

# Reasoning with hard and soft constraints to repair and query inconsistent data

## Basic Information

**Duration:** 4-6 months, ideally between February - July 2023

**Supervisors:** Meghyn Bienvenu (LaBRI, Bordeaux), Camille Bourgaux (DI ENS, Paris)

**Host lab:** LaBRI (Laboratoire Bordelais de Recherche en Informatique)

## Context

It is widely acknowledged that real-world data is plagued with quality issues, such as inconsistencies (due to false or outdated information) and incompleteness (missing information). A prominent approach to handling inconsistent data is to use declaratively specified knowledge in the form of logical constraints to identify errors and define a space of possible "clean" databases (called repairs). By reasoning over these repairs, it is possible to obtain meaningful query answers from contradictory data and to classify answers according to their confidence. This general approach was first explored for relational databases equipped with integrity constraints [1,2]. More recently, it has been extended to ontology-enriched databases (see survey [3] and references therein), where the domain knowledge provided by the ontology not only serves to identify errors but also to infer missing information [4,5] (helping to tackle the incompleteness issue).

The Master's internship is part of the INTENDED Chair on Artificial Intelligence, whose aim is to develop intelligent, knowledge-based methods for handling imperfect data. The chair is funded by the French National Research Agency (ANR) and the University of Bordeaux.

***The internship can lead to a funded PhD position within the INTENDED project.***

## Research Topic

In both the pure database and ontology settings, logical constraints are typically assumed to be hard (absolute), which means repairs must not contain any constraint violations. However, there exist natural constraints that are mostly true but nevertheless admit rare exceptions. For example, while we may assert (absolutely) that a person has precisely one birthplace, we cannot reasonably state that a person has only one residential address, even though this is true for most people. Clearly, such soft constraints can provide valuable guidance for handling inconsistent data (suggesting potential errors), but their integration into repair-based approaches remains little explored.

The aim of this internship is to investigate notions of repairs in the ontology setting that allow for both hard and soft constraints. After reviewing the state-of-the-art [6,7] and formally defining suitable repair notions, the main task will be to study the computational properties of such repairs. This will involve pinpointing the computational complexity of the main computational tasks: recognizing and generating repairs, and answering queries

under repair-based inconsistency-tolerant semantics. It is likely that many of these problems will prove intractable, so it will be necessary to find ways to circumvent the high complexity, either by identifying relevant tractable subcases or devising pragmatic algorithms that can be expected to behave well in practice.

## Candidate Profile

This internship is best suited to candidates who have prior experience with knowledge representation and reasoning (especially: description logics, non-monotonic reasoning), database theory, or Semantic Web (ontologies).

Candidates must demonstrate familiarity with propositional and first-order logic and basic notions of computational complexity. Strong English skills are desired.

## How to Apply

Candidates for the Master's internship should contact the two supervisors by email:

- Meghyn Bienvenu ([meghyn.bienvenu@labri.fr](mailto:meghyn.bienvenu@labri.fr))
- Camille Bourgaux ([camille.bourgaux@ens.fr](mailto:camille.bourgaux@ens.fr))

The email should include a CV, course transcripts (last two years), and a short description of how the internship topic relates to their prior experience and research interests.

## Bibliography

- [1] Marcelo Arenas, Leopoldo E. Bertossi, Jan Chomicki: Consistent Query Answers in Inconsistent Databases. In Proceedings of the 18th ACM Symposium on Principles of Database Systems (PODS), pages 68-79, 1999.
- [2] Leopoldo E. Bertossi: Database Repairing and Consistent Query Answering. Synthesis Lectures on Data Management, Morgan & Claypool Publishers, 2011.
- [3] M. Bienvenu, C. Bourgaux: Inconsistency-Tolerant Querying of Description Logic Knowledge Bases. Tutorial Notes of the 12th International Reasoning Web Summer School (LNCS 9885), pages 156- 202, 2016.
- [4] M. Bienvenu, M. Ortiz: Ontology-Mediated Query Answering with Data-Tractable Description Logics. Tutorial Notes of the 11th International Reasoning Web Summer School (LNCS 9203), pages 218-307, 2015.
- [5] G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati, and M. Zakharyashev. Ontology-based data access: A survey. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), pages 5511–5519, 2018.
- [6] Nofar Carmeli, Martin Grohe, Benny Kimelfeld, Ester Livshits, Muhammad Tibi: Database Repairing with Soft Functional Dependencies. In Proceedings of the 24th International Conference on Database Theory (ICDT), pages 16:1-16:17, 2021.
- [7] Thomas Eiter, Thomas Lukasiewicz, Livia Predoiu: Generalized Consistent Query Answering under Existential Rules. In Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR), pages 359-368, 2016.