Presentation of the semantic technologies for searching of information

Mariya Evtimova-Gardair

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Dr. Mariya Evtimova-Gardair

- University Paris 8, IUT Montreuil, Paris, France 2021

email:m.evtimovagardair@iut.univ-paris8.fr

- Avignon University, Avignon, France, 2020- 2021

- INRAE (French National Institute of Agricultural Research), France 2019- 2021

- Technical University- Sofia, Bulgaria, 2010- 2019

Research projects

- European Reform University Alliance
- Action research mental health
- Distributed knowledge graphs
- Des Données aux Connaissances en Agronomie et Biodiversité
- INSYLVA France infrastructure for management of forest resources
- Multi- agent system for searching of medical information in internet
- Large Scale Industrial Structural Optimisation for Advanced Applications
- The development of a simulation model of administrative services and procedures in order to develop e-services and e-management

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Bibliography

- Mariya Evtimova- Gardair, Evangelos Pallis. Intelligent agents system for medical information. Intelligent Wireless Communications, **Chapter 5**, IET, 2021
- Evtimova-Gardair M.,"Multi-agent searching system for medical information", WSEAS Transactions on information science and applications, E-ISSN: 2224-3402, vol.16,2019, pp.140-145
- Evtimova- Gardair M., Tashev T.,"Semantic system for searching of employees", International Journal of Economics and Management Systems, ISSN:2367-8925, Vol.3,2018, pp.132-136
- Evtimova M., "Big data healthcare system to improve healthcare information searching in the Internet", **Chapter 7**,e-ISBN:9781785612121, 2017, IET, UK, p.161-182

Web search engines





Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

The Problem???

Increase the quality of the returned results

✓ *short and vague requests*

✓ *increase the volume of the data*

- Semantic technologies
- Agent technologies
- Personalization



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Objective of the project

Development of a model of the system

- BIG DATA from vague and uncertain information
- Use technologies for personalization, semantic ontologies, intelligent agents, and fuzzy and probability theory



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Tasks

- Investigation of the problems in the reasoners
- To propose a model of a reasoner and ontology
- To integrate mobile agent for extraction of information from uncomplete and contradictory information environment from internet sources
- To create a model of user profile

Tasks

- To create new conceptual model of personalized semantic searched system
- To create a program- partial realization of separate modules
- Evaluation of the results

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Ontology technologies include:

- Reasoners
- Instruments for annotation
- Ontology based robots
- Instruments for extraction of information

Knowledge graphs and ontologies

KGs can represent complex structured information about entities, their properties and their relationships with other entities.

1. CovidGraph (https://covidgraph.org/)
2. KGTK COVID-19 KnowledgeGraph (Hope, Portenoy, Vasan et al., 2020)
3. Blender Lab COVID-KG (Wang, Li, Wang et al., 2020)
4. COVID-19 KnowledgeGraph (Wise et al., 2020)
5. COVID-KOP (Korn, Bobrowski, Li et al., 2020)

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

2 part. Conclusion

- User profile
- Using mobile agent
- A reasoner for Big Data , uncertain and vague query
- Ontology- fuzzy and probability logic for BIG DATA

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

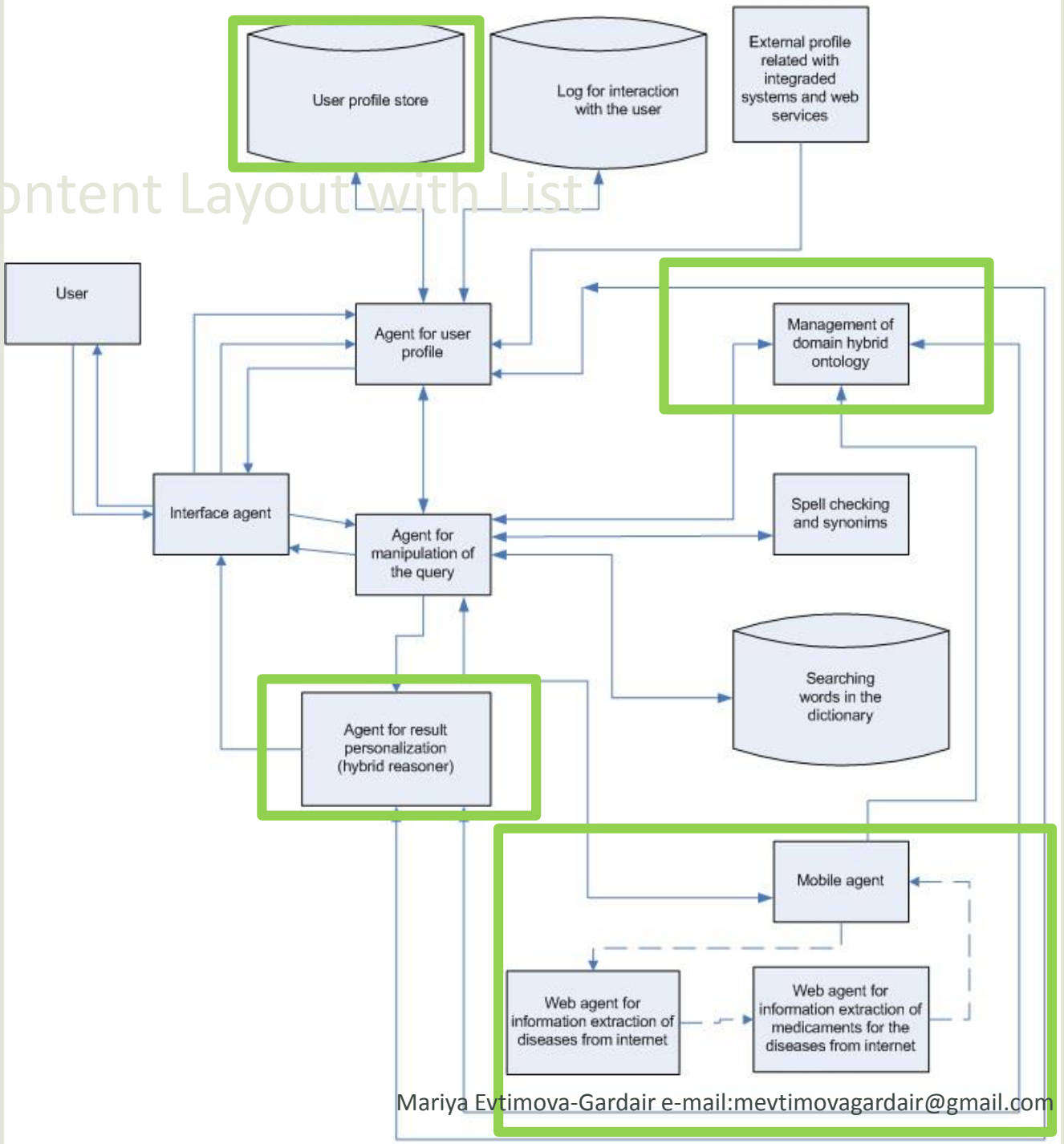
3 Part- Modeling of semantic search system

Semantic search system:

- big volume of data
- requested vague and uncertain information



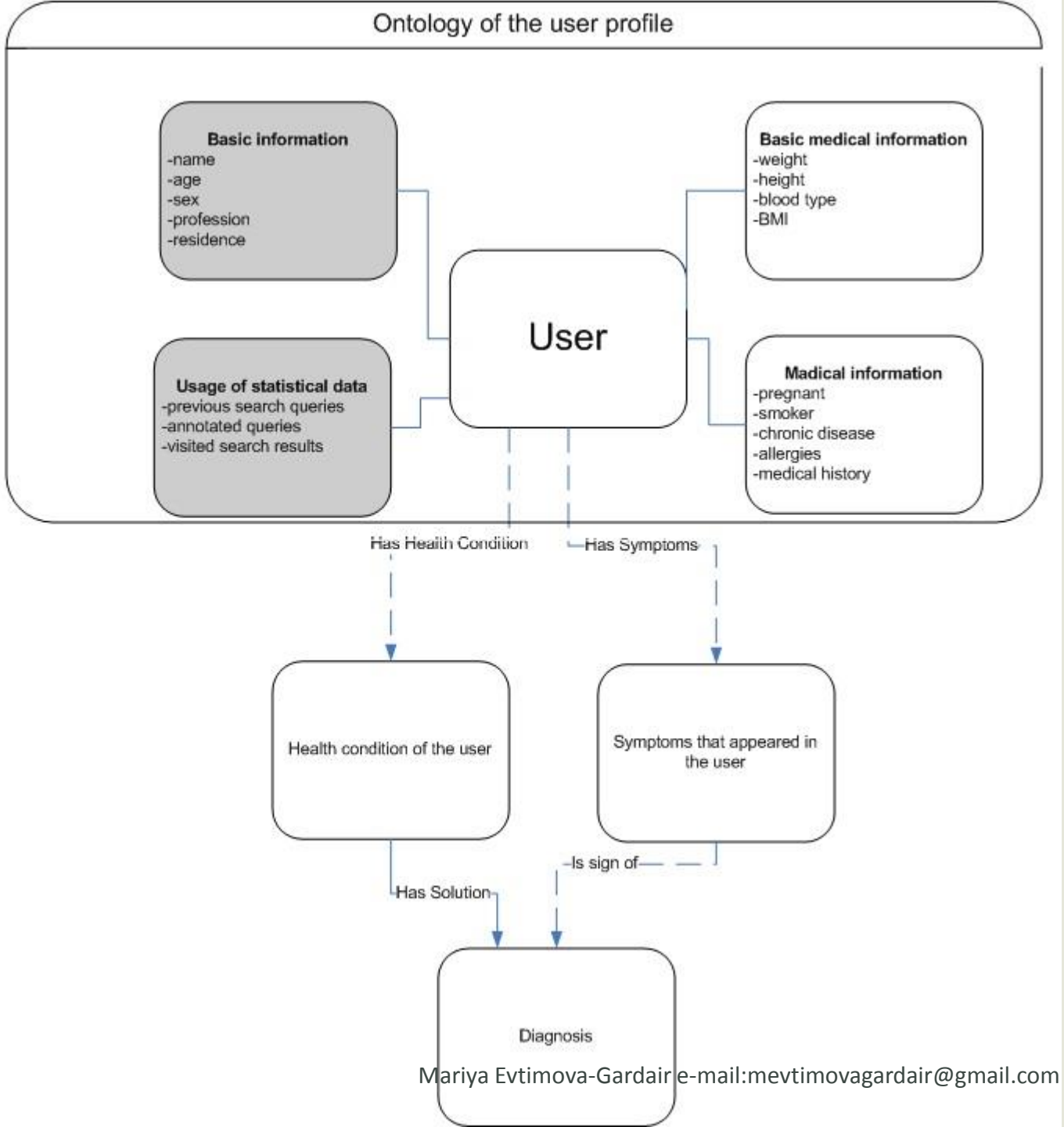
Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



Personalized semantic searched system



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



Personalized user profile



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Hybrid ontology

- Fuzzy logic
- Probability logic
- Case and rule based ontology



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

What is fuzzy logic?

- Many- valued logic
- Truth value [0,1]
- Fuzzy set theory
- Membership function

Mariya Evtimova-Gardair e-mail:mebtimovagardair@gmail.com

What is probability theory?

- Probability measure [0,1]

New model of reasoner

1. Algorithm for Big Data reasoning from vague and uncertain request

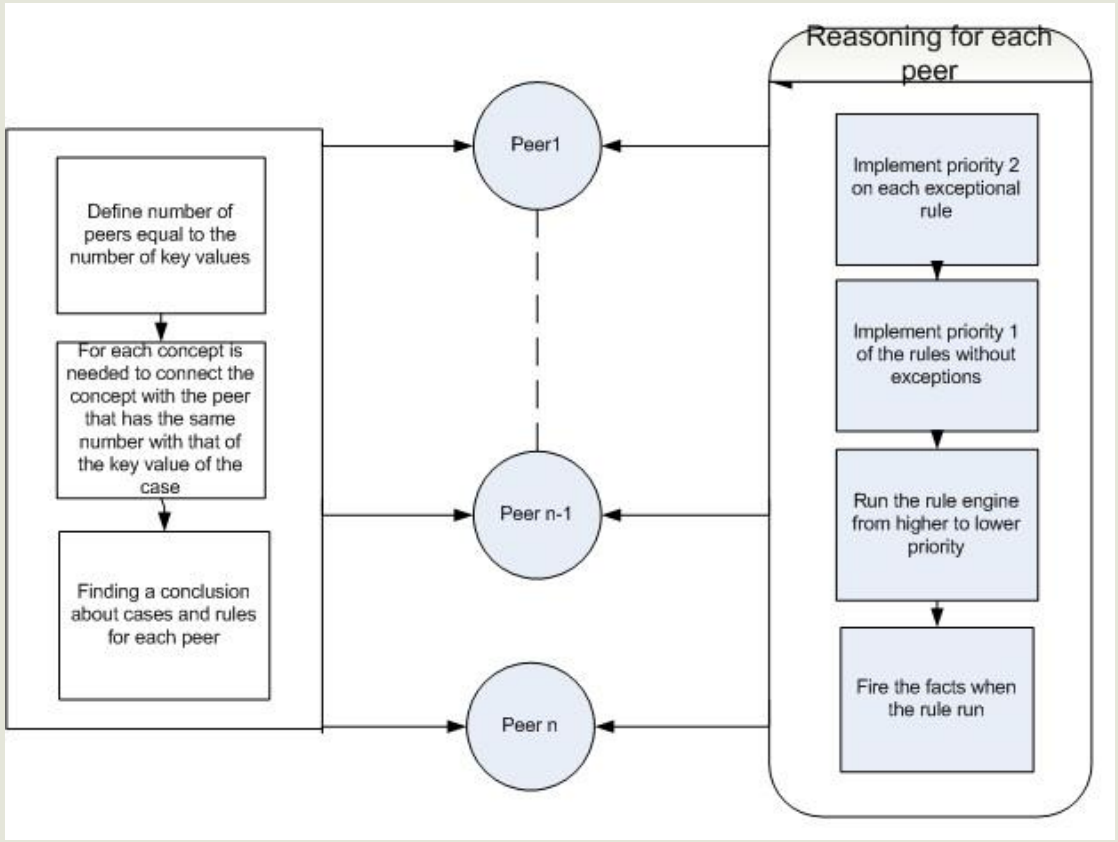
- Define the distribution of the data
- Separate the data into groups

2.Algorithm for reasoning for big data from standard ontology and measurement of semantic similarities



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

1.Algorithm



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

4 part- Partial realization of the system. Experimental evaluation of the results

- Description of the created prototype of the system

- Evaluation and the results of the system



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

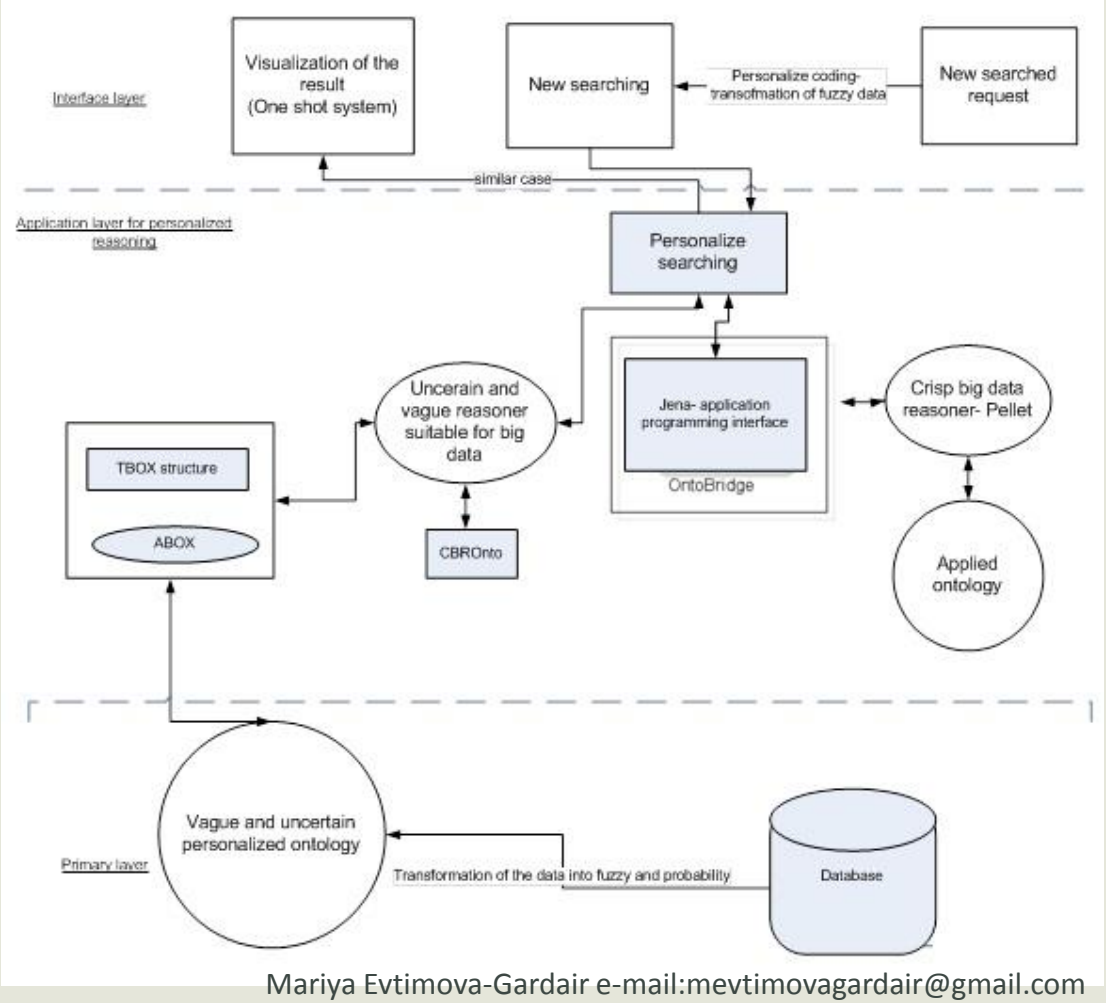
Usage of software

- jColibri- ALADIN
- JADE
- Protege



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

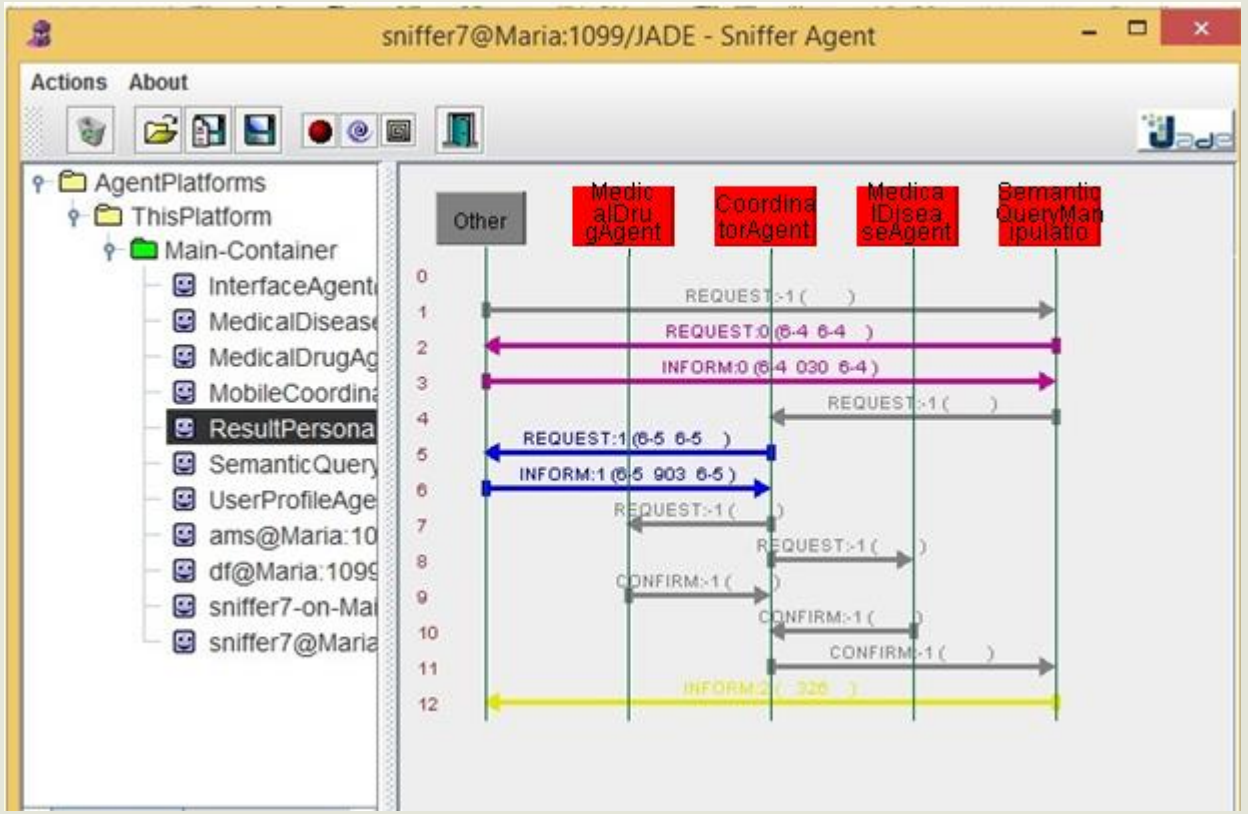
Architecture of the personalized searching system



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



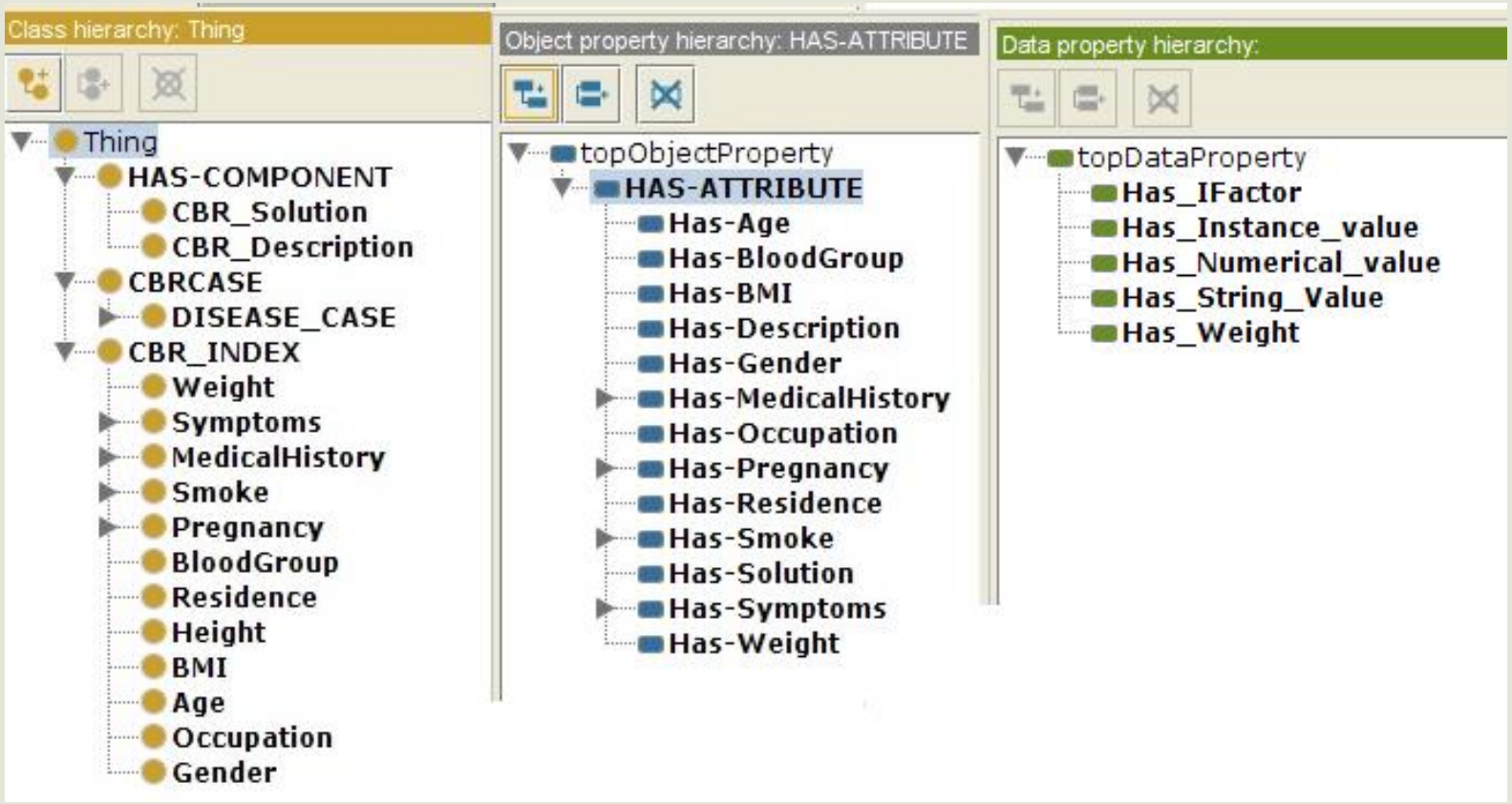
Functional diagram – with static agent



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



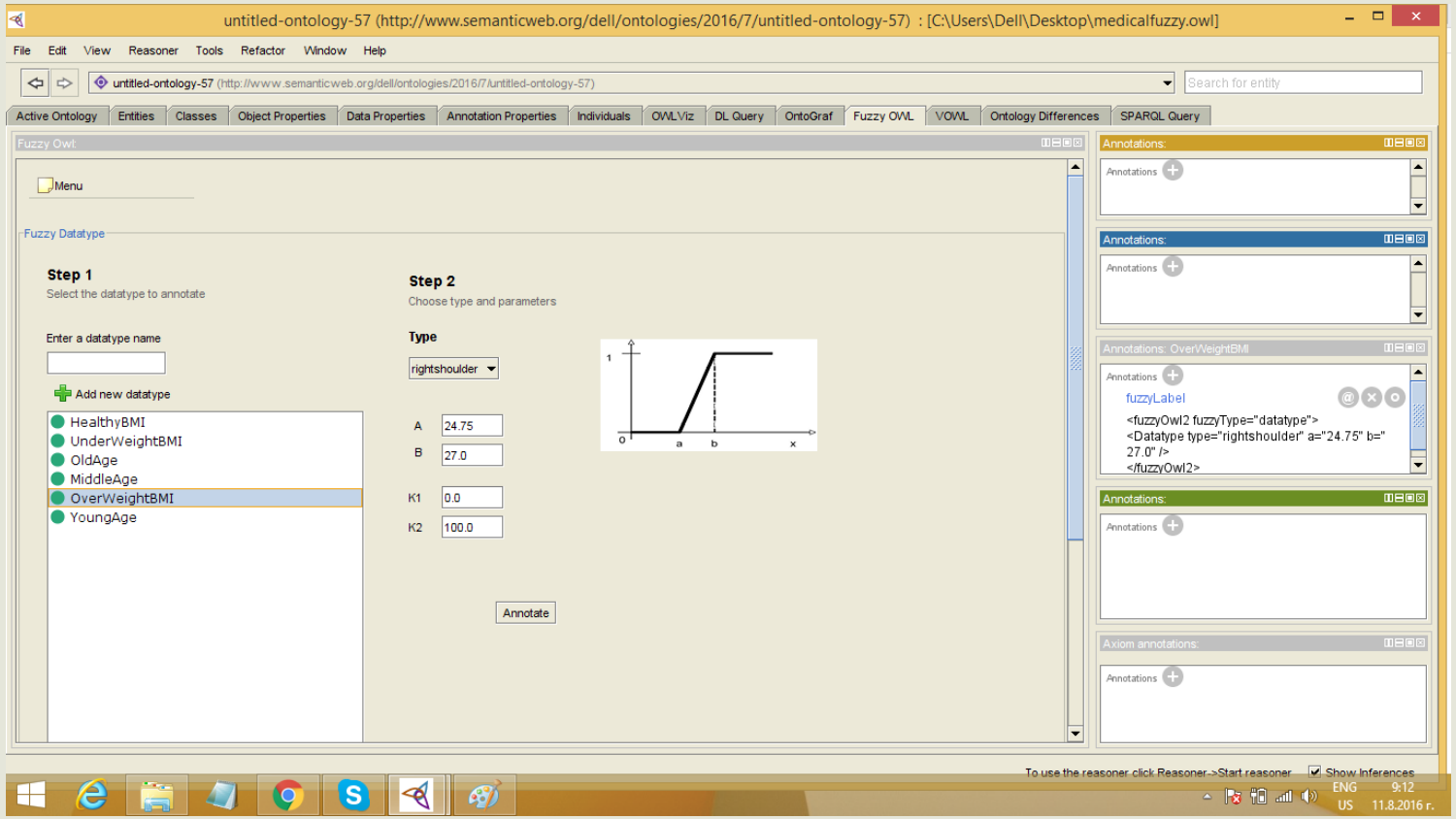
Creation of ontology





Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

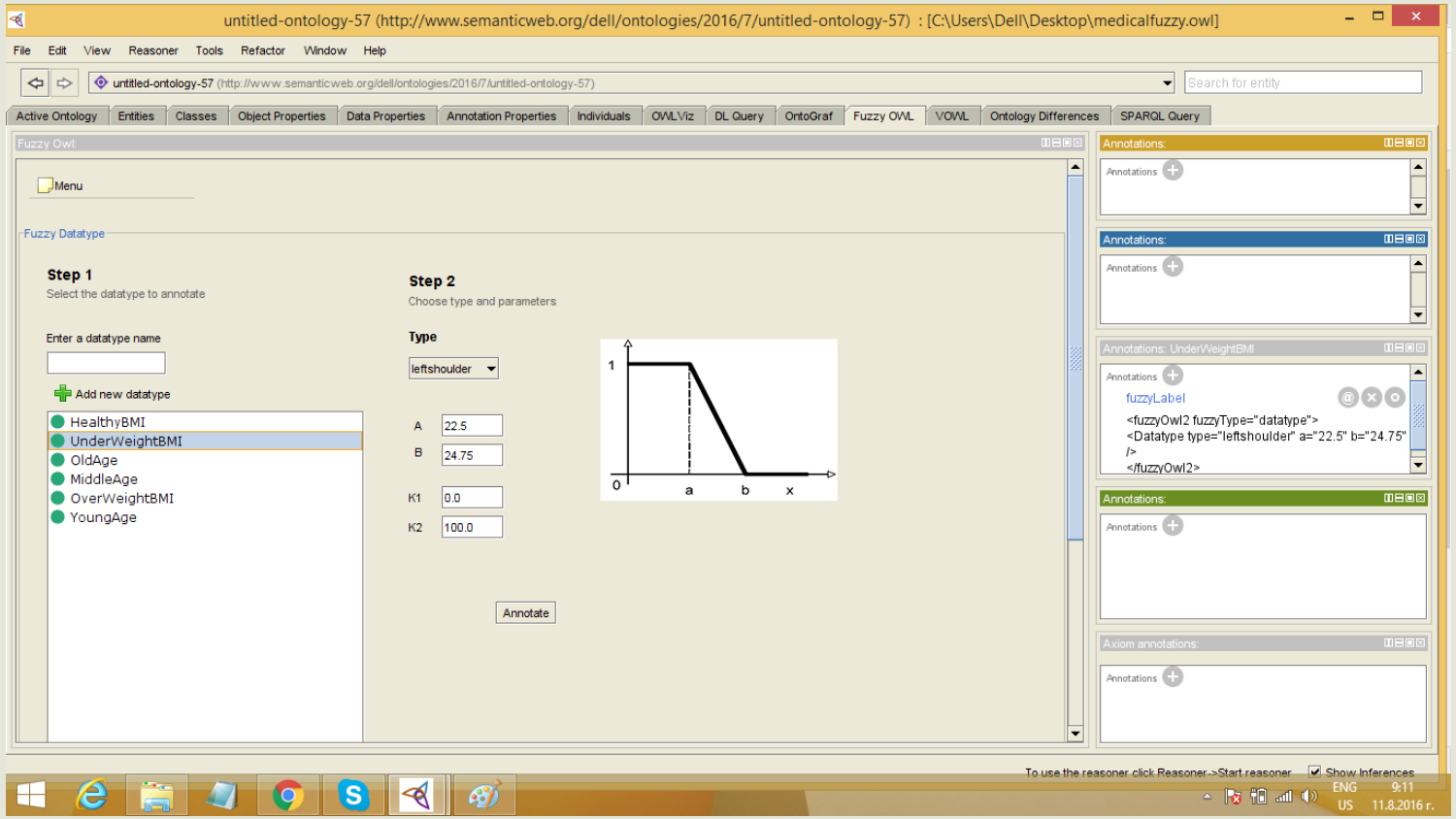
Fuzzy function of BMI



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



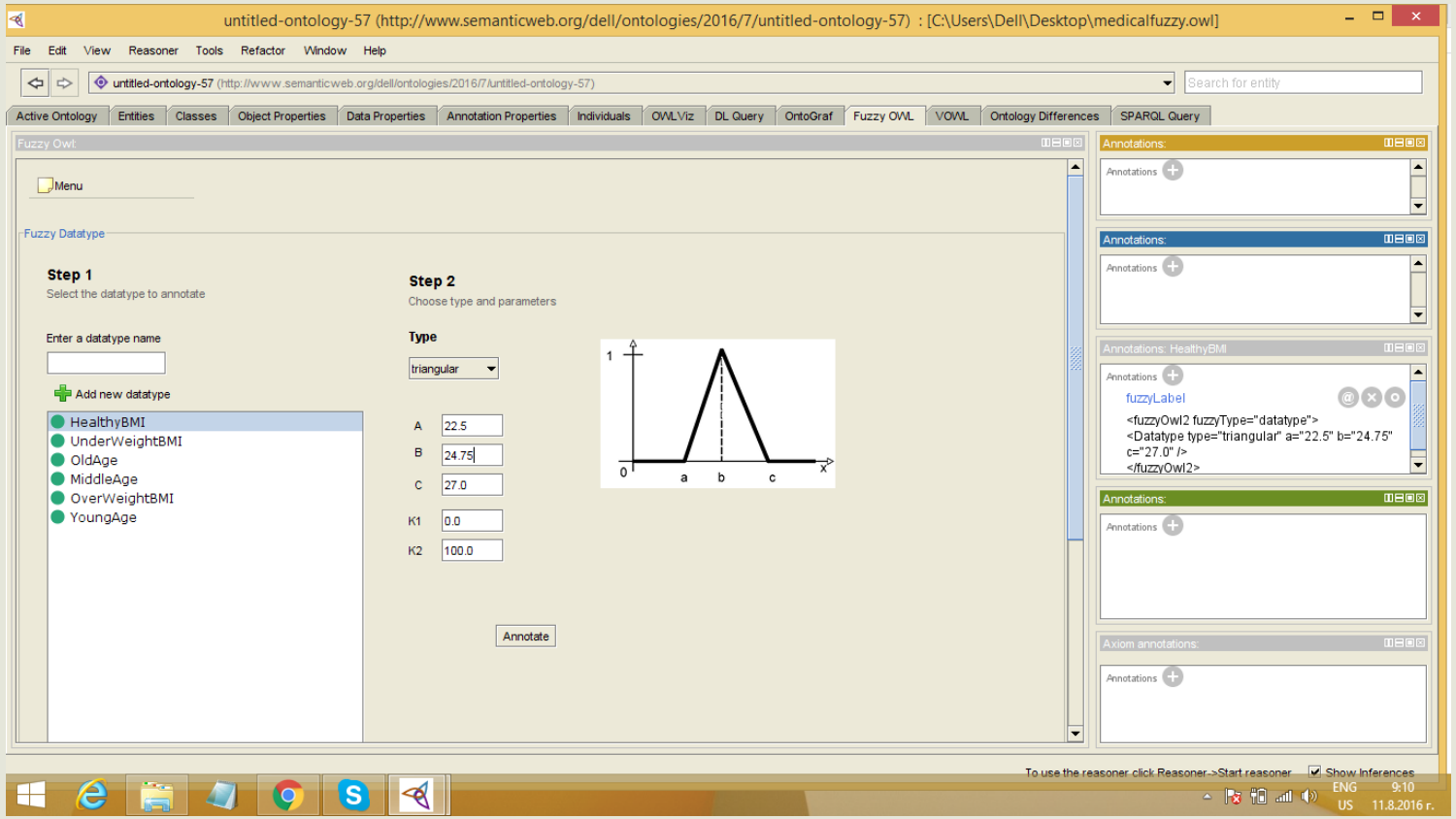
Fuzzy function of BMI



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



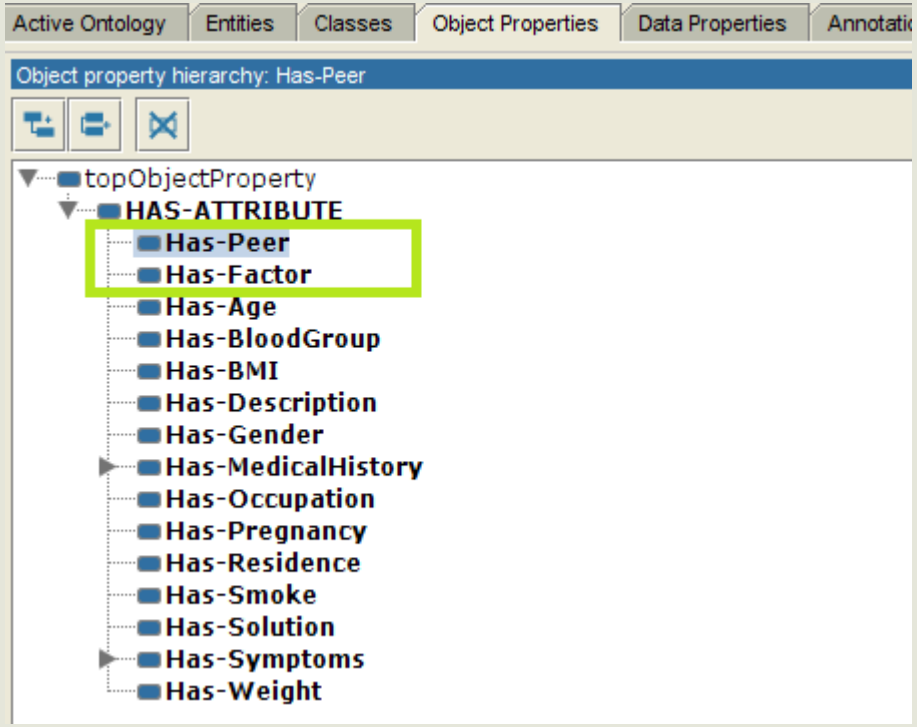
Fuzzy function of BMI



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



Expand the ontology to work with uncertain and Vague information and with Big Data



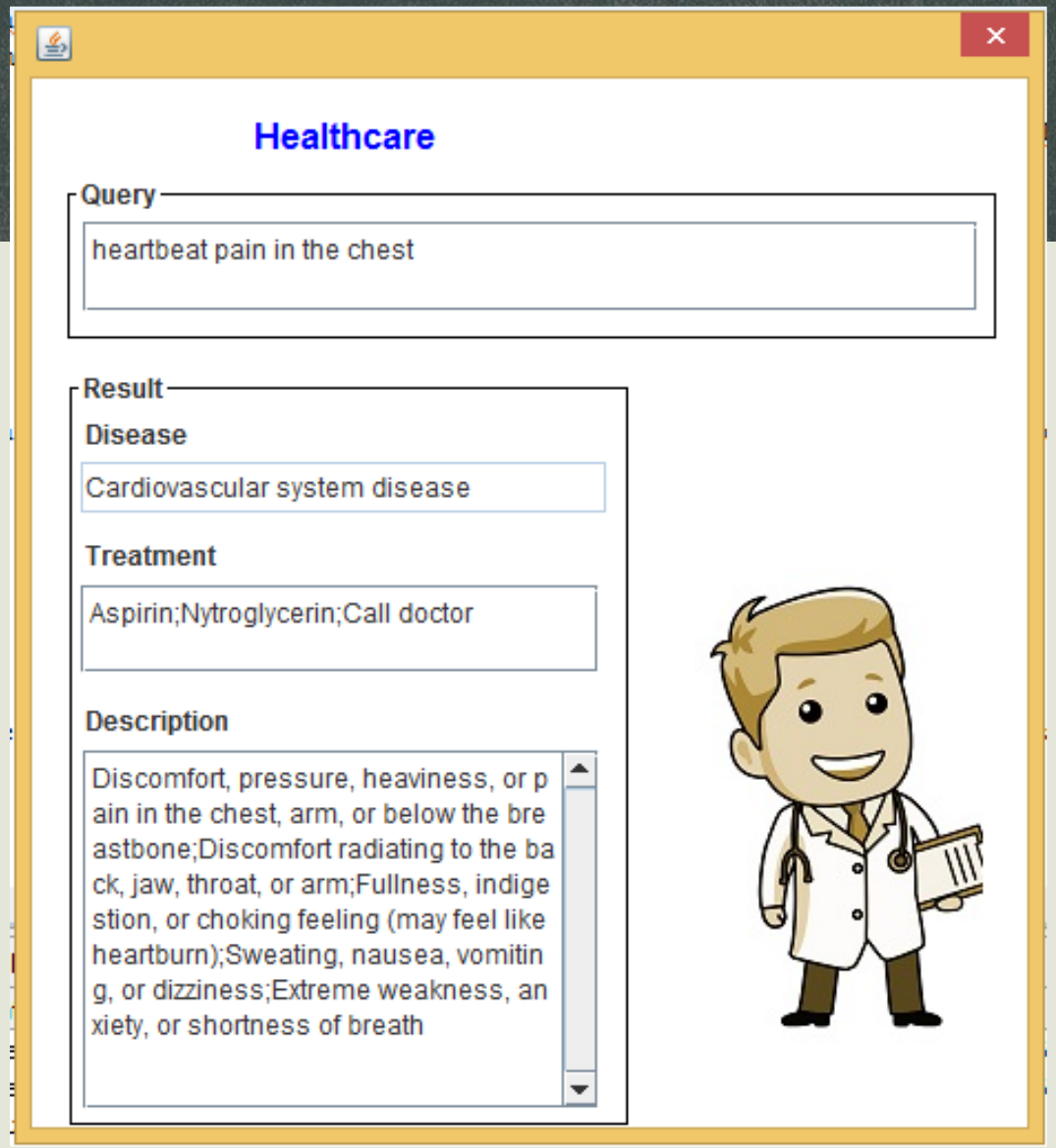


Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Experimental results

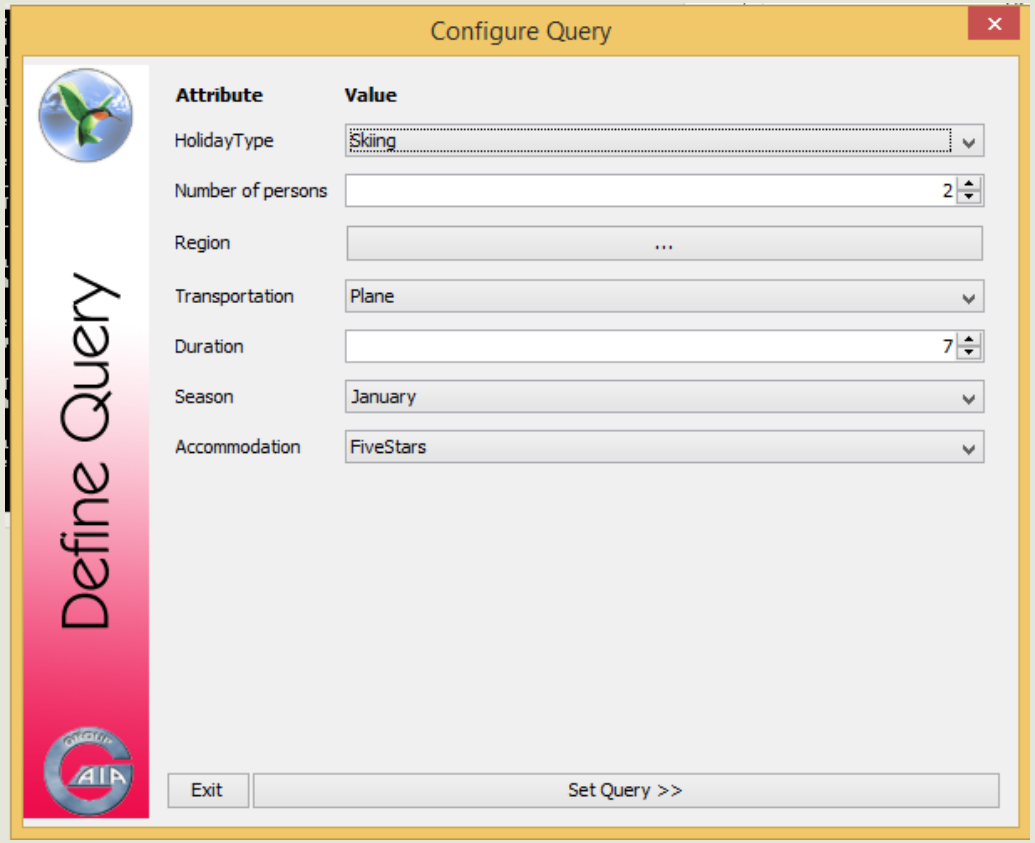
Category symptoms	Total Queries	Diagnosed by the system	Diagnosed by an expert	A1	A2	A3	Precision	Recall	F-measure
Abdominal & Digestive symptom	35	34	36	31	1	3	0,96875	0,911764706	0,939393939
Cardiovascular system symptom	22	22	22	20	0	2	1	0,909090909	0,952380952
Head and neck symptom	19	19	19	18	0	1	1	0,947368421	0,972972973
Hemic and immune system	9	9	8	8	1	0	0,888888889	1	0,941176471
Musculosceleton system symptom	8	10	8	8	0	0	1	1	1
Nervous system symptom	27	26	27	25	0	2	1	0,925925926	0,961538462
Nutrition, metabolism and development symptom	30	32	30	28	1	1	0,965517241	0,965517241	0,965517241
Reproductive system symptom	21	20	20	18	1	2	0,947368421	0,9	0,923076923
Respiratory system and chest symptom	32	32	33	31	0	1	1	0,96875	0,984126984
Skin and intergumentary tissue symptom	13	12	13	12	0	1	1	0,923076923	0,96
Urinary system symptom	9	9	9	8	0	1	1	0,888888889	0,941176471
Total	225	225	225						
Average							0,98	0,94003482	0,96
Legend:									
A1- The system and the expert decide the same									
A2- The system and the expert have different opinions, experts believe that the disease is in another category									
A3-Expert decide that is in this category, but the system not									

Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

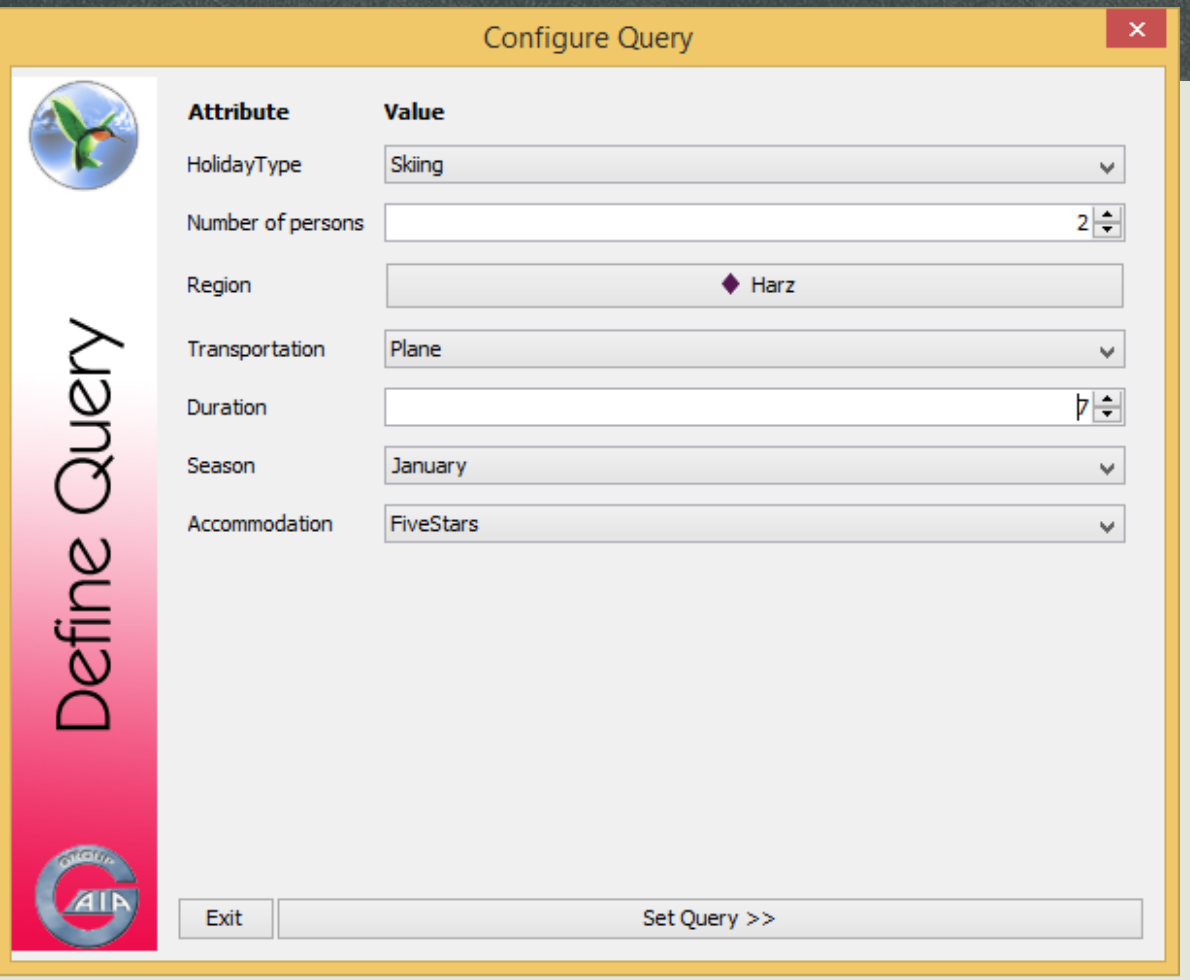


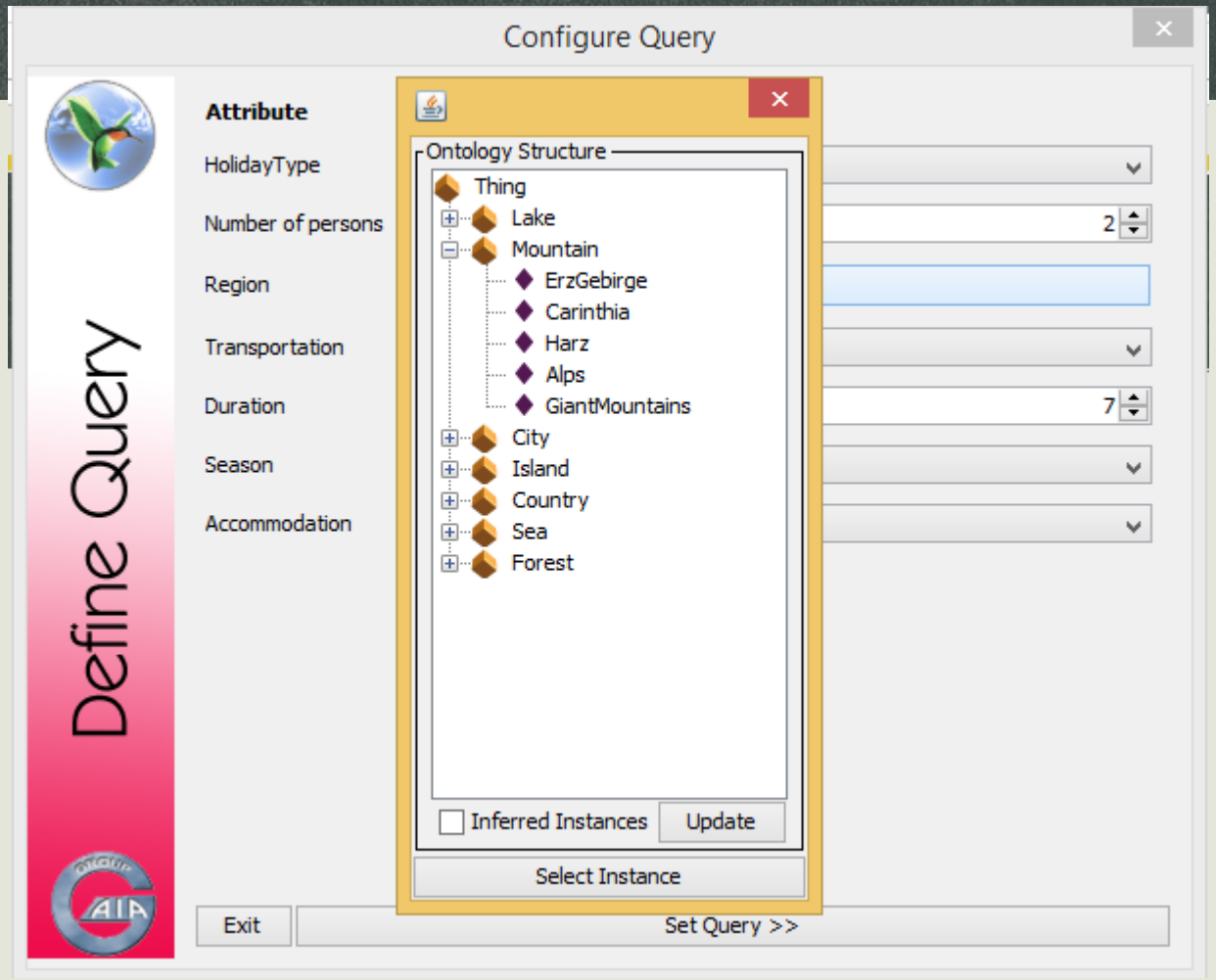
Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

JColibri

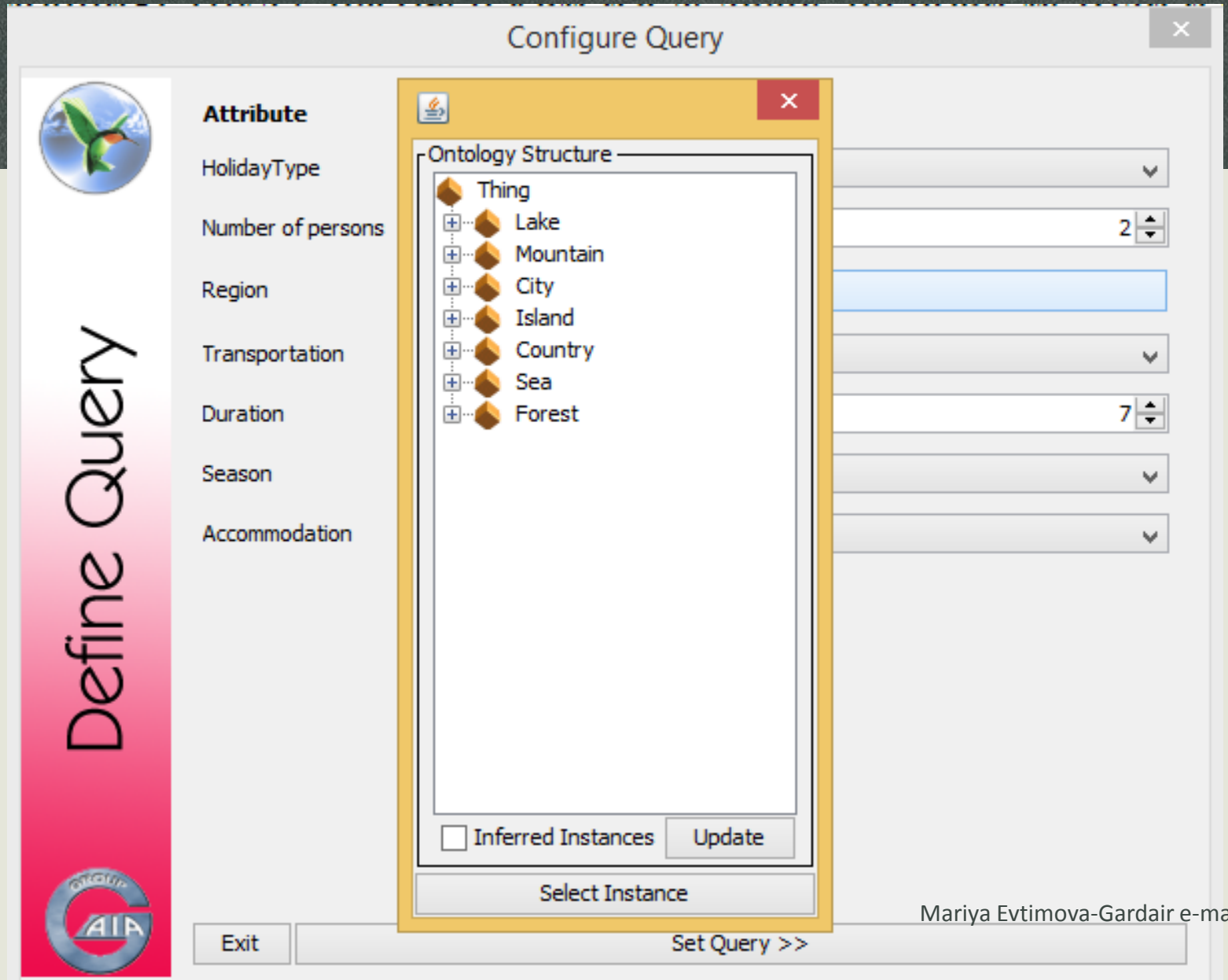


Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

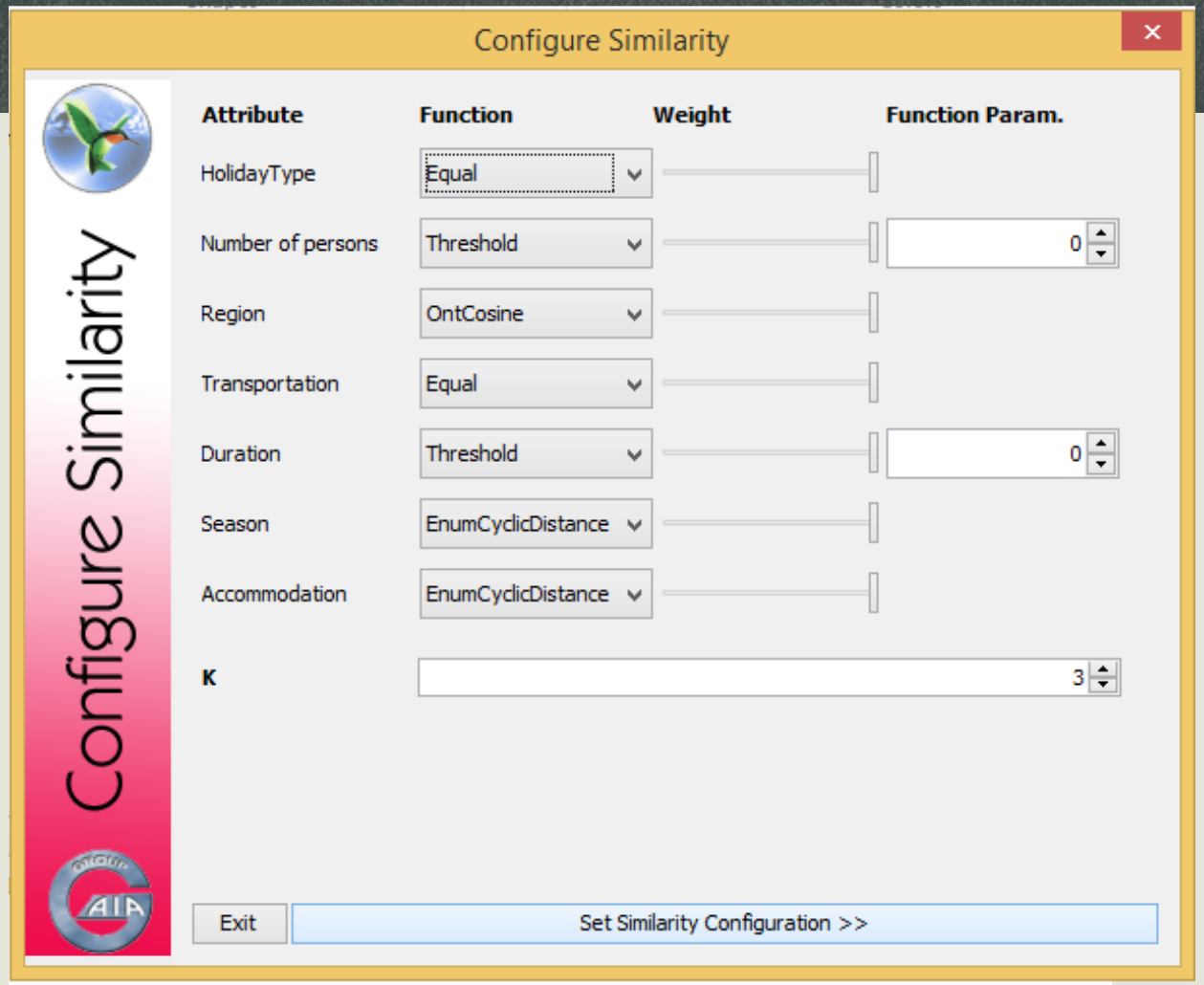




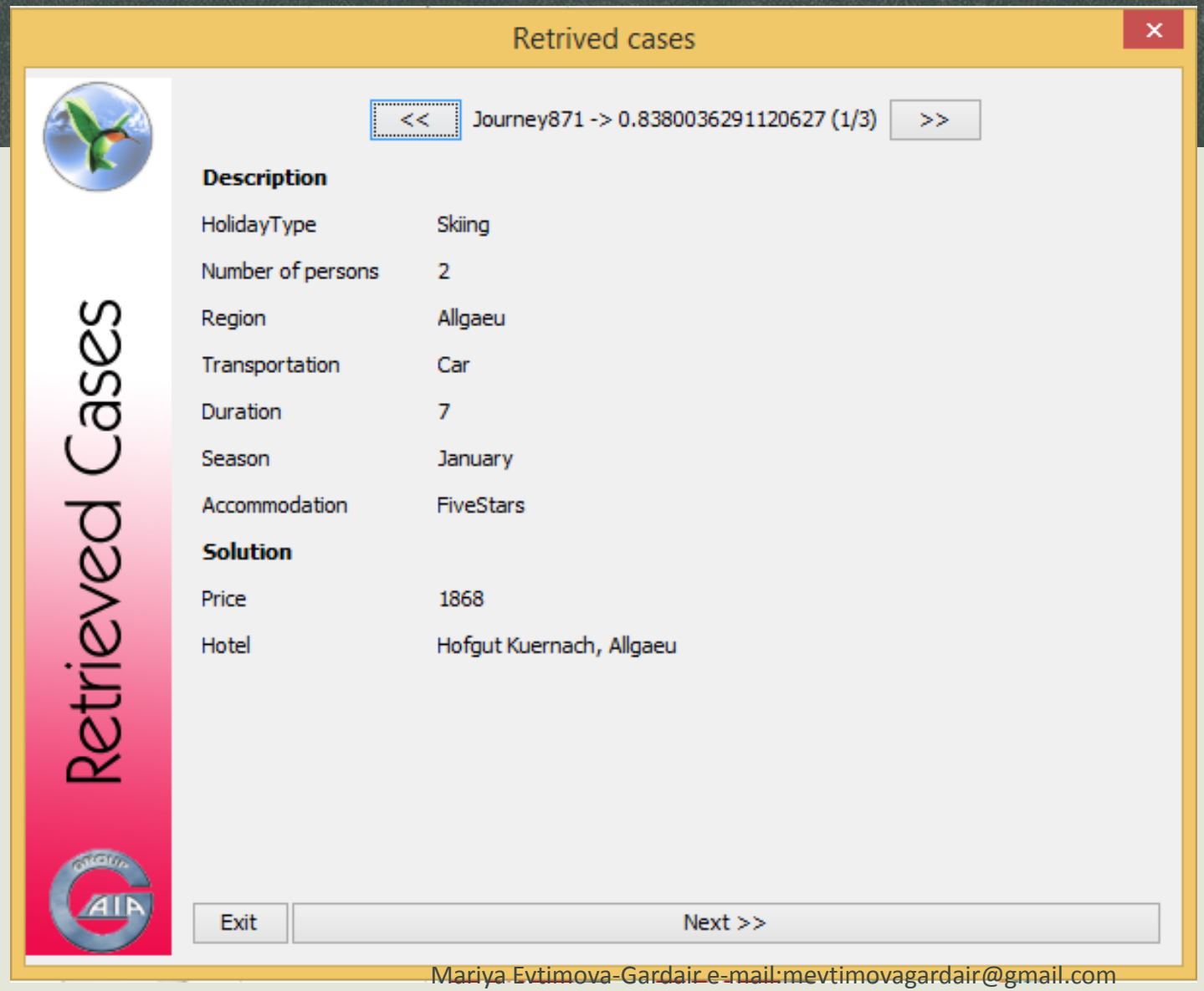
Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



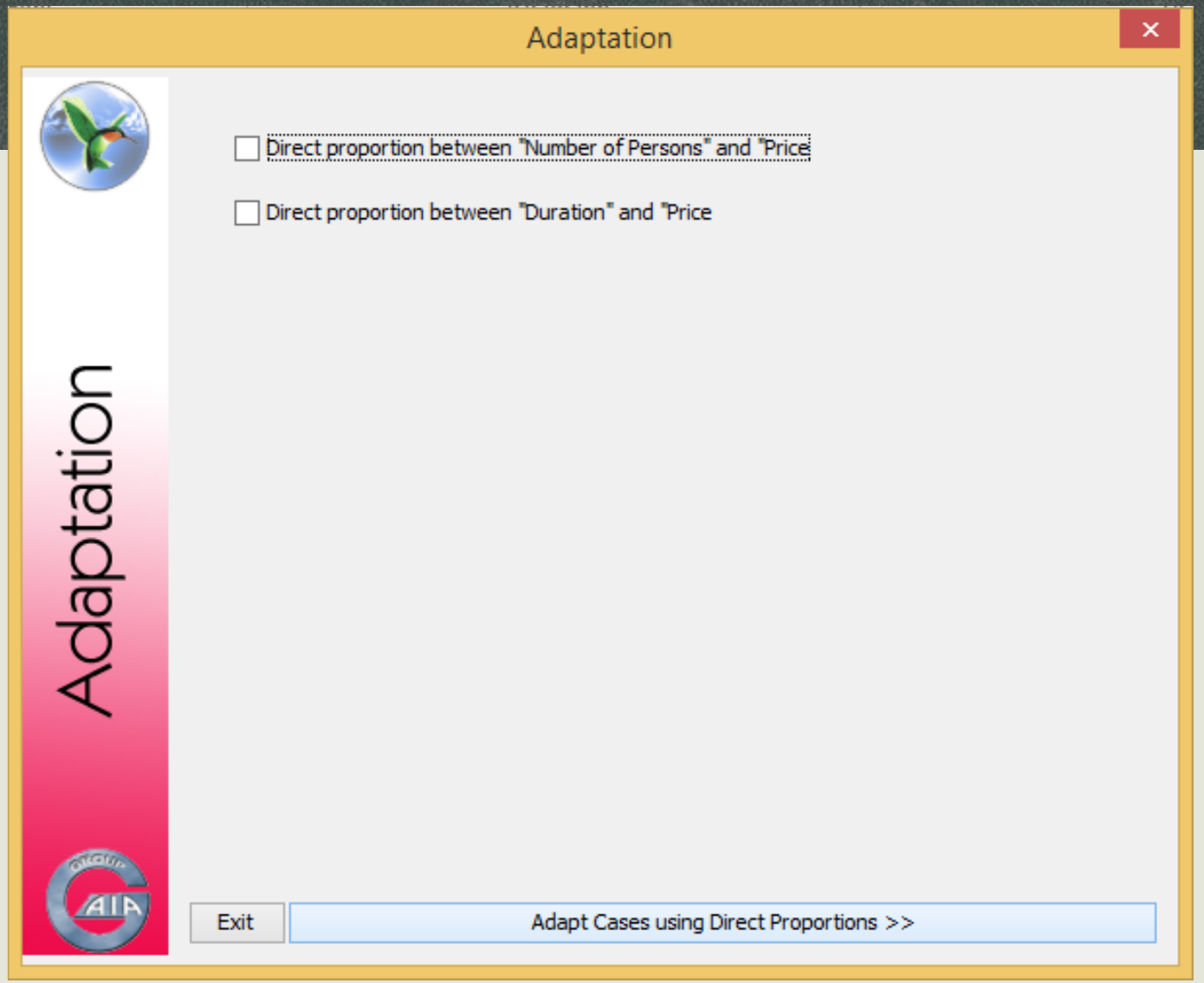
Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

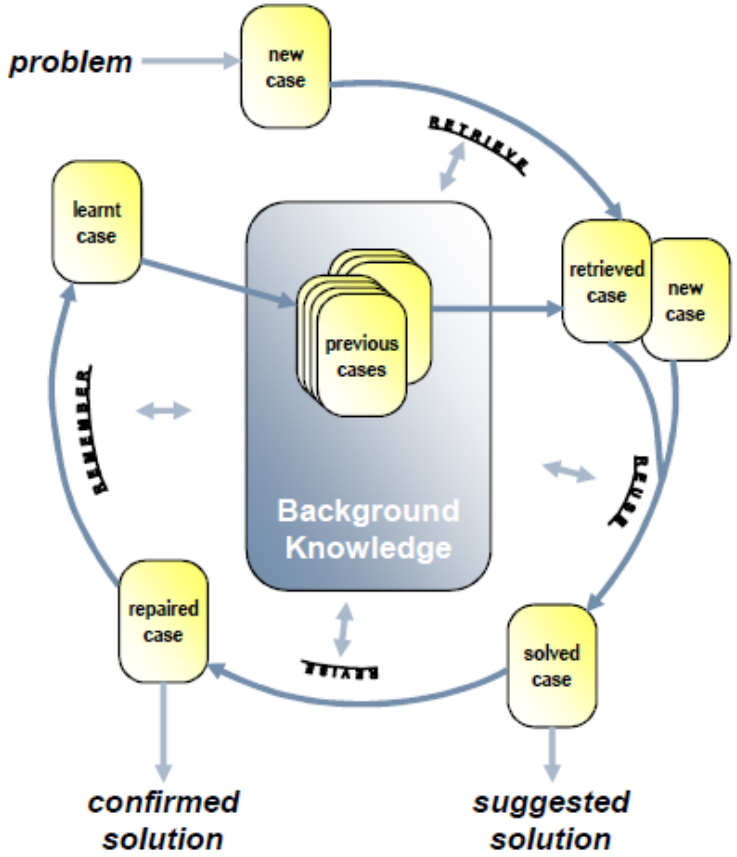


Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com



Mariya Evtimova-Gardair e-mail:mevtimovagardair@gmail.com

Motivation



CBR is:

- ✓ Successful AI subfield
- ✓ Mature & established

Building CBR applications requires:

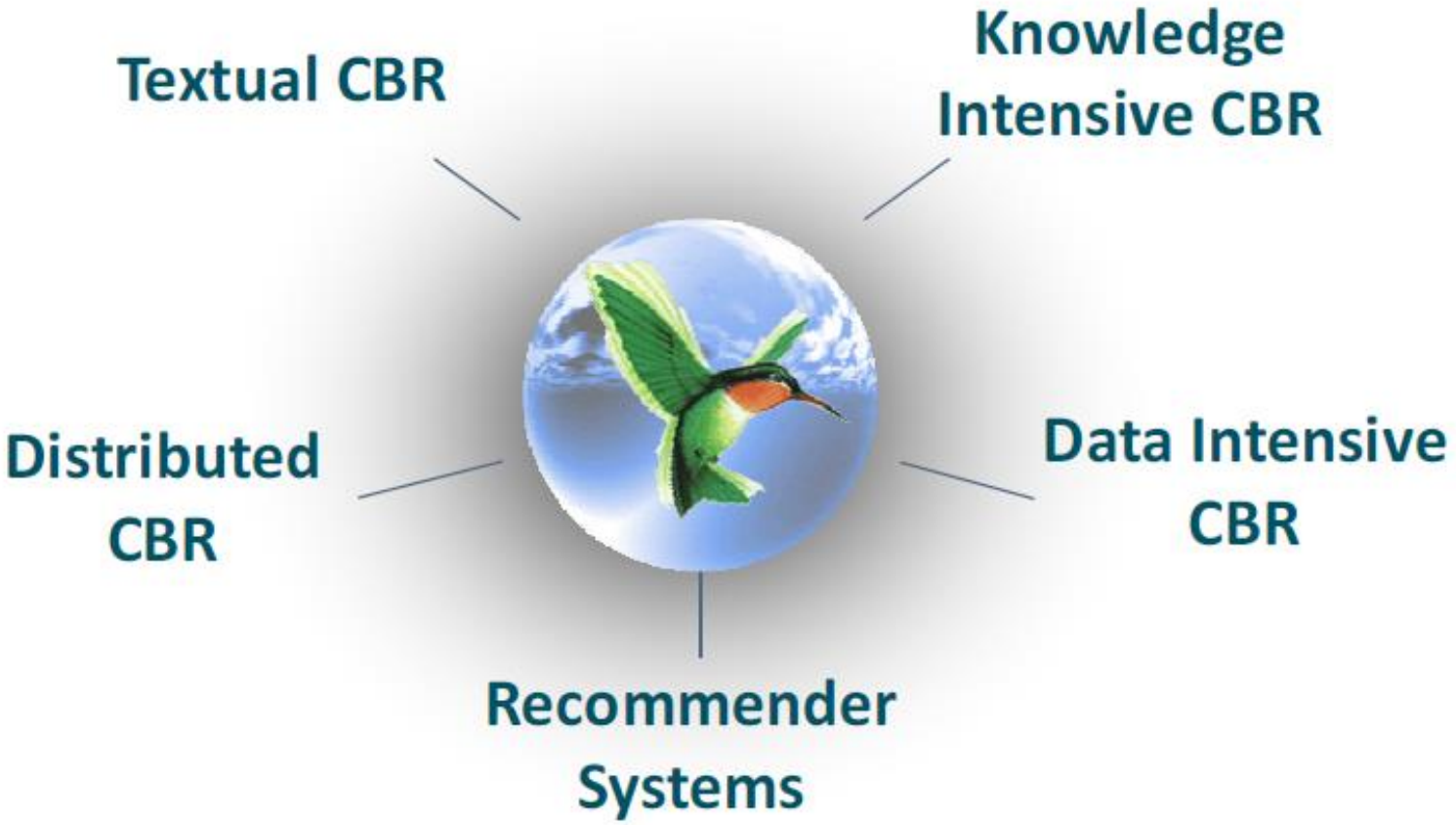
- ➢ Formalization
- ➢ Methodologies
- ➢ Implementation assistance

Main features of jCOLIBRI



- **Extensible**
- **Reusable**
- Different types of **users** and different **purposes**:
  - Development, Research and/or Teaching
- Compatible with **commercial** applications and suitable for developing **large scale** applications
- Supporting different **types of CBR systems**
- It is just a .jar file → **Suitable for web applications**

Extensions


Textual CBR
Knowledge Intensive CBR
Distributed CBR
Data Intensive CBR
Recommender Systems