



# **Imageline Scope and Feasibility Report**

prepared by:

**Branka Kosovac**

for

**IFLA Section of Art Libraries**

**Vancouver, B.C., July 2002**

<u>EXECUTIVE SUMMARY</u> .....	4
<u>WHAT</u> .....	4
<u>HOW</u> .....	6
<u>WHO</u> .....	8
<u>1 INTRODUCTION</u> .....	11
<u>1.1 Background</u> .....	11
<u>1.2 Scope and purpose of the study</u> .....	11
<u>2 INTERESTED PARTIES AND THEIR NEEDS</u> .....	11
<u>2.1 End users</u> .....	11
<u>2.2 Member libraries as end user intermediaries</u> .....	13
<u>2.3 Member libraries as collection providers</u> .....	15
<u>2.4 Other collection providers</u> .....	15
<u>3 STATE OF PRACTICE</u> .....	15
<u>3.1 Access to images on the Internet</u> .....	15
<u>3.2 Existing sources of images</u> .....	16
<u>3.3 Wider context</u> .....	18
<u>3.3.1 Integration and interoperability</u> .....	18
<u>3.3.2 Standardization</u> .....	19
<u>4 SCOPE SUGGESTIONS</u> .....	20
<u>4.1 Objective</u> .....	20
<u>4.2 Precedents</u> .....	20
<u>4.3 Strategy</u> .....	21
<u>4.4 Target audience</u> .....	21
<u>4.5 Information coverage</u> .....	22
<u>4.5.1 Disciplines</u> .....	22
<u>4.5.2 Geographical coverage</u> .....	23
<u>4.5.3 Temporal coverage</u> .....	23
<u>4.5.4 Types and accessibility of resources</u> .....	23
<u>4.5.5 Level of granularity</u> .....	24
<u>4.5.6 Sources of information</u> .....	24
<u>4.5.7 Level of “difficulty”</u> .....	24
<u>4.5.8 Languages</u> .....	25
<u>4.5.9 Evaluative comments</u> .....	25
<u>4.5.10 Criteria for selection</u> .....	25
<u>5 PROCESS</u> .....	26
<u>5.1 Opportunities, obstacles, and risks</u> .....	26
<u>5.2 Providing content</u> .....	27
<u>5.2.1 Discovering Resources</u> .....	27
<u>5.2.2 Evaluating and selecting</u> .....	28
<u>5.2.3 Controlling editorial access</u> .....	28
<u>5.2.4 Submitting</u> .....	28
<u>5.2.5 Coordinating additions</u> .....	29
<u>5.2.6 Storing/persisting resource descriptions</u> .....	29
<u>5.3 Providing access to content</u> .....	29
<u>5.3.1 Describing resources</u> .....	29

5.3.2	<a href="#">Data structure</a>	30
5.3.3	<a href="#">Vocabulary</a>	32
5.3.4	<a href="#">Cataloguing rules</a>	34
5.3.5	<a href="#">Data entry tool</a>	34
5.3.6	<a href="#">Search mechanism</a>	34
5.3.7	<a href="#">Browsing</a>	35
5.3.8	<a href="#">User interface</a>	36
5.4	<a href="#">Collection Management and Maintenance</a>	36
5.4.1	<a href="#">Updating</a>	37
5.4.2	<a href="#">Obtaining and incorporating user feedback</a>	37
5.4.3	<a href="#">Link checking</a>	37
5.4.4	<a href="#">Technical maintenance and support</a>	38
5.5	<a href="#">Planning and Control</a>	38
5.5.1	<a href="#">Project management and coordination</a>	38
5.5.2	<a href="#">Communication and meetings</a>	39
5.5.3	<a href="#">Promotion, dissemination, and reporting</a>	39
5.5.4	<a href="#">Training</a>	40
5.5.5	<a href="#">Quality control</a>	40
5.5.6	<a href="#">Monitoring</a>	41
5.5.7	<a href="#">(Re-)evaluation</a>	41
5.5.8	<a href="#">Intellectual property rights</a>	41
5.5.9	<a href="#">Future proofing</a>	42
6	<a href="#">Technical Options</a>	42
6.1	<a href="#">Summary of requirements for <i>Imageline</i> hardware and software</a>	42
6.2	<a href="#">Hardware and operating system OPTIONS</a>	42
6.3	<a href="#">Options</a>	43
6.4	<a href="#">Evaluation Criteria</a>	44
7	<a href="#">ORGANIZATIONAL MODEL</a>	44
8	<a href="#">BUSINESS MODEL AND SUSTAINABILITY</a>	45
9	<a href="#">SUMMARY OF EXPENSES</a>	46
	<a href="#">REFERENCES:</a>	47
	<a href="#">APPENDIX 1: METHODOLOGY</a>	51
	<a href="#">WHAT</a>	51
	<a href="#">HOW</a>	51
	<a href="#">WHO</a>	52

## EXECUTIVE SUMMARY

This document reports the results and findings of a study that assessed the feasibility of *Imageline*—an international Web gateway to be collaboratively developed by the members of the IFLA Section of Art Libraries. The purpose of the envisioned gateway is to provide quick and efficient access to sources of art images, selected based on quality and relevance to the primary audience, and to prevent the duplication of efforts by the IFLA Section of Art Libraries members towards satisfying this need.

The summary provides a quick overview of facts, options, and recommendations. The main body of the report details the process, reasoning, and criteria used to select between the options. The discussion follows a methodology designed for potential reuse in other domains. The background information used to support the recommendations is provided in the appendices.

### WHAT

An analysis of user needs and the state-of-practice led to the following conclusions:

The **need for a new service exists** because:

- there is no single point providing quick and efficient access to sources of art images;
- there is no limited set of complementary services providing that function;
- efforts currently invested into creating lists of collections are substantial, costly, and redundant;
- results of these efforts are of little value to the users.

Looking from a larger perspective, however, the feasibility of a traditional gateway cannot be guaranteed in a longer run. Although the value of subject gateways has been unanimously acknowledged, the fast paced Internet environment raises general concerns regarding the longevity of the concept in its existing form and imposes the need for the development of future-proofing strategies. Recommended **future-proofing** strategies include: factorization across multiple dimensions allowing modular and incremental development, use of standards, regular re-evaluation of the continuing value to the user, continuous watch of related new technologies and initiatives and evaluation thereof in terms of their potential use in the gateway and vice versa, partnership with related projects. The current developments worth watching include primarily initiatives related to harvesting metadata from original sources, item level access to either virtual or physical mega-collections of art images, hybrid libraries and clumps.

The **primary audience** whose needs should be considered first in defining the scope and features for the gateway should match the sum of the primary audiences of the IFLA Section of Art Libraries members. The assumption, which should be verified by the membership, is

that this category encompasses: researchers, teachers, students, and practitioners in the fields of art, design, and architecture, museum curators, and librarians. The gateway should also be available to the general public, including art lovers, researchers, students and teachers from other fields, publishers, and anyone else who may want to explore art images, since these groups typically represent the secondary audience of art libraries and a considerable portion of their actual users.

The needs of these users are extremely diverse; breadth and depth are equally required. Undergraduate students and the general public are served the best by the existing services.

**Major gaps and missing functionality** required to meet the needs of researchers, teachers, and practitioners encompass: truly international scope, uniform coverage of periods and styles, deep coverage of topics, quality of selection, descriptions, maintenance procedures, and accessing mechanisms, standardization, and interoperability with other services.

The **scope in terms of disciplines** should be guided by the scopes of the member libraries. Fine arts in the strictest sense and architecture are recommended as starting points. These disciplines are most widely covered by the existing services; however, it is assumed that they represent the intersection of member libraries' scopes and must be used as a central core covered by a high-quality gateway. Additional modules, which can be developed in parallel, if the sufficient need and resources exist, or in subsequent phases should be identified and prioritized starting from the suggestions in the section 4.5.1 of the report.

There should be no restrictions in terms of **geographical, temporal, and language coverage**; a well-balanced representation of regions, periods, and languages should be aimed for. The possibility to limit searches to resources in certain languages should be provided.

Due to the frequent need for good quality graphics, there should be no **constraints in terms of technical requirements**, such as broadband connection or special viewers. However, any special requirements or features should be stated in the resource descriptions and accessibility guidelines, as prescribed by W3C, should be followed.

In terms of **types of collection**, sources providing free online access to images should be given preference and included first. Sources providing access to image descriptions and fee-based access to images should be considered next, if providing any additional value to the user. Offline digital and analogue collections and fully commercial services can be added in the subsequent phases, if satisfying the selection criteria and filling gaps in coverage. The description format should provide for the potential addition of different collection types in the future.

**Quality criteria** for selecting resources can be based on corresponding guidelines developed by other gateways. Examples are listed in Section 4.5.1. Project specific extensions can include the following. No types of sources of image collections should be *a priori* excluded from the scope. However, only those image collections that represent items from permanent third party collections assembled by museums or similar institutions should

be included. Well-defined collections based on a clear collecting principle, such as artist, style, culture, theme, should be preferred to collections based on owner or location.

## HOW

Resource descriptions should be stored and administered in a **central repository**. The alternative model, consisting of distributed catalogues cross-searched via broker-services from a single user interface, although attractive, is excluded for the following reasons: not enough content to justify a dedicated server at each node, difficulties involved in having control over databases on mixed-use servers owned by different member libraries, importance of synchronizing content across more than one dimension. The gateway should follow the model of a union catalogue or "physical clump" with the possibility to provide different views of its content and to be integrated into other services. A mirror site should be set up in an institution on the opposite hemisphere for faster access and as a backup in case of a breakdown.

An Intel Pentium based server running Linux operating system is recommended as a suitable **platform** that gives the best value for investment. Windows NT should be considered as a viable alternative only if a significant technical expertise is readily available exclusively for that platform. General **software** options include: develop from scratch using open source development tools, develop combining several open source components, partner with related research projects and use their tools, buy commercial off-the-shelf product tested on subject gateways, outsource to a commercial solution provider. These options vary widely in terms of the ratio between the initial financial investment and investment of effort, overall costs, and the time needed for implementation. The "build option" is cheaper, even including the cost of human work, and can be better moulded to the specific needs; buying is easier and faster to implement but more expensive and may require compromises in functionality. The use of research-developed tools can be convenient and inexpensive if suitable project can be identified, but it can make it difficult to stick to the schedule and ensure continuing support, once the funding for the research is over. The final choice of **technology** will depend on the availability of resources for system development and maintenance among the members and on the possibility to obtain external funding. As sufficient input information was not available at the time of writing, specific recommendations are not made but lists of requirements, options, and selection criteria are provided in Section 6.

A fully suitable **metadata schema** is not available. The recommendation is to develop an application profile based on Dublin Core Metadata Element Set, extended by elements from RSLP Collection Level Description and custom elements where required. The Dublin Core Collection Description Schema, scheduled for release in May 2002, should be reviewed and assessed for potential adoption once it gets published.

The key issues in defining the schema involve the relationship between the collection of art objects, collection of images representing items from that collection, and virtual collection and agents related to each, the meaning of the term “subject” in this particular context, and synchronization with schemas for describing related resources, such as art images and art history texts. It should be possible to extract a subset of basic unqualified metadata elements needed for integration with other services. The schema will also need to include some housekeeping, project-specific fields, which can be fully defined at a later date.

A widely used, multilingual, domain specific **vocabulary** would be a preferred option for the data content standard. Two closest solutions are the Getty Research Institute vocabularies: Art and Architecture Thesaurus (AAT), Thesaurus of Geographic Names (TGN), and Union List of Artists Names (ULAN) and the deliverable of the ongoing MACS project that provides mapping between three subject heading lists—the U.S. LCSH, French RAMEAU, and German SWD/RSWK.

AAT is domain specific, translated partly into French and Dutch (some work was planned/started on Spanish and Italian translations, as well), widely used for managing art information, and convenient for faceted browsing of content. Getty vocabularies are freely available on the Web to support limited research and cataloguing efforts. Regular or extensive use of the vocabularies involves licensing costs, which were not known at the time of writing. Integration of files is not supported and requires some programming.

MACS includes general subject headings in English, French, and German and is planning further expansion to include other languages and vocabularies. As MACS is still in the prototype stage, the art-specific terminology has not been covered yet and the terms of using it are not known. The use of the full vocabulary would require a powerful database and a corresponding platform, which involves substantial extra costs.

The recommendation is to investigate the possibility to set up a research project with suitable partners, or join one of the existing initiatives to develop an art-specific multilingual vocabulary. In the meantime, as Dublin Core allows the use of multiple schemes, indexing can be done using either the freely available Getty vocabularies, or any of those mapped in MACS, without integrating any into the system. The fact that most AAT terms are mapped to LCSH’s should enable its mapping to other vocabularies included in MACS. Although such solution is clumsy and would hinder the development of subject-heading browsing option, it might be very useful for assessing the usability and the required subset of each vocabulary for this purpose.

*Imageline* should establish its own set of **cataloguing rules**, starting from the guidelines developed in other gateway projects. A list is available in Section 5.3.4.

The ideal of allowing each user to access resources in the **language** of their choice and each contributor to describe the resources in their first language should be strived for at all times. The basic user interface should ideally be available in official languages of all

participating countries; it should be the responsibility of each national node to provide translations and any other customization of the user interface for their own audience. Free text fields (i.e. title and evaluative comment) should be provided in the original language of the resource and in English. National nodes are encouraged but not obliged to add additional translations, as an added value for their audience. Controlled-value fields should ideally automatically be mapped to as many languages as possible. Where this is not feasible, values should be provided in English or in a small subset of languages identified by the members as most widely spoken among the primary audience. The fields using authority files for their values are discussed in the foregoing paragraphs. Until multilingual vocabularies are integrated, terms should be provided in the language of the resource and in English.

Although there is no explicit requirement to support any **technical communication standards**, the gateway should preferably provide support for Z39.50, and plan to support OAI-MHP or other relevant technical protocols in the future. The need for actual support should be reassessed on a regular basis. The capability to import and export XML and RDF is strongly recommended.

## WHO

An appropriate **host institution** should be identified by the members, based on the following criteria: experience with related projects, degree of commitment to the project, and availability of staff capable of administering and maintaining the server, troubleshooting, and software development.

The **organizational model** should be based on a two-dimensional grid of national nodes and topic centres. Preferred candidates for national nodes include national libraries, centres of national virtual libraries, gateways, or union catalogues, or any library having enough influence, respect, and leadership in its country and specific expertise relevant for the project. Alternatively, for regions underrepresented in the membership or where required resources are not available, regional nodes can be created instead. Each node should be responsible for resources originating from their country or region. In general, all IFLA Section of Art Libraries members should be able to contribute resources to the gateway. The allocation of resource-selection and cataloguing duties and workflow procedures within a country should be at the discretion of national nodes, designed internally to match the specific circumstances.

At the same time, each node should serve as a topic centre for a specific sub-domain or for a specific aspect of the project. In that way, each national node representative would have a dual function, acting either as an editor for both national and specific discipline content or as an editor for national content and an officer on the project management team.

Intellectual property rights over resource description should be dual—owned by both IFLA and the contributing node.



The *Imageline* project should be run by a **project team** responsible for the efficient and smooth operation of the project. The team should consist of a project manager, product manager, technical, access, and content officers. The structure of the team can be changed once the project reaches the maintenance phase. The representatives of nodes/topic centres should make an **editorial board**. In addition, an external **advisory board** should be appointed to mould the goals of the project, to make sure that it remains on course over time, or change the course if needed, to warn about emerging trends, and address emerging issues. The advisory board should include representatives of user community and specialists for various aspects of the project, such as subject specialists, specialists for indexing and classification, for digital libraries, Internet technologies etc., all leaders and authorities in their fields.

The **business model** should be a combination of “public investment as part of the role of cultural, educational or scholarly institutions” and “collective activity through membership.” In-kind member contributions should form the basis of all operations but that approach requires sensible delegation of tasks and responsibilities and policies for giving proper credits for efforts. Compensation for extra efforts and external expenses should be planned, however, and potential sources of funding identified. If opportunities to obtain research grants are to be explored, best bets lie in partnering with related projects in the areas of multilingual domain-specific vocabularies, multilevel access to information, and integration into various broker architectures.

**Quality control** requires explicit specification of all requirements. These include written documents specifying: scope statement, selection criteria, description standards, design and branding guidelines, language policy, workflows, link checking and content review schedule, communication and co-ordination procedures, evaluation criteria and procedures. Quality control should be carried out at different levels and at each step of the process within a total quality management framework. All contributors should use the documentation and make sure that their selection and descriptions conform to the specified requirements. National and regional editors should make sure that resource descriptions within their scopes are consistent and collections well balanced, while topic centres need to take care that specific disciplines are well covered and the gateway conforms to the requirements in all its aspects. The product manager and the editorial board should ensure the conformance at the gateway level and analyze usage statistics and review the gateway on a quarterly basis. Yearly evaluation procedures should include: evaluation reports by the members of the advisory board, online user survey encompassing both end users and intermediaries, user log analysis, and usability testing if major changes have been implemented. All participants should commit to continuously provide comment, feedback, suggestions, and support to each other.

Annual **operational costs** can be expected to amount to \$US 200,000 in the first year and \$US 40,000 in the subsequent years, according to a conservative and rough estimate. The factors to be taken into account include:

Planning: role assignment, documentation development, metadata schema definition, vocabulary decisions, technology evaluation, equipment acquisition, identification of potential partners, co-ordination, meetings, communication, travel, consumables.

Development: software development or configuration, information architecture definition, user interface design and translation, vocabulary integration or development, logo design, promotional materials, private and public Web site development, resource selection, and description, editorial work, usability testing, co-ordination, meetings, communication, promotion, travel, consumables.

Maintenance: resource addition, updating, weeding, editorial work, server maintenance, troubleshooting, data backups, link checking, use log analysis, evaluation, improvement, co-ordination, meetings, communication, promotion, travel, consumables.

The assumption is that equipment such as personal computers, printers, Internet connection can be used and not included in project expenses.

## 1 INTRODUCTION

This document reports the results and findings of the scope and feasibility study for the *Imageline* project. The main part of the document discusses the problem, issues involved, observable trends, available technologies, potential risks, proposed organizational model, and related costs. The Appendices provide detailed background information used to support the recommendations. The document should serve as a basis for further discussion and decision-making by the IFLA Section of Art Libraries regarding the future of the *Imageline* project.

### 1.1 BACKGROUND

At the 66th IFLA Council and General Conference held in Jerusalem in August 2000, the main theme for the Art libraries section of IFLA, *I'm looking for an image of...* led the delegates to look at the work currently being invested into preparing lists of resources to help in the search for images. Further discussion at the 67<sup>th</sup> IFLA Council and General Conference in Boston in 2001 confirmed the need for rationalization. It was concluded that, in order to prevent the existing duplication of efforts, it is essential to explore possibilities to improve the efficiency and scope of this work at an international level. Following the discussions, the IFLA Section of Art Libraries obtained funding and commissioned a study to assess the feasibility of establishing an international Web gateway providing access to sources of images.

### 1.2 SCOPE AND PURPOSE OF THE STUDY

This scope and feasibility study aims to:

- assess the potential value to users, the cost and the general feasibility of establishing an international Web gateway providing access to sources of images,
- be sufficiently open in scope to serve as a potential basis for future application in other specialized domains covered by the IFLA Special libraries division,
- be structured in such a way as to provide a format and methodology applicable to similar needs in other domains, but itself be limited to the field of Art libraries.

The study answers the questions what, how, and by who needs to be done in order to achieve an international Web gateway providing access to sources of images.

## 2 INTERESTED PARTIES AND THEIR NEEDS

### 2.1 END USERS

*"The satisfaction of user needs is the ultimate, and only, justification for the existence of any information service."*  
[Bawden, p. 40]

As a project-specific user study was beyond the scope of the project, the assessment of user needs has been based on studies conducted within related projects and those found in library literature, studying users of Web gateways, Fine Arts information, and image collections in particular. The findings from these studies served as the initial partial answers to some of the more specific questions this study is trying to answer, listed as sub-headings further in this section. The answers are further refined by matching the needs to the existing services and to the capabilities of the existing technologies.

It is assumed in this study that the end users of the proposed gateway coincide with the existing users of the IFLA Section of Art Libraries members. The actual user body would with no doubt be substantially larger and more diverse (unless the access to the gateway is restricted); however, this subset is considered as the proper target audience whose needs should be considered first for the definition of features and functionality of the gateway. The underlying assumption is that the IFLA SAL members are funded to satisfy the needs of these specific audiences and thus responsible to their fund providers for performing that function.

As the actual profiles of the member libraries were not available for the study, it is recommended that any final decision be examined with respect to members' target users as defined in their mission statements. A random Internet search showed that specialized art libraries mention two basic purposes in their mission statements: (a) support for the work of their parent organization, i.e. members of the art community—researchers, educators, students, and practitioners and (b) access to that information to the general public. Art departments of national and public libraries, according to the mission of their parent institutions, do not mention any audience more specific than "all citizens."

Following the user profiles revealed by the studies, the potential users include: researchers (primarily art historians), educators, undergraduate and research students, art, design, and architectural practitioners, art lovers, librarians (discussed separately in the subsequent sections), museum curators, publishers, and, practically—anyone else. The latest ADAM survey [Monopoli and Nicholas] revealed that the percentage of "other" users match that of practitioners (15%) while the comparison with their 1996 survey revealed a significant decrease in the use by the educational sector over time. AMICO survey at Cornell University found that only 47% of the surveyed users identified Fine Arts as their major field of study, 10% reported Humanities, 4% Social Sciences, 1% Science, while the remaining 38% selected "Other". Similarly, the respondents in the MESL user survey [Sandore et al.] classified themselves as 43% Fine Arts/Humanities, and 57% Social Science/Other.

The "general public" obviously makes a significant portion of actual users and even target audience of some major member libraries. For practical reasons, however, as the needs of such a broad category are difficult to predict, it is recommended to follow as a general rule: mould the gateway to the needs of "professional users" but make it available to anyone else.

This suggestion may require special consideration and discussion before any final decisions are made.

Information needs of this diverse population vary widely in terms of what, how, and why they need to find, how they want to use it, and how badly they want to find it. According to some general studies [e.g. Headline], the users have a strong preference for one-stop-shops, i.e. they want to locate and get the actual resource using the same service. However, the studies specific to the domain of art images found that the users want online images to be linked with local slide libraries [AMICO], that some queries do not require access to the actual image but can be answered from its catalogue entry [Hastings, AMICO-Cornell], while some questions cannot be answered by a digital image but require access to the real object of art [Hastings]. The need to simultaneously access documentary resources was ranked quite low [AMICO]. Respondents in AMICO survey [AMICO-Cornell] rated the usefulness of "large presentation image" at 4.51, "full display" at 4.37, "offline ordering" at 3.5. This means that although online accessible images with full rights are strongly preferred, sources containing only descriptions of images and pointers to offline collections will find their audience in the gateway.

Similar logic can be applied to the dilemma whether to include only free sources or to add subscriber services. Free sources are certainly strongly preferred but, if they are not available, some users will be interested to access commercial ones. It is recommended to exclude commercial sources in the first phase, unless they are providing quality descriptions with no charge. These resources can be added in subsequent phases, if evaluation procedures indicate that they would bring some missing value for the user.

An important feature of the *Imageline* potential audience is that it is international and multilingual. The frequent claim that English language has become a *de facto lingua franca* of the Internet obscures the fact that it prevents a major portion of the world's population from accessing valuable information. Art images are language independent, however their descriptions and accessing mechanisms are not. The ideal situation would be that all contributors describe collections in the language of their choice and each user accesses the descriptions in their own language. The experiences of many recent projects showed that it is not an easy goal to achieve. The actual distribution of languages among the target population can be used as a starting point but a truly international scope should be kept in mind at all times as the ultimate goal of *Imageline*.

## 2.2 MEMBER LIBRARIES AS END USER INTERMEDIARIES

The consulted user studies list librarians as major users of subject gateways and image collections. Their specific uses include:

- Answering reference questions
- Creating guides and bibliographies

The former use does not impose any requirements in addition to those required by the *Imageline* end-users, except a stronger need for the inclusion of searching aids, such as searching thesauri or advanced search options, which librarians are more apt to use than the others [Monopoli and Nicholas]. The latter use might benefit from the breadth of scope, detailed metadata, and automatic filtering.

An important issue is the difference in the scopes of the member libraries. Following the criteria used for the inclusion in the IFLA Section of Art Libraries International Directory of Art Libraries [IFLA], this section includes art, architecture, and design libraries; archaeological, ethnological and anthropological libraries are included as long as their collections support research on visual artifacts of art-historical interest (including architecture and applied arts), while film libraries and slide, photographs and other exclusively visual resource collections are excluded.

As individual profiles of the Section actual membership were not available, an informal search of art library web sites was used as an alternative source of information. The results showed that the standard core includes painting, drawing, printing (depending on period coverage), sculpture, and architecture. Extensions vary from different areas of design, applied arts, to landscape architecture, performing arts, archaeology, numismatics, urban planning, film and so on.

There are also differences in classification schemes used by art libraries, which should be considered if envisioning integration into hybrid libraries but also for providing a familiar or, at least, conceptually compatible browsing structure.

Differences in cataloguing formats and vocabularies used by different libraries in different countries have been reviewed within the Renardus project [Becker et al.].

The language differences are also present between the librarians. The attitude and policies related to the use of English as a *lingua franca* of the Internet, vary between countries rather than between individual libraries; in some countries there is an existing practice of providing all Web content in English. Differences in language skills needed either to assist the users to locate sources or to create resource descriptions in different languages are more likely to vary between individual librarians.

An important advantage of *Imageline* is that its creators and contributors are at the same time its users. This provides an opportunity to design the gateway according to the known user needs and to continuously fine-tune it to better meet these needs.

The fact that the creators of *Imageline* are at the same time its major users is very convenient, as it provides for continuous and direct user feedback. In order to capitalize on this advantage, mechanisms need to be established for continuous feedback and adaptation of the system.

### 2.3 MEMBER LIBRARIES AS COLLECTION PROVIDERS

The project proposal regards art libraries exclusively as users of image collections. Therefore, their role of potential image collection providers [Graham] is not covered by this study. In making final decisions and in any future reviews, the exclusion of this role may need to be re-examined. A substantial number of collections provided by the IFLA Section of Art Libraries members would most likely require a fundamentally different approach, such as harvesting of distributed descriptions created by collection providers.

### 2.4 OTHER COLLECTION PROVIDERS

Two opposite attitudes may be expected from the providers of resources candidate for inclusion in any gateway.

Majority of providers will see the gateway as an opportunity to promote their collections. Commercial providers or aspiring artists, eager to benefit from the opportunity for free advertising, may try to put pressure on editors to include their collections. Such pressure forced many gateways to remove their public "Suggest a resource" option [e.g. Kidsclick]. However, such option accessible to "serious" providers and colleagues is highly desirable. It is recommended to provide the option for suggesting resources but to visibly publish a clear scope statement and selection criteria.

Other providers may fear from being swamped by excessive requests. This group includes small providers of offline collections as well as those serving their collections online and being concerned with the excessive traffic on their servers. This is not expected to be a major concern.

A survey of image suppliers have been conducted within ELISE II project and summary of findings is available [ELISE II] for consultation, should any further concerns regarding this party be raised.

## 3 STATE OF PRACTICE

### 3.1 ACCESS TO IMAGES ON THE INTERNET

General, encyclopaedic, quality controlled information gateways [e.g. BUBL, Lii, DMOZ] provide poor support for accessing art images. Although the gateways include links to art image collections, these are difficult to locate via the home-grown, rigid, organically developed classification/browsing schemes. An exception is Infomine, the only gateway that relies on a good search mechanism and well elaborated and exhaustive metadata. The progress of the announced Fiat Lux alliance of portals should be watched, as it promises to bring some improvement [Schneider].

Image search engines, either independent or included in general search services are often more useful for finding images. This holds true for known-item searching and the general

audience or undergraduate students; however, the lack of quality selection criteria makes these services unsuitable for more serious research and professional use.

Lists of links, developed by librarians from authoritative institutions and routinely included in other such lists are of surprisingly poor quality. They either lack any classification or use home-grown very inconsistent schemes; broken links, the lack of any kind of descriptions or total inconsistency, complete omission of newly added collections, important collections in other countries and especially in other languages, very limited scope are common features. Observed collectively, these lists repeat a limited set of collections and miss some important resources; they are including other lists of the basically same resources, thus generating self-referencing loops, so typical for the Internet chaos. Surprisingly, one-man efforts are often of better quality.

Kirriemuir points to the harms of the coexistence of the large numbers of such lists:

“As well as the somewhat pointless activity of producing a system that is inferior in every way to an existing system, there is also the problem of the large number of these indexes that obscure the users path to decent systems. I noticed more than once that lists of links often had an annoying habit of appearing much higher in the ranking produced by the search engines, than more refined quality-based services in the same subject area.”

And he points to the consequent need to promote a quality gateway:

“This raises another action for high quality resource discovery systems to undertake, namely that of promotion not just to end users, but to people who may be thinking of setting up a gateway or some kind of resource discovery system (no matter what the quality criteria is) that largely overlaps their own system in terms of subject coverage.” [Kirriemuir]

Access to offline image collections is extremely poor.

### 3.2 EXISTING SOURCES OF IMAGES

The state of practice is rapidly changing in this area; at this particular moment probably even faster than the rest of the Internet. Providing art images online is heavily dependent on technology; it had to wait for a substantial decrease in the costs of storage space and broadband connections with a simultaneous increase in processing power and graphical capabilities of computers to become a wide-spread practice.

Currently, there is a proliferation of local projects aimed at digitization and online serving of art image collections directly by their owners. The discovery of new online resources is a real challenge. The international and national initiatives to create clearinghouses or registries of digitization projects, relying on self-submission of new projects, such as the University of Arizona Clearinghouse of Image Databases and National Library of Canada Inventory of Canadian Digital Initiatives, cannot scale up to provide a sufficient world-wide overview of new online image collections.

Another observable trend is the development of virtual mega-collections providing access to items owned by different institutions and stored at different locations. These projects come



from very different parties, both commercial and educational, and use very different approaches but the implication for the *Imageline* project are the same.

All these projects aspire to provide item-level access to a substantial number of images, which sometimes brings into debate the need for a service like *Imageline*. It has been increasingly recognized that the collection-level description and the hierarchical access model that it provides have an important role in the content-rich Internet environment. Collection-level descriptions can provide an overview of what is available, reduce the number of hits, assist in the selection of resources to be searched, and provide access to the "invisible Web" (i.e. items stored in databases and thus not accessible to general search engines). [Miller; Dunn; Ardö A. et al.]

Another implication is that the virtual mega-collections further complicate collection description by imposing the need for an additional class in the data model and the clarification of the relationship between the primary collection and the service that is making it available. In other subject gateways, the service provider is typically irrelevant and therefore not included in resource description. In this case, a mere URL is not sufficient, since before following it, the user wants to know what else they can expect to find there.

An important issue is the wide scope of the concept of a "collection." The IFLA *Multilingual Glossary for Art Librarians* provides the following definitions for the term:

1. "A number of separate works or parts of works, not forming a treatise or monograph on a subject, combined and issued together as a whole.
2. An accumulated group of materials or works having a focal characteristic." [IFLA]

*ArtLex* art dictionary elaborates on the diversity of focal points:

"collection -- An accumulation of objects. [...] Collections can be formed around any of a variety of parameters. They may be centered upon a medium or technique, a certain period or group of artists, or a subject, for instance; or they may be encyclopedic, as can be the entire collection of a large museum. Museums typically have both permanent collections and traveling collections. [...]". [ArtLex]

A collection may equally mean the entire holdings of a particular institution or a discrete part of an institution's collection grouping items sharing a certain property. In addition to collections based on their subject, i.e. encompassing the works of a particular artist, group of artists, particular medium, technique, or discipline, institutions frequently define collections in administrative terms, e.g. the collection of a particular donor, or suited for a particular purpose or audience. In networked environments, the scope of the concept is even broader, as a collection may include items from many different institutions, grouped based on any shared property or intended purpose, regardless of their physical location and ownership.

The collections available online belong to all these categories and users may want to access any.

"Researchers wishing to access information about these collections on-line may be interested in any of the above interpretations of "collection". Museums cannot predict what their users will consider to be a collection, in what

language or terminology they will request the data, or what level of information they need (is it for an elementary school project, or a Ph.D thesis?) Ideally, collection-level description for resource discovery should provide access to many interpretations of "collections" that are dynamically created by the user." [Dunn]

This diversity has serious implications for the design of metadata schemas and selection and description guidelines.

### 3.3 WIDER CONTEXT

#### 3.3.1 Integration and interoperability

Seeking sources of images on the Internet is rarely an isolated process. The user will often need a simultaneous access to local offline image collections, art history sources, other images related to cultural heritage, or any kind of "inspirational sources", such as natural shapes, for example. Efforts aimed at satisfying such needs include the development of:

- "hybrid libraries," which bring together in a seamless way a range of different information sources—printed and electronic, local and remote;
- union catalogues (or "physical clumps"),
- "portals", aggregations of services needed by a particular category of users. Typically, a subject gateway makes the core of a portal, while additional services can include access to complementary databases, current awareness, personalization, discussion groups and so on. Several gateways (e.g. EEVL) have developed into 'portals'.
- broker-based services (or "virtual clumps"), which allow seamless access to the content of distributed gateways, catalogues, or any other resources, from a single user interface (e.g. Renardus and Isaac Network).

Such trends impose the need to regard *Imageline* as a component in a global network of interoperable information services, which can be assembled in different ways to meet particular user needs. All these projects rely on the use of different types of shared standards, which enable interoperability between different systems.

According to Koch,

"Subject gateways and services who want to contribute to global discovery services can accomplish this

- a) through adaptation to standard metadata and exchange formats,
- b) through detailed service profile descriptions,
- c) through adaptation to internet retrieval and query routing protocols,
- d) through support of the full indexing of rich metadata descriptions by metadata aware search services
- e) through mapping of the classification/browsing structure to a common switching system
- f) through adaptation to the requirements of co-operative subject gateway services like Renardus and their broker architectures." [Koch]

### 3.3.2 Standardization

Interoperability and integration are made possible by the use of shared standards. Technical protocols used in the aforementioned projects include Z39.50, CIP, OAI, LDAP, Whois++, and SOAP.

#### 3.3.2.1 Z 39.5

Z39.50 [<http://lcweb.loc.gov/z3950/agency/document.html>] is a communications standard which describes the rules and procedures for communicating between two computer systems for searching and retrieving information from databases. It allows the searcher to search many different databases accessible through Z39.50 servers and retrieve results using the same search interface and without having detailed knowledge of the system being searched.

Because of differences in interpretations and implementations of the standard by different software developers, different Z39.50 profiles have been developed for use in different domains. Those of potential interest to *Imageline* include Bath profile [<http://www.nlc-bnc.ca/bath/bp-current.htm>] used by libraries, CIMI profile for cultural heritage information [[http://www.cimi.org/public\\_docs/HarmonizedProfile/HarmonProfile1.htm](http://www.cimi.org/public_docs/HarmonizedProfile/HarmonProfile1.htm)], Zthes for thesaurus navigation [<http://lcweb.loc.gov/z3950/agency/profiles/zthes-04.html>], Renardus profile [[http://www.renardus.org/about\\_us/deliverables/d2\\_2/D2\\_2\\_final.html](http://www.renardus.org/about_us/deliverables/d2_2/D2_2_final.html)] used in the EU Renardus project. The list of existing profiles can be found at [<http://lcweb.loc.gov/z3950/agency/profiles/profiles.html>].

*Imageline* can serve as an interesting Z39.50 target but, in the future, a need may arise to provide access to other related sources as a Z39.50 client. The Z39.50 is a powerful integration tool but its implementation is not easy and it costs money, especially the implementation of Z39.50 target.

#### 3.3.2.2 OAI-PMH

The Open Archives Initiative (OAI) [<http://www.openarchives.org/>] provides interoperability agreements for archiving materials so that they can be accessible to mediator services. The OAI Protocol for Metadata Harvesting (OAI-PMH) defines an interoperability framework with two classes of participants. Data providers administer systems that support the OAI protocol as a means of exposing metadata about the content in their repository. Service providers issue OAI protocol requests to the systems of data providers and use the returned metadata to build value-added services.

The OAI technical framework is intentionally simple and easy to implement but its functionality is inferior to other protocols. The OAI technical framework is not intended to replace other approaches but to provide an easy alternative for different constituencies or different purposes than those addressed by other interoperability solutions.

The initiative has been embraced by the museum community; CIMI announced the version 1.0 for the CIMI repository code [<http://www.cimi.org/publications.html#oai>]. The potential wide use of the protocol by museums and other relevant organizations to expose descriptions of their collections might bring the need to re-implement *Imageline* as an OAI service provider.

### 3.3.2.3 CIP

The Common Indexing Protocol (CIP) [<http://www.faqs.org/rfcs/rfc2651.html>] is used to pass indexing information from server to server in order to facilitate query routing. Query routing is the process of redirecting and replicating queries through a distributed database system towards servers holding the desired results.

The most important encoding standards are RDF, XML and the whole suite of XML satellite standards.

Overviews of technical and encoding standards:

*The IMesh Toolkit Review of Protocols and Standards, June 2002.*

[http://www.imesh.org/toolkit/work/tech\\_review/protocols/](http://www.imesh.org/toolkit/work/tech_review/protocols/)

Renardus D2.1-- Technical Standards and Solutions, June 2000.

[http://www.renardus.org/about\\_us/deliverables/d2\\_1/D2\\_1\\_final.html](http://www.renardus.org/about_us/deliverables/d2_1/D2_1_final.html)

In the simplest implementation of the gateway, these standards can be disregarded, as they, without exception, involve distributed metadata collection. The most important ones are briefly mentioned here in order to provide information about context that might be relevant for longevity of the project.

## 4 SCOPE SUGGESTIONS

The analysis of user needs and its comparison with the available services, provided the initial answers to the WHAT group of questions. Due to the observed remarkable diversity of user needs and the lack of specific information about target audiences and user needs of the existing and envisioned IFLA SAL membership, this section recommends the strategy to be used, rather than a specific definition of scope. The suggestions listed below should, therefore be regarded as starting points for further discussion and refinement by the members, which should follow the recommended criteria and strategy.

### 4.1 OBJECTIVE

To provide quick and efficient access to sources of art images, selected based on quality and relevance to the primary audience.

### 4.2 PRECEDENTS

There has been no existing gateway identified that meets this objective and there is no service in another field that can be directly re-applied for this purpose.

### 4.3 STRATEGY

All decisions regarding the features and functionality of *Imageline* should be based on identified needs of the member libraries' target audience. User needs should be continuously monitored throughout the lifecycle of the project and the scope and selection criteria should continuously be adapted to match these needs. The suggestions provided by this study should be discussed as soon as possible in order to define and document the scope policy and selection criteria. These documents should be based on "must", "should", and "may" statements, as used in Internet Engineering Task Force (IETF) documentation.

Flexibility and willingness to change direction where necessary are essential in the rapidly changing environment. These requirements are best met using a modular approach and continuous feedback, monitoring of relevant developments, and re-evaluation.

An important lesson learned within AMICO project should be kept in mind:

"Collections development for networked cultural heritage will remain a challenge for some years. The size and scope of the problem – collections comprising millions of works are not uncommon – requires any solution to be incremental. At about 55,000 works there are more gaps in the AMICO Library than there are strong points – more holes than cheese – and opinion differs as to the appropriate strategy for collection development. In theory, Strategies to construct content by "pulling" what users request or editors suggest, is complicated by the differences in opinions between different users as to what constitutes critical mass or desirable content. In focus group studies conducted by AMICO in 1999, we found tensions between users who want breadth of coverage, and many representative works, and those who want depth of coverage, many images of a particular work along with detailed documentation and accompanying multimedia.

We found some potential users wanted works that they were familiar with while others thought it would be more valuable to have obscure works and works from hard to get to collections.

In short, the expectations and desires of digital library users will color their perception of networked cultural heritage resources. Collectively they will not seek the same content, for the same purposes.

AMICO has accepted that the AMICO Library will grow over time. Vehicles to facilitate communication between users and AMICO Members to identify pockets of broadly useful content may be a way forward." [Trant et al.]

### 4.4 TARGET AUDIENCE

Criteria:

Target audience should be defined as cumulative primary audiences of the IFLA Section of Art Libraries members. User profiles should be defined in terms of disciplines and job roles.

Suggestion to be verified through discussion:

Primary audience includes: researchers, educators, undergraduate and graduate students, practitioners, and museum curators in the disciplines as defined in the following section, and librarians.

Secondary audience includes art lovers, students, educators and researchers from other fields, publishers, and anyone else who, for any reason, may want occasional access to visual art images.

#### 4.5 INFORMATION COVERAGE

##### 4.5.1 Disciplines

Criteria:

Coverage of disciplines should be defined as cumulative subject matter scope of the IFLA SAL members. The overlap of the coverage of the member libraries should be defined as a core domain of *Imageline*, which should be included first and with the maximum number of contributors. Disciplines suggested for the inclusion into the core are marked with "(1)"; those that might be of equal importance but have contributors, depending on the current membership are marked with "(2)", while "(0)" denotes disciplines to be excluded based on the scope of the IFLA directory, and "(?)" those that are not mentioned in the directory scope but might be included in the scope of some libraries.

Suggestion to be verified through discussion:

- fine arts (1) including painting, drawing, printing, sculpture, new media, and conceptual art,
- architecture (1), including landscape architecture (?), naval architecture (?), and urban design (?),
- design (2), including graphic, fashion, interior, costume, industrial, packaging, and theatre design
- applied or decorative arts (2), including textiles, ceramics, glass, metals, jewellery, furniture
- archaeological, ethnological and anthropological collections to be included as long as they support research on visual artifacts of art-historical interest (including architecture and applied arts) (2),
- collections of images compiled with the intention to serve as "sources of inspiration" for artists and designers (e.g. sources included in THAIS) (3),
- film (0),
- performing arts (0),
- photography and video only if created with artistic intention and collected because of artistic rather than documentary value (2).

The scope should be limited to collections of images representing objects that have artistic, as opposed to documentary value. Although some partner institutions, in particular

those dealing with cultural heritage, may have both types of collections within their scopes, it is deemed that the exclusion of the latter is important in order to achieve the modular architecture. Expansion of the scope would require different vocabularies, similarly as the expansion to include art history resources would require a different metadata schema. These extensions may be considered as separate projects, investigating problems of linking to other domains (e.g. history) in a modular architecture. It may be worth noting here that the presence of artistic value in objects represented in the collections should not be assessed by the contributors, since it is not their field of expertise. The fact that a museum or another institution decided to assemble the items into a permanent collection should be considered as an indicator of their artistic value.

#### 4.5.2 Geographical coverage

The inclusion of collections from all parts of the world is desirable.

Collections in all world languages should be included as long as the person selecting the resource has the language skills sufficient to evaluate it.

Geographical and language distribution of the IFLA SAL membership should be identified in order to see how this facet can be best handled by the organizational model. There should be no priorities in terms of geographic coverage.

#### 4.5.3 Temporal coverage

The gateway should cover the full range of time periods and styles, from prehistory to contemporary art, each having equal value for different users and user tasks.

There should be no priorities in terms of periods or styles covered but any specialization in temporal or style coverage by the member libraries should be noted.

#### 4.5.4 Types and accessibility of resources

Preferences and priorities in terms of types of resources should be stated and used for planning the project schedule.

- Public domain sources, allowing free, full-rights access to online images (1),
- Collections with any kind of added value—structured item descriptions and evaluative critical comments from authoritative sources (1),
- Resources that require free registration (1) but with information about the registration provided in the resource description,
- Due to the significance of quality graphics and associated technical requirements, there should be no restraints on selection of resources that use advanced WWW

technology (1). Any specific technical requirements, such as connection speed, special viewers, plug-ins, fonts etc. should be stated in the resource description.

- Collections providing free image descriptions and thumbnail images with pay-per-use access to large size images (2),
- Collections of digital or hard copy images that provide descriptions of items and allow ordering or loan (2),
- Resources requiring subscription fees for any access if satisfying the other selection criteria and providing collections and images important to the target audience and not otherwise accessible. (2) The resource description should include this requirement and the gateway should allow filtering of such resources.
- Resources with access restricted to local, or otherwise limited groups of users (0) unless available through a license to several member libraries (2).
- Resources which consist exclusively of links to other resources (0),
- Resources that consists for a considerable part of publicity materials (0).

#### 4.5.5 Level of granularity

Both primary sources and services providing images from several collections online should be included. The distinction and relationships should be made explicit by the description format.

Records should describe and point to collections, rather than to museums or other institutions providing them. Relationship with and required information about the owner institution should be stated in the collection description.

#### 4.5.6 Sources of information

There should be no restrictions on sources of information—academic, government, commercial, non-profit, and individual sources are all acceptable provided that they meet the quality selection criteria.

#### 4.5.7 Level of “difficulty”

The gateway should be at a level suitable for higher research and education. Resources should be scholarly rather than popular. The scope needs not to be exhaustive but sufficient to support research.



#### 4.5.8 Languages

The ideal of allowing each user to access resources in the language of their choice and each contributor to describe the resources in their first language should be strived for at all times.

The basic user interface should ideally be available in official languages of all participating countries, if possible.

Free text fields should be provided at least in the original language of the resource and in English. Controlled-value fields should ideally automatically be mapped to as many languages as possible. A mandatory subset of languages should be agreed upon. Integrated multilingual vocabularies should be used for the fields using authority files. If that is not possible, subject terms should be provided in the language of the resource and in English.

#### 4.5.9 Evaluative comments

Added value is the key factor for the high appreciation of subject gateways. The key elements of added value include: selection, description, and organization of resources. The exact contribution of evaluative critical comments cannot be precisely assessed but they unquestionably represent a major factor in satisfying the third function of a bibliographic record, especially bearing in mind the non-standard definition and frequent lack of subject in the case of art image collections. Any decision should also take into the account the effort required to create and translate the descriptions. It is recommended that evaluative descriptions be up to two sentences long, rich with keywords, and describing specific features of a resource not covered by its metadata, but not evaluating artistic merit of the works within the collection.

#### 4.5.10 Criteria for selection

Selection criteria should be based on target audience needs. A clear scope statement and quality criteria should be developed in the preparation phase, starting from the existing standard guidelines and those used in specific gateways.

Hofman P. and E. Worsfold, *Selection Criteria for Quality Controlled Information Gateways: Work Package 3 of Telematics for Research project DESIRE (RE 1004) 1997*  
<http://www.ukoln.ac.uk/metadata/desire/quality/>

Internet Detective: interactive online tutorial [<http://www.sosig.ac.uk/desire/internet-detective.html>]

Pitschmann, L. A. Building Sustainable Collections of Free Third-Party Web Resources, Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, June 2001. Available at: <http://www.clir.org/pubs/reports/pub98/pub98.pdf>

Internet Detective. Creating a Scope Policy for Your Service. Available at: <http://www.netskills.ac.uk/TonicNG/content/detective/56.html>.

DESIRE Information Gateways Handbook. DESIRE Consortium, 2000. Available at: <http://www.desire.org/handbook/2-2.html>.

ADAM: Guidelines for the Selection of Resources for Inclusion in the ADAM Database. Available at: <http://adam.ac.uk/adam/reports/select/>

DutchESS. Dutch Electronic Subject Service. DutchESS Scope Policy. Available at [http://www.kb.nl/dutchess/manual/scope\\_eng.html](http://www.kb.nl/dutchess/manual/scope_eng.html).

Humbul Humanities Hub. Available at <http://www.humbul.ac.uk/about/colldev2.html>.

Jennings, Simon. 2000. RDN Collections Development Framework.

Version 1.1 (May). Available at <http://www.rdn.ac.uk/publications/policy.html>.

Library of Congress. BEOnline: Selection Criteria for Resources to be included on the BEOnline+ Project. Available at <http://lcweb.loc.gov/rr/business/beonline/beonsel.html>.

SOSIG Social Science Information Gateway. Training Materials and Support. Available at [http://www.sosig.ac.uk/about\\_us/user\\_support.html](http://www.sosig.ac.uk/about_us/user_support.html).

## 5 PROCESS

Once it is identified what should ideally be done, the entire process required to achieve that goal needs to be analyzed—broken down into steps or specific tasks performed by both providers and users, available options for each of the tasks identified and evaluated. Each task should be examined in terms of the possibility and desirability of automation, decisions that need to be made, start-up and long-term requirements. Available options typically involve human efforts versus various levels and types of automation. In the light of the available options, the feasibility of specific features defined in the scope may need to be re-assessed. Any special circumstances, both advantages and challenges specific to the project need to be identified at the outset.

### 5.1 OPPORTUNITIES, OBSTACLES, AND RISKS

- Opportunities to build on:
  - o enormous human expertise potential—professional contributors,
  - o member libraries both developing and using the gateway,
  - o world-wide effort coordinated by a well-established, permanent, well-organized body,
  - o available new technologies
  - o related initiatives and research
  
- Special challenges:
  - o unclear boundaries of the domain,
  - o diverse collections in terms of scope, nature, purpose,
  - o diverse information management approaches: library, archival, museum

- diverse data standards used
  - diversity of user needs,
  - multiple languages,
  - non-standard domain to which widely used general standards are not easily applicable
- Risks and risk mitigation strategies:
- Lack of continuous funding
  - Lack of buy-in / commitment
  - Loss of momentum (enthusiasm over time)
  - New technologies allowing cheaper and better service in the near future, for example, efficient global searching/browsing of images at item level

## 5.2 PROVIDING CONTENT

### 5.2.1 Discovering Resources

The first task is to locate resources suitable for inclusion into the gateway, making sure that newly developed resources and those of highest quality are not missed. This is a time consuming process; the very *raison d'être* of subject gateways is to spare the end-user from that effort. A judicious division of resource discovery responsibilities among contributors can eliminate duplication of effort and should be used as a basis for the organizational model. Two primary candidates to be used as a basis for the allocation of responsibility domains in *Imageline* are language/region and subject.

Internet Scout uses three basic techniques for discovering new resources: information trawling—monitoring mailing lists and newsgroups, information scouting--searching the Internet for quality resources, using the "site-ation" pearl growing technique, where new sites are located by searching known quality information meta-sites, and user submissions. [Solock]

Many gateways allow the user community to suggest new resources using a special Web form or an e-mail address available on the site. There is a danger, however, that the submission mechanism gets abused by commercial image providers and aspiring artists. Publishing clear selection criteria is suggested as a countermeasure but many gateways were forced to remove submission forms after being swamped by excessive suggestions [e.g. KidsClick]. Promoting the project and requesting suggestions from colleagues and subject specialists from local communities through personal and professional networking can be invaluable.

The use of automation is desirable in this task, as massive processing is required. The ideal solution would involve the use of human controlled, easily customizable Web crawlers. Combine and Harvest\_NG are popular harvesting tools used in subject gateways. Although these tools can also create indexes, ELVIL showed that they are not good enough for that purpose. [ELVIL]

[DESIRE Handbook] provides sources, guidelines, and strategies to be used by gateway managers and staff.

### 5.2.2 Evaluating and selecting

Evaluation and selection criteria have been discussed in the previous sections; these processes are observed here from the organizational point of view. Following the very definition of a subject gateway, the use of automation should not be considered for this task, since the use of human judgement for selecting high quality resources gives the irreplaceable value to this type of service. While resource discovery benefits from multiple contributors, evaluation and selection should be more centralized

### 5.2.3 Controlling editorial access

Access control system needs to be in place, allowing different access rights to different groups of users. This requires a user registry and login forms. Roles of contributors, editors and administrators need to be assigned allowing different rights across different scopes. Administrators responsible for assigning these roles and controlling access to the system need to be identified. User access would best be administered centrally although certain administrative rights can be assigned to national centres.

### 5.2.4 Submitting

Direct submission of resources by end users should be excluded from consideration, due to the high quality requirements. Two options to be considered include:

- Submission by all section members
- Only by selected editors.

The latter option can ensure better consistency of records, while the former puts less burden on individuals. Due to the very different circumstances across the world, in terms of infrastructure and skills, it is recommended that the most feasible distribution of responsibilities be defined within each specific country.

Contributors need to access submission forms allowing entry of information needed to describe a particular resource. These forms should be accessible through a wide range of Web browsers. The tool selected for this purpose should not require licenses or per-seat royalties.

### 5.2.5 Coordinating additions

A mechanism needs to be put in place to ensure that there are no redundant records. This can be achieved in two ways : by clearly dividing domains of responsibility or by technically preventing addition of duplicates. Both approaches involve some challenges. Despite the extensive work on standardizing resource identifiers, there is no standard way yet to unambiguously identify a unique Internet resource. The same resource can exist at several mirror sites with different URLs, there is no standard naming or numbering scheme for Internet resources, and standardized or consistent local metadata are not always available; making it impossible to automatically prevent duplicate entries. Setting clear boundaries between sections of the Internet is equally difficult. National division stops short because of international sites, collaboratively developed, mirrored, and translated, there are sites in one country serving collections from another ; subject or disciplines are combined in very different ways. Therefore, manual checking cannot be avoided and clear guidelines should be provided for contributors.

### 5.2.6 Storing/persisting resource descriptions

Descriptions of the resources should be stored as database records and used to generate HTML pages on-the-fly. An alternative is to store them in a file system as static HTML files and thus make the content accessible to general search engines. The database option, however, allows a considerably greater flexibility, including searching by different parameters, alternative displays of content, easy management of changing data, filtered views, personalization, customization etc. A database-indexing tool, which allows the database engine to find information faster should be provided.

Although the Z39.50 target functionality is not specified in the project proposal, it is commendable to include it in order to enable access to the content to various kinds of brokers.

## 5.3 PROVIDING ACCESS TO CONTENT

### 5.3.1 Describing resources

Quality descriptions of resources make another distinguishing feature of subject gateways. Descriptions should allow the user to find, identify, select, and access resources [IFLA 1998]. Currently there are no established and widely used standards for describing visual resources and collections in art libraries; a variety of local formats have been developed and used [Graham; Medina and Coso]. A review of metadata formats [Dempsey et al., 1997] identified and described over twenty formats for describing Internet resources that were in use or under development in 1996. There have been significant advances towards standardization since that time but there is no standard that can be directly used in a new gateway yet. The use of

standard description formats is important for two reasons. One is that it provides for interoperability with other related systems and allows machine processing by many standard software applications. The other is that standard formats have been developed over a long time period by a number of knowledgeable contributors and therefore hold promise that all issues have been thoroughly considered and best solutions found.

According to VADS/TASI guidelines [Grout et al.], “the following three considerations should be looked at when choosing a standard:

- Fitness for purpose
- Reputation
- Existing experience.”

DESIRE handbook also mentions the ease of use and maintenance, technological and social factors. [DESIRE]

Quality of descriptions should be controlled with respect to specificity, completeness of fields, syntactic correctness, semantic correctness, and consistency.

### 5.3.2 Data structure

The selection of the data structure format should be based on the nature of items to be described and on the environment in which the system is implemented and used. The *Imageline* metadata format should be suitable for describing art image collections available on the Internet well enough to allow the user to find, identify, select, and access them, should be easy for the cataloguer to use, should not require an overly complex data structure and code, and should be interoperable with formats used in related services.

An important feature of this gateway is that it provides links to image collections and not to individual images. An overview of collection description standards is available in [Powell et al.] and a recent survey of formats actually used for this purpose in [Collection Description Focus / CIMI]. EAD and ISAD are developed specifically for describing archival collections and appear to be too complex for other purposes.

RSLP Collection Level Description (RSLP CLD) is an effort to standardize relatively simple standard archival descriptions encoded using RDF/XML syntax, for use in a variety of contexts. The RSLP team also developed an accompanying Web-based collection description tool. However, the schema may still be too complex for direct use in *Imageline*. It is important to note that a semantically correct model of the *Imageline* domain might be even more complex, however, very few of its elements are actually needed to provide access to sources of images.

Dublin Core is widely accepted and successfully tested for describing a wide range of Internet resources and there are numerous tools available for generating DC-based descriptions. The standard enumerates “collection” as one of possible resource types but it

lacks many fields needed to describe collections. One of the key features of the DC schema is its extensibility and the development of an application profile based on DC is certainly an option to be considered. That approach has been used in Renardus, where collection description schema includes nine DC elements, one RSLP CLD, and six project-specific ones.

The DCMI Collection Description Working Group is currently working on the development of a Dublin Core-based application profile for collection description, including a metadata schema, a set of guidelines for the usage of that schema, an enumerated list of collection types and exploring possible mechanisms for mapping between different collection description standards, including the development of appropriate crosswalks. The release, scheduled for May 2002, was not available yet at the time of the writing.

The [Collection Description Focus / CIMI] survey also revealed a surprisingly wide use of MARC for describing collections in various institutions. Therefore, and especially if the scope of *Imageline* is to include collections not available online, MARC format should be considered as an alternative.

Numerous schemas that have been developed specifically for the art domain address description of individual items within art collections and, as such, are not suitable for describing collections. JIDI does add collection description to the descriptions of 'Works' or 'Visual Documents,' which are based on VRA. A coordinated approach to the description of these three levels is definitely an advantage; however, the format is intended for use only within the JIDI project and not tested in practice yet. All these standards should certainly be extensively consulted, however, while developing an application profile. There is a crosswalk between major standards, developed by the Getty Research Institute [Getty].

Major problems encountered using various schemes, as reported in [CD Focus/CIMI] survey, include:

- "unfamiliar concepts and terminology
- relationship between item-level and collection-level description
- legacy data embodies a wide range of (possibly informal) approaches to collection description
- requirement to describe collections for specific audiences
- complex relationships between collections, services and agencies
- granularity issues - defining collections appropriately
- use of standards across domains (e.g. applying MARC to archival description)"

Major problems that can be expected in describing image collections for *Imageline* include:

- what is being described--a collection of images, collection of works shown by these images, or the service that is providing access to the images,

- which "agents" are relevant: collection owner, administrator, and location, digital collection owner, creator, and provider, provider and location of the value-added service making the collection searchable.

RSLP CLD “takes the view that ownership, administration and location are relevant to the definition of a collection,” [Heaney] but the question is whether these attributes are relevant to provide access to these collections in *Imageline*. Some trade-offs between semantic correctness and simplicity and compatibility with other services will need to be made.

Another set of issues involves the non-standard meaning of “subject” in the context of art image collections and frequent lack of that attribute, which has been identified as the most important access point in all surveys. Related are the issues involving synchronization with other services. An artist’s name will be the value of “creator” attribute in services providing item level access to images; here it will be the value of “subject.”

All decisions should primarily be based on the value to the user—whether it helps them to locate, identify, evaluate and access the resource. Secondary, but still very important factors include familiarity and intuitiveness to both end user and cataloguer, simplicity, and compatibility with formats used by other services.

In addition, there is a need for project-specific housekeeping metadata describing the record itself, such as cataloguer, date added, date modified, and so on. This data is needed for creating lists of older records that need to be checked, generating lists of new additions to be displayed to the user, and for other maintenance and quality control purposes. Some of the attributes may need to be assigned for each metadata element. This subset of required elements should be defined once the organizational model, procedures and the system are fully defined. The goal should be to minimize cataloguer’s effort by relying on automatic extraction wherever possible—using timestamps, user login information etc.

### 5.3.3 Vocabulary

The next standard to be considered involves allowed values of data elements defined by the data standard. The value of controlled vocabularies has been widely recognized. In the user studies consulted, subject heading browsing and use of thesauri in searching have been reported as strongly preferred means for accessing art information.

The first requirement in the selection of vocabularies is their suitability for the domain. The [CD Focus/CIMI] study reports that “[a]lmost all of the difficulties identified related to the fact that general-purpose subject vocabularies like LCSH and UNESCO were not specific enough for use in specialised subject areas.”

Another issue is related to the international nature of the gateway. The ideal is that every user can access resources using their first language and every contributor can catalogue them in their own. Providing for searches and record displays in different languages, however, requires elaborate solutions and substantial effort and investment.

In terms of language requirements, three types of metadata elements or fields can be distinguished. The elements restricted to limited lists of allowed values, such as collection



type or language of the resource, can easily be automatically mapped from the cataloguer's to the user's language. Lists of allowed values for all appropriate fields need to be established in the planning phase. Another subset of metadata includes those fields that should preferably be populated from authority files. While some terms, such as place and artists' names can be usable in different countries, even if only the original name is provided, terms denoting iconographic content of collections, for example, would require the use of a multilingual thesaurus. Free-text elements, such as description, would need manual translation into each language.

A widely used, multilingual, domain specific vocabulary would be a preferred option for the data content standard. Two closest solutions are the Getty Research Institute vocabularies: Art and Architecture Thesaurus (AAT), Thesaurus of Geographic Names (TGN), and Union List of Artists Names (ULAN) and the deliverable of the ongoing MACS project that provides mapping between three subject heading lists—the U.S. LCSH (Library of Congress Subject Headings), French RAMEAU (Répertoire d'Autorité-Matière Encyclopédique et Alphabétique Unifié), and German SWD/RSWK (Schlagwortnormdatei/Regeln für der Schlagwortkatalog).

AAT is domain specific, translated partly into French and Dutch (some work was planned/started on Spanish and Italian translations, as well), widely used for managing art information, and convenient for faceted browsing of content. Getty vocabularies are freely available on the Web to support limited research and cataloguing efforts. Regular or extensive use of the vocabularies involves licensing costs, which were not known at the time of writing. Integration of files is not supported and requires some programming.

MACS includes general subject headings in English, French, and German and is planning further expansion to include other languages and vocabularies. As MACS is still in the prototype stage, the art-specific terminology has not been covered yet and the terms of using it are not known. The use of the full vocabulary would require a powerful database and a corresponding platform, which involves substantial extra costs.

The recommendation is to investigate the possibility to set up a research project with suitable partners, or join one of the existing initiatives to develop an art-specific multilingual vocabulary. In the meantime, as Dublin Core allows the use of multiple schemes, indexing can be done using either the freely available Getty vocabularies, or any of those mapped in MACS, without integrating any into the system. The fact that most AAT terms are mapped to LCSH's should enable its mapping to other vocabularies included in MACS. Although such solution is clumsy and would hinder the development of subject-heading browsing option, it might be very useful for assessing the usability and the required subset of each vocabulary for this purpose.

Potential partners for the work on vocabularies include: MACS (Multilingual access to subjects) project, Getty Research Institute, CHIN (Canadian Heritage Information Network), RDK (Rijksbureau voor Kunsthistorische Documentatie), ArtSTOR project, AMP (Accès Multilingue et Patrimoine), and University of Laval.

#### 5.3.4 Cataloguing rules

Standards for the selection, organization, and formatting of terms used as values of metadata elements (data content standards) must be used in conjunction with other standards in order to achieve a high level of consistency and efficient retrieval. Such standards (Internet counterparts of AACR) have attracted less attention than data structure and vocabularies and currently there is no widely accepted standard for this purpose.

*Imageline* should establish its own set of rules, or rather, adapt guidelines developed for other gateway projects. Examples include:

- National Library of Australia and the State Library of Tasmania, "Guidelines for the Creation of Content for Resource Discovery Metadata."  
[<http://www.nla.gov.au/meta/metaguide.htm>]
- Day, M. and P. Cliff, "RDN Cataloguing Guidelines: Version 1.0 - final draft"  
[<http://www.rdn.ac.uk/publications/cat-guide/>]
- Day, M. ROADS cataloguing guidelines.  
[<http://www.ukoln.ac.uk/metadata/roads/cataloguing/cataloguing-rules.html>]

The Visual Resources Association (VRA), sponsored by Digital Library Federation and Getty Grant Program, is currently reviewing and evaluating existing data content standards and current practice in order to compile a manual that may be used to describe, document, and catalog cultural objects and their visual surrogates. The guidelines, which should be available in 2003, should also be reviewed. [Project Proposal]

#### 5.3.5 Data entry tool

The tool used for describing resources should be able to accommodate a Dublin Core-based application profile and allow customization of metadata schema. It should be able to validate conformance to the schema, identify and prevent duplicates, to record identity of the user, and generate timestamps. It should be possible to configure default values. The tool should be able to extract metadata values from HTML files to assist the user. The pick lists for appropriate fields should be available and a thesaurus integrated. Online instructions and help should be provided.

#### 5.3.6 Search mechanism

All the surveys showed a strong user preference for using simple search. However, advanced search features, such as field search, Boolean logic, truncation, phrase and proximity searching, date comparisons, thesaurus integration, combined browsing and

searching, query refinement are irreplaceable in some cases and should certainly be added unless involving substantial extra costs. A customizable ranking mechanism with a sophisticated relevance ranking option should be a preferred choice. Record elements to be shown in display results should be configurable. The system should be able to record searches and create statistical reports.

The language of the search interface and language(s) used for searching should match the overall language strategy of the service.

### 5.3.7 Browsing

Browsing can be implemented in two ways: using a classification scheme or as a metadata-based faceted browsing system. The choice will primarily depend on users, user needs, and the actual content of the gateway. The heterogeneous, international user community calls for a widely known scheme, the diversity of user needs would best be satisfied by a faceted one, while the scope of *Imageline* content suggests a domain specific scheme. Unfortunately, there is no single scheme that satisfies all these requirements.

Universal schemes, such as DDC, UDC, LCC have the advantage of being used worldwide and thus familiar to both the user and cataloguer, convenient for multilingual implementations because of numerical coding and existing translations, and allowing easy integration with hybrid libraries and cross-domain broker services like *Renardus*. Unfortunately, none of these schemes is truly suitable for browsing art image collections. Dewey Decimal Classification (DDC) would probably be the first choice among the three best known schemes, being most widely used, translated into most languages, and used in most gateways, including *Renardus*. However, the lessons learned from ADAM are discouraging. After years of trying to use DDC for browsing art information, they "were forced to concede defeat—[they] had failed to produce a logical user-friendly browsing interface that was based upon the Dewey Decimal Classification," stating that "[...] Dewey is simply too inconsistent and complex, particularly in cases where classmarks are *synthesised* from multiple notations taken from different parts of the scheme." [Wallace and Web] *Imageline* presents even more challenging problems due to the non-standard definition of "subject".

Fortunately, even if another scheme is used for browsing, it is possible to use DDC classification codes for subject indexing in Dublin Core, which would allow for integration into hybrid libraries or cross-domain portals using DDC.

National schemes, especially those general in scope, such as SAB, BIBSYS, BC, or RVK should not be considered, due to the international scope of the gateway. Again, it is always possible to add additional codes locally, if needed.

The only standard, subject-specific classification scheme is ICONCLASS. It has been translated into several European languages [<http://www.iconclass.nl/texts/faqfr.htm>] and well

known in the domain. However, ICONCLASS is based on iconographic content of images and, as such, may not really be suitable for classifying image collections, which are rarely based on iconographic content.

It is recommended to use a faceted browsing scheme based on the selected metadata schema, which would allow multiple ways for accessing information required to satisfy the diversity of user needs. It should also be possible to filter displays by collection type, language of the resource or another attribute.

#### 5.3.8 User interface

The possibility to create different user interfaces for different audiences should be considered. It can involve interfaces in different languages, adaptation for different cultures, possible limitation of default scopes or integration with local resources. The customized user interfaces should be the responsibility of national nodes. Guidelines for branding need to be agreed upon in that case.

The design component of the user interface should be carefully considered due to the high aesthetics criteria of the audience. Competitions for logo and general design can be considered across parent institutions or wider community, which can at the same time contribute to the promotion of the gateway. Design guidelines need to be established and followed.

User support should include help with searching and a detailed overview of the gateway content and structure.

### 5.4 COLLECTION MANAGEMENT AND MAINTENANCE

*Imageline*, having quite specific resources in its scope, can be envisioned as a smaller gateway, which, according to Koch's classification, can include a few hundred to 2,000 resources [Koch]. It can be estimated that after the first year, it will include 1,000 records, as a maximum. The maximum monthly growth rate of 10% can be expected in the subsequent years. The targeted level of completeness and balanced coverage of disciplines, cultures and periods are difficult to measure.

Due to the ephemeral and dynamic nature of Web resources, and the especially fast pace of change in this specific domain, the maintenance of *Imageline* can be expected to be a labour-intensive and costly task, which should be properly accounted for in advance. Nielsen warns:

"As a rule of thumb, the annual maintenance budget for a website should be about the same as the initial cost of building the site, with 50 percent as an absolute minimum. Obviously, ongoing costs are even higher for news sites and other projects that depend on daily or real-time updates. If you simply spend the money to build a glamorous site but do not keep it up to date, your investment will very rapidly turn out to be wasted." [Nielsen]

#### 5.4.1 Updating

The proliferation of digitizing projects can be expected to continue in the foreseeable future. Therefore, the addition of new resources in the maintenance phase, updating and replacing the existing ones will most likely be equally intensive as the setup. A recent survey showed that "continual updates" of Web sites make one of top two consumer expectations, expressed by 96% respondents [Gaudin]. Following the DESIRE Handbook, a record needs to be edited

- "if the information content of the resource has changed so that the resource description and keywords need to be updated
- if any of the factual details of the resource have changed (e.g. new admin email, new short title)
- to correct any errors made in the original record."

The software should be able to create lists of older resources that need to be reviewed and send them to the appropriate person.

According to the Handbook, the resource should be removed:

- "if the resource is no longer available
  - if the currency or reliability of the resource has lessened
  - if another Internet site or resource offers more comprehensive coverage."
- [DESIRE Handbook]

Due to the dynamic developments in the domain, it can be expected that *Imageline* will need regular weeding and replacement of older resources with better ones as they appear. Appropriate weeding policies need to be formalized. Koch's review has not identified examples of such policies in the existing gateways. [Koch]

#### 5.4.2 Obtaining and incorporating user feedback

One of the key sustainability factors is the continuing adaptation to user needs. The gateway should provide a mechanism for obtaining continuous, direct user feedback and efficient cross-node communication for obtaining feedback through user intermediaries. Yearly evaluation procedures should include online user surveys and focus groups. Procedures for incorporating that feedback need to be established.

#### 5.4.3 Link checking

Regular link checking and updating is of the uttermost importance in the unstable Internet environment. Luckily, the major portion of this task can easily and efficiently be automated. Most tools can be scheduled to check URLs of the included resources at regular intervals. It is recommendable to program the tools to run at least once a week, preferably at 'quiet' times, e.g. overnight, to reduce the load on the network. The link checker will normally produce a report on broken links, classified by error code, which should then be manually reviewed and decided which URLs need to be fixed and which records should be removed. It is advisable to keep contact information for Web masters of sites on which resources reside since it is often

the easiest way to find out the real reason for a broken link; otherwise special skills, knowledge and effort is needed to detect the reason for a broken link.

There are many link checking tools available for free download or integrated into Web development software but they are frequently not powerful enough for a gateway.

DESIRE handbook gives the example of SOSIG gateway, which at that time contained 7,000 records and had one member of staff assigned to link checking, who spent approximately one day a week going through the reports and updating or deleting records as appropriate [DESIRE 2.6].

#### 5.4.4 Technical maintenance and support

The need for maintenance and technical support will depend on the technology used for the gateway. Many server maintenance tasks can be easily automated but there is still a need for someone to monitor the system performance and deal with technical problems that may arise. This role should require no more than one day of human effort a week, even if a dedicated server is used, but the skilled staff should be available at all times. The same person should ideally have appropriate programming, scripting, and user interface design skills and understanding of networked environment and related issues, such as data integrity and security. The selection of the institution to host the server should be based on the availability of such skilled staff. Additional equipment to consider includes a mirror site server, and local tape drive or other archiving facility for backing up data, configuration and software.

In addition, the need for software and hardware support need to be taken into account. Selecting equipment from providers with good support plans is essential, especially if there is no enough internal expertise available. Public license software, on the other hand, gives the advantage of excellent community support.

### 5.5 PLANNING AND CONTROL

#### 5.5.1 Project management and coordination

There is a need to form a team that will oversee the entire project, coordinate the work of different nodes, and ensure the smooth operation of all teams.

Although collocation of members would be beneficial for the efficiency, the distributed nature of the gateway imposes the need for remote collaboration, and physical proximity of members should not be a factor in assigning the roles.

In the planning and development phases, all project team members should be available at least 1/3 of time. The need for their involvement will decrease in the maintenance phase, while the involvement of editors should remain constant through all phases, estimated at 4

hours/week. It is important to ensure that the staff who will be running the project in the later stages gets involved from the very beginning.

The project team and the editorial board will need to agree on:

- quality selection criteria and scope policies
- areas of responsibility
- organizational and political issues
- communication and workflow procedures
- metadata formats, cataloguing rules, and vocabularies.

### 5.5.2 Communication and meetings

Standard modes and frequency of communication and coordination need to be established. Regular meetings at all levels within the project are required to ensure that project work is progressing in the right directions. In the geographically distributed model, face-to-face meetings will probably have to be kept down to once a year. Weekly conference calls should be scheduled during the set-up phase; the frequency can later be adjusted according to the needs. The existing facilities for conference calls or video conferencing at participating nodes need to be assessed in order to find optimal mode of communication and assign funds for that purpose if required.

Mailing lists and discussion boards for all distinct teams should be extensively used as a liaison between the virtual meetings. A system for documenting and archiving discussion of project activities should be established.

A team Web site is needed for allowing team members password-protected access to shared documents.

Communication between the project team and the wider community can be established through mailing lists, a public project Web site, and dissemination events.

### 5.5.3 Promotion, dissemination, and reporting

The need for publicity and promotion of the gateway should not be overlooked. The methods may include: news releases, articles published in art and library journals, announcements on mailing lists and newsgroups, on libraries' Web sites, printed leaflets and posters, submitting to search engines, general directories and gateways, presentations at conferences and other events in both art and library communities, project Web site and so on. Although some of these methods are free, publicity and promotion need to be budgeted and the responsibility assigned. The plan and budget should include design of a logo, development of a branding policy and design guidelines.

Informal monthly progress reports should regularly be published and announced to all identified stakeholders. The most convenient way is probably posting to the internal project Web site and announcing through appropriate mailing lists.

A formal end-of-year report should be more widely published: posted on the project public Web site, announced through a news release, and presented at the annual IFLA conference and other appropriate events.

The responsibility for developing and maintaining the public and private Web sites needs to be assigned.

#### 5.5.4 Training

All contributors need training for the selection of resources, cataloguing and classification. Although face-to-face workshops are typically recommended as the best training method, due to the wide geographical dispersion of contributors, the options for remote training such as online tutorials and Webcast sessions should be planned instead. Similarly, the documentation needed for continuing support should be in the form of an online help system integrated into the editing interface. Printed copies may be needed for nodes without permanent Internet connection. Setting up a mailing list for discussion of issues that may arise and mutual support, regular editorial meetings and efficient communication channels should play an important role.

Time, funds, and responsibility for training need to be assigned. *Imageline* training may need to be two-level, editors in national nodes to be trained first in order to be able to train contributors within their countries, if broadly distributed cataloguing is envisioned.

#### 5.5.5 Quality control

Quality control should ensure that the gateway conforms to all specified quality requirements. This implies

- (a) the existence of written specifications for all aspects of the gateway including: scope statement, selection criteria, description standards, design and branding guidelines, language policy, workflows, link checking and content review schedule, communication and co-ordination procedures, evaluation criteria and procedures, system performance
- (b) delegation of duties to ensure conformance in all these areas.

Quality control should be performed at different levels: national editors, topic specialists, product manager, external reviewers.

The total quality management framework should be followed, at least informally, by ensuring quality at each step in the process. All participants should commit to continuously



provide comment, feedback, suggestions, and support to each other. All contributors should use the documentation and make sure that their selection and descriptions conform to the specified requirements. National and regional editors should make sure that resource descriptions within their scopes are consistent and collections well balanced, while topic centres need to take care that specific disciplines are well covered and the gateway conforms to the requirements in all its aspects. The product manager and the editorial board should ensure the conformance at the gateway level and analyze usage statistics and review the gateway on a quarterly basis. Yearly evaluation procedures should include: evaluation reports by the members of the advisory board, online user survey encompassing both end users and intermediaries, user log analysis, and usability testing if major changes have been implemented.

#### 5.5.6 Monitoring

It is necessary to monitor the use of the gateway: frequency of use over time, search logs, growth rate, usage across countries, and any other automatically trackable data. Product manager should also continuously monitor user feedback, suggestions for adding resources, performance of teams, and system performance.

#### 5.5.7 (Re-)evaluation

Due to the fast-paced developments in related areas, the project needs to be re-evaluated on a regular basis. Quality, usability, value to the user and perceived threats should be assessed. Aside from internal evaluation and user surveys, an advisory board with the «role to shape the overall goals and objectives of the collection, to confirm that the project remains on course over time, and to address emerging issues» [Pitschmann] should be appointed. The one-year span between IFLA conferences can be regarded as an appropriate interval—allowing enough time to accomplish some measurable results and not too long to allow any important trends or omissions to pass unnoticed, without enough time to react. Another advantage is the opportunity to have together experts from various areas, who can serve as members of the advisory board. A formal End of Year Report should be submitted at these occasions.

Recently, the Dlib Metrics Working Group (MWG) developed quantitative performance measures [Larsen 2002] that can be used for the evaluation of the system.

#### 5.5.8 Intellectual property rights

Intellectual property rights over resource descriptions may become important in case of the cessation of membership, desire to re-use in other context or in relationships with broker-type services. It is recommended to apply a dual ownership model for records, which would thus be owned by both IFLA and the contributing node, each party having full rights over them.

### 5.5.9 Future proofing

Future proofing should involve definition of change management strategy, identification of threats, potential partnerships, interoperability, and scalability requirements. These documents should be annually reviewed.

## 6 TECHNICAL OPTIONS

### 6.1 SUMMARY OF REQUIREMENTS FOR *IMAGELINE* HARDWARE AND SOFTWARE

- Database to store metadata records—no special requirements to store huge number of records, large file sizes, or complex data structures
- Database indexing software
- User interface for remote editing of records through any Web browser on any platform
- Access control mechanism for restricting editorial rights to a limited group of contributors. No special security requirements.
- User interface for remote searching and browsing through any Web browser
- Enough memory to allow concurrent searching by a large number of end-users
- Capability to hide certain records from end user's view
- Support for the qualified Dublin Core metadata model
- Possibility to extend the model to include custom elements and qualifiers
- Possibility to specify constraints on the metadata records, such as mandatory fields, restricted values etc.
- XML/RDF import/export capability
- Z39.50 support
- Possibility to integrate authority files
- Capability to generate different reports, usage statistics, log files etc.
- Link checking,
- Data backup utilities
- It should be possible to set up a pilot within three months and fully operational system within the next three

### 6.2 HARDWARE AND OPERATING SYSTEM OPTIONS

The choice of the operating system should be made between Windows NT and UNIX-based operating systems, since the Macintosh is rarely used as a database platform.

The gateway software is typically installed on a computer running a proprietary UNIX variant, such as a Sun system running a proprietary Solaris, Digital UNIX, or HP-UX, or on an Intel-based server running Linux or Windows NT. UNIX has a reputation of not being secure enough but its security performance is not inferior to Windows' to that extent that it would outweigh all its other benefits, since Linux runs faster, is more scalable, and more affordable. A Windows NT system should be considered only in case there is no expertise available for administering and maintaining a UNIX-based system.

An Intel Pentium-based server running Linux will probably give the best value for money. There are no special requirements for any of the features of an entry-level Sun system, for example, which costs approximately US\$ 11,000. A powerful Intel-based DELL server, for example, with all software utilities (Apache Web server, MySQL database, user management, replication, remote administration etc.) installed and configured costs less than \$US 4,000, although a functional configuration can be assembled for as low as \$US 1,000-2,000.

Data backup options: RAID architecture, automatically copying data to a second disk partition, regular archiving of data to a tape.

### 6.3 OPTIONS

As usual, there are two basic options to choose from—buy vs. build. Currently, there are many diverse open source products on the market, which could be combined to provide the required functionality with no investment in commercial software. Such systems can be fine tuned to meet the very specific requirements of the project. However, such approach requires programming effort, time, and expertise. The «buy» option can be implemented faster but typically requires compromises in functionality.

From the technical point of view, subject gateways do not differ much from any database-backed Web site, meaning that specialized tools are not a must. The options include:

- Building from scratch using a freely available relational database system, such as MySQL or PostgreSQL--possible if programming resources are available.
- Using a general web publishing tool; there are free tools available in this category as well, (e.g. Zope or OpenACS). This option also requires some software development resources but is easier to setup and run.
- Research developed subject gateway software and partnering with other projects (e.g. ROADS, iMesh toolkit, OCLC SiteSearch, ELVIL nexus, Infomine iVia: this option can mean
- Commercial subject gateway software (Blue Angel Technologies MetaStar, System Simulation's Index+)
- Outsourcing consulting service to assemble, implement, and maintain a functional system (Indexdata MetaCat)

## 6.4 EVALUATION CRITERIA

General criteria for evaluating the options include:

- affordability
- ease of implementation and maintenance
- availability of required expertise
- flexibility,
- reliability
- responsiveness
- efficiency
- scalability.

The main factor influencing the decision will be the availability of software development resources among the membership. Although there are real costs associated with the development, maintenance, and use of open source software, these are always lower than those associated with reliance upon commercial software. Continuing maintenance and support are also crucial. Dependence on a single commercial provider, consulting service, or research project may be risky, open source community is providing efficient and continuous support. On the other hand, the ease and speed of implementation can be more important. Several commercial packages have been successfully tested in subject gateways and may not require compromises in functionality.

## 7 ORGANIZATIONAL MODEL

An organizational model based on a two-dimensional grid of national nodes and topic centres is suggested as the most suitable for the project. Each country represented in the IFLA Section of Art Libraries should select a national node which would coordinate submission of resources originating from the respective country by contributors from that country. National libraries are typically recommended as the best candidates for that role because of their general role of bringing together national resources, respect and leadership. However, in many cases, leaders of national consortia, union catalogues, digital libraries, institutions responsible for national heritage may be more appropriate, depending on specific circumstances, such as availability, commitment, experience, expertise and so on. Representatives of the national nodes should act as both editors and co-ordinators within their scopes, making sure that the coverage of national resources is reasonably complete, well balanced, and conforming to all set quality criteria. Distribution of contributors responsibilities within a country should be at the discretion of the national node.

At the same time, each national node would act as a topic centre. Some topic centres should be responsible for cross-national coverage of specific disciplines, making sure that the

coverage of the discipline is well covered in terms of styles/periods, regions, most important artists and works etc. Other topic centres would be responsible for specific aspects of the project, such as technical development and support, collaboration and interoperability with related initiatives, indexing standards and quality and so on. Depending on circumstances, a national node may want to assign the topic node role to another member in the same country, however, the primary responsibility for the topic would still remain with the national node.

National nodes should be responsible for the provision of a user interface in the language(s) of respective countries and any kind of its customization that may be needed, translations of free text fields, and vocabularies. It should be possible for members from other countries to add translation of free-text fields and additional classification codes.

National nodes should be coordinated by a Project Team. The team should include:

- Project manager, responsible for co-ordination of work
- Product manager, responsible for overall quality of the gateway
- Technology officer, responsible for its technical performance
- Indexing and cataloguing officer responsible for providing access to content
- Content officer, responsible for the well balanced high quality collection

Representatives of national/regional nodes would make an Editorial Board, headed by the content officer. The board should collaborate closely with the project team and serve as liaison with all contributors.

An Advisory Board should also be formed with the mission to mould the goals of the project, to make sure that it remains on course over time or change the course if needed, to warn about emerging trends, and address emerging issues. The board should include representatives of user community and specialists for various aspects of the project, such as subject specialists, specialists for indexing and classification, for digital libraries, Internet technologies etc., all leaders and authorities in their fields.

## 8 BUSINESS MODEL AND SUSTAINABILITY

Sustainability should be the primary concern in selecting the business model for the gateway. The most likely model would be a combination of "public investment as part of the role of cultural, educational or scholarly institutions" and "collective activity through membership," as defined in Renardus deliverables [Day et al.]. The model should be based on the perception of the gateway as an integral part of each library's collection, so that its development makes part of each member's collection management. Thus, each member library should be willing to contribute either in-kind or financially towards its development and maintenance. An initial funding will be needed to start up the project, primarily for the

acquisition of hardware and software. Potential sources of funding need to be identified (e.g. member pledges, research and other funding agencies).

The option involving in-kind member contributions requires sensible delegation of tasks and responsibilities as well as clear policies for giving proper credits for efforts. Compensation for extra efforts and any external expenses should be planned, however, and potential sources of funding identified.

Reliance on monetary contributions would require fair calculation of dues. These may be based on member budgets, per capita of served population, per collection size but should take into account the differences in economic conditions across the world.

If opportunities to obtain research grants are to be explored, best bets lie in partnering with related projects in the areas of multilingual domain-specific vocabularies, multilevel access to information, and integration into various broker architectures.

## 9 SUMMARY OF EXPENSES

The factors to be taken into account include:

Planning phase:

role assignment, documentation development, metadata schema definition, vocabulary decisions, technology evaluation, equipment acquisition, identification of potential partners, co-ordination, meetings, communication, travel, consumables.

Development phase:

software development or configuration, information architecture definition, user interface design and translation, vocabulary integration or development, logo design, promotional materials, private and public Web site development, resource selection, and description, editorial work, usability testing, co-ordination, meetings, communication, promotion, travel, consumables.

Maintenance:

resource addition, updating, weeding, editorial work, server maintenance, troubleshooting, data backups, link checking, use log analysis, evaluation, improvement, co-ordination, meetings, communication, promotion, travel, consumables.

Annual operational costs, according to a conservative rough estimate, can be expected to amount to \$US 200,000 in the first year and \$US 50,000 in subsequent years.

The assumption is that equipment such as personal computers, printers, Internet connection can be used and not included in project expenses.

## REFERENCES:

- Ardö A. et al. "Technical requirements for participating organizations: Renardus deliverable D2.2," 2001. Available at: [http://renardus.lub.lu.se/about\\_us/deliverables/d2\\_2/D2\\_2\\_final.html](http://renardus.lub.lu.se/about_us/deliverables/d2_2/D2_2_final.html)
- ArtLex [<http://www.artlex.com>]
- Bawden, D. *User-oriented evaluation of information systems and services*. Aldershot: Gower, 1990.
- Becker, H. J. et al., "Evaluation report of partner subject gateways," 2000. Available at: [http://www.renardus.org/about\\_us/deliverables/d6\\_1/D6\\_1\\_final.pdf](http://www.renardus.org/about_us/deliverables/d6_1/D6_1_final.pdf)
- Besser, H. and R. Yamashita, *The Cost of Digital Image Distribution: The Social and Economic Implications of the Production, Distribution, and Usage of Image Data*, 1998. Available at: <http://sunsite.berkeley.edu/Imaging/Databases/1998mellon>
- Best Practice Checklist for Australian Subject Gateways [<http://www.nla.gov.au/initiatives/sq/bestpractice.html>]
- BUBL LINK [<http://bubl.ac.uk/link>]
- CIMI. Guide to Best Practice: Dublin Core, Version 1.1, 21 April 2000. Available at: [http://www.cimi.org/public\\_docs/meta\\_bestprac\\_v1\\_1\\_210400.pdf](http://www.cimi.org/public_docs/meta_bestprac_v1_1_210400.pdf)
- Collection Description Focus / CIMI "Summary Report on Survey Questionnaire," April 2002. Available at: <http://www.ukoln.ac.uk/cd-focus/questionnaire/report.html>
- Cole, Timothy W., Joanne Kaczmarek, Paul F. Marty, Christopher J. Prom, Beth Sandore, and Sarah Shreeves, "Now That We've Found the 'Hidden Web,' What Can We Do With It?: The Illinois Open Archives Initiative Metadata Harvesting Experience", University of Illinois at Urbana-Champaign, USA 2002, *Museums and the Web 2002*. Available at: <http://www.archimuse.com/mw2002/papers/cole/cole.html>
- Combine [<http://www.lub.lu.se/combine/>]
- Day, M., *Metadata: Mapping Between Metadata Formats*, UKOLN, 2002. Available at: <http://www.ukoln.ac.uk/metadata/interoperability/>
- Day, M. et al. "Business issues which impact the functional model: Renardus deliverable 8.1," September 2000. Available at: [http://www.renardus.org/about\\_us/deliverables/d8\\_1/D8\\_1\\_final.pdf](http://www.renardus.org/about_us/deliverables/d8_1/D8_1_final.pdf)
- Delivering Digital Images: Cultural Heritage Resources for Education*, Christie Stephenson and Patricia McClung, eds. Getty Information Institute 1998. (has some survey. Fix the bib entry)
- Dempsey, L. "The subject gateway: experiences and issues based on the emergence of the Resource Discovery Network," *Online Information Review*, 24(1), 2000, 8-23.
- Dempsey, L., R. Heery, M. Hamilton, D. Hiom, J. Knight, T. Koch, M. Peereboom & A. Powell, *A review of metadata: a survey of current resource description formats (DESIRE deliverable D3.2 (1), March 1997)*. Available at: <http://www.ukoln.ac.uk/metadata/desire/overview/>
- DESIRE Information Gateways Handbook. DESIRE Consortium, 2000. Available at: <http://www.desire.org/handbook/>.
- Digital Library Federation (DLF), "A New Approach to Finding Research Materials on the Web," June 2001. Available at: <http://www.diglib.org/architectures/vision.htm>
- Digital Library Foundation (DLF), "Shared Cataloguing Tool for Visual Resources," April 2001. Available at: <http://www.diglib.org/collections/vrtool/vrtgetty.htm>
- Dublin Core, <http://purl.oclc.org/dc>.
- Dunn, H. "Collection Level Description - the Museum Perspective," *D-Lib Magazine* 6 (9), September 2000. Available at: <http://www.dlib.org/dlib/september00/dunn/09dunn.html>
- EEVL [<http://www.eevl.ac.uk/>]

ELISE II: Electronic Library Image Service for Europe - Phase II: Supplementary Information: Summary findings of the image-supply survey analysis. Available at: [http://severn.dmu.ac.uk/elise/el2\\_dels/d1\\_9sum.htm](http://severn.dmu.ac.uk/elise/el2_dels/d1_9sum.htm)

ELVIL: Public Reports [<http://elvil.sub.su.se/sam/elvil.htm>]

Ferry, Alison, "ADAM 1996 Survey of User Information Needs and Search Methods Results," <http://adam.ac.uk/adam/reports/survey/>

Gaudin, S. "Is Your Site Getting The 'Internet Death Penalty'?", *E-Commerce Archives* May 30, 2002. Available at: [http://itmanagement.earthweb.com/ecom/article/0,,11952\\_1145241,00.html](http://itmanagement.earthweb.com/ecom/article/0,,11952_1145241,00.html)

Getty Research Institute, "Metadata Standards Crosswalk." Available at: [http://www.getty.edu/research/institute/standards/intrometadata/3\\_crosswalks/crosswalk1.html](http://www.getty.edu/research/institute/standards/intrometadata/3_crosswalks/crosswalk1.html)

Graham, M. E., The Description and Indexing of Images: Report of a survey of ARLIS members, 1998/99. University of Northumbria at Newcastle, Institute for Image Data Research, May 1999. Available at: <http://www.unn.ac.uk/iidr/ARLIS/>

Grout, C. et al. "Creating Digital Resources for the Visual Arts: Standards and Good Practice, Visual Arts Data Service and Technical Advisory Service for Images." Available at : [http://vads.ahds.ac.uk/guides/creating\\_guide/contents.html](http://vads.ahds.ac.uk/guides/creating_guide/contents.html)

Harvest\_NG [<http://webharvest.sourceforge.net/ng/>]

Hastings, Samantha K., "[Evaluation of Image Retrieval Systems: Role of User Feedback](#)," *Library Trends* 48 (2), fall 1999, "Progress in Visual Information Access and Retrieval", Edited by Beth Sandore.

Hastings, S. K., "A Comparison of Retrieval Problems for Digital Images in a Distributed Network versus a Closed System." Available at: <http://courses.unt.edu/shastings/homepage/research/ichim97.html>

Hastings, S. K., "Query Categories in a Study of Intellectual Access to Digitized Art Images," *Proceedings of the ASIS Annual Meeting* v32 p3-8 1995.

Headline Newsletter: Hybrid electronic access and delivery in the library networked environment. What users want from the hybrid library? 1999. Available at: <http://www.headline.ac.uk/newsletter.pdf>

Heaney M., "An Analytical Model of Collections and their Catalogues," January 2000. Available at: <http://www.ukoln.ac.uk/metadata/rsip/model/amcc-v31.pdf>

Heery, R. "Information gateways: collaboration on content" *Online Information Review*, 24(1), 2000, 40-45.

Heery, R. and M. Patel. "Application profiles: mixing and matching metadata schemas," *Ariadne* 25. September 2000. Available at: <http://www.ariadne.ac.uk /issue25/app-profiles/intro.html>

Hofman, P. and E. Worsfold, Selection Criteria for Quality Controlled Information Gateways: Work Package 3 of Telematics for Research project DESIRE (RE 1004) 1997 <http://www.ukoln.ac.uk/metadata/desire/quality/>

IFLA Section of Art Libraries, *International Directory of Art Libraries*. Available at: <http://iberia.vassar.edu/ifla-idal/>

IFLA Section of Art Libraries, *Multilingual Glossary for Art Librarians* [<http://www.ifla.org/VII/s30/pub/mg1.htm>]

IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records*. München: Saur, 1998. Available at: <http://www.ifla.org/VII/s13/frbr/frbr.pdf>

Images Online: Perspectives on the Museum Educational Site Licensing Project, Patricia McClung and Christie Stephenson, eds. Getty Information Institute 1998.

Infomine [<http://infomine.ucr.edu/>]



- Isaac Network [<http://scout.cs.wisc.edu/research/isaac/>]
- Jones M.L.W., Gay G.K., Rieger R.H., "Project Soup: Comparing Evaluations of Digital Collection Efforts," *D-Lib Magazine* Volume 5 Number 11, November 1999. Available at: <http://www.dlib.org/dlib/november99/11jones.html>
- Kidsclick [<http://www.kidsclick.org/URLsubmit.html>]
- Kirriemuir, J., "A Brief Survey of Quality Resource Discovery Systems," Resource Discovery Network Centre, September 1999. Available at: <http://www.rdn.ac.uk/publications/studies/survey/>
- Koch, T. "Quality-controlled subject gateways: definitions, typologies, empirical overview," *Online Information Review* 24:1, February 2000.
- Koch, T. "The Renardus broker: a "meta subject gateway," 2001 Available at: <http://www.lub.lu.se/tk/renardus/tokyoren.html>.
- Larsen R., "The DLib Test Suite and Metrics Working Group: Harvesting the Experience from the Digital Library Initiative," April 2002. Available at: [http://www.dlib.org/metrics/public/papers/The\\_Dlib\\_Test\\_Suite\\_and\\_Metrics.pdf](http://www.dlib.org/metrics/public/papers/The_Dlib_Test_Suite_and_Metrics.pdf)
- LII (Librarians' Index to the Internet) [<http://www.ii.org/>]
- Medina, Alicia García and Teresa Coso. "Graphic resources in the Spanish Art Libraries," 67th IFLA Council and General, Conference, August 16-25, 2001. Available at: <http://www.ifla.org/IV/ifla67/papers/123-169e.pdf>
- Miller P., "Collected Wisdom: Some Cross-domain Issues of Collection Level Description," *D-Lib Magazine* 6(9), September 2000. Available at: <http://www.dlib.org/dlib/september00/miller/09miller.html>
- Mitchell, S. and M. Mooney, "INFOMINE: A Model Web-based Academic Virtual Library," *Information Technology and Libraries* 3/96, v.15 (1). Available at: [<http://www.lita.org/ital/index.htm>]
- Monopoli, M. and D. Nicholas. "A user evaluation of Subject Based Information Gateways: Cast study ADAM," *Aslib Proceedings*; London; Jan 2001; v.53 n.1, p.39.
- National Library of Canada. *Inventary of Canadian Digital Initiatives* [<http://www.nlc-bnc.ca/initiatives/erella.htm>]
- De Nederlandstalige Art & Architecture Thesaurus*. <http://www.rkd-db.nl/aat/index.html>
- Nielsen, J. "Top Ten Mistakes of Web Management," *Jakob Nielsen's Alertbox*, June 15, 1997. Available at: [<http://www.useit.com/alertbox/9706b.html>]
- Pitschmann, L. A. Building Sustainable Collections of Free Third-Party Web Resources, Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, June 2001. Available at: <http://www.clir.org/pubs/reports/pub98/pub98.pdf>
- Powell, A. et al. "Collection Level Description: A review of existing practice," August 1999. Available at: <http://www.ukoln.ac.uk/metadata/cld/study/>
- Project Proposal for Guide to Good Practice: Cataloging Standards for Describing Cultural Objects and Images [<http://www.diglib.org/standards/vrwork.htm>]
- Renardus [<http://renardus.lub.lu.se/cgi-bin/egwcgi/egwrtcl/screen.tcl/name=Start&lang=eng&service=ren>]
- RSLP Collection Level Description, May 2000. Available at: <http://www.ukoln.ac.uk/metadata/rsdp/schema/>
- Sandore, B., N. Shaik and Y. Zhang "Using Museum Images In the Classroom and Beyond: *The Museum Educational Site Licensing Project*," 1997. Available at: <http://images.grainger.uiuc.edu>
- Schneider, K. "A Yahoo with Values and a Brain." *American Libraries* , April 2002. Available at: <http://www.ala.org/online/netlib/il402.html>.
- Solock, Jack. "Anatomy of a Scout Report: Resource Discovery in the Information Age, or How We Do It," March 1997. Available at: <http://scout.cs.wisc.edu/addserv/toolkit/enduser/archive/1997/euc-9703.html>

THAIS [<http://www.thais.it>]

Trant, J., D. Bearman, and K. Richmond, "Collaborative Cultural Resource Creation: the example of the Art Museum Image Consortium," 2001. Available at:

<http://www.amico.org/docs/papers/2000/amico.mw2000.pdf>

Troll, Denise A., How and Why Are Libraries Changing? Digital Library Federation, January 9, 2001. Available at: <http://www.diglib.org/use/whitepaper.htm>

University of Arizona *Clearinghouse of Image Databases*

[<http://www.library.arizona.edu/images/clearinghouse/clearinghouse.html>]

Wallace, I and V. Webb, "ADAM Subject Headings: ASH: A classification scheme for the Art, Design, Architecture and Media Information Gateway," 22 October 1998. Available at:

<http://adam.ac.uk/adam/public/SF3/SF3f9807.rtf>

## APPENDIX 1: METHODOLOGY

### WHAT

1. State the mission/purpose/raison d'être of the proposed service
2. Identify and compare the need and the existing services ("demand and supply")
  - a. Identify primary end-users and their information needs
  - b. Identify secondary users and their needs
  - c. Identify service providers, their missions and interests
  - d. Identify original source providers, their missions and interests
  - e. Assess the level to which the need is currently satisfied
    - i. Are sufficient required resources available on the Internet and only need to be made accessible?
    - ii. Are those resources accessible through any of the existing gateways or lists of Internet resources?
    - iii. By general search engines and directories?
    - iv. By locally available non-Internet resources?
  - f. Analyze the needs across all dimensions (subject matter, types of resources, languages, currency, depth etc.)
  - g. Identify the most frequent and most widely spread needs
  - h. Identify the least satisfied needs.
3. Identify relevant related developments—new technologies, types of services etc. and potential implications thereof
  - a. risks and risk mitigation strategies
  - b. opportunities and ways to capitalize on them
4. State whether and why there is a real need for a new service
5. Refine or modify mission statement if needed.
6. Identify distinguishing features of the proposed service and specific challenges.
7. Define the scope and identify priorities.
8. Define scope and content related "future-proofing" strategies

### HOW

1. Identify obstacles and opportunities specific to the project.
2. Identify steps/tasks involved in the development of the service starting from the following list:
  - Providing content
    - o Harvesting
    - o Evaluating and selecting
    - o Controlling editorial access
    - o Submitting

- Coordinating additions
  - Storing/persisting
- Providing access to content
  - Describing resources
  - Search mechanism
  - Browsing mechanism and classification
  - Ranking
- Maintenance
  - Updating
  - Link checking
  - Technical maintenance and support
  - Obtaining and incorporating user feedback
- Control
  - Project management and coordination
  - Meetings
  - Communication
  - Quality control
  - (Re-)evaluation
- 3. For each task identify available options in terms of level of automation.
- 4. For each task and option assess staffing needs, technology/equipment requirements, and associated costs
  - a. Start-up requirements
  - b. Long-term operational requirements.
- 5. Summarize technical requirements, summarize options, and define evaluation criteria
- 6. Evaluate the options, if possible, and make recommendations or basic guidelines.

## WHO

1. Summarize staffing needs.
2. Define organizational model.
3. Identify potential modules—subdivisions of subject areas, types of resources, processes, or other categories that can be used for incremental or distributed development.
4. Recommend criteria for populating the model.
5. Define cost elements and estimate costs.
6. Define business model.
7. Identify sources of funding.
8. Define strategy for sustainability.

## APPENDIX 2: CLASSIFICATION SCHEMES

[ART IN GENERAL CLASSIFICATION SCHEMES](#)

[ART SPECIFIC SCHEMES](#)

[EXAMPLES OF ART AREAS IN LOCAL CLASSIFICATION SCHEMES](#)

[ADDITIONAL SOURCES OF INFORMATION](#)

[CHECKLIST FROM DESIRE HANDBOOK](#)

## ART IN GENERAL CLASSIFICATION SCHEMES

DDC	UDC
<p> <b>700</b> <a href="#">The arts, Fine and decorative arts</a>  <b>710</b> <a href="#">Civic and landscape art</a>  <b>720</b> <a href="#">Architecture</a>  <b>730</b> <a href="#">Plastic arts, Sculpture</a>  <b>740</b> <a href="#">Drawing and decorative arts</a>  <b>750</b> <a href="#">Painting and paintings</a>  <b>760</b> <a href="#">Graphic arts, Printmaking and prints</a>  <b>770</b> <a href="#">Photography and photographs</a>  <b>780</b> <a href="#">Music</a>  <b>790</b> <a href="#">Recreational and performing arts</a> </p>	<p>           7 The arts. Recreation. Entertainment. Sport  <b>71</b> <a href="#">Physical planning. Regional, town and country planning</a>  <b>72</b> <a href="#">Architecture</a>  <b>73</b> <a href="#">Plastic arts</a>  <b>74</b> <a href="#">Drawing. Design, applied arts and crafts</a>  <b>75</b> <a href="#">Painting</a>  <b>76</b> <a href="#">Graphic art. Graphics</a>  <b>77</b> <a href="#">Photography and similar processes</a>  <b>78</b> <a href="#">Music</a>  <b>79</b> <a href="#">Recreation. Entertainment. Games. Sport</a> </p>
<p style="text-align: center;"><b>LCC</b></p>	<p>Special auxiliary table</p>
<p> <b>N</b> Visual Arts  <b>NA</b> Architecture  <b>NB</b> Sculpture  <b>NC</b> Drawing. Design. Illustration  <b>ND</b> Painting  <b>NE</b> Print Media  <b>NK</b> Decorative Arts  <b>NX</b> Art in General         </p>	<p>           .01 Theory and philosophy of art            .02 Technique. Craftsmanship            .03 Artistic periods and phases. Schools, styles, influences            .04 Subjects for artistic representation. Iconography. Iconology            .05 Applications of art (in industry, trade, the home, everyday life)            .06 Various questions concerning art            .07 Occupations and activities associated with the arts and entertainment            .08 Characteristic features, forms, combinations etc. (in art, entertainment and sport)            .091 Performance, presentation (in original medium)         </p> <p>Common Auxiliaries</p> <p>Table</p> <p> <b>1a</b> Co-ordination. Addition  <b>1b</b> Relation. Subgrouping. Order-fixing  <b>1c</b> <a href="#">Common auxiliaries of language</a>  <b>1d</b> <a href="#">Common auxiliaries of form</a>  <b>1e</b> <a href="#">Common auxiliaries of place</a>  <b>1f</b> <a href="#">Common auxiliaries of ethnic grouping and nationality</a>  <b>1g</b> <a href="#">Common auxiliaries of time</a>  <b>1h</b> Specification by non-UDC notation (e.g. 1/9, A/Z)  <b>1k</b> </p> <p>Common auxiliaries of</p> <p> <a href="#">properties</a>  <a href="#">materials</a>  <a href="#">persons and personal characteristics</a> </p>

**ART SPECIFIC SCHEMES**

<b>ADAM Subject Headings</b>	<b>AAT Facets</b>	<b>AAT Hierarchies</b>	<b>Iconclass</b>
<a href="#">Applied Arts</a> <a href="#">Architecture</a> <a href="#">Design</a> <a href="#">Fine Art</a> <a href="#">General Arts Resources</a> <a href="#">Media</a> <a href="#">Museum Studies</a>	ASSOCIATED CONCEPTS	Associated Concepts	Abstract, Non-representational Art Religion and Magic Nature Human Being, Man in General Society, Civilization, Culture Abstract Ideas and Concepts History Bible Literature Classical Mythology and Ancient History
	PHYSICAL ATTRIBUTES	Attributes and Properties Conditions and Effects Design Elements Color	
	STYLES AND PERIODS	Styles and Periods	
	AGENTS FACET	People Organizations	
	ACTIVITIES	Disciplines Functions Events Physical Activities Processes and Techniques	
	MATERIALS	Materials	
	OBJECTS	Object Groupings and Systems Object Genres Components Settlements and Landscapes Built Complexes and Districts Single Built Works OpenSpaces and Site Elements Furnishings Costume Tools and Equipment Weapons and Ammunition Measuring Devices Containers Sound Devices Recreational Artifacts Transportation Vehicles Visual Works Exchange Media	

**EXAMPLES OF ART AREAS IN LOCAL CLASSIFICATION SCHEMES**

<b>AGORA</b>	<b>FINISH VIRTUAL LIBRARY</b>
<p><b>ART AND PHOTOGRAPHY</b></p> <p>&gt;&gt; <a href="#">General</a></p> <p>&gt;&gt; <a href="#">Architecture</a></p> <p>&gt;&gt; <a href="#">Photography</a></p>	<p><b>Culture, language &amp; education</b></p> <p><a href="#">Alcohol, drugs and other substance abuse</a></p> <p><a href="#">Anthropology and folkloristics</a></p> <p><a href="#">Applied linguistics</a></p> <p><a href="#">Architecture, interior design and furniture design</a></p> <p><a href="#">Art education</a></p> <p><a href="#">Community and Regional Planning</a></p> <p><a href="#">Education</a></p> <p><a href="#">Fashion and clothing design</a></p> <p><a href="#">Film, stage design, and photography</a></p> <p><a href="#">Finnish art music, jazz and folk music</a></p> <p><a href="#">Gerontology</a></p> <p><a href="#">Graphic design and new media</a></p> <p><a href="#">Hydrobiology and limnology</a></p> <p><a href="#">Industrial design and art handicraft</a></p> <p><a href="#">Literary research</a></p> <p><a href="#">Psychology</a></p> <p><a href="#">Sexology</a></p> <p><a href="#">Sport science</a></p> <p><a href="#">Textile art</a></p> <p><a href="#">Theatre and dance</a></p> <p>While:</p> <p><b>Society and economy</b></p> <p>Community and regional planning</p>



## **ADDITIONAL SOURCES OF INFORMATION**

- Koch, T. "The role of classification schemes in Internet resource description and discovery: Work Package 3 of Telematics for Research project DESIRE (RE 1004)," February 1997. Available at: [http://www.ukoln.ac.uk/metadata/desire/classification/class\\_tc.htm](http://www.ukoln.ac.uk/metadata/desire/classification/class_tc.htm)

Comparison of different types of classification schemes covering:

- extent of usage in Internet services
- extent of usage in traditional and other online services
- multilingual capability
- strengths and weaknesses of the scheme
- integration between classification scheme and other systems e.g. controlled subject headings
- linking to third party classification data
- digital availability
- copyright
- extensibility and development effort of the scheme

- ADAM. "Strategy for the Use of Vocabulary in the ADAM Database," 1996. Available at: <http://adam.ac.uk/adam/public/SF1/SF1f9605.rtf>

Evaluation of available schemas and vocabularies and recommendations.

- Wallace, I and V. Webb, "ADAM Subject Headings: ASH: A classification scheme for the Art, Design, Architecture and Media Information Gateway," 22 October 1998. Available at: <http://adam.ac.uk/adam/public/SF3/SF3f9807.rtf>

Describes ADAM's experiences with using DDC for browsing, why they decided to create a scheme of their own and shows its full structure.

## **CHECKLIST FROM DESIRE HANDBOOK**

- do I want to use a classification scheme?

- What are the pros and cons?
- Which schemes are available?
- How do I decide which one is the most appropriate scheme for my service?
- Is it better to design my own scheme instead of using an existing scheme?
- Can I adapt or extend existing schemes?
- is it useful to adopt a keyword system as well as a classification scheme?
  - What are the pros and cons of using controlled and uncontrolled vocabularies
- will my users require both searching and browsing facilities?
  - Is there an existing classification scheme which might be the best basis for a browsing structure or could a thesaurus or keyword system be adapted for this purpose?
  - How do I create a browsing structure from a classification scheme?
- how will my choices affect interoperability issues?
- how will my choices affect multilingual issues?

## APPENDIX 3: VOCABULARIES

### [Art & Architecture Thesaurus \(AAT\)](#)

[<http://www.getty.edu/research/tools/vocabulary/aat/index.html>]

- Partly translated into Dutch [<http://www.rkd.nl/prjcts/aat-e.htm>]
- Partly translated into French  
[[http://daryl.chin.gc.ca/Artefacts/f\\_MasterLayout.cgi?la=f&db=1&style=5&realm=5](http://daryl.chin.gc.ca/Artefacts/f_MasterLayout.cgi?la=f&db=1&style=5&realm=5)]  
CHIN willing to share the translation
- No details about the progress of the collaborative project made between the Getty Information Institute and the Centro de Documentación de Bienes Patrimoniales to translate it into Spanish  
[[http://severn.dmu.ac.uk/elise/el2\\_dels/d42\\_1d.htm#\\_Toc413726824](http://severn.dmu.ac.uk/elise/el2_dels/d42_1d.htm#_Toc413726824)].
- Developed from numerous existing terminologies with terms pointing to corresponding terms in other terminologies, such as: LCSH, RIBA Architectural Keywords, RILA (International Repertory of the Literature of Art Subject Headings) etc.
- Used in numerous relevant projects worldwide [e.g. [ADAM](#), [AMICO](#), [CHIN/RCIP](#), [HELIX](#), [EOLE](#), [NAL](#), [OhioLink](#) etc.], all tests proved it to be very suitable for the purpose, even for domains broader than Fine Arts [ELISE II, CHIN].
- Can be implemented using free tools (Zebra index and retrieval engine, which is free for commercial use implemented in ELISE II)
- A potential problem is the size of the thesaurus, which may require trimming to the subset required for indexing *Imageline* content. CHIN reported some more specific technical problems that may need to be solved in specific implementations.

### [MACS \(Multilingual Access to Subjects\)](#)

[<http://infolab.kub.nl/prj/macs/>]

- provides mapping between three general subject heading lists--U.S. LCSH, French RAMEAU, and German SWD/RSWK.
- planning extension with other languages and other vocabularies
- still in development, art domain not covered yet
- technical and licensing details regarding possible use in subject gateways not known
- huge, majority of terms not relevant to *Imageline* domain

### [Le Répertoire de vedettes-matière \(RVM\)](#)

[<http://www.bibl.ulaval.ca/adele/rvm.html>]

- partially bilingual (French and English) encyclopaedic thesaurus

- started from creating French equivalents to parts of Library of Congress Subject Headings (LCSH) and Canadian Subject Headings (HSC), later started adding MeSH and AAT terms
- licensing fees vary from 500\$ to 2970\$ for the first year and from 400\$ to 2250\$ per annum for subsequent years

#### [Thesaurus of Geographic Names \(TGN\)](#)

[<http://www.getty.edu/research/tools/vocabulary/tgn>]

- covers all continents and nations of the modern political world, historical places, both physical features and administrative entities, such as cities and nations, with an emphasis on places important for art and architecture
- The terms in the TGN are multilingual, but the database interface is available only in English
- Widely used, in combination with AAT and with other subject vocabularies (e.g. [ETB](#)).

#### [Thesaurus of Graphic Materials I: Subject Terms \(TGM I\)](#)

[<http://lcweb.loc.gov/rr/print/tgm1>]

- intended for subject indexing of **graphical materials**, including historical photographs, architectural drawings, artwork, etc.
- no translations, not much used outside U.S.

#### [Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms \(TGM II\)](#)

[<http://lcweb.loc.gov/rr/print/tgm2>]

- provides headings for categories of material - by genre (portraits, etc.), by vantage point, method of representation, production technique or version, marking, shape or size, intended purpose, characteristics of the image's creator, or publication status.
- no terms denoting artistic movements and styles
- no translations, not much used outside U.S.

#### [UNESCO Thesaurus](#)

[<http://www.ulcc.ac.uk/unesco/>]

- general thesaurus, intended for indexing UNESCO resources
- trilingual French, English, Spanish
- widely used, mostly in archives
- can be purchased in a package including UNESCO databases
- small, easily manageable but may not be sufficient for the needs of *Imageline*

## [Union List of Artist Names \(ULAN\)](#)

[<http://www.getty.edu/research/tools/vocabulary/ulan/index.html>]

- Contains artists' names in vernacular, English, other languages, natural order, inverted order, or nicknames and some additional information about artists.
- The coverage is from Antiquity to the present, and the scope is global.
- The artist names in ULAN are multilingual, but the database interface is available only in English.

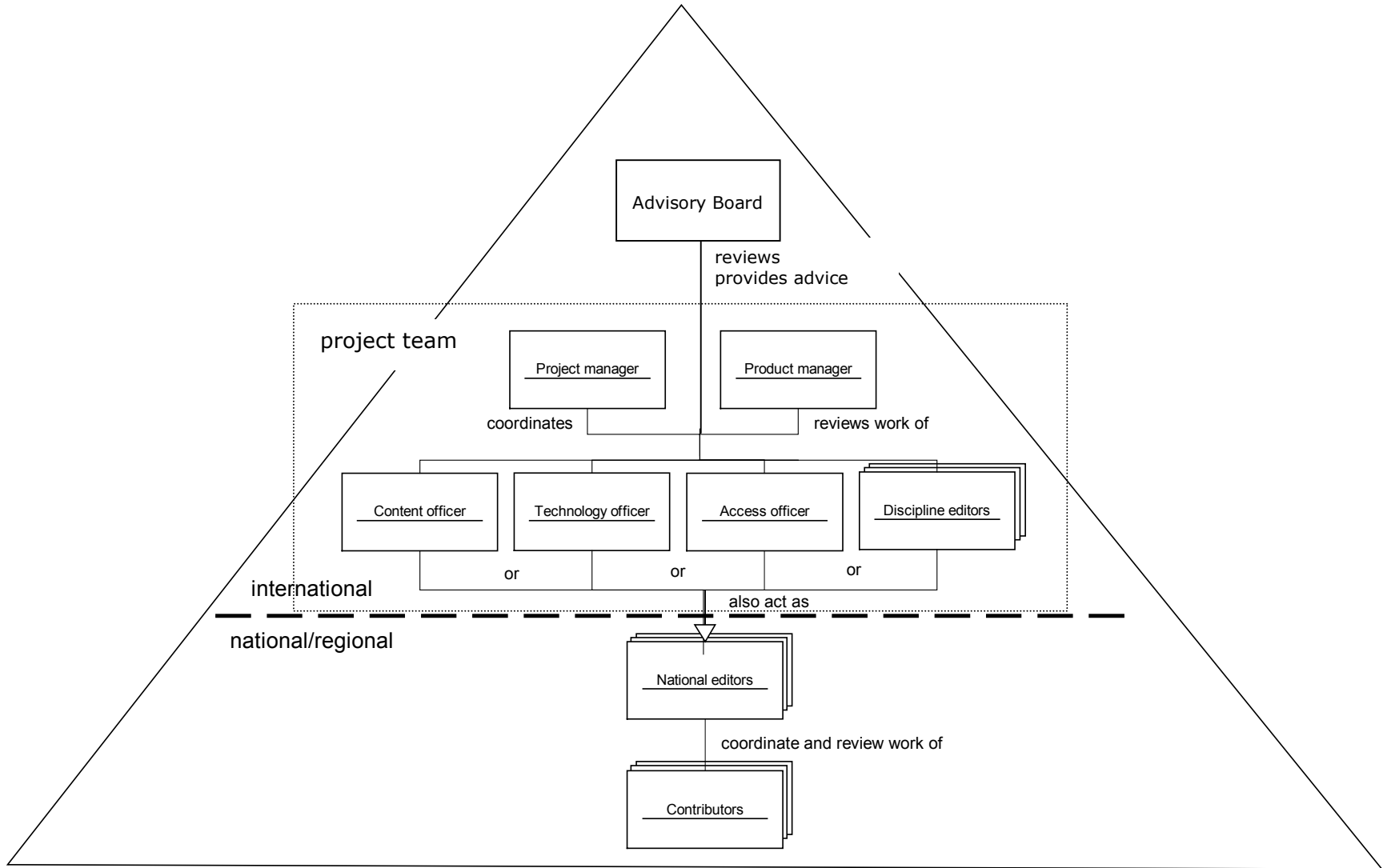
### Specialized vocabularies

- architecture [[Base de données Thésaurus](#)],
- museum objects [[British Museum Object Names Thesaurus](#)],
- man-made objects [[The Revised Nomenclature for Museum Cataloging](#)],
- religious objects [[Religious Objects –User's Guide and Terminology](#)],
- costume [[Vocabulary of Basic Terms for Cataloguing Costume](#)],
- materials [[British Museum Materials Thesaurus](#), [Building Materials Thesaurus](#)],
- types of monuments [[Thesaurus of Monument Types](#)],
- archaeological objects [[mda Archaeological Objects Thesaurus](#)], etc.

## [IFLA Section of Art Libraries Multilingual Glossary for Art Librarians](#)

English with Indexes in Dutch, French, German, Italian, Spanish and Swedish

# APPENDIX 4: IMAGELINE ORGANIZATIONAL MODEL



## APPENDIX 5: ESTIMATED EXPENSES

<b>YEAR 1</b>	<b>Total: \$172,240.00</b>
---------------	----------------------------

<b>Salaries</b>	<b>rate</b>	<b>\$102,240.00</b>
-----------------	-------------	---------------------

project coordination 2	20 hrs/week for 8 weeks	40	\$6,400.00
managers	10 hrs/week for 44 weeks	40	\$17,600.00
documentation	70 personhours	30	\$2,100.00
vocabularies & schema	500 personhours	40	\$20,000.00
adding resources	500 personhours	20	\$10,000.00
editorial work	960 personhours	40	\$38,400.00
technical	4 hrs/week for 52 weeks	30	\$6,240.00
UI design & information architecture	50 personhours	30	\$1,500.00

<b>Equipment</b>	<b>\$53,000.00</b>
------------------	--------------------

hardware w/OS	2 servers, Linux, mSQL	\$8,000.00
software	w/tweaking & implementation	\$45,000.00

<b>Other</b>	<b>\$17,000.00</b>
--------------	--------------------

training	\$7,000.00
logo	\$500.00
Web sites	\$2,000.00
posters, flyers	\$500.00
consumables	\$500.00
communication expenses phone charges, mail etc.	\$500.00
travel	\$4,000.00
evaluation expenses usability testing, online survey, focus groups, expert reviews	\$2,000.00

<b>YEAR 2</b>	<b>Total: \$56,400.00</b>
---------------	---------------------------

<b>Salaries</b>	<b>rate</b>	<b>\$40,400.00</b>
-----------------	-------------	--------------------

project coordination 1			
manager	2 hrs/week	40	\$4,160.00
adding/editing resources	500 manhours @ \$20	20	\$10,000.00
editorial work	500 manhours	40	\$20,000.00
technical	4 hrs/week for 52 weeks	30	\$6,240.00

<b>Equipment</b>	<b>\$9,000.00</b>
------------------	-------------------

maintenance and support up to 20% of purchase price	\$9,000.00
---	------------

<b>Other</b>	<b>\$7,000.00</b>
--------------	-------------------

consumables	\$500.00
-------------	----------

communication expenses phone charges, mail etc.	\$500.00
travel	\$4,000.00
evaluation expenses usability testing, online survey, focus groups, expert reviews	\$2,000.00

**NOTES:**

- hourly rates based on average pay rates for U.S. university libraries, where librarian annual salaries range from 35-100K depending on experience
- rates most likely vary widely across the world, annual rates for Australian government librarians, for example, 19K-45K (converted into \$US)
- according to the same sources, programmers' annual salaries range from 36-95K, computer resources managers' from 63-126K
- software price based on the most expensive option of buying all software and having the provider implement and adapt it (example of Index+ s/w option and implementation by System Simulation Ltd.)
- to provide for complex multilingual implementation, the highest price of 45K (converted to \$US) is used

**Information on software prices and conditions received up to date:**

**MetaSuite**

The normal commercial price is US\$5500, but we do offer a 50% discount to accredited non-profit and academic and research institutions. Maintenance is optional at 20% of the initial license fee, per annum. This provides unlimited email support with a 24 hour response time, based on normal Australian business hours, no-charge upgrades and Early Access releases.

Contact:  
 Ron Chernich  
 Engineering Director, DSTC  
[corba-product@dstc.com](mailto:corba-product@dstc.com)

**Index+**

vanilla installation of a Dublin Core repository, webserver and Z39.50 server is 3,500 UK Pounds.

In practice we configure these to customer requirements and the price ends up typically between 7,500 UKP and 30,000 UKP.

non-vanilla features: vocabulary integration, multilingualism, OAI support, MS-Windows client for data entry etc.

Contact:  
 Mike Stapleton  
[mike@ssl.co.uk](mailto:mike@ssl.co.uk)

**MetaStar**

no prices given

Contact:  
 Jim Mathias  
 Blue Angel Technologies, Inc.  
[Jim.Mathias@BlueAngelTech.com](mailto:Jim.Mathias@BlueAngelTech.com)  
 610-917-9200 x4124 / ICQ# : 164104471  
[www.BlueAngelTech.com](http://www.BlueAngelTech.com)



## **Overviews of software used in subject gateways:**

UKOLN. Metadata: Other Software Tools [<http://www.ukoln.ac.uk/metadata/software-tools/non-ukoln.html>]

The IMesh Toolkit. Products: Review of Software for Subject Gateways [[http://www.imesh.org/toolkit/work/tech\\_review/products/](http://www.imesh.org/toolkit/work/tech_review/products/)]

Renardus. Evaluation report of existing broker models in related projects (D1.1) [[http://www.renardus.org/about\\_us/deliverables/d1\\_1/D1\\_1\\_final.pdf](http://www.renardus.org/about_us/deliverables/d1_1/D1_1_final.pdf)]  
includes information about software that has been used in different subject gateways

## APPENDIX 6: METADATA SCHEMAS

[Principles](#) | [Renardus](#) | [RLSP CLD](#) | [ROADS](#) | [JIDI](#) | [Domain specific description formats](#) | [Suggested Imageline scheme for discussion](#)

### Principles

- Providing access as the primary purpose of descriptions  
The record should allow the user to find, identify, select, and obtain access to resources
- 1:1 principle  
Only one object or resource may be described within a single metadata set. How the element sets are linked to form a single record is a local database implementation issue.
- Dumb-down principle  
For interoperability purposes, any description needs to be usable in unqualified form. A client should be able to ignore any refinement and use the description as if it were unqualified. While this may result in some loss of specific meaning, the remaining element value (minus the qualifier) must continue to be generally correct.

### Renardus Collection Level Description

[<http://renardus.sub.uni-goettingen.de/renap/rcld.html>]

Attribute	RDF property	Definition
<b>Dublin Core (based) elements:</b>		
<b>Title</b>	dc:title	The name of the collection.
<b>Identifier</b>	dc:identifier	An unambiguous reference to the collection within a given context (encoding scheme: URI).
<b>Description</b>	dc:description	An account of the content of the collection.  <i>Comment:</i> Renardus will provide a standardized structure of the content of description with information about granularity of collected resources, type of subject indexing, etc. in context of D6.5.
<b>Language</b>	dc:language	The main language(s) of the metadata in the collection with quantitative indication.  <i>Syntax:</i> Free text.
<b>Publisher</b>	dc:publisher	An entity responsible for making the collection available. <i>Comment:</i> The organization etc. who is responsible for the intellectual (not technical) distribution of the collection.
<b>Format.Extent</b>	dc:format dcq:extent	The size of the collection.  <i>Comment:</i> It is recommended to provide the number of records as follows: about x records.
<b>Date.Issued</b>	dc:date dcq:issued	Date of formal issuance (e.g. publication) of the collection.
<b>Subject</b>	dc:subject	The topic of the content of the collection.  <i>Syntax:</i> Main DDC captions for the

		subjects represented in the Subject Gateway.
<b>Relation</b>	dc:relation dcq:hasPart dcq:isPartOf	A reference to a related resource.  <i>Syntax:</i> Acronym followed by empty character must precede other describing text for every related subject gateway.  <i>Comment:</i> At the moment only used by RDN and its member Subject Gateways.
<b>Collection Level Description elements based on RSLP schema:</b>		
<b>Country</b>	clد:country	The country in which the collection is physically located.  <i>Syntax:</i> Free text.
<b>Renardus specific Collection Level Description elements:</b>		
<b>Subject Notation</b>	rclد:subjectNotation	The topic of the content of the collection.  <i>Syntax:</i> Main DDC notations and captions for the subjects represented in the Subject Gateway: DDC notation1 – DDC caption1; DDC notation2 – DDC caption2 etc.  <i>Comment:</i> Element content not displayed in human readable Collection Level Descriptions. For technical and license reasons this element is declared as a Renardus CLD element instead of a DC element.
<b>Acronym</b>	rclد:acronym	The acronym of the collection.  <i>Comment:</i> Use only the letters of the 128 ASCII characters 0 to 127 and no special characters (see e.g. <a href="http://www.cs.mun.ca/~michael/c/asciitable.html">http://www.cs.mun.ca/~michael/c/asciitable.html</a> ). It is allowed to use a hyphen for the acronym.
<b>Resource Language</b>	rclد:resourceLanguage	Language(s) of the described resources.  <i>Syntax:</i> Free text.
<b>DDC mapping URL</b>	rclد:ddcMapping	URL of local DDC mapping information in Renardus format.  <i>Comment:</i> Element content not displayed in human readable Collection Level Descriptions.
<b>Z39.50 Location</b>	rclد:Z3950Location	The online location of the Z39.50 server of the subject gateway  <i>Syntax:</i> machine name; port number; database name  <i>Comment:</i> Element content not displayed in human readable Collection Level Descriptions.

<b>Logo URL</b>	rclid:logoURL	The URL of the logo (image) of the subject gateway.  <i>Comment:</i> Element content not displayed in human readable Collection Level Descriptions.
-----------------	---------------	---

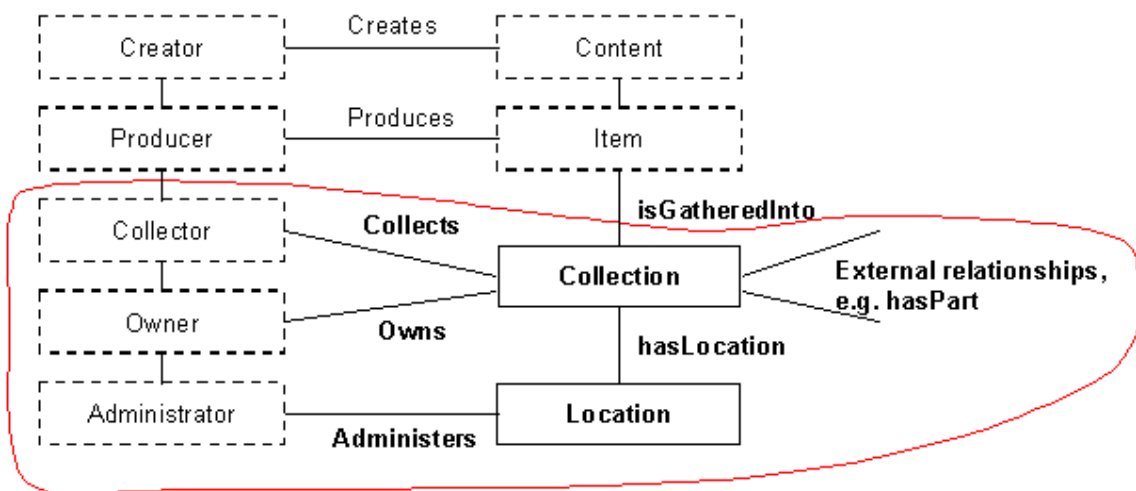
### RSLP Collection Level Description

<http://www.ukoln.ac.uk/metadata/rslp/>

Collection		
Attribute	RDF property	Definition
General attributes		
Title	<b>dc:title</b>	The name of the collection.
Identifier	<b>dc:identifier</b>	A formal identifier for the collection.
Description	<b>dc:description</b>	A description of the collection.
Strength	<b>cld:strength</b> (sub-property of dc:description)	An indication (free text or formalised) of the strength(s) of the collection.
Physical Characteristics	<b>dc:format</b>	The physical or digital characteristics of the collection.
Language	<b>dc:language</b>	The language of the items in the collection.
Type	<b>dc:type</b>	The type of the collection.
Access Control	<b>cld:accessControl</b> (sub-property of dc:rights)	A statement of any access restrictions placed on the collection, including allowed users, charges, etc.
Accrual Status	<b>cld:accrualStatus</b> (sub-property of dc:description)	A statement of accrual policy (closed, passive, active, partial/selective), accrual method (purchase, deposit) and accrual periodicity (closed, irregular, periodic).
Legal Status	<b>cld:legalStatus</b> (sub-property of dc:description)	A statement of the legal status of the collection.
Custodial History	<b>cld:custodialHistory</b> (sub-property of dc:description)	A statement of any changes in ownership and custody of the collection that are significant for its authenticity, integrity and interpretation.
Note	<b>cld:note</b> (sub-property of dc:description)	Any general information about the collection.
Location	<b>cld:hasLocation</b> (sub-property of dc:relation)	The identifier for the physical or online (digital) location of the collection.
Subject		
Concept	<b>dc:subject</b>	A concept (keyword) of the items in the collection.
Object	<b>cld:objectName</b> (sub-property of dc:subject)	An object name associated with the items in the collection.
Name	<b>cld:agentName</b> (sub-property of dc:subject)	An personal or corporate name associated with the items in the collection.
Place	<b>dcq:spatial</b> (sub-property of dc:coverage)	The spatial coverage of the items in the collection.
Time	<b>dcq:temporal</b>	The temporal coverage of the items in

	(sub-property of dc:coverage)	the collection.
Dates		
Accumulation Date Range	<b>cld:accumulationDateRange</b> (a sub-property of dc:date)	The range of dates over which the collection was accumulated.
Contents Date Range	<b>cld:contentsDateRange</b> (a sub-property of dc:date)	The range of dates of the individual items within the collection.
Associated agents		
Collector	<b>dc:creator</b>	The identifier for an agent who gathers (or gathered) the items in a collection together.
Owner	<b>cld:owner</b>	The identifier for an agent who has legal possession of the collection.
External relationships		
Sub-collection	<b>dcq:hasPart</b> (sub-property of dc:relation)	The identifier or name of a second collection contained within the current collection.
Super-collection	<b>dcq:isPartOf</b> (sub-property of dc:relation)	The identifier or name of a second collection that contains the current collection.
Catalogue or description	<b>cld:hasDescription</b> (sub-property of dc:relation)	The identifier or name of a second collection that describes the current collection (for example, the catalogue for the current collection).
Described collection	<b>cld:isDescriptionOf</b> (sub-property of dc:relation)	The identifier or name of a second collection that is described by the current collection.
Associated collection	<b>cld:hasAssociation</b> (sub-property of dc:relation)	The identifier or name of a second collection that is associated by provenance with the current collection.
Associated publication	<b>cld:hasPublication</b> (sub-property of dc:relation)	The identifier or name of a publication that is based on the use, study, or analysis of the collection.
<b>Location</b>		
<b>Attribute</b>	<b>RDF property</b>	<b>Definition</b>
General attributes		
Name	<b>dc:title</b>	The name of the location.
Identifier	<b>dc:identifier</b>	A formal identifier for the location.
Access Conditions	<b>cld:accessConditions</b>	Hours of access, classes of permitted user, etc.
Held collection	<b>cld:isLocationOf</b> (sub-property of dc:relation)	The identifier for a collection held at this physical or online (digital) location.
See also	<b>cld:seeAlso</b> (sub-property of dc:relation)	The identifier of a resource that provides further information about this location (typically the URL for an organisational home page).
Associated agents		
Administrator	<b>cld:administrator</b> (sub-property of dc:publisher)	The identifier for an agent who has responsibility for the physical or electronic environment in which the collection is held.
Physical location		

Postal address	<b>cld:address</b> (sub-property of dc:identifier)	The full postal address for the physical location of the physical collection.
Post/zip code	<b>cld:postcode</b>	The post code or zip code for the physical location of the collection.
Country	<b>cld:country</b>	The country in which the collection is physically located.
Online location		
Locator	<b>cld:locator</b> (sub-property of dc:identifier)	The online location (URL) of an online (digital) collection.
<b>Agent</b>		
<b>Attribute</b>	<b>RDF property</b>	<b>Definition</b>
<b>General attributes</b>		
Name	<b>vcard:fn</b>	The name of the agent.
Identifier	<b>dc:identifier</b>	A formal identifier for the agent.
Organisation name	<b>vcard:org</b>	The organisational name of, or affiliated with, the agent.
Role	<b>vcard:role</b>	The role (typically an organisational role) fulfilled by the agent.
Telephone number	<b>vcard:voice</b> (sub-property of vcard:tel)	The telephone number of the agent.
Fax number	<b>vcard:fax</b> (sub-property of vcard:tel)	The fax number of the agent.
Email address	<b>vcard:email</b>	The electronic mail address of the agent.
Agent History	<b>cld:agentHistory</b>	An administrative history of, or biographical details on, the agent.



### ROADS Collection Template

<http://www.ukoln.ac.uk/metadata/roads/templates/collection.html>

Attribute	Notes
Template-Type	COLLECTION
Handle	Assigned automatically
Template-Version	0.2
Title	The name given to the collection, usually by

	the Admin, Owner or Publisher.
Subject-v*	The topic of the collection. Typically, subject will be expressed as keywords or phrases that describe the subject or content of the collection. The use of controlled vocabularies and formal classification schemes is encouraged.
Description-v*	A textual description of the content of the collection.
Owner- (AGENT CLUSTER) Owner-Name-v* Owner-Type-v* Owner-Email-v* Owner-Identifier-v*	The person or organization that owns the collection. See <a href="#">AGENT</a> cluster.
Date-v*	A date associated with the creation or availability of the collection. Recommended best practice is defined in <a href="#">a profile of ISO 8601</a> that includes (among others) dates of the forms YYYY and YYYY-MM-DD. In this scheme, for example, the date 1994-11-05 corresponds to November 5, 1994.
Language-v*	The language of the intellectual content of the collection. Recommended best practice is defined in <a href="#">RFC 1766</a> .
Source-v*	Information about the resources from which the collection is derived.
Coverage-v*	The spatial or temporal characteristics of the intellectual content of the collection. Spatial coverage refers to a physical region (e.g., celestial sector) using place names or coordinates (e.g., longitude and latitude). Temporal coverage refers to what the collection is about rather than when it was created or made available (the latter belonging in the Date element). Temporal coverage is typically specified using named time periods (e.g., neolithic) or the same date/time format as recommended for the Date element.
Relation-v*	An identifier of a second resource or collection and its relationship to the present collection. This element is used to express linkages among related resources and collections of resources.
Type-v*	The category of the collection. For the sake of interoperability, Type should be selected from an <a href="#">enumerated list</a> .
Notes-v*	Specialized information that cannot be accommodated in any of the other areas.
Purpose-v*	A statement of why the collection is offered and the origin and lineage of the resources held. This attribute may also describe any collection policy associated with the collection.
Identifier-v*	A string or number used to uniquely identify the collection.
Rights-v*	A rights management statement, an identifier that links to a rights management statement, or an identifier that links to a service providing

	information about rights management for the collection.
UseConstraints-v*	A description of any constraints or legal prerequisites for using the information resource or its component products or services. This includes any use constraints applied to assure the protection of privacy or intellectual property and any other special restrictions or limitations on using the information resource.
Logo-v*	The URI of a logo associated with the collection or the Owner of the collection. If present, this logo must be displayed by any service that provides access to the collection.
Resource- (ACCESS CLUSTER) Resource-Admin-Name-v* Resource-Admin-Type-v* Resource-Admin-Email-v* Resource-Admin-Identifier-v* Resource-Publisher-Name-v* Resource-Publisher-Type-v* Resource-Publisher-Email-v* Resource-Publisher-Identifier-v* Resource-Identifier-v* Resource-AccessTimes-v* Resource-Location-v* Resource-AccessPolicy-v* Resource-ChargingPolicy-v*	Information about access to the collection. See <a href="#">ACCESS</a> cluster.
To-Be-Reviewed-Date	
Record-Last-Verified-Email	
Record-Last-Verified-Date	
Comments	
Destination	

## JIDI Metadata Structure for Collection Description

[<http://www.tasi.ac.uk/advice/delivering/jidi-metadata.html> - jm2]

<b>Collection:</b> A Collection is an aggregation of one or more Works or Visual Documents that have some common property or provenance that is likely to be of interest to the researcher. This data need only be entered once for each collection. Sites may hold more than one collection (or sub-collections).		
Field	Description of field	Mandatory
<b>COLLECTION TITLE</b>	<b>Definition:</b> The title or identifying phrases given to a Collection.	Yes
<b>COLLECTION DESCRIPTION</b>	<b>Definition:</b> A textual description of the nature, history etc. of the collection.	Yes
<b>COLLECTION SIZE</b>	<b>Definition:</b> An estimate of the number of items (either digital or non-digital) in the collection. <b>Guidelines:</b> Store a numeric value.	No
<b>COLLECTION OWNER</b>	<b>Definition:</b> Name of a person or organisation owning the collection <b>Guidelines:</b> Alternatively can be used to store a link to a separate PERSON or ORGANISATION record.	No
<b>COLLECTION COPYRIGHT</b>	<b>Definition:</b> Name of a person or organisation describing the owner of the copyright in the	No



<b>OWNER</b>	Collection. <b>Guidelines:</b> Alternatively can be used to store a link to a separate PERSON or ORGANISATION record.	
<b>COLLECTION COPYRIGHT STATEMENT</b>	<b>Definition:</b> A statement regarding the ownership of copyright that will be displayed whenever images from the collection are displayed. <b>Guidelines:</b> As detailed in the JIDI license; to be completed by the sites.	No, but advisable to include
<b>COLLECTION COPYRIGHT STATUS</b>	<b>Definition:</b> The copyright status of the Collection, where appropriate. <b>Guidelines:</b> This field will primarily be used to differentiate between the different types of licenses for disseminating resources, e.g. UK HE only. Should be selected from a short list of options.	Yes
<b>COLLECTION NOTES</b>	<b>Definition:</b> Used for recording other pertinent information about either the collection described by the record, or the record itself, but that cannot be recorded in any other field.	No

### Domain specific description formats

[Visual Resources Association's Core Categories](http://www.vraweb.org/vracore3.htm) [http://www.vraweb.org/vracore3.htm]

[Object ID](http://www.object-id.com/) [http://www.object-id.com/]

[JIDI](http://www.tasi.ac.uk/advice/delivering/jidi-metadata.html) [http://www.tasi.ac.uk/advice/delivering/jidi-metadata.html]

[Categories for the Description of Works of Art \(CDWA\)](http://www.getty.edu/gri/standard/cdwa/index.html)  
[http://www.getty.edu/gri/standard/cdwa/index.html]

[CIDOC Information Categories](http://www.cidoc.icom.org/guide0.htm) [http://www.cidoc.icom.org/guide0.htm] ???

[The CIMI Profile Release 1.0H: A Z39.50 Profile for Cultural Heritage Information](http://www.cimi.org/public_docs/HarmonizedProfile/HarmonProfile1.htm)  
[http://www.cimi.org/public\_docs/HarmonizedProfile/HarmonProfile1.htm]

[SPECTRUM: The UK Museum Documentation Standard](http://www.mda.org.uk/spectrum.htm)  
[http://www.mda.org.uk/spectrum.htm]

[Canadian Heritage Information Network \(CHIN\) Data Dictionary](http://www.chin.gc.ca/Artefacts/RULH/e_sample.html)  
[http://www.chin.gc.ca/Artefacts/RULH/e\_sample.html]

[Méthode d'inventaire informatique des objets beaux-arts et arts décoratifs](http://www.culture.fr/documentation/joconde/methode.htm#peinture)  
[http://www.culture.fr/documentation/joconde/methode.htm#peinture]

[Museum Educational Site Licensing Project \(MESL\) Data Dictionary](http://lists.village.virginia.edu/uvamesl/search_help/datadict.html)  
[http://lists.village.virginia.edu/uvamesl/search\_help/datadict.html]

[REACH Element Set](http://www.rlg.org/reach.elements.html)  
[http://www.rlg.org/reach.elements.html]

[Van Eyck Core Record Structure: Proposal for a Standard](http://www.rkd.nl/pblctns/stndr-e.htm) [http://www.rkd.nl/pblctns/stndr-e.htm]

[AMICO Data Dictionary](http://www.amico.org/AMICOLibrary/dataDictionary.html) [http://www.amico.org/AMICOLibrary/dataDictionary.html]

### Suggested Imageline scheme for discussion