papers, reports, books) and case files with a fairly strict format (e.g. the insurance account file where a foreseeable limited number of transactions will take place against any one account number). Other organizations, such as the U.S. State Department, the Swiss and Foreign Office and the EEC are dealing with this same problem. At the present time, there are no fully implemented systems which do not have drawbacks from the point of view of this Department's particular needs. Nevertheless, each has some experience of value which has been of assistance in the selection of what is believed to be the most suitable approach.

126. In order to form a manageable framework within which to examine the state of the art the various essential elements of the system were identified and studied individually and in relation to the total system. These elements are:

* Information input
* Information storage
* Information processing
* Information retrieval
* Indexing/thesaurus techniques

Information Input

127. The system will be expected to handle in the order of 300,000 substantive items per year, of which perhaps 90,000 will be telegraphic communications. The rest will be an assortment of memoranda, letters, documents, Minutes, situation reports and policy statements. With such diverse categories of material, problems are encountered if full text storage in other than original "hard copy" is desired. The telegraphic communications will pose no difficulty since the new message switch will retain them for a magnetic tape transfer from the communications to the substantive information system in machine-readable full text.

128. One aspect of dealing with material in machine-readable form which should not be overlooked is that at one stage the text was converted from original "hard copy" by a second physical act of striking keys for each letter. In theory, a similar approach could be taken for all items received in the system by keying their contents on to magnetic tape. This is not a practical proposition because of the length of many items; furthermore, it duplicates, stroke for stroke, the work of the typist who prepared the text initially. Because many organizations have faced this problem, industry has applied itself to finding a solution and has come forward with the technique of Optical Character Recognition (OCR) in order to place plain language text into the computer. The limitations as to format and font type imposed by OCR would seem to confine use of this technique to material originating within the control of the Department. Well structured items, such as Memoranda to the Minister would lend themselves to OCR more readily than inter-divisional memoranda and letters, while the possibilities for telegram input are amply demonstrated by the success of the U.S. State Department and the German Foreign Office. Further exploration of this technology will be made in Phase II, in conjunction with a probable rival in the shape of Word Processing (Power Typing). The use of special typewriters with magnetic tape attachments, or links to a central controller with disk storage, provides two benefits. The editing and multi-copying feature is already