ADP methods used in documenting the transforming of meteorology through the work of the Bergen school 1917-1920

by

Ralph JEWELL
Arhus SIGBJORN
University of Bergen - SWEDEN

481

1.- The Problem

In this paper we describe how ADP methods were was put to use to serve the needs of a specific investigation in the History of Science. The investigation had been prompted by vigorous philosophical debate about scientific revolutions. What was their "nature" ? What was their "structure" ? Was it characteristic of the growth of scientific knowledge that it periodically underwent "revolutionary change" ? Could one find examples of "true" revolution in science ?

To clarify issues like these I resolved to make a thorough study of a specific example of what might fairly be taken to be a case of revolutionary science. The example I chose belonged to the history of Meteorology. The modern era in the development of this science was inaugurated by the new methods and ideas worked out by the Bergen School of Meteorology, mainly in the period 1917-1920. Based on the concept of "fronts" in the atmosphere the meteorological system developed in Bergen came to provide new foundations for weather forecasting and dynamic meteorology. Through this new system meteorologists acquired a consensus which previously they had lacked. The historical processes by which this new situation arose are processes which illustrate the character of revolution in a science.

From the outset of this investigation of the Bergen School my aim was to be as free as possible from the special restraints and injunctions of any particular given philosophical analysis of scientific revolutions. The case of the Bergen School appeared to have a rich and suggestive historical plot, and I wanted to allow themes and details to evolve freely in the course of widening and deepening the inquiry into it.

From the very start there seemed to be an obvious connection between the *emergency* conditions of a world at *war* and the *extraordinary* character of the meteorological initiatives of the Bergen School. For one thing the international telegraphic meteorological reports which had normally served as the basis for forecasting were now available only sporadically - from the most important observation stations they were not available at all. The Bergen School's meteorology *had* to find a new basis. Further, Norway faced a severe threat of *food* shortages during the winter 1917-1918 and the Government was disposed to encourage any effort from any quarter which could assist in the effort to improve food production. The Bergen group sought and won financial support from the Government on the understanding that they could devise a scheme for meteorological forecasts for farmers in the summer of 1918. Later this was extended to provide a gale-warning service for fishermen through the risky months of Autumn and Winter.

Aviation had developed rapidly during the War and in Norway moves were afoot to establish civil aviation routes. The Bergen School took part in the planning of the meteorological services which these required. As a Neutral, Norway saw that it could perform a mediating role in the building of *peace* at the close of the war. One of the ways in which it considered doing this was through *Science*. In these national plans the Bergen School of Meteorology was allotted a major role.

483

As the study progressed great practical problems were posed not only by the proliferation of research themes like these but also by the increasing amounts of documentary material which became available. It was the discovery of one archive in particular, in May 1976, which spurred the consideration of using ADP methods as a means of making the best use of the sources. This particular archive was found on an out-of-doors site and many of its thousands of letters, maps and reports etc. were written or drawn by members of the Bergen School at or around the time of the main meteorological discoveries.

At this stage of the work the most sensible course seemed to be to combine the ordering of the mass of material with its thematic exploitation for he purpose of the research.

All in all the material was so rich in suggesting new angles from which to approach the Bergen School that I had to find a systematic method for dealing simultaneously with the three functions of salvaging, archiving and the research exploitation of the documents. Above all I wanted to satisfy two demands in my choice of such a system :

1) To avoid "freezing" the thematics of the investigation by a rigid classification system.
2) To find an efficient method for allowing the retrieval of documents connected with any given theme.

With the guiding themes of the research in a healthy state of flux and with a very large amount of highly suggestive material awaiting processing it seemed important to find the means of avoiding having to make too many sweeps over the whole material in order to reclassify items in the light of themes which emerged into significance only later, as the work of processing the papers proceeded. With these requirements in mind I decided to invest time and effort in exploring the opportunities available to me at the NAVF Norwegian Computing Centre for the Humanities.

Through discussions with the staff of the NAVF Centre it became clear that my demand for thematic openness in the classification of the documents could best be satisfied by using a free-text retrieval system, in which documents would be represented not by deliberately chosen key words but by text unrestricted by a particular format design.

2.- The ADP Connection

The Norwegian Computing Centre for the Humanities is a national institution serving the whole of Norway. The purpose of the Centre is to provide for a sensible use of computing in research and development work within the Humanities. In cooperation either with individual researchers or with research institutions it undertakes all phases of general program development.

In principle the Centre is open to researchers from within all branches of the humanities, both from the universities and from elsewhere. At present most of the research and development work of the

484

Centre is in the fields of text processing and archive maintenance.

In this History of Science project we have a machine-readable archive of about 4000 documents. Every individual document of the archive, (which might be e.g. a letter, a report, a memorandum or a lecture-manuscript) was represented in the data base by a number of fixed fields of information, as well as by freely chosen keywords and either a direct quotation or a paraphrase.

This numbering system gave us ample freedom to define various classes of keywords.

    001 : Date
    002 : Sender
    003 : Recipient
    096 : Key words
    099 : Persons named in document
    100 : Free text quotation
          paraphase

Eg. To facilitate the listing of persons named in the document, their names were placed in the 099 rubric.

By this means we combined the economies of structured information with the flexibility of unstructured information.

Without entering into a detailed discussion of how we organized the project we may outline the plan of our task as follows.

We wanted to use the data in the following two ways :

    a) To produce theme-lists, i.e. adp-lists of all those letters, reports etc. which deal with a given theme;

    b) Automatic document-retrieval in the archive.

It was clear that a full-text retrieval system would be appropriate for our purpose.

NOVA*STATUS (NOrsk Versjon Av Status) is such a full-text search system, and is the result of further development in Norway of the English system STATUS 1.

Text retrieval is retrieval of documents in natural language, where the words of documents themselves identify the document, without the necessity for additional data.

In the search for documents having a particular content we have to find those words and combinations of words which such documents would probably contain.

485

A particular search query will involve the two elements :

1) Freely selected words which we judge to be characteristic of the problem or topic-field we wish to explore;

2) Operators which specify the logical relationships between the chosen search words.

The search query, or question, will thus be a definition of the criteria for the possibility that a document will be relevant for our purpose.

NOVA*STATUS offers the possibility of composing complex search-arguments by the use of operators. Text retrieval is traditionally based on Boolean operators AND, OR and NOT.

The relations operators

less than, less than or equal to, greater than or equal to are also available. These four operators are particularly useful when the material contains numbered fields.

The question

002 : Bjerknes AND 001 : < 1910 ?

will give us all documents written by Bjerknes dated earlier than 1910. By using several of the operators one might compose a question having the complexity one requires. By means of these operators and certain commands for sorting and printing, it has been possible to fulfill the requirements of this project in the History of Science.

3.- Assessment

The 4.000 documents represented in the machine-readable data-base were considered to be particularly rich in relevance for the study. It was hoped that such a core archive, when fully processed by the ADP facility, would be a "headquarters" for research into the large number of documents which had been left out of the core archive.

It was never envisaged as a practical possibility to represent *all* the available documents in the data-base, nor, of course, was it intended that subsequent research would be *limited* to those documents which had been given such a representation.

An example of the use of the ADP facility : Dating the "discovery" of the Polar Front

We must distinguish between the various phases in the growth of the *concept* of the Polar Front from the actual coining of the expression "Polar Front". From the maps I could deal with the task of showing the cartographical extrapolations which suggested the idea of the Polar Front. From various other documents it was possible to plot the way in which the Bergen School became aware of the potential international interest which such a line might engender in the planning for future forms

486

of cooperation in Meteorology. But what about the actual naming of this line as the "Polar Front" ?

The concordance showed that the term made its first occurrence in a letter written by a leading member of the Bergen School on January 16th, 1920. But after this occurrence there was no further ones for several months, when suddenly it became quite frequent. From the maps and the correspondence I had already formed the view that it was not until February/March that the first fumbling attempts were made to make the kind of drawings which could reveal the Polar Front. Since this issue lay at the heart of the Bergen School's achievement I had to resolve the question. The occurrence in January 1920 of the technical name for a line which all my other research showed to have been first drawn in February/March definitely stood out as an anomaly. I decided to return to the full text of the original letter, only part of which had been quoted in the data-base representation. It turned out that it was a letter which dealt with issues which could have been discussed only after the summer of 1920 in spite of the fact that it bore the January 16th, 1920 date. The true date of the letter was *January 16th 1921*, the mistake being the familar error of getting the year wrong in dating a letter written in the first days of the New Year. In this way I became even more convinced of the reliability of the conjecture that the expression "Polar Front" was coined not earlier than March 1920.

In this example I have tried to show two things :

1) That the concordance can show at a glance the existence of an apparent anomaly in the development of the Bergen School's terminology, and

2) that the ADP products, such as the concordance, are used in conjunction with the full texts of the original documents. The ADP products turned out to be invaluable research aids, but they do not replace working with the full text of the documents themselves.

The system provided a very effective means of satisfying the original demand for the creation of a research-sensitive archive. Because any word occurring in the data base could be used as the basis for the selection of documents it was an easy matter to shed light on even the most subtle research questions. To document the way in which the scientific milieu of the Bergen School was characterized by great *enthusiasm* in combination with a heightened sense of personal *duty* it was an easy matter to locate abundant relevant material by using the concordance which displayed all occurrences of "enthusiasm" and "duty", and their synonyms.

The likelihood of being able to locate a sufficient number of documents for bringing into focus just those details and connections which were needed for resolving a given research query, made it possible to pose increasingly subtle questions as the investigation proceeded. Re-searches of the archive for documents relating to a topic which emerged into significance rather late in the investigation, could be effected in minutes instead of weeks.

This facility for making it feasible to combine the systematic use of a large research archive with the

487

desire to allow the research to be open to surprising new issues and themes seems to be just what is required for enquiry into the subtle human origins of new theory in science.

PHASES OF THE NOVA* STATUS
REGISTRATION

488