Development of suitable information systems for early warning systems (EGIFF)

Breunig M. (1), Reinhardt W. (2), Ortlieb E. (2), Mäs S. (2), Boley C. (3), Trauner F. X. (3), Wiesel J. (4), Richter D. (4), Abecker A. (5), Gallus D. (5), Kazakos W. (6)

- (1) Institute for Geoinformatics and Remote Sensing (IGF), University of Osnabrück, Kolpingstr. 7, D-49069 Osnabrück, Email: mbreunig@uni-osnabrueck.de
- (2) Geoinformatics Working Group (AGIS), University of the Bundeswehr Munich, Heisenberg Weg 39, D-85577 Neubiberg, Email: Wolfgang.Reinhardt@unibw.de
- (3) Institute for Soil Mechanics and Foundation Engineering, University of the Bundeswehr Munich, Heisenberg Weg 39, D-85577 Neubiberg, Email: Cornad.Boley@unibw.de
- (4) Institute of Photogrammetry and Remote Sensing (IPF), University of Karlsruhe, Englerstr. 7, D-76128 Karlsruhe, Email: Wiesel@ipf.uni-karlsruhe.de
- (5) Research Centre for Information Technologies (FZI) at University of Karlsruhe, Haid-und-Neu-Str. 10–14, D-76131 Karlsruhe, Email: Andreas.Abecker@fzi.de
- (6) disy Informationssysteme GmbH, Stephanienstr. 30, D-73133 Karlsruhe, Email: Kazakos@disy.net

Abstract

At present the analysis and preparation of information are particularly critical points of an early warning chain. The responsible decision makers are usually confronted with huge amounts of structured and unstructured data. To enable reliable early warning, the available data must be pre-selected, analysed and prepared. The decision makers should be provided with a reliable and manageable amount of information for the warning decision and for taking preventive measures. In the introduced joint project »Development of suitable information systems for early warning systems« (EGIFF) components of an information system for early recognition of geological hazards will be developed and examined in a mass movement scenario. In particular the improvement of the information analysis and preparation is researched. Therefore techniques like GIS, numerical simulations, Spatial Data Mining, geodatabases and the application of linguistic methods will be combined. A particular focus will be put on innovative methodical investigations, which will serve to merge new technologies into the operational workflow of early warning and to

establish an efficient and reliable medium for distributing information to the responsible hazard/crisis managers.

1. Overall Objectives of the Project

The central components of an early warning system are the recognition of the threats, the assessment and evaluation of the danger, the dissemination and communication of the warning, as well as the public reaction to the warning (Smith, 2004). The effectiveness of an early warning system largely depends on the transformation of the event recognition into the report of warning to the population (Dikau and Weichselgartner, 2005). Obviously the analysis and information preparation are particularly critical points of the early warning chain. They provide a basis for the warning decision and the risk estimate for the extent of the natural event.

The objective of the joint project EGIFF is to improve the early warning chain through the conception and development of appropriate components for early warning systems. Hereby the analysis and preparation of information are the basis of the warning and the risk estimation of the forthcoming geological natural

event. The main research objectives of the project include:

- Conception of a distributed component architecture of an information system for early warning systems;
- 2. Geotechnical evaluation of mass movements with the help of available events;
- 3. Coupling of numerical simulations and GIS;
- Modelling and visualization of spatial relationships and their uncertainty as well as their extraction from textual and numerical data;
- Transfer of Data Mining and analysis methods to spatially referenced data (Spatial Data Mining), as well as processing of structured and unstructured data (free texts of disaster and damage messages);
- 6. Web-based alerting;
- Geodatabase support for handling data of different scenarios as decision basis for the geotechnical evaluation of mass movements;

- 8. Development of concepts and methods for the support of the evaluation of risks and their prototypically implementation and finally
- 9. Evaluation of the developed prototype on the basis of concrete geological data and application scenarios.

2. Application Area

A number of places in the German alpine upland and the Bavarian Alps seemed to be suitable investigation areas for exemplary land-slide simulations. Main selection criteria were the availability of detailed geographical data and sensor measurements for a longer period of time, a coherent and comprehensive geology to verify the gained methodology in a generalized way and a potential risk. After investigations of different landslide areas a part of the Isar valley in the south of Munich, next to Pullach and Neugrünwald (see Figure 1), has been selected for the further studies.

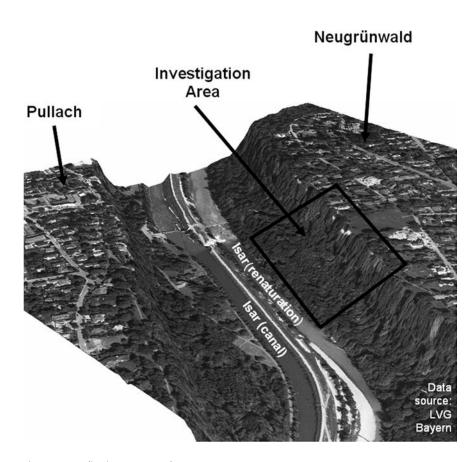


Figure 1: Application Area »Isarhänge«



Figure 2: Slopes in the Application Area »Isarhänge«

In this area, the height difference of the slope reaches up to 40 meters and the potentially endangered human infrastructure is located directly at the edge of the slope. Because of the high risk potential, the area is observed by the responsible authorities (Bayerisches Landesamt für Umwelt), and corresponding inclinometer and extensometer measurements and geodetic surveys are available.

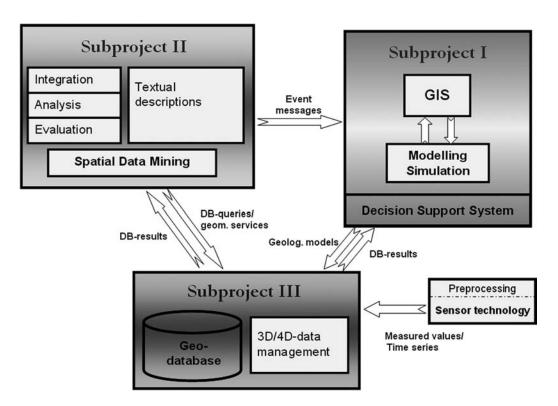


Figure 3: Overall Architecture

3. Description of the subprojects

The project is subdivided into three subprojects. Figure 3 illustrates these subprojects and their interaction. The following subchapters provide an insight into the objectives and applied methodologies of the subprojects.

3.1. Subproject I »Development of an interconnected information and simulation system«

Responsible: Conrad Boley, Institute for Soil Mechanics and Foundation Engineering (IBGN) and Wolfgang Reinhardt, Geoinformatics Working Group (AGIS) both University of the Bundeswehr Munich

3.1.1. Objectives

The main aim of this subproject is the conception, prototypically implementation and evaluation of an interlinked three dimensional Numerical Simulation System (SIMS) and a Geographical Information System (GIS). Its research contributions are, in particular, the development of coupled information and simulation models for mass movements with a user-friendly control of the complete system. Another focus is put on a more precise predictability and a more exact determination of exposure of slopes. Therefore a transferable numerical simulation algorithm for landslide movements will be realised within a simulation system. Thus, it is possible, based on existing field data (sensor- and GIS data, for example DTM) and virtual exogenous scenarios, to assess the slope stability, future system behaviour and potential risk scenario. As a result, a »communication« between the sensor measurements, GIS and the geotechnical model is enabled.

The interconnection between SIMS and GIS is schematically shown in figure 4. The process starts with the selection of relevant parameters (Input Data) which may have influence on the occurrence of landslides. These parameters include basically temperature, precipitation and slope geometry. The parameter transfer is controlled by the GIS. In the simulation system, the slope stability and deformation is calculated and, if possible, used for prediction of

the future deformation behaviour. Results of the simulation are parameters such as hazard potential, deformation and movement vectors and stability indices. These results are transferred to the GIS for visualization and, as far as possible, enrichment with other information. Furthermore, the data can be checked against rules to support the decision making process of the user.

The developed concepts for the interconnection of the GIS and the simulation system will be prototypically implemented and evaluated. With this prototype of a Decision Support System (DSS), it can be demonstrated how the linking of numerical simulations and GIS can improve predictions of landslide hazards and calculation of vulnerability, and how the uncertainty of and lacks within the data can be considered. Additionally, it will be verified how the output of simulations, enriched with GIS analytic methods, can optimize the decision making process for the prevention of catastrophic hazards.

Within the overall system architecture (see Figure 3), this DSS is integrated with the geodatabase (Subproject III) and the data mining system components (subproject II). The geodatabase provides the input data for the simulation and the modelling of slope stability and stores the simulation results. The event messages detected by the data mining system will also be incorporated into the decision support.

3.1.2. Concepts and Methodology In the following the main work packages of subproject I are described:

- 1. Concept of an overall architecture
 In cooperation with the project partner University of Osnabrück, a suitable architecture of the overall system consisting of SIMS, GIS, geodatabase, in-field installed sensors as well as other components will be developed. The architecture should enable a seamless communication between individual system components.
- Data integration and modelling
 The data models required for the numerical simulation and GIS will be designed. The

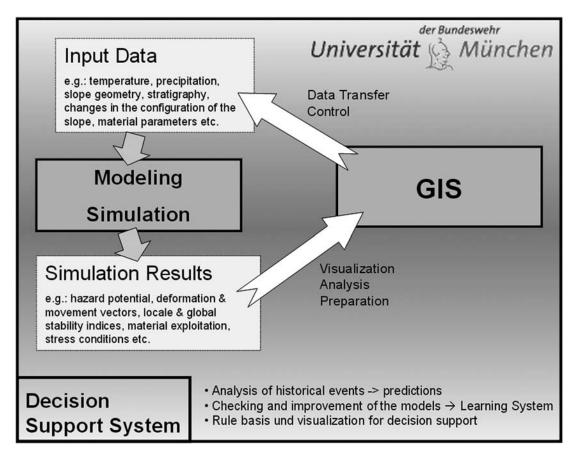


Figure 4: Interconnection between simulation system and GIS

main focus of the modelling process lies in the combination of stratigraphical, lithological and geotechnical data with geometric sensor data, long term physical series of measurements, as well as the requirements of visualisation and further processing of the simulation results. Subsequent to modelling, the available data of the test area will be adopted and integrated.

3. Conception of geotechnical models and their interpretation

Natural slopes often consist of several different soil types, whose mechanical properties have to be described with different constitutive equations. For a realistic characterisation of the deformation behaviour of different soil types, the validity of distinct constitutive equations for each relevant material has to be analyzed and assessed. Therewith the application of appropriate constitutive equations for dif-

ferent soil types will be enabled. To apply the best suitable constitutive equations, the required material parameters have to be identified and a concept for the determination of these parameters developed. The material parameters will be stored in a database, so that the final numeric model can access the required parameters.

For numeric modelling of landslides initialand boundary conditions have to be defined. These include the geometry of the slope as well as superimposed load and meteorological influences, etc. Due to the complexity of the actual initial- and boundary conditions, a simplification of the numerical model is inevitable.

Numeric modelling demonstrates not only the deformation of the slope but also the slope stability. For the determination of the slope stability failure criteria have to be developed, which may vary for different materials and different constitutive equations respectively. With the definition of these criteria the material properties, the state of stress and strain and the desired security level have to be considered, as well as the strain rate and stress level. Based on a time variant 4D-simulation, calibrated with backwards analysis for past events, probabilities for future scenarios can be calculated. By comparing predicted deformations with measured displacements, the model can be adjusted and improved.

4. Conception of the Decision Support System

The criteria for determining the stability of slopes enable the mapping of landslide prone areas and vulnerable zones. This means the slopes can be divided into homogenous areas according to the degree of actual or potential hazard. Furthermore, methods will be developed for refining and linking the simulation results with other information (e.g. from subproject II) and for visualizing the simulation events. In addition, the simulation results can be linked with decision rules in order to support and guide the user. For example, dangerous situations and adverse constellations that may be predefined and formalized in the system can be automatically revealed to the user as advice. The main goal is to establish an efficient and reliable medium for distributing information to the responsible hazard managers. Furthermore, it will be investigated if and how uncertainties of the data used in the simulation and subsequent processes can be modelled. In particular, for visualization and for the support of the user in the decision-making process these uncertainties should be recognizable, in order to allow for validation of the results by the user. Moreover it will be investigated if and how this validation of uncertainties can be done automatically. Another emphasis is put on the convenient presentation of the extensive simulation results to the user and the development of a user friendly and intuitive system control.

5. Definition of the interfaces between GIS/ geodatabase and the simulation component The main interfaces that need to be defined are for the transfer of the input data and the simulation results between the GIS and the simulation component, and for the connection of the GIS to the geodatabase. Therefore available international standardized interfaces will be examined and, if necessary, restricted, extended or adjusted. For the bidirectional access to the geodatabase, standardized interfaces of the OGC (Open Geospatial Consortium) will be taken into consideration. For the linking between the simulation system and the GIS also standards like HLA (High Level Architecture) will be analysed.

6. Prototype development

The developed concepts will be prototypically implemented and verified with data of the above mentioned application area. Practical tests will be carried out to determine the system feasibility, performance and reliability. Furthermore a transferability study shall analyse the applicability of the concepts and solutions to areas with similar and also varying geologic and boundary conditions and data availability.

3.2. Subproject II »Integration, analysis and evaluation of vague textual descriptions of geoscientific phenomena for early warning system support«

Responsible: Peter Lockemann, FZI Karlsruhe; Joachim Wiesel, Institute of Photogrammetry and Remote Sensing (IPF), University of Karlsruhe; Wassilios Kazakos, disy Informationssysteme GmbH, Karlsruhe

3.2.1. Objectives

The goal of subproject II is the collection, preprocessing and analysis of relevant structured and unstructured data (numerical and textual measurements, observations and descriptions of specialists and laymen (citizens)) related to natural hazards in form of disastrous mass movements, with a focus on development and application of novel computational methods aimed at a combination of results of analyses from heterogenic data sources, followed by a prototypical implementation as suitable components of an early warning system.

For this goal, the following central tasks have been identified:

- development and application of techniques and methods aimed at an automatic extraction of early warning-relevant information and spatial references from textual messages,
- transfer of suitable techniques and methods for automated data analyses and an application thereof in the domain of applied geosciences,
- development and application of novel methods aimed at a combination of results of analyses of structured and unstructur-ed data related to natural hazard phenomena of the type »disastrous mass movement«
- design of ergonomic user interfaces for the task of complex data analyses in context of early warning systems, with a focus on users with none or no profound IT knowledge

Besides concentrating on disastrous mass movements as the context for development and application of appropriate techniques and methods, a high degree of transferability to other natural hazard types or even analytical tasks in the domain of applied geosciences is regarded as desirable. The latter is also seen as part of a general objective to maximize scientific connectivity as well as a prospective commercial value of the project's results.

3.2.2. Concepts and Methodology

In the following the three main work packages of subproject II are described. They will be worked on collaboratively by FZI, IPF and disy Informationssysteme GmbH, Karlsruhe.

1. Conceptual work

In the first year of the project, the more geology-oriented and the more computer science-oriented project partners will come to a common understanding of the project's goals, and develop well-defined, effective and efficient ways of collaboration. Based on a deeper understanding of the application domain with its specific requirements, characteristics and potential problems, and in close cooperation with the other subprojects, a comprehensive system vision and architecture will be developed, including a set of accepted interfaces and – where necessary – shared data structures. On this basis, example-driven techniques and algorithms will be specified to be implemented later (year 2). In detail, we aim to focus on the following tasks:

1.1 Analysis of the application domain

In a first step of the analysis of the application domain, a glossary of central concepts and terms for the subproject partners will have to be developed in order to facilitate communication with other subprojects and with end users. At the same time, relevant functional and non-functional requirements from end users' perspective and constraints for the systems to be developed have to be identified.

1.2 Investigation of relevant prior work

In this task, focus is set on an investigation of relevant prior work within and outside the domain of applied geosciences and an analysis with respect to an applicability and usefulness in the given problem setting.

A first part of the work will include a review of SOKRATES, a system developed by FGAN/FKIE in Wachtberg (Schade and Frey, 2004). Relying on a combination of methods from natural language processing and techniques of knowledge representation, the system can be used to display events on a tactical map by means of limited automated interpretation of battle field reports. The combination of techniques employed by SOKRATES can serve as a methodological basis for parts of the system envisioned in subproject II. Due to differences in application context, a number of adjustments will have to be made, whereas parts of the system may be replaced by Semantic Web de facto standards or Open Source developments. A parallel thread of work will focus on an investigation of data mining techniques which could be transferred into the domain of applied geosciences in order to devise, develop and select suitable techniques, algorithms and tools (Lindner, 2005). In this context, assessment of algorithms and tools will also comprise an evaluation of the transferability/applicability of selected representation formalisms, methods and systems to other real-world applications, as well as their interoperability with established tools.

1.3 Collection and detailed analysis of input data

The next step in task chain will concern the collection and detailed analysis of available data for data mining algorithms to be developed. Available structured and unstructured data (text) will have to be collected, formally described and stored for application and test. Additionally, further work will address the issue of acquisition of expert knowledge needed for the interpretation of information as the basis for planning actions to avert damages of people and property.

1.4 Design of interfaces and specification of requirements for shared data structures and system components
In this task, focus is set on the design of interfaces and requirements for shared data structures and system modules in order to ensure interoperability with other subprojects, in particular with respect to specific requirements introduced with computational representation of three-dimensional spatial data and the representation of vagueness and uncertainty resulting from an analysis of textual spatial references.

1.5 Specification of systems to be implemented (end user perspective) The next steps, subsumed in subtask 1.5,

will deal with a specification of systems

and methods in the form of use cases, including a coordination of use cases with the other subprojects' use cases and an evaluation of the use cases by end users. Another task will be concerned with a general assessment of transferability, scientific connectivity and a commercial exploitability of the systems to be developed, with respect to different applications in the context of disaster management (early warning systems) or data analysis in the domain of applied geosciences.

1.6 Specification of algorithms to be implemented

In this task, three parallel threads of work will focus on the following subtasks:
In a first subtask, necessary changes and modifications to the SOKRATES system with respect to the given setting will be defined. For the task of the analysis of structured spatial data, suitable data mining techniques will be specified. Additionally, the task of a rough specification of a graphical user interface of the system will be addressed.

The tasks in the first year of the project (conceptual work) will be worked on collaboratively by all partners.

2. Implementation-technical realization

In the second year, the focus of subproject II will be set on an implementation of the algorithms previously specified and the necessary system adjustments. Depending on the progress and state of work in other subprojects, further functional aspects will be defined which will be implemented later. It is seen as reasonable to delay the specification of these functional aspects into this phase of the project until a condition is met assuming basic work in the other subprojects is completed and a first feedback concerning the results of implementation etc. is received. In particular, the following packages will be processed:

- 2.1 Implementation of text analysis methods In this work package, previously specified adjustments, modifications, and enhancements to the SOKRATES system will be prototypically implemented. Subsequently, the system will be tested, making use of the data collected.
- 2.2 Implementation of spatial data mining/ suitable methods for the analysis of structured spatial data In this work package, the techniques and methods previously specified will be implemented, adapted to, and integrated into the project's common software framework. Adaptations of existing techniques may also include handling of three dimensional spatial data, methods for efficient processing and techniques for processing data characterized by inherent vagueness or uncertainty.
- 2.3) Specifying mixed analysis methods and the integrated user interface Relying on results from 2.1 and 2.2, specifications of methods to be implemented in the third year of the project will be developed. A first part of the work will consist of devising methods for combined processing of structured data and postprocessed unstructured data (i.e. results of analyses of textual messages). In this task, central focus will be set on a more detailed investigation of the concept of »spatial coincidence« and certain aspects of methods for combined processing of heterogenic data, in particular related to a quantitative treatment of results under a common decision-theoretic framework. Secondly, an integrated user interface will be specified which will allow non-IT-experts to efficiently make use of the multitude of data sources and algorithms. Here, conceptual work related to problems of suitable representation of spatial data will be put into practice, including, but not limited to, a pragmatic choice of tools for formulating and testing hypotheses or for querying complex data and knowledge bases.

In context of visualization of spatial data, characteristics of the data related to reliability and precision of spatial references will have to be presented in a way satisfying requirements for adequate usability. In particular, previously established qualities like priority or urgency of information and their possible interpretations as well as scale-dependencies will be major aspects.

3. Evaluation

In the third project year, two major threads of work will be pursued. In the first subtask, implemented algorithms and systems will be documented, tested, evaluated with real-world data and validated against different defined use cases. This will be done in coordination with other subprojects, and in touch with international scientific community. In the second subtask, a parallel thread of work will be started, focussing on an implementation of combined analysis algorithms for structured and unstructured data. Additional work will include an implementation of ergonomic user interfaces, test-driven removal of errors and deficiencies based on results of parallel evaluation work, and finally an integration of the implementation with the other subprojects' results.

3.3. Subproject III »Geodatabase support for the geotechnical evaluation of mass movements«

Responsible: Martin Breunig (Consortium coordinator), Institute for Geoinformatics and Remote Sensing (IGF), University of Osnabrück

3.3.1. Objectives

In this subproject, the primary geological data, provided by the Bayerisches Landesamt für Umwelt, will be modelled and managed in a geodatabase management system. The management of the geological spatial and timerelated primary data enables its usage for analysis, simulation, and 3D visualization at any time. In a second step, finite element models and their results will be stored in the

geodatabase. This supports the geotechnical evaluation of mass movements through the assistance of suitable database queries to be developed during the project. Finally, the development and implementation of geometric/topological and time-related operations as input for spatial data mining methods shall provide new spatio-temporal patterns for the recognition of mass movement hazards. The developed methods will be evaluated on the basis of concrete geological application data from the selected area of investigation.

3.3.2. Concepts and Methodology

Geodatabases are used for archiving and for providing fast and efficient access to geodata. This enables for the reuse of geodata, e.g. for upcoming new natural events. In addition, a geodatabase can also manage models and their results and even become active, e.g. by computing geometric intersection queries between a set of geoobjects.

The subproject pursues three objectives:

- 1. Geometric and topological management and 3D visualization of existing geological primary data and time-dependent data.
- Management of FE models and their results for the use in geotechnical evaluations of mass movements.
- Development and implementation of geometric/topological and time-dependent operations as input for spatial data mining methods.

For the first objective, selected test data from the investigation area (see section Application Area) will be examined and managed in a geodatabase in close cooperation with the geotechnical group Boley (subproject I). This includes a first interpretation of the primary data. In particular the representation of discontinuities (e.g. with already existing faults in the slope) will be taken into account. This goal can use the developments implemented in (Breunig et al., 2005) and benefit from this development of geodata management for complex geological models (surface layer and volume models). The available data structures and operations have to be examined for the suitability of the special requirements by mass

movements. Where necessary these need to be adapted and extended to fit the new special requirements.

The second goal basically concerns the management of the input data and results from 3D model calculations within the geodatabase. This will be done in close cooperation with the geotechnical group of Boley. With the FE model calculation a second kind of data interpretation will be realized. The management in a geodatabase creates new possibilities of further use of such modelling results and for overlaying with further information like utilization data. For example the information »How is the area under the endangered slope used?« is relevant for early warning and therefore needs to be coupled with the model calculation. Homogenized and/or converted geometric data models from the model calculations and the geodatabases have to be examined. Also the information computed for each cell (scalars and vectors) of the FE model has to be considered.

Furthermore, the results of a larger number of model calculations can be stored in the database and on later demand they can be compared and queried from different points of view. For example, versions or scenarios of model calculations can be stored and compared with each other. Additionally, taking the use of a »4D model« into consideration, i.e. different time steps of the FE model, the data for the temporal analysis of mass movements can be managed in the geodatabase.

It is also of importance to store error margins and their accuracy in the database. Here we refer to subproject II. Intervals need to be stored in the database. Passing over or falling short of these intervals will cause an action of the database. From the database side, further geometric/topological constraints and/or integrity conditions can be realized and included in the analysis of mass movement endangerment. The thematic information connected with the geometry has to be extended in a way that the model calculation receives information about the different semantics of geometrically equal geometries (e.g. fault surfaces versus boundary surfaces). In this context we

speak of »thematic colouring« of the model, which indicates to the user, for example, at which locations danger zones exist. Afterwards the results are used in the data mining component (subproject II).

The third goal consists of developing geometric/topological and time-dependent database operations, which support the geotechnical analysis of mass movements. Of concern are not only simple range calculations between point geometries (like it is the case in classical 2D GIS buffer operations). Also the movement/speed of complex 3D geometries and the direction of the movement have to be considered. The implemented operations can be used afterwards in the data mining methods, in order to determine a priori unknown spatial patterns from the combination of spatial and non-spatial attributes for upcoming mass movements.

Methodically, one of the principal purposes of object-oriented software technology, the reusability of individual components, will be used during the software development, in order to react as flexible as possible to new requirements of future natural events.

4. Outlook

The concrete research results of the joint project will be:

- Geotechnical evaluation of mass movements;
- Reusability of the geological and geotechnical primary data and models relevant for early warning through management in a geodatabase;
- Employment of spatial data mining methods as support of the warning decision for natural events;
- Development of new methods for the visual representation and decision support of early warning relevant information from alphanumeric data and texts including fuzzy and incomplete data;
- Prototypical implementation of the developed methods and concepts and evaluation with the application scenario »mass movements« (e.g. landslides);
- Improvement of the reliability of alarms for the early warning of natural events/hazards by employment of the developed methods.

The evaluation of the project results on the basis of concrete application scenarios and their economic utilization by the partner company are part of the project. The close cooperation with the Bayerisches Landesamt für Umwelt guarantees the direct application of the research results.

Acknowledgements

The funding of the research project »Development of suitable information systems for early warning systems« (Entwicklung geeigneter Informationssysteme für Frühwarnsysteme) by the German Ministry of Education and Research (BMBF) by grant no. 03G0645A et al. within the framework of the geotechnology initiative (http://www.geotechnologien.de) is gratefully acknowledged. The responsibility for the contents of this publication is by the authors. We also thank the Bayerisches Landesamt für Umwelt (LfU, www.lfu.bayern.de) and the Bayerisches Landesamt für Vermessung und Geoinformation (LVG, www.lvg.bayern.de), for providing data for the application area.

References

Breunig, M.; Bär, W.; Häussler, J.; Reinhardt, W.; Staub, G.; Wiesel, J. (2005): Advancement of Mobile Spatial Services for the Geosciences. In: Data Science Journal, International Council for Science (ICSU)

Dikau, R., Weichselgartner, J. (2005): Der unruhige Planet – Der Mensch und die Naturgewalten, Wissenschaftliche Buchgesellschaft, Darmstadt, 191 S.

Lindner, G. (2005): Algorithmenauswahl im KDD-Prozess. Dissertation an der Universität Karlsruhe, Institut AIFB, 2005.

Schade, U.; Frey, M. (2004): Beyond information extraction: the role of ontology in military report processing. In: Bachberger, E. (Ed.): Proc. of KONVENS 2004, Vienna, Austria.

Smith, K. (2004): Environmental Hazards: Assessing Risk and Reducing Disaster, London.