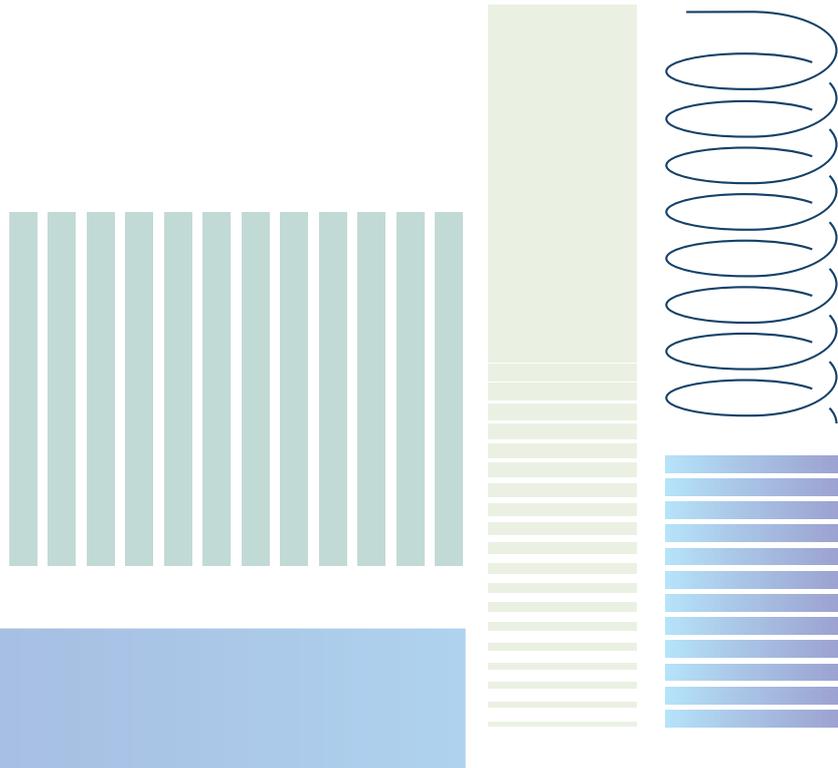


D7.2. Testing report in a real scenario



Deliverable number and name: D 7.2 – Testing Report in a Real Scenario

Due date: 30/03/2021

Work Package: WP7

Deliverable leader: UNIPA

Authors: Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella

Reviewers: Mar Gaitan, Javier Sevilla Peris

Approved by: Cristina Portalés

Dissemination level: Public

Version: 1.1

Document revision history			
Version	Date	Contributor	Comments
0.1	19/01/2021	Valeria Seidita, Georgia Lo Cicero	Initial document with TOC
0.2	25/01/2021	Valeria Seidita, Georgia Lo Cicero	Test cases added
0.3	30/01/2021	Valeria Seidita	Methodology and Discussion added
0.3	5/02/2021	Valeria Seidita, Georgia Lo Cicero	The whole document completed
0.4	10/02/2021	Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella	First revision
0.7	11/03/2021	Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella	Version submitted to reviewers
0.8	12/03/2021	Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella, Mar Gaitán, Javier Sevilla	First reviewer comments met
1.0	18/03/2021	Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella	Final revision and proofreading
1.1	23/03/2021	Valeria Seidita, Georgia Lo Cicero, Maurizio Vitella	Quality checks implemented

List of acronyms	
CH	Cultural Heritage
ICT	Information and Communication Technology
KG	Knowledge Graph
TA	Target Audience

Table of contents

1. INTRODUCTION.....	4
2. RELATIONSHIP WITH OTHER DELIVERABLES.....	5
3. TARGET AUDIENCE DESCRIPTION	5
4. METHODOLOGY	7
5. TEST CASES.....	10
5.1. Scenario 1 - Cultural heritage and leisure.....	10
5.2. Scenario 2 - Research and Education sector.....	11
5.3. Scenario 3 - Media.....	12
5.4. Scenario 4 - Creative Industries	13
5.5. Scenario 5 - Tourism	13
6. GENERAL FEEDBACK.....	13
6.1. Results on random inputs	13
6.1.1. Unexpected results from a domain expert point of view.....	14
6.1.2. Repeated objects	23
6.1.3. Results not adequately filtered.....	27
6.2. Stress test.....	28
6.3. Dead page or invalid redirects	28
7. RESULTS OF THE TESTING ACTIVITIES	28
7.1. List of errors	28
7.2. Needed improvements and how they have been solved.....	29
8. CONCLUSION.....	30
9. REFERENCES	31

This deliverable reports the results of testing the system in real scenarios, which were previously defined as part of D2.4. The test aims to evaluate the performance of the system by domain experts, who can validate if the outcomes are correct. The document illustrates the methodology used to perform the test phase, the list of errors found and reports feedback on how to improve the system.

1. INTRODUCTION

In this report, we present the work in Task 7.2. The purpose of the task is to test the SILKNOW integrated system. It should be noted that ADASilk is still under development at the time of conducting the tests in Task 7.2 and therefore some modules are not yet fully integrated. However, the goal of these tests is to obtain relevant conclusions to improve ADASilk and therefore the results of this deliverable will be used for the final integration of all modules into ADASilk.

The main purpose of this task is to provide feedback to the project partners and developers on the results of the text search integrated in ADASilk in order to improve the final product in the final integration phase.

A test of some modules of the system has already been performed in Task 7.1. This test was performed in parallel on the three main modules of the system: the text analysis module, the knowledge graph module and the image classification module. These were tested by experienced users and experts in the field of silk and cultural heritage. The aim of Task 7.1 was to ensure the quality and consistency of the product with the objectives of the SILKNOW project, for which specific scenarios were used. The objective of Task 7.2 is the same as 7.1 and therefore a team of people from the ICT sector and experts from Cultural Heritage were also involved. The team performed tests in real scenarios inspired by the scenario reported in Deliverable D2.4.

The system was primarily tested using a black box technique. The system was accessed through the ADASilk web application (we used the <https://ada-preprod.silknow.org/> application provided for internal use). The test team simulated the use of the system by users belonging to the Target Audiences identified in D2.4. The team entered specific inputs resulting from the scenarios identified in D2.4 and observed the output. The output was observed from both a technical and semantic perspective to determine if the results were consistent with the SILKNOW project goals. This was possible due to the joint presence of ICT and CH experts in the team. No judgements were made about the reason for some search results, only highlighting what was considered to be incorrect or anomalous behavior. The list of found potential anomalies was presented to partners and included in this document. Partner feedback will be helpful in improving the final product.

The rest of the document is organized as follows: Section 2 reports the relationships with the other deliverables, Section 3 briefly illustrates the user profiles included in Target Audience from D2.4 to better understand the testing process, Sections 4 and 5 describe the methodology used for testing and the associated test cases, Section 6 describes the black-box test performed on the system using random inputs from CH, Section 7 reports the results of the testing activities followed by a discussion, and finally, Section 8 draws some conclusions.

2. RELATIONSHIP WITH OTHER DELIVERABLES

In Deliverable 7.1, we created a test report in a controlled scenario, while in Deliverable 7.2 we consider a real-world scenario. In Deliverable 7.1, we reported how some domain experts mainly performed a semantic annotation of the text analytics module and a verification of the correct mapping of the data available in the museum collections to those in Knowledge Graph (KG). In this deliverable, the previous tests were extended to the already integrated part of the system, and real scenarios were used. For this purpose, museum data available in KG were considered and a series of workshops with different scenarios were conducted, involving a number of domain experts simulating individuals from the TAs in Deliverable 2.4.

In Deliverable 2.4, 26 different scenarios were identified. These scenarios included different work and cultural environments to create hypothetical archetypes and provide the user's cultural identity to predict possible targets of interest. The scenarios were divided into five different types derived from some of the seven Silknow Target Audiences (see 2.21 in the Description of the Action and deliverables 8.3, 8.4, 8.5). The areas of the scenarios are: 1) Cultural Heritage, 2) Research and Education Sector, 3) Creative Industries, 4) Tourism, and 5) Media. In Deliverable 7.2, the test team emulates domain experts from the different scenarios, tests the functionality of the system, and validates whether the results are appropriate.

The results of this report will mainly feed into WP6, as all software modules are integrated in this work package (including text and image analysis). We have also developed ADASilk, the exploratory search engine, as part of WP6, which we have used in real-world scenarios for validation by domain experts.

3. TARGET AUDIENCE DESCRIPTION

In this section, we describe the types of users in the TAs identified in Deliverable 2.4. Here we briefly summarize these user profiles to better understand the inputs to the test cases in Section 5, and to explain why we reported some results as incorrect or strange or not in line with the goal of the system. The test team had these different profiles in mind when they ran the tests.

Individuals from Cultural Heritage audience typically visit cultural institutions such as museums, are interested in furthering their research in texts and internet searches, and are generally interested in historical and productive information about silk. They are often people

SILKNOW

who work in museums or in the cultural heritage field in general (collector, museum curator, museum curator, fashionista, middle-aged museum visitor, staff member of international organizations, museum director). They can be considered both experienced and non-experienced users of the SILKNOW system (from now on we will refer to the system under test by the name of the connected web application, namely ADASilk). In this case, ADASilk and its associated tools can be used to obtain specific information and share it through social media, etc.

The Research and Education sector includes people who can benefit from all the tools of SILKNOW (search, thesaurus, virtual loom, space-time map). Primarily, they will use the results of ADASilk as an educational support. They can be considered as normal users interested in analyzing all the results to get new information and improve their own background.

Creative Industries is another basic area that is part of the target audiences. It includes people from industries that are particularly interested in finding information about the application of new technologies. Typical users include the CEO of a silk manufacturer, fashion designers, textile designers, photographers, hardware manufacturers (such as 3D printers). They usually perform a specific search and expect precise and detailed results.

The valorization of the textile world is also relevant for the tourism sector, so we wanted to analyze the point of view of users belonging to it; the analyzed characters are local guides, museum marketing professionals, regional associations, young museum visitors.

The media sector includes all fashion journalists looking for information to share in blogs or promote in publications. These TAs emphasize the importance of sharing information on a daily basis.

D2.4 presents all these user needs and the final uses of information and digital materials that emerged from the ADASilk consultation. The user scenarios describe what and how users search for, find, use and share ADASilk results.

We assume that people from Cultural Heritage and Creative Industries are experienced users who perform specific searches and expect specific results. They are likely to use filters and typically provide very specific input. Everyone else is a normal user who is likely to use mainly the main bar Text Search (ADASilk and all its features are described in D6.5). Anyway, we cannot predict the sequence of actions that users will perform, and Task 7.3 is in some sense devoted to exploring this issue. Indeed, during the evaluation phase we will gain information about the nature of the users and their opinion about the usability of all the tools provided. Therefore, the test team provided input for each user type in different ways, as explained in the next sections.

4. METHODOLOGY

Before describing the testing methodology used in this task, it is necessary to make a premise about the type of software we are testing. ADASilk, the web-based platform that emerged from the SILKNOW project, provides text, images, and printable patterns to enhance users' knowledge of silk textiles. ADASilk was designed and built as an exploratory search engine. An exploratory search engine is a specialized search engine that aims to perform search tasks for people who are not familiar with certain application domains or who are not sure how to reach their goal. In a standard search engine, the user has a specific goal in mind and searches for that goal. In exploratory search, on the other hand, information retrieval is aimed at the type of user mentioned above and is focused on research, comparison and analysis. In SILKNOW, the core of the system that provides this type of functionality is in the ontology, how it is structured and how it is used for queries, as explained in D6.5. In this task, we have taken this into account when choosing the type of test that will be described later in this section.

Testing is the process of analyzing a system or some of its components to determine differences between expected and observed behavior [1][2][3]. In general, it is not possible to test a complex system completely and simply. Testing is defined as the systematic attempt to find bugs in the developed software, instead of the concept that sees testing as the phase that proves that there are no bugs. The approach we used in Task 7.2 is based on the first definition of testing.

Our explicit goal was to prove the existence of bugs and suboptimal behavior. To this end, we referred to the scenarios identified in D2.4, which describe the expected operation of the software from the perspective of different users. Each of these scenarios represents the users' goals and what they want to achieve. So we (the testing team - ICT and CH experts) derived some possible inputs from the goal of each scenario and examined the related outcome as if we were users belonging to different audiences. We gradually sent the results of our tests to the developers and other partners to make the necessary improvements to the system. In this report, we list the results of the tests, the improvements made after feedback with the developers, and what still needs to be done.

The work on this task was done in parallel with the development of the software. When we did the testing, some modules were not quite ready, so we focused on the part of the software that displays the elements of the Thesaurus in the Knowledge Graph to perform simple searches. We used ADASilk to explore the Knowledge Graph (described in [4]) and performed searches that implemented the above scenarios.

Among the techniques used for testing software, we find fault detection, which allows finding faults in the system. Generally, it is done in the development phase using black box techniques. Fault detection can be well applied to the work required in Task 7.2, and we implemented it using so-called test cases.

SILKNOW

A test case defines a set of input data and its expected results when using a component or part of the software. The main purpose of using test cases is to force the system to produce incorrect results in order to identify bugs. A test case usually has five attributes: a name, a location, an input, an oracle (the expected output), and a log (the actual output).

We constructed all test cases starting from use cases or scenarios that have a close relationship to the functional requirements of the software identified in the analysis phase. In our case, the project partners identified the requirements during the development of WP2. The test case name is usually derived from the use case or scenario to which the test case relates. The location describes where the test case should be found for execution when automatic tools are used for testing. In this task we did not use automatic tools. The test was performed by interacting with the ADASilk web application on the user side, so it was not necessary to include this field in the test case. The input describes a set of data or commands to be given to the system. The oracle describes the correct behavior expected when the test case is executed. The log is a set of temporal correlations between the expected and actual behavior. It provides information about the extent to which the operation deviates from that specified in the scenarios, and also identifies any errors.

The next section presents the test cases used for this task. To identify the inputs, the test team identified a set of inputs for each scenario and the expected output. To identify the inputs, we looked at the users' descriptions of "research interest and use" (see Figure 1). To identify the expected output, we analyzed the "Expectation and Sources" field. We excluded all scenarios that included image comparison because, as mentioned earlier, this module was not yet integrated at the time of testing.

For example, let us consider the first scenario. For clarification, a screenshot is shown below (see Figure 1). It describes the interaction of a user of Cultural Heritage, who wants to use ADASilk to enhance his knowledge about fashionable textiles from the 1800s.

Scenario 1. Cultural heritage and leisure	
Case 2	
Fernando, 42, comes from Argentina. He speaks both Spanish and English and is professional writer. He is a big fan of fashion and defines himself as a ' <i>fashionista</i> ', in fact he runs a blog and is very active on Instagram. Together with his friends he often participates in historical re-enactment festivals, so he is very interested in historical textiles	
Personal key facts	Fernando, 42, <i>fashionista</i> and passionate about historical costumes, uses his blog and Instagram daily
Scenario	
Search/Research Interest	Fernando gathers information on SILKNOW about the most fashionable textiles in the early '800, to create costumes inspired by this historical period, as he has to take part in a Festival of historical re-enactment
Use	Flipping through the catalog of images on SILKNOW, Fernando is able to understand the colors, the decorative patterns and the types of fabrics that were fashionable in that particular historical period.
Expectation and Sources	Fernando looks for images to be downloaded and searches for them by historical period. He looks specifically for decorative patterns to be replicated to carry out his costumes. Once he finds them he will share his research on his blog and on Instagram

Figure 1. Scenario description from D2.4

1) a list of historical information that the user can navigate through to analyze and acquire new knowledge.

Attributes	Description
Name	Searching for historical information
Input	19th century and brocade
Oracle	A list of historical information on 19th century brocade

log	A list of objects - by clicking on an object ADASilk should show a complete description from which to take historical information, it is not so for every object.
-----	---

Commands in ADASilk:

Text Search: Brocade

Production Time: Nineteenth century

The previous table shows the test case related to the scenario in Figure 1; here the last row contains the log. As you can see, the real behavior matches the expected one. In fact, ADASilk displays a list of objects (images of objects) that the user can click on to get more information and expand his knowledge about brocade. The same thing happens when we change the input values. We created several different test cases with several different inputs for almost every scenario. In the next section, we report only one or two of them per category. We noticed that for some objects the full description is not yet available. This is because the software is still under development and not all the data from the museums' historical records has been entered into the system. However, we found it helpful to record all these situations to help other partners and speed up their work.

In addition to the test cases, we performed another type of test based on the scenarios and functionalities listed in D2.4, and performed a simple test phase, that might be called a brute force test. Mainly we inserted random inputs to observe the behavior of the system. Section 6 contains the results of this last part of the test, as well as some evaluation and discussion. In each test, it is worth highlighting that Text Search was used when the desired item was not available in the filters (e.g., in the case of Doublet) or more often to simulate the specific willingness of a normal user to use text search, which we consider to be easier and more immediate for a user.

5. TEST CASES

In this section, we list the test cases we identified for each scenario in D2.4.

The test cases were identified with the help of people involved in defining the scenarios, as well as domain experts from the test team in the role of end users. We have performed about a hundred tests with different inputs for each scenario. The behavior of the system can be described considering the list below.

5.1. Scenario 1 - Cultural heritage and leisure

Attributes	Description
Name	Searching for historical information
Input	19th century and brocade
Oracle	A list of historical information on 19th century brocade

log	A list of objects - by clicking on an object ADASilk should show a complete description from which to take historical information, not so for all the objects.
-----	--

commands in ADASilk:

Text Search: Brocade

Production Time: Nineteenth century

Attributes	Description
Name	Searching for historical information_bis
Input	18th century and chinoiserie
Oracle	A list of historical information on 18th century chinoiserie
log	A list of objects - by clicking on an object ADASilk should show a complete description from which to take historical information, not so for all the objects.

commands in ADASilk:

Text Search: Chinoiserie

Production Time: Eighteenth century

Attributes	Description
Name	Searching for historical information_ter
Input	18th century and flowers
Oracle	A list of historical information on decorative pattern in 18th century containing flowers
log	A list of objects - by clicking on an object ADASilk should show a complete description from which to take historical information not so for all the objects.

commands in ADASilk:

Text Search: 18th century

Depiction: Floral Motif

5.2. Scenario 2 - Research and Education sector

Attributes	Description
Name	Information about general option for education
Input	Chasuble of fifteenth century
Oracle	A list of chasubles of the fifteenth century
log	A list of chasubles of fifteenth century

commands in ADASilk:

Text search= casulla (in this case we used the Spanish term)

Production time: fifteenth century

Attributes	Description
Name	Materials for education
Input	Animal Fibre
Oracle	A list of fabrics produced in animal fibre
log	A list of objects of different material, mainly silk, and a lot of repetitions

commands in ADASilk:

Materials: Animal Fibre

Attributes	Description
Name	Materials for education (an English teacher)
Input	thesaurus>Groups>techniques>
Oracle	A list of weaving techniques in the Thesaurus
log	A list of weaving techniques in the Thesaurus

commands in ADASilk:

main page: link on thesaurus, groups and then techniques

5.3. Scenario 3 - Media

Attributes	Description
Name	Fashionista
Input	1800 and doublet
Oracle	A list of images of doublets
log	A list of images (not only doublets)

commands in ADASilk:

Production time: 19th century

Text Search: Doublet

Attributes	Description
Name	Fashionista_bis (search for fabric with parrots)
Input	Parrot
Oracle	A list of objects with patterns of parrots
log	none

commands in ADASilk:

Depiction: Parrot

5.4. Scenario 4 - Creative Industries

Attributes	Description
Name	Searching a pattern to reproduce in 3D
Input	Bizarre motif
Oracle	A list of fabrics with bizarre patterns
log	A list of fabrics with bizarre patterns

commands in ADASilk:

Depiction: Bizarre Motif

Running Virtual Loom

Note: All instances of the Creative Industries scenarios fall into the simple search. Task 7.2 does not need to test the Virtual Loom, only whether the search engine returns the correct results for a given query, so we did not perform many tests for this scenario.

5.5. Scenario 5 - Tourism

Attributes	Description
Name	Organizing a tour in Valencia
Input	Valencia
Oracle	A list of images of various objects (fans, chairs, fragments of fabric, etc.) that can be found in Valencia or have some relation with Valencia
log	A list of images of various objects – not all of them are related to Valencia

commands in ADASilk:

Text search=Valencia

Attributes	Description
Name	Collection of ancient textiles
Input	Ancient textiles
Oracle	Data on the history and use of the weave of interest
log	A list of seven object without description

commands in ADASilk:

Text search= Ancient Textiles

6. GENERAL FEEDBACK

6.1. Results on random inputs

In this section, we illustrate the results of the random inputs with a series of screenshots and associated comments on the behavior of ADASilk from the experts' point of view.

6.1.1. Unexpected results from a domain expert point of view

Below, for each input given to ADASilk, we describe the result and the reason why the ICT or the CH experts in the test team want to bring it to the attention of the developers.

Figure 2 shows the first of the results that appear when we enter the term Giubba (Italian synonym for jacket).

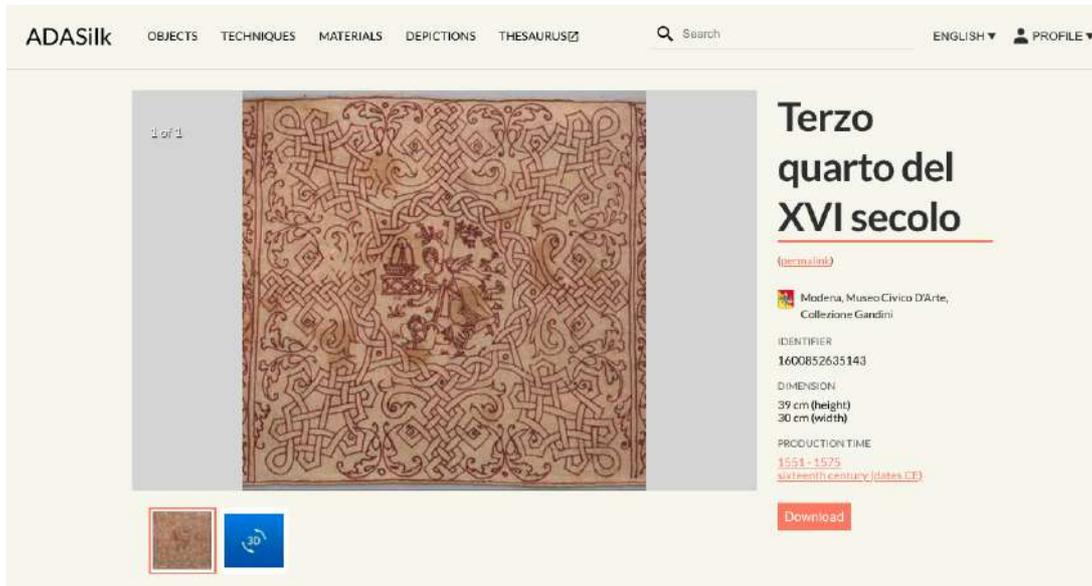


Figure 2. One of the results for Giubba.

The result does not represent a "giubba" but is a piece of fabric with a drawing representing a woman and a man. The man is wearing a jacket (giubba in Italian). Probably the result is due to the fact that the word giubba appears in the object description "l'uomo indossa giubba e pantaloni rigonfi corti al ginocchio". The experts expected the visualization of a list of images representing jackets.

Figure 3 shows the list of results when we type the term "doublet" in the main search bar. Figure 4 shows the details of the second object in the third row.

Since this result did not seem to be relevant to the search, the test team decided to check the permalink (<https://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debcb>).

The Permalink is the reference to the knowledge graph. By analyzing the values of each attribute, we can see that the object in Figure 4 is in the result list because it contains part of the term "doublet" as part of another term, namely doublet . This fact occurs because each term inserted in the Text Search is searched (in whole or in part) in the whole description of all the objects in the KG and this causes the results not to match completely the search performed.

16787 search results

Sort by [Select...](#)

Man's ensemble, 1630 - 1639, England

Tapestry fragment, 1600, Warwickshire

Doublet, 1615 - 1620, England

Doublet, 1620 - 1625, England

Doublet and breeches, 1625 - 1635, England

Picadil, 1600 - 1615, England

Suit, 1635 - 1640, England

Stage costumes, disfraces, 1888, London

Doublet, 1625 - 1630, Italy

Doublet, 1625 - 1630, Italy

Doublet and breeches, 1630 -

fleurs

bouquet de fleurs

semis de fleurettes

semis de fleurettes

Previous 1 2 3 4 5 6 ... 839 840 Next

Figure 3. The list of results for Doublet.

ADASilk OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS

Search ENGLISH PROFILE

1 of 1

fleurs

Joconde Saint-Etienne; musée d'Art et d'Industrie

IDENTIFIER: 10230001672

MATERIAL: Silk

DIMENSION: 183.5 cm (height), 5.2 cm (width)

PRODUCTION TIME: 1855 - 1900, 1901 - 1925, nineteenth century (dates CE), twentieth century (dates CE)

PRODUCTION PLACE: Saint-Etienne

TYPE OF OBJECT: fabric, passementerie, fabrics

Download

Figure 4. A specific object in the list found for Doublet.

About: <http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc> [Goto](#) [Sponge](#) [NotDistinct](#) [Permalink](#)

An Entity of Type : [ecrm:E22_Man-Made_Object](#), within Data Space : [data.silknow.org](#) associated with source [document\(s\)](#)

Type: [E22_Man-Made_Object](#) [New Facet based on Instances of this Class](#)

Attributes	Values
rdfs:type	E22_Man-Made_Object
rdfs:label	fleurs (fr)
rdfs:comment	Ruban en soie, satin damassé jaune, broché en couleurs, blanc, rose et doré, bord royale avec décor par doubleté (fr)
P102_has_title	fleurs (fr)
P138_has_representation	http://data.silknow.org/image/9afe060a-sce5-3fe7-bb68-e2731d693e14
P1_is_identified_by	95.71.74 Tg.74
P3_has_note	Ruban en soie, satin damassé jaune, broché en couleurs, blanc, rose et doré, bord royale avec décor par doubleté (fr)
P43_has_dimension	http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/dimension/h http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/dimension/w
P62_depicts	représentation végétale (rinseau, fleur, figure ondulante) (fr)
dc:identifier	10230001672
is P104_applies_to	http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/right
is P108_has_produced_of	http://data.silknow.org/production/e80a5abf-aaf6-3b26-ae20-62e919821d6c
is P129_is_about_of	ID_10230001672_filename_10230001672.json http://data.silknow.org/object/1cad017c-9b80-3674-94aa-731a8420fab http://data.silknow.org/object/84975139-449f-3eed-a33c-5267c1c389c9
is P30_transferred_custody_of	http://data.silknow.org/event/4271c4b9-b7c3-387f-87e2-102eed1ca4a3
is P39_measured_of	http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/dimension/measurement
is P41_classified_of	http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/type_assignment/1 http://data.silknow.org/object/524a6282-f29d-3445-a72f-a71ac44debc/type_assignment/2

Figure 5. Content in the KG related to the object in Figure 4

Figure 6 and Figure 7 respectively show the results when "wall hanging" and "arazzo" are entered. They are synonymous, so we expected the same number of results.

ADASilk OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS ENGLISH

[CLEAR](#) [FILTER](#)

Text search:

Production time:

Production place:

Material:

Technique:

Depiction:

Type of object:

Museum:

18277 search results

Sort by

Suzanni (embroidered wall hanging), April 200...	Wall hanging, 1710 - 1720, England	Wall hanging, 1710 - 1720, England	Panel, 1701 - 1900, Guangzhou	Hanging, 1580 - 1590, England
Wall panel, 1710, France	Wall panel, 1710, France	Wall panel, 1710, France	Wall panel, 1710, France	Valance, 1560 - 1570, France
Wall panel, 1710, France	Wall panel, 1710, France	Wall panel, 1710, France	Wall panel, 1710, France	Wall panel, 1710, France

Figure 6. Results for "Wall Hanging".

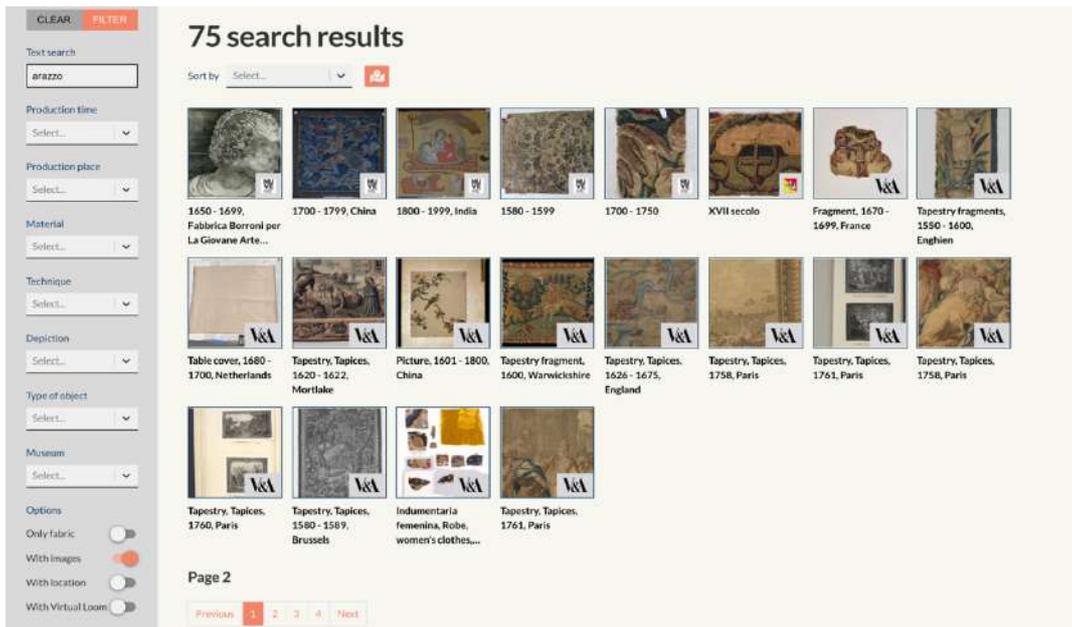


Figure 7. Results for "Arazzo"

Figure 8 shows the result when the word "Valencia" is entered into the text search. In Figure 9 we show the first result in the list.

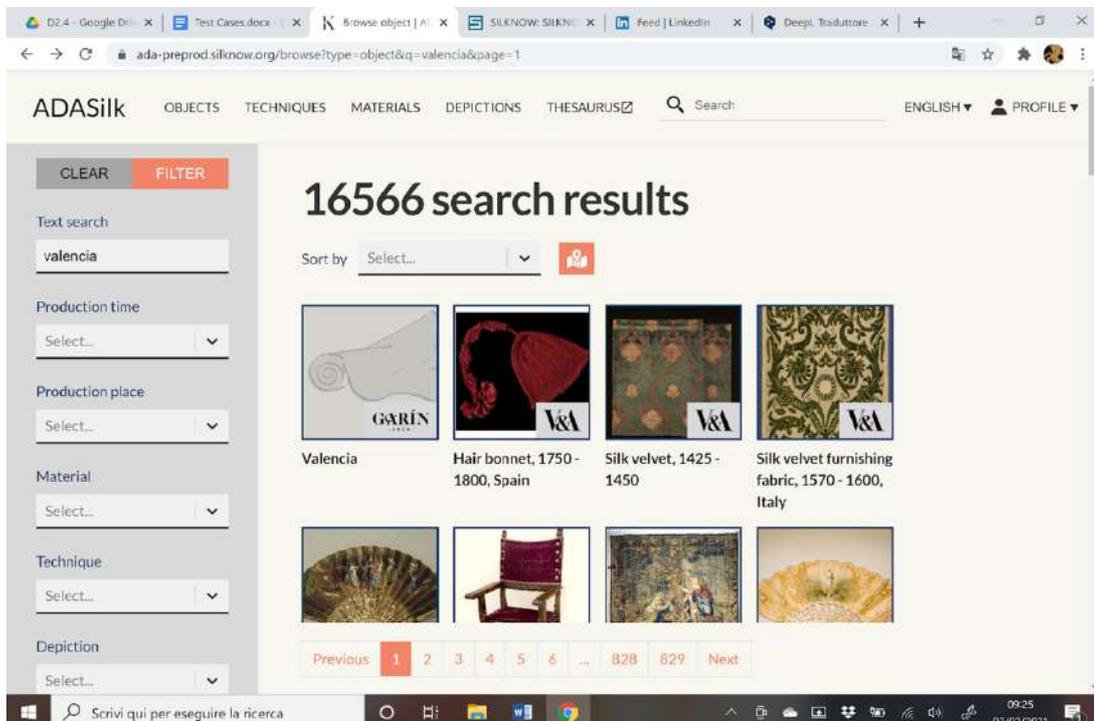


Figure 8. Results for "Valencia"

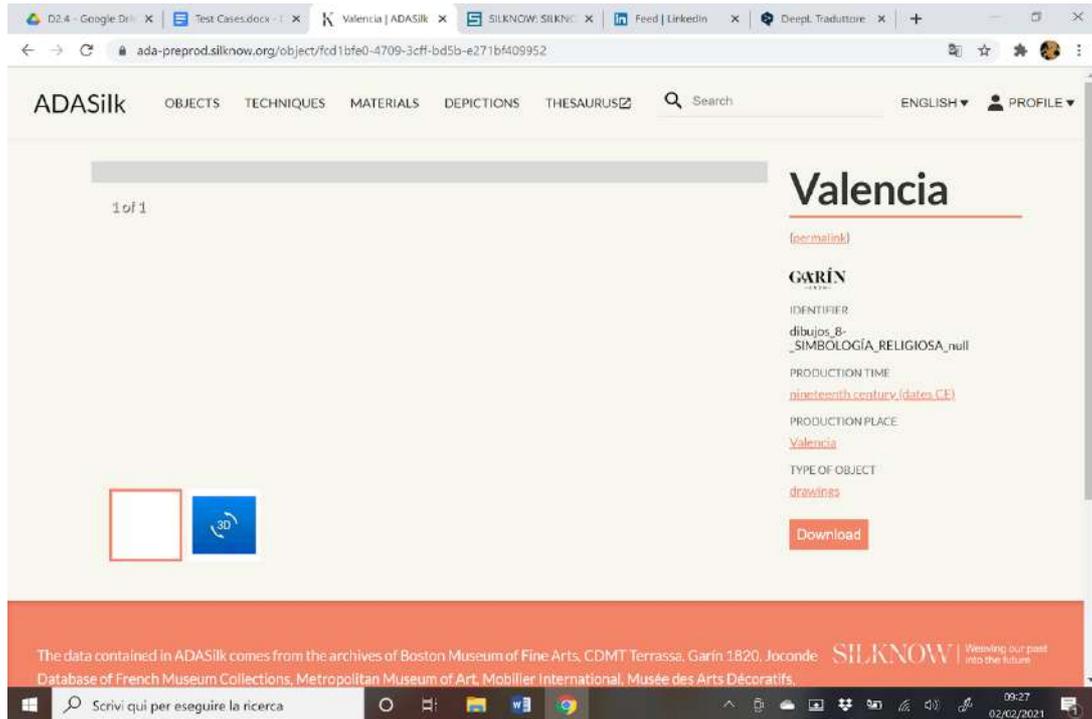


Figure 9. Details on the first result of Figure 8.

The result is incomplete, there is no description and there are no pictures.

Figure 10 shows the sixth result in the list. From what appears in the resulting object, it is not possible for a domain expert to grasp the significance of this result. For example, it might be important to an art historian to report on a similar work in Valencia, but in fact any useful information about the object can only be found by accessing KG via the permalink. Also, the word "Valencia" only appears in the description (see the permalink) because it is compared to another similar example kept in Valencia. The result does not seem to be pertinent to the search for Valencia.

Figure 11, Figure 12, shows other results from the list in Figure 8, nothing in these objects seems to be related to Valencia either. In the first one we did not find any text (or anything else like place of production, etc.) related to Valencia, while in the second result the word Valencia appears only in the bibliography.

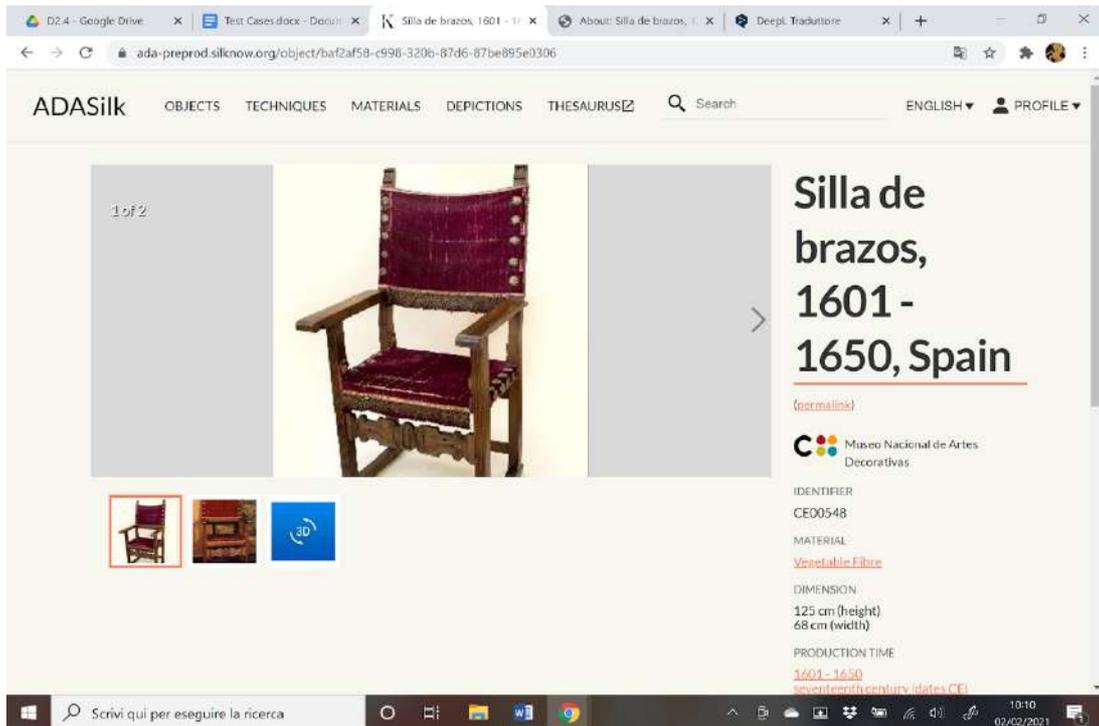


Figure 10. Details on the sixth result of the list in Figure 8.

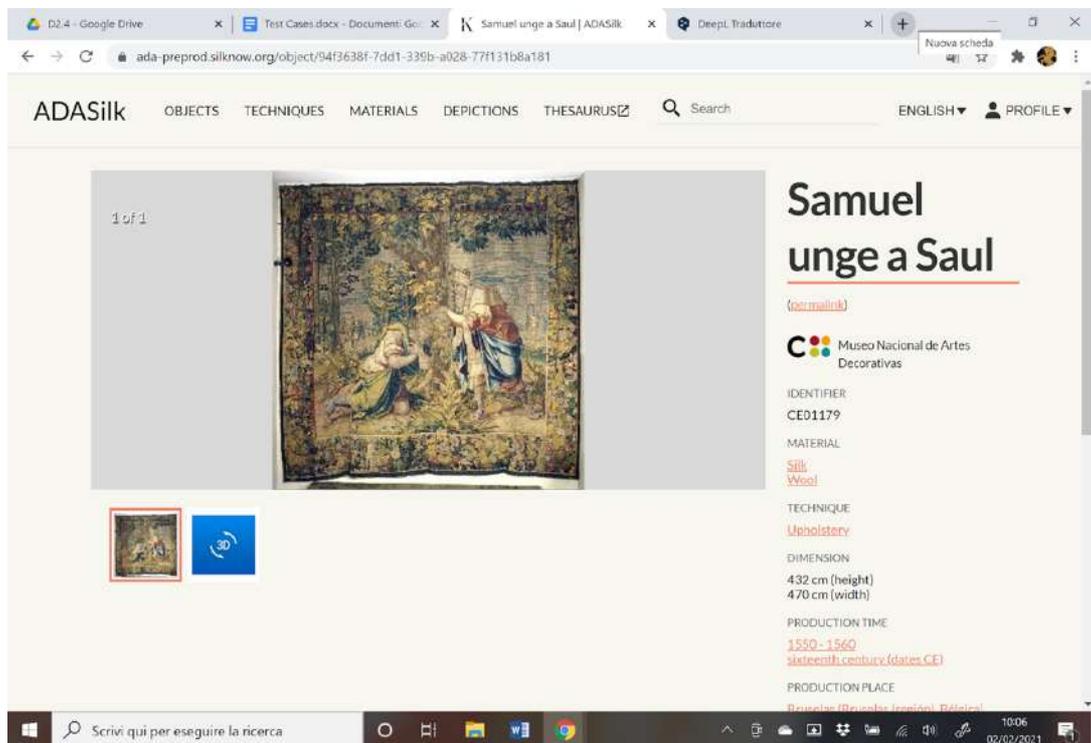


Figure 11. The seventh result of the list in Figure 8.

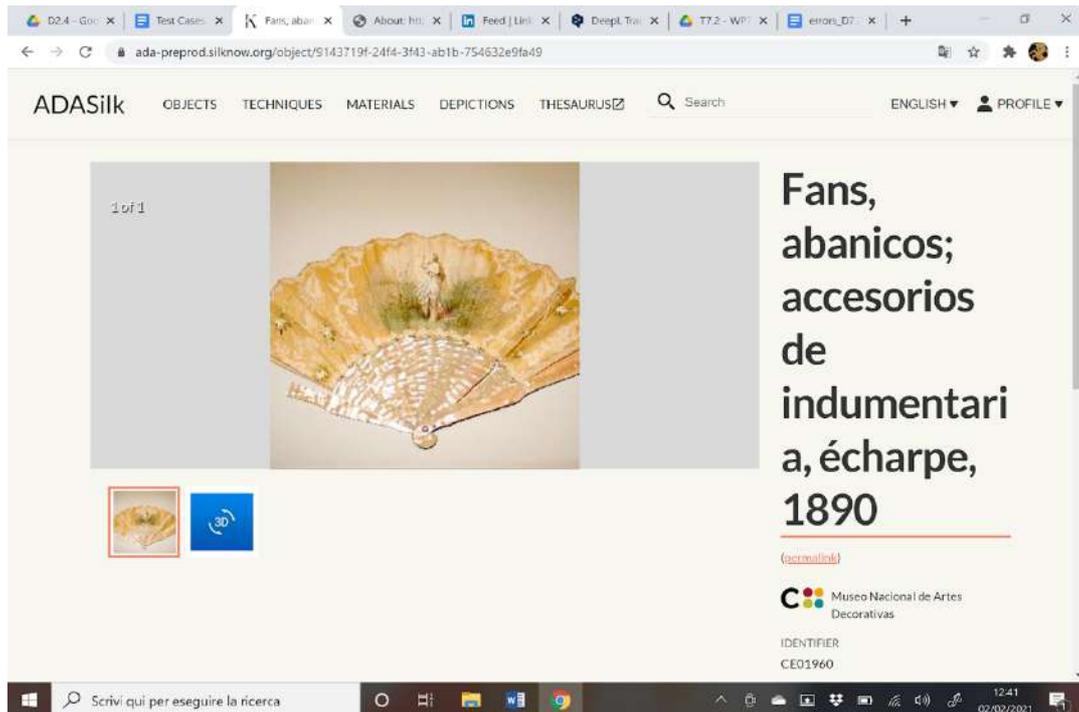


Figure 12. The eighth result in the list of Figure 8

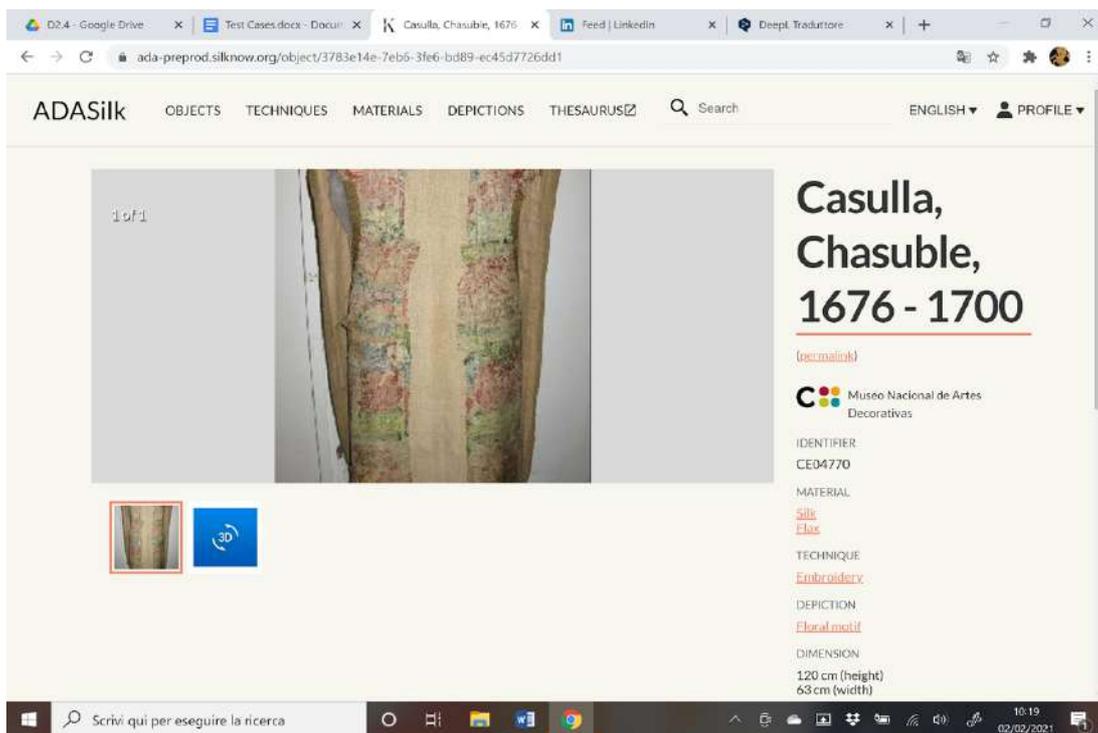


Figure 13. The eleventh result of the list of Figure 8.

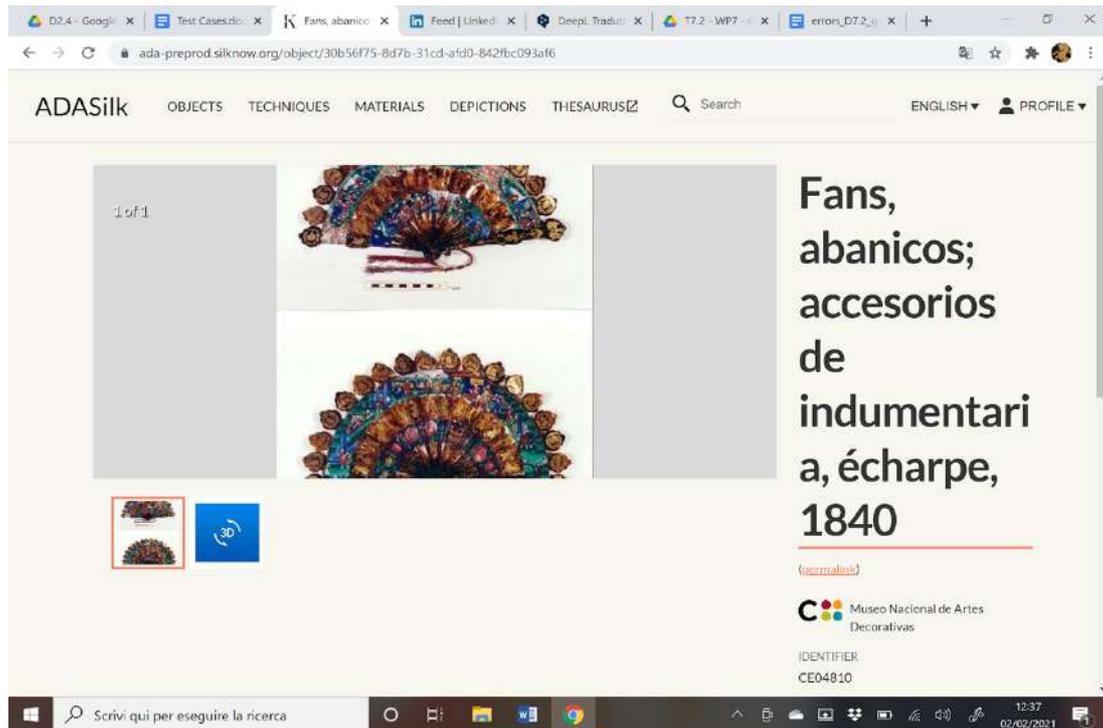


Figure 14. The twelfth result of the list of Figure 8

Figure 15 and Figure 16 show the list of results when "fan" or its plural "fans" is entered. There are no results for the first search, while there are 18911 results for the second search. This seems strange for the same reason as mentioned before (see Figure 4 and Figure 5). The search engine also displays partial results. We think that this different behavior needs to be investigated and possibly fixed.

In addition, the search for "fans" also returns some incorrect objects such as parasols and footstools. An example of "parasol" is shown in Figure 17, where it can be seen that the word "fans" is mentioned in the history of the use of parasols within the description of the dataset. The same for the other false result is shown in Figure 18.

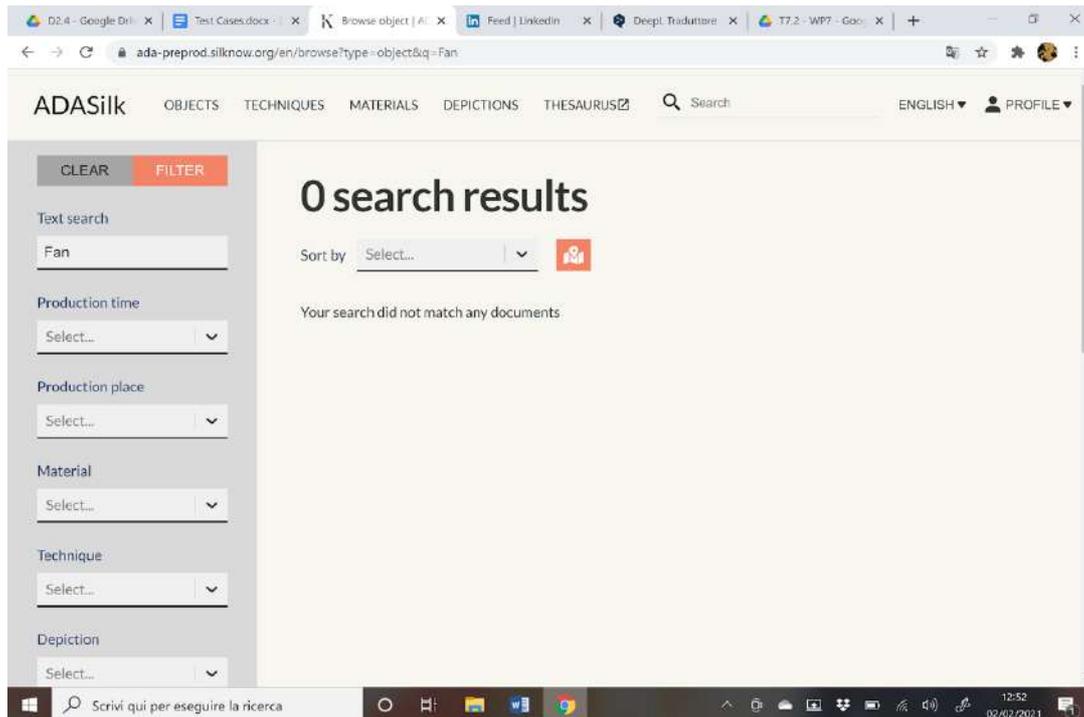


Figure 15. Results for "Fan".

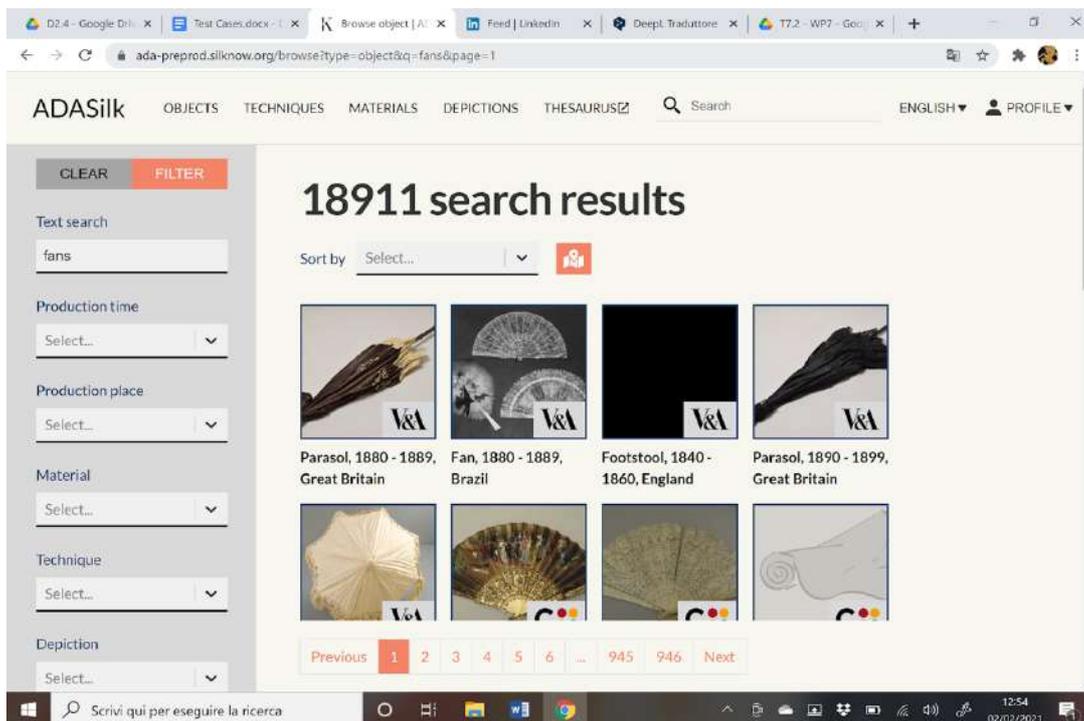


Figure 16. Results for "Fans".

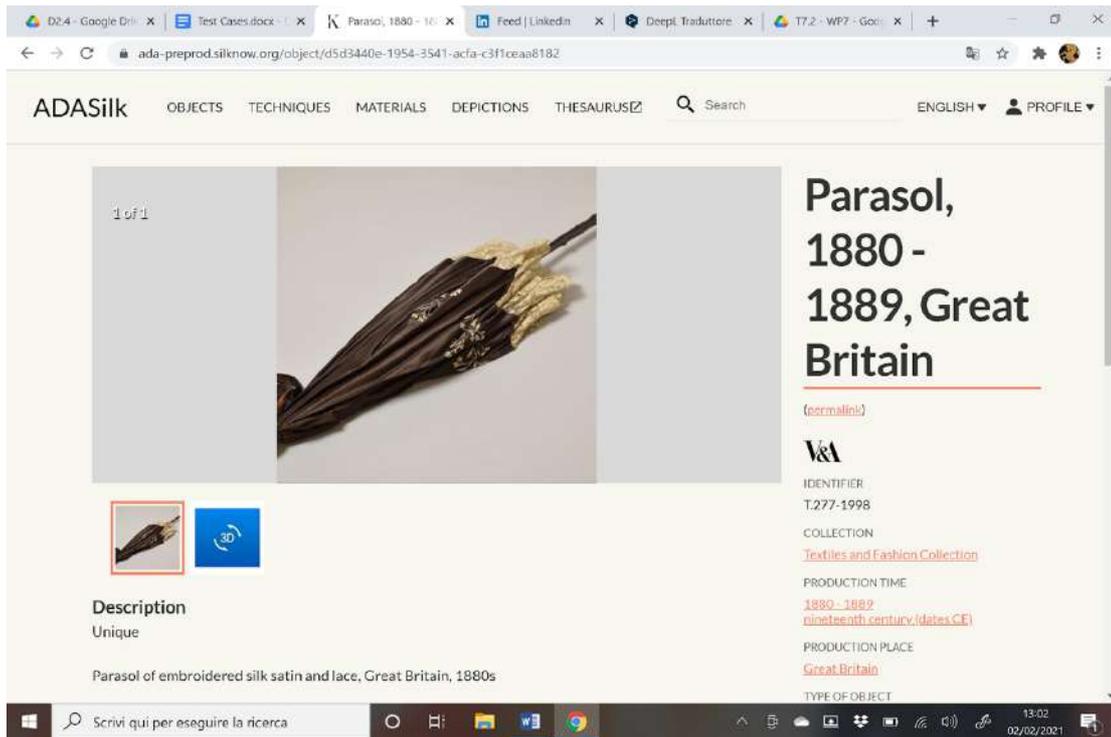


Figure 17. A specific object (Parasol) in the list of "Fans".

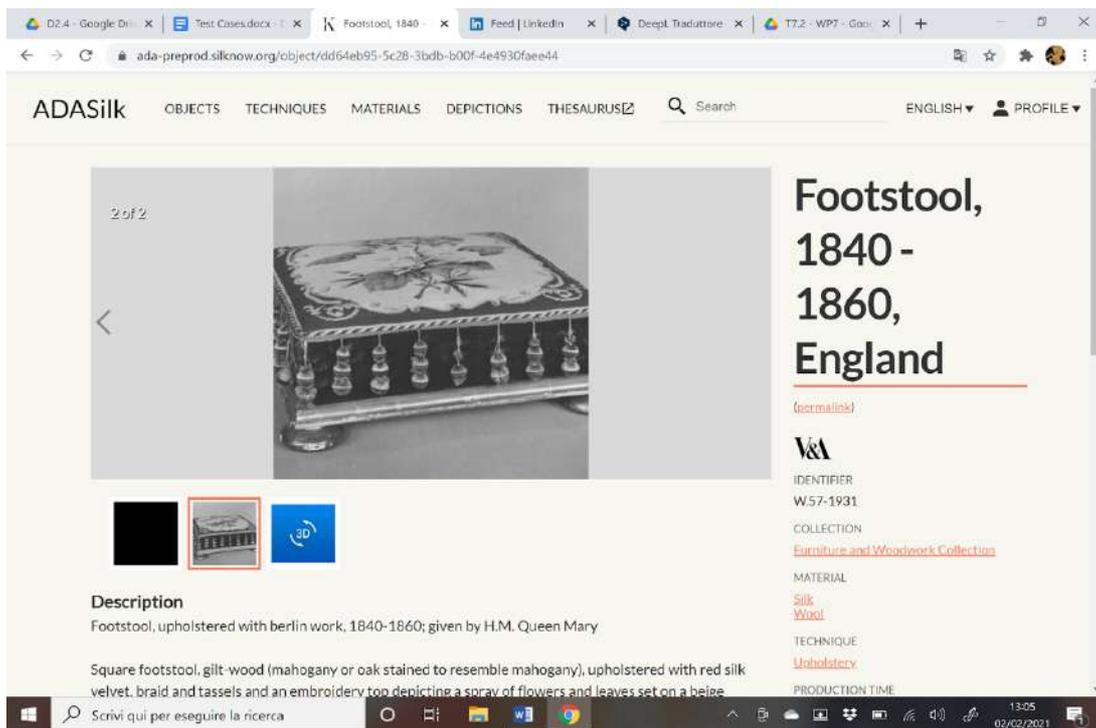


Figure 18. A specific object (Footstool) in the list of "Fans".

6.1.2. Repeated objects

Sometimes we found among the search results the same object several times with different identifiers. Three examples are shown in the following figures.

ADASilk OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS

1 of 1

Alb decoration, 1848, England

[View image](#)

V&A

IDENTIFIER
T.3021-1989

COLLECTION
[Textiles and Fashion Collection](#)

MATERIAL
[Gold thread](#)

TECHNIQUE
[Brocaded Weaving](#)

PRODUCTION TIME
[1848](#)
[19th century \(United Kingdom\)](#)

PRODUCTION PLACE
[England](#)

TYPE OF OBJECT
[textile fabric](#)

[Download](#)

Description

Alb decoration of woven silk brocaded with purple and gold threads.

Alb decoration of woven silk brocaded with gold threads, designed and made by A.W.N. Pugin, England, 1848.

The data contained in ADASilk comes from the archives of Boston Museum of Fine Arts, CDMT Terraza, Garin 1820, Jacondi Database of French Museum Collections, Metropolitan Museum of Art, Mobilier International, Musée des Arts Décoratifs, Musée des Tissus, Paris Musée, Red Digital de Colecciones de Museos de España, Rhode Island School of Design, Sicily Cultural Heritage, Smithsonian American Museum of Natural History, Victoria and Albert Museum.

SILKNOW Weaving our into the future

Figure 19. An example of repeated results – first.

ADASilk OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS

1 of 1

Alb decoration, 1848, England

[View image](#)

V&A

IDENTIFIER
T.3021-1989

COLLECTION
[Textiles and Fashion Collection](#)

MATERIAL
[Gold thread](#)

TECHNIQUE
[Brocaded Weaving](#)

PRODUCTION TIME
[1848](#)
[19th century \(United Kingdom\)](#)

PRODUCTION PLACE
[England](#)

TYPE OF OBJECT
[textile fabric](#)

[Download](#)

Description

Alb decoration of woven silk brocaded with purple and gold threads.

Alb decoration of woven silk brocaded with gold threads, designed and made by A.W.N. Pugin, England, 1848.

The data contained in ADASilk comes from the archives of Boston Museum of Fine Arts, CDMT Terraza, Garin 1820, Jacondi Database of French Museum Collections, Metropolitan Museum of Art, Mobilier International, Musée des Arts Décoratifs, Musée des Tissus, Paris Musée, Red Digital de Colecciones de Museos de España, Rhode Island School of Design, Sicily Cultural Heritage, Smithsonian American Museum of Natural History, Victoria and Albert Museum.

SILKNOW Weaving our into the future

Figure 20. An example of repeated results – second.

ADASIik OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS

1 of 1



Alb decoration, 1848, England

(original)

V&A

IDENTIFIER: T.302M-1989

COLLECTION: [Textiles and Fashion Collection](#)

MATERIAL: [Gold thread](#)

TECHNIQUE: [Brocaded](#) [Woolen](#)

PRODUCTION TIME: [1848](#) [19th century \(dates CE\)](#)

PRODUCTION PLACE: [England](#)

TYPE OF OBJECT: [textile arts](#)

[Download](#)

Description
Alb decoration of woven silk brocaded with purple and gold threads.

Alb decoration of woven silk brocaded with gold threads, designed and made by A.W.N. Pugin, England, 1848

Figure 21. An example of repeated results – third.

ADASIik OBJECTS TECHNIQUES MATERIALS DEPICTIONS THESAURUS

1 of 1



Alb decoration, 1848, England

(original)

V&A

IDENTIFIER: T.302K-1989

COLLECTION: [Textiles and Fashion Collection](#)

MATERIAL: [Gold thread](#)

TECHNIQUE: [Brocaded](#) [Woolen](#)

PRODUCTION TIME: [1848](#) [19th century \(dates CE\)](#)

PRODUCTION PLACE: [England](#)

TYPE OF OBJECT: [textile arts](#)

[Download](#)

Description
Alb decoration of woven silk brocaded with purple and gold threads.

Alb decoration of woven silk brocaded with gold threads, designed and made by A.W.N. Pugin, England, 1848

Figure 22. An example of repeated results – fourth.

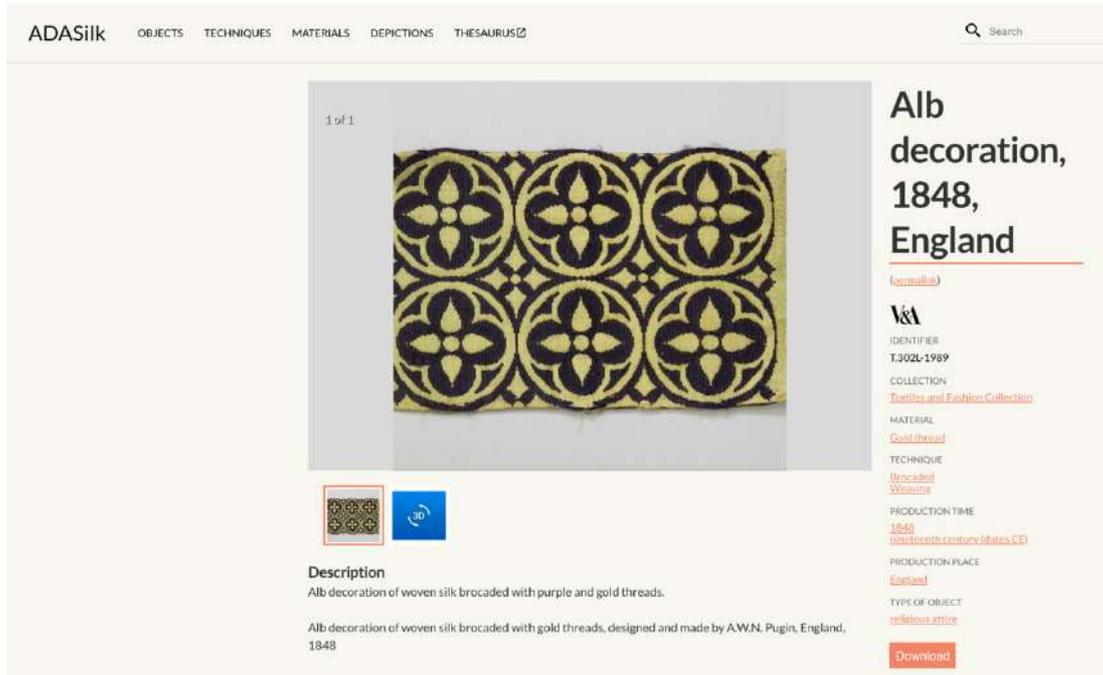


Figure 23. An example of repeated results – fifth.

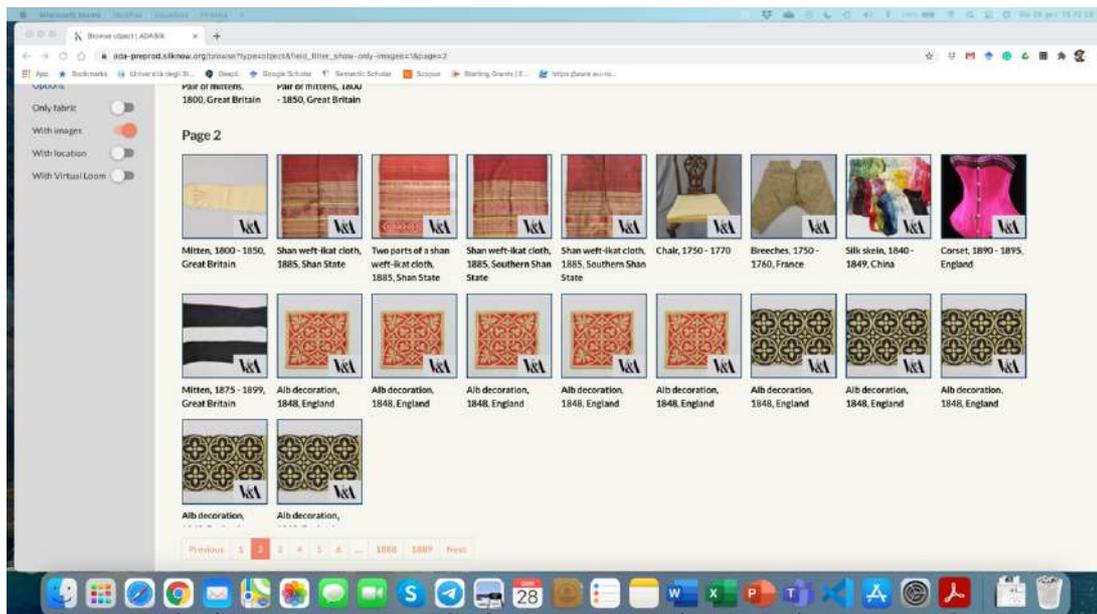


Figure 24. Second example of repeated results. Overview (list of results when selecting “costume accessories” in the Type of Objects’ filter.

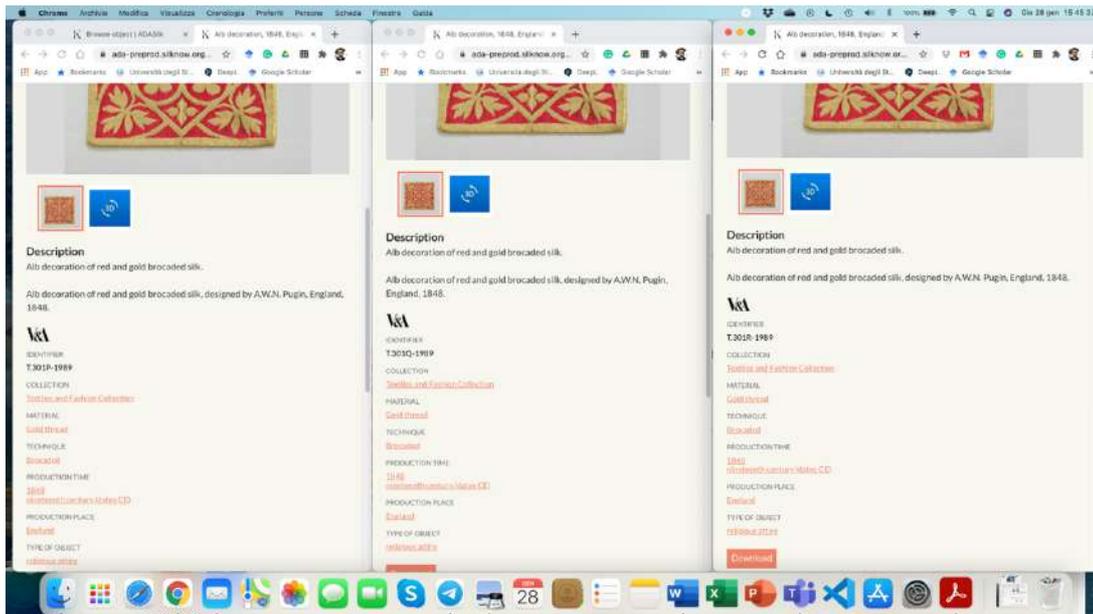


Figure 25. Details on repeated results.

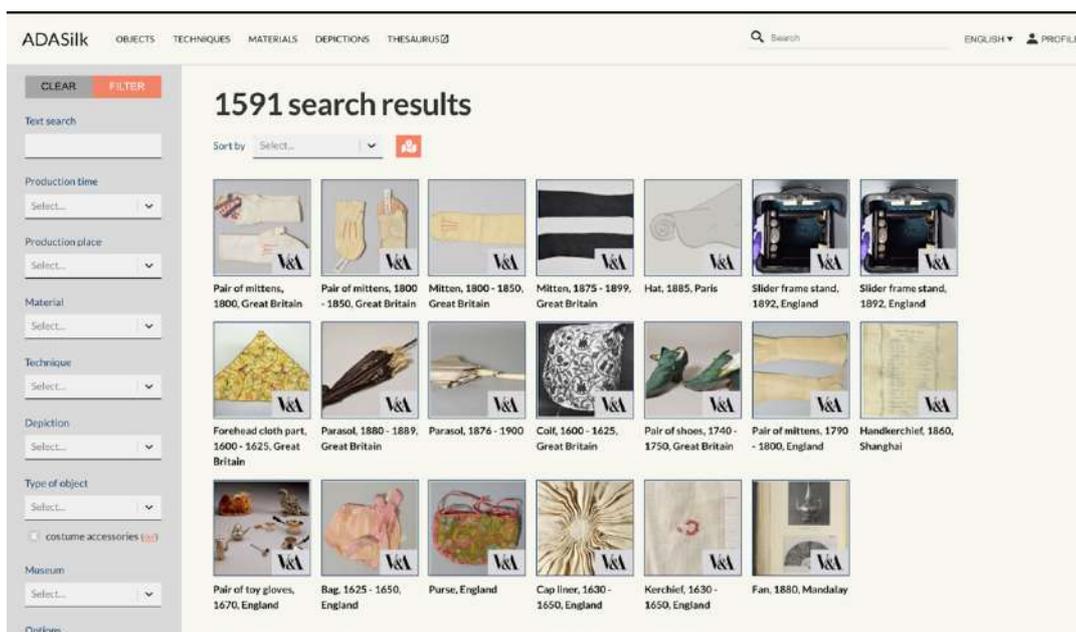


Figure 26. Third example of repeated results.

6.1.3. Results not adequately filtered.

Several times we found that the results were not sufficiently filtered, in the sense that the results list showed objects that did not match the search.

Figure 27 shows the results when one wants to find information about brocade chasubles and enters the word "chasuble" in the Text Search and "brocade" in the technique filter.

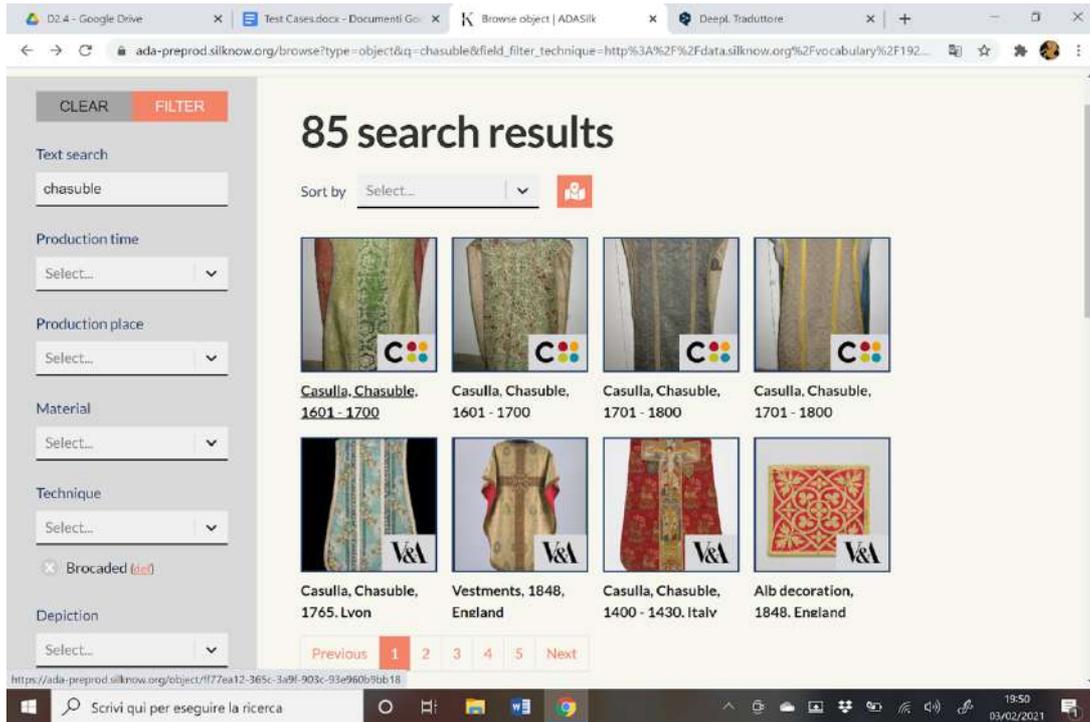


Figure 27. An example of results not adequately filtered.

In the list of results we find brocade chasubles, but also objects that are not chasubles (but made of brocade).

6.2. Stress test

Although not directly related to testing the adequacy of the data in the search results, we performed a stress test to complete the process of improving the ADASilk application. We found that in some cases the application could not handle a load of concurrently connected users.

We noticed that when more than five people are connected at the same time and ADASilk is displaying the main page, if one clicks on any button or enters a search, the loading process starts but then stops indefinitely. Most likely, the fact that testing was performed during the development, integration and maintenance of the software modules prevented the proper functioning of the web application as the developers were constantly working. This issue will be considered and fixed once the software is completed.

6.3. Dead page or invalid redirects

For the same reasons as the stress test, we also checked for dead pages and invalid redirects and found no problems.

7. RESULTS OF THE TESTING ACTIVITIES

7.1. List of errors

From the analysis of all tests performed with ADASilk, we summarize the most common errors or problems in the following list:

- **Unexpected results:** This type of error affects free text search. Every word you enter in Text Search is searched in the entire text of the museum entry in the database. This

leads to results that domain experts may not see results as matching the search performed. We have communicated this fact to other partners and developers and there is an ongoing discussion.

- Repeated objects: Some results appear to be repetitive - same description, same objects, but different identifiers.
- Results not adequately filtered: This is somehow also related to the first of the above bugs, plus some searches return results that do not match the actual search, in the sense that the results contain objects that have nothing to do with the original search.
- Returned numbers of items are different when we use a translated version of the same term. Ex: giubba, doublet.
- Inserting a word followed by a space returns zero results. A user may accidentally insert a space after a word. The web application should reply that the space was not inserted or report the error, otherwise the user will not understand why zero results.
- Some combined words do not return results at all, when they reasonably should. Others, which are incorporated in the thesaurus, such as "wall hanging", do work properly. This could be considered an user experience issue.
- Incomplete descriptions: Some resulting objects do not have complete descriptions (see Figure 28 for an example of two objects, the first has no field description while the second has one). So from the point of view of a domain expert, but also from the point of view of an ordinary citizen, this may be unusable. It is worth considering whether the problem lies in the search process or in the description of the object in the record provided by the museum.
- For some result objects, no images are shown, although we can see that images are available if we click on the permalink and go to the KG page. For example, when searching for Italy in the Production place, one of the resulting objects "CattedralePA1" (which is the value of the identifier), the description in the KG, accessible through the permalink, shows the attribute P138_has_representation with a value indicating the image "Splendori di Sicilia - 35.jpg".
- In some searches for dates it was not possible to enter the year in the filters, but only the century, so if the user wants to enter the year he/she has to use the text search; in the two cases different results are obtained.

7.2. Needed improvements and how they have been solved

From the above list it is evident that the main work that remains to be done is to match the KG with the data from the museums' archives and to complete the description of the objects. Some of these problems have already been solved. As mentioned above, the software is still under development and integration. It should be noted that the database of records from the museums is very large, so the work of mapping and inserting all the descriptions is very demanding and takes a long time. This solves problems with unexpected results, repeated results, incomplete descriptions and insufficiently filtered results and everything related to how the datasets from the museums are constructed.

Images, synonyms, dates, and other minor technical issues will be resolved during final system integration and evaluated in Task 7.6.

Additionally, all the issues highlighted in the above list will be solved in future iterations of the software module development and finally tested during the final validation at the end of WP7.

As for the first of the highlighted problems in the list, this will be further investigated and then solved with the full integration of the text search and image processing modules.

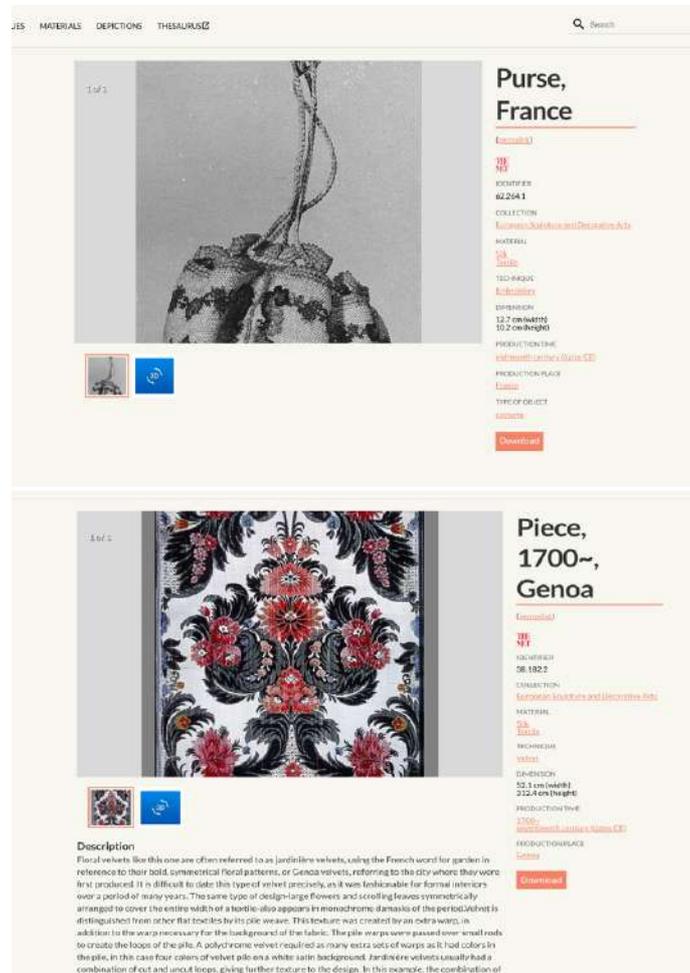


Figure 28. Objects with incomplete and complete description.

8. CONCLUSION

The purpose of Task 7.2 is to perform a test on real-world scenarios of the SILKNOW system (ADASilk), which is currently under development, to investigate whether the results of a simple search are meaningful and consistent with the project goal.

The tests were performed by a team consisting of ICT and CH experts using well-known software testing techniques. Although some modules are not yet fully integrated, it can be stated that the system provides the expected results. Only a few bugs were found from a technological point of view. Most of the errors are semantic in the sense that the domain experts consider some results as not significant. However, we expect that after the integration of the text and image prediction module, all these problems will be solved.

It should be noted that the object descriptions reported by the museum catalogs are not yet complete, which affects the functionality of ADASilk. As the integration and development work is still ongoing, this is not a problem or bug and the descriptions will all be completed soon.

SILKNOW

For the same reason, any results that are not complete will be corrected as soon as possible. Moreover, all repeated results will be corrected after double-checking with records from museums. In many cases the repeated results are due to the way in which museums, e.g. Victoria and Albert Museum, present objects in their collection.

9. REFERENCES

- [1] Sommerville, Ian. Software Engineering 10. Harlow: Pearson Education Limited, 2016.
- [2] Luo, Lu. "Software testing techniques." Institute for software research international Carnegie Mellon University Pittsburgh, PA 15232.1-19 (2001): 19.
- [3] Abran, Alain, et al. "Software engineering body of knowledge." IEEE Computer Society, Angela Burgess (2004).
- [4] SILKNOW Deliverable 6.5 - Integrated System.

