<<<Back to Table of Contents
On to Part D: Familiarization with IT applications to support the workgroup>>>

Geoscience after IT: Part C

Familiarization with IT applications to support the individual geoscientist

T. V. Loudon
British Geological Survey, West Mains Road, Edinburgh EH9 3LA, U.K.
e-mail: v.loudon@bgs.ac.uk

Postprint of article in Computers & Geosciences, 26 (3A) April 2000, pp. A15-A19

Abstract - Familiarization with IT proceeds best by first learning the basic skills, using them, and considering the consequences; thus developing a mindset to take advantage of future opportunities. A desktop computer is a good starting point. Word processing can be seen as a route from tangled thought to immaculate presentation. Spreadsheets can help to build, analyze and plot datasets. Data can be collected with forms and spreadsheets, with on-line instruments in the laboratory or in the field, or by scanning and maybe OCR. Standards are needed to communicate between systems.

Key Words - Desktop computer, word processing, spreadsheets, data capture.

1. The route to IT familiarization

Not all geoscientists are familiar with available methods of IT, although these influence all phases of a project and every type of information. This review looks for underlying principles, looking at individual support in this part, moving on to project, then global requirements (Parts D - H). It tracks the process of familiarization, from ubiquitous tasks like word processing through statistical analysis, spatial analysis and computer visualization to the management of databases and repositories.

You learn about computers by using them, just as you get to know a town by living in it. A guidebook can help by pointing out features you might otherwise miss, and explaining the background to improve your understanding and give an overview to tie it all together. I see no reason to repeat sections of computing manuals here, when you can readily find the real thing, often as online help and demonstrations of software on your computer. Instead, I offer suggestions about what to do and where to find things - a means to an end, not an end in itself.

This guidebook also has a theme: that information technology is changing the way we conduct geoscience, and to control that change we need to understand, not just the technology but also the way our science works. As I am unsure of what you already know, I start with basic concepts, and hope you will skip ahead if they offer nothing new. A search of the World Wide Web will provide details of many of the topics mentioned.

Many geoscientists have a good basic knowledge of computing, and add to it as required. The advantage of learning by experience is that the information may be more relevant and so more easily remembered. The disadvantage is that broader issues may be forgotten and better methods neglected in favor of the familiar, but periodic review of similar applications elsewhere helps to overcome bad habits. Formal instruction through books and courses should lead to a wider appreciation of the possibilities, and to learning the best approach rather than one that was stumbled on by accident.

Those who are less familiar with the basics of IT or its geoscience applications have different priorities. They need to gain the basic skills of operating with keyboard and mouse. They must learn how simple tasks are handled on the computer before they are in a position to learn about more complex applications, where these skills are taken for granted. A step-by-step approach where skills are learned and immediately put to good use is strongly recommended. Word processing is one skill which is at once useful and also leads to familiarity with the keyboard, mouse and screen, and the idiosyncrasies of the graphical user interface (GUI). Handling electronic mail (e-mail), on-line public access catalogs (OPACs), and surfing the Web are then less daunting.

2. Desktop hardware

A good starting point for learning about computing applications is your own desktop computer, where skills can be learned by experiment, making mistakes in the privacy of your own office. Here is a place to learn of the user interface, word processing, spreadsheets, database and communication features. Desktop computers also support many small geoscience applications that are helpful in individual investigations.

The computer screen on the desktop is a window into information technology. It may be part of a personal computer or a workstation or a simple network computer with little computing capacity of its own, but with the ability to download instructions from a remote server. It may be static or portable. It may be for your exclusive use or it may be one of many machines which get their individuality from downloaded software. Thus you may be able to go to one of many different computers, enter your identifier and password and operate it as your own machine.

If you would be seen as technically aware, you would know the make and model of your desktop machine and its main server, if any. You would know the name of its operating system and the company that produced it. You would know whether the computer is free-standing, that is, unconnected to others, or networked. If the latter, you would know what type of network it is connected to and by what means, how extensive the network is, and the bandwidth of the connection. The protocols for graphics on screen, the memory size, the processor speed, the disk capacity, and any peripheral attachments including the printing facilities (type, manufacturer, resolution, color availability, paper sizes), are additional points of interest. Some computing terms are reviewed in E. The details for your own machine can usually be discovered by searching through the menus on the machine, by looking at the manuals, or by asking a well-informed colleague.

Most of us have little say in selecting a computer at work. If we do, the important points are likely to be how it fits in with existing equipment and planned developments. At home, other considerations come into play, but compatibility with one’s employers' plans may still be crucial. For mainstream, general-purpose computers, you should be able, as when renting an automobile, to get in and drive, assuming, of course, some training and experience. If you plan to select and buy your own personal computer, there are many periodicals that offer advice. The vendors are likely to be biased, and it is well to precede any discussions with them with a little background research and to have in mind some figures for the details mentioned in the last paragraph. There is no point in offering further advice on purchase here, as it would be long out-of-date before publication. Remember, however, that a computer soon becomes obsolete, and the cost of purchase may have to be written off over as little as three years. Remember also that the main costs will lie in training and upkeep rather than the initial purchase.

Faced with an unfamiliar machine, progress can be made with a combination of experimentation, reading books and manuals, on-line help systems and demonstrations, advice from colleagues and vendors, and from training courses if they are readily available. Skills should be learned with a purpose in view, not studied and forgotten. An obvious starting point in computer use is word processing.

3. Word processors

Word processing skills are not difficult to acquire, and are a vital part of most geoscientists' work. Furthermore, the skills are fundamental to most other computer applications, and few of the present readers should be without them. A very small number of word processor packages dominate the market, all with similar features, and all suited to a geoscientist's needs. A choice between them might depend on the availability of local support, cost, their integration with other relevant packages, and perhaps the ease of handling mathematical or chemical formulas. They should be seen as part of the overall computing facility, not as a self-contained item.

There is no point in offering advice on their usage here, as this is best obtained from their on-line tutorials or manuals. However, we need to review their function, to see how it contributes to the information system. Think first of the earlier alternatives, which have shaped the information system as we know it. A familiar analogy is writing with pen and paper, starting at the top of the page and writing line by line until the page is full. Corrections can be made by striking out unwanted text and adding more wherever space is available. As the hand-written document can become rather illegible, the next step is to pass it to a professional typist who, by reading and reentering all the material, can prepare an equally inflexible document on a typewriter. Corrections can be marked by pen as before, at the expense of legibility, and retyping is inevitable for a document of any significance. Publication again requires reentry from a keyboard, with all the consequent possibilities for mistakes.

Word processing changes these procedures in several ways. The output from the computer printer is more legible than most handwriting. More fundamentally, ink is not put on paper until required. It is not fixed during the recording process. The record on the computer can be viewed on screen and altered at will. The repeated keyboard entry of the manual system is replaced by altering only what must be changed. Updated versions of a document, such as a list of references, can be prepared without reentering old material. The document need not be prepared in a set sequence, but can be built up from any starting points, and the contents rearranged whenever required. This is a major advantage for those whose thoughts seldom follow a straight line.

Content is separated from form of presentation, which can be altered separately. The computer record can generate a draft, a final document, or camera-ready copy for offset-litho printing. Parts of the text can be extracted for use in other contexts. With descriptions of, say, fossils or borehole records, items can be subdivided or expanded and automatically renumbered, without affecting the rest of the document. If the document is appropriately designed, tables of contents and indexes can be prepared and brought up to date automatically to match the current page numbering.

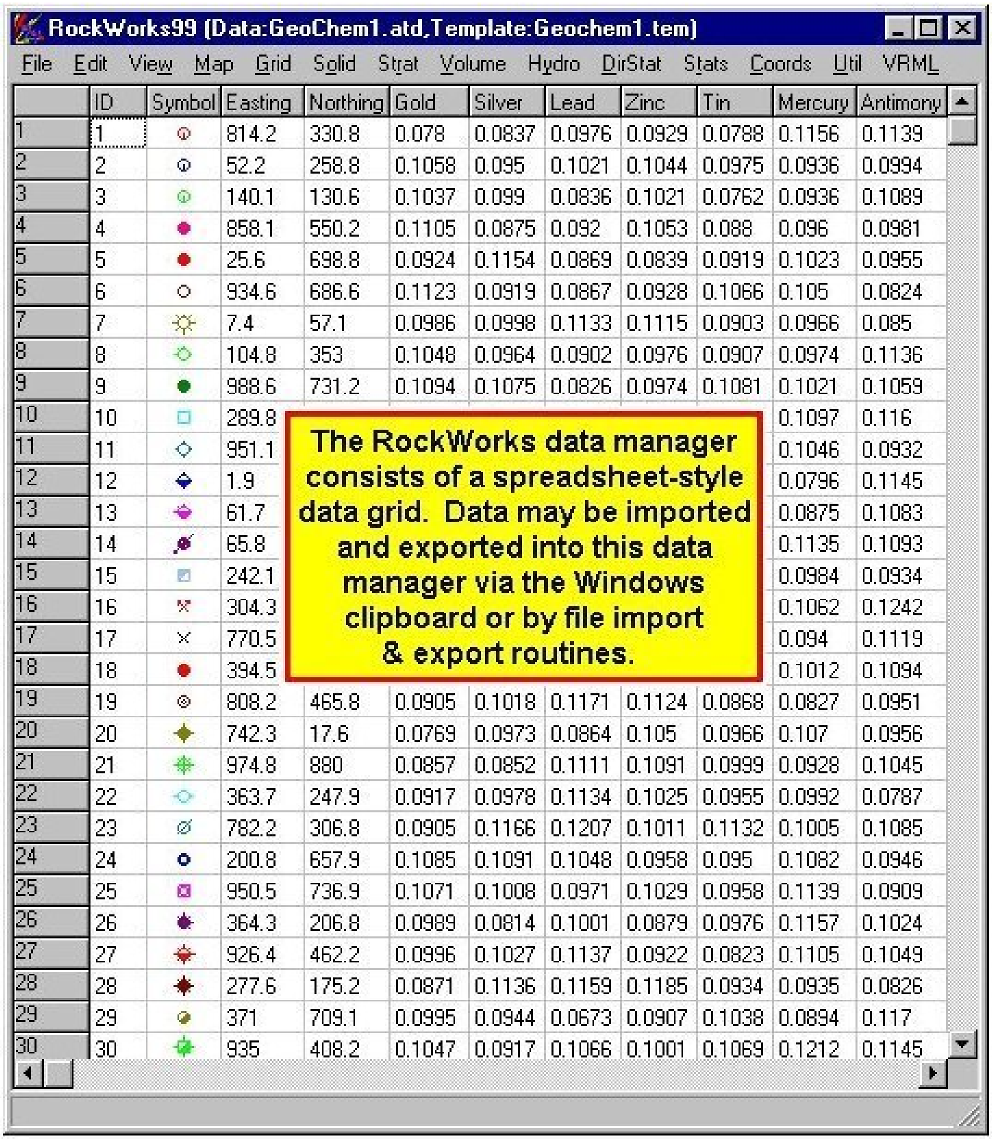
Typing skills, including high keyboard accuracy, good spelling, and the ability to visualize and plan the layout of the final product before it is typed, are less necessary. Scientists can record their own information, helped by tools such as a spelling checker, and adjust the layout as the document develops. This may be easier than preparing a hand-written draft for later typing. Although clear handwriting can be read mechanically, simple keyboard skills are not difficult to acquire and are likely to provide a better solution. Pen and paper remain more robust, versatile and cheaper in some circumstances. Despite the availability of rugged hand-held notebook computers, for example, it may be easier to use a conventional notebook in the field, and later decide what should be transferred to a computer record.

Word processing is a step on the way to the preparation of compound documents, which may incorporate data tables and diagrams as well as text (L 3). You can include placeholders to indicate where external information, perhaps selected from a database, is to be inserted later. You can then print circular letters or reports where the placeholders are replaced by information specific to the recipient, such as their name and address or paragraphs matching their interest profile.

The digital record can be accessed remotely, searched for keywords, and combined with images. Links can be inserted to other documents and to points in the same document. These need not refer to text, but can link to any electronic information, such as an image or data. They can be included in a compound document produced by a word processor, but require special facilities for manipulation and editing. Such concepts are greatly extended in the World Wide Web (E 4).

4. Spreadsheets and business graphics

Most desktop computers offer systems for preparing and handling spreadsheets, which are of value in many areas of geoscience. Because each step of the calculation is clearly visible, spreadsheets are of particular interest in learning, teaching and exploring new ways of analyzing data. The spreadsheet is a table, or array, of numbers arranged as rows and columns (see Fig. 1). The array can be large, but because the program offers simple procedures for entering and adjusting data, spreadsheets are also useful for small illustrative tables to insert in a report. New rows and columns can be created by adding new data or by inserting a formula stating how the new entries are to be calculated from the existing data. Because the program stores the formulas, the entire spreadsheet can be recalculated automatically when entries are added or amended.



	ID	Symbol	Easting	Northing	Gold	Silver	Lead	Zinc	Tin	Mercury	Antimony
1	1		814.2	330.8	0.078	0.0837	0.0976	0.0929	0.0788	0.1156	0.1139
2	2		52.2	258.8	0.1058	0.095	0.1021	0.1044	0.0975	0.0936	0.0994
3	3		140.1	130.6	0.1037	0.099	0.0836	0.1021	0.0762	0.0936	0.1089
4	4		858.1	550.2	0.1105	0.0875	0.092	0.1053	0.088	0.096	0.0981
5	5		25.6	698.8	0.0924	0.1154	0.0869	0.0839	0.0919	0.1023	0.0955
6	6		934.6	686.6	0.1123	0.0919	0.0867	0.0928	0.1066	0.105	0.0824
7	7		7.4	57.1	0.0986	0.0998	0.1133	0.1115	0.0903	0.0966	0.085
8	8		104.8	353	0.1048	0.0964	0.0902	0.0976	0.0907	0.0974	0.1136
9	9		988.6	731.2	0.1094	0.1075	0.0826	0.0974	0.1081	0.1021	0.1059
10	10		289.8							0.1097	0.116
11	11		951.1							0.1046	0.0932
12	12		1.9							0.0796	0.1145
13	13		61.7							0.0875	0.1083
14	14		65.8							0.1135	0.1093
15	15		242.1							0.0984	0.0934
16	16		304.3							0.1062	0.1242
17	17		770.5							0.094	0.1119
18	18		394.5							0.1012	0.1094
19	19		808.2	465.8	0.0905	0.1018	0.1171	0.1124	0.0868	0.0827	0.0951
20	20		742.3	17.6	0.0769	0.0973	0.0864	0.105	0.0966	0.107	0.0956
21	21		974.8	880	0.0857	0.0852	0.1111	0.1091	0.0999	0.0928	0.1045
22	22		363.7	247.9	0.0917	0.0978	0.1134	0.1025	0.0955	0.0992	0.0787
23	23		782.2	306.8	0.0905	0.1166	0.1207	0.1011	0.1132	0.1005	0.1085
24	24		200.8	657.9	0.1085	0.1091	0.1048	0.0958	0.095	0.1082	0.0946
25	25		950.5	736.9	0.1071	0.1008	0.0971	0.1029	0.0958	0.1139	0.0909
26	26		364.3	206.8	0.0989	0.0814	0.1001	0.0879	0.0976	0.1157	0.1024
27	27		926.4	462.2	0.0996	0.1027	0.1137	0.0922	0.0823	0.1105	0.1049
28	28		277.6	175.2	0.0871	0.1136	0.1159	0.1185	0.0934	0.0935	0.0826
29	29		371	709.1	0.0995	0.0944	0.0673	0.0907	0.1038	0.0894	0.117
30	30		935	408.2	0.1047	0.0917	0.1066	0.1001	0.1069	0.1212	0.1145

Fig. 1. Table or "flat file" of geological data. The array of data is held in a spreadsheet where it can be edited, manipulated or transferred to other programs for databasing, analysis or display. This spreadsheet is reproduced by permission of Rockware. More at http://www.rockware.com/

Spreadsheet systems generally include full documentation and on-line demonstrations, a good way to appreciate their characteristics and learn how to use them. Their applications may lie in administrative tasks, such as keeping track of expenditure and staff time on various projects, week by week. Their geoscience applications, where many data are collected as tables, are surprisingly varied. A number of properties or characteristics are used as column headings and their values are recorded in a row for each item. In this way, the data are collected consistently. Many computations (see F 3, Fig.1), including statistical and geophysical calculations, can be set out as sequences of formulas relating successive columns in a spreadsheet. This leads to rapid programming, since no complex coding is required. As intermediate steps in the computation are held as separate columns, they can be inspected to check on unexpected results, or to get a feel for the influence of different factors on the final result.

Simple business graphics are widely available for desktop computers. If data have been collected in the form of a spreadsheet, it is not difficult to experiment with their display. Simple pie charts, barcharts and x-y plots are a straightforward means of displaying the distribution of values and the relationships between variables, and are suitable for including in word processor documents. They are as relevant to science as to business, and you should certainly be aware of their existence.

5. Capturing data and images

Spreadsheets and database systems provide means of data entry and editing, including the possibility of creating your own form to enter and check data on screen (D Fig.3). Field or other observations can be manually recorded on printed forms, designed to ensure that all the necessary data are recorded systematically. They make it easier to transfer the data to the computer at a later date, particularly if they match a form on screen. It is also possible, though not necessarily cost-effective, to take rugged and portable computers into the field, where they can store records as the observations are made (Briner et al., 1999). This simplifies review of the information and may help with initial analysis. The course of the investigation can then proceed on the basis of what has already been discovered.

Some surveying instruments and positioning systems can plot locations directly to a computer map. Up-to-date accounts of the satellite-based Global Positioning System (**GPS**) and trials of surveying instruments can be found on the Web, and a complete overview in Hofmann-Wellenhof et al. (1997). Expensive data-collection instruments, such as geochemical equipment in the laboratory or some geophysical equipment in the field, are usually linked to a computer or to a digital recording device. A computer has the advantage that some processing of the data can be done at the time of capture, and it may be possible to adjust or control the instrument by feedback reflecting the incoming data.

Images can be generated from graphic programs, as mentioned in the last section. Existing images, such as diagrams, maps or photographs can be **captured**, that is recorded for use in the computer, with a scanner. The **scanner** captures the image as a raster, a set of colored or monochrome dots on an evenly-spaced grid, typically at a resolution of 300 dots per inch or 15 dots per mm, and perhaps four times as many for a high-quality image. Images use a considerable amount of storage space, from 1 to 24 or more bits for each dot, depending on color resolution. Once captured, image editing and enhancement is possible with appropriate software. Color and density can be modified, the resolution can be changed, the size and shape of the image can be adjusted, or part of the image can be selected by cropping. Images can be combined, for example by overlaying small pie charts of lithology ratios on a scanned map.

Scanning is an important means of capturing data. It can capture images of text on a printed page. **Optical character recognition (OCR)** systems can convert from the image to a word-processor representation of the characters. As this is a moderately expensive and error-prone process, it would not normally be used if the text had already been keyed in and was available in computer-readable form. It would probably, however, cost less than rekeying the document.

6. Information delivery and presentation

A surprisingly important ingredient in the success of computers is the widespread availability of good-quality printers. It is surprising because the rapid availability of information on the computer screen might seem to make the paper copy less necessary. However, paper remains a most convenient medium for reading and studying documents of any significant length. The ability to receive documents from a distant source and print them locally does, however, mark a significant change. Documents can be maintained by the originator, and obtained, edited and printed by the reader. The consequences (M 2.1) amount to a fundamental change in the procedures of publication.

Hypertext documents (E 4) can similarly be delivered and viewed on the screen. In print, of course, it is possible to arrange the document only as a single sequence, and the network structure is lost. Multimedia insertions, such as audio, video or computer programs, cannot be transferred to paper. Electronic documents can, however, be cited in a paper document (see IFLA(1998) and ISO (1999) for style guides). It is worth remembering also that presentations to an audience can be made with a suitable projector, and software such as Microsoft PowerPoint. The images and video from the computer screen can be projected to a large screen and sound to a loudspeaker system.

The ability to pass information between systems depends on shared standards. At a basic level, most systems can accept ASCII characters and thus a character string can be passed between them. Different versions of the same word processing system can generally exchange more detailed information, including adjustment of lines to left, right or center, font, point size, and italic, bold or underlined. By saving the document in rich text format (**RTF**), this information may also be exchangeable between different types of word processor. Similarly, images and multimedia can be transmitted in various standard forms using hypertext transmission protocol (**http**). The appearance of a page can be represented in the Postscript language (E 6), including both images and text. However, this is not appropriate if the text is to be edited by the recipient, as it might require redesign of the entire page layout. Portable data format (**PDF**) is a possible compromise, allowing full access to the text in order, for example, to search for a keyword, while preserving the appearance of the page. Where the recipient is more concerned with content than appearance, particularly where text searching and editing is required, a simpler format, such as RTF, is more effective and efficient.

Some familiarity with the general use of desktop computers is a good starting point for anyone intending to make serious use of computer methods. They should then be able to create simple and compound documents on the computer, and understand the graphical user interface. It is argued later (L 6.3) that the ability to integrate information types will have important consequences for the process of publication. The ability to manipulate and analyze quantitative data will continue to transform the ways in which scientists express and exchange ideas. Most of the features of IT which are set to have a huge impact on geoscience can be seen in embryo in systems on the humble desktop machine.

7. References

Briner, A.P., Kronenberg, H., Mazurek, M., Horn, H., Engi, M., Peters, T., 1999. FieldBook and GeoDatabase: tools for field data acquisition and analysis. Computers & Geosciences, 25 (10), 1101-1111.

Hofmann-Wellenhof, B., Lichtenegger, H., Collins, J., 1997. Global Positioning System: Theory and Practice. Springer-Verlag, New York, 389pp.

7.1 Internet references

Graham. L.A., 1997. Land, sea, air: GPS/GIS field mapping solutions for terrestrial, aquatic and aerial settings. GIS World, January 1997. http://www.geoplace.com/gw/1997/0197/0197feat.asp

IFLA, 1998. Citation guides for electronic documents (Style guides and resources on the Internet). International Federation of Library Associations and Institutions, The Hague, Netherlands. http://www.ifla.org/I/training/citation/citing.htm

ISO, 1999. ISO 690-2, Bibliographic references to electronic documents. Excerpts from International Standard ISO 690-2. http://www.nlc-bnc.ca/iso/tc46sc9/standard/690-2e.htm

<<<Back to Table of Contents
On to Part D: Familiarization with IT applications to support the workgroup>>>