International Cartographic Association
Association Cartographique Internationale

A formal description of topographic space – maps contributing to AI

Markus Jobst [a,b]*, Georg Gartner [a]

[a] *Research Unit Cartography, Department of Geodesy and Geoinformation, Vienna University of Technology, Austria*
markus@jobstmedia.at, georg.gartner@tuwien.ac.at

[b] *Federal Office of Metrology and Surveying, Austria*

* Corresponding author

Keywords: map production, spatial knowledge infrastructure, artificial intelligence, geo-communication, foundational topographic ontology

Abstract:

Maps play a key role for describing and relating information to space [Kitchin et al 2011]. Several rules within map production aim at the effective and efficient communication of spatial related knowledge [Brodersen 2017]. Beside a number of "technical" rules to enhance perception, also different parameters of communication, linguistics and information encoding need to be considered. Most of these rules are embedded in semi-automated procedures for industrial map production [Wabiński 2019] and still make use of human design and composition decisions [Kalbach 2020]. By observing the developments of artificial intelligence (AI) and by evaluating examples like ChatGPT [Deng et al 2022], LaMDA [Thoppilan 2022] or DALL-E [Kapelyukh 2022], its relevance for a fully automated map compilation becomes clear. The same methods used in natural language processing (NLP), grammar compilation and image rendering could easily be adopted to cartography. As soon as the amount of available and appropriate spatial information is sufficient and the basic requirements are fulfilled, algorithms could decide in an automated production procedure. Maps could then be created with a human command like "create a topographic map in scale 1:25k of region A with highlighted flooding areas and destroyed transport infrastructure".

A sufficient amount of appropriate spatial information allows for an extensive training of algorithms for "true" and "false" results [Tracewski 2017] and the fine-tuning of decisions within algorithms with the help of machine learning [Bonaccorso 2020]. This continuous learning results in a formal knowledge stack for the set of algorithms for specific situations and tasks, e.g. map creation in our purpose. Therefore training and the creation of a "statistical knowledge stack" is the most essential part that leads to success of an artificial intelligence, which depends on the existence of a large amount of appropriate (spatial) information chunks.

Basic requirements in order to automatically process digital information for AI purposes in spatial related use cases cover spatial vocabularies, taxonomies and the formal description of spatial knowledge [Bader et al 2020]. Some requirements for automated spatial information processing were realized within spatial data infrastructures (SDI) and their service-oriented architecture approach [Tripathi et al 2020], where each component provides sufficient information about itself, its content and its relations to other components. Although for many SDI the content and schemas are defined [Chojka 2020], the implemented approach still lacks information depth and knowledge structures that are needed for an automated processing in AI.

The way data scientists name as well as define information depends on the use case, their experiences and the specific situation, where this information becomes part of expression and communication. This contextualized knowledge utilization by humans may lead to different names for the same real-world object. For this reason controlled vocabularies play an important role in making the name/definition precise. Automated procedures of machines will hardly recognise any context or relation if the naming is not precise. In addition any classification and aggregation of vocabularies has to be fixed in an extensible schema, which is mainly called taxonomy [Sibolla et al 2016].

Abstracts of the International Cartographic Association, 6, 110, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-110-2023 |

Actually spatial data scientists make use of controlled vocabularies and taxonomies on one hand to describe geospatial data and their schemas and on the other hand to name features, constellations, relations and situations. In addition they could think of extending these vocabularies and taxonomies for the styling in maps, their production procedures (e.g. generalization rules) or scale dependent simplifications. All together becomes structured as ontology to represent the comprehensive and formalized knowledge [Ristoski et al 2016]. In general ontologies and its specific knowledge description is built according to an upcoming question (finding an answer for a specific use case).

Prominent examples for the application of ontologies, vocabularies and taxonomies are search engines, which need to find relevant content according to our needs. Therefore schema.org [Guha et al 2016] collects all relevant vocabularies, taxonomies and builds one central ontology for searching content on the Web. A small part of it describes spatial information in the category "Place". A more detailed look shows that in terms of cartography, the existing categorization may of course not be sufficient for automated tasks in map production. The category "Place" consists of Accommodation, AdministrativeArea, CivicStructure, Landform, LandmarksOrHistoricalBuildings, LocalBusiness, Residence, TouristAttraction and TouristDestination. From a topographic point of view, the category "Landform" is of main interest. It splits into BodyOfWater, Continent, Mountain and Volcano. "BodyOfWater" is the only category that has subcategories and contains Canal, LakeBodyOfWater, OceanBodyOfWater, Pond, Reservoir, RiverBodyOfWater, SeaBodyOfWater and Waterfall. Map legends as representation of vocabularies and taxonomies that are used in maps easily show that more classes seem to be useful for the formal description of topographic space.

In the continuous work of the authors they highlight the role of maps for controlled topographic vocabularies, taxonomies and spatial knowledge representations. Existing map legends contribute to semantic structures in spatial knowledge infrastructures. The description of space within topographic maps establishes a foundational topographic ontology [Jobst 2022], that helps to address geospatial features, relations, situations and label its semantics.

The technology stack that is used to make a foundational topographic ontology work, follows the Semantic Web, interfaces of linked data and methods of service-oriented architectures. Latest developments, like OGC API´s, consider a massive processing of geoinformation for machine learning or artificial intelligence. Their API´s interfaces differ from the geographic markup language (GML) and often make use of GeoJSON. Especially Comma Separated Values (CSV) are important for data science applications and gain importance in the spatial open data domain and modern processing. For example search engines like elastic or solr making use of this geo-shape datatype [https://www.elastic.co/guide/en/elasticsearch/reference/6.2/geo-shape.html#input-structure]. In these datatypes the spatial encoding is the done via WKT (well known text), a standardized text markup language for representing vector geometry objects [https://www.ogc.org/standards/sfa]. Linked Data brings in the Web paradigm, which connects different data, structures and semantics, to the spatial data storage and processing tier. As result we are establishing Spatial Data on the Web [van den Brink et al 2019] in near future realisations of our SDI. The trigger for cross-domain collaboration begins to live with the growing amount of accessible sources, which enables the training for algorithms, but also makes it impossible, due to their big data characteristic (volume, variety, velocity), to store the information redundantly on one single node. Any identification and selection of features in this network of big data nodes may request a foundational topographic ontology, for which parts of it can already be represented by the content and structure of map legends, to correctly extract information chunks and relate to their semantics. When maps follow this common approach for a formal description of space and provide a foundational topographic ontology, they could directly contribute to and support applications in AI.

Selected References

Bader, S., Pullmann, J., Mader, C., Tramp, S., Quix, C., Müller, A. W., Lange, C., 2020. The International Data Spaces Information Model–An Ontology for Sovereign Exchange of Digital Content. In International Semantic Web Conference, Springer, Cham, , pp. 176-192.

Chojka, A., 2020. Semantic and Syntactic Interoperability Issues in the Context of SDI. Geomatics and Environmental Engineering, 14(3).

Deng, J., & Lin, Y., 2022. The Benefits and Challenges of ChatGPT: An Overview. Frontiers in Computing and Intelligent Systems, 2(2), 81-83.

Abstracts of the International Cartographic Association, 6, 110, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-110-2023