New trends in scientific and technical information

1. Topicality

The access to scientific and technical information (STI), its circulation and its availability are key factors for the effectiveness of scientific research, of industrial competitiveness and of social progress. This explains the amount and the spectacular growth of the investments authorized in its favour in all the developed countries. However, in France and abroad, experts and actors do agree in realizing that the situation of STI is not satisfactory and, worse still, that it is steadily degrading despite (or due to?) the so-called new technologies in information and communication.

This short contribution aims at identifying the principal stakes of this problem, first by passing in review the different positions, not to say forces, which are involved.

The actors

• On a first side, **major publishers, most of them of European origin**, such as ACS, Blackwell, Reed Elsevier, Springer, Thomson, Wiley, etc, as well as **national actors**, are **alarmed at the appearance of the new form of competition** that arises from non profit, co-operative and free modes and supports of diffusion and access. According to them, these modes and supports put in danger their trade and they **weaken the economy of the sector**, which has just faced the advent of the Web and adapted to it. The industrial and **financial movements and merges** that the various branches of the educational, scientific and technical edition are submitted to presently are, to a certain extent, related to these concerns.

Obviously most of these publishers are conscious of the need **of providing new tools and services** in order to answer the changes in progress: they speak of search engines, data mining, indexing, intermediation and even of bibliometric evaluation. At the same time, they readily admit that the **profitability** of these new tools and services is not clear yet. Moreover, they hardly imagine which new economic models would coexist with the old ones without, inevitably, replacing them.

The users

• On their side, **the final users** of STI, in the world of **research, either public or industrial**, complain about the explosive increase of the **cost of the access to scientific journals**. Due to the lack of transparency of the publishers regarding the marginal costs of digitalizing the data and making them available on the Web, the **tariffs of the main publishers appear unjustified**. Moreover it is observed that the **"bundle practice"** of the publishers oblige the users to subscribe a pack of titles although many journals of the pack do not correspond to their proper needs, and also that, without any explanation, they suddenly shift some favourite titles of the bundle into separate marketing more profitable to them.

• Among the users, the **authors** and, even more, the **public research organizations** (*e.g.* in France, CNRS, INRA, INRIA, INSERM, universities) are more and more **reluctant** in accepting the constraints imposed by the large publishers. The director of one of those organisations used to say that, as a whole balance, the research community **is paying at least three times** for a research: paying the author, paying the reviewer, paying for the bundle and, maybe, paying for downloading.

• Industrial research organizations express the same reluctance regarding the "bundle" practice. They negotiate contracts with publishers privileging ***"core and pay-per-view" packages***. The tendency is even towards contracts where a given amount of money gives full access to the whole catalogue of the publisher not being linked to a list of titles and leaves the possibility of buying up to this amount. Negotiations are handled by professionals of STI and the ***purchase department*** of the concerned company.

• Another source of sharp concern is that the interoperability of the systems are not always guaranteed and that apart from initiatives coming from private publishers which deserve being acknowledged, such as the deposit of the numerical files to the Royal Library of the Netherlands, there are good reasons to think that certain **patrimonial data of priceless value may disappear or become inaccessible** because of **obsolescence of the standards and material supports**.

• Finally an ultimate cause of dramatization is that the discussion may take a **political turn** opposing a concept privileging **public service** with the support of the State and self-management by the scientific community to the liberal approach, free enterprise

and market regulation. One easily guesses the virulence of the *pros* and *cons*, in particular when realising that most major publishers are, as already said, of European origin.

Then the question to be answered is:

How to give access to scientific and technical information, to the maximum of users at a minimum cost?

2. The context

The context is determined by the considerable changes that scientific and technical information undergoes, because of digitalization. Started at the beginning of the eighties, with the development of the Web, these changes affect three essential elements in the practice of access to knowledge, with major consequences on the services provided by the publishers.

Publishing without a publisher

A first element comes out from the fact that, by now, the development of numerical technologies makes it possible to **individuals to carry out themselves the tasks** which they used to hand out to specialists formerly in just the same way as one reserves his plane tickets without calling upon a travel agency. As a matter of fact, the **main parts** played by the publishers in the publishing process have been **impacted by digitalization**.

Text processing software such as TeX or even Word enables one to design and print scientific or technical documents without the services of a specialized personnel, and sometimes **reaching professional standards**. Publishers followed these evolutions and, as a counterpart of putting on the authors the burden and effective costs of typing and design, they argue that they provide them with the possibility of keeping the control of their publication from beginning to end. As a matter of fact it appears that, despite the above mentioned obligations, **most authors prove to be satisfied**. This is one reason why the leading practices in STI changed so rapidly, notwithstanding, in certain cases, a lower quality of the final product.

The Web also offers new and **diversified possibilities for the distribution of articles** to the scientific community: a researcher can proceed individually, using his personal home page for posting his preprints and/or his publications; he may even

post successive versions of the same paper according to their degree of elaboration. Another possibility is to post them in a **repository** such ArXiv or HAL or to refer to the sites of the learned societies whose he is member.

Even the function of validation of content through peer-reviewing, which the researchers have always been in charge of, sometimes with the support of the publisher, has been modified by the Internet notably.

This shows how the functions of typing, making up, printing publishing and distributing, which used to be the monopoly of the publishers and appeared as the core of their business, are now being taken over by the researchers themselves. Isn't it worth noticing that this revolution that we may call "publication without a publisher" only took two decades? This may explain why the principal actors of the sector, the authors, the users, the publishers, the librarians, the custodians are somewhat disappointed and look for the conditions of a new balance.

The publishers

It is observed that the **considerable reduction – if not the disappearance – of the costs of distribution** makes it possible to the large publishers to defer on the upstream (production, organization, presentation) a larger part of the expenditure in order to cover the **increasing amount of the investments** and the need to acquire **increasingly complex know-how** for the maintenance and management of their Web-site.

Contrary to what might be expected this **does not imply the disappearance of less concentrated actors**: according to a well-known mechanism, the technical facilities and the lower "entrance fee" encourage **small actors** to try their chance: by restraining the costs of operation to a minimum since they **do not have much to lose and they have much to gain** if they manage to occupy a profitable "slot".

Thus, concentration in the publishing industry is the fact of those of the publishers who, thanks to financial and industrial winning strategies, succeed in surviving the changes, and benefit from them, in particular by acquiring a quasi monopoly *de facto*; but these same publishers are weakened by the existence of a fringe of challengers and by the presence of "amateurs", nonprofessional publishers and leading structures of great institutions, who are ready to compete with them effectively.

Data bases

Scientific and technical information also refers to data bases. The **providers** of such bases are now offering **ready made solutions** for industrial research organizations, helping them to define their needs and providing them with the relevant content. It appears that such solutions that are seldom accepted by the big companies, which prefer to rely on their own approach, can **benefit to small companies**.

The access to information

An increasing number of users of STI, either in public or in industrial research, do not wish any more to have to go to **a library or to a resource centre** to consult the scientific and technical publications; these users want to reach information directly through their own computer, **on the spot** of their work, **wherever and whenever** they need. As a matter of fact, this phenomenon is not specific to STI. Anyhow, as in other areas, it is the origin of new problems that are still waiting for a solution, just to mention, as an example, the necessity of an identification of the **authorized** recipients, the question of **copyrights** and the **validation** of content.

This brings us to the important question:

What will be the role of the libraries in the future?

3. Open Archives and Open Access

As already stated, the Internet opened new means of communication between researchers: the personal page is one of them but ***Open Archives*** is another one that needs being examined by now, due to its connexion with recent changes in the policy of important research organizations, particularly in the US.

The original term *Open Archives Initiative (OAI)* was introduced in July 2001 for a technical standard, the *Protocol for Metadata Harvesting,* **without any reference** to a concept of free access. This latter idea only appeared gradually among the researchers, once this technical tool was available and it was explicited in the Budapest statement of *Open Access Initiative* in February 2002 (under the impulse of George Soros). On an institutional level in Europe, the Berlin Joint Statement *Open Access to Knowledge in the Sciences and Humanities* was signed in October 2003. Then the scientific community developed the term *Open Access* to cover two types of actions:

- The **"free access"** on the Internet to journals published by the scientific community, directly or via a commercial publisher.
- "Open Archives", data bases of scientific articles either in pre-publication or published.

This issue turns out to be one of highest importance and we will focus our attention first onto the related policy adopted by **prominent scientific institutions in the United States.** As a matter of fact the United States did react very quickly and positively to the Free Access strategy. Let us quote the NIH (*National Health Institute)* policy:

Starting from April 2008, all the articles that proceed from work financed by NIH funding and that are published in journals must be transferred on Open Access in PubMed Central in the specific line dedicated to NIH. The contracts with the publishers must explicitly allow it. There is even a list of publishers for which this transfer is automatic without any intervention by the researcher. The law (Consolidated Act Appropriations, 2008) **specifies that the transfer must be carried out at the latest 12 months after the effective publication in a review.**

From May 2008, any submission of a research project to the NIH must mention, in the articles quoted in reference, the URL of the article on the PubMed site.

This policy of the NIH was taken according to the "SPARC amendment" that had been **voted by the American Senate** in May 2006, imposing the 11 research agencies with an annual budget superior to 100M$ to set up a policy of free access on the Internet.

And what about Europe?

The European Commission seems to be "sensitive" to the economic positions and lobbying of the large European publishers, which dominate the world market. Several directions of the Commission are neutralizing each other: direction of research (which defends the interests of the researchers), the economic direction (which defends the interests of the publishers). It follows that the conclusions of the *Council of the European Union on scientific information in the numerical era*, on November 22, 2007, remain somewhat timorous as regards Open Archives: the Council invites the Member States to promote

"*As of 2008, free access on the Internet to the results of any research financed by the authorities, through economically viable means, including a delayed free access*".

4. Conclusion

Although, as I said this revolution started two decades ago, it is not over yet. **Changes will go on** that will impact many features of science and engineering among others.

Just to list a few examples:

- It seems evident to me that **"publishing without a publisher"** is an unavoidable trend **provided the validation** of contents of such "Open Journals" is achieved.

- Researchers and engineers will require **certified data bases** and **certified information** to be extracted from the ever-growing bulk available.

- As it is already the case in some large companies I happened to know of, **librarians and custodians** will be in charge of providing **permanently improving facilities** and some kind of **continuing education** to researchers and engineers in order that they derive the best profit of them.

- Also, I wish to emphasize the point that great attention should be paid to the **preservation of scientific and technical documents and data** of high patrimonial values. Unfortunately, the new material supports we are using now will not last as long as the stones the Mayas were using. We won't complain because they are, by far, easier to handle. But to a certain extent this adds to the complexity of the problem since the historical value of any document is difficult to anticipate.

Obviously such an issue should be taken care of at an **international level**.

- Finally, although it was not included within the scope of the report I am in charge of, I wish to address two issues which I consider very important as long as information implies communication namely: **languages** and **books**.

When one focuses on **scientific research**, there is no doubt that **English,** whatever its form, is the unavoidable language for information and communication presently. Therefore, maybe with the exception of some areas where it is still possible to publish and be read in another language, such as French for Mathematics or social sciences, it is clear that international reviews and journals are English-speaking.

When it comes to **technical** aspects I believe things should be considered differently: one should write in the language that is the **most understandable** to the **concerned reader** in order to save time and prevent any confusion. When speaking of languages I would say, as in electrical engineering, that **no transformer** should be necessary to

pass from the writer's voltage to the reader's one! Then both the **language** and the **words** will be those of the reader.

A similar rationale applies to **books**, whatever their future material support, which should not be any surprise as it is related to the same issue, namely: **communication**. Since we already mentioned the ever-growing amount of information, books that bring together, classify, assess and coordinate different pieces will be more and more valuable. Here again, as long as a book is concerned with the **front edge of science** its language will be English and the publishers will not give the writer any choice on that point. But I believe that when **basic scientific knowledge** is concerned, which develops and advances continuously, there is an absolute necessity that books and text-books be written in the mother tongue of the reader in order to provide him with the best chances of access to the intimate content.

These necessities, not to say these duties, should be taken into account, depending on the discipline involved, when **assessing researchers and teachers**.

¡Muchas gracias por su atención!