New initiatives in map library automation. Workshop held at BGS, Keyworth on 23 September 1992.

Summary of paper given by Dr T V Loudon of the British Geological Survey's Information Systems Group.

BGS Geoscience Index System

It is appropriate, at a map library automation workshop on the British Geological Survey's home ground, to consider a new initiative that has a bearing on BGS's map indexing activities.

The geological map is a primary means of communicating geological information, especially that collected by methodical geological survey. BGS has therefore been a publisher of maps since 1835, mostly at 1:50 000 and smaller scales. In addition, BGS staff are users of geological and other maps world-wide, and the BGS Library holds a reference collection of some 200 000 maps. For these reasons, BGS undertakes map cataloguing and indexing. Conventional procedures are followed, but are not adequate for all purposes. It is timely to look afresh at the user requirement, the nature of the map, and changes in survey methods.

The geology of the UK is surveyed, not at the publication scale of 1:50 000, but at a larger scale, generally 1:10 000. The survey-scale maps contain more detail than the published maps, and are held in a reference archive. The raw data may be selectively reused, possibly with additional data collection, to prepare maps on specific themes, such as sand and gravel resources, geological hazards in construction work, engineering geology characteristics, geological aspects relevant to land-use planning, and so on. BGS also collects and publishes a wide range of other geoscientific information in such fields as geophysics, geochemistry, hydrogeology, marine geology and geophysics, some at smaller scales, such as 1:250 000. The geology of the UK continental shelf has been mapped in some detail, and the BGS International Division has prepared maps for many areas overseas.

Users of BGS information come from many fields: civil engineering, land-use planning, mineral exploration, water resources, academic geology, geophysics and many more. The map user may need to know what maps are available for a particular area at any scale, published or unpublished, and may wish to know of earlier maps, for historical reasons or because recent coverage is incomplete. The user also needs to know about other map-like documents, and other spatially-referenced information. Examples are mine plans, site-investigation reports (referring perhaps to one single building or to a lengthy section of a motorway), coalfield reports, and areas where surveys are planned or in progress. Such reports deal with defined areas, but do not fit neatly into recognised sheet boundaries. In addition, line data may be of interest, such as seismic lines, or published cross-sections (analogous to geological maps of a vertical slice through the earth's crust); or point data such as boreholes, geophysical recording stations, outcrops, or collection sites for samples and specimens. They do not fit easily into a conventional catalogue.

The user's requirements are likely to refer to an irregular area, perhaps a structure plan, or a mineral lease, or a corridor of possible road construction. The user would prefer not to have to consult a large number of catalogues, even if they were readily available and up to date. The perceived requirement may be redefined as the user discovers what is available. An interactive system is thus required, in which the user can define and refine the area and topics of interest. Graphical presentation is essential, with the data accurately positioned against a clear background of relevant topographic features.

One route to a solution is to use the structure and concept of the map itself. The map is, after all, a predictable and reversible geometrical transformation, not of the real world, but of the surveyor's model of that world. Distances are transformed, but most of the spatial relationships of the original are retained - adjacent areas remain adjacent, points within an area remain within it, and so on. The conformance of spatial characteristics between map and model is what makes maps useful, and applies to index data as well as to any other.

The BGS geoscience index system uses the computer equivalent of the map, namely, the Geographic Information System (GIS). Building a spatial index within a GIS enables the user to perceive and make use of spatial relationships. For example, it is easy to see which site-investigation reports are nearest to a proposed development, or to make a selection scattered evenly around it. Having located relevant data, the user needs sufficient catalogue information to determine whether the data are relevant, and where they can be obtained. Thus he might need to know the date and source of a map, its availability and its sheet number. The index accesses non-spatial data through a relational database management system (Oracle) to which the GIS (Arc/Info) is linked.

Having located data in the index, the user may then wish to obtain more detail and carry out some processing. For example, having determined which are the nearest boreholes to a site, the user might wish to select those penetrating a particular rock unit, and then draw a contour map of the formation. Although this goes beyond the capabilities of the present index system, such needs must be kept in mind in future developments. The word "index" is derived from the Latin word for forefinger - a digit pointing to information. Geological data are like the layers of an onion. It is always possible to move to a more generalised layer, or to go in the other direction to more specific information. Indexes could point from a layer of any degree of generality to layers with more detail. Ultimately, no sharp dividing line can be drawn between an index level and a data level.

Returning specifically to the indexing of maps, another issue should be mentioned. BGS currently produces geological maps by computer, and plans, in due course, to prepare maps from a primary record held as a GIS. It is intended that the database should be seamless, so that maps could be drawn for any selected area. It is also envisaged that there should be continuous revision of the map data. When data arrive from, say, a new borehole or seismic survey, the geological lines in the GIS would be redrawn to reflect the new data. The actual maps would then be plotted on demand, from the most recent information, for the required area and selected topics. A map with no sheet boundaries nor fixed contents poses problems for the cataloguer. The map, in these circumstances, is not the appropriate cataloguing unit.

The BGS Geoscience Index can be seen as providing an overview of corporate activities and products. It covers a wide range of topics where individual groups have responsibility for the detailed data. It may take its place in the future as part of a front-end system giving access to a range of application programs. No clear distinction is drawn, perhaps none can be drawn, between published and unpublished records, between maps and other forms of spatial information, nor between the index and the data.

The Geoscience Index is designed for casual and occasional users, with the menus and the usual point-and-click features of a graphical user interface. The user defines his area of interest by pointing on a screen map, or entering a place name from a gazeteer, or by specifying coordinates or a map sheet number. Data topics are selected from a menu, and the retrieved features are displayed against a base map, the components of which, such as place names, roads, rivers, railways, settlements, coastlines, geological formations, are chosen by default or selected by the user as being relevant without causing clutter.

The user can query items by pointing and clicking on a borehole, map area, or other feature. The attributes of the selected feature are listed within a window on the screen. Alternatively, an area can be outlined on the screen and all records for the selected topic within that area are displayed in a window, and the user can scroll through the items. Images of documents related to points, such as scanned images of borehole logs, can be retrieved and displayed by clicking on the borehole location. There are many other features, but they are better shown than described. The author of the system, Keith Adlam, demonstrated the system during the Workshop.

The programs for the Geoscience Index are written in Arc Macro Language (AML). They are implemented using the Arc/Info GIS and the Oracle relational database management system. The demonstration was mounted on a Sun Sparcstation 2, and the index is installed on standalone Sun IPX workstations at BGS regional offices. At BGS headquarters in Keyworth, and in the BGS Scottish office at Edinburgh, the index database is held on a Sun 630 server, accessed over the Local Area Network from Sun IPC workstations. This is an attempt to achieve a balance between cost and efficiency. BGS data has been scanned and digitised for the index on Intergraph equipment. Arc/Info and Bartholomew's topographic data have been obtained for research purposes through the Combined Higher Education Software Team (CHEST). Other topographic information has been obtained from the Ordnance Survey.

BGS sees it as part of its role to make its products available for others. The Geoscience Index System is tailored for BGS use, but BGS can advise and assist with the development of similar systems. Our consultancy rates are competitive, and BGS would be happy to discuss how you can obtain best value for money.

T V Loudon
British Geological Survey
Murchison House
West Mains Road
Edinburgh EH9 3LA
Scotland

21 December 1992