OSCARS
Open Science Clusters' Action for Research & Society

Funded Project

FASCA - Facilitating Scholarly Communication Analysis through GoTriple Pipeline

Presenter: Sy Holsinger, OPERAS AISBL, https://orcid.org/0009-0005-8978-225X
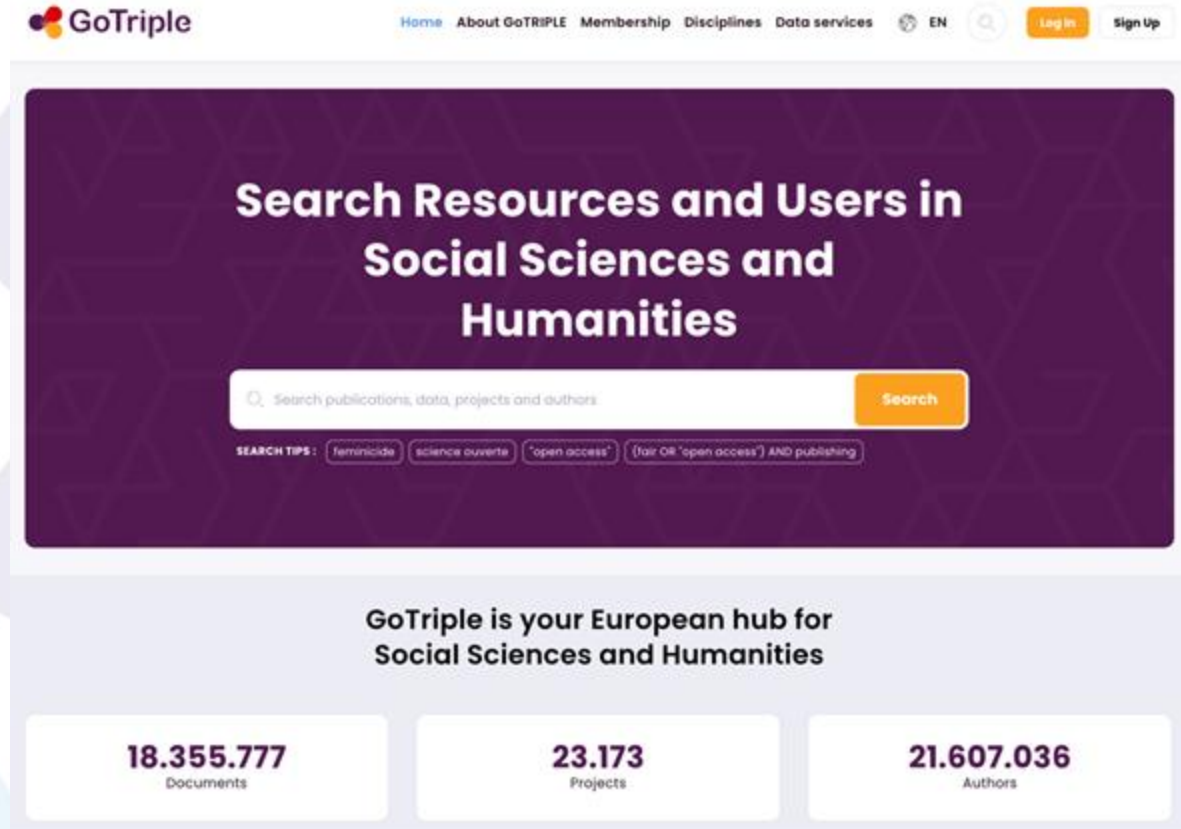
Implemented by

What problem(s) are you going to solve?

- **Traditional methods fail** to fully leverage **large datasets** and **diverse research outputs**, such as survey data, interviews, blogs, articles and books.

- **Navigating multilingual research resources** remains a challenge for SSH scholars, limiting access to global publications, datasets, and archives.

- There is a growing need to **align scholarly communication with global challenges** like the SDGs, including climate change, social inequality and cultural preservation.

- **Lack of efficient data-driven tools** hinders cross-disciplinary collaboration and Open Science practices, restricting the integration of AI, text mining, and semantic search in SSH research.

GoTriple is a **multilingual** platform tailored for the in-depth exploration of Social Sciences and Humanities (SSH) research. It facilitates the **discovery** of scholarly content from a single access point. Researchers can find millions of **publications** and **authors** and thousands of **projects** which are currently scattered across local repositories.

https://www.gotriple.eu/

What are you planning to do to solve the problem?

🔧 **GoTriple Enhancements:**
- Improve API and Pundit annotations.

🛠️ **GoTriple Data Science Pipeline:**
- Integrate: JupyterLab, Neo4j, GitHub, and Zenodo.

🤖 **GoTriple Chatbot:**
- Expand with 20,000 docs and AI-powered multilingual support.

🧪 **Research Pilots:**
- Support 2 pilots to implement and test pipeline along with ensuring FAIR outputs.

FASCA

FACILITATING **SCHOLARLY COMMUNICATION ANALYSIS** THROUGH GOTRIPLE PIPELINE

Final Output: Establish a GoTriple Data Science Support service for multilingual data-driven research

What will be the results and how do you plan to make them available to the broader community?

FASCA has identified four target groups it aims to benefit:

- **SSH researchers**: broadly defined, covering all disciplines in SSH and career levels (from students, to established scholars), and different European Countries.
  - Support for performing data-driven research based on the GoTriple data, novel use of existing tools, and tailored training and collaboration.
- **Data scientists**: scholars working on data science with new digital solutions (e.g. LLMs, Graphs).
  - Opportunity to collaborate on novel research projects, alongside SSH scholars who may lack specific technical skills. Opportunity to enter the GoTriple Data Science team after FASCA's completion (based on the models developed in FASCA's WP4).
- **Business & Industry**: SMEs, startups delivering new content and solutions in e.g. tourism, gaming, education and any other SSH applied area.
  - Innovation support through tailored training and better access to the GoTriple data via API, to deliver novel content or imagine novel solutions for their business needs. Possibilities of knowledge and technology spillovers.
- **Research Performing Organisations (RPOs):**
  - Access to novel solutions, data and training for their researchers.

What will be the results and how do you plan to make them available to the broader community?

- **Publishing & Open Access Dissemination**
  - GoTriple Platform: Access to established community + dedicated webpage showcasing results.
  - Zenodo & GitHub: Research data, publications and software available for reuse.
  - SSHOC Marketplace: Workflows and methodologies documented for SSH researchers.
- **Training & Knowledge Sharing**
  - 30 Free Data Science Consultations for SSH researchers.
  - Workshops & Webinars to train scholars on using the GoTriple Data Science tools.
- **Community Engagement & Outreach**
  - Social Media & Blogs: OPERAS communication channels to reach a broad audience.
  - Conferences & Events: Participation in major SSH and EOSC conferences.
- **Sustainable Impact & Future Plans**
  - Integration into EOSC: Ensuring the GoTriple Data Science Support service remains accessible beyond FASCA.
  - Business Model Development: Exploring sustainability strategies to maintain long-term support for SSH researchers.

What risks could limit the success of the project, and how can they be mitigated

- **Data Processing Challenges**
  - Risk: Research data (publications/papers) are mostly in PDF format, making it difficult to parse and analyze.
  - Mitigation: Develop FASCA's data pipeline and leverage the GoTriple API to support efficient data extraction and parsing.
- **Complexity & Technical Expertise Gaps**
  - Risk: Many SSH researchers lack data science skills (e.g., programming, analytics).
  - Mitigation: Offer technical support, training materials, and active collaboration with data scientists.
- **Usability of FASCA Services**
  - Risk: Integrating new tools into GoTriple might lead to user experience issues.
  - Mitigation: Conduct co-design workshops and iterative usability testing with researchers to refine the interface.
- **Limited Researcher Engagement During Pilots**
  - Risk: SSH researchers involved in pilot projects may face time constraints or technical challenges.
  - Mitigation: Implement agile research processes, break tasks into manageable steps, and provide constant support from FASCA partners.
- **Lack of Collaboration Between SSH & Data Science Experts**
  - Risk: Limited interaction between SSH researchers and data scientists could hinder project success.
  - Mitigation: Organize regular checkpoint meetings, facilitate direct collaboration sessions, and allow role adjustments if bottlenecks occur.

OSCARS

Who is doing it?

Sy Holsinger
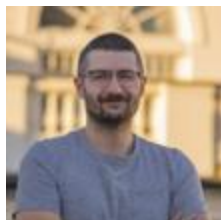
Tomasz Umerle   Cezary Rosinski

Luca De Santis

Julien Homo

Stefano De Paoli

Suzanne Dumouchel   Nikodem Wolczuk   Barbara Wachek

Kevin Darty

Alex Fawzi