

Preface to the proceedings of the 20th conference on Information and Research science Connecting to Digital and Library science (IRCDL 2024)

Eleonora Bernasconi¹, Andrea Mannocci², Antonella Poggi³, Angelo Salatino⁴ and Gianmaria Silvello⁵

¹University of Bari, Italy

²CNR-ISTI, Pisa, Italy

³“La Sapienza” University of Rome, Italy

⁴KMi, The Open University, UK

⁵University of Padova, Italy

Abstract

The IRCDL 2024 conference, marking its 20th edition since its inception in 2005, celebrates two decades of advancements in the field of Digital Libraries (DL). Originating at the University of Padua, the conference has traversed various locations, embodying the evolution of DL over time. The 20th-anniversary edition featured a special panel titled “20 Years of IRCDL,” where Prof. Maristella Agosti explored the history of Information Retrieval in the DL landscape. Prof. Floriana Esposito emphasized the pivotal role of Machine Learning in DL and IRCDL. At the same time, Prof. Domenico Saccà provided insights into the significance of databases in Italy and within DL, focusing on structured, semi-structured, and unstructured data.

The conference addressed diverse topics, including applications of DL, machine learning in research data, cultural heritage analysis, data citation and provenance, digital preservation, document analysis, knowledge acquisition, user experience, and more. These topics underscored the multidisciplinary nature of IRCDL and its role in shaping the future of DL.

Keywords

Artificial Intelligence, Cultural Heritage, Digital Libraries, Information Retrieval, Databases

Preface

Since its inception in 2005, the Italian Research Conference on Digital Libraries (IRCDL) has annually served as a pivotal gathering for scholars engaged in Digital Libraries and associated

20th conference on Information and Research science Connecting to Digital and Library science (IRCDL 2024, <https://ircdl2024.dei.unipd.it/>). 22-23 February 2024, Bressanone-Brixen, Italy

Corresponding author.

✉ eleonora.bernasconi@uniba.it (E. Bernasconi); andrea.mannocci@isti.cnr.it (A. Mannocci); antonella.poggi@uniroma1.it (A. Poggi); angelo.salatino@open.ac.uk (A. Salatino); gianmaria.silvello@unipd.it (G. Silvello)

🌐 <https://www.uniba.it/it/docenti/eleonora-bernasconi> (E. Bernasconi); <https://andremann.github.io> (A. Mannocci); <https://salatino.org> (A. Salatino); <http://www.dei.unipd.it/~silvello/> (G. Silvello)

🆔 0000-0003-3142-3084 (E. Bernasconi); 0000-0002-5193-7851 (A. Mannocci); 0000-0002-4030-3458 (A. Poggi); 0000-0002-4763-3943 (A. Salatino); 0000-0003-4970-4554 (G. Silvello)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

domains. Organized by the Italian Research Community, IRCDL has evolved into a prominent national symposium addressing technical, practical, and social facets. Furthermore, it has garnered contributions from international researchers across Europe and beyond. To enhance its global and interdisciplinary character, IRCDL has transformed and now adopts the thematic title “Information and Research Science Connecting to Digital and Library Science.”

IRCDL comprehensively explores the diverse facets of digital libraries, encompassing emerging forms of information institutions, operational information systems housing various digital content, innovative methodologies for selecting, collecting, organizing, and disseminating digital materials, as well as theoretical frameworks for information media, including document genres and electronic publishing. Digital Libraries (DLs) are perceived as novel information institutions and extensions of libraries’ services. The conference invites participation from academia, government, industry, research communities, research infrastructures, and other stakeholders. It draws upon a wide-ranging and multidisciplinary spectrum of research areas: computer science, information science, librarianship, archival science and practice, museum studies and practice, technology, social sciences, cultural heritage and humanities, and scientific communities.

In response to profound transformations within the realm of DLs, IRCDL has recently expanded its scope to embrace these changes and strategically shape the future of DLs. The objective is to engage new research communities and stakeholders, establishing connections between the expansive Research and Information Science fields and DL-related disciplines. Historically, IRCDL has embraced the overarching theme of “Digital Libraries,” encompassing three key areas of scholarly communication (e.g., research data, research software, digital experiments, digital libraries) and e-science/computationally. This strategic approach positions IRCDL at the forefront of addressing the evolving landscape of DLs and fostering collaborative endeavors across diverse domains.

IRCDL positions itself as a pivotal national component within the Digital Library (DL) ecosystem, which encompasses renowned events like the European DL conference known as *Theory and Practice of Digital Libraries* (TPDL), the *International Conference on Asia-Pacific Digital Libraries* (ICADL), the *ACM/IEEE Joint Conference on Digital Libraries*, and the *International Journal of Digital Libraries*. It also maintains affiliations with various initiatives, including the *Italian Workshop on Artificial Intelligence for Cultural Heritage* (IAI4CH), organized by the Italian Association for Artificial Intelligence.

The **IRCDL 2024** conference, accessible at <https://ircdl2024.dei.unipd.it/>, represents a milestone as it commemorates the 20th edition of the conference, which originated in 2005 (<http://ims.dei.unipd.it/websites/ircdl/home.html>). The inception of IRCDL was spearheaded by Maristella Agosti (University of Padua) and Costantino Thanos (ISTI-CNR, Pisa). The conference was initially hosted at the Department of Information Engineering of the University of Padua from 2005 to 2010, after which it transitioned to various locations, including Pisa, Bari, Rome, Bolzano, Firenze, Modena, and Udine.

In commemoration of IRCDL’s 20th anniversary, the steering committee orchestrated a special panel titled “20 Years of IRCDL,” skillfully moderated by Costantino Thanos (ISTI-CNR) of Pisa. The distinguished panelists included Prof. Maristella Agosti from the University of Padua, who provided insights into the history and the significance of Information Retrieval, especially within the Digital Library (DL) landscape. Prof. Floriana Esposito from the University

of Bari delved into the realm of Machine Learning, elucidating its pivotal role in both DL and IRCDL. Lastly, Prof. Domenico Saccà from the University of Calabria offered a comprehensive perspective on the role of databases in Italy and within the DL context, with a particular focus on the significance of structured, semi-structured, and unstructured data.

The topics of interest for IRCDL 2024 were:

- Applications of Digital Libraries
- Applications of Machine Learning Techniques to Research Data and DL
- Citation Analysis and Scientometrics
- Cultural Heritage Access and Analysis
- Data Citation, Provenance, and Pricing
- Data and Information Lifecycle (creation, store, share, and reuse)
- Data and Metadata Quality
- Data Repositories and Archives
- Digital Preservation and Curation
- Document Analysis (Layout, Text, Images)
- Knowledge Acquisition from Scientific Papers
- Knowledge Discovery and Representation in Digital Libraries
- Human-Computer Interaction and User Experience
- Information Extraction from Tables and Figures in Scientific Literature
- Information Retrieval and Access
- Metadata (definition, management, curation, integration)
- Multi-media Handling
- Ontologies
- Open Data and FAIR Data
- Open Science: Models, Practices, Mandates, and Policies
- Quality and Evaluation of Digital Libraries
- Research Infrastructures
- Scholarly Communication
- Semantic Web Technologies and Linked Data for DLs
- Services for Digital Arts and Humanities
- Standards and Interoperability
- User Participation

IRCDL 2024 accepted full papers (8-10 pages) and short papers (6 pages) suitable for presenting work in progress, software prototypes, or extended abstracts of doctoral theses. At least two program committee members reviewed all the 21 accepted papers, of which 10 long and 11 short.

The first day commenced with an opening session featuring two presentations on Nanopublications. The first, titled *Building and Exploiting a Web of Machine-Readable Scientific Facts to Make Discoveries*, was delivered by *Nicola Raffaele Di Matteo* (Dalhousie University), *Andrea Schimmenti* (University of Bologna), and *Fabio Vitali* (University of Bologna). The second presentation, *Publishing CoreKB Facts as Nanopublications*, was given by *Fabio Giachelle* (University

of Padua), *Stefano Marchesin* (University of Padua), *Laura Menotti* (University of Padua), and *Gianmaria Silvello* (University of Padua).

Following a coffee break, the day continued with an invited talk by *Sara Tonelli* (Bruno Kessler Foundation) on *Toxic Language Detection: Why should computers decide when humans disagree?*

The afternoon session explored LLMs and NLP-related techniques, covering topics such as exploiting large language models, resolving duplicates in large multiple-choice question repositories, evaluating differential privacy approaches for information retrieval, and addressing file format conversion issues in digital libraries.

The afternoon continued with a panel on *20 years of IRCDL*, moderated by *Costantino Thanos*, featuring panelists *Maristella Agosti*, *Floriana Esposito*, and *Domenico Saccà*.

The final session on Knowledge Representation and Knowledge Management included presentations on *A Platform To Communicate Knowledge: K-COM* by *Giovanni Addamo* (ENEA), *Fabio Cignini* (ENEA), *Enrico Cosimi* (ENEA), *Vittoria Cozza* (University of Verona), *Giorgio Ponzio* (ENEA), *Maria Salvato* (ENEA), and *Veronica Tomassetti* (ENEA).

Additionally, *Evaluation of Expressing without Asserting Approaches in RDF. The Case of Conjectures* was discussed by *Valentina Pasqual* (University of Bologna), *Gerald Manzano* (University of Bologna), *Eduart Uzeir* (University of Bologna), *Francesca Tomasi* (University of Bologna), and *Fabio Vitali* (University of Bologna).

The session concluded with *A Holistic Ontology for Digital Libraries* presented by *Stefano Ferilli* (University of Bari), *Eleonora Bernasconi* (University of Bari Aldo Moro), *Davide Di Pierro* (Università degli Studi di Bari), *Domenico Redavid* (University of Bari), and *Liza Loop* (Loop Center).

The second day featured sessions on Cultural Heritage (CH) Case Studies, Open Science & Education, and Challenges in Library Science. In the CH Case Studies section, the presentations included “A tool for empowering Symbol Detection through Technological Integration in Library Science: A case study on the Voynich manuscript” by *Eleonora Bernasconi* (University of Bari) and *Stefano Ferilli* (University of Bari), “Ontofest: An Ontology to Integrate and Retrieve Data from the Locarno Film Festival Archives” by *Alessandro Cosentino* (Università della Svizzera italiana), *Webert Júnio Araújo* (Universidade Federal de Minas Gerais), and *Fabio Crestani* (Università della Svizzera italiana), “A Novel Methodology for Topic Identification in Hadith” by *Sania Aftar* (University of Modena and Reggio Emilia), *Luca Gagliardelli* (University of Modena and Reggio Emilia), *Amina El Ganadi* (University of Modena and Reggio Emilia), *Federico Ruoizzi* (University of Modena and Reggio Emilia), and *Sonia Bergamaschi* (Università di Modena e Reggio Emilia), and “The MAGIC project: first research results” by *Stefania Conte* (University of Naples Federico II), *Gian Marco Di Domenico* (SA Documents S.r.l.), *Andrea Mazzei* (SA Documents S.r.l.), *Andrea Mazzucchi* (University of Naples Federico II), *Guido Russo* (Università di Napoli Federico II), *Alessandro Salvi* (Netcom Group S.p.a.), and *Augusto Tortora* (University of Naples Federico II).

After a coffee break, the Open Science & Education session featured presentations on “An Overview of Open Science in Italy” by *Alessia Bardi* (CNR-ISTI), *Leonardo Candela* (CNR-ISTI), *Dario Mangione* (CNR-ISTI), and *Gina Pavone* (CNR-ISTI), “Research and Teaching Public Communication of Science and Technology on Digital Data” by *Emanuele Di Buccio* (University of Padua) and *Federico Neresini* (University of Padua), “Mapping bibliographic metadata collections: the case of OpenCitations Meta and OpenAlex” by *Elia Rizzetto* (University of Bologna) and

Silvio Peroni (University of Bologna), “What motivates university students to engage in learning computer skills?” by *Andrea Micheletti* (University of Padua) , *Nicola Orio* (University of Padua), and *Daniel Zilio* (University of Padua), and “Paving the Way for Personalized Museums Tours in the Metaverse” by *Alex Falcon* (University of Udine) , *Beatrice Portelli* (University of Udine) , *Ali Abdari* (University of Udine), and *Giuseppe Serra* (University of Udine).

After a lunch break, the Challenges in Library Science session included presentations on “HeRO: A Semantic Framework for Heritage Risk Assessment in the SIRIUS Project” by *Sebastian Barzaghi* (University of Bologna) , “A Geospatial and Time-based Reconstruction of the Venetian Lagoon in a 3D Web Semantic Infrastructure” by *Ludovica Galeazzo* (Università degli Studi di Padova) , *Remo Grillo* (I Tatti - The Harvard University Center for Italian Renaissance Studies), and *Gianmarco Spinaci* (Villa I Tatti), and “The role of libraries for PDA” by *Adele Gorini* (Università di Bologna). The day concluded with a closing session.

Conference Organization

General Chair

Antonella Poggi, “La Sapienza” University of Rome, Italy

Program Chairs

Eleonora Bernasconi, University of Bari A. Moro, Italy

Andrea Mannocci, ISTI-CNR, Italy

Proceedings Chair

Angelo Salatino, The Open University, Milton Keynes, UK

Local Organization

Davide Lanti, University of Bozen, Italy

Gianmaria Silvello, University of Padua, Italy

Publicity and Web Chair

Fabio Giachelle, University of Padua, Italy

Program Committee

Massimiliano Assante, CNR-ISTI

Miriam Baglioni, CNR-ISTI

Nicola Barbuti, University of Bari – A. Moro

Valentina Bartalesi, CNR-ISTI

Miguel Ceriani, CNR-ISTC

Marilena Daquino, University of Bologna

Giorgio Maria Di Nunzio, University of Padova
Guglielmo Faggioli, University of Padova
Alex Falcon, University of Udine
Stefano Ferilli, University of Bari – A. Moro
Amel Fraise, University of Lille
Donatella Firmani, “La Sapienza” University of Rome
Fabio Giachelle, University of Padova
Ornella Irrera, University of Padova
Davide Lanti, Free University of Bozen-Bolzano
Francesco Leotta, Sapienza University of Roma
Paolo Manghi, CNR-ISTI
Simone Marinai, University of Firenze
Riccardo Martoglia, University of Reggio Emilia
Alessandro Mosca, Free University of Bozen-Bolzano
Viviana Patti, University of Torino
Beatrice Portelli, University of Udine
Domenico Redavid, University of Bari A. Moro
Rossana Damiano, University of Torino
Martin Ruskov, University of Milano
Giuseppe Serra, University of Udine
Luigi Siciliano, Free University of Bozen-Bolzano
Angelo Salatino, Open University
Gianmaria Silvello, University of Padova
Manuel Striani, University of Piemonte Orientale
Anna Maria Tammaro, University of Parma
Maria Theodoridou, ICS-FORTH
Thanasis Vergoulis, IMSI, “Athena” RC