Matteo Ciastellardi, Ugo Eccli, and Fabio Perotta

IN3 – Universitat Oberta de Catalunya, Università degli Studi di Milano
mciastellardi@uoc.edu

Abstract

Libraries are the traditional depot of culture and knowledge. The concept of digital libraries (D-Lib) emerged consistently with the growth of the web society and the broader diffusion of ICT systems and services. This idea empowers both the aspects of traditional editorial products and the way of conceiving digital education: on the one hand it implies a transformation in the cultural heritage of libraries, increasing the presence of digital contents; on the other hand it shifts the habitual approach to knowledge by offering a more sustainable distribution, a fastest access and a different user-centered concept to better satisfy the needs of the people.

Considering this scenario the paper would present the project Open Archipelago (OA) as a final delivery of a set of research activities oriented to offer an open source framework to innovate some aspects of fruition, distribution and management of editorial contents in libraries and in academic environments.

The project offers to the users a different cognitive approach to a wide variety of electronic materials and a more sustainable way to distribute, share and organize knowledge especially in the public spaces of libraries, research centers and universities.

Open Archipelago is based on a open source central system to index and to catalogue (and to store, in particular cases) Open Access materials, like digital versions of magazines, papers, books, etc. Connected to this main platform there are different end-user platforms called “islands” (clustered platforms as “islands” that creates the “archipelago”), in which such materials can be distributed through different typologies of devices (iPads, e-book readers, multimedia cards, USB keys, etc.) or consulted directly “on screen” (multimedia tactile screen) and partially printed according to the policies of the materials and of the hosting institution.

The main idea behind this kind of network is to create an “archipelago” of platforms (kiosks with different features) to put each user in a participative, interactive and immersive environment based on digital contents and to empower the heritage of the
knowledge encouraging the institutions to adopt and to distribute Open Access products with web 2.0 frameworks.

Such configuration allows also to design specific guidelines in order to offer a low-cost, sustainable, scalable and modular solution to implement a system based on new low-consumption devices, on online-trusted Open Access resources and on self-automated settings of distribution based on web 2.0 platforms.

The research focus started analyzing the issues related to the different approaches between digital natives and digital migrant in specific environments which offer editorial contents. One of the main topics that emerged was how to re-modulate the social inclusion for both these categories in the academic context giving a powerful and affordable solution in the hand of the institution. After a grounded analysis based on virtual ethnography research and on-field surveys and interviews, emerged mostly the aspect to reframe the practices to approach resources, not only in relation to the actual systems which offer the main index or/and an overall resume, but to access directly the whole content in few steps.

The answer to these feedbacks, taking also into account the need not to overlap with the best practices in the international context, was to design a framework oriented towards a methodological and structural innovation in the field of D-Lib cultural heritage based on well-profiled Open Access resources.

Introduction

Open Archipelago (OA) is a cooperative project in collaboration of two research institutes (Università degli Studi of Milano and IN3 – Internet Interdisciplinary Institute of Universitat Oberta de Catalunya). It is a framework prototype to collect and distribute Open Access materials in a web 2.0 perspective. The project started in 2010 and it is still in development, although some deliveries and the very first applications of the framework permit to offer practical outcomes for the scientific community.

The core idea of OA is a central system to index and to catalogue (and to store in particular cases) Open Access materials, like digital version of magazines, papers, books, etc. Connected to this main platform are different end-user platforms called “islands” (clustered platforms as “islands” that create the “archipelago”), in which such materials can be distributed through different typologies of devices (iPads, e-book readers, multimedia cards, USB keys, etc.) or consulted directly on screen (multimedia tactile screen) and partially printed according to the policies of the materials and of the hosting institution. In this sense, the name “Open Archipelago” defines the growing decentralized network that could be considered as an archipelago of platforms (kiosks

---

1 Open Access refers to unrestricted online access to articles published in academic journals. Open access products are mostly in digital format, online, free of charge, and free of most copyright and licensing restrictions.

2 In this sense Web 2.0 perspective means to provide the user with more user-interface, software and storage facilities, all through web browser. This has been called also “Network as platform” computing (O’Reilly, Tim, What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software, 2005, September 30th, URL: http://oreilly.com/web2/archive/what-is-web-20.html, last accessed 2011, January 10th)
with different features) and which is able to put each user in a participative, interactive and immersive environment based on open digital contents. Furthermore, the aim of OA is also to enrich the heritage of the knowledge encouraging academic institutions to adopt and to distribute Open Access products following the social and connective dimension of the web 2.0.

**Present scenario**

Libraries are the traditional depot of culture and knowledge. The concept of digital libraries emerged consistently with the growth of the web society and the broader diffusion of ICT systems and services. This idea enriched both the aspects of traditional editorial products and the way of conceiving digital education: on the one hand it implies a transformation in the cultural heritage of libraries, increasing the presence of digital contents; on the other hand it shifts the habitual approach to knowledge by offering a more sustainable distribution, a fastest access and a different user-centered concept to better satisfy the needs of the people.

In a moment in which university courses adopt a digital curricula of studies the libraries need to appropriate a new role as institutions oriented to support also in this direction both learning and research. This process is actually so very effective in public libraries as in academic libraries, where the digital heritage started growing some years before. On the other hand, during the same years, editors changed to digital process of publishing and many scientific publications switched to digital format with electronic version of journals. This change was the main reason of the birth of the Open Access movement and of its resulting system of publishing.

The number of journals obtained a remarkable growth and many institutional repositories of research products were opened in several universities. Considering the Italian scenario for example, the Telethon project, a fund raising program to help research on genetic diseases, adopted the Open Access policy to diffuse and communicate the research deliveries.

Today Open Access Journals are often defined as journals that use a funding model that does not charge readers or their institutions for access: from the Budapest Open Access Initiative definition of “Open Access” users have the right to “read, download, copy, distribute, print, search, or link to the full texts of these articles”. Existing Open Access

---

3 Digital libraries (D-Lib) are libraries in which collections are stored in digital formats and accessible by media devices (computer, portable readers, etc.). The digital contents may be stored locally, or accessed remotely.


5 For a more exhaustive perspective on the Open Access initiative: Hall, Gary, *Digitize This Book!: The Politics of New Media, or Why We Need Open Access Now*, Univ Of Minnesota Press, 2008.

6 Telethon Italy is an Italian NPO organization which goal is the aim of funding research on muscular dystrophies and genetic diseases. Telethon considers supporting unrestricted access to the published output of research a fundamental part of its mission. For this purpose Telethon has joined in 2010 the UK PubMed Central (UKPMC), a free-to-access digital archive of peer-reviewed biomedical and life sciences research.

7 The Budapest Open Access Initiative arises from a meeting convened in Budapest by the Open Society Institute (OSI) on December 1-2, 2001. The purpose of the meeting was to accelerate progress in the international effort to make research articles in all academic fields freely available on the internet. What emerged from the meeting was at once a statement of principle, a statement of strategy, and a statement of commitment (Budapest Open Access initiative, URL: http://www.soros.org/openaccess/index.shtml; last accessed: 2011, January 7th).
Journals directories, like DOAJ, consider this definition as mandatory for a journal to be included in the directory.

According to this idea, Open Archipelago could be understood also as the final delivery of a set of research activities oriented to offer an open source framework to innovate some aspects of fruition, distribution and management of editorial contents in libraries and in academic environments.

In order to reach this objective, OA project during its development challenged with some particular issues connected to the social and the institutional limits of Open Access initiative. One critical element in this view is that the Open Access products are not well known in “soft sciences” environment. If benefits of Open Access are evident even for the Humanities, the Open Access model is historically less known by researchers. In order to reach the necessary critical mass to consolidate the Open Access model in this disciplines would be necessary a reframed path for an immediate approach to the products. Libraries could be the very promoter able to respond to this goal of meta-dissemination: they have ever had cultural heritage goals, and a user-centered perspective. In this sense Open Archipelago would try to become a supporting tool to help to broadly diffuse the open-culture model, as seen in Open Source model for software or in cultural contents for licenses.

On another front, the growth of e-ink technologies with e-book reader devices has allowed to think a different way to deliver “paperless” digital contents. Considering OA as a framework to distribute open contents on these mobile and portable devices, we must distinguish between lean back reading and lean forward reading\(^8\). The first model is entertainment reading while the second one is more aligned to study or work purposes. Today e-book readers are mostly used as only lean back reading devices. To be useful for research and study purposes e-book readers need to have a wider diffusion in public spaces and improved accessibility features (technical, like a bigger dimension of the reading surface and a more scalable reading software – compatible with the most diffused formats – and commercial, like a more affordable price and a broader public diffusion).

Considering these aspects, the adoption and the diffusion of this new devices that OA project offers to the final users (student, researcher, professors, etc.) and to the platform managers (librarians and researchers as well) is a first step towards an overall framework able to cross-distribute in many ways digital contents and to encourage the adoption of a new reading style.

Approaching best practices

In this scenario, there are many best cases that have adopted technologies and models aligned to the aims of Open Archipelago project. The existence of overlapping technologies and similar objectives shows that the goals of OA are widely considered in different scenarios and the same OA can be partially designed by a mash up of existent experiences.

Among the different case analysed the most interesting practices for the OA development could be considered the following:

MedialibraryOnLine: an italian case. Horizons Unlimited (Bologna) has developed MedialibraryOnLine (ML), a platform aimed to share digital resources between

---

different institutions, especially public libraries. Public libraries, after a subscription and with an annual fee, can access a lot of digital resources. They can also buy collections of resources and offer directly to their users by ML platform. In this way ML can be seen as a model of remote digital lending. However one of the limit noticed in ML is that Open Access documents or Public Domain books, like the resources delivered by Project Gutenberg or Italian LiberLiber, can be downloaded only on subscription. This project is not oriented to index research documents. ML makes it possible to use an advanced search form that sends a single query at a time to other search engines like Google Books, Google Scholar or DOAJ, but it does not index directly Open Access resources.

PLEIADI Project: a second Italian institutional initiative. Pleiadi (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Deposti Istituzionali) is a project developed by Italian CILEA and CASPUR. It offers as a service the centralized access to the scientific literature archived in Open Access Institutional archives of Italian universities and in other Italian research centres. Pleiadi works as a service provider to collect and to index the metadata from Open Access Italian archives. In this way Pleiadi allows a simultaneous search from a single web interface over all indexed archives. Pleiadi covers only institutional archives: many digital resources published in Open Access journals are not collected.

DOAJ: a directory of journals. DOAJ (Directory of Open Access Journals) is a directory that provides access to referred Open Access Journals. The directory aims to be comprehensive and covers all Open Access scientific and scholarly journals that use an appropriate quality control system, and it will not be limited to particular language or subject area. The aim of the DOAJ is to increase the visibility and the diffusion of Open Access scientific and scholarly journals. Since DOAJ indexes only free, full text, referred scientific and scholarly journals, there are a lot of resources that have not been indexed because of a “lower” level of accuracy.

OpenDOAR: a directory of Open Access repositories. OpenDOAR is a reliable source of academic Open Access repositories. OpenDOAR has been identified as a key resource for the Open Access community\(^9\). The OpenDOAR in-depth approach to repository does not rely on automated analysis and gives a quality-controlled list of repositories. It appears to be for repositories the counterpart of DOAJ for Journals. Both have a stakeholders’ communities that contribute to the growth of Directories.

RePEc (Research Papers in Economics): a volunteer-driven initiative. RePEc aims creating a public-access database that promotes scholarly communication and at enhancing the dissemination of research in economics disciplines. The heart of the project is a decentralized database of working papers, journal articles and software components. All RePEc material is freely available. The participation in RePEc as a provider only involves the time of volunteers to prepare and to maintain metadata describing publications of institutional repository. But RePEc does not contain full-text journal articles. RePEc services provide links to many full-text articles, but a personal or institutional subscription is needed to follow these links.

OpenAIRE: a system for research. OpenAIRE (Open Access Infrastructure for Research in Europe) provides a network of open repositories providing free online access to knowledge produced by scientists receiving grants from the Seventh Framework programme (FP7) and European Research Council (ERC). This is the limit of OpenAIRE: a meta engine with a domain limited to FP7 and ERC researches.

Considering these initiatives among the others, it was possible to define specific guidelines and to adopt some of the models analysed to design the core project of OA. The aim was to define which were the best solutions to browse repositories, which were the most used tools to collect and distribute resources and to embed, when possible, the technologies used in other platforms in order to produce a sort of meta-repository for Open Access products.

**The Open Archipelago initiative**

Open Archipelago started in 2010 as a research initiative to innovate the traditional practices of distribution and collection of academic materials. The project focus moved from the idea to offer a parallel answer to research communities’ needs, not only offering another tool for libraries, but primarily an opportunity to broadly open knowledge in institutional situations where exists a considerable gap between digital and traditional resources.

The whole initiative, supported by an international team of researchers, has several aims, but the research focus was specially oriented to: facilitating and empowering the approach to the librarian heritage adopting an Open Access policy; decreasing the digital divide inside the academic institution by the introduction of resourceful system able to be easily used by digital natives and easily approached by digital migrants; allowing a sustainable access to the information, taking particular care of usability, low-consumption and saving policies promoted in the SDIs European Union indicators\(^\text{10}\); promoting an innovative and immersive approach to the cultural heritage, mashing-up previous technologies, in a web 2.0 vision.

**Concept**

The OA concept is based on two levels: the first one consists on a basic research to provide a framework capable to adopt, to share and to deliver collections of scientific Open Access resources; the second level consists in an applied research to propose a multi-platforms solution (the “Islands”) to manage these resources. The model that emerges is an open access network architecture which could manage resources as a traditional repository as well as a meta-crawler indexing system. The framework is based on a semantic hybrid database\(^\text{11}\) and on a

\(^{10}\) The Sustainable Development Indicators (SDIs) are parameters used to monitor the EU Sustainable Development Strategy in a report published by Eurostat every two years.

\(^{11}\) The concept of hybrid database is based on a bottom-up approach to define the relations and the tags shared between the different resources by the final users (Ciastrillardi, Matteo; Cruciani, Andrea; Miranda de Almeida, Cristina and de Kerckhove, Derrick, *The space between. Designing bottom-up knowledge in an interconnected society*, in Aa. Vv., Design Connexity Proceedings, Aberdeen, EAD Publishing, 2009, attached).
collection of resources that could be consulted, downloaded and tagged. The process follows the following steps (fig. 1):

A. Part of the people in a traditional library access to the core application (web-based) and start crawling the resources, according to their needs and to the suggestions visually offered after every research.

B. The core system (web-based) elaborates the queries and:
   B.1. it answers as a traditional database offering the most coherent information;
   B.2. it offers semantic driven answers, suggesting pertinent contents based on overlapping areas of interest, defined by the resources' tags and relations;
   B.3. it stores the queries analyzing the tags, the path of each research and the users' choices.

C. The database process the queries and it grows during the users' researches, adding new tags and reframing the existing classification.

D. The hybrid engine permits to have a dynamical map of the resources that change partially after each research, considering the communities those are using the system (in a local or in a networked distribution) and the distribution settings.

The hybrid core is a database powered by a metacrawler. This configuration allows to store resources and to index contents from different outside repositories, like DOAJ. This approach permits to have a clustered system for information retrieval, that means more resources with less expenses (time, human costs, database complexity, etc.) and with an automatic update sustained by the connection of different self-fed repositories.

Once resources are indexed, it is possible to distribute them in several ways (fig. 2).
Fig. 2 – Contents distribution from the mainframe

a) Directly, on site (in the library, on the screen of the different isles, with the possibility to print some parts of the contents)
b) Directly on personal/mobile device (downloading on usb keys, memory cards, tablets, smartphones, ebook readers, etc.)
c) Remotely by networks (accessing with personal computer to the full source via Internet connection)
d) Remotely by subscription (with feed and RSS distribution)
e) On demand by bookmarking (acquiring a QRcode of the resource with any device and downloading on demand in any synchronized platform)

Each “isle” (an interconnected web kiosk with touch screen) allows to search the Open Access resources, to explore each one and the related suggestions (matching and refining the research) and to transfer/bookmark to a specific device or to use on-site (read, print, etc.) (fig. 3).

Fig. 3 – Three moments of a research on an “isle”: search, match/ refine and transfer.

The pilot project at the Università degli Studi of Milano offers different isles connected with some ebook readers and tablets (fig. 4). Approaching the new devices people can also experiment different possibilities of fruition and the benefits related to portable systems, e-ink technology, on-demand resources, always-on/24-7 content access.
There are two kinds of “isles” platforms in the OA project:

a) The SEELE platform: main kiosk.
SEELE (Smart Electronic Environment for Learning Experiences) is a kiosk with a wide interactive full touch monitor to offer a reading experience directly on the screen. It simulates the book-format adopting a digital variant of the text, with pages to browse and to flip through with fingers.

b) The PEOPLE platform: satellite kiosk.
PEOPLE (Paperless Electronic Open Public Library Environment) is a kiosk optimized for Open Access materials delivery. It has a smaller screen than the SEELE version but with all the connections with the external devices to transfer directly all the resources offered.

In both platforms users can download the digital resources in many different ways. It is possible to download files simply “drag’n’dropping” with a finger the cover of the resource on an available device. There are so many possibilities as many devices are connected to the kiosk, from traditional USB keys or SD cards, to more advanced tablets (like IPad) or ebook readers.

When users find resources they can also acquire the QRcodes related to them: this way proceeding they do not download the contents, but only a link as a bookmark. In a second moment they will be able to manage this bookmark on their personal computer or on their mobile devices to download the resource connected to the QRcode.

**Development notes**

In order to accomplish to the mashing-up approach, which aim is to connect together, re-apply and empower previous existing (and broadly diffused) technologies, the OA project has moved from the best practices to isolate and to
define some particular open source technologies and some model to diffuse and distribute contents.

The research has analyzed the relation between the best cases previously indicated, scheduling the different layers that constitute these initiatives, and classifying the main features oriented toward a sustainable and innovative framework. As a second step the research has shifted to a deeper level of analysis, in order to better understand the missteps and the blindspots related to the different approach between digital natives and digital migrants in specific environments which offers editorial contents. The deliveries of this processes have permitted to design a preliminary framework of intervention and different guidelines on how to apply the OA model to different environments. The whole analysis was based on a virtual-ethnographic approach to understand which is the behavior of the people during the use of digital contents, especially in the four moments of:

a) Information retrieval
b) Information browsing
c) Information management
d) Information distribution

The evaluation of all the behavioral aspects before and after the use of the technological platforms has been investigated by traditional surveys and with the application of a RPA (Replication Protocol Analysis)\(^\text{12}\) to collect much information as possible in terms of feedbacks and motivational factors.

One of the main issue emerged has been how to re-modulate the social inclusion of the two principal categories of users: digital natives and digital migrants. The OA project has reframed some synesthetic strategies\(^\text{13}\) in order to develop informative system for unpaired people (audiovisual and tactile approach) and to offer a more reliable environment for both the users’ categories. It also moved from some basic principles of ergonomics in design\(^\text{14}\) to encourage the use of the platform with everyday’s life objects.

After grounded analysis based on virtual ethnography research and on-field surveys and interviews, emerged mostly the aspect to reframe the practices to approach resources, not only in relation to the actual systems which offer the main index or/and an overall resume, but to access directly to the whole contents in few steps.

**Technical aspects**

OA needs housing on server Linux (LAMP), BSD or Unix-like with this features: PHP 4.2.X or higher with functionality MySql, XML and Zlib; MySQL database 3.23.X or higher, 500 MB minimum space, minimum 5 db; Apache web server 1:13:19 or higher; PERL; htaccess configuration.


The housing should also include a minimum amount of web space of 5 Tb and no limit or a minimum bandwidth guaranteed daily, with possibility to extend it according to future needs of the system. All the kiosks can incorporate a free WiFi Hot Spot. There are network/WiFi policies that limit the navigation to the Open Access resources.

Conclusions

OA Projects is a young framework of activities with several aims. As any project with a multi-purpose direction, OA presented some positive deliveries but also blindspots and missteps to correct. The very first deliveries, related to the pilot experiment in Milano, are the partial change of the traditional librarian environment (by the introduction of the “islands” platforms) and the correlated change of users’ behavior interacting with new tools to discover information related to their own researches. This result allowed to change partially the perspective of the library as simple repository of resources, introducing the idea of a more flexible and helpful user centered environment. Traditionally the library was approached with a clear idea of the contents, authors and resources needed. Researchers were able to deepen their analysis using the resources (books, catalogues journals and so on). Now they have one chance more by the semantic connective interface.

In the early stage the main trouble expected, the “technological gap” using new devices and kiosks, has not been a real issue: the easily approach to the resources and the intuitive interface allow a friendly approach to the overall framework. Considering the blindspots, the project is now trying to deal with the issue related to the education of librarians as manager of the system. Archipelago is considered a set of parallel activities that challenge with the traditional approach to libraries. This causes a complicate first impact for the librarians: they have to take care of another technology with another dimension of use. Fortunately after the first steps the platforms are mainly self feeded by users and could be self feeded also in terms of updating by the connection to the database of the different repositories.

The experience of Archipelago is only at the beginning, but the possibilities are many and completely oriented toward a scenario of open knowledge and sustainable connective growth.

References


