

Marco Console

Curriculum Vitae (Ai Fini della Pubblicazione)

Part I – General Information

Full Name	Marco Console

Part II – Education

Type	Year	Institution	Notes (Degree, Experience,...)
University graduation	2009	Sapienza, University of Rome	Laurea in Ingegneria Informatica
Post-graduate studies	2012	Sapienza, University of Rome	Laurea Magistrale in Ingegneria Informatica
PhD	2017	Sapienza, University of Rome	Dottorato in Ingegneria Informatica

Part III – Appointments

- From 16/10/2016 to 30/09/2020. Research associate. School of Informatics, University of Edinburgh. Edinburgh, Scotland, UK.
- From 05/10/2020 to date. Ricercatore a Tempo Determinato di Categoria A. Dipartimento di Ingegneria Informatica, Automatica e Gestionale “Antonio Ruberti”. Sapienza, University of Rome. Rome, Italy.

Part IV – Teaching Experience

- A.Y. 2020/21. Course Title: “Laboratorio di Architetture Software e Sicurezza Informatica”. Bachelor’s Degree in “Ingegneria Informatica”. Sapienza, University of Rome. Course Length: 30hrs.
- A.Y. 2020/21. Course Title: “Large Scale Data Management”. Master’s Degree in “Engineering in Computer Science”. Sapienza, University of Rome. Course Length: 30hrs.
- A.Y. 2020/21. Course Title: “Incomplete Data – Theory in Practice”. Course for the PhD program in “Engineering in Computer Science”. Sapienza, University of Rome. Course Length: 15hrs.
- A.Y. 2021/22. Course Title: “Basi di Dati”. Course for the Bachelor’s Degree in “Ingegneria dell’Informazione (Sede di Latina)”. Sapienza, University of Rome. Course Length: 30hrs.

Part V - Society Memberships, Awards and Honors

- 2018. Ray Raiter Best Paper Award. Principles of Knowledge Representation and Reasoning (KR). KR is one of the major conference on the topic of artificial intelligence and is ranked A* on CORE.

Part VI - Funding Information [grants as PI-principal investigator or I-investigator]

Part VII (A)– Research Activities

Ontology-Based Data Access and Information Integration. Ontologies are conceptualizations of a domain of interest, written in terms of some logical formalism. Using ontologies, one can provide a reconciled view of a set of heterogeneous data sources. I started my study of Ontology-Based Data Access and Information Integration during my PhD. My contribution to the field spans several different topics.

1. **Ontology-Based Data Quality.** I worked extensively on algorithms and techniques to assess the quality of data using ontologies. My work on this topic is in collaboration with Prof. Maurizio Lenzerini. This novel paradigm, that we dubbed Ontology-Based Data Quality, gave rise to several interesting research questions connected to both ontological reasoning and data management. Part of our work on this topic was accepted for publication in AAAI 2014 and ECAI 2014, two of the most prestigious international conferences on the topic of AI.
2. **Ontology-Based Data Access and Management.** Ontology-Based Data Management Systems implement the Ontology-Based Data Access paradigm in real world scenarios. Some of my most recent contributions to the development of such systems are the following. 1) Integrity Constraints. I studied a language of integrity constraints to express requirements on the underlying data sources using only the reconciled view of the ontology. A paper detailing the language was accepted for publication in AAAI 2020. 2) Views for Ontology-Based Data Management Systems. I worked on the development of a language to define database-like views for ontology-based data access systems. A paper detailing the language was accepted for publication in ICJAI 2021.
3. **Query Abstractions.** Queries to Ontology-Based Data Management systems are expressed using the symbols in the ontology, thus abstracting from the actual structure of the data sources. In different contexts, however, one would like to do the opposite, i.e., express a query using the schema of the data sources and then find an ontological query that best captures its meaning. We call such queries “abstractions”. Abstractions can be used, for example, to provide human-understandable, semantic explanations of services defined over the source data. I recently started working on the problem of abstraction jointly with Prof. Maurizio Lenzerini (Sapienza, University of Rome), Prof. Antonella Poggi (Sapienza, University of Rome), and Doc. Gianluca Cima (CNRS and University of Bordeaux). In particular, we focused on abstractions expressed as monotonic queries and defined an ontological query language suited to express relevant forms of monotonic abstractions. Our work on this topic was submitted for publication in one of the major international conferences on artificial intelligence and is currently under review.

Databases with Incomplete Information. During my post-doctoral years at the School of Informatics of the University of Edinburgh, I worked on several topics connected to incomplete databases. My work on this topic, mainly in collaboration with Prof. Leonid Libkin (University of Edinburgh and Ecole Normal Supérieur, PSL) and Prof. Paolo Guagliardo (University of Edinburgh), spans several different problems.

1. **Bag-Semantics.** We studied the algorithmic complexity of the bag semantics data model in the presence of incomplete information. Intuitively, the bag semantics data model allows multiple occurrences of tuples in database relations, thus, capturing a fundamental feature of real-world database systems. In this context, we studied the complexity of computing a very natural form of answer, i.e., *certain answers*, for queries defined in several different fragments of relational algebra. This work was accepted for publication in ICDT 2019 and invited to Information System Journal in the special issue of the “Alberto Mendelzon Workshop” 2019.
2. **SQL query language.** In real-world databases, the standard way to express queries is via SQL, i.e., the ISO committee standard for database queries. While the expressive power of SQL is clear in the case of complete databases, the picture becomes much more complex in the case of incomplete information. In our work, we tried to shed some light on the problem, studying the logic underlying incomplete SQL databases. This work was accepted for publication in KR 2018 and won the Best Paper Award.
3. **Numerical Queries.** An important feature of real-world databases is the ability to evaluate arbitrary arithmetic expressions over the data. Unfortunately, several natural problems become intractable when we consider queries with arithmetic. To mitigate this issue, we defined a novel form of answers suited to work in efficient approximation schemes. Preliminary results on this topic have been accepted for publication in IJCAI 2019. Works discussing several algorithmic techniques related to this problem has been accepted for publication in ACM PODS 2020 and KR 2020 conferences.

Consistent Query Answering. Consistent Query Answering is the problem of computing meaningful answers to queries over inconsistent datasets, i.e., datasets that violate a given set of integrity constraints. I worked on this topic in collaboration with Prof. Andreas Pieris (University of Edinburgh) and Prof. Marco Calautti (Università di Trento). My contribution to the field was both theoretical and practical. We studied the complexity of counting database repairs that satisfy a given query considering integrity constraints in the form of primary keys. Our work on this topic was accepted for publication in ACM PODS 2019. More recently, we developed a benchmark for Consistent Query Answering in the presence of key constraints and carried out an experimental evaluation some of the most popular approximation algorithms for Consistent Query Answering. A paper detailing the benchmark and the results of the evaluation was accepted for publication in ACM PODS 2021.

Model Theory. Model theory is the field of mathematical logic that studies the interaction between the syntax of formulae and their semantics. Model theory provides powerful and elegant tools to study several popular problems in databases and artificial intelligence. In this context, I developed a model-theoretic characterization of a formalism commonly used to define database constraints and ontologies, namely Tuple-Generating Dependencies. I worked on this topic in collaboration with Prof. Andreas Pieris (University of Edinburgh) and Prof. Phokion Kolaitis (University of California, Santa Cruz and IBM Almaden Research Center). One of our papers on the topic was accepted for publication in ACM PODS 2021.

Part VII(B)–Participation to Research Groups and Research Projects

I participated to the activities of different research groups.

- (Oct, 2016 – Sept, 2020) I worked in the “Data Management” group led by Prof. Leonid Libkin at the School of Informatics of the University of Edinburgh.
- (Oct, 2020 – To Date) I have been working in the “Data Management and Service-Oriented Computing” led by Prof. Maurizio Lenzerini at “Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Antonio Ruberti” (Sapienza, University of Rome).

I have been part of the following research projects.

- Optique (Scalable End-User Access to Big Data). FP7 EU Grant n. FP7318338. My work focused on data quality aspects of ontology-based data management systems and the Ontology-Based Data Quality paradigm.
- VADA: Value Added Data. EPSRC grant M025268. My work focused on the problem of querying databases under bag semantics with incomplete information.
- EQUID: Efficient Querying Inconsistent Data. EPSRC grant EP/S003800/1. My work focused on the notion of relative frequency for consistent answers and the definition of a benchmark for consistent query answering.
- MAGIC: Managing Incomplete Data. EPSRC grant N023056. My work focused on the semantics of SQL queries over databases with incomplete information.
- PRIN 2017 project HOPE. PRIN prot. 2017MMJJRE. My work focused on the notion of query abstraction in the context of open data.
- TAILOR: Foundations of Trustworthy AI. H2020 EU2.1.1 grant id 952215. I am the coordinator of Sapienza’s local unit of WP 4 “Integrating AI Paradigms and Representations”, task 4.3 “Learning and Reasoning with Embeddings, Knowledge Graphs, and Ontologies”. My work focusses on the notion of virtual data graphs, ontologies, and ontology-based data management.

Part VIII – Summary of Scientific Achievements

Number of international papers: 31 (Scopus). From 2012 to 2021.

Total Impact Factor 2.047; Average Impact Factor per Product: 0.06.

Total Citations: 178 (Scopus); Average Citations per Product: 5.7 (Scopus)

Hirsch (H) Index: 7 (Scopus); Normalized H Index: 0.875 (Scopus), over 8 years of academic activity.

Part IX– Selected Publications (In Chronological Order)

1. **Title:** Benchmarking approximate consistent query answering
Authors: Calautti, M., Console, M., Pieris, A.
In: Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 2021, pp. 233–246
2. **Title:** Model-theoretic characterizations of rule-based ontologies
Authors: Console, M., Kolaitis, P.G., Pieris, A.
In: Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 2021, pp. 416–428
3. **Title:** Queries with Arithmetic on Incomplete Databases
Authors: Console, M., Hofer, M., Libkin, L.
In: Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 2020, pp. 179–189
4. **Title:** Coping with Incomplete Data: Recent Advances
Authors: Console, M., Guagliardo, P., Libkin, L., Toussaint, E.
In: Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 2020, pp. 33–47
5. **Title:** Epistemic integrity constraints for ontology-based data management
Authors: Console, M., Lenzerini, M.
In: Proceedings of AAAI 2020 - 34th AAAI Conference on Artificial Intelligence, 2020, pp. 2790–2797
6. **Title:** Fragments of bag relational algebra: Expressiveness and certain answers
Authors: Console, M., Guagliardo, P., Libkin, L.
In: Information Systems, 2020, 101604
7. **Title:** Reasoning about measures of unmeasurable sets
Authors: Console, M., Hofer, M., Libkin, L.
In: Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, 2020, 1, pp. 263–272
8. **Title:** Measuring the likelihood of numerical constraints
Authors: Console, M., Hofer, M., Libkin, L.
In: Proceeding of IJCAI International Joint Conference on Artificial Intelligence, 2019, 2019-August, pp. 1654–1660
9. **Title:** Counting database repairs under primary keys revisited
Authors: Calautti, M., Console, M., Pieris, A.
In: Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 2019, pp. 104–118
10. **Title:** Propositional and predicate logics of incomplete information
Authors: Console, M., Guagliardo, P., Libkin, L.
In: Proceedings of Principles of Knowledge Representation and Reasoning: Proceedings of the 16th International Conference, KR 2018, 2018, pp. 592–601
11. **Title:** On querying incomplete information in databases under bag semantics
Authors: Console, M., Guagliardo, P., Libkin, L.
In: Proceedings of IJCAI International Joint Conference on Artificial Intelligence, 2017, 0, pp. 993–999
12. **Title:** Data quality in ontology-based data access: The case of consistency
Authors: Console, M., Lenzerini, M.
In: Proceedings of AAAI 2014 - 28th AAAI Conference on Artificial Intelligence, pp. 1020–1026

Part X– Other Academic Activities and Service to the Community

- PC Member of the following international conferences
 - ACM Principles of Database Systems (PODS). 2022.
 - AAAI Conference on Artificial Intelligence (AAAI). 2022.
 - AAAI Conference on Artificial Intelligence (AAAI). 2021.
 - AAAI Conference on Artificial Intelligence (AAAI). 2020.
 - International Joint Conference on Artificial Intelligence (IJCAI). 2020.
 - International Conference on Scientific and Statistical Database Management (SSDBM). 2020.
 - Principles of Knowledge Representation and Reasoning (KR). 2020.
- Reviewer for the following international journals
 - Information Systems (2019);
 - International Journal of Information Management (2019).
 - Journal of Computer System Science (2021)
- Conference Chair
 - Virtual Conference Arrangement Chair for the Principles of Knowledge Representation and Reasoning (KR) conference. 2021.
- Member of the Editorial Board of Scientific Journals
 - Guest Editor of ACM Journal of Data and Information Quality (JDIQ) for the Special Issue “Quality Aspects of Data Preparation”. 2021.

28/09/2021

Marco Console