

**Selected Curriculum of SONIA BERGAMASCHI**  
**(from 2006 up to now)**  
**8 November 2023**

**Personal Data & Academic Career**

- SONIA BERGAMASCHI was born in \_\_\_\_\_, she graduated in Mathematics (Laurea degree ) at the “Università degli Studi di Modena” (now “Università degli Studi di Modena e Reggio Emilia” - UNIMORE) on \_\_\_\_\_
- since 1/11/1999 she is Full Professor of Computer Science, serving in this position at the UNIMORE Engineering Department “Enzo Ferrari” ([www.ingmo.unimore.it](http://www.ingmo.unimore.it));
- from 1 Novembre 1992 to 31/10/1999 she was Associate Professor at the same institution;
- from 1/3/1984 to 31/10/1992 she was senior researcher at CIOC-CNR research center in Bologna;
- e-mail: \_\_\_\_\_ UNIMORE site: \_\_\_\_\_  
research site: <http://dbgroup.unimore.it/site/home.html>

address: Dipartimento di Ingegneria “Enzo Ferrari”, Via P. Vivarelli n.10, building 27, office: IM-1-11 41125 MODENA (Italy) – tel: \_\_\_\_\_ ; fax: + \_\_\_\_\_ – mobile: \_\_\_\_\_

skype: \_\_\_\_\_

- She is member of ACM, member number: 4925475 since 1998 and has been appointed as Distinguished Researcher on November 2016.
- She is Member of IEEE, member number: 21614495 since 1992 and has been appointed as Senior Member on August 2016.
- from 1 November 2019 to 31 October 2023 she has been appointed as delegate for Information Technology at UNIMORE
- From 22 May 2017 to 31 October 2023 she was the dean of the ICT International Doctorate at UNIMORE ([www.ict.unimore.it](http://www.ict.unimore.it)), renewed for three years starting on 1 November 2020.
- On 1<sup>st</sup> November 2023 \_\_\_\_\_ and was appointed as Senior Professor at UNIMORE (validity of appointment three years)

## Significant Professional Leadership

### Project Leadership

On years 2005-2006, she was national coordinator of the PRIN project WISDOM (Web Intelligent Search based on DOMain ontologies). The aim of the project was to overcome the "information overloading" problem of traditional search engines.

On years 2006-2009, she was local coordinator for UNIMORE of the IST european project “STASIS (SoftWare for Ambient Semantic Interoperable Services)”, whose aim was to enable SMEs and corporates to participate in the eEconomy, by offering services and applications based on the open SEEM registry and repository network.

On years 2006-2009, she was national coordinator of the Italian basic research (MIUR FIRB) project “NeP4B (Networked Peers for Business)”. NeP4B goal was to contribute innovative ICTs solutions for SMEs, by developing an advanced technological infrastructure to enable companies of any nature, size and geographic location to search for partners, exchange data and services, negotiate and collaborate without limitations and constraints.

On years 2009-2010 she was the coordinator of the Keymantic project founded by the “Fondazione Cassa di

Risparmio di Modena (<http://www.fondazione-crmo.it/>), whose aim was the development of a keyword-based query engine supporting users in querying the “deep web”.

On years 2010-2011 she was the local coordinator for UNIMORE of the P6-IST-4 Project FACIT-SME (Facilitate IT-providing SMEs by Operation-related Models and Methods) project. The project addresses SMEs operating in the ICT domain. The goal was to facilitate the use of Software Engineering (SE) methods and to systematize their application integrated with the business processes. [IJ38]

On 2012-2014 she was coordinator of the UNIMORE strategic research project “ICT- SMART ICT FOR SMART SOCIAL WORLDS”, involving about sixty researchers of Computer Science, Telecommunication and Mathematicians of UNIMORE aiming to show the training role of the ICT not only for supporting the industrial and service sectors, but also as a driving and innovative social force.

On 2014 she was the coordinator of the “ [Open linked data dell'Osservatorio Giovani della Regione Emilia-Romagna](#)”, a data integration project developed with the MOMIS system. The project goal was to identify and integrate the data sources, both proprietary and open, of interest for the youth politics and to make the resulting value-added information public and searchable on the Web as Linked Open Data.

On years 2015-2019 she was the recipient of a research grant for the management of Big Scientific Data by CINECA (<http://www.cineca.it/en>), the italian supercomputing infrastructure supporting the research activities of the italian scientific community.

Since 2017 she is the referent person for UNIMORE within the Italian CINI Big Data Lab, member of the european Big Data Value Association (BDVA) and collaborates with the BDVA I-Space “Data Innovation Space” at CINECA which received the Silver Label by BDVA association on November 2017.

On 2016-2018 she was the UNIMORE coordinator of the Emilia-Romagna region project “Creazione di valore per imprese e società con la gestione e l’analisi di Big Data (Creating value for Industry and Society by Managing and Analyzing Big Data)”. She obtained the funding of a doctorate scholarship on Big Data Integration and 4 research grants on Big Data in co-development with Modena’ companies. Moreover, she obtained the approval and co-financing of an “Academy on Management and Analysis of Big Data” in collaboration with the CINECA department of “Supercomputing, Applications and Innovations”.

On 2020 she was the UNIMORE coordinator of the two-year research project: “ Design of an application and a data architecture for the enhancement of energy virtuosity and integration of the application and data architecture for enhancement of energy virtuosity, grant from ENEA.

On 2022 -2023 she was the coordinator of the workpackage” Privacy Preserving Temporal Record Linkage” within the b"Recidivism Data Mart and Criminal Data Warehouse" project, funded by the [CRUI Foundation](#), which is part of the broader "[PNRR Justice Plan Datalake](#)" project.. The research activity aims to deliver a Proof of Concept of Privacy Preserving Temporal Record Linkage techniques.

On 2023, she became member of two (three years) **PNRR** research project: **ECOSISTER** and **ITSERR**. Within ECOSISTER she participate in the Spoke3: “ Green Manufacturing for a sustainable economy”, contributing at deliverable: D2.3.2 “Definition of technologies and key components, including mapping of regional existing technologies " and in particular at the section Digital Twins for hydrogen plants. The research activity coordinated by ITSERR, funded by the PNRR and designed to strengthen RESILIENCE (the European research infrastructure for religious sciences) aims to design and implement a system to manage, catalogue, and share historical heritage in non-Latin alphabets. Within ITSERR she plays the role of the co-coordinator of WP5 “Maktaba”.

### **Database Research Group (DBGROUP) Leadership**

Since 1992 she leads the database research group (DBGROUP) at UNIMORE ([www.dbgroup.unimore.it](http://www.dbgroup.unimore.it)). The group grew up in the years from one person (myself) to nine person (6 professors: Sonia Bergamaschi, Domenico

Beneventano, Maurizio Vincini, Francesco Guerra, Laura Po, Giovanni Simonini); 1 researcher: Luca Gagliardelli, 1 senior technician: Alberto Corni & 1 phd: Mirko Orsini CEO of the DATARIVER start-up she founded.

- She was the master and phd thesis supervisor of all the members of the DBGROUP and of many other master (about seventy) and phd students (about twenty) who achieved relevant positions in the professional carrier.
- PHD: Domenico Beneventano (now associate professor at UNIMORE), Alberto Corni (now high profile IT technician at UNIMORE), Jean Paul Ballerini (now IT cyber security expert at IBM), Jonathan Gelati (now IT analyst at Varian Medical Systems), Ilario Benetti (now CEO of APEX srl at Modena <http://www.apex.it/solutions/en/home>), Maurizio Vincini (now associate professor at UNIMORE), Mirko Orsini (now CEO of DATARIVER srl at [www.datariver.it](http://www.datariver.it)), Francesco Guerra (now full professor at UNIMORE), Laura Po (now associate professor at UNIMORE), Serena Sorrentino (now member of the Digital Technologies and Digital Engagement team at BARILLA group), Matteo Interlandi (now researcher at Microsoft Research- Seattle - <https://www.linkedin.com/in/minterlandi/>), Marius Olaru (now senior Manufacture Project Engineer at SCM group, Rimini), Giovanni Simonini (now associate professor at UNIMORE), Fabio Benedetti (CTO at Cambiomarcia, Gruppo UNIPOL, Bologna), Zhou Song (now at TESLA -Shangai), Luca Gagliardelli (now researcher at DBGROUP- UNIMORE), Luca Magnotta (now IT Architect at DATARIVER, Modena), Giuseppe Fiameni (now at NVIDIA AI Technology Center Italy). (see the DBGROUP site & follow the link published thesis).

### Management Experiences at Universities/ Research Labs

- From 2004 to 2010 she was the vice-dean of the DII (Dipartimento di Ingegneria dell'Informazione) at UNIMORE
- From 1 January 2006 to December 2010 she was the dean of the International Doctorate School in ICT at UNIMORE ([www.ict.unimore.it](http://www.ict.unimore.it)) and from 7/5/2009 to December 2010 she was the coordinator of the Council of the Doctorate Schools deans at UNIMORE.
- From 1 November 2011 to 30 June 2012 she was the dean of the DII department.
- From 22 May 2017 up to now she is the dean of the ICT International Doctorate at UNIMORE ([www.ict.unimore.it](http://www.ict.unimore.it)), renewed for three years starting on 1 November 2020.
- From 2 December 2016 to June 2020 she was member of the Scientific-Strategic Committee of AlmaLaurea, a Interuniversity Consortium established in 1994 and currently counting 74 Universities as members and representing about 90% of Italian graduates (<http://www.almalaurea.it/en>).
- Since 1 November 2019 she has been appointed as Information Technologies delegate for UNIMORE.

### Teaching Record Innovations

#### Courses for Computer Engineering, at the Engineering Department "Enzo Ferrari" UNIMORE (from 2010 to now)

- **Databases (basic - for bachelor students)**, which aims to provide the fundamental notions of: conceptual and relational models and languages; database design and web databases. The main innovation, starting from 2010, consists in the agreements with the major database systems vendors as to provide the students a true database work environment. The exam is performed at the lab to prove the acquired skills.
- **Big Data Management and Governance (advanced for master students)**, which aims to provide an advanced knowledge of: the full standard Relational DBMS language SQL92 (trigger, store procedures, embedded SQL); the RDBMS technology: architecture, concurrency and recovery of database transactions; storage, organization and access techniques for persistent database; optimization; relational

- normalization theory; distributed management of Massive Data (RDBMS and NOSQL DBMS); General notions on architectures, policies, practices and procedures to properly manage the full data life cycle of an enterprise; general notions on data security, privacy and quality management.
- The impact on this course, due to the incoming of the Internet companies was very strong. Starting from 2012, the new NOSQL systems have been introduced, as well as the new technological achievements.
    - A new lab including hpc computers, in collaboration with the CINECA (<http://www.cineca.it/en>), the italian supercomputing infrastructure supporting the research activities of the italian scientific community, has been activated since 2015.
    - Since 2017 the course is taught in English.
    - Moreover, a new complementary course on “Big Data Analysis” has been activated since 2015 (held by professor Francesco Guerra of the DBGROUP).
  - On 2015/2016 I activated a course named “Not only Technology” to improve the engineering students knowledge on soft-skills, where the students are exposed to professors and domain experts of Law, Business, Organization, Sociology and their vision/needs with reference to information technology.
  - More information on courses held in previous years at [www.dbgroup.unimore.it](http://www.dbgroup.unimore.it).

### Experience of Technological Transfer of Research Results and professional awards

- Sonia Bergamaschi was the promoter of the academic (UNIMORE) start-up “DataRiver” ([www.datariver.it](http://www.datariver.it)), founded on 2009 with the aim of engineering and delivering an open source version of the MOMIS data integration system (see in [IBCH8] the system architecture).
- From then to now she plays the role of scientific responsible of the company. DATARIVER developed in the last years another important activity track on Clinical Data Management. Now DATARIVER is a successful small innovative company employing 11 person full time, 2 consultants for the activities related to privacy, security and national & international contractual rules and the five founding partners of the company. It was accredited as Research Innovation Laboratory by the Emilia Romagna region and certified on 2017 as a CRO (Contract Research Organization), i.e. an organization that provides support to the pharmaceutical, biotechnology and medical devices industries in the form of research services **outsourced** on a contract basis.

The DataRiver start-up was the winner of relevant grants for innovative Start-Ups:

- Bando "FIT Start-Up" Fondo per l'Innovazione Tecnologica of the Ministero dello Sviluppo Economico, L. 46/82. Research grant of 684.000 € in 3 years (2012-2014) to support the research and development of an advanced integration system for traditional and multimedia data;
- Bando POR FESR 2007-2013 “Sostegno allo start-up di nuove imprese innovative” of the Emilia Romagna Region, Activity I.2.1. Research grant of 44.000 € to support the development of the MOMIS Data Integration System.
- On April 2017 she delivered a course on “Big Data Management and Analysis” for the “Ordine degli Ingegneri”, association of Engineers at Modena.
- On August 2017 she obtained by the Emilia Romagna region the approval and co-financing of an Academy on “Methodologies, Techniques and Tools for Big Data Analysis” in collaboration with CINECA.
- Professional Awards: she was appointed as ACM Distinguished Researcher on November 2016 and IEEE Senior Member on August 2016.
- On 2019 she has been appointed (for three years) as member of the committee of experts of Computer Science for the Heideberg Laureate Forum.

## Significant Technical Contributions

Sonia Bergamaschi research activity has been mainly devoted to Knowledge Representation and Management in the context of Very Large Databases and Knowledge Bases facing both theoretical and implementation aspects as described in the following.

### Databases & Description Logics

Since 1985 she was very active in the area of coupling artificial intelligence (Description Logics) and database techniques to develop Intelligent Database Systems. On this topic very relevant theoretical results have been obtained [IJ3, IJ12].

### Intelligent Information Integration

Since 1997 (and up to now), her research efforts have been devoted to the Intelligent Information Integration (I3) topic. An I3 system, called **MOMIS**, providing an integrated access to structured and semistructured data sources and allowing a user to pose a single query and to receive a single unified answer has been designed and developed. Description Logics, Automatic Annotation of schemata plus clustering techniques constitute the theoretical framework. The seminal papers were: [IJ8, IC35].

**MOMIS** development started on 1999 and was carried out with the contribution of the DBGROUP members and of many phd and master students with the financial support up to 3million euros of many EU and MIUR Italian research projects (for details on projects see the site: [www.dbgroup.unimo.it](http://www.dbgroup.unimo.it) ).

On 2009 Sonia Bergamaschi founded the academic start-up **DataRiver** ([www.datariver.it](http://www.datariver.it)), whose aim was the engineering and delivering of an open source version of the **MOMIS system** (see in [IBCH8] the system architecture).

### Semantic Web, Automatic Annotation, Linked Open Data

Since 2002 (and up to now) her research activities have been extended to the **Semantic Web** research area, including MOMIS as the main architectural component of Semantic Search engines [IBCH6]. On Automatic annotation of schemata and inter-schema mapping discovery, very relevant theoretical results have been obtained [IJ24, IC84, IC101, IJ42, IJ43, IJ45].

Since 2013 (and up to now) her research activities have been extended to the Linked Open Data research area. A semi-automatic methodology for facilitating resource providers in publishing public data into the Linked Open Data (LOD) cloud, and for helping consumers (companies and citizens) in efficiently accessing and querying them was designed. To significantly increase the number of consumers of LOD data, a new method and a tool, LODeX have been devised. LODeX is able to automatically provide a high level summarization of a LOD dataset, including its inferred schema, and a powerful visual query interface to support users in querying/analyzing the dataset. (see <http://www.dbgroup.unimo.it/site2012/index.php/component/content/article/2-uncategorised/97-lod> for more details)

### Keyword Search

Since 2009 (and up to now) her research activities has been extended to the Keyword Search research area.

On Keyword Search very relevant theoretical results have been obtained [IC85, IJ25, IJ39, [IJ40] and a tool, Keymatic (<http://www.dbgroup.unimo.it/keymantic>), performing rewriting of keyword queries over relational databases has been implemented. On years 2014-2016 she is participating at the IC1302 COST Action "KEYSTONE - Semantic keyword-based search on structured data sources" (2014-2017) (See <http://www.dbgroup.unimo.it/site2012/index.php/component/content/article/2-uncategorised/96-keywordsearch>)

### Big Data & Big Data Integration

Since 2013 (and up to now) her research activities has been extended to the Big Data Research Area and very relevant theoretical results have been obtained, She is member and reference person for UNIMORE of the CINI Big Data Lab.

On 2014 and 2015 she was a co-organizer and teacher of the "Emerging Tools and techniques for massive data

analysis" schools promoted by Cineca (<https://www.cineca.it/en>).

She was invited speaker on Big Data Management and author of papers in International Conferences (see <http://www.dbgroup.unimo.it/site2012/index.php/component/content/article/2-uncategorised/94-bigdata> for more detailed information)

On 2019 She was the promoter with CINECA of the “Academy Big Data” promoting the higher training course “Tools and Techniques for massive data” (<https://dbgroup.ing.unimore.it/AcademyBigData/>).

Since 2017 she and the DBGroup research addresses issues related to the storage, management, querying, and analysis of streaming, heterogeneous, and distributed data. Big Data Integration techniques and entity recognition - the process for identifying the same real-world entity in different data sources - are the main current research interests [IJ41, IJ46, IJ47, IJ48, IJ4, IJ52, IJ53, IJ54]. The approach to data analysis is based on AI techniques, in particular the so-called Data-Centric AI. We are part of the new international research movement that proposes a new perspective of using Machine Learning (ML), that is MLOps. The goal of MLOps is to make high quality data available through all stages of the ML project lifecycle. MLOps tools **are needed to make** Data-Driven AI an efficient and systematic process.

### **Digital Maktaba**

Since 2022 she was involved in the design and implementation of a system to manage, catalogue, and share historical heritage in non-Latin alphabets, within the ITSEERR PNRR. Significant research results have been obtained so far [IJ51, IC30, IC35].

She published more than three hundreds of international journal and conference papers and her researches have been funded by the Italian MURST, CNR, ASI institutions and by the European Community.

**For a detailed description of the research activity see: [www.dbgroup.unimore.it](http://www.dbgroup.unimore.it)**

For a complete list of publications see:

DBLP <http://dblp.uni-trier.de/pers/hd/b/Bergamaschi:Sonia>

Google Scholar

<https://scholar.google.it/citations?user=AECQn1MAAAAJ&hl=it>

Scopus

<https://www.scopus.com/authid/detail.uri?origin=resultslist&authorId=7006782238&zone=>

Bibliometric Indexes at 8 November 2023

SCOPUS H-Index: 26

Google Scholar H-Index: 36

### **Substantial professional contributions**

She has served on the committees of international and national Database and AI journals and conferences (SIGMOD, VLDB, IJCAI, KR, ER, OODBASE, SEBD, ACM TODS, ACM TOIS, Information Systems, IEEE-ICDE, IEEE TKDE, ESWC, CIKM) and was co-chair in several workshops and conferences (DBISP2P, AP2PC, SWAE), Phd symposium chair for the 2012 E/R conference and Program Committee Chair of the ODBASE 2012 conference. General Chair of SEBD 2018 and treasurer of CIKM 2018.

### **International and National Research Evaluation**

- On research evaluation, she has a long record of local experiences as a supervisor of many phd students and a national/international experience as member of the final examination commissions of PHD students in Italy (Roma, Bologna, Trento), Spain, England.
- She supported: the Ireland Ministry of Research on 2008, 2012 and 2015, 2017 for evaluating their 5-years research projects on Computer Science and Engineering, the Belgian Ministry on 2007 and the Cech Republic Ministry on 2010 and 2015. She was member of the Computer Science Reviewer Panel for The Hildeberg Laureate Forum on 2019, 2020, 2021, 2022.

- She supported the Finland Academy of Science for evaluating Computer Science projects on March 2017 and January 2020.
- She was reviewer of the IST projects “SEVENPRO” and “SMARTMUSEUM” for the European Commission.
- She served as MIUR evaluator for the 2001/2003 Italian VQR and for PRIN 2009.
- She was reviewer for the Evaluation of Industrial projects for the Provincia Autonoma di Trento on 2015 and 2017
- On 2018 she was appointed member of the Commission (composed of three) of Experts for the evaluation of the MIUR PNR 2015-2020 Industrial Research projects in the specialization area "Smart Secure and Inclusive Communities"
- On 2018 she participated in the peer review of the scientific proposals submitted to the ERC Consolidator Grant 2018 Call
- On 2018 she was appointed by the Minister of the Italian Public Administration as member of the Commission (composed of three) of Experts for the evaluation of nominations for the position of Director General of the Agency for Digital Italy.

**Selected Publication**            IJ = Intern. Journal      IC= Intern. Conference

[IJ3] S. Bergamaschi, C. Sartori, "On taxonomic reasoning in conceptual design", ACM Transaction on Database Systems, Vol. 17, No. 3, Settembre 1992, pagg. 385-422.

[IJ12] D. Beneventano, S. Bergamaschi, C. Sartori: "Description Logics for Semantic Query Optimization in Object-Oriented Database Systems", ACM Transaction on Database Systems, Volume 28: 1-50 (2003).

[IJ8] S. Bergamaschi, S. Castano e M. Vincini “Semantic Integration of Semistructured and Structured Data Sources”, ACM SIGMOD Record Special Session on Semantic Interoperability in Global Information, Vol. 28, n.1, March 1999.

[IC35] D. Beneventano, S. Bergamaschi, S. Castano, A. Corni, R. Guidetti, G. Malvezzi, M. Melchiori e M. Vincini: “Information Integration: the MOMIS Project Demonstration”, International Conference on Very Large Data Bases (VLDB'2000), Cairo, Egypt, Settembre 2000. & ACM SIGMOD Digital Symposium Collection 2000 (DiSC2000) (<http://www.cs.wpi.edu/~ifc/DISC>)

[IBCH6] D. Beneventano, S. Bergamaschi: "Semantic search engines based on data integration systems". In Semantic Web Services: Theory, Tools and Applications. Idea Group Publishing, 2007.

[IBCH8] S. Bergamaschi, D. Beneventano, F. Guerra, M. Orsini, “Data Integration”, In: D. W. Embley, B. Thalheim, Handbook of Conceptual Modeling. pp. 441-476, 2011, Berlin Heidelberg: Springer Verlag.

[IJ24] S. Sorrentino, S. Bergamaschi, M. Gawinecki, L. Po “Schema Label Normalization for Improving Schema Matching”, Data & Knowledge Engineering, vol. 69, no. 12, pp. 1254-1273, 2010, ISSN: 0169-023X.

[IC84] S. Bergamaschi, M. Gawinecki, Sorrentino S. (2011). NORMS: an automatic tool to perform schema label normalization. In: ICDE- International Conference on Data Engineering- 2011. Hannover, Germany, 11-16 Aprile 2011, p. 1344-1347, Los Alamitos:IEEE Computer Society 2011, ISBN: 9781424489589.

[IC85] Sonia Bergamaschi, Elton Domnori, Francesco Guerra, Raquel Trillo Lado, Yannis Velegrakis (2011): Keyword search over relational databases: a metadata approach. SIGMOD Conference 2011: 565-576.

[IJ25] S. Bergamaschi, E. Domnori, F. Guerra, M. Orsini, R. Trillo Lado, Y. Velagrakis, “Keymantic: Semantic Keyword-based Searching in Data Integration Systems”, Proceeding of the VLDB Endowment, vol. 3, no. 2, pp. 1637-1640, 2010, ISSN: 2150-8097.

[IJ38] Sonia Bergamaschi, Riccardo Martoglia, Serena Sorrentino: “ Exploiting semantics for filtering and searching knowledge in a software development context”. Knowl. Inf. Syst. 45(2): 295-318 (2015)

[IJ39] Sonia Bergamaschi, Francesco Guerra, Matteo Interlandi, Raquel Trillo Lado, Yannis Velegrakis: “Combining user and database perspective for solving keyword queries over relational databases”. Inf. Syst. 55: 1-19 (2016).

[IC101] Lorenzo Albano, Domenico Beneventano, Sonia Bergamaschi: “Multilingual Word Sense Induction to Improve Web Search Result Clustering”. WWW (Companion Volume) 2015: 835-839.

[IJ41] Giovanni Simonini, Sonia Bergamaschi, H.V. Jagadish (2016): “BLAST: a Loosely Schema-aware Meta-blocking Approach for Entity Resolution”. PVLDB Vol 9.