The Internet and the Information Environment

V. B. Borshchev

Received December 8, 2008

Abstract—The Internet has changed the information environment in which people involved in scientific work live in slightly over ten years. The purpose of this article is to trace the vector of these changes using several examples, first of all, by the example of linguistics, using the LINGUIST List system as an example. In this paper, different principles of organizing information interaction in science are discussed, which are “vertical information” principles, being implemented in our country by VINITI RAS and other information services, and the information self-organization of the scientific community, which relies on the network information infrastructure of the Internet.

Key words: information environment, metainformation, comfort, information resources, information servicing, information self-organization of the scientific community.

DOI: 10.3103/S0005105509010051

1. INTRODUCTION

The Internet has changed the information environment in which people involved in scientific work live in slightly over ten years. I have previously written about this along with R.S. Gilyarevskii [2–7]. These changes still endure, the character of scientific communications and its tools are changing, and the style of information interaction in science and its methods are changing as well.

The purpose of this paper and its fundamental idea is to trace the vector of these changes using several examples. The fact that I consider to be the most important is that the Internet promotes the self-organization of information of the scientific community and “information self-service,” as it may be called.

The concept of the information environment of science was introduced in 1976 by Yu.A. Shreider in the paper “Information processes and the information environment” [8]. He discovered important peculiarities of this environment and contrasted (or compared) this concept with the existing ones of an information system and information resource. Shreider’s fundamental work will be mentioned in section 2.

In the main part of the paper (section 3) I will illustrate the changes that are occurring in the information environment, basically by the example of linguistics. There are many different linguistic information resources on the Internet. These include descriptions of books and journals (increasingly often online versions are presented), archives of articles and other data, blogs, personal websites of scientists, websites of organizations and arrangements, Wikipedia, text corpuses and other linguistic databases and tools, specialized software, all kinds of educational resources, etc.

It wouldn’t be easy to deal with this variety of information resources, but metasystems appeared, which help users (mostly linguists) to get oriented in the existing information infrastructure, and help the developers of new information resources to build them into the infrastructure freely. First of all, the LINGUIST List system is remarkable.

All these factors have essentially changed the character and style of scientific work, mainly from the information point of view. The information infrastructure itself and, first of all, its formalized part, which exists now mostly in electronic (digital) form, has changed radically. It has become incomparably richer than before, when it was, roughly speaking, a system of scientific publications, files of abstracts and library catalogues.

In section 4, this information infrastructure, represented on the Internet, is compared with the centralized system of scientific information, the development of which started in our country in the 1950s and which is represented by the VINITI institute where I work. We will discuss principles that are in some sense non-overlapping; the Vertical information implemented (in our country) by VINITI and other information services, and the information self-organization of the scientific community, which relies on the network information infrastructure of the Internet.
2. INFORMATION ENVIRONMENT

2.1 Shreider: The Information Environment, Metainformation and Comfort

As discussed above, the concept of information environment was introduced in 1976 by Yu.A. Shreider. At that time neither the Internet, nor even PCs existed. But science did not live in a vacuum and Shreider inherently created the idea of the information “ether,” in which science existed, introducing the concept of the information environment. Shreider didn’t give any definitions, and, more likely, he considered the concept self-evident, relying on natural associations connected with the word environment².

Shreider discussed such properties of information environment as metainformation and comfort. These properties are closely connected with each other. Metainformation is information about information. Shreider defined informatics as the science about “properties and methods of information retrieval and representation.” This means different means of classification and rubrication in scientific and abstracts journals, catalogues, and information systems (hyperlinks and navigation paths in the present-day Internet).

Comfort may be understood in different ways and “two opposite user’s purposes may be distinguished”:

(1) “The first purpose means that the metainformation that describes a user’s information requirements in the most accurate and complete (so-called targeted) way is considered the best one. Such a user feels comfort, because the acquired metainformation removed the complexity of decisions about the utility of any additional information. A student gets the textbook recommended by the program, a designer gets a recommended list of books, a director gets information about the latest achievements in science and engineering within the branch that he is responsible for, etc.” At the time of the first information systems, such terms as selective dissemination of information and information support of scientific research were popular. In particular, it was proposed to regularly distribute information that corresponded to the (appointed) user’s information requirements³.

(2) “The opposite purpose is that the user feels comfort as an opportunity to investigate information by himself. Metainformation that is a good guide to scientific information and far enough from the user’s direct interests, but not the type that emphasizes the scope of the user’s requirements, is helpful for this type of user. For such a user, the model of a comfortable information environment means a library with open access to its shelves, a group of scientists of contiguous specialties, a dialogue information system providing advanced semantic associations, etc. (...)”

In the first case, the comfort of the information environment may be achieved by improving targeted searching (in this case the information service decreases a user’s opportunities for choice in cutting off information noise), in the second one, by providing the necessary metainformation for scientific information. In this last case, comfort means increasing a user’s opportunities for choice. In such an approach the quality (comfort) criterion of information environment is determinative.”

2.2 The Internet

New mechanisms: hyperlinks and search engines

The Internet strengthened the opportunities for the approaches described above significantly. The metainformation structure was supplemented with the mechanism of hyperlinks, which is organic to the Internet. Through this, the information environment in the Internet becomes hypertext. In the Internet, methods of accessing information resources and navigation are supported mostly by the mechanism of hyperlinks.

Search engines and such systems as Google and Yandex, which use such engines (it is worth mentioning that the first information systems, which carried out searches using keywords, were the basis for the appearance of search engines), became another powerful tool. It is a powerful tool indeed, although it is inflexible.

Its inflexibility is caused, first of all, by the penury of the applied query language. If a retrieval is carried out directly on the Internet, then it hardly can be called targeted, it looks more like a scattershot approach. It is natural, to use such a search in combination with other methods of accessing information resources.

As well, there are many information tasks that can’t be formulated using query language. For example, you want to read news or a certain newspaper on the Internet. If you were offered the choice of formulating queries for news instead, that would be absurd. The same reasons make it impossible for any query language to replace paging viewing scientific journals. However, there are systems in which query language is almost the only communication tool.

Of course, these remarks are not a criticism of search engines, which are a wonderful invention. As any other tool, they are just not universal.

Information environment in the Internet

With the appearance of the Internet, the information environment has changed radically and is still changing. The PC screen has really become an intermediary between the scientist and the world of information. The character of scientific communication itself has changed radically. Many “old” resources, including journals, books, library catalogues, etc., are now avail-

---

² In an article from 1987 written in cooperation with Yu.A. Shreider for a general audience, we provided some explanations: “The environment has a complicated structure. It includes personal contacts and correspondence, manuscripts, preprint and reprint exchange, seminars and conferences, scientific journals and other publications, and the system of scientific information providing researchers with an expanding range of information services” [9].

³ In this subsection, quotation marks indicate quotes from Yu.A. Shreider’s work in [8].