

# Francesca Bugiotti

---

## PERSONAL INFORMATION

**E-mail**  
francesca.bugiotti@centralesupelec.fr

## EDUCATION

**Università Roma Tre** (*November 2008 - April 2012*)  
PhD in Computer Science - Computer Science and Automation department.

- Thesis: “A model oriented approach to heterogeneity”
- Advisor: Prof. Paolo Atzeni

**Università Roma Tre - IBM - Formit**<sup>1</sup> (*January 2008 - March 2009*)  
Post lauream degree in IT governance: development, management and monitoring.

**Università Roma Tre** (*October 2008*)  
Qualifying examination to exercise the engineering profession.

**Università Roma Tre** (*October 2005 - December 2007*)  
Master degree in Computer Engineering (“Laurea Specialistica in Ingegneria informatica”).

- Title of the thesis: “Tools and methodology for model management problems”.
- Advisor: prof. Paolo Atzeni
- Final grade: 110/110 lode (maximum honors)

**Università Roma Tre** (*October 2002 - July 2005*)  
First level degree in Computer Engineering (“Laurea in Ingegneria informatica”).

- Thesis: “Datalog rules management for data and schema translation”.
- Advisor: Prof. Paolo Atzeni
- Final grade: 110/110 lode (maximum honors)

## AWARDS

**Accenture - Università Roma Tre** (*March 2009*)  
“Accenture” Outstanding Engineering Graduate Award.

**IBM** (*July 2007*)  
Participant in “IBM EMEA Best Student Recognition Event”, Nice.

## RESEARCH ACTIVITY

**CentraleSupélec**  
(*March 2015 - present*)  
Research activity on Data Modeling and Data integration.

**Inria - Institut National de Recherche en Informatique et en Automatique**  
(*November 2013 - present*)  
Research activity on RDF datasets indexing in the Amazon Cloud Computing Environment.

**Università Roma Tre** (*April 2012 - October 2013*)  
Post graduate research activity on model management in databases and No-SQL data stores integration.

**Inria - Institut National de Recherche en Informatique et en Automatique**  
(*April 2011 - July 2011*)  
Research activity on RDF datasets indexing in the Amazon Cloud Computing Environment.

**Università Roma Tre** (*January 2008 - November 2008*)  
Research activity on model management problems in databases.

---

<sup>1</sup>FORMIT is a Foundation that performs activities of scientific research, technical support, analysis and industrial, financial and socio-economic evaluation to sustain migration processes and integration of technological systems in every field of society.

**ISA s.p.a.** (*December 2008 - January 2010*)

Part-time research activity on data mining applied to clinical data.

TEACHING  
ACTIVITY

**Software Engineering**, course at CentraleSupélec (2015), with Paolo Ballarini. 35 ETD per year of exercises, test grading, office hours. The course covers Java Programming, Object Oriented Analysis and Design, introduces to UML and concurrent programming.

**Database**, course for **2nd year Bachelor students** at Università Roma Tre (2008, 2009, 2011), with Paolo Atzeni. 20 ETD per year of exercises, test grading, office hours. The course introduces databases design and data base management systems (DBMS). I curated the practical part that requires to perform experiments in some DBMss (IMB DB2, PostgreSQL, Oracle, SimpleDB) and in some environments for data warehousing (Pentaho).

**Object-Oriented Analysis and Design**, course for **3rd year Bachelor students** at Università Roma Tre (2009, 2011, 2012), with Luca Cabibbo. 30 ETD per year of exercises, test grading, office hours. The course covers Object Oriented Analysis and Design, introduces to agile object oriented programming and to the iterative and incremental software developing approach.

**IT Governance**, course for **2nd year Master students** at Università Roma Tre (2012), with Paolo Atzeni. 15 ETD of lectures, exercise sessions, consultation hours, grading of the exam. The course introduces the principles governing the lifecycle of an IT systems describing the methodological issues, the economic rules and the juridical aspects.

**Java Programming and Algorithms**, course for **3rd year Bachelor students** at Università Roma Tre (2007), with Luca Cabibbo. 30 ETD of laboratory exercise lessons and test grading for the exam. The course introduces the Java programming methodologies.

**Remedial Mathematics**, course for **1st year Bachelor students** at Università Roma Tre (2009), math lessons. 30 ETD of lectures and exercise sessions. The course introduces math concepts to students. It covers linear algebra, geometric, trigonometry and function analysis.

**Contribution to a database book**, “Databases - Models and query languages” (“Basi di dati - Modelli e linguaggi di interrogazione”), McGraw Hill 2009, under contract with the publisher. I wrote 100 exercises accompanying the book chapters and I provided the solutions (accessible on-line at <http://www.ateneonline.it/atzeni3e/areastudenti.asp>).

TECHNICAL  
SKILLS

- *Programming languages*: Java, C/C++, Prolog, OCaml, Assembly.
- *Databases*: Relational Databases, SQL, No-SQL data stores (MongoDB, Redis, Oracle NoSQL, HBase, DynamoDB).
- *Operating Systems*: Linux, Mac Os, Windows.

RESEARCH  
PROJECTS

**Università Roma Tre**

**SOS**

Save Our Systems (SOS) is a common programming interface [D10] to NoSQL systems. Its goal is to support application development by hiding the specific details of the various systems.

I contributed to the definition to the architecture of the platform, the operations it exposes and the query strategies it implements. I have been involved in the definition of the strategies for integrating the NoSQL data stores into the system. I also participated in the definition of the data storage techniques that are used in each datastore in order to perform operations the interface exposes [IC7,IJ4].

**NOAM and ONDM**

NoSQL Abstract Model (NOAM) is a logical approach to the NoSQL database design problem [N15] and aims at exploiting the commonalities of various NoSQL systems. It is based on an intermediate, abstract data model where aggregates are units of distribution (to

support scalability) and consistency (to the extent it is needed). Some intermediate representations can be implemented in target NoSQL datastores, taking into account their specific features and providing an effective support for scalability, consistency, and performance [M16]. ONDM (Object-NoSQL Datastore Mapper), is the framework [M17] that supports NOAM approach. It provides application developers with a uniform programming interface, as well as the ability to map application data to different data representations and can be used, in an effective way, for performing the experiments during the design of a NoSQL database.

### **MIDST**

Model-Independent Schema and Data Translation (MIDST) is a platform for model-independent schema and data translation based on a meta-level approach over a wide range of data models (Relational, OR, OO, ER, XML).

I contributed to the extension of MIDST *supermodel* (a general model handled by the platform that describes the various data models in terms of a small set of *basic constructs*) and I also implemented some core software components like the Datalog-SQL translation engine giving some ideas about the evolution of the platform [N12].

### **MIDST-RT**

Model-Independent Schema and Data Translation-RunTime (MIDST-RT) is a platform based on MIDST but that implements a runtime approach.

I contributed to the definition, design and implementation of the MIDST-RT algorithm that, given the schema of the source database and the model of the target one, generates views on the operational system that expose the underlying data according to the corresponding schema in the target model. The implemented approach generates views in an automatic way, on the basis of the Datalog rules for schema translation [IJ2,IC6,N14].

### **MISM**

Model Independent Schema Management (MISM) is a platform for model management that offers a set of operators to manipulate schemas.

I designed and implemented the algorithm that gives one solution to the round-trip engineering problem considering the main model management operators (merge, diff, and modelgen) implemented according to a model-independent and model-aware approaches based on MIDST supermodel [IJ3,IN13].

### **MATRIX and EXL**

I collaborated with the Bank of Italy, supporting the implementation of EXLEngine, a tool that manipulates statistical data at high level, in terms of entities of statistical models such as time series. We proposed (i) a language, EXL, has been defined for the declarative specification of statistical programs, (ii) an approach for the translation of EXL code into executables in various target systems has been developed, and (iii) a concrete implementation, EXLEngine. The approach leverages on schema mappings as an intermediate specification step, in order to facilitate the translation from EXL towards several target systems [IC5].

### **GENDATA**

The work regarding data models continues within the GENDATA European project <http://gendata.weebly.com/> that aims at building the abstractions, models, and protocols for supporting a network of genomic data, making them available for genome servers located in the major biologist laboratories in the world. I started to collaborate to the project within the working packages involving Università Roma Tre investigating about the model design, the query language and the model standardization.

### **INRIA**

#### **AMADA**

During my internship I contributed to the AMADA project: a platform [D9,N11] for storing Web data (XML documents and RDF graphs) based on the Amazon Web Services (AWS) cloud infrastructure.

I provided a solution for the problem of indexing RDF datasets by using SimpleDB, a

key-value store provided by AWS. I contributed to the definition and development of four indexing strategies [B1,IC8].

### **ESTOCADA**

A novel system capable of exploiting side-by-side a practically unbound variety of DMSs, all the while guaranteeing the soundness and completeness of the store, and striving to extract the best performance out of the various DMSs. Our system leverages recent advances in the area of query rewriting under constraints, which we use to capture the various data models and describe the fragments each DMS stores. [N15].

#### **PUBLICATIONS**

All my publications can be found at:

<http://www.bugiotti.it/publications.html>

The general practice within my scientific community is to *list all authors in alphabetic order* unless one author has contributed very significantly more than her or his share.

---

### **Books and contribution to books**

---

- [B1] Francesca Bugiotti, Jesús Camacho-Rodríguez, François Goasdoué, Zoi Kaoudi, Ioana Manolescu, and Stamatis Zampetakis. SPARQL query processing in the cloud. In Andreas Harth, Katja Hose, and Ralph Schenkel, editors, *Linked Data Management: Principles and Techniques*, pages 49–72. CRC Press, 2013. To appear.

---

### **Articles in international journals**

---

- [IJ2] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, Fabrizio Celli, and Giorgio Gianforme. A runtime approach to model-generic translation of schema and data. *Information Systems*, 37(3):269–287, 2012.
- [IJ3] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme. MISM: A Platform for Model-Independent Solutions to Model Management Problems. *Journal of Data Semantics*, 14:133–161, 2009.
- [IJ4] Paolo Atzeni, Francesca Bugiotti, and Luca Rossi. Uniform Access to NoSQL Systems. *Information Systems*, 2013.

---

### **Full articles in international conferences and workshops**

---

- [IC5] Paolo Atzeni, Luigi Bellomarini, and Francesca Bugiotti. EXLEngine: executable schema mappings for statistical data processing. In *International Conference on Extending Database Technology (EDBT)*, pages 672–682, 2013.
- [IC6] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme. A runtime approach to model-independent schema and data translation. In *International Conference on Extending Database Technology (EDBT)*, pages 275–286, 2009.
- [IC7] Paolo Atzeni, Francesca Bugiotti, and Luca Rossi. Uniform Access to Non-relational Database Systems: The SOS Platform. In *International Conference on Advanced Information Systems Engineering (CAiSE)*, pages 160–174, 2012.
- [IC8] Francesca Bugiotti, François Goasdoué, Zoi Kaoudi, and Ioana Manolescu. RDF Data Management in the Amazon Cloud. In *Workshop on Data analytics in the Cloud (DanaC 2012)*, Berlin, Allemagne, February 2012.

---

## Demonstrations in international conferences

---

- [D9] Andrés Aranda-Andújar, Francesca Bugiotti, Jesús Camacho-Rodríguez, Dario Colazzo, François Goasdoué, Zoi Kaoudi, and Ioana Manolescu. AMADA: Web Data Repositories in the Amazon Cloud. In *ACM CIKM - International Conference on Information and Knowledge Management*, Maui, États-Unis, 2012.
- [D10] Paolo Atzeni, Francesca Bugiotti, and Luca Rossi. SOS (Save Our Systems): a uniform programming interface. for non-relational systems. In *EDBT*, pages 582–585, 2012.

---

## Articles and demos at national database conference

---

- [N11] Andrés Aranda-Andújar, Francesca Bugiotti, Jesús Camacho-Rodríguez, and Zoi Kaoudi. AMADA: Web Data Repositories in the Amazon Cloud. In *Journées des Bases de Données Avancées (BDA)*, 2012.
- [N12] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme. From Schema and Model Translation to a Model Management System. In *British National Conference on Databases (BNCOD)*, pages 227–240, 2008.
- [N13] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme. A platform for model-independent solutions to model management problems. In *Italian Symposium on Advanced Database Systems (SEBD)*, pages 310–317, 2008.
- [N14] Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme. A runtime approach to model-independent schema and data translation. In *Italian Symposium on Advanced Database Systems (SEBD)*, pages 245–252, 2009.
- [N15] Francesca Bugiotti, Damian Bursztyn, Alin Deutsch, Ioana Ileana, and Ioana Manolescu. Toward Scalable Hybrid Stores. In *Italian Symposium on Advanced Database Systems (SEBD)*, pages 312–319, 2015.
- [N16] Francesca Bugiotti and Luca Cabibbo. A Comparison of Data Models and APIs of NoSQL Datastores. In *Italian Symposium on Advanced Database Systems (SEBD)*, pages 63–74, 2013.
- [N17] Francesca Bugiotti, Luca Cabibbo, Paolo Atzeni, and Riccardo Torlone. How I Learned to Stop Worrying and Love NoSQL Databases. In *Italian Symposium on Advanced Database Systems (SEBD)*, pages 216–223, 2015.

---

## Manuscripts

---

- [M18] Francesca Bugiotti and Luca Cabibbo. A logical approach to nosql databases. 2014.
- [M19] Francesca Bugiotti and Luca Cabibbo. An object-datastore mapper supporting nosql database design. 2014.