

# Artificial Intelligence Applications in Geology: A Case Study with EXPLORER

M.D. Mulvenna, C. Woodham & J.B. Gregg

## ABSTRACT

The modern exploration geologist deals with large amounts of data. This paper describes a system, EXPLORER, which was designed to help manage data collected in gold exploration in Northern Ireland. The system uses geologists' heuristics of exploration and may be expanded to search for base-metals and precious-metals in regions of glaciated terrain. EXPLORER generates reports on each square km of the licence. It utilises a forward-chaining inference strategy, where the rules are fired from licence data held in a Prolog database, to produce the reports. The data includes information on geology, geomorphology, geophysics; and empirical data is analysed for both target and pathfinder elements.

## INTRODUCTION

Modern exploration techniques can present the geologist with a plethora of data which he or she may have difficulty representing and interpreting using traditional presentation methods. As computing technology is making inroads into geoscience, and generating large databases, data management will occupy an increasingly greater percentage of most geologists' time. To quote an example, "over a five week period, a survey involving the collection of 500 samples per week with ten field observations at each site and an analytical suite of only 30 elements will provide 750,000 items of information" [1].

Increasingly, because of the large data sets that are generated, the field geologist will need to rely on "exploration assistants" [2]. These expert decision-support systems will ensure consistent treatment of data, and importantly, when fully developed, tutor the inexperienced geologist in under-developed nations.

## EXPLORATION BACKGROUND

### Review

Traditionally, gold prospecting in Ireland has been carried out using the tried and

tested method of stream panning. This can be advanced to sampling the soil around the streams in search of the bedrock source of gold. Nowadays, these methods are still used, but are complemented with a wide range of other methods (Table 1). These empirical, or field methods, when combined into an integrated database, provide a comprehensive suite of data for analysis.

Another approach to gold exploration is to construct geological models which can aid the geologist enormously. For example, a model which uses information on the genesis of a geological structure which may contain gold (metallogenic models), could then be used in conjunction with what the geologist knows about a particular prospect. The model could perhaps reveal much about the prospect, and allow both qualitative and quantitative assessments to be made.

Prospector [3], was one of the first expert systems in geology, and it used a form of model-based reasoning, as outlined above, to evaluate a prospect. It “matched (volunteered) information with deposit models stored in a knowledge base”.

Models developed for gold exploration in Ireland are too weakly focused to restrict the spatial extent for exploration sufficiently. However, they can target areas for recce evaluation (Figure 1).

### Exploration Detail

The building of any expert system to aid gold exploration in Ireland must employ the geologists heuristics of exploration. The system should try to emulate the geologists interpretation of the field data.

Once designed, the system could then be generalised to explore for most minerals in glaciated regions.

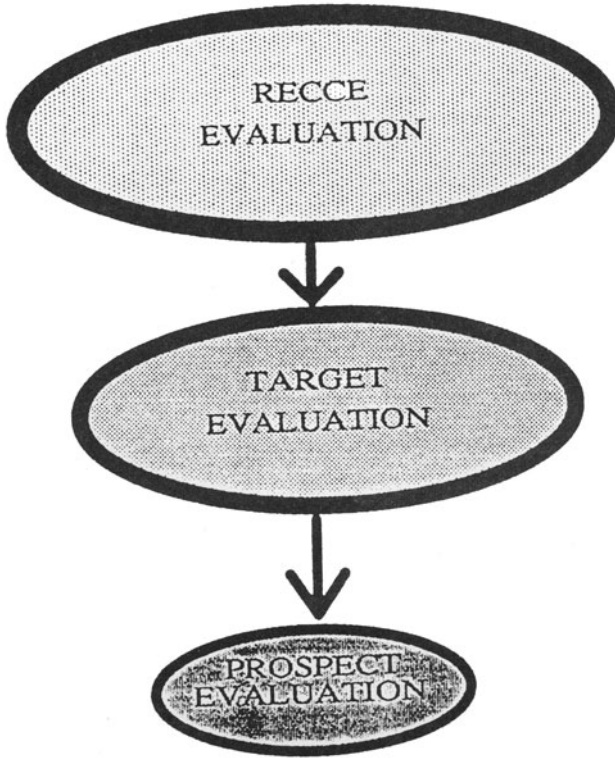
Geological criteria from known gold mineralization can be built into the expert systems rule base to define favourable areas for further follow-up. The targeting criteria can assist at the ground selection and prospecting stages of exploration.

There are an abundance of empirical methods available to the exploration geologist, and results from these field methods, i.e. geological mapping, soil and drift geochemistry, outcrop geochemistry and geophysics, must be accessible to the expert system. Ranking of geochemical anomalies [4], where a score is accumulated for each area according to the results of the data, may be incorporated into future systems.

Generally the geochemical data includes readings (in parts per million/billion) for target elements and pathfinder elements. Pathfinder elements are so-called because they can lead to the target. They are frequently of higher concentration and more widely dispersed than the target element. In this version of the EXPLORER

Table 1. The fifteen data sets used for gold exploration in Northern Ireland

Data	Sampling Method	Analysis/Information
1. Drift	Minimum of 1 sample per square km	Information on drift type and thickness
2. Geology	Minimum of 1 sample per square km	Information on bedrock geology
3. Recce Soil	Carried out at 1 sample per square km (generally at roadsides)	Analysis for gold and arsenic
4. Recce Stream Sediment	Carried out at all road/stream intersections	Analysis for gold and arsenic
5. Stream Sediment HMC	Involves panning of stream bank sediment at all road/stream intersections in licence	Gold grains in the pan are counted and the concentrate is analysed for gold and arsenic
6. Till HMC	Involves panning of stream bank till sediment	Gold grains are counted and the concentrate is analysed for gold and arsenic
7. Grid Soils	Involves shallow soil sampling; generally on a 100m x 50m grid	Analysis for gold and arsenic
8. Grid Deep Sampling	Involves deep overburden sampling on a 100m x 50m grid	Analysis for gold and arsenic
9. Detailed Roadside Soils	Involves shallow soil sampling along roadsides at 0.1 mile intervals	Analysis for gold and arsenic
10. Detailed Roadside Deep Sampling	Involves deep overburden sampling along roadsides at 0.1 mile intervals	Analysis for gold and arsenic
11. Detailed Traverse Deep Sampling	Carried out as follow-up to grid deep sampling at 5 m intervals	Analysis for gold and arsenic
12. Outcrop Geochemistry	Involves geochemical analysis of outcrop samples	Analysis for gold and arsenic
13. Float Geochemistry	Analysis of boulders/cobbles, on surface or in stream sections	Analysis for gold and arsenic
14. Till Geochemistry	Analysis of the fine fraction of glacial overburden	Analysis for gold and arsenic
15. Geophysical Methods	Generally targeted to interesting areas	Induced Polarization/Resistivity collected



**Fig. 1. How the exploration geologist seeks to constrain the area of exploration.**

system, gold (Au) is the target element and arsenic (As) is a pathfinder.

The current generation of geoscience database systems have many benefits, including complex statistical procedures, graphics, plotting, and windowing facilities. Windowing is the constraining of data sets to a given area. These facilities can allow complex spatial modelling to be carried out. However, to cover every square km of a licence area for every data set, with various statistical routines and different window sizes can prove extremely time consuming. Also, consistency is not guaranteed. In short, complex quantitative spatial reasoning is required [5]. This is a task at which most humans are not naturally adept, especially when the licence area can be 1000 square km in size. The procedure on how to carry out exploration varies from geologist to geologist and company to company, but the data set is finite to an extent; usually changing only by modification of an existing technique or old techniques becoming redundant and superceded through time.

The goal of exploration is to find mineral deposits by successfully restricting and testing the target area(s) (Figure 1), therefore methods which cover a larger area are tried first, and so on, until the area is either rejected or the target is well enough delineated to allow small scale exploration, and possibly, core drilling.

This brings up the question of cost. Exploration is very expensive, and when an area is rejected early, monies can be allocated instead to more promising prospects. Therefore, if a method existed to consistently examine the total licence area and assist in target definition and evaluation, resources would be optimised, and planning of future exploration could be more concisely pinpointed.

## THE EXPLORER SYSTEM

A rule-based expert system to implement such a method was considered feasible (Figure 2). Benefits could include:

- rapid report generation and target definition for licence areas.
- consistent follow-up recommendations for target evaluation, optimising objectivity.
- a flexible exploration method able to be updated as new successful techniques are introduced or discovered and able to be used in different terrains for different minerals given new expert exploration criteria.

Reports are built up for each square km and kept on file for reference and comparison with updated reports generated when new data is added. As the exploration method is stored as rules with an understandable syntax, these can be modified easily by an exploration geologist using a text editor. Thus methods which represent the knowledge and expertise of a geologist (or company) can be captured.

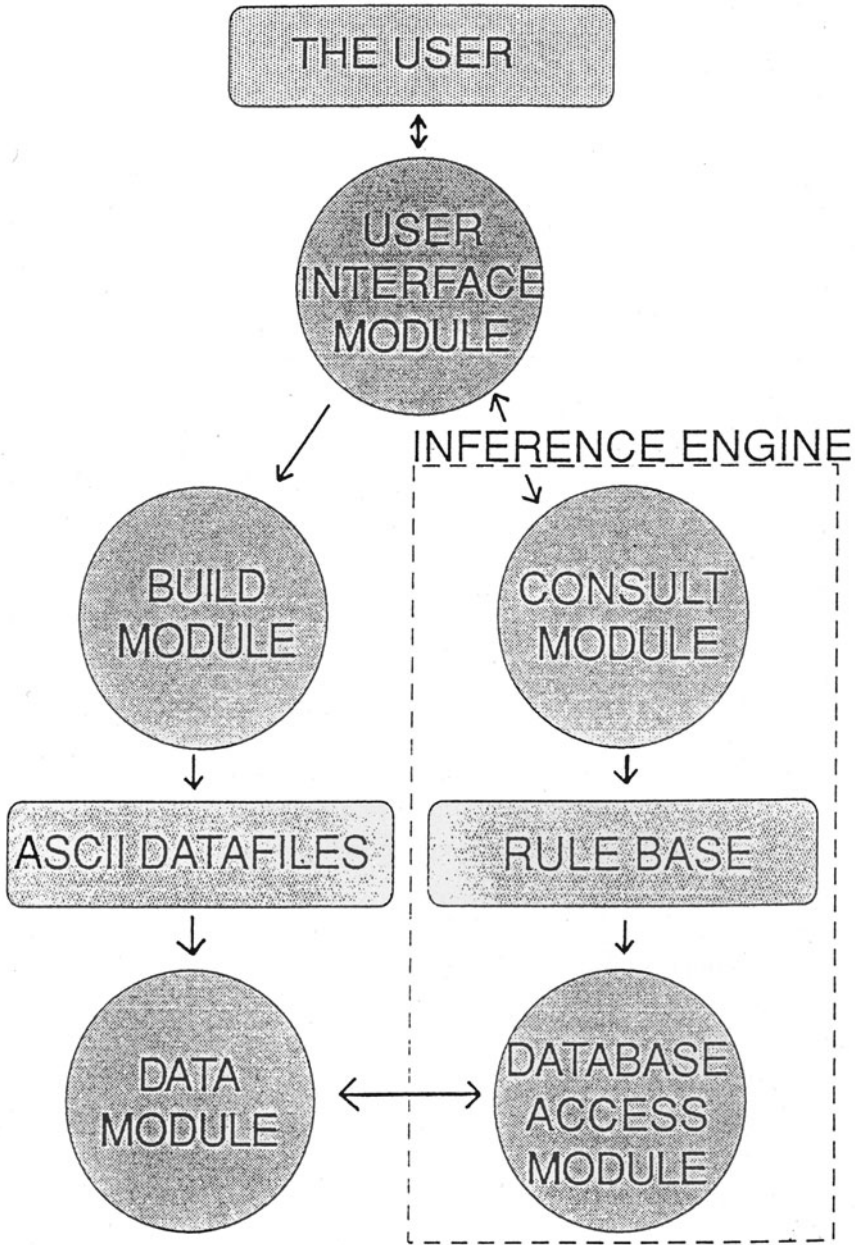


FIGURE 2. EXPLORER SYSTEM ARCHITECTURE

Initially the system was to link up to some commercially available database product, preferably in the geoscience field. Unfortunately the database which held the test licence data was unsuitable in its present form, due to an inflexible menu system, and no other geoscience products were available at the time. Rather than use a commercial database product such as dBase IV, where it was unrealistic to expect the geologist to maintain two separate database applications on one busy PC, the exploration databases were dumped to standard ASCII text datafiles. These datafiles could then be input to the BUILD module of the EXPLORER system.

### The Build Module

When converting from ASCII files to Prolog clauses, the BUILD module (Figure 3) creates an imaginary 100 x 100 km grid over the National Grid and calculates the 1 km square in which the sample is located. It then adds this 'squares' information to the Prolog fact for that sample. This method considerably shortens access time, as the X and Y values of the square are instantiated when the system looks for a sample in the current square. Because the X co-ordinate is hashed, only those clauses with the same X square co-ordinate are considered by the ACCESS module.

The BUILD module compiles the data into 'heap' format, a machine representation which does not need to be syntax-checked or tokenised, and so loads quickly into memory. The DATA module created by BUILD is stored in virtual memory, which allows the size of the module to be much greater than the memory capacity of PC's running DOS, up to a limit of one megabyte. In practise, data sets for licence areas are less than this 1 megabyte limit. If more than 640 kilobyte RAM is present on the PC, the EXPLORER system can automatically use this to store the data as a virtual module on a virtual RAM disk, so optimising speed. This storage method bypasses the problems of large Prolog database storage in main memory, although ultimately the intention would be to revise the EXPLORER system to form a tightly coupled expert database system [6].

### INFERENCE

The system forward chains through the rules initially from the data and then from inferred facts and data. The rule base has fifteen rule groupings each representing a data set. The inferencing method emulates the geologists thinking by first looking at general rules within the chosen 1 square km window. All the general rules are

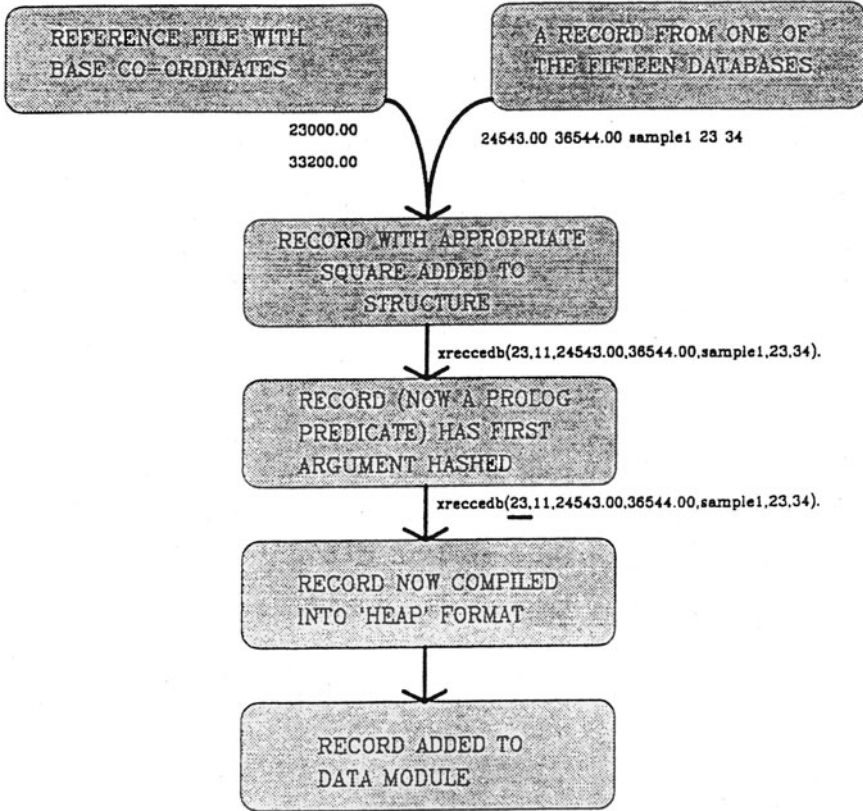


Figure 3. How the BUILD module works



gathered up and processed through conflict resolution. Conflict resolution is determined by refactoriness, recency and specificity, in that order, and the system picks the first rule in the list that emerges from the process.

### Meta-rules

Meta-rules, embedded within each of the fifteen rule groups flag two situations:

- the system should look at the eight squares adjacent to the current square, and apply a selected rule grouping to those squares;
- the system should apply a different rule set to the current square.

The meta-rules proved necessary because the geologist often decides to look around the area currently being investigated. An example of this would be to expand the window, i.e. look at adjacent squares. If data from one rule group is positive, the geologist must look around for support in adjacent squares. One example would be to look for support as geochemical dispersion trains located in a 'down ice' direction perhaps the result of glacial smearing from a bedrock source. Also, if one rule group showed nothing for a square then the geologist looks at the data from another source for that area.

When the system has processed all applicable rules and returns to the main menu, all items flagged are dumped into two lists, the first for applying flagged rules groups to the current square, and the second for applying flagged rule groups to adjacent squares (windowing). These lists are displayed to the user as sequences of two menu sets. Each menu gives the user the option to carry out the flagged option, ignore it and look at the next flagged option, or see why that item has been flagged.

The main output from the system is a report generated from text and variables tagged onto individual rules. Explanation facilities in the form of 'why' are implemented for flagged items. This gives the user an explanation generated from the rules to aid decision-making, and guide the EXPLORER exploration process.

### The ACCESS Module

A subset of the inference manager is the database ACCESS module. This looks at the database for each square and tries to find the best value for that square. For example, with the outcrop geochemistry database, it searches for a sample with gold and arsenic above their respective thresholds. If this fails, it attempts to find a sample with gold above threshold. In turn, if this fails it searches for arsenic above threshold. The bottom line is that, if all the above fail, it attempts to find any

reading for the current square. If this does not succeed, then the ACCESS module fails to find a reading and the conditional part of the rule which called the access fails. This causes the rule to fail, and the inference manager tries the next rule in the queue, or cycles again.

## USER INTERFACE

The EXPLORER system was designed to be user-friendly. This entailed a design with almost total menu control. The pull-down menus are mouse-driven. Context-sensitive help is provided in the form of a help screen displayed with associated help nodes shown as an additional menu on an adjacent screen. These items can, in turn, be selected allowing the user to peruse the help information.

While the user is navigating through the system, all menu items are dynamically checked to see if their selection is valid. This style of dynamic menu-checking traps many potential conflicts in the system and therefore reduces errors.

## FUTURE DIRECTIONS

EXPLORER has been applied to a data set representing exploration primarily in a prospecting licence (DE5) covering 130 km<sup>2</sup> in Co. Tyrone, N. Ireland. This licence has been held by Celtic Gold plc since 1987 and has known bedrock gold mineralization thus enabling the development of the EXPLORER rule base [7]. The system is currently, as of May 1990, being evaluated by a consulting geologist attached to the company. Future work is planned to integrate a more general system for exploration, perhaps employing blackboard architecture, and using empirical results, with a relational database, and graphics system to provide an Intelligent Geological Information System (IGIS)[8], a step further than Geological Information Systems [9]. Initially, the scope of the IGIS will be to reason spatially about precious metal and base-metal deposits overlain by glaciated terrain.

It is recognised that geological reasoning systems will be in demand to help understand and filter the large data sets which cover many disciplines, and areas of the earth [1]. It may be that by developing IGIS's, we can explore and more fully integrate field data, satellite data and qualitative data to provide a richer working environment for the exploration geologist.

## SYSTEM IMPLEMENTATION

The system was built on an IBM 80286 PS/2 DOS computer. The Prolog system used was Prolog-2 Professional Plus. The computer was chosen because it was compatible with that used by the geologist, while Prolog-2 was picked primarily for its virtual memory management, window facilities and modular architecture, which enabled large database construction, a good user interface development environment and incremental program development, respectively.

Other methods of knowledge representation, including frame- and object-based systems, were not used.

Prolog-2, PS/2 and dBase IV are trademarks of Chemical Design, IBM, and Ashton-Tate, respectively.

## REFERENCES.

1. Plant J. A. Hale M. and Ridgway J. Developments in regional geochemistry for mineral exploration. *Trans. Instn. Min. Metall.* 1988 Sect. B;97:116-140.
2. Martin L. Expert systems and their use as exploration assistants. Unpublished paper presented at 'Exploration 87' the Third decennial international conference on geophysical and geochemical exploration for minerals and groundwater, Toronto, Canada.
3. Gaschnig J. G. PROSPECTOR: An expert system for mineral exploration. In: Michie D (ed) *Introductory readings in expert systems*, Gordon and Breach science publishers New York, 1982, pp 47-64
4. Bonnefoy D. Jebrak K. M. Rousset M. C. and Zeegers H. SERGE: an expert system to recognise geochemical anomalies. *J. Geochem. Explor.* 1989; 32:343-344.
5. Yatabe S. M. and Fabbri A. G. Putting AI to work in geoscience. *Epiodes*, 1989;12,1:10-17.
6. Torsun I. S. and Ng M. Y. Tightly-coupled expert database systems interface. In: Kelly B & Rector A (eds) *Research and Development in Expert Systems V*. Cambridge University Press, 1988, pp 211-223
7. Woodham C. Finlay S. and Holman R. Gold Exploration in the Dalradian of Northern Ireland. *Trans. Instn. Min. Metall.* 1989; Sect. B;98:63-65.
8. Smith T.R. Artificial Intelligence and its Applicability to Geographical Problem Solving. *The Professional Geographer.* 1984;26:147-158
9. Coupez Y. A micro-computer based geological information and processing system for exploration. In: Farrell L & Jones G Ll (eds) *Irish Assoc. Econ. Geol. Annual review. IAEG*, 1989 p 48