Data Concretization States as Metadata of Parameterized Regional Futures in a WebSDSS Development Context

Rico Vogel and Marco Neubert and Axel Sauer

Leibniz Institute of Ecological Urban and Regional Development (IOER)
Dresden, Germany
Email: {r.vogel, m.neubert, a.sauer}@ioer.de

***Abstract*—The development of a web-based spatial decision support system (WebSDSS) aims at the transfer of knowledge that arises from the operationalization of a scenario approach that focuses on so-called 'parameterized regional futures' (PRFs). Because the operationalization of the PRF approach is an ongoing process, a gap exists between data required and currently available. To overcome this gap and to allow specific software testing, test data sets can substitute expected PRF data. Based on the nature of spatial and non-spatial data and its role in different WebSDSS development phases, this article proposes an approach to distinguish concretization states as metadata of original PRF data and test data counterparts. This contribution accordingly addresses topics like modeling or managing spatial data, geospatial domain applications, and digital cartography.**

Keywords–Software development; WebSDSS development; Metadata information; Data concretization states; Test data; Parameterized regional futures; Scenario approach.

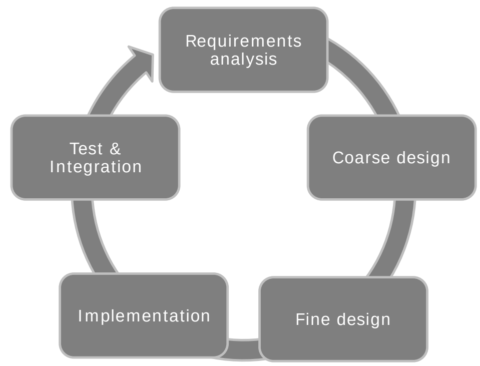

Figure 1. Iterative incremental development and method steps (modified according to [2]).

I. Introduction

Currently, a web-based spatial decision support system (WebSDSS), characterized by Rinner [1] for example, is being developed by Vogel [2] and the related research group. This tool aims at enabling the immediate transfer of knowledge, which arises from the operationalization of a scenario approach of Schanze and Sauer [3] that focuses on so called 'parameterized regional futures' (PRFs, see also [4]). This software accordingly provides steered access to the complex pool of spatial and non-spatial PRF data by using web-based geographical information system (WebGIS) technology, which facilitates thematical, temporal, and spatial selection, preparation, and presentation of such data via the Internet.

Within a PRF, a *future* is the core component composed of one specific *scenario* and one specific *strategic alternative*. A *scenario* consists of a set of *projections* derived from a narrative storyline addressing climate change, demographic change, technological change, economic change, as well as land-use change (e.g., increasing temperatures, aging society). A *strategic alternative* comprises different *intervention options* that are single and partly site-specific measures [4]. Building a dike or improving flood resilience of buildings are examples for such options. Both, *projections* and *intervention options* are defined by further sub components. A formalization of the PRF approach capturing the component hierarchy has been prepared by [5].

In the case it serves as the leaf of the PRF hierarchy, each sub-component is commonly expected to be represented by a spatial data set, which may be derived according to [6]. Dependent on the number of *futures* to be operationalized, a number of slightly varying data sets is expected for each theme (leaf sub-component). In the case of the PRFs, about a dozen of *futures* are operationalized [3].

This article proposes an approach to distinguish concretization states of original PRF data and types of test data. First, the role of spatial and non-spatial data in the WebSDSS development process is explained (Section II). Rainardi [7] (pp. 477–489) emphasize the importance of test data in data warehouse development, and [8] exemplifies its value in conformance testing in the context of standardization. Afterwards, the nature of PRF data and associated test data is captured in Section III. Relevant data concretization states are identified in Section IV. Test data of such states can support different phases of WebSDSS development (e.g., [9], p. 299). Therefore, in user tests, data concretization states are important metadata information provided by web services like Web Mapping Services (WMSs). Finally, Section V concludes this article.

II. WebSDSS Development and Data Availability

The WebSDSS development intended by Vogel [2] follows an iterative-incremental approach. According to Kleuker [11] (pp. 30 f.) in particular the *requirements analysis*, *coarse design*, *fine design*, *implementation*, as well as *test & integration* phases may be distinguished. As shown in Figure 1, these phases are repeated; and, according to Rumbaugh, Jacobson, and Booch [12] (p. 319), each iteration results with an executable system that can be executed, tested, and debugged. Nevertheless, each of the phases may have certain requirements with regard to the availability of PRF data. While some like *coarse design* may be proceeded based on simple theoretical

 ISBN: 978-1-61208-383-4

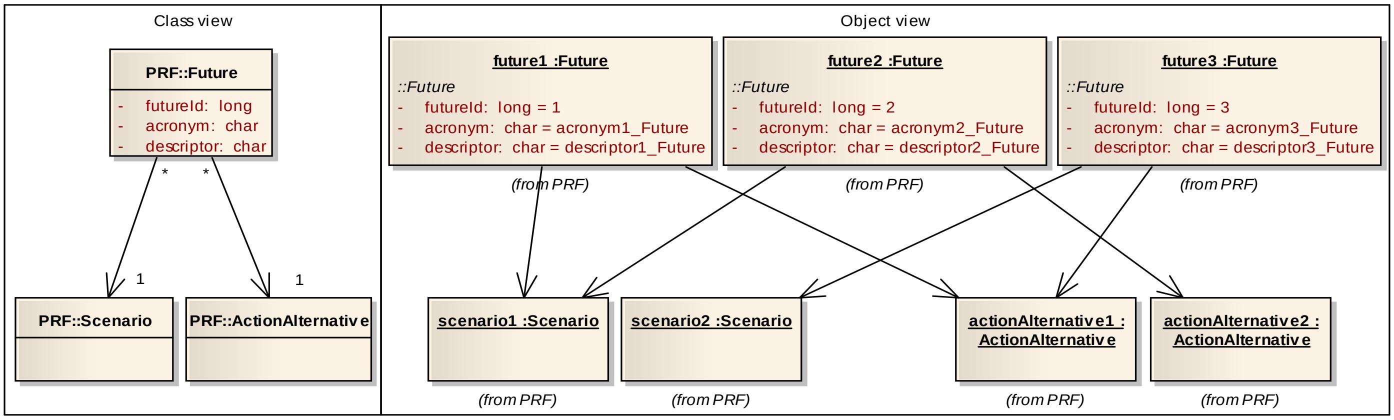

Figure 2. Concept of the *Future* component (Class view) and instances of the *abstract* data conretization state (Object view) used by *software developers* in early iterations of (conventional) software development (inspired by [10]).

examples, others like *test & integration* may require original data sets or at least such that approximate them.

Since PRF operationalization is an ongoing process that captures recent (e.g., [13] contributes to the 'Integrated Regional Climate Adaptation Programme for the Model Region Dresden', REGKLAM, research project) and current research (e.g., [6] contributes to the 'Vulnerability Study Saxony' (VusS) and to the 'Cross-Process Analysis of Vulnerabilities and Risks of Urban Regions with Respect to Climatic Factors and their Changes – Conceptualisation and Modelling' (RegioRisk) research projects), the complete set of multi-dimensional data (e.g., [14]) is not yet available for all PRFs to be operationalized. Rather, certain PRF aspects have already been analysed (e.g., [13]) or are work in progress (e.g., [6]).

Besides the intended database of PRF data, for the WebSDSS development there are several reasons to build up a test database (see Figure 2). For example, in early iterations simple use case examples are sufficient, while during *test & integration*, test data are required that at least embraces a sub area by a selection of PRF themes. Further data sets improve later iterations to *design* and *test* the system behavior for extreme data values, which are not expected from causal PRF operationalization (e.g., increasing average temperatures by 20 K in 5 years).

To fill the gap between the availability of real PRF data and test data (see Figure 3), an approach needs to be developed that takes into account the data requirements for each development phase as well as the opportunities to provide sufficient phase-related test data. Therefore, in a first step, important criteria have to be derived that enable the differentiation of (test) data requirements. Afterwards, valid data concretization states are suggested that seem most important in the context of PRF-specific WebSDSS development.

III. Capturing the Nature of PRF Data and Associated Test Data

From a development viewpoint, this section exemplifies the excerpt of important criteria, which enable the capturing of the nature of PRF data. These findings will be used to discuss proposed categories of *futures* in the next section. To capture the criteria, social, spatial, development, and scientific scopes

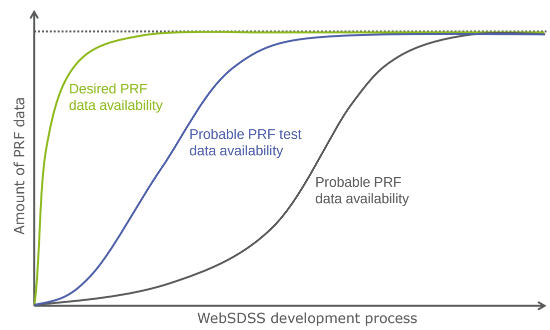

Figure 3. Schema of gaps between desired and probable availability of PRF data in WebSDSS development and supporting role of test data.

were differentiated. These metadata may be the source for characterizing different data concretization states.

From a social perspective, the main addressees of certain PRF data are distinguished. *Practitioners* and *planners* may be typical addressees of real PRF data. In contrast, test data are applicable in discussions during *requirements analysis* including further stakeholders such as the *scientists*, which developed the PRF approach (i.e., [3]).

Taking the spatial dimension into account leads to a distinction of *spatial* and *non-spatial* aspects of PRF data. Furthermore, the scale of a pilot region compared with the whole study region is differentiated. In particular, selecting *micro*, *meso*, or *macro* levels influences the amount of data that has to be processed during the *test & integration* phase for example.

During the technical WebSDSS development, it may be of importance to which extent the data sets cover the study area or if the data sets have limited spatial coverage (*complete*, *incomplete*). Additionally, data protection is an important issue. *High*, *moderate*, *low*, as well as *zero* protection levels can be distinguished.

Finally, scientific aspects can be taken into consideration. In particular, the scientific origin is of importance. For PRFs, the projects REGKLAM, VusS, and RegioRisk are itemized leading to *regklam*, *vuss*, and *regiorisk* identifiers. The data source can be differentiated in terms of scientific products

 ISBN: 978-1-61208-383-4

such as internal documents (*report*), publications (*article*, *proceeding*) or data sets (*dataset*, *map*).

IV. Data Concretization States

Based on the criteria identified in Section III, certain concretization states for data sets of the PRF approach are determined, which have importance for the WebSDSS development. Against the backdrop of the characterization of PRF data and test data in the previous section, four different data concretization states can be distinguished: *abstract*, *arbitrary*, *mockup*, as well as *genuine* (see Figure 4). These can shortly be described as follows:

- *Abstract data sets* are used to outline concepts of components and relationships among components in a simple manner. Often, simple characters are used to mark objects (e.g., by *a*, *b*, *A*, *B*) or such characters are combined with the type name of an object, for example *aTown* and *bTown* are used to distinguish two towns. In certain development phases, abstract data may be sufficient; in particular in *fine design* realized by a *software developer* simple examples are needed. In contrast, *test & integration* involving local and regional *planners* and *scientists* will be impractical if use case examples are merely based on *abstract data*.
- *Arbitrary data sets* are likewise valuable especially if spatial data is required in early development phases. Comparable with 'lorem ipsum' text used in typography, *arbitrary* data sets may be used as placeholders, for example to test map components by consuming map services. Data of this concretization state will primarily be helpful for the *software developer*. The current prototypical implementation [16], for example, simply associates accessed ArcGIS representational state transfer (REST) services, also used by Strode [17], to certain PRF component instances, which itself are merely of *abstract* state (e.g., sub-components of a *scenarioC* and a *strategicAlternativeB*).
- Unlike data of the previous concretization states, *mockup* data sets approximate (e.g., Benenson and Torrens [18], p. 16) or even consider the conditions of the study region. Although, they are still fictional, *mockup* (or dummy/ synthetic) data sets are designed to approximate *genuine* ones (comparable with 'potemkin villages'), since they are often the result of the same or similar models like used for the *genuine* PRF data. Accordingly, *mockup* data are assumed to be close to reality which therefore enable nearly-realistic software *test & integration*, involving stakeholders such as local and regional *planners*. Furthermore, *mockup* data sets may be created (e.g., simulated according to [18]) to offer (extreme) data values that are not expected to appear within *genuine* data sets.
- Finally, *genuine* data sets are the intended results gathered from the realization of the PRF approach. From the development point of view, data of this concretization state is valuable in each of the development phases. But in the subsequent production phase, merely *genuine* data sets are valid to build up the WebSDSS data component.

V. Conclusion

The article demonstrated an approach for capturing data concretization states of PRF data. Within WebSDSS development, these states are influencing each phase. For example, during *implementation* and associated unit *testing* by a *developer*, the state of PRF data, consumed by a web service may be registered by a logging mechanism. Of equal importance is the fact that such metadata information should be visualized by a graphical user interface (GUI) during *test & integration* to inform involved *planners* and other stakeholders about the concretization states of the data displayed.

Therefore, it is highly recommended to mark the data sets by their concretization state. Especially with a huge amount of data, this marking allows for the fast investigation of the required degree of replacement of test data by *genuine* one. Table I summarizes concretization states of PRF data, which at least have to be available in early software development iterations (x) or even in the production phase. Data sets may be swapped in the course of time, for example, by replacement with more concrete ones (o). To ease the access on such state information, metadata extensions or profiles according to [19] may be appropriate. To avoid misinterpretation of data content, induced by lacked information of the states of data consumed, especially in the case of mockup data, it is strongly recommended to automatically add the data concretization state information during data set production, or during the creation of related web services to the data and web service metadata.

TABLE I. Data concretization states in WebSDSS development and production (x – minimum requirement / o – aimed substitutions in subsequent iterations).

Development phase	*abstract*	*arbitrary*	*mockup*	*genuine*
Requirements analysis	x	o	o	o
Coarse design	x	o	o	o
Fine design	x	o	o	o
Implementation		x	o	o
Test & integration			x	o
Production				x

Nevertheless, considering aggregates on upper concretization levels requires more specific distinctions among aggregates, because the concretization states of an aggregate depends on the ones of its sub components. In upcoming work, the presented approach will be refined and adapted for such aggregates. In this context, the taxonomy defined by the domain model of [5] may be used to identify the aggregates and its members. As already mentioned, each *future* references exactly one *scenario* and one *strategic alternative*. Accordingly, its concretization state results from those of both sub components. For example, while a *scenarioC* might be marked with the *arbitrary* state, the state of a *strategicAlternativeB* is of type *mockup*. For such cases, intermediate operationalization states are currently under development. Vogel [20] proposes appropriate alternatives such as *abstract prevailed*, *arbitrary prevailed*, as well as *mockup prevailed*.

The concept presented in this article, is also applicable to data of further scenario approaches. Beyond the scope of WebSDSS development, it can also improve conventional (non-spatial) software development. Even in offline development like spatial analysis, data concretization states can be attached to metadata of geoprocessing workflow results, for

 ISBN: 978-1-61208-383-4

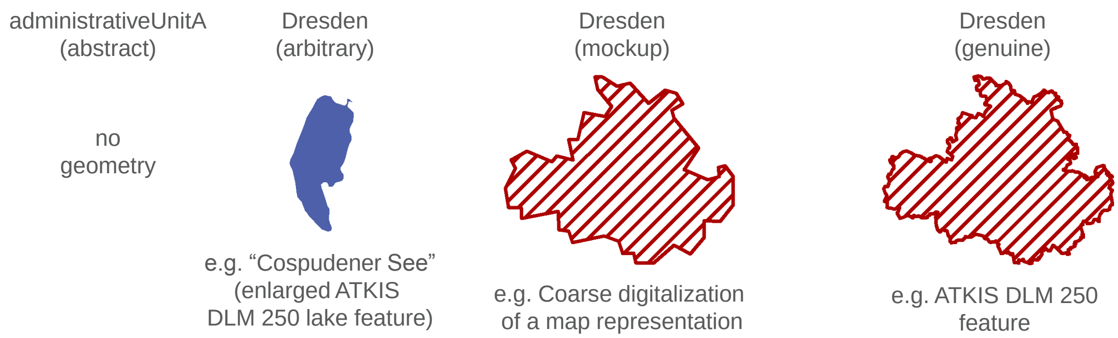

Figure 4. Data concretization states at the example of possible instances of an *AdministrativeUnit* feature type according to INSPIRE [15].

example to improve interpretation of exchanged data by project members.

ACKNOWLEDGMENT

The research described in this paper emerged from developments in the 'Design, Technologies and Usability of web-based Spatial Decision Support Systems for regional assessment of climate change impacts and adaptation – for the Model Region of Dresden' (ZENON; sCENarios ONline), a research project funded by the Leibniz Institute of Ecological Urban and Regional Development (IOER), Dresden, Germany. Furthermore, the authors thank colleague Robert Hecht and several anonymous reviewers for their advice and time.

REFERENCES

[1] C. Rinner, “Web-based Spatial Decision Support: Status and Research Directions,” Journal of Geographic Information and Decision Analysis (GIDA), vol. 7, no. 1, 2003, pp. 14–31, URL: http://www.ryerson.ca/~crinner/pubs/Rinner_published_GIDA-01-2003.pdf [accessed: 2015.01.08].

[2] R. Vogel, “ZENON – Entwicklung eines WebSDSS zur Abschätzung der Folgen des Klimawandels und des gesellschaftlichen Wandels [ZENON – Development of a WebSDSS for assessment of impacts of climate change and societal change],” in Angewandte Geoinformatik 2013: Beiträge zum 25. AGIT-Symposium Salzburg, J. Strobl, T. Blaschke, G. Griesebner, and B. Zagel, Eds. Berlin, Offenbach: AGIT Salzburg, 2013, pp. 227–232, URL: http://gispoint.de/fileadmin/user_upload/paper_gis_open/537533082.pdf [accessed: 2015.01.08].

[3] J. Schanze and A. Sauer, “Dokumentation des fragestellungs- und gebietsspezifischen Szenarioansatzes [Documentation of the topic-related and region-specific scenario approach],” Leibniz-Institut für ökologische Raumentwicklung, Dresden, REGKLAM Ergebnisbericht, Produkt 2.4a, Version: 1, Status: Konzept, unpublished, 2011.

[4] J. Schanze, J. Trümper, C. Burmeister, D. Pavlik, and I. Kruhlov, “A methodology for dealing with regional change in integrated water resources management,” Environmental Earth Sciences, vol. 65, 2012, pp. 1405–1414.

[5] R. Vogel, A. Sauer, and J. Schanze, “Object-Oriented Domain Modeling of a Regional Climate Change Scenario Approach as a Central WebSDSS Development Artifact,” N. N., in preparation.

[6] M. Neubert and R. Vogel, “Integrierte räumliche Schadensanalysen zum Aufbau der Datenhaltungskomponente eines webbasierten Entscheidungsunterstützungssystems [Integrated spatial damage analysis for establishing the database component of a web-based decision support system],” in Angewandte Geoinformatik 2014: Beiträge zum 26. AGIT-Symposium Salzburg, J. Strobl, T. Blaschke, G. Griesebner, and B. Zagel, Eds. Berlin, Offenbach: AGIT Salzburg, 2014, pp. 375–380, URL: http://gispoint.de/fileadmin/user_upload/paper_gis_open/537543055.pdf [accessed: 2015.01.08].

[7] V. Rainardi, Building a Data Warehouse With Examples in SQL Server. Berkeley: Apress, 2008.

[8] OGC, “OpenGIS® Implementation Specification for Geographic information – Simple feature access – Part 2: SQL option,” Open Geospatial Consortium, OpenGIS® Implementation Specification OGC 06-104r4, Version 1.2.1, 2010, URL: http://www.opengeospatial.org/standards/sfs [accessed: 2015.01.08].

[9] D. Haywood, Domain-Driven Design: Using Naked Objects. Raleigh, Dallas: The Pragmatic Bookshelf, 2009.

[10] R. Vogel, “Testanwendung eines technologieintegrierenden WebSDSS-Prototyps – Dokumentation des 2. Tests [Test application of a technology integrating prototype of a WebSDSS – Documentation of the 2nd test],” ZENON AP5 project report, unpublished, 2014.

[11] S. Kleuker, Grundkurs Software-Engineering mit UML: Der pragmatische Weg zu erfolgreichen Softwareprojekten [Basic course software engineering with UML. The pragmatic way to successful software projects], 2nd ed., ser. Studium. Wiesbaden: Vieweg + Teubner, 2011.

[12] J. Rumbaugh, I. Jacobson, and G. Booch, The Unified Modeling Language Reference Manual. Reading, Massachusetts: Addison-Wesley, 1999.

[13] O. Kretschmer, “GIS-basierte Projektionen der Siedlungsflächenentwicklung mittels multikriterieller Bewertungsfaktoren und zellulärem Automatenmodell DINAMICA [GIS-based projections of residential area dynamics with multi-criteria evaluation factors and the cellular automata model DINAMICA],” Diploma thesis, IOER Dresden and Philipps-Universität Marburg, unpublished, 2012.

[14] A. Yeung and G. Hall, Spatial Database Systems: Design, Implementation and Project Management, ser. The GeoJournal Library 87. Dordrecht: Springer-Verlag, 2007.

[15] INSPIRE, “Inspire data specification on administrative units – guidelines,” INSPIRE Thematic Working Group Administrative units, Data Specification D2.8.I.4, 2009-09-07 2009, URL: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AU_v3.0.pdf [accessed: 2015.01.08].

[16] R. Vogel, “Prototypische Technologieintegration für ein WebSDSS zur Folgenabschätzung des Klimawandels und des regionalen Wandels [Prototypical Technology Integration for a WebSDSS for Assessment of Impacts of Climate Change and Regional Change],” FB R-Kolloquium, presentation, 2012.08.08, IOER Dresden, 2012.

[17] G. Strode, “The GIS Behind iMapInvasives: The ”Open Source Sandwich”,” in Online Maps with APIs and WebServices, ser. Lecture Notes in Geoinformation and Cartography, M. Peterson, Ed. Berlin, Heidelberg: Springer, 2012, ch. 6, pp. 73–90.

[18] I. Benenson and P. M. Torrens, Geosimulation: Automata-based Modelling of Urban Phenomena. Southern Gate, Chichester, West Sussex: John Wiley & Sons, 2004.

[19] ISO 19115, “Geographic information – Metadata,” International Organization for Standardization, ISO 19115:2005-05, 2005.

[20] R. Vogel, “Zukünftedaten im Kontext der Entwicklung eines WebSDSS zur Analyse von Szenarien des Klimawandels und des regionalen Wandels [Future data in the context of the development of a WebSDSS for analysing scenarios of climate change and societal change],” ZENON AP4 & AP5 project report, unpublished, 2014.