Proceedings of the 9th International Workshop on Semantic Web Solutions for Large-scale Biomedical Data Analytics - SeWebMeDa-2026

Rita T. Sousa[1], Michel Dumontier[2] and Ali Hasnain[3]

[1]Data and Web Science Group, University of Mannheim, Germany
[2]Maastricht University, Maastricht, Limburg, Netherlands
[3]Royal College of Surgeons, Ireland

Abstract

This preface summarises the 9th International Workshop on Semantic Web Solutions for Large-scale Biomedical Data Analytics (SeWebMeDa-2026), a co-event with The ESWC 2026: Extended Semantic Web Conference, held on May 10 2026, in Dubrovnik, Croatia.

1. Introduction

The explosive growth of complex biomedical and life sciences data presents both unprecedented opportunities and significant challenges for data analysis. Researchers now have access to a vast spectrum of data sources, including high-throughput sequencing, multi-omics platforms, electronic health records, imaging data, wearable devices, among others. However, given the heterogeneous nature of these data, meaningfully integrating, analyzing, and interpreting them remains a significant hurdle. Linked data and semantic web technologies offer a powerful solution to address these challenges, providing the tools to effectively harmonize and integrate diverse data sources. By leveraging linked data, ontologies, and knowledge graphs (KGs), biomedical data can be transformed into structured knowledge that supports advanced queries, AI-driven discovery, and decision-making. Furthermore, recent developments in machine learning (ML), generative AI, and neurosymbolic approaches offer powerful new tools to uncover hidden patterns, provide predictive insights, and support explainable applications in healthcare and life sciences.

The goal of the ninth edition of the International Workshop on Semantic Web solutions for large-scale biomedical data analytics (SeWeBMeDA) is to bring together researchers working at the intersection of biomedical sciences and semantic web technologies. SeWeBMeDA provides a forum for exploring novel approaches that utilize semantic web technologies and linked data for integrating, representing, and analyzing biomedical data, with a focus on scalable infrastructure, knowledge representation, and advanced analytics that support both research and clinical applications.

This workshop invites original contributions that address both technical infrastructure and practical applications. It covers the distributed infrastructure for consuming, storing, and querying large volumes of heterogeneous linked data, utilizing indexes and graph aggregation to better understand large linked data graphs, query federation to integrate internal and external data sources, and linked data visualization tools for healthcare and life sciences. It further covers topics around data integration, data profiling, data curation, querying, knowledge discovery, ontology mapping / matching / reconciliation and data / ontology visualization, applications / tools / technologies / techniques for life sciences and the biomedical domain. Beyond its technical scope, SeWeBMeDA also seeks to address emerging societal and ethical challenges. This workshop aims to provide researchers in biomedical and life sciences

SeWebMeDA-2026: 9th International Workshop on Semantic Web Solutions for Large-scale Biomedical Data Analytics, May 10, 2026, Dubrovnik, Croatia
alihasnain@rcsi.ie (A. Hasnain)
https://www.rcsi.com/people/profile/alihasnain (A. Hasnain)
0000-0002-7241-8970 (R. T. Sousa); 0000-0003-4727-9435 (M. Dumontier); 0000-0003-4014-4394 (A. Hasnain)

with an insight and awareness about large-scale data technologies for linked data, which are becoming increasingly important for knowledge discovery in the life sciences domain.

To guide submissions, the workshop is structured around (but not limited to) the following thematic areas and topics of interest:

1. **Data Integration**
   - 1.1 Dataspaces, data warehouses, and database solutions in healthcare and life sciences
   - 1.2 Large-scale curation, integration, processing, and analysis of heterogeneous biomedical data
   - 1.3 Cleaning, quality assurance, and provenance tracking for life sciences data
   - 1.4 Biomedical data quality assessment and improvement
   - 1.5 Implementation, governance, and assessment of FAIR data principles in life sciences
   - 1.6 Data and metadata publishing, profiling, and discovery of new biomedical datasets
   - 1.7 Data streams, Internet of Things, mobile platforms, cloud environment in life sciences
2. **Knowledge Representation**
   - 2.1 Biomedical ontology creation, mapping/ matching/ translation and reconciliation
   - 2.2 Building and maintaining KGs in healthcare and life sciences
   - 2.3 KG enrichment using text mining and natural language processing techniques
   - 2.4 Visualization and exploration of linked data, ontologies, and KGs in life sciences
3. **AI/ML/Reasoning over Ontologies and KGs**
   - 3.1 ML and relational learning over biomedical KGs
   - 3.2 Neurosymbolic AI and hybrid reasoning approaches for biomedical data
   - 3.3 Large language models integrated with biomedical KGs
   - 3.4 Explainable AI approaches leveraging semantic technologies in life sciences
4. **Applications and Use Cases**
   - 4.1 Semantic technologies supporting research and clinical applications in life sciences
   - 4.2 Querying and federating data over heterogeneous data sources
   - 4.3 Virtual and augmented reality in life sciences education and applications
   - 4.4 Generative AI and conversational AI applications in life sciences
5. **Ethical, Social, and Practical Considerations**
   - 5.1 Risks and opportunities of using Semantic Web and AI technologies in life sciences
   - 5.2 Bias and fairness in AI systems for healthcare and life sciences
   - 5.3 Challenges of explainability and accountability in AI systems for healthcare and life sciences
   - 5.4 Responsible development of neuro-symbolic and generative AI in clinical workflows

The workshop also aligns closely with the main conference's topics, fostering cross-community collaboration. In this edition, in addition to established topics, SeWeBMeDA introduces emerging topics, including explainable AI and generative AI, ensuring that participants engage with the latest trends and practical solutions.

2. Workshop Format

Due to the emerging interest in the application of semantic technology in the biomedical domain, we expect that the semantic web community will be interested in this workshop and anticipate between 30 and 40 attendees.

The workshop will be organized as either a half-day or full-day event, depending on the number of accepted papers. The program will include a keynote presentation by an invited, established researcher from the Semantic Web/Bioinformatics community. This will be followed by oral presentations of the accepted papers, with 15 minutes allocated for each full paper and 5–7 minutes for short paper

presentations. At the end of each presentation, an open Q&A will provide an opportunity for participants to exchange ideas and insights. The workshop will conclude with an interactive session designed to foster collaboration and networking. If possible, this final segment may take the form of a panel discussion involving the speakers or a poster session.

The workshop will accept the following types of submissions:

- **Scientific Papers**: Long or short papers presenting novel research relevant to the workshop topics. Long papers should be 10-15 pages in length, while short papers should be 8-10 pages.
- **Demo Papers**: Submissions showcasing technical advances or innovative tools in the biomedical and life sciences domains. Demo papers must not exceed 4 pages.
- **Position Papers**: Submissions that identify open challenges, or propose new directions relevant to the workshop topics. Position papers should be 4–6 pages in length.
- **Work-in-progress Papers**: Papers reporting preliminary, ongoing, or negative results that can foster discussion and feedback. These papers should be 5–10 pages in length.
- **European/ International Project Showcase**: Papers presenting and disseminating results from ongoing or recently completed projects (within the last 24 months). These submissions should be 6–8 pages in length.

All submissions will be peer-reviewed by three to four members of the Program Committee. Accepted papers will be included in the workshop proceedings. The call for papers will be announced on December 15, with the submission deadline set for March 3. Authors will be notified of acceptance by March 31, and the camera-ready versions of accepted papers will be due by April 15.

3. Organisation

3.1. Workshop Chairs

- Ali Hasnain, Royal College of Surgeon, Ireland.
- Michel Dumontier, Maastricht University, Maastricht, Limburg, Netherlands.
- Rita T. Sousa, University of Mannheim, Germany.

3.2. Programme Committee

- Michel Dumontier, Maastricht University, Netherlands
- Qurratal Ain Fatimah, University Hospital Galway, Ireland
- Maria-Esther Vidal, Leibniz Universität Hannover, Germany
- M. Scott Marshall, The Netherlands Cancer Institute, Netherlands
- Pierre-Yves Vandenbussche, eBay Inc, Virginia Tech, USA
- Holger Stenzhorn, Saarland University Hospital, Germany
- Antoine Zimmermann, École des Mines de Saint-Étienne, France
- Muhammad Saleem, AKSW- University of Leipzig, Germany
- Mikel Egaña Aranguren, University of Basque Country, Spain
- Helena F. Deus, Elsevier Labs, USA
- Adrien Coulet, Université Paris Cité, France
- Ratnesh Sahay, Novartis, UK
- Dietrich Rebholz-Schuhmann, Information Centre for Life Sciences, Cologne University, Germany
- Sujan Perera, Amazon, USA
- Claudia d'Amato, Università degli Studi di Bari, Italy
- Robert Hoehndorf, King Abdullah University of Science and Technology, Saudi Arabia
- Jodi Schneider, University of Illinois at Urbana-Champaign, USA

- Alasdair Gray, Heriot-Watt University, Edinburgh
- Alba Morales Tirado, The Open University, UK
- Rita T. Sousa, University of Mannheim, Germany