Mixing the Physical and Digital Libraries :

Algorithmic, Visual and Interactive Tools of Book Discovery

Eric Briys – Cyberlibris (Founder)

Laure Delrue – Lilliad learning center Innovation (Head of the Collections Department) – University of Lille

Clément Roux – Cyberlibris (Data scientist)

To be presented at **The Fiesole Retreat Series**
"Serving Learning and Scholarship"
Barcelona, April 25-27, 2018

"... l'exemplarité de la bibliothèque publique repose sur son aptitude à accompagner les transitions dans lesquelles sont entraînées les pratiques culturelles et tout spécialement celles concernant la lecture."

Luigi Failla, La bibliothèque comme espace public:
Du livre à la ville, p10, MētisPresses, 2017

Introduction

It takes two to tango, a fortiori to e-tango. With the advent of ebooks and the tsunami of digital uses, libraries have to cope with an odd dancing couple, order and disorder. They are indeed home not only to (print) atoms[1] but also to usage bits. Order is part of the DNA of libraries. In the seventeenth century, Claude Clément, a French Jesuit, recommended that libraries shelved their books according to twenty-four broad categories matching the curricula of studies in the universities (Lerner (2009)). Melvil Dewey introduced the eponymous classification to enable the arrangement of books on shelves. Whatever the method, the aim is the same: Order. This order rhymes with the limited physical space in which it is deployed. It tries to optimally fill it and, in the course of doing so, to make it highly affordable to patrons.

Nowadays, usage bits are gushing forth at digital speed, and it is fair to say that their apparent disorder is at odds with the library system of order. Does this imply that although they are two, order and disorder cannot tango/e-tango? Has the library become an oxymoron or are there ways in which print atoms and usage bits can enrich each other? Will patrons truly benefit from a hybrid experience? The truth is that the library shall not ponder whether it is physical and/or digital, whether order shall take precedence over disorder or vice versa. The library long

[1] Print is in bracket to emphasize both the print atoms and the building atoms. Brackets are lifted hereafter.

history, its resistance to the passage of time are testimonies of its ability to handle these questions. The so-called digital age is no exception.

This paper explores a real world case, namely that of the library of the University of Lille. It analyzes how this library has tackled the challenges brought by the juxtaposition of print atoms and usage bits, which steps, among other things, it has taken to enhance the book discovery process and how these changes have been perceived by patrons. The exploration path is structured as follows. A first section goes back to basics. It contrasts the physical library and the digital library in terms of a simple and straight trade-off between fidelity and convenience. A second section builds on this trade-off to draw its architectural implications. A third section looks at the cultural implications of (big) data, more specifically of databases. A database can be viewed as a cultural object whose structure can and shall be deciphered. In the library case, it may for instance contain the aggregation of hundred of millions of microscopic reading patterns from which, thanks to algorithms, a complex and useful macroscopic library pattern may emerge. A fourth section details how this macroscopic pattern has been extracted and converted into a visual and interactive platform, namely LilliadVis.io, accessible to University of Lille patrons. A fifth section is devoted to a first experiment that was recently conducted among a group of students in order to assess their evaluation of the new LilliadVis.io platform. A conclusion summarizes the main findings.

1. Physical libraries, digital libraries: Their geometries and embedded trade-offs

Physical libraries are remarkable tributes to Euclidian geometry. Indeed, a straight shelf can always be "drawn" between any two books[2]. Given a point and a shelf not crossing that point, there is only one shelf which can be built through that point and which is parallel to the first shelf. As a result Euclidian libraries are remarkable interfaces facilitating the discovery of the works they contain. Musing

[2] Although some designers have created curvy shelves, these are usually limited to a few shelves hanging on a domestic wall.

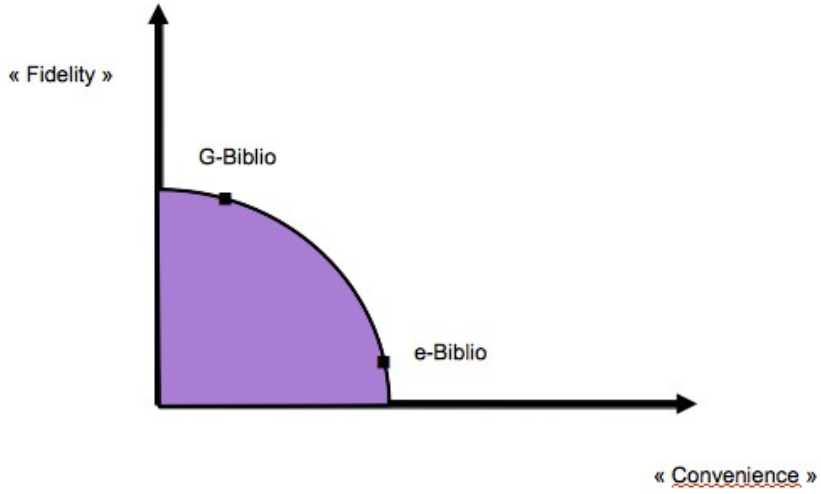
in a physical library is indeed a highly structured and organized experience. Books sit (and wait) on shelves. Not any shelves though. These shelves are meticulously filled with books by librarians who have spent years assigning metadata to books. As a result, libraries are remarkable interfaces where eyes can collect a wealth of information in one go. To use a term borrowed from psychology, the physical space is structured so that its affordance is maximized for patrons musing through the shelves. But this affordance has a price. Under Kevin Maney's insightful fidelity-convenience framework[3] (2010), the library Euclidean geometry provides patrons with a high degree of fidelity that comes at the cost of a lower degree of convenience, words that he defines as follows:

« *Fidelity is the experience of something – not just how good it is, but how it makes you feel or what it lends to your personal identity. Convenience is how easy it is to get something. So if a product or service is ubiquitous and cheap, it's pretty convenient. Think Wal-Mart or McDonald's*. »

While a physical library constitutes an hospitable space and represents an interface to which time has conferred a remarkable patina, it does impose constraints on patrons. Indeed, when a book is borrowed, the next patron cannot access it as long as the book has not been returned. The whole library itself cannot be taken home. It is bounded in space, and its storage capacity is far from unlimited. With the emergence of e-books, the library has become digital. Its physical anchor has disappeared. The ensuing loss in fidelity (and quality of experience) is compensated by a gain in convenience. Indeed, the patron no longer benefits from a dedicated physical space in which she can easily cast a gaze on a complete collection, query the librarian, move around the shelves and sit down for several hours of comfortable perusal. Musing in a digital library is radically different, almost an oxymoron. Even though it is easily accessible through any computer anywhere, it can be a frustrating experience. The reader is limited

[3] He also calls it the fidelity swap.

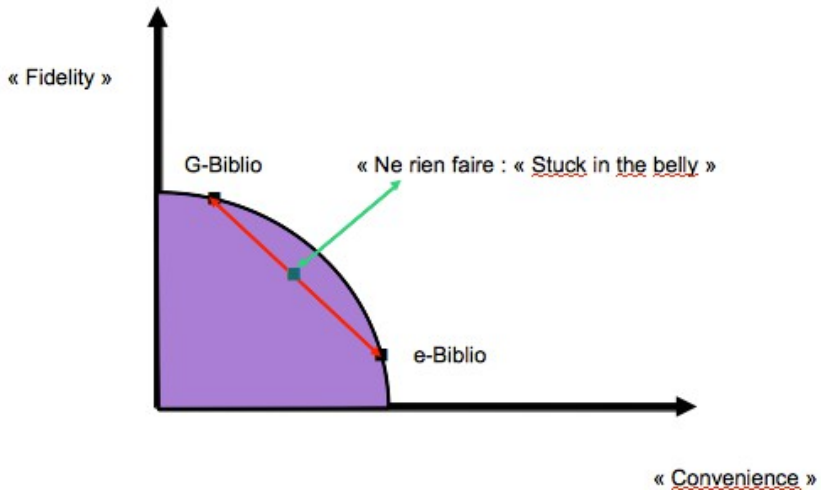
to a small 2D space where not much can be displayed in one go. This is a significant constraint. However, the patron can take the digital library home with her through an Internet connection. She has the possibility to full-text search books, read any book online even though two thousand patrons are also reading the same book at the same time. The following graph visually summarizes the advantages and drawbacks of the two libraries:



G-Biblio represents the physical library, and e-biblio the digital library. G-Biblio offers its users a high degree of fidelity and experience but is not without drawbacks that do decrease its convenience. While highly practical, e-Biblio is an interface devoid of the riches associated to its physical alter ego. All in all, what patrons agree to lose in terms of fidelity (convenience) is something they try to regain in terms of convenience (fidelity). Provided that the result stays on the black perimeter, patrons are willing to accept the trade-off. None of the dots inscribed in the violet area are right for patrons, the reason being that all these dots are outperformed by the dots of the black perimeter, which delineates the available compromise between fidelity and convenience, compromise on which patrons are ready to sign off.

These days most libraries deliver both physical and digital options. They offer the two interfaces to their patrons. However, this dual display is more often than not

limited to a straight juxtaposition of the two interfaces. As we write and to the best of our knowledge, there is very few full-fledged exhibition of spaces in which the two interfaces are merged into novel ones, not to say hybrid ones. The current simple cohabitation is summarized in the following graph:

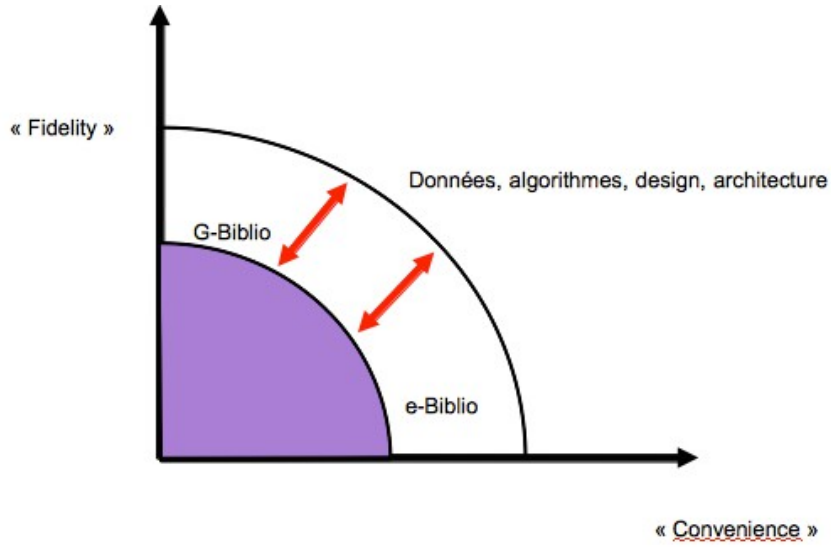


The red line illustrates the different set of “linear” combinations of the two libraries according to the respective weights they are given. For instance, the green dot at the middle of the red line indicates that each library is granted the same importance. The key factor underlying the geometric locus of library combinations is that mixtures of the two libraries are “linear” mixtures. Nothing is accomplished other than allowing the physical and digital libraries to coexist and to be accessed. Users will at times use one, and at times its counterpart. The introduction of the digital library does not modify the design and organization of the physical library. The physical library does not alter its digital counterpart either.

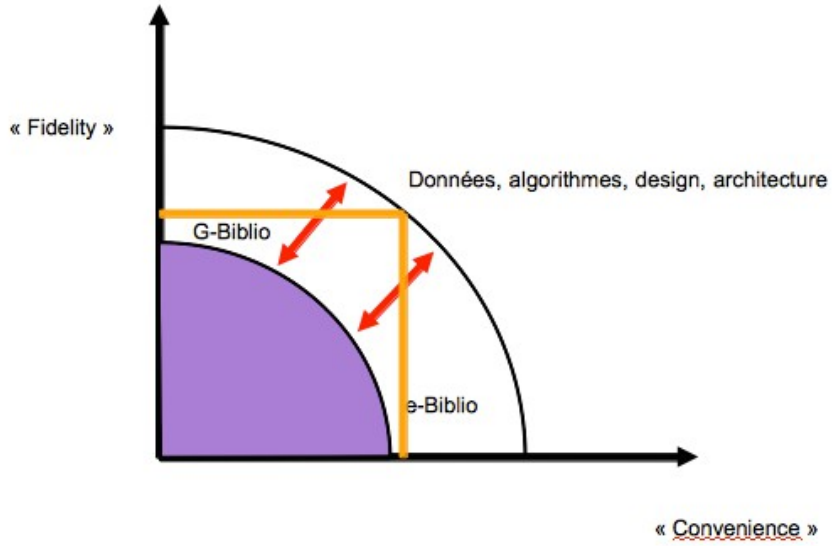
Situated within the violet zone, the green dot is « stuck in the belly ». It constitutes a weak spot away from the trade-off perimeter. It is overhung by the dots scattered on the latter. It represents a weak compromise. This fallback onto an unsatisfactory middle ground is hardly astonishing. At this stage of the game not even the slightest synergy between the two libraries has been explored, much less put into play. Withdrawal into the belly cannot however be the last word. The

two libraries are bound to enrich one another. Geometrically speaking, the hope is that the initial perimeter can be moved towards the northeast, thereby freeing up new and more fruitful opportunities to effectively connect fidelity and convenience.

Can such a felicitous displacement be carried out? The response is affirmative. It echoes Luigi Failla's quote at the beginning of this paper. In order to successfully make this move, numerous professions have to be called upon to contribute. The skills of the librarian, the teacher, the patron, the architect, the engineer, the designer, the data scientist, the web designer etc....are all needed to efficiently capture the new library “zeitgeist”.



These coordinated efforts will enable the emergence of a new fidelity/convenience perimeter with enhanced appeal for each and every user. Graphically, the objective translates into reaching a point (such as the one indicated below) at the intersection of the two orange segments. Maney calls it the mirage point:



This new point concomitantly offers more fidelity and more convenience than would either the physical or the digital library on a stand-alone basis. The patron is better off in all dimensions. She can benefit from a library experience that is at both and the same time of enhanced quality and improved convenience. Is this point a chimera? What does it imply for the current architecture of the physical library? How does it translate into a brand new fidelity-convenience space accessible to patrons and librarians?

2. From print atoms to usage bits : Architectural implications

In his remarkable latest book Luigi Failla (2017) ponders whether libraries will become « museums shelving a technology (the book) whose time has gone », whether they will mutate into some kind of « closed server rooms with remote access » or, last but not least, whether « they could become an urban device able to cope with the loss of public spaces in today's metropolitan areas. » It is no coincidence that these questions are asked by an architect. At the core the issue is an authentic architectural one. Even though some authors argue that « one is no longer at a time when the virtual was in direct opposition to the real and was threatening its stability. » (Picon (2015), again an architect!), one still wonders

how one can mingle bits and atoms to deliver hybrid services, hybrid experiences, in other words how one can pave a seamless way between waves of bits triggered by patrons usages and atoms. Lilliad learning center founded by the University of Lille is a bold step into that direction. It echoes French librarian Michel Melot's statement (2004) : « The library is less and less about collections and more and more about architecture. »

Luigi Failla notes that in the history of architecture the library belongs to some of the oldest architectural typologies in the same way as the domestic house and religious buildings. This observation has far-reaching consequences. It means among other things (the obvious) : Atoms have a longer history than bits. Library atoms are embedded into rituals, rituals polished and centered around the print book and its reading. This patina is a healthy sign. It means that the library ages well, so well that one can dare to say that it ages according to the Lindy effect. The Lindy effect is a puzzling explanation as to why some businesses survive way longer than others. The Lindy Effect states that the current lifetime of a non-perishable item is most likely to be at its half-life. In other words, a business that has been around for the last 1000 years should expect to be around another 1000 years. In short, its mortality decreases with time. To understand why this effect comes into play, it is useful to think of time as a proxy for disorder. Things that have been long with us are things that have best resisted the teeth and ravages of time. They have handled and resisted the ups and downs as time flew by. This capacity is a testimonial to their antifragility (Taleb (2012)). The library has been long with us, a sign of its ability to take advantage of disorder.

If one follows Failla's historical record, the recent past (last 40 years) provide additional evidence of the library antifragility. In the late seventies up to the nineties, the library was (still) the print book locus. It was shaped by it. The library building was architecturally structured along three levels. Level 1 handled the reception desks and general cultural activities (in the wake of the print book).

Level 2 was structured around more specific content information required by patrons and individual reading. Level 3 was dedicated to specialized research, complex library services and the preservation of documents. To put it in Failla's words, « the conception of the three level library stems from the idea of a path, a progression towards the proper use of the locus, a use still marked by silent reading. » Failla concludes that there is a clear distinction in the three level library between the spaces where the print book dictates its rules and the spaces where the patrons impose new behaviors which impact on the library organization. The print book rules dominated the three level library. They shaped its architecture.

During the first ten years of the 21st century (2000-2010) a two level library emerged. The third level tended to disappear as a specific entity. It was often merged with the second level. According to Failla, this move corresponded to a shift from a three level library that was fully centered around the document to a two level library in which the first level tended to be structured more and more around the patrons and their new usages. To put it in a nutshell, the bits started to shake the atoms. The level demographic transition was not over. Failla observes a new trend towards the merging of level 1 and level 2 (Failla, p 123). The impact of bits on atoms has become a lot stronger as over the last ten years digital uses have truly exploded (outside and inside the library).

These last 40 years offer a vivid testimony of the ability of the library to handle the increasing disorder brought by blossoming digital usages. To cope with this disorder, to take advantage of it, the library has changed gear. Its focus has shifted from a print atoms one to a (usage) bits one. This evolution is aptly summarized by the architects of University of Lille Lilliad learning center :

« The university presents itself as a modern institution, able to connect innovation with tradition. The Learning Center offers many different learning environments. Boundaries dissipate in the digital age; the classical library is obsolete; digital

media and interaction are gaining significance and cause the dissolution of the spatial boundaries in the learning areas. The building sketches a landscape of knowledge, surrounding a central hall as a communicative area. Various learning institutions flow into each other and create spaces with multifunctional uses, providing the users with spaces to be used flexibly. »[4]

This shift to usage focus begs an immediate question: What do we precisely know about usages beyond the usual discourse ? The very fact that atoms do converge to bits, that boundaries between them are blurring does not mean that we are all set. Indeed, bits are not yet at the same ritual stage as atoms. We are still at a stage where the hybridization of bits with atoms is far from complete and accessible. For instance, even though the notion of a smart city is on many lips, its « smart library » equivalent is still an infant concept in search of a proof. And, as the Anglo-Saxon common saying goes, « the proof is in the pudding ».

Our good fortune is to have an impressive pudding at hand : The data pudding.

3. From the data pudding to hybrid library experiences through complexity

There is perhaps no better place/space than a library to understand what we mean by data pudding and what its implications are. In his book titled The Language of New Media, Lev Manovich (2001) observes that the novel, and subsequently cinema, privileged narrative as the key form of cultural expression of the modern age. Libraries are agents of this narrative predominance as they are places where, among other things, novels are shelved and lent to patrons. The novel (and cinema) narrative mode is compact : Stories have a beginning and an end. So has their container, the print book. In the computer age (to use Manovich words), things have changed. As surprising as it may be, the dominant form of cultural expression is the database. A database is a structured collection of data,

[4] https://www.archdaily.com/791427/lilliad-learning-centre-innovation-auer-weber

a collection of items that can be fast searched and retrieved. Thanks to a database of images one can for instance organize a virtual museum in which the user is able to take a digital tour of the museum collections. In this respect, a book is also a database. Mining its text (database of words) allows to define new ways to interact with the book content beyond its traditional narrative. As a result, while the database is an authentic form of cultural expression, it opens a Danaides' barrel. Manovich notes that «as a cultural form, database represents the world as a list of items and it refuses to order this list. » He adds that « In contrast, a narrative creates a cause-and-effect trajectory of seemingly unordered items (events). » Database and narrative seem to be born natural enemies. Manovich also notes that « competing for the same territory of human culture, each claims an exclusive right to make meaning out of the world. »

Interestingly enough a similar tension is to be found in the library, a tension between what is ordered and what is not. A physical library is an ode to order. Librarians are the custodians of this order. Print books are shelved according to this authoritative order. The library thus creates a narrative. One could even argue that the physical library itself becomes the narrative. In the same way as words printed on paper becomes a story, shelved print books in a library create a well-defined and unique path[5]. But, at the same time in the current digital age, the library mutates. It becomes a database. Usage data, be they physical (loaned books) or digital (reading behavior data), can indeed be collected to constitute a database, a data pudding. They accumulate with no limit at digital speed. Taken by digital storm, the library tends to become an oxymoron. A new (digital) cultural form challenges its cherished order and narrative. An order backed by a strong physical narrative now coexists with a growing digital disorder with no narrative[6] whatsoever.

---

[5] Unique in the physical constraint sense.

[6] « All this further contributes to the anti-narrative logic of the Web. If new elements are being added over time, the result is a collection, not a story. Indeed, how can one keep a coherent narrative or any other development trajectory through the material if it keeps changing? » Manovich p 196

This tension, when properly understood and mastered, is more of an opportunity than a threat. In a sense, more can be more, in an algorithmic sense. An algorithm is a sequence of operations which are carried out by a computer. When applied to a given database the algorithm transforms the initial database into a new one. The mash-up of geographic coordinates and people urban movements into a visual map is an example of this transformation. Mixing two sets of collected data creates a new set of data deemed to be of help to end users. At the risk of oversimplifying, data base and algorithms are the two sides of the same coin. In Manovich words « algorithms and data structures have a symbiotic relationship. »

This brings us back to the library as a data pudding. Pudding is the right word to describe the kind of database one assembles when one gathers both physical and usage reading data. These data stem from physical usage when they relate to the borrowing of print books by patrons. They are linked to patron digital usage when they capture books read online, books assigned to digital bookshelves, annotated books, commented books etc... Looking at these massive data (literally) is fascinating. It is a dive into a complex system in which the microscopic components consist of patrons reading books, in which the macroscopic component is the collective, complex, and noisy behavior of the library as a whole. Whereas the physical library obeys an order predicated on a top-down approach driven by librarians (down to patrons), the data pudding seems to obey some kind of bottom-up spontaneous order, of self-organization, and emergence in which the total is in the end more than the sum of the parts. According to Geoffrey West (2017) « a typical complex system is composed of myriad individual constituents or agents that once aggregated take on collective characteristics that are usually not manifested in, nor could easily be predicted from, the properties of the individual components themselves. » As « a city is much more than the sum of its buildings, roads and people » (West) the data pudding is much more than the sum of its data. Like colonies of ants, the colony of patrons build an (up to now) invisible digital network, an invisible digital library. This observation is the reason

why this text now (slightly) departs from Manovich's framework, a framework in which the notion of complexity (in the scientific sense) was not explored. To be fair to Manovich's book, it was published in 2001 at a time when the algorithms, complexity and artificial intelligence trio was not full steam ahead.

The beauty of the (yet to be unveiled) structure of the data pudding is that there is no central control. It stems from an emergent behavior in which the components (say patrons, books or both) agglomerate to form the emergent whole (say a reading network or a library “ordered” according to usages). A bottom-up self-organization grows and constantly adapts itself to changes in its environment, changes like for example the addition of new patrons, new books, new comments, the emergence of best-sellers, the arrival of new publishing houses etc... In contrast, the traditional physical library is a top-down system in which the librarian control/authority is at work. Without the help of algorithms though, there is little hope to make, one way or the other, the emergent whole either visible or actionable. As stated earlier the data pudding and algorithms are thankfully the two sides of the same shiny coin. The algorithm we refer to here are machine learning algorithms, a different breed than the one Manovich was alluding to in his book. These algorithms will help us learn and unveil the emergent “wholes”. We write wholes on purpose as there are many wholes depending on the angles one wishes to take. In that respect, Manovich is right in stressing out that when a database is considered as a cultural object there is no longer one single legitimate narrative. When we started building LilliadVis.io we knew we needed different angles to mesh several narratives into one convincing story. LilliadVis.io was imagined to unveil the invisible.

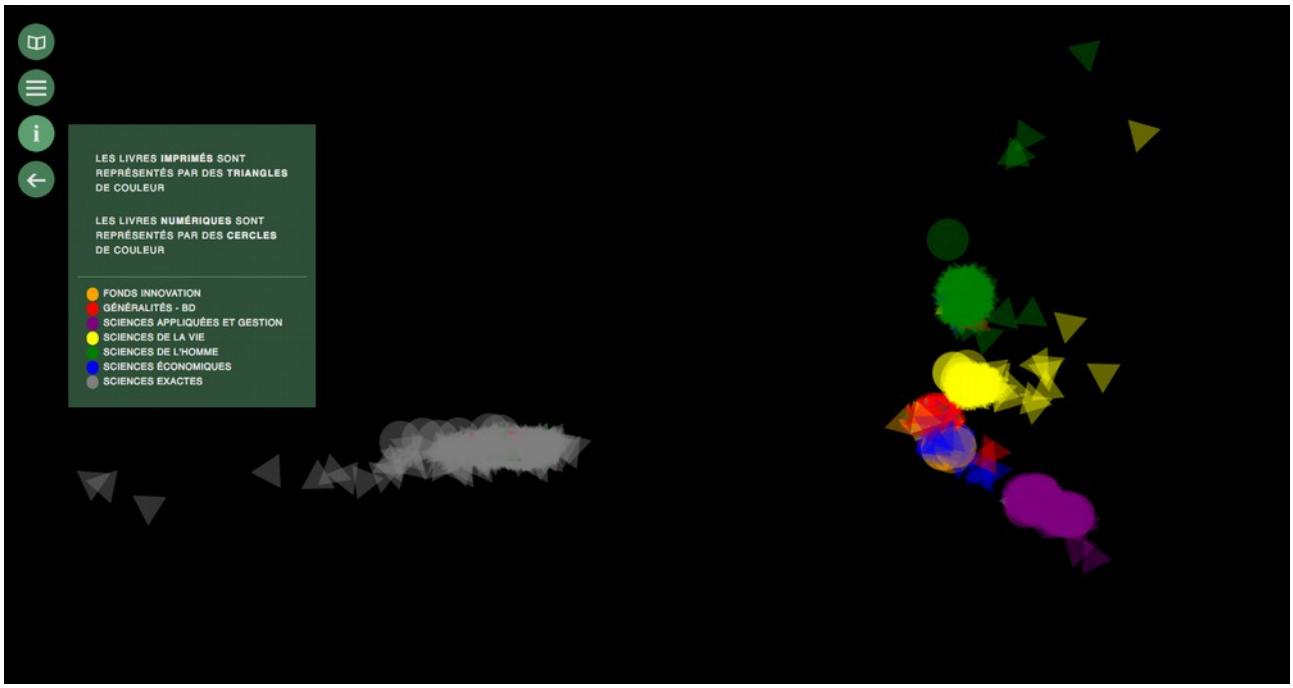
4. LilliadVis.io : Unveiling the invisible

In the data pudding there is more than meets the eye. To understand why it is time to be more specific about the data pudding. In machine learning parlance, the

data pudding is the so-called adjacency matrix (Newman (2017), Barabási (2017)). An adjacency matrix is a matrix representation of a network emerging from the data and made of edges (links) and vertices (nodes). Vertices are joined by edges. The elements of the matrix obey the following rule : They are equal to one if there is an edge between two given vertices, to 0 otherwise. The adjacency matrix is one of the simplest ways to represent a network on a computer. Indeed, storing the (data pudding) network in the form of an adjacency matrix is convenient as it allows to develop formulas and computations using mathematics, especially linear algebra. One should also bear in mind that there are multiple adjacency matrices that one can use to represent a network. This is what we meant when we talked about the different angles one can take to decipher a network. In our case for instance, the network can be represented either in terms of patrons (each patron is a vertex) or in terms of books (each book is a vertex) : Two patrons are linked to each other if they « share » a given book, two books are linked to each other if they have readers in common.

LilliadVis.io has been built thanks to a massive amount of reading data and a robust story meshing four basic narratives. The reading data were not restricted to University of Lille patrons only. On the contrary. This is why they are so numerous. Anonymized reading data stemming from all academic institutions accessing http://univ.scholarvox.com were used (ScholarVox is one of the digital libraries designed by Cyberlibris). There is indeed no reason why one should not take advantage of the collective reading wisdom embedded into all these institutions. As far as the story is concerned, it has been written according to the following scenario. First, we wanted to start from a library, not any library though. The physical library is an interface designed according to a top-down pattern. We needed one interface too. Hence we wondered how the digital library would be organized if one were only relying on an algorithmic bottom-up approach, more specifically on usage data.

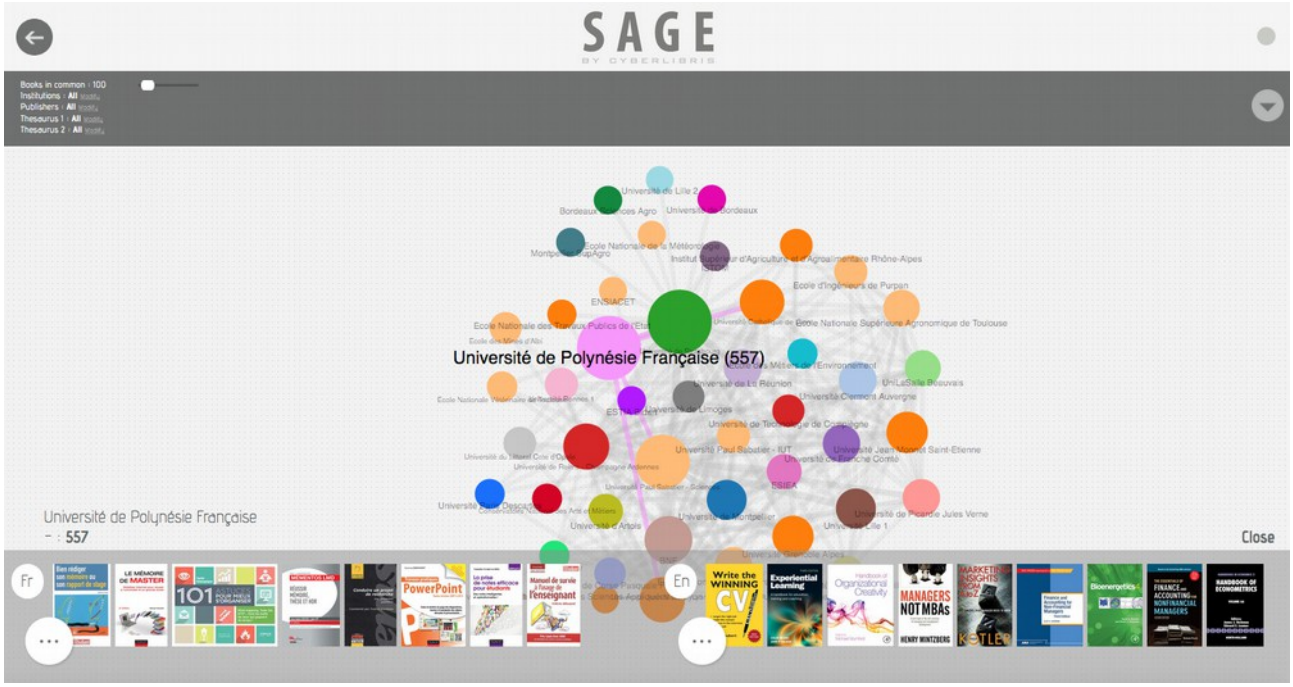
The **DICE** (Digital Content Explorer, Portulan in French) is born from such an attempt. It has a clear objective : To be able to compute the social book graph embedded into the data pudding, to visualize it and to interact with it. The following screen capture is what patrons and librarians look at when they play with the DICE.



The DICE's algorithm is predicated on a simple observation : Books can either be considered as bags of words – based on traditional librarian keywords such as the Dewey classification – or as bags of readers – based on the readers' usage data. As a result, the DICE gives the opportunity to switch from a standard reading recommendation relying on classifications to a more serendipitous recommendation balancing usage data and classification keywords (For more details refer to Nock, Nielsen & Briys, 2013 and Briys (2017)). In LilliadVis.io's case, as can be seen from the open panel in the above screen capture, the ambition was to include not only e-books but also print books. Print books are indeed borrowed by patrons. As such they emit data points that have been added to e-books data points. In the DICE, print books are represented by colored triangles, e-books by colored circles. In order to truly anchor the DICE into its

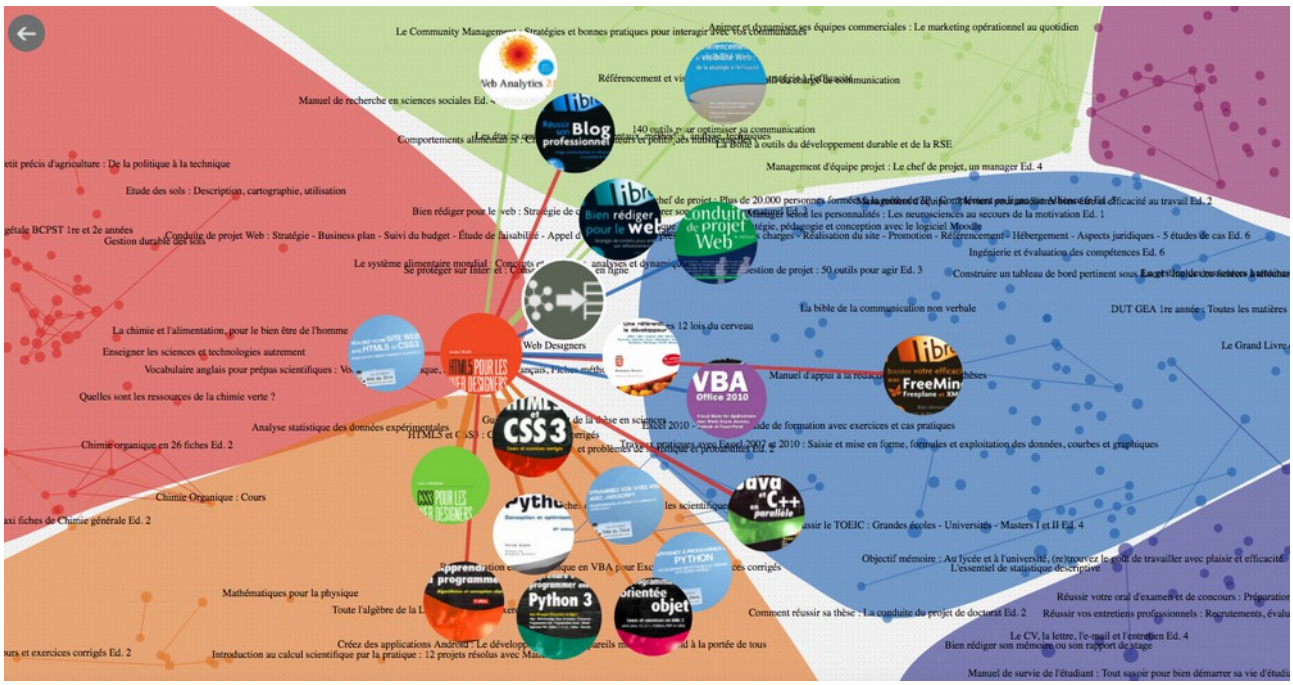
Lilliad environment, colors were chosen so as to match Lilliad librarians' wishes. One strong signal was captured when LilliadVis.io's DICE first appeared on screen. Indeed, the computed spatial organization suggests that exact sciences (in grey) behave like an island. If one had to build a library based on the current DICE plans, books in that area should be remotely shelved together. This strong signal, albeit an interesting information, may not last « forever » though. One may conjecture that exact sciences patrons will be attracted by the other colored areas. They may start crossing their habits boundaries and seamlessly explore areas they so far did not venture into.

The next narrative is called SAGE (Smart Academic Graph Explorer, Boussole in French). With SAGE we tried to answer the following question : Now that we can browse the DICE library is there something more we can learn from expert (smart) users, namely professors and/or librarians ? To answer the question, we narrowed the adjacency matrix down to the reading data of librarians and faculty members and computed the corresponding graph linking the institutions to which librarians and professors belong to. The following screen capture delivers a display of SAGE as it is currently viewed and used by Lilliad patrons.



The vertices represent institutions and the edges show the number of shared books between each pair of institutions. The higher the number of books the bigger the vertex's radius or the edge's thickness. In the first instance the graph highlights institutions with similar curriculum but it also allows users to drill down to each vertex and edge details. Clicking on the latter unveils the set of covers either corresponding to the institutions' curriculum or shared between both institutions (bottom panel). The experience complies with Shneiderman's mantra: "Overview first, zoom and filter, then details-on-demand" (1996). SAGE is a graphical and interactive overview in the form of a network of authority, from which the user is invited to drill in, by means of linearization (the covers are linearly presented). The SAGE also provides four filters for graph reduction (top panel): minimum number of shared books, institutions, publishers and keywords.

The third narrative is called **SCOPE** (Smart Communities of Patrons Explorer, Phare in French). Our objective was to go beyond the experts choices assembled in SAGE, to identify and map active reading communities. To do so we boosted the adjacency matrix to include any reader, be they patrons, librarians or faculty members, to discover how and what they read. The two following screen captures display these communities as computed by our algorithms and the network of books they each contributed to.



Multiple algorithms and data treatments are at play in the SCOPE. Each book needs to be assigned with a spatial position and a community. Spatialization and community finding based on large adjacency matrices can be very performance-consuming which is the reason why the matrix was filtered based on two rules. Firstly, a book needs to gather a given amount of readers (40 persons for LilliadVis.io), secondly the proportion of shared readers between two books needs to reach at least a given percentage (30% for LilliadVis.io). Both thresholds vary depending on the initial matrix size and configuration. The Fruchterman Reingold algorithm handles spatialization (Fruchterman, T. M., & Reingold, E. M., 1991) and the fast-greedy modularity optimization algorithm defines the communities (Clauset, A., Newman, M. E., & Moore, C., 2004) by looking at densities of edges among vertices. Several techniques are then used to provide users with an overview of the communities, as well as recommendations for each book. Each community is highlighted using colored convex-hulls tagged with the five most representative books' classifications. The user is invited to zoom into the SCOPE from which books and edges uncover by increasing the zoom's scale. Finally, the user is invited to click on any node to reveal the corresponding sub-graph composed of books with the biggest proportions of shared readers. At any time,

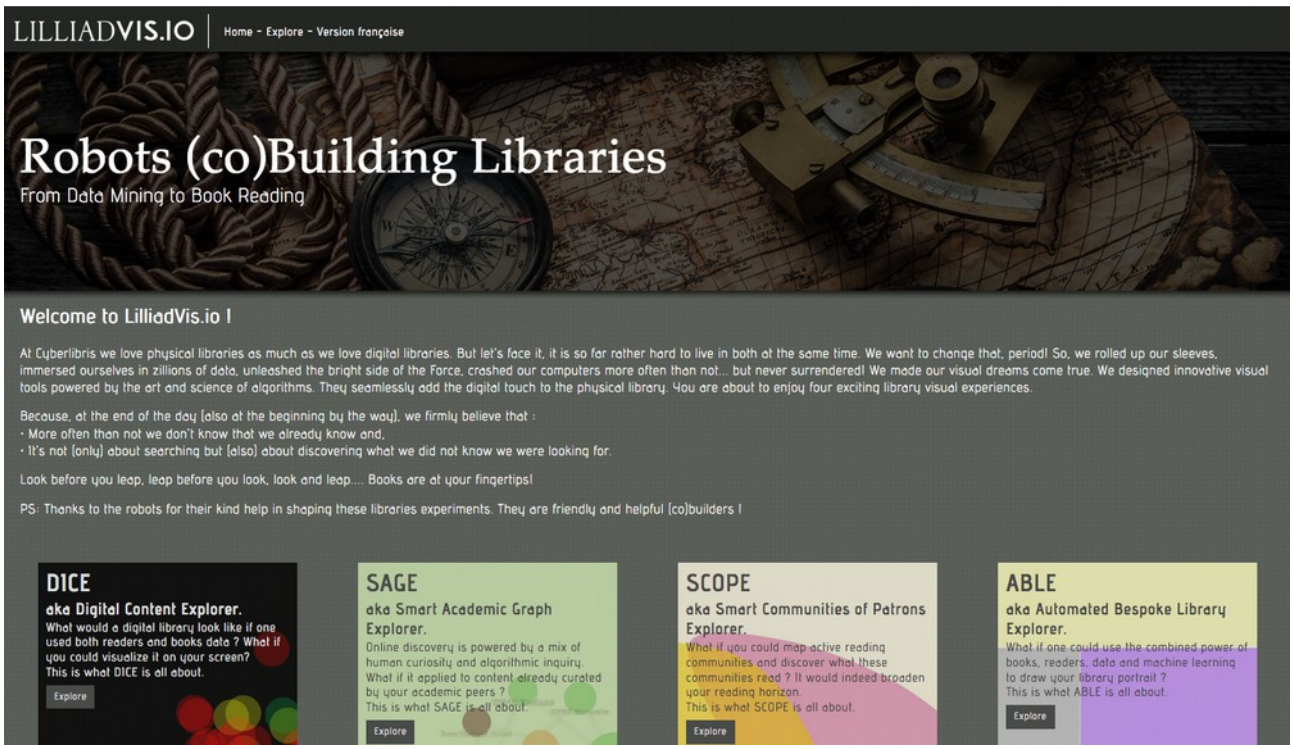
bunches of books can be harvested and stored in personal bookshelves (by clicking on the grey node).

Last but not least, the fourth narrative tackles a final task which any patron would most probably like to see completed. Now that the library can be browsed (DICE), that experts knowledge can be curated (SAGE), that the knowledge of communities can be put at work (SCOPE), is there a way for each patron to see what his/her bespoke library would look like? This task is fulfilled by **ABLE** (Automated Bespoke Library Explorer, Gouverne in French). Algorithms are “able” (hence the acronym) to extract from the adjacency matrix the library that should make sense to each patron. ABLE first displays each patron's current library:

It then delivers the new customized library whose content can be in part or in total added to the patron current library:

ABLE provides books' recommendations based on the user's bookshelves. The recommendations are therefore uniform among users, not leveraging personal data. All books appearing at least once in a shelf therefore need to be assigned a reading path. Indeed, should a user drop a book on a shelf, the corresponding recommendation should be made available in ABLE. The books' similarity calculation is done by means of manipulation of the adjacency matrix. Once computed, each cell of the symmetric matrix indicates the number of shared readers between the corresponding pair of books. For each book, the set of books shared with the most number of readers is extracted (top 10 in LilliadVis.io). Depending on the initial user's shelf, the corresponding selection is then combined and rendered in the form of a tree-map made of covers.

These four narratives are meshed in a full-fledged story that patrons can visualize as follows:



5. LilliadVis.io: Experiment and results

Most of Lilliad's collections are now digital, a trend observed in many science libraries. This digital move is a testimony of Lilliad's ambition to be a place to live, study and socialize while seamlessly leveraging physical and digital assets. Switching to e-books has an initial unfortunate consequence: Not much is known about patrons apart from dry COUNTER figures. Paradoxically, more is known about 1st year students who still have the ability to borrow print books. Lilliad's goal is to close this information gap while improving both patrons reading and discovering experiences. More specifically there is a simultaneous need to know how patrons are foraging contents, from print to digital, and to make e-books more visible.

The second co-author of this article met the first one a few years ago when she was working in Lille public library. The library had just subscribed to Bibliovox (www.bibliovox.com), an ebook platform tailored by Cyberlibris for public libraries. In 2014 a first research project stemmed from this collaboration, namely the

Bibliomobi experiment, led with the city of Lille, an innovation cluster dedicated to RFID and NFC technologies (CITC) and Lille public transportation network (Transpole). The goal was to give commuters the opportunity to download curated short stories during their travel by bus or subway. The first co-author was already fascinated by maps, by digital visualization and by the social aspects of reading. While developing Bibliomobi with the second co-author, he gave her a first insight into what were the very early steps of the DICE.

The second co-author joined the University of Lille after having spent five years at Lille public library. Lilliad learning center was not opened at the time. She was immediately struck by the fact that the university library did nothing to showcase its collections, be they printed or electronic. She also observed a large discrepancy between human and financial resources dedicated to the print materials compared to the ones dedicated to the electronic resources. It was necessary to reallocate human resources (staff and skills) to showcasing, cataloguing and valorizing electronic resources. It was indeed crucial to work on the library's engagement towards its patrons to ensure their needs were met. As a result, nine cataloguer profiles have been converted into liaison librarians.

The focus was put on informing and educating students about the electronic resources they could use during their first years. Indeed, electronic resources are heavily used by researchers whereas students information literacy skills remain poor. Most of the first years students are convinced that all what the library can provide them with sits on the open stacks - where they often feel lost and led to conclude that Google is easier and faster. A major effort was undertaken to make the electronic collections visible to the students, through bibliographies, exhibitions, stickers put on book covers linked to the electronic version of the book. A recommendation system linking examination questions (the most consulted of Lilliad databases!) to a curated list of print and electronic resources (mainly to handbooks available on ScholarVox) is also being developed. The

creativity bus did not stop there. The second co-author reminded the first one of the discussion they had about the DICE. She suggested an extension to it, namely to add print books to it, namely to add book loans data points to the e-books data points. The LilliadVis.io project was born.

A lot of work has been put into the building of a workable anonymized database ("data pudding") where the emphasis was put on usage instead of classification (order) only. This crucial step echoes Thomas Edison famous split, one percent inspiration, ninety nine percent perspiration[7]. For 2017 the primary usage metrics boil down to the following: 57 450 books were loaned, and 4500 ScholarVox ebooks were consulted, for a total of 350 000 pages read. The data pudding step is a step for which librarians are not that comfortable to start with. It is tantamount to let the fox (a disorderly band of usage data) into the poultry (an orderly band of neatly classified book metadata). As if one were trying to reconcile Dewey and Amazon. As David Weinberger (2007) nicely puts it : "Dewey created a single way to cluster books; Amazon finds as many ways as it can. Melvil Dewey took the design of the system upon himself; Amazon lets anyone create her own category, give it a fun name and publish it. Dewey prized neatness and order, bowing to the metric gods when he created a system based on multiples of ten; Amazon likes a friendly disorder, stuffing its pages with alternative ways of browsing and offbeat offers peculiar to each person's behavior."

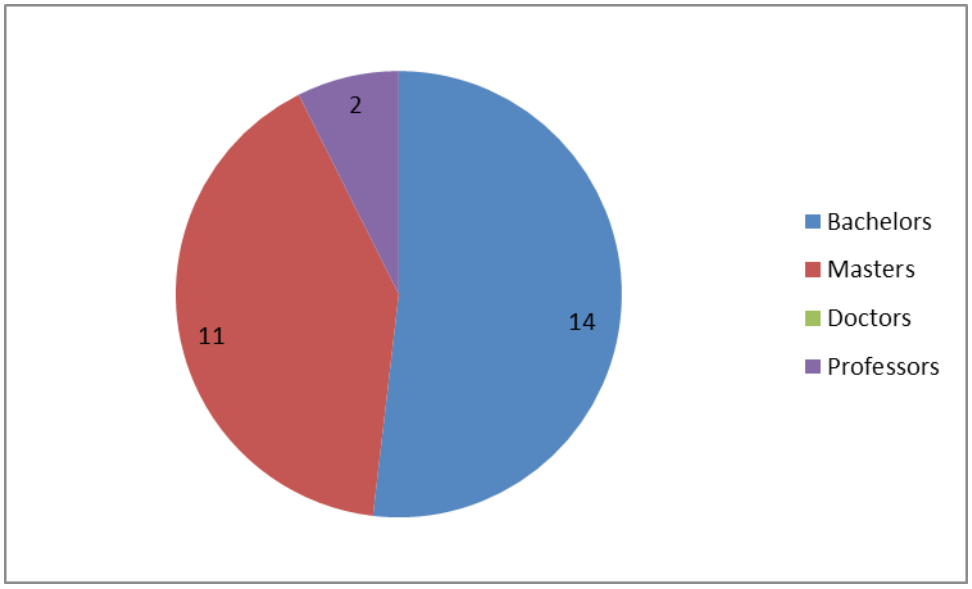
Once the data pudding was up and running and hours had been spent tinkering with data, algorithms and visualization techniques, the first and the third co-author came back to the second one. To her surprise they delivered not only the DICE, but also three other visual discovery tools, namely the above mentioned SAGE, SCOPE and ABLE. These four connected tools are what makes LilliadVis.io which is displayed on a 46 inches tactile screen located within Lilliad building. Time had come to experiment, to understand how an innovation vision was received by

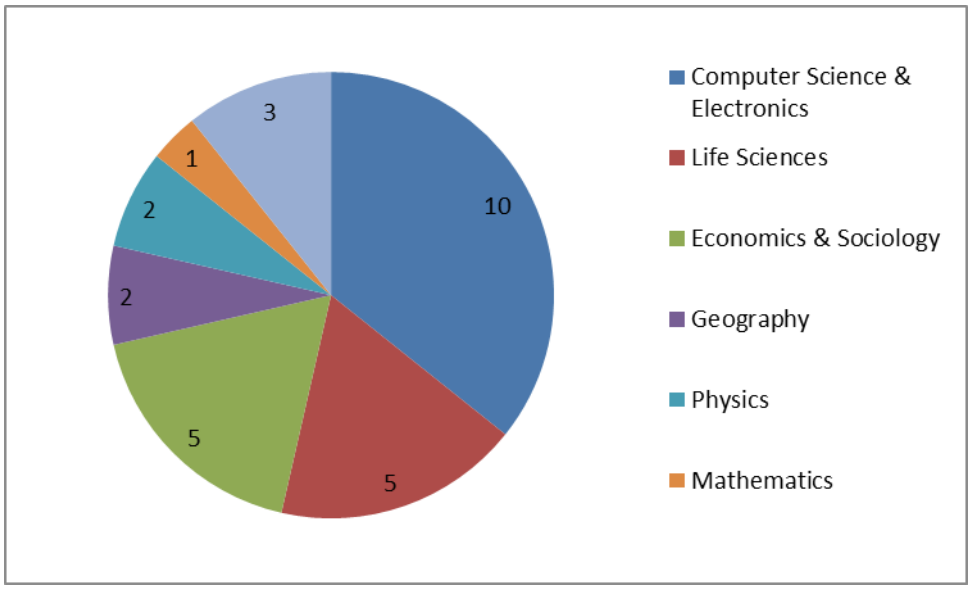
[7] One shall never underestimate the time and energy it takes to prepare the data.

patrons.

A very simple study, consisting in interviews, usability tests and comparison of search results was put in place. Its aim was to investigate how students, especially bachelor students, were apprehending the library print and electronic collections now they had access to LilliadVis.io. Can LilliadVis.io help students in their discovery of electronic resources? Can it stimulate ebooks reading? Is the celebrated “wisdom of crowds” truly useful in the book discovery process? Do usage data really foster a serendipitous discovery of books? Can LilliadVis.io, especially SAGE, be used as additional material to courses, in the spirit of course shelves ? And, last but not least, can LilliadVis.io be used at the reference desk for librarians to subtly teach students some information literacy principles ?

In order to ignite the evaluation process surveys were organized to collect information about students e-reading habits and to gather their feelings about LilliadVis.io. The sessions were presented as a mix between a presentation of the visualization device and a study. Students were attracted with chocolates and sweets. A raffle was put in place so that some of the students could win a snack at the library Café. Twenty-eight patrons were interviewed. Interviews lasted from 9 to 31 minutes. The following piecharts describe the surveyed student population:





Only three students already knew ScholarVox, and only one had already an account on the platform. 70 % percent were familiar with the use of a tactile screen. One librarian was conducting the survey. Another one was observing the way the patrons were using the interface. After having been given proper explanations about the study, students were asked to make a research on the different screens (DICE, SCOPE, and when time permitted, SAGE) and to comment it. In the following the results are detailed.

Most of the students were more impressed by the ScholarVox electronic library in than by the discovery tool itself. But, the truth is that they would not have paid attention to the electronic library content had they not been attracted by the discovery tool in the first place. This is an important observation as it shows that once a “state of the art” discovery interface is put in place, the associated electronic resource gains in traction. Students really enjoyed using the SCOPE interface, which most of them found easier to understand and to operate than the DICE. Still, one of the student mentioned she would never have thought of reading a book about Cistercian architecture, a book she discovered while looking via the DICE at a book on the geometry theories of Anaximander. She said she probably

would have remained in the mathematics shelves and would not have thought about searching for art history books. The reception of the SAGE interface was impressive. Students were very receptive to professors reading recommendations, even master students. Most of them were truly enthusiastic about professors shelves. Several students mentioned they would like their own professors to use such electronic course shelves and to provide links to them in the e-learning environment. A professor of physics decided he would link to ScholarVox in order to offer a more lively bibliography than the printed one. He appreciated the possibility to recommend a specific chapter even a diagram of a handbook, for instance.

Another professor (geography) confessed that he was only looking for books he already knew in the library catalogue, either because he had seen them cited in another book or in an article, or because he had been advised a new publication in his area. He was really seduced by the idea of seeing what other patrons had read or recommended in his discipline, and by the idea of being "surprised". Five students mentioned this also (they were 3rd and 5th year students). A student in Biology noticed that the visualization tool could help discovering in one go books that were physically displayed in several areas of the physical library.

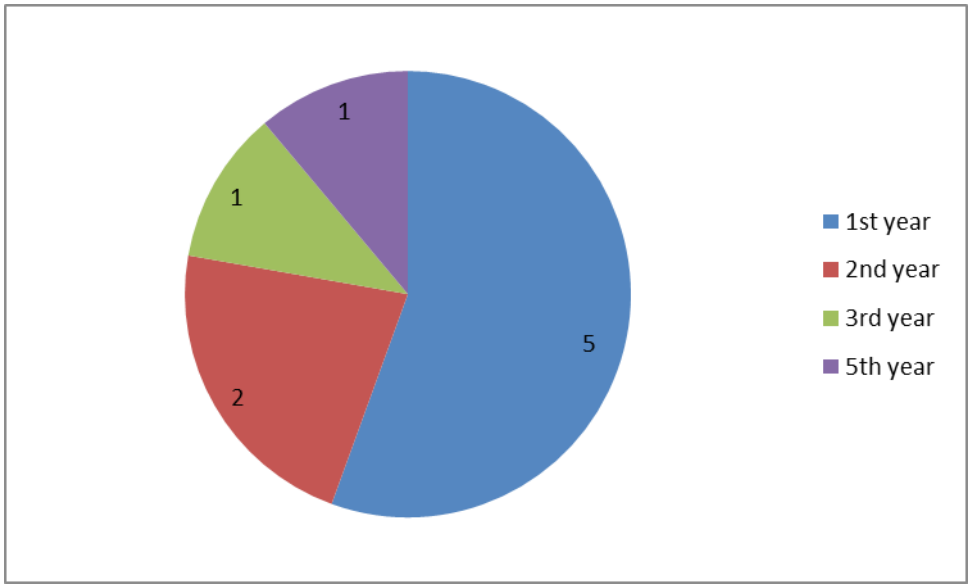
Half of the first years students mentioned the library as huge and impressive. They added that discovering the books with their cover and a few lines about their content was very helpful. Two students even said that they had the impression that there was a larger choice of books on LilliadVis.io than on the shelves and that, at the same time, it was easier to find the right book on the SCOPE than through browsing the shelves (a minority of them do use the Primo[8] discovery tool).

A significant number of first year students still prefer print books to e-books,

[8] Primo is a resource discovery solution marketed by ExLibris.

especially for visual comfort. – and, as odd as it may seem, students in computer sciences were particularly adamant on this point. Master students usually mentioned they were reading articles online, but books on paper. Several students said that they preferred not to read online because they were afraid of being distracted by facebook, snapchat, instagram, etc... Nevertheless, students enjoyed having the opportunity to e-read without any limit of neither time nor of the number of simultaneous readers of the same title, to have the opportunity to find and immediately read a books any time anywhere. One of the students mentioned the interest of linkages between ebooks. All of them made very interesting remarks on the usability, and on the searching tools. They would like to use the DICE with a keywords search. They sometimes found the titles difficult to read on the SCOPE, they did not like the keywords selection mode on SAGE (they would prefer a hierarchical structure to browse keywords and concepts). They also would like librarians to show this tool and the ebooks library to their professors!

The usability tests were led with nine students (see background piechart below), who did not know ScholarVox and who had never “played” with the surface table.



They were asked to use three of the four applications without any explanation nor help. Only three of them had already used ebooks. For each interface, there were three steps :

- Manipulation and observation (students on their own)
- Students were asked questions about the interface to check what the student had understood
- Tasks to be accomplished in order to assess the interface usability

The results of the usability tests are described hereafter. As far as the DICE is concerned, only two of them understood that circles and triangles were representing books, electronic and print. None of them understood why shapes were close or distant one from another, and even 2 of them thought this screen was a wallpaper... One imagined that the links between the books (they were told at this stage that the shapes were representing books) corresponded to books recommended for the same course. Only one student understood the whole thing because he took time to read the legend and to click, test and compare the results listed.

As far as the SCOPE is concerned, no one tried to zoom first on the colored areas (an observation also made during the interviews). Some of them did not understand that the texts appearing upon zooming were books titles : As they had seen disciplines or subjects on the colored areas, they believed that these titles were course or lecture chapters, for instance. Everybody understood that there is a disciplinary link between the books, and five out of nine succeeded in checking the references of a single title.

As far as the SAGE is concerned, students did not click on the circles corresponding to the different universities but tried to move them (which by the way is feasible). Then, six out of eight managed to find the universities recommending the same books as in Lille University, but only four out of eight were able to achieve a multiple keywords selection.

Some comparison tests related to discovery results were also run. The relevance of a list of search and discovery results was thus tested. A comparison was run to investigate how the DICE, the SCOPE and Primo were performing similar requests. For instance, a book on optical physics in the DICE was selected and a simple search was launched with the same words (“physique optique”) in Primo simple search box, with a search restriction on books. Three results appeared in Primo, two of them being research books. In the DICE, students could muse in a fifteen books list related to the “physique optique” book selected, which main subjects were relativity theory, general courses in physics, integral calculation, organic chemistry and analytical chemistry. These themes are adequate with the broad scientific culture a 1st year student should acquire or be interested in. Another selection on the SCOPE from a plant biology atlas draws to a visual memo in Biology, a visual memo in Earth sciences, a neurosciences handbook, a molecular biology handbook... All titles needed by 3rd year students preparing the competitive exam to become high school teacher in life and earth sciences. The request “plant biology” on Primo gave 423 books results, which is at the same time too many books to check at, and poorer than the list found on the Scope because all these 423 deal with the same topic.

Conclusion

It does after all take more than two to e-tango. When bits shake atoms and eventually start to compete with them, the librarian-patron couple, as skilled as it may be, is no longer enough on the traditional library ballroom floor. Another couple, database-algorithm, has to join the dance. But, letting it join the dance has serious implications. It modifies both the dance rules and the dance floor. In this paper, several angles are blended to figure out what this new dance and this new floor may turn out to be all about. The first angle view is very down to earth. It looks at things through a simple fidelity-convenience lens. This lens is powerful enough to identify tremendous synergies between the physical library and its

digital counterpart. How to harvest these synergies is however the hundred million dollar question. To answer it a historical angle is useful. Indeed, the observation of the architectural evolution of libraries is a vivid testimonial of their ability to handle the disorder embedded into the passage of time. Disorder is the right word: The digital economy in which we all try to strive is made of myriads of microscopic data, spreading at computer/smartphone speed and filling huge data tanks. A short detour through the sciences of complexity is worth the effort. The complexity angle teaches that the visible disorder (at the micro level) may hide an (yet) invisible (bottom up) order at the macro level. This fascinating observation helps reconcile order and disorder within the library. As a matter of fact, it does a lot more than that.

Indeed, in this paper, we show how we rolled up our sleeves. We guide the reader into the nitty-gritty of our hands-on approach, an angle that is predicated on a triplet. This triplet is made of an innovative library locus, Lilliad learning center, of a reputable digital library, ScholarVox, and of a state-of-the-art interactive visual platform, LilliadVis.io. This platform is articulated along four intertwined narratives backed by a powerful story. Thanks to these physical and digital assets in place, experiments were feasible. This paper details the first experiment to date that was put in place using a patron angle. The experiment aim was to survey 25 to 30 students, and to collect usability tests results from 6 to 10 students, during one month. Apart from the results analyzed in the previous section, some final observations can be proposed.

The first one is a caveat. It relates to the choice of words. In the initial interviews the patrons had difficulties understanding the meaning of certain French words used in LilliadVis.io. Most of the patrons were unfamiliar with the names that had been chosen for the DICE (Portolan, in French Portulan) and for the ABLE interface (steering, in French gouverne). Foreign students had difficulties with the French names of both SAGE and SCOPE (compass, “boussole” for the SAGE, and

lighthouse, “phare” for the SCOPE). This implied a preliminary work to ensure that used words are known, and indeed properly understood by patrons.

The second observation relates to an interesting comment made by American writer and poetess Johanna Drucker which we quote hereafter: “Images embody information through three different models, each of which has a different structural relation to the referent. They can work 1) through offering a visual analogy or morphological resemblance, 2) through providing a visual image of non-visible phenomena, or 3) by providing visual conventions to structure operations or procedures.” In a sense this more than welcome “mise en garde” is reminiscent of the Heisenberg uncertainty principle. What if data were produced by users complying with what is being expected from them? What if visualization schemes were made out of pre-established conception of their prospective use? On this issue we take a very pragmatic stance: Are we (patrons and librarians for that matter) happy about what we do and how we do it? An affirmative answer suffices to make our day.

A third observation is that, at its very core, LilliadVis.io is a tool that is predicated on human interaction. Most of the students, even when “bribed” with chocolates, were reluctant at first, slightly on the defensive. They were maybe dreading that the objective was to sell them something. All of them nevertheless left having understood that the aim was to support them.

A fourth observation is that most of the surveyed students are keen on the “wisdom of crowds”. But this tropism may be due to the fact we interviewed “relaxed” and volunteer students several weeks before the exams. They might not be as sensible to discovery and serendipity if and when urged by a deadline.

A fifth observation relates to time. The time it takes to develop and promote a platform like LilliadVis.io shall not be underestimated. The promotion itself is time

consuming. Up to now the tactile table is rarely used when no librarian is around to help and show. Moreover, in its current set-up, this table requires an interaction between no more than four students. This begs the question of the scenography around the table. This issue is on the agenda but, as of today, has not been addressed yet. Merging bits, brick and mortar ("e-tangoing") implies the obvious: The user experience cannot be limited to a tactile screen waiting to be touched. It takes architectural and interior design work to set up a space with a strong and attractive identity, a space that prompts the desire to visit it. In a sense what has been done and achieved at the macro Lilliad level has now to be undertaken at the micro LilliadVis.io level. This is for instance why it is useless to roadshow LilliadVis.io to freshmen when they visit the library at the beginning of the academic year. These students come from high schools where they were used to small and under-equipped libraries. They have to first adapt to their new environment. Moreover they are overwhelmed by the information amount they are provided with upon joining the university.

A sixth observation was made by the people from the information literacy department. They would enjoy using LilliadVis.io on tablets, in a "sniper mode". Indeed, when they walk through the library it would help them engage "hic et nunc" with students musing in the shelves.

A last observation is related to faculty members. The course-pack/bookshelf tradition (based on liaison librarian / professor teamwork) that is so popular in the USA hardly exists in France. Although the experiment was not targeted to faculty members, evidence mounted that students viewed LilliadVis.io as a tool enabling their professors to seamlessly build course shelves and link them to the local e-learning environment.

Bibliography

Barabási, Albert-László. Network Science, Cambridge University Press, 2017

Briys, Eric. D'or et d'airain: Penser, cliquer, agir, Les Belles Lettres, 2017

Clauset, Aaron, Newman, Mark EJ, and Moore, Cristopher. Finding community structure in very large networks. Physical Review E, vol. 70, no 6, 2004

Failla, Luigi: Du livre à la ville: La bibliothèque comme espace public. Métis Presses, 2017

Fruchterman, T. M., and Reingold, E. M. Graph drawing by force-directed placement. Software: Practice and experience, 21(11), 1129-1164, 1991

Johanna Drucker. Graphesis : Visual knowledge production and representation. Poetess Archive Journal. Vol. 2, n°1, pp. 1-50, 2010

Lerner, Fred. The Story of Libraries: From the Invention of Writing to the Computer Age. Continuum, 2009

Maney, Kevin. Trade-Off: Why Some Things Catch On, and Others Don't. Broadway Books, 2009

Manovich, Lev. The Language of the New Media, MIT Press, 2001

Melot, Michel. La sagesse du bibliothécaire, L'Oeil neuf éditions, 2004

Newman, M.E.J. Networks: An Introduction. Oxford University Press, 2017

Nock, R., Nielsen, F., and Briys, E. Non-linear book manifolds: Learning from associations the dynamic geometry of digital libraries. In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries, p 313-322, 2013

Picon, Antoine. Smart Cities: Théorie et critique d'un idéal auto-réalisateur. Editions B2, 2015

Shneiderman, Ben. The eyes have it: A task by data type taxonomy for information visualizations. Visual Languages, Proceedings, IEEE, 1996

Taleb, Nassim Nicholas. Antifragile: Things that Gain from Disorder. Penguin Books, 2012

Weinberger, David. Everything Is Miscellaneous, Holt Paperbacks, 2007

West, Geoffrey. Scale. Penguin Press, 2017