most cases be unsought terms as far as the chain index was concerned. This matter will be investigated in the test programme, but if it had been decided not to prepare chain index entry for any of these terms the number of cards inserted would have been reduced from 18,083 (as shown above) to 16,475. On the other hand, as discussed in Chapter 4, we did not do full chain indexing up to the containing head. If this had been done, it is estimated that the number of inserted cards would have risen to over 30,000.

CHAPTER 7

SUPPLEMENTARY INDEXING

For the testing programme we estimated that we should require 1600 questions which were to be based on documents in the collection, and we sought the co-operation of a large number of organisations and individuals. Discussion of the questions will be given in a later report dealing with the test programme, and it is sufficient to say here that we obtained the required number of questions.

We decided that at the same time we would invite people to index documents for us. The original purpose for doing this was so that we might have a comparison with which the standard of indexing of the project staff could be checked. The following is an extract from the notice which we sent to those whom we thought might be interested enough to help.

> "Assistance of two kinds is requested from individuals or organisations who are interested in this project. Firstly, we wish to have a number of the project documents indexed by persons or groups of persons who are familiar with the subject content,

and/or the particular indexing system used. Arising from such indexing, entries would be made in the appropriate project catalogues and should help to provide a check on the indexing done by the project staff.

"It would be expected that the selected indexing system would be that used by the individual or organisation, but should anyone wish to try another, or all the systems, there is no objection to this. To assist with the indexing we will provide the following items:

(1) Universal Decimal Classification

"An alphabetical index, compiled by the project staff, of all terms used in the indexing of the first 10,000 documents, and copies of any appropriate schedules that are not generally available.

(2) Alphabetical Subject

"A list of the main headings, sub-headings and all cross references originated during the indexing of the first 10,000 documents together with the project rules for the formation of new headings.

(3) Facet

"A set of the special schedules, the alphabetical index and the rules for using the schedules.

(4) Uniterm

"A list of the terms originated in the indexing of the first 10,000 documents."

"It would be most satisfactory if those who are willing to co-operate would agree to index twenty papers or multiples of twenty.

"Approximately half of the papers to be indexed will deal with high

speed aerodynamics while the remainder will range over a broad field of subjects that are relevant to aeronautical engineering. The papers will be in technical periodicals or reports from various research establishments. On the reverse side of the accompanying form is a list of the periodicals and reports series which have been used in the project and those co-operating are asked to indicate which periodicals or report series are readily available to them. Copies of the documents to be indexed can, if necessary, be made available in England from the College of Aeronautics Library, while aeronautical librarians in other countries have also offered to lend the papers to anyone co-operating."

The response which we received exceeded all reasonable expectations, and to those who agreed to assist, we sent the relevant items as mentioned in the above extract, together with instructions regarding the methods to be used in the indexing. Selected lists of documents were prepared from which the indexers could select any document they wished. These lists were set out in such a way that we had external indexing for documents in all groups throughout the programme. In Appendix A is given a complete group of documents for the project numbers 14001 - 14100, and Appendix F is a copy of the documents extracted from this group from which the indexer was able to make his own selection. Each indexer received twenty such sheets, and if they agreed to index twenty documents, they were asked to select one document from each sheet. If indexing forty documents they would select two from each sheet, and so on.

Altogether 3793 items were indexed, the indexing being fairly equally spread over all four systems. We shall not only in our test programme be able to check our indexing against that done by many other people, but

in addition it seems that we might also be able to obtain some quite interesting information by analysing the different methods which have been used to index the same report by people of widely differing qualifications and interests. A list of the organisations and individuals who helped with this indexing is given in Appendix G. Table 10 gives data on the indexing done by outside organisations in regard to the number of postings and the times taken for the four systems. From a comparison with the indexing done by the project staff it will be seen that the average time considered necessary to index the type of document used in the project, fell reasonably near to the middle of the various fixed times which we used, and it therefore appears that our range from 16 minutes to 2 minutes covered the conditions very adequately. The average entries requested by external indexers are, in the main, in line with project indexing, the only significant difference being with facet, where we thought it necessary to make a greater number of entries than did external indexers. It is interesting to note, as can be seen from Table 5B, that in the early stages the facet indexing of the project staff was in line with that done by external indexers, and that only with greater experience of using the system did they come to the opinion that additional entries were more often required.

Where the same document was indexed by three or more organisations, we have recorded in Table 11 information concerning the duplication of notational elements or headings. Column 1 gives the document number and Column 2 shows the number of elements and headings used by each indexer. These are totalled in Column 3, while Column 4 shows the number of different elements or headings used by all the indexers. This number can never be less than the number used

by the indexers using most elements or headings, so in Columns 5 and 6 we show the result when this number is deducted from the figures in Columns 3 and 4. The first two columns therefore represent the possible and the actual number of different notations or headings that were used.

It appears to be a reasonable assumption that in an ideal system the same document would have the same notation or heading whoever may do the indexing. The nearer this consistency is approached in the index, the greater the chance that the searcher will locate the required information in the expected place. One might assume, therefore, that the lower the ratio of Column 6 to Column 5, the more likely will the system prove to be successful. If we take the results given in Table 12 at their face value, it would seem that the systems come out, in descending order of efficiency as :-

Uniterm	21%
Facet	32%
U.D.C.	52%
Alphabetical	78%

There are, however, various factors which might modify the basic figures, and a fuller analysis of this matter will be considered in conjunction with the test results.

	No. of Documents indexed		Average Time	Average Entries	Average Elements	
U.D.C.	1,115		10.8 mins.	3.6	3.2	
		Project	12 mins.	4.9	3.5	
			8 mins.	3.9	3.1	
FACET	902		11.3 mins.	1.1	5.0	
		Project	12 mins.	1.8	6.9	
			8 mins.	1.5	5.9	
UNITERM	775		10.8 mins.		12	
		Project	12 mins.		10.2	
			8 mins.		9.2	
					Main Headings	Sub Headings
ALPHABETICAL	1,001		8.0 mins.	3.6	2.1	1.6
		Project	12 mins.	3.5	2.6	2.2
			8 mins.	3.2	2.3	2.0

TABLE 10 DATA ON EXTERNAL INDEXING WITH COMPARATIVE PROJECT INDEXING FOR 12 MINUTES AND 8 MINUTES

1	2							3	4	5	6
Document No.	External Indexing							Total Elements	Different Elements	Total Less 2A	Difference Less 2A
	A	B	C	D	E	F	G				
420	5	5	4	4	3	3		24	10	19	5
6082	5	3	3					11	10	6	5
6846	4	3	3					10	9	6	5
6905	2	2	1					5	2	3	0
7606	4	4	1	1	1			11	9	7	5
8003	4	2	2	1	1	1		11	8	7	4
8011	7	3	2	1				13	11	6	4
8020	6	3	1					10	8	4	2
8858	5	4	3	3	2			17	12	12	7
9651	8	2	2	1	1			14	11	6	3
9889	3	3	2					8	8	5	5
10819	5	3	2					10	7	5	2
11696	4	4	3					11	8	7	4
12018	5	4	4	3	3			19	6	14	1
12063	3	2	2					7	4	4	1
12086	3	3	3	1	1			11	8	8	5
12093	8	4	4	3				19	11	11	3
12095	3	3	2	2				10	6	7	3
12428	5	4	4	4	3	2	2	24	17	19	12
12436	4	4	3					11	5	7	1
12462	4	1	1					6	6	2	2
12465	7	3	3	3	1	1	1	19	12	12	5
12482	7	5	3					16	12	9	5
12654	7	4	3					15	14	8	7
12810	7	6	3					16	11	9	4
12831	3	2	2	2				9	7	6	3
12870	8	5	3					16	14	8	6
12895	3	3	1	1				8	8	5	5
12910	4	4	3					11	8	7	4
12979	6	3	2					11	8	5	2
						Totals		385	270	234	120

TABLE 11A ANALYSIS OF U.D.C. INDEXING OF 30 DOCUMENTS INDEXED BY THREE OR MORE ORGANISATIONS

1	2						3	4	5	6
Document No.	External Indexing						Total Postings	Different Headings	Total Less 2A	Difference Less 2A
	A	B	C	D	E	F				
565	4	3	2				9	5	5	1
1771	3	3	2	1			9	8	6	5
4487	9	6	4				19	16	10	7
5779	6	5	1	1			13	11	7	5
5781	8	5	3				16	16	8	8
5785	11	10	1				22	20	11	9
6105	2	2	1				5	2	3	0
6905	3	3	3	2			11	9	8	6
7782	7	6	6	2	1		22	15	15	8
8104	5	3	2				10	10	5	5
8114	6	2	1				9	9	3	3
8995	4	3	1				8	6	4	2
9759	3	2	1				6	5	3	2
12143	5	3	1				9	8	4	3
12163	3	3	2				8	7	5	4
12186	5	2	1				8	6	3	1
12190	10	9	3				22	22	12	12
12514	2	2	1				5	5	3	3
12576	6	1	1				8	8	2	2
12671	6	5	3	3	2		19	17	13	11
12910	16	6	5				27	25	11	9
12918	3	1	1				5	3	2	0
12939	6	4	3	2	1		16	14	10	8
12979	7	6	1				14	12	7	5
13264	8	4	2				14	13	6	5
13321	4	3	1				8	8	4	4
13363	3	3	2				8	7	5	4
13367	3	2	1				6	6	3	3
13369	4	2	2				8	8	4	4
13763	15	4	4	3			26	25	11	10
						Totals	370	326	193	149

TABLE 11B ANALYSIS OF ALPHABETICAL INDEXING OF 30 DOCUMENTS INDEXED BY THREE OR MORE ORGANISATIONS

1	2							3	4	5	6
Document No.	External Indexing							Total Elements	Different Elements	Total Less 2A	Difference Less 2A
	A	B	C	D	E	F	G				
670	9	6	2					17	15	8	6
675	4	3	3					10	5	6	1
685	9	3	3					15	9	6	0
1607	4	4	3					11	6	7	2
4607	6	6	5	4	3	3	2	29	13	23	7
4678	8	3	2	1				14	11	6	3
4681	8	5	5	4	1			23	15	15	7
5819	4	4	4	3				15	7	11	3
6231	4	4	3					11	5	7	1
6249	5	5	2	2	2	1	1	19	7	14	2
7028	8	5	4	3				20	14	12	6
7096	8	8	6					22	19	14	11
8221	7	3	3	3	3			19	10	12	3
9016	6	4	3					13	7	7	1
9854	10	5	4	3	3	2		27	13	17	3
9863	7	4	3					14	8	7	1
9874	5	5	3	3	2	1		19	10	14	5
10245	11	4	4	3				22	12	11	1
11037	7	3	1					11	7	4	0
12095	4	2	2					8	6	4	2
12257	7	5	4	4	4			24	15	17	8
12270	6	4	2					12	7	6	1
12619	7	6	6	6				25	13	18	6
12654	6	4	4	3	3			20	14	14	8
12667	4	4	3	2				13	5	9	1
12671	4	4	3	3				14	11	10	7
12675	4	3	3	3	3	2	2	20	4	16	0
13073	5	5	3					13	7	8	2
13074	5	3	3					11	7	6	2
13080	4	1	1	1				7	4	3	0
					Totals			498	286	312	100

TABLE 11C ANALYSIS OF FACET INDEXING OF 30 DOCUMENTS INDEX BY THREE OR MORE ORGANISATIONS

1	2					3	4	5	6
Document No.	External Indexing					Total Terms	Different Terms	Total Less 2A	Difference Less 2A
	A	B	C	D	E				
670	14	6	4			24	15	10	1
1915	14	11	10	4		39	15	25	1
4487	19	7	7			33	24	14	5
5929	50	17	12	12		91	55	41	5
7028	27	21	3			51	34	24	7
8221	15	13	7			35	20	20	5
9148	16	15	10	8		49	25	33	9
10121	15	10	7			32	19	17	4
12671	18	11	9			38	24	20	6
13153	14	12	6	6		38	16	24	2
13156	13	8	4			25	14	12	1
13451	20	9	7			36	25	16	5
13987	11	9	6			20	14	15	3
14384	67	28	15			113	77	46	10
14565	15	9	4			28	18	13	3
14758	13	10	10			33	14	20	1
15058	15	9	8			32	19	18	4
15079	20	9	8			37	23	17	3
15504	20	15	10			45	33	25	13
15506	20	17	10	5		52	23	32	3
15553	15	12	11	9	6	53	18	38	3
15590	13	9	6			28	20	15	7
16395	21	8	7			36	28	15	7
16678	12	9	7			29	16	17	4
16763	26	17	10			53	29	27	3
16985	9	8	6			23	11	14	2
16988	8	7	6			21	10	13	2
17107	15	11	9			35	20	20	5
17462	24	20	10			54	36	30	12
17983	9	4	4			17	11	8	2
					Totals	1,206	706	639	138

TABLE 11D ANALYSIS OF UNITERM INDEXING OF 30 DOCUMENTS INDEXED BY THREE OR MORE ORGANISATIONS