



# 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>

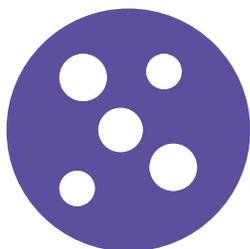
Implemented by



Funded by  
the European Union

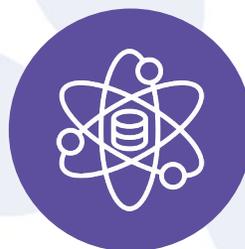
## What problem does FASCA address?

### Scattered multilingual resources



SSH publications, datasets and archives remain fragmented across local repositories, limiting access for scholars.

### No integrated data science tools for SSH



AI, text mining, knowledge graphs and semantic search remain largely out of reach for researchers without technical skills.

### Growing need to connect research



with societal challenges like the SDGs, climate change, social inequality, cultural preservation.

### Barriers to Open Science



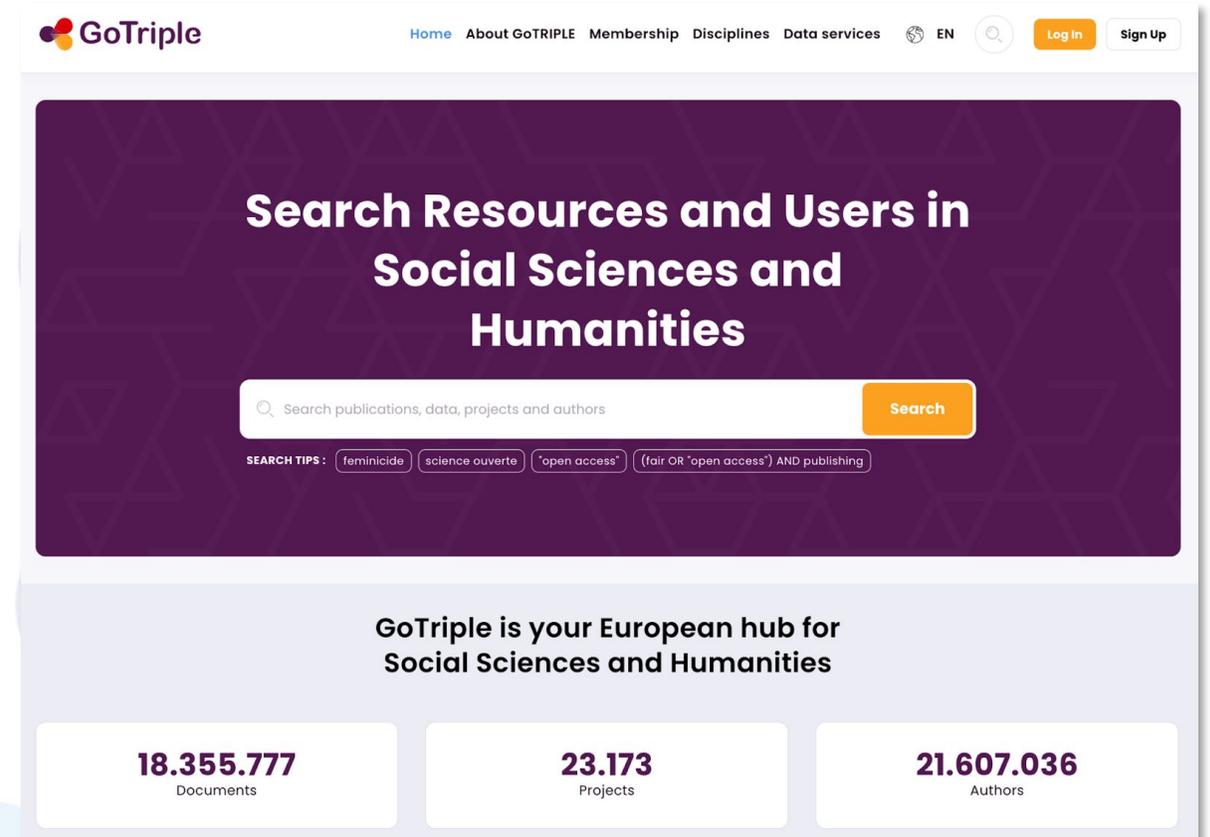
lack of reproducible, data-driven workflows hinders cross-disciplinary collaboration in SSH.

---



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.

FASCA extends GoTriple with integrated data science services for SSH researchers.



The screenshot shows the GoTriple website homepage. At the top, there is a navigation bar with links for Home, About GoTRIPLE, Membership, Disciplines, Data services, EN, and a search icon. There are also buttons for Log In and Sign Up. The main content area has a dark purple background with the text "Search Resources and Users in Social Sciences and Humanities". Below this is a search bar with the placeholder text "Search publications, data, projects and authors" and a "Search" button. Under the search bar, there are search tips: "femicide", "science ouverte", "open access", and "(fair OR 'open access') AND publishing". At the bottom of the page, there is a section titled "GoTriple is your European hub for Social Sciences and Humanities" with three statistics: 18.355.777 Documents, 23.173 Projects, and 21.607.036 Authors.

Category	Count
Documents	18.355.777
Projects	23.173
Authors	21.607.036

<https://www.gotriple.eu/>

## What have you done to solve the problem?

### GoTriple Enhancements

Improved API and Pundit annotations to better support data-driven workflows.



### GoTriple Data Science Pipeline

Deployed the FASCA Data Science Toolbox JupyterHub, Neo4j knowledge graphs and Git integration for FAIR publication.



# FASCA

FACILITATING SCHOLARLY  
COMMUNICATION ANALYSIS  
THROUGH GOTRIPLE  
PIPELINE

### GoTriple Chatbot

Developed a RAG-based multilingual chatbot allowing researchers to interact with GoTriple documents via AI, with citation transparency and data extraction (JSON/CSV).



### Research Pilots & Support

Selected 2 research pilots from an open call and launched free data science consultations for SSH researchers.



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



## Zoom on GoTriple Data Science Toolbox

### JupyterHub workspace

Each user gets an isolated notebook environment, no local installation required. Pre-installed libraries (pandas, numpy, networkx, neo4j...) and ready-to-use templates.

### Neo4j knowledge graphs

Integrated graph database with Graph Data Science module. Each user works in a dedicated graph, full isolation, free experimentation.

### Interactive visualisation

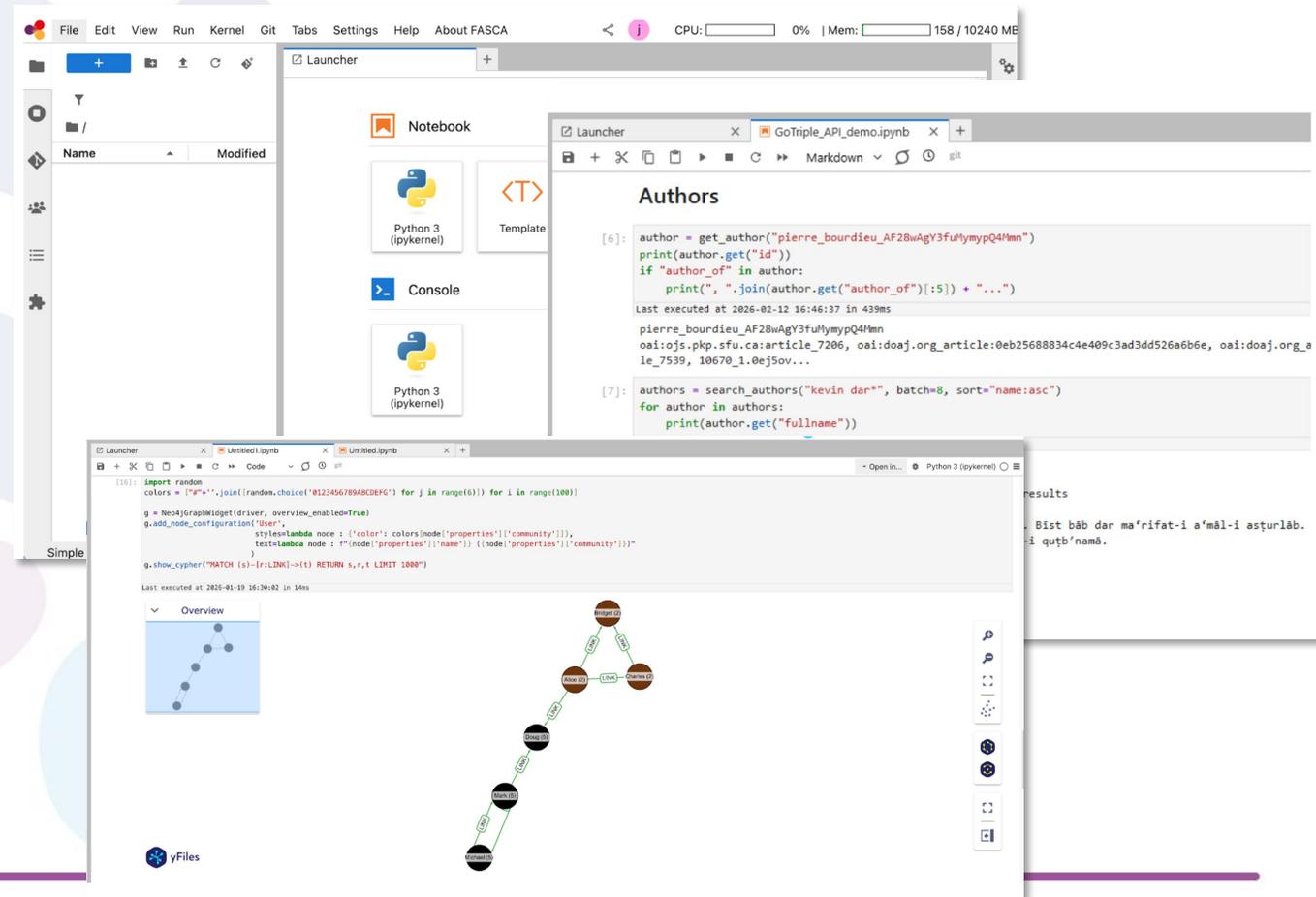
Explore nodes, relationships and patterns with yFiles graph widgets directly in notebooks.

### FAIR publication workflow

Push notebooks to GitHub/GitLab, then to Zenodo

### Free to use

For all GoTriple registered users



The screenshot displays a JupyterHub interface with a notebook titled 'GoTriple\_API\_demo.ipynb'. The notebook contains two code cells:

```
[6]: author = get_author("pierre_bourdieu_AF28wAgY3fuUymypQ4Mnn")
print(author.get("id"))
if "author_of" in author:
    print(", ".join(author.get("author_of")[:5] + "..."))

Last executed at 2026-02-12 16:46:37 in 439ms
pierre_bourdieu_AF28wAgY3fuUymypQ4Mnn
oai:ojs.pkp.sfu.ca:article_7206, oai:doaj.org:article:0eb25688834c4e409c3ad3dd526a6b6e, oai:doaj.org_article_7539, 10670_1.0eJ5ov...
```

```
[7]: authors = search_authors("kevin dar*", batch=8, sort="name:asc")
for author in authors:
    print(author.get("fullname"))
```

Below the code, a graph visualization is shown using the yFiles widget. The graph displays a network of nodes and edges, with a central node labeled 'pierre\_bourdieu' connected to other nodes like 'Axe (2)', 'Diana (2)', and 'Diana (1)'. The graph is rendered in a dark theme with red and black nodes and edges.



## Zoom on GoTriple Chatbot Allowing GoTriple users to interact via AI with documents found on the platform

### A Retrieval Augmented Generation (RAG)

**chatbot** on documents selected by the users

Documents metadata plus the full-text of the publications if available

### Increased transparency

Possibility to check the text excerpts used to formulate the answer

### Data extraction

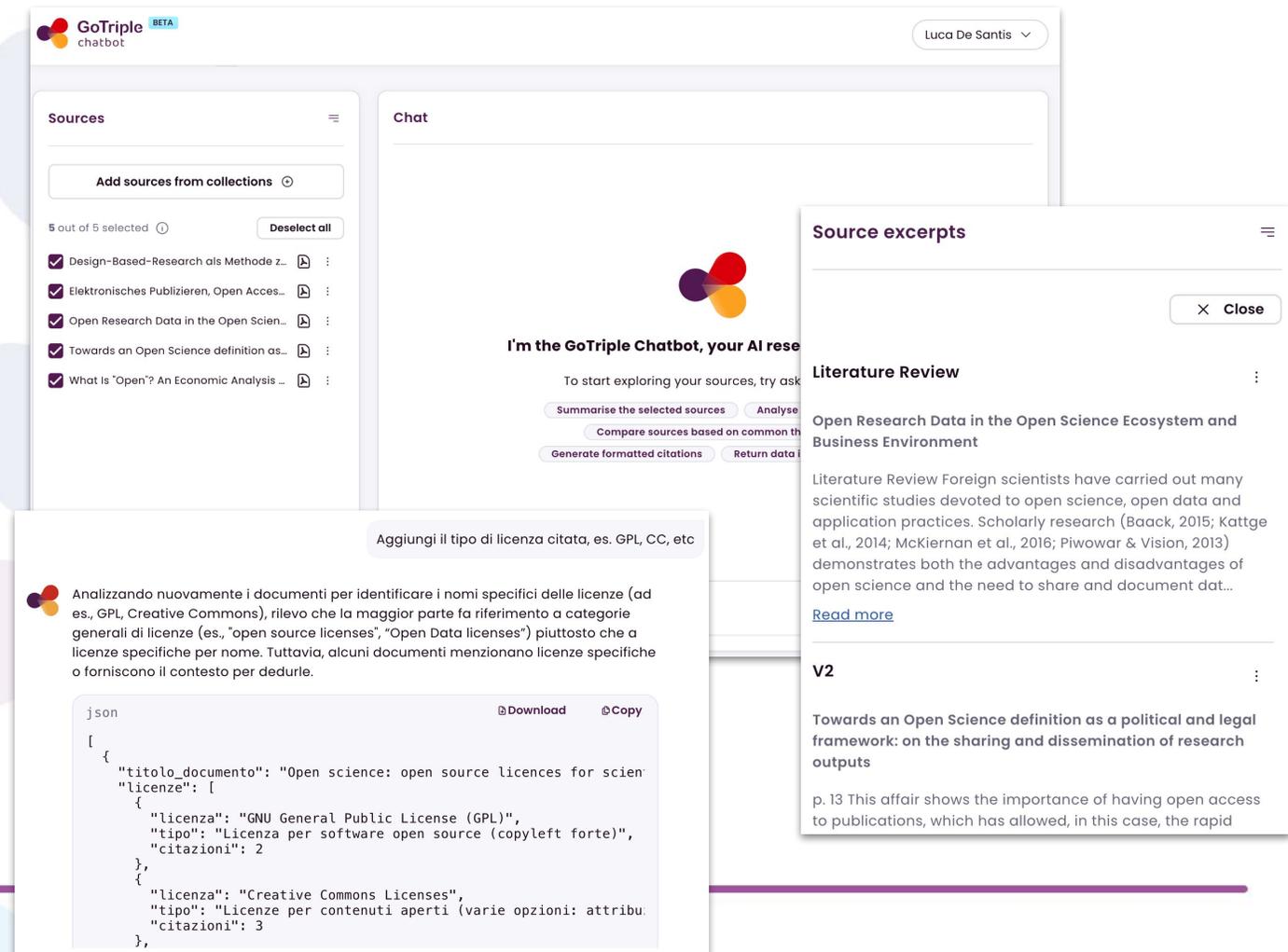
Possibility to obtain answers as data, formatted in JSON and CSV, ready to be downloaded or copied

### Multilingual

Users can chat in their native language to analyse documents written in other languages

### Free to use

For all GoTriple registered users



The screenshot displays the GoTriple chatbot interface. On the left, a 'Sources' panel lists five selected documents. The main chat area shows a message from the chatbot: 'I'm the GoTriple Chatbot, your AI research assistant. To start exploring your sources, try asking: Summarise the selected sources, Analyse the selected sources, Compare sources based on common themes, Generate formatted citations, Return data as JSON or CSV.' Below the chat, a 'Source excerpts' panel shows a snippet from a 'Literature Review' document. A separate window displays a JSON data extraction result for license information.

**Source excerpts**

**Literature Review**

Open Research Data in the Open Science Ecosystem and Business Environment

Literature Review Foreign scientists have carried out many scientific studies devoted to open science, open data and application practices. Scholarly research (Baack, 2015; Kattge et al., 2014; McKiernan et al., 2016; Piwowar & Vision, 2013) demonstrates both the advantages and disadvantages of open science and the need to share and document data...

[Read more](#)

**V2**

Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs

p. 13 This affair shows the importance of having open access to publications, which has allowed, in this case, the rapid

**JSON Data Extraction:**

```

{
  "titolo_documento": "Open science: open source licences for science",
  "licenze": [
    {
      "licenza": "GNU General Public License (GPL)",
      "tipo": "Licenza per software open source (copyleft forte)",
      "citazioni": 2
    },
    {
      "licenza": "Creative Commons Licenses",
      "tipo": "Licenze per contenuti aperti (varie opzioni: attribuzione)",
      "citazioni": 3
    }
  ]
}

```

## What are the key results achieved to date and how have you made them available to the broader community?

### 4 target groups it aims to benefit



**SSH Researchers** Data science support via the Toolbox and Chatbot, tailored training, and personalised consultations. 2 research pilots underway: multilingual corpus building (Translation Studies) and metadata exploration via Jupyter notebooks.



**Data Scientists** Opportunity to collaborate with SSH scholars on novel research using LLMs and knowledge graphs. Pathway to join the GoTriple Data Science team after FASCA.



**Business & Industry** Better access to GoTriple data via API for SMEs and startups delivering solutions in tourism, gaming, education and other SSH-applied areas.



**Research Performing Organisations** Access to new tools, data and training for their researchers — integrated within the EOSC ecosystem.

---

## How will make your results sustainable over time? How will the scientific community/-ies further exploit them?

### **Publishing & Open Access Dissemination**

GoTriple portal: dedicated FASCA services page. Zenodo & GitHub: notebooks and software available for reuse in a FAIR-compliant workflow.

### **Training & Knowledge Sharing**

Free personalised consultations for SSH researchers (30 sessions available, first consultations open call applicants).

### **Community Engagement & Outreach**

Dissemination via OPERAS/IBL PAN channels and LinkedIn. Poster accepted at the OPERAS Co LUMEN chatbot). Participation in SSH & EOSC events.

### **Sustainable Impact & Future Plans**

GoTriple Data Science Support Service designed to persist beyond FASCA, within OPERAS infrastructure and the EOSC ecosystem.



## Who is doing it?



Sy Holsinger



Tomasz Umerle



Cezary Rosinski



Luca De Santis



Julien Homo



Stefano De Paoli



Suzanne Dumouchel



Nikodem Wolczuk



Barbara Wachek



Alessandro Bertozzi



Kevin Darty



Georgia Rubidge



# OSCARS

Open Science Clusters' Action  
for Research & Society

## Funded Project

# Thanks! Questions?



Funded by  
the European Union