

## [2] Fundamental Functions and Required Features of the System

### [2.1] Use cases

It was decided to formulate the fundamental functions of the system by defining an all encompassing set of situations faced by users or 'use cases'. These use cases are intended to define the ways a user would want to interact with the system, the situations within which they would wish to do this and how they would like the results presented. The use cases were requested from all partners within the project and also from a selection of other interested parties as defined on the 'twiki' (<http://www.bernstein.oeaw.ac.at/twiki/bin/view/Main/TWikiGroups>).

The use cases have been split into the following sections:

- Integration - use cases related to data sources and methods of integration of these sources.
- Search - types of search that the system will be required to support
- Display - display mechanism and formats required
- Statistics - various methods for processing the results of a search to provide clarity or further information

Most of the use cases submitted are included in this integration blue print document. Those that could not be supported were removed. Those that were removed were: searches for medieval symbolism and searches for watermark counter-marks: this information is not available in the current databases and provision to add this data is not covered in any of the Bernstein work packages.

In addition, it is important to stress that some of the included use cases will only be supported if the underlying data is available in the original databases, for example, in use case 20, searching for water mark twins is only possible if the twin information has been recorded in a field in the database.

This project is intended to integrate each of the watermark databases, to link to other related databases, to provide a suitable mapping between the same terms in different databases independent of language and to integrate useful tools to provide expertise.

To fulfill all aspects of all use cases, the watermarks would need to be entirely re-catalogued; this is outside the scope of the project. In recognition of this need, as highlighted in the use cases, a new method of watermark description will be developed - the 'component model'. This is discussed further in section [3.2].

### [2.1] Use Case List

#### The Integration should:

- 1 Provide results at a reasonable speed
- 2 Be accessible to the world outside Bernstein via a machine2machine link
- 3 Perform image scaling
- 4 Have appropriate error treatment
- 5 Have a ISTC link
- 6 Have a Briquet Search
- 7 Use a Bernstein Code for each watermark

**The Search should:**

- 8 Provide Multilingual access
- 9 Be reliable
- 10 Provide combined searches with logical operators
- 11 Avoid ambiguity due to subjective classification
- 12 Account for synonymy – database side
- 13 Account for synonymy – user side shape ambiguity
- 14 Search for specific elements within the databases
- 15 Search for elements within a watermark
- 16 Inform users if they are searching for watermarks not in the databases
- 17 Allow for the heterogeneous content provided by the different databases and draw the user's attention to this
- 18 Use terminology that is understood by users
- 19 Be able to combine other factors such as measurements to narrow down search results
- 20 Be able to search for watermark twins
- 21 Be able to search for bibliographical references

**Display should:**

- 22 Incorporate all functions into one workspace
- 23 Provide multilingual interfaces
- 24 Provide a method for a user to store results

**The statistics that should be provided are:**

- 25 General statistics on the results of a search
- 26 The facility to export results
- 27 To provide information on the dating of paper
- 28 To provide indications of document authenticity
- 29 Plotting results sets on to maps
- 30 Plotting bibliographic references on a maps

**Off Line Tools:**

- 31 Image Processing

**[2.3] Search / Display**

At present there is three databases which provide textual search (WILC, Piccard, NIKI) and three which provide textual hierarchy/ image based browsing (WZMA, WILC, Piccard). It is felt that successful implementation of this project will need to offer both of these facilities. It has not, as yet, been discussed and decided as to which of the fields in the present databases will be made available to search from within the Bernstein workspace.

It has been decided that the architecture would need to support three types of searches:

- **Text.** Searches of varying complexities, combinations of visual and quantifiable parameters combined via boolean operators the user will be able to perform complex textual searches which can be linked by logical operators – use case 10.
- **Constructed image.** It will be possible to create a representation of the query watermark for use in finding similar or exact matches in the target databases.

- **Combined.** Using text to construct an image. Text based queries will either be mapped silently into this query description as above, or used as input to automatically construct an image in the graphical representation, which can be viewed by the user before being submitted as a search.

From the use case it is important that the searches provide multilingual access as discussed in [2.5]. Many of the other use cases are self explanatory, for example, 10 which provides combined searches with logical operators' and 22 which incorporates all functions in one work space and so on; more detail is provided below for more complex use cases.

Use cases 11, 12 and 13, all discuss the problems of ambiguity of shape and synonyms, for example, a user may be searching for a shape which he thinks is a ribbon. In one of the databases it might be classed as a rope (similar shape) whereas in another it might be classed as a snake. In the Bernstein workspace there should be a way to map these synonyms and remove the ambiguity. The terminology used in the Bernstein Project will be at a suitable level so that it is of use to watermark experts and still understandable to novel users as highlighted by use case 18.

In an additional search, the user may want to look at all watermarks that contain this ribbon / snake / rope element but this could be more likely to occur within a watermark that contains more prominent elements such as a snake with a dagger or a rope with an anchor. In the present databases these watermarks could be classified under the motifs: anchor or dagger, not ribbon / snake / rope. Therefore, it is impossible to search directly for the ribbon or snake or rope. This issue was highlighted in use case 14 – a desire to search by elements. This issue is reconfirmed in use case 15 – if a user has an incomplete watermark, a common occurrence for the art historian, he may only have a section of the watermark with a snake on it but the dagger is missing. The watermark may be classed as a dagger watermark; therefore, he has no way to find it as he does not know that the dagger exists.

In use case 16 the user wishes to know if his watermark does not exist in the database, for example if a user searches for all watermarks that contain both a bull's head and an anchor, he wants to be told that there is no record of this watermark rather than be given all bulls head and all anchor watermarks. This relates to the boolean operator as discussed in use case 10.

Use case 21 describes two methods for accessing bibliographical data. In the first the user wishes to input a search term within one of a number of defined fields, for example, author, title, publisher, subject, classification, he expects to receive data relating to paper history and watermarks. In the second case a user already has a list of watermarks from a search and is interested in bibliographical context relating to these watermark locations or motifs. The results for both of these methods would be displayed in a sortable list.

The user will wish to store their results – use case 24. This use case means that the Bernstein workspace will support the ability for the user to log in to the workspace where they will locate their own area which will contain a record of information they have stored from past visits.

## [2.4] Statistics

The use cases in the statistic section revolve primarily around post-processing of search results. For each statistical package, information is derived from a search and fed into appropriate statistical processes and data added, where appropriate, from other sources such as the bibliographic databases. The results will be displayed within the Bernstein workspace.