Back to Contents Page

CHAPTER I

INFORMATION MODEL OF THE PROCESS OF THE DEVELOPMENT OF SCIENCE

Science is a certain process which is developed with time, and this process, naturally, can undergo qualitative investigation accurately by the same method by which processes in biology, chemistry or physics developed with time are studied.

Qualitative methods of investigation can be used successfully when the problem is formulated well. It is necessary to propose a model of the phenomenon to be studied in order to clearly determine the approach to the problem. Here one must not forget that any model is always narrower than the phenomenon to be studied - it covers only any of its aspects.

Phenomenologically science can be examined as the process of the obtaining of substantially new information. This process has a sequential and collective nature: any scientific work is based on a certain multitude of ideas expressed earlier. New scientific works appear as a result of the further development or reconsideration of works carried out earlier. Science is the self-organizing system the development of which is controlled by its informational flows. External conditions - assignments released for the development of science, organizational forms which are formed in a certain country, ideological pressure, secrecy - all these are only elements of that medium in which science is developed. The medium can be favorable or unfavorable for the development of science, but it cannot force science to be developed in any direction organically alien to it. Here a comparison with biology is appropriate. The development of an organism is determined by the hereditary informational flows; the environment can only promote or retard the development of an organism, but it cannot direct its development according to any specific means alian to this organism.

We will call the quantitative methods of the research of the development of science as an informational process the *measurement of science*. This is the *cybernetic* approach. It is known that the study of complex systems can be conducted by means of an investigation of those informational flows by which they are controlled. One of such examples is the study of heredity by means of the interpretation of the DNA code.

The model examined by us allows clearly outlining the range of the phenomena studied. Immediately we can exclude from the examination everything which refers to the so-called "closed" science, since the development of the "closed" works do not affect the informational flows in science. Let us explain our thought in the following example. Let us assume that scientist A makes a report and communicates in a lecture a concept which has become obsolete and already inaccurate. Let us assume that in the same lecture there sits scientist B, who knows a new solution to the problem but cannot report it, since this result was obtained in a closed work. Which of these two scientists affects the development of science? Only scientist A, although in this particular case he "moves science backwards." Scientist B has an effect on the development of science only when he can report on his results in any form, even obscure, and make it so that they merge into the general informational flow.

The model accepted by us also allows excluding from the examination everything connected with the inculcation of science into technology. Technology refers to the sphere of the material activity of man, which although it is based on informational flows coming from science it is not assigned entirely by them. The goal of technology is the creation of material values and not ideas, and this process, naturally, cannot be examined within the framework of the informational model formulated above. The development of technology can be examined as the development of a system which is the daughter with respect to science. The interaction of these systems is a subject of special investigation, which he merges outside the limits of the problem examined by us. Here it is interesting to note that there exists an informational barrier between science and technology - informational flows of science, as a rule, are not very accessible to the engineer. On the boundary of this barrier special transformers exist which convert informational flows of science into form available for engineers. The first transformer is the higher educational institution, which prepares the engineer at the level of the contemporary development of science. (Note: a scientist working in the field of technology, let us say, a metallurgist, works nevertheless not at a plant but at a higher educational institution.) The second transformer is a different kind of specialized courses, which serve to acquaint the engineer with new scientific ideas. The third transformer is the publication of such literature in which new ideas, let us say, ideas of mathematics, are treated from positions of their possible use in technology. The front of technical developments, probably, is assigned not as much by the front of scientific investigations as by the state and organization of systems of the transformation of information. Specialized and highly mechanized information centers should become a new form of such transformers, but as yet we do not have them available to us.

From positions of the definition accepted by us, beyond limits of our interests there are also questions as gnoseology, methodology of scientific investigations, logic of the development of science, psychology of scientific work, and many sociological and economic aspects of the development of science. One should not think that we deny the importance of these questions. The fact here is simply that now in the first stages of investigation it is difficult to propose a single model which would cover the process of the development of science and its adversity. The formulation of the problem is determined, naturally, by the selection of the model of the process, and the methodological problem fundamental for any exact science is solved. It immediately becomes clear what and how it is necessary to measure and to interpret results of measurements from any positions.

If we examine science as an informational process, then it is natural, first of all, to observe with time the growth of the number of scientific publications, considering them as information carriers. Very interesting results in this trend were obtained by Derek Price; they are given in his book "Little Science, Big Science" [12]. Here, by the way, let us note that for us and for Price there is a very similar relationship to science as to the object of investigation. Its formulations, true, are somewhat paradoxical: "...the main finite goal of work of the scientist is the article which he publishes," "...the scientist needs to write and not read; the engineer needs to read and not write," "...it would be naive to examine technology as an applied science..." [13]. From works of Derek Price it follows that in our time science has entered into a new phase of development (little science has turned into big science), which is connected with a certain crisis of growth. He explained the mechanism of this crisis (adaptive retardation of the development of science) and noted the appearance of new forms of the organization of science directed at its overcoming. All these questions will be thoroughly examined in the following chapter.

In certain divisions of knowledge in the investigation of informational flows, it is possible to study not only the growth of publications - carriers of information, but also the growth of *separate indices* characterizing the direct results of certain investigations, and even to produce *a formal statistical analysis of the content of the publications*. Thus, for example, it is possible to study the growth of the energy of elementary particles attainable in accelerators, the growth of operating speed and storage capacity of computers, and so on. This approach to the problem will be examined in Chapter III. In such an approach, of course, it is possible to obtain only certain particular results having values only for certain separate fields of knowledge. These data prove to be useful in the control of the process of a development of science, for example, in the selection of the most promising and most fruitful trends of investigation being developed.

It is possible to propose one more approach - the study of scientific journals as channels of communication. This trend recently has been intensively developed by the Moscow group of researchers. In such an approach it is possible to obtain certain concepts on the internal (logical) structure of the front of scientific investigations; it is possible to evaluate the efforts expended by individual countries on the development of science and, which is especially interesting, the effectiveness of efforts expended by individual countries on the development of world informational flows.

The statistical study of the language of scientific publications is of special interest. With its isolation and deepening each field of knowledge develops its special language - unique scientific "slang," which makes it bearly available for specialists in other fields. This process is quite natural, since in each even narrow field of knowledge there appear concepts having their special and very deep meaning - with their help whole scientific concepts are coded. Such a system of the coding of scientific concepts facilitate the exchange of information within separate scientific trends, but hampers the contact with adjacent disciplines. Explanatory dictionaries here are of no benefit - they cannot replace manuals which explain scientific concepts. It is of interest to us to examine by means of statistical analysis the history of the development of specific languages in separate fields of knowledge and evaluate their role in the development of the system of scientific information. Thus it is possible, apparently, to solve separate particular problems. One of the similar problems is the study of the penetration of any methods of investigation (and thought) into adjacent fields. For example, it is possible quantitatively to determine the degree of penetration of methods of mathematical statistics into different fields of knowledge by evaluating the frequency of the appearance of specific terms of the given discipline in publications which are devoted to other divisions of science. Such an analysis allows evaluating not only the latitude but also the depth of penetration of new ideas and methods. The latter will be assigned by the frequency of appearance of words by which complex concepts are coded. By this method, apparently, it will be possible to observe the development of separate ideas or scientific schools within one field of knowledge. By using it, it would be possible, probably, purely formally, on the basis of a frequency-response analysis, to observe, let us say, the penetration of the Bayes approach into various divisions of mathematical statistics. The importance of such a linguistic and statistical analysis of scientific publications were repeatedly pointed out to us in the discussion of our reports (seminars at VINITI and the symposium "Investigation of Operations and Analysis of the Development of Science," which was conducted under the leadership of A. A. Lyapunov in Moscow in 1967). Here, true, there appear very great difficulties of both a technical nature, for example, the requirement for specialized computers, and of a fundamental nature - in certain fields of knowledge terms of adjacent "fashionable" fields without sufficient understanding of their content. In our book, unfortunately, there are not any chapters devoted to such an analysis. Work in this direction has only begun. Certain extremely modest results, obtained by the domestic method, will be given in Chapter III (page 62) with a discussion of the question of the penetration of methods of mathematical statistics into analytical chemistry.

Finally, in the last case, if we take the informational model of the process of the development of science, then it is natural to pay attention to the analysis of the specific code language of scientific references in informational flows of science. We propose to examine the system of bibliographical references in scientific publications as a special specific language of scientific information. Abroad the greatest contribution to the study of the system of bibliographical references or citations[1] has been made by E. Garfield.

---

[1]Here one must introduce a certain clarity into the terminology. In this work we will understand by the word citation as the bibliographical reference in a publication and not any words given literally.

One of his articles is called: "Citation Indexes for Science - A New Dimension In Documentation" [14].

Why do we examine the citation index system as a special language? It was stated above that each publication is based on a certain set of ideas expressed earlier. These ideas can be quite new and hardly known to the reader. Nevertheless, the author of the publication is usually not stopped by any account given in detail on them - he is limited by references to publications which appeared earlier in which these ideas were first expressed. The system of scientific citations proves to be that coded language which allows writing articles in compact form without repetitions. The ideas on which the author bases his publication are transmitted by this language. Following after Kessler [15] we can confirm that bibliographical references *reflect that intellectual atmosphere* in which the creation of the publication occurred. Many scientists so well understand and use this coded language that they obtain a certain concept of the work without reading it. Let us assume, for example, that you examine any mathematical journal and encounter great difficulty in the content of the article; let us assume, further, that the brief content (summary) is written extremely incomprehensibly. Then it is natural to pay attention to the system of references - if among them you find familiar names and publications known to you, then you will easily obtain the concept about that range of questions which here can be examined. Colleagues of VINITI must constantly encounter a similar problem - how to understand, even if generally, the topic discussion in an article written in one of the "exotic" languages such as Hungarian or Japanese. The analysis of the system of references immediately gives an answer to this question.

If the system of information references is examined as a special language of scientific information, then there appear two problems: 1) to study the structure of this language and 2) to learn to use it in order to follow the developments of science.

In this chapter we examine only the first problem. First of all, we want to conduct a comparison of language of scientific references with the standard language of man and with artificial descriptor languages used in documentation. The contribution which was introduced by the English Linguistic Philosophy School into the study of natural language is well-known. Here we must pay attention to two self-evident theses of linguistic philosophy, the appearance of which is connected with the name of Ludwig Widgenstein during the last period of his creativity[1].

1) "The language is polymorphic." This means that the concepts do not have clearly outlined facets - in different contexts they can acquire a different meaning. The same thesis can be formulated differently: "the concepts are not names of essence" - there does not exist a unique connection between concepts and essences.

2) "The natural language is refined"; it is not necessary to invent any special best language.

Hardly anyone will argue with the confirmations given above - here we associate them with the English Linguistic Philosophy School only because there they were exquisitely formulated. From them it follows, in particular, that artificial descriptor languages with clearly outlined facets cannot be fundamentally good for the concepts. We assume that everyone could be convinced of this, even though they had little dealings with such languages. It is difficult to transmit the contents of the publication, enriched with new ideas, by means of a great number of descriptors assigned beforehand with clear word facets. This multitude proves to be insufficient for the expression of new ideas. All the difficulties are slowly removed if one turns to the natural language with its ill-defined outlined concepts. In the natural language you can easily impart a completely different meaning to old words.

---

[1] A certain representation about ideas of the English Philosophy Linguistic School can be obtained from the book of Hellner [16].

The language of scientific references possesses an even greater polymorphism than the ordinary language. In it there are no concepts, and it is associative. The reference in a publication is a sign which refers the reader to a certain work published earlier, which is enriched with interesting ideas. The multitude of ideas is associated with one sign - the bibliographical reference[1].

The content of scientific publications is coded in bibliographical references. The author himself produces coding in the process of writing an article. This coding in a certain sense reflects more completely the saturation of the publication with ideas than does the coding by descriptors. In one of his works, E. Garfield cites the following example [18]: he took the article of Siley, which is well-known and enriched with ideas, on the general adaptive syndrome. In a survey of all numbers of one medical journal during five years, there were selected 23 publications in which references to the article of Siley were contained. The themes of the selected articles varied greatly, but in all of them the influence of the work of Siley is felt, although in some cases it is not even mentioned in the text of the publication itself. It proved to be remarkable that in the well-known manual "Quarterly Cumulative Index Medicus" not one of the 23

---

[1]Here it is interesting to pay attention to the well-known analogy in the method of describing a meaning to signs in ordinary language and in citation language. Contemporary linguistics pays attention to different features of words [17]. One can speak about the concept of reference. The word refers to a definite object or several objects. This property of the word is determined more or less clearly. The relationship of the reference creates only a poor language - people go further and assign special meaning to words. It is confirmed that the meaning of a word is drawn out from within the consciousness of man. The word is a certain "scoop," single for all but for different people the contents of this scoop is far from being identical. In the language of scientific references the bibliographical reference serves for referring to another publication. In order to impart a meaning to the reference as to the sign, it is necessary "to scoop" ideas from this publication. The language of references can be used only by the well prepared reader. In exactly the same manner only a highly educated man can ascribe deep meaning to words of the ordinary language.

articles was in the section "Adaptation," which was referred to by the initial article of Siley. It is difficult to produce a search of publications by the generality of ideas with respect to descriptor indexing, which encompasses a formal but not an idea content of publications.

Thus, we propose to examine a system of bibliographical references as a specific informational language. Its statistical analysis proved to be one of the means for studying the development of informational flows in science. By using this method, it is possible to observe the development of separate scientific trends and the penetration of new methods of investigation and to adjacent fields, to evaluate the effectiveness of the work of scientists, proceeding from the influence of their works on the development of informational flows, and so on. All these questions are given in Chapter VI.

Back to Index

To Chapter 2