Improving Data Quality of Programme of Measures for the Water Framework Directive in Saxony

Friedhelm Hosenfeld [iD] [1], Roland Dimmer[2] and Christoph Mattes[3]

Abstract: A web application is presented to support the responsible authorities in the management of measures for the WFD (Water Framework Directive) in Saxony. The web application enables the maintenance of WFD measures data by different authorities in a common database. The central data management supports the tasks of implementing the WFD of the LfULG for the fulfilment of the EU reporting obligations. A key requirement deals with the improvement of data quality implemented by comprehensive consistency and completeness checks, input rules and support functions for geometry creation. The spatial data are verified for consistency with the attribute data during data acquisition.

Keywords: WFD (EC Water Framework Directive), WFD measure management, Disy Cadenza, data quality, quality assurance, web application, improving spatial data

Addresses Sustainable Development Goal 6: Clean water and sanitation

1.   Introduction and overview

1.1   Overview

The Water Framework Directive (WFD) of the European Community [Di00], which came into force in 2000, pursues an integrated water protection policy in Europe, which also brings about coordinated management of the waters within the river basins across state and national borders. In order to achieve the central objective, a good status of as many water bodies as possible, plans and programmes of measures were created and will be updated in further management periods.

This paper describes a web application operated by the State Agency for the Environment, Agriculture and Geology (LfULG) for the management of measures planned and implemented under the WFD in the Free State of Saxony and to support reporting to the EU. One main goal focusses on facilitating central data management improving data quality and quality assurance methods. The application was developed by the companies Disy and DigSyLand and integrated into the Saxon data portal iDA, which provides

[1] DigSyLand - Institute for Digital System Analysis & Landscape Diagnosis, Zum Dorfteich 6, D-24975 Husby, hosenfeld@digsyland.de, [iD] https://orcid.org/0000-0002-0538-7684

[2] Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie, Referat 44, Oberflächenwasser, Wasserrahmenrichtlinie, Zur Wetterwarte 11, D-01109 Dresden-Klotzsche, Roland.Dimmer@smekul.sachsen.de

[3] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, D-76131 Karlsruhe, Deutschland, christoph.mattes@disy.net

interdisciplinary access to environmental data and maps on the web.

In the following sections, the tasks of measure management in Saxony, the initial situation and the requirements for the application are discussed. Section 2 deals with the design of the application and explains the components of the software architecture. Section 3 is dedicated to the implementation of data quality improvement methods in the web application, where the special challenges due to the technical requirements are mentioned. The article concludes with a summary and an outlook in section 4.

1.2   Initial situation and general conditions

The LfULG maintains a state-wide database on the implementation of the WFD and already makes much of these data available in a variety of ways (like map services, fact sheets and downloads), because the LfULG is the central authority in Saxony for informing the authorities involved and the public about the state-wide data and activities for the implementation of the Water Framework Directive.

Of particular importance are the measures to improve the status of water bodies, for which the data are maintained locally by other water authorities and about 18 additional institutions according to their responsibilities. In particular, information on the preparation and updating of management plans and on programmes of measures is essential for reporting to the EU, as well as for interim reports on the progress of measure implementation.

The data on WFD measures maintained by other water authorities and further stakeholders were kept by them in different standardized Microsoft Excel tables, which were merged manually at certain dates. The LfULG manually compiled the data in a database for reporting purposes at state level. Working with the Excel tables proved to be error-prone, inflexible and time-consuming.

Especially the non-optimal data storage and management with all consequences resulted in the need to develop a web application that both reduces the workload and supports the optimisation of data quality.

1.3   Requirements for a data management web application

The aim of the data management application for WFD measures was to create a possibility for web-based creation and editing of WFD measure data by the users in the Saxon authorities, whereby all relevant data should be available in a central database.

The LfULG developed a technical concept that formed the basis for the conception and development of the web application, which mentioned the following requirements, among others:

- Web-based processing of the WFD measure data, so that in particular the creation, editing, deletion, display, research and export of these data are possible.

- In order to continue to support also local data management, import of measure records into the central database should be offered in addition to interactive editing.

- With a suitable user management, the technical and spatial responsibilities should be appropriately mapped within the application.

- The ability to integrate the solutions into the existing IT infrastructure had to be guaranteed. This included in particular the integration into the iDA data portal, which is based on the Disy Cadenza[4] evaluation and GIS platform.

- The technical requirements included extensive consistency conditions, which are to be ensured by the data management, as well as conditions for the analysis and optimisation of WFD measure geometries in relation to the water bodies important for the WFD and their catchment areas.

2. Design of the application

2.1 Software architecture

The software architecture of the web application complies to the general conditions of the LfULG and uses the components that are already in use in the data and analysis platform iDA (see section 2.2):

- Data storage in the central Oracle database (section 2.3)

- Analysis, export and GIS functionality in Cadenza Web (section 2.4)

- PHP for implementation of data processing and management functions (section 2.5).

2.2 Saxony's data portal iDA - interdisciplinary data and analysis

Saxony's web portal iDA (interdisciplinary data and analysis) offers a comprehensive access to environmental data and maps, which originate from measuring and research programmes of the LfULG and further information systems of the Free State of Saxony[5].

The uniform portal platform is implemented with Cadenza Web and combines the structured access to data of the different departments of the state. In addition to public access, further access options are offered for registered users, who thus can also access non-public thematic data according to their area of responsibility.

2.3 Database design and data management

The thematic data available in iDA are managed in the Oracle RDBMS, supplemented by

---

[4] Disy Cadenza: Software for data analysis and business intelligence (https://www.disy.net/en/products/disy-cadenza/overview/, accessed: 10/07/2022)

[5] iDA: https://www.umwelt.sachsen.de/umwelt/infosysteme/ida/, accessed: 10/07/2022

spatial data that are integrated via services such as ArcGIS Server REST services.

Concerning the WFD measures management, data modelling in the Oracle database included the representation of geometries. This means that the geometries can be used directly by common GIS software with Oracle support without additional interfaces, especially by Cadenza but also by PHP applications.

The data model requires that the WFD measures are edited exclusively using the data management application in order to guarantee consistency and access rights. Due to complexity and specific requirements, as well as better configurability, most consistency checks have to be handled by the data management application, except for referential integrity constraints on the database level (see section 3.2). By default, each WFD measure data set has exactly one geometry (polygon, line or point), except for conceptual measures which have no concrete geometry.

Special challenges for the data modelling included the mapping of multiple assignments of properties to WFD measures, which cause both content-related and technical consequences. For example, several water bodies and also several entries of the nationally defined LAWA[6] catalogue of measures [LA20] can be assigned to a WFD measure. This catalogue controls the validity of geometry types, but also of water body categories. In addition, it must be possible for analyses to unambiguously evaluate these multiple links, so that, for example, WFD measures are not considered multiple times in calculations and map representations.

Other important technical and content-related aspects that had to be appropriately considered in the data modelling include:

- Ownership (right to change) of a WFD measure, which can also be passed on to other institutions.

- Flags for technical correctness and completeness (factual data and geometries)

- Different statuses with different consistency requirements, also ensuring that legacy data are also available for evaluation in the same system.

2.4   Cadenza Web as Evaluation and GIS Component

The evaluation and GIS platform Disy Cadenza forms the main component of the iDA web portal, which provides access to environmental data and map inventories in Saxony for the public and other user groups.

---

[6] LAWA: Bund/Länder-Arbeitsgemeinschaft Wasser: German Working Group on water issues of the Federal States and the Federal Government represented by the Federal Environment Ministry (https://www.lawa.de/English-About-LAWA.html, accessed: 21/05/2022)
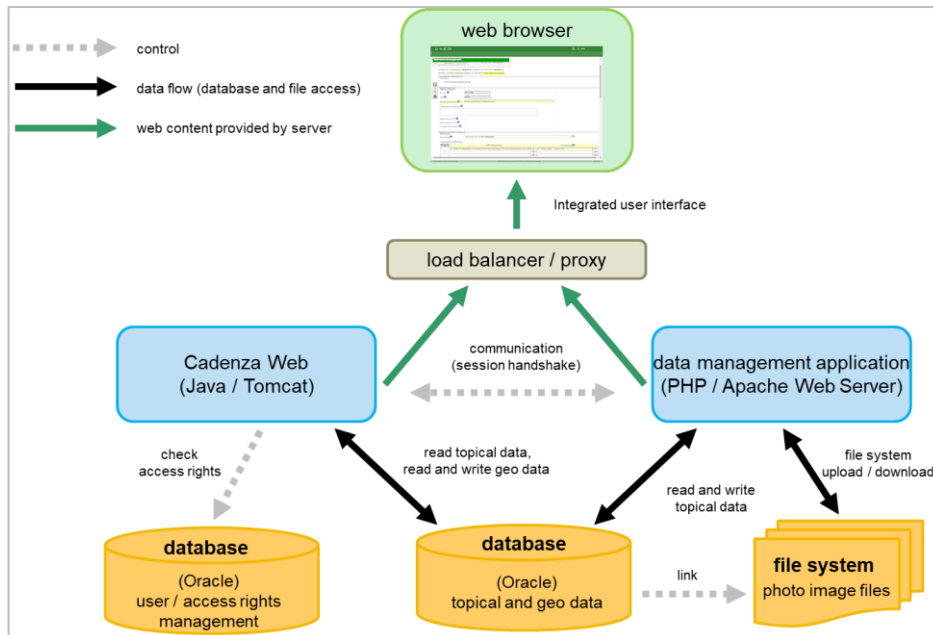
Fig. 1: Overview of the interaction of data sources, Cadenza Web and PHP data management application

The Cadenza Web variant includes numerous evaluation functionalities of topical and spatial data as well as the acquisition and maintenance of geometries.

Cadenza is continuously developed by the company Disy in close cooperation with several federal and state authorities, especially from the environmental sector, and is therefore tailored to the current requirements of these cooperation partners (for example applications, also for WFD applications, see e.g. [HTT16, Ho12b, Ho21a, Ho21b]).The so-called Cadenza repository forms an intermediate layer for the individual application areas as an integrating view of the underlying data sources, which can be database-based as well as services and geodata files (for details see e.g. [Ho21a] p16f.).

With filter forms Cadenza offers the combined query of topical data and spatial data according to all required criteria with presentation of the results in exportable tables, but also as reports and interactive maps.

While the focus of the Cadenza component is on data analysis, reporting and GIS, for the realisation of thematic data management systems in which data are edited and processed, it offers the integration of data management applications via a programming application interface with which, for example, data acquisition functions can be added [Ho12a], which can be tailored to the respective application purpose. In this way, the generic, configurable standard functionalities can be supplemented with customised acquisition and data management functionalities.

The companies Disy and DigSyLand have not only developed data management

information systems following this architecture in Saxony on behalf of the LfULG, but also for the environmental state authorities in Schleswig-Holstein [Ho12b, Ho21a, Ho21b]. While earlier developments of WFD supporting web applications in Schleswig-Holstein were successfully realized only based on PHP [Ho08, HLT11], the additional standard functions offered by Cadenza, especially concerning spatial data turned out to be essential.

Therefore, it was obvious to use this combination of components of Cadenza Web and PHP also for the realisation of the WFD measures management.

2.5   PHP management application for data editing and maintenance

For the management of the WFD measures in Saxony a data management application was designed to fulfil the requirements of the technical concept regarding the acquisition and maintenance of the measures.

This component was realised using the scripting language PHP[7] and seamlessly integrated into the Cadenza Web environment via its application programming interface (see Fig. 1).

Both Cadenza Web and the PHP application access the Oracle database directly, while Cadenza Web only performs read accesses except for editing geometries of the WFD measures, while the management application is responsible for editing the topical data (see Fig. 1).

3.   Implementation of data quality improvement methods in the web application

3.1   PHP-Framework as the basis for the data management application

The data management application is based on a PHP-framework developed by DigSyLand, which was optimised for the integration into the Cadenza Web application programming framework, so that a mapping of the application-related access functions to the user groups of the Cadenza user administration is possible (see also [Ho12b, Ho21a, Ho21b]).

3.2   Geometry adjustment and consistency checks

While the WFD measure input form could be designed largely by configuration on the basis of the PHP framework, some special functions, which particularly concern the complex consistency checks and also the checking, adjustment and intersection of the WFD measure geometries, had to be programmed separately.

The important technical requirement that linear measures should be adapted to the actual

---

[7] PHP stands for "PHP: Hypertext Preprocessor" (https://www.php.net/)

course of the waterbody route was implemented in PHP based on Oracle geometry processing functions: After a line is drawn on the river water using the Cadenza geometry acquisition function and then transferred to the data management application, the latter forms a new line geometry from the line segments of the watercourse from the starting point to the drawn end point of the WFD measure, which lies exactly on the watercourse line.

For all measure geometries (points, lines, or polygons) it is checked in which water bodies or catchment areas and in which municipalities they are located. These assignments are automatically added as attributes in the attribute data for evaluations and consistency checks. In addition, it is checked whether the measure lies within the user's area of responsibility. Thus, the lower water authorities can only create geometries in their district area, while other authorities can edit WFD measures Saxony-wide.

Invalid measure geometries are marked accordingly, but are initially kept in the system to facilitate subsequent corrections without having to completely create the geometry again.

Konsistenz der Attributangaben: **Attributfehler**, Konsistenz der Geometrien: **Geometrie ok**

Der Datensatz wurde gespeichert. Bitte beachten Sie folgende Hinweise:

- Das Feld 'Feinbelastung' ist ein Pflichtfeld. Bitte tragen Sie dort einen Wert ein!
- Das Feld 'Maßnahmenträger' ist ein Pflichtfeld. Bitte tragen Sie dort einen Wert ein!
- Das Feld 'Maßnahmenkosten in Euro, ohne Nachkommastellen' ist ein Pflichtfeld. Bitte tragen Sie dort einen Wert ein!
- Das Feld 'Kostenträger' ist ein Pflichtfeld. Bitte tragen Sie dort einen Wert ein!
- Für den Status *nicht umsetzbar* muss eine **Begründung der Nicht-Umsetzbarkeit** angegeben werden!
- Für den LAWA-Katalog-Eintrag 2 *Ausbau kommunaler Kläranlagen zur Reduzierung der Stickstoffeinträge* muss noch eine Feinbelastung angegebenen werden!
- Für den Wasserkörper DESN_VM-1-2-1 gibt es keine LAWA-Katalog-Zuordnung mit passender Gewässerkategorie!
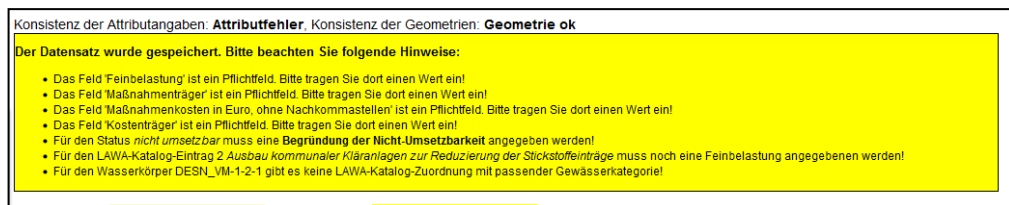
Fig. 2: Results of comprehensive consistency checks are displayed after saving a record

This also applies to WFD measures whose attribute data were entered inconsistently or incompletely: These are marked as incorrect with regard to the attributes, so that a subsequent completion and correction can be made (see Fig. 2).

In addition to the obligatory filling of mandatory fields, the consistency checks include the checks of content-related correlations, such as the LAWA catalogue entry [LA20] matching the water body category and the geometry type with the specification of the associated pressures and other specifications (see Fig.2).

As far as possible, the data are already checked during input or controlled by limiting selection options. Due to the dependencies of different attributes, the data can only be analysed completely after saving, so corresponding hints based on the consistency checks are displayed (see Fig. 2).

For measures marked as legacy data, which were initially transferred to the Oracle database, slightly relaxed consistency conditions apply to enable inclusion in the application.

3.3 Import interface

The institutions responsible for WFD measure management are subject to different local general conditions. While the central data management in the newly developed WFD

measure management application is favoured by the LfULG, local measure management solutions are also still in use. In order to also be able to transfer the measures maintained locally into the central database, a standardized import format based on Excel files was specified, which is based on the original Excel exchange format.

The data management application offers an import interface for the transfer of externally maintained data. Both new measures can be introduced and existing measures can be updated.

During the import, the same consistency conditions are checked as during the interactive input. If an automatic correction is possible, it is carried out while otherwise incorrect data records are rejected during the import. After the import, a detailed log is generated with notes on errors and corrections. Point and line geometries are defined by specifying the coordinates of the (end) points. Recently an additional import option was implemented for polygon geometries based on shape files.

3.4   Evaluations and map representations

Using the Cadenza standard functions the search and evaluation of the managed measures were developed and integrated into a new Cadenza repository, so that all attributes of WFD measures are available as criteria for the definition of the search. The results can be displayed and exported as a list, but can also be displayed on the map. Several map display options of the measures were implemented, symbolizing the measures according to different properties like status, responsibilities, waterbody category and more.

4. Summary and outlook

With the web application the central data management of the WFD measures was implemented which offers access for all stakeholders for interactive processing and evaluation of the common data pool. Additionally external third-party data management solutions were connected via an import interface. Using the data management application all organisations responsible for WFD measures in Saxony have a common up-to-date view on all relevant data. Especially in areas where different organisations are planning WFD measures a better exchange is facilitated.

While formerly the spatial representations of the measures in the manually compiled central data pool were very limited, the new web application offers suitable map representations of the WFD measures according to different criteria which can be combined with several related map layers which were centrally managed by the LfULG.

The realisation of comprehensive consistency checks and completeness rules help to ensure the data quality, thus achieving a key requirement which is especially important for EU reporting. The improved data quality also enables better evaluations on the state level and reduces the amount of effort to work efficiently with the data.

In autumn 2021 a pilot phase started when interested participants of different organisations tested the application and exchanged their experience with the LfULG.

As a result of the pilot phase, it was found that the application improves the management of measure data in the desired way, but that there was still a need for supplementing certain functionalities.

While these optimizations are currently being added which include among others the extension and performance improvements of the import interface, archiving of WFD measure records, more evaluation options, additional data fields and improvements of usability, users can test the application on a test system and deliver their feedback.

In autumn and winter 2022 the final data migration into the new system is planned so that productive operation of the WFD measure management web application can start at the end of 2022. Support for filling EU reporting templates is planned for future development phases.

Bibliography

[Di00]   Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy (EU Water Framework Directive). In: Official Journal of the European Communities, 22/12/2000 pp. 0001-0073, 2000, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0060&qid=1653152811078&from=EN, accessed: 21/05/2022.

[Ho08]   Hosenfeld, F. et al.: Database of measures supporting creation of management plan and programme of measures for implementing Water Framework Directive. In: Möller, A.;

Page, B.; Schreiber, M. (eds): Environmental Informatics and Industrial Ecology. 22th International Conference on Informatics for Environmental Protection, pp. 541 – 548, 2008.

[Ho12a]   Hofmann, C. et al: Cadenza - Raumbezogene Berichtssysteme und Fachanwendungen auf der Basis von Cadenza. In: Weissenbach, K.; Ebel, R.; Weidemann, R. (eds.): UIS BW Umweltinformationssystem Baden-Württemberg, F+E-Vorhaben MAF-UIS, Phase I 2011/2012, Report-Nr.. KIT-SR 76616, Karlsruhe, pp. 81-92, https://pudi.lubw.de/detailseite/-/publication/38498-Cadenza_-_Raumbezogene_Berichtssysteme_und_Fachanwendungen_auf_der_Basis_von_Cadenza.pdf, accessed: 21/05/2022.

[Ho12b]   Hosenfeld, F. et al.: Development of a Database Supporting the Management of Nature Conservation Measures in Schleswig-Holstein. In: Arndt, H.-K.; Knetsch, G.; Pillmann, W. (eds.), Light up the Ideas of Environmental Informatics, Dessau, Shaker Verlag Aachen, pp. 623–629, 2021, http://enviroinfo.eu/sites/default/files/pdfs/vol7793/0623.pdf, accessed: 21/05/2022.

[Ho21a]   Hosenfeld, F. et al.: Neuentwicklung einer Intranet-Web-Anwendung für das Geotopkataster Schleswig-Holstein. In: Freitag, F. et al. (eds.): Umweltinformations-systeme - Wie verändert die Digitalisierung unsere Gesellschaft? Tagungsband des 27. Workshops „Umweltinformationssysteme (UIS2020)" des Arbeitskreises „Umwelt-informationssysteme" der Fachgruppe „Informatik im Umweltschutz" der Gesellschaft für Informatik (GI). Springer Vieweg, Wiesbaden, pp. 13-27, 2021, https://doi.org/10.1007/978-3-658-30889-6_2, accessed: 21/05/2022.

[Ho21b]   Hosenfeld, F. et al.: Governmental Information Systems Supporting Nature Conservation in Schleswig-Holstein. In: Wohlgemuth, V. et al. (eds.): Environmental Informatics. A bogeyman or saviour to achieve the UN Sustainable Development Goals? Adjunct Proceedings of the 35th edition of the EnviroInfo – the long standing and established international and interdisciplinary conference series on leading environmental information and communication technologies Berlin, Germany, September 27-29, pp. 133-140, 2021.

[HLT11]   Hosenfeld, F.; Lietz, J.; Trepel, M.: Assessment Database Supporting Implementation of Water Framework Directive in Schleswig-Holstein. In: Pillmann, W.; Schade, S.; Smits, P. (eds.): Innovations in Sharing - Environmental Observations and Information, Ispra, pp. 460 – 466, 2011.

[HTT16]   Hosenfeld, F.; Tiffert, J.; Trepel, M.: Waterbody and Nutrient Information System Schleswig-Holstein. In: Wohlgemuth, V.; Fuchs-Kittowski, F.; Wittmann J. (eds.): Environmental Informatics – Stability, Continuity, Innovation. Adjunct Proceedings of the 30th edition of the Enviroinfo. Berlin, pp. 407-412, 2016.

[LA20]   LAWA (Bund/Länder-Arbeitsgemeinschaft Wasser): LAWA-BLANO Maßnahmenkatalog (WRRL, HWRMRL, MSRL), 60pp., 2020, https://www.lawa.de/documents/lawa-blano-massnahmenkatalog_1594133389.pdf, accessed: 21/05/2022.