



EMBRC
EUROPEAN
MARINE
BIOLOGICAL
RESOURCE
CENTRE

Plans for digital library

Deliverable D11.10

January 2014

The EMBRC preparatory-phase



D11.10 – Plans for digital library

Summary

4 – 5 sentences with main conclusions and contents of the report

Contents

1. Introduction.....	2
1.1 Definition of the e-library.....	2
1.2 Copyright and Digitization	2
2. Plans for the construction of marine biological virtual libraries	3
2.1 Questionnaire.....	3
2.2 Processes of digitization.....	3
2.3 Selection criteria.....	4
3. Options for integration and expansion	4
4. Plans for digitalization of information.....	5
5. Operational links with EMBRC E-infrastructure.....	6
6. Conclusion	7

Annexes

ANNEX 1 - Questionnaire EMBRC Library resources

ANNEX 2 - Summary of the questionnaire results

Executive Summary

Objectives

The main objectives of this report on digital library are to identify the available resources in order to build an accessible online portal for library resources. An important objective is to plan for digitalization of information in old books and monographs, which are rare and/or no longer easily accessible to the wider community. After identifying the existing library resources (on-site and remote) at ppEMBRC partner institutes through a questionnaire, some recommendations are made for their integration and development. This digital library will be continuously updated and lead to more accessible online resources and services.

Background

To prepare the digital library plan for the implementation phase of EMBRC, an inventory of existing library resources (annex I) at each of the ppEMBRC partners has been conducted. Based on this inventory, the experience gained during ppEMBRC and after internal advices from experts, the WP11 team has written a proposal (D11.10) of the EMBRC Plans for digital library.

1. Introduction

In collaboration with EMBRC partners' libraries and other partners, EMBRC wish to assemble a comprehensive digital research library and to facilitate the ways that universities, students, and researchers may discover and have access to information. The recent development of new digital technologies offers new opportunities to enhance the ability of libraries to fulfill their missions in terms of access, preservation, research, and education.

This online library catalog will provide our users with a vast collection of journals, books, old books and historical archives. By facilitating the co-investment and sharing of materials and services used by libraries across the partnership, EMBRC will build a unique and comprehensive portal, in collaboration with their associated university libraries to bring the treasures of our collections and cultural heritage institutes to a wide diversity of publics. The innovative interface will allow users to visit online collections, search collection databases, access images of collection items, and historical archives such as manuscripts and maps. EMBRC will explore how new services such as data management, social media or virtual exhibition can support research throughout the information lifecycle.

Our unique position within the Universities allows us to provide the infrastructure and support commonly needed by the libraries, allowing them to focus their resources on the needs of their users. Looking ahead, EMBRC will continue to use innovative technology to connect contents and communities in ways that promote teaching, learning, and research.

EMBRC is aware that digitization raises the possibility of copyright restriction and that copyright instruction programs need to be done to ensure that staff knowledge of copyright is up to date.

1.1 Definition of the e-library

An electronic library is a virtual portal *“that collects, manages and preserves for the long depth of time rich digital content, and offers to its targeted user communities specialized functionality on that content, of defined quality and according to comprehensive codified policies”* (The Digital Library Reference Model, April 2011). The goal of EMBRC's e-library should be to ensure that a wide diversity of users (scientists, professors, historians, artists, teachers, etc.) have access to relevant historical collections and journals that are generated in all marine stations and laboratories in order to make durable and enhance these valuable collections.

1.2 Copyright and Digitization

The copyright term enables us to digitize all the on-site collections which are already available online through the universities' electronic access. Copyright term is based on the life of the author, and extends to 50 or 70 years beyond the death of the author. As rights are country-based, a work may be subject to rights in one country and not in another. Copyright assessments will play a key role in this digitization project and will need to be addressed early in the selection process. If the proposed digitizing project involves materials that are not in the public domain, permissions must be secured and appropriate fees paid.

2. Plans for the construction of marine biological virtual libraries

2.1 Questionnaire

In October 2013, EMBRC partners were asked to complete a survey designed to collect information on existing library resources (on-site and remote) and on current and planned projects that might include digitalization of books and monographs. The questions were designed to determine each partner's current resources and to anticipate their needs: 10 partners responded. A full summary of the results can be found in annex 2 at the end of this report.

Marine biology libraries are located either at the stations (SZN, SAMS, MBA, SBR, OOB, OOV, AWI) or hosted at associated university libraries. All libraries have e-access to journals and digitized publications, whereas those at marine station also host precious historical monographs, journals and reports on-site.

The main conclusions from the questionnaire were that most of the partners are already engaged in digitalization projects. Many partners are already able to access on a web portal their on-site resources. They have a common documentary collection and some complementary resources that could be exchanged within the stations. The questionnaire also showed that EMBRC partners hold a wide range of biological and historical collections that are not very well catalogued.

Many marine stations publish their own journals, which contain relevant information on the work done at the station by the staff and visiting investigators. The historical archives cover documents relating to the activity of the Institution from its foundation. It includes institutional documents, private and public correspondences between the pioneers, photographs, scientific drawings, and others. The libraries holds also further historical documents, donated to the Institutes during the years. The administration of these historical archives will lead to new actions for managing and preserving the documentation, for both local and international communities.

2.2 Processes of digitization

Selection for digitization is a complicated process. The conversion of textual, visual, and numeric information to electronic form, from preparation and conversion to presentation and archiving, encompasses a range of procedures and technologies with widely varying implications and costs. The following factors need to be considered: the intellectual and physical nature of the source materials; the number and location of current and potential users; the current and potential nature of use; the format and nature of the proposed digital product and how it will be described, delivered, and archived; how the proposed product relates to other digitization efforts; and projections of costs in relation to benefits.

2.3 Selection criteria

There is a common conception that digitization is for old and rare material. This is understandable as many of the projects undertaken at libraries, have placed their focus of digitization there. But maybe more modern material could also be appropriate. Not only rare/unique items need to be identified, material that is in constant/high demand, such as internal periodic journals, need to be identified too.

EMBRC should particularly encourage digitization projects that:

- Increase the availability, and consequently the use of a collection of demonstrated historical significance
- Identify collections (old books, historical archives, manuscripts, pictures, maps) from the libraries that are important and should be accessed online by a large community
- Demonstrate strong interest within the academic community including people involved in human sciences for access to the collection
- Contribute significantly to the existing digital collections such as the ones included in the future EMBRC's digital library
- Support the EMBRC's subject strengths

3. Options for integration and expansion

The EMBRC Central Hub needs to define a team that will have the responsibility for implementing an agreed plan or strategy to achieve EMBRC's Virtual Library objectives. The team should be composed of colleagues from our libraries who are best-placed and best-trained to develop and update this virtual library. The Communications Officer will coordinate the project and will have to find a name for this digital library in order to facilitate its promotion and develop an interface, linked to the e-infrastructure, based on the current brand guidelines.

According to the questionnaire results and internal expertise, the main duties would include:

- Coordinating the online inventory of EMBRC publications
- Managing the digitization (with Optical Character Recognition) of old books, historical archives (grey literature) and maps. The Optical Character Recognition (OCR) is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed online, and used in machine processes such as machine translation or key data extraction for example.
- Working with the marine stations' libraries to make available their internal journals and publications

- Preparing the standardization of the collections. All the collections should be identified and listed online. A catalog (with a search interface) could give access to the collections and their localization throughout the stations.
- Optimizing web interface by analyzing other libraries' websites and stay aware of new digital technologies (through networks and events)
- Presenting and promoting EMBRC Virtual Library by identifying key conferences, events, websites, journals through which messages can be disseminated
- Creating new online services (virtual exhibitions, news about recent acquisitions, lectures, publications, etc., and links with social media such as Flickr or Pinterest to disseminate and promote images' collections)
- Creating collaterals/communication tools (online and printed materials) for dissemination (e.g. the logo (linked with the Digital Library's name), factsheet, posters, postcards, bookmarks, and even a PowerPoint presentation about the project)

4. Plans for digitalization of information

A few H2020 work programmes offer interesting opportunities for libraries. “**Future and Emerging Technologies**” (FET) seeks innovation proposals in support of the integration of data. “**Inclusive, Innovation and Reflective Societies**” offers a number of opportunities in the area of cultural heritage, under **VI. Inclusive, innovative and reflective societies**, 6.3 Reflective societies – Cultural heritage and European Identity: *“aims at contributing to an understanding of Europe’s intellectual basis: its history, traditions, and regional and national identities. In this context, the important role of collections in libraries, museums and archives is highlighted, as well as the need to make this material accessible to researchers and citizens, also through new technologies, to enable a look to the future through the archive of the past. Accessibility and preservation of cultural heritage in these forms is needed for the vitality of the living engagements within and across European cultures now and contributes to sustainable economic growth.”*

Two other relevant programmes are the inter-related **European Research Infrastructures, including e-infrastructure**s and **Science with and for Society**. In addition, **Industrial leadership** seems significant, 1.1.4. Content technologies and information management: ICT for digital content and creativity: *“The objective is to provide professionals and citizens with new tools to create, exploit and preserve all forms of digital content in any language and to model, analyse, and visualise vast amounts of data, including linked data. This includes new technologies for language, learning, interaction, digital preservation, content access and analytics; intelligent information management systems based on advanced data mining, machine learning, statistical analysis and visual computing technologies.”*

As the demand for Research Data Management support is increasing through H2020, INFRASUPP-3-2014 – Strengthening the human capital of research infrastructures which seeks to promote the role of the data librarian by developing education and training – may benefit to our library community.

This may also require interaction with relevant organizations such as Europeana* and/or companies such as Google (*The Europeana Libraries project is working to make 5 million digital objects from 19 leading European research libraries freely accessible (with free access) on The European Library and Europeana websites). EMBRC could lead an integrated activity with other professional organizations, whose focus is more centered targeted.

Some important events (conferences, workshops), devoted to Cultural Heritage, are going to be held in 2014:

- **EuroMed 2014** (November) brings together researchers, policy makers, professionals and practitioners to explore some of the more pressing issues concerning cultural heritage. In particular, this year the conference will focus on interdisciplinary and multi-disciplinary research on tangible and intangible Cultural Heritage, on the use of cutting-edge technologies for protection, restoration, preservation, massive digitization, documentation and presentation of the Cultural Heritage content. At the same time, the event is intended to cover topics of research ready for exploitation, demonstrating the acceptability of new sustainable approaches and new technologies by the user community, SMEs, owners, managers and conservators of cultural patrimony.
- **EGI Community Forum** (May 2014): The main theme of the Forum is Advancing Excellent Science. In preparation for the Horizon 2020 program, the Community Forum 2014 will focus on the EGI contribution to advance excellent science in the European Research Area through the use of innovative services for data and computing.
- **TPDL 2014**: The International Conference on Theory and Practice of Digital Libraries constitutes a leading scientific forum on digital libraries that brings together researchers, developers, content providers and users in the field of digital libraries
- **DATeCH** (May 2014): The DATeCH international conference brings together researchers and practitioners looking for innovative approaches for the creation, transformation and exploitation of historical documents in digital form.

These events could be an opportunity for EMBRC Library team to meet experts and to have access to the most up-to-date informations on new technologies and projects around digital libraries.

5. Operational links with EMBRC E-infrastructure

Many partners are already able to access on a web portal their on-site resources. EMBRC Virtual Library should not archive the collections, such as Europeana, but serves only as a search catalog. The collections will be available through links towards the library institutes having digitized it.

This library portal will be an interface to access library resources and services through a single access and management point for users. Most digital libraries provide a search interface which allows resources to be found. These resources are typically deep web (or invisible web) resources since they frequently cannot be located by search engine crawlers. This web portal should be linked to the EMBRC e-Infrastructure with a comprehensive and original search interface.

Some examples of innovative interface could be [Digital Public Library of America](#) and [Biodiversity Library](#) websites. The interfaces are well-organized, allowing the users to search into the system via titles, authors, place, date, and collection. Next events and exhibits (recent acquisitions, exhibitions, lectures, and publications) are mentioned on the homepage. The tools, such as virtual exhibitions or images collections linked to Flickr media (e.g. [Biodiversity Library Flickr page](#)), bring an added-value to the Digital library objectives and help promoting the value of historical collections.

6. Conclusion

This first work deserves to be pursued and finalized. This project seems relevant regarding opportunities offered by H2020 WPs and will bring our collections accessible to researchers and citizens, widely unexploited, through these new technologies. *“Accessibility and preservation of cultural heritage in these forms is needed for the vitality of the living engagements within and across European cultures now and contributes to sustainable economic growth”* (Reflective societies – Cultural heritage and European Identity - H2020 work programme).

ANNEX I – Questionnaire EMBRC Library resources

In order to plan the construction of a comprehensive and original marine biological virtual library for EMBRC (D11.10), it is required to survey existing library resources (on-site and remote) at ppEMBRC partner institutes. With your answers, we will assess options for integration and expansion of these, specifically through creation of accessible online resources.

There are 22 items in this questionnaire. For each, please choose the answers that best suit your institute and leave a comment each time you think it's necessary. If an item is not applicable, leave that response blank.

Thank you in advance for your contribution.

Identification

1) Respondent Name:

2) Respondent e-mail and phone contacts:

Presentation of the library

3) Accessibility of the collections to in-house users

- Yes
- No
- Partially

4) Accessibility of the collections to external users

- Yes
- No
- Partially

5) Description of the buildings (m2, reading room, learning centre, etc.)

6) What are the main missions of the Library?

- Management & distribution of information
- Bibliographic research
- Bibliometry
- Documentary digitization
- Promotion of the station publications
- Preservation and promotion of the cultural heritage
- Open archives warehouse management

- Storage in a national portal (example: OpenAIRE)
- Internal training
- External training
- Interlibrary loans
- Debate and meeting spaces
- Other

Edition

7) What kind of documents are you editing inside your station?

- Journals
- Books
- Maps
- Other

8) According to you, would it be possible to exchange these documents within the institutes?

Digitisation

9) How many (number of) collections do you have on-site (print)?

- Journals
- Publications
- Books
- Thesis
- Maps
- Old books collections
- Historical archives of the station (manuscript, pictures, etc.)
- Other

10) Among these collections, which ones are digitized (done or in progress)?

- Journals
- Publications
- Books

- Thesis
- Maps
- Old books collections
- Historical archives of the station (manuscript, pictures, etc.)
- Other

11) According to you, which ones should be available online?

- Journals
- Publications
- Books
- Thesis
- Maps
- Old books collections
- Historical archives of the station (manuscript, pictures, etc.)
- Other

Online access

12) Do you have specific electronic access to the following collections?

- Journals
- E-books
- Databases
- Old books collections
- Other

13) Do you provide electronic access to external users?

- Yes
- No

Catalog/Software

14) Are your collections computerized ?

- Yes

- No

15) If yes, what is the software?

16) Are you adding information to the national university library catalog? (Example: [SUDOC France](#))

- Yes
- No

Website

17) Does the library have a specific website (with search option)?

- Local
- University

18) Is the website frequently updated?

- Yes
- No

19) Do you have other communication tools (eg. Blog, social media, netvibes)?

Networking

20) Are you involved in networks with relevance for library resources/services? If yes, which one(s)?

- Thematic
- Regional
- National
- International

21) Do you participate in collective catalogs other than a national library catalog (example: SUDOC)? If yes, which one(s)?

- Yes
- No

Other relevant topics

22) Have we forgotten to ask any questions? Do you see any other topics that should be relevant for the construction of a virtual library?

Thanks for your time!

ANNEX II – Results of the EMBRC Library resources Questionnaire

Respondents to the online survey: SBR, SZN, SAMS, AWI, MBA, HCMR, OOV, OOB, EMBL, CCMAR

Identification

1) Respondent Name: NA

2) Respondent e-mail and phone contacts: NA

Presentation of the library

3) Accessibility of the collections to in-house users

Yes (8), No (0), Partially (1)

4) Accessibility of the collections to external users

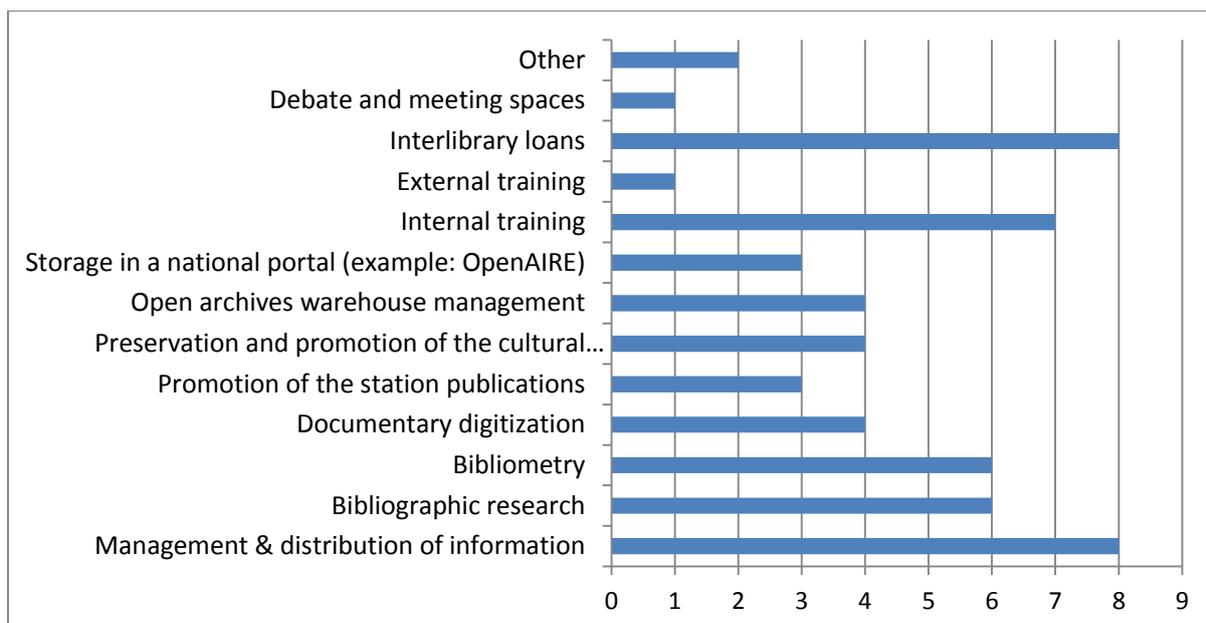
Yes (2), No (0), Partially (7)

5) Description of the buildings (m², reading room, learning centre, etc.)

The 9 respondent partners have a library with reading spaces and study rooms.

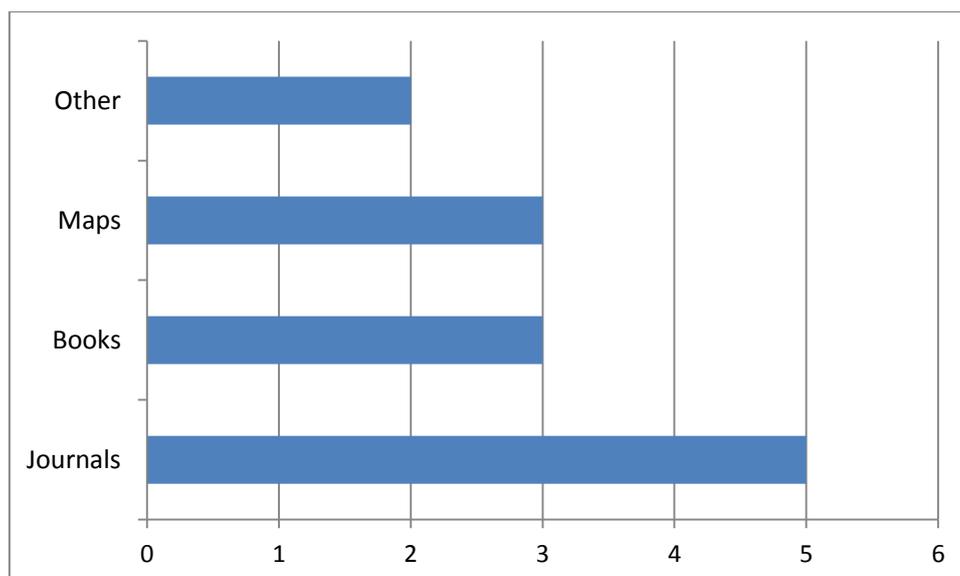
- 2 libraries in 2 buildings 10 reading places 1500 linear meter
- The library's deposit is 800mq (5 floors); the collections count over 100.000 issues (more than 3 kilometres of shelves); the office and the reading room are approximately 35 mq
- Research library: 2 open plan rooms on 2 floors in the main building Undergraduates' library: 1 room in the teaching building
- Holdings mainly in compact shelves, some reading places for single users
- total floor space 530 m², reading room, (learning centre in separate, main building)
- 200 m²
- 200m² Reading room
- The library resources are the university library. It is a modern library with a reading space of 3300m², 368 places and 26 study rooms, computer terminals and wireless access (eduroam and visitor passes as required)
- Reading/studying room
- 500 m² with a reading space

6) What are the main missions of the Library?



Edition

7) What kind of documents are you editing inside your station?



8) According to you, would it be possible to exchange these documents within the institutes?

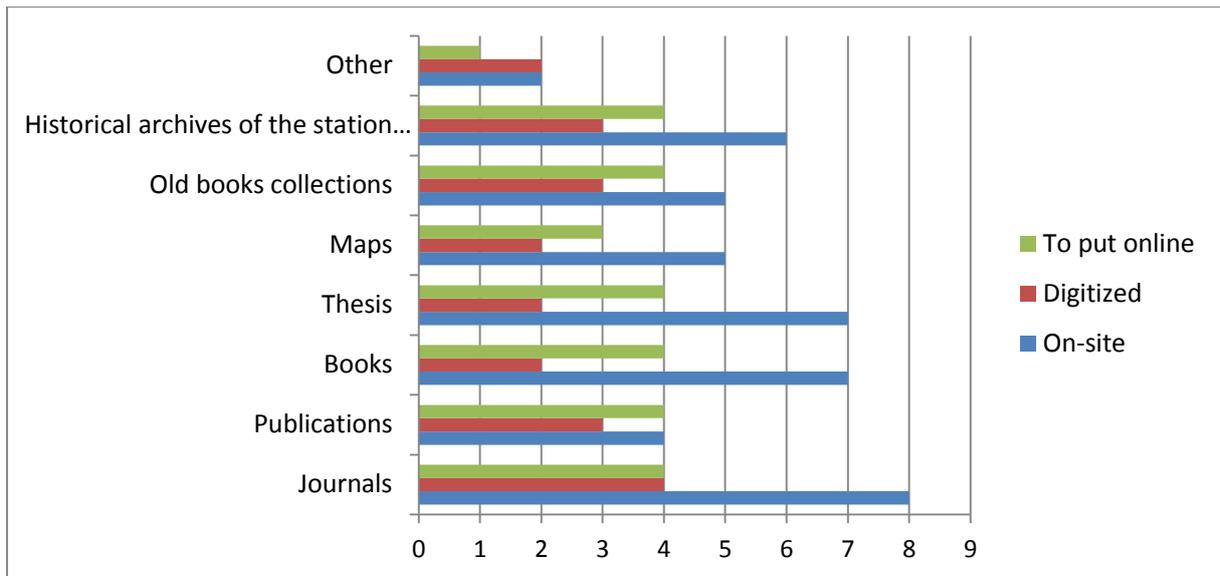
The respondents (6/9) think that it would be possible to exchange totally or partially these documents within the institutes. No answers from 3 respondents.

Digitisation

9) How many collections do you have on-site (print)?

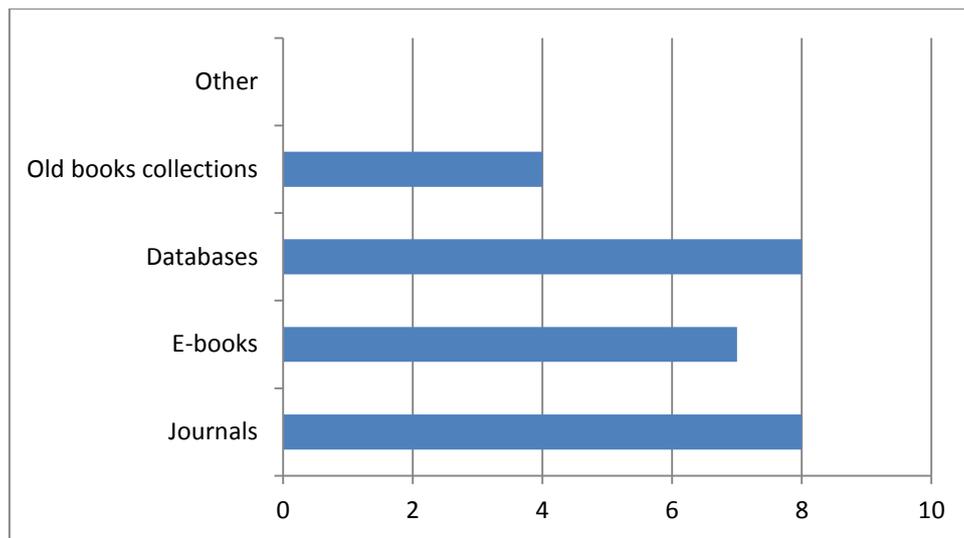
10) Among these collections, which ones are digitized (done or in progress)?

11) According to you, which ones should be available online?

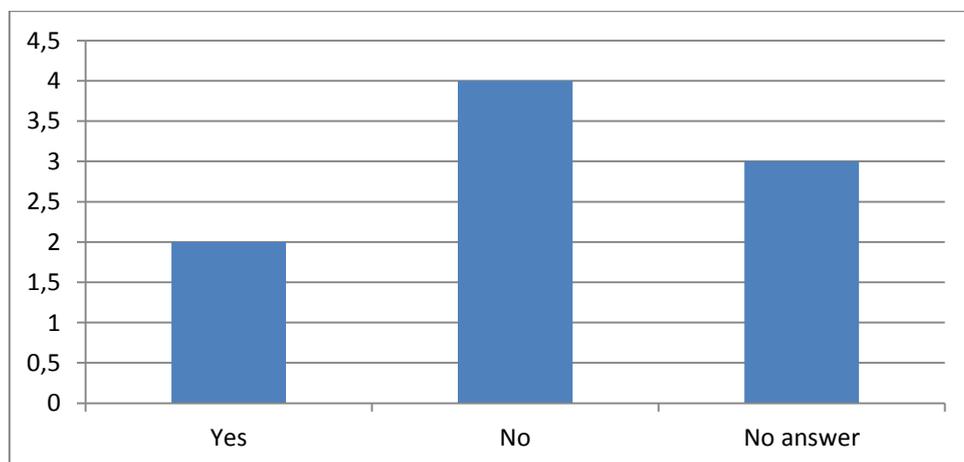


Online access

12) Do you have specific electronic access to the following collections?



13) Do you provide electronic access to external users?



Catalog/Software

14) Are your collections computerized ?

Yes (2), No (4), No Answer (3)

15) If yes, what is the software?

Answer (4), No Answer (5)

- We are leaving Aleph 500 to use Koha (open source ILS)
- Pica (OCLC)
- CDS Isis (to be up-graded)
- Logiciel Symphony

16) Are you adding information to the national university library catalog? (Example: [SUDOC France](#))

Yes (3), No (3), No Answer (3)

Website

17) Does the library have a specific website (with search option)?

Local (7), University (3)

18) Is the website frequently updated?

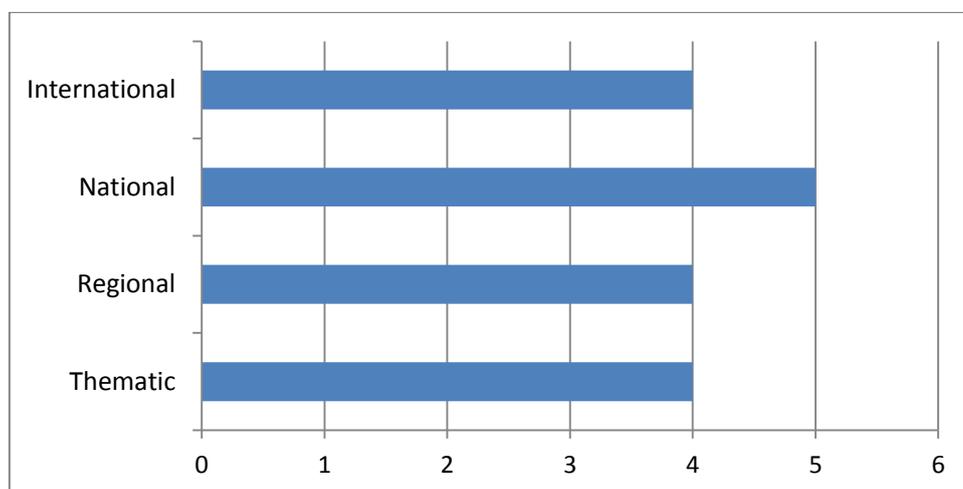
Yes (4), No (1), No Answer (4)

19) Do you have other communication tools (eg. Blog, social media, netvibes)?

Answer (5), No Answer (4)

Networking

20) Are you involved in networks with relevance for library resources/services? If yes, which one(s)?



21) Do you participate in collective catalogs other than a national library catalog (example: SUDOC)? If yes, which one(s)?

Yes (4), No (2), No Answer (3)

Examples: GBV, catalog UPMC LIBRARY

Other relevant topics

22) Have we forgotten to ask any questions? Do you see any other topics that should be relevant for the construction of a virtual library?

Answer (5), No Answer (4)

- The current subjects are: marine ecology, marine environment, oceanography, marine biology, cell biology, botany and zoology. Our collections are mostly in English and German, but also in small part Japanese and Russian. The catalog's language is English.
- Currently we are digitizing our author catalogue from 19th century to current
- We must build a virtual EMBRC Library so we need in the future to have specific new subscriptions to publications and reviews
- Suggest we try a Horizon 2020 partnership if available
- Access for external users has up to recently been more limited, though there is some experience among university libraries to provide services based on agreements for “pay-per-use”. Such agreements may form the basis for providing library services also in the EMBRC consortium.
One item of particular interest is access to scientific journals which are subscribed for: Such will often be restricted solely to users within the subscribing institution – as I would assume is the case for most other EMBRC partners.