

Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782

Start Date of Project: 01/01/2019

Duration: 40 months

Deliverable 7.4 Marketplace – Data population & curation

Dissemination Level	PU
Due Date of Deliverable	31/12/2021 (M36)
Actual Submission Date	02/12/2021
Work Package	WP7 - Creating the SSH Open Marketplace
Task	Task 7.4 Governance: Population, Curation & Sustainability of the SSH Open Marketplace
Type	Report
Approval Status	Waiting EC approval
Version	V1.0
Number of Pages	p.1 – p.47

Abstract:

This report summarises the state of the Marketplace regarding the data available in it (a snapshot at the end of the project). It describes the sources used to populate the Marketplace and the curation steps taken.

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. This deliverable is licensed under a Creative Commons Attribution 4.0 International License.



History

Version	Date	Reason	Revised by
0.0	30/03/2020	first draft of the table of content	Frank Fischer, Laure Barbot
0.1	13/07/2021	second draft of table of contents	Edward Gray, Nicolas Larrousse
0.2	01/08/2021	Drafting	Edward, Gray, Nicolas, Laure Barbot, Matej Ďurčo, Cesare Concordia, Klaus Illmayer, Alexander Konig, Dieter Van Uytvanck
0.3	22/10/2021	Version ready for Peer Review	Erzsébet Tóth-Czifra (DARIAH) Péter Király (GWDG)
0.4	24/11/2021	Integration of Peer Review Comments	Edward Gray, Nicolas Larrousse
0.5	30/11/2021	Work Package Leader Review	Laure Barbot
1.0	01/12/2021	Final version for submission	

Author List

Organisation	Name	Contact Information
CNRS	Edward Gray	edward.gray@huma-num.fr
CNRS	Nicolas Larrousse	nicolas.larrousse@huma-num.fr
CNRS	Clara Petitfils	clara.petitfils@huma-num.fr
DARIAH	Laure Barbot	laure.barbot@dariah.eu
DARIAH	Frank Fischer	frank.fischer@dariah.eu
DARIAH/OEAW	Matej Ďurčo	matej.durco@oeaw.ac.at
DARIAH/OEAW	Klaus Illmayer	klaus.illmayer@oeaw.ac.at
CNR	Cesare Condordia	cesare.concordia@isti.cnr.it
CLARIN	Alexander Konig	alex@clarin.eu
CLARIN	Dieter Van Uytvanck	dieter@clarin.eu
DARIAH/UGOE	Stefan Buddenbohm	buddenbohm@sub.uni-goettingen.de

Executive Summary

The SSH Open Marketplace is a discovery portal for digital tools, services, and methods in the Social Sciences and Humanities (SSH), with an emphasis placed on placing these resources in their research context. It is the SSH component of the EOSC Marketplace. Previous initiatives have shown that curation of content is the key element to maintain the attractiveness and the utility of a discovery portal like the SSH Open Marketplace.¹ Thus at each stage of the development of the Marketplace, from the first ingestion of existing catalogues to the curation sprint, curation has been at the heart of conception. This emphasis, based on previous examples and the work done throughout the SSHOC project, has led to profound changes in the data model, definition of quality criteria, standardization of the content as the team adapts the machinery of the Marketplace based on lessons learned. This continuous work has laid the foundation of the development of Editorial Guidelines which will be used by future curation actors as well as the definition of their different curation roles and the organisation of their relationship.

Two different methods are being used to populate the Marketplace: integration of a batch of entries, called “mass ingestion”, or individual contributions. Even though general criteria remain similar, the curation processes are different and the tools used are not the same. For instance, to be able to curate a large number of entries, it is necessary to rely on automatized controls. In both cases, a specific module, called “curation dashboard”, has to be developed to monitor and facilitate the curation process.

A key added value of the Marketplace is the possibility of linking different entries, to provide context. For instance, a tool can be linked to a training or a publication that helps researchers understand how they can use that tool to advance their research. The ability to link Marketplace entries allows the creation of special entries called Workflows. The idea is to create a step by step procedure to accomplish a research task, like a transcription of a digitized corpus, and to be able to associate entries from the Marketplace to each step (e.g. tools). As this kind of object has some specificities regarding curation, a section is dedicated to them.

Another important part of the curation process is to maintain over time “semantic artefacts” used to access to Marketplace entries like concepts and more generally vocabularies.

This document describes all these curation components (criteria, tools, organisation of actors, management of vocabularies etc.) developed during the SSHOC project into task 7.4 in close coordination with other tasks from WP7 and beyond (e.g WP3). It represents the basis for the future functioning of the curation of the Marketplace after the end of the project in complementarity to the general governance of the Marketplace expressed in SSHOC D7.5 *Marketplace - Governance*².

¹ Public instance of the SSH Open Marketplace (in beta at the time of writing): <https://marketplace.sshopencloud.eu/> [20.10.2021]

² Clara Petitfils, Suzanne Dumouchel, Nicolas Larrousse, Edward J. Gray, Laure Barbot, Arnaud Roi, Matej Ďurčo, Klaus Illmayer, Stefan Buddenbohm, & Tomasz Parkola. (2021). D7.5 Marketplace - Governance. <https://doi.org/10.5281/zenodo.5608487>

Abbreviations and Acronyms

ACDH-CH	Austrian Centre for Digital Humanities and Cultural Heritage
API	Application Programming Interface
CRUD	Create, Read, Update, Delete
DACE	Data Aggregation and proCessing Engine
DiRT	Digital Research Tools
EOSC	European Open Science Cloud
FAIR	Findable Accessible Interoperable Reusable
GDPR	General Data Protection Regulation
GLAM	Galleries, Libraries, Archives, Museums
PARTHENOS	Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies
PSNC	Poznan Supercomputing and Networking Center
SHE	Social science and Humanities Entity
SKOS	Simple Knowledge Organization System
SSK	Standard Survival Kit
TAPoR	Text Analysis Portal for Research

Table of Contents

Introduction	6
Data Population	8
Mass Ingestion of Existing Catalogues	8
Individual Contributions	9
The Special Case of Workflows	9
Curation	12
Curation Actors	12
User roles	12
Editorial Board	13
Curation Guidelines and Workflows	14
Curation criteria	16
Curation Tooling	17
Automated quality control with curation notebooks	17
Curation dashboard	18
Vocabulary Curation	19
Conclusions	21
References	22
Annex 1 - SSH Open Marketplace - Editorial Guidelines	23

1. Introduction

Context of the deliverable

The SSH Open Marketplace aims to provide information about useful resources for SSH research communities. Good maintenance of the Marketplace's content over time is the key to ensure its success, based on what happened previously to similar initiatives.³ This requires the implementation of different components in order to achieve a good level of curation. This work began in earnest after the first data population from different sources was completed and will continue through the end of the project and throughout the Marketplace's lifetime. This iterative work was done by a "Curation Task Force" including people coming from several WP7 tasks in order to be able to work at different levels, from the modification of the data model to the specifications of the curation dashboard. Other SSHOC Work Packages were also involved in that process, especially WP3.

This deliverable therefore represents a synthesis based on the feedback gained from various tasks accomplished during the construction of the Marketplace.

Other elements were taken into account, in particular the links between the SSH Open Marketplace and the EOSC Marketplace⁴ which is also in development: one of the desired goals is to avoid duplication of work and inconsistencies. The SSH Open Marketplace will be one of the major components of the EOSC for SSH communities and will showcase the vitality of SSH.

Objective of this deliverable

This deliverable summarizes the work done on the different aspects of curation of items for the Marketplace in task 7.4. The spectrum covered by this document is quite wide as it deals with quality criteria but also provides a clear definition of the roles needed at each step of the curation life cycle and the associated tools. It is closely related to the Editorial Guidelines, which can be found in the appendix, that realizes the practical implementation of the curation criteria described in this document.

These different elements will feed the future operating structure dedicated to the content management of the Marketplace which is one of the outcomes of the SSHOC project that will be sustained after its end.

On the other hand, this document is not intended to describe the general governance of the Marketplace itself, which is described in the D7.5 *Marketplace Governance*, but of course these different levels of management have been designed in close cooperation to be compatible. Equally, this document focuses only briefly on the process of mass ingestion of sources, which is described in D7.3 *Marketplace Interoperability* and D7.2 *Marketplace Implementation*⁵, in particular the pipeline built for mass ingestion of existing catalogues based on diverse technologies.

³ Quinn Dombrowski, What Ever Happened to Project Bamboo?, *Literary and Linguistic Computing*, Volume 29, Issue 3, September 2014, Pages 326–339, <https://doi.org/10.1093/llc/fqu026> [19.10.2021] and "Who Needs Directories: A DH 2020 Forum", DH 2020 Conference, <https://dh2020directoriesforum.hcommons.org/> [21.10.2021]

⁴ <https://marketplace.eosc-portal.eu/> [19.10.2021]

⁵ Matej Ďurčo, Laure Barbot, Klaus Illmayer, Sotiris Karampatakis, Frank Fischer, Yoann Moranville, Joshua Tetteh Ocansey, Stefan Probst, Michał Kozak, Stefan Buddenbohm, & Seung-Bin Yim. (2021). 7.2 Marketplace – Implementation (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.5749465>

Comparison with other existing portals: curation

Other similar portals have implemented a curation process.⁶ Even if the needs are the same, there are variations in the implementation of the curation. It is not always easy to understand the process from an external point of view, but it seems that the main differences are in the definition of the roles and their organisation in particular on how individuals submit new content: some require submitters to be a registered user, others use a simple contact by email or a form and some do not give that possibility as they deal only with mass ingestion.

For the Open SSH Marketplace, the choice was to try to be as flexible as possible and offer the possibility for individual contributors to suggest a new single content as well as to evaluate new sources for mass ingestion. Indeed, this is a key part of the SSH Open Marketplace's plan to rely on and be of service to the community of SSH researchers, by allowing them to take ownership of the materials and improve the SSH Open Marketplace based on their knowledge and needs. The different roles and their missions are described in section 3.1.

Another variation discovered in the survey of the previous tool registries is the number of people involved in the curation process, as it can range from zero in the case of a fully automated process to 30 persons in a real editorial team which takes decisions based on more than one review. For the Open SSH Marketplace, the goal is to have a limited number of moderators, between 10 to 20, and use automated controls to support the work of the editorial team as much as possible, but also to ensure contributions from the scientific community thanks to targeted events and an appropriate set of incentives and rewards. Additionally, based on previous experiences it is necessary to ensure that there are people who have dedicated working time towards animating the community of contributors and maintaining the marketplace. Data population and curation are too important of tasks to be cumulated on top of a normal work schedule. These considerations are addressed in D7.5 *Marketplace Governance*.

⁶ The portals considered in our survey of other curation options were TAPoR (<http://tapor.ca/home>), DiRT (formerly <http://dirtdirectory.org/>, now offline), OpenMethods (<https://openmethods.dariah.eu/>), and OpenAire Connect (<https://connect.openaire.eu/>). The DiRT Directory is now offline, though several landing pages can still be found, for instance at the University of California Berkeley. <https://digitalhumanities.berkeley.edu/resources/digital-research-tools-dirt-directory>. DiRT itself grew out of Project Bamboo, whose demise was described in Dombrowski, Dombrowski, "What Ever Happened to Project Bamboo?". All accessed on [19.10.2021]

2. Data Population

The SSH Open Marketplace contains resources that can be broadly declined across five categories: tools & services, training materials, publications, data sets, and workflows. There are two main ways to populate the SSH Open Marketplace with these resources, the first being the mass ingestion of existing sources, and the second being individual suggestions, be they for improving existing items or for wholly new items. While the bulk of the Marketplace's content will come from ingested sources, it is important to allow the community to contribute to the data population of the Marketplace via individual suggestions, as it allows them to engage with their scholarly community and ensures that the SSH Open Marketplace remains relevant to its communities of practice. In addition to these two data population methods, this section also considers the special case of Workflows, as they are a special case of marketplace entry.

2.1. Mass Ingestion of Existing Catalogues

To populate the SSH Open Marketplace a list of possible sources was created which then was filtered to create a shortlist of sources that were mass ingested into the SSH Open Marketplace. Selection of the sources that were ingested was based on a variety of factors, chief among them being the quantity and quality of the items in the source and the ease of ingestion. The latter factor meant that, for example, such sources were preferred that provided an API that could be used to feed them directly into an ingestion pipeline. In case of some very high-quality sources that did not provide an API or another way for automatic mass ingestion it was decided to ingest them manually, but this was only possible for a very limited number of sources that were both of high quality (to be worth the extra effort) and also of a reasonably low quantity (to be possible to manually ingest with the available limited human resources).

The actual ingestion process was done using two different pipelines, the first being PoolParty⁷, an existing tool that was available free of charge for the duration of the project. The second ingestion pipeline was DACE⁸, a tool that was developed by the partners at DARIAH/PSNC in the course of the project. The reason for using two pipelines was a practical one. Using an existing tool meant that the actual ingestion work could start right away (as soon as the internal Marketplace data model was in place), but as this tool would no longer be available (that is, affordable) after the end of the project, it was decided to create a second pipeline that could still be used after the project has ended to continuously reingest some sources to make sure that external updates will also be reflected in the SSH Open Marketplace. The second pipeline was ready towards the beginning of the final year of the project and during this final year all ingestion workflows for the various sources that had already been ingested using the first pipeline were migrated to the second pipeline. For more details on these pipelines, see *D7.2 Marketplace Implementation*.

The actual ingestion of a source was a multi-step process. First the source's metadata was mapped to the SSH Open Marketplace's data model. This mapping was then imported into the ingestion pipeline tool and fine-tuned until the ingestion worked well. In this phase any necessary extra properties within the MP data model could be created as well. Indeed, the heterogeneity of the chosen sources, and

⁷ PoolParty Semantic Suite: <https://www.poolparty.biz/> [19.10.2021]

⁸ Data Aggregation and proCessing Engine: <https://gitlab.pcass.pl/dl-team/aggregation/dace> [20.10.2021]

sometimes even within sources, led to varied completeness for their metadata, and thus mapping proved to be one of the more challenging aspects of ingestion. Then the source was ingested into the development instance of the MP where it then was manually and automatically checked for consistency and any problems with the ingestion mapping that might have been overlooked. This review of the ingestion was then fed back into the ingestion pipeline to create an improved ingestion workflow and the source was ingested again. Further details related to mass ingestion are covered in D7.3 *Marketplace Interoperability*.

2.2. Individual Contributions

Individuals may contribute to the SSH Open Marketplace by suggesting improvements to existing items, or by submitting wholly new entries, once authenticated. This is an important feature of the SSH Open Marketplace, as it allows the research community to engage with Marketplace content, which ensures continued relevance for the SSH research community. These users must be identified via the authentication service in order to do so. Once authenticated, these contributors can then use the edit forms to suggest changes or create new items. Contributors can save drafts of their contributions, and once they are happy with their content, can submit them to approval for moderators. Moderators can then approve or deny the contribution, at which point it will become available to view on the SSH Open Marketplace.

2.3. The Special Case of Workflows

Workflows are a key added-value of the SSH Open Marketplace. They are a direct legacy of a previous project - the PARTHENOS project⁹ - in which the Standardization Survival Kit (SSK)¹⁰ was developed. The SSK showcases research scenarios divided into different steps, guiding the users through best research practices and recommended standards and methods. The 29 scenarios coming from the SSK¹¹ represent the initial data population for what was renamed “workflows” in the SSH Open Marketplace, and future new scenarios will be created directly in the Marketplace, thanks to the user-friendly edit form. For different reasons we are detailing below, a workflow is a very special content type in the Marketplace:

- First of all, it is the only content type that can be **hosted** in the Marketplace. While all the other items in the portal are records pointing to an external object, workflows can be created in the Marketplace, meaning that the object “workflow” itself can be hosted by the Marketplace.
- Workflows are an opportunity to increase the **contextualisation** layer of the SSH Open Marketplace. Indeed, while “basic” context is given by the relations between the items indexed in the registry, workflows are structured recipes, composed of steps, bringing together other items in order to not only increase serendipity in digital methods, but also to provide a companion for the other entities in the Marketplace.

For these reasons, the curation needed to maintain and develop the workflow content type is slightly different from the rest of the Marketplace’s items. Based on the SSK experience - whose “intellectual

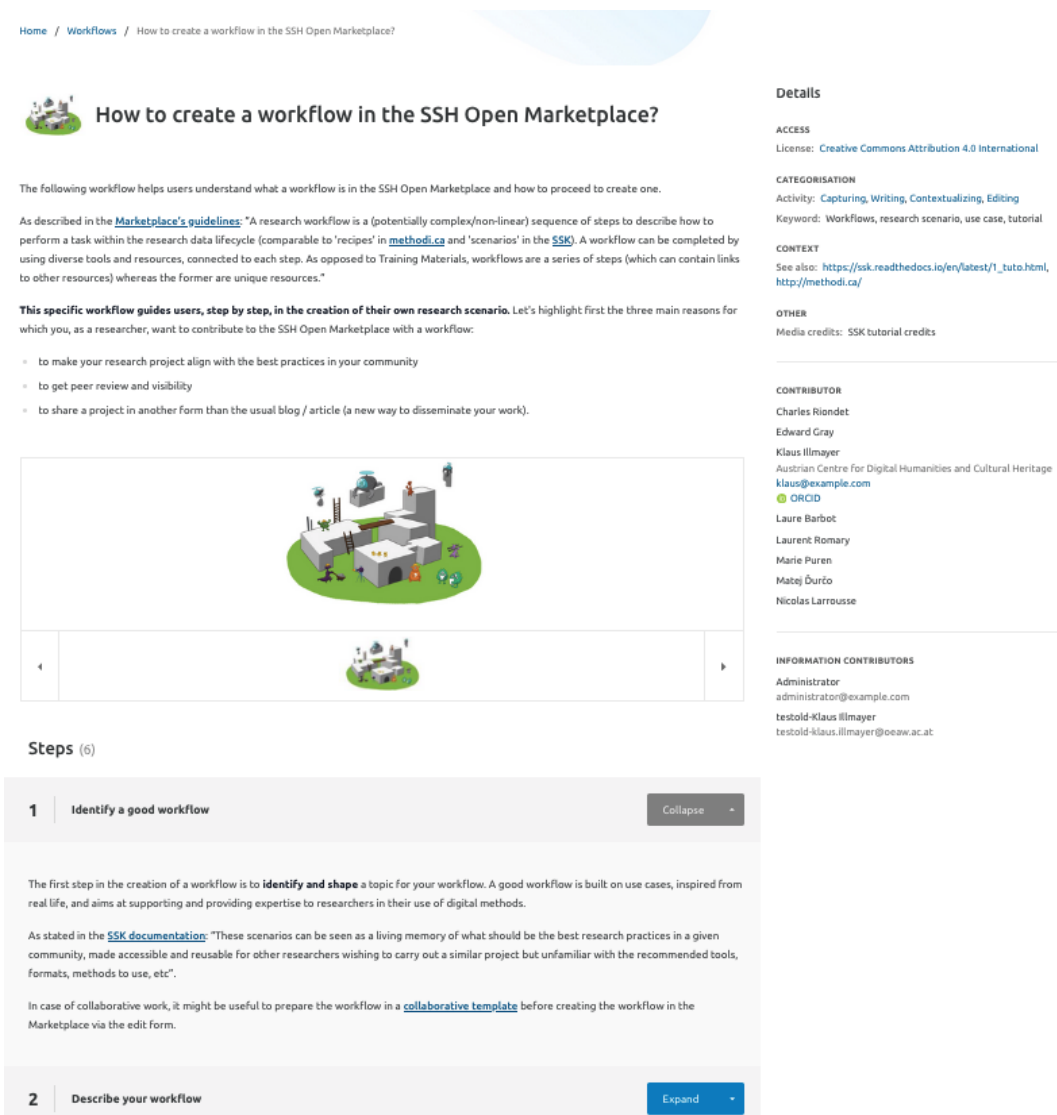
⁹ PARTHENOS project website: <https://www.parthenos-project.eu/> [16.09.2021]

¹⁰ Standardization Survival Kit website: <http://ssk.huma-num.fr/#/> [16.09.2021]

¹¹ SSK scenarios in the SSH Open Marketplace: <https://marketplace.sshopencloud.eu/search?categories=workflow&order=label&f.source=SSK> [02.12.2021]

maintenance” is ensured by the DARIAH Guidelines and Standards Working Group¹² - several key points can be highlighted:

- Creating workflows (or steps composing the workflow) is more efficient when done during hands-on sessions and datathons. That is the reason why workshops and webinars are foreseen as the main drivers for creation of new workflows. That said, because user documentation and support is one of the keys to ensure collaborative curation, a tutorial providing guidance on how to create workflows in the SSH Open Marketplace was also designed.



Home / Workflows / How to create a workflow in the SSH Open Marketplace?

How to create a workflow in the SSH Open Marketplace?

The following workflow helps users understand what a workflow is in the SSH Open Marketplace and how to proceed to create one.

As described in the [Marketplace's guidelines](#): "A research workflow is a (potentially complex/non-linear) sequence of steps to describe how to perform a task within the research data lifecycle (comparable to 'recipes' in [methodi.ca](#) and 'scenarios' in the [SSK](#)). A workflow can be completed by using diverse tools and resources, connected to each step. As opposed to Training Materials, workflows are a series of steps (which can contain links to other resources) whereas the former are unique resources."

This specific workflow guides users, step by step, in the creation of their own research scenario. Let's highlight first the three main reasons for which you, as a researcher, want to contribute to the SSH Open Marketplace with a workflow:

- to make your research project align with the best practices in your community
- to get peer review and visibility
- to share a project in another form than the usual blog / article (a new way to disseminate your work).

Steps (6)

- 1 Identify a good workflow** Collapse

The first step in the creation of a workflow is to **identify and shape** a topic for your workflow. A good workflow is built on use cases, inspired from real life, and aims at supporting and providing expertise to researchers in their use of digital methods.

As stated in the [SSK documentation](#): "These scenarios can be seen as a living memory of what should be the best research practices in a given community, made accessible and reusable for other researchers wishing to carry out a similar project but unfamiliar with the recommended tools, formats, methods to use, etc".

In case of collaborative work, it might be useful to prepare the workflow in a [collaborative template](#) before creating the workflow in the Marketplace via the edit form.
- 2 Describe your workflow** Expand

Details

ACCESS
License: [Creative Commons Attribution 4.0 International](#)

CATEGORISATION
Activity: [Capturing, Writing, Contextualizing, Editing](#)
Keyword: [Workflows](#), [research scenario](#), [use case](#), [tutorial](#)

CONTEXT
See also: https://ssk.readthedocs.io/en/latest/1_tuto.html, <http://methodi.ca/>

OTHER
Media credits: [SSK tutorial credits](#)

CONTRIBUTOR
Charles Riondet
Edward Gray
Klaus Illmayer
Austrian Centre For Digital Humanities and Cultural Heritage
klaus@example.com
 ORCID
Laure Barbot
Laurent Romary
Marie Puren
Matej Durčo
Nicolas Larrousse

INFORMATION CONTRIBUTORS
Administrator
administrator@example.com
testold-klaus.illmayer@oeaw.ac.at

Fig.01: Screenshot of the workflow creation tutorial

¹² DARIAH Guidelines and Standards Working Group
<https://www.dariah.eu/activities/working-groups/guidelines-and-standards/> [19.10.2021]

- Since workflows are based on real research use cases, it can be challenging to find the right balance between generic definition and project specific description. This is why steps' descriptions, references to standards used/recommended and pointers to resources are of utmost importance to ensure reusability of the workflows. An interesting element could be further developed in the future to reinforce pertinence of the workflows. Indeed, when several digital methods are in play with many options or interpretations possible, it could be interesting to show alternative steps and to show different ways of how to do something but also allow the re-use and integration of other steps/workflows.
- Ensuring that workflows will remain up-to-date, especially the more technical workflows where context changes fast is also a challenge that shouldn't be underestimated. Lessons learnt from previous experiences is that the SSH Open Marketplace should try to motivate original authors to further evolve their described workflows, but should also try to bring together experts during the datathon sessions in order to share and discuss the must-have references and recommendations for a given topic.

3. Curation

With the Marketplace populated with the necessary data, the next task is to ensure their proper curation according to a set of standards that ensure the quality that researchers expect. The curation processes and workflows should ensure:

- An effective update of the items, based on their topicality, vividness and frequency;
- The quality and consistency of contents' description - i.e., enrichment - be it relations between records (e.g. workflow) or additional metadata (e.g. illustrations);
- The content's reliability, in terms of integrity and correctness (e.g. licencing information).

This section aims to collect and describe the guidelines (quality criteria and workflows) for curating the content of the marketplace (i.e. the metadata about the entities represented in the marketplace). They should help answer the question: "how does one set about making a useful, rich, and shareable collection description that will facilitate discovery?"¹³. Because most of the curation literature and experiences comes from the GLAM sector, this is where t7.4 team members looked for guidance and best practices, and the next sections reflect the path taken for curating the SSH Open Marketplace collection.

3.1. Curation Actors

A good and frequent curation of the content of the Marketplace is a key element to ensure quality of the content and thus user trust in the service. This section describes the different roles (3.1.1) required to implement the curation workflow in order to reach the desired level of quality. The approach chosen gives a central role to the Editorial Board and allows the integration of community aspects for the marketplace (see 3.1.2).

3.1.1 User roles

The curation workflow of the SSH Open Marketplace currently defines four different user roles with different permissions as described below.¹⁴ These roles are described in the Editorial Guidelines (annex 1) with more details.

Role	Function	Description & Rights
Viewer	Visitor of the MP; without any authentication	Read-only mode for the MP; only anonymised audit trails available; can report issues; can authenticate and then become a contributor

¹³ cf. Collections development principles

(<https://archive.mith.umd.edu/dhcurator-guide/guide.dhcurator.org/contents/digital-collections-and-aggregation-s/>) [19.10.2021]

¹⁴ Note that in the system five user roles are implemented: Contributor, Moderator, Administrator, System Contributor and System Moderator. The last two ones are used to perform mass ingest and mass approval operations. A viewer is not registered in the system, even though statistics about page views are collected and can give an idea of the number of visitors of the website (regardless of their user role).

Contributor	Suggests new content or updates to the MP; authentication on a login basis	Main target audience of the MP; active member of a research community; can create or update individual items; contributors names recorded and displayed; authenticated viewer.
Moderator	Keeps track of editing, ingesting and updating the MP's content; responsible for the management of curation workflow components; community engagement	All rights for adding new items or sources and for moderating contributors' suggestions; mapping and reviewing new sources ingestion; taking care of the curation workflow and the curation criteria; updating all necessary components in this regard such as vocabulary management; in charge of the community management. Member of the Editorial Team.
Administrator	Maintenance of the MP; user management; ingesting of sources	Close collaboration with the moderators; processing of ingesting new sources; handling of the automated processes such as automated curation enrichment; user management. Member of the Editorial Team.

Table 01: User roles overview

3.1.2 Editorial Board

The Editorial Board is the coordinating body ensuring the proper execution of the responsibilities and policy related to all curation and editorial aspects. These aspects of governance are covered in D7.5 *Marketplace Governance*, though the relevant points for curation are highlighted here. The Editorial Board is acting under the supervision of the Governing Board that is in charge of the strategic decisions on editorial orientations and policy.¹⁵ The Editorial Board is consulted by the Governing Board for any relevant decisions and strategic orientations on scientific, technical and managerial matters.

Moderators and Administrators are members of the Editorial Board. They are editorial and/or technical experts appointed by the Governing Board of the SSH Open Marketplace. The Editorial Board is led by a Coordinator, selected by the governing board, whose mission is to coordinate the moderators and administrators where necessary and to lead the discussion in the editorial board. Depending on the Governing Board decisions, members of the Editorial Board are foreseen to be chosen either among the service providers' workforce or their network. The idea being that Moderators and Administrators involvement could be officially declared as in-kind contributions by their employer.

The Editorial Board is the body dedicated to the daily running of operation through:

- **technical maintenance.** Administrators are in charge of the technical infrastructure of the system.
- **editorial and curation-related tasks.** Moderators follow the Editorial Guidelines (annex 1) to ensure the Marketplace data quality (see Annex - Editorial Guidelines).
- **community engagement.** Moderators maintain a close and strong link with the Contributors. As described in D7.5 *Marketplace Governance*, the Editorial Board organises and animates

¹⁵ See D7.5 *Marketplace Governance* for more details about the relations between the three entities identified for the service management

engagement activities, relating to the curation and content of the Marketplace. Incentives and rewards are offered to support the work of contributors and to limit voluntary work as much as possible.

3.2 Curation Guidelines and Workflows

Now that roles are defined, the next step is to establish guidelines and workflows for the curation of the SSH Open Marketplace. Several conceptual distinctions presented in this section have led to the elaboration of the Editorial Guidelines presented in Annex 1.

First of all, a distinction between criteria/guidelines needed to ensure the two different data population avenues need to be introduced:

- a) **guidelines for sources** - i.e. does an existing catalogue generally fit into the scope of the SSH Open Marketplace, so that it should be considered as a potential source.
- b) **guidelines for items** - i.e. is information about an item (from a source, or suggested manually) sufficient to be presented in the marketplace. These can be found in the *Editorial Guidelines*.

Furthermore, in applying the guidelines (for items), we need to distinguish between automatic and manual actions, i.e. which criteria can be operationalized into some algorithmic check, and which need to be checked manually. In general, the SSH Open Marketplace relies on a hybrid semi-automatic approach, where automatic checks are run on the ingested data, followed by manual review of the identified problems as well as of aspects that cannot be checked automatically. The results and overview of these checks are presented to the moderator via an editorial dashboard.

Finally, one can distinguish between the criteria that are set as **hard constraints**, like mandatory fields implemented in the system and **curation criteria** that decide, if an item is displayed to the users or not, but still can exist (hidden) in the system. In general, very little criteria are set as hard constraints. This is to allow any potential “dirty data” to come into the system and then be improved, so that this data can be presented to users. This implies a three-fold distinction¹⁶:

1. **deprecated content** - content that do not fulfill the hard constraints and thus is deprecated from the system - this type of content can be seen only by the authorised accounts and administrators are allowed to revert the status of such items.
2. **to-be-reviewed/curated items** - content that has passed the ingestion or has been suggested by an individual user, but is not reviewed yet
3. **visible/published content** - content considered good/rich enough to be made visible in the Marketplace

Based on these conceptual distinctions, the following diagram represents an overview of **how items suggested by a contributor** are processed in the SSH Open Marketplace.

¹⁶ For the detailed list of possible statuses for items (and for item versions), see the Entry Lifecycle section of the Editorial Guidelines.

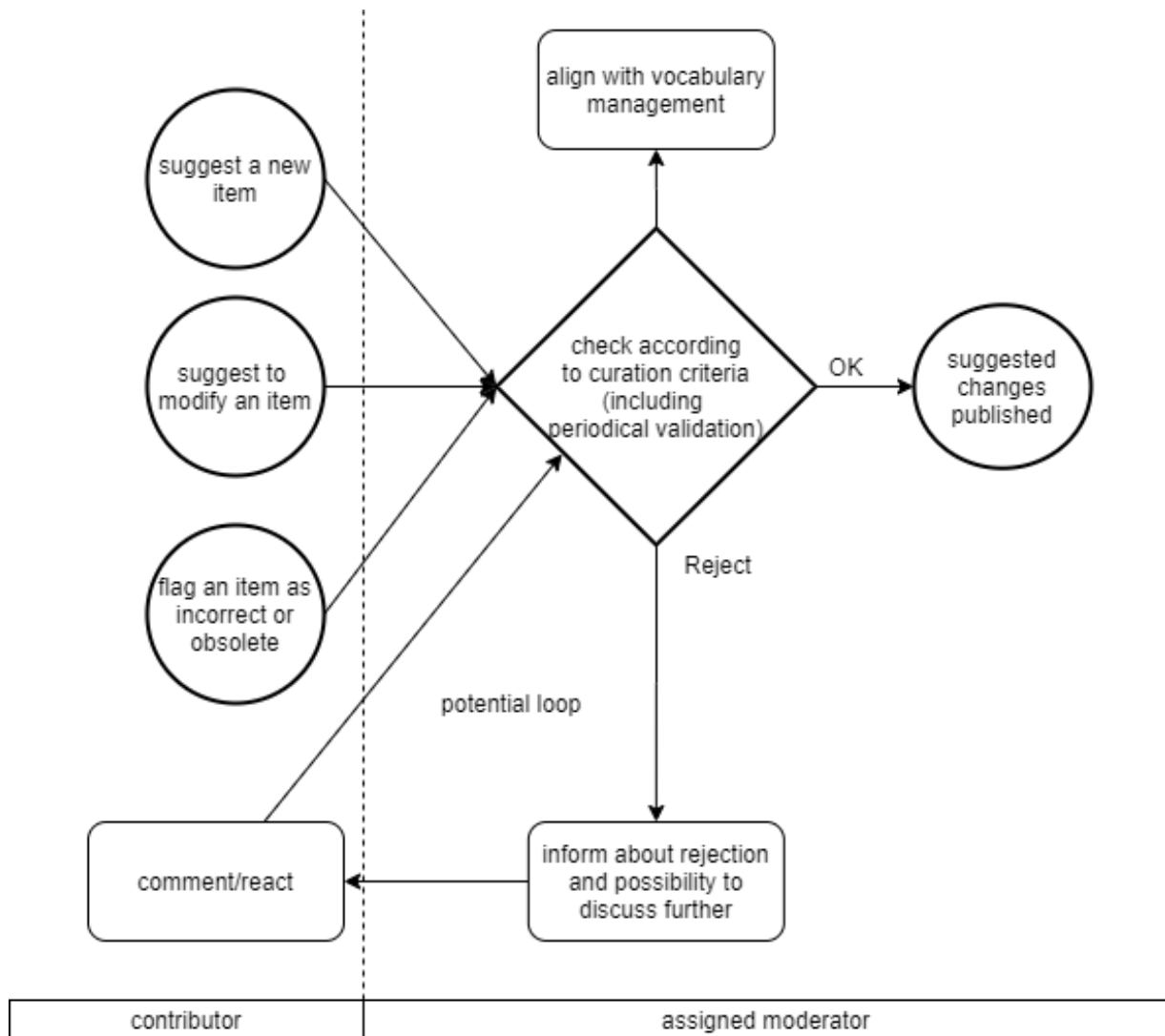


Fig. 02: Processing an item in the SSH Open Marketplace: overview

It should be noted that the Editorial Guidelines are intended to be a sort of “user manual” for the curation of the SSH Open Marketplace, and as such are mainly focused on the curation workflow initiated by individual contributions. These guidelines are to be considered dynamic and evolving over time, informed by the experience the team gathers in the curation process. The initial set of criteria is based both on the end-user requirements as they were defined in the D7.1 *System Specification of the SSH Open Marketplace*. They were further refined via different consultations (see MS43 report for example), and via the mappings exercises between MP data models and sources’ data models that led to new iteration of the MP data model, introducing new properties for each curation criteria needed to be determined. A first stable version of this criteria and guidelines is provided with Annex 1 of the present deliverable, but further developments based on the continuous data ingestions, consultations, and real use of the service are expected. As explained in D7.5 *Marketplace Governance*, the Editorial Board is in charge of updating the Guidelines, but any major changes affecting the core parameters of the service should also be submitted to the Governing Board for approval.

3.3 Curation criteria

Three different but complementary levels/layers of curation criteria are intertwined in the SSH Open Marketplace to ensure data quality and trust.

1. The **generic entry requirements** are described and explained in detail in the Editorial Guidelines and only an overview is added below:

Generic entry requirements	Overview of what is covered
Scientific requirements	<ul style="list-style-type: none"> - Relevance for research and researchers/scholars - Pertinence to the digital methodologies used within the SSH landscape
Technical requirements	<ul style="list-style-type: none"> - Current and supported - Open
Contextualization Requirement	<ul style="list-style-type: none"> - Datasets and publications will only be presented IF they can be linked to another item in the Marketplace
Legal/Ethical requirements	<ul style="list-style-type: none"> - FAIR compliance - GDPR compliance

Table 02: Generic Entry Requirements

2. The **quality criteria for sources** are an indirect contribution to the curation criteria in the sense that:
 - a. The choice/prioritisation of Marketplace's source contributes to the final data quality and can also be seen as a guarantee of trust. These criteria are detailed in D7.3 *Marketplace Interoperability*.
 - b. The review of sources ingestion is a "proto-curation" exercise, allowing the identification of (meta)data weaknesses for future curation in the Marketplace
3. The **quality criteria for items** are described and explained in detail in the Editorial Guidelines. In addition to the generic entry requirements detailed above, a condensed version of the quality criteria is added below:
 - Non-redundancy
 - Completeness
 - Verification of conformity and relevance of metadata
 - Interlinking

Based on these qualitative criteria for items and their definition, some quantitative requirements have been extracted (e.g. the number of characters in description) and served as basis for the automatic checks in the notebooks, which are described in the next chapter.

3.4 Curation Tooling

The curation procedures are supported by a few technical components:

- **Item Edit Form** - This form guides users through the Marketplace Item submission or editing process. It pre-populates the mandatory and required fields, and provides short descriptions of what these fields require, in order to aid the users in the submission of edits or wholly new Item entries.
- **Curation Notebooks** - These are Python notebooks that serve as the basis of the automatic checks of the SSH Open Marketplace by searching for errors. For instance, these notebooks can check to see if the `AccessibleAt` property, which provides a link to the Marketplace Entry, is a functioning link. Once a curation notebook finds an error, it is signalled with a curation property (visible only to moderators and administrators), which results are then fed through the Curation Dashboard.
- **Curation Dashboard** - This is an interface to allow moderators and administrators to address the curation flags drawn up by the Curation Notebooks. It allows for moderators and administrators to select the type of error they wish to process and work more efficiently.

For more details on the implementation aspects see also D7.2 *Marketplace Implementation*, section Curation components.

3.4.1 Automated quality control with curation notebooks

A series of Jupyter Notebooks have been designed to run automatic checks and inform the Marketplace moderators of problematic items.¹⁷

A Jupyter Notebook is a free, open-source, interactive web tool known as a computational notebook, which researchers can use to combine software code, computational output, explanatory text and multimedia resources in a single document. The notebooks developed in this task are used to explore the Marketplace data in order to show relevant information about it and to perform analysis to check the quality requirements/criteria defined in the SSH Open Marketplace Editorial Guidelines. Two kinds of Notebooks have been designed and developed:

- **analysis** Notebooks that download the Marketplace's descriptive metadata, apply to that metadata a number of analysis procedures, and show results (descriptive statistics) to moderators. Given the nature of the Notebooks, these scripts can be easily extended to refine analysis and obtain more details on demand.

- **check** Notebooks that download the Marketplace data, perform the data quality controls defined in the Editorial Guidelines, and write back to the Marketplace curation information about items not respecting quality criteria. The curation information will be shown in the Curation dashboard and will be used by moderators/curators to fix data quality problems.

¹⁷ The curation notebooks can be found at the following link on the SSH Open Marketplace Gitlab. However, it should be noted that as of the due date of this deliverable, the curation section of the Gitlab is not yet public, even though the plan is for it to be public by the end of the SSHOC Project. <https://gitlab.gwdg.de/sshoc/curation>

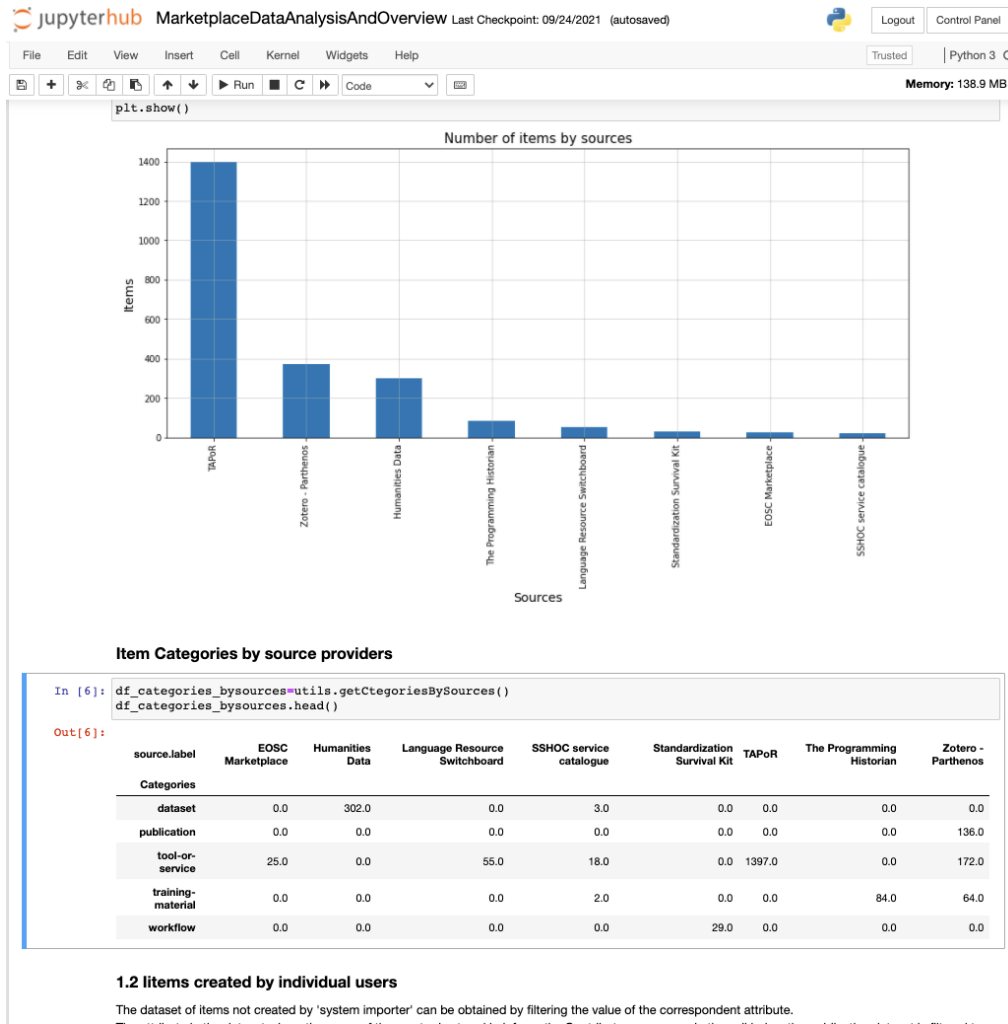


Fig. 03: An example analysis notebook

The interactions of the Notebooks with the system are based exclusively on the Marketplace API, much in the same way as it is used by the frontend component. The programming language used for the development of the curation Notebooks is Python.

3.4.2 Curation dashboard

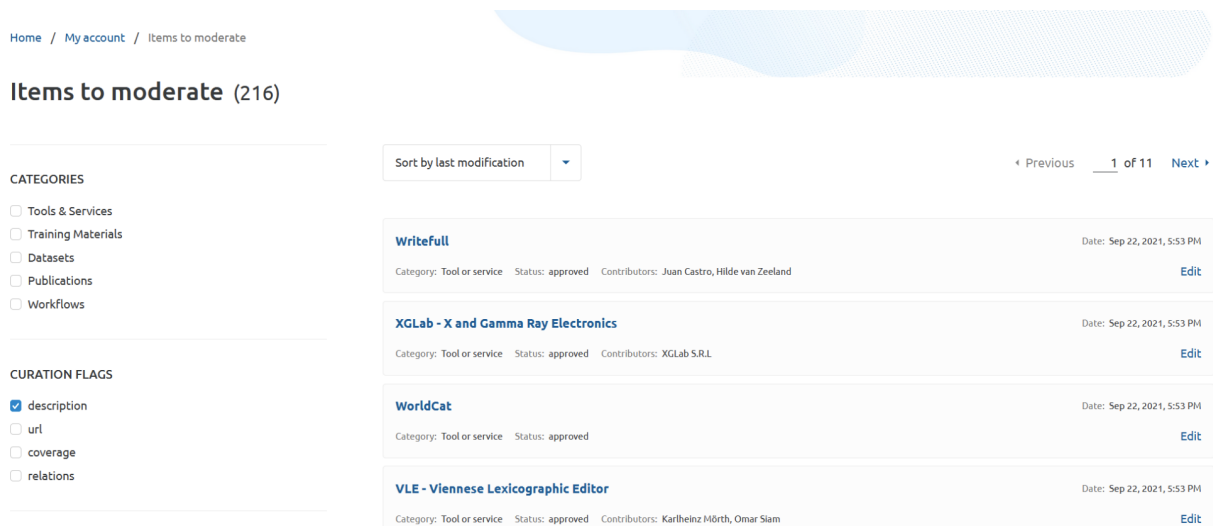
The curation dashboard is integrated as a module of the frontend component and is available to users with elevated permissions (moderators), providing an overview of the status of the items in the Marketplace with focus on the quality of the information and the ensuing curation tasks.

It allows moderators and administrators to view and filter the items in the marketplace based on the outcomes of the curation checks performed by the Jupyter Notebook scripts. These checks identify problems and they set a flag for further curation. Curators are then able to tackle the open issues, e.g.

looking for a better description. Additionally, the curation dashboard also allows moderators to approve items that are suggested by contributors.

In summary, the key functionalities of the curation dashboard are the following:

- integration into the marketplace frontend
- permission control so that only moderators see the open issues
- connection to the results of the automated controls (curation flags)
- integration of approval workflow for suggested items
- visual solution to filter curation issues based on curation flags delivered by the automated curation controls
- easy ways to edit the concerned items



Home / My account / Items to moderate

Items to moderate (216)

Sort by last modification

← Previous 1 of 11 Next →

CATEGORIES

- Tools & Services
- Training Materials
- Datasets
- Publications
- Workflows

CURATION FLAGS

- description
- url
- coverage
- relations

Writefull	Date: Sep 22, 2021, 5:53 PM
Category: Tool or service Status: approved Contributors: Juan Castro, Hilde van Zeeland	Edit
XGLab - X and Gamma Ray Electronics	Date: Sep 22, 2021, 5:53 PM
Category: Tool or service Status: approved Contributors: XGLab S.R.L	Edit
WorldCat	Date: Sep 22, 2021, 5:53 PM
Category: Tool or service Status: approved	Edit
VLE - Viennese Lexicographic Editor	Date: Sep 22, 2021, 5:53 PM
Category: Tool or service Status: approved Contributors: Karlheinz Mörth, Omar Siam	Edit

Fig. 04: Example of Curation Dashboard showing items to moderate due to missing description text

3.5 Vocabulary Curation

As described in D7.2 *Marketplace Implementation*, section 3 “Data Model”, concept-based properties and controlled vocabularies are central elements of the Marketplace design and interoperability. For many of the properties describing entities in the SSH Open Marketplace dedicated controlled vocabularies are employed to restrict the value ranges. These vocabularies are recruited from various sources: whether they be established, external vocabularies which can be reused, or purpose-built vocabularies tailored to the data in the Marketplace. A complete list of these vocabularies can be found in D7.2 *Marketplace - Implementation*. Therefore, as opposed to the original plan, there is no tight integration with an external vocabulary management component. Rather the primary storage for the vocabularies is the Marketplace-core component and the primary interoperability mechanism is the Marketplace-core component capable of importing and exporting vocabularies in SKOS format. At the same time, the Marketplace-core API supports basic CRUD-operations on vocabularies and concepts. These two together form the technical basis for the vocabulary curation. Simple operations like adding or removing a concept can be performed via the API, and/or a simple user interface in the Curation Dashboard of the front-end component. In case of major curation effort for a vocabulary, it can be

exported in SKOS format, and processed either in a separate vocabulary management tool, or through batch processing, or even in a simple text editor. A new version of the vocabulary can be reimported to Marketplace-core in SKOS-format via the API.

Main challenge in vocabulary curation, is the dynamic evolution of the vocabularies, i.e. including new terms. New concepts can be introduced by contributors/moderators in need for a new term when editing a Marketplace item, but most often new terms are encountered during the automated ingest of data from existing sources. It should be up to the moderators to decide, if the new term indeed represents a new concept, or if it is just a different lexical representation (alternative label) for an existing concept, or else if the term should not become part of vocabulary for a given field at all. To allow oversight over newly introduced concepts, these are marked as “candidate” concepts, and are made available to the moderators in the curation dashboard or via the python notebooks.

4. Conclusions

Over the course of the SSH Open Marketplace's development, data population and curation took shape guided by the goal of long-term data quality and accessibility. As (meta)data catalogue curation can be challenging, the T7.4 team built on previous experiences in the Digital Humanities, current trends in the EOSC context, as well as on input from other SSHOC WPs - especially WP3 - and other WP7 tasks to come up with the framework presented in this deliverable.

Based on the important difference between a data population performed by mass ingestion and individual contributions, the SSH Open Marketplace has been developed to ensure robust mass ingestion pipelines, as described in D7.2 *Marketplace Implementation*, and to provide a user-friendly and well documented interface for individual contributions and enrichment highlighted in the present document. For similar reasons, a distinction was made for quality criteria which include differences as there are different needs to evaluate a single contribution compared to an existing catalogue.

To minimize the human curation effort, the goal was to support curation through automatic data ingestion reviews and data analysis. The Python Jupyter Notebooks checks developed answer this need and have been set up following quality criteria and their related quantitative thresholds. The flexibility of this approach is also a guarantee for the long-term sustainability of the data quality tasks.

A central resource to ensure consistent curation procedure are the editorial guidelines, annexed to this deliverable. They will become a main building block of the user documentation on the SSH Open Marketplace website.

Throughout the data ingestion phase, T7.4 team ensured that the point of view of the end-users (future contributors and moderators) was taken into account by ensuring continuous review of the data model changes and translating them in the appropriate sections of the user documentation and interfaces. Additionally, a curation sprint was organized¹⁸ in order to refine the process with a selection of real end-users. These experiences informed the team's approach to establishing these guidelines, so they could match the needs of users.

Based on this strong curation framework, and alongside the governance model developed in D7.5 *Marketplace Governance*, the steps ahead for T7.4 until the end of the project and in order to prepare the sustainability of the Marketplace and its content are clear. Chief among these tasks is to set up an Editorial Team, ideally composed of 15/20 persons with a complementary set of skills and ability to take over the moderation and administration of the SSH Open Marketplace. Moreover, curation workflows and components will also be refined, by employing live tests and hands-on sessions with end-users. During this time, new potential sources will be evaluated for a future integration into the Marketplace.

¹⁸ Curation sprint article, SSHOC Website: <https://sshopencloud.eu/news/curation-sprint-ssh-open-marketplace> (Nov 2021)

5. References

Quinn Dombrowski, What Ever Happened to Project Bamboo?, *Literary and Linguistic Computing*, Volume 29, Issue 3, September 2014, Pages 326–339, <https://doi.org/10.1093/lc/fqu026>

Clara Petitfils, Suzanne Dumouchel, Nicolas Larrousse, Edward J. Gray, Laure Barbot, Arnaud Roi, Matej Ďurčo, Klaus Illmayer, Stefan Buddenbohm, & Tomasz Parkola. (2021). D7.5 Marketplace - Governance. <https://doi.org/10.5281/zenodo.5608487>

Matej Ďurčo, Laure Barbot, Klaus Illmayer, Sotiris Karampatakis, Frank Fischer, Yoann Moranville, Joshua Tetteh Ocansey, Stefan Probst, Michał Kozak, Stefan Buddenbohm, & Seung-Bin Yim. (2021). 7.2 Marketplace – Implementation (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.5749465>

List of Figures

[Figure 1: Screenshot of the workflow creation tutorial](#)

[Figure 2: Processing an item in the SSH Open Marketplace: overview](#)

[Figure 3: An example analysis notebook](#)

[Figure 4: Example of Curation Dashboard showing items to moderate due to missing description text](#)

List of Tables

[Table 1: User Roles Overview](#)

[Table 2: Generic Entry Requirements](#)

Annex 1 - SSH Open Marketplace - Editorial Guidelines

This document aims at gathering concrete and practical information to help users - mainly contributors and moderators, both curators of the Marketplace - of the SSH Open Marketplace (MP) to correct and refine the metadata of MP items (or even add completely new items). These editorial guidelines of the SSH Open Marketplace are added as an annex to the D7.4 *Marketplace – Data population & curation* deliverable, to be published in November 2021, and to the SSH Open Marketplace website, in “Contribute” pages, as user documentation.

Structure of the Editorial Guidelines

I. Design Principles, Entry Requirements, and Quality Criteria for SSH Open Marketplace Entries 25

What the SSH Open Marketplace is and what it isn't	25
Item Types in the SSH Open Marketplace: Definitions	26
Tools & Services	26
Training Materials	26
Publications	26
Datasets	27
Workflows	27
Quality Criteria for SSH Open Marketplace Items	27

II. Roles & Missions in the SSH Open Marketplace 29

Roles in the SSH Open Marketplace	29
Contributing to the SSH Open Marketplace	30
The SSH Open Marketplace Entry Lifecycle	30
Creation of an individual item	31
Mass ingestion from (new) trusted third party sources	33
Metadata status	33

Moderating the SSH Open Marketplace	34
Priority List of Editorial Actions	34
Vocabulary management	35
Managing Actors	35
Mass validation	35
III. Guidelines on individual fields	36
Guidance for metadata fields	39
Guidance on curation properties/metadata	46

I. Design Principles, Entry Requirements, and Quality Criteria for SSH Open Marketplace Entries

What the SSH Open Marketplace is and what it isn't

Above all, the SSH Open Marketplace serves researchers in the social sciences and the humanities, therefore, there must be an established link between the content added into the Marketplace and the SSH. As well, the SSH Open Marketplace is meant to provide a discovery portal for tools and services, while placing these tools and services in context via publications, training materials, datasets, and workflows. As such, these last four categories are indexed in the SSH Open Marketplace insofar as they can be placed in relation with tools and services, and if no relation to a tool or service exists, they should not be accepted.

Based on this scope, the following **Generic Entry Requirements** can be highlighted:

1. **Scientific requirements**

The broad scope of the SSH Open Marketplace means that to be selected, any resource must fulfill at least two criteria:

- Relevance for SSH research and researchers
- Pertinence to the digital methodologies used within the SSH landscape

Selection decisions are made by Moderators. Particularly for resources such as technical services or software, the SSH link could be very loose. In such cases the relevance may be assessed in terms of the uptake or prominence of a resource within the SSH community.

2. **Technical requirements**

To be published two technical curation criteria should be met by any resource:

- Current and supported
- Open

The SSH Open Marketplace favours the uptake of Open Science workflows and open research practices. Software resources are preferably built upon open source solutions.

Note: Given that the SSH Open Marketplace seeks to mirror actual research practices, commercial or non-current resources are also referenced where these are relevant for the scientific community.

3. **Contextualization Requirement**

The SSH Open Marketplace is a discovery portal for researchers to find tools, services, and methods that help complete their research. Thus, Tools and Services are placed in context with relevant Training Materials, Publications, Datasets, and Workflows. Workflows are central items of the SSH Open Marketplace.

- Contextualization is at the heart of the Marketplace's design, as it permits researchers to discover materials related to their search. Therefore, datasets and publications will only be presented IF they can be linked to another item in the Marketplace.

4. **Legal/Ethical requirements**

By design, the SSH Open Marketplace maximises the findability and re-use of data in line with the **FAIR data principles** – Findable, Accessible, Interoperable and Re-usable – for research data, and guides users towards tools, services or training materials that can help them in their FAIRification of workflows.

Of necessity, the SSH Open Marketplace is GDPR compliant. This is reflected in the manner of presentation of the resources and – more obviously – in the management of users.

Item Types in the SSH Open Marketplace: Definitions

Five main [types of items](#) are used to classify the resources of the SSH Open Marketplace. When creating a new entry in the SSH Open Marketplace, a type needs to be chosen first.

Concerning Entry Requirements for each type of Marketplace Entry, the general criteria outlined above apply to each subsection.

1. Tools & Services

Definition: Tools & services are resources used to perform digital activities. Under this type of item, SSH Open Marketplace users will find desktop client solutions (to be installed locally), browser-based or command-line based resources, mobile apps, programming libraries or APIs, but also data catalogues.

See also Social Science and Humanities Entity (SHE) Service and SHE Tool classes definitions in the [SSHOC Reference Ontology](#)

Example: Gephi is a visualisation and exploration software for all kinds of graphs and networks.

2. Training Materials

Definition: Tutorials, lessons or didactic resources explaining how to perform an action or highlighting the learning outcomes one would gain from engaging with the material. As opposed to workflows, training materials are a unique resource or ensemble of related resources whereas the former are a series of steps.

Example: The lesson “Beginner's Guide to Twitter Data” coming from the Programming Historian website.

3. Publications

Definition: Pieces of scholarly outputs, published as books, articles in an academic journals or proceedings, or pre-prints. In the SSH Open Marketplace, only publications that can be connected to other resources in the Marketplace will be shown. We don't aim at presenting all the research papers in

SSH (which is the goal of other services like the OpenAIRE Explore discovery system), but only to highlight the contextualisation of our resources. For example, the SSH Open Marketplace could demonstrate how a tool has been used and cited in a paper.

Example: “PoetryLab. An Open Source Toolkit for the Analysis of Spanish Poetry Corpora” is a Conference Paper presented during the DH2020 Conference in which you can find an example of use of the SpaCy library (referenced as a tool in the Marketplace)

4. Datasets

Definition: Dataset is one digital object or a set or collection of data, records or information that is kept as a persistent unit of information in the knowledge generation process. The Marketplace does not compete with genuine data repositories or data discovery aggregators like OpenAire Explore, but aims at holding only a few hundred of the best-practice research datasets and contextualising them with other Marketplace items, mainly for illustration.

See also *SHE Dataset Class* in the [SSHOC Reference Ontology](#)

Example: The 88milSMS corpus contains more than 88,000 authentic French text messages received from the public during the “sud4science LR” project for sociolinguistic purposes (See http://88milsms.huma-num.fr/corpus_en.html)

5. Workflows

Definition: A research workflow is a (potentially complex/non-linear) sequence of steps to describe a digital research method (comparable to 'recipes' in [methodi.ca](#) and 'scenarios' in the [SSK](#)). A workflow can be completed by using diverse tools and resources, connected to each step. As opposed to Training Materials, workflows are a series of steps (which can contain links to other resources) whereas the former are unique resources. Unlike other entries, Workflows can be hosted by the SSH Open Marketplace, and do not necessarily need to link to an external resource.

Example: “Extract textual content from images” is a workflow composed of 13 consecutive steps coming from the Standardization Survival Kit.

Quality Criteria for SSH Open Marketplace Items

Because findability and discoverability can only be ensured with good (meta)data quality of entries for the SSH Open Marketplace, quality criteria are of utmost importance. Different levels of criteria used to improve the SSH Open Marketplace content are presented/summarised below:

1. **General Entry Requirements** described above: research requirements, technical aspects, contextualisation and legal/ethical requirements are considered here.
2. **Non-redundancy:** the same entity should only be referenced once in the SSH Open Marketplace. Duplicate items are merged to ensure the coherence of the items showcased in the portal.

3. **Completeness** of item description. This criterion concerns the number of metadata filled out to describe an item and answer to the question of how an entity is described comprehensively. Even if we have not yet established a minimum viable number of metadata, contributors should strive for completeness.

4. **Verification of conformity and relevance** of metadata. Information value checks depend on the type of content expected.
 - if a **verbose description** is expected, quality criteria will revolve around correctness, understandability and concision (i.e. length),
 - while for **URLs based fields**, the quality criteria relate to the link accessibility.
 - In order to support discoverability of the Marketplace content, a lot of fields are also (re)using controlled vocabularies¹⁹, thus the validation checks of **vocabulary concepts** is also supported by the Marketplace.
 - Quality of the **media objects** is yet another dimension, e.g. ensuring a good picture quality.

5. **Interlinking**. To ensure serendipity and reinforce the browsing experience of the Marketplace, the quality of the links between items is an important criterion. The number of related items for a given entry, and the pertinence of the links are some of the checks performed to ensure that this criterion is met (e.g. whether the link goes to a resource internal or external to the Marketplace).

These criteria for data quality in the SSH Open Marketplace can only be met by mixing a set of methods - automatic checks and manual actions - and the following section “editorial tasks” explains the methodology used, the technological choices made and the associated tasks supporting the data quality improvement of the SSH Open Marketplace.

¹⁹ Add reference to definition/paper explaining why using controlled vocabularies matter

II. Roles & Missions in the SSH Open Marketplace

Roles in the SSH Open Marketplace

There are four different user roles for people involved in the SSH Open Marketplace.

Viewer: A viewer (= an unregistered visitor) can access the public part of the Marketplace anonymously (no login or barrier) in a read-only mode. This role will represent the majority of all visitors. If a viewer wants to suggest new information or changes to existing information he or she needs to authenticate thus becoming a contributor, or can use the contact form.

Contributor: The contributor is a user suggesting new information or changes to existing items. Users are identified via the EOSC AAI, which can be accessed with a series of pre-existing accounts, such as DARIAH AAI, EGI, OpenAire, Google, and many more. A contributor can perform the following actions:

- Suggest an individual new item
- Suggest to add or change information of an existing item; including to add a relation between two items
- Flag an item as incorrect or obsolete to alert the moderator

Moderator: The moderator has all rights with regard to editing the content and filtering the "noisy input" from the contributors or from harvesting third party sources. Moderators will handle the bulk of content related work of the marketplace.

A moderator has to fulfil the following tasks:

- Accept/Dismiss information:
 - Accept/Dismiss the ingestion of a source (to be decided by the team of moderators)
 - Accept/Dismiss the addition of an item suggested by a contributor
 - Accept/Dismiss suggestions on individual fields within an item
- Define mappings for a new source followed by test ingest, reviewing the results of the test ingest and potentially refining the mapping, until the test ingest is approved (by team of moderators)
- Invoke production ingest- this Includes/implies automatic checks and automatic curation
- Check ingest of a new source according to curation criteria
- Merge information. When two items are not recognised as duplicate in the ingestion phase, a moderator has the possibility to merge two (or more) items.

Administrator: The admin can perform all the actions of a moderator, plus the following:

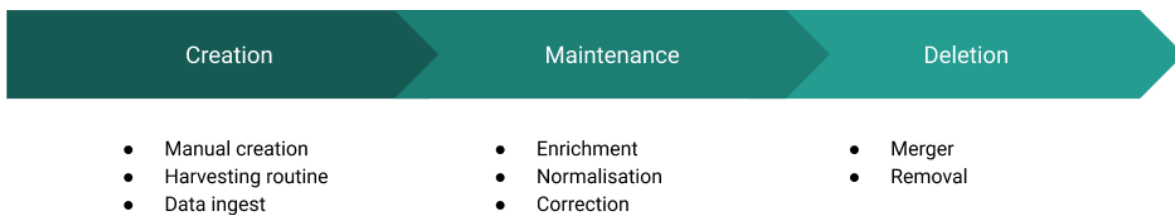
- Processing of ingesting new sources together with the moderators

- User management: deactivating contributors and moderators, assigning the elevated moderator role

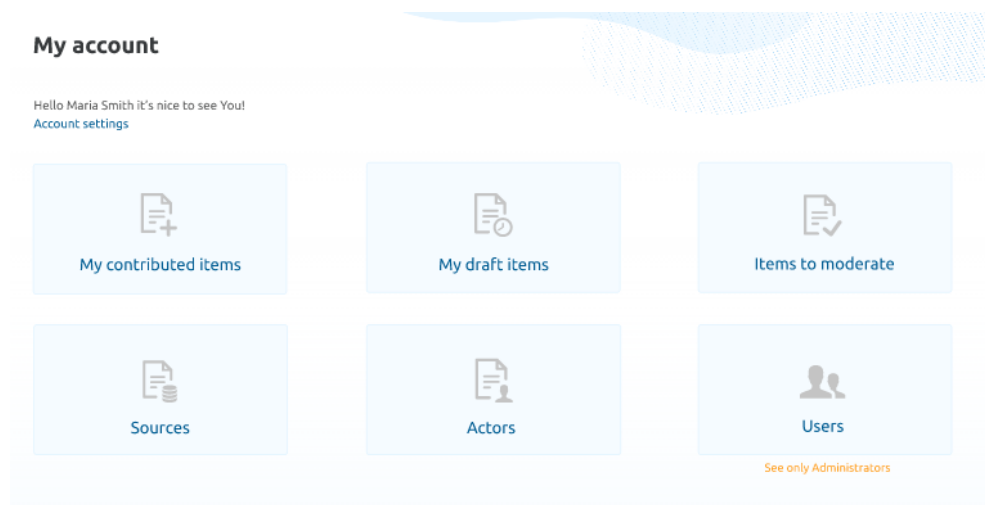
Contributing to the SSH Open Marketplace

The SSH Open Marketplace Entry Lifecycle

The Entry Lifecycle consists of a series of actions taken from the submission of a new item until the final publication of a new item.



These editorial actions will be completed through the Editorial Dashboard.



Each entry into the marketplace will have a status as it goes through its lifecycle in the SSH Open Marketplace. The possible statuses are listed below:

- DRAFT - instead of immediately publishing an item on the Marketplace users can create a draft version of an item which indicates some information is missing that needs further research. A draft item is accessible only for the user who creates it and cannot be shared with other users. The user can add further information and publish the item when it is ready, thus setting the new state "SUGGESTED" (if the user is a contributor) or "APPROVED" (if the user is a moderator/administrator).
- SUGGESTED - items created or enriched/updated versions of an item by contributors get this status. These items/item versions are submitted to Moderators/Administrators for approval/rejection.

- **INGESTED** - items coming from an ingestion pipeline get this status. This allows moderators/administrators to differentiate them from items that are “SUGGESTED” by contributors. These ingested items are passed on to the Moderators/Administrators for approval/rejection.
- **APPROVED** - items or item versions with the status “INGESTED” or “SUGGESTED” receive the status “APPROVED” when such an item/item version has been reviewed and accepted by a Moderator/Administrator. If Moderators/Administrators create or update an item it automatically gets the status “APPROVED”. If an item version is approved then the previous approved item version becomes DEPRECATED.
- **DISAPPROVED** - items or item versions with the status “INGESTED” or “SUGGESTED” do get the status “DISAPPROVED” if they have been rejected by a Moderator/Administrator.
- **DEPRECATED** - items that are deleted by Administrators or item versions that are superseded by a newly approved item version become deprecated. They are not visible to the public in the front-end anymore. Deprecated item versions can be found in the history of an item, either the creator of such an item version or Moderator/Administrator can see them. Deleted items are still available for authorized accounts - e. g. an ingestion pipeline needs to be aware that an item is deprecated so that it is not ingesting this item again - and administrators are allowed to revert the status of such items to “APPROVED”. Items once ingested or created but not meeting the quality criteria should be deprecated.

STATUS

- ingested
- suggested
- approved
- deprecated

Screenshot of the item statuses as displayed in the curation dashboard

Creation of an individual item

Contributors thus can add to the marketplace in two ways, once by adding details to an existing marketplace item, and secondly by creating an entirely new marketplace item. In this second case, the new item will have “draft” status until it is submitted. In either case, Moderators will approve these additions and enrichments, changing their status from “suggested” to “approved.” Finally, Moderators can deprecate entries when they become obsolete or if they are judged irrelevant.

Contributors will suggest individual entries for admission into SSH Open Marketplace via an Edit Forms tool.

Label * **Version**

Ethnic and Migrant Minorities (EMM) Survey Registry beta

Description *

The EMM Survey Registry—which has been co-created by COST Action 16111 - ETHMIGSURVEYDATA, the H2020 SSHOC project, and the ANR project FAIRETHMIGQUANT—is a free online tool for discovering and learning about existing quantitative surveys undertaken with EMM (sub)populations in Europe and beyond. This tool not only brings together previously scattered information about these surveys but also offers detailed and structured metadata for each of the surveys. Designed with the end-users in mind, this tool can also easily be leveraged by anyone interested in doing research or policy-related work using quantitative surveys on EMMs' integration and inclusion.

The EMM Survey Registry is currently live and available in beta version. This version is fully functional and displays metadata about 870+ surveys from 21 different countries. In the coming months, the tool will not only expand its metadata offerings (to cover at least an additional 100-200 surveys from roughly 10 additional countries), but also enrich its various user-friendly features and functionalities.

Accessible at [x Delete](#)

<https://ethmigsurveydatahub.eu/emmregistry/>

[+ Add URL](#)

[+ Add external ID](#)

Actors [+ Create new actor](#)

Actor role	Name	x Delete
Contributor	Laura Morales	x Delete
Contributor	Ami Saji	x Delete

Properties

Property type	Value	x Delete
See also	https://ethmigsurveydatahub.eu/ethmig-survey-data-hub/	x Delete
Media	https://gitlab.gwdg.de/sshoc/sshoc-marketplace/-/raw/master/images/Screenshot%20of%20the%20EMM%20Survey%20Registry%E2%	x Delete
Select an option		x Delete
Processed at		x Delete
License		x Delete
Terms of Use URL		x Delete
Terms of Use		x Delete
Access Policy URL		x Delete
Geographical Availabilities		x Delete
Select an option		x Delete
Select an option		x Delete

[+ Add property](#)

Related items

Relation type	Item	x Delete
Relates to	Surveys to Ethnic and Migrant Minorities across Europe. Identifying Knowledge Strengths and Gaps Using Survey Metadata	
Relates to	COST Action 16111: ETHMIGSURVEYDATA Zenodo Community	

For more details about the creation of workflows, see the tutorial *“How to create a workflow in the SSH Open Marketplace?”* (which is itself a workflow in the SSH Open Marketplace).

Mass ingestion from (new) trusted third party sources

The other method of adding items to the marketplace, via mass ingestion, will give each ingested marketplace entry an “suggested” status until it can be approved by moderators.

These Editorial Guidelines do not consider the process of ingestion of new sources, which should be decided upon by a group of moderators and involve extensive mapping exercises. They should, however, be useful for the curation of ingested materials. For more information on ingestion, see SSHOC D7.2 *Marketplace Implementation*.

Metadata status

When creating a Marketplace entry, contributors will be confronted with a series of metadata fields that they must fill out. Marketplace entries each have a series of metadata, which are necessary for good indexing and proper discovery. These metadata fields have varying levels of necessity, according to the entry type of the marketplace item. This is more fully described in the next section, though the definitions of these metadata status are below.

Automatic - These are metadata fields that are automatically generated by a submission to the marketplace and are not filled out by Contributors. Indeed, marketplace users will not often interact with these fields, even though they are necessary for the database behind the Marketplace to function. These include metadata such as the entry ID.

Mandatory - These are metadata fields that are required for the entry to be submitted to review to the Moderators. These include metadata fields such as Label and Description.

Recommended - These are metadata fields which are not required to be submitted, but may lead to an entry draft being rejected by Moderators if they are lacking. Whenever possible, recommended metadata fields *should* be filled out by the contributor who submits an entry for review, as the more complete this metadata is, the more likely the entry is to help users of the Marketplace. These include metadata such as AccessibleAt, Actors, or Keyword.

Optional (unlisted) - These are metadata fields which are not required at any step of the approval process, but can help enrich the understanding of the Marketplace entry. Often, “optional” metadata fields are recommended for some entry types, but are not relevant for other entry types. These include metadata fields such as publisher (recommended for Publications, yet optional for other item types) or user manual-url (recommended for Tools and Services, optional for other item types).

Moderating the SSH Open Marketplace

The Moderation and Curation of the SSH Open Marketplace can be considered in two ways, the first being the moderation and approval of newly submitted items, and the second being the sustained curation and assurance of data quality for already-existing items in the Marketplace. While these will be handled in similar ways, there are differences.

The approval of new Marketplace items will revolve around ensuring that the submitted items are up to the standards of the Marketplace described in these guidelines. At a minimum, this will mean ensuring that all required fields are properly filled out in a standard way and that they are relevant to the source material.

The continued control of existing Marketplace items will revolve around ensuring that the information is still valid and up to date, as well as encouraging the enrichment of items by adding properties that are currently recommended. Moderators will complete these tasks with both automatic and manual checks.

In order to gain an overview of the SSH Open Marketplace data and to perform some analysis to prioritise the curation tasks and improve the Marketplace data quality, Python Jupyter notebooks have been developed. These flexible scripts allow moderators and administrators to query the SSH Open Marketplace with advanced parameters and filters and, in some cases, to write back to the system to flag some items for curation in the editorial dashboard. Details about this process will be given in D7.2 *Marketplace Implementation*.

Priority List of Editorial Actions

Moderators will be responsible for a series of actions. These actions are ordered based on a priority list, which should appear as a guide so that Moderators work efficiently through the tasks. In the Editorial Dashboard, there will be both automatic and manual checks. These checks are complementary, as while the automatic checks can signal a problem, only a human moderator can solve it. Moderators should therefore view the automatic checks as helpful time savers, yet should also understand that an automatic check cannot do all the work for them. For instance, an automatic check can verify that a Description's length is between a certain number of characters, but it cannot guarantee that this description is relevant to the tool.

Based on the quality criteria defined above, some quality metrics have been established (e.g. media quality for images or number of characters in a description) and are automatically checked. These checks are run via a series of Python notebooks. Results of the checks are recorded in a set of metadata specially created for curation purposes and only visible for moderators. Flagged metadata (i.e. fields needing improvement) are signaled and listed in the editorial dashboard, helping Moderators to see what errors must be corrected.

As moderators work through the editorial tasks highlighted in the Editorial Dashboard, they should follow a hierarchy of tasks, which is described below.

1. Moderation and validation of Suggested SSH Open Marketplace Entries - Suggested Entries should be evaluated based on the Quality criteria described above. As well, Moderators should

- make sure that metadata fields are as fully described as possible, including those that are recommended.
2. Continuous Curation of already existing items. This continuous work is organised and prioritised in the system, accessible via the editorial dashboard. As well, when examining a flagged entry, Moderators should briefly look over the other metadata fields to ensure that they are correct.
 - a. Example : Dead links: malfunctioning links should be treated with the next highest priority, as without these links our users cannot access the described tool. While an automatic check can confirm whether the link works or not, moderators will have to manually find a functioning link to replace the old one.
 3. Random Checks of pre-existing entries: Never trust a computer. Moderators should, every few months, do a random selection of pre-existing and non-flagged entries to ensure that these are up to Marketplace quality criteria, as described above.

Vocabulary management

Some of the metadata fields used in the Marketplace are using (controlled) vocabularies, meaning that contributors can select concepts among a pre-existing list. For some fields this list is open and new terms can be added (i.e. keywords) if they don't appear yet, for some others no new concepts can be added (i.e. activity is using the TaDIRAH taxonomy that is developed and maintained outside of the Marketplace).

One of the central roles the moderators have is vocabulary management in the SSH Open Marketplace: they are responsible for the vocabularies coherence and update. When a contributor suggests a new concept to a vocabulary, a moderator should decide, if the new term indeed represents a new concept, or if it is just a different lexical representation (alternative label) for an existing concept, or even if the term should not become part of vocabulary for given field at all and should be discarded.

While the functionality for basic curation of the vocabularies is integrated as part of the curation dashboard, it remains possible to effect major vocabulary management changes. The chosen vocabulary can be exported in SKOS format, imported into a separate vocabulary management tool, where the curation is performed and the new version of the vocabulary is again exported in SKOS and reimported into MP via the API. Similarly, if there is an update of the closed (external) vocabularies introduced by the authorities managing these vocabularies, this new version can be introduced to MP via SKOS import. This specific curation task will be further described in a dedicated vocabulary curation notebook.

Managing Actors

Moderators will ensure the consistency of the actors database (=> dashboard): one actor should only be registered once.

Mass validation

Though not an everyday occurrence, at times Moderators will, in concert with the SSH Open Marketplace Administrators, choose and approve newly-ingested materials. In such cases, moderators will have to firstly assess the quality and pertinence of the data source before its ingestion and once this is complete, perform "mass validation" of (meta)data before the Marketplace population. Details about this process will be given in *D7.2 Marketplace Implementation*.

III. Guidelines on individual fields

This section of the guidelines includes recommendations on metadata fields to be used while creating and/or reviewing an item entry in the SSH Open Marketplace.

The first table presents the fields that are mandatory and recommended depending on the type of items you want to create. A second table gives some explanations about the fields themselves - in case the title is not self-explanatory - how to fill in a field if you create an item and what you should check if you are a moderator.

We distinguish following types of properties:

- Mandatory fields are clearly indicated in the edit forms and you won't be able to submit an item if these fields are empty

i Label is required.

- Recommended fields are present by default in the forms, depending on the item types you want to create (i.e. if you want to create a publication, all the bibliographic metadata are included by default in the edit forms).
- Optional fields can be added via the edit forms
- There are also hidden properties that are used by the moderation team in the process of curation, and these are only visible to moderators that have logged in.

[+ Add property](#)

Property Field	Tools& Services	Training Material	Publication	Datasets	Workflows	Steps (of a workflow)
Generic Metadata						
label	mandatory					
description	mandatory					
actor	recommended					
accessibleAt	recommended					

externalID	recommended					
media	recommended					
thumbnail	recommended					
related Items	recommended					
Categorisation metadata						
activity	recommended					
keyword	recommended					
discipline	recommended					
language	recommended					
tool-family	Recommended					
service-type	Recommended					
mode-of-use	recommended					
object-format		recommended		recommended		
extent		recommended	recommended	recommended		
IntendedAudience	recommended					
standard	recommended					
resource-category	recommended					
Context metadata						
see-also	recommended					
usermanual-url	recommended					
helpdesk-url	recommended					
service-level-url						
Access metadata						

license	recommended					
terms-of-use-url	recommended	recommended			recommended	
terms-of-use						
accesspolicy-url						
geographical-availabilities						
privacy-policy-url						
authentication						
Bibliographic metadata						
publication-type			recommended			
publisher			recommended	recommended		
publication-place			recommended			
year		recommended	recommended	recommended		
journal			recommended			
conference			recommended			
volume			recommended			
issue			recommended			
pages			recommended			
Technical metadata						
life-cycle-status						
technical-readiness-level	recommended					
version	recommended					

The list of recommended fields is also [recorded in the front-end code](#).

Guidance for metadata fields

Property Field's name	How to fill in? Some recommendations for contributors	Quality checks: what and how moderators can validate content "Dashboard" = problematic items are automatically flagged and prioritised for curation within the editorial dashboard "Manually" = moderator has to assess coherence and relevance of the info
Accessible at = "Go to..." button	URL of the resource. Landing page of the resource which could be the persistent identifier. Ideally, only one URL should be added as accessibleAt. For other URLs, see "externalID".	Check that link is working (=> dashboard)
Access Policy URL	URL pointing to information about the access policy that applies to the Resource. Follows EOSC Portal Providers Documentation Related to `Authentication`	Moderators will check for coherence of access policy url with linked Marketplace entry. (=> manually) Check that link is working (=> dashboard)
Activity	Contributors will select a term, drawn from the controlled vocabulary TaDIRAH , that describes the activities you can perform using the resource showcased in the SSH Open Marketplace Entry.	Moderators will check for consistency between activity selected and the resource described. (=> manually)
Actors	An Actor is a person or an organisation who contributed in some way to the creation of the linked resource described in the SSH Open Marketplace. Name of an actor is mandatory: while creating an actor, you need at least to add a name.	Moderators will ensure the consistency of the actors database (=> dashboard): one actor should only be registered once.
	Define role as one of: contributor author reviewer director programmer editor provider contact funder helpdesk	It is recommended that the Actor(s) be identified by an externalID (esp. ORCID). Anything that can help to identify the contributor is recommended

	<p>By default, the Actor's role is contributor.</p> <p>Add ExternalID for actors: ORCID, recommended.</p> <p>Add ROR IDs for research organisation.</p>	
Authentication	<p>Describe the conditions or restrictions for a user to access the resource: "yes" when an authentication is needed, "no" when the resource is accessible without authentication.</p> <p>Related to `Access Policy URL` which provides more details about access to the resource.</p>	<p>Moderators will check for coherence of authentication with linked MP entry. (=> manually)</p>
Conference	<p>If the publication concerns a paper presented during a Conference, specify the name of the Conference</p>	<p>Moderators will check for coherence of conference selected with linked Marketplace entry. (=> manually)</p>
Description	<p>Check the original resource, and if possible, reuse the text description of the source. This description should be short (3 sentences recommended) and should aim at raising interest Example: "Gephi is a Photoshop alternative for the humanities...."!</p> <p>Workflows specificities: longer descriptions may be needed for a workflow and its steps. ~1500 characters are recommended.</p> <p>Description field supports Markdown, although to include links we recommend using the `see also` property.</p>	<p>Ensure that description corresponds to and is up to date with the Marketplace entry described (=>manually)</p> <p>Every entry in the Marketplace should have a description between 25 and 1500 characters (=> dashboard)</p>
Discipline	<p>Contributors will select a term, drawn from the controlled vocabulary Austrian Fields of Science and Technology Classification 2012, that describes the discipline covered by the resource showcased in the MP.</p> <p>If a given resource is pertinent to too many disciplines, use the top-level concepts (such as "social sciences" or "humanities". If the resource is discipline-agnostic or too generic, then leave blank.</p>	<p>Moderator will check for consistency between discipline selected and the resource described. (=> manually)</p>

Extent	If an item covers other sub-items, use extent to describe the granularity. Example: you create an item for a training-material course that has 50 lessons (extent = 50 lessons)	Moderator will check for consistency between extent indicated and the resource described. (=> manually)
external ID	<p>External identifiers (or labeled URLs) can be added to any items. At the moment, Wikidata, Github and DOIs identifiers can be added to the Marketplace items.</p> <p>These external identifiers are also used for actors (see actors recommendations) and ORCID iD, Wikidata or DBLP identifiers can be added to an actor's profile.</p>	Check that link is working (=> dashboard)
Geographical Availability	<p>Locations where the Resource is offered to be chosen among the list of controlled values coming from the EOSC Resource Geographical Availability.</p> <p>Follows EOSC Portal Providers Documentation</p>	
Helpdesk URL	<p>The URL to a webpage to ask more information from the Provider about this Resource.</p> <p>Link to the helpdesk where users with incidents and requests will be directed. It should allow private communication with providers (not a public forum).</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of helpdesk url with linked Marketplace entry. (=> manually)</p> <p>Check that link is working (=> dashboard)</p>
Intended audience	Contributors will select a term, drawn from a controlled vocabulary, that describes the intended audience targeted by the resource showcased in the MP.	Moderators will check for coherence of audience selected with linked Marketplace entry. (=> manually)
Issue	A particular published issue of a journal, one or more of which will constitute a volume of the journal.	Moderators will check for coherence of issue selected with linked Marketplace entry. (=> manually)

Journal	The journal or magazine the work was published in. Please use ExternalID to add a PID (persistent identifier) for the journal, such as an ISSN.	Moderators will check for coherence of journal selected with linked Marketplace entry. (=> manually)
Keyword	Contributors will select an existing keyword or create concepts/terms for categorizing items to be used only if the concept does not fit in the TADIRAH activities used in the activity field. The recommended number of tags per item is 5.	Moderator will check for consistency between keywords selected and the resource described. (=> manually) Need to check for duplicates of terms regularly (=> dashboard & vocabulary manager)
Label	Name of the MP entry, ideally using the original name, spelling and case sensitivity. Multilinguality supported. No line break Max length recommended 255 characters	Go to website “accessibleAt” and compare label with the name of the tool on the website (=> manually) Is label already used by another MP entry (=> dashboard duplication & merging of entries)
Language	Language(s) of content of a resource, or language in which an interface or a resource is available. Contributors will select a term, drawn from the controlled vocabulary ISO 639-3.	Regular checks for non-English content (=> dashboard) Moderator will check for consistency between language(s) selected and the resource described. (=> manually)
License	Contributors will select a license, drawn from the controlled vocabulary SPDX License List . If no matching, leave empty and fill in the terms of use.	Moderator will check for consistency between the license selected and the resource described. (=> manually)
Life Cycle Status	Status of the Resource life-cycle. List of controlled values coming from the EOSC Resource Life Cycle Status Follows EOSC Portal Providers Documentation	Moderators will check for coherence of life-cycle-status with linked MP entry (=> manually).
Media	Media can be images or videos. Some recommendations: Images: 1. Look for freely available screenshots or visualisations of a tool (CC licence or similar). 2. If that doesn't yield good results, look for material available on the original website.	Automatic checks will determine presence or not of media (=> dashboard). Moderators will check for coherence of media with linked MP entry. (=> manually) More media-sources can be added by moderators using the API call:

	<p>3. Consider making own screenshots.</p> <p>4. NOTE: You can add the URL of an image, if you choose to upload, image must be 1MB or smaller in size. Format supported: JPG, TIFF, PNG, GIF</p> <p>Videos: 1. If on YouTube or Vimeo: just deliver the URL (title/caption of the video might be retrieved automatically). 2. If elsewhere: provide link to video and descriptive caption.</p> <p>Media-caption Every media should include a descriptive caption, i.e verbose information mentioning license and credits about the media used.</p>	<p>/api/media-sources. More information in the moderators guidelines.</p>
Mode of use	<p>Contributors will select a mode of use type, drawn from an open controlled vocabulary. Examples: desktop; client (install locally); browser-based; mobile app; programming library; API; (REST-API). New concepts can be suggested to the vocabulary if needed.</p>	<p>Moderators will check for coherence of mode of use with linked MP entry. (=> manually)</p>
Object format	<p>File format of the linked resource. Choose from the controlled vocabulary IANA Media Types.</p>	<p>Moderators will check for coherence of format selected with linked MP entry. (=> manually)</p>
Pages	<p>Page numbers, separated either by commas or double-hyphens.</p>	<p>Moderators will check for coherence of pages selected with linked MP entry. (=> manually)</p>
Privacy policy URL	<p>Link to the privacy policy applicable to the Resource. A Privacy Policy is a document that explains how and for what purposes the linked tool or service will utilize, disseminate, and stock user's data.</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of privacy policy url with linked MP entry. (=> manually)</p> <p>Check that link is working (=> dashboard)</p>
Publication type	<p>Contributors will select a publication type, drawn from the controlled vocabulary Bibliographic Ontology BIBO.</p>	<p>Moderators will check for coherence of publication type selected with linked MP entry. (=> manually)</p>

Publication Place	The place of publication	Moderators will check for coherence of publication-place selected with linked MP entry. (=> manually)
Publisher	An entity responsible for making the resource available.	Moderators will check for coherence of publisher selected with linked MP entry. (=> manually)
Related items	<p>Relations between SSH Open Marketplace items help to create context. If you know about a training material describing or citing a tool or service for example, use this to provide more context.</p> <p>To relate to items external to the Marketplace, use the “See also” property.</p>	<p>Relations between items are central to ensure contextualisation and quality of the Marketplace entries!</p> <p>Check the coherence of the relations suggested (=> manually)</p> <p>Automatic checks will determine presence or not and numbers of relations (=> dashboard).</p>
	<p>Relation type:</p> <ul style="list-style-type: none"> -relates-to/is-related-to: to mark a simple relation between items -documents/is-documented-by: use the document relation when a resource provides information on another -mentions/is-mentioned-in: when one resource refers to another <p>By default, the relation-type is “related-to”</p>	
	Item: search among the existing items in the Marketplace.	
Resource category	<p>Contributors will select a resource category type, drawn from the EOSC categories controlled vocabulary. This helps refining the resource’s nature and function.</p> <p>Follows EOSC Portal Providers Documentation</p>	Moderator will check for consistency between categories added and the resource described. (=> manually)
See also	<p>This is a URL-based field. It references additional materials linked to the resource, but external to the SSH Open Marketplace. These are interesting resources we don’t want to add as entries in the MP, but that can be useful for the users.</p>	<p>Moderator will check for consistency between references added and the resource described. (=> manually)</p> <p>Automatic checks with other URLs-based fields (=> dashboard)</p> <p>Check that link is working (=> dashboard)</p>

	For relations between MP entries, see `related items`.	
Service Level URL	<p>Webpage with the information about the levels of performance of the Resource that a Provider is expected to deliver. For example, link to a Service Level Agreement (SLA) or Service Level Specification (SLS) applicable to the Resource.</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of service level url with linked MP entry. (=> manually)</p> <p>Check that link is working (=> dashboard)</p>
Standard	<p>Contributors will select a standard used by the tool or service, drawn from a controlled vocabulary built within the PARTHENOS project. If there is no match in the list of concepts provided, contributors can suggest a new standard.</p>	<p>Moderators will check for coherence of the standard selected with linked MP entry. (=> manually)</p>
Technology Readiness Level	<p>The Technology Readiness Level of the Resource. List of controlled values from the EOSC Resource TRL</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of TRL level with linked MP entry. (=> manually)</p>
Terms Of Use	<p>If the license is not known among the choices provided by the controlled vocabulary, Terms Of Use provides a free text field to describe the type of use (terms of use, copyright...)</p>	<p>Field to check regularly and change into license if matches are identified.</p> <p>Need to check for duplicates of terms regularly (=> dashboard & vocabulary manager)</p>
Terms of Use URL	<p>Webpage describing the rules, resource conditions and usage policy which one must agree to abide by in order to use the resource.</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of Terms of use url with linked MP entry. (=> manually)</p> <p>Check that link is working (=> dashboard)</p>
Thumbnail	<p>Please try to provide a logo related to the MP Entry, if we are authorized to use it. If no logo is submitted, a generic icon is automatically provided.</p> <p>Recommendations in terms of quality: 200x200 pixel.</p>	<p>Moderator will check for consistency between image added and the resource described, and for quality of the image chosen. (=> manually)</p>

Tool Family	This field can be used to further define the type of tool, or if the tool belongs to a suite of tools for example.	Moderators will check for coherence of the tool family indicated with linked MP entry. (=> manually)
User Manual URL	<p>Link to the Resource user manual, documentation, or other basic user instructional information about the Resource.</p> <p>Follows EOSC Portal Providers Documentation</p>	<p>Moderators will check for coherence of user manual url with linked MP entry. (=> manually)</p> <p>Check that link is working (=> dashboard)</p>
Version	<p>This refers to the version of the tool or service that is referenced in the Marketplace. This refers to the names, often in the form of numbers (e.g. 1.2), that describe unique states of a tool or service.</p> <p>Recommendations in case of several versions: creating a new item or updating version field?</p>	
Volume	The volume of a journal or multivolume book.	Moderators will check for coherence of volume selected with linked MP entry. (=> manually)
Year	The year of publication (or, if unpublished, the year of creation)	Moderators will check for coherence of year selected with linked MP entry. (=> manually)

Guidance on curation properties/metadata

These properties are hidden from users and contributors but are available for logged in moderators. The “curation-” metadata are used by the curation notebooks as well as to filter the “items to moderate” in the dashboard.

Field's name	Quality checks: what moderators should know
source-last-update	records the last modification of the information at the source (if such information is available at the source). This field is useful for moderators to understand the time of last change to the entry at source.
processed-at	Property used to record the date at which the relation between items based on the tool extraction module was processed.
deprecated-at-source	records the fact that an item ingested in the Marketplace was deprecated at source.

curator-detail	hidden property of type string that contains all curator issues in a human readable structured format. For example, this field describes when the curator-flag-url is TRUE where the url has problems and the http status code as a json, e.g. {"AccessibleAt": "404"}
curator-flag-url	hidden property of type boolean, is "TRUE" when an error on at least one of the URL-based fields of an item has been identified
curator-flag-description	hidden property of type boolean, is "TRUE" when there is a problem with the description of an item (for example when the description is "no description provided" and should be changed).
curator-flag-coverage	hidden property of type boolean, is "TRUE" when coverage ratio of recommended properties is lesser as expected
curator-flag-relations	hidden property of type boolean, is "TRUE" when the contextualisation quality is low (based on interlinked items)