ff

Knowledge Representation meets Databases at LUB

Marijke Keet

Faculty of Computer Science
Libera Università di Bolzano (Free University of Bozen-Bolzano)
Piazza Domenicani 3, 39100 Bolzano
http://www.inf.unibz.it/krdb

Gent, 18 April 2007

Outline

1. The KRDB group

2. Research at KRDB
   - Semantic-based information management
   - Data and information integration
   - Formal bio-ontologies

3. LUB Contributions to AROBIO

Members of the KRDB group

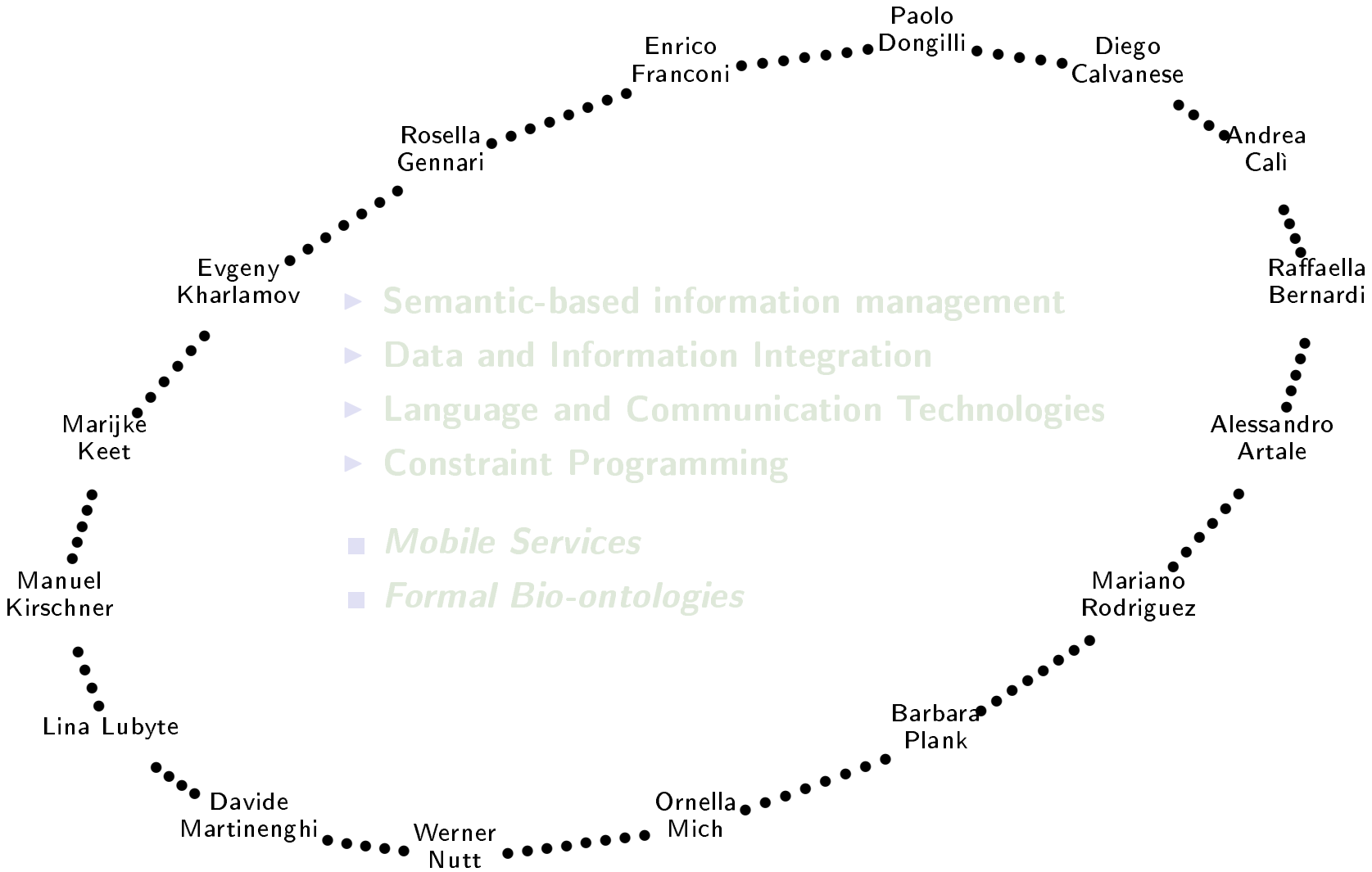
Professors:
Diego Calvanese,
Enrico Franconi,
Werner Nutt,
David Toman

Researchers:
Alessandro Artale,
Raffaella Bernardi,
Jos de Bruijn,
Andrea Calì,
Rosella Gennari,
Davide Martinenghi,
Sergio Tessaris

Research Assistants:
Paolo Dongilli,
Daniele Gobbetti,
Barbara Plank

PhD Students:
Marijke Keet,
Evgeny Kharlamov,
Manuel Kirschner,
Lina Lubyte,
Ornella Mich,
Mariano Rodríguez,
Vladislav Ryzhikov,
Camilo Thorne

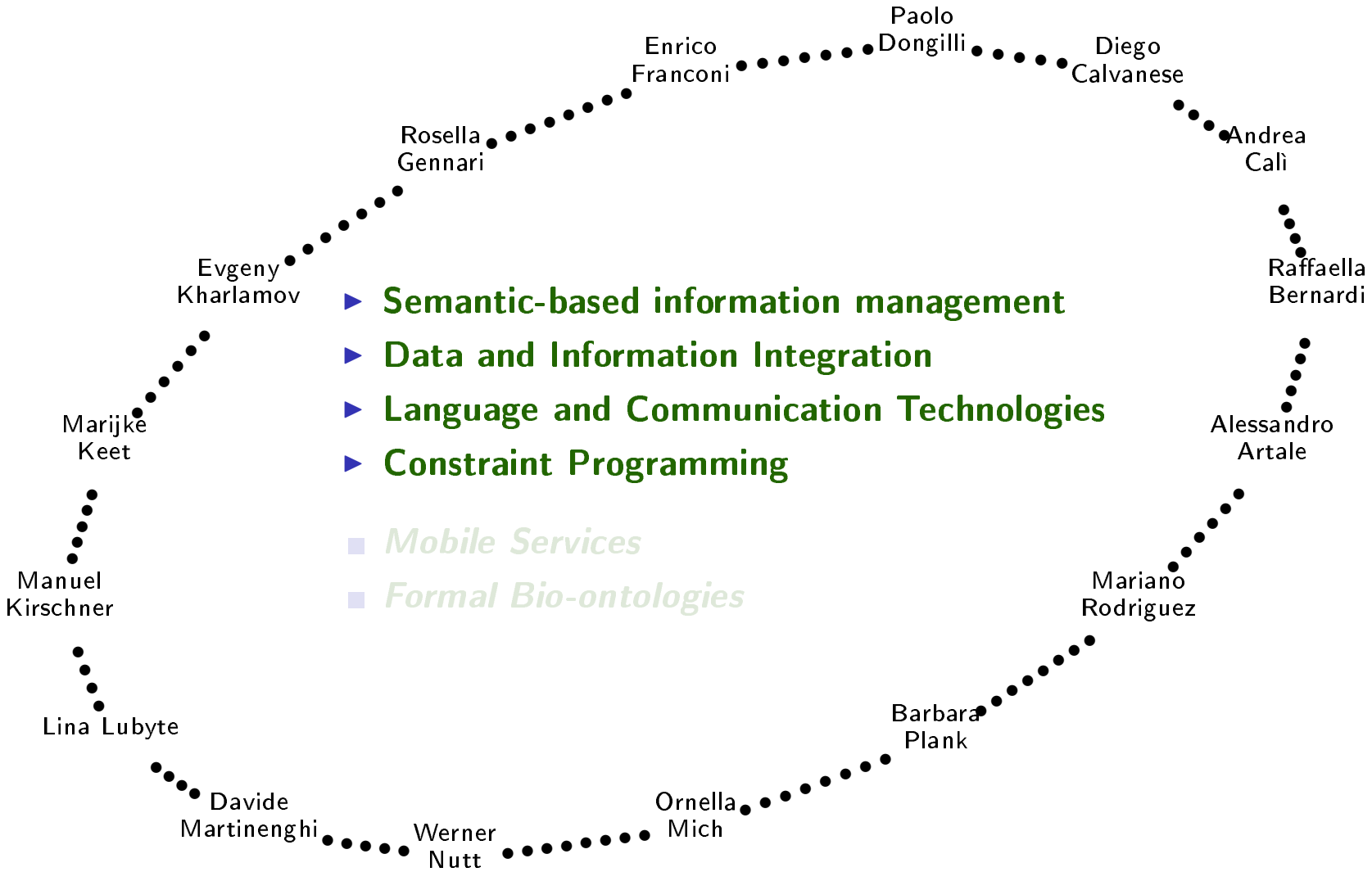
Research topics in the KRDB group



- Semantic-based information management
- Data and Information Integration
- Language and Communication Technologies
- Constraint Programming

- *Mobile Services*
- *Formal Bio-ontologies*

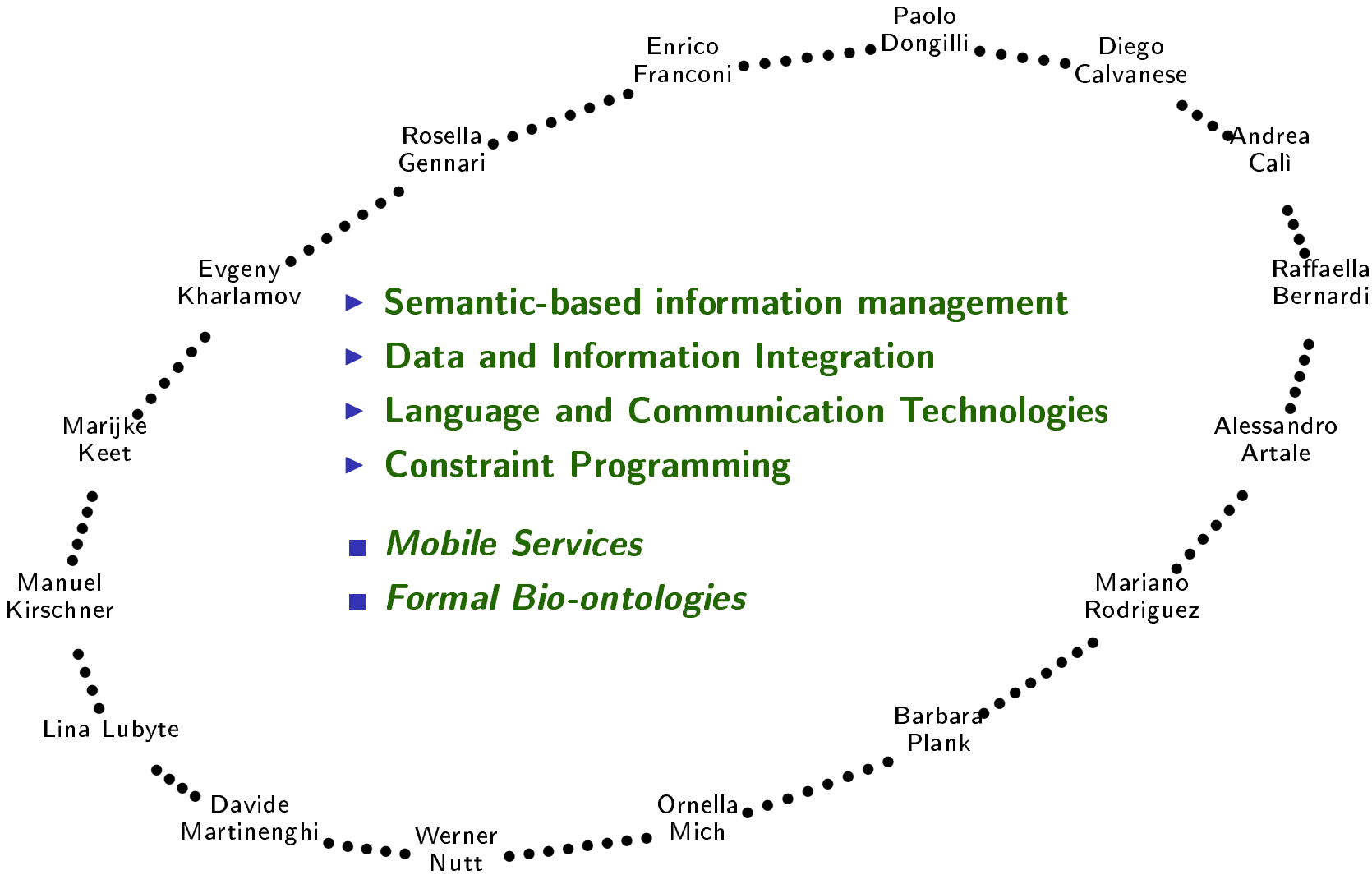
Research topics in the KRDB group



- **Semantic-based information management**
- **Data and Information Integration**
- **Language and Communication Technologies**
- **Constraint Programming**

- *Mobile Services*
- *Formal Bio-ontologies*

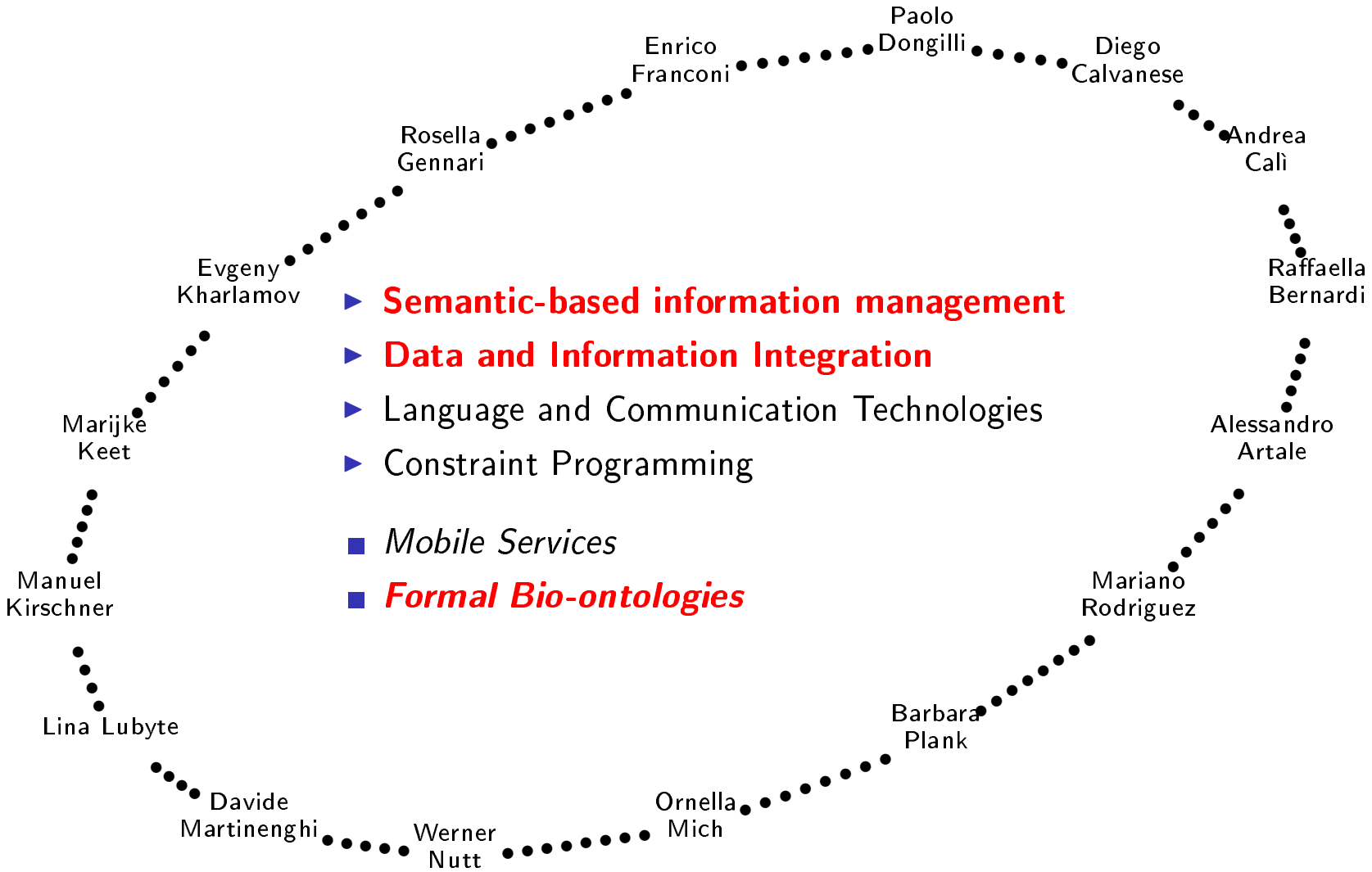
Research topics in the KRDB group



- **Semantic-based information management**
- **Data and Information Integration**
- **Language and Communication Technologies**
- **Constraint Programming**

- ***Mobile Services***
- ***Formal Bio-ontologies***

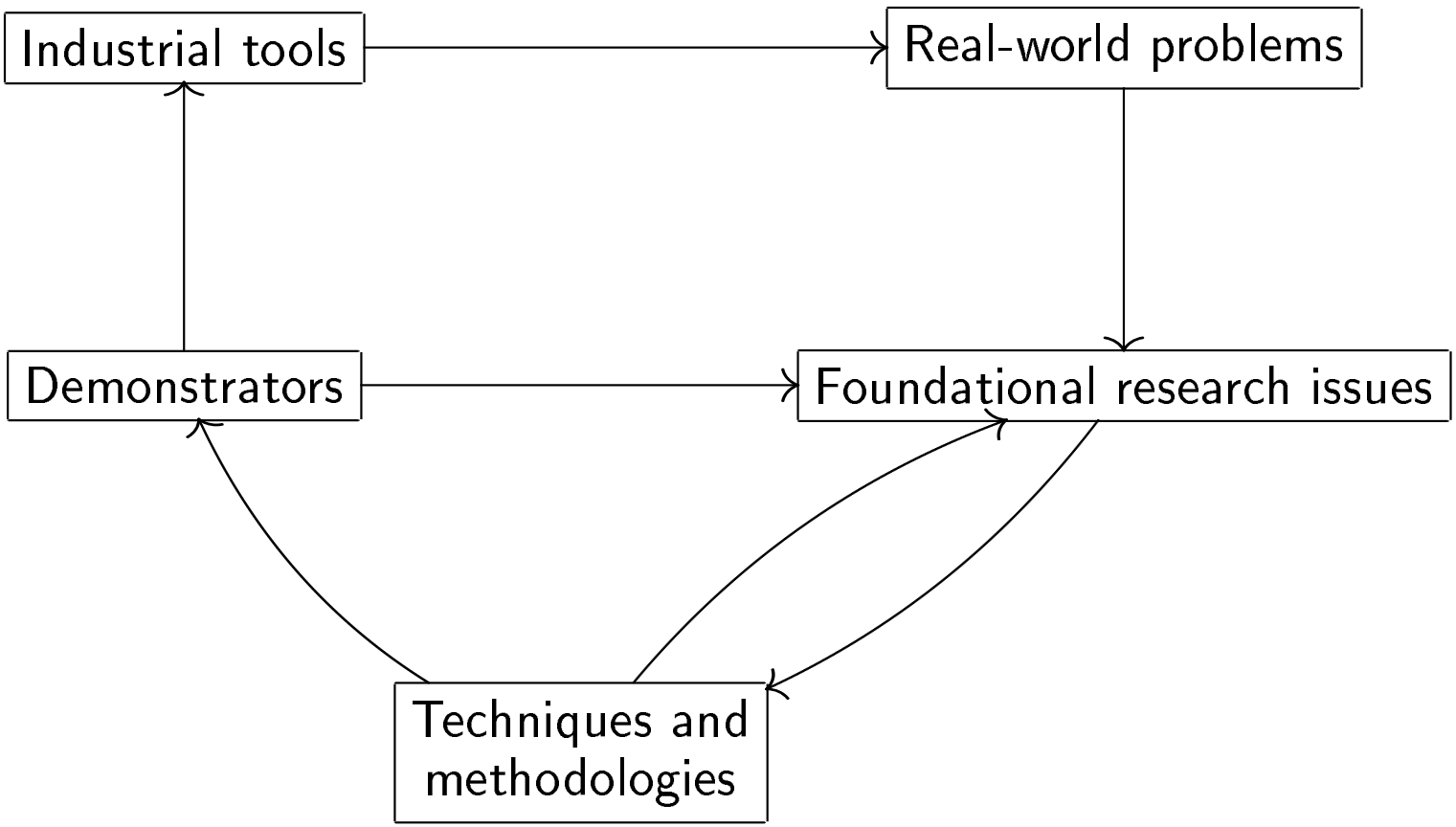
Research topics in the KRDB group



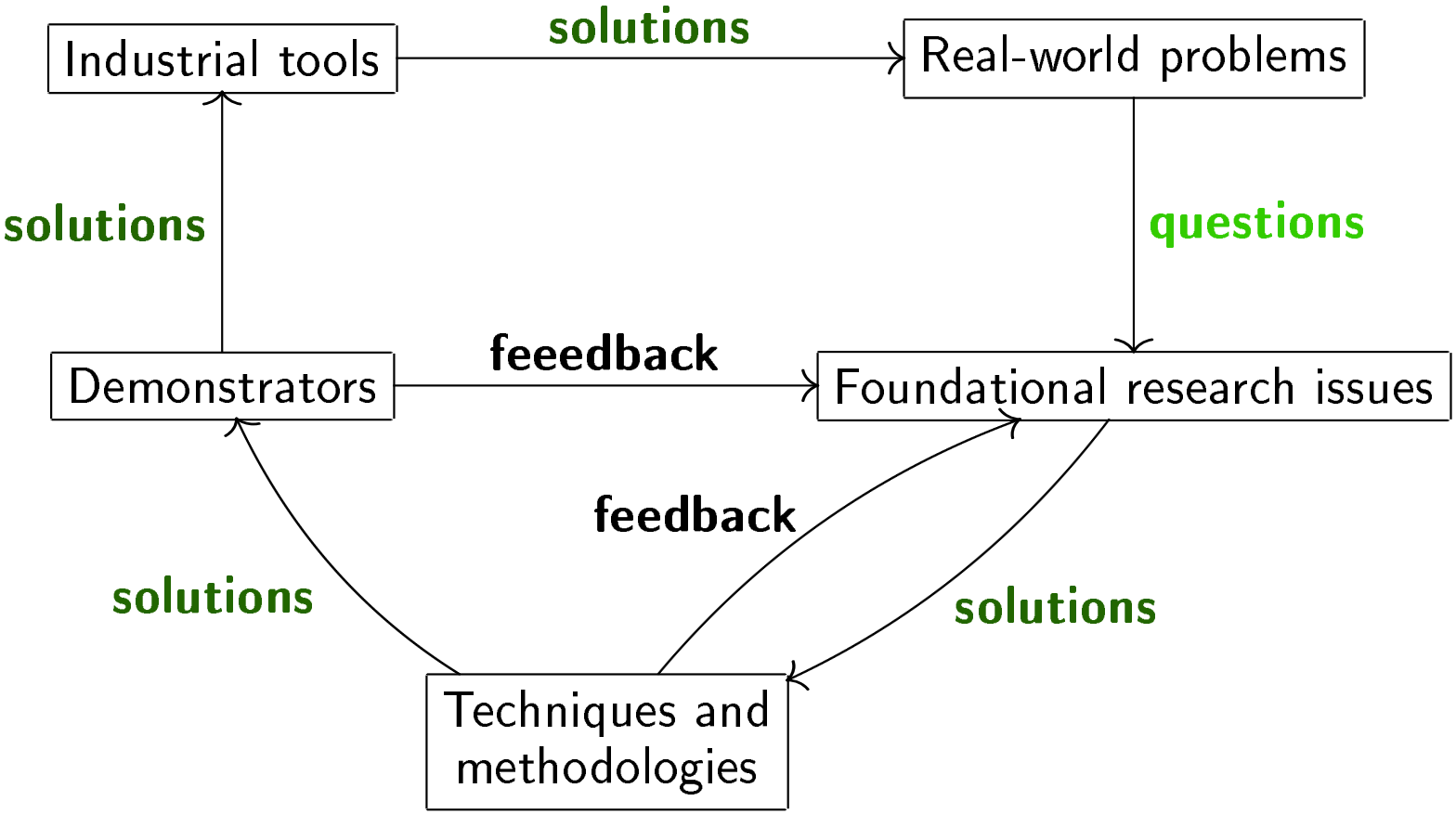
- **Semantic-based information management**
- **Data and Information Integration**
- Language and Communication Technologies
- Constraint Programming

- *Mobile Services*
- ***Formal Bio-ontologies***

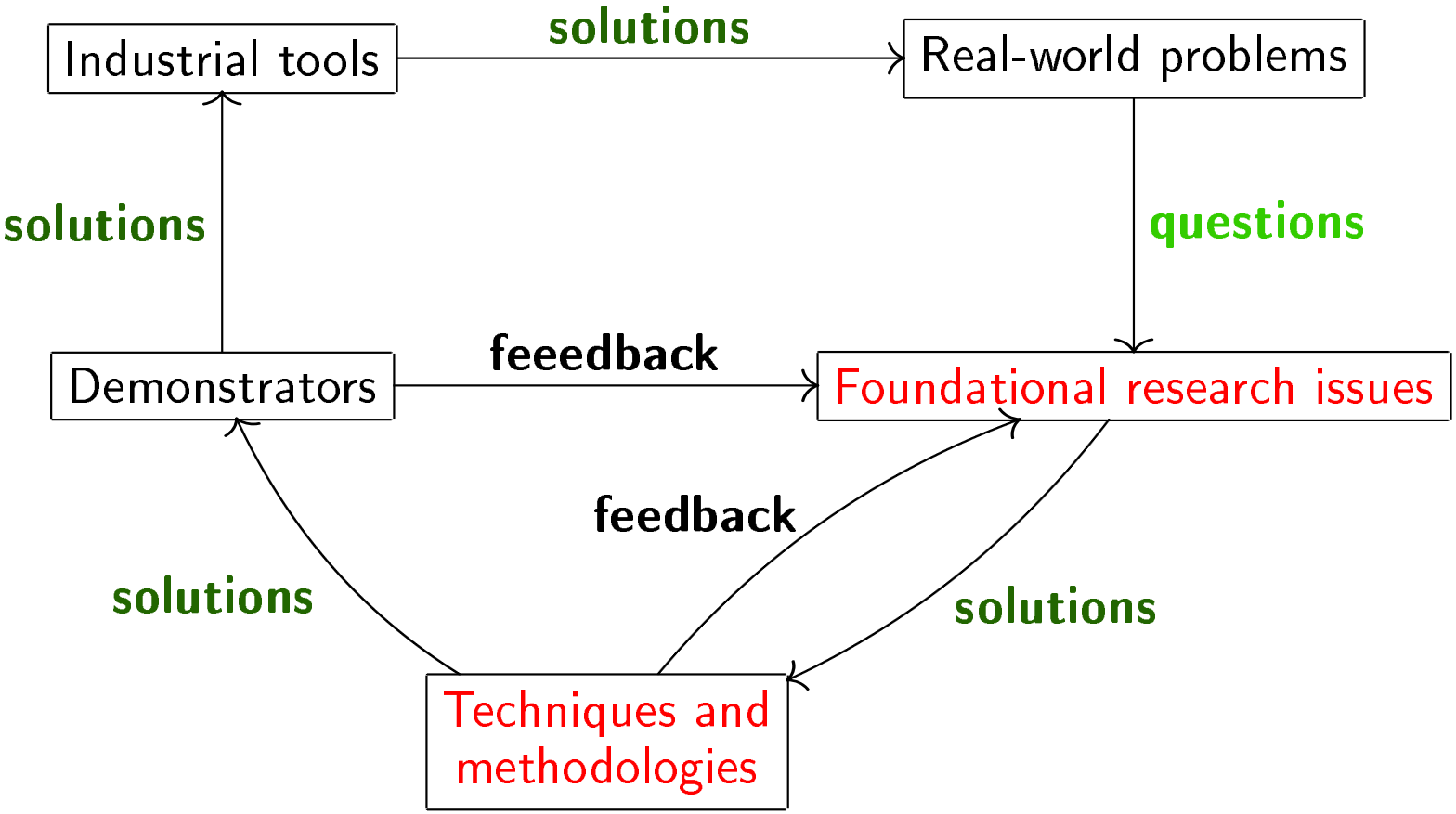
Research Methodology



Research Methodology



Research Methodology



Research topics

- **Semantic-based information management**
- **Data and Information Integration**
- Language and Communication Technologies
- Constraint Programming

- *Mobile Services*
- ***Formal Bio-ontologies***

Semantic-based information management

Use of semantic-based methodologies and techniques to manage information explosion and complexity of data

Problems and foundational research issues

- How to pinpoint the right piece of information?
- How to query and use the available data?
- How to deal with information sources built *ad-hoc*?
- How to bridge the gap between information and data? (Semantic Web)

Adopted methodologies and techniques

- Model information sources at the conceptual level
- Relationship between conceptual models and logic-based formalisms
- Automated reasoning techniques to support the whole life-cycle of data management ("intelligent" conceptual modeling tools, ontology-supported query formulation, ontology-based data access)

Semantic-based information management at KRDB

Ontologies: Main activities and research topics

- Ontology-based data management
- Languages for ontologies and for the Semantic Web
- Reasoning over ontologies
- Ontology design and development (bio-ontologies)

Involved KRDB people: Artale, De Bruijn, Calì, Calvanese, Franconi, Keet, Lubyte, Martinenghi, Ryzhikov, Rodríguez, Tessaris, Toman

Semantic-based information management at KRDB

Developed tools and systems

- iCom: UML based intelligent Conceptual Modelling
  - graphical user-interface for conceptual modelling
  - logic-based semantics enables automated support to the modeller
  - verifies the specification, infers implicit constraints, etc.
- QTool: semantics driven query formulation support
  - support in formulation of information needs
  - exploits the schema describing the semantics of the data
- QuOnto: Ontology-based access to data sources (developed in collaboration with the University of Rome)
  - processes queries over ontologies by relying on relational data
  - queries ultimately handled by relational engine (scalability)

Semantic-based information management at KRDB

Developed tools and systems

- iCom: UML based intelligent Conceptual Modelling
  - graphical user-interface for conceptual modelling
  - logic-based semantics enables automated support to the modeller
  - verifies the specification, infers implicit constraints, etc.
- QTool: semantics driven query formulation support
  - support in formulation of information needs
  - exploits the schema describing the semantics of the data
- QuOnto: Ontology-based access to data sources (developed in collaboration with the University of Rome)
  - processes queries over ontologies by relying on relational data
  - queries ultimately handled by relational engine (scalability)

Semantic-based information management at KRDB

Developed tools and systems

- iCom: UML based intelligent Conceptual Modelling
  - graphical user-interface for conceptual modelling
  - logic-based semantics enables automated support to the modeller
  - verifies the specification, infers implicit constraints, etc.
- QTool: semantics driven query formulation support
  - support in formulation of information needs
  - exploits the schema describing the semantics of the data
- QuOnto: Ontology-based access to data sources (developed in collaboration with the University of Rome)
  - processes queries over ontologies by relying on relational data
  - queries ultimately handled by relational engine (scalability)

Semantic-based information management at KRDB

External partners and projects

- Work carried within several European projects
  - TONES (Thinking ONtologiES): 3-year Basic Research project coordinated by KRDB — Development of basic inference techniques and tools for ontology management.
    http://www.tonesproject.org/
  - KnowledgeWeb (Realizing the Semantic Web): 4-year Network of Excellence project — Support the transition process of Ontology technology from Academia to Industry.
    http://knowledgeweb.semanticweb.org/
  - InterOp (Interoperability Research for Networked Enterprises Applications and Software): 3-year Network of Excellence project – http://interop-noe.org/
- Industrial partnerships through collaborations within projects
- KRDB is a member of the W3C (World-Wide-Web Consortium)

Data and information integration

Deals with the problem of providing uniform access to multiple, heterogeneous data and information sources

Problems and foundational research issues

- How to model an integration system? (global schema, mappings to and from sources, P2P systems, streaming data)
- How to process queries over integration systems?
- How to deal with incompleteness of information?
- How to deal with inconsistencies among sources?
- How to scale with number and size of information sources

Adopted methodologies and techniques

- formal tools (based on logic) to model of data integration systems
- query processing via query rewriting and view-based query answering
- query processing under constraints (database chase techniques)
- integration in P2P systems

Data and information integration at KRDB

Developed tools and systems

- iCom:
  - enables also the modelling and integration of multiple data sources
  - logic-based mappings to establish semantic relationships
- QuOnto:
  - handles multiple data sources through mappings from a global schema
- coDBz: peer to peer data management system
  - Based on a robust logical and computational characterisation of peer-to-peer data management systems
  - Implements distributed algorithms for robust data sharing and updates

Involved KRDB people: Artale, Calì, Calvanese, Franconi, Kharlamov, Martinenghi, Nutt, Tessaris, Toman

Formal bio-ontologies at KRDB

Foundational research issues: **Ontologies in bio-informatics**

- Focus on the conceptual analysis stage (what, why, how), types of relations in ontologies
- Different languages (expressiveness & limitations of languages)
- Granularity

Adopted methodologies and techniques

- Methodology for bottom-up ontology development of ontologies (in ecology mining extant eco-tool)
- Improving (bio-)ontology management with granularity (e.g., for infectious diseases, granular queries)
- Foundational relations for biology, such as parthood and transformation
- Commenced project: *In silico* bio-hypothesis testing using conceptual models

Formal bio-ontologies at KRDB

Foundational research issues: **Ontologies in bio-informatics**

- Focus on the conceptual analysis stage (what, why, how), types of relations in ontologies
- Different languages (expressiveness & limitations of languages)
- Granularity

Adopted methodologies and techniques

- Methodology for bottom-up ontology development of ontologies (in ecology mining extant eco-tool)
- Improving (bio-)ontology management with granularity (e.g., for infectious diseases, granular queries)
- Foundational relations for biology, such as parthood and transformation
- Commenced project: *In silico* bio-hypothesis testing using conceptual models

Formal bio-ontologies at KRDB (2)

Involved KRDB people: Artale, Franconi, Keet, Rodríguez

External partners

- IFOMIS (biomedicine & ontologies)
- MAD-IBU (bioinformatics, Uni of Amsterdam)
- LOA-CNR Trento (ontology)
- Uni Washington (FMA)
- Ochman lab (molbio & HGT), Evolutionary Genomics Group Uni Tarragona (HGT, bioinformatics)

Organized events

- CSBio Seminar Series: "Computer Science & IT with/for Biology" http://www.inf.unibz.it/krdb/biology/
- Hosting co-founder of the OBO Foundry Prof. Barry Smith

Research topics

- WP6
  - Temporal representation & reasoning: apply & improve latest results on temporal DL language, $\mathcal{DLR}_{\mathcal{US}}$, and reasoning to adapt it for OWL DL and/or OWL 1.1.
  - Granularity: develop computationally usable variant of the theory of granularity for use with ontologies/KBs, expand on current types of granular queries to more complex reasoning tasks.
  - Ontolome analysis: focus on the automated reasoning/querying, feasibility & trade-offs for implementation; comparisons and validating hypothesis against lots of data does not scale well, but should be possible with *DL-Lite*, which also can be used to link the ontology to the data. (latter can be linked to WP5 activities regarding simper ontology languages)
  - Quantitative aspects: expand upon modelling & querying role values (and ternary relations)

Research topics

- **WP2** – Semantic power: collaborate from WP6 to making theory on temporal & quantitative representations usable, ontology development methodology (principles, design criteria), link with top-level ontologies
- **WP5** – Heterogeneity of data structures: optimisation of mapping between OBO and OWL flavours, including mappings with *DL-Lite*.

- WP3 – visualisation: depends
- WP5 – Persistence of user comments: contribute to requirements engineering modelling (CA phase) for the storage of 'personal' knowledge ?
- WP xx: Any training & dissemination activities ?

Research topics

- **WP2** – Semantic power: collaborate from WP6 to making theory on temporal & quantitative representations usable, ontology development methodology (principles, design criteria), link with top-level ontologies
- **WP5** – Heterogeneity of data structures: optimisation of mapping between OBO and OWL flavours, including mappings with *DL-Lite*.

- WP3 – visualisation: depends
- WP5 – Persistence of user comments: contribute to requirements engineering modelling (CA phase) for the storage of 'personal' knowledge ?
- WP xx: Any training & dissemination activities ?