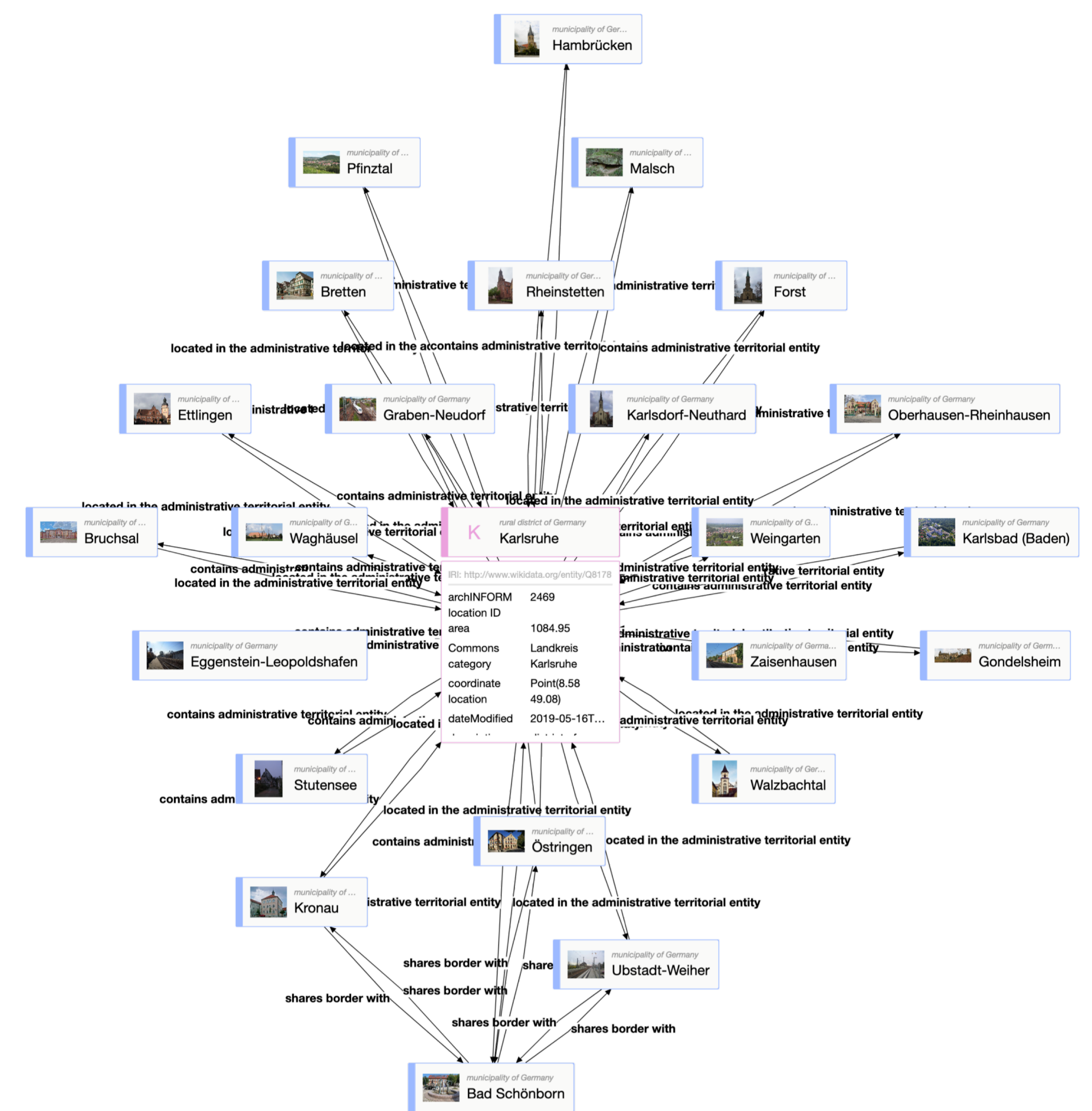


Enriching Knowledge Graphs with Geo Data from OpenStreetMap and Wikidata

Wolfgang Schell Andreas Schwarte Johannes Trame Peter Haase

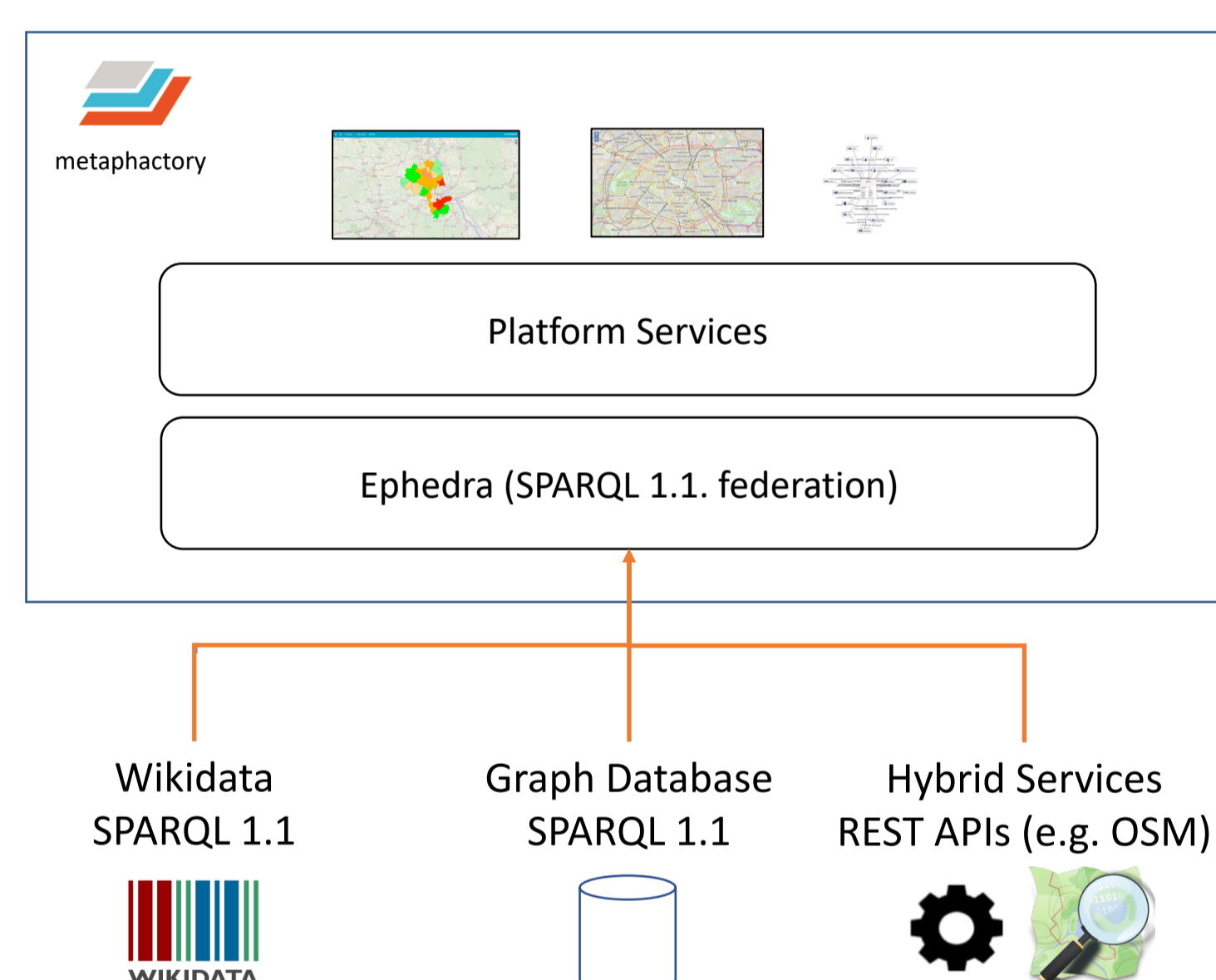
Knowledge Graphs: a modern data model for integrated and heterogeneous data

- A **Knowledge Graph** contains entities with properties and relations to other entities. Typically it integrates information from multiple data sources. This is often used for use cases around social networks, IT infrastructure, financial transactions, Internet of Things, life science, etc.
- The Resource Description Format (**RDF**) is used for data modelling and **SPARQL** (a W3C **standardized query language**) can be used for data retrieval. **Federation** (accessing multiple graphs or repositories at the same time) allows joining independent data sources in a single query
- **Graph analytics** such as clustering, proximity or centrality analysis, shortest path, etc. give powerful and actionable insights into your graph data
- A knowledge graph can be enriched by or used as a base for information mining using **machine learning**
- There are public knowledge graphs such as **Wikidata** and private ones, often called **Enterprise Knowledge Graphs**



Example Knowledge Graph of the rural district of Karlsruhe with data from Wikidata visualized using metaphactory

Geo-enriched Knowledge Graphs: bring location to your data




```
SELECT (?city as ?link) ?wkt ?color (round(?density) as ?description) WHERE {
  ?city wdt:P131+ wd:Q8178 .
  ?city wdt:P402 ?osmCode .
  OPTIONAL { ?city wdt:P2046 ?area }
  OPTIONAL { ?city wdt:P1082 ?population }

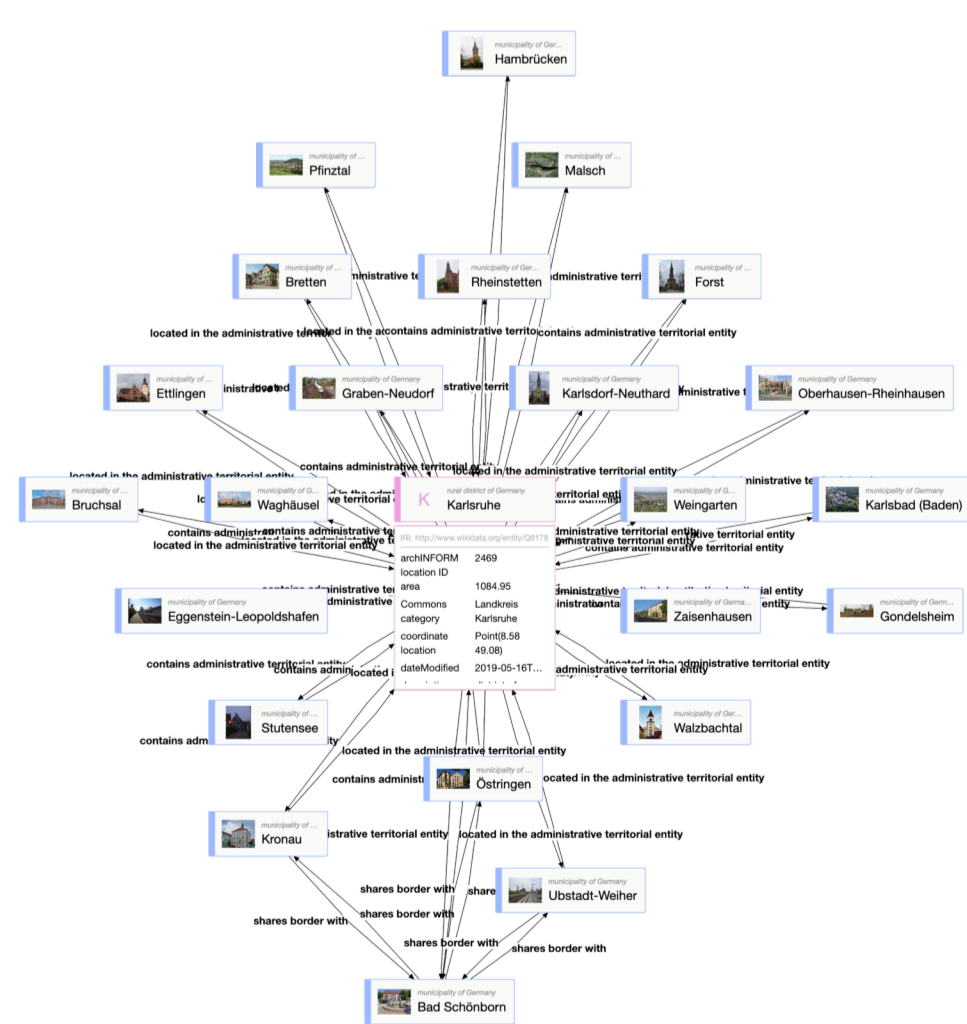
  SERVICE mp-federation:openstreetmap-lookup {
    ?results osm:osmId ?osmCode . ?results rdfs:label ?label .
    ?results osm:polygon ?wkt . ?results osm:wikidataID ?wikidataID .
  }

  BIND(?population/?area as ?density)
  BIND(IF(?density<300, "rgba(0,255,0,0.5)", IF(?density<600, "rgba(255,165,0,0.5)", "rgba(255,0,0,0.5)")) as ?color)
}
```

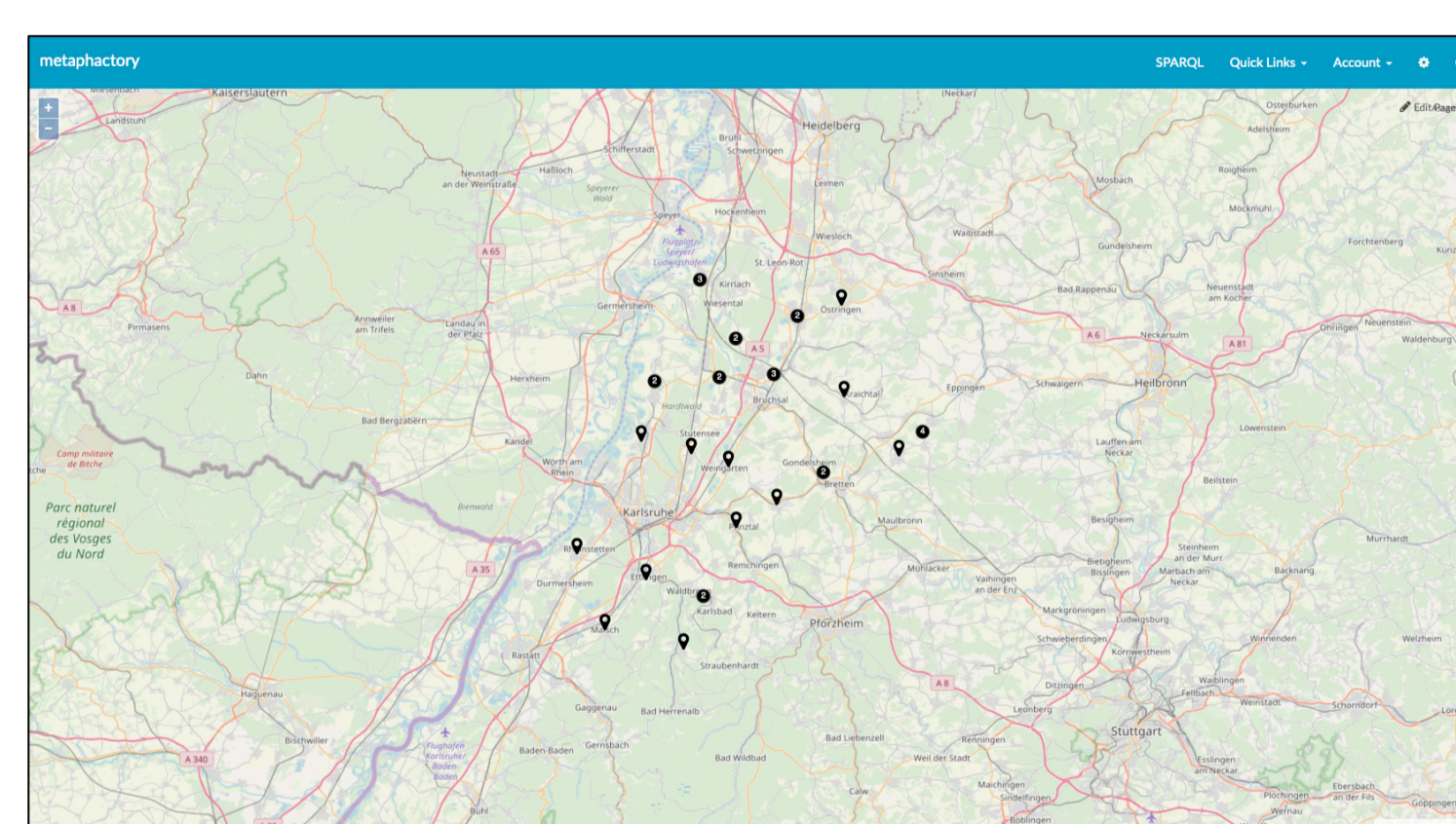
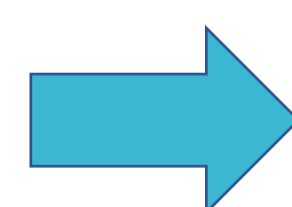
Technical details (based on *metaphactory* platform):

- **Integration of heterogeneous data** using *Ephedra* federation services
 - Structured graph data available as RDF (in this case: RDF data from Wikidata)
 - access OpenStreetMap through REST API for geocoding and to fetch shape information for areas
 - Semi- or unstructured data can be integrated as well using additional adapters or importers
- **Visualization of data** is possible using a rich set of UI components (e.g. map or graph) in the *metaphactory* platform
- The integrated use case can be packaged and **published as app**, e.g. for the Internet or as an internal application. This allows to **make the app available** to a wider audience and/or distribute it
- The **Geiser research project** evaluates concrete use cases to build intelligent location-based services 

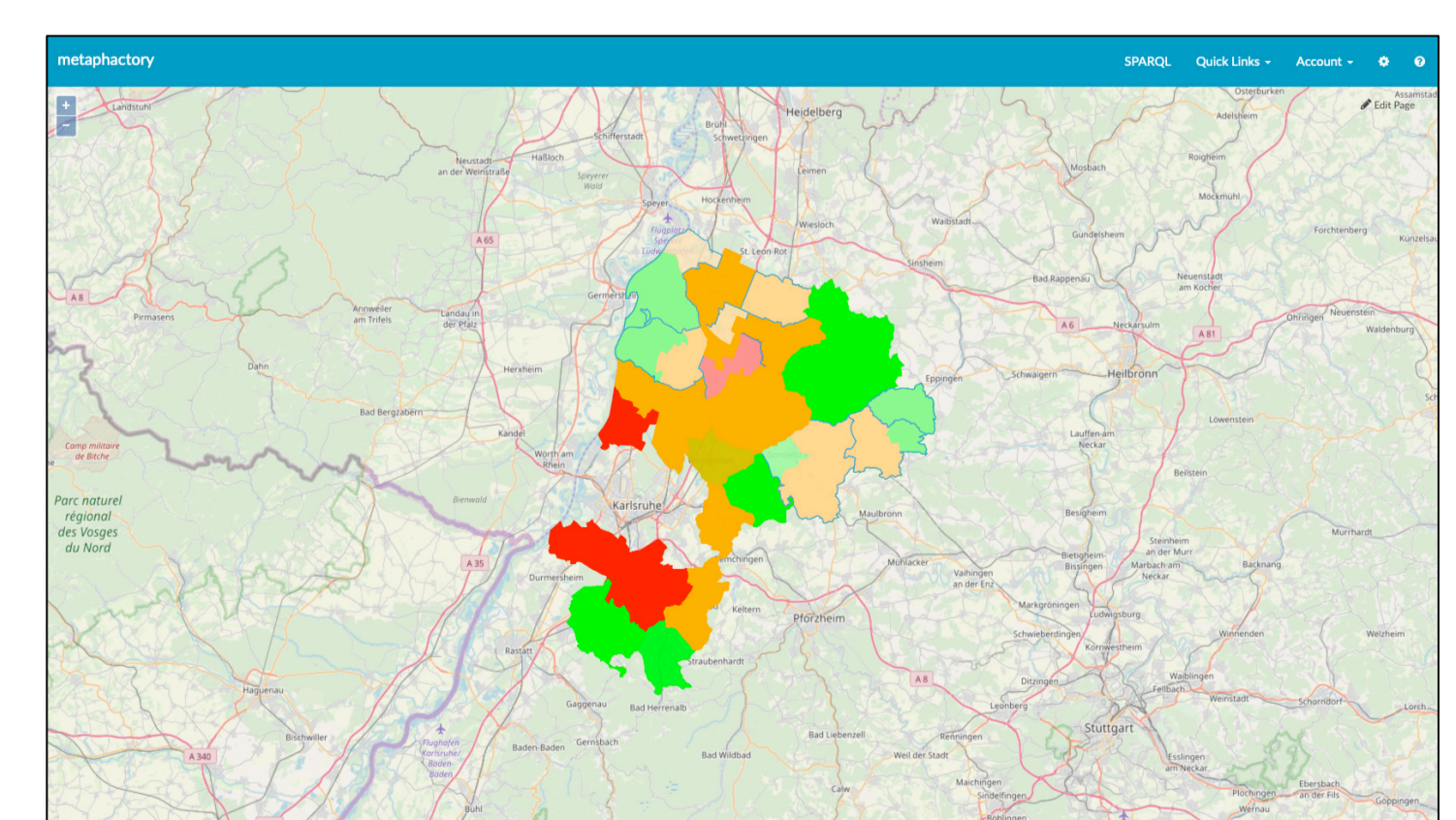
Example use case: Visualizing the density of cities in the district of Karlsruhe



Explore the Wikidata knowledge graph and formulate the SPARQL query: identify locations in the rural district of Karlsruhe



Visualize the locations on a map in the *metaphactory* platform using geo coordinates from Wikidata



Enrich the visualization with geo information from OpenStreetMap: retrieve WKT polygon data and area data for cities to show the color-coded population density