

Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782

Start Date of Project: 01/01/2019

Duration: 40 months

User Stories for the SSH Open Marketplace

Dissemination Level	PU, Public
Due Date of Deliverable	not applicable
Actual Submission Date	04/05/20
Work Package	WP7 – Creating the SSH Open Marketplace
Task	T7.1
Type	Report
Approval Status	not applicable
Version	V1.0
Number of Pages	p.1 – p.14

Abstract

To include the community in the design and creation of the SSH Open Marketplace, the project SSHOC pursues various means of participatory design. One element of this efforts are user stories, that describe researchers' reception and use of the marketplace. This document included the four initial user stories as of April 2020.

Author List

Organisation	Name	Contact Information
UGOE	Stefan Buddenbohm	buddenbohm@sub.uni-goettingen.de
UGOE	Raisa Barthauer	barthauer@sub.uni-goettingen.de
DARIAH ERIC	Laure Barbot	laure.barbot@dariah.eu
UGOE	Andrea Bertino	bertino@sub.uni-goettingen.de
DARIAH ERIC	Frank Fischer	frank.fischer@dariah.eu
CNRS	Clara Petitfils	clara.petitfils@huma-num.fr
ÖAW	Klaus Illmayer	Klaus.Illmayer@oeaw.ac.at

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.



SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the EU Horizon 2020 Research and Innovation Programme (2014-2020); H2020-INFRAEOSC-04-2018, under the agreement No. 823782

Table of Contents

Aim of this document	4
Use case A - Humboldt	5
The use case researcher: field of study, prior knowledge, ground assumptions	5
The use case exemplary research project	5
Use of the SSH Open Marketplace	6
Use case B - Spatial visualisation of historical data	8
The use case researcher: field of study, prior knowledge, ground assumptions	8
The use case exemplary research project	8
Use of the SSH Open Marketplace	9
Use case C - Ethnic and Migration Studies	10
The use case researcher: field of study, prior knowledge, ground assumptions	10
The use case exemplary research project	10
Use of the SSH Open Marketplace	11
Use case D - Young Researcher	12
The use case researcher: field of study, prior knowledge, ground assumptions	12
The use case exemplary research project	12
Use of the SSH Open Marketplace	12
Conclusions	14

Aim of this document

We want to create user stories which explain to potential users how to use and how to benefit from the SSH Open Marketplace as a researcher. On the other hand the user stories - if commented by the readers - may help us to evaluate and enrich the MP's contents. This is the most important point at this stage of the MP's development: to ensure that our coverage of content maps with the researchers' requirements.

The user stories are structured into logical sections: a persona (=the researcher) and its knowledge, the research question, and the persona's process of using the marketplace.

As a proof of concept the user stories include real content items, which will be edited/enriched manually in the MP, so that a clickable demo would be possible. This is an important part of the user stories, as this bridges the gap between a theoretical presentation of the ideal marketplace to the current development status of the platform. We want to induce feedback regarding the functions, the presentation and the contents of the marketplace.

Other potential user stories are of course possible but at this stage of the development and just before the alpha release of the marketplace we focus on four initial user stories described below. These four initial user stories contain roughly 40 items, which will be in the end "clickable" from the presentation to the marketplace.

Use case A - Humboldt

Wilhelm von Humboldt's works on educational reforms - Commented edition and textual analysis of a correspondence of the German 18th/19th century scholar.

a. The use case researcher: field of study, prior knowledge, ground assumptions

Marianne is a researcher interested in the textual analysis of text corpora of German philosophers dating from the 19th century. Marianne has studied German literature and philosophy at European universities. She is familiar with the conventional methods in the field of comparative literature and linguistics, but also a curious adaptor of the recent developments under the label Digital Humanities. In doing so she has acquired some knowledge of new methods, standards and technologies, for instance applying XML transcriptions of text for mining and analysing.

She has heard of the SSH Open Marketplace via a tweet from DARIAH ERIC¹. A recent tweet highlighted the SSH Open Marketplace as a growing resource for information on resources for textual analysis. The tweet was pointing out the contextualized approach of the marketplace. This means that the individual entities are - wherever possible - embedded in a logical context, called a scenario (or workflow).

b. The use case exemplary research project²



Although Wilhelm von Humboldt³ is sufficiently studied, there are some works of him on the subject of educational reforms in Prussia resting in archives, which haven't been explored yet. Marianne wants to investigate a small text corpus of a correspondence of Wilhelm von Humboldt on educational reform with other colleagues over the course of fifteen years. She has access to the text material, which is - fortunately - available as digitized material in an archive. The material consists of high resolution scanned images of the text corpus.

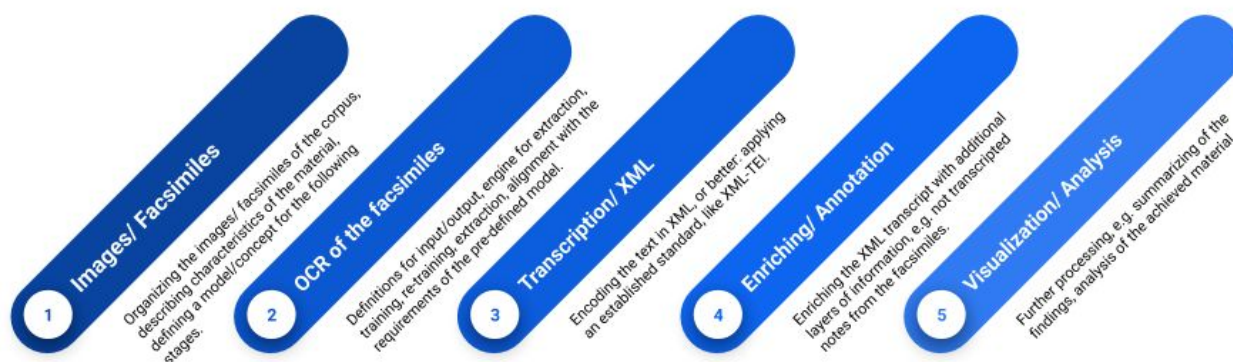
Marianne would be interested to easily search for specific concepts in the correspondence and to digitally annotate the corpus. It would help her to gain an overview of the corpus and to use it to clarify Humboldt (and his colleagues) position(s) about educational reforms over the 15 years covered. Depending on her results, she might also be interested to work on a digital edition of this specific subset of Humboldt correspondence. Although Marianne has been a curious follower of recent developments, for instance textual analysis of text corpora in the Digital Humanities, this would be her first project applying such methods and technologies for the whole process. This could serve her as groundwork for future research work.

¹ <https://twitter.com/DARIAHeu>

² Source: https://commons.wikimedia.org/wiki/File:DBPB_1952_100_Wilhelm_von_Humboldt.jpg

³ https://en.wikipedia.org/wiki/Wilhelm_von_Humboldt

Which is one of Marianne's urgent questions: she wants to understand the different steps she should follow to work on her corpus because she is sure that decisions for certain tools or standards may have repercussions later on. Her interest was aroused by the tweet mentioning scenarios or research workflows and she hopes to gain inspiration from other research processes and projects.



c. Use of the SSH Open Marketplace

Marianne access the SSH Open Marketplace via the standard landing page⁴ and her initial keywords for searching the Marketplace are: textual analysis, XML transcription, digital edition, OCR, facsimile.

The most interesting topics to her seem:

- [Extract textual content from images](#) (SSK, not yet included in the MVP⁵)
- [Generating an Ordered Data Set from an OCR Text File](https://sshoc-marketplace.acdh-dev.oeaw.ac.at/training-materials/1420) (Programming Historian, MVP: <https://sshoc-marketplace.acdh-dev.oeaw.ac.at/training-materials/1420>)
- [Oxygen](https://sshoc-marketplace.acdh-dev.oeaw.ac.at/tools/420) (TAPoR, MPV: <https://sshoc-marketplace.acdh-dev.oeaw.ac.at/tools/420>)
- [Voyant](https://sshoc-marketplace.acdh-dev.oeaw.ac.at/search?query=voyant) (TAPoR, MVP: <https://sshoc-marketplace.acdh-dev.oeaw.ac.at/search?query=voyant>)
- [Deutsches Textarchiv](#) (or SSK⁶, not yet included in the MVP)
- [DTABf](#) (SSK, not yet included in the MVP)
- [Das DTABf in neuem Gewand](#) (SSK, not yet included in the MVP)

She starts with the [Hypotheses-Article describing the DTA-Basisformat](#) as it seems to hit her use case, is written in German language, and is published on Hypotheses.org, a known and well-reputed source for her. The Hypotheses-article also proves as useful as it links to other sources, either on hypotheses.org or other resources in the web.

⁴ <https://sshoc-marketplace.acdh-dev.oeaw.ac.at/>

⁵ The items are manually curated and included in the minimum viable product of the SSH Open Marketplace. The highlighted items have to be created completely new.

⁶ http://ssk.huma-num.fr/#/scenarios/SSK_sc_DTABf

The SSH Open Marketplace is not hosting its resources directly but rather harvests them from third party sources, enriches and structures them. These harvested sources can be registries like TAPoR but also a blog/publishing platform like hypotheses.org.

After this she pays a closer look at the [Deutsche Textarchiv](#), which is already known to her and proves useful for Marianne as it presents examples of corpora. She uses this source to align her knowledge with the DTA as a standard or best practice. She is even able and delighted to find some [Wilhelm von Humboldt related sources](#). The other results from the SSH Open Marketplace are addressing certain steps in the process or present tools which may be useful for her. What is easing her decision making process: each resource comes with additional information on how to use it (tutorials, licencing) and provenance, which including if the resource is up to date.

Use case B - Spatial visualisation of historical data

a. The use case researcher: field of study, prior knowledge, ground assumptions

Lorenzo is a young researcher in linguistics and Italian studies. He is already quite familiar with digital humanities tools and methods such as NPL, Textometry, Stylometry. He would like to improve his skills in Mapping and Data Visualization.



DARIAH-DE Geo-Browser

Load Data

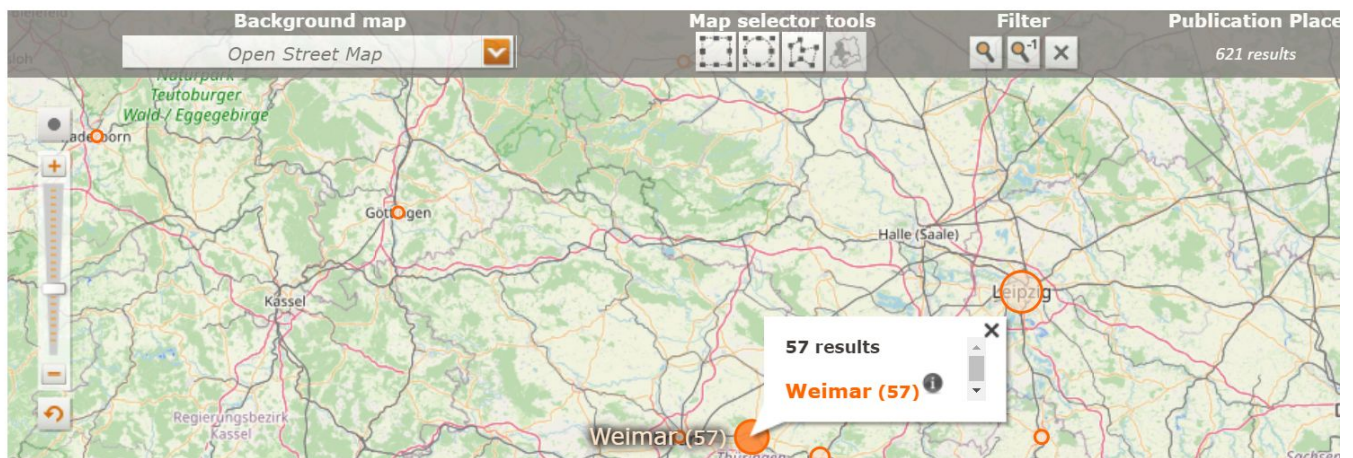
Static Data ▾
Friedrich Schiller ▾ **load**

Load Overlay

KML File URL ▾
 load KML

Dataset Information

Magnetic Link
schiller.kml [↗](#)



7

b. The use case exemplary research project

Lorenzo is starting a project based on several datasets representing 3.500 letters of people sentenced to death during the First and Second World War in Italy thanks to which he can perform linguistic and thematic analysis to contribute to historiography. He wants to use a combination of network analysis and spatial analysis to understand what categories of people were mainly repressed and where and when this happened.

Lorenzo heard about the SSH Open Marketplace during a recent conference, and he is curious to see if other tools or methods than the one he is using are referenced in the portal. Especially for the spatial

⁷ Source: DARIAH-DE GeoBrowser as tool to visualize geospatial data. <https://geobrowser.de/dariah.eu/>

visualisation, he would be interested to change the tools he starts using and find an alternative free and open solution.

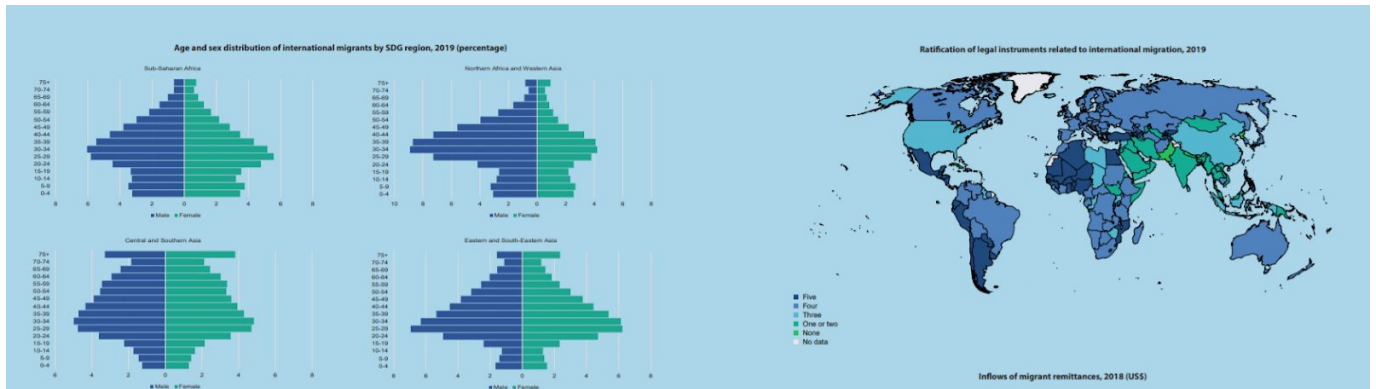
c. Use of the SSH Open Marketplace

Consulting the SSH Open Marketplace, Lorenzo is first trying to search the tools he is using. Searching for [Gephi](#), he also sees the alternative tools [Cytoscape](#) and [Palladio](#) mentioned. As he never used Palladio, but wanted too, he is quite curious about it and consults the Palladio entry in the Marketplace, where he notices the training material [From Hermeneutics to Data to Networks: Data Extraction and Network Visualization of Historical Sources](#) (coming from the "[Programming Historian](#)" website) which he finds quite interesting (even if too basic for its needs).

As he is more interested in the spatial visualisation aspects, he decides to browse the contents of the Marketplace directly with this category, and the most relevant results of his search are the following tools: [StoryMap Js](#), [Map Warper](#), Palladio, [QGIS](#), [Mapbox](#), [ArcGIS](#), [Carto](#) (these two he is already using and would like to replace them). And more interesting for him, the lesson [Using JavaScript to Create Maps of Correspondence](#) (source: [Programming Historian](#)). He finds in this lesson what he needed to understand how he can work with the mapping visualisation of his data.

Use case C - Ethnic and Migration Studies

FAIRification of data in Comparative Politics: what does it mean and how to proceed?



8

a. The use case researcher: field of study, prior knowledge, ground assumptions

Matéo is the assistant of Milena, Professor in Political Science who works among other things on the political consequences and implications of migration processes. Matéo and Milena already heard about standards used to facilitate the reuse of resources, and the Research Data Alliance and the Data Documentation Initiative are not unknown to them, but they are not active members of these organisations. They would like to understand the concrete steps and consequences of applying standards recommended by these institutions on the resources they are working with.

b. The use case exemplary research project

Matéo was hired after the end of a research project conducted by Milena about political representation of citizens of immigrant origin to work on files produced during the project and make data created available. In the meantime, Milena started a new research project about the political incorporation of migrants' minorities and their descendants in European countries, and she would like to avoid having to wait until the end of the project to deal with the data collected during the surveys that will be conducted. Milena heard about the FAIR principles and she suggested to Matéo to investigate what it would mean for both their projects to follow these principles. Matéo is looking for recommendations or workflow examples suited to their needs. Some of their colleagues involved in the SSHOC project/ a COST action recommend them to consult the SSH Open Marketplace.

⁸ Source: United Nations Population Division, International migration wallchart 2019
https://www.un.org/en/development/desa/population/migration/publications/wallchart/docs/MigrationStock2019_Wallchart.pdf

c. Use of the SSH Open Marketplace

Matéo starts his search with the FAIR acronym and refined it with the “training materials” type of resources. The first interesting result he finds is the [CESSDA Data Management Expert Guide](#). It’s a good source of information for him, because it covers all the data life cycle. This resource will be useful for the project he is working on and for Milena as well. In the subsection dedicated to FAIR data, he especially likes the [FAIR checklist](#) recommended.

Coming back on the SSH Open Marketplace website and using the disciplines browsing (or activity types or even keywords linked to migration), he discovers under the “Political science” topic very useful resources: the [SSHOC Workshop: Caring for Sharing – Data Management and FAIRness of Migration Data](#) training material, in which the [Ethnic and Migrant Minorities \(EMM\) Survey Registry](#) that bring together the currently scattered EMM survey data into a centralised and systematised database. One of the tools providing an “health check for numeric data” mentioned in the webinar slides is also of interest for him: [QAMyData](#).

Of course, he also recognises in the list of results some of the tools and software he is using to perform statistical analysis like [Stata](#), [R](#), [SPSS](#) and what he likes as he is browsing to have a look at the different tools entries, is that the most active forum for the R community for example ([R Stack Overflow](#) and [R Studio Forum](#)) are referenced in the description of the software in the SSH Open Marketplace.

Use case D - Young Researcher

1. The use case researcher: field of study, prior knowledge, ground assumptions

Marie is a young researcher who has just finished her studies in cultural studies and now plans to write a doctoral thesis. So far she has rarely worked with digital methods but she is aware that the DH should also be a progressive subject in her field. Therefore she wants to get acquainted with new methods and familiarize herself with the discipline. She is wondering how she can use existing offers in a meaningful way for herself. She has become aware of the Marketplace during her Internet research. She would like to get an overview of the opportunities and possibilities that exist for her.

2. The use case exemplary research project

The subject of the researcher's doctoral thesis has not yet been specified. She would like to work on early comic books, so her data will be both texts and images. During the process of defining her topic of research, she wants to systematise her work, find ways to access data and find out what possibilities she has to analyze and organize her data. An important step for her is to get in contact with the community. She is looking for tools and services, tutorials and instructional videos that she would like to try out in a first playful way.

3. Use of the SSH Open Marketplace

In order to get a general overview of how the site works, and in search of an easy introduction to the subject, Marie initially looks for tutorials and webinars from which she would like to learn. The [training module "Introduction to Digital Humanities"](#) is a useful source of information for her, giving her a good introduction on the DH and an idea of the possibilities of using it. As a guideline on the site, Marie uses the keywords and activities already suggested there. While browsing around and clicking through the different keywords, she finds some interesting tools that she would like to test for her research like tools for the comic images she will be studying ([TILE](#), [Sumo Paint](#), [UVic](#) and [Virtual Lightbox](#)) and for the work with the texts in the comics ([CATMA](#), [Annotation Studio](#), [Quirkos](#) and [Text Analysis Online](#)). For her data management, Marie also finds several possibilities ([LitBlitz](#), [DEVONthink](#)). To get a first contact with the community, she finds the suggested [Twitter](#) hashtag very helpful.



While exploring the MP, Marie is excited to accidentally come across tools she didn't even know existed and could be useful to her like [ScrapBook](#), [BibServer](#) or [eLaborate](#). On the one hand, she is looking forward to trying out all the tools she found. However, on the other hand she feels a little intimidated by the variety of results. For her as a beginner and young researcher, she could use a best-practice guide, instructions or helpful tips for the site. She still has to learn how to search the site properly.

Conclusions

User stories serve as a vivid and usable means to present the concept of a service to its potential users apart from a theoretical sketch. By describing the contact of a researcher with the SSH Open Marketplace it becomes possible to identify gaps or pitfalls in the concept and to align it with the real research practices. However this will likely stay on an exemplary level therefore user stories are not a substitute for an elaborated concept but merely an additional means bringing life into the concept of participatory design.

The vividness of the user stories becomes even more striking when bridging the gap to the prototype of the marketplace. The above described user stories contain real search results that are to be found in the MVP⁹ of the SSH Open Marketplace. Without question still a prototype of future platform but very suitable to provoke specific feedback within the context of the user story such as:

- What else would I expect as a researcher to find in the SSH Open Marketplace?
- Am I satisfied with the way the items are presented or linked to each other?
- Is the SSH Open Marketplace covering “my” main sources, is it fitting to the state of the art of research?

Of course our four initial user stories are only a small sample and by way not representative. Other user stories are possible and we expect them to emerge when discussing the material above and when looking at the SSH Open Marketplace together with users. In this sense, the user stories form only as a spark to initiate - hopefully - heated and productive discussions on the SSH Open Marketplace. Additional user stories and/or use cases (including also some functional requirements for the SSH Open Marketplace) will be added to this list. Some topics are already identified: one focus on Standards and Scenarios in Digital Humanities and Heritage research; one describing the connections with the SSHOC Switchboard (developed within SSHOC WP3). A priority will be given to user stories or use cases including the results of SSHOC specific WPs.

Furthermore, several ideas are also under investigation to improve the quality of these future user stories and use cases:

- alignment with a generic research workflow, e.g. https://nfdi4culture.de/images/us/NFDI4Culture_UserStoryAll.pdf
- inclusion of the local academic and institutional environment. These user stories present researchers disconnected from their institutional environment (no institutional repositories mentioned, no data stewards, archivists or librarians to whom they could ask their question...).
- more detailed explanations of how users discovered the SSH Open Marketplace. This has been shortened here to keep a reasonable length, but it's also an important dimension we should keep in mind if we want to improve these use cases.

⁹ Minimum viable product.