

RESEARCH CENTRE

**Inria Branch at the University of  
Montpellier**

IN PARTNERSHIP WITH:

Université de Montpellier, INRAE

2024

ACTIVITY REPORT

Project-Team

**BOREAL**

## **Knowledge Representation and Rule-Based Languages for Reasoning on Data**

IN COLLABORATION WITH: Laboratoire d'informatique, de robotique et  
de microélectronique de Montpellier (LIRMM), Ingénierie des  
Agropolymères et Technologies Emergentes (IATE)

**DOMAIN**

**Perception, Cognition and Interaction**

**THEME**

**Data and Knowledge Representation and  
Processing**

The Inria logo is a stylized, cursive script in red, positioned in the bottom right corner of the page.

# Contents

<b>Project-Team BOREAL</b>	<b>1</b>
<b>1 Team members, visitors, external collaborators</b>	<b>2</b>
<b>2 Overall objectives</b>	<b>3</b>
<b>3 Research program</b>	<b>3</b>
<b>4 Application domains</b>	<b>4</b>
4.1 Agronomy and agroecology . . . . .	4
<b>5 Highlights of the year</b>	<b>5</b>
<b>6 New software, platforms, open data</b>	<b>5</b>
6.1 New software . . . . .	5
6.1.1 InteGraal . . . . .	5
6.1.2 TreeForce . . . . .	6
6.1.3 B-Runner . . . . .	7
<b>7 New results</b>	<b>7</b>
7.1 Preliminaires about Knowledge-Based Data Management with Existential Rules . . . . .	7
7.2 Contributions . . . . .	8
7.3 Fine-Grained Complexity of Database Queries . . . . .	8
7.4 Repairing Databases . . . . .	9
7.5 Foundations of Reasoning with Logical Languages . . . . .	10
7.5.1 Reasoning with Existential Rules . . . . .	10
7.5.2 Non-Monotonic Reasoning . . . . .	10
7.6 Systems for Reasoning on Integrated Data . . . . .	11
7.6.1 Benchmarking Reasoning Systems . . . . .	11
7.6.2 Exploring the Application of Ontology Based Data Access in Agroecology . . . . .	11
<b>8 Partnerships and cooperations</b>	<b>12</b>
8.1 International research visitors . . . . .	12
8.1.1 Visits of international scientists . . . . .	12
8.1.2 Visits to international teams . . . . .	14
8.1.3 Visits to national teams . . . . .	14
8.2 National initiatives . . . . .	15
<b>9 Dissemination</b>	<b>15</b>
9.1 Promoting scientific activities . . . . .	15
9.1.1 Scientific events: organisation . . . . .	15
9.1.2 Scientific events: selection . . . . .	16
9.1.3 Journal . . . . .	16
9.1.4 Scientific expertise . . . . .	16
9.1.5 Research administration . . . . .	16
9.2 Teaching - Supervision - Juries . . . . .	17
9.2.1 Teaching . . . . .	17
9.2.2 Supervision . . . . .	17
9.2.3 Juries . . . . .	17
9.3 Popularization . . . . .	17
9.3.1 Productions (articles, videos, podcasts, serious games, ...) . . . . .	18
9.3.2 Participation in Live events . . . . .	18

<b>10 Scientific production</b>	<b>18</b>
10.1 Major publications	18
10.2 Publications of the year	19
10.3 Cited publications	20

## **Project-Team BOREAL**

*Creation of the Project-Team: 2022 June 01*

### **Keywords**

#### **Computer sciences and digital sciences**

- A3.1. – Data
- A3.2. – Knowledge
- A7.1.3. – Graph algorithms
- A7.2. – Logic in Computer Science
- A9. – Artificial intelligence
- A9.1. – Knowledge
- A9.8. – Reasoning

#### **Other research topics and application domains**

- B3.5. – Agronomy
- B6.5. – Information systems

## 1 Team members, visitors, external collaborators

### Research Scientists

- Federico Ulliana [Team leader, UNIV MONTPELLIER, Associate Professor Detachement]
- Jean-Francois Baget [INRIA, Researcher]
- Pierre Bisquert [INRAE, Researcher]
- Nofar Carmeli [INRIA, Researcher]
- David Carral Martinez [INRIA, Researcher]

### Faculty Members

- Michel Chein [UNIV MONTPELLIER, Emeritus]
- Michel Leclère [UNIV MONTPELLIER, Associate Professor]
- Marie-Laure Mugnier [UNIV MONTPELLIER, Professor, Part-time (80%) in 2024]

### Post-Doctoral Fellow

- Guillaume Perution-Kihli [INRIA, Post-Doctoral Fellow, from Mar 2024]

### PhD Students

- Akira Charoensit [INRIA]
- Lucas Larroque [ENS PARIS, Student based at the ENS Paris; co-supervised by David Carral]

### Technical Staff

- Guillaume Perution-Kihli [INRIA, Engineer, until Feb 2024]
- Florent Tornil [INRIA, Engineer, until Nov 2024]

### Interns and Apprentices

- François Colin De Verdier [ENS DE LYON, Intern, from Jun 2024 until Jul 2024]
- Noah Collinet [INRIA, Intern, until Jun 2024]
- Maksym Lytvynenko [INRIA, Intern, from Jun 2024 until Aug 2024]
- Clément Rouvroy [INRIA, Intern, from Jun 2024 until Aug 2024]
- Nicolas Valayannopoulos-Akrivou [INRIA, Intern, from May 2024 until Aug 2024]

### Administrative Assistant

- Maeva Jeannot [INRIA]

### External Collaborators

- Patrice Buche [INRAE]
- Maxime Buron [UNIV CLERMONT AUVERG]
- Alain Gutierrez [CNRS, from Feb 2024]

## 2 Overall objectives

Current information systems are grounded on the exploitation of data coming from an increasing number of heterogeneous sources. Today, coping with the variety of data requires novel paradigms for effectively accessing and querying information that adapt to the different types of sources, as well as declarative high-level languages to drive the data processing and data quality tasks.

BOREAL is a team working at the crossroads of *knowledge representation and reasoning* and *database theory*. The team focuses on the study of foundational and applied issues of *reasoning* in a context of *data variety*. More specifically, the team aims at deriving a better understanding of the logical fragments that are at the foundations of the frameworks used for exploiting corporate and Web data - and in particular *rule-based languages*. This will pave the way to novel automated-reasoning and graph-based techniques that can be put at service of data-centric applications exploiting heterogeneous and federated data. The team also aims at combining solid foundational and algorithmic work with software development and applications, with an emphasis on the field of agronomy.

## 3 Research program

The BOREAL team pursues a *knowledge-based data management* (KBDM) approach for tackling the grand challenges posed by data variety, with an important focus on the framework of *existential rules*. The idea of knowledge-based data management is to orchestrate the access to a complex information system made by federated databases through a three-layer architecture - also common to data-integration and ontology-based data access (OBDA). Under this prism, a set of *heterogeneous data sources* is connected to a *knowledge base* via a layer of *mappings*. The idea of KBDM is to define the business logic for data-centric applications at the knowledge base level, and then automatically translate the data-services towards the heterogeneous sources - through reasoning. This approach paves the way to a more principled use of complex information systems, with benefits to both data scientists, data curators, and administrators. What really characterizes the KBDM approach is the leveraging on formalized *domain-specific knowledge*, for abstracting on heterogeneous data and achieving high-quality of data-integration, and on expressive *rule-base languages* like existential rules (and extensions thereof), to drive the effective exploitation of data through reasoning.

Our project focuses on a set of topics related to knowledge-based data management, which we now describe.

**Foundations of rule languages** A great deal of the power of a KBDM system comes from its rule base. A prominent research direction for the team is the analysis and design of rule languages for reasoning on data. It is well understood that enriching a language with novel features can sensibly increase the complexity of the reasoning tasks. Our goal is hence to identify rules featuring decidable query answering and static analysis, and at the same time find good tradeoffs between their expressivity and complexity, so as to devise novel and practically useful rule-based frameworks.

**Algorithms and optimizations for query answering** Reasoning-driven data management needs optimization to effectively exploit large data. We target the design of efficient and scalable algorithms for query answering. Our goal is to devise novel hybrid approaches that combine materialization and virtualization strategies and account for the interplay between the components of the KBDM system (data, mappings, rules). Our ambition is also to build new bridges between knowledge representation and data-management by exploring the range of possibilities opened by the reuse of existing database technology to develop new reasoning systems.

**Fine-grained complexity of query answering** The query answering problem is at the heart of many reasoning tasks in KBDM. From a complexity analysis point of view, since the database to query can be voluminous, it is not always enough to know that a certain task can be done in polynomial time. Hence, an important goal for us is to study the fine-grained complexity (that is, to find the degree of the polynomial that bounds the number of operations required) as well as the enumeration complexity of

the query answering problem. The aim of this research direction is to obtain the theoretical knowledge required for practical query optimization.

**Architectures for knowledge-based data integration** The realm of possibilities in heterogeneous data integration leads to the offspring of a family of KBDM architectures, one for each applicative context. Our goal is to study architectures inspired from emerging practical use-cases, including *federations* of independent sources as well as *multi-level* architectures where KBDM systems are stacked to progressively distill information and achieve high-value data. We also focus on the type of mappings required to cope with heterogeneity, because data may differ along several dimensions such as its format, refinement, dynamicity, and certainty; this is required to build a unified view of a complex information system.

**Quality of knowledge-based data integration** Knowledge-based data management can result in high quality data for users and applications. Yet, they also need mechanisms to assist data curators to constantly evaluate and improve all of their components towards the ultimate goal of matching the desired data integration level. Our aim is to investigate explanation mechanisms able to justify answers to queries and to point out inconsistencies in the data. We are also interested in techniques for deriving, within a knowledge-base, equivalent formulations of queries that are expressed outside of it, at the source level; these are critical for the verification of mappings and rules.

## 4 Application domains

### 4.1 Agronomy and agroecology

Agronomy is more and more at the center of important debates around questions of environmental impact related to the practice of intensive agriculture, especially at large scale. Through our research collaborations with INRAE (National Research Institute for Agriculture, Food and Environment) and DFKI (German Institute for Artificial Intelligence) our goal is to contribute and to define new models, techniques, and applications, enabling a better exploitation of data generated in these fields so as to put it at the service of decision-making processes.

Agronomy is a strong expertise domain in the area of Montpellier. And indeed, BOREAL is a joint team with INRAE, and the team has established closed collaborations with two Montpellier research laboratories (UMR, “Unités Mixte de Recherche”), namely IATE and ABSys. These collaborations can also reach a larger extent, for example, in the context of the #DigitAg (Institute Convergences Agriculture Numérique, Section 8.2) our team participated to the joint Inria-INRAE “White Book” on digital agriculture which can be considered a manifesto of the current challenges posed by digital agriculture [22].

A major issue for IATE (Engineering of Agro-polymers and Emerging Technologies) is to model the transformation of products in agrifood chains (i.e., the chain of all processes leading from some raw material, such as plants, to the final products, including waste treatment). This modeling has several objectives. It provides better understanding of the processes from start to finish, which aids in decision making, with the aim of improving the quality of the products and decreasing the environmental impact (e.g., reducing waste, choosing right food packaging). There is a need for tools for making easier for data scientists to integrate and analyze the heterogeneous data resulting from agrifood chains.

A major issue for ABSys (Biodiversified Agrosystems) is the study of sustainable farming systems. It is now established that the restoration of sustainable farming systems requires the adoption of agroecological practices supporting the reintroduction of biodiversity in agroecosystems. Indeed, an agroecosystem should provide not only cash crops but also ecosystem services that support the durability of the farming systems itself. This leads to more complex agroecosystems including a higher number of plant species. There is thus a crucial need for tools that would assist users in the design of such new agroecosystems, from researchers in agronomy to agricultural advisors and farmers.

Beside INRAE, our team collaborates with two DFKI teams located in Osnabrück and Kaiserslautern in the context of a bilateral project Inria-DFKI (“R4Agri”, Section 8). From an applicative perspective, the major issue targeted by this project is the development of monitoring tools based on reasoning which can equip robotic or mechanic devices used in agricultural farms. This can be used to enhance the agricultural processes but also to enforce regulations, for instance by assessing that the spraying of

chemicals remains at a safe distance from river borders. In this context, there is a need for tools allowing one to interpret and analyze the number of types of sensor data that are generated.

## 5 Highlights of the year

- The team published 4 papers in top-tier venues (Core Ranking A\*) in the fields of knowledge representation and reasoning (KR, IJCAI) and database theory (ACM Transactions on Database Systems).
- Furthermore, we also published 2 papers in high-quality venues (Core Ranking A) in the fields of artificial intelligence (ECAI) and database theory (ICDT).

## 6 New software, platforms, open data

**Participants:** Jean-François Baget, Pierre Bisquert, Akira Charoensit, Michel Leclère, Marie-Laure Mugnier, Guillaume Pérution-Kihli, Florent Tornil, Federico Ulliana.

**InteGraal** is the main (Java) platform developed by the team to reason on heterogeneous data with existential rules. This year, InteGraal has been used in the context of our collaborations with INRAE as well as in the context of the bilateral project INRIA-DFKI “R4Agri” focusing on reasoning on agricultural data. InteGraal has been used in the context of the internship of Maksym Lytvynenko (L3 Univ. Montpellier) as well as for experimental analysis in the PhD thesis of Akira Charoensit (funded by INRIA-DFKI “R4Agri”), see Section 9.2.2. In general, InteGraal is a federating tool for the team, which follows a monthly-based software development cycle for its advancement. This year, to strengthen our project, we introduced B-Runner : a library for collaborative benchmarking on knowledge and rule-based reasoners. The motivation for this project was to systematize testing and experimental analysis, both on InteGraal and other reasoners.

### 6.1 New software

#### 6.1.1 InteGraal

**Name:** InteGraal : Knowledge-Representation and Reasoning for Data Integration

**Keywords:** Knowledge Bases, Data integration, Knowledge representation, Automated Reasoning, Heterogeneous Data, Knowledge Graphs

**Scientific Description:** InteGraal is a tool for integrating and reasoning on heterogeneous and federated data. The tool embodies algorithms and techniques developed at the crossroads between the fields of knowledge representation and reasoning and data management. From the historic point of view, this tool is the result of a complete re-engineering of the Graal tool, whose API and functionalities have been completely updated. Also, with respect to Graal, the tool is very much oriented towards data integration.

**Functional Description:** InteGraal has been designed in a modular way, in order to facilitate software reuse and extension. It should make it easy to test new scenarios and techniques, in particular by combining algorithms. The main features of Graal are the following: (1) internal storage to store data by using a SQL or RDF representation (Postgres, MySQL, HSQL, SQLite, Remote SPARQL endpoints, Local in-memory triplestores) as well as a native in-memory representation (2) data-integration capabilities for exploiting federated heterogeneous data-sources through mappings able to target systems such as SQL, RDF, and black-box (eg. Web-APIs) (3) algorithms for query-answering over heterogeneous and federated data based on query rewriting and/or forward chaining (or chase)



**Release Contributions:** 2024. Added reasoning with stratified negation. Advanced usability and features for mappings for integrating heterogeneous data. Extended the command line interface.

2023. Mappings for integrating heterogeneous data. Compilation-based query rewriting. Command line interface.

2022: First release, software deposit with Apache 2 licence.

2021: Functional specification, design and development of a major improved version of the tool. Started refactoring of the API, and of several modules for knowledge base representation, data storage, query answering and forward-chaining reasoning (chase). Started the development of new modules for handling heterogeneous data: mappings and federations.

**News of the Year:** This year we continued working on the major module including mappings for integrating heterogeneous data. We also added the support for reasoning with negation. Finally, we also improved our command line interface to interact with the platform. Moreover, the tool has been used in several internship and thesis of the team, as well as for collaborations with our research partners.

**URL:** <https://gitlab.inria.fr/rules/integraal>

**Publication:** <https://hal.science/INRIA2/lirimm-04304601v1>

**Contact:** Federico Ulliana

**Participants:** Jean-Francois Baget, Pierre Bisquert, Guillaume Perution-Kihli, Michel Leclère, Marie-Laure Mugnier, Florent Tornil, Federico Ulliana

### 6.1.2 TreeForce

**Keywords:** JSon, Databases, Knowledge Bases, Automated Reasoning, Rewriting, NoSQL, Data integration, Knowledge representation, Heterogeneous Data

**Scientific Description:** TreeForce is a java tool for reasoning on tree data. It leverages on query rewriting techniques and NoSQL document oriented key-value stores. This library can be seen as a general toolbox for implementing reasoning techniques tailored for tree-shaped data and rules. It is composed of two main modules. The first includes generic data structures and algorithms for trees and tree-automata. The second includes automata-based query rewriting techniques as well as efficient evaluation techniques for large sets of rewritings.

**Functional Description:** TreeForce is a java tool for reasoning on tree data. It leverages on query rewriting techniques and NoSQL document oriented key-value stores. This library can be seen as a general toolbox for implementing reasoning techniques tailored for tree-shaped data and rules. It is composed of two main modules. The first includes generic data structures and algorithms for trees and tree-automata. The second includes automata-based query rewriting techniques as well as efficient evaluation techniques for large sets of rewritings.

**Release Contributions:** 2023. ArangoDB wrapper. Code improvement. 2022. Novel instance-aware rewriting and evaluation algorithms. Introduced summarization, partitioning and parallelization techniques. 2021: First version of TreeForce. Automata for unordered tree languages. Automata-based query-rewriting algorithms. MongoDB wrapper.

**Contact:** Federico Ulliana

**Participants:** Olivier Rodriguez, Federico Ulliana

### 6.1.3 B-Runner

**Name:** B-Runner

**Keywords:** Benchmarking, Experimentation, Java, Automated Reasoning, Databases, Knowledge Graphs

**Scientific Description:** B-Runner is a Java tool for the conduction of experimental analysis.

**Functional Description:** Runner is a library for collaborative benchmarking on knowledge and rule-based reasoners. The motivation for this project was to systematize testing, both on on InteGraal and other reasoners. The goal of B-Runner is to enable benchmarking for reasoning tools with a small cost, high robustness, and repeatability guarantees. The tool can be used as a best-practice for realizing and communicating on experimental analysis.

**Release Contributions:** Core module for experiment conduction

**News of the Year:** First release of the tool.

**URL:** <https://gitlab.inria.fr/rules/brunner>

**Contact:** Federico Ulliana

**Participants:** Federico Ulliana, Pierre Bisquert, Florent Tornil, Renaud Colin, Quentin Yeche, Akira Charoensit

## 7 New results

Before presenting this year's results, we first introduce some general preliminary notions in Section 7.1 to provide context for the results discussed later in this section. Moreover, we provide a summary of this year's contributions in Section 7.2 and then discuss each of them in a dedicated section.

### 7.1 Preliminaires about Knowledge-Based Data Management with Existential Rules

This broad topic encompasses research areas such as ontology-mediated query answering (OMQA), data integration (DI), and ontology-based data access (OBDA) because of the expressivity of existential rule languages and the complexity of integration architectures it embraces.

**Existential rules.** Existential rules are first-order-logic formulas representing implications of the form  $\forall X \forall Y [Body(X, Y) \rightarrow \exists Z Head(X, Z)]$  where *Body* and *Head* are positive conjunctions of atoms without functional symbols, and *Head* can have existentially quantified variables. These rules allow one to model complex relationships over the domains of interest, and at the same time dispose of a value invention mechanism through existentially quantified variables. This makes them suitable for many data and knowledge tasks on both open and closed domains. As a result, existential rules are ubiquitous in many fields. They are used to model dependencies, schema mappings, and expressive queries in databases. They are used as ontological languages as a valid complement to Description Logics, and at the same time as a generalization of so-called Horn Description Logics which lay at the foundations of important Semantic Web standards.

**Rule-based query answering** Given a query  $Q$ , a database  $D$ , and a set of rules  $R$ , query answering asks to determine whether  $D, R \models Q$  (where  $\models$  denotes standard first-order logic entailment), that is if the query  $Q$  is a logical consequence of the knowledge base made by the database  $D$  and the rules  $R$ . In the field of knowledge representation and reasoning, rule-based query answering is studied for rules expressing ontologies and referred as ontology-mediated query answering (OMQA). Formalisms such as Description Logics and Existential Rules (a.k.a, Tuple-Generating-Dependencies, or Datalog $_{\pm}$ ) are typically targeted for expressing ontologies. Overall, the main emphasis of this topic is in the study of *rule languages* and the role they play in query answering.

**Rule-based query answering over heterogeneous and federated data** In this context, the problem formulation remains similar, however the database  $D$  is replaced by a more complex notion of *federation*  $(\mathcal{D}, \mathcal{M}, \mathcal{S})$  where  $\mathcal{D}$  is a collection of heterogeneous data sources,  $\mathcal{S}$  is a global integration schema, and  $\mathcal{M}$  is a set of mappings linking the datasources in  $\mathcal{D}$  to the global schema  $\mathcal{S}$ . This framework is at the foundations of data integration (DI) in databases and of ontology-based data access (OBDA) in knowledge representation and reasoning. OBDA focuses on global integration schemes and rules built on ontologies enabling query rewriting, while DI is more concerned with rules representing data-dependencies. Overall, both give more emphasis to *heterogeneous and federated data* in rule-based query answering.

**Reasoning strategies for query answering** The two prominent strategies for rule-based query answering are *materialization* (also known as saturation, or forward-chaining) and *virtualization* (also known as query rewriting, or backward chaining). Both can be seen as ways of reducing query answering (which involves reasoning) into classical query evaluation. Materialization amounts to storing the inferences enabled by rules, thereby obtaining an extended database, on which queries are evaluated. Query rewriting amounts to compiling relevant rules into the query, thereby obtaining a rewritten query (usually a union of queries), which is evaluated on the (unaltered) database. Both approaches have their own strengths, and at the basis of this duality is the fact that while materialization is independent of queries, rewriting is independent of the database. Hence, each strategy better suits certain applicative scenarios, and both can possibly be combined thereby resulting in hybrid approaches.

## 7.2 Contributions

This year, we studied a number of theoretical, algorithmic, and applied questions of knowledge-based data management and database theory. Our main contributions cover the following topics:

- Fine-grained complexity of database queries, notably optimizing retrieval in databases with direct access (Section 7.3), and repairing databases (Section 7.4).
- Foundations of reasoning with logical languages, notably existential rules (Section 7.5.1) and non-monotonic languages (Section 7.5.2).
- Systems for reasoning on integrated data, notably concerning benchmarking of reasoning systems (Section 7.6.1) and the application of OBDA in agroecology (Section 7.6.2).

Beside our main publications, it is worth noting that the team also supervised a number of student internships from ENS, MIT (Boston), and Montpellier University (Section 9.2.2) investigating other foundational and applied issues of reasoning and database theory. Finally, complementing methodological work, it is worth noting that we also pursued an important team effort in the development of tools for rule-based query answering over heterogeneous and federated such as InteGraal (see Section 6.1).

## 7.3 Fine-Grained Complexity of Database Queries

**Participants:** Nofar Carmeli, David Carral.

We studied the task of lexicographic direct access to the answers of a relational algebra query to a database. This amounts to simulating an array containing the answers of a join query sorted in a lexicographic order chosen by the user. Once this array is simulated, we can directly access a specific value. For instance, this would allow us to access the median of a certain attribute in a single step. We determined the preprocessing time needed to achieve polylogarithmic access time for all join queries and all lexicographical orders. To this end, we proposed a decomposition-based general algorithm for direct access on join queries. We then explored its optimality by proving lower bounds for the preprocessing time based on the hardness of a certain online Set-Disjointness problem, which shows that our algorithm's bounds are tight for all lexicographic orders on join queries. Then, we proved the hardness of Set-Disjointness based on the Zero-Clique Conjecture which is an established conjecture

from fine-grained complexity theory. Interestingly, while proving our lower bound, we were able to show that self-joins do not affect the complexity of direct access (up to logarithmic factors).

- Published at ACM Transactions on Database Systems [10] with Karl Bringman (Max Planck Institute for Informatics) and Stefan Menghel (CRIL)

We studied the same task of lexicographic direct access to query answers also for the more involved conjunctive queries with grouping and aggregation. Specifically, we investigated the ability to evaluate such queries by constructing in log-linear time a data structure that provides logarithmic-time direct access to the answers. For some common aggregate functions (e.g., min, max, count, sum), such a query can be phrased as an ordinary conjunctive query over a database annotated with a suitable commutative semiring. We showed that the past results about conjunctive queries without aggregation and annotation continue to hold for annotated databases, assuming that the annotation itself is not part of the lexicographic order. On the other hand, we showed infeasibility for the case of count-distinct that does not have any efficient representation as a commutative semiring. We then investigated the ability to include the aggregate and annotation outcome in the lexicographic order. Among the hardness results, standing out as tractable is the case of a semiring with an idempotent addition, such as those of min and max. Notably, this case captures also count-distinct over a logarithmic-size domain.

- Published the International Conference on Database Theory (ICDT-2024) [16] with Idan Eldar and Benny Kimelfeld (Israel Institute of Technology)

Apart from these results, Nofar Carmeli and David Carral co-supervised two students from ENS Paris and Lyon on topics related to database theory. More precisely, they studied the problem of checking if the answers of a given conjunctive query can be enumerated with linear preprocessing and constant delay. Nofar Carmeli also supervised a student from MIT (Boston) on the study of conjunctive queries that efficiently support direct-access with updates.

## 7.4 Repairing Databases

In addition to the study of the fine-grained complexity of query answering, we also investigated another relevant database issue, namely that of repairing data in the presence of soft constraints.

**Participants:** Nofar Carmeli.

Soft constraints penalize the database for every violation of every constraint, where the penalty is the cost (weight) of the constraint. A computational challenge is that of finding an optimal subset: a collection of database tuples that minimizes the total penalty when each tuple has a cost of being excluded. When the constraints are strict (i.e., have an infinite cost), this subset is a “cardinality repair” of an inconsistent database; in soft interpretations, this subset corresponds to a “most probable world” of a probabilistic database, or a “most likely intention” of a probabilistic unclean database. Within the class of functional dependencies, the complexity of finding a cardinality repair is thoroughly understood. Yet, very little is known about the complexity of finding an optimal subset for the more general soft semantics. In addition to general insights about the hardness and approximability of the problem, we presented algorithms for two special cases (and some generalizations thereof): a single functional dependency, and a bipartite matching. For these special cases, we also investigated the complexity of additional computational tasks that arise when the soft constraints are used as a means to represent a probabilistic database in the case of a probabilistic unclean database.

- Published at ACM Transactions on Database Systems [11] with Martin Grohe (Aachen University), Ester Livshits (Edinburgh University), Benny Kimelfeld and Muhammad Tibi (Israel Institute of Technology).

## 7.5 Foundations of Reasoning with Logical Languages

**Participants:** Jean-Francois Baget, Pierre Biquert, David Carral, Michel Leclère, Marie-Laure Mugnier, Guillaume Perution-Kihli, Akira Charoensit, Federico Ulliana.

### 7.5.1 Reasoning with Existential Rules

Ontology-based (or rule-based) query answering is a problem that takes as input an ontology  $R$  (i.e., a set of existential rules in our context), a set  $D$  of facts, and a Boolean conjunctive query (CQ)  $Q$ , and asks whether  $R, D \models Q$ , where  $\models$  denotes standard first-order logic entailment. This problem is undecidable in general, and a widely investigated approach to tackle it in some cases is query rewriting: given some “rule query”  $\langle R, Q \rangle$ , we compute a Boolean query  $Q_R$  such that, for any fact set  $D$ , it holds that  $R, D \models Q$  if and only if  $D \models Q_R$ . Insofar, previous work has mostly focused on output queries  $Q_R$  expressed as union of Boolean conjunctive queries (UCQs), and an effective algorithm that computes such a query  $Q_R$  whenever it exists has been proposed in the literature. However, UCQ-rewritability is not a very general notion and many real-world interesting rule queries do not admit UCQ-rewritings. This raises the question whether such a generic algorithm can be designed for a more expressive target language, such as datalog. We solved this question by the negative, by studying the difference between datalog-expressibility and datalog-rewritability. More precisely, we showed that query answering under datalog-expressible rule queries is undecidable.

- Published at the International Conference on Principles of Knowledge Representation and Reasoning (KR-2024) [13] with Michaël Thomazo (Inria VALDA)

We studied the problem of computing explanations for queries under datalog knowledge bases. One of the major benefits of rules (and symbolic AI methods in general) is explainability. When new knowledge is obtained via a reasoning process, it is possible to determine precisely all elements of the knowledge base that yield this knowledge. Typically, one would use a SAT (Boolean satisfiability problem) solver to compute the explanations. However, SAT-solving is computationally expensive, and as the knowledge base grows, the time required increases exponentially. In our work, we presented a method for filtering a datalog knowledge base to optimize the time used by a SAT solver. This is achieved by creating a hypergraph representing the grounded knowledge base and pruning the nodes that are not reachable from the fact that one wants to explain. The approach proved to be time-effective. Interestingly, one additional benefit of using this hypergraph is that it is possible to encode more information about the rules used in the reasoning process. By using an off-the-shelf group-SAT solver, this extra information allows us to find specific explanations that would be missed if we only considered facts.

- Published at the International Joint Conference on Rules and Reasoning (RuleML+RR 2024) [14].

### 7.5.2 Non-Monotonic Reasoning

Answer set programming (ASP) is a non-monotonic logic programming formalism used in various areas of artificial intelligence like combinatorial problem solving and knowledge representation and reasoning. It is known that enhancing ASP with function symbols makes basic reasoning problems highly undecidable. However, even in simple cases, state of the art reasoners, specifically those relying on a ground-and-solve approach, fail to produce a result. Therefore, we reconsidered consistency as a basic reasoning problem for ASP. We showed reductions that give an intuition for the high level of undecidability. These insights allow for a more fine-grained analysis where we characterize ASP programs as “frugal” and “non-proliferous”. For such programs, we were not only able to semi-decide consistency but we also proposed a grounding procedure that yields finite groundings on more ASP programs with the concept of “forbidden” facts.

- Published at the International Joint Conferences on Artificial Intelligence (IJCAI 2024) [17] and the International Workshop on Nonmonotonic Reasoning (NMR 2024) [18] with Lukas Gerlach (Technische Universität Dresden) and Markus Hecher (MIT, Boston).

We investigated the evaluation of queries in non-monotonic propositional logic. In order to avoid ambiguity and be efficient, the context in which a query is made can help to better target the relevant pieces of information from the knowledge base to be processed by the inference system. We studied a notion of "dynamical compartmentalization" where the knowledge base that will be used for reasoning is dynamically extracted from the original base. Compartmentalization is a selection of a sub-base which is done according to a function, called "refiner", and depending on this function some properties are satisfied. We introduced a particular "syntactic refiner" that uses a similarity symbol-based distance between a context (a multiset of variable symbols) and a formula of a knowledge base. We proved that the inference operator based on this refiner, called "contextual inference", satisfies a series of desirable axioms.

- Published at the European Conference on Artificial Intelligence (ECAI 2024) [15] with Florence Dupin de Saint-Cyr (IRIT).

## 7.6 Systems for Reasoning on Integrated Data

**Participants:** Jean-Francois Baget, Pierre Biquert, Akira Charoensit, Michel Leclère, Marie-Laure Mugnier, Guillaume Perution-Kihli, Florent Tornil, Federico Ulliana.

### 7.6.1 Benchmarking Reasoning Systems

Conducting experimental analysis on rule reasoners is a mainstream task for validating novel algorithms and systems. Nevertheless, providing robust, verifiable, and reproducible experiments can still raise a sensible challenge. We introduced B-Runner, an open library for *collaborative benchmarking* focusing on the deployment of articulate tests for *knowledge and rule-based systems* with low cost and high robustness. B-Runner reduces the benchmarking setup time while guaranteeing experiment repeatability. At the same time, it improves the scrutability of experimental protocols thereby enhancing their robustness as well as fairness of system comparisons.

- Published at the International Joint Conference on Rules and Reasoning (RuleML+RR 2024) [19] and at Bases de Données Avancées (BDA-2024) [20] with Quentin Yeche (Inria EVERGREEN).

### 7.6.2 Exploring the Application of Ontology Based Data Access in Agroecology

As part of Elie Najm's PhD thesis (funded by #DigitAg) and in collaboration with the INRAE ABSys research lab in agroecology, we have investigated issues related to the design of sustainable agroecosystems based on agroecology. Agroecology leads to more complex agroecosystems including a higher number of plant species, whether cash crops or service crops (i.e., plants providing various ecosystem services). The design of these new agrosystems requires to integrate various data as well as unstabilized scientific knowledge. We more specifically considered the issue of selecting service plant species according to their potential to provide ecosystem services. To tackle that issue, we adopted an approach based both on a formalized representation of domain knowledge and on the exploitation of available data, collected independently from the targeted application. More specifically, we rely on the one hand on recent scientific results in agronomy linking functional traits (i.e., measurable characteristics of plant species) to ecosystem services, and on the other hand on data about functional traits collected by the research community in ecology. The architecture of our system reprises that of rule-based query answering over heterogeneous and federated data we already presented. We provide a methodology to acquire scientific knowledge in the form of diagrams linked to data sources, as well as a formalization in a logical rule-based language. Importantly, our rules are independent from specific diagrams and data, to ensure better genericity and system's ability to evolve. We carried out an experimental evaluation of our system on the use case of vine grassing, i.e., installing herbaceous service plants in vineyards, which showed that very satisfactory results can be obtained as long as the proportion of missing data values is not too high.

- Published in *Computers and Electronics in Agriculture* [12] as a joint work with Christian Gary, Raphael Metral, and Léo Garcia (INRAE).

Following this work, and as part of Guillaume Pérution's postdoc, we have performed an extensive rework of this application with the aim of improving the quality of both the code (bug fixes, extensibility, genericity) and the modeling (declarativity, genericity). In addition, the new implementation essentially relies on Integraal, which has also led to the evolution of Integraal's functionalities.

## 8 Partnerships and cooperations

### 8.1 International research visitors

#### 8.1.1 Visits of international scientists

**Other international visits to the team** Except for the first person listed in this section (i.e., Markus Hecher), every other researcher visited our group in the context of the Formal Logic At Montpellier AND database Theory (FLAMANT) workshop that we organized in February 2024.

##### **Markus Hecher**

**Status** post-Doc

**Institution of origin:** Massachussets Institute of Technology

**Country:** United States of America

**Dates:** from the 06/01/2024 (Monday) to the 10/01/2024 (Friday)

**Context of the visit:** meet the group in preparation for the Inria 2024 CRCN/ISFPO interview

**Mobility program/type of mobility:** research stay

##### **Lukas Gerlach**

**Status** PhD

**Institution of origin:** TU Dresden

**Country:** Germany

**Dates:** from the 26.02.24 (Monday) to the 08.03.24 (Friday)

**Context of the visit:** FLAMANT 2024

**Mobility program/type of mobility:** research stay

##### **Lucia Gomez Alvarez**

**Status** researcher

**Institution of origin:** Gomez Alvarez

**Country:** France

**Dates:** from the 26.02.24 (Monday) to the 08.03.24 (Friday)

**Context of the visit:** FLAMANT 2024

**Mobility program/type of mobility:** research stay

**Philipp Hanisch****Status** PhD**Institution of origin:** TU Dresden**Country:** Germany**Dates:** from the 27.02.24 (Tuesday) to the 05.03.24 (Tuesday)**Context of the visit:** FLAMANT 2024**Mobility program/type of mobility:** research stay**Markus Krötzsch****Status** researcher and professor**Institution of origin:** TU Dresden**Country:** Germany**Dates:** from the 27.02.24 (Tuesday) to the 29.02.24 (Thursday)**Context of the visit:** FLAMANT 2024**Mobility program/type of mobility:** research stay**Lucas Larroque****Status** PhD**Institution of origin:** Valda team at Inria Paris**Country:** France**Dates:** from the 26.02.24 (Monday) to the 01.03.24 (March)**Context of the visit:** FLAMANT 2024**Mobility program/type of mobility:** research stay**Timothy Stephen Lyon****Status** post-Doc**Institution of origin:** TU Dresden**Country:** Germany**Dates:** from the 26.02.24 (Monday) to the 01.03.24 (Friday)**Context of the visit:** FLAMANT 2024**Mobility program/type of mobility:** research stay**Piotr Ostropolski-Nalewaja****Status** post-Doc**Institution of origin:** TU Dresden**Country:** Germany**Dates:** from the 26.02.24 (Monday) to the 15.03.24 (Friday)**Context of the visit:** FLAMANT 2024**Mobility program/type of mobility:** research stay



**Sebastian Rudolph**

**Status** researcher and professor

**Institution of origin:** TU Dresden

**Country:** Germany

**Dates:** from the 26.02.24 (Monday) to the 08.03.24 (Friday)

**Context of the visit:** FLAMANT 2024

**Mobility program/type of mobility:** research stay

**Michaël Thomazo**

**Status** researcher

**Institution of origin:** Valda team at Inria Paris

**Country:** France

**Dates:** from the 05.03.24 (Tuesday) to the 08.03.24 (Friday)

**Context of the visit:** FLAMANT 2024

**Mobility program/type of mobility:** research stay

**8.1.2 Visits to international teams****Nofar Carmeli**

**Visited institution:** Leipzig University

**Country:** Germany

**Dates:** 22/04/2024 - 26/04/2024

**Context of the visit:** Collaboration on evaluating conjunctive queries in the presence of tuple generating dependencies

**Mobility program/type of mobility:** research stay

**8.1.3 Visits to national teams****Nofar Carmeli**

**Visited institution:** CRIL (Centre de Recherche en Informatique de Lens)

**Dates:** 21/05/2024 - 24/05/2024

**Context of the visit:** Collaboration on direct access to conjunctive query answers in the presence of functional dependencies

**Mobility program/type of mobility:** research stay

**David Carral**

**Visited institution:** Valda Inria research team at ENS Paris

**Dates:** 16/12/2024 - 19/12/2024

**Context of the visit:** Collaborating with Lucas Larroque and Michaël Thomazo

**Mobility program/type of mobility:** research stay

## 8.2 National initiatives

### CQFD (ANR PRC, Jan. 2019-Dec. 2024)

**Participants:** Jean-François Baget, Pierre Bisquert, Nofar Carmeli, David Carral, Michel Leclère, Marie-Laure Mugnier, Guillaume Pérution-Kihli, Florent Tornil, Federico Ulliana.

CQFD (Complex ontological Queries over Federated heterogeneous Data), coordinated by Federico Ulliana (BOREAL), involves participants from Inria Saclay (CEDAR team), Inria Paris (VALDA team), Inria Nord Europe (SPIRALS team), IRISA, LIG, LTCI, and LaBRI. The aim of this project is to tackle two crucial challenges in OMQA (Ontology Mediated Query Answering), namely, heterogeneity, that is, the possibility to deal with multiple types of data-sources and database management systems, and federation, that is, the possibility of cross-querying a collection of heterogeneous datasources. By featuring 8 different partners in France, this project aims at consolidating a national community of researchers around the OMQA issue.

[www.lirmm.fr/cqfd/](http://www.lirmm.fr/cqfd/)

### Convergence institute #DigitAg (2017-2026)

**Participants:** Jean-François Baget, Marie-Laure Mugnier, Federico Ulliana.

Located in Montpellier, #DigitAg (for Digital Agriculture) gathers 17 founding members: research institutes, including Inria, the University of Montpellier and higher-education institutes in agronomy, transfer structures and companies. Its objective is to support the development of digital agriculture. BOREAL is involved in this project on the issues of designing data and knowledge management systems adapted to agricultural information systems, and of developing methods for integrating different types of information and knowledge (generated from data, experts, models). A PhD thesis (Elie Najm, 2019-2022) investigated knowledge representation and reasoning for the design of new agroecological systems, in collaboration with the research laboratory ABSys - Biodiversified Agrosystems (formerly UMR SYSTEM).

[www.hdigitag.fr/eng/](http://www.hdigitag.fr/eng/)

## 9 Dissemination

### 9.1 Promoting scientific activities

**Participants:** Nofar Carmeli, David Carral, Michel Leclère, Marie-Laure Mugnier, Federico Ulliana.

#### 9.1.1 Scientific events: organisation

##### General chair, scientific chair

- Area chair of the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR 2024): Marie-Laure Mugnier
- Proceedings chair of the Symposium on Principles of Database Systems (PODS 2025): Nofar Carmeli
- Test-of-time award committee member of the International Conference on Database Theory (ICDT 2024): Nofar Carmeli

### 9.1.2 Scientific events: selection

#### Member of the conference program committees (PC)

- PC member of the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR 2024): David Carral (Main track) and Marie-Laure Mugnier (Recently Published Research track)
- PC member of the Symposium on Theoretical Aspects of Computer Science (STACS 2025): Nofar Carmeli
- PC member of the Symposium on Principles of Database Systems (PODS 2024): Nofar Carmeli
- PC member of the Ingénierie des Connaissances 2024 (IC 2024): Michel Leclère
- PC member of the Gestion de Données – Principes, Technologies et Applications (BDA 2024): Federico Ulliana
- PC member of the Workshop on Enumeration Problems and Applications (WEPA 2024): Nofar Carmeli

#### Reviewer

- Reviewer at the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR 2024): David Carral (Main track) and Marie-Laure Mugnier (Recently Published Research track)
- Reviewer at the International Conference on Database Theory (ICDT 2024): Nofar Carmeli
- Reviewer at the Ingénierie des Connaissances 2024 (IC 2024): Michel Leclère
- Reviewer at the Gestion de Données – Principes, Technologies et Applications (BDA 2024): Federico Ulliana

### 9.1.3 Journal

#### Reviewer - reviewing activities

- Reviewer at the Knowledge-Based Systems Journal of Elsevier: Marie-Laure Mugnier
- Reviewer at the Artificial Intelligence Journal (AIJ) of Elsevier: David Carral
- Reviewer at the ACM Transactions on Database Systems (TODS): Nofar Carmeli
- Reviewer at Logical Methods in Computer Science (LMCS): Nofar Carmeli

### 9.1.4 Scientific expertise

- Marie-Laure Mugnier was a member of the selection committee for a Professorship position at the University of Montpellier.

### 9.1.5 Research administration

- President of the “Section 27 Committee” (Computer Science) of the University of Montpellier (since July 2021): Marie-Laure Mugnier
- Member of the Council and Human Resources Commission of the Scientific Pole MIPS (Mathematics Informatics Physics and Systems) of the University of Montpellier (since its creation): Marie-Laure Mugnier

## 9.2 Teaching - Supervision - Juries

**Participants:** Jean-François Baget, Pierre Bisquert, David Carral, Nofar Carmeli, Michel Leclère, Marie-Laure Mugnier, Guillaume Pérution-Kihli, Federico Ulliana.

### 9.2.1 Teaching

- Michel Leclère and Marie-Laure Mugnier teach at the Computer Science department of the Science Faculty. They are in charge of courses in Programming and Logics (Licence), as well as Symbolic Artificial Intelligence, Semantic Management of Data, and Theory of Data and Knowledge Bases (Master).
- Concerning full-time researchers in 2024, Jean-François Baget, Nofar Carmeli, David Carral, and Federico Ulliana (on secondment from Montpellier University) taught in the Computer Science Master about topics including Theory of Data and Knowledge Bases, Datawarehouses, Big-Data and NoSQL systems.
- Guillaume Pérution-Kihli has been coaching students from University of Montpellier for the SWERC (Southwestern Europe Regional Contest)<sup>1</sup> competition since 2020.

### 9.2.2 Supervision

- Pierre Bisquert, David Carral, and Federico Ulliana continue to co-supervise Akira Charoensit, who is a PhD student based in our group and is about to finish her second year.
- David Carral continues to co-supervise Lucas Larroque with Michaël Thomazo. Lucas is a PhD student based at the ENS Paris who has recently started the second year of his degree.
- Nofar Carmeli and David Carral supervised Noah Collinet, who was an M2 student of the University of Montpellier, during her master's project.
- Michel Leclère and Jean-François Baget supervised Maksym Lytvynenko, who is a master's student at the University of Montpellier. Namely, they worked together on an M1 research project.
- Nofar Carmeli and David Carral supervised François Colin De Verdier and Clément Rouvroy, who are two L3 students from the ENS of Lyon and Paris, respectively. They visited our group to do a research internship on topics related with database theory that lasted 6 weeks.
- Nofar Carmeli supervised Nicolas Valayannopoulos-Akrivou, who is an L2 student from MIT on a research internship on topics related with database theory.

### 9.2.3 Juries

#### PhD / HDR juries

- Member of the PhD committee for Ali Ballout at Nice University in June 2024: Federico Ulliana

## 9.3 Popularization

**Participants:** Michel Chein, Marie-Laure Mugnier.

---

<sup>1</sup>[www.swerc.eu](http://www.swerc.eu)

### 9.3.1 Productions (articles, videos, podcasts, serious games, ...)

- Michel Chein wrote the article "Et si certains scientifiques, en Intelligence Artificielle, avaient une responsabilité dans le défiance vis à vis de la raison ?", which appeared in the 54th volume of the Bulletin de l'Académie des Sciences et Lettres de Montpellier. For more info, see this [link](#).
- Michel Chein wrote the article "Penser l'Intelligence Artificielle (les agents conversationnels)", which will appear in the 55th volume of the Bulletin de l'Académie des Sciences et Lettres de Montpellier. For more info, see this [link](#).
- Michel Chein wrote the article "Émergence de l'Informatique à Montpellier" in the brochure for the "45 ans du Centre National de Calcul Scientifique à Montpellier (CINES)". This publication is an extension of his previous article "Sur la science informatique et son installation à Montpellier", which was published in the 48th volumen of the Bulletin de l'Académie des Sciences et Lettres de Montpellier. For more info, see this [link](#).

### 9.3.2 Participation in Live events

- Marie-Laure Mugnier gave an interview about symbolic and hybrid AI to the magazine LUM of the university of Montpellier. This led to an article in the special issue about AI published on 07/10/2024 (LUM n°22) and to a podcast. For more info, see this [link](#).

## 10 Scientific production

### 10.1 Major publications

- [1] K. Bringmann, N. Carmeli and S. Mengel. 'Tight Fine-Grained Bounds for Direct Access on Join Queries'. In: *ACM Transactions on Database Systems* (6th Dec. 2024). DOI: [10.1145/3707448](https://doi.org/10.1145/3707448). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04834303>.
- [2] N. Carmeli, M. Grohe, B. Kimelfeld, E. Livshits and M. Tibi. 'Database Repairing with Soft Functional Dependencies'. In: *ACM Transactions on Database Systems* 49.2 (2024), pp. 1–34/8. DOI: [10.1145/3651156](https://doi.org/10.1145/3651156). URL: <https://hal.science/hal-04617777>.
- [3] D. Carral, L. Larroque and M. Thomazo. 'Ontology-Based Query Answering over Datalog-Expressible Rule Sets is Undecidable'. In: *KR 2024 - 21st International Conference on Principles of Knowledge Representation and Reasoning*. Hanoi, Vietnam, 2nd Nov. 2024. URL: <https://hal.science/hal-04710719>.
- [4] A. Charoensit, D. Carral, P. Bisquert, L. Rouquette and F. Ulliana. 'Rule-aware Datalog Fact Explanation Using Group-SAT Solver'. In: *RuleML+RR'24: Companion Proceedings of the 8th International Joint Conference on Rules and Reasoning*. RuleML+RR 2024 - 8th International Joint Conference on Rules and Reasoning. Bucarest, Romania, 16th Sept. 2024. URL: <https://hal.science/hal-04706324>.
- [5] F. Dupin de Saint-Cyr and P. Bisquert. 'The Form and the Content: Non-Monotonic Reasoning with Syntactic Contextual Filtering'. In: *ECAI 2024*. ECAI 2024 - 27th European Conference on Artificial Intelligence. Santiago de Compostela, Spain: IOS Press, Aug. 2024, à paraître. URL: <https://hal.science/hal-04669429>.
- [6] I. Eldar, N. Carmeli and B. Kimelfeld. 'Direct Access for Answers to Conjunctive Queries with Aggregation'. In: *ICDT 2024 - 27th International Conference on Database Theory*. Leibniz International Proceedings in Informatics. Paestum, Italy: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 4:1–4:20. DOI: [10.4230/LIPIcs.ICDT.2024.4](https://doi.org/10.4230/LIPIcs.ICDT.2024.4). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04617768>.
- [7] L. Gerlach, D. Carral and M. Hecher. 'Finite Groundings for ASP with Functions: A Journey through Consistency'. In: *IJCAI 2024 - 33rd International Joint Conference on Artificial Intelligence*. Jeju, South Korea, 2024, pp. 3386–3394. DOI: [10.24963/ijcai.2024/375](https://doi.org/10.24963/ijcai.2024/375). URL: <https://hal.science/hal-04710694>.

- [8] E. Najm, M.-L. Mugnier, C. Gary, J.-F. Baget, R. Métral and L. Garcia. ‘Integrating data and knowledge to support the selection of service plant species in agroecology’. In: *Computers and Electronics in Agriculture* 217 (Feb. 2024), p. 108594. DOI: [10.1016/j.compag.2023.108594](https://doi.org/10.1016/j.compag.2023.108594). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04393663>.
- [9] F. Ulliana, P. Bisquert, A. Charoensit, R. Colin, F. Tornil and Q. Yeche. ‘Collaborative Benchmarking Rule-Reasoners with B-Runner’. In: *RuleML+RR 2024 - 8th International Joint Conference on Rules and Reasoning*. Vol. 15183. Lecture Notes in Computer Science. Bucarest, Romania, 11th Sept. 2024, pp. 21–31. DOI: [10.1007/978-3-031-72407-7\\_3](https://doi.org/10.1007/978-3-031-72407-7_3). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04670543>.

## 10.2 Publications of the year

### International journals

- [10] K. Bringmann, N. Carmeli and S. Mengel. ‘Tight Fine-Grained Bounds for Direct Access on Join Queries’. In: *ACM Transactions on Database Systems* (6th Dec. 2024). DOI: [10.1145/3707448](https://doi.org/10.1145/3707448). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04834303> (cit. on p. 9).
- [11] N. Carmeli, M. Grohe, B. Kimelfeld, E. Livshits and M. Tibi. ‘Database Repairing with Soft Functional Dependencies’. In: *ACM Transactions on Database Systems* 49.2 (2024), pp. 1–34/8. DOI: [10.1145/3651156](https://doi.org/10.1145/3651156). URL: <https://hal.science/hal-04617777> (cit. on p. 9).
- [12] E. Najm, M.-L. Mugnier, C. Gary, J.-F. Baget, R. Métral and L. Garcia. ‘Integrating data and knowledge to support the selection of service plant species in agroecology’. In: *Computers and Electronics in Agriculture* 217 (Feb. 2024), p. 108594. DOI: [10.1016/j.compag.2023.108594](https://doi.org/10.1016/j.compag.2023.108594). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04393663> (cit. on p. 12).

### International peer-reviewed conferences

- [13] D. Carral, L. Larroque and M. Thomazo. ‘Ontology-Based Query Answering over Datalog-Expressible Rule Sets is Undecidable’. In: *KR 2024 - 21st International Conference on Principles of Knowledge Representation and Reasoning*. Hanoi, Vietnam, 2nd Nov. 2024. URL: <https://hal.science/hal-04710719> (cit. on p. 10).
- [14] A. Charoensit, D. Carral, P. Bisquert, L. Rouquette and F. Ulliana. ‘Rule-aware Datalog Fact Explanation Using Group-SAT Solver’. In: *RuleML+RR’24: Companion Proceedings of the 8th International Joint Conference on Rules and Reasoning*. RuleML+RR 2024 - 8th International Joint Conference on Rules and Reasoning. Bucarest, Romania, 16th Sept. 2024. URL: <https://hal.science/hal-04706324> (cit. on p. 10).
- [15] F. Dupin de Saint-Cyr and P. Bisquert. ‘The Form and the Content: Non-Monotonic Reasoning with Syntactic Contextual Filtering’. In: *ECAI 2024*. ECAI 2024 - 27th European Conference on Artificial Intelligence. Santiago de Compostela, Spain: IOS Press, Aug. 2024, à paraître. URL: <https://hal.science/hal-04669429> (cit. on p. 11).
- [16] I. Eldar, N. Carmeli and B. Kimelfeld. ‘Direct Access for Answers to Conjunctive Queries with Aggregation’. In: *ICDT 2024 - 27th International Conference on Database Theory*. Leibniz International Proceedings in Informatics. Paestum, Italy: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 4:1–4:20. DOI: [10.4230/LIPIcs.ICDT.2024.4](https://doi.org/10.4230/LIPIcs.ICDT.2024.4). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04617768> (cit. on p. 9).
- [17] L. Gerlach, D. Carral and M. Hecher. ‘Finite Groundings for ASP with Functions: A Journey through Consistency’. In: *IJCAI 2024 - 33rd International Joint Conference on Artificial Intelligence*. Jeju, South Korea, 2024, pp. 3386–3394. DOI: [10.24963/ijcai.2024/375](https://doi.org/10.24963/ijcai.2024/375). URL: <https://hal.science/hal-04710694> (cit. on p. 10).

- [18] L. Gerlach, D. Carral and M. Hecher. ‘Finite Groundings for ASP with Functions: A Journey through Consistency (Extended Abstract)’. In: *CEUR Workshop Proceedings*. NMR 2024 - 22nd International Workshop on Nonmonotonic Reasoning. Vol. CEUR-3835. Nonmonotonic Reasoning 2024 : Proceedings of the 22nd International Workshop on Nonmonotonic Reasoning (NMR 2024). Hanoi, Vietnam, 2nd Nov. 2024, pp. 183–186. URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04823986> (cit. on p. 10).
- [19] F. Ulliana, P. Bisquert, A. Charoensit, R. Colin, F. Tornil and Q. Yeche. ‘Collaborative Benchmarking Rule-Reasoners with B-Runner’. In: *RuleML+RR 2024 - 8th International Joint Conference on Rules and Reasoning*. Vol. 15183. Lecture Notes in Computer Science. Bucarest, Romania, 11th Sept. 2024, pp. 21–31. DOI: [10.1007/978-3-031-72407-7\\_3](https://doi.org/10.1007/978-3-031-72407-7_3). URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04670543> (cit. on p. 11).

#### National peer-reviewed Conferences

- [20] F. Ulliana, P. Bisquert, A. Charoensit, R. Colin, F. Tornil and Q. Yeche. ‘Collaborative Benchmarking Rule-Reasoners with B-Runner’. In: *BDA 2024 - 40e Conférence sur la Gestion de Données – Principes, Technologies et Applications*. Orleans, France, 21st Oct. 2024. URL: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-04646842> (cit. on p. 11).

#### Reports & preprints

- [21] P. Bisquert, P. Buche, A. Hecham and M. Koptelov. *Justified collective decision-making in agri-food systems: an approach based on computational social choice, argumentation methods and tools (Supplementary Material)*. 10th June 2024. URL: <https://hal.science/hal-04607020>.

### 10.3 Cited publications

- [22] N. Mitton, L. Brossard, T. Bouadi, F. Garcia, R. Gautron, N. Hilgert, D. Ienco, C. Largouët, E. Lutton, V. Masson, R. Martin-Clouaire, M.-L. Mugnier, P. Neveu, P. Preux, H. Raynal, C. Roussey, A. Termier and V. Bellon Maurel. ‘Foundations and state of the art’. In: *Agriculture and Digital Technology: Getting the most out of digital technology to contribute to the transition to sustainable agriculture and food systems*. White book Inria 6. Acknowledgements (contribution, proofreading, editing) – Isabelle Piot-Lepetit. INRIA, 2022, pp. 30–75. URL: <https://hal.inrae.fr/hal-03609470> (cit. on p. 4).