AN APPLICATION OF INFORMATION RETRIEVAL AND ANALYSIS IN EGYPTOLOGY

N. Strudwick University of Liverpool

By way of introduction to this paper, it may be of interest to summarise the ways in which the computer has already been employed in the field of Egyptology.

Applications may be divided into three broad areas: text processing, documentation of objects, and methods of drawing hieroglyphic signs. A general discussion of the uses of the computer in the subject may be found in the recent article by Crozier-Brelot (1981).

1. Text Processing: This is the oldest and widest application in the field. The pioneering project is Maschinelle Analyse AltÄgyptische Texte (M.A.A.T.), begun in Germany in the early 1960's by Gundlach and Schenkel (Crozier-Brelot 1981, 51). This system dissects Egyptian Texts input in transliteration, not only producing a concordance of words, but also analysing the grammatical structure of the text. As yet, only one major textual analysis has been published (Gundlach & Schenkel 1970). This is a listing of all words and forms appearing in a spell from the corpus of Ancient Egyptian Funerary Literature, and has met with some hostility among Egyptologists (Mueller 1972, Faulkner 1973), some of which criticism, it must be admitted, is justified. The book is produced using reduced copies of the original computer output which, because of the system of transliteration and the small size, renders its consultation uncomfortable. A similar lexical/grammatical list of negative sentence patterns was made by Davis (1973; discussion of method on pp. 203-17), but here the system of transliteration developed for the computer has been manually retranscribed after analysis into forms recognisable to Egyptologists. Most other work of this type has also been in connection with concordances of texts. Computerised bibliographies are almost non-existant in this subject.
2. Documentation of Objects: This implementation is perhaps the most widely spread, and primarily concerns museum collections (Crozier-Brelot 1981, 49-50). One rather unusual use in connection with objects is the project to reassemble blocks from a destroyed temple with the

aid of a computer (Smith & Redford 1976).

3. Drawing of Hieroglyphs: This is the application which most catches the imagination of the non-specialist. However, the relatively small number of signs that can be usefully recorded in comparison with the almost infinite range of variant signs in hieroglyphic script, somewhat restricts its usefulness (Crozier-Brelot 1981, 50).

Of these applications, that concerning objects will potentially be of most use to the field. The other applications have been very limited, and few (if any) attempts have been made to establish systematic data-sets in specific branches of the subject which may be made freely available to relevant institutions around the world. It is hoped that it may prove possible to do this in the future with the present project. The computer is here used in conjunction with research for the thesis, 'The Administration of the Memphite Region of Egypt in the Old Kingdom (c. 2600-2150 B.C.)'. Such a topic is centered around the study of titles of contemporary officials, their functions, when the titles appeared and disappeared, and the nature of the other titles with which they are found. The ultimate aim is to shed light on the government and institutions of the time. This material is culled primarily from the available archaeological publications, with additional material coming from unpublished sources both inside and outside Egypt. This material is thus not the result of any one systematic excavation.

It was decided to use the computer because of the size of the data to be analysed - currently this consists of approximately 1600 officials with 9000 titles. For each official is also recorded a reference number, his provenance, a bibliographical reference, and a selection of possible dates; this amounts to a very large quantity of mainly textual data.

The obvious problem of representing this material on the computer is that of the script. The method used by all workers in the field where the actual forms of the written hieroglyphs are not important is to use the conventional transliteration system, which employs the following alphabet:

ꜣ ỉ ꜥ w b p f m n r h ḥ ḫ ẖ z s š q k g t ṯ d ḏ

Apart from signs foreign to roman script, there are several accents, all of which fall beyond the range of the normally available typeface. Three solutions are possible: one may define a new character set, and this is now feasible with the use of sophisticated graphics terminals. Secondly, the awkward characters may be replaced by others not used in the above alphabet - obvious substitutions are 3 for ꜣ , ' for ꜥ , i for ỉ - and the

accented letters may be replaced by others according to a convention, such as X for ḫ, and so on. Thirdly, one may make the obvious substitutions as above and then ignore the differences between the accented and non-accented versions of the same basic letter.

The last solution is currently used here. As it is for the use (for the present) of one person only, familiarity with the vocabulary and grammar of the language makes it obvious in most cases as to which of the possibilities is intended; it also restricts the selection of letters to those in the above alphabet and avoids the need to learn a new system solely for the benefit of the computer (the German system M.A.A.T. uses the second of the three methods and is not particularly pleasing to the egyptological eye). One drawback of keeping the letter selection as simple as possible is that the data is consequently sorted in roman and not egyptian alphabetical order, although this presents no problem in practice. It is not felt worth the trouble of obtaining sorting in egyptian order for the current purposes. The data for the thesis uses modern words (for example in provenances and dates) as well as hieroglyphic transliteration, and so the former is set in mixed case, the latter in upper. Many of the more common titles have been abbreviated for simplicity of input; such abbreviations occasionally employ extra roman characters.

The major operations performed on the data are not complex. The programs in use are written in the language ALGOL 68, which gives complete control over these operations, handles text well, allows complex data-structures, is efficient on space with its ability to work on a reference level, and encourages modern structured programming techniques. The programs take their data input from a variable number of files, thus keeping the average size of data-sets low, and are able to take control instructions from either a terminal or a file. Every effort has been made to arrange the output in a clear and readable form so as to facilitate consultation. The only moderately complex feature is the data-structure, as it has to cope with a varying number of officials, each with a minimum of one and a maximum of ninety titles. Other problems have been mainly with the external linkage, notably in relation to the size of input and output files, and the amount of time and core memory necessary for a successful run of the larger programs.

Development was begun on the University of Liverpool's ICL 1906S computer, but has now been transferred to a new IBM 4341 machine (VM/CMS), which allows development and analysis to be carried out online, with the

added advantage of upper and lower case alphabets.

Numerous types of analysis are possible, but in practice the following have proved most relevant to the current research:

1. Listings of the data in numerical (or other) order of the officials (fig. 1).
2. Listings of titles in alphabetical order with the names of holders.
3. Concordances of words present in titles, together with the original title and its holder, with cross-references where necessary (fig. 2).
4. Lists of Bibliography and Dates.
5. Presence-absence charts for particular titles (fig. 3).

The latter analysis is the most complex, and, once the relevant officials have been sorted into their chronological order, illustrates how the other titles (represented by the letters A B C ... at the top) held by these officials changed over a period of time. All of these analyses may of course be applied to any set of officials selected on some relevant criterion.

Conclusions:

The advantages of a computer in respect of time, accuracy and legibility are well known. It is of immense value to work with a clear printout of material rather than to consult continually a vast number of semi-legible slips. Manual construction of such a presence-absence chart as shown in fig. 3 may take the best part of a day; the computer will produce the same (without errors) in as little as three minutes.

It will be clear that the computer-generated output is not the analysis central to the thesis (as may be the case in numeric disciplines and certain archaeological applications), but rather an aid to the more efficient realisation of conclusions, which is no less significant.

The production of large data-sets on a computer as done here raises the possibility of their transfer simply and cheaply between centres with similar interests. Research would be greatly expedited if one did not have to spend at least a year collecting the basic information by a laborious search through publications when much of it may duplicate an effort previously made elsewhere. The raw material of certain subjects (such as the study of titles) lends itself very easily to representation on a computer; if a tape or disc of this could be passed on to an individual with kindred interests, he would only have to alter the details to conform with his own particular requirements.

Rather than creating data specific to one's requirements, archaeology

relies on the use of the same material produced by others in many different ways. To have as much of it as possible conveniently available would remove much of the drudgery from research, leaving one free to concentrate on those aspects of one's work where the computer cannot (yet) match the abilities of the researcher.

School of Archaeology
and Oriental Studies,
University of Liverpool.

Nigel Strudwick

References:

Crozier-Brelot, C. 1981 — 'Égyptologie et informatique'. CHRONIQUE D'EGYPTE 56: 49-54.

Davis, V.L. 1973 — SYNTAX OF THE NEGATIVE PARTICLES bw AND bn IN LATE EGYPTIAN. Deutscher Kunstverlag, Munich.

Gundlach, R. & Schenkel, W. 1970 — LEXICALISCH-GRAMMATISCHE LISTE ZU SPRUCHE 335a DER ALTÄGYPTISCHE SARGTEXTE LL/CT. 335A. Deutsches Rechenzentrum, Darmstadt.

Faulkner, R.O. 1973 — Review of Gundlach & Schenkel 1970. JOURNAL OF EGYPTIAN ARCHAEOLOGY 59: 237.

Mueller, D. 1972 — Review of Gundlach & Schenkel 1970. BIBLIOTHECA ORIENTALIS 29: 300-1.

Smith, R.W. & Redford, D.B. 1976 — THE AKHENATEN TEMPLE PROJECT. Aris & Phillips, Warminster.

Non-obvious Abbreviations used in headings of Figures.

Date-1, -2, -3	Alternative possible dates.
Dif Tits	Different Titles.
PM3	Reference to B. Porter & R.L.B. Moss, TOPOGRAPHICAL BIBLIOGRAPHY OF ANCIENT EGYPTIAN HIEROGLYPHIC TEXTS, RELIEFS AND PAINTINGS, volume III, second edition. Oxford University Press, 1974-82.
Ref	An integer in the range 1 to 60 corresponding approximately to the officials's relative position in Old Kingdom Chronology, to facilitate ordering officials in correct sequence by date.

Number	Name	Site	PN3	Date-1	Date-2	Date-3	Ref	Titles
5	3HI	Giza	137	B.Henk	K.L Henk	Me.Snd 9L	12	HRP NSTY HST IR KJT NBT NT NZWT IR PRWY-HD IR SNWT NZWT WMS ZAH
19	3HT-HTP	Saqqara	549	B.5D	K.M Wen	Me.EM Wen	34	HN M3'T HRP WSHT HRP ZS HRP ZS NB HTN IK IR GP IR NIWT Pyr Djed IR NIWT Pyr Menkaunor IR NIWT Pyr New IR PRWY-HD IR S3'W IR SNWTY IR WD'-MDW NB IR WSHT IR ZAH MR NST HNTT SHD HN Pyr Djed SHD HN Pyr Menkaunor SHD HN Pyr New SW TZT WMS ZAH ZAH M3'
22	3HT-HTP;HMI	Saqqara	627	B.5b	K.M Wen	Me.M Wen	34	AH DHRP HA HH HHHT HST HSTPD HTNB IM3-' INH IP IR GP IR PRWY-HD
...	22 3HT-HTP;HMI							

Fig. 1

Word	Original Title	Number	Name	Site	PMJ	Rat	Word Summary Occurs	Dif Tits	Word
'	IR '	P 7	INI-ITIF;MHW	Aswan	0	49	1	1	'
'*	See under relevant part of HRIW-'								
'H	HRP 'H	P 10	INTI	Deshasha	0	47			
	IR 'H	2	3d-NB	Saqqara	741	1	2	2	'H
'HJW	IR PR 'H3W	11	3HT-MHW	Giza	87	34	1	1	'HJW
'W	IR 'W	1	3BBI	Saqqara	768	46	1	1	'W
'3	SD3WTY NTR M IMW2 '3	P 8	INI-K3F	Zawayda	0	57	1	1	'3
'3*	See under relevant part of R-'3-G3W								
'3T*	See under relevant part of HWT-'3T								
B3WY	HN B3WY HNTY PR HW	6	3HI	Saqqara	690	57	1	1	B3WY
Djed	IR NIWT Pyr Djed	19	3HT-HTP		599	34			
	SHD HN Pyr Djed	19	3HT-HTP		599	34	2	2	Djed
DD	DD NRW HR M B3SWT	P 8	INI-K3F	Zawayda	0	57			
		P 9	INI-KJF		0	57	2	1	DD
Goat	SSM-T3 NIWTY N Goat	P 15	ITTI;SDW	Deshasha	0	53	1	1	Goat
GP	IR GP	19	3HT-HTP	Saqqara	599	34	1	1	GP
G3W*	See under relevant part of R-'3-G3W								
H-T	H-T MDW N WD'-HDW ST3 N HWT-WRT	11	3HT-MHW	Giza	87	34			
	H-T MDW ST3	11	3HT-MHW		87	34			
	H-T MDW ST3 N HWT-WRT	11	3HT-MHW		87	34	3	3	H-T
HA*	2 examples								
HFT-HR	ZAN HFT-HR	P 14	ITI	Z e Mayitin	0	57	1	1	HFT-HR
HH*	4 examples								
HKR	INN HKR NZWT M HJSWT RSYT	P 8	INI-K3F	Zawayda	0	57	1	1	HKR
HMWT	IR HMWT	3A	3BDW	Giza	1	99	1	1	HMWT
HN	HN B3WY HNTY PR HW	6	3HI	Saqqara	690	57			
	HN Khufu	12	3HT-HTP	Giza	49	55			
	HN M3'T	11	3HT-MHW		87	34			
		19	3HT-HTP	Saqqara	599	34			
	IR HN	P 11	IaY	Z e Mayitin	0	57			
		P 6	INI-ITIF	Thebes	0	58			
	SHD HN HWT-K3 Pepy	P 14	ITI	Z e Mayitin	0	57			
	SHD HN M3' HWT-K3 Pepy	P 14	ITI		0	57			
	SHD HN Pyr Djed	19	3HT-HTP	Saqqara	599	34			
	SHD HN Pyr Menkaunor	19	3HT-HTP		599	34			
	SHD HN Pyr New	19	3HT-HTP		599	34	11	9	HN
		19	3HT-HTP		599	34			
HNTT	NSI HNTT	P 12	IHY	Thebes	0	46			
		12	3HT-HTP	Giza	49	55			
		12	3HT-HTP		49	55	4	1	HNTT
HNTY	HN B3WY HNTY PR HW	6	3HI	Saqqara	690	57	1	1	HNTY
HQH*	5 examples								
HR	DD NRW HR M H3SWT	P 8	INI-K3F	Zawayda	0	57			
		P 9	INI-K3F		0	57	2	1	HR
HR*	See under relevant part of HFT-HR								

Fig. 2

23/MAR/82
Presence-Absence Chart

Officials holding title of "Overseer of Royal Document Scribes"

Part One Page 1

		1																								2																								3			
Number	Name	A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
251	W3S-PTH;IZI	.		.		X	X	.		.		.	X	.		X		.	X	X	X	X		.		.		.		X		.		X		X		.		X		.		.		.		.		.		.	
1018	K3I (non vizier)	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		.		X		.		X	X	.		.		.		.		.		.	
1018	K3I (vizier)	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.	X	X		X		2	X	X		.		.		.	X	.		1		.	
960	SSM-NPR II	.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		X		.		X		.		.		.		.		.		H	
318	PH.N-WI-K3I	.		.		X		X		X		.		.		X		.		X	X	.		X		.		.		.		.		X		X	2	2		X		.	1	.		.		.		.		.	
934	SHM-'NH-PTH	.		.		X	X	X		X		.		.		X		.		.		.		.		.		.		.		.		X		X		.		X		.		.		.		.		.		H	
858	ZTW	.	X	.		X	X	X		X		.		.		.		.		X	X	.		.		.		.		.		.		.		.		.		X		.		.		.		.		.		.	
1135	IY	.		.		.	X	.		.		.		.		.		.		.		X		.	X	.	X	.		X		.		.		X		.		X		.		.		.	X	.		1		N	
634	NTR-WSR	.		.		.		.		.		.		X		X		.		.		.		.		.		.		.		.		.		X		.		X		.		.		.		.		.		.	
336	PTH-HTP DSR	.		.	X	X		.		.		.		.		.		.		.		.		.		.		.		.		.		X		X		.		X		.		.	X	X		.		.		.	
337	PTH-HTP	.		.	X	X		.		.		.		.		.		.		.		.		.		.		.		.		.		X		X	2	2		X		.		.	X	X		.		.		.	
961	SSM-NPR III (n-vz)	.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		X		.		.		.		X		.		N	
669	R'-SPSS	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.	X	X		X	2	.		X		.		.	X	.		.		.		.	
339	PTH-HTP	.		.	X	X	X	X		X		.	X	.		.		.		.		X		.		.		.		.		.		X		X	2	2	X	X		.		.	X	.	X	.		.		N	
766	H'-MRR-PTH	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		.		2		.		X		.	1	.		.		.		.		.	
1037	K3I-PW-INPW	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		.		.		.		X		.		.		.		.		.		.	
919	SNDM-IB;INTY	.		.	X	X	X	.		.		.		.		X		.		.		.		.		.		.		.		.		X		X	2	2		X		.		.		.	X	X		.		.	
363	M3-NFR	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		.		.		.		X		.		.	X	.		.		.		.	
936	SHM-'NH-PTH	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		.		.	2	.		X		.		.		.		.		.		.	
19	3HT-HTP	.		.		.	X	.		.		.		.		X		.		.		.		.		.		.		.		.	X	X	X	.	2	2		X		.	1	.		.		.		.		.	
22	3HT-HTP;HMI	.		.	X	X	X	X	X	.		.		.		.		.		X	X	X		.	X	.		.		X		.		X		.	2	2	X	X		X	1	X		.		.		.		H	
920	SNDM-IB;MHI	.		.	X	X	X	.		.		.		.		X		.		.		.		.		.		.		.		.		X		X	2	.	X	X		.		.		.	X	X		.		N	
52	IY-NFRT;S3NF	.		.		X	X	.		.		.		.		X		.	X	X	X	.		.		.		.		.		.		X		.		2		X		.		.		.		.		.		H	
128	IHY	.		.	X	X	X	.		.		.		.		X		.		.		.		.		.		.		.		.		X		X	2	2		X		.	1	.		.		.		.		.	
836	HNMW-NTI	.		.	X	X	X	.		.		.		.		X		.		.		.		.		.		.		.		.		X		X	2	2	X	X		.		.		.		.		.		N	
1103	K3I-GM.NI;MMI	.		.	X	X	X	X	X	X		.		.		.		.	X	X	X	X		.	X	.		.		X	X	.	X	X	X	X	2	2	X	X	X	.		.	X	X		.		1		H	X
602	NFR-SSM-R';SSI	.		.	X	X	X	.		.		.		.		X		.		.		.		.		.		.		.		.	X	X		X	2	2	X	X	X	.	1	.	X	X		.	X	1		.	
409	MRRW-K3I;MRI	.		.	X	X	X	X	X	X	X	.	X	.		.		.	X	X	X	X	X	X	X	X		X	X	.	X	X		X		X	2	2	X	X		.		X		X	X	X	X	.		2	X
809	HNTY-K3I;IHHI	.		.	X	X	X	X	X	.		.		.		.	X	.		.		.	X	.		.		.		.		X		X	X	X	2	2	X	X		X		X		X		.		X		N	
417	MHW	.		.	X	X	X	X	X	.		.		.		X		.		.		X		.		.		X		X	X	.	X	X		.	2	2	X	X	X	.	X	.		X		.		.		H	X
343	PTH-HTP	.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		X		.	2	2		X		.		.		.	X	.		.		.	
395	MRY-TTI	X		.	X	X	X	X	X	X	X	.	X	.		.		.		X	X	X	X	.	X	.		X		.	X	X		X		X		.		X		.		.		X		.		.	X	N	X
171	IDW I;NPR	.		.		.		.		.		.		.		X		.		.		.		.		.		.		.		.		X		X	2	2		X		.	2	.		X		.		.		N	
404	MRRI	.		.	X	X	X	X	X	.		.		.		.		.		.		.		.		.		.		.		.	X	X	X	X	2	2		X		.		X		X		.		.		.	
605	NFR-SSM-SS3T	.	X	.		X	X	X		.		.		.		X		.		.		.		.		.		.		.		.		X		X		.		X		.		.		X		.		.		.	
221	'NH-MRY-R'	.		.		X	X	X		.	X	.	X	.		.		.		.		.		.		.		.		.		.		X		X	2	2		X		.	1	.		.		.		.		.	
820	HNMW;HNTI	.		.		X	X	.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		.		X		.		.		.		.		.		.	
477	NY-HRSD-NFRK3R'	.		.	X	X	X	X		.		.		.		.		.		.		.		.		.		.		.		.		X	X	.		2		X		.	X	.		.		.		.		.	
1128	TTI	X		.	X	X	X	X		.		.	X	.		.		.		.		.		X		.		.	X	.		.		X	X	.	2	2		X		.	X	.		.		.		.		.	X

Fig. 3