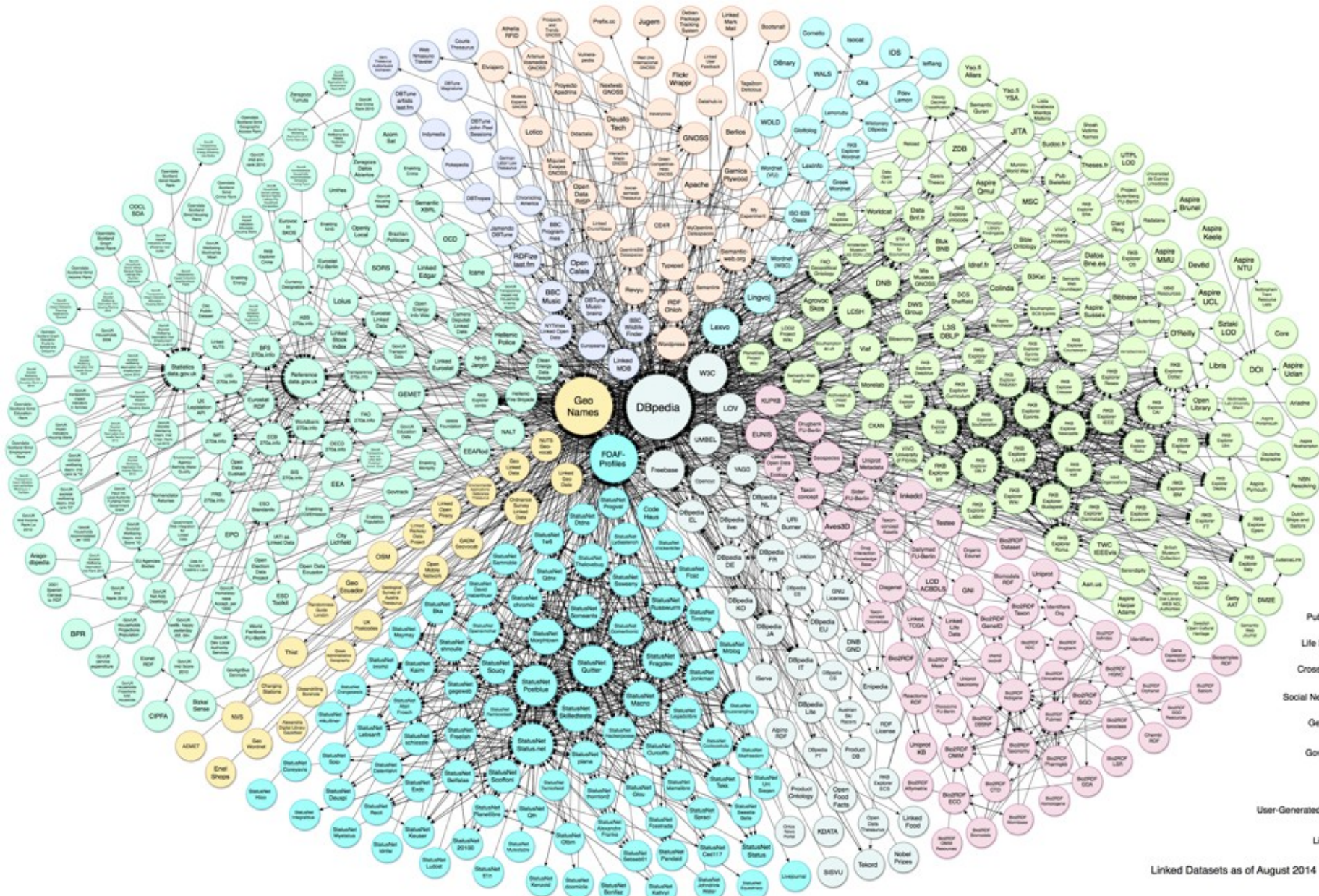
# Background

- Difficulties of authoring queries in SQL
  - ✓ Language (SQL)
  - ✓ Schema



# Background

- Linked Data (in RDF)
  - ✓ Is based on an extreme join model
  - ✓ Data pieces in heterogeneous schema are interlinked
    - Makes it even more difficult to figure out the schema.
- To author a SPARQL query, one needs
  - ✓ to learn the language (SPARQL), and
  - ✓ to figure out terms (URI) and schema (graph structure)



- Publications
- Life Sciences
- Cross-Domain
- Social Networking
- Geographic
- Government
- Media
- User-Generated Content
- Linguistics



# Natural Language QA

- To author queries in natural language, one **does not** need
  - ✓ To learn the language, and
  - ✓ To figure out the schema of data sets.

# Natural Language QA

- To author queries in natural language, one **does not** need
  - ✓ To learn the language, and
  - ✓ To figure out the schema of data sets.

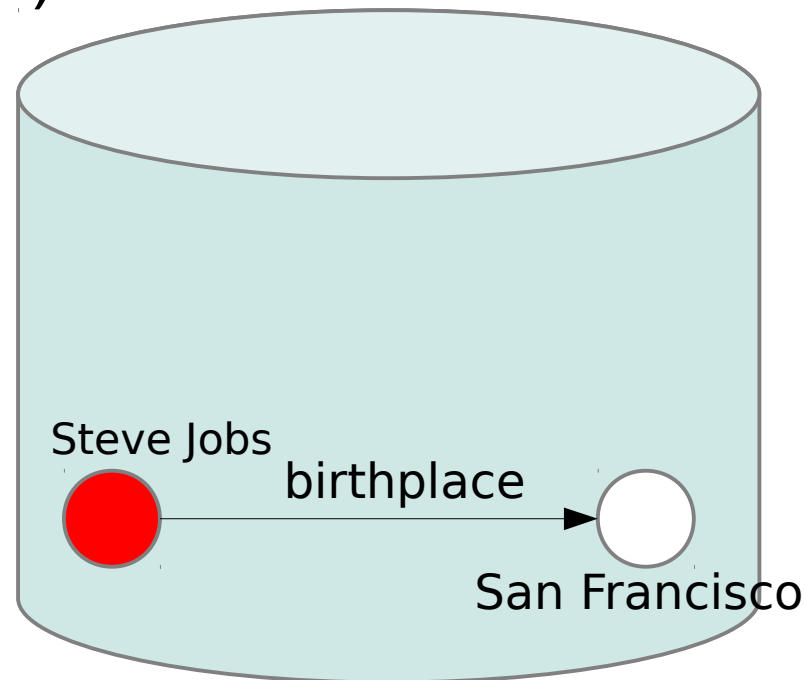
Queries are authored without considering how data is represented in database.

# Challenges

- Discrepancy

- ✓ Model representation (in NL)
- ✓ Data representation (in DB)
  
- ✓ Lexical discrepancy
- ✓ Structural discrepancy

What is the birthplace of Jobs?

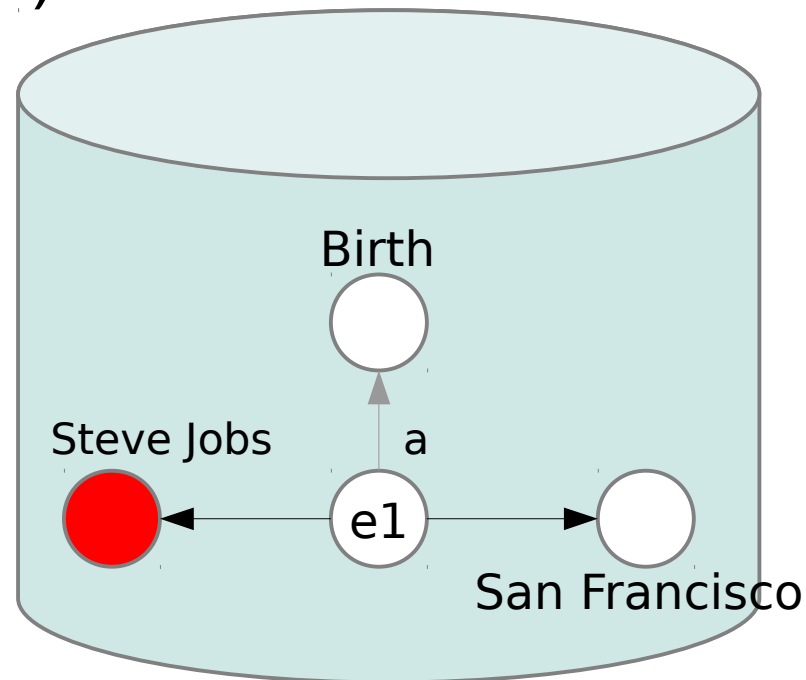


# Challenges

- Discrepancy

- ✓ Model representation (in NL)
- ✓ Data representation (in DB)
  
- ✓ Lexical discrepancy
- ✓ Structural discrepancy

What is the birthplace of Jobs?





# LODQA

- Demo system (<http://lodqa.org>)
- Documentation (<http://www.lodqa.org>)
- WS (REST)

# LODQA

- Linked Open Data Question-Answering
  - ✓ Input: natural language query
  - ✓ Intermediate output: SPARQL queries
  - ✓ Final output: solutions

# LODQA

- Architecture

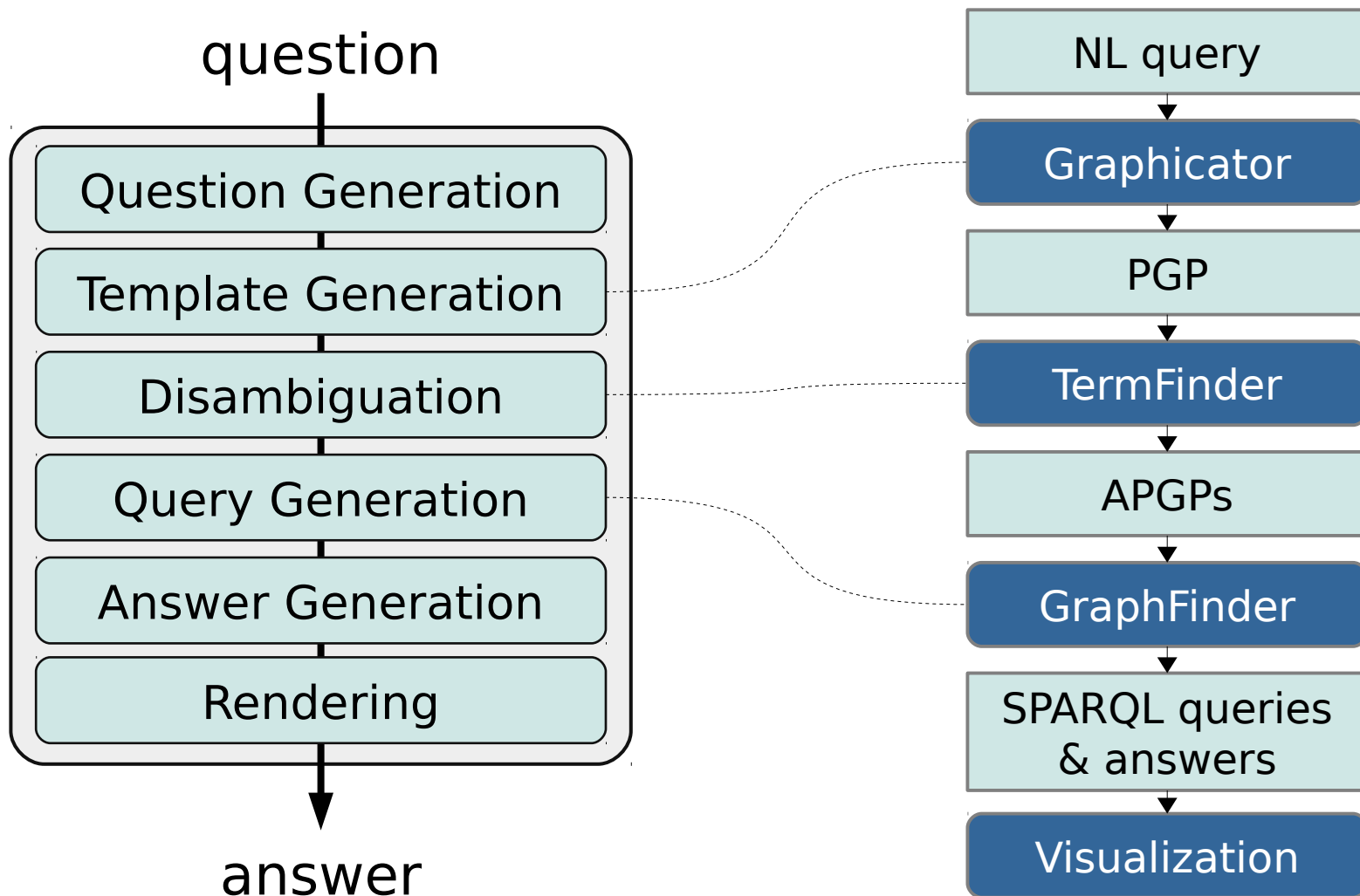
- ✓ <http://www.lodqa.org/docs/architecture/>

- Demo

- ✓ <http://lodqa.org/>

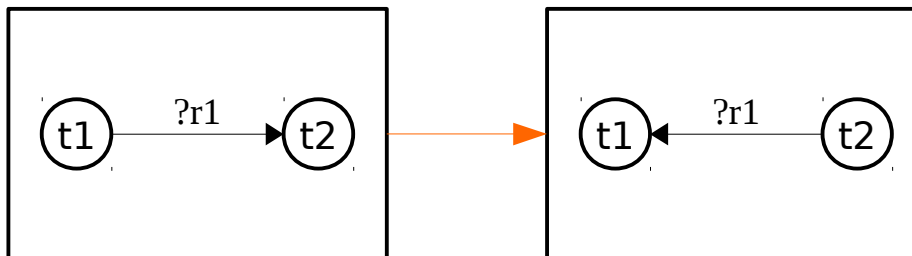


# Mapping to OKBQA framework

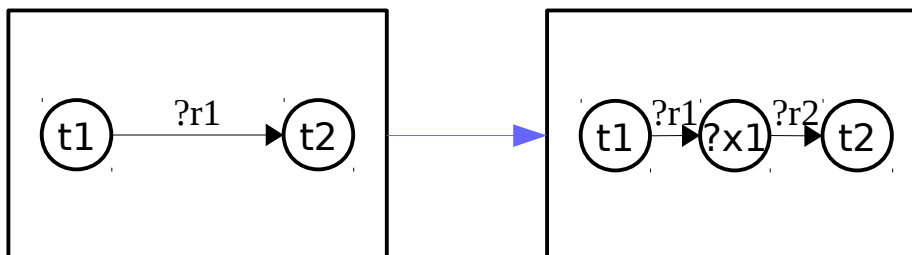


# GraphFinder operations

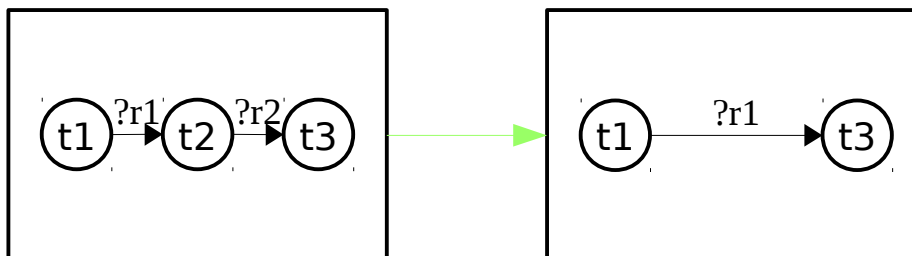
inversion



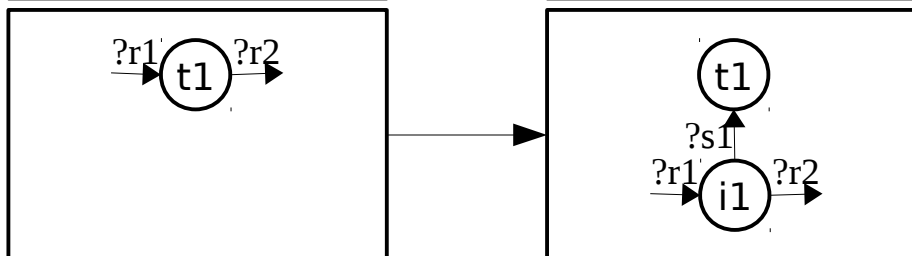
split



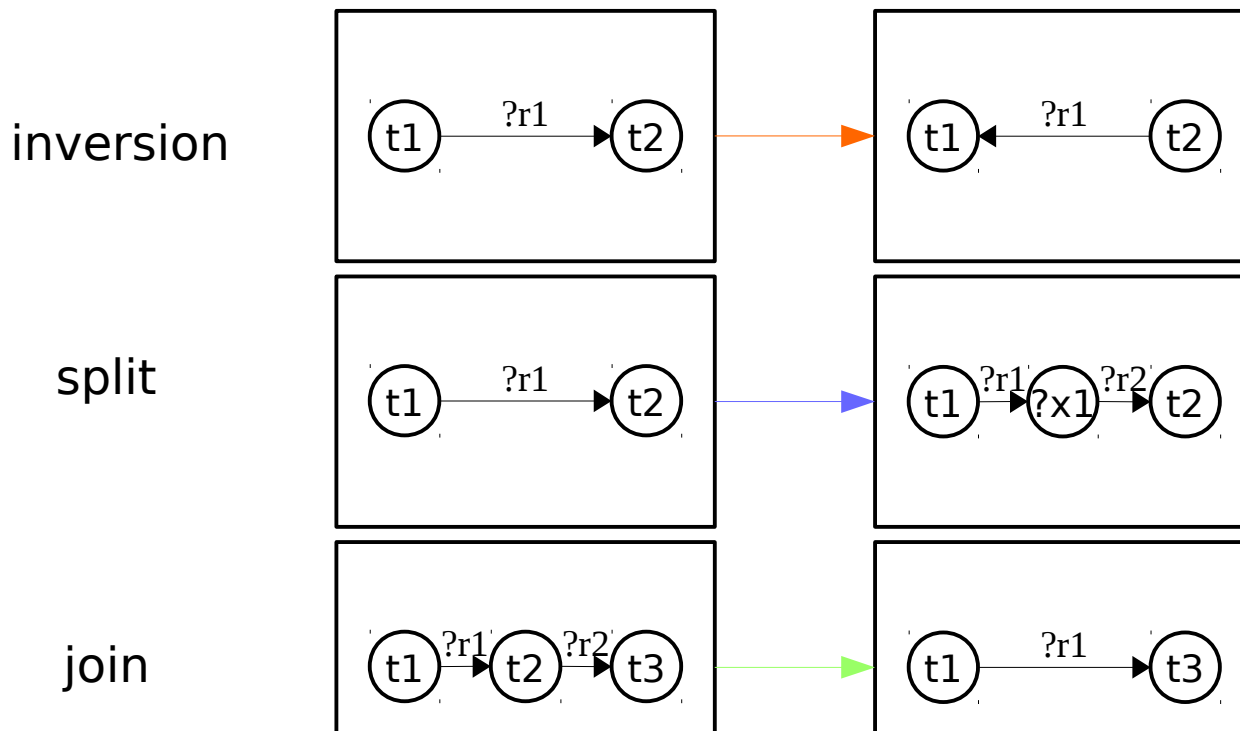
join



instantiation

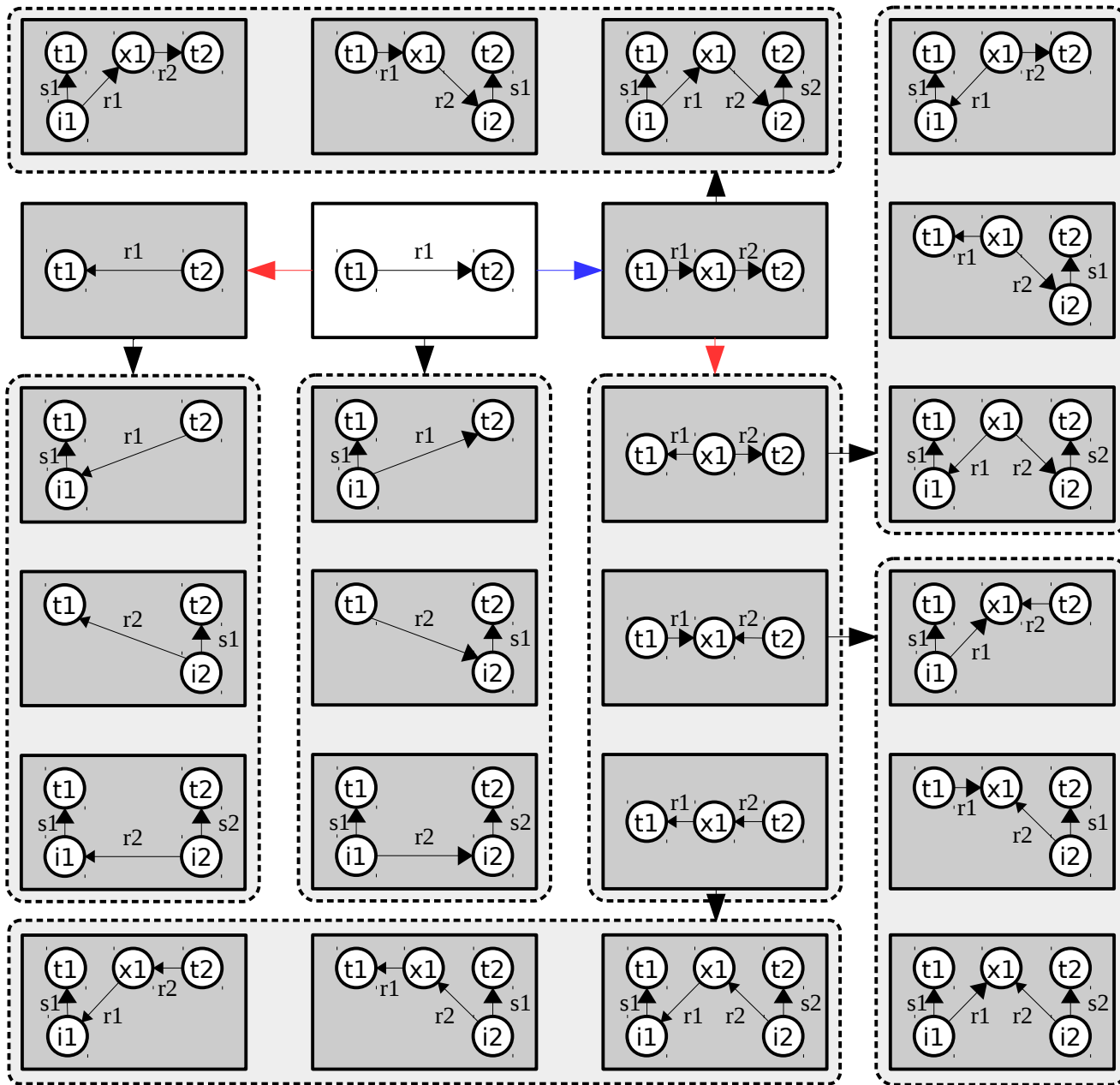


# GraphFinder operations

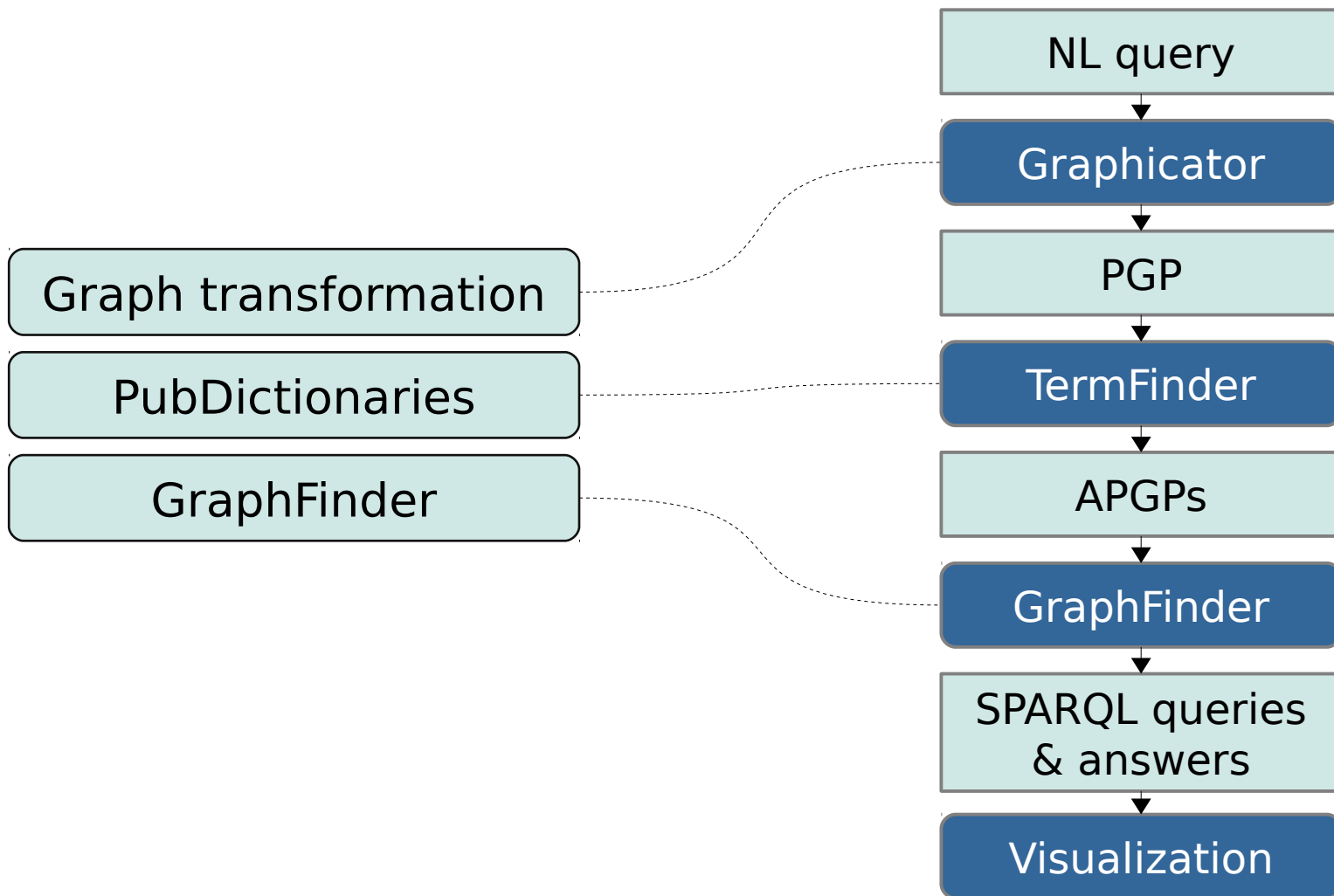


Jin-Dong Kim and Kevin Bretonnel Cohen,  
 "Triple Pattern Variation Operations for Flexible Graph Search",  
*Proceedings of the 1st international workshop on  
 Natural Language Interfaces for Web of Data (NLIWoD), 2014.*





# Techniques for each LODQA module



# Available Resources (1/2)

- LODQA
  - ✓ A highly adaptable QA system
  - ✓ <http://lodqa.org>
- GraphEditor
  - ✓ A web-based graph editor
- PubDictionaries
  - ✓ A crowd-sourcing web-based dictionary system
  - ✓ <http://pubdictionaries.org>



# Available Resources (2/2)

- PubAnnotation
  - ✓ A crowd-sourcing literature annotation system
  - ✓ <http://pubannotation.org>
- TextAE
  - ✓ A web-based text annotation editor
  - ✓ <http://textae.pubannotation.org>