GEODATA INTELLIGENCE FOR INTEROPERABLE DATA MANAGEMENT WITH THE ENVVISIO METHOD - ENVVISIO-GI

P. Poßner 1, J. Großmann 2, H. Rudolf 3, R. Kaden 4

1 Chair of Surveying and Geoinformatics, Erfurt University of Applied Sciences, Erfurt, Germany - pascal.possner@fh-erfurt.de
2 Simplex4Data GmbH, Dresden, Germany - janik.grossmann@simplex4data.de
3 Simplex4Data GmbH, Dresden, Germany - heino.rudolf@hrd-consulting.eu
4 Chair of Surveying and Geoinformatics, Erfurt University of Applied Sciences, Erfurt, Germany - robert.kaden@fh-erfurt.de

KEY WORDS: Data Management, Data Modeling, Data Integration, Interoperability, INSPIRE, BIM

ABSTRACT:
The objective of the envVisio method is to create generic tools and a flexible data backbone for the preparation, storage, linking and provision of complex and multi-thematic municipal geodata. Although geographic base data of the federal states in Germany are represented and processed in standardized and established data models and systems, there is also a large number of additional geodata such as environmental data and mobility data that are not managed in a uniform manner. Specifically, the envVisio method is about the further development of the model approaches for geospatial and environmental data provision. It is a new type of modeling method that has already been tested in individual state environmental authorities for interdisciplinary cooperation between the specialist departments. The development of technical solutions is intended to reduce the effort required to provide municipal geospatial data and to promote higher usability of public data. The aim of this article is to present the methods and the use cases to be implemented.

1. INTRODUCTION

Due to their high level of informative value, municipal mobility and environmental data serve as a fundamental basis for regional, intermodal mobility concepts and the implementation of environmental sustainability goals. At the moment, however, there is no platform that enables the integrated provision of both municipal and federal mobility and environmental data in a standardized way in Germany. The particular challenges are briefly outlined below.

1.1 Motivation

During the INSPIRE implementation process, municipal mobility and environmental data are to be prepared in a standardized manner by the municipalities themselves or by the competent state administrations, depending on the legal requirements and municipal impact (European Commission 2007). However, municipalities currently must make a great deal of personnel and material expenditure to offer their extensive and multi-layered geodata in accordance with the currently valid standards (including INSPIRE). These efforts often cannot be met, so that municipal data are not available for potential users or data can only be obtained with increased effort for integration into IT systems. This also significantly hinders the digitization process in Germany.

Mobility is one of the priority issues of both cities and districts. The district roads, for example, which are subject to road construction in the German districts, make up around 40% of the supra-local road network and thus make an indispensable contribution to maintaining traffic in the rural regions. In contrast, the cities in particular are faced with the challenge of reducing the impending traffic collapse and the resulting inner-city environmental damage (BMVI 2018).

Given this initial situation, many municipalities are currently in the process of developing and implementing multimodal transport concepts. Possible partial solutions are the bundling of transport offers using networked digital timetable information, innovative intermodal mobility offers (networking of different modes of transport), digital payment systems (e.g., electronic ticket for public transport, mobile phone ticket, mobile apps and online platforms), intelligent parking space management and an intelligent one Traffic control using real-time data from the modes of transport. In this context, the German District Association is currently developing a handbook on “Digitization and Mobility”.

The multitude of mobility-related infrastructure data forms the basis for developing networked mobility concepts. However, the required mobility data is currently only available at different locations due to the different responsibilities and is collected in inconsistent data formats.

The main challenge is therefore to link, harmonize and use this data across all levels. In essence, it is about the further expansion of the federal spatial data infrastructure with regard to the flood of available mobility data (municipal, state, federal data), in which the data of the Federal Ministry of Transport and Digital Infrastructure should also be integrated.

1.2 Objectives

The main objective of the envVisio method and the ongoing envVisio-GI project is to create tools for the preparation, storage, linking and provision of relevant mobility and environmental data. This information serves as a foundation for regional intermodal mobility concepts and the implementation of environmental sustainability goals. In the medium to long term, the project aims to help secure mobility in rural areas,

---

1 Infrastructure for Spatial Information in Europe, Directive of the European Commission
improve the planning of transport and the public transport network, and make an important contribution to climate protection.

In contrast to the governmental geospatial base data, which are managed and available in standardized and established data models and formats, there are no standardized structures for the harmonized management and provision of municipal environmental data (e.g. forest, climate, water). The current data provision for this information is therefore often insufficient to answer socially relevant questions (e.g. on the subject of climate change), such as analyzes of forest conditions, across municipal borders. The envVisio method aims to harmonize thus data and provide interoperable data provision.

Another objective of the project is to support the municipalities in the implementation of reporting obligations and the fulfillment of relevant requirements of consumers in a substantial way. In this context, it is planned to create extensive documentation during project implementation, consisting of guidelines, processes, (partial) concepts and methods.

With regard to the technical implementation, the envVisio-GI project is about the further development of the model approaches for geospatial and environmental data provision with the help of the envVisio method. The project addresses both current and future challenges of environment and mobility issues in urban areas as well as the provision of geodata for Building Information Modeling (BIM) and the INSPIRE reporting (Annex 3 topics).

With a special communication and data processing system, the data is evaluated in a standardized manner and made available to appropriate user groups. The project aims to provide data efficiently and reduce the operating costs of municipalities by linking information via a technical platform. The solutions will significantly reduce the effort required to provide municipal data and thus promote the better valorisation of public data. For this purpose, the documentation of the project will also describe the processes and methods of data management (EnvVisio-GI Project Team 2020).

2. RELATED WORK

2.1 INSPIRE as a starting point

In addition to the increased effort for standardized preparation of data, the complicated data structures of INSPIRE also make the effective use of the data and their interoperability more difficult. A lot of effort is required on the part of the municipalities to review, prepare and process data content. The individually separated annex topics are often not sufficient to answer complex questions that go beyond the horizon of a single topic. An example from Chapter 6.1 Ecosystem Models would require data from the INSPIRE Annex topics Geographical Names, Hydrography, Protected Sites, Land Cover, Geology, Soil, Land Use, Human Health and Safety, Environmental Monitoring Facilities, Natural Risk Zones, (European Commission 2021) and so on. Bringing all these complex INSPIRE data structures into a singular system would mean considerable expenditure of time and money for a local administration, provided that the data is actually available in the corresponding data formats. This problem inspired the development of the envVisio method, as a possibility to represent all real-world objects and their related processes in a common system as described in the previous chapters.

2.2 Linked Data and Semantic Web Approaches

The approach of linking different information and data domains with one another is not a unique to the envVisio method. One possibility for establishing and using such connections is found in so-called “linked data” approaches. These use the semantic web technologies OWL (W3C Web Ontology Lanugage 2012), RDF (W3C Resource Description Framework 2014) and RDF Schema (W3C RDF Schema 2014). Information domains are described in the form of ontologies in OWL. Using these ontologies, information and data are translated from the original into domain and data format in RDF format and thus into so-called triples of subject, predicate and object. To create a higher information potential, links are created between data from different information domains, which serve as a bridge between the originally heterogeneous source data. The combined information content then enables complex analyzes and data queries, for example using the query language SPARQL (W3C SPARQL 2013).

There are currently examples of the use of these linked data approaches for semantic linking of BIM and GIS data. The technical report “A Semantic web platform for BIM-GIS integration” (Hor A. 2015) explains the generation of an integrated BIM-GIS linked data model. The linkage between BIM-RDF and GIS-RDF is achieved through SPARQL and results in a unified BIM-GIS RDF graph.

The Article “Linking BIM and GIS Standard Ontologies with Linked Data” (Hbech E., Roxin A. 2020) discusses possible standardized ontologies for the BIM and GIS domains to ensure knowledge continuum between these domains and to enable seamless reasoning for complex applications such as combining indoor and outdoor navigation.

In comparison to these approaches the envVisio method is more of a “closed off linked data light” concept. Similar to the RDF of linked data, the envVisio method aims to provide a “neutral ground” for different domains of information outside of proprietary, domain-specific software products and data structures and thus to minimize the loss of information during data transformation and mapping processes between the original domains. Likewise, the envVisio method tries to create a correct mapping of the real world in the envVisio Simplex component, similar to how ontologies are described in the OWL format. In contrast to the linked data and semantic web approach, envVisio is thematically specified, namely environmental and geographic information. A large spectrum of data can also be mapped within this thematic direction, as can be seen from the use cases. Any linking with resources that are not yet supported by the envVisio method is not possible to the same extent as the Semantic Web offers.

3. THE ENVVISIO METHOD

The concept of the envVisio method is not the classification of spatial object classes, as is the case with most spatial data or environmental information systems, but the understanding of the processes within and between these objects (Rudolf 2018). In terms of system analysis, this means that an attempt is made to name procedures and processes that determine the interrelationships between the system elements. The subject of these processes is, for example, substances, energy, money, knowledge / insights, information (i.e. categories that are stored, converted or transported).
3.1 EnvVisio Double Tetrahedron

From the system analysis it is now necessary to derive an approach for the formation of object classes. In terms of graph theory, we find nodes and connections in a wide variety of data systems. The nodes are the components in the systems to be mapped, here referred to as “entities” (real world objects). The connections describe processes in the system. Since data has to be processed again and again for such processes, a further base class “environmental processes” is defined. By deriving the object classes to be specifically considered for the respective task from these two basic classes and creating any links between them, it is possible to map the processes and relationships in any system. Since the environmental processes also include anthropogenic processes, human activities (e.g. use of natural resources, supply and disposal, construction and operation of technical systems / facilities) are included in this approach. An important aspect is still missing in the basic classes: We humans control and regulate these processes. A third basic class is therefore defined: human actions to assess, measure, control or regulate the ecosystem (“AMCR processes”).

![Figure 1: EnvVisio Model](Simplex4Data Image Copyright 2021)

Data is often collected on these actions and entities: target values, measured values, limit values and other characteristics of entities. Another base class “Values” is defined for these values. If the AMCR processes are directed towards the entities, then in general state-describing values are determined; if they target environmental processes, process-describing values are obtained. Above this triangle of entities, environmental processes and AMCR processes are metadata, such as centrally provided spatial references for geo-data objects, or information on the time and methods of the data collection. The base classes of entities, environmental processes and AMCR processes inherit basic structures from a base class “object” to ensure easy interoperability between the classes. The result of the modeling forms a double tetrahedron (Figure 1) (Rudolf 2018).

3.2 Real World Objects and their Images in envVisio

There is a consensus among most developers that a real object should be represented in a model by exactly one actual instance of an object class. This is the only way to ensure data consistency and save us from overgrown collections of data. This is often not in conformity with the practically common GIS, because these manage objects according to their geometry. This means that a real object is stored for its various geometrical characteristics in separate object classes (and thus multiple times), e.g.:

- A street as a line or area object class, depending on the scale range
- Buildings and other structures as point, line or area object classes, depending on how it is displayed

In the envVisio model, object definitions are not determined by maps or other representations of reality. An object class is understood as an image of a set of real objects and not as an image of images. Therefore, a consistent distinction is made between homogeneous data storage and the views on it (Rudolf 2018).

3.3 Modelling in 4D

In principle, the following applies to the envVisio model:

- Environmental processes and AMCR processes happen in a certain place and at a certain time
- Objects are managed historically

This means that changes to the object (e.g., of attributes) are managed, which leads to historical data records for an object. The historization problem is only solved when the connections between the data are clear. This means that all links that are valid at that point in time are to be shown for a historical instance.

In addition to the structure of the model, functions are used for the historization that continue concrete instances historically, i.e., generate successors, allow these to refer to the predecessors and take over all assigned links. Then both the attributes and the links to the instance can be edited and changed, and everything is retained with the predecessor (Rudolf 2018).

4. SYSTEM ARCHITECTURE

envVisio is a data warehouse that brings together and harmonizes data from a wide variety of sources, formats, specialist processes and applications. This data warehouse and its components form part of the system architecture specified for envVisio-GL.

The underlying data model is tailored to environmental and geodata, whereby the entire spectrum of our environment can be mapped: objects under consideration (e.g. geo-referenced data, traffic, supply and disposal networks, (protected) areas, development plans, buildings), processes in our environment (e.g. emissions, immissions, malfunctions, accidents, production and consumption), actions for evaluation, measurement, control and regulation (e.g. permits, measurements, planning) and values (e.g. measured value series, Grid maps, evaluation
envVisio does not compete with specialist applications, but places the data of the individual specialist processes in a common context with mutual references. This creates a new quality of interdisciplinary harmonized data. envVisio can also be integrated into existing infrastructures. In the background, data can be imported from specialist applications or made available for them. envVisio can act as a straightforward and robust interface for open data.

The core of the envVisio platform is data storage. It is structured in two stages; Simplex for data management according to the double tetrahedron structure and Scenario for creating user views. The necessary tools are organized around them; the Import for importing the data, Visualization component for viewing the imported data and Service for providing the data.

**envVisio Import:** The import module is used to load documents in a number of formats onto the server and import them into the database. (Geo) JSON, CSV, shapefiles and XML / GML are currently supported. The import module also controls API features services (such as the Open Data Portal Dresden) and automatically query their data. The data is transformed/converted into the double tetrahedron structure via an interface and loaded into the Simplex. This is where the first content-related discussion of the data takes place. Since the envVisio data model works with simple, recurring structures, the assignment (often also referred to as mapping) is clear and repeats itself. When converting, it is also important to define the connections between the data records, which is implemented via the identification of foreign keys.

**envVisio Simplex:** It is the core component of envVisio for data storage in the object-oriented and graph-based structures of the double tetrahedron. It ensures the consistency of the model and enables efficient and long-term management of the data and metadata.

**envVisio Scenario:** Views and images of the data managed in the Simplex can be defined via an interface. The views can implement technical specifications (e.g. for INSPIRE), represent the specific perspective of an actor on the data, or be of a purely exploratory nature. The uniform basic structure of the data model guarantees easy-to-learn access to the data based on just a few operations, which can, however, process the data flexibly. Several views can be bundled into a scenario. Scenarios represent use cases.

**envVisio Visualization:** The visualization component allows to get an overview of the existing data and their links and to gain suggestions for the definition of views or scenarios. It also offers graphic support for creating new scenarios. It works differently than conventional geospatial data viewers, because it is for example to no longer define and select layers. All affected object classes can be displayed directly. It is possible to switch between the object classes along their connections (e.g. from a body of water to the assigned measuring points, to the address, further to the edge of the street, from there to the traffic volumes measured in the street, to protected areas ...). The visualization is designed to be as accessible as possible and, in particular, does not require any specialist knowledge. In this way, a broader group of interested parties (e.g. citizens, visitors to the municipalities, seekers of objects) can be invited to use the data.

**envVisio Service:** All data in envVisio can be selected via the service and are available for further processing. The service implements API features (core), the first of the new service standards of the OGC, which is essentially based on REST. It opens up opportunities to open up the geospatial special area to a broader spectrum of applications and actors. This is particularly interesting for the further processing of environmental data, because it makes the data accessible to amateurs, volunteers and professional users from other disciplines. More details on the envVisio Service can be found in the following chapter.

5. **ENVVISIO SERVICE**

EnvVisio Service was designed to provide access to the envVisio modelled data to clients (Großmann 2020). A main goal was to translate the genuine universal structure of envVisio Model from database level to service level. Because of this, envVisio Service is by design connected to a data pool and regulated by its content: the resources of the service dynamically map the state and content of the data pool. The service is – essentially – serving the data, imposing as few overlays as possible.

EnvVisio Service extends the new OGC API-Features (Core) interface standard (Open Geospatial Consortium Inc. 2019). Compared to Web Feature Service (WFS 2.1) (Open Geospatial Consortium Inc. 2014), this new OGC standard...
marks a turn away from specialization on geodata towards an extendable approach, based on well-known technologies and practices of mainstream IT. It can and surely will be used as a “Web Feature Service 3”, but it has the potential to reach a broader audience than specialists of geoinformatics.2

Rather than an exclusive set of interactions, API-Features (Core) defines a modular structure of linked resources. EnvVisio Service extends this structure with additional resources and parameters, as table 1 shows. Some of the most important additions are discussed below.

### Table 1. URI-Templates of API-Features (black) and envVisio Service (green)

| Resource | Template |
|----------|----------|
| Scenarios | /scenarios |
| Scenario | /scenarios/{scenarioId} |
| Collections | /scenarios/{scenarioId}/collections |
| Collection | /scenarios/{scenarioId}/collections/{collectionId} |
| Items | /scenarios/{scenarioId}/collections/{collectionId}/items |
| Links | /scenarios/{scenarioId}/collections/{collectionId}/items/links |

In a RESTfull manner, the structure of resources in the interface directly reflects the structure of the underlying data. Each single object, be it a tree, an agency or a river, can be found and requested under its unique identifier. Requests on resources representing lists (e.g., /collections/{collectionId}/items) offer complete datasets that can be downloaded en bloc or filtered with a set of parameters.

An additional level of hierarchy is added as scenarios. envVisio Model allows a fast and nearly unrestricted creation of arbitrary views on the data. These views are made accessible as scenarios. Each scenario represents an own, independent view on the data pool, e.g., a defined use case for client software, or the data model of a whole INSPIRE annex theme. Updates and changes to the data pool are instantly reproduced in the scenarios. As a fixed point, scenario/ represents all data of the pool, represented in envVisio Model.

Special attention is drawn on the resources /links and /neighbours. They reflect the emphasis envViso Model puts on the links between objects. Requests on /neighbours of a single feature (e.g., a tree) return all features connected to the specified one: no matter if they are part of the same collection (e.g. other trees in the same street) or not (e.g. the public office responsible for this tree).

As links are themselves objects in the envVisio Model, they can be reached via /links. In the mentioned example, a request on the /links of the tree will be responded with a list of connections, e.g., the connection tree-to-tree and the connection tree-to-public office. Both resources are also available for whole collections, so tree-collection/links will show at one glance that there is a link between trees and public offices. Last to mention, /links can be used to filter the objects of a collection so tree-collection/links will help to to find all objects – independent of the collection they are in - that have a link to any of the tree objects that match the request.

### 6. USE CASES

The envVisio method was already used successfully in various domains for data management before the envVisio-GI project. In the following, some already implemented use cases are briefly presented. Furthermore, use cases are explained that are to be implemented in the envVisio-GI project.

#### 6.1 Completed Use Cases

**Road network and traffic:** It is very complex to combine topological and traffic-related aspects across all types of traffic in one model. As part of the Cartox² research project, a model was developed in order, on the one hand, to capture the dynamic data generated during automated driving in simple structures with high performance and, on the other hand, to manage data in a precise manner in order to also be able to describe the traffic flows. Figure 3 shows the distribution of road network information in the envVisio model.

![Figure 3: Road network in envVisio](image)

© Simplex4Data Image Copyright 2021

---

2 An example for this is the open data portal of the city of Dresden that uses an API-Features service serving GeoJSON to provide access to a range of hundreds of collections.(see https://opendata.dresden.de/DreiD/ and https://kommisidd.dresden.de/net4/public/ogcapicollections), last seen 11.05.2021.
Other use cases that have already been implemented with envVisio include:

- Ecosystem models
- Models for air quality monitoring, management of soil parameters and for visualizations on the subject of climate and energy
- Information systems for the management of water bodies, water body sections based on network models, as well as associated systems and structures
- Environmental law and monitoring
- Environmental planning
- Reporting in environmental information systems

6.2 Ongoing use cases

Noise mapping: As a metropolitan area with more than 250,000 inhabitants, the city of Leipzig creates noise maps, which are to be updated every 5 years. The basis is the EU Directive 2002/49 / EC of 2002 for the assessment and control of environmental noise. At this time, the data must be updated again and transmitted to the State Office by June 30, 2022 (European Commission 2002).

A large number of geodata is necessary for the implementation of the noise mapping (these include the road network, building models and environmental data). These data are processed in a very labor-intensive process and serve as the basis for the subsequent calculation of the noise situation and the determination of those affected.

The aim in the envVisio-GI project is to develop a partially automated process for:

- More efficient data preparation, continuation and updating of data
- Description of requirements for data structure
- Identification of added value for further applications

Preparation of construction planning data: This use case deals with the development of a guideline for the more efficient preparation of plans for land use (fully vectorised), considering the upcoming data exchange format "XPlanGML" and interoperable provision for INSPIRE, for disclosure in planning processes and further processing in the building application and building permit process. In addition, there is the preparation of further municipal geodata (school locations, school catchment areas, natural monuments, protected areas) for the provision according to the INSPIRE specifications and, if necessary, for processing in digital specialist processes. Furthermore, it should be tested whether existing zoning and development plans can be transferred to the BIM standard IFC 3 in order to be able to support new construction projects in the region on the part of the municipality (Fritsch et al. 2020).

Post-fossil mobility: The express bus axes project for the Lippe district in Germany emerged from a funding project. In the continuation, these are to be converted to electric drive. Additional partners are required to implement this step. In addition to the transport associations and their organizations, there are also parties involved in the approval process, energy providers, municipal facilities, etc. For this application in the envVisio-GI project, they are potential users of information to be provided. These can be the existing and to be created e-charging points, broadband infrastructure, traffic control, movement data, availability, status displays of the infrastructure and the like. From the point of view of the existing processes, envVisio-GI should help to bundle existing resources and simplify the possibilities of internal and external use.

7. CONCLUSION AND FUTURE RESEARCH

The presented envVisio method has already been used successfully for the preparation, storage and linking of data in different professional domains such as environmental data and road networks. The aim of the envVisio-GI project is to expand the use of the envVisio method to other applications of the associated partners from municipalities, cities and associations. Part of this is the establishment of a new type of data management for the provision of municipal data: standardized, interoperable, universally processable and inexpensive. The processed specialist information should then be available in other standardized data formats in addition to the envVisio service presented and support the municipalities in fulfilling the obligation to provide INSPIRE data and other reporting obligations. Added value for the user is potentially expected cost savings of up to 90% compared to the currently used methods and tools for the delivery of INSPIRE-compliant data, as well as better visibility of municipal data and better valuation of data by third parties.

In addition to the use cases presented, it is also planned to implement the IFC standard of “Building Information Modeling” (BIM). The processing with BIM as well as planning-relevant processes and methods is worked out. Further interfaces for external application developers are planned. These are based on contemporary industry standards for data exchange. An important application for such interfaces is the interoperability between GIS and BIM (Kaden et al. 2020). Particularly efficient methods of georeferencing BIM planning models could be an important contribution to improving the data exchange between these professional domains (Kaden and Clemen 2017).

Work on the research project is planned until summer 2022. Until then, all use cases of the associated partners of the project will be processed and the envVisio service will be available for use.

REFERENCES

BMVI, 2018: Stufenplan Digitales Planen und Bauen – Einführung moderner, IT-gestützter Prozesse und Technologien bei Planung, Bau und Betrieb von Bauwerken. Bundesministerium für Verkehr und digitale Infrastruktur, Berlin, Germany. https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/stufenplan-digitales-bauen.pdf?__blob=publicationFile (2019-05-08)

EnvVisio-GI Project Team, 2020: GeodatenIntelligenz für interoperables Datenmanagement mit der envVisio-Method

3 Industry Foundation Classes = Standard date exchange format for BIM
