Towards a Dedicated Data Model for Seamless Pedestrian Navigation

Lucia Tyrallová, Carolin Kucharczyk, Lasse Scheele, Hartmut Asche

University of Potsdam, Department of Geography, Geoinformation Research Group

Abstract. Awareness of one's location at any given time is a fundamental human need. To navigate vehicles or persons, geodata are indispensable for calculating the route. At present, there is a lack of a proper database that meets the specific requirements for spatially inclusive and comprehensive pedestrian navigation. Therefore, firstly, real-world objects relevant for pedestrian navigation need to be determined. The selected and captured data, secondly, need to be structured and stored in an application-specific database to make proper use of them for orientation and navigation purposes. This paper deals with a pedestrian-focused data model and feature catalogue and selected aspects of data quality relevant for pedestrian navigation.

Keywords. Pedestrian-focused data model, Feature catalogue, Data quality, Pedestrian navigation

1. Background and Objectives

Awareness of one's location at any given time is a fundamental human need. This mental and geodetic determination of position is facilitated through mental maps and landmarks or maps and navigation systems, respectively (Asche 2014).

Navigation is a movement resulting from positioning and route selection (Bollmann & Koch 2002). To navigate vehicles or persons, geodata are indispensable for calculating the route. At present, there is a lack of a proper database that meets the specific requirements for spatially inclusive and comprehensive pedestrian navigation. The divergent claims notwithstanding, navigation today is still mainly based on data for vehicle navigation,

Published in "Proceedings of the 11th International Symposium on Location-Based Services", edited by Georg Gartner and Haosheng Huang, LBS 2014, 26–28 November 2014, Vienna, Austria.

e.g. GDF data (Geographic Data Files). In recent versions (ISO 2011), a number of pedestrian-relevant objects have been included. Nevertheless, a uniform, seamless geospatial database dedicated to pedestrian navigation is not available for the time being, at least not in Germany.

To compensate for this situation, the investigation of databases with similar geospatial data is an appropriate R&D starting point. In Germany, the following databases, among others, can be analysed for the identification of spatial objects relevant to pedestrian navigation and inclusion in a dedicated database:

(1) Topographic data of the mapping agencies of the German federal states (FMAs). Existing data of the ATKIS or the recent AAA system[1], termed Digital Landscape Models (DLM), suffer the following shortcomings:

- Landscape objects captured by geodetic data acquisition are included in the DLM database only.
- Cadastral data of the ALK[2] database, recently integrated in the AAA system, and ATKIS data vary in granularity of geometric and attribute features due to differences in original data acquisition.
- ATKIS or AAA, respectively, lacks a uniform database for the whole of Germany in large to medium scales (< 1:200,000), since DLM data models are constructed and updated for the federal states only.

As a consequence, a seamless geospatial database that might be used as a basis for pedestrian navigation is not available from the FMAs to date. Hence, selected objects only can be relevant for setting up a database for pedestrian navigation.

(2) Topography-like data of the OpenStreetMap project (OSM), a free global geodatabase created by non-professional volunteers in large to small scales. As any user-generated content, OSM data display the following deficiencies:

- Both on a global and regional scale, completeness, topicality and quality of the OSM data pool varies considerably.

---

[1] AAA system describes (AFIS-ALKIS-ATKIS-model). ATKIS- *Amtliches Topographisch-Kartographisches Informationssystem* of the *Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland* (AdV), ALKIS- Amtliches Liegenschaftskatasterinformationssystem, AFIS- Amtliches Festpunktinformationssystem

[2] ALK -Automatisierte Liegenschaftskarte, which has been recently integrated in the AAA system, specifically in ALKIS.

- OSM operates on a less detailed geospatial data model and rather loosely defined data acquisition rules to make it possible for volunteers to contribute to the database.
- Depending on region or scale under investigation, reliable geotopographic data relevant to pedestrian navigation may or may not be included in the database.

For these reasons, the OSM database, like its FMA equivalents, can only be made limited use of as a homogenous, seamless database for pedestrian navigation. Figure 1 depicts an extract of the OSM dataset for a study area – the city centre of Potsdam (Germany). It can be seen that several spatial objects and relations relevant to pedestrian navigation, in this particular case from Jaegerstrasse (left) to Friedrich-Eberst-Strasse (right), are lacking. The red line shows the route calculated on the basis of OSM data between the two points in the centre of Potsdam. The green line represents an optimal pedestrian route through inner courtyards. The respective data are not available in the OSM database.

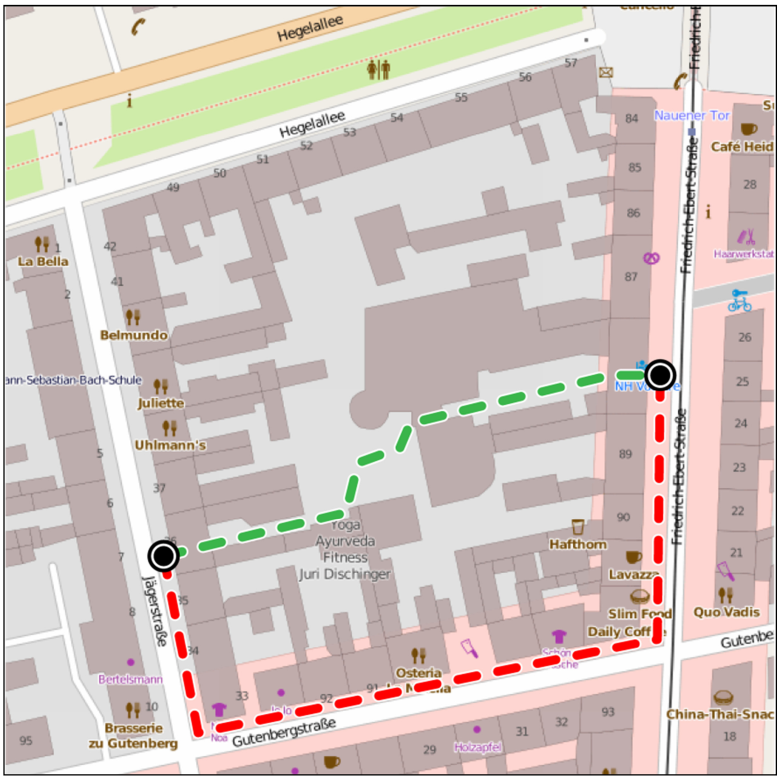

Figure 1. Reality vs. data base – the red dashed line shows the route calculated on the basis of OpenStreetMap data between two points in the centre of Potsdam, the green dashed line represents a shorter way through inner courtyards a pedestrian may take in reality.

What can be seen from the above examination of existing geospatial databases is that a reliable, uniform large-scale geospatial database for seamless pedestrian navigation is not available. Using databases of vehicle navigation systems instead is no alternative, as, e.g. Kluge & Asche (2008) have shown, since, in general, they do not include pedestrian-specific objects. One significant difference between pedestrian and vehicle navigation is manifest in the routing functions available. While cars drive on defined roads only, pedestrians are able to move freely in the whole of public space. This includes movement on/in, e.g., pavements, footpaths, public squares, shopping arcades, publicly accessible courtyards, public greens, parks, all inaccessible to automobile traffic (Figure 1). Pedestrian moving and navigation space, on the other hand, is much lesser restricted. This, however, only applies when pedestrians are considered a homogenous group, which, of course, they are not. Restrictions, such as barriers, become particularly apparent when considering mobility-impaired persons like wheelchair users or persons reliant on walking frames. Relevant barriers include stairs, steep paths and specific surfaces of roads, footpaths or pedestrian overpasses, such as cobblestones.

Both individual and requirement-matched pedestrian navigation therefore requires the definition of different user groups. For these groups as well as for the entire pedestrian community criteria need to be determined which real-world objects will have to be acquired, processed and included in a geospatial database specifically dedicated to pedestrian navigation. As with vehicle navigation systems, binding standards do not exist as guidelines for building and populating such database.

2. Concept and feature model for a pedestrian navigation database

In the following, we present an elaborated concept for a pedestrian navigation data model. This data model is conceived as a generic, modular, scalable system. Its structure is based on a hierarchy of objects and relations relevant to pedestrian navigation. In a first step, data model components can be sepatated into thematic data specific for pedestrian orientation and navigation, such as landmarks, points of interests, footpaths, and topographic data for a precise geospatial reference. In a second step, the data model can be broken down into a granular feature catalogue (FC), detailing objects, structures and relations.

A FC is a repository for geographic features the creation of which is covered by the ISO standard 19110 *Methodology for feature cataloguing* by the *International Organization for Standardization*'s *Technical Committee* (ISO/TC) 211 *Geographic information/Geomatics* (Kresse & Fadaie 2004).

Accordingly, a FC is a classification of real-world phenomena usable as a standard in many applications: a standardised and modular framework that enables the transferability to multiple applications and therefore minimises the time-consuming creation of a FC. At the same time, a FC should be an application-orientated abstraction of the real world accessible to users of the corresponding data. All in all, a FC is to provide a clear description of the contained features to facilitate a differentiation or combination of different FCs.

Data models and corresponding FCs do already exist for topographic data. It has been mentioned that the majority of those include, at most, a limited range of features of interest (FOI) relevant for a database of a pedestrian-focused navigation system. Among the existing FCs are those of FMAs, e.g. ATKIS in Germany and the OS MasterMap of the Ordnance Survey (OS), the National Mapping Agency (NMA) for the UK (OS 2006, Kuhn 2005, AdV 2008). The ATKIS data model, e.g., classifies areas and structures in six primary categories or domains: settlement, traffic, vegetation, water, relief and administrative divisions. These domains are subdivided in groups of objects, resulting in a structured model of the topographic landscape. The OS MasterMap is based on a similar data model with nine so-called top-level themes.

A comparable structure can be found in the OSM FC, which categorises features according to their thematic classification (Haklay & Weber 2008, Ramm et al. 2010). Of particular interest within the OSM FC is the *points of interest* (POI) category which includes many pedestrian-related FOI.

Investigation of existing FCs, like those of ATKIS and OSM, suggests a general structure of geospatial data models that a pedestrian navigation-orientated data model and FC should be based on. This includes the thematic domains of *traffic*, *settlement*, *water* and *vegetation* supplemented by a separate category for POIs (compare Feature Identifiers (FIDs) 1000-8000 in Figure 2). Further structuring in feature groups might be useful, e.g. by specifying the different means of traffic (FIDs 2100-2400) or kinds of POIs (FIDs 8100-8900). Single features can then be assigned to these feature groups (e.g. FID 2101). For instance, the domain *2000 Traffic* and the group *2100 Pedestrian traffic* would include the object class *2101 Path* with the attributes (e.g. function, surface or gradient) associated with it. Because of its modular design the FC is easily adjusted when other domains or new features need to be included.

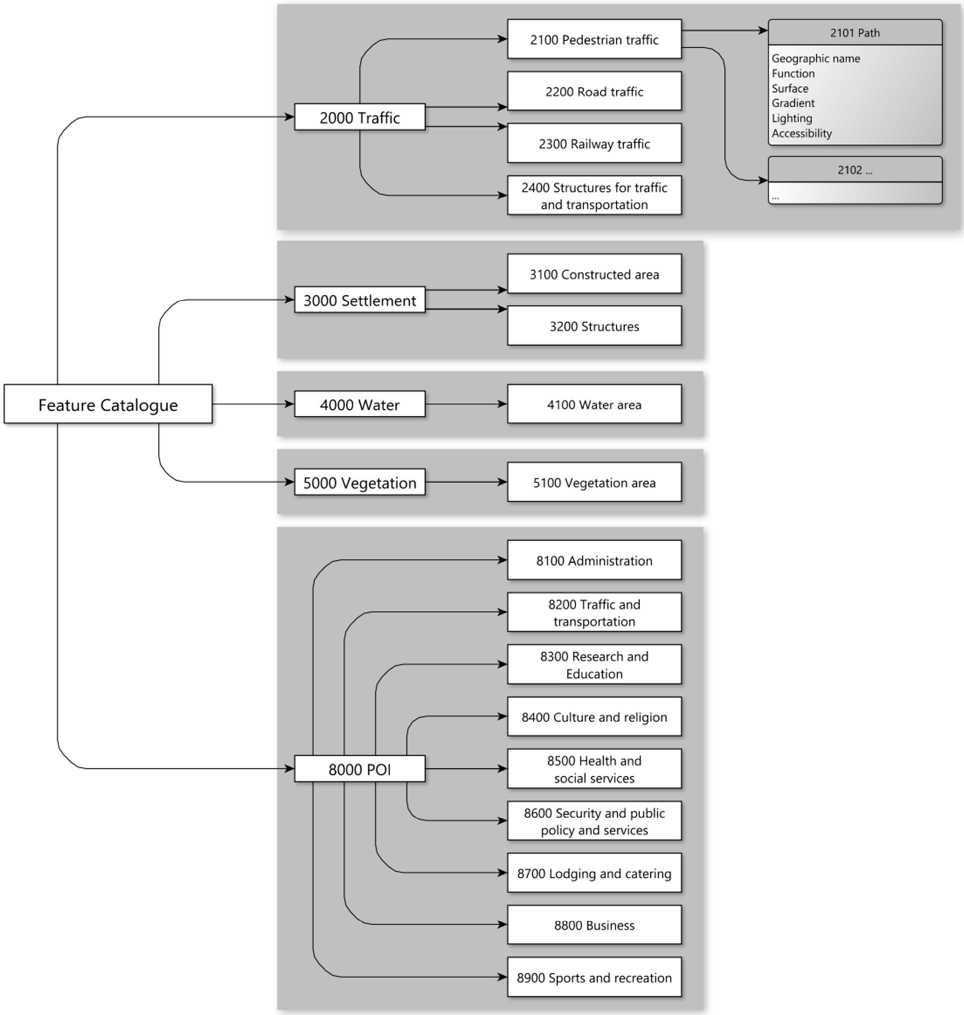

Figure 2. General structure for a FC detailing a data model for pedestrian navigation systems.

Based on the FC and the underlying data model data requirements for the FC objects can be defined. Data requirements are mandatory when it comes to assessing potential data sources and data acquisition techniques relevant to populate a pedestrian navigation database. Paramount in evaluating data suitability for a pedestrian navigation FC is the concept of data quality. Some elements of this concept are detailed below.

3. A tentative concept for data quality

For any spatial analysis, decision or application, the quality of input data is crucial, since the quality of a solution is directly dependent on the quality of the data available. Hence, pedestrian navigation (from calculating an optimal route to the design of communication media like map visualisations) can only be as good as the geodata on which it is based. For this reason it is necessary to formalise and standardise spatial data quality. ISO 19113:2002 defines quality as the "totality of characteristics of a product that bear on its ability to satisfy stated and implied needs" (ISO 2002). There are different parameters and indicators to describe and assess the quality of geodata. In principle, the selection of possible objective quality criteria and its associated quality measures has to be performed according to the respective application purpose (van Oort 2006). Selected quality criteria of significance for pedestrian navigation are described below (Laufer 2011, PAS 1071 2007, van Oort 2006).

Positional accuracy: Positional accuracy is defined as the accuracy of the coordinate values measured, compared to the real position of the object. A distinction can be made between absolute (comparison with another dataset) and relative positional accuracy (internal comparison in the same dataset).

Thematic accuracy: Thematic or attribute accuracy refers to the accuracy of quantitative attributes, the correctness of qualitative attributes as well as the classification of the features.

Completeness: Completeness defines the presence or absence of spatial objects, attributes or relationships. Data may exhibit a loss of quality not only with regard to incompleteness (error of omission) but also overcompleteness (error of commission).

Logical consistency: Logical consistency is the degree of compliance with logical rules concerning the data structure. Aspects can be valid values, conceptual, format, geometric, thematic and topological consistency.

Temporal quality: In addition to the spatial and attributive information geodata may include temporal information, too. Temporal quality can be described by the accuracy of time measurements, temporal consistency and temporal validity.

Timeliness: Navigation in general requires up-to-date data. Timeliness defines the degree of conformance between the information and the time-varying conceptual reality.

Availability: Availability indicates the level of existence of the information at a given point in time and a defined location.

Inconsistent, inaccurate or incomplete data as well as imprecise location information lead to low-quality data resulting in errant datasets effecting faulty decisions. For this reason, a definition of objectives and application-dependent quality criteria and related quality measures and their subsequent formalisation in a criteria catalogue is needed. This catalogue can be used to decide if an existing dataset is suitable for the intended purpose, or to ensure the demanded quality when capturing new data.

An evaluation of data quality is based to the application of direct or indirect methods (ISO 2005). Direct methods are applied when the available data are compared with the real world or a reference dataset of higher quality. Evaluating data based on descriptive parameters (e.g. metadata) is the application of indirect evaluation methods. Generally, a data quality model consists of two components: data quality and model quality (Joos 2000). Hence consideration of both parameters is important to conceptualise and implement an effective system for pedestrian navigation.

4. Conclusion and Outlook

This article focuses on the importance and requirements of pedestrian specific geodata for orientation and navigation. It has been demonstrated that there is a lack of spatial information adapted to the need of pedestrian navigation as well as missing standards concerning the collection of these data.

Establishment of a geospatial database adjusted to pedestrian-specific requirements is of paramount importance. It will fill the existing data gap and minimise the redundancy of already existing data sets.

The purpose of data processing is to complement existing spatial data by expansion of user-group-related attributes. One method of obtaining data for pedestrian navigation can be the identification and extraction from remote sensing imagery, if possible, in an automated process (Tyrallová 2013). The extraction of, e.g., pedestrian paths from operational remote sensing data can be used to enhance the existing OSM database. Geometric and spectral resolution of the remote sensing images has been greatly improved in the last 30 years. For the identification of pedestrian-specific objects and their attributes the current data resolution is sufficient.

Additional services such as detection and mapping of traffic signals, traffic signs and pedestrian and bicycle paths can be effected by crowd-sourcing. Volunteers are provided with detailed data capture rules and are remunerated for the objects acquired. Before inclusion in a pedestrian navigation, these data will be subjected to a rigorous quality assessment. One option is to these newly gathered data with existing data in automated data fusion. In this conflation process two (or more) suboptimal datasets are merged into a new optimal data set with enhanced geometric and/or thematic quality (Stankute & Asche 2012). This method can also be employed to minimise data redundancy.

Next steps towards a dedicated data model for seamless pedestrian navigation include the specification of a granular FC on feature and attribute levels. The FC will detail a pedestrian navigation data model that has to be finalised and validated. The existence of such data model is a prerequisite for target-oriented data acquisition, such as high-resolution operational remote sensing imagery or crowd-sourcing acquisition techniques applied by volunteers or micro-jobbers. Concept and content of a data model dedicated to pedestrian navigation have been presented and discussed in this paper.

Acknowledgements. This paper is based on R&D for a feasibility study on seamless pedestrian navigation. The support of this research by the German Federal Ministry of Education and Research (BMBF) is gratefully acknowledged.

References

Asche H, (2014) Ringvorlesung Grako 1539 | Sichtbarkeit und Sichtbarmachung, Universität Potsdam

AdV (2008) Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). ATKIS-Objektartenkatalog Basis-DLM Version 6.0

Bieber R, Gerlach P, Klinke R, Scheibner K, Schulze M, Ziegler S, Kluge M (2011) Fußgängerbezogene Datenaufbereitung in OpenStreetMap. In: FOSSGIS e.V. (Hrsg.) FOSSGIS. Heidelberg, pp. 168-177.

Bollmann J, Koch W-G (2002) Lexikon der Kartographie und Geomatik in zwei Bänden, zweiter Band Karto bis Z. Spektrum Akademischer Verlag, Heidelberg, Berlin

Haklay M (2008) How good is Volunteered Geographical Information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B: Planning and Design, vol. 37 no. 4, pp. 682-703

Haklay M, Weber P (2008) OpenStreetMap: User-Generated Street Maps. Pervasive Computing, IEEE, vol. 7, no. 4, pp. 12-18, Oct.-Dec. 2008, doi:10.1109/MPRV.2008.80

Helbich M, Amelunxen C, Neis P, Zipf A (2012) Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata. In: Jekel T, Car A, Strobl J & G Griesebner (eds.), GI_Forum 2012. Conference Proceedings. pp. , 24-33. Wichmann Verlag, Berlin, Offenbach

ISO (2002) DIN ISO 19113:2002 Geographic information – Quality principles, 29 p.

ISO (2005) DIN ISO 19114:2005-05: Geoinformation – Verfahren zur Ermittlung der Datenqualität (ISO 19114:2003); Englische Fassung EN ISO 19114:2005. Beuth-Verlag, Berlin

ISO (2011) DIN ISO 14825:2011: Intelligent transport systems – Geographic Data Files (GDF) – GDF5.0

Joos G (2000) Zur Qualität von objektstrukturierten Geodaten. Schriftenreihe des Studiengangs Geodäsie und Geoinformation der Universität der Bundeswehr München, Heft 66

Kluge, M. Asche H (2008). 3D pedestrian navigation - Reality view. In: Gartner, G., Leitinger, S. & K. Rehrl. Keynotes, Extended Abstracts, LBS Showcases: LBS & TeleCartography. 5th International Symposium on LBS & TeleCartography. Salzburg, Austria 26-28 Nov 2008. Salzburg research/Technical University of Vienna, Salzburg. [online]

Kresse W, Fadaie K (2004) ISO Standards for Geographic Information. Springer, Berlin

Kuhn W (2005) Geospatial Semantics: Why, of What, and How? In: Spaccapietra S, Zimányi E (eds.) Journal on Data Semantics III. Springer, Berlin, Heidelberg

Laufer R (2011) Prozedurale Qualitätsmodellierung und -management für Daten - ingenieurgeodätische und verkehrstechnische Anwendungen. Universitätsbibliothek der Universität Stuttgart, Stuttgart

Neis P, Zipf A (2012) Analyzing the Contributor Activity of a Volunteered Geographic Information Project – The Case of OpenStreetMap. ISPRS International Journal of Geo-Information, 1, pp. 146-165

OS (2006) The OS MasterMap. Part 1: user guide. http://www.centremapslive.co.uk/files/osmmuserguide.pdf. Accessed: 1 October 2014

PAS 1071 (2007) Qualitätsmodell für die Beschreibung von Geodaten. Deutscher Dachverband für Geoinformation (DDGI)

Ramm F, Topf J, Chilton S (2010) OpenStreetMap: Using and Enhancing the Free Map of the World by UIT Cambridge, 386 pp.

Stankute S, Asche H (2012) A Data Fusion System for Spatial Data Mining, Analysis and Improvement. In: Murgante B, Gervasi O, Misra S, Nedjah N, Rocha A M, Taniar D & B Apduhan (eds.) 12th International Conference on Computational Science and Its Applications, Salvador da Bahia, Part II LNCS 7334. Springer, Berlin, Heidelberg, pp. 439-449

Tyrallová L (2013) Automatisierte Objektidentifikation und Visualisierung terrestrischer Oberflächenformen. PhD Thesis, Universität Potsdam

van Oort P (2006) Spatial data quality: from description to application. PhD Thesis, Wageningen Universiteit

Zielstra D, Zipf A (2010) A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. AGILE 2010. The 13th AGILE International Conference on Geographic Informatione Science. Guimarães, Portugal