



fédération de données et de ConnaissancEs
Distribuées en Imagerie BiomédicaLE

*Distributed Biomedical Imaging Data and
Knowledge federation*

Johan Montagnat



What is CrEDIBLE?

- National project funded by CNRS (French National Research Center)
 - Part of CNRS inter-disciplinary mission
 - “MASTODONS” call (**Challenge in Big Scientific Data**)
 - Five partners involved in France (CNRS, INSERM, INRIA, U. Picardie, U. Lyon)
- Started in 2012, extended on a yearly basis
- Targets **biomedical image resources federation**
 - Data fusion, mediation, semantic alignment, querying, link with processing

What is CrEDIBLE?

- Scientific networking initiative
 - Budget for networking and dissemination
 - Opening new contact opportunities
- Follow-up on various prior initiatives from project partners
 - Medical data sharing
 - Distributed computing
 - Large-scale data analysis

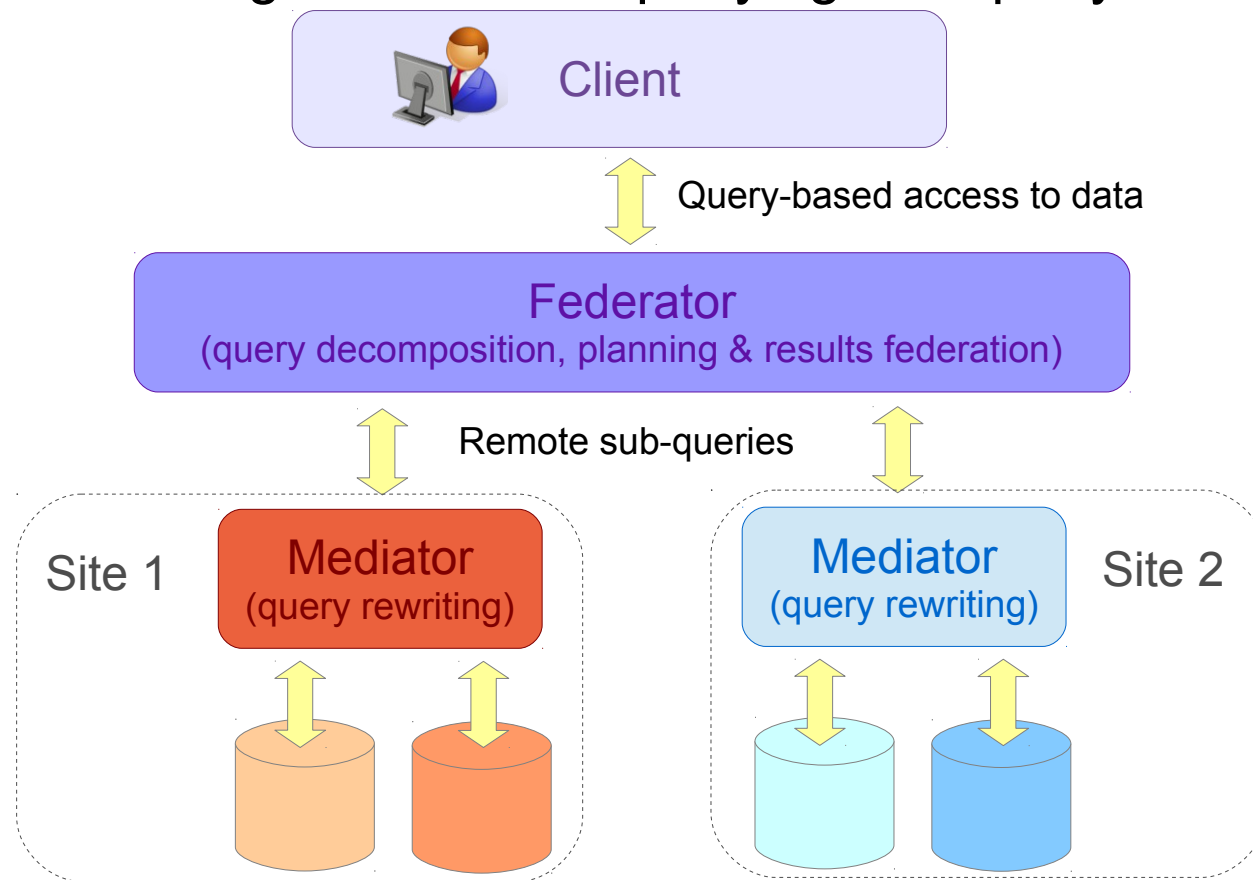
What is this workshop?

(https://credible.i3s.unice.fr/doku.php?id=2014_workshop)

- A multi-communities “think tank” opportunity
 - Intent to raise community discussions
 - 30 min. allocated per talk + panel for each session
 - Multi-disciplinary
 - Place for discussions
- How was it organized?
 - By invitation
 - Considering scientific themes of interest
 - Considering scientific challenges identified
- This year is the second edition
 - 2012 and 2013 reports:
<https://credible.i3s.unice.fr/doku.php?id=rappports>

Biomedical data mediation & federation

- Heterogeneous databases schema mediation
- Data federation through distributed querying and query rewriting



- Scientific experiment support platform:
 - raw data + models + processing results + models + provenance...

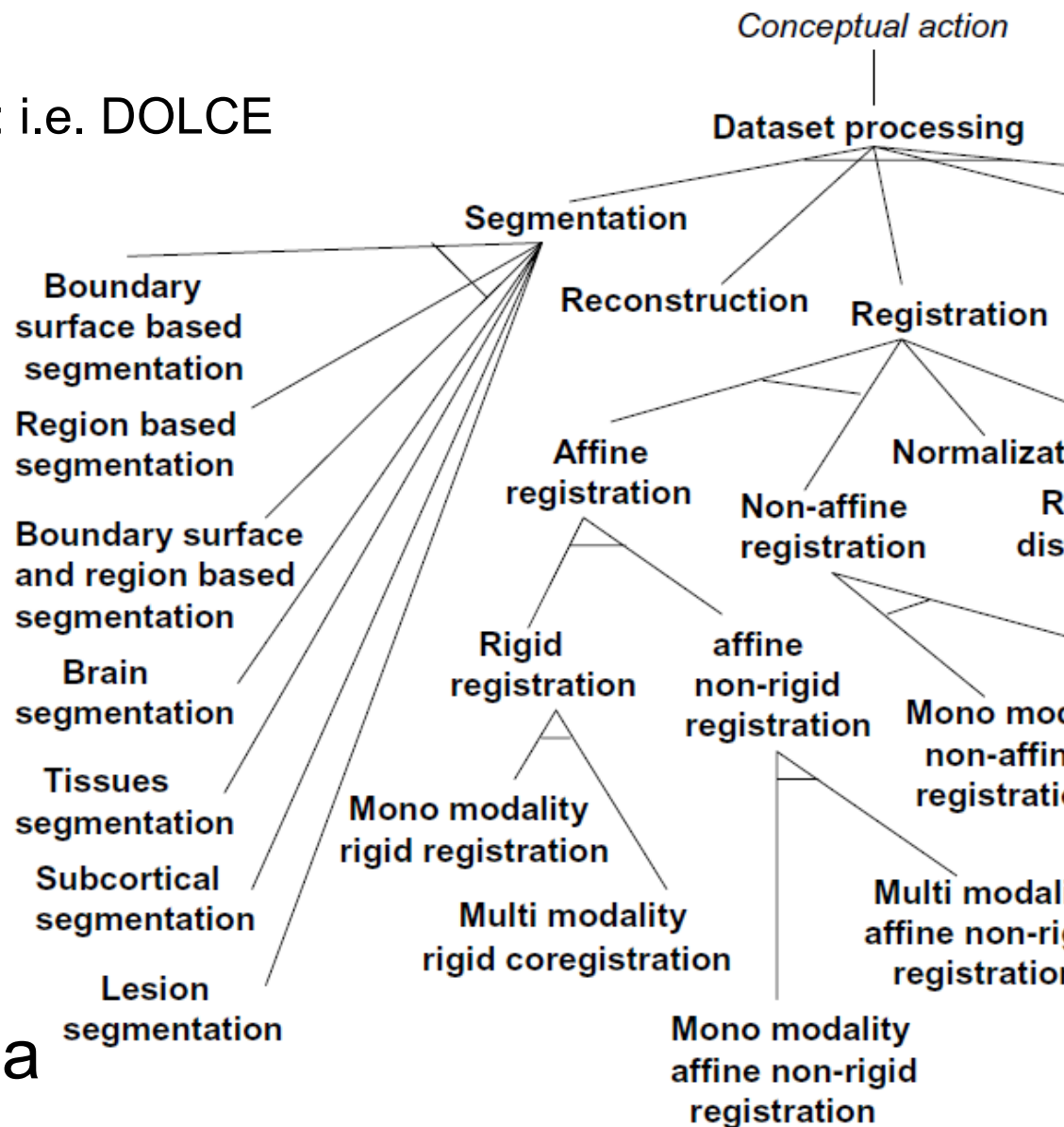
Semantic reference

- Application ontology
 - 3-levels structure
 - one Foundational ontology: i.e. DOLCE
 - Several Core ontologies
 - Several Domain ontologies

- Covering
 - Data sets
 - Data processing tools
 - Scientific measure
 - Medical context
 - Data provenance
 - ...

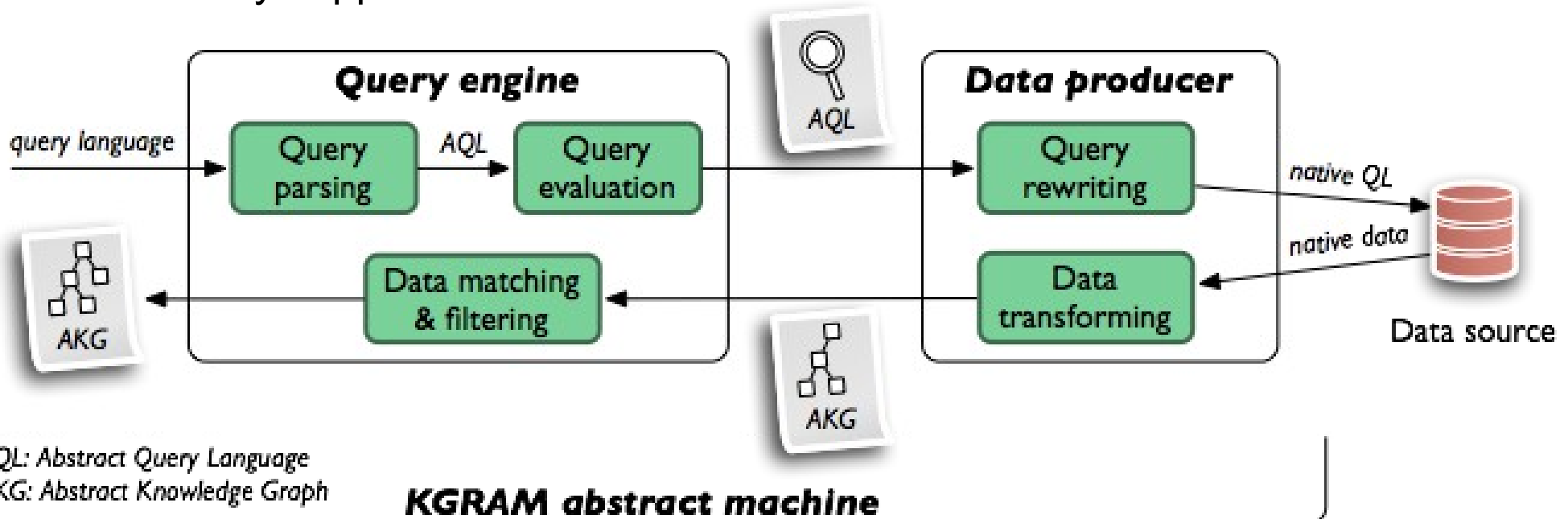
- Domain-specific rules
 - Inference abilities

- Derived relational schema



From relational to semantic medical stores

- Databases federation
 - Requires relational model mediation
 - Based on a semantic reference to derive the federated schema
- Based on KGRAM (Knowledge Graph Abstract Machine)
 - Semantic query engine enabling:
 - Heterogeneous data models
 - Fully supports SPARQL v1.1



Distributed Query Processing

- KGRAM query processing

```
Q SELECT ?name ?date
   WHERE { ?x foaf:name ?name . ?x dbpedia:birthDate ?date .
           FILTER (CONTAINS (?name, 'Bob')) }
```

- Asynchronous execution

Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution

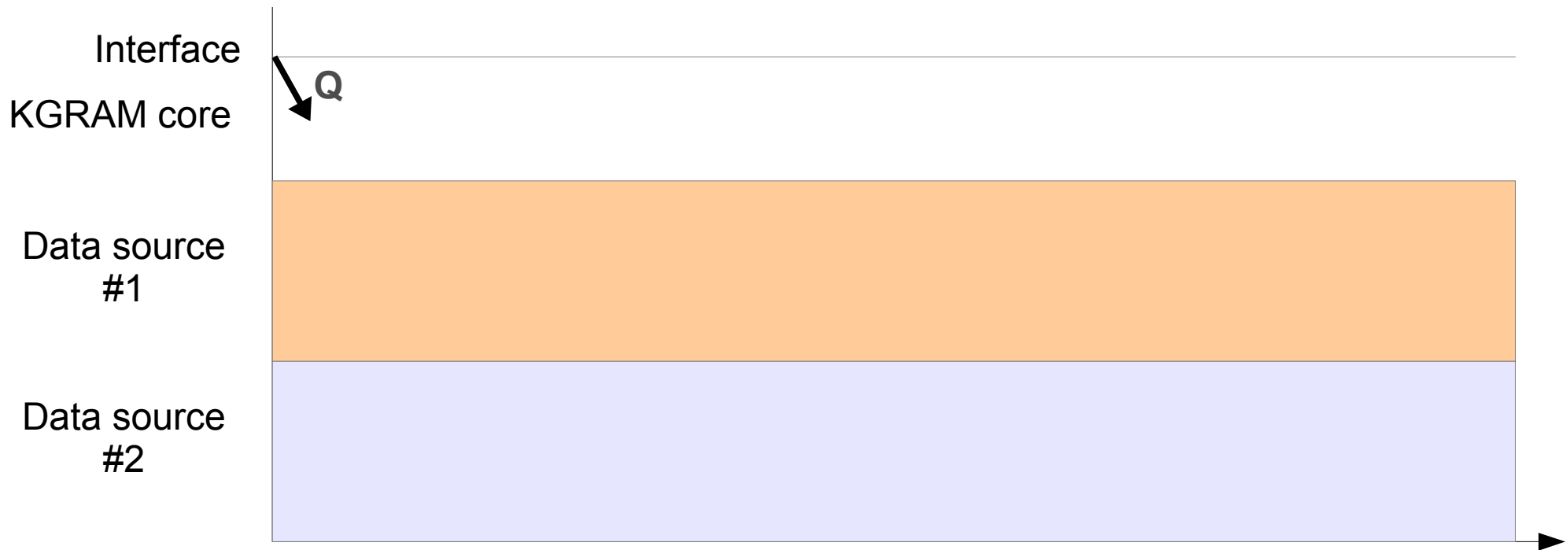
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



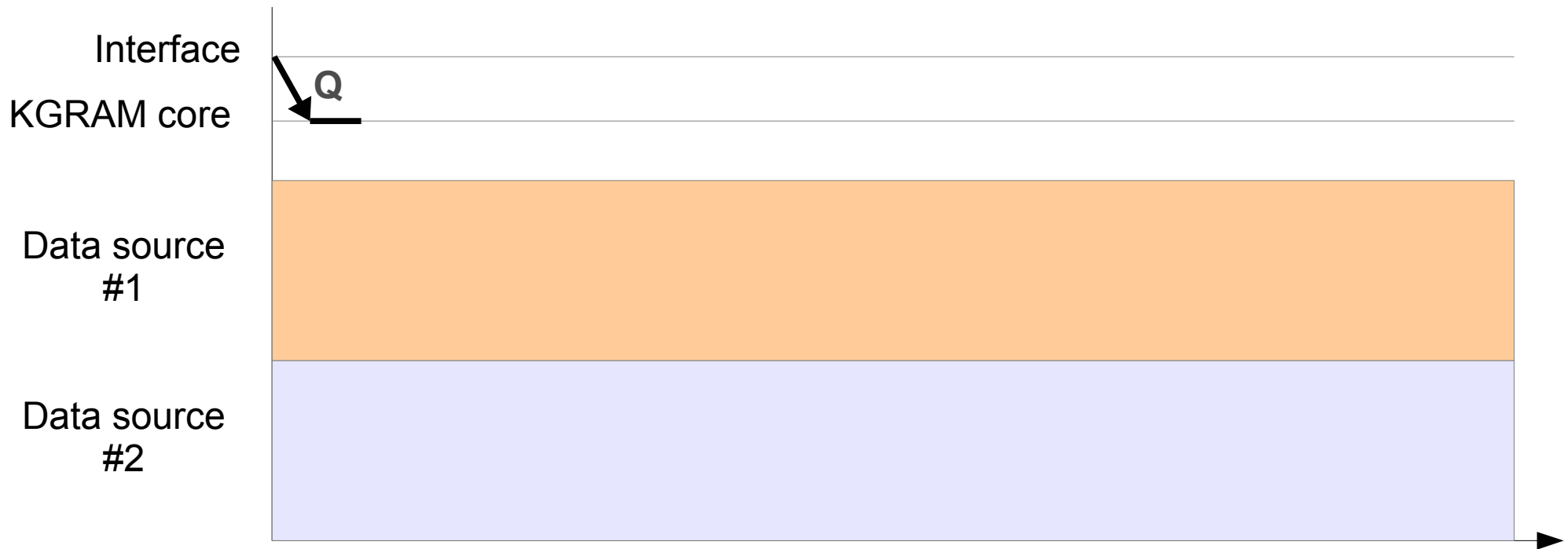
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



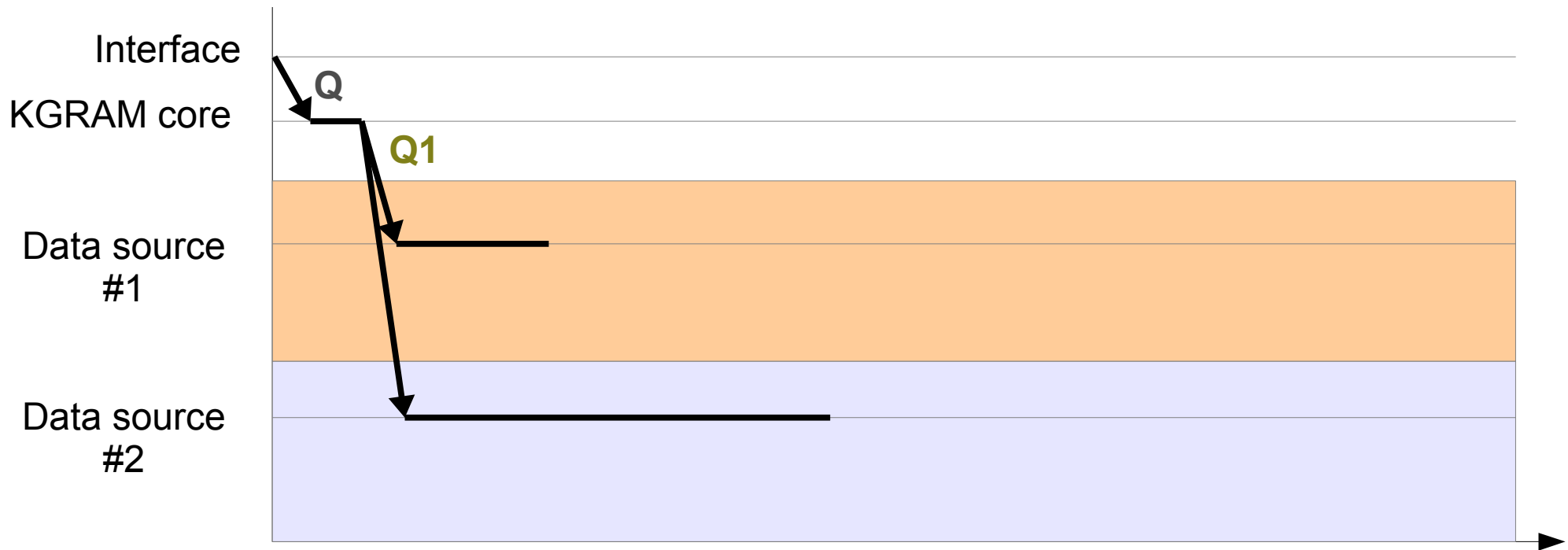
Distributed Query Processing

- KGRAM query processing

```

Q  SELECT ?name ?date
    WHERE {
        ?x foaf:name ?name .
        ?x dbpedia:birthDate ?date .
        Q1 FILTER (CONTAINS (?name, 'Bob')) }
        Q2
    
```

- Asynchronous execution



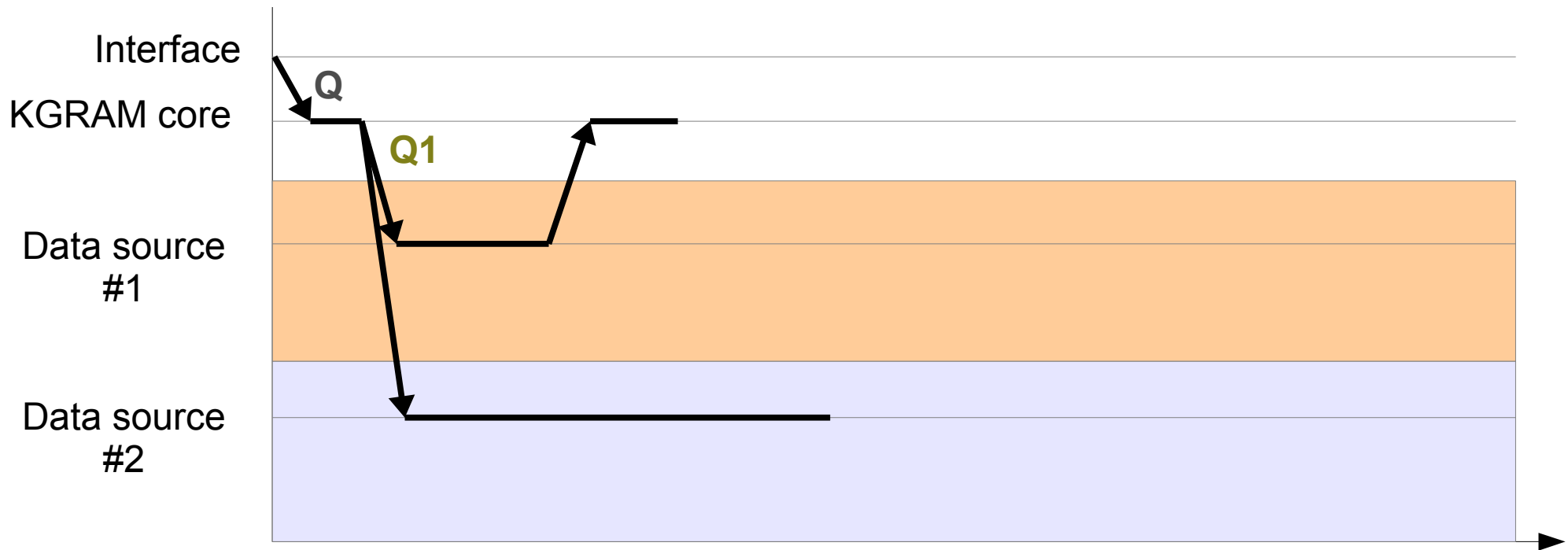
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



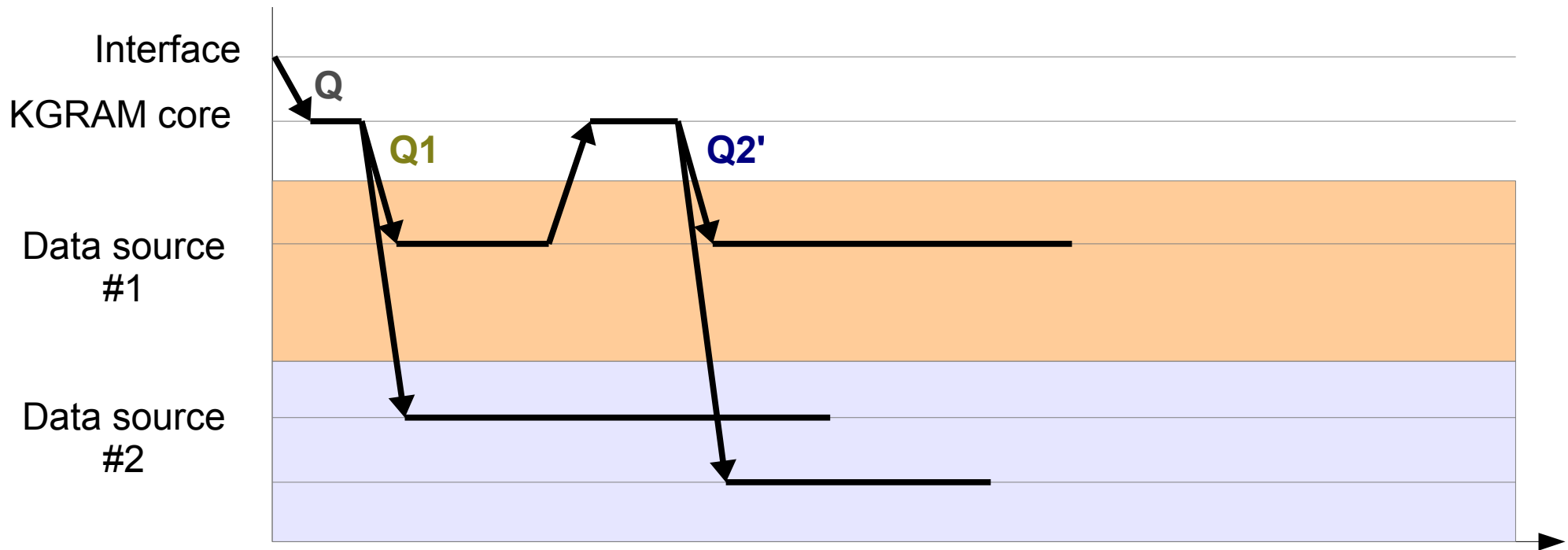
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



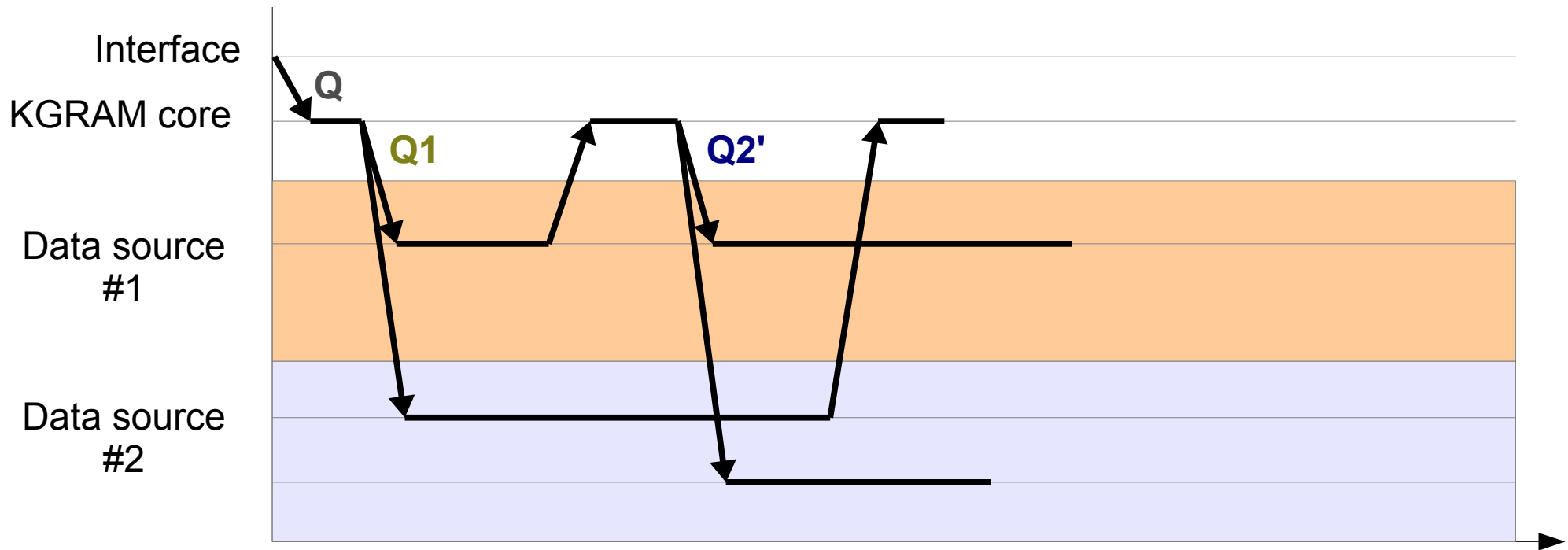
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



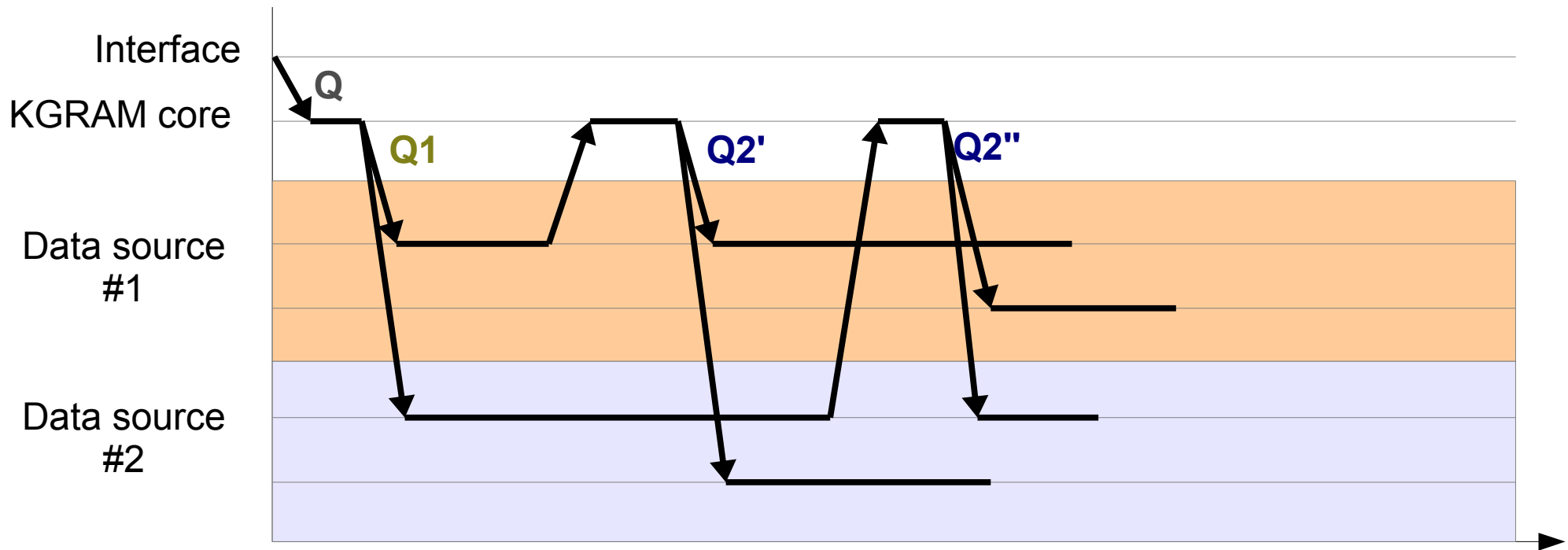
Distributed Query Processing

- KGRAM query processing

```

Q  SELECT ?name ?date
    WHERE {
        ?x foaf:name ?name .
        ?x dbpedia:birthDate ?date .
        Q1 FILTER (CONTAINS (?name, 'Bob')) }
        Q2
    
```

- Asynchronous execution



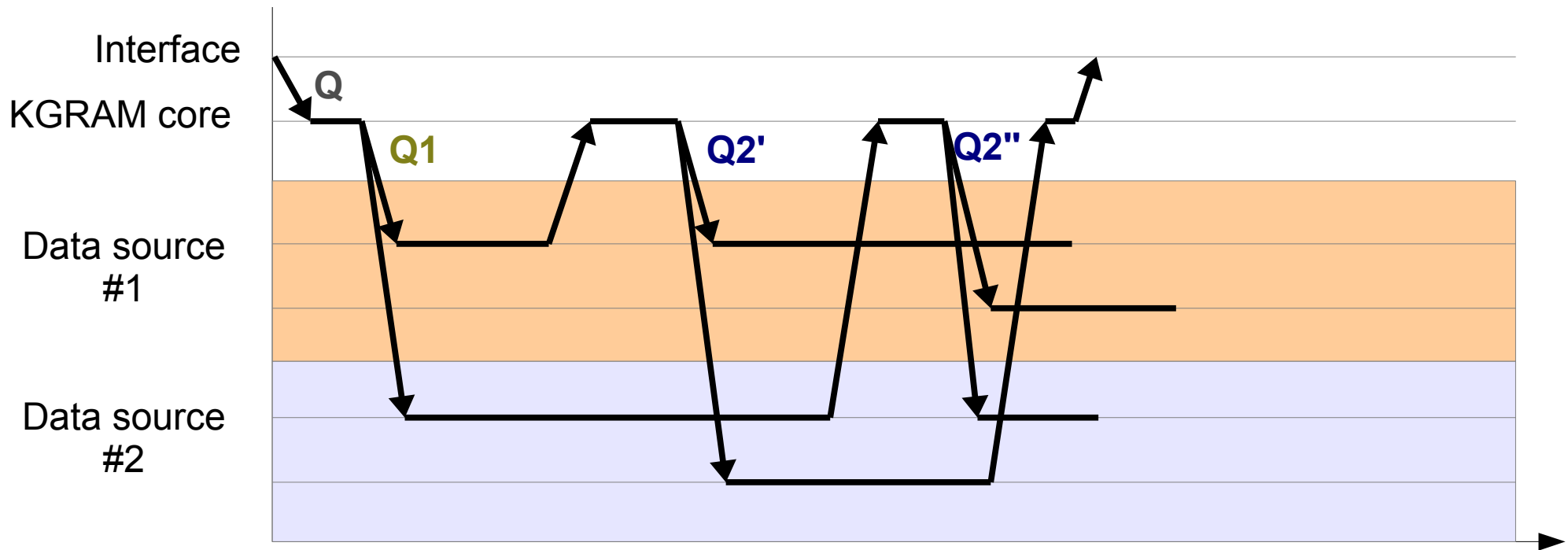
Distributed Query Processing

- KGRAM query processing

```

Q  SELECT ?name ?date
    WHERE {
        ?x foaf:name ?name .
        ?x dbpedia:birthDate ?date .
        Q1 FILTER (CONTAINS (?name, 'Bob')) }
        Q2
    
```

- Asynchronous execution



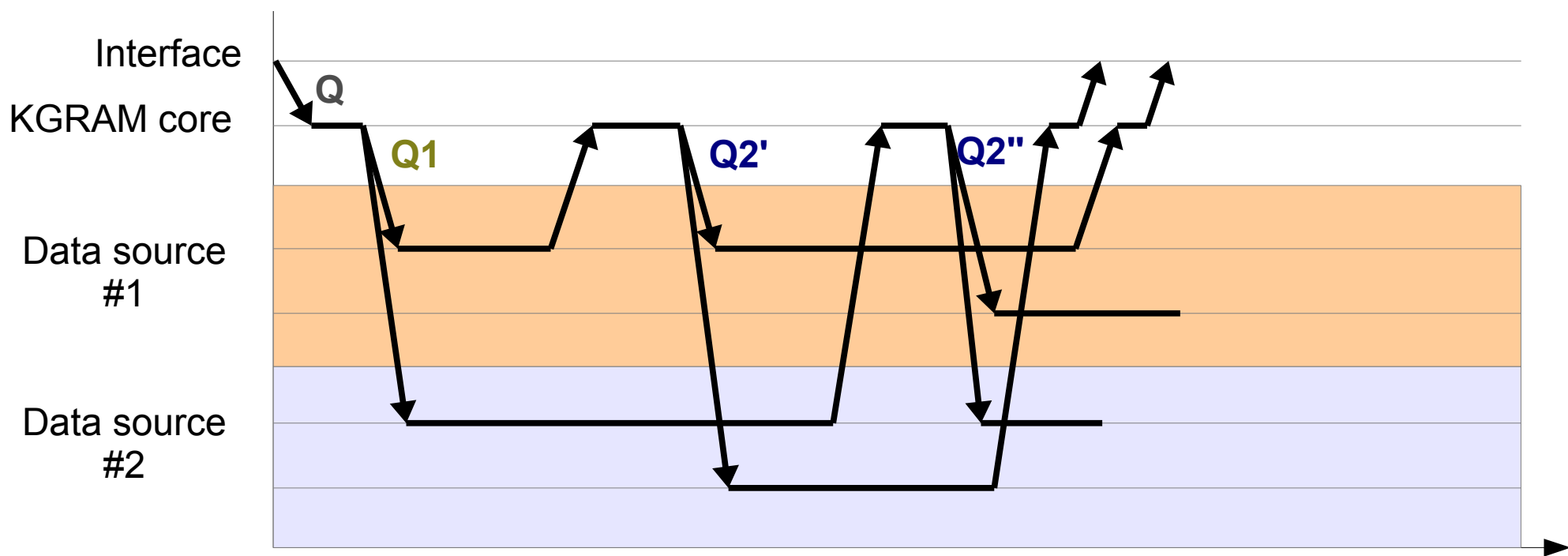
Distributed Query Processing

- KGRAM query processing

```

Q SELECT ?name ?date
  WHERE {
    ?x foaf:name ?name .
    ?x dbpedia:birthDate ?date .
    Q1 FILTER (CONTAINS (?name, 'Bob')) }
    Q2
  
```

- Asynchronous execution



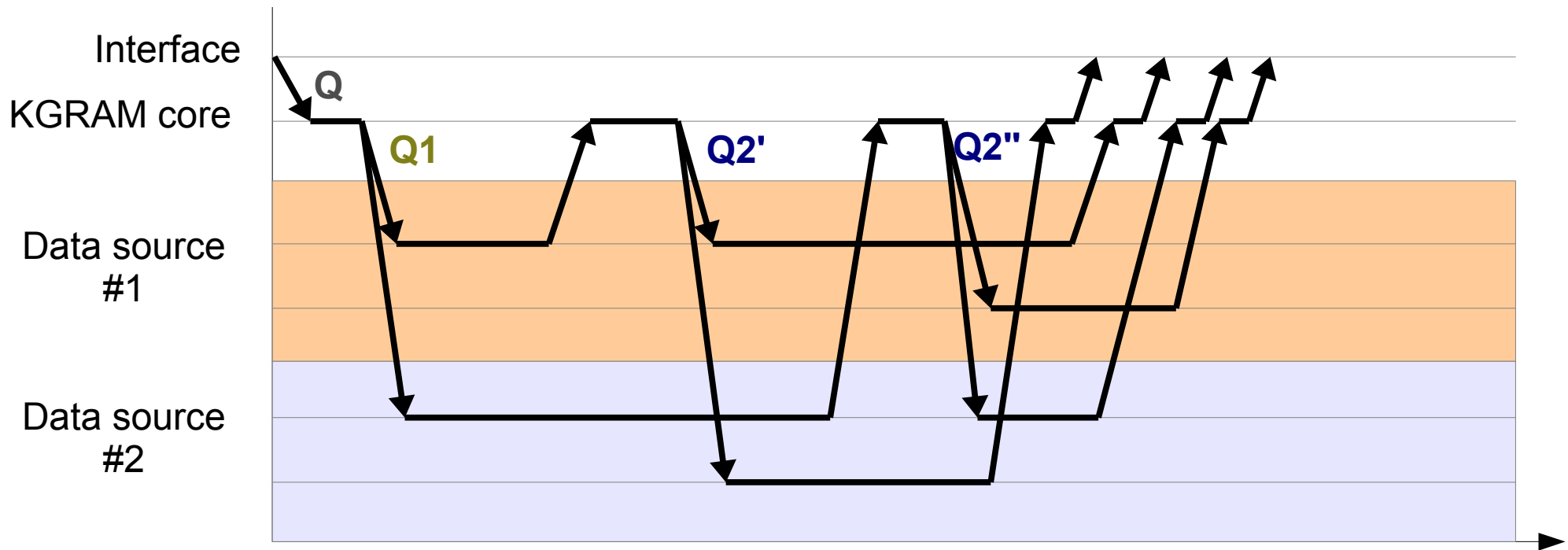
Distributed Query Processing

- KGRAM query processing

```

Q  SELECT ?name ?date
    WHERE {
      ?x foaf:name ?name .
      ?x dbpedia:birthDate ?date .
      Q1 FILTER (CONTAINS (?name, 'Bob')) }
      Q2
    
```

- Asynchronous execution



Workshop schedule

- Wednesday 8th (afternoon)
 - 14:30 Session on biomedical data federation in practice
(Conveners: B. Gibaud, J. Montagnat)
 - 19:30 Diner at hotel Omega
- Thursday 9th
 - 9:30 Session on Data mediation (Conveners: F. Michel, J. Montagnat)
 - 13:00 Session on Data federation (Conveners: A. Gaignard, O. Corby)
 - 15:30 Session on Knowledge graphs and reasoning
(Conveners: C. Faron Zucker, O. Corby)
 - 19:30 Cocktail at hotel Omega
- Friday 10th (morning)
 - 9:00 Session on Biomedical ontologies

Workshop schedule

- **Wednesday 8th (afternoon)**
 - 14:30 Session on biomedical data federation in practice
(Conveners: B. Gibaud, J. Montagnat)
 - 19:30 Diner at hotel Omega
- **Thursday 9th**
 - 9:00 Session on Knowledge graphs and reasoning
(Conveners: C. Faron Zucker, O. Corby)
 - 13:00 Session on Data federation (Conveners: A. Gaignard, O.corby)
 - 15:30 Session on Data mediation (Conveners: F. Michel, J. Montagnat)
 - 19:30 Cocktail at hotel Omega
- **Friday 10th (morning)**
 - 9:00 Session on Biomedical ontologies

Let us get started

- Session 1: Biomedical data federation in practice: feedback on existing approaches (14:30 – 17:30)
 - [Camille Maumet](#) (University of Warwick, UK): Supporting image-based meta-analysis with NIDM: Standardised reporting of neuroimaging results
 - [Silvia Olabbarriaga](#) (Amsterdam Medical Center, NL): Challenges for services integration into science gateways: the local story of the AMC
 - [Pascal Neveu](#) (INRA Montpellier, FR): From Gene To The Bottle
 - [Frédric Segond](#) (Viseo, FR): Language and knowledge technologies to properly model textual medical data and support better reasoning
 - [Panel discussion](#). Moderators: [Bernard Gibaud](#), [Johan Montagnat](#)

Session 2: Reasoning on Graphs

- [Harald Sack](#) (U. Potsdam, DE), The Journey is the Reward - Exploratory Semantic Search based on Linked Data
- [Martin Peters](#) (U. Dortmund, DE), Rule-based reasoning using GPUs
- [Guy Mélançon](#) (U. Bordeaux), Graphs visualisation
- [Christian de Sainte Marie](#) (IBM)
- Panel discussion. Moderators: [Catherine Faron Zucker](#), [Olivier Corby](#)

Session 3: Data Federation

- [Ruben Verborgh](#) (Ghent University): Querying data on the Web – client or server?
- [Ester Pacitti](#) (U. Montpellier 2 / INRIA, FR): Profile Diversity for Query Processing using Users Recommendations
- [Axel-Cyrille Ngonga Ngomo](#) (U. Leipzig): Hibiscus project, sources selections
- [Panel discussion](#). Moderators: [Alban Gaignard](#), [Olivier Corby](#)

Session 4: Data Mediation

- [Pascal Molli](#) (U. Nantes, FR): Towards Writable and Scalable Linked Open Data
- [Freddy Priyatna](#) (U. Polytechnic Madrid, SP): An Overview of the Research Carried Out at Data Integration Group - OEG
- [Franck Michel](#) (CNRS, FR): An R2RML extension for the translation of non-relational databases into RDF
- [Panel discussion](#). Moderators: [Franck Michel](#), [Johan Montagnat](#)

Session 5: Biomedical ontologies

- [Simon Jupp](#) (EBI, UK): BiomedBridges project
- [Heiner Oberkamp](#) (Siemens, DE): Expressing Medical Image Measurements using Open Biological and Biomedical Ontologies
- [Alan Ruttenberg](#) (University at Buffalo, School of Dental Medicine, USA): What should the role of biomedical ontology in imaging be?
- [Bernard Gibaud](#) (INSERM, FR): ontology modelling needs for image biomarkers description
- [Panel discussion](#). Moderators: [Gilles Kassel](#), [Bernard Gibaud](#)