

# THE *gein*<sup>®</sup> 2.0 INFORMATION BROKER FOR ENVIRONMENTAL AND GEOSPATIAL DATA

Thomas Vögele<sup>1</sup>

Fred Kruse<sup>1</sup>

Oliver Karschnik<sup>1</sup>

<sup>1</sup>UDK/GEIN Coordination Group  
Ministry for the Environment  
Lower-Saxony, Germany  
kug@numis.niedersachsen.de

## ABSTRACT

Within the framework of a technical and organizational re-design, the German Environmental Data Catalog (Umweltdatenkatalog – UDK) and the German Environmental Information Network (*gein*<sup>®</sup>) are in the process of being integrated into one new information system, *gein*<sup>®</sup> 2.0. One of the primary goals of this new system is to support government agencies in Germany to fulfil the obligations defined by EU Directive 2003/4/EC on public access to environmental information. Because a large percentage of environmental data are geospatial by nature, the UDK is already used as a data catalog within the national SDI. Implementation of relevant international standards (e.g., ISO 19115, ISO 19119 and ISO 19139) and specifications (e.g., OGC CSW 2.0) will further improve the capabilities of the system and be the basis for a full integration of *gein*<sup>®</sup> 2.0 into the emerging national and international geodata infrastructures.

**KEYWORDS:** Metadata, catalog service, portal, data harmonisation, geodata infrastructure

## INTRODUCTION

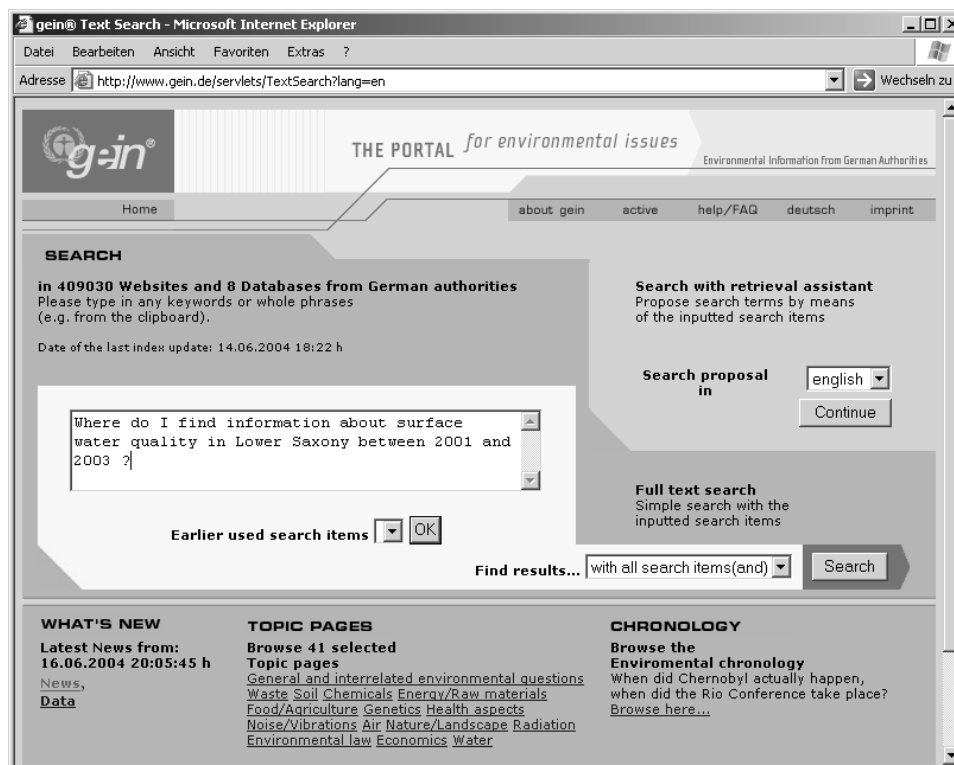
The *German Environmental Information Network gein*<sup>®</sup> (KUG 2004) is the official information broker for environmental data held by German government agencies. The system was developed between 1998 and 2000 within the framework of a research and development project hosted by the German Federal Environmental Agency (Umweltbundesamt - UBA) in Berlin (Bilo and Streuff 2000). Version 1.0 of *gein*<sup>®</sup> was introduced in the year 2000 at the world trade fair Expo'2000 in Hannover. Currently, the system is in a process of re-design. The focus of this re-design is on the implementation of EU Directive 2003/4/EC on public access to environmental information (EC 2003). On an organizational level, this implies a better integration and synchronization of *gein*<sup>®</sup> within the emerging national and European frameworks for environmental and geospatial data. On a technical level, this means among others the implementation of the relevant standards and specifications, including ISO standards (e.g., ISO 19115, ISO 19119, ISO 19139) and OGC specifications (e.g., CSW-2).

## GEODATA IN *gein*<sup>®</sup>

As of spring 2004, the *gein*<sup>®</sup> 1.1 information broker (Figure 1) provides access to more than 390.000 individual web-pages hosted by 90 information suppliers. In addition, the system has access to a number of environmental (meta)databases with a total of over 500.000 information objects. Among them are the UBA's *Geographic Information System for the Environment (Geografisches Informationssystem Umwelt, GISU)*, the *Database of Environmental Publications (Umweltliteraturdatenbank, ULIDAT)*, the *Catalog of*

*Environmental Objects (Umweltobjektkatalog, UOK)* run by the federal state of Bavaria, and the *Catalog of Environmental Data Sources (Umweltdatenkatalog, UDK)*.

All information sources in *gein*<sup>®</sup> have in common that they are “official”, i.e., managed and maintained by government organizations responsible for environmental information. Membership in *gein*<sup>®</sup> is restricted to state and federal agencies, but it is planned to also include agencies on a community level in the future. This rather selective membership policy ensures a high quality of information providers, and consequently the high quality of environmental data accessible through *gein*<sup>®</sup>.



**Figure 1:** The *gein*<sup>®</sup> (version 1.1) start page

*gein*<sup>®</sup> was not designed as an information broker specifically for geospatial data. Nevertheless, most of the web pages and database entries that are accessible through *gein*<sup>®</sup> do have a geo-reference, and many of them refer explicitly to geospatial data, in most cases digital maps. With the GISU (UBA 2004a), a metadata information system explicitly targeted at geospatial data is accessible through *gein*<sup>®</sup>. In addition, almost one third of the information objects stored in the UDK refer to either geospatial data or geo-services.







## THE UMWELTDATENKATALOG (UDK)

Among the metadata bases included in *gein*<sup>®</sup>, the UDK is probably the most important one. In the next and completely re-designed version of *gein*<sup>®</sup> (*gein*<sup>®</sup> 2.0, expected to be released by the end of 2005), both *gein*<sup>®</sup>

and UDK will be merged into one integrated application. The UDK will become the Meta Information Broker (MIB) and core module of the new system.

The UDK in its current version is a metadata catalog that has been used since 1991 as a metadata registry for environmental information (Swoboda, Kruse et al. 2000). It was developed as an information system to be used mainly by government agencies in Germany and Austria. The purpose of the system is to give information providers a tool to create registries of relevant data holdings and key responsible parties. The metadata records were designed to answer questions like: *Who holds where what kind of data in which format?*

From the beginning, the thematic focus of the UDK was on environmental data. Because adequate international standards for environmental metadata were not available at the time, a specific UDK metadata model was developed. In this model, individual metadata records are stored as *addresses* (describing individuals and organizations) and the so-called *UDK objects*. Among the main attributes of an UDK object are an object identification number, an object name, a general description of the object, descriptive key words (taken from an integrated environmental thesaurus), as well as a description of the temporal and spatial coverage of the object.

	1.) data collection / data base
	2.) service / application / information system
	3.) document / report / publication
	4.) geodata / digital map
	5.) organisational unit / assignment
	6.) proposal / project / programme

*Table 1: UDK object classes*

## INTEGRATION INTO THE NATIONAL SDI

The UDK uses six different object classes to categorize UDK metadata records (*Table 1*). Only two of these object classes, namely class 2 (*service/application/information system*) and class 4 (*geodata/digital map*) are intended to hold information about geospatial information and related services. This shows that the UDK was initially not designed to reference geospatial data and web services in particular. However, with the ongoing development of regional and national spatial data infrastructures in Germany, it became obvious that the UDK could provide a valuable tool for this purpose. For one, the UDK is well established in almost every German federal state. This includes the necessary organizational infrastructure and the funding necessary to maintain and manage the catalog. Secondly, the UDK object model proved to be flexible enough to handle geospatial data and geo-services. In fact, as of spring 2004, mapping agencies in three federal states (Hamburg, Hesse, and Lower Saxony) already do use the UDK to reference their geospatial data holdings. Therefore it was decided to include the

UDK into the evolving national spatial data infrastructures in Germany, namely the GeoMIS.Bund<sup>®</sup> (IMAGI 2004a) and the GDI-DE<sup>®</sup> (IMAGI 2004).

However, to enable interoperability with other data catalogs in the GDI-DE<sup>®</sup>, and with other spatial data catalogs in general, two modifications were necessary:

1. The UDK metadata model had to be modified to be compatible with the relevant international standards, i.e. ISO 19115, ISO 19119 and ISO 19139, and
2. the system had to be equipped with an OGC compatible catalog service interface.

The necessary modifications of the UDK data model are under way and will be finalized with the next release of the UDK (UDK Version 5.0), which is expected in fall 2004. To minimize the effort necessary to migrate large amounts of legacy data from UDK version 4.3 to version 5.0, only the two “geospatial” object classes (classes 2 and 4 in Table 1) were modified.

In the near future, version 5.0 of the UDK (and consequently, *gein*<sup>®</sup> 2.0) will also be equipped with an OGC compatible catalog service interface. The interface will be based on OGC’s CSW 2.0 specification and on an application profile currently being developed by a German special interest group. This group brings together representatives from government agencies and industry on a national level. The purpose of the application profile is to develop an implementation-level specification for catalog services that can be used on a national, and potentially an international level.

## **ACTIVE DISSEMINATION OF GEOSPATIAL DATA**

Directive 2003/4/EC calls for the “active and systematic dissemination to the public” of environmental information. In the context of *gein*<sup>®</sup> and the UDK, we interpret “active dissemination to the public” as providing the technical means to directly access environmental information and data through the Internet. With respect to geospatial data, both systems are so far limited to providing indirect access through metadata descriptions of data sources, or through pointers to the URLs of the respective web-applications. In the new integrated system *gein*<sup>®</sup> 2.0 it will be possible to directly access the actual digital map behind a respective metadata reference. For this purpose, an OGC-compatible Web Mapping Service (WMS) and the respective Map Viewer client will be included in the system. For metadata objects that refer to services implemented as OGC compatible WMS, the *gein*<sup>®</sup> 2.0 WMS will then be able to both visualize the digital maps and to access the attribute data behind them. The general architecture of this WMS will use the specification developed by the OGC for cascading web mapping services. By implementing this architecture, the *gein*<sup>®</sup> 2.0 WMS will be able to access multiple digital maps simultaneously and to visualize them in a single, integrated view.

## **ENHANCED INFORMATION RETRIEVAL SERVICES**

*gein*<sup>®</sup> is equipped with enhanced search tools that help users to find relevant information faster and more efficiently. For this purpose, *gein*<sup>®</sup> makes use of the Semantic Network Services (SNS), a web-based service provided and managed by the UBA (UBA 2004). SNS integrates an environmental thesaurus with almost 40.000 terms, a geo-thesaurus with nearly 50.000 geographical objects, and a synopsis of historical and contemporary events that have affected the environment into a topic map. Based on the SNS and with the help of an integrated research assistant, *gein*<sup>®</sup> supports the specification of information requests that have a thematic, spatial, and temporal component. The SNS is also used to analyse the web pages registered with *gein*<sup>®</sup>, and to automatically create the respective context-based thematic, spatial and temporal metadata.

## SUMMARY AND OUTLOOK

Both the UDK metadata catalog ([www.umweltdatenkatalog.de](http://www.umweltdatenkatalog.de)) and the *gein*<sup>®</sup> information broker ([www.gein.de](http://www.gein.de)) play an important role for the access to and the distribution of environmental and geospatial data held by government agencies in Germany and Austria. The first version of the UDK was introduced in 1991, and *gein*<sup>®</sup> 1.0 went online in the year 2000. With the next release of *gein*<sup>®</sup> (*gein*<sup>®</sup> 2.0), both tools will be merged into one integrated system. This new system will be specifically designed to help government agencies to fulfil the requirements of EU-Directive 2003/4/EC on public access to environmental information.

A significant percentage of the data referenced in the UDK metadata catalog and *gein*<sup>®</sup> are geospatial data. In addition, a number of mapping agencies in Germany already use the UDK as a registry for their geospatial data holdings. Three of these registries (i.e., the catalogs of the federal states of Hesse, Hamburg, and Lower Saxony) are already linked to emerging national spatial data infrastructure (i.e., GeoMIS.Bund<sup>®</sup> and GDI-DE<sup>®</sup>). Our intention is to fully integrate the UDK (and ultimately *gein*<sup>®</sup> 2.0) into the national SDI. The necessary modifications (i.e., implementation of the relevant ISO standards and OGC specifications) are under way, with completion expected for the end of 2004.

## BIBLIOGRAPHY

- Bilo, M. and H. Streuff (2000). Das Umweltinformationsnetz Deutschland - GEIN2000 - Fachliche Anforderungen an ein Forschungs- und Entwicklungsvorhaben. 3rd workshop Hypermedia im Umweltschutz, Ulm.
- EC (2003). "Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on Public Access to Environmental Information." Official Journal of the European Union L 41: 26 - 32.
- Günther, O., H. Lessing, et al. (1996). UDK: A European Environmental Data Catalog. Third International Conference in Integrating GIS and Environmental Modeling, NCGIA, Santa Barbara (USA).
- IMAGI (2004). *GDI-DE® Geodateninfrastruktur Deutschland*. Interministerieller Ausschuss für Geoinformationswesen (IMAGI), Frankfurt. Last access: 14.06. 2004. [http://www.imagi.de/de/gdi\\_de/f\\_gdi\\_de.html](http://www.imagi.de/de/gdi_de/f_gdi_de.html)
- IMAGI (2004a). *GeoMIS.Bund*. Interministerieller Ausschuss für Geoinformationswesen (IMAGI), Frankfurt. Last access: 14.06. 2004. <http://www.geomis.bund.de>
- KUG (2004). *German Environmental Information Network - gein*. Koordinierungsstelle UDK/GEIN (KUG), Hannover. Last access: 14.06. 2004. <http://www.gein.de>
- Swoboda, W., F. Kruse, et al. (2000). Harmonisierter Zugang zu Umweltinformationen für Öffentlichkeit, Politik und Planung: Der Umweltdatenkatalog UDK im Einsatz. 14. Internationalen Symposium "Informatik für den Umweltschutz", Umweltinformatik '00, Bonn, Germany, Metropolis Verlag.
- UBA (2004). *Semantic Network Services (SNS)*. Umweltbundesamt (UBA), Berlin. Last access: 14.06. 2004. <http://www.semantic-network.de/home.html>
- UBA (2004a). *GISU - Geografisches Informationssystem Umwelt*. Umweltbundesamt (UBA), Berlin. Last access: 14.06. 2004. <http://osiris.uba.de:8081/gisu/start>