

Knowledge Isles in an Open Access World: The Open Archipelago Project

Matteo Ciastellardi¹, Ugo Eccli², Fabio Perotta²

¹ IN3 – Universitat Oberta de Catalunya

² Università degli Studi of Milano

Introduction

Every culture has been characterized by different systems to transfer and reproduce knowledge overtime, generation after generation. Our culture is discovering the capacity to create and transmit knowledge in a shared, participative way. If the collective and connective approach is growing faster mostly in online communities, thanks to the wide diffusion of social networks, in the academic environments (Universities, libraries, research centers, etc.) sometimes there is a lack of cooparticipation between the different actors.

Concerning this attitude, it is interesting to remember the words of McLuhan:

The university and school of the future must be a means of total community participation, not in the consumption of available knowledge, but in the creation of completely unavailable insights. The overwhelming obstacle to such community participation in problem solving and research at the top levels, is the reluctance to admit, and to describe, in detail their difficulties and their ignorance. There is no kind of problem that baffles one or a dozen experts that cannot be solved at once by a million minds that are given a chance simultaneously to tackle a problem. The satisfaction of individual prestige, which we formerly derived from the possession of expertise, must now yield to the much greater satisfactions of dialogue and group discovery. The task yields to the task force.

Marshall McLuhan, 1971, convocation address at the University of Alberta

According to McLuhan idea, if a first step is to sustain and promote the “task force”, a second mandatory step must be to offer the most reliable tools and to design the most sustainable environments to trigger the processes of sharing, distributing and accessing knowledge.

The aim of this paper is to analyse the contemporary situation starting exactly from specific traditional learning environments: the libraries as the first depot of culture and knowledge, and then to reframe these social spaces according to the transition towards the 2.0 (digital) panorama.

One of the milestone in our scenario is the key concept of digital libraries (D-Lib), that 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 fast-

est access and a different user-centered concept to better satisfy the needs of the people.

In recent years there has been some important changes in the way to access to digital resources on three trends: growth of online access to resources, development of paper-less reading devices, open culture diffusion. This last trend has been crossed the boundaries of the software development world, the Open Source movement, and has involved the production of cultural contents through the deployment of a lot of types of free licenses, such as Creative Commons.

Considering this scenario, the paper would present a specific action oriented to design sustainable and more interactive environments inside traditional spaces like libraries, offering also an open framework to access and distribute resources.

This action is focused in the Open Archipelago project (OA): it represents a final delivery of a set of research activities to innovate some aspects of fruition, distribution and management of editorial contents in libraries and in academic environments with an open source approach.

Open Archipelago is based on a open source central system to index and to catalogue Open Access materials and to redistribute them in "islands" (end-user clustered platforms 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.

Open Archipelago has the aim to create a community of librarians that develops a directory of well profiled resources. In this way we want to stimulate the cooperation between two communities: researchers and librarians. In this sense OA will try to define an active and participative role for librarians, putting them in charge of the transformation from traditional environments to digital libraries.

The OA system of platforms allows also to follow specific design 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 new web 2.0 technologies.

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.

The Open Archipelago project

Open Archipelago (OA) is a cooperative project under development by Università degli Studi di Milano and IN3 – Internet Interdisciplinary Institute of Universitat Oberta de Catalunya. It is a framework prototype to collect and distribute Open Access⁹⁹ materials in a web 2.0¹⁰⁰ 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 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. In this sense, the name “Open Archipelago” defines the growing decentralized network that could be considered as an archipelago of platforms (kiosks 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.

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 (Roncaglia, 2010). 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

99 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.

100 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 2005).

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 Depositi 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 (Oliver & Swain 2006). 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.

Concept

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 indicators¹⁰¹; promoting an innovative and immersive approach to the

cultural heritage, mashing-up previous technologies, in a web 2.0 vision.

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 database¹⁰² and on a collection of resources that could be consulted, downloaded and tagged. The process follows the following steps (fig. 1):

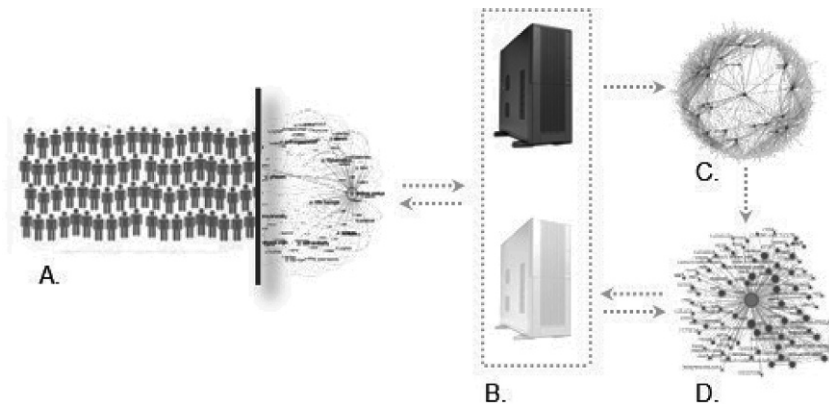


Fig. 1 – Information process

Part of the people in a traditional library (A) 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.

- The core system (web based) elaborates the queries (B) and:
- it answers (C) as a traditional database offering the most coherent information;
- it offers semantic driven answers, suggesting pertinent contents based on overlapping areas of interest, defined by the resources' tags and relations;
- it stores (B') the queries analyzing the tags, the path of each research and the users' choices.
- The database process (D) the queries and it grows during the users' researches, adding new tags and reframing the existing classification.
- The hybrid engine permits to have a dynamical map of the resources (A') 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

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

102 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 (Ciastellardi et al. 2009).

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-fedded repositories.

Once resources are indexed, it is possible to distribute them in several ways (fig. 2).



Fig. 2 – Contents distribution from the mainframe

Distribution	
Directly, on site in the library.	on the screen of the different isles, with the possibility to print some parts of the contents.
Directly, on personal/mobile device.	downloading on usb keys, memory cards, tablets, smart-phones, ebook readers, etc.
On site, by WiFi network.	accessing the full source via WiFi connection by a personal device.
Remotely, by networks.	accessing the full source via Internet connection.
Remotely, by subscription.	with feed and RSS distribution.
On demand, by bookmarking.	acquiring a QRcode of the resource with any device (smart-phone, for example) and downloading on demand in any synchronized platform.

Table 1 – Distribution system

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.



Fig. 4 – An “isle” with two devices remotely connected to transfer Open Access resources.

There are two kinds of “isles” platforms in the OA project:

1. 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.
2. 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 blind-spots 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:

- Information retrieval
- Information browsing
- Information management
- 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) (Galle & Kovács 1996, pp. 181-200) 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 (Anceschi & Riccò 2000) 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 to encourage the use of the platform with everyday's life objects (Cooke & Salas 2008).

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, was not 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

Aneschi, G & Riccò, D 2000. Research of Communication Design: a synesthetic approach, *Proceedings of the «Design plus Research» international conference*, Politecnico of Milano.

Ciastellardi, M, Cruciani, A, Miranda de Almeida, C & de Kerckhove, D 2009. *The space between. Designing bottom-up knowledge in an interconnected society*, in: Malins, J (Ed), *Design Connexity*. Proceedings of the Eighth International Conference of the European Academy of Design, Aberdeen, EAD Publishing.

Comeaux, DJ 2008. Usability Studies and User-Centered Design in Digital Libraries, *Journal of Web Librarianship* 2, no. 2: 457.

Cooke, NJ & Salas, E 2008. *Best of Human Factors: Thirty Classic Contributions to Human Factors/ Ergonomics Science and Engineering*, Human Factors & Ergonomics Society, Santa Monica, CA.

Galle, P & Kovács, LB 1996. *Replication protocol analysis: a method for the study of real-world design thinking*, *Design Studies* 17, no. 2.

Hall, G 2008. *Digitize This Book!: The Politics of New Media, or Why We Need Open Access Now*, University Of Minnesota Press.

Oliver, KB & Swain, R 2006. Directories of Institutional Repositories: Research Results & Recommendations, *72nd IFLA General Conference, Proceedings*. Available: http://archive.ifla.org/IV/ifla72/papers/151-Oliver_Swain-en.pdf [Accessed 7th January 2011].

O'Reilly, T 2011. *What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software*, September 30th 2005. Available: <http://oreilly.com/web2/archive/what-is-web-20.html> [Accessed 10th January 2011].

Roncaglia, G 2010. *La quarta rivoluzione. Sei lezioni sul futuro del libro*, Bari , Laterza.