Sara El-Gebali, PhD

Metadata Specialist, Research Data Strategist, and Community Builder

✉ _selgebali@opencider.org_
📞 _+46-(0)70 712 79 55_
🌐 _Nationality: **Swedish**_
📍 _Location: **Gothenburg**_
🏠 Personal Website
FAIRPoints
ⓘD ORCID

DataCite, Germany

Metadata Specialist
Present

- Champion the promotion, refinement, and integration of the DataCite metadata schema across broader PID infrastructures.
- Conduct quality assurance checks and metric analysis to elevate metadata quality.
- Contribute to the strategic planning and concept development within the PID4NFDI project.
- Provide expert training, guidance, and support.

FAIRPoints

Co-founder

- Initiated and currently co-lead FAIRPoints, a pivotal platform fostering discussions and training on FAIR principles globally.
- Organize international webinars and workshops aimed at advancing data management standards implementation.

Current Professional Affiliations

FAIR Digital Objects Forum

Steering Committee Member

- Engage in strategic planning and initiative execution to promote FAIR Digital Objects' adoption in research and data management.
- Develop guidelines and standards to boost digital objects' interoperability, accessibility, and reusability.

Open Life Science

Governance Committee Member

- Advising on policy and direction within Open Life Science to promote ethical and inclusive open science practices.
- Mentor and guide early-career researchers in implementing effective open science principles within their projects.

Africa PID Alliance

Technical Advisor

- Offer strategic and technical advice for PID management and integration across African research institutions, enhancing discoverability and access to the continent's cultural heritage and innovation.

Research Data Alliance (RDA) and European Open Science Cloud (EOSC)

Ambassador

- Represent the interests of the RDA and EOSC, promoting the integration of their standards and practices within various scientific communities across Europe.
- Organize and lead community engagement events that discuss and promote the implementation of FAIR principles across European research communities.

FAIR Impact for Signposting and ROCRATE

Awardee

- Contributed to initiatives aimed at enhancing the findability and accessibility of research outputs through improved signposting and the use of ROCRATE metadata standards.

Previous Experience

SciLifeLab-DataCentre, Sweden

Project Lead, Metadata & Curation

August 2021 - October 2023

- Directed the implementation of multiple open-source technical solutions, notably the Digital Research Hub for Biosamples metadata handling.
- Promoted Open & FAIR practices extensively through strategic outreach and training programs.
- Managed a dynamic team as Agile Product Owner, comprising developers and data stewards.

NASA Transform to Open Science

Open Data Module Lead

June 2022 - November 2022

- Led the development of the Open Data module within the NASA OpenCore project, enhancing data accessibility and openness.
- Courses are in the official offerings for professional development available to the staff of the Executive Office of the President.

Max Delbrück Center for Molecular Medicine (MDC), Berlin

Unit Head- Research Data Management

June 2020 - July 2021

- Established and led the research data management unit, crafting policies and strategies that strengthened research data governance and support.
- Supported and advised researchers, providing training and workshops.

European Bioinformatics Institute (EMBL-EBI), Cambridge, UK

Scientific Database Curator

June 2016 - August 2019

- Managed bioinformatics resources for protein biology at Pfam.
- Curated extensive protein sequence datasets, and engaged in regulatory processes.
- Awarded the Wellcome Genome Campus Award for Best Practice in Supporting Equality and Diversity in Science 2019.

EMBO (European Molecular Biology Organization), Germany

Scientific Data Editor

January 2015 - June 2016

- Conducted in-depth analysis and interpretation of scientific content to create structured and machine-readable metadata.
- Coordinated activities with software and bioinformatics teams to develop databases and user interfaces.
- Developed workflows for editorial and production processes and created new avenues for promoting the SourceData project.

Education

University of Bern, Switzerland

PhD in Molecular Biology

2010 - 2014

Research Focus: Investigated metabolic cell signaling pathways and amino acid transporters in colon cancer progression.

Queen Mary University of London, UK

MSc in Molecular Pathology and Genomics

2007 - 2008

Focus: Explored cancer-related inflammation in skin carcinogenesis. Recipient of the Cancer Research UK bursary.

Lund University, Sweden

BSc in Molecular Biology

2004 - 2007

Coursework: Covered essential topics in cell biology, genetics, microbiology, and molecular biology techniques.

Skills

- **Advanced**: Python, Bash, GIT, ElasticSearch, Kibana, APIs
- **Intermediate**: MySQL, HTML/CSS
- **Project Management Tools**: Proficient in Jira, Confluence
- **Certifications**: Certified Scrum Product Owner (scrum.org), Certified Carpentries Trainer

Languages

- Fluent: English, Swedish, Arabic
- Conversational: German

Conferences and Public Speaking

- **AAAS Conference Panelist:** Actively participated as a panelist, discussing strategies to enhance data inclusivity in scientific research. This involved presenting innovative methods to integrate diverse data sources and promote equitable access to scientific information.
- **Keynote Speaker at Bioinformatics Open Source Conference (BOSC):** Delivered the keynote address on the impact of community-driven development and the potential of open-source resources in fostering innovation and collaboration in the field.
- **Regular Participant at RDA and EOSC Conferences:** Engaged in multiple international conferences organized by the Research Data Alliance and the European Open Science Cloud, speaking on topics such as data management, bioinformatics, and the implementation of FAIR data principles.
- **FOSDEM 2023 Presenter:** Highlighted pragmatic measures for implementing FAIR data principles, discussing their importance in the context of global data sharing and interoperability.

Teaching and Mentoring

- **Bioinformatics Hub of Kenya Initiative (BHKI)**: Led training sessions data management, helping to build research capacity within the Kenyan research community.
- **Research Data Management (RDM) Material on Zenodo:** Developed and shared a comprehensive set of RDM training materials on Zenodo, which are used globally by researchers and data managers to improve their data stewardship practices. These resources include guidelines, best practices, and templates that facilitate the adoption of FAIR principles.
- **Open Life Science (OLS) Mentoring:** Actively mentored several cohorts through the OLS program, guiding participants through the process of establishing and maintaining open science projects. This mentorship focused on empowering researchers with the tools and knowledge needed to effectively implement open science practices.
- **Cambridge University,** Guest lecturer, MSc Courses: Delivered lectures and practical sessions on protein sequence analysis to MSc students at Cambridge University.
- **The Carpentries, Instructor:** Conducted multiple certified Carpentries and Carpentries-style workshops in Basic Bash and Research Data Management.

Volunteer Work and Community Engagement

- **Volunteer at Refugee Camps in Lesvos**, Greece (2020): Contributed to humanitarian efforts by volunteering at refugee camps, providing support and aid to displaced individuals.
- **Establishing Open Science Community Egypt**: Played a pivotal role in establishing the Open Science Community in Egypt. Provided strategic advice and support, facilitating workshops and seminars to train local researchers in open science methodologies, thereby promoting open science practices among Egyptian researchers.

Honors and Awards

- eLife Innovation Leaders 2020 Mentorship Program: Recognized as a mentor in the eLife Innovation Leaders program, providing guidance and support to up-and-coming researchers in open science practices.
- Wellcome Genome Campus Award for Best Practice in Supporting Equality and Diversity in Science (2019): Awarded for outstanding contributions to promoting equality and diversity within the scientific community, highlighting efforts to create an inclusive environment that fosters diversity in scientific research and education.
- Brave the Shave Campaign (2017): Participated in the Brave the Shave campaign, raising significant funds for MacMillan Cancer Support, and also contributed to the Little Princess Trust by donating hair for wigs for children undergoing cancer treatment.

Podcasts and Media Appearances

- **FAIR Data Podcast:** Engaged in a deep dive into the evolving landscape of FAIR data practices, discussing their pivotal role in advancing scientific research. The conversation addressed how these practices enhance the accessibility and usability of research data across various scientific fields. Listen to the episode.

- **Podbean Interview on Open Science:** Explored the crucial aspects of data interoperability and envisioned the future of scientific collaboration. This discussion emphasized the need for seamless data exchange and integration to foster collaborative research efforts globally. [Listen on Podbean](#).
- **RSEng Dev Stories:** Shared personal insights and professional experiences regarding the challenges and triumphs in the realm of data management. This narrative highlighted how strategic data management practices contribute to the efficiency and effectiveness of scientific research. [Read the story](#).
- **Science Business Article:** Discussed the pressing challenges and innovative solutions in data sharing aimed at enhancing open access. The article covers strategic approaches to overcoming barriers to data sharing that can stifle scientific innovation and collaboration. [Read the article](#).
- **RDA EOSC Future Spotlight:** Highlighted my significant contributions to the research data community, emphasizing the impact of my work on promoting and implementing FAIR principles across various research settings. [Read the feature](#).

Selected Publications

- El-Gebali, S.; Stathis, K. (2024, March 7). Harmonizing metadata across disciplines – bioschemas and the datacite metadata schema. DataCite. https://datacite.org/blog/bioschemas_2024/
- Liechti, R., George, N., Götz, L., El-Gebali, S., Chasapi, A., Crespo, I., Xenarios, I., & Lemberger, T. (2017). SourceData: A semantic platform for curating and searching figures. Nature Methods, 14(11), 1021–1022. https://doi.org/10.1038/nmeth.4471
- Stathis, K., Chen, X., & El-Gebali, S. (2024, January 1). DataCite metadata training. Zenodo. https://zenodo.org/doi/10.5281/zenodo.10950234
- Adams, A., Almarzouq, B., Lacerda, A., Alwood, J., Alzahawi, S., Azevedo, F., Barry, R. K., Batalha, N., Bayer, J. M. M., Bell, T., Black, M., Blanchette, K., Bremner, P, Brown, S. V., Bugbee, K., Burt, J., Toro, C., Campitelli, E., Chandler, M., … Yuen, K. (2023, December 6). NASA TOPS open science 101. Zenodo. https://zenodo.org/doi/10.5281/zenodo.10161527