Golden Agents: Creative industries and the making of the Dutch Golden Age

A. Jan de Mooij, A. Can Kurtan, Jurian Baas, and Mehdi Dastani
Intelligent Systems, Department of Information and Computing Sciences, Utrecht, The Netherlands
{A.J.deMooij, A.C.Kurtan, J.Baas, M.M.Dastani}@uu.nl

Utrecht University

Introduction

- Querying Cultural Heritage Linked Open Data (LOD)
- Data sources are federated, distributed, and separately maintained
  - Too volatile to all join together
  - Potential intellectual property
- Overlapping topics in data
- Cross-domain research questions:
  The whole is more than the sum of the parts

Problem Statement

Querying as a novice requires knowledge of:

1. The described domain
2. The schema of the specific data
3. The query language

Querying multiple heterogeneous databases as one is even harder because:

- Schemes differ from database to database
- No uniform semantics
- Potential (and likely) inconsistencies

Proposed Solution

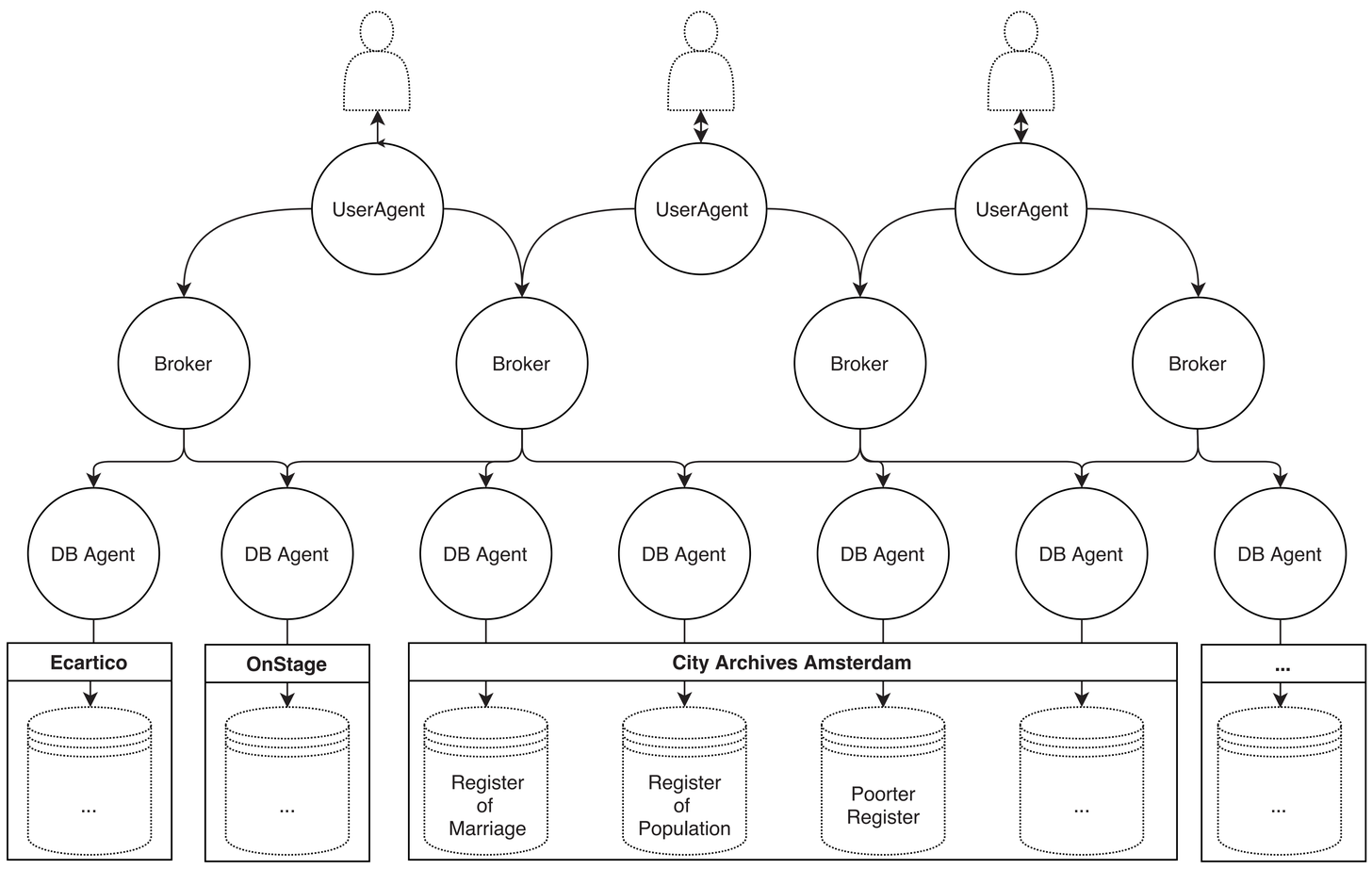
Multiagent System for querying distributed data sets, where autonomous intelligent agents:

1. act on behalf of – and in the interest of – users
2. help find users what they are looking for
3. protect the interests of the data providers

In addition, search interfaces which assist the user in:

1. exploring (new) data
2. posing the right questions
3. translating these questions to a system language (e.g. SPARQL)
4. interpreting results
5. downloading results for further processing

Multiagent Architecture



User Agent

- Interface with the user
- Learns the user's preferences and interests
- Negotiates with other agents on user's behalf

Database Agent

- Encapsulates data source
- Translates data to common vocabulary
- Pads data with interesting, related information
- Removes private or obscured information dynamically

Broker Agent

- Develops expertise on general topics
- Receives a query and propagates it to knowledgeable database agents
- Aggregates results

Example Scenario

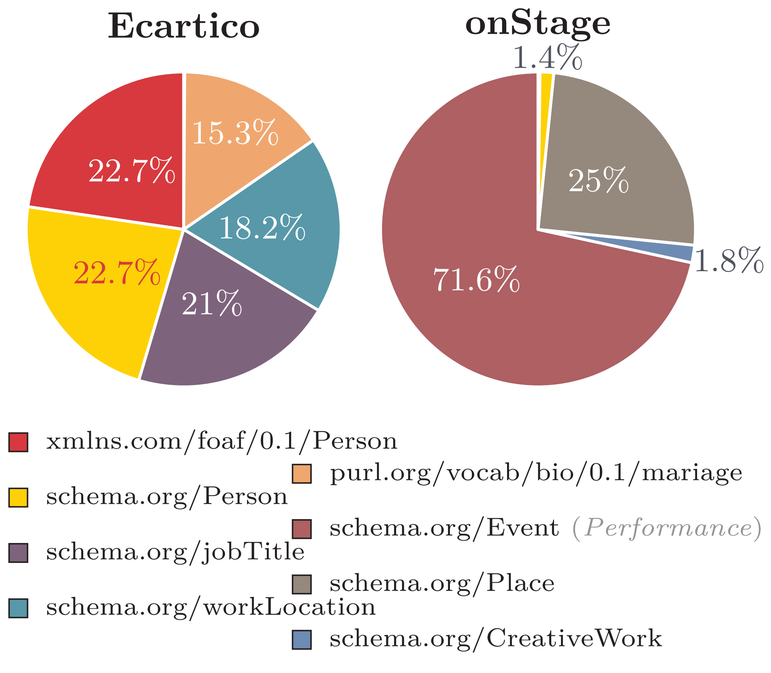
- **Question:** What was the age of an artist at the time of their first performance?
- User writes the whole query, which has nine `GA` ontology based statements in it
- One data source, *Ecartico*, only has data related to people
- The system generates a subquery with four person-related statements for Ecartico
- Another data source, *Onstage*, has data of performances and their authors
- The system generates a subquery with seven statements for Onstage
- It collects 2388 statements from Ecartico and 57923 statements from Onstage
- The broker uses a link set to match people with different ID's
- Finally, the broker finds the following nine different artists that match the query:

Name	Age
Jacob Cats	79
Gilbert de Flines	30
Pieter Cornelisz Hooft	57
Lescailje, Katharina	36
Geerardt Brandt	18
Samuel Coster	59
Joannes Antonides van der Goes	31
Vondel, Joost van den	51
Alewyn, Abraham	38

System Overview

- **Golden Agents (`GA`-)Ontology** defines a common taxonomy of the domain
- **Link sets** link entities between pairs of databases
- **Agents** simulate the inner workings of a knowledge institute: Answering outsider questions by delegating to divisions or individuals with the required expertise
- **Graphical User Interface** allows users to interact with the system by formulating queries

Prevalent Classes



References

[1] https://www.goldenagents.org/.

[2] Mehdi Dastani. 2apl: a practical agent programming language. *Autonomous agents and multi-agent systems*, 16(3):214–248, 2008.