Gianluca Cima

Curriculum Vitae - ai fini della pubblicazione

Place: Bordeaux, France
Date: 31/01/2023

I - General Information

Full Name	Gianluca Cima
Citizenship	Italian
Spoken Languages	Italian (Native Speaker) & English (Proficient)

II - Education

Type	Year	Institution	Notes (Degree, Experience, . . .)
University Graduation BSc	2014	Sapienza University of Rome	Ingegneria Informatica e Automatica - 110/110 *Cum Laude*
University Graduation MSc	2016	Sapienza University of Rome	Engineering in Computer Science - 110/110 *Cum Laude*. Completion of the *Excellence Path*
PhD	2020	Sapienza University of Rome	Engineering in Computer Science - Doctor Europaeus *Cum Laude*. Visiting PhD student at University of Oxford from March 2018 to August 2018

III - Academic Appointments

Start	End	Institution	Position
01/02/2020	31/01/2021	Sapienza University of Rome	Temporary Research Fellow
12/02/2021	Current position	CNRS - University of Bordeaux	Postdoctoral Researcher

IV - Teaching Experience

Year	Institution	Lecture/Course
2021-2022	University of Bergamo	5 hours of the 10 hours PhD course "Advanced Data Management 2"

2020-2021	Sapienza University of Rome (sede di Latina)	3 credits of the 9 credits course "Basi di Dati"
2020-2021	Sapienza University of Rome (sede di Latina)	6 credits of the 9 credits course "Progettazione del Software"
2019-2020	Sapienza University of Rome	40 hours of tutoring for the course "Fondamenti di Informatica 1". I have managed and prepared most of the lab activities
2018-2019	Sapienza University of Rome	30 hours of tutoring for the course "Basi di Dati". I have managed and prepared most of the lab activities
2018-2019	Sapienza University of Rome	20 hours of tutoring for the course "Linguaggi e Tecnologie per il Web". I have managed and prepared most of the lab activities
2017-2018	Sapienza University of Rome	40 hours of tutoring for the course "Basi di Dati". I have managed and prepared most of the lab activities

V - Society memberberships, Awards and Honors

My PhD thesis, **Abstraction in Ontology-based Data Management**, has been identified as being of outstanding quality during the evaluation process of the *2020 EurAI Doctoral Dissertation Award*. In light of this, my PhD thesis has been invited to be published in the subseries *Dissertations in Artificial Intelligence* of the *Frontiers in Artificial Intelligence and Applications (FAIA)* book series. The book can be found in the IOS Press Catalogue at the following link: https://www.iospress.com/catalog/books/abstraction-in-ontology-based-data-management.

VI - Research Projects

I have been involved in the following projects:

- *INTENDED: Intelligent handling of imperfect information.* The huge wealth of data available nowadays holds tremendous potential to improve our lives, whether it be by advancing scientific knowledge, improving patient care, or supporting more informed policymaking. However, obtaining relevant and reliable information from real-world data is difficult due both to the need to integrate data from heterogeneous sources and the ubiquity of data quality issues (e.g., missing or incorrect facts). The aim of the project is to develop intelligent, knowledge-based methods for handling imperfect data, thereby enabling confident and informed decision making. My contribution in this project concerns the development of new techniques to handle data quality issues.

- *HOPE - High quality Open data Publishing and Enrichment*. HOPE aims at overcoming the main technical problems that current open data solutions suffer from, by developing a methodology and associated tools for a new way of producing, publishing, maintaining, accessing and exploiting privacy-preserving open data. The envisioned result of the project is a complete web-based semantic open data manager that an organization can use for governing the whole lifecycle of its open data, and the final users can access for effectively consuming the information provided by the organization. My contribution in this project is in the help of the management of one of the workpackages.

- Collaboration with *ACI Informatica*, the in-house IT company of ACI - Automobile Club of Italy, dealing in particular with all aspects of mobility at the national level. The collaboration has the goal of integrating all national data sources regarding the register of vehicles and their taxation, as well as to design and develop an Open Data portal. My contribution in this project has been designing a portion of the mapping component between the existing ACI source data and the ontology that models the ACI domain.

VII - Research Activities

My research topic is located in the *Artificial Intelligence (AI)* field, and in particular on *Knowledge Representation and Automated Reasoning*. My principal focus is on the *Ontology-Based Data Management (OBDM)* paradigm. In this paradigm, an ontology provides an high-level, logic-based conceptual integrated view of a domain of interest, linked through declarative mapping assertions to existing and autonomous data sources. More formally, an OBDM specification is a triple $\langle O, S, M \rangle$ of ontology O, source schema S, and mapping M, and an OBDM system is a pair $\langle \langle O, S, M \rangle, D \rangle$, where D is a database for S, also called source database. The major contributions provided by my research, carried out together with my collaborators, can be summarized as follows:

Keyword	Brief Description
Abstraction	I have introduced the notion of *Abstraction* in OBDM, a new reasoning task whose goal is to automatically associate formal semantic descriptions to data services through ontologies. Given a data service expressed as a query q_S over the source schema S of an OBDM specification, Abstraction is the task of finding a query q_O over the ontology O, expressed in some specific target query language of interest, that captures the semantics of q_S with respect to all the axioms composing the OBDM specification, independently from the current source database.

Controlled Query Evaluation	I have worked on *Controlled Query Evaluation* in the OBDM context, introducing a novel framework, called *Policy-Protected Ontology-Based Data Management (PPOBDM)*, which extends OBDM with a data protection policy component, specified in terms of logical statements over the ontology. Such data protection policy declares confidential information that must not be revealed to the users, and a function, called *censor*, alters standard query answering to users' queries in order to avoid the disclosure of protected data and enforce confidentiality.
Separability	I have worked on *separability* in the OBDM context. Given an OBDM system $\langle\langle O, S, M\rangle, D\rangle$ and two datasets λ^+ and λ^-, i.e. two sets of tuples of constants, representing positive and negative examples, separability is the reasoning task of finding a query q_O over the ontology O, expressed in some specific target query language of interest, whose set of *certain answers* contains all the tuples in λ^+ and none of the tuples in λ^-. Although separability has been intensively studied over the years, the paper I worked on is the first to study it in the OBDM context.
Entity Resolution	In the relational database theory, I have worked on *entity resolution*, proposing a novel logical framework, called LACE, that has been designed to satisfy three main desiderata, namely, being (*i*) *collective*, i.e. supports complex interdependencies between merges of different entities; (*ii*) *declarative*, i.e. adopts a declarative language with logical rules and constraints; and (*iii*) *justifiable*, i.e. is able to justify why two constants have been deemed to represent the same entity. LACE utilizes hard and soft rules to specify conditions under which pairs of entity references must or may be merged, together with denial constraints that enforce consistency of the resulting instance.
Bag Semantics	I have studied an extension of the recently proposed *bag* (i.e. *multiset*) *semantics* for OBDM, by studying the impact of adding *functionality axioms* on the computational complexity of conjunctive query answering.
Inequalities	In the context of Description Logics (DL) Knowledge Bases (KBs), I have investigated the problem of adding forms of *inequalities* to unions of conjunctive queries, the most popular class of queries for both database and DL KBs. My principal focus on this topic has been a thorough study of how inequalities affect the computational complexity of query answering over lightweight KBs.

Metamodeling and Meta-querying	In the OWL 2 QL context, I have worked on *Metamodeling* (i.e. allowing for *metarelations*) and *Metaquerying* (i.e. allowing for variables that can be bound to relations). I have introduced the *Metamodeling Semantics Entailment Regime (MSER)*, which generalizes the *Direct Semantics Entailment Regime (DSER)* by (*i*) adopting a Metamodeling Semantics for OWL 2 QL, and (*ii*) relaxing the *typing constraint* of DSER.

VIII – Summary of Scientific Activities

As of today's date (31/01/2023), the following two tables summarize my scientific activities relevant to the competition sector 09/H1:

Product Type	Number	Data Base	Start	End
Papers [international]	20	Scopus	2017	2023
Papers [national]	5	Scopus	2017	2023
Books [scientific]	1	Scopus	2017	2023

Total Citations	86
Average Citations per Product	3.3
Hirsch (H) index	5
Normalized H index*	0.7

*H index divided by the academic seniority (2017–2023).

As of today's date (31/01/2023), here is the list of my scientific publications relevant to the competition sector 09/H1 (sorted by year), according to Scopus:

- Gianluca Cima. **Abstraction in Ontology-based Data Management**. Volume 348 of *Frontiers in Artificial Intelligence and Applications (FAIA)*. IOS Press, 2022.

- Meghyn Bienvenu, Gianluca Cima, and Víctor Gutiérrez-Basulto. **LACE: A Logical Approach to Collective Entity Resolution**. In *Proceedings of the 41st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (PODS 2022)*, pages 379–391, 2022.

- Piero Bonatti, Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, Luigi Sauro, and Domenico Fabio Savo. **Controlled Query Evaluation in OWL 2 QL: A "Longest Honeymoon Approach"**. In *Proceedings of the 21st International Semantic Web Conference (ISWC 2022)*, volume 13489 of Lecture Notes in Computer Science, pages 428–444, 2022.

- Gianluca Cima, Marco Console, Maurizio Lenzerini, and Antonella Poggi. **Investigating Monotone Abstractions**. In *Proceedings of the 30th Italian Symposium on Advanced Database Systems (SEBD 2022)*, volume 3194 of CEUR Electronic Workshop Proceedings, pages 522–529, 2022.

- Gianluca Cima, Federico Croce, and Maurizio Lenzerini. **Query Definability and Its Approximations in Ontology-based Data Management**. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM 2021)*, pages 271–280, 2021.

- Gianluca Cima, Marco Console, Maurizio Lenzerini, and Antonella Poggi. **Abstraction in Data Integration**. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2021)*, pages 1–11, 2021.

- Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, and Domenico Fabio Savo. **Controlled Query Evaluation over Prioritized Ontologies with Expressive Data Protection Policies**. In *Proceedings of the 20th International Semantic Web Conference (ISWC 2021)*, volume 12922 of Lecture Notes in Computer Science, pages 374–391, 2021.

- Mauro Mandorino, António J. Figueiredo, Gianluca Cima, Antonio Tessitore. **A Data Mining Approach to Predict Non-Contact Injuries in Young Soccer Players**. *International Journal of Computer Science in Sport*, 20(2):147–163, 2021.

- Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, and Domenico Fabio Savo. **Controlled Query Evaluation over Ontologies through Policies with Numerical Restrictions**. In *Proceedings of the 4th IEEE International Conference on Artificial Intelligence and Knowledge Engineering (AIKE 2021)*, pages 33–36, 2021.

- Gianluca Cima, Domenico Lembo, Riccardo Rosati, and Domenico Fabio Savo. **Privacy preserving query answering in description logics through instance indistinguishability**. In *Proceedings of the 29th Italian Symposium on Advanced Database Systems (SEBD 2021)*, volume 2994 of CEUR Electronic Workshop Proceedings, pages 490–497, 2021.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Answering conjunctive queries with inequalities in DL-Lite$_R$**. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020)*, pages 2782–2789, 2020.

- Gianluca Cima, Domenico Lembo, Riccardo Rosati, and Domenico Fabio Savo. **Controlled query evaluation in description logics through instance indistinguishability**. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI 2020)*, pages 1791–1797, 2020.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Non-monotonic ontology-based abstractions of data services**. In *Proceedings of the 17th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2020)*, pages 243–252, 2020.

- Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, and Domenico Fabio Savo. **Controlled Query Evaluation in Ontology-Based Data Access**. In *Proceedings of the 19th International Semantic*

Web Conference (ISWC 2020), volume 12506 of Lecture Notes in Computer Science, pages 128–146, 2020.

- Federico Croce, Gianluca Cima, Maurizio Lenzerini, and Tiziana Catarci. **Ontology-based explanation of classifiers**. In *Proceedings of the Workshops of the EDBT/ICDT 2020 Joint Conference*, volume 2578 of CEUR Electronic Workshop Proceedings, 2020.

- Gianluca Cima, Domenico Lembo, Riccardo Rosati, and Domenico Fabio Savo. **Controlled query evaluation in description logics through instance indistinguishability**. In *Proceedings of the 33rd International Workshop on Description Logics (DL 2020)*, volume 2663 of CEUR Electronic Workshop Proceedings, 2020.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Semantic characterization of data services through ontologies**. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019)*, pages 1647–1653, 2019.

- Gianluca Cima, Charalampos Nikolaou, Egor V. Kostylev, Mark Kaminski, Bernardo Cuenca Grau, and Ian Horrocks. **Bag Semantics of DL-Lite with Functionality Axioms**. In *Proceedings of the 18th International Semantic Web Conference (ISWC 2019)*, volume 11778 of Lecture Notes in Computer Science, pages 128–144, 2019.

- Gianluca Cima, Charalampos Nikolaou, Egor V. Kostylev, Mark Kaminski, Bernardo Cuenca Grau, and Ian Horrocks. **Bagging the DL-Lite Family Further**. In *Proceedings of the 32nd International Workshop on Description Logics (DL 2019)*, volume 2373 of CEUR Electronic Workshop Proceedings, 2019.

- Gianluca Cima, Federico Croce, Maurizio Lenzerini, Antonella Poggi, and Elian Toccacieli. **On queries with inequalities in DL-Lite$^=_R$** . In *Proceedings of the 32nd International Workshop on Description Logics (DL 2019)*, volume 2373 of CEUR Electronic Workshop Proceedings, 2019.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Exploiting ontologies for explaining data sources semantics**. In *Proceedings of the Discussion and Doctoral Consortium Papers of the 18th International Conference of the Italian Association for Artificial Intelligence (DDC@AI*IA 2019)*, volume 2495 of CEUR Electronic Workshop Proceedings, pages 33–35, 2019.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Reverse engineering of data services**. In *Proceedings of the 27th Italian Symposium on Advanced Database Systems (SEBD 2019)*, volume 2400 of CEUR Electronic Workshop Proceedings, 2019.

- Gianluca Cima, Giuseppe De Giacomo, Maurizio Lenzerini, and Antonella Poggi. **On the SPARQL Metamodeling Semantics Entailment Regime for OWL 2 QL ontologies**. In *Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics (WIMS 2017)*, pages 10:1-10:6, 2017.

- Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Semantic technology for open data publishing**. In *Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics (WIMS 2017)*, pages 1:1, 2017.

- Gianluca Cima. **Preliminary results on ontology-based open data publishing**. In *Proceedings of the 30th International Workshop on Description Logics (DL 2017)*, volume 1879 of CEUR Electronic Workshop Proceedings, 2017.

- Gianluca Cima, Giuseppe De Giacomo, Maurizio Lenzerini, and Antonella Poggi. **Querying OWL 2 QL ontologies under the SPARQL Metamodeling Semantics Entailment Regime**. In *Proceedings of the 25th Italian Symposium on Advanced Database Systems (SEBD 2017)*, volume 2037 of CEUR Electronic Workshop Proceedings, pages 165, 2017.

I served as a reviewer for the following international journals:

- Artificial Intelligence (AIJ);

- The VLDB Journal;

- Information Systems;

- Journal of Web Semantics (JoWS);

- Journal of Data and Information Quality (JDIQ).

I served in the following program committees:

- Program committee member of the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2023);

- Program committee member of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-ECAI 2022);

- Program committee member of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR 2022);

- Program committee member of the 35th International Workshop on Description Logics (DL 2022)

- Program committee member of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI 2021);

- Program committee member of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-PRICAI 2020).

IX - Selected Publications

1. Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Semantic characterization of data services through ontologies**. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019)*, pages 1647–1653, 2019.

2. Gianluca Cima, Domenico Lembo, Riccardo Rosati, and Domenico Fabio Savo. **Controlled query evaluation in description logics through instance indistinguishability**. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI 2020)*, pages 1791–1797, 2020.

3. Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Answering conjunctive queries with inequalities in DL-Lite$_R$**. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020)*, pages 2782–2789, 2020.

4. Meghyn Bienvenu, Gianluca Cima, and Víctor Gutiérrez-Basulto. **LACE: A Logical Approach to Collective Entity Resolution**. In *Proceedings of the 41st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (PODS 2022)*, pages 379–391, 2022.

5. Gianluca Cima, Charalampos Nikolaou, Egor V. Kostylev, Mark Kaminski, Bernardo Cuenca Grau, and Ian Horrocks. **Bag Semantics of DL-Lite with Functionality Axioms**. In *Proceedings of the 18th International Semantic Web Conference (ISWC 2019)*, volume 11778 of Lecture Notes in Computer Science, pages 128–144, 2019.

6. Piero Bonatti, Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, Luigi Sauro, and Domenico Fabio Savo. **Controlled Query Evaluation in OWL 2 QL: A "Longest Honeymoon Approach"**. In *Proceedings of the 21st International Semantic Web Conference (ISWC 2022)*, volume 13489 of Lecture Notes in Computer Science, pages 428–444, 2022.

7. Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, and Domenico Fabio Savo. **Controlled Query Evaluation over Prioritized Ontologies with Expressive Data Protection Policies**. In *Proceedings of the 20th International Semantic Web Conference (ISWC 2021)*, volume 12922 of Lecture Notes in Computer Science, pages 374–391, 2021.

8. Gianluca Cima, Domenico Lembo, Lorenzo Marconi, Riccardo Rosati, and Domenico Fabio Savo. **Controlled Query Evaluation in Ontology-Based Data Access**. In *Proceedings of the 19th International Semantic Web Conference (ISWC 2020)*, volume 12506 of Lecture Notes in Computer Science, pages 128–146, 2020.

9. Gianluca Cima, Maurizio Lenzerini, and Antonella Poggi. **Non-monotonic ontology-based abstractions of data services**. In *Proceedings of the 17th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2020)*, pages 243–252, 2020.

10. Gianluca Cima, Federico Croce, and Maurizio Lenzerini. **Query Definability and Its Approximations in Ontology-based Data Management**. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM 2021)*, pages 271–280, 2021.

11. Gianluca Cima, Marco Console, Maurizio Lenzerini, and Antonella Poggi. **Abstraction in Data Integration**. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2021)*, pages 1–11, 2021.

12. Gianluca Cima, Giuseppe De Giacomo, Maurizio Lenzerini, and Antonella Poggi. **On the SPARQL Metamodeling Semantics Entailment Regime for OWL 2 QL ontologies**. In *Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics (WIMS 2017)*, pages 10:1-10:6, 2017.

Firma..