Nationwide collection, recording, and provision of geo-scientific data. Examples from Namibia, Germany, and Kosovo. An experience report.

Andreas Barth

Beak Consultants GmbH, Am St. Niclas Schacht 13, D-09599 Freiberg, E-Mail: andreas.barth@beak.de

1 Introduction

Over hundreds of years geo-scientific data were collected and fixed on paper. Handmade descriptions, bore hole logs, maps, exploration reports etc. have filled up public and non-public archives with materials of inestimable value. The systematic capture, storage and distribution of that information is an extremely important, long-term and very expensive activity.

State managed geological surveys and similar organisations as well as large private companies make a broad use of modern data storage systems, such as text files, spreadsheets, databases, CAD and GIS, and electronic images (scans). Many systems grew independently from each other even in one organisation. The result is a jungle of information, many redundancies, problems with coding, networking and distribution of information.

Traditionally, geo-scientific spatial information was collected on maps. Maps were the most important final product of any geo-scientific work. This principal methodology was continued to use even after the introduction of GIS into the normal geo-scientific working process. Now, we are on a threshold of a principal alteration of the management and distribution of geo-scientific information. More and more, our attention is being directed to the creation of flexible, and redundancy – free information systems, which allow the production of many different user defined final products.

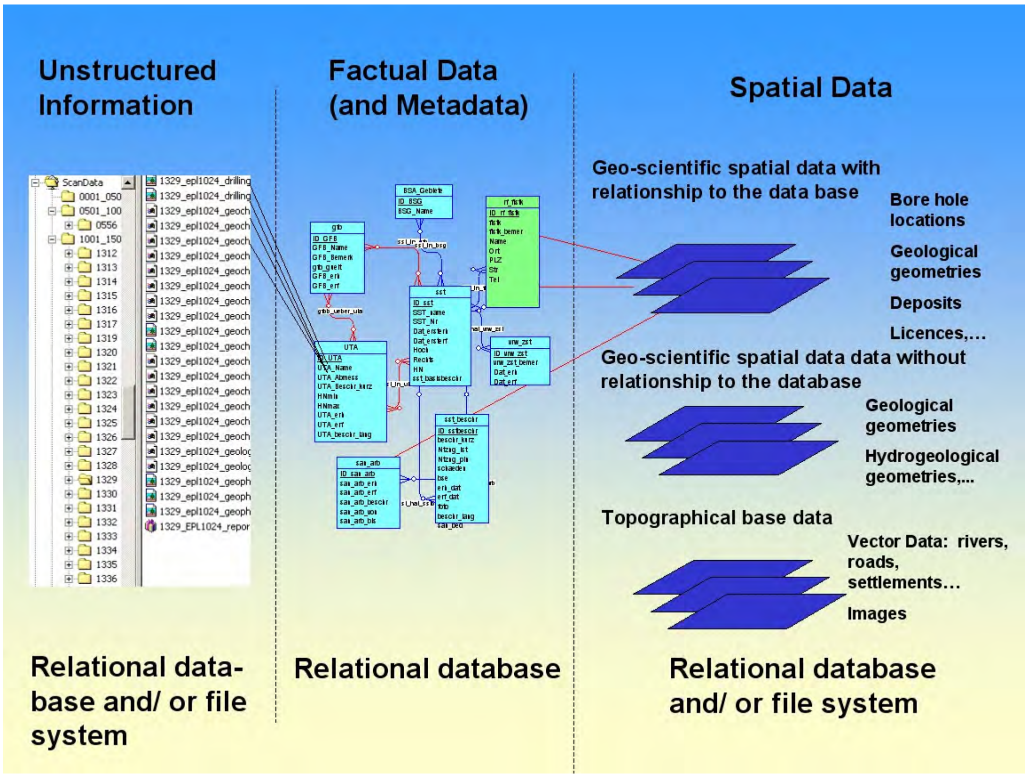

Figure 1: Principle data structure.

2 Principles of data storage

The principal components of the data base are storage containers for **factual data, spatial data** and **unstructured information** (such as scanned maps; Figure 1). Although many variants of data storage principles exist, the core of the system is always a **relational data base**. Here the main entities (such as bore holes, sampling points and analyses, legends, mineral occurrences and deposits, mines and licences) are described in systems of tables. This entities are linked with each other by spatial and logical principles (e.g. sampling points related to a certain licence site, licence applicant to a certain licence). The linkages might be created by (link)- tables or by application of SQL-queries (only when spatial data is stored in the relational data base). This linkages are the prerequisite for a comfortable use of the data base (inquiries).

The introduction of web based services for the provision of geographical base data makes sense for a web based system only.

3 Important functions

Most important functions are complex **inquiries** with spatial and/or logical background (type: show all Cu occurrences of potential economic importance in a certain part of the country; show all sampling points and related Au values for a certain mineral licence/ for all licences of a certain applicant; show all reports and maps with regard to geochemical sampling done by a certain company, in a certain area; etc.).

The **presentation** and **export** of the inquiry results is another important function. It consists of different functions for the automatic creation of **user defined maps** (content, scale, size), the generation of **text-files** (tables) for further customized processing (e.g. with other software).

Important **administration functions** are tools for a customized **user management** (setting of access rights to different modules of the system) and for data security reasons (backup system and protection against the loss of data and unauthorized data manipulation).

The creation of (fixed) **reports** is important in some cases only, e.g. the creation of deposit passports, mineral licence documents etc.. Principally, the graphical user interface (GUI) can be implemented as a **client-server application** (Figure 2) or as a **web-able application** (Figure 3).

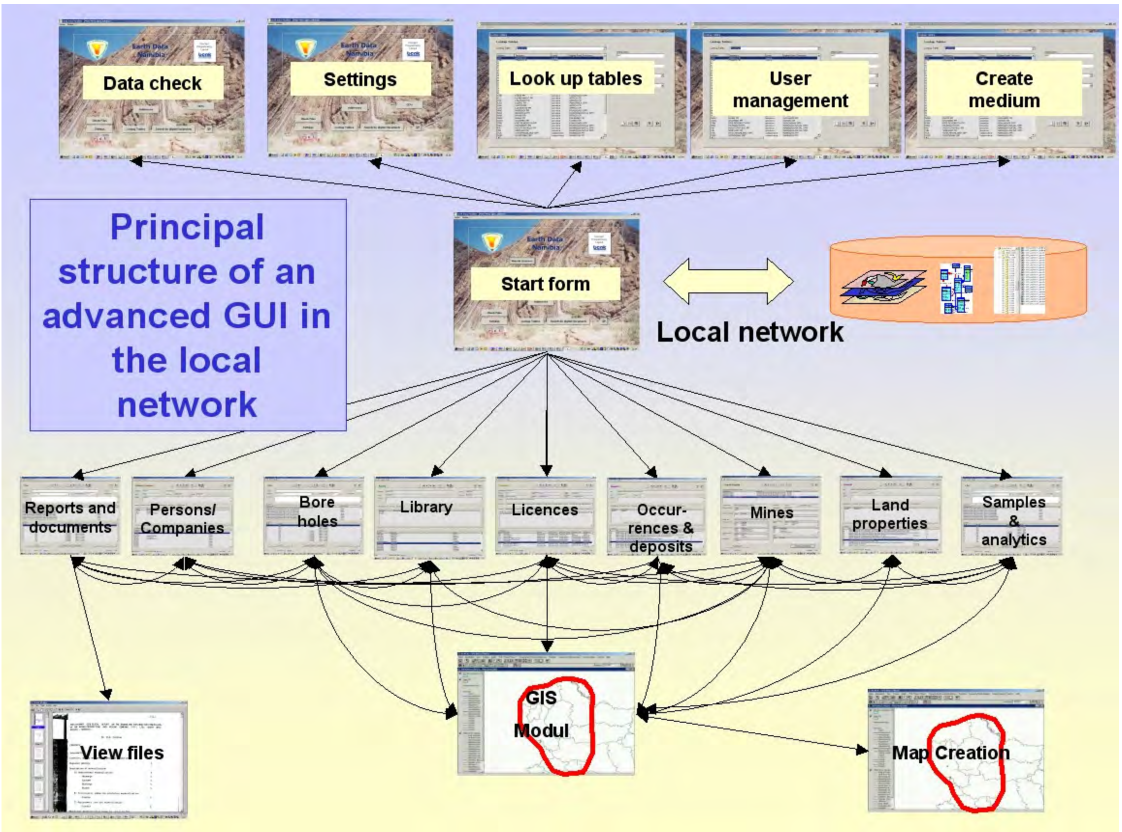

Figure 2: Principle structure of an advanced GUI in the local network.

4 Data capture process

The **value** of a complex geo-scientific information system consists not only of its easy to use and fast GUI, but of course of the **correctness and completeness of its data,** too. Therefore, the **organisation of the data capturing process** (Figure 4) **and its quality management** (Figure 5) are of high importance. The quality insurance process consists of different constituents: e.g. the storage of primary paper documents for further checks, plausibility checks during the data entry process, data check by third persons, recording of the date and the person of any data modification.

Good experience was made both with a **continuous data capturing process** (data is entered as it arrives), and with a **project-like organised process**. The latter one is important for the capture of old information from different sources, such as mineral occurrences and deposits, bore holes, reports and documents.

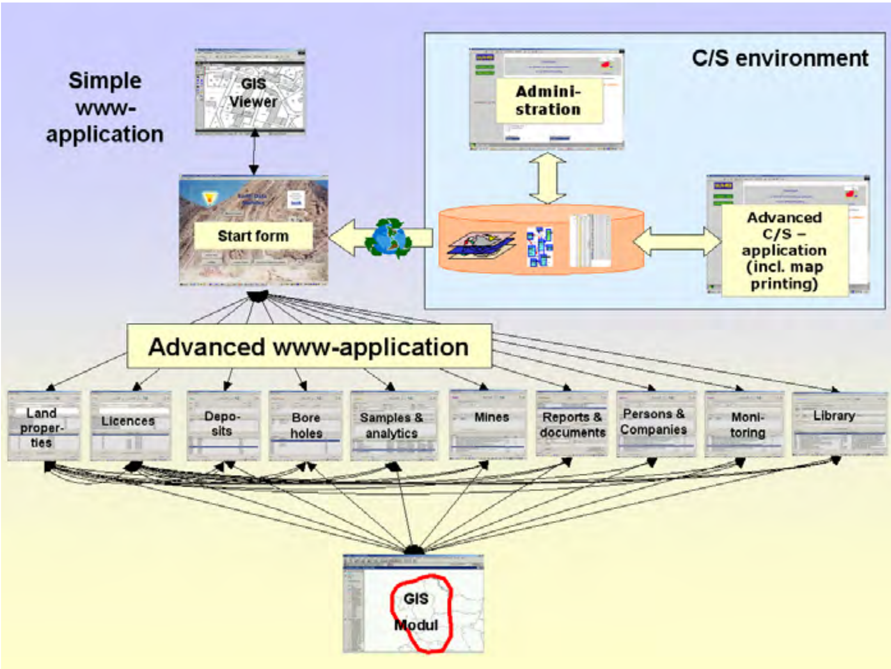

Figure 3: WWW- based application.

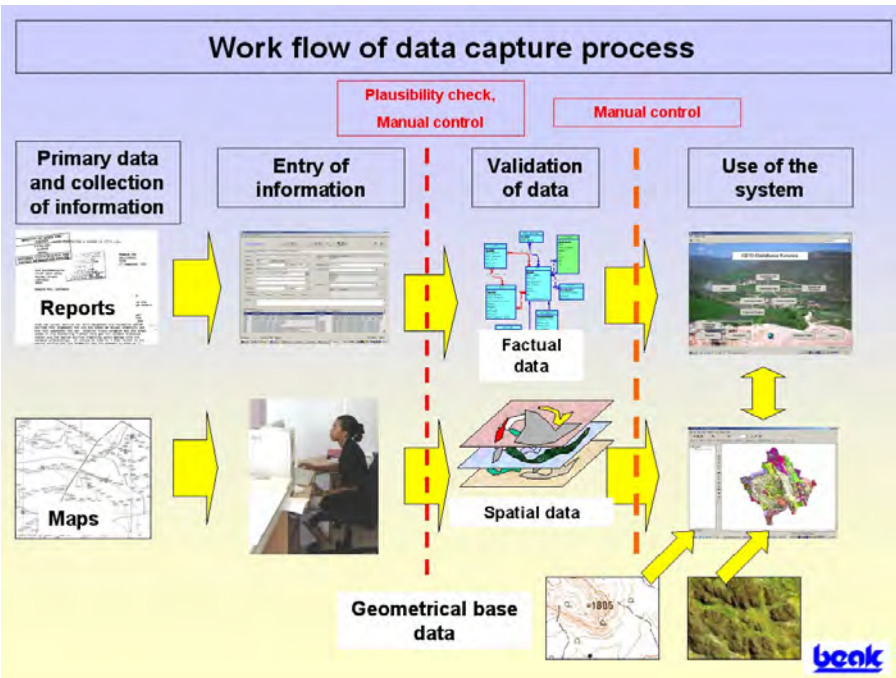

Figure 4: The data capture process.

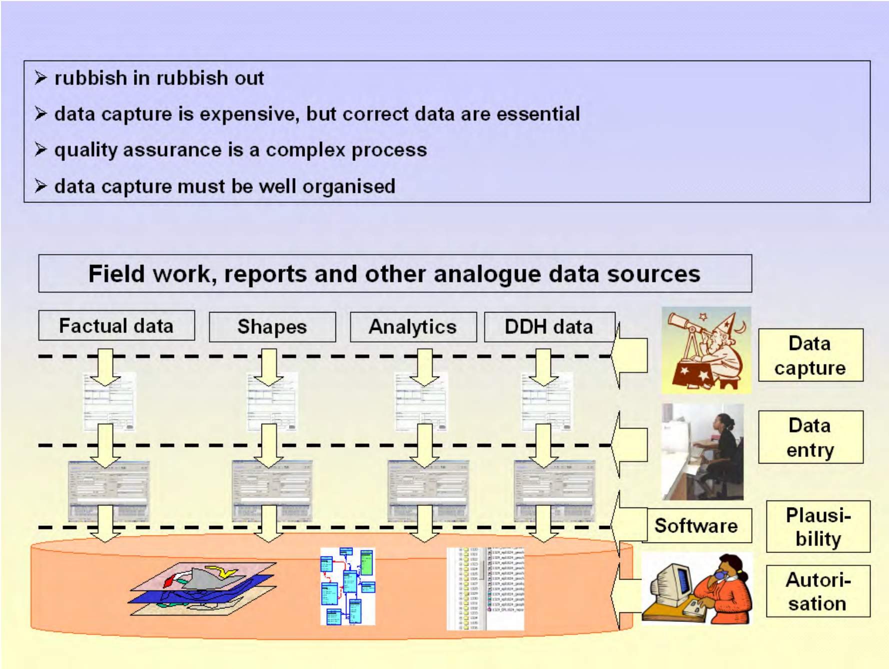

Figure 5: Multi-level quality assurance process of data capture

5 Software

Although the principle structure of the systems are equal everywhere, the creation of customer designed tailor-made software is a necessary. Due to local requirements (e.g. laws, regulations, responsibilities) and different prerequisites (hardware, software, languages, knowledge, coding systems), a multiple use of exactly the same database structure and front end software in most cases is impossible. Moreover, the fast development of hard and software demand the adjustment of the special system to new conditions.

We have made the best experience with high end products from **Microsoft, Oracle** and **ESRI**. This products are not cheap, but they guarantee investment security, good support and a comparable long term use of the systems.

6 Experiences

Our experience shows, that Information Systems for an organisation like a Geological Survey or a large company should be planned strategically. The best “basement” of a future-oriented system is a modular, network like, and redundancy-free organised **data model**.

The creation of seamless datasets demands the strict use of **general legends** (organised flexibly and open for new entries) and **look up tables**. A restricted access to the look up table management including the legends is essential for a long-term useable data base.

The **data capture** process is very expensive, but saving up measures will press on the quality of the data. As experience shows, the retrieval of wrong data and its correction is much more expensive than the entry of correct data.

An easy to use and **self explaining GUI** will contribute a lot to the acceptance of the system.

7 Presentation

The presentation includes the following systems:

- Non-metallic Minerals of Saxony (Germany): ACCESS97, ARCVIEW 3.2 (1998), Saxonian Authority of Geology and the Environment.
- Earth Data Namibia: ORACLE 10, ARCVIEW 3.2, Visual Basic 6.0, Map Objects (2001-2004), Geological Survey of Namibia
- Geo-Database Kosova: SQL-Server, ARC-GIS 8.3, VB.Net, Map Objects (2003-2005), Independent Commission of Mines and Minerals of Kosovo
- GROWAS (Ground water information system): SQL-Server, VB.Net (2004-2005), Department of Water Affairs, Ministry of Agriculture, Water Affairs, and Rural
- Development of Namibia