

# Assuring data quality in Posiva's biosphere assessment

Thomas Hjerpe<sup>1</sup>, Dmitry Hofman<sup>2</sup>, Per-Gustav Åstrand<sup>2</sup>

<sup>1</sup> Saanio & Riekkola Oy, Laulukuja 4, FI-00420 Helsinki, Finland. <sup>2</sup> Facilia AB Gustavslundsvägen 151C SE-16751, Bromma, Sweden.

## INTRODUCTION

In Finland, Posiva Oy (Posiva) is responsible for implementing the final disposal programme for spent nuclear fuel, and is currently compiling a Safety Case to provide the safety basis for the construction license application for a geological nuclear waste repository at the Olkiluoto site. One main component in the Safety Case is the biosphere assessment (BSA). The BSA in the Posiva Safety Case aims at describing the past, present and future conditions of the surface systems of the Olkiluoto site, track the fate of hypothetical releases of radionuclides from the repository reaching the biosphere, and assess possible radiological consequences on humans and other biota.

The Safety Case is a project spanning over a long time period and has an iterative nature, i.e., the Safety Case is to be updated several times. In addition to this, the BSA involves vast amount of data, as well as assumptions and expert judgements, thus a clear strategy how to handle all this information is of great importance in orders to produce confident results from the BSA. In addition, the Finnish regulation (STUK, 2001) requires that the Safety Case work shall be based on an appropriate quality management. The regulation also states that a safety analysis shall include:

- functional description of the disposal system by means of conceptual and mathematical modelling and the determination of the input data needed in these models
- uncertainty and sensitivity analyses and complementary discussions on the significance of such phenomena and events which cannot be assessed quantitatively
- documentation of the safety analysis

According to regulations, documentation shall target to *transparency*, so that the approaches, methods, results and the coupling to the entirety in each part of the analysis can easily be discovered, and *traceability*, so that justifications for the adopted assumptions, input data and models can easily be found in the safety assessment report or its reference reports.

This contribution describes the strategy for managing information in general to carry out a coherent and consistent BSA that produces confident results, and the quality assurance procedure implemented in the BSA database (BSADb) in particular.

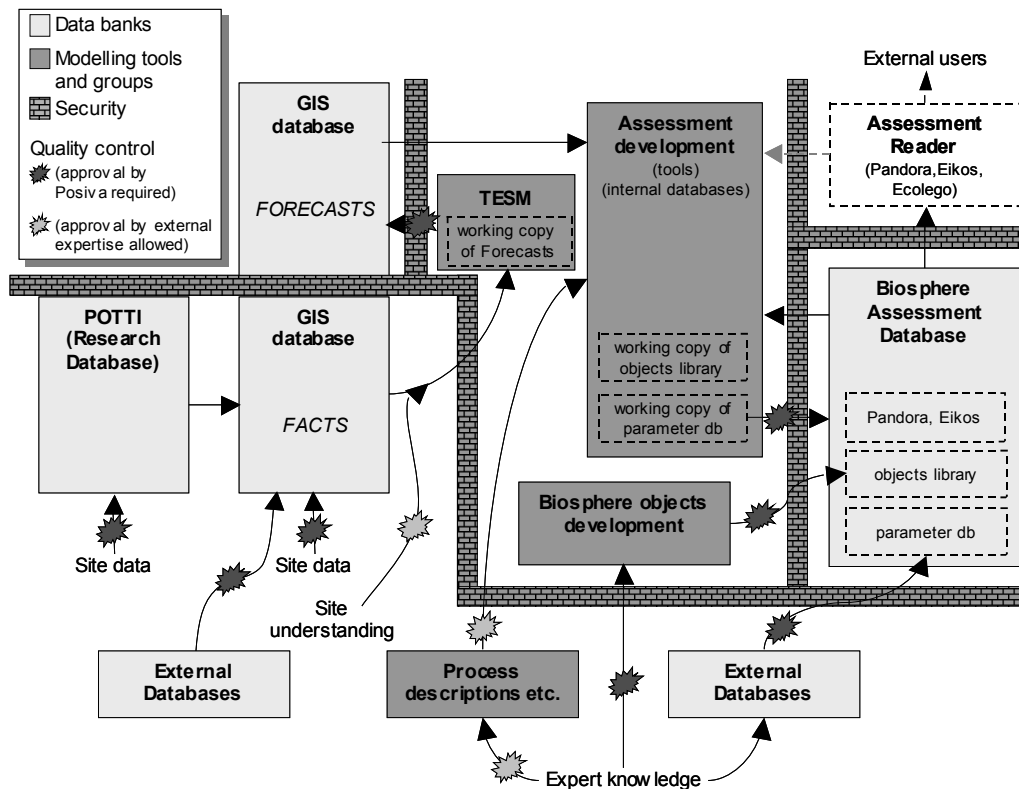
## METHODOLOGY

### Information infrastructure

From an information (in this context information means: scientific knowledge, assumptions, parameter data, processed data, calculation and modelling results) point of view, the BSA can be divided into three major components: the biosphere (site) description, prediction of future terrain and ecosystem (referred to as “forecasts” below), and modelling the temporal and

spatial distributions of radionuclide concentrations in the biosphere from potential releases from the repository. The forecasts are produced by the terrain and ecosystems development model (TESM), and radionuclide concentrations are the output from landscape modeling (LSM), which is a time-dependent linked transport model containing multiple biosphere objects represented by ecosystem-specific compartment models. The overall strategy of handling all information, and assuring the quality, used in these three components is based on employing the infrastructure presented in Fig. 1. The infrastructure regulates both flow and storage of information, and is the platform for assuring that the information used is quality checked (and approved), retrievable, un-compromised, and unambiguous.

The infrastructure is based on a network of databases and assessment development components (e.g., the LSM tools), and defines how information is allowed to be transferred, and the type of quality control required in the movement of information. In the whole Safety Case, a large part of the information will be managed by a number of specialised databases. Only in the BSA, several databases will be utilised, managed by Posiva or by external parties. Three databases managed by Posiva can be identified in Fig. 1: (i), the POTTI research database, (ii), the GIS database, and (iii), the BSADb.



**Figure 1.** Schematic presentation of the information infrastructure implemented in the BSA.

Principally, all data and results from the individual investigation or research tasks will be delivered and stored in POTTI. The GIS databases contain the information possible to present as maps, or the location data with relevant attributes. These two databases are intended for the different processing levels of site data. However, throughout the BSA, these data need to be post-processed. In order to maintain the difference, the site data used in the BSA is referred to as environmental data and the derived forecasts as biosphere data. In the BSA workflow, data from POTTI can be used either directly to assign parameter values or transported into GIS system of environmental data and co-processed, and further developed manually, or by using specific tools, into biosphere forecasts with attached data. Whilst the information in POTTI and the GIS database mainly will be used as input to supply the TESM with data, the BSADb will mainly supply the LSM with data. In addition to supply data, the purpose of the databases is also to store the results and ensure that the data used is quality assured.

### **The biosphere assessment database**

For the BSADb, detailed requirements on security, data integrity and (data) quality assurance have been outlined (Hjerpe 2006) and implemented. The requirements on security handle issues such as physical protection, redundancy and backup strategy. For data integrity, it is required that access to the database should be divided into a five-level user profile system. The five levels range from an anonymous “read-only” access profile, to a personal profile with full access to add and modify data entries. In addition to the five-level system, there is a user profile intended for reviewers involved in the quality assurance procedure. The quality assurance of data to be used in the BSA is a vital part of the Knowledge Quality Assessment (KQA) strategy, described in Ikonen (2006), and Hjerpe (2006). To be able to be confident in the overall quality of the BSA it is important to ensure that appropriate and reliable data has been used. To facilitate this, two features have been implemented in the database: *usability status* and *data history*.

**Usability status** – each data entry in the BSADb has a usability status (US) value, corresponding to a usability class, ranging from rejected-in-quality-check to the quality-approved data to be used in the assessment simulations (Table 1). Permission to change US is dependent on the profile of the user, e.g., level 1 and 2 users are not permitted to do any changes. The approach to populating the database is to permit a relatively large number of users to add data entries, resulting in a large amount of data at the US2 level. A smaller group of users are permitted to change the US to US3, after checking that the data entry is relevant and that the information given is complete. Sending the data for quality checking (US4) will be permitted by an even smaller group of users, mainly the members of the BSA Team, and approving data in the quality check (US5) can only be made by a reviewer. And finally, the formal quality approval of a data entry (US6) can only be performed by the BSA Coordinator (project manager of the BSA). The default requirement in the LSM is that only data from the BSADb with US6 will be allowed for usage.

**Data history** – When a field value for a data entry is changed (e.g., US, numerical value, numerical uncertainty, added references, etc.) the data entry get an inherent data history. All changes in the data history are tagged with time and full affiliation of the user performing the action. The full data history is retrievable.

**Table 1.** Denotation system for the usability status of data entries in the BSADb.

<b>Usability Status</b>	<b>Assigned to</b>	<b>Approved for Assessment</b>
US0	- Deleted <sup>(1)</sup> data	No
US1	- Data rejected in quality check	No
US2	- Default status for new and altered <sup>(2)</sup> data - Draft data - Data with insufficient information	No
US3	- Complete data, awaiting quality check procedure	No
US4	- Data undergoing quality check	No
US5	- Data accepted in the quality check awaiting QA decision (accept/reject)	No
US6	- QA approved data	Yes

<sup>(1)</sup> The information is not physically deleted from the database.

<sup>(2)</sup> When a data entry with US>2 is altered, the default action is to automatically decrease the data entry to US2.

### **Quality check procedure**

The quality check will be conducted by a reviewer approved by the BSA Coordinator, and will preferably be performed via a web-interface to the database. A special user profile has been created for reviewers, showing only the data entries for the reviewer to quality check (and data previously checked by the same user).

### **CONCLUSIONS**

This contribution has presented the BSADb, which is a solution for integrating the storage of data with a systematic quality assurance procedure. The BSADb will facilitate that all data used in the landscape modelling in future assessments are quality assured, fully traceable and transparent.

### **REFERENCES**

- Hjerpe, T., 2006. Overall strategy for handling parameters and data in the Posiva biosphere assessment. Posiva Working Report 2006-109. Posiva Oy, Olkiluoto, Finland.
- Ikonen, A. 2006. Posiva Biosphere Assessment: Revised structure and status. Posiva Report POSIVA 2006-07. Posiva Oy, Olkiluoto, Finland.
- STUK, 2001. Long-term safety of disposal of spent nuclear fuel. Guide YVL 8.4. Radiation and Nuclear Safety Authority (STUK).