DOI: 10.5281/zenodo.7701286

Triple

Transforming Research through Innovative
Practices for Linked Interdisciplinary Exploration

[January 2023] Advancing Open Scholarship

D5.7 – Additional services updated

Version 1.0 – Final
PUBLIC

H2020-INFRAEOSC-2019
Grant Agreement 863420

The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 863420

Disclaimer- "The content of this publication is the sole responsibility of the TRIPLE consortium and can in no way be taken to reflect the views of the European Commission. The European Commission is not responsible for any use that may be made of the information it contains."

Deliverable Name

Project Acronym:	**TRIPLE**
Project Name:	**Transforming Research through Innovative Practices for Linked Interdisciplinary Exploration**
Grant Agreement No:	**863420**
Start Date:	**1/10/2019**
End Date:	**31/03/2023**
Contributing WP	**WP5**
WP Leader:	**Net7 Srl**
Deliverable identifier	**D5.7**
Contractual Delivery Date: 01/2023	**Actual Delivery Date: 01/2023**
Nature: Report	**Version: 1.0 Final**
Dissemination level	**PU**

Revision History

Version	Created/Modifier	Comments
0.0	Luca De Santis (Net7)	First version of the TOC
0.1	Maxi Schramm (OKMAPS), Christopher Kittel (OKMAPS), Jan Konstant (OKMAPS), Peter Kraker (OKMAPS), Simone Kopenik (Know-Center), Dieter Theiler (Know-Center), Gael Van Weyenbergh (MEOH), Andrew Pomazanskyi (Nuromedia), Luca De Santis (Net7)	Chapter 2 & 6 (Net7), Chapter 4 (MEOH, Nuromedia), Chapter 3 (Know-Center), Chapters 5 & 7 (OKMAPS)
0.2	Luca De Santis (Net7)	Publishable Summary, introduction and conclusions added; general revision
0.3	Sy Holsinger (OPERAS), Julien Homo (Foxcub)	External review
1.0	Luca De Santis (Net7)	Final version

Table of Contents

Table of Figures

Acronyms

AAK	Annotating All Knowledge
API	Application Programming Interface
AWS	Amazon Web Services
BASE	Bielefeld Academic Search Engine
CBF	Content-based Filtering
CF	Collaborative Filtering
CI	Corporate Identity
CSL	Citation Style Language
DIH	EOSC Digital Innovation Hub
DOI	Digital Object Identifier
EOSC	European Open Science Cloud
EOSC-DIH	EOSC Digital Innovation Hub
GUI	Graphical User Interface
HMS	Harvesting Management Platform
MP	Most Popular
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
PSVS	Parameterized Search and Visualisation Service
REST	REpresentational State Transfer
RS	Recommender System
ScaR	Scalable Recommendation-as-a-service
SSH	Social Sciences and Humanities
TBS	Trust Building System
TLS	Transport Layer Security
TF-IDF	Term frequency–inverse document frequency
UI	User Interface
UX	User eXperience
VIPER	The Visual Project Explore

Publishable Summary

In TRIPLE, WP5 is devoted to the integration of Innovative Services in the GoTriple platform. The work has been organised in the following 6 tasks:

- T5.1: Third-party applications integration
- T5.2: Recommender system
- T5.3: Trust building system
- T5.4: Visualisation
- T5.5: Open annotation tool
- T5.6. Discovery system

The structure of this document, which describes the work done in WP5 after September 2021, reflects the organisation of the WP, with a dedicated chapter for each of its tasks.

Chapter 2 presents T5.1, which is devoted to the identification of innovative services to integrate in GoTriple, beyond those developed by the project partners in the other tasks of WP5. After the successful integration of the Crowdfunding platform, in the current period the work has been concentrated on integrating two other services identified in the previous phases of the project: an extendible mechanism to export documents' metadata for easily importing them in bookmarking reference management apps and services (e.g. Zotero, Mendeley or EndNote) and the OPERAS Metrics service.

In Chapter 3, the work in T5.2 on the Recommender system is described, with new functionalities and services that this platform exposes to GoTriple. In particular, new algorithms for suggesting documents, projects, and authors have been developed, together with the personalised recommendations service, which allows registered users to automatically receive in their personal page dedicated documents suggestions, according to their previous history of interactions with GoTriple.

Chapter 4 presents how the Trust Building System (TBS) has been expanded within T5.3, with the implementations of the Public groups, and the definition of new strategies to ease the onboarding of new users in the TBS platform.

The significant improvements in T5.4 for the interactive Visualization system are described in Chapter 5, spanning from the enhancements on the User Interface side to the development of new features, together with refinements to the existing ones.

Chapter 6 discusses the progress of the Open Annotation Tool Pundit, which is part of T5.5. The chapter highlights the achievements and ongoing efforts, including Pundit's successful integration into the EOSC Marketplace, which can be attributed to its participation in an EOSC Digital Innovation Hub (DIH) pilot. Additionally, the team is currently working on incorporating new features into the web annotation tool during the remaining months of the project.

Finally, Chapter 7 describes the work done in T5.6 on the Discovery System, which is the back-end component that powers the advanced visualisations of T5.4. Several goals have been accomplished here, both on the infrastructure side (the finalisation of the transition to the Custom Services pipeline, improvements to stability and scalability) and on the functional side, in particular with the development of the author and project visualisations, the former already integrated in GoTriple, while the latter will be included soon.

1. Innovative Services in GoTriple: a final update

We describe in this document the work done in WP5 after September 2021, where the state of the art of the six tasks of this work-package has already been described in dedicated deliverables.[1]

The team has made noteworthy enhancements to the relevant services, as demonstrated in this document. These improvements include the successful integration of third-party applications, which were not developed by TRIPLE consortium partners. For each of the WP5 tasks, a detailed description of the achievements, and in some cases of the remaining work to do in these final two months of the project, is provided.

2. Third-party applications integration

T5.1 is the task of the TRIPLE project devoted to the identification of innovative services to integrate in GoTriple, beyond those developed by the project partners in the other tasks of WP5. The deliverable D5.1 [1] fully described the selection process of these services, where in the end, was decided to integrate:

- a crowdfunding platform, operated by the Swiss company WeMakeIt
- bookmarking services, that GoTriple users can use to keep track of the publications found on the platform
- the OPERAS Metrics service, which, as the name implies, presents indicators regarding the impact of a publication indexed in GoTriple.

Their integration is described in the chapters that follow.

2.1. The crowdfunding platform

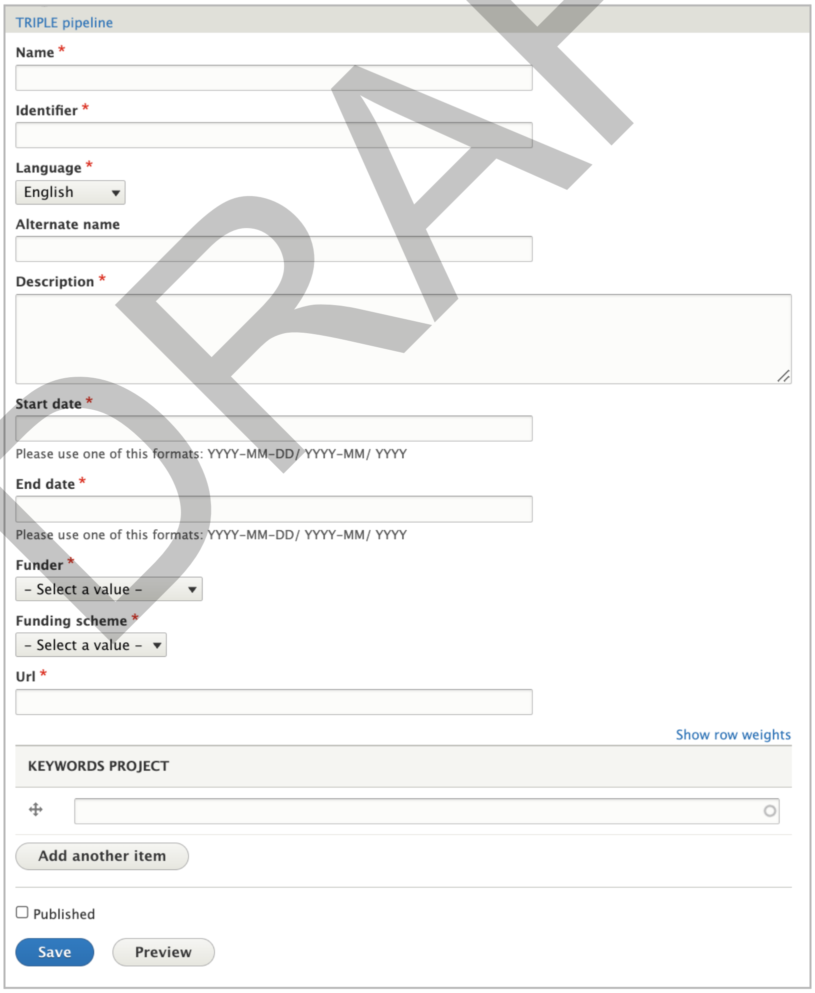
For the integration of WeMakeit, a light technical approach was chosen: it is a crowdfunding platform completely independent of GoTriple and only a link to it appears in the top menu of our discovery platform.

This choice was taken both for pragmatic and technical reasons. In the first case, we wanted to verify the appeal and the success of this kind of service amongst the SSH community. In this respect, the results so far have been limited. Also, the use of WeMakeIt comes at a cost, so OPERAS, together with the GoTriple Committee board should verify at the end of TRIPLE whether to renew the agreement with the Swiss company or find a possible alternative.

[1] In particular D5.1 [1], D5.2 [2], D5.3 [3], D5.4 [4], D5.5 [5], and D5.6 [6].

As far as the technical aspects are concerned, WeMakeIt doesn't provide ready-to-use APIs that would have allowed the Triple team to develop a connector to automatically import OPERAS' crowdfunding projects in GoTriple. The only option would have been to directly access their database, a choice that was not advisable, also for security reasons.

The TRIPLE development team has therefore decided to follow a completely different approach, given on one hand the small number of crowdfunding projects to import, and on the other considering a possible generalization of the need to insert projects in GoTriple. For example, this might happen when the effort to develop a dedicated connector is technically too high or not convenient because the amount of the available data is small. What has been developed is an extension of the Harvesting Management Platform (HMS), that allows a GoTriple administrator to manually insert the description of a project. The HMS, as described in deliverable D2.6 (*Report on Global Data Retrieval*) [7], has been developed by using the Drupal 9 framework. Administrators can access the HMS Drupal back-end and insert the project description as shown below.



TRIPLE pipeline

Name *

Identifier *

Language *
English ▾

Alternate name

Description *

Start date *

Please use one of this formats: YYYY-MM-DD/ YYYY-MM/ YYYY

End date *

Please use one of this formats: YYYY-MM-DD/ YYYY-MM/ YYYY

Funder *
- Select a value - ▾

Funding scheme *
- Select a value - ▾

Url *

Show row weights

KEYWORDS PROJECT

Add another item

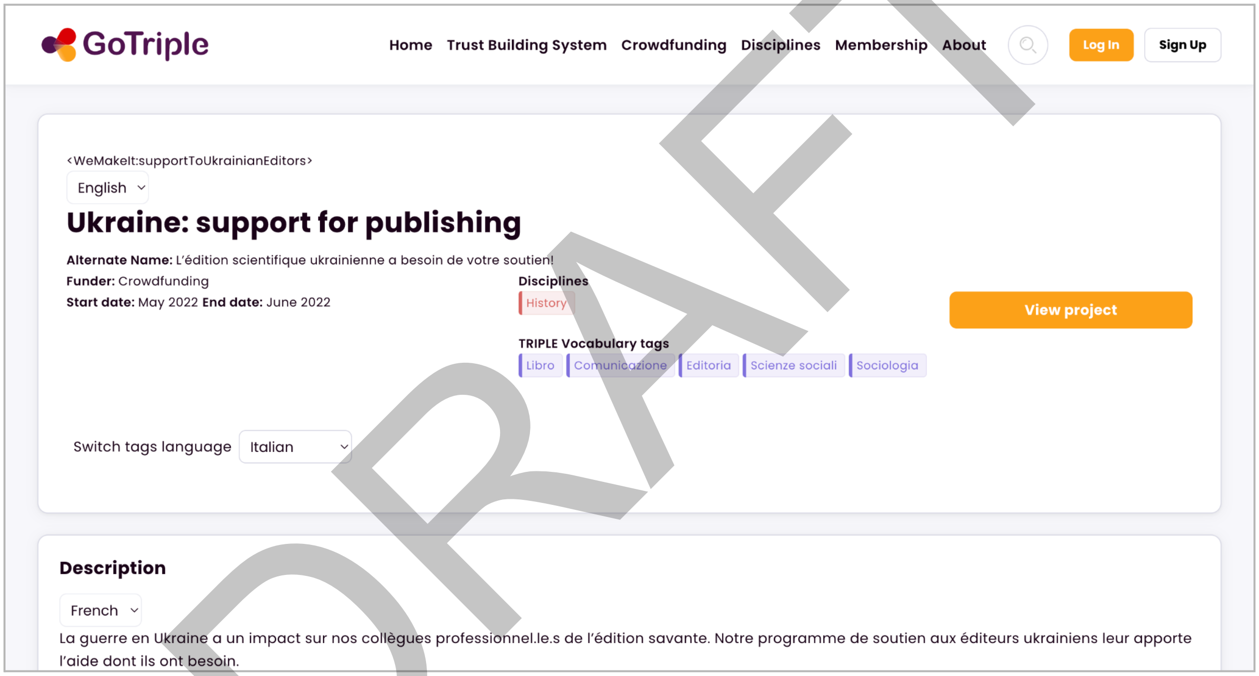
□ Published

Save
Preview

Figure 1 - Project entry form in the HMS

In particular, they must insert the language in which the textual descriptions are written and the specific "funding program" of the project. The latter is based on a Drupal taxonomy that can be expanded by the administrator to indicate all possible project sources to manage with this manual interface.

Once the project is "published", its data are automatically sent to the SCRE pipeline, which processes them as described in chapter 3 deliverable D2.5 (*Report on Data Enrichment* [8]). In particular, the project metadata passes the enrichment phase: if the textual descriptions are not in English, an automatic translation is provided; also the automatic classification and annotation with the TRIPLE vocabulary keywords are applied. The image that follows shows a WeMakeIt crowdfunding project imported in GoTriple through the procedure described above.



<WeMakeIt:supportToUkrainianEditors>

English

Ukraine: support for publishing

Alternate Name: L'édition scientifique ukrainienne a besoin de votre soutien!
Funder: Crowdfunding
Start date: May 2022 **End date:** June 2022

Disciplines
History

TRIPLE Vocabulary tags
Libro | Comunicazione | Editoria | Scienze sociali | Sociologia

View project

Switch tags language: Italian

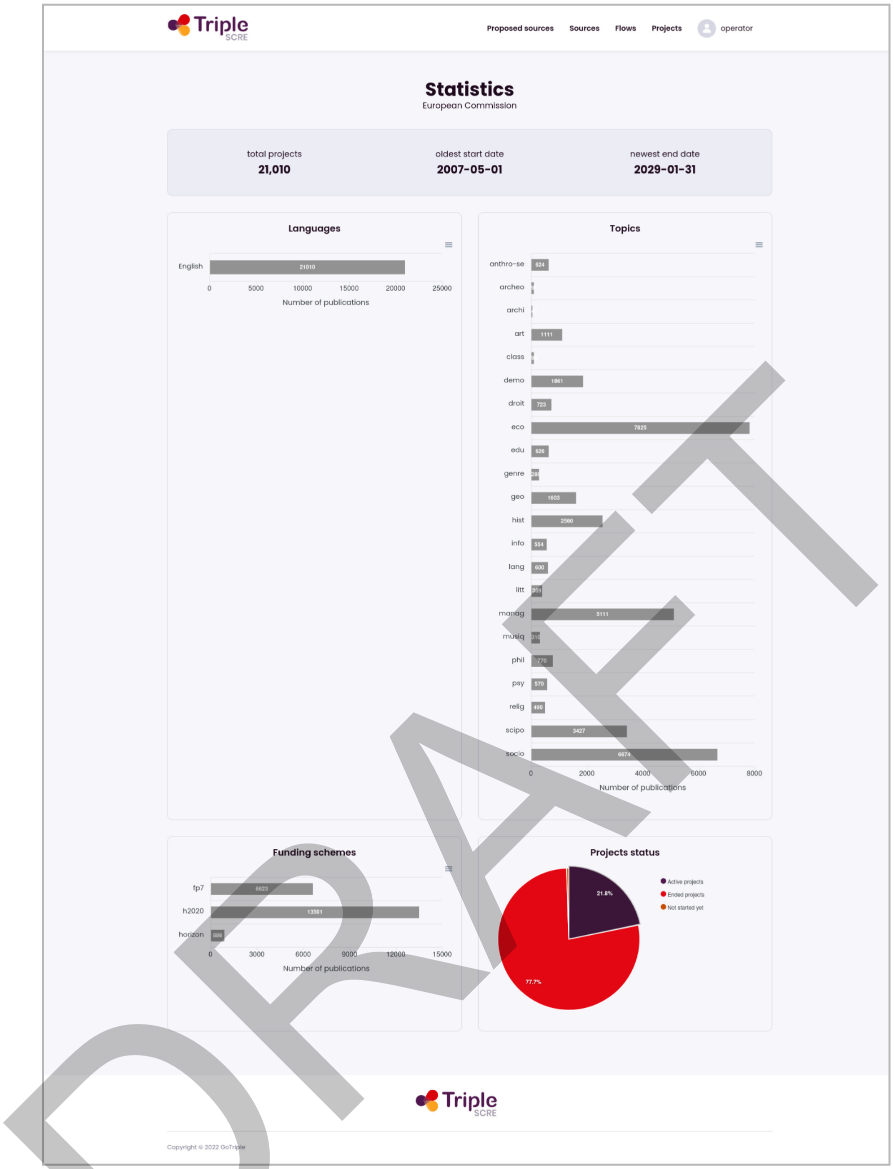
Description

French

La guerre en Ukraine a un impact sur nos collègues professionnel.le.s de l'édition savante. Notre programme de soutien aux éditeurs ukrainiens leur apporte l'aide dont ils ont besoin.

Figure 2 - A crowdfunding project in GoTriple

Finally, albeit not directly related to crowdfunding, it is worth mentioning another HMS development done in the period related to projects, that is the statistics. One example, for CORDIS projects, is shown in the image below.



Statistics

European Commission

total projects	oldest start date	newest end date
21,010	2007-05-01	2029-01-31

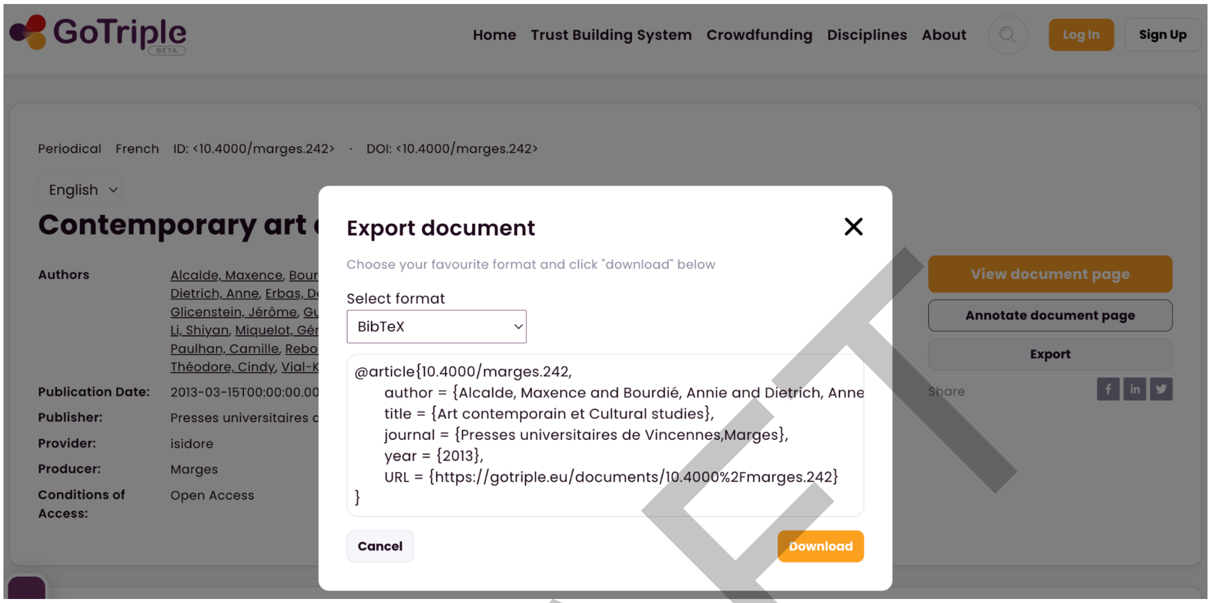
Copyright © 2022 GoTriple

Figure 3 - The project statistics in the HMS

2.2. Bookmarking services

As indicated in D5.1 [1], the possibility to export the description of a document found in GoTriple to be imported in a bibliographic management app or service, was on top of the wishlist of the GoTriple prospect users.

Instead of choosing the integration with a specific tool, a more general approach, based on standards, was followed, which is providing an export feature in GoTriple document's landing pages in BibTeX format. The latter is a widely used standard, supported by practically any bibliographic tool. The use of this feature is shown in the image below.



Periodical French ID: <10.4000/marges.242> · DOI: <10.4000/marges.242>

Contemporary art

Authors: Alcalde, Maxence, Bour… Dietrich, Anne, Erbas, D… Glicenstein, Jérôme, Gu… Li, Shiyan, Miquelot, Gér… Paulhan, Camille, Rebo… Théodore, Cindy, Vial-K…

Publication Date: 2013-03-15T00:00:00.00

Publisher: Presses universitaires c

Provider: isidore

Producer: Marges

Conditions of Access: Open Access

Choose your favourite format and click "download" below

Select format: BibTeX

```
@article{10.4000/marges.242,
      author = {Alcalde, Maxence and Bourdié, Annie and Dietrich, Anne
      title = {Art contemporain et Cultural studies},
      journal = {Presses universitaires de Vincennes,Marges},
      year = {2013},
      URL = {https://gotriple.eu/documents/10.4000%2Fmarges.242}
}
```

Figure 4 - Exporting document's metadata in BibTeX format

From a document detail page, a new window appears when clicking on the "Export" button. Here, the user can select the data format for the export of the document's metadata, in particular BibTeX. The export is directly shown in the window or can be downloaded on the user's desktop computer. This way, the user can import the document's details in a tool such as Zotero, Mendeley or EndNote.

It is important to point out that this strategy was also chosen for its extensibility. Publications' metadata are in fact offered in multiple formats, including JSON and JSON-LD, simplifying the reuse of GoTriple's data through software programs.

2.3. Operas Metrics Service

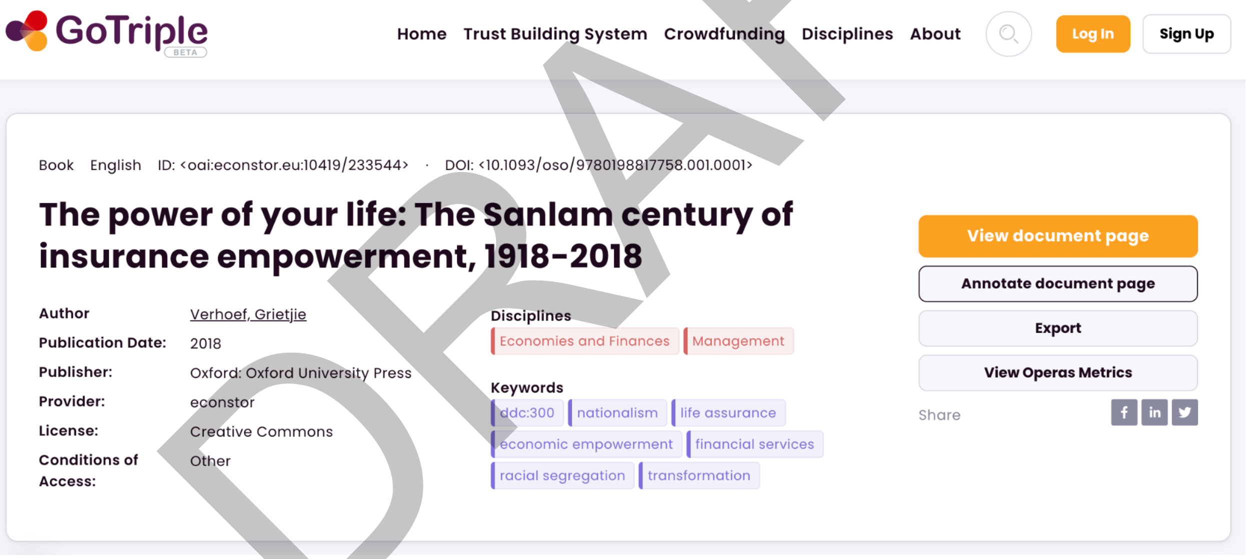
One of the other top choices of prospect users for useful services to integrate in GoTriple, as documented in D5.1 [1], was a metrics measurement for publications.

It was therefore natural to choose the OPERAS Metrics Service for this purpose: on one hand this goes in the direction of encouraging synergies amongst OPERAS services, while on the other, the integration with Metrics is technically straightforward. The service provides a convenient JavaScript-based interactive widget that can be easily embedded in an external website such as GoTriple. The identification of the publication is done through its DOI.

At the time of writing, the amount of publications for which Metrics provides statistics is somewhat limited (mainly open access monographs), albeit their number is destined to grow significantly as the service is expanded (for example, this is one of the goals of the OPERAS-PLUS research project[2], currently ongoing).

In order to make sure that the metrics are visualisable for a publication, we first verify if its DOI is present on Metrics by dynamically calling an API via JavaScript in the publication’s landing page.

Only if the answer is positive the "View Operas Metrics" button is displayed. By clicking on it, the widget appears (see the two images below).



Book English ID: <oai:econstor.eu:10419/233544> · DOI: <10.1093/oso/9780198817758.001.0001>

The power of your life: The Sanlam century of insurance empowerment, 1918–2018

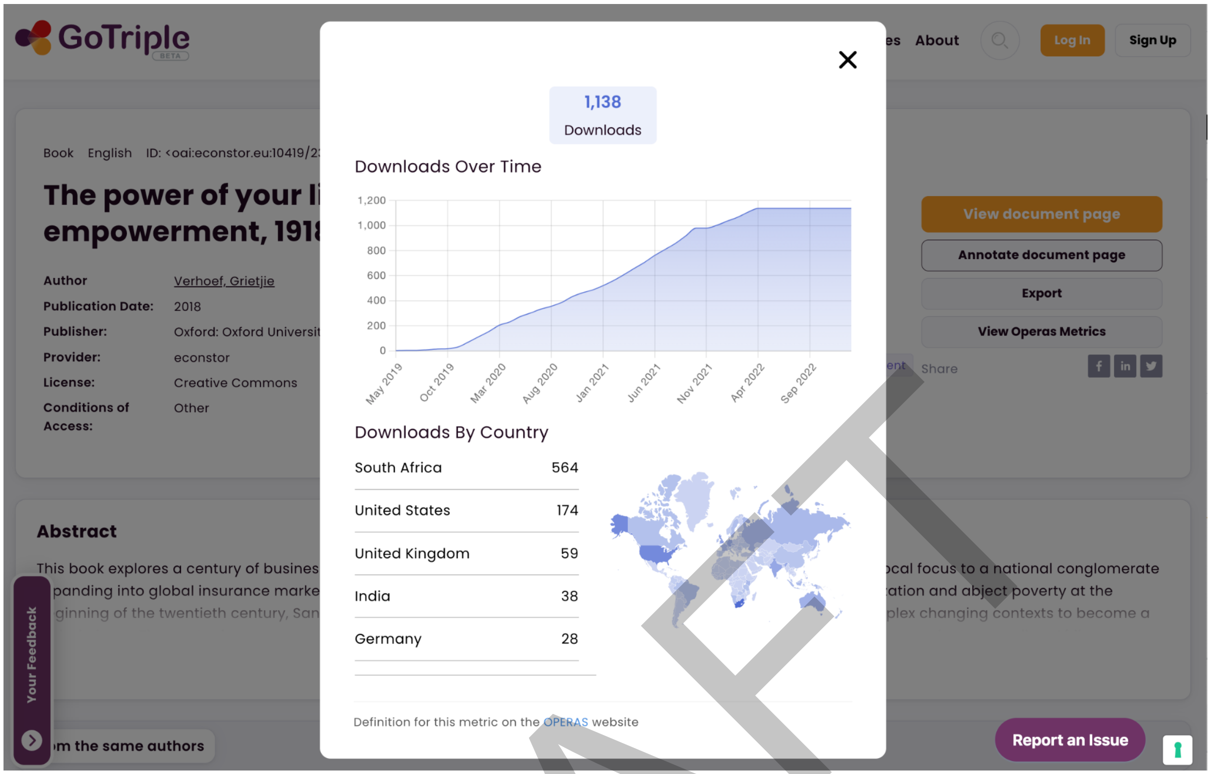
Author: Verhoef, Grietjie
Publication Date: 2018
Publisher: Oxford: Oxford University Press
Provider: econstor
License: Creative Commons
Conditions of Access: Other

Disciplines
Economies and Finances | Management

Keywords
ddc:300 | nationalism | life assurance | economic empowerment | financial services | racial segregation | transformation

Figure 5 - A document with metrics available

[2] The OPERAS-PLUS project (https://operas-eu.org/projects/operas-plus/) runs from September 2022 to July 2025. OPERAS was selected in 2021 as a new infrastructure on the ESFRI Roadmap and it is now on its path to becoming operational as an ERIC in 2028. The project aims at supporting this process, together with the development of the needed infrastructure.
Also it supports the development of OPERAS' services portfolio, providing both required technology and a monitoring system for services development.



Book English ID: <oai:econstor.eu:10419/2

**The power of your li
empowerment, 191**

Author	Verhoef, Grietjie
Publication Date:	2018
Publisher:	Oxford: Oxford Universit
Provider:	econstor
License:	Creative Commons
Conditions of Access:	Other

Abstract

This book explores a century of busines … ocal focus to a national conglomerate
panding into global insurance marke … ation and abject poverty at the
ginning of the twentieth century, San … plex changing contexts to become a

Downloads Over Time

Downloads By Country

South Africa	564
United States	174
United Kingdom	59
India	38
Germany	28

Definition for this metric on the OPERAS website

m the same authors

Figure 6 - OPERAS Metrics widget in GoTriple

3. Recommender system

Recommender Systems (RSs) are software components that predict and suggest items of interest to users or user groups. Task 5.2 is adding to the platform's user friendliness, focusing on the exploration and integration of tailored information through ranking algorithms implemented in the form of a Recommender System.

To date, recommendation algorithms mostly use data about users, items, and user-item interactions, where users' historical data traces form the main information source of the recommendation logic [9]. GoTriple's Recommender System (RS) is a service component that supports researchers in finding their items of interest and, with this, enhances their discovery experience. Based on the Scalable Recommendation-as-a-service (ScaR) framework, we offer GoTriple a tailored bundle of recommendation services, extending on collaborative and content-based recommendation approaches. To this end, we collect, model, and exploit user and content data to suggest interesting research documents, research projects, and research peers.

The RS is based on static data available in the system (i.e., meta-data of documents, research projects, and research peers) and in addition, is continuously informed with user interaction data. GoTriple user data is collected in the platform and sent to the RS on occurrence in an event-based manner. Triggered by the navigation of the user in the platform, GoTriple calls respective recommendation service endpoints to retrieve a list of suggested items that are then displayed in the GoTriple Graphical User Interface (GUI). As the GoTriple RS is based on ScaR, communication is implemented with REST-based web services.

The Recommender services are implemented in the Java programming language and use either the open-source framework Spring or Dropwizard. For deployment, they are packaged as Docker images, which enables operating system level virtualization.

To support the development cycle executed in TRIPLE, most recent functionalities and bug-fixes can be accessed via the development server[3]. To make sure only stable versions of the service are used for the GoTriple production environment, the production-ready service can be accessed via the production URL[4]. Both instances are secured with the help of HTTP Basic Authentication requiring username and password to login on top of HTTP connections that are encrypted with TLS.

3.1. Implemented Recommendation Algorithms

Currently, three types of algorithms are deployed for the GoTriple platform: (i) Collaborative Filtering (CF) [10], which is a personalized algorithm that analyses items' interaction data to find similar users. Subsequently, items of these similar users (i.e., nearest neighbours) are extracted and recommended. (ii) Content-based Filtering (CBF) [11], which calculates item similarities based on content features and recommends a ranked list of these similar items. (iii) Most Popular (MP) [12], which is an un-personalized algorithm recommending items that are weighted and ranked by their popularity.

- To enable CF-based recommendations, which are based on users' interaction traces extracted from the GoTriple platform, different interactions are considered, e.g., accessing and opening research items, projects, and profiles. To find research items, projects and profiles of interest, entities from within the traces of other users which share the same interests in research items and projects are considered.
- To be able to offer users with content-based documents recommendations, we calculate similarities to a user's currently viewed research item or project. These CBF recommendations are provided based on the similarity (i.e., using TF-IDF) of research entities' headlines, abstracts, and sets of categorizing keywords attached to them. Our implementation of the content-based recommender utilizes similarity queries directly supported by the database backend and allows to tune parameters for respective built-in similarity measures, if needed.

[3] https://triple.know-center.at/sp/swagger-ui.html

[4] https://prod-triple.know-center.at/sp/swagger-ui.html

- To enable MP-based recommendations, which are based on the users' interaction traces, to be recommended research items, and projects are weighted and ranked by their popularity (e.g., the item which was accessed the most).

For the instantiation of ScaR in the context of TRIPLE, the RS needs to customise the algorithms to the needs of the domain and the application context. The instantiation of an algorithmic configuration is done via configuration files following the YAML syntax. The configuration for each algorithm, so called “recommender profile”, can look different for each domain and can be replicated for various application scenarios. To instantiate the algorithms for the recommendation of research documents, projects, and peers (i.e., users of the GoTriple platform), we created multiple instances of recommender profiles.

The following profiles have been created for the first GoTriple release and were periodically updated to the now final version:

- triple_document_cf.yaml: Profile for recommending personalized research items.
- triple_document_item_cb.yaml: Profile for recommending similar research items.

New recommendation profiles that have been added with the availability of data for projects and research peers:

- triple_project_cf.yaml: Profile for recommending personalized research projects.
- triple_project_item_cb.yaml: Profile for recommending similar research projects.
- triple_user_cf.yaml: Profile for recommending individually matched research peers.

Additionally, the following profiles are provided to meet new requirements that arose during the project:

- triple_document_mp.yaml: Profile for recommending most popular research items.
- triple_document_trending_mp.yaml: Profile for personalized recommending of most popular research items from within a recent timeframe.
- triple_project_mp.yaml: Profile for recommending most popular research projects.
- triple_project_related_item_cb.yaml: Profile for recommending similar research projects given a research document.
- triple_project_trending_mp.yaml: Profile for personalized recommending of most popular research projects from within a recent timeframe.

3.2. Recommender Extension for Notification System

The notification system established in GoTriple is now also used by the RS to suggest trending research documents to users who sign up to have notifications delivered to their profile on a regular basis. For that, the RS allows registering users of the platform who want to receive notifications based on recommendations. Currently, this configuration allows defining the user receiving the notification and the interval in which notifications should be generated. The interval in which notifications are received can be “daily”, “every three days”, “weekly”, and “bi-weekly”. The registration for the user to receive notifications can also be removed so that the user is not included in future notifications anymore. If needed, it is also possible to clean all previously generated recommendations used for a particular user. The sequence diagram in Figure 7 shows the components involved in the process of adding a notification configuration to the RS, and in turn, receiving recommendations on a regular basis, which can be subsequently aggregated to be shown as notifications in the users’ profiles.

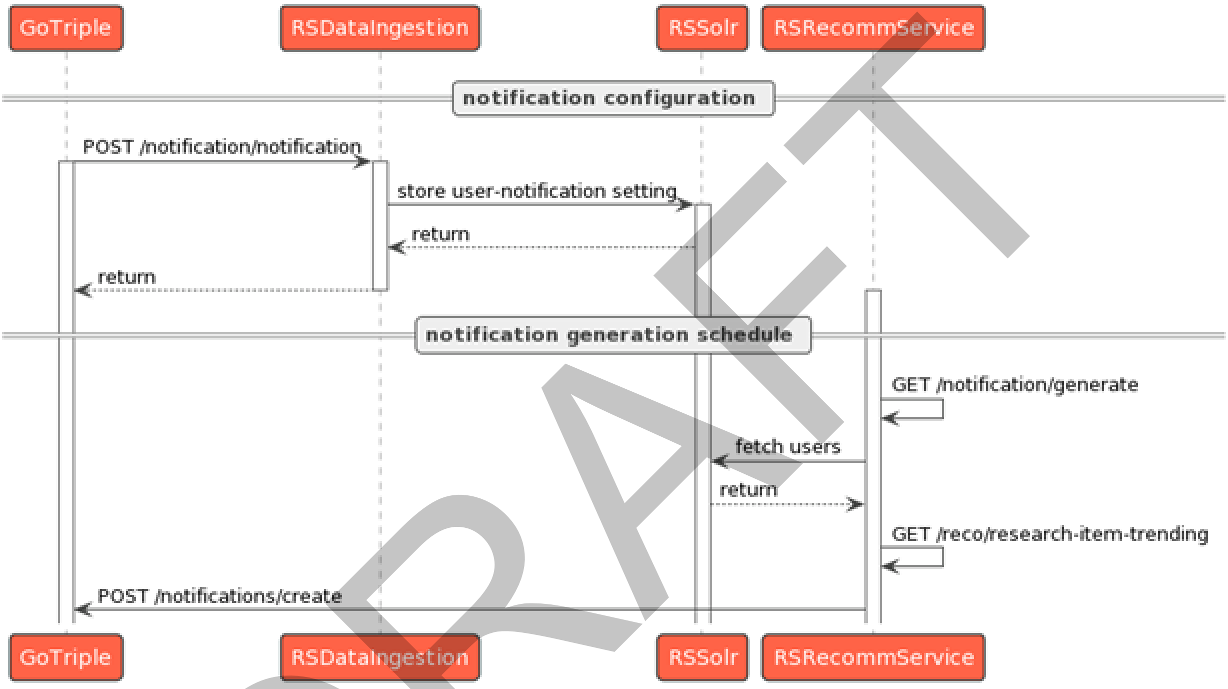

GoTriple
RSDataIngestion
RSSolr
RSRecommService
notification configuration
POST /notification/notification
store user-notification setting
return
return
notification generation schedule
GET /notification/generate
fetch users
return
GET /reco/research-item-trending
POST /notifications/create
GoTriple
RSDataIngestion
RSSolr
RSRecommService

Figure 7 - Notifications configuration and generation logic

The recommendation profile used for generating recommendations for users on an individual basis is “triple_document_trending_mp.yaml” - it considers the recency of research documents “being popular” in the system to in turn recommend the most popular ones. Additionally, the notification approach implemented in the RS considers whether a particular user was already presented with certain research documents in the last notification intervals.

3.3. Recommender REST API – Final Version

The REST API can be divided into Data Ingestion Services and Recommendation Services. The Data Ingestion Service includes a resource for storing user interaction with the RS, a resource to trigger the data synchronisation process, and now a newly introduced resource, “notification-controller”, to handle needed operations for notifications.

The “data-ingestion-controller" was updated so that it now also allows the storage of profiles and projects besides documents. The import process is similar to the one of document import but uses dedicated indexes offered by the GoTriple database backend.

The usage of “interaction-ingestion-controller" from within GoTriple was extended to reflect interactions based on profiles and projects.

To support notifications, the new resource “notification-controller” was added to allow the configuration of recommendations used for notifications. Here, the interval to send recommendations as notifications can be set for each user individually, users can be removed from receiving recommendations, and previously generated recommendations for users can be cleared.

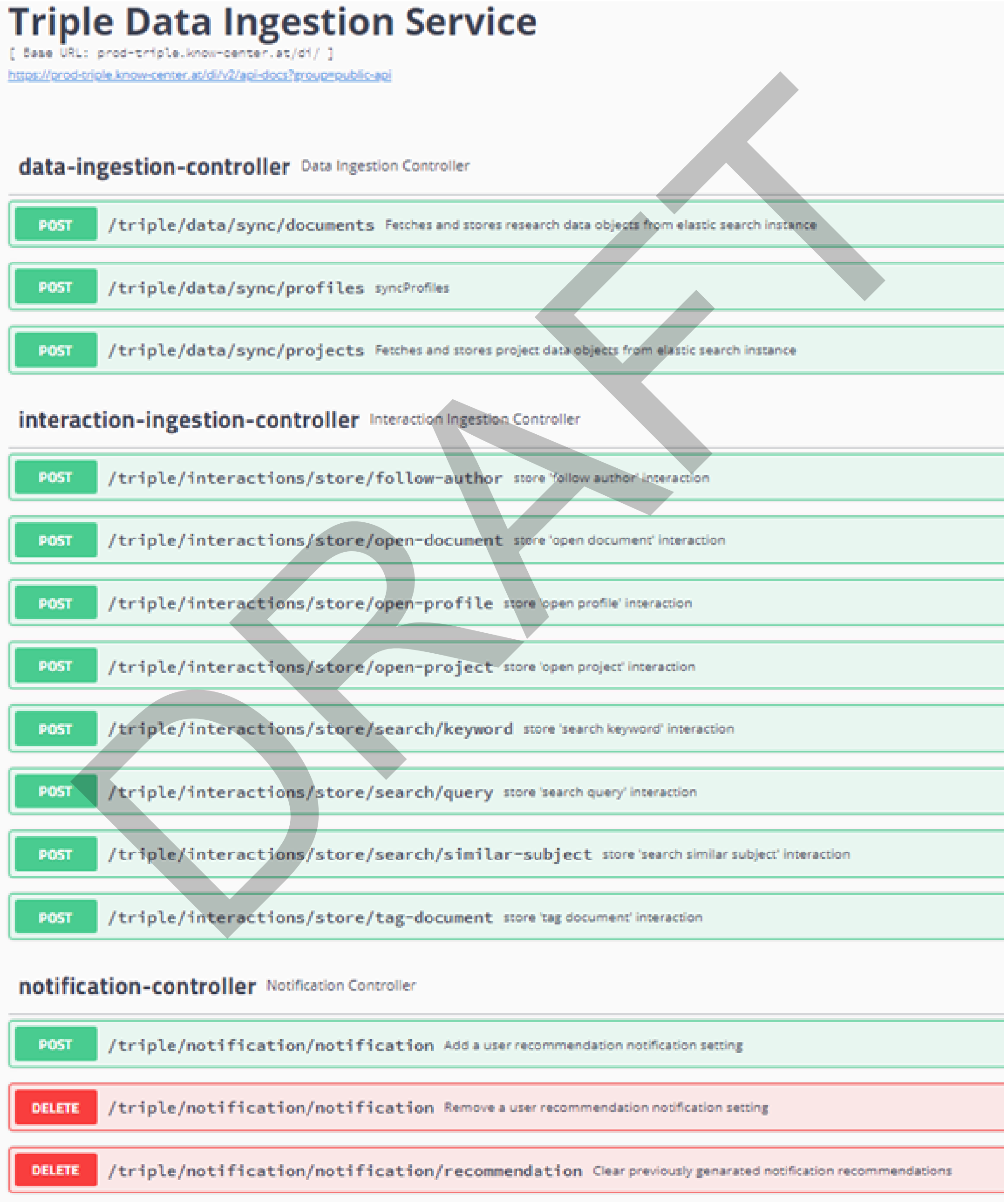
The Recommendation Services Resource provides services for requesting recommendations from the system and now also provides a resource to generate recommendations to be used for notifications.

The “notification-controller” allows triggering the generation of recommendations that are sent to the respective endpoint in GoTriple that stores notifications for all users. Here, the number of recommendations to be generated and the interval that defines the corpus of users configured to receive recommendations in the chosen interval, can be set. This endpoint normally is only instantiated from within the new scheduler of the Recommendation Service, which takes care of sending notifications in configured intervals as set for each individual user.

The “recomm-controller” offers various endpoints to be requested from within GoTriple to generate proper recommendations for logged in as well as anonymous users at different locations in the frontend:

- “research-item-most-popular" recommends documents to a user using the most popular algorithm
- “research-item-personalized" recommends documents to a user using a collaborative filtering algorithm
- “research-item-similar" recommends documents to a user that are similar to the input document using a content-based filtering algorithm
- “research-item-trending" recommends documents to a user using a collaborative filtering algorithm that considers the recency of users’ interactions with documents
- “research-peer” recommends users to a user using a collaborative filtering algorithm
- “research-project-most-popular" recommends projects to a user using the most popular algorithm
- “research-project-trending" recommends projects to a user using a collaborative filtering algorithm that considers the recency of users’ interactions with projects

- “research-projects-personalized" recommends projects to a user using a collaborative filtering algorithm
- “research-projects-related" recommends projects to a user that are similar to the input document using a content-based filtering algorithm
- “research-projects-similar" recommends projects to a user that are similar to the input project using a content-based filtering algorithm



Triple Data Ingestion Service

[Base URL: prod-triple.know-center.at/di/]

https://prod-triple.know-center.at/di/v2/api-docs?group=public-api

data-ingestion-controller Data Ingestion Controller

POST /triple/data/sync/documents Fetches and stores research data objects from elastic search instance

POST /triple/data/sync/profiles syncProfiles

POST /triple/data/sync/projects Fetches and stores project data objects from elastic search instance

interaction-ingestion-controller Interaction Ingestion Controller

POST /triple/interactions/store/follow-author store 'follow author' interaction

POST /triple/interactions/store/open-document store 'open document' interaction

POST /triple/interactions/store/open-profile store 'open profile' interaction

POST /triple/interactions/store/open-project store 'open project' interaction

POST /triple/interactions/store/search/keyword store 'search keyword' interaction

POST /triple/interactions/store/search/query store 'search query' interaction

POST /triple/interactions/store/search/similar-subject store 'search similar subject' interaction

POST /triple/interactions/store/tag-document store 'tag document' interaction

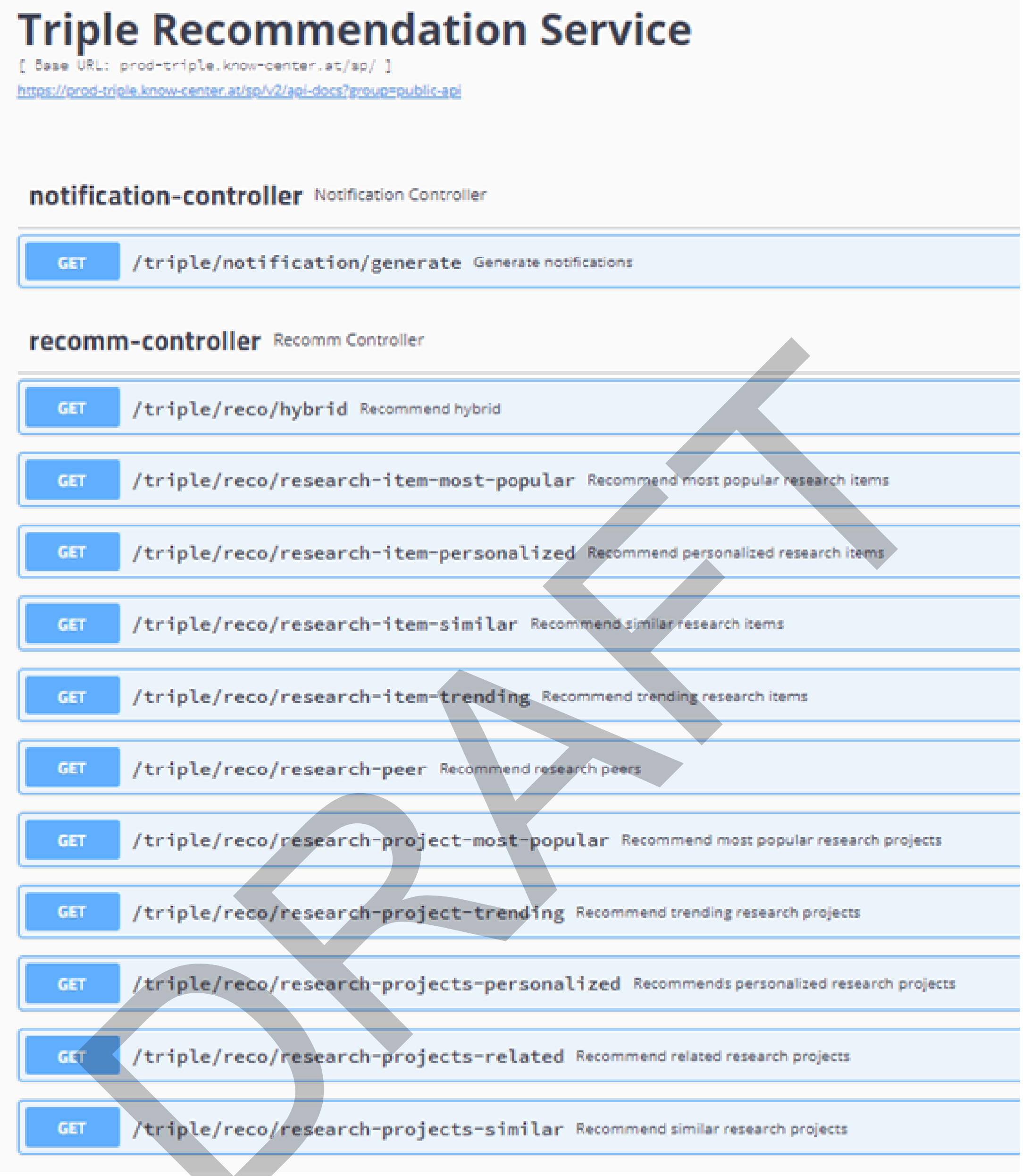
notification-controller Notification Controller

POST /triple/notification/notification Add a user recommendation notification setting

DELETE /triple/notification/notification Remove a user recommendation notification setting

DELETE /triple/notification/notification/recommendation Clear previously genarated notification recommendations

Figure 8 - Recommender Data Ingestion APIs



Triple Recommendation Service

[Base URL: prod-triple.know-center.at/sp/]

https://prod-triple.know-center.at/sp/v2/api-docs?group=public-api

notification-controller Notification Controller

GET /triple/notification/generate Generate notifications

recomm-controller Recomm Controller

GET /triple/reco/hybrid Recommend hybrid

GET /triple/reco/research-item-most-popular Recommend most popular research items

GET /triple/reco/research-item-personalized Recommend personalized research items

GET /triple/reco/research-item-similar Recommend similar research items

GET /triple/reco/research-item-trending Recommend trending research items

GET /triple/reco/research-peer Recommend research peers

GET /triple/reco/research-project-most-popular Recommend most popular research projects

GET /triple/reco/research-project-trending Recommend trending research projects

GET /triple/reco/research-projects-personalized Recommends personalized research projects

GET /triple/reco/research-projects-related Recommend related research projects

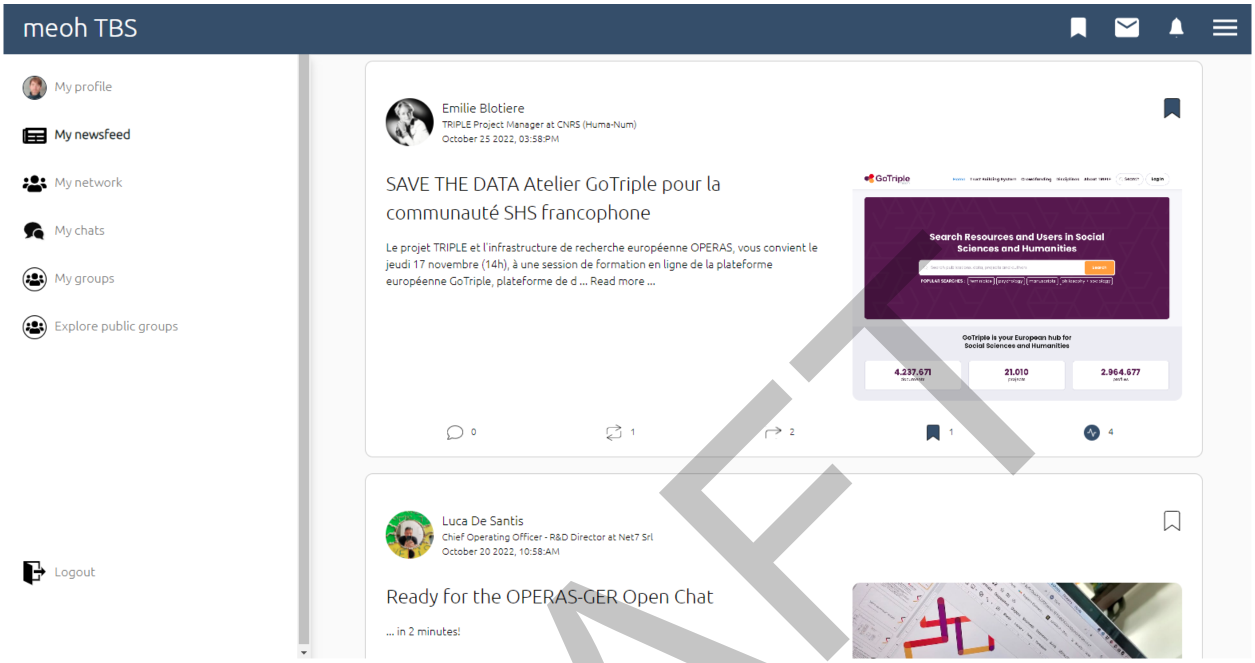
GET /triple/reco/research-projects-similar Recommend similar research projects

Figure 9 - Recommendation Service APIs

4. Trust building system

Focusing on enabling effective networking and strong partnerships among the SSH community, the Trust Building System (TBS) serves as a third party application that helps to establish and maintain adequate level of trust between various stakeholders within the community.

Following the iterative prototyping of the application and the guidelines established during the workshops conducted within WP3 at the beginning of the project, the system is advancing towards a more robust product, with the purpose to enable cooperation and synergies across the SSH community and beyond. The final development cycle delivered new end-user interfaces based on the feedback from the community as well as additional APIs offering yet new functionality within the TBS system to allow an easier onboarding process of the GoTriple users.



Emilie Blotiere
TRIPLE Project Manager at CNRS (Huma-Num)
October 25 2022, 03:58:PM

SAVE THE DATA Atelier GoTriple pour la communauté SHS francophone

Le projet TRIPLE et l'infrastructure de recherche européenne OPERAS, vous convient le jeudi 17 novembre (14h), à une session de formation en ligne de la plateforme européenne GoTriple, plateforme de d ... Read more ...

Luca De Santis
Chief Operating Officer - R&D Director at Net7 Srl
October 20 2022, 10:58:AM

Ready for the OPERAS-GER Open Chat

... in 2 minutes!

Figure 10 - Trust Building System user News Feed

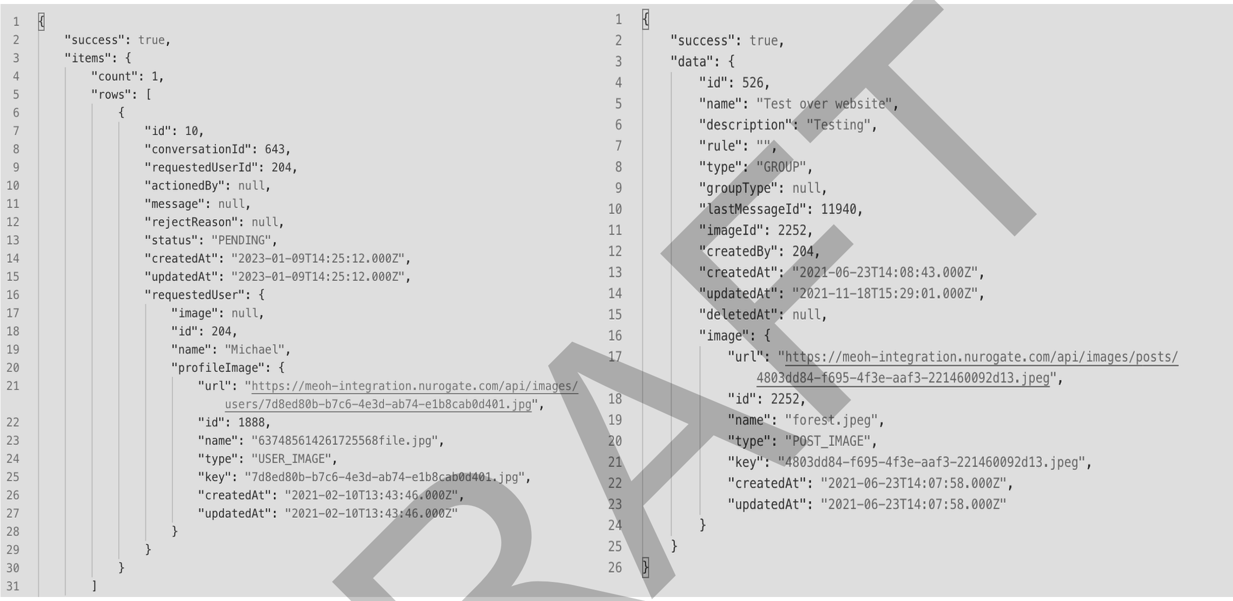
4.1. Public groups within Trust Building System

TBS, as a federated system functioning outside of the core GoTriple platform, was responsible for the onboarding, user management, and participation in groups (e.g., related to SSH topics) within its own ecosystem exclusively. To allow for a better onboarding process and by implication integration with the GoTriple ecosystem, additional incentives to register profiles within TBS had to be established while at the same time opening up the network for a wider variety of stakeholders. A concept of Public Groups has been introduced with the goal of providing external users of GoTriple the ability to see the content of the subject matter before engaging into the discussion or registering on the platform. Additional role management features were introduced so that the admins can review the request of an external user before allowing them to engage into conversations. The following steps of interaction are envisioned within the TBS platform:

- Following the redirection link from GoTriple to the TBS platform, users are able to browse public info pages of individual groups
- Users will be able to send a request to the group administrators with the request to join the group

- The system prompts GoTriple users to fill in a profile page, if they do not have a TBS account yet
- The system sends a request with a profile page to the group admin for review
- Admins can accept or decline the request
- Upon acceptance, users can post and interact in the chat conversation page
- From there on, users can use further functionalities to get invited in private networks by clicking on “add me to your network” or by sending invites to other group members directly from the group info page

The respective endpoints have been provided to support the above described functionality:



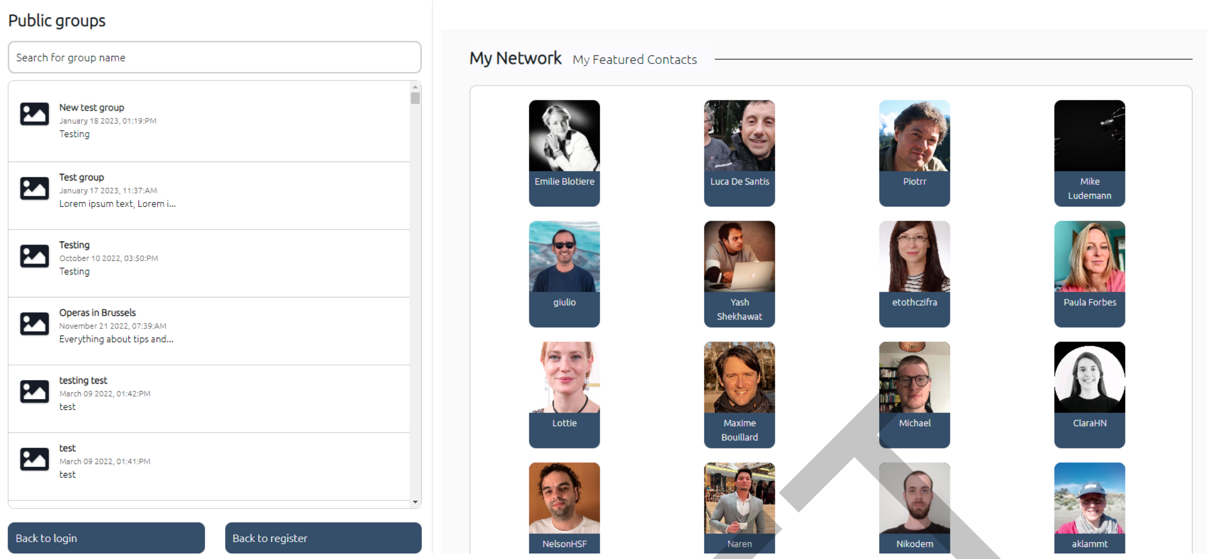
```
{
    "success": true,
    "items": {
        "count": 1,
        "rows": [
            {
                "id": 10,
                "conversationId": 643,
                "requestedUserId": 204,
                "actionedBy": null,
                "message": null,
                "rejectReason": null,
                "status": "PENDING",
                "createdAt": "2023-01-09T14:25:12.000Z",
                "updatedAt": "2023-01-09T14:25:12.000Z",
                "requestedUser": {
                    "image": null,
                    "id": 204,
                    "name": "Michael",
                    "profileImage": {
                        "url": "https://meoh-integration.nurogate.com/api/images/
                            users/7d8ed80b-b7c6-4e3d-ab74-e1b8cab0d401.jpg",
                        "id": 1888,
                        "name": "637485614261725568file.jpg",
                        "type": "USER_IMAGE",
                        "key": "7d8ed80b-b7c6-4e3d-ab74-e1b8cab0d401.jpg",
                        "createdAt": "2021-02-10T13:43:46.000Z",
                        "updatedAt": "2021-02-10T13:43:46.000Z"
                    }
                }
            }
        ]
```

```
{
    "success": true,
    "data": {
        "id": 526,
        "name": "Test over website",
        "description": "Testing",
        "rule": "",
        "type": "GROUP",
        "groupType": null,
        "lastMessageId": 11940,
        "imageId": 2252,
        "createdBy": 204,
        "createdAt": "2021-06-23T14:08:43.000Z",
        "updatedAt": "2021-11-18T15:29:01.000Z",
        "deletedAt": null,
        "image": {
            "url": "https://meoh-integration.nurogate.com/api/images/posts/
                4803dd84-f695-4f3e-aaf3-221460092d13.jpeg",
            "id": 2252,
            "name": "forest.jpeg",
            "type": "POST_IMAGE",
            "key": "4803dd84-f695-4f3e-aaf3-221460092d13.jpeg",
            "createdAt": "2021-06-23T14:07:58.000Z",
            "updatedAt": "2021-06-23T14:07:58.000Z"
        }
    }
}
```

Figure 11 - TBS public groups requests endpoints

With the current approach, additional interfaces were created to allow managing such interactions. A dedicated menu offers an overview of all the public groups the user has created or has been able to join.

On the administrator side, additional fields were added to review the requests and accept or reject those requests. These options are available to the group admins also after the user has been accepted to a given public group.



Public groups

Search for group name

New test group
January 18 2023, 01:19:PM
Testing

Test group
January 17 2023, 11:37:AM
Lorem ipsum text, Lorem i...

Testing
October 10 2022, 03:50:PM
Testing

Operas in Brussels
November 21 2022, 07:39:AM
Everything about tips and...

testing test
March 09 2022, 01:42:PM
test

test
March 09 2022, 01:41:PM
test

Figure 12 - TBS public groups and networks overview

4.2. Trust Building System integration

The TBS took a loose integration strategy by interlinking websites, while striving to include additional functionalities requested during the iterative prototyping. It will further test processes to make it easier for SSH stakeholders to connect and collaborate with each other, providing a feature-rich interaction space. It is nevertheless acknowledged that notification linking between the accounts of GoTriple and TBS can serve as an added value and can be integrated as a feature moving into the sustainable support of the products and services stemming out of a TRIPLE project.

5. Visualisation

In this task led by Open Knowledge Maps (OKMAPS), we have developed two sets of visualisations that are integrated in GoTriple as part of the innovative services. The first set are complex, interactive visualisations, knowledge map, and streamgraph. They are included in the longstanding open source framework Head Start of OKMAPS, which was adapted and extended within TRIPLE for the needs and requirements of the SSH domain. While these two visualisations can be embedded within other interfaces, they are usually used on their own. The second set are less complex diagram components. They are not meant to be used as standalone interfaces. Instead, they are meant to enhance the information included on other pages within the GoTriple platform such as search results and landing pages.

For more information on the basics of these two sets of visualisations, in particular the diagram components that have been finalised by M24, please refer to Deliverable D5.4 [4].

The work done on the visualisations since M24 focused on the knowledge maps and the streamgraph. For these visualisations, we implemented a redesign, developed new features, and improved existing ones.

5.1. User interface redesign

We have decided to do a gentle redesign of the user interfaces for knowledge maps and the streamgraph for various reasons. First of all, we needed space for the new features such as export and cite features. We also wanted to take the opportunity to simplify some of the existing UI elements to make them more easily accessible and to unify and enhance the overall look and feel. Finally, we made it easier to integrate the two visualisations into third-party systems like GoTriple by implementing a look and feel that is devoid of OKMAPS specific corporate identity (CI) and that can more easily fit into any existing CI.

The following UI elements were improved: the context line, the layout of the list items (document entries), the search field and filters section, the credit line, and the additional functionalities toolbar. We harmonised the look and feel of the modal windows by unifying the design and behaviour of the windows and their control elements, such as the copy button. We also took the opportunity to improve our tooltips and interface texts.

5.2. New features

We developed several new functionalities to accommodate user needs and to elevate the researcher workflow through our services. We chose features that are in high demand by our users and that were identified as important during the user research carried out in WP3.

For each of the features, we followed a structured process that included four steps:

1. Research/evaluation of existing features in other discovery services
2. Research for suitable libraries (if necessary)
3. Design and specifications for the new feature
4. Implementation and review cycle for the new feature

The features we developed are described in more detail below.

5.2.1. Export feature and citation feature

We implemented two new features that are related to references and reference management:

1. An export feature that enables users to easily export any document to their reference manager
2. A citation feature that enables users to easily cite any document.

To realise these two features, we evaluated existing export and citation features in other discovery services. The goal was to analyse the functionalities from a UX design, usability, and technical point of view. We evaluated what these features have in common, what works really well, and what doesn't. The following services were taken into consideration: Google Scholar, Lens, Zenodo, Semantic Scholar, Dimensions, and ConnectedPapers.

Design choices and specifications (e.g., call to actions, which citation formats to choose) were based on (i) the outcome of our findings, (ii) careful consideration of what would work really well in our own UI, and (iii) technical feasibility.

For the export feature, we currently provide the BibTeX format.

The citation feature provides four widely used citation styles: APA, MLA, Chicago, and ACM.

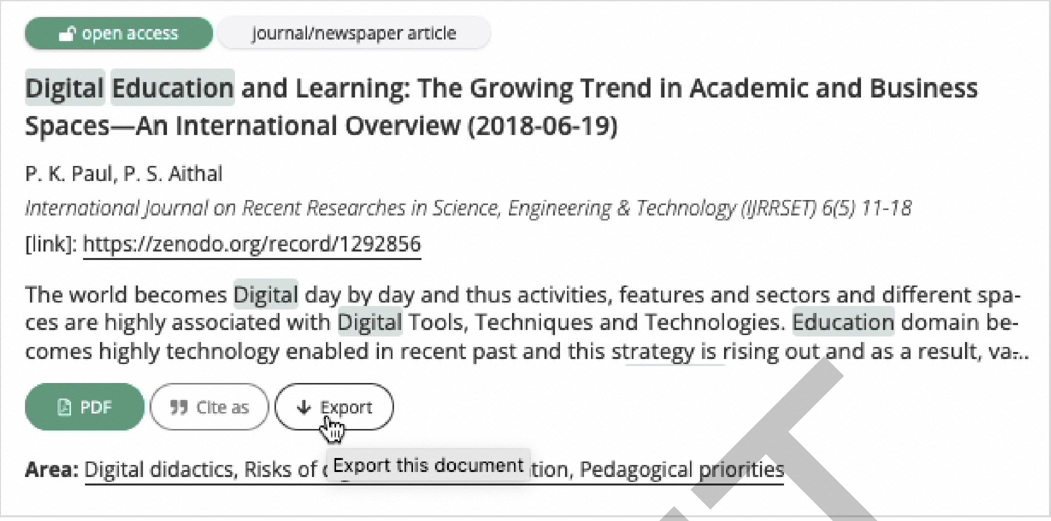
For the export feature, we started the implementation phase with a detailed analysis of the BibTeX specification. The format itself is unfortunately specified very loosely and therefore the support of its features (the document types and their attributes) differs between various services. We decided to primarily focus on compatibility with Zotero, Mendeley, and Endnote. Based on the analysis, we defined a mapping between our metadata and the BibTeX attributes.

In the next step, we used the mapping to implement a backend endpoint that accepts a single document's metadata and returns a formatted BibTeX file. For the file creation, we used a Python library, bibtexparser[5]. The endpoint works both as a data source for the BibTeX content and as a download link for the file.

In the final step, we used the new endpoint in our frontend (see figure 13). We added a new button to each list item (document entry). When the user clicks the button, a new modal window is opened and the BibTeX file preview is displayed (see figure 14). The user can proceed to download the file by clicking the Download button or to copy the file content by clicking the Copy button. The feature was also designed in a way that we can easily add additional export file formats in the future.

The export files created via our interfaces were successfully tested in Zotero, Mendeley, and EndNote.

[5] https://github.com/sciunto-org/python-bibtexparser



Digital Education and Learning: The Growing Trend in Academic and Business Spaces—An International Overview (2018-06-19)

P. K. Paul, P. S. Aithal

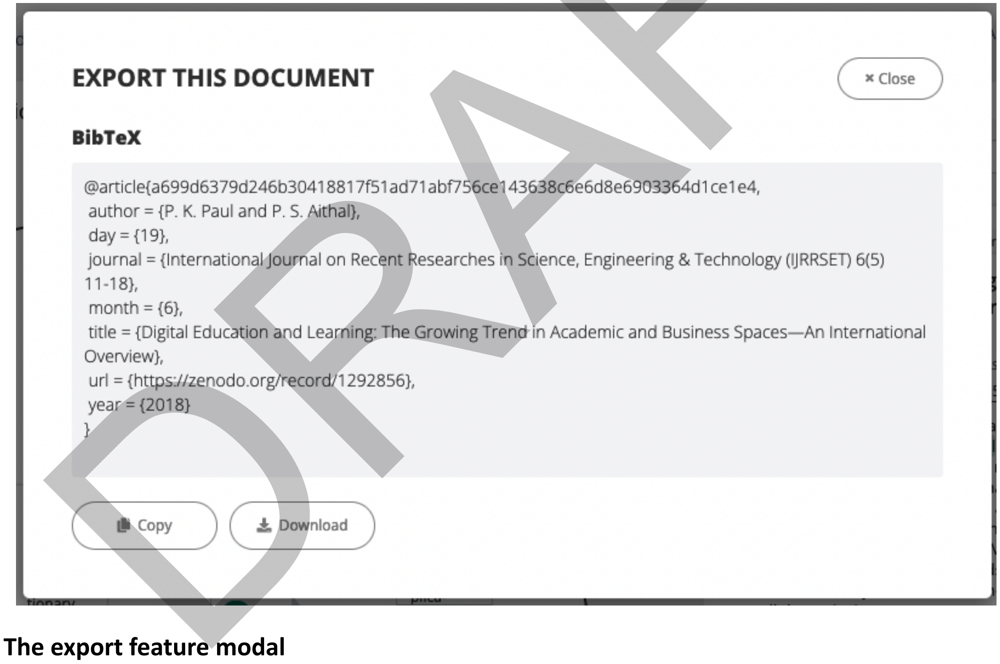
International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET) 6(5) 11-18

[link]: https://zenodo.org/record/1292856

The world becomes Digital day by day and thus activities, features and sectors and different spaces are highly associated with Digital Tools, Techniques and Technologies. Education domain becomes highly technology enabled in recent past and this strategy is rising out and as a result, va-..

Area: Digital didactics, Risks of (tion, Pedagogical priorities

Figure 13 - Access to the export feature and cite-as feature via the list element



```
@article{a699d6379d246b30418817f51ad71abf756ce143638c6e6d8e6903364d1ce1e4,
 author = {P. K. Paul and P. S. Aithal},
 day = {19},
 journal = {International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET) 6(5)
11-18},
 month = {6},
 title = {Digital Education and Learning: The Growing Trend in Academic and Business Spaces—An International
Overview},
 url = {https://zenodo.org/record/1292856},
 year = {2018}
}
```

Figure 14 - The export feature modal

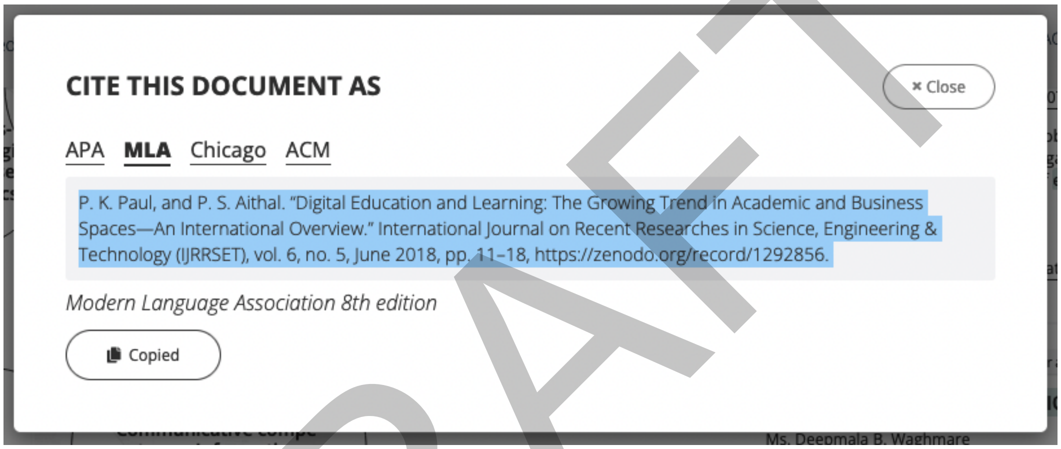
For the citation feature implementation, we decided to use the JavaScript library citation-js[6]. This library, together with its CSL (citation style language) plugin, takes a CSL template and the cited paper metadata as an input and produces the citation string as an output. As a source of the CSL templates, we used the GitHub repository citation-style-language/styles[7].

[6] https://citation.js.org/

[7] https://github.com/citation-style-language/styles

While reviewing the first prototype of the feature, it turned out that the metadata from BASE needs further processing. Information on the document source, such as publication venue, issue, volume, and number of pages are included in a single, combined attribute (dc:source), which often includes additional information such as the publisher as well. This caused the generated citations to be incorrect. We solved the issue by parsing the data attribute. The parsing is performed during the initial data gathering in the backend.

The frontend then gets the sanitised data and displays the citation when the user clicks on the "Cite as" button in the list (see figures 13 and 15). In case the data could not be parsed, the frontend displays a warning message. In the citation feature modal, the user can select between four citation styles mentioned above (APA, MLA, Chicago, and ACM) and copy the citation with a click.



CITE THIS DOCUMENT AS

P. K. Paul, and P. S. Aithal. "Digital Education and Learning: The Growing Trend in Academic and Business Spaces—An International Overview." International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET), vol. 6, no. 5, June 2018, pp. 11–18, https://zenodo.org/record/1292856.

Modern Language Association 8th edition

Figure 15 - The citation feature modal

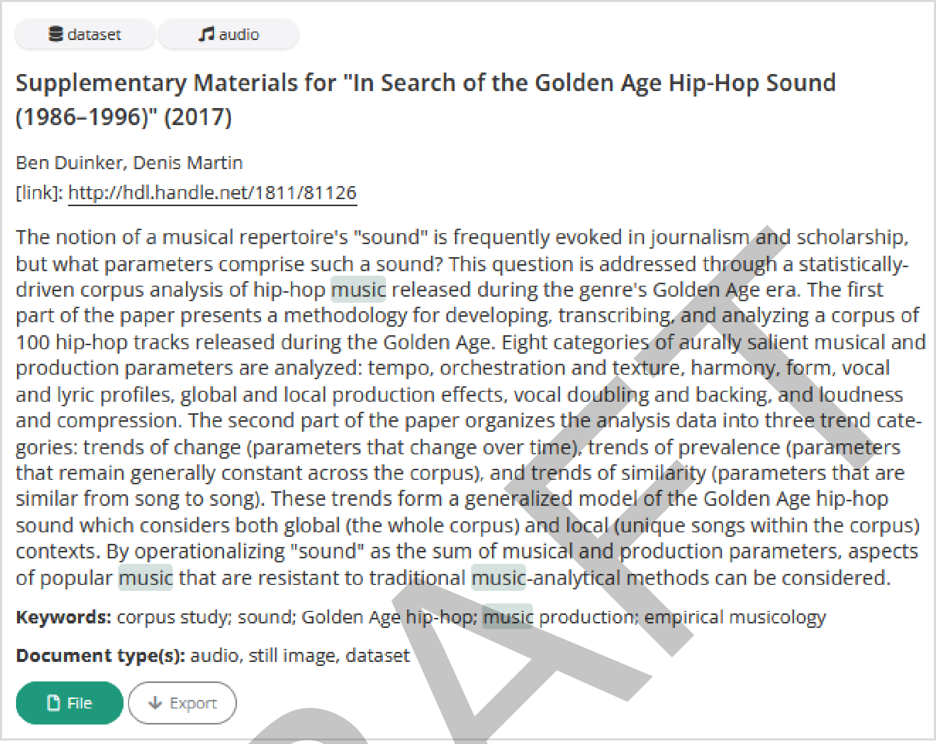
5.2.2. Document types feature

In the implemented knowledge maps, we consider 25 different document types provided by BASE[8], including text, data sets, software, images, audio, and video. Users can create knowledge maps for individual document types or multiple document types.

In the past, users didn't know that we offer such a wide variety and assumed our visualisations only include journal articles. For that reason, we implemented a feature that enables users to easily identify the document type. Furthermore, we decided to highlight the following document types by adding little icons: dataset, map, audio, video, software, still images, and moving image/video (see figure 16). In some cases, the document types are unknown or were not specified in the repository. These items do not have a document type tag.

[8] See BASE Interface Guide - Version 1.25 (May 2022) at https://www.base-search.net/about/download/base_interface.pdf

For documents that are annotated with more than one document type, we display up to two document types as tags and the rest of the document types are mentioned in a separate line underneath the abstract of the document (see figure 14).



Supplementary Materials for "In Search of the Golden Age Hip-Hop Sound (1986–1996)" (2017)

Ben Duinker, Denis Martin
[link]: http://hdl.handle.net/1811/81126

The notion of a musical repertoire's "sound" is frequently evoked in journalism and scholarship, but what parameters comprise such a sound? This question is addressed through a statistically-driven corpus analysis of hip-hop music released during the genre's Golden Age era. The first part of the paper presents a methodology for developing, transcribing, and analyzing a corpus of 100 hip-hop tracks released during the Golden Age. Eight categories of aurally salient musical and production parameters are analyzed: tempo, orchestration and texture, harmony, form, vocal and lyric profiles, global and local production effects, vocal doubling and backing, and loudness and compression. The second part of the paper organizes the analysis data into three trend categories: trends of change (parameters that change over time), trends of prevalence (parameters that remain generally constant across the corpus), and trends of similarity (parameters that are similar from song to song). These trends form a generalized model of the Golden Age hip-hop sound which considers both global (the whole corpus) and local (unique songs within the corpus) contexts. By operationalizing "sound" as the sum of musical and production parameters, aspects of popular music that are resistant to traditional music-analytical methods can be considered.

Keywords: corpus study; sound; Golden Age hip-hop; music production; empirical musicology

Document type(s): audio, still image, dataset

Figure 16 - Document list item with document tags and a supplementary line including all document types beneath the abstract

The technical implementation of this feature was straightforward: we simply displayed the metadata within UI components that we have already implemented for other tags (e.g., the open access tag). The icons originate from the FontAwesome icon library.

5.2.3. Navigation via browser back & forward button

Another often demanded usability improvement that we implemented is the ability to navigate the visualisations using the browser's back and forward buttons. Before, it was only possible to navigate the visualisations using the mouse. For some users, however, the back and forward buttons are the most intuitive way to navigate between the overview and individual areas. This feature also allows the user to easily retrace their interactions.

When a user clicks the browser’s back or forward button, the browser switches to the previous/next URL from the browsing history. Therefore, the first step we had to implement was changing the visualisation URL parameters based on the current state. Since we wanted to enable the back/forward navigation for zooming-in/-out and selecting/deselecting papers, this meant we had to introduce two new parameters into the URL: one representing the zoomed-in area (bubble) and the other representing the selected paper.

With this logic prepared, we could then implement the navigation. Head Start listens on the back/forward button usage. When the user clicks on one of the buttons, Head Start is alerted and updates the visualisation according to the current URL parameters.

5.3. Updates/improvements of existing features

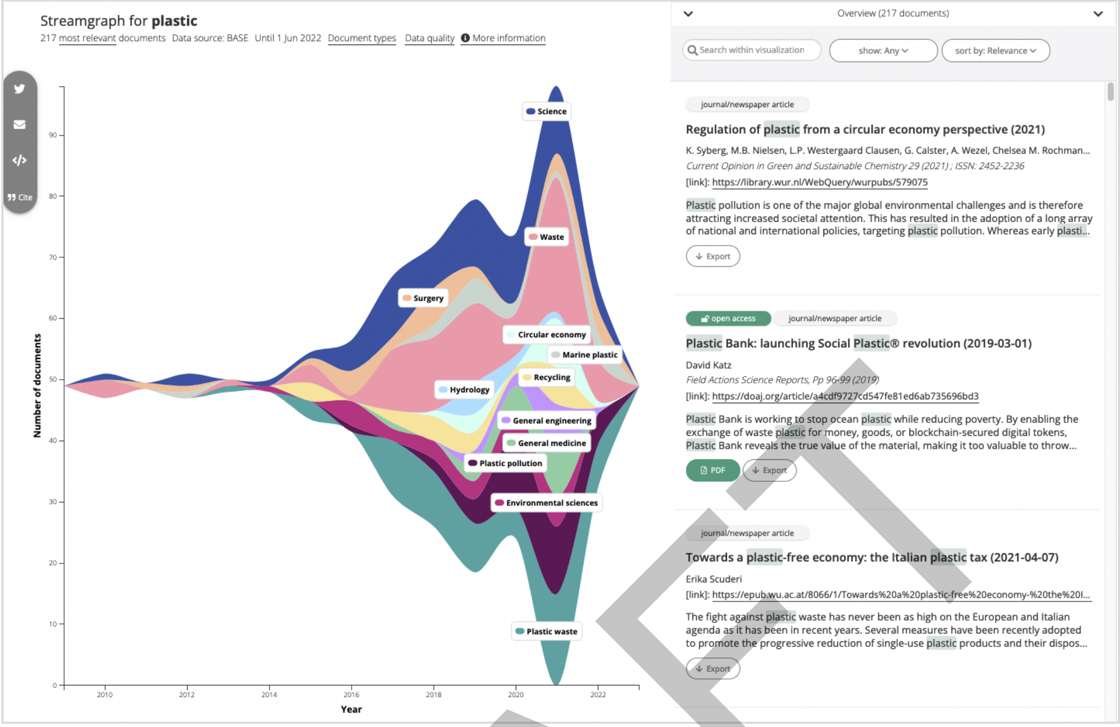
5.3.1. Improvements to the streamgraph visualisation

The streamgraph visualisation is the newest visualisation type in the Head Start library. Based on feedback from our users, we improved the streamgraph considerably (see figure 17).

First, we improved the readability of the visualisation. The stream labels positioning algorithm was upgraded. It is now capable of preventing the labels from overlapping and makes sure no label overflows out of the visualisation. The labels order also respects the order of the streams.

In addition, we improved the streams tooltip. The tooltips overall appearance was harmonised with the rest of the application. Several UX behaviour bugs were fixed that impacted the readability, e.g., previously the tooltip disappeared under certain conditions.

The way the tooltip is positioned relative to the mouse cursor was improved. We now calculate the distance between the tooltip’s bottom and the cursor, whereas previously, we used the distance between the tooltip’s top and the cursor. This allows us to position the tooltip more accurately, even though the tooltip height changes based on the content.



Streamgraph for **plastic**

217 most relevant documents Data source: BASE Until 1 Jun 2022 Document types Data quality More information

Overview (217 documents)

journal/newspaper article

Regulation of plastic from a circular economy perspective (2021)

K. Syberg, M.B. Nielsen, L.P. Westergaard Clausen, G. Calster, A. Wezel, Chelsea M. Rochman...

Current Opinion in Green and Sustainable Chemistry 29 (2021) ; ISSN: 2452-2236

[link]: https://library.wur.nl/WebQuery/wurpubs/579075

Plastic pollution is one of the major global environmental challenges and is therefore attracting increased societal attention. This has resulted in the adoption of a long array of national and international policies, targeting plastic pollution. Whereas early plasti...

open access journal/newspaper article

Plastic Bank: launching Social Plastic® revolution (2019-03-01)

David Katz

Field Actions Science Reports, Pp 96-99 (2019)

[link]: https://doaj.org/article/a4cdf9727cd547fe81ed6ab735696bd3

Plastic Bank is working to stop ocean plastic while reducing poverty. By enabling the exchange of waste plastic for money, goods, or blockchain-secured digital tokens, Plastic Bank reveals the true value of the material, making it too valuable to throw...

journal/newspaper article

Towards a plastic-free economy: the Italian plastic tax (2021-04-07)

Erika Scuderi

[link]: https://epub.wu.ac.at/8066/1/Towards%20a%20plastic-free%20economy-%20the%20I...

The fight against plastic waste has never been as high on the European and Italian agenda as it has been in recent years. Several measures have been recently adopted to promote the progressive reduction of single-use plastic products and their dispos...

Figure 17 - The improved streamgraph visualisation

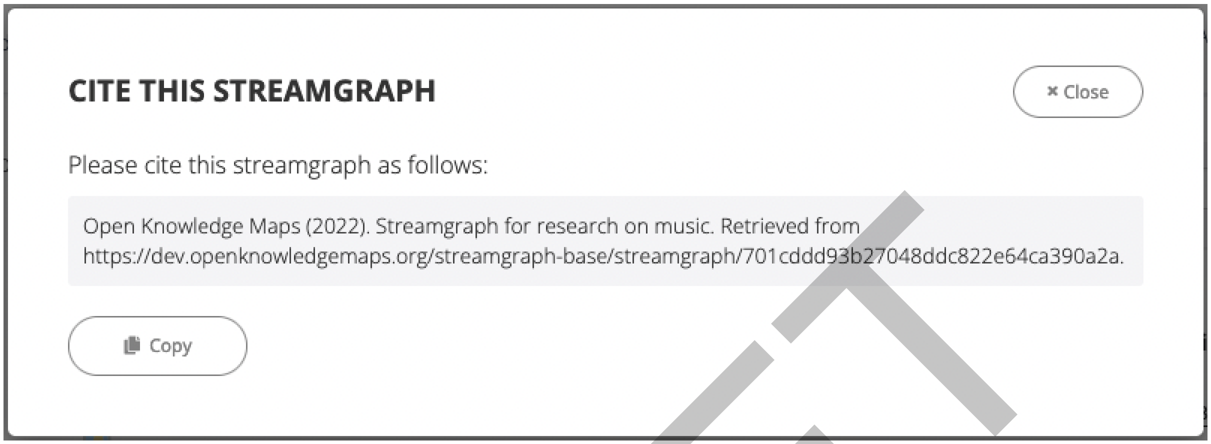
We have also created an extensive FAQ page for the Streamgraph. The link was added to the “more information” pop-up and to the credit line within the visualisation interface. The FAQs describe how the streamgraph is computed and provides guidance on how to interpret the results in more detail.

The changes did not stop there but went beyond the visualisation itself. We also adapted the search flow for the streamgraph. For example, we adjusted the descriptions and labels in the search waiting page and failed page with streamgraph-specific content. We also customised some of the Head Start control elements (e.g., the modal buttons) so that they fit with the streamgraph visualisation.

5.3.2. Improved citation of knowledge map and streamgraph

We enhanced our existing feature “Cite this visualisation”. This feature is distinct from the citation feature described in 5.2.1 and automatically generates a citation for the visualisation itself. It was previously hard to find and only available for knowledge maps. In addition, users had to copy the citation manually.

We have added the functionality to our toolbar. A mouse click on "cite" triggers a pop-up message (see figure 18) that allows the user to copy the citation with a single click. The feature is now available for both visualisations, knowledge map, and streamgraph.



CITE THIS STREAMGRAPH

Please cite this streamgraph as follows:

Open Knowledge Maps (2022). Streamgraph for research on music. Retrieved from https://dev.openknowledgemaps.org/streamgraph-base/streamgraph/701cddd93b27048ddc822e64ca390a2a.

Figure 18 - Cite this streamgraph modal

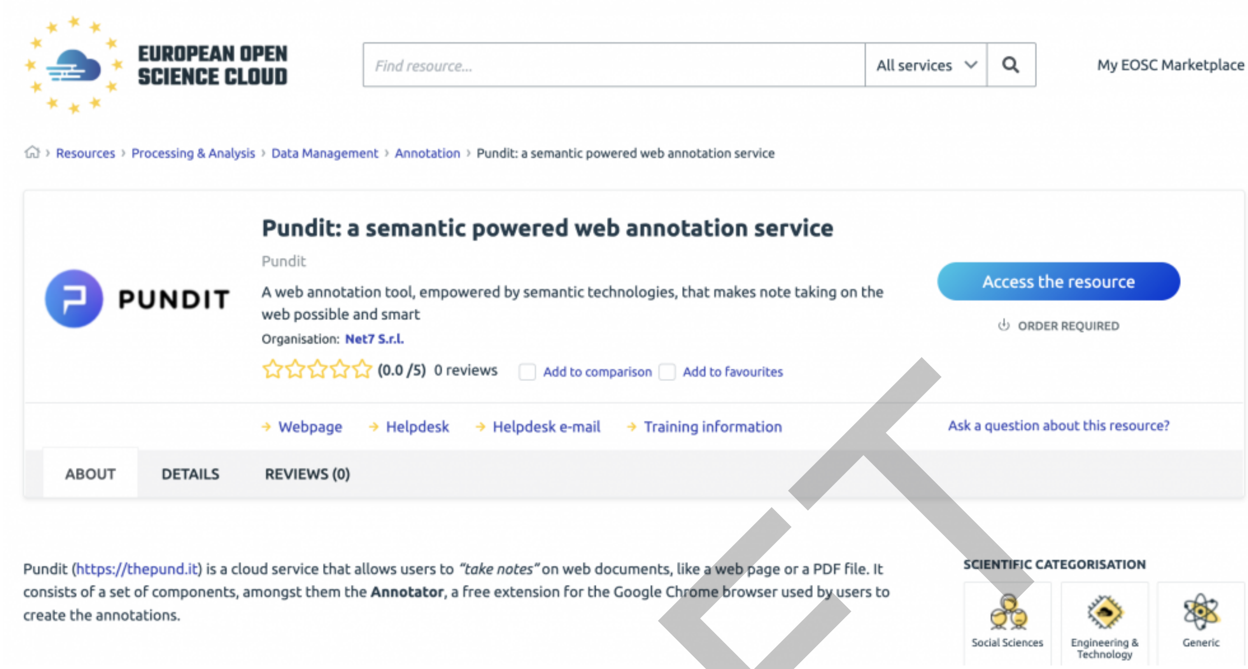
For the implementation, we re-used the modal window component and the copy button component that we had created for other features. The citation follows a fixed scheme and is capable of distinguishing between the two visualisation types.

Finally, we also improved the citation of visualisations using third-party software such as the Zotero browser extension. These extensions read the page metadata in order to retrieve citation information. To accommodate these extensions, we added metadata entries in the form of meta tags to the visualisation's HTML such as author, title, and the creation date of a visualisation.

6. Open annotation tool

In these last months of TRIPLE, the Net7 team working in T5.5, focused on two aspects: on one hand the onboarding of the Open Annotation Tool Pundit in the EOSC marketplace; while on the other on adding new features to the tool, a work that will continue until the end of the project.

As far as the former point is concerned, Pundit has been accepted in the EOSC Marketplace in November 2022. This was one of the results of an EOSC DIH pilot on Pundit that started a few months before.



Pundit: a semantic powered web annotation service

Pundit

A web annotation tool, empowered by semantic technologies, that makes note taking on the web possible and smart

Organisation: **Net7 S.r.l.**

Pundit (https://thepund.it) is a cloud service that allows users to *"take notes"* on web documents, like a web page or a PDF file. It consists of a set of components, amongst them the **Annotator**, a free extension for the Google Chrome browser used by users to create the annotations.

Figure 19 - Pundit in the EOSC Marketplace

The EOSC Digital Innovation Hub (EOSC DIH) is an international and multi-partner cooperation that supports companies, with a focus on start-ups and SMEs, in easily accessing the digital technologies and services offered by the European Open Science Cloud (EOSC).

The Pundit pilot has the following objectives:

(1) supporting Net7 to onboard Pundit in the EOSC Marketplace, which can give to this tool a significant boost in visibility, especially amongst European researchers;

(2) studying other possible use cases and business exploitation options for Pundit, including the release of the service with a Freemium model (free to use for basic functionalities while for more advanced and professional features the user must pay a recurrent subscription);

(3) analysing the technical feasibility of deploying Pundit, which is currently hosted at Amazon AWS, into an EOSC infrastructure.

The on-boarding of Pundit in the EOSC is of course an important achievement, not only for Net7, but for the TRIPLE project as a whole as well.

As said, also the technical side of Pundit hasn't been neglected.

First of all, the EOSC DIH pilot itself has an implementation aspect, that is verifying the feasibility of deploying Pundit in an infrastructure run by EGI. At present, Pundit runs on the Amazon AWS cloud and heavily exploits some of its services (e.g., Lambda functions, API Gateway, OpenSearch server,…). The pilot, which is currently ongoing, is aimed at refactoring some of Pundit's software components for using open-source alternatives to AWS' services, in particular, rewriting the business logic implemented under the serverless paradigm, that are then deployed as AWS' lambda functions. All of them have been developed in TypeScript, a strongly typed JavaScript version, and by using the Node.js and the Serverless libraries.

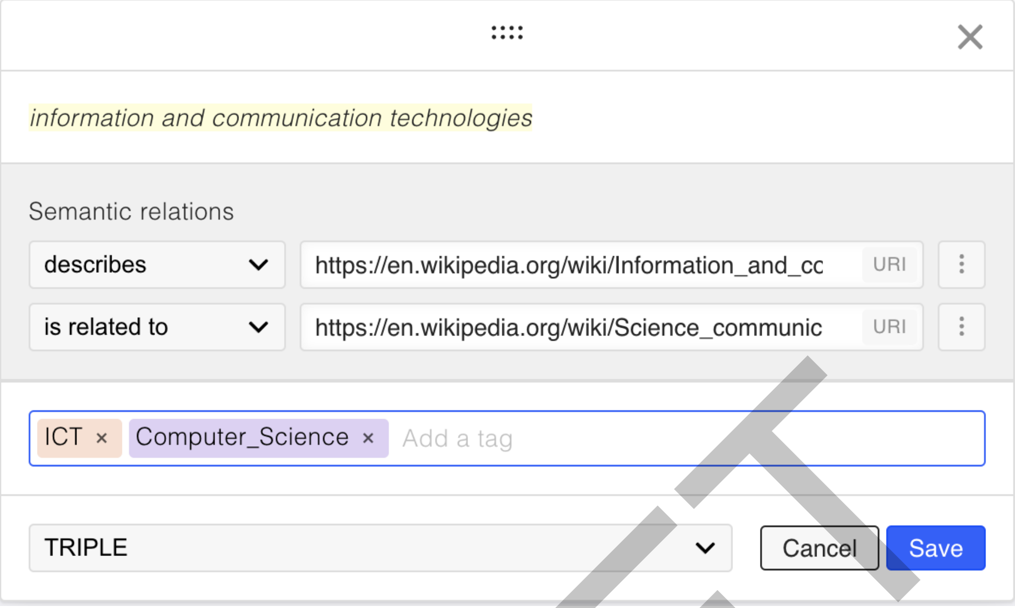
The refactoring consists of getting rid of the Serverless dependencies and to deploy the code as Node.js functions by using the Express framework.

This development is linked to understanding a possible sustainability model for Pundit, an analysis that Net7 is conducting also as a part of the EOSC DIH pilot. The possibility to run Pundit in a public infrastructure instead of Amazon's AWS would of course lower significantly its whole operational costs, an important component to ensuring its availability to the academic and research community.

The rest of the technical activities done in the period focused on adding four main functionalities for users:

- support for semantics
- notebook sharing
- better interoperability with other annotation tools
- the federation of Pundit's and GoTriple's accounts.

The support for semantic annotations has been introduced in Pundit in the course of TRIPLE. As shown in the image that follows, it allows to create an annotation as a semantic web "triple", with the subject being the selected text (or the whole web document), a predicate and an object, specified either as a literal or as a URI.



information and communication technologies

Semantic relations

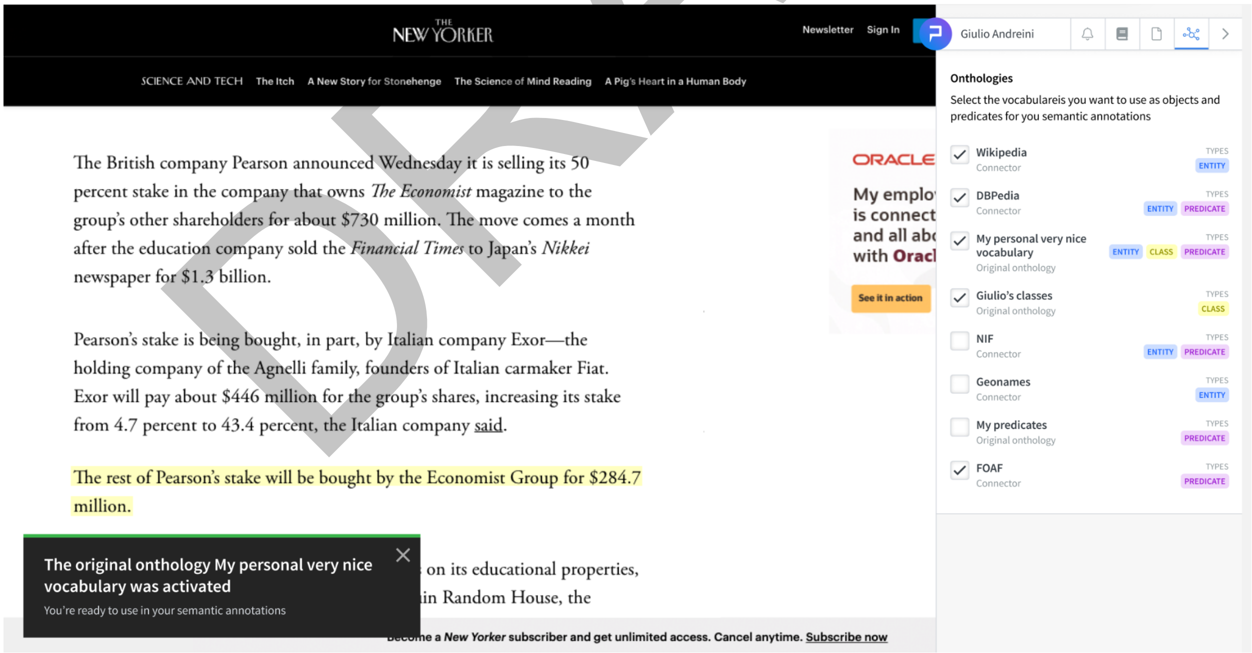
describes	https://en.wikipedia.org/wiki/Information_and_cc	URI
is related to	https://en.wikipedia.org/wiki/Science_communic	URI

ICT × Computer_Science × Add a tag

TRIPLE

Figure 20 - Semantic annotation in Pundit

The next step related to semantics, currently in development, gives users the possibility to select multiple ontologies in the annotation process.



The British company Pearson announced Wednesday it is selling its 50 percent stake in the company that owns *The Economist* magazine to the group's other shareholders for about $730 million. The move comes a month after the education company sold the *Financial Times* to Japan's *Nikkei* newspaper for $1.3 billion.

Pearson's stake is being bought, in part, by Italian company Exor—the holding company of the Agnelli family, founders of Italian carmaker Fiat. Exor will pay about $446 million for the group's shares, increasing its stake from 4.7 percent to 43.4 percent, the Italian company said.

The rest of Pearson's stake will be bought by the Economist Group for $284.7 million.

on its educational properties,
in Random House, the

The original onthology My personal very nice vocabulary was activated

You're ready to use in your semantic annotations

Become a *New Yorker* subscriber and get unlimited access. Cancel anytime. Subscribe now

Onthologies

Select the vocabulareis you want to use as objects and predicates for you semantic annotations

			TYPES
☑	Wikipedia	Connector	ENTITY
☑	DBPedia	Connector	ENTITY PREDICATE
☑	My personal very nice vocabulary	Original onthology	ENTITY CLASS PREDICATE
☑	Giulio's classes	Original onthology	CLASS
☐	NIF	Connector	ENTITY PREDICATE
☐	Geonames	Connector	ENTITY
☐	My predicates	Original onthology	PREDICATE
☑	FOAF	Connector	PREDICATE

Figure 21 - Multiple ontologies support in Pundit

Ontologies will be specified on the Pundit back-end through an easy to use web interface (available for administrators, although in the future we might also allow Premium users to define their own ontologies).

The back-end interface for defining the ontology will be very simple: for each ontology, it only allows to specify predicates, classes, and entities, defined as:

- For predicates and classes:
    - a label (e.g., “rdf:type” or “has_class”)
    - an optional description
    - a fully qualified URI (e.g., https://www.w3.org/1999/02/22-rdf-syntax-ns#type)
- For entities
    - a label: e.g., “NIN”
    - an optional description
    - a fully qualified URI: e.g., https://dbpedia.org/resource/Nine_Inch_Nails

The data model to support this functionality is pretty straightforward, with only two entities, one for the Ontologies, and one for their components, as described in the diagram that follows.

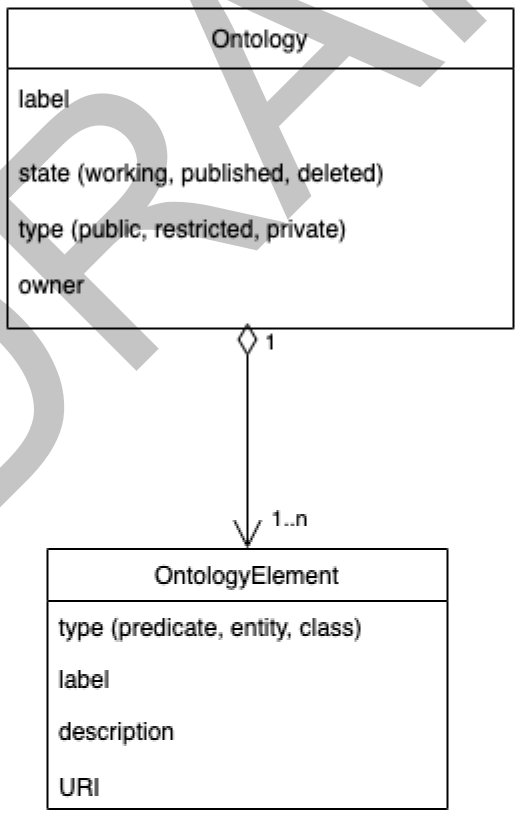

Ontology
label
state (working, published, deleted)
type (public, restricted, private)
owner
1
1..n
OntologyElement
type (predicate, entity, class)
label
description
URI

Figure 22 - Data model to manage ontologies support in Pundit

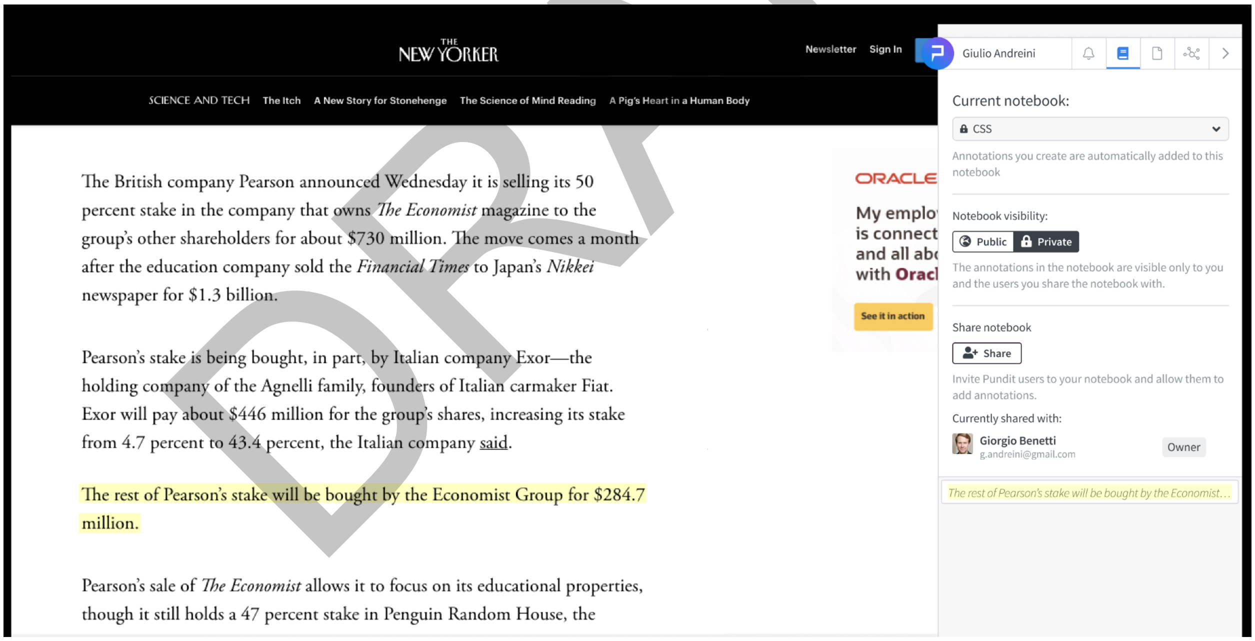
Notebook sharing on the other hand is one of the most requested features by Pundit users. This functionality will allow them to share with selected people a "notebook".

In Pundt, a notebook stores the annotations of a user. Up to now, notebooks could only be either private or public. In the latter case, all of their annotations are freely visible by every Pundit user.

Sharing it with other persons extends the logic for retrieving and visualizing the annotations. When accessing a specific web content in fact - html or PDF alike - with a specific URL, users see:

- their annotations associated to that URL in their notebooks
- all annotations done by other users on that URL contained in public notebooks
- all annotations for that URL that are contained in shared notebooks.

The steps to share a notebook with a user from the Pundit extension are exemplified in the mock-ups that follow:



The British company Pearson announced Wednesday it is selling its 50 percent stake in the company that owns *The Economist* magazine to the group's other shareholders for about $730 million. The move comes a month after the education company sold the *Financial Times* to Japan's *Nikkei* newspaper for $1.3 billion.

Pearson's stake is being bought, in part, by Italian company Exor—the holding company of the Agnelli family, founders of Italian carmaker Fiat. Exor will pay about $446 million for the group's shares, increasing its stake from 4.7 percent to 43.4 percent, the Italian company said.

The rest of Pearson's stake will be bought by the Economist Group for $284.7 million.

Pearson's sale of *The Economist* allows it to focus on its educational properties, though it still holds a 47 percent stake in Penguin Random House, the

Current notebook: CSS

Annotations you create are automatically added to this notebook

Notebook visibility:

The annotations in the notebook are visible only to you and the users you share the notebook with.

Share notebook

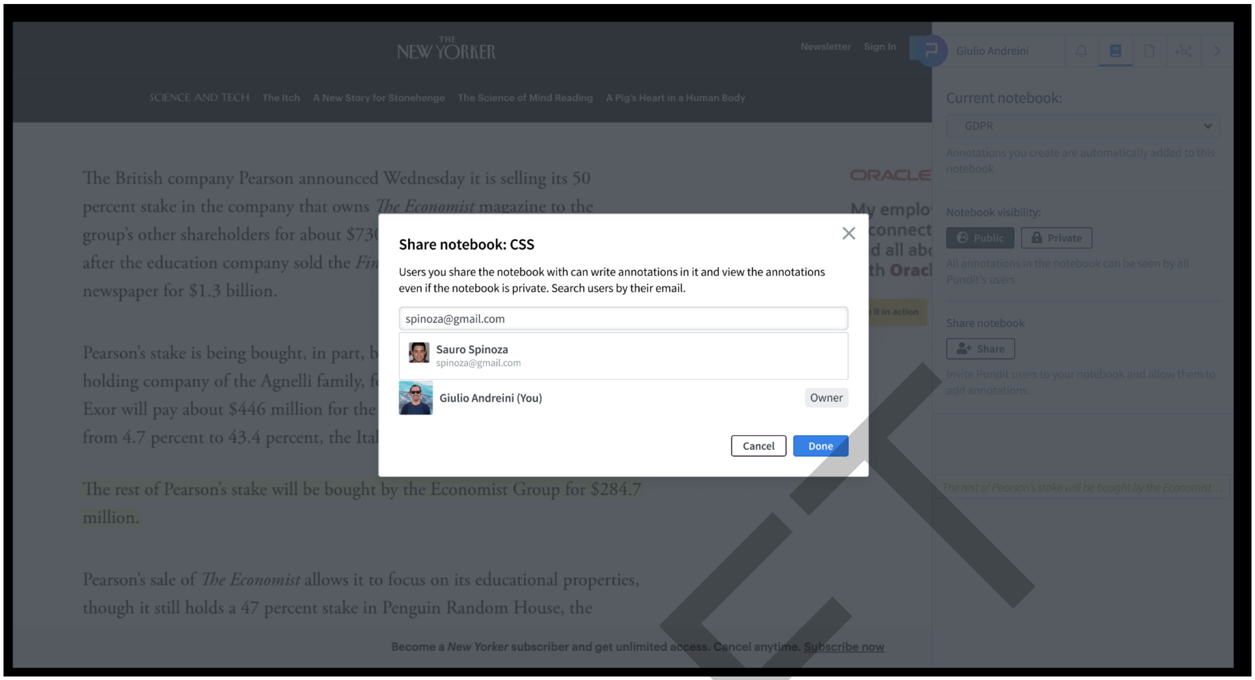
Invite Pundit users to your notebook and allow them to add annotations.

Currently shared with:

Giorgio Benetti
g.andreini@gmail.com

The rest of Pearson's stake will be bought by the Economist…

Figure 23 - Notebook sharing in Pundit


THE NEW YORKER
Newsletter
Sign In
SCIENCE AND TECH
The Itch
A New Story for Stonehenge
The Science of Mind Reading
A Pig's Heart in a Human Body
ORACLE
Become a New Yorker subscriber and get unlimited access. Cancel anytime. Subscribe now

Share notebook: CSS

Users you share the notebook with can write annotations in it and view the annotations even if the notebook is private. Search users by their email.

spinoza@gmail.com

Sauro Spinoza
spinoza@gmail.com

Giulio Andreini (You) Owner

Cancel Done

Giulio Andreini

Current notebook:

GDPR

Annotations you create are automatically added to this notebook

Notebook visibility:

Public Private

All annotations in the notebook can be seen by all Pundit's users

Share notebook

Share

Invite Pundit users to your notebook and allow them to add annotations.

The rest of Pearson's stake will be bought by the Economist…

Figure 24 - Notebook sharing in Pundit

The third line of development for Pundit in the course of TRIPLE is about allowing users to see public annotations done on a page with the popular Hypothes.is[9] annotation tool.

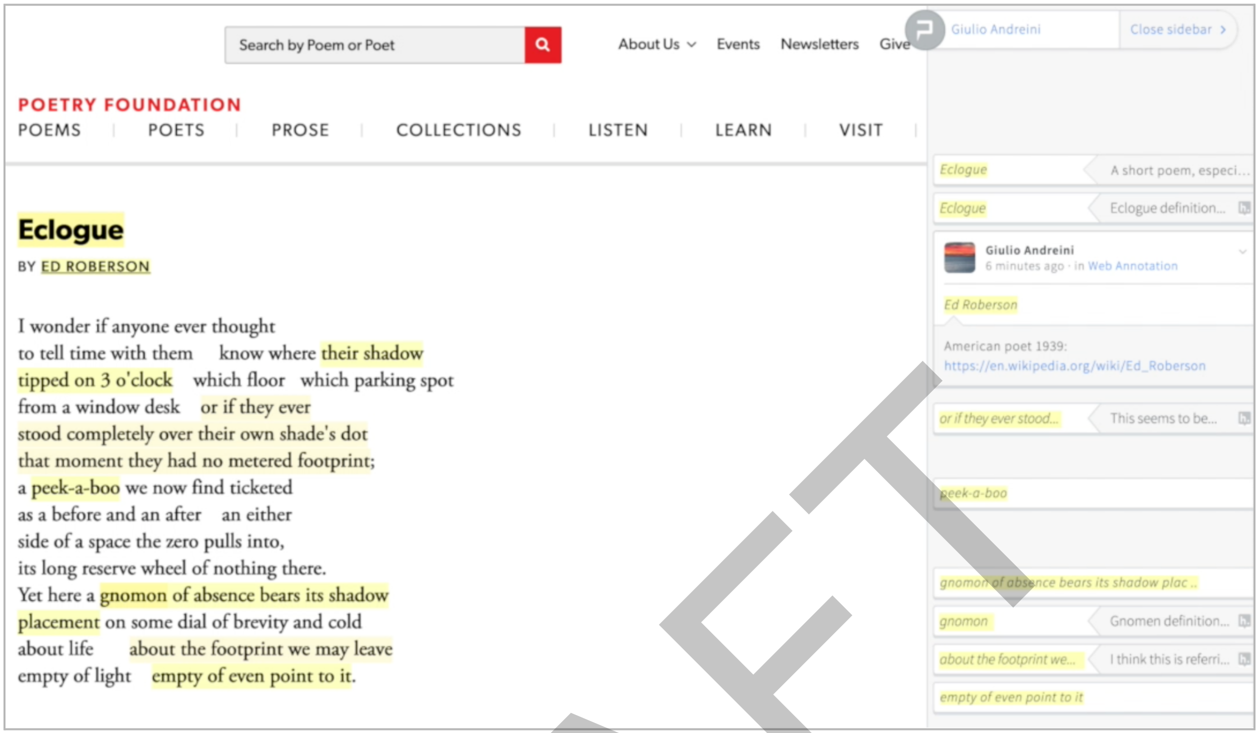
The goal is to encourage the development of an open and interoperable annotation layer on top of web documents.

It should be considered therefore as a proof of concept, which also makes sense given the presence of Pundit in the Annotating All Knowledge (AAK) Coalition[10]. AAK is a coalition of scholarly publishers, platforms, libraries, educational institutions, and technology organizations whose aim is to create an open, interoperable annotation layer over web content.

The Hypothes.is integration feature, mocked in the image that follows, is currently in development and will be released before the end of the TRIPLE Project.

[9] https://web.hypothes.is/

[10] https://hypothes.is/annotating-all-knowledge/



Eclogue

BY ED ROBERSON

I wonder if anyone ever thought
to tell time with them   know where their shadow
tipped on 3 o'clock   which floor   which parking spot
from a window desk   or if they ever
stood completely over their own shade's dot
that moment they had no metered footprint;
a peek-a-boo we now find ticketed
as a before and an after   an either
side of a space the zero pulls into,
its long reserve wheel of nothing there.
Yet here a gnomon of absence bears its shadow
placement on some dial of brevity and cold
about life   about the footprint we may leave
empty of light   empty of even point to it.

Eclogue — A short poem, especi...

Eclogue — Eclogue definition...

Giulio Andreini
6 minutes ago · in Web Annotation

Ed Roberson

American poet 1939:
https://en.wikipedia.org/wiki/Ed_Roberson

or if they ever stood... — This seems to be...

peek-a-boo

gnomon of absence bears its shadow plac ..

gnomon — Gnomen definition...

about the footprint we... — I think this is referri...

empty of even point to it

Figure 25 - Interoperability with Hypothes.is

Finally, the T5.5 team will finalise the integration of the federation between Pundit and GoTriple accounts, to allow Pundit notifications to be shown on the MyGoTriple page of the user. This feature has been already described in the D5.5 (*Report on the Open Annotation Tool*) deliverable [5].

7. Discovery system

In this task, we developed an innovative discovery system that brings together textual and visual interface components. This discovery system is based on Head Start, an open source framework that has been developed by Open Knowledge Maps over more than six years. Head Start provides interactive, web-based overviews of research topics, authors, and projects. The overviews are based on related resources, for example, the most relevant publications on a topic, or the articles and datasets published by an author or project. Head Start is one of the innovative services that is integrated with GoTriple in order to support discovery.

For more information on the basics of the discovery system and its integration in GoTriple, please refer to Deliverable D5.6 (*Discovery System)* [6].

The work done on the discovery system since M24 focused on the transition to the Custom Services, the development of author maps and project maps, and improvements to stability and scalability of the service.

7.1. Transition to the Custom Services pipeline

Based on the positive experiences that we have made with providing Head Start as a hosted solution in TRIPLE, as well as in other projects, we have made a proposal to the TRIPLE consortium to move the GoTriple integration to a new infrastructure run by OKMAPS, the so-called Custom Services. The Custom Services provide Head Start's visualisation capabilities as customizable cloud components on a shared, open, and not-for-profit platform. There are several advantages to the Custom Services regarding the sustainability and maintenance of the service. For one, the Custom Services enable any future improvements of Head Start to automatically be available to GoTriple, even beyond the end of the TRIPLE project. The Custom Services also help to reduce integration and maintenance costs on the GoTriple side as the Custom Services are not running on GoTriple servers, but are provided as cloud services by OKMAPS.

The TRIPLE consortium agreed to this plan and subsequently, the transition to the Custom Services began. In the Custom Services model, GoTriple data is not directly accessed by Head Start, but rather consumed via BASE, the Bielefeld Academic Search Engine. Therefore, it was necessary to create a mapping between the data schemas of GoTriple and BASE as well as implementing an OAI-PMH endpoint to the GoTriple data. This work was carried out in WP4 with strong support from WP5.

On the OKMAPS side, it was necessary to improve the Parameterized Search and Visualization Service so that it can be used within GoTriple (further described in the following section). Additionally, we have extended the retrieval and analytics containers so that streamgraphs can be created on top of BASE data. Finally, we provided the TRIPLE implementation partners in WP4 with updated instructions for integration of the Custom Services.

7.2. Improvements to the Parameterized Search and Visualisation Service

The Parameterized Search and Visualisation Service (PSVS) is an entry point into the automatic pipeline that leads users directly to a visualisation. The entry point is a parameterized search/waiting page that accepts URL-encoded parameters.

On the frontend-side, it provides an integration framework for direct link-based integration, and for an integration with our customisable search box. On the backend-side, it converts user queries (coming through a link or a search-box) into requests to the backend-server. It then handles the response from the backend server and forwards the user to the visualisation.

On a technical level, the PSVS is implemented as a PHP middleware between the frontend and the containerised backend API. The PSVS is functionally similar to a HTTP GET interface.

However, the service is not returning a raw data representation of a visualisation, but rather a web page that triggers the creation of the visualisation. Therefore, the service entry point does not conform to the strict definition of a HTTP GET interface.

We improved the PSVS for use in GoTriple in several ways. First, we migrated the language filter from the old pipeline to the Custom Services. This filter is crucial to the multilingual approach of GoTriple and allows to create visualisations in specific languages.

We also updated the error-handling mechanisms for GoTriple. Error handling is a crucial part of our services that we pay a lot of attention to. Our goal is to lower user frustrations and keep the user on our page. This concerns technical problems, but also searches that were unsuccessful due to user input. We achieve this by providing accurate feedback to the user, so they can easily understand what went wrong and how to fix the issue or how to proceed. In case of unsuccessful searches, the backend applies error detection heuristics and returns a likely response reason to the PSVS and the frontend such as too specific queries, typos, or a too narrow date range. Users are informed about the outcomes and where possible are offered a solution or mitigation to the error.

For GoTriple, we added two improvements to the error handling system in our search flow. First, we have customised our 404 error. It's now also displayed when an invalid or no map ID is provided in the URL. Second, we have added a new error message that is displayed when our services experience too many user requests. In this case, the user is informed about the cause and a countdown appears. Once the countdown runs to zero, the page is refreshed and another attempt to create the visualisation is made automatically.

Finally, the Custom Services pipeline has been updated to be fully embeddable within iframes. The waiting and failed pages of the search-flow are now also embeddable through the use of the embed-parameter in request URLs.

These changes have been reflected in the technical documentation for the Custom Services, in particular, the interactive API descriptions implemented in Swagger.

7.2.1. Addition of Streamgraph to Custom Services

As part of the transition to the Custom Services pipeline, the Streamgraph visualisation has been made available in the Parameterised Search and Visualisation Service. To enable the integration of the new visualisation mode, the PSVS has been extended with a parameter that lets integrators choose between knowledge maps and streamgraphs.

Additionally, our BASE data client was improved and we added functionality for a pagination of search results beyond the first top 100 results. This was necessary to retrieve the higher amount of documents required for producing streamgraph visualisations.

7.3. Development of author and project visualisations

In the first 24 months, we developed knowledge maps and streamgraphs for publications. In the past 16 months, our focus has been on the other two main resources in the GoTriple platform: authors/profiles and projects.

7.3.1. Visualisations for authors

We adapted our pipeline so that it is now possible to create both knowledge maps and streamgraphs for the works of a given author. This gives users a topical overview of the oeuvre of an author and lets them explore the evolution of topics in an author's work over time.

The creation of single-author maps is possible through the introduction of the new search parameter "q_advanced". This parameter allows for a more fine-grained search in the index of our data provider BASE, and within GoTriple, it can be used to search for an author's ORCID. This will produce a knowledge map of all available documents within BASE that are associated with this ORCID. The only precondition is that the ORCID is known by GoTriple beforehand (e.g., as part of an author profile).

On a technical level, q_advanced is implemented as a separate query term within a request to the BASE-API. In the case of an author-maps via ORCID, the parameter works as a filter on the *dcorcid* metadata field.

We also updated relevant elements of the search-flow (waiting and fail pages), the Head Start visualisation client, and associated documentations. For GoTriple, we have also provided integration instructions with certain adjustments to other search parameters, with the goal to improve the results of ORCID-based visualisations.

7.3.2. Visualisations for projects

Similar to the author-based visualisations, the project-based visualisations give an overview of the output of single projects. This enables users to understand what topics have been addressed in the project and which resources have been produced.

The creation of single-project maps is enabled through the introduction of a new search endpoint for the OpenAIRE data source. The endpoint requires two parameters (project_id and funder) and creates a knowledge map of all documents and datasets associated with the project_id. It is a precondition that both parameters are known by GoTriple (e.g., as part of a project overview page).

For this feature, we integrated an existing data client developed in the project VIPER - The Visual Project Explorer[11] into the Custom Services and adjusted the pipeline (backend-API, search-flow and Head Start client) accordingly. The content of the knowledge map is created by retrieving the project output from the OpenAIRE database via our data client. Two requests are made from our side, one for publications and one for datasets. Both results are combined, and the combined metadata set is forwarded to the Custom Services machine learning and natural language processing pipeline.

In the search-flow, we updated the relevant error handling functionality for the new API endpoint, and adjusted parts of the user interface regarding the new set of search parameters.

Lastly, we migrated an existing frontend client for the OpenAIRE data source into the Custom Services pipeline, and modified it to accommodate project-related metadata.

7.4. Improvements to stability and scalability of the service

7.4.1. Distributed processing of machine learning tasks

The work done regarding machine learning tasks is a continuation and completion of a backend refactoring that was carried out over the first 24 months of the project. Previously, the backend was refactored from a monolithic architecture into a microservice architecture where data retrieval, data processing, and persistence operations are clearly separated and can be scaled/distributed across servers on a more granular level.

Scaling in this context means adding additional hardware resources in the form of server nodes, on which identical versions of a microservice are running in parallel, and computing power is allocated more efficiently to tasks. When creating a knowledge map, the computing demand is larger for machine learning and natural language processing tasks, which are happening on-the-fly for each unique user request. They can take 60-80% of the total wait time, even under full utilisation of an individual server.

For this reason, we made use of the capabilities gained by containerizing the R-data processing service in Docker to distribute work across multiple servers (“worker nodes”). Worker nodes are exclusively running a data processing container and are connected to the rest of the system via the Redis-message queue. Worker processes are actively looking for new tasks appearing on the queue, and whenever a user request enters the system and input data is ready for processing, workers pick up the data, do the data processing, and return the result back to the queue. From here, the result is persisted to a database by another service, and returned back to the user by the API.

[11] For a full description see [13]

The architecture is a “hub and spoke” system (see figure 26), with a highly-scalable Redis node in the centre (“hub”) and a variable number of worker nodes (“spokes”) connected to it. Worker nodes are securely connected to the hub node via SSH tunnels.

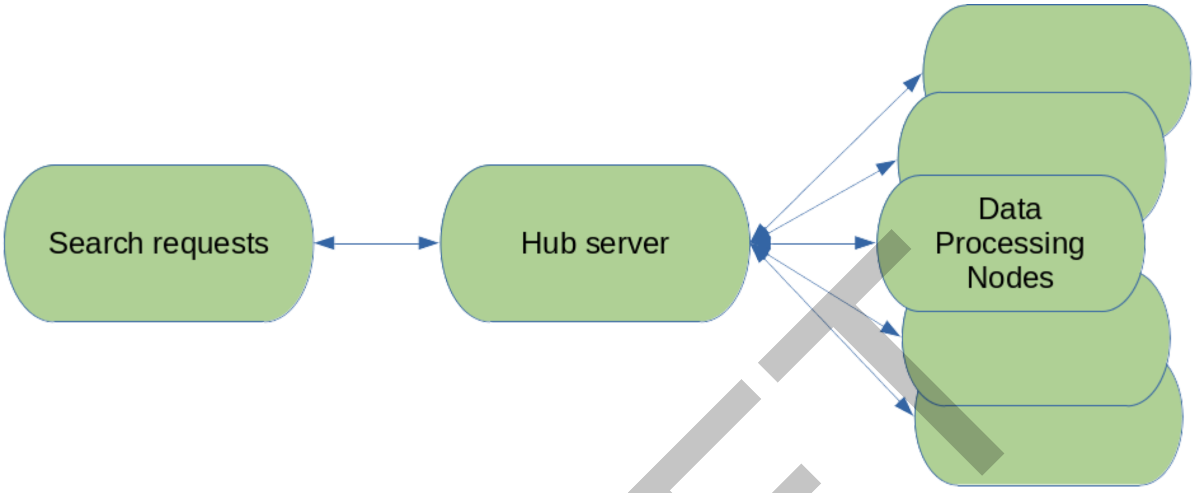


Figure 26 - Simplified overview of the new scalable backend

Adding additional worker nodes is automated via Ansible playbooks. This tool automates all steps for setting up a fully functional worker node on a fresh root server, and adding it to the network. Ansible playbooks are also used to update containers with new versions of the backend across all involved servers.

The benefits of the distributed processing are already measurable in our production system:

- The restrictions on maximum average and peak search loads are removed by running on a one-node system, and it is now possible to handle an average search load beyond ~200 knowledge maps/hour.
- Additionally, creating a knowledge map takes less time because individual processing steps can now fully utilise server resources for parallel processing, without competing with other requests. This means wait times for users are reduced considerably, by about 10-15 seconds for each map (a 25%-30% reduction).

7.4.2. Upgrades and maintenance of R data processing backend

As part of the upgrades to distributed processing, the R data processing services underwent a code cleanup and updates of dependencies. A performance test identified minor bottlenecks that were improved, and a code review revealed some removable natural language processing functions that consumed computing power without purpose. Overall, the upgrades and maintenance yielded small, but noticeable, additional performance improvements.

7.4.3. Upgrades and maintenance of PHP-service layer

A small update was done within the PHP-service layer that intermediates between user requests and the backend service API. A lightweight API wrapper was created to improve code structure and readability, and a new functionality to improve peak-load error handling was added.

7.5. Further improvements

7.5.1. Data processing

The frontend of our application needs to handle large and complex data provided by our backend. The documents in our visualisations are represented by data entries that have many attributes - from metadata provided by the data source, such as the document title or publication date, to internal information used for the computation of the visualisation. The metadata transformation and rendering process is therefore very sensitive to data tidiness - any deviation from the expected data scheme can cause the application to crash.

That is why we decided to add a new layer to our frontend data processing: a data validation and sanitization layer. Its purpose is to check if the data that came from our backend fulfils the expected scheme, and if not, to try to correct it. Possible corrections include data format conversions and inference (e.g., testing if a string is actually an array), or adding fallback values in case of missing data. If no correction can be applied, the metadata remains empty, and an error is logged. This helps us with reducing metadata-related bugs to a minimum. It also results in a more stable and robust frontend environment, which allows us to add new features more easily. Another important benefit is that we have improved backwards compatibility with knowledge maps that were created in the past.

7.5.2. Development environment

In order to simplify and speed up the development process, we decided to improve our development environment. This work was carried out in two steps: first, we cleaned up the codebase repository and removed dead code and unused files. Second, we added more local examples - what we call sandbox versions of our app frontend running on localhost. The local examples make development easier, because we can quickly test the new features of all of our data integrations without having to deploy them anywhere.

These changes mostly result in a faster feature development. We can now also test our frontend features across different data integrations more easily, which increases the chance of spotting and fixing data-integration-related bugs early on. And finally, our repository is now tidier as well as better maintainable.

8. Conclusions

As emerged by the descriptions of the previous chapters, we can affirm that the goal of WP5 to integrate Innovative Services in GoTriple has been completely achieved. All of the services are available to the users of the Discovery Platform, some of them since the first alpha release of GoTriple.

The progress doesn't stop here, as the WP5 team is still at work, in close collaboration with WP4, to complete and improve their services for these remaining months of TRIPLE.

While the work in T5.1 is completed, for T5.2, the next integration regards the inclusion of the "Most Popular" content (documents and projects) in the home page of GoTriple.

For T5.3, the team will refine the logic to ease the onboarding of new users in the Trust Building System.

Together with further improvements and refinements, the work of T5.4 and T5.6 will be devoted to the integration of the projects' knowledge maps in GoTriple.

Finally, together with the development of new functionalities of Pundit currently in the works, the T5.5 team will finalise the federation between Pundit's and GoTriple's accounts, to allow Pundit notifications to be shown on the MyGoTriple page of the user.

9. References

[1] TRIPLE Deliverable D5.1 - Report on Third-Party Applications: https://doi.org/10.5281/zenodo.5702398

[2] TRIPLE Deliverable D5.2 - Report on the Recommender System: https://doi.org/10.5281/zenodo.5702426

[3] TRIPLE Deliverable D5.3 - Report on the Trust-Building System: https://doi.org/10.5281/zenodo.5702559

[4] TRIPLE Deliverable D5.4 - Report on the Visualisations: https://doi.org/10.5281/zenodo.5702575

[5] TRIPLE Deliverable D5.5 - Report on the Open Annotation Tool: https://doi.org/10.5281/zenodo.5653218

[6] TRIPLE Deliverable D5.6 - Discovery System: https://doi.org/10.5281/zenodo.5702656

[7] TRIPLE Deliverable D2.6 - Report on Global Data Retrieval: https://zenodo.org/record/7359654#.ZAMeyuzMI40

[8] TRIPLE Deliverable D2.5 - Report on Data Enrichment: https://doi.org/10.5281/zenodo.73596538

[9] Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-based systems, 46, 109-132.

[10] Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In The adaptive web (pp. 291-324). Springer, Berlin, Heidelberg.

[11] Lacic, E., Kowald, D., Eberhard, L., Trattner, C., Parra, D., & Marinho, L. B. (2013). Utilizing online social network and location-based data to recommend products and categories in online marketplaces. In Mining, Modeling, and Recommending 'Things' in Social Media (pp. 96-115). Springer, Cham.

[12] Jäschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2007). Tag recommendations in folksonomies. In Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, September 17-21, 2007. Proceedings 11 (pp. 506-514). Springer Berlin Heidelberg.

[13] Kraker, P., Schramm, M., Kittel, C., Chamberlain, S., & Arrow, T. (2018). VIPER: The Visual Project Explorer. Zenodo. https://doi.org/10.5281/zenodo.1248119