IT'S A REPOSITORY, IT'S A DEPOSITORY, IT'S AN ARCHIVE...: OPEN ACCESS, DIGITAL COLLECTIONS AND VALUE[1]

ES UN REPOSITORIO, ES UN DEPÓSITO, ES UN ARCHIVO...: OPEN ACCESS, COLECCIONES DIGITALES Y VALOR

Jean-Claude Guédon

Département de Littérature Comparée
Université de Montréal
B. P. 6128, succ. Centre-ville
Montréal QC H3C 3J7, Canadá
jean.claude.guedon@umontreal.ca

RESUMEN: *En el contexto del Acceso Abierto a la literatura científica y académica, los repositorios, tanto los institucionales como los disciplinares o temáticos, van a jugar un papel importante. No obstante, resulta difícil caracterizar la naturaleza de los repositorios debido a que cada categoría de personas relacionadas con ellos parecen tener diferente perspectiva. La cambiante interpretación que han teniendo los repositorios ha constituido una fuente de debilidad para la promoción de este instrumento. Aquí se sugiere que esta situación no se da sólo en el caso de los repositorios; al contrario, como muestran algunos importantes estudios provenientes de la escuela de pensamiento que aboga por la "Construcción social de la tecnología", todos los objetos sociotécnicos pasan por la misma fase. Esta escuela sugiere que los objetos técnicos tienen éxito cuando los grupos sociales relevantes interpretan el significado y la función de una tecnología particular. Examinando un conjunto de eventos acerca de los repositorios, en particular las luchas alrededor de la posibilidad de obligar a autoarchivar los artículos, es posible identificar un buen número de grupos sociales, así como examinar el modo en que pueden aliarse entre sí o manifestar la conflictiva falla que los separa. Este tipo de análisis debería ayudar a desarrollar estrategias que permitan desarrollar los repositorios.*

PALABRAS CLAVE: *Acceso abierto, repositorios, acceso, editores, editoriales, bibliotecas, investigadores, administradores.*

ABSTRACT: *In the context of Open Access to scientific and scholarly literature, repositories, both institutional and subject-based, have come to play an important role. However, the nature of repositories appears to be difficult to pin down as each category of people involved seems to have a different vision. The shifting interpretation of repositories has been a source of weakness in the promotion of these instruments. It is suggested here that this situation is not unique to repositories; on the contrary, all socio-technical objects go through such a phase if we are to follow some of the important studies coming from the "Social construction of technology" school of thought. This suggests that technical objects succeed when relevant social groups interpret the meaning and function of a particular technology. By examining a number of events around repositories, in particular struggles around the possibility of mandating deposits, it is possible to identify a number of relevant social groups, as well as examine how they can either ally with each other or are displaying conflictual fault-lines between them. Using this form of analysis should help develop strategies to develop repositories.*

KEY WORDS: *Open access, repositories, access, publishers, libraries, researchers, administrators, interpretive flexibility, relevant social groups, social construction of technology.*

1. FRAMING THE ISSUE OF REPOSITORIES

Repositories are very good examples of what historians and sociologists of technology call socio-technical devices. With such a term, they intend to keep the links between social networks and technical knowledge at the center of their analyses. A technology is neither purely technological nor the direct emanation of a social structure: it belongs to networks that are partly human, partly technical. As a society, human beings form a collective cyborg. Technology studies often label this conclusion as "social construction of technology" (SCOT).

Repositories ought to be examined in the light of the concepts and ideas encountered in the context of SCOT debates. Because this approach has been largely ignored by Open Access discussions, it may be useful to revisit it briefly, if only to bring to the fore some interesting theoretical tools better to understand the repository as a socio-technical object. It should be added that if repositories appear simple, this is largely an illusion.

One of the earliest examples of SCOT-inspired studies was a fascinating examination of the emergence of the modern bicycle (Pinch and Bijker, 1984). In several ways, the present situation of repositories, either institutional or subject-based, is reminiscent of the situation of the bicycle before it stabilized in the now familiar form that has remained the same for over a century: everyone wanted some sort of individual vehicle propelled by human muscle but neither the shape, nor the primary function of these vehicles was clear. Moreover, many partial solutions were in competition. For example, was it better to protect the rider by putting some sort of elastic layer between the wheel and the road, or was it better to build some kind of suspension system in the very frame of the machine?

Historians and sociologists of technology have demonstrated that the bicycle was designed and shaped by interactions between various stakeholders. More precisely, the kind of analysis first put forth by Bijker and Pinch, rested on a number of fundamental ideas that included:

• The notion of "interpretive flexibility". This means that technological design is a process open to various possibilities according to the social context where it emerges;

• The idea of "relevant social groups" (RSG). Nowadays, the buzzword would be "stakeholders", except that the earlier form of expression was clearer as to its essentially collective nature;

• The idea of closure or stabilization: the process of technological design does not go on forever. At some point, it must stop according to some mechanism that will involve some social dimension such as consensus, a vote, etc.[2].

These basic insights have been either criticized or complemented in a variety of ways. For example, more importance has been attributed to the wider social context. Others have striven to clarify the nature of the interactions between the RSG's, in particular by showing greater sensitivity to power relationships and their consequences. Attempts have also been made to move beyond what is perceived to be the agency approach of Bijker and introduce a more structural approach which, of course, dovetails nicely with concern for the "wider context". However, the three basic ideas above remain sufficiently important to be singled out and used to approach the issue of repositories.

Nowadays, the system of scientific and scholarly publishing also involves a wide variety of stakeholders and the debates about Open Access have provided the means by which they have identified themselves. However, knowing which RSG's are involved in the shaping of repositories does not mean that we have a clear perception of what is really at stake from the perspective of each one of these groups. This situation is well reflected in the vocabulary used: everyone seems to know what a repository/depository/archive is and is meant to do; meanwhile the vocabulary has vacillated between several terms and it is only gradually that the word "repository" is becoming dominant in English. In other languages, the issue, paradoxically, emerges more clearly and more quickly, probably because the need to translate an English-language terminology forces the translator to make conscious choices. If we recast this reasoning in SCOT terms, it can be said that, in languages other than English, the translators form a RSG of their own, albeit a minor one, that may or may not overlap with other RSG's.

The idea of power has been introduced more recently within SCOT analyses and it adds important insights into

the social shaping of technologies. RSG's harbor hierarchies of their own, and even oligarchies. Power struggles certainly account for their history and evolution. Also, RSG's display various degrees of social concentration or dispersion. For example, consumers will tend to be quite dispersed and, because of this, will find it relatively more difficult to mount coordinated actions than the tight executive of a well-run company. The more dispersed a RSG is, the more likely its power will be diffuse and more difficult to marshall.

RSG's also compete with each other and, to do so, they rely on access to various resources. Economic resources are always crucial; political resources can be important, especially when policy issues emerge. Less obvious perhaps are what some analysts call "cultural resources", also referred to as "discursive legacies". More simply put, these resources deal with the extent to which a particular technological artifact may resonate with a given cultural theme and how this theme is perceived. An example of this situation is provided by Sport Utility Vehicles: nowadays, with economic worries and a great deal of talk about global warming, SUV's are often described as going against the grain of environmental concerns. Not so long ago, particularly in advertisements, they were portrayed as symbols of power and adventure. Publicity constantly promotes cultural meanings of technical objects to stimulate sales and in so doing reveals the power of cultural resources.

Cultural resources are found at all scales of society: in particular, they can work at macro-social levels as in the SUV example above, but they can also be linked to the "culture" of a particular institution: for example, setting up a computer network will lead to concerns, in particular security concerns, that will vary greatly from one institution (e.g. a Church) to another (e.g. a police precinct or a bank).

Repositories, like any other socio-technical entity, find themselves under the stewardship of various RSG's. Understanding what these groups are, how they relate, what their implicit or explicit agendas are, and what resources they have at their disposal, will clarify the situation greatly. If, occasionally, repositories appear as fuzzy and incoherent objects, this is because they are examined from a variety of viewpoints that are not clearly brought to the fore, and it is also because they are not yet stabilized. Although various stakeholders or RSG's may have a very precise and clear

idea of what they would like repositories to be, the end result may remain in the balance for a period of time that can extend over several years. In short, repositories are technical artifacts and applying some theories developed to analyze such objects should be helpful. In particular, for those committed to Open Access, such an approach would provide a more systematic turn to strategic planning.

2. How are repositories presently portrayed in the literature?

Two very recent texts provide an excellent entry point to this question. The first one, by Charles Bailey, is designed to introduce "the reader" to "key aspects of institutional repositories" (Bailey Jr., 2008). Unsurprisingly, Bailey begins by a quest for a definition, but leaves us with three distinct possibilities: for Clifford Lynch, it is a "set of services that a university offers"; for Mark Ware, it is "a web-based database… of scholarly materials which is institutionally defined", in contrast to a subject-based repository. Finally, for Raym Crow, it is made up of "digital collections capturing and preserving the intellectual output of a single or multi-university". All three definitions place the preservation function front and center. Lynch and Ware also underscore organization and access issues related to repositories but Lynch places a greater accent on stewardship as a generic goal while Ware adds concerns for interoperability. Crow's document presents a different perspective by emphasizing the "reforming [of] the system of scholarly communication…" and by mentioning that an institutional repository (IR) can be useful for the ways in which a university presents itself to the world.

In themselves, these three definitions are good examples of what SCOT defenders would call "interpretive flexibility". They also show how strongly situated forms of discourse can be. Cliff Lynch, for example, is head of the Coalition for Networked Information (CNI) and, as such, has played an important role in trying to chart a path for the professional evolution of librarianship in an increasingly digital environment. His emphasis on services, organization, and access stand at the heart of the library profession, independently of print or digitization. Mark Ware works at the interface between publishers (in this case the Association of Learned and Professional Society Publishers (ALPSP)

and the Publishers Association) and the Joint Information Systems Committee (JISC) anchored within the British system of higher education. His reference to the neutral term "data-base", and his insistence on interoperability certainly fit well with his need to chart a ground common to two worlds, and identify some division of labor acceptable to both sides.

Raym Crow framed his definition in a position paper designed for an OA advocacy group: the Scholarly Publishing and Academic Resources Coalition (SPARC). His talk of reform and his attempt to gain the support of university managers show strong strategic motivation. Interestingly, both CNI and SPARC are offshoots of the Association of Research Libraries (ARL), but their concerns are quite different: the former studies the "transformative promise of networked information technology"[3], while the latter deals with the politics of scientific and scholarly electronic publishing. It was very engaged in the recent passage of the law mandating the archiving of all the researchers' publications supported by the National Institutes of Health in the USA[4].

Charles Bailey's article provides more examples of "interpretive flexibility": for example, the reasons supporting the development of a local IR range from increasing the visibility and impact of the institution's scholarship, especially if OA is attached to the IR, to providing unified access to the institution's scholarship. The former point largely reflects the concerns of the researchers; the second point corresponds more to the concerns of university managers[5].

In another article, Alma Swan and Leslie Carr discuss "Institutions, their Repositories and the Web"[6]. They too underscore the role repositories can play to "maximize the visibility of the institution's research outputs". While they equate the issue of visibility with providing Open Access, they also stress the importance of "a mandatory policy on the use of the repository for collecting outputs". In defending IR's, Swan and Carr have chosen the point of view of the institution's managers: thanks to their survey of European institutions, the authors can state that "the primary reason for establishing a digital repository is to increase the visibility of the institution's research output by making it Open Access". They also add that "... a repository is a tool that enables senior management in research institutions to collate and assess research, to market their

institution, to facilitate new forms of scholarship and to enable the tools that will produce new knowledge". A concrete example is then briefly analyzed in which it is shown that three rankings of the University of Southampton's differ drastically. While the *World University Rankings* of the *Times Higher Education Supplement* and the *Shanghai Jiao Tong University Academic Ranking of World Universities* place Southampton between the 100th and the 200th places, the *G-Factor International University Ranking* places the same university in the 25th position. The difference is simply that the last ranking system relies on university-to-university web links. If Southampton shows up well there, it is because it has a very rich and useful OA IR.

If we look now at IR's from the perspective of publishers, we come across very different perspectives, including negative or even panicked reactions[7]. At best, they emphasize the role that IR's can play for gray literature and for data, perhaps for long-term preservations[8], but the basic reaction remains very cautious and even hostile. In short, in the literature, IR's are always described from the perspective of some RSG.

3. UNDERSTANDING THE "INTERPRETIVE FLEXIBILITY" ATTACHED TO INSTITUTIONAL REPOSITORIES

a) Identifying the RSG's

The definition and design of IR's is contested, as the SCOT tells us it should be. All socio-technical entities go through an uncertain phase when their designs have yet to stabilize. During that phase, RSG's and their respective visions become visible and they become all the more visible in moments of divergence and even opposition. In such moment, success or failure will depend on the amount of resources available to RSG's, including their ability to create alliances with other RSG's.

The RSG's affecting the shaping of IR's are actually quite diverse:

1. Researchers and advanced students.
2. Librarians, library consortia and associations.
3. University and research center administrators.

4. Large commercial publishers, some society publishers and publishing professional associations.
5. Most society publishers and university presses.
6. Public and private research granting agencies.
7. Education funding agencies (ministries or governmental departments).
8. Foundations.
9. Law making bodies.
10. Public interest advocacy.

To these RSG's could be added infrastructural elements that can heavily affect the shaping of repositories. Copyright laws certainly weigh in on the deposit and access rules of a repository, as do the ready availability (or not) of licensing schemes such as Creative Commons. Some possible RSG's have been left aside, for example computer services and programmers.

This list would of very limited usefulness, were it not for the possibility of creating some hierarchy within it so as to identify the most important groups. Let us remember that the ability of a group to influence the design of a socio-technical object rests on a number of resources, but that the status of a group within an institution can severely constrain the same group. For example, librarians will be considered as service providers within a university, and will generally find themselves at some disadvantage in comparison with faculty.

Overall, repositories are often under the responsibility of libraries, but they can also be under the responsibility of computer services. Sometimes, repositories regroup a number of institutions and can even be national in scope. In Europe, with the DRIVER[9] project, supra-national coordination of repositories begins to appear. This suggests that repositories exist within variable scales involving different kinds of RSG's. In short, repositories cannot be taken as a homogeneous set and their social shaping varies greatly according to the size, status and type of their RSG's.

b) Summit RSG's

Publishers generally are the owners of the materials that ought to go into the repositories. As can be expected, they contemplate the rise of repositories with some anxieties; at the same time, they have been careful, for the most part, to avoid direct confrontation with two important

interlocutors: researchers and librarians. As a result, they have carried out a dual strategy at two different levels: on the one hand, they have aimed at confusing the issues around repositories; on the other, they resist any quest for mandatory deposits that tends to be national in scope or involve a high-level national institution. To refer back to the early discussion on the SCOT approach to technology, publishers know that the dispersion of potential opponents can turn to their collective advantage. By making some issues confusing, they manage to generate some degree of bewilderment and thus stand in the way of concerted action. For example, when publishers allow self-archiving, but each of them places a different set of constraints on the process, they will discourage many researchers from going down a route whose meaning is not entirely clear to them[10]. Maintaining control over the "reference" version of any given article (the one that can be cited) minimizes concessions while offering the public the facade of a sensible attitude. Publishers generally know well the rhetorical injunction to ensure *captatio benevolentiae*. In short, publishers have some very powerful tools to divide and conquer:

1. With regard to researchers, publishers have the privilege of intervening in the selection of editors-in-chief if they own the journal. It must be remembered that promotion to gatekeeper status is very prestigious for any researcher and comes with a good deal of implicit power.

2. With regard to librarians, the digital context has led to licensing, rather than selling, documents. This transformation has also led to a concentration of forces where nowadays groups of libraries working within a consortium negotiate "Big deals". As a result, publishers often work with consortia leaders and generally manage to assign them to a procurement role with relatively few possibilities for free exchange of information among consortia[11]. This may explain why some consortia have not supported repositories or Open Access journals very vigorously[12]. In fact, with consortia and their peculiar relationship with publishers, the library voice appears to have grown more hesitant, less assured, less unified. For some consortial leaders, maintaining good relations with vendors is more important than solidarity with other consortia. Of course, behaving this way can have its rewards too,

such as invitations to conference events sponsored by publishers[13].

In short, the richest and most powerful among the publishers have maneuvered very efficiently to divide researchers and librarians and coopt a fraction of each group. They have also managed to cloud the issues sufficiently to make researchers indecisive and librarians less militant. In particular, library associations encounter difficulties in keeping the consortial groups and the advocacy groups on the same page. The International Coalition of Library Consortia (ICOLC) went through difficult discussions before adopting recommendations in 1998 in favor of sharing the results of individual negotiations. Publishers, of course, voiced artful opposition[14]. When the ICOLC principles were revised in 2004, not all consortia signed.

The publishers oppose Open Access but sometimes settle for slowing down its progress sufficiently to discourage its supporters. However, OA moves forward, as the recent law passed in the United States testifies: researchers financed by the National Institutes of Health must deposit their papers in PubMed Central within twelve months[15]. The composition of the alliance needed to obtain such a result will tell a great deal about the relative strengths of the RSG's that have been shaping IR's at the highest level.

The recent political success of Open Access in the United States relies on an institution that is part of a very specific group of RSG's: the funding agencies. Some are private, such as the Wellcome Trust in the United Kingdom, while a majority of them in a majority of countries are public. In the United States, medical research is supported both by private (e. g. Howard Hughes Medical Institute (HHMI)) and public (National Institutes of Health) organizations. In many countries, such funding agencies have taken the lead in favor of Open Access. The reasons are clearly laid out by the Wellcome Trust:

> The Wellcome Trust has a fundamental interest in ensuring that the availability and accessibility of this material is not adversely affected by the copyright, marketing and distribution strategies used by publishers (whether commercial, not-for-profit or academic)[16].

Funding agencies such as the Wellcome Trust want to improve the system of scientific communication because a better system will lead to the better and faster scientific progress. In the particular case of the Wellcome Trust, we are talking about health issues so that the benefits are very easy to comprehend as well as concrete. The mixture of fundamental research needed for progress and the familiarity with painful health situations around all of us makes the position of the private charity easy to grasp.

Similar arguments are relevant for public agencies funding medical research, but they are couched differently. While NIH and its sister institutions elsewhere all share the simple, yet lofty, goals of the Wellcome Trust, they must also do so in such a way as to convince lawmakers. For this reason, value for money has to be shown, as well as concern for the fairness of the decisions taken and the public (voter) good. Symmetrically, publishers have pushed back, particularly in the United States, with arguments having to do with the fundamental importance of market mechanisms, and warnings about misdirected governmental interventions. Large publishers and some powerful scientific societies have lobbied members of Congress to try blocking any legislation favoring the mandating of article deposits in suitable repositories, in particular in PubMed Central[17].

The difficulty of the battle around the NIH deposit mandate as well as its intensity can be evaluated by the length of time to pass the legislation[18]. The long and tedious path through committees, the false starts, lost votes, vetoed laws and, sometimes, the less than felicitous solutions proposed all bear witness to the incredible work that was accomplished. But who did it? What were the RSG's involved?

The answer to this question is relatively simple but it does include a surprising element: an advocacy group within a library professional organization, the Scholarly Publishing and Academic Resources Coalition (SPARC), most ably led by Heather Joseph, was constantly supported by a motley crew of researchers that had been involved in all of these battles ever since 2000-1. The people behind PloS, some Nobel Prize winners such as Harold Varmus and a number of other first-rank scientists were involved. The majority of researchers, however, heard the sound and fury of this epic battle only vaguely, if at all. Administrators, for the most part, remained indifferent as well. But another, somewhat unexpected, RSG came to the fore and began to play an

extremely important role. Made up of ordinary people with ordinary concerns, the Alliance for Tax Payer Access began to weigh in[19]. This organization is self-described as follows: "The Alliance for Taxpayer Access is a coalition of patient groups, physicians, researchers, educational institutions, publishers and health promotion organizations that support barrier-free access to taxpayer-funded research".

RSG's, let us remember, also draw upon cultural resources to influence the shaping of socio-technical objects. In this case the Alliance occupied an enviable position in the sensitive issue of access to health. In the United States, with the largely privatized system of health insurance and the very high costs of both medicine and drugs, this issue has been very much at the forefront of political debates for nearly two decades and it generates a great deal of anxiety among voters. Although the Alliance was not speaking directly about institutional repositories, it was clearly addressing issues of access to medical knowledge for everyone. The way to obtain this result was to get a law passed that would mandate the deposit of publicly-supported research results into a suitable repository or a set of repositories. For their part, library advocacy groups had access to some financial and political resources, but, with the Alliance, they had a great deal more of both and it is with this particular combination that the lobbying efforts of SPARC got the added traction it needed.

The United States omnibus law that contained the language needed to mandate the deposit of articles in a repository within twelve months of their publication was ultimately obtained through the convergence of librarians, a well-placed and highly-motivated minority of researchers, an advocacy branch of the Association of Research Libraries and an institution allocating billions of dollars each year for medical research (NIH).

A number of lessons can be drawn from this example:

1. The battle waged by large commercial publishers to neutralize the institutional repositories was lost because of the coming together of a number of RSG's, some of whom were not obvious.

2. While publishers succeeded in keeping many librarians and most researchers out of the push in favor of Open Access, they could not prevent some powerful

minorities in both of these groups to continue working actively for Open Access and even join forces.

3. The emergence of funding agencies as strong supporters of Open Access was a crucial moment, and a very dangerous one for the publishers. They had to stop very powerful institutions, with deep pockets, that have a constant, positive, relationship with researchers. The mixture of private and public institutions in that group of funding agencies or charities also prevented publishers from designing a single, simple strategy. In particular, the presence of public institutions meant that the battle had to shift level and instead of being waged at the lower level of universities or below, it had to be fought at the highest levels of government.

4. The Alliance for Taxpayers Access suddenly provided librarians, militant researchers and the funding agency with a very powerful cultural resource. From that point on, the Open Access informal coalition had all the resources it needed: political, cultural and even financial. By contrast, publishers had only their financial resources. They could use those resources to try buying some political resources, but these are not always of the best quality. Finally, publishers had almost no viable cultural resource.

This episode also shows that, ultimately, the battle ground shifted in scale and ended up reaching the highest level of government. This shift should not come as a surprise as it had already been observed in Britain with the debates surrounding the Select Committee on Science and Technology in 2004. In this case, the attempt to move the government itself failed despite a favorable report from the Select Committee[20]. However, when the UK Government directed the question of Open Access to the Research Councils, it gave the Councils the opportunity to choose what they preferred. As research funding agencies, they obviously were sensitive to the arguments that had already moved the Wellcome Trust and nudged NIH in the United States. Not too surprisingly by June 2006, Research Councils UK had drafted and made public a document supporting access to publicly supported research. The press release of 2006 contained an interesting sentence worth quoting in full: "Ideas and knowledge derived from publicly-funded research must be made available and accessible for pub-

lic use, interrogation and scrutiny, as widely, rapidly and effectively as practicable"[21]. Although the British scene witnessed a to and fro trajectory from Select Committee to Government back to Councils, it is clear that the parameters that have been involved in the United States debates were also at work in Britain: the granting agencies in these two countries (as in France with INSERM, and in Germany with the Max-Planck Institute) stand for the widest access to the research they finance, especially when that financing is public. Researchers and librarians played their role too in Britain. In the UK, however, there was no Alliance for Taxpayers Access and this may have been the missing element that would have allowed the government to accept the Select Committee's recommendations.

In summary, the large commercial publishers, confronted with the growing push of funding agencies in favor of Open Access, carried the battle to the highest legislative and executive levels only to find themselves either thrown back into the hands of the Councils, as in Britain, or, ultimately defeated in Congress in the United States. The lessons learned were crucial and must be carefully meditated for similar battles in other countries. However, this analysis deals with only one layer of the whole repository scene. Other, more modestly scaled, debates have also taken place and are being waged at this very moment. To those we will now turn.

c) Inside Research Institutions

For most authors, depositing their research results in a repository does not fulfill their publishing needs: both validation and branding are lacking. Peer review and the prestige of the journal harboring a particular article are perhaps even more important to the researcher than knowing that the journal is going to be available in a number of libraries around the world. Various bureaucratic forms from universities, or granting agencies, will ask researchers about articles appearing in peer-reviewed journals, not about the runs (and costs) of these journals. In publishing, researchers often respond more to institutional pressures ("publish or perish"), than to a deep-seated desire to produce and disseminate new knowledge. For this reason, the ritual passage through the journal phase is essential to researchers because many evaluative procedures are linked to their presence: tenure and promotion committees pay

great attention to the journal variable, as do juries helping to allocate grants. Repositories come only later, if at all.

Except where mandates have been instituted at some institutional level (a department, a faculty or a whole university), the lack of obvious links between the act of depositing one's papers in the local institutional repository and the evaluative procedures of the institution makes the gesture appear largely irrelevant. It is all the more irrelevant that depositories are generally not on the radar screen of researchers when they look for information, except if their libraries are poorly stocked. As for the Open Access advantage[22], if one compares it with something like the "impact factor" of journals, it simply is not treated as a valid or useful argument in most administrative evaluation procedures. In short, many researchers do not pay attention to institutional repositories, do not know what they can really do for them, and do not feel they are losing anything in terms of their careers when they benignly neglect them. One of the fundamental paradoxes of the institutional repositories is that, until now, researchers have not placed them at the center of their preoccupations, be they career promotion or knowledge acquisition (with the exception, once again, of researchers working in developing countries: they know the value of Open Access because they are not subsidized readers like their richer colleagues).

Inside the universities and the research laboratories, a movement largely led by librarians (again) has driven the growth of institutional repositories in hundreds of institutions[23]. Here, the impetus is on collecting the production of a local institution and displaying it to the world. The alliances that have allowed this situation to develop are very specific to each institution but almost always involve the library and the administration. The library, guided by recommendations coming from professional library associations worldwide such as the International Federation of Library Associations and Institutions (IFLA)[24] have developed guidelines, best practices, institutional strategies and a number of other documents to grow repositories. In some universities, much energy has been and is being spent on filling these repositories, often with unconvincing results.

Let us look at some statistics. Repository holdings vary from a few dozen records to several thousand. In the "OpenDoar" directory of repositories, a search engine al-

lows us to search by category of documents and offers fourteen possibilities. Dissertation sites (526) are the most frequent and unpublished works (488) follow. Neither are very convincing categories from the viewpoint of researchers looking for authoritative information. Then come publications without any definition. They are found in 345 repositories but we do not know what they are. Preprints occur in 62 cases, and postprints in 256 cases. Even assuming that the two groups do not overlap and deal only with refereed articles, this would mean that less than a third of repositories hold peer-reviewed literature. Other categories are not very helpful: they include software, learning objects, references, multimedia and even patents. In short, while the Open Access movement often reiterates the need to collect the refereed journal literature of the world, the repositories seem to include about anything that can be gathered inside a university. Obviously, this state of affairs will not be terribly attractive to the researcher looking for reliable and authoritative literature. The noise level is much too high. Neither does this situation incite researchers to deposit their articles in such a repository.

What are the SRG's that have shaped repositories in this very peculiar manner in a great many institutions? The answer is quite obvious: librarians. They probably decided to build repositories as part of their general positive feelings for Open Access[25]. However, in order to proceed, they checked with their administrator and probably were asked for a justification, especially if they requested help to build the repository. It is not difficult to imagine the form of the argument taken: our repository will expose the intellectual output of our institutions to the whole world and it will be much easier to monitor what people are producing within the university. It will also, with time, build the intellectual memory of this institution. With such a starting point, repositories came to collect all the local intellectual output, from courses to articles, from student theses to research published in the most prestigious journals. At the same time, because the librarian did not have the power to bring about a deposit mandate, and because the administrators were (and always are) worried about creating tensions within the institution, most especially by requiring that faculty do something, the repository limps along with deposit rates that rarely exceed 20%[26]. And publishers are delighted because institutional repositories, and with it Open Access, have been marginalized.

It is also worthy of note that researchers are largely absent from this effort with few exceptions: occasionally, a librarian may organize a colloquium to advertise the repository to faculty. Typically, and we write here from personal experience, only a few members of the faculty come. Some e-mail campaigns may go on for a while, including visits to various departments, again with mitigated success.

The real needs of the researchers appear to be misconstrued or totally neglected. As readers, they need to find good literature fast, and without too much noise. This is the strength of traditional bibliographies and specialized journals in their particular fields. So far, the tools linked with repositories, while giving access to a great deal of literature, do so in a rather unreliable way[27].

Symmetrically, we have seen that researchers largely ignore or neglect repositories. For tenure and promotion, or for grant proposals, the repositories do not offer anything yet judged valid by administrators and juries. At the level of the whole institution, the metrics used to assess the research of research institutions have created a situation where branding by journals and impact factors dominate the scene: this is how "excellence" is "measured" nowadays[28]. Such methods leave little room for Open Access even though IR's hold the promise of improved impact. However, most researchers are but dimly aware of this so-called "OA advantage" and administrators appear even more oblivious to it.

This leads to a fairly absurd paradox: institutional repositories have been designed for researchers to help them both in their work and their career. However, as the rest of the institution does not integrate IR's in its working practices, researchers do not see their importance and neglect them. On the other hand, librarians design their repositories to capture the research output of the university, but often end up designing a device that gathers all forms of intellectual output, not just research. Moreover, researchers are not consulted very efficiently or very well.

Institutions appear to organize their institutional repositories and set up mandates only when they become part of national research assessment exercise. Such a situation can be observed in Australia with the Asher program[29] that lies at the center of the proposed Research Quality framework procedures. In short, it makes the evaluation

of the research in a given university depend on what is actually collected within the local institution[30]. The Asher program provides help to universities so that they can set up their repository.

This survey of the local forces shaping repositories reveals some striking elements. If we think about the way in which researchers are supposed to be at the center of repositories and their *raison d'être,* the result appears contradictory: researchers are absent both in terms of preoccupations and ability to influence the design of these new institutional devices. Administrators appear more reactive than pro-active: they tend to go along with the librarian's desire to build a depository, but remain skeptical. Meanwhile, the real issues, such as integrating the new device into important evaluative practices are not touched. Students, in particular doctoral students, can also be added to the list of important, yet neglected, partners. As a result, they too practice neglect, not because they are mean spirited, but because the issues do not appear clearly to them.

When progress is noted, it generally signals that more central organizations intervene, as the Australian example shows. The hierarchy of institutions and the action of power centers reappear in this context. It is clear that the financing arm of higher education can be determinant. In some ways, ministries of higher education and granting agencies, although not identical, may find themselves adopting congruent attitudes. Globally, the intervention of ministries, of funding agencies and similarly positioned national institutions points to the fact that policy changes seem to require movement from the very top authorities of a country or administrative region. At the same time, local actors within institutions must also act, and that is where the minorities of researchers advocating Open Access, the militant segments in the library community and the odd administrator can begin to agitate to make the institutional "middleware" (i.e. the cadres of higher education and research) move in useful directions.

Having now completed a quick tour of the main RSG's involved in defining and designing IR's, we can now turn to the last part of this paper which will deal with the possible ways to bring about a positive closure to this long debate. This is the point where analysis and strategy begin to converge in earnest.

4. THE PATH TO CLOSURE: DESIGNING THE RIGHT INSTITUTIONAL REPOSITORIES

Anyone surveying the repository landscape for the first time would be immediately struck by its diversity. And anyone doubting the validity of the "interpretive flexibility" concept would have to explain why such a wide variety of designs exist. Many types of repositories stand beside institutional repositories: national repositories (such as "Hyper-articles en ligne" or HAL in France)[31], subject-based repositories (such as PubMed Central at NIH or ArXiv at Cornell University)[32] or repository consortia (such as White Rose for the Universities of York, Leeds and Sheffield in the UK or, on a full national scale, DAREnet in Holland)[33]. Contents also vary enormously from one repository to another, as noted earlier, including non-research materials, and materials that are not in Open Access. Some digital collections have been designed in part to create value, as in the case of the "Cream of Science"[34] project, a national repository located within DAREnet in Holland. The scope of these variations may signal the fact that the stabilization of the repository format is still some distance away.

More cohesive is the movement of most repositories as they use a common metadata standard, OAI-PMH[35], and this standard is being completed by the Object Reuse and Exchange (ORE) specifications[36]. These are important tools that offer the promise of well ordered repositories with the possibility of searching through specified subsets of documents such as peer-reviewed articles or dissertations. Protocols such OAI-PMH are indifferent to proprietary or limited-access digital collections, and to OA materials.

How can repositories and their contents move toward some stabilized, coherent framework? To bring back the vocabulary of SCOT, let us call this stabilization process "closure". Common sense immediately tells us that closure does not mean erasing diversity. Far from being a problem, diversity brings some very positive elements in support of creativity and innovation. Closure should lead to developing a framework in which various types of repositories can function in a complementary and mutually supportive manner.

The vocabulary just used did not emerge by chance: it is inspired by history, in particular the history of computer networks. With computers, diversity could neither be ignored

by the relevant RSG's, nor controlled. A "good" solution in emerging computer networks had to incorporate diversity and even make use of it. The Internet, as we know, was built upon a network of networks where diversity could contribute to the robustness of the whole.

The Internet analogy demonstrates the need for diversity in the design of repositories. It also underscores the potential importance of a networked solution which means thinking about more than a single repository and thinking about their possible relationships. Networked solutions are increasingly being favored because they protect individual autonomy and diversity while ensuring some degree of co-ordination at a more global level. The result generally benefits everyone (Benkler, 2006). In short, repositories should be conceived as nodes within networks of networks.

The DRIVER project in Europe gives some hints that this networking process has begun. DRIVER seeks to create a "confederation of repositories" across Europe. A recent overview of European repositories brings out a number of interesting results (Van der Graaf, 2007). It identifies some "important stimuli" for the development of IR's and helps understand how repositories are evolving. Among the stimuli, increased visibility and citations of the publications as well as a simple depositing process were favored by nearly 50 % of the respondents; increasing interest and awareness among faculty and administrators were mentioned by about 33 % of the repositories. Conversely, the lack of a depositing mandate was seen as an inhibitor by over 50 % of the respondents. Nearly a third deplored the failure of funding agencies to mandate deposit.

The stimuli and inhibitors just mentioned have one point in common: they pay attention to the concerns of researchers and administrators. The problem is to engage these two important RSG's. Both groups must help shape the repositories, in direct and frequent dialog with librarians and/or computer services, but this is the detail that has been missing in most cases. Libraries try spreading the good word to faculty and administrators but, as they remain largely outside the development process, the librarians' message often remains unheard and the strategy largely ineffective. The institutional hierarchies and the traditional division of labor through delegation appear in this case to work against true collaborative engagement.

Faculty should take the lead in the design of the repository characteristics that serves them most directly, i.e. with the functions that support research and increase their impact in their own peer communities. Administrators should help develop repositories in such a way that they enrich and improve procedures they use or need to administer. As a result, they would understand better the value and importance of mandating deposits. They would also have a clearer view of the relationship between career management (tenure and promotion for example), evaluation metrics and the search for excellence.

Researchers and faculty members, if invited, will quickly point out their needs. It can be expected that these will include clearly separating the peer-reviewed documents from the other digital collections. They will point out that they work in specific subjects and, therefore, need subject-based sets of documents. They will add that sets must be of a sufficient size to attract any researcher. Consequently, they will most likely suggest that their university should team up with other institutions to reach the size necessary to attract the attention of scholars. The beauty of digitized documents is that they can be indefinitely repackaged to suit the most obscure needs; Open Access documents lend themselves even more readily to this reconfiguration of content.

Researchers also want to make themselves more visible and increase their impact. They will, therefore, ask for tools that make harvesting easier and tools that create some forms of branding. The harvesting issue is essentially being solved with OAI-PMH and with ORE, as well as with various search engines, including Google Scholar and Oaister[37]. The branding issue is much newer to the repository world, but it is not entirely unknown. Repositories can lean upon the experience of the "Faculty of one thousand" of Biomed Central[38]. They can also pay attention to the Dutch project, "Cream of Science" mentioned earlier. More generally, subject-based repositories, especially if they are the result of networked institutional repositories, can powerfully contribute to complementary forms of branding. There is no question that researchers will deposit their articles far more enthusiastically if they understand that this simple gesture adds to their impact, visibility and, ultimately, authority.

The branding capacity of repositories should also attract the administrators' attention. New forms of branding can

go a long way toward correcting some of the most glaring deficiencies of scientific evaluation nowadays, including peer-review. They include quality rating by users[39] or new metrics (for example downloads and webometrics). The point here is that Open Access repositories can become very important means to enrich and improve existing forms of scientific and scholarly evaluations.

Creating alliances between researchers, administrators and librarians will inevitably lead to new questions and new solutions. For example, the repackaging of peer-reviewed papers within subject-based collections will lead to redefining relationships between subject-based and institutional repositories. Not only can subject-based repositories harvest institutional repositories, but they can provide some added services that would serve researchers. Symmetrically, institutional repositories can endeavor to build prestigious subsets of articles within various subjects or disciplines.

This argument requires a working alliance between administrators, researchers and librarians beside or even beyond the hierarchies and divisions of labor that have stood in the way of these collaborations so far. In many ways, the present stage of development of repositories brings us back to the Renaissance: then, printers were trying to imagine how to make humanists and metallurgists work together despite deep social divisions and strong hierarchical biases. In this, they were helped by various commands and demands from the power structure that provided both rewards and some degree of enforcement. Likewise, our modern funding agencies are in a position to open opportunities, but also to express specific demands, and so are educational agencies that conduct various evaluation exercises, as documented by the Australian example mentioned earlier.

One final remark may help better to understand the potential of repositories when they begin to provide evaluation and branding services. If we change perspective for a moment and look at Open Access journals, some interesting consequences follow. Perhaps because so much has been made of sustainable models for OA journals and perhaps because the major form of sustainability found so far has been the so-called "author-pays" model (in actuality, the model should be called "author-proxy-pays" model), little notice has been made of the majority of OA journals

that do not require any fees from any authors because they rely on other forms of support. The SciELO collection of journals[40] in Latin America, Spain and Portugal is supported by governments. Governments do so because they support research and they simply include the cost of publishing in the research budgets. For this reason, the only requirement to publishing in a SciELO journal (there are 528 titles in the SciELO collection as of January 2008) is acceptance by peer review. Once the peer-review process is completed, the collections of articles from these journals (over 160,000 articles) form a repository, but a repository endowed with the ability to manage peer review and enriched by a variety of branding tools that are associated with it. In other words, looking at repositories from this angle reveals that OA journals can begin to look suspiciously like subject-based repositories with new functions and services added. Repositories have the potential to present themselves by subject and to add some peer review and branding capacity. And that begins to look suspiciously like a journal. In short, repositories and OA journals are not two entities that shall never meet; on the contrary, they appear perfectly capable of converging, overlapping and even of merging.

Because of this convergence phenomenon, it becomes possible to bring to the design of repositories notions borrowed from journals. For example, journals are important elements in building intellectual communities. Repositories can then be envisioned as part of the architecture of various communities. Librarians who have worked on building communities within universities should feel at ease in this role. But, to repeat, this approach cannot succeed without a strong sense of networking beyond the walls of the local institution. Knowledge is woven across the whole world. Consequently, designing knowledge tools on a parochial basis will fail. Bringing closure to repositories requires harmonizing these devices with the requirements of truly globalized knowledge. Repositories will then become part of a "knowledge society".

In the end, it does not matter if repositories are depositories, collections or archives; in fact they should be all of these. When bicycles stabilized, they fulfilled a whole range of functions. They were tools to go to work, to race, to explore the countryside with the family. They were also toys for children and they adapted one way to men and another to women. Repositories can be just

as encompassing in their functions at the moment of stabilization, but this versatility should not work against a full insertion within host institutions; neither should it inhibit networking with repositories in other institutions.

Finally, it should not stand in the way of a vibrant dialog with their relevant readers. Texts cannot live without communities of readers. A text resting in a repository is no exception to this rule.

NOTAS

1 Many thanks are due to Frances Groen, my wife, who kindly saved me from the worst gallicisms that were initially present in this text. Dr. Karim Gherab Martín should also be thanked for his infinite patience with my delays.

2 This summary of SCOT's main concepts relies heavily on Klein and Kleinman (2002).

3 http://www.cni.org/.

4 See http://publicaccess.nih.gov/. It took several years to achieve this result. Peter Suber provides a timeline of the "saga" leading to the recent NIH policy at http://www.earlham.edu/~peters/fos/newsletter/08-02-07.htm#nih.

5 One should add that it is a very useful article to review a number of issues and questions related to IR's. In particular, it assembles an interesting set of articles discussing the ways and means to populate the IR's, along with the need for mandates to deposit. The latter issue clearly belongs to institutional politics.

6 This paper is slated to appear in *Serials Review* in 2008. A preprint is available at http://eprints.ecs.soton.ac.uk/14965/.

7 Sally Morris, former Executive Director of ALPSP, provides an interesting example of this viewpoint: On February 1st, 2007, Morris was reported to have said: "'I think that this is actually much scarier for publishers than the open access publishing mo-

del. While self-archiving may coexist with the subscription model, it has the potential to parasitize it to the point that it actually kills it'. Institutional repositories may influence librarians to cancel subscriptions and 'could bring about the demise of a lot of journals very quickly'". Michael Mabe, the CEO of the International Association of Science, Technology & Medical Publishers added: "I think the increasing use, in some quarters, of repositories as an alternative to scholarly dissemination will potentially create great damage in terms of the ability for journals to remain viable...". See Drake (2007).

8 See, for example, the CLOKSS project, http://www.clockss.org/clockss/Home.

9 http://www.driver-repository.eu/.

10 See http://www.sherpa.ac.uk/romeo.php.

11 See the ICOLC guidelines at http://www.library.yale.edu/consortia/statement.html. In particular, it states: "Non-disclosure language, if necessary, should not preclude library consortia from sharing pricing and other significant terms and conditions with other consortia". The response by Elsevier can be found at http://www1.elsevier.com/homepage/about/esi/issue1101.pdf.

12 For an early discussion on this issue, see Peters (2002). See also Kohl and Sanville (2006). An example of consortium repository is OhioLink (see http://www.ohiolink.edu/). "White Rose Research Online" is a reposi-

tory shared by three UK universities (Leeds, Sheffield and York) but this "consortium is not a licensing consortium, unlike OhioLink".

13 For examples related to the Canadian consortium, CRKN, see the Elsevier publication *Library Connect,* vol. 4, n.º 3 (August 2006), p. 12 (available on-line at http://www.elsevier.com/framework_librarians/LibraryConnect/LCN0403/LCN0403.pdf and vol. 5, n.º 2 (April 2007), p. 1 (available on-line at http://libraryconnect.elsevier.com/lcn/0502/lcn050213.html).

14 As early as March 1998, ICOLC issued a "Statement of Current Perspective and Preferred Practices for the Selection and Purchase of Electronic Information" in which it was stated that "Non-disclosure language, if necessary, should not preclude library consortia from sharing pricing and other significant terms and conditions with other consortia". Elsevier responded in the Fall of 1998 (http://www1.elsevier.com/homepage/about/esi/issue1101.pdf) with careful language asking not to be pushed into identical arrangements with all consortia.

15 See the text of H.R.2764 at http://thomas.loc.gov/cgi-bin/query/D?c110:9:./temp/~c110Eb0YRh:: [the double colon is part of the URL].

16 http://www.wellcome.ac.uk/doc_WTD002766.html.

17 http://www.pubmedcentral.nih.gov/. For an example of lobbying language, see http://www.publishers.org/main/PressCenter/PRISMLaunch.htm.

18 See note 5.

19 http://www.taxpayeraccess.org/.

20 http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/1200/120002.htm.

21 The full report can be found at http://www.rcuk.ac.uk/news/20060628openaccess.htm.

22 On the OA advantage, there exists a large collection of texts. For a bibliography of these, as well as comments, see http://opcit.eprints.org/oacitation-biblio.html.

23 The "Roar" site (or Registry of Open Access Repositories) lists 989 repositories while "Opendoar" (directory of open access repositories) holds 1,035 repositories (January 23rd, 2008). See http://www.opendoar.org and http://roar.eprints.org.

24 http://www.ifla.org.

25 In OpenDoar, we find that 836 are institutional repositories, 128 are disciplinary, 53 aggregate their documents from subsidiary repositories and 18 are governmental repositories.

26 One of the first institutions to achieve a deposit mandate is the University of Minho in Portugal. See Ferreira, Baptista, Rodrigues and Saraiva (2008).

27 Oaister (http://www.oaister.org/), a search engine harvesting from 934 repositories gives access to over 14 million documents. Oaister harvests repositories that hold materials in Open Access but in so doing, it regularly points to toll-gated materials that are also kept in a repository.

28 On a related subject, William L. Cheung shows how evaluation methods and the peculiar needs of post-doc students intersect each other and affect the building of what he calls a "portfolio of objectives". See Cheung (2008).

29 See http://www.dest.gov.au/NR/rdonlyres/3765571A-6B53-4E39-90A1-0057245509C6/16873/FactsheetASHER30May07.pdf.

30 With the recent change of government in Australia, this project may undergo significant changes.

31 http://hal.archives-ouvertes.fr/.

32 http://www.pubmedcentral.nih.gov/ and http://arxiv.org/.

33 http://www.darenet.nl/en/page/language.view/search.page.

34 http://www.creamofscience.org/en/page/language.view/keur.page.

35 http://www.openarchives.org/. OAI-PMH stands for Open archive initiative - Protocol for Metadata Harvesting.

36 See http://www.openarchives.org/ore/documents/ORE-Announcement.html. A succinct objective of ORE might be as follows: "Object Reuse and Exchange (ORE) will develop specifications that allow distributed repositories to exchange information about their constituent digital objects".

37 http://www.oaister.org/.

38 http://www.facultyof1000.com/.

39 *PloS One* has been a most interesting pioneer in this kind of development. See http://www.plosone.org/home.action for further information.

40 See http://www.scielo.org/php/index.php?lang=en.

BIBLIOGRAFÍA

Bailey Jr., Charles W. (2008): "Institutional Repositories, Tout de suite", available on-line at http://www.digital-scholarship.org/ts/irtoutsuite.pdf.

Benkler, Yochai (2006): *The Wealth of Networks: How Social Production Transforms Markets and Freedom,* New Haven, Yale University Press, *passim.*

Cheung, William L. (2008): "The Economics of Post-Doc Publishing", *Ethics in Science and Environmental Politics,* vol. 8, Preprint 2008, available online at http://www.int-res.com/articles/esep2008/8/e008pp4.pdf.

Drake, Miriam A. (2007): "Scholarly Communication in Turmoil", vol. 24, n.º 2,

Feb., available on-line at http://www. infotoday.com/it/feb07/Drake.shtml.

Ferreira, M.; Baptista, A. A.; Rodrigues, E. y Saraiva, R. (2008): "Carrots and Sticks. Some Ideas on How to create a Successful Institutional Repository", *D-Lib Magazine,* vol. 14, n.os 1-2, available online at http://www.dlib.org/dlib/january08/ferreira/01ferreira.html.

Klein, Hans J. y Kleinman, Daniel Lee (2002): "The Social Construction of Technology: Structural Considerations", *Science, Technology & Human Values,* vol. 27, n.º 1, Winter, pp. 28-52.

Kohl, David and Sanville, Tom (2006): "More Bang for the Buck: Increasing the Effectiveness of Library Expenditures Through Cooperation", *Library trends,* vol. 54, n.º 3, pp. 394-410.

Peters, Thomas (2002): "Digital repositories: individual, discipline-based, institutional, consortial, or national?", *Journal of Academic Librarianship,* vol. 28, n.º 6, Nov.-Dec., pp. 414-417.

Pinch, Trevor J. y Bijker, Wiebe E. (1984): "The Social Construction of Facts and Artifacts: or, How the Sociology of Science and the Sociology of Technology Might Benefit Each Other", *Social Studies of Science,* vol. 14, pp. 399-441, available on-line at: http://arno.unimaas.nl/show.cgi?fid=1270.

Van der Graaf, Maurits (2007): "DRIVER: Seven Items on a European Agenda for Digital Repositories", *Ariadne,* vol. 52, July, available on-line at http://www.ariadne.ac.uk/issue52/vandergraf/.