Review

Geospatial Data Management Research: Progress and Future Directions

Martin Breunig [1], Patrick Erik Bradley [1,*] , Markus Jahn [1], Paul Kuper [1,*], Nima Mazroob [1], Norbert Rösch [1], Mulhim Al-Doori [2], Emmanuel Stefanakis [3] and Mojgan Jadidi [4]

[1] Geodetic Institute, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany; martin.breunig@kit.edu (M.B.); markus.jahn@kit.edu (M.J.); nima.mazroob@kit.edu (N.M.); norbert.roesch@kit.edu (N.R.)

[2] College of Information Technology, University of Fujairah, P.O. Box 1207 Fujairah, UAE; maldoori@uof.ac.ae

[3] Department of Geomatics Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada; emmanuel.stefanakis@ucalgary.ca

[4] Department of Earth & Space Science & Engineering, Lassonde School of Engineering, York University, Toronto, ON M3J 1P3, Canada; mjadidi@yorku.ca

* Correspondence: patrick.bradley@kit.edu (P.E.B.); paul.kuper@kit.edu (P.K.)

check for updates

Abstract: Without geospatial data management, today's challenges in big data applications such as earth observation, geographic information system/building information modeling (GIS/BIM) integration, and 3D/4D city planning cannot be solved. Furthermore, geospatial data management plays a connecting role between data acquisition, data modelling, data visualization, and data analysis. It enables the continuous availability of geospatial data and the replicability of geospatial data analysis. In the first part of this article, five milestones of geospatial data management research are presented that were achieved during the last decade. The first one reflects advancements in BIM/GIS integration at data, process, and application levels. The second milestone presents theoretical progress by introducing topology as a key concept of geospatial data management. In the third milestone, 3D/4D geospatial data management is described as a key concept for city modelling, including subsurface models. Progress in modelling and visualization of massive geospatial features on web platforms is the fourth milestone which includes discrete global grid systems as an alternative geospatial reference framework. The intensive use of geosensor data sources is the fifth milestone which opens the way to parallel data storage platforms supporting data analysis on geosensors. In the second part of this article, five future directions of geospatial data management research are presented that have the potential to become key research fields of geospatial data management in the next decade. Geo-data science will have the task to extract knowledge from unstructured and structured geospatial data and to bridge the gap between modern information technology concepts and the geo-related sciences. Topology is presented as a powerful and general concept to analyze GIS and BIM data structures and spatial relations that will be of great importance in emerging applications such as smart cities and digital twins. Data-streaming libraries and "in-situ" geo-computing on objects executed directly on the sensors will revolutionize geo-information science and bridge geo-computing with geospatial data management. Advanced geospatial data visualization on web platforms will enable the representation of dynamically changing geospatial features or moving objects' trajectories. Finally, geospatial data management will support big geospatial data analysis, and graph databases are expected to experience a revival on top of parallel and distributed data stores supporting big geospatial data analysis.

Keywords: geo-data management; big geospatial data; *nD* geo-database; topology; graph database; artificial intelligence

1. Introduction

The massive use of geo-referenced data sets in many fields of science and economy including earth observation [1], environmental sciences [2], city planning [3], BIM [3,4], real-time processing [5,6], and analytics for geospatial data [5] makes geospatial data management increasingly a central task in the workflow of geospatial data processing [1,5,7]. Recent approaches consider open data platforms and containers to handle big geospatial raster and vector data efficiently. However, little benchmarking has been done to compare the efficiency of the different big data stores especially in respect of their spatial functionalities [7], and big data support is not the whole story. For example, geospatial data management is also responsible for the cumbersome task of data integration from heterogeneous data sources [3,4,8].

In this article, we use the term "geospatial data management" for the insertion, deletion, update, and retrieval of geo-referenced geometric, i.e., vector or raster data sets, as well as topological data sets. However, the close connection of geospatial data management with other research fields within geo-information science such as data modelling, data visualization, and data analysis is the focus of our interest. The main contributions of this study are first to give an overview of past research during the last decade, coming up with five milestones, and secondly to highlight future directions of geospatial data management research that have the potential to become key research fields in the next decade.

Geospatial data management is also heavily influenced by the fusion of application fields such as BIM and GIS. As reported by the Web of Science and presented in Figure 1, the number of publications dealing with BIM and GIS integration has increased substantially since 2015, which reflects the significance of BIM and 3D GIS integration by both AEC and OGC communities. Alongside, a new term for integrated BIM and GIS data is coined as GeoBIM, and a benchmark (https://3d.bk.tudelft. nl/projects/geobim-benchmark/) was initiated by ISPRS Commission IV communities to foster the emergence of BIM and GIS integration in research and development engaging academia, stakeholders, government, and industries.
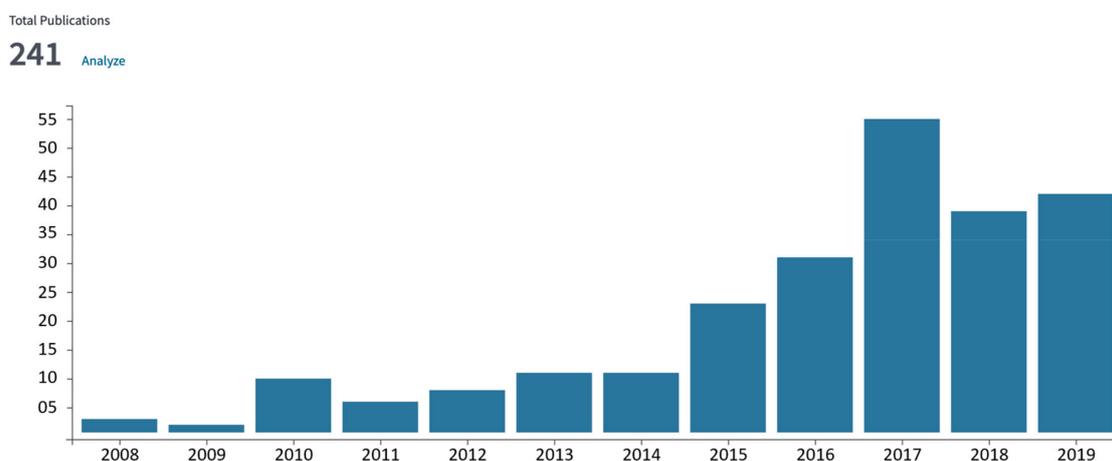
Figure 1. Publication report by the Web of Science for BIM and GIS integration (exerted from https://apps.webofknowledge.com/).

This paper is structured as follows: In Section 2, we refer to research in the field of geospatial data management closely referring to ISPRS progression during the last decade, followed by future directions of geospatial database research in Section 3. Finally, Section 4 presents the conclusions of the paper.

2. Progress during the Last Decade

During the last decade, significant progress has been achieved in geospatial data management research. This includes application-near research as well as theoretical results. In the following we present five milestones that reflect the wide spectrum of this research.

2.1. Milestone 1: Advancing GIS/BIM Integration at Data, Process, and Application Levels

For geospatial data management, the fusion of GIS and BIM [9–12] means that two different "philosophies" of spatial data representations are growing together: composition models such as triangle networks used for surfaces and tetrahedron networks used for solids on the one (GIS) side, and constructive solid geometry (CSG) and sweep models on the other (BIM) side. Thus, both kinds of spatial data models have to be considered as geometric and topological data types in geospatial databases. Additionally, point clouds from terrestrial laser scanning [12] have to be considered as important data sources for BIM. They additionally have to be handled in geospatial databases.

Recently, 3D virtual city space has been noticed as a key characteristic of smart sustainable cities [4]. To foster dynamic city applications such as business growth and solutions for monitoring infrastructure, buildings, and the built environment life cycle, many cities around the world have launched 3D open data to the public [3]. Geospatial communities have been highly active in the collection of 3D data using laser point cloud and imaging and in providing 3D GIS models of building envelopes and city spaces [13]. On the other hand, construction communities have offered BIM and rich 3D geometric and semantic information of internal building structure. Also, BIM is being considered as a rich intelligent digital repository that uses an object-oriented (OO) approach to describe the characteristics, i.e., semantics, geometry, and relationships in the architecture, engineering and construction (AEC) domain [4]. Although there are many open BIM standards such as BIMXML and COINS, industry foundation classes (IFC) have become the most popular open standard BIM format for interoperability purposes in the AEC industry developed by the BuildingSMART community [8]. In practice, building elements in BIM are represented and stored as CSG for simple objects, sweep volumes for paths to define a solid, and boundary representation (B-Rep) for its bounding surfaces with implicit topology either in EXPRESS-based or XML-based formats [8,14]. In the industrial foundation classes, BIM models are divided into five levels of development (LOD) marking the course of the construction within the life cycle of each building information model: LOD 100 stands for "conceptual", 200 for "approximate geometry", 300 for "precise geometry", 400 for "fabrication", and 500 for "as-built". Thus, LOD 100 is a simpler model than LOD 200, etc. This means that the resolution of the model increases during its construction process. Figure 2a shows some of the different LODs for a special building part. In contrast to this, 3D geospatial information systems (GIS) focus on collecting, storing, and analyzing geospatial data at a small scale. More than two decades ago, the Open Geospatial Consortium (OGC) developed the City Geography Markup Language (CityGML) standard data model, an Extensible Markup Language (XML-based) standard for geospatial data interoperability and exchange using the BRep geometry model. This geometry model can also be used for the integration of GIS and BIM geometries [15]. Similar to IFC for BIM, CityGML divides 3D GIS data into four levels of detail (LoDs), from LoD 1 to LoD4 as presented in Figure 2b. Here, the building or city model usually is provided at the highest resolution (e.g., LoD4), and abstractions from this high-resolution model are made by changing the level of detail from LoD4 to LoD3, etc. Furthermore, LoD1-3 describes the model from the outside, whereas LoD4 presents the interior of the model (indoor model).
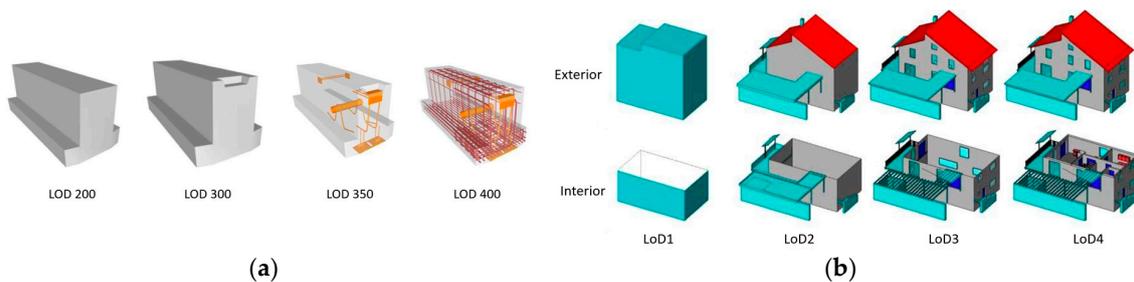
Figure 2. Different LODs in IFC for a precast structural inverted T beam (extracted from https://bimforum.org/lod/) (**a**) and LoDs for CityGML of a residential building (courtesy of Karlsruhe Institute of Technology) (**b**).

It is obvious that both BIM and 3D GIS are developed with fundamentally different purposes and sometimes contradictory geometries, semantics, and relationship data models. Although these differences between BIM and GIS became the target of many studies, they do share common characteristics such as data error checking (clash detection and topology analysis), *nD* simulation, data contents, and extensibility [8]. Therefore, over the last decade, the requirements for linking cross-domain specific information have increased [4]. Obviously, multiple applications in the AEC domain require adjoining information for pre- and post-construction phases, involving GIS data. On the contrary, GIS is enhancing from world and city level information to specific entity or detailed building level information which was originally in the AEC domain [15]. However, a building in the GIS domain even in LoD4 is less complete than in BIM. Thus, these vital conditions have accelerated the focus of bringing the BIM and GIS domains together, and during the last decade, a large number of studies have been conducted to achieve an integration of these heterogeneous systems [4,8,15–18]. In addition, BIM and GIS integration can provide essential capabilities of quantitative and semantics analysis as well as visualization opportunities for knowledge discovery and informed decision making in smart sustainable cities and the built environment.

Based on the recent review by [8], the integration of BIM and GIS has been conducted at the fundamental level to the application level. The fundamental level focuses on data exchange standards and interoperability at the data level, while the application level deals with developing new methods with full potential to exploit the BIM and GIS systems together [16]. Also, this can be divided into five groups such as schema-based, service-based, ontology-based, process-based, and system-based methods [18]. Furthermore, a comprehensive three-level data, process, and application level framework was developed by [17] as presented schematically in Figure 3. At the data level, models and structures are converted, modified or extended to meet the requirements by using extract, transform and load (ETL) techniques. Process-level integration encompasses data standards from BIM and GIS to be simultaneously adopted in a cooperative way stipulating the flexibility of data interoperability among systems. Application-level integration deals with the development of new applications with integrated BIM-GIS or extending existing applications such as plugins. The application approach is stated to be costly and time consuming due to its inflexibility in different domains where no software so far can read and process both BIM and GIS data at once without having required plugins [18].
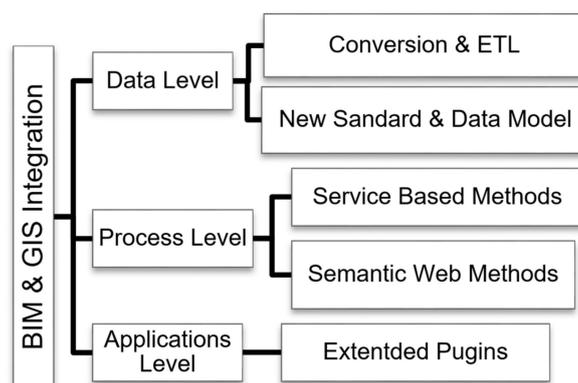
Figure 3. BIM and 3D GIS integration approaches at three levels: data, process, and application.

However, a wide range of studies has been conducted at the data level to achieve interoperability among BIM-GIS data formats with promising results [19]. These include linking, translation/conversion, extension, and meta-models (mediation) and can be further divided into geometric level and semantic level integration [18]. The geometric-based integration addressed the problem of reference systems (local vs. geodetic coordinate systems), 3D geometry (B-Rep, CSG, sweep volumes vs. B-Rep or ESRI$^{®}$ multi patch), and level of details (LOD vs. LoD). Although there are huge achievements for such conversion or integration even with commercial software packages such as Feature Manipulation Engine (FME$^{®}$) from Safe Software$^{®}$, BIMServer$^{®}$, and IfcExplorer$^{®}$, all these studies focused only on a single building structure rather than the context of the standard GeoBIM having multiple built structures at the city scale [20]. Also, most studies focused on IFC to CityGML exchange where the opposite direction is important to be considered [20]. Indeed, the exchange from CityGML to IFC seems more complex since there are more classes of objects in IFC than in CityGML. Thus, the transformation from CityGML leads to many empty classes in the IFC domain. The semantic-based integration methods include schema extension by ADEs, simplification, or new intermediate schema creation. Among the modification and introducing new models at semantic-level integrations, Semantic Web Technology has proven to be promising with its flexibility of integrating heterogeneous data formats.

Recent studies by [15] suggested that integrating BIM and GIS requires an understanding of the ontology of each data model first, then studying the similarity of the geometry and semantics for each data item. The core concept of the semantic web methods relies on the ontologies of two domains via a hierarchical graph structure. This usually consists of three steps: Ontology generation for Open Building Information Management (OBIM) and Open Geospatial Open Geodata Interoperability Specification (OGIS), ontology mapping to link the similar relationships or concepts, OBIM-GIS, and querying OBIM-GIS in the application domain using the SPARQL protocol and RDF query language (SPARQL) in order to retrieve the information needed from the model [4,15]. The main challenge in semantic web methods is ontology generation where there is no unified and accepted method to achieve this with a solid quality measurement matrix. Also, a lack of automatic generation of ontology is another burden in this method. Therefore, this leads to potential opportunities for future research in BIM and GIS integration using the Web Ontology Language (OWL).

Finally, topology has been proven to be an important concept for unifying GIS and BIM concepts as well as to describe GIS and BIM data structures and methods.

2.2. Milestone 2: Advancing Topology as a Key Concept for Geospatial Data Management

Topology, as it is used in the GIS/BIM world, models the relationships between spatial entities, e.g. it can express the fact that a certain object is part of the border of another object. This example relationship can be labeled bounded by. Certain types of spatial analysis, e.g., such that rely on connectivity queries or path finding, require topology to be modeled. Certain relationships bear a "topological" notion, such as contains, covers, bounded by, and touch. On the other hand, as

Alexandrov already observed in 1937, any binary relationship defines a topology in the mathematical sense, and any topology of a finite set of entities comes from a binary relation [21]. Consequently, the topology of the GIS/BIM world is a part of the mathematical domain named topology. Furthermore, this kind of topology can model the relationships between any kind sof entities, whether they are spatial or not. The geo-information community has not seemed to be aware of this fact yet.

A famous example of the aforementioned kinds are Egenhofer's 9-intersection matrices which are nothing but a means of classifying spatial relationships between regions [22] and which were extended to raster data in [23]. However, due to Alexandrov's result, topology is not restricted to relationships which are viewed as "spatial". A special case is given by binary relations which form a partial order, i.e., which do not allow cyclic iterated relationships between distinct objects. Obviously, GIS/BIM topologies exclusively deal with topologies given by partial orders.

An important example for topology in GIS or BIM is given by the adjacency graph whose nodes model GIS/BIM objects such as regions or rooms, and an edge is defined whenever two objects are adjacent. The first time that this concept from GIS was used in BIM is probably the article [24], cf. also [25] for the three-dimensional case. [26] contains an application to emergency response in buildings. Also, indoor networks from IFC models rely on correctly capturing the topology in buildings [27].

Undirected graphs are well-known examples of one-dimensional Alexandrov topological spaces. The introduction of simplicial complexes has been an important step towards generalizing graph topology to higher dimensions for GIS/BIM. These are used in [28] to model spatial or space-time scenarios. Here, there is the important relation face of, as simplices have faces which are also simplices. This defines the so-called incidence graph. Its reflexive and transitive closure is a partial order of all simplexes. A natural generalization of a simplicial complex is given by cell complexes or cw-complexes. They have a boundary operator between the vector spaces spanned by the cells (where each cell produces a new dimension). This boundary operator has a natural representation in relational databases, called relational chain complex, and the face topology (i.e., incidence graph) can be easily derived from this representation [29].

Manifold-like spaces are often modeled as so-called generalized maps (G-maps) [30] whose entities are the darts which are nothing but maximal relationship chains in the underlying partially ordered set. These darts are also called cell tuples, as their elements are cells at the boundary of other cells. A certain structure is imposed on the set of darts with the help of a family of involutions. In [31], it was observed that G-maps for higher dimension tend to be verbose, as the worst-case number of maximal chains in a partially ordered set is exponential to its size. This is in contrast with the worst-case storage complexity of the incidence graph itself, which is quadratic in the number of objects, and this is also a worst-case lower bound for storing topology [32]. In other words, the incidence graph is storage efficient. Nevertheless, combinatorial maps which are similar to G-maps have been recently used for the topological reconstruction of CityGML data [33].

Topology, geometry, and semantics come together in 3D plus time, i.e., 3D/4D applications that play a major role in planning scenarios and geo-scientific applications.

2.3. Milestone 3: Advancing 3D/4D Geospatial Data Management

In the last decade there have been significant advancements in the field of 3D/4D geospatial databases. The parallelization of queries using n-dimensional space-filling curves [34] has been examined [35] and applied to massive point clouds [35–38]. Furthermore, following the concepts of Langran and Stuart [39], spatio-temporal data models including moving surface and solid geometries have been implemented in object-relational [40–43] and object-oriented [44,45] geospatial database management systems. The latter have been implemented by time-dependent net components containing simplicial complexes [46,47]. As applications for these spatio-temporal data models, geological strata and faults of open mines have been used extensively [44,45,47].

One approach to classify the various lines of development for spatio-temporal data models is to distinguish them by their data models [47]. For 3D/4D geospatial databases, four different prominent

data models can be identified depending on if they are focusing on geometry, topology or semantics: point clouds, simplicial complexes, other topological approaches, and CityGML.

For an efficient management of massive point cloud data, space-filling curves [34] are frequently used on top of object-relational databases [35,36]. Massive data sets were tested within different environments and point cloud benchmarks were developed [37]. This research includes solutions to handle the level of detail (LoD) of various point cloud properties [38] by discrete global grid systems (DGGSs) [48].

Usually, simplicial complexes [46] are used as a data model for the geometric modelling of natural structures [49,50] in the geosciences because they allow a better approximation of the geometry in 3D space than grids and ease geo-computation, as the algorithms can be broken down to basic geometric operations on triangles and tetrahedra. Whenever the data are modelled over a period of time, multiple time steps can be combined and form a spatio-temporal object. The so-called snapshot model is a well-known concept for the handling of such data [51]. Nearly three decades ago, Worboys introduced an object-oriented approach for handling simplicial complexes in a geospatial database [52]. Based on the data models of GIS pioneers such as Worboys, Le et al. developed an approach for a spatio-temporal DBMS that is based on the object-relational geospatial database PostGIS® [40–43]. This approach is capable of handling spatio-temporal data based on the snapshot model and is linked to the 3D modelling software GOCAD® [53] that is rampant in the oil industry. The geospatial database architecture DB4GeO has been realized as an object-oriented approach [44,45,54]. Both approaches focus mainly on the management of 3D models, but also manage the composition and storage of simple 4D (3D space plus 1D time) models. Recently, an approach based on integrating the concept of point tubes with net components has been developed [47]. This approach can be used for data-saving and high-performance management of spatio-temporal data based on simplicial complexes.

Finally, the management of digital twins based on CityGML within a DBMS has been examined [55,56]. With 3DCityDB® [57] a robust DBMS solution based either on Oracle or PostgreSQL is available as free and open source software [45]. Latest developments focused on the support of CityGML application domain extensions (ADEs) and the 3D web viewer 3DCityDB-Web-Map-Client [57]. For both, subsurface and city applications, 3D/4D geospatial data management is indispensable, as for planning purposes it is no more sufficient to store only the current city model in the geospatial database, but the history of the 3D city model as well [55,56]. This includes the integration of subsurface and city 3D/4D objects [47].

2.4. Milestone 4: Modelling and Visualization of Massive Geospatial Features on Web Platforms

With the increasing popularity of web-mapping applications and the rapid growth of map data availability, the pre-computation and caching of map image tiles (a.k.a. tile maps) has become a common practice in online map services as these processes use far fewer server resources than maps rendered on demand [58,59]. Online map service providers, such as Google Maps, Bing Maps, ArcGIS® Online, and OpenStreetMap, organize and deliver their content in tile maps that correspond to different areas (multi-location mapping), are displayed at various zoom levels (multi-resolution representation), and include various themes (multi-theme mapping) in different annotation languages (multi-language mapping) [59]. Obviously, rendering these products on the fly is not an option due to the cartographic limitations, such as the lack of a reliable mechanism for automated map generalization [60,61].

Google was one of the first major mapping providers to adopt map tiles [59]. Others, such as Bing and OpenStreetMap, followed the same practice. Geographic information systems (GIS) software vendors, such as Esri® and Oracle®, provide functionality for map tiling and caching of both vector layers and raster images. They also support single fused and multi-layer tiles. In the former, a group of layers is combined into a single image per tile, while in the latter, these layers appear at the client as a collection of layers with enabled feature selection and controllable visibility. Obviously, the consistent mapping of geospatial features, such as roads, rivers, lakes, forests, seashores or vessel trajectories on

all these tiles and zoom levels involves a huge investment of resources, provided that the simplification of linear features is a semi-automated process [61].

Discrete global grid systems (DGGS) offer an alternative geospatial reference framework with increasing popularity over the last few years [62]. The growth in the number of ground, airborne, and satellite sensors has led to a significant increase in the amount, variety, and rate of collection of geospatial data. However, combining huge volumes of heterogeneous geospatial data on the geographic grid is computationally expensive and time consuming [60,61]. A DGGS enables the rapid assembly of geospatial data without the difficulties of working with projected coordinate reference systems. It offers a base information multi-resolution grid capable of efficient management, storage, integration, exploration, mining, and visualization of geospatial big data.

In October 2017, OGC officially created the DGGS Abstract Specification [63], which provides a concise definition of the DGGS conceptual model and essential characteristics of a conformant DGGS. With the recent explosion in volume and variety of geospatial data, DGGSs consisting of equal area uniform grids provide a way to store and manage such data while enabling efficient integration, analysis, and visualization [64].

Another emerging method for delivering web maps is vector tiles. Vector tiles are packets of geographic data in vector format that have been clipped to the boundaries of predefined tiles, instead of pre-rendered map images. In summer 2018, OGC launched the Vector Tiles Pilot [65] in an effort to standardize vector tiles and promote interoperability. The combination of vector tiles with map (raster) tiles and DGGS data has also become an emerging need. Obviously, rendering these products on the fly is not an option due to the cartographic limitations, such as the lack of a reliable mechanism for automated map generalization [60,61]. The visualization of cartographic features at various scales (also known as multiple representation) has been an open issue in the digital era. Over the last decade, many advantages have been made in effective and meaningful visualization of mapping content in various environments and over different devices [66].

2.5. Milestone 5: Extensive Use of Geosensor Data Sources

Over the last decade, sensor technology developed rapidly, which has many advantages concerning the availability of geo-referenced data, but also includes privacy issues to be researched [67,68]. Devices have gotten smaller, cheaper, and need less energy. A sensor with a high interest for the geospatial community is the smartphone [67,69]. Being equipped with global navigation satellite system (GNSS) sensors, the location of the device is fairly well known, depending on the quality of the sensor and the signal [68,69]. Location in this context implicitly means geo-reference, but even if the phone has no GNSS sensor or the sensor is switched off, at least the phone cell delivers raw knowledge about the position of the device [69].

A GNSS track is an excellent example to explain the possibilities given by these data [70]. At first, a track represents nothing but a collection of points, each initially defined by a 3D coordinate and a time stamp; hence, a 4D data set. This set would be needless if we ignore the track as a whole because based on the consecutive points, further information can be derived, such as velocity, acceleration or azimuth, to name but a few. All these are valuable input to get information about the travel mode of citizens, e.g., walk, bicycle, etc., which is a highly interesting issue in urban planning [71]. Based on the properties of the trip (velocity, acceleration, length, etc.) in combination with sophisticated algorithms, the pattern can be evaluated to get an idea how the user moved [71]. Furthermore, singular point positioning ("navigation solution") can be used for applications where the GNSS signal is not available for tracking [69].

Almost all smart devices communicate with the net. The term Internet of Things (IoT) was coined for this [72–74], and data mining in the IoT has become a field of research on its own [75,76]. The purposes of the sensors are various. Some of them have been dedicated to security purposes such as fire, gas or humidity [77]. Others have been installed to get additional data such as private meteorological stations [78], or so-called wildlife traps have been placed in remote areas to get information about shy

or rare animals, etc. [79]. All these examples have in common that they produce a huge amount of data which have to be managed in geospatial databases, and, most of all, these data can no longer be analyzed in the traditional way. This has led immediately to approaches such as machine learning or artificial intelligence and approaches to combine them with parallel data stores such as SpatialHadoop® [80], ST-Hadoop® [81] or Hadoop-GIS® [82]. These approaches, as mentioned before, have developed to the key science for big data [83].

As a consequence of the available geosensor data, "LinkedGeoData" has been established as a concept to add the spatial reference, i.e., coordinates, to the Web of Data/Semantic Web [73]. As an example, LinkedGeoData has used data of the OpenStreetMap project [84,85] and thus has made the OpenStreetMap data available as a resource description framework (RDF) knowledge base following the linked data principles [73]. Furthermore, OpenStreetMap started to interlink this data with knowledge bases of the Linking Open Data initiative [86].

It is obvious that the collection of geosensor data has led to thousands of gigabytes in a short time period. Without geospatial data management, these large data sets would be no more manageable, especially if the replicability of the examinations is required. For big geosensor data applications, parallel data storage platforms such as SpatialHadoop® [80], ST-Hadoop® [81] or Hadoop-GIS® [57] have been used to support geospatial data analysis. This facilitates the use of artificial intelligence (AI) algorithms [81].

Following Michael Goodchild´s concept of "citizens as sensors" [87], which opened the world of "volunteered geography" and "volunteered geographic information" (VGI) [88,89], researchers analyzed large crowd source data sets such as the OpenStreetMap project [85,90]. Furthermore, online data sources of citizens generated by social media have been included in the examinations [88,89,91]. As an example, the collection of VGI from social media and other VGI sensors [88,89,91] has been identified as having enormous potential to get a detailed real-time picture of a scene after hazards such as earthquakes or tsunamis [92].

These improvements in geospatial data-collecting technology have helped non-experts to share and aggregate geospatial data in a progressively accessible way [88,89]. However, this "revolution" of data collection has also posed new questions for geospatial data management: Opposite to traditional spatial data production processes, data can be easily generated by each citizen, i.e., non-experts. This has led to badly structured data and has caused serious quality issues as observed in the OpenStreetMap project [85,90]. Hitherto, traditional geospatial database management systems are not well prepared to manage data with heterogeneous schemes for the same data type or data with dynamically changing schemes. The same is true for the management of unstructured data, i.e., data that neither have a pre-defined data model nor are organized in another pre-defined manner. Unstructured data usually are found in texts or text messages and have been analyzed e.g., in the context of hazard monitoring [91,92]. Using today´s parallel data stores such as Spatial Hadoop® [80], ST-Hadoop® [81] or Hadoop GIS® [57], the storage of these data has not been the problem. Opposite to the traditional retrieval of structured data in geospatial data management systems, the retrieval of unstructured data often has focused on the retrieval of spatial or spatio-temporal patterns or clusters. Thus, the retrieval of the data has not been realized directly in geospatial database management systems, but by statistical methods and methods of artificial intelligence such as unsupervised neural networks [91].

3. Future Directions

Based on the five milestones of geospatial data management achieved during the last decade presented in Section 2, in this section, fields of research are outlined that have the potential to become key research fields of geospatial data management in the next decade.

3.1. Geospatial Data Science

Data science has been developed as an interdisciplinary field of science that enables scientifically sound methods, processes, algorithms, and systems to extract knowledge, patterns, and conclusions

from both unstructured and structured data [93]. Thus, in future geospatial data science will have to deal with the integration of unstructured and structured data extracting geospatial knowledge, i.e., data and patterns, from unstructured data such as text messages or tweets. On the one hand, methods for the extraction of knowledge from unstructured data have to be improved concerning data quality and the comparability of different data sources. On the other hand, for structured data, the heart of data science will retain the model or the structure of the data [94]. With knowledge about the structure of the data and how it should be available, a data scientist will be able to plan a suitable system to handle the needs. Without some sort of, e.g., relational schema, the data themselves are only a string of bits and bytes without any meaning. One of the bottlenecks for the moment, the amount of data transfer between different software components, has to be eliminated in future. Cloud-based solutions are reaching their limits shown by the problems raised by the Internet of Things [73]. Edge- and fog-based solutions mentioned by Cisco® tend to solve those problems, but they are still not implemented widely [95].

It is the duty of geospatial data management to solve those problems by bridging the gap between modern information technology concepts and the geo-related sciences such as geography, geo-sciences or civil engineering to provide tools and models to ease geo-related work across all disciplines [56]. Geospatial data themselves are by nature large in quantity and extremely heterogeneous in quality. Although there are optimized and specialized solutions for sub-disciplines of geo-related sciences such as CityGML for city models [55], there is a lack of general solutions. To process geospatial data quickly and to generate added value (information) in foreseeable time, the models have to adapt to parallel and distributed network systems on the one hand, and on the other hand, they have to enrich a common communication medium. The integration of unstructured with structured geospatial data using parallel and distributed network systems, realized by ontologies, linking, and a common communication medium, is likely to become the greatest challenge geospatial data management is facing. All together, there is a need for concepts and systems consisting of related and linked ontologies to describe the semantic contents of all the intervening components such as geospatial, topological, thematic, and temporal. Finally, a communication standard has to be developed for n-dimensional data exchange between systems processing the different components.

3.2. Topology

Topology is a powerful and general concept to analyze GIS and BIM data structures and spatial relations that will be of great importance in emerging applications such as smart cities [96] and digital twins [97]. Besides semantics and geometry, topology is likely to become a key concept to support geospatial data modelling and management efficiently [98].

For example, the use of several levels of details leads to the storage of aggregation maps which are nothing but a binary relation between objects at different levels of detail [60]. Including urban infrastructure, such as water pipes, leads to relationships between buildings or land parts and infrastructure parts in a combined topological model [99]. DB4GeO allows the management of topological space-time data. All objects are stored as triangulated or polytope meshes in different dimensions at different levels of detail [45].

A spatial model consisting of points, (open) segments, (open) polygons, (open) polytopes, etc. (called objects) is topologically consistent if all objects are pairwise disjoint. This definition appears first in [100] and differs from previous definitions of topological consistency in the literature. The advantage is that it relates geometry and topology in a meaningful way.

In [100], it was observed that this notion of consistency is equivalent with specifications in the ISO 19107 standard, and that CityGML (Version 2.*) does not comply with this standard, as there are correctly modeled data sets which are not topologically consistent. This is remarkable, as CityGML claims to be based on OGC standards, cf. Figure 4.
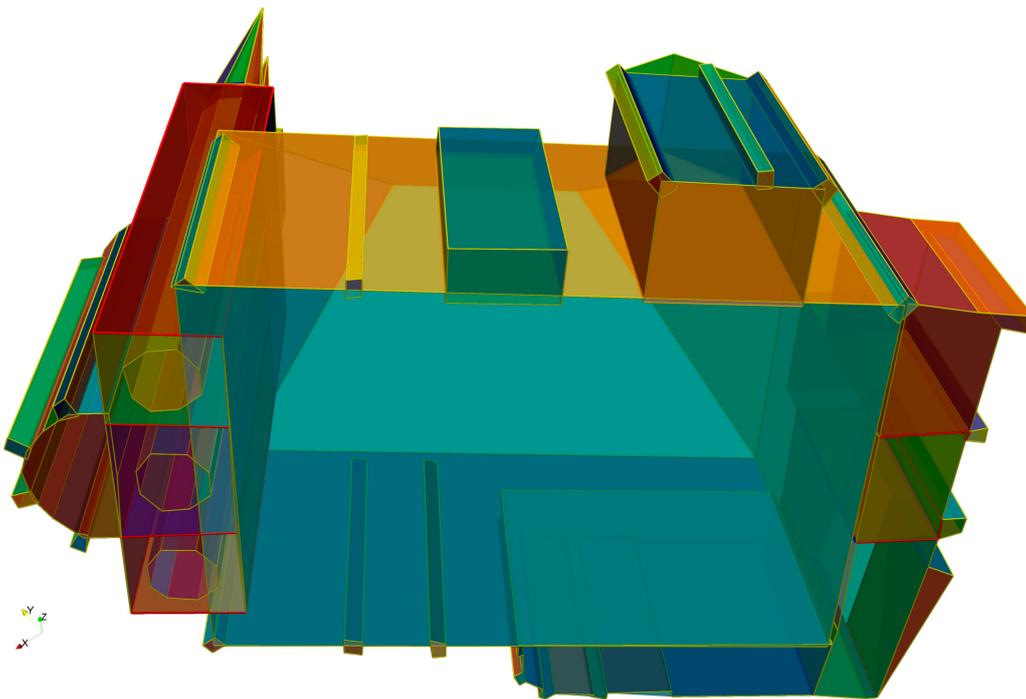
Figure 4. Example of non-consistent CityGML data.

Wireframe/Linestring-based boundary representations used in CityGML are not trivial to be triangulated in order to retrieve higher dimensional geo-objects (surface boundary representations or true volumes, e.g., tetrahedron nets). If the focus to create the model is analysis or even simulation rather than only visualization, all n-dimensional geometrical solids and their topological relations must be declared or calculated in an efficient and unequivocal way. Figure 4 shows some of the topological inconsistencies. The multi-surface of the bottom plane shows intersections which also lead to penetrations within the defined multi-volume. Recalculating all topological inconsistencies by finding all n-dimensional intersections in runtime is a complex task which should be avoided. The model is also not unequivocal due to different ways of triangulating Wireframe/Linestring-based boundary representations to create surfaces for a surface boundary representation and different ways of triangulating surface-based boundary representations to create the final volumes as a tetrahedron net. The problem gets more complex by adding further dimensions such as time coordinates or any other thematic/semantic coordinates.

Topological consistency is a requirement for analysis and simulation of geospatial data models, as cumbersome geometric operations such as intersections can be avoided, and the simulation can be run on the topological model itself. In order to obtain a topologically consistent model, an overlay method can be applied, which breaks the overlapping objects into disjoint atomic objects and retains the boundary relationships. A precise mathematical definition of this is deferred to a future publication.

The topology model has a minimal representation in the form of the so-called Hasse diagram, a graph in which only direct boundary relationships are explicitly modeled. This can be used as the core model for analyzing flows or running simulations. This means that in future work, partial differential equations need to be solved numerically on graphs and the solutions compared with corresponding solutions for the underlying 3D model. The mathematical theory for this approach is still quite recent, cf. e.g., [101] and the references therein, but lacks a thorough numerical treatment. Its potential lies in the applicability of the topology of geospatial data stored in NoSQL databases and possible parallelization of numerical algorithms.

The tasks of parallelisation and streaming of big topological data sets are the consequent next steps for the future, as by now the means for this are at hand.

3.3. Bridging Geospatial Data Management with Data Streaming Libraries and "In-Situ" Geo-Computing

In the future, different kinds of network architectures which involve edge- and fog-based solutions [95] will have to be tested to connect online transaction processing (OLTP) and online analysis processing (OLAP) systems [102,103] in order to reduce data transfer and to establish efficient geo-computing solutions. Recent programming paradigms support functional-style operations on streams of elements such as map-reduced transformations of collections, e.g., the java streaming api. Data streaming is not only a technology to move data from *A* to *B*, but it allows us to filter and analyze the data as well. While some functional-style operations work within one OLAP-system to retrieve certain added value, it is also possible to use streaming libraries across a largely distributed network to set up some real-time streaming pipeline or application, e.g., apache kafka®. Examples for geospatial streaming data are point streams that represent continuously moving points, but also line and region streams representing complex spatio-temporal phenomena derived from geo-referenced data [104].

Sensor networks, geospatial data warehouses or high-performance computing clusters for geospatial data may differ in geospatial data modelling and network architectures. Each network/cluster is optimized to handle certain collection, simulation, processing, analysis, and presentation tasks and depends strongly on its geospatial data model. In the future, the modelling steps from sensor networks to high-performance computing could be managed through a real-time streaming pipeline or application. If some of those networks or clusters are able to translate a geospatial dataset into another geospatial data model in parallel or even shares the same geospatial data model, then "in-situ"-approaches could become feasible and tested. "In-situ" geo-computing of objects executed directly on the sensors will revolutionize geo-information science and bridge geo-computing closer to geospatial data management.

3.4. Geospatial Data Visualization on Web Platforms

The representation of geospatial features (e.g., roads, rivers), outlines of areas (e.g., municipal boundaries, lake banks) or moving objects trajectories (e.g., of humans, vehicles) on paper or digital maps comprises a massive number of vertices. To facilitate the processing, analysis or mapping of these geometries at a small scale, beyond semi-automated cartographic methods, new developments in Geospatial Web and Digital Earth have to be introduced to model and map voluminous geospatial data. To deliver the content of map services in a standardized tile format [64] for high-speed dissemination of voluminous data over the Web and the adaptation of discrete global grid systems (DGGS) to understand the planet model by offering an analysis-ready-information-grid [63,65] are only the first steps in this way of development. The rapidly growing use of these reference frameworks (i.e., tile maps and DGGS) has urged the development of new approaches to an efficient, consistent, and compliant mapping of massive polyline geometries representing geographic features or trajectories.

Future research will aim to introduce new approaches to the fundamental problem of multiple representation in geospatial data handling. These approaches will offer sophisticated modelling and visualization of massive linear geospatial features in modern geospatial reference frameworks. Industry and government map service providers will be able to generate massive map products in a faster and more accurate mode.

Finally, geospatial data management and data visualization have to grow together to enable more efficient solutions for geospatial data analysis. This means that the visualization of large amounts of data should be executed by visualization clusters. A good presentation often needs some analysis to filter information which will be visualized. Those filters in extremely large datasets need to work on distributed memory computing resources. The visualization of simulation results or field data in real-time analysis of extremely large datasets is still challenging. The amount of data to be transferred to the visualization cluster from data-warehouses and/or a sensor networks is as large as the extremely large datasets which should be visualized. It is in question if a visualization cluster can be merged with data warehouses or even sensor networks. One approach of merging simulation clusters and visualization clusters at a large scale is implemented by Kitware within ParaView® and its

"in-situ"-technology called "catalyst$^{®}$" [105]. The visualization cluster is able to calculate simulations if the source code of the simulation has been provided. After a simulation step has been calculated, the presentation calculations can be done in memory without moving the data to a different cluster. With this approach, it is possible to visualize each simulation result at simulation/presentation-calculation time. Certainly, new research will be based upon this research and will adopt it for new applications.

3.5. Database Support for Big Geospatial Data Analysis

Obviously, big geospatial data analysis will be increasingly supported by artificial intelligence (AI) for geospatial data [106], also called geoAI [107], and the retrieval of the data will be facilitated by parallel geospatial database architectures. GeoAI will closely combine geo-information science, AI methods in machine learning, data mining, and parallel computing to extract knowledge from big geospatial data [107]. To give an example, in future, geoAI methods will be used for geospatial data cleansing, i.e., to detect and learn errors in geospatial data sets. These methods will complement the existing data consistency checking mechanisms of geospatial database management systems. Parallel geospatial database management systems will serve as a "preliminary stage" to pre-select geospatial and spatio-temporal data before AI-supported data analysis will take place. Thus, geospatial analysis tools will be supported by geospatial data management systems that split the data into geospatial or spatio-temporal partitions to organize parallel database accesses.

Many approaches for geospatial data management have been implemented on top of object-relational [55,108] or on object-oriented databases [45,109]. However, graph databases promise to be well suited for this task, as has been first proven by [110]. For future research, we expect that graph databases will experience a revival in the context of distributed and parallel geospatial data analysis. One of the major topics within this research will be how to integrate known geospatial-, spatio-temporal- or *nD*-access methods into the property graph system.

A further research topic will be how to use graph algorithms to quickly solve standard geo-tasks known from standard GISs on the feature graph.

Increasingly, software and methodologies will be imported primarily from intelligent structures [107] for the aim of developing shrewd GIServices. It is the ability to understand, plan, analyze, and behave in an innovative way that is crucial to intelligent structures. Intelligent geospatial services will serve as a kind of intelligent structure from the AI point of view [111–113]. Intelligent geospatial services will be intelligently aware, net-connected, and regulated by IoT software use [114]. The service-oriented architecture (SOA) strategy with relation to the IoT [73] will have to be efficiently implemented, which allows support sharing and recycling through services-oriented characteristics [115]. The creation of smooth ontology will help smart geospatial services to resolve confusion [116]. One of the key cutting-edge gaps in geospatial data transformation is an "intelligent" mechanism that facilitates the discovery of facts and integration in the Internet [117]. Ontologies and geospatial metadata will play a central role in this process by linking data from different data sets. Thus, the dissemination of such mechanisms will lead to a better understanding and integration of heterogeneous geospatial data sources.

4. Conclusions

This overview article has made two contributions: to highlight research in geospatial data management during the last decade and to formulate five future directions of research that have the potential to become key research fields of geospatial data management research in the next decade.

Five milestones of research in the last decade have been identified that range from application to theory. On the application side, especially the progress in interdisciplinary research has been emphasized, such as the growing together of the BIM and the GIS geospatial data management. This has opened hitherto closed doors on the data, process, and application side on the way to integrated BIM/GIS models. This three-level approach also considers for the first time the lifecycle of buildings and cities. Looking at the other (theoretical) spectrum of the milestones, the theory of geospatial data

modelling and management has been improved by considering topology as a sound mathematical framework which is not limited to topological relationships as formulated in the 4- or 9-intersection model. Thus, topology is seen as a key concept of geospatial data management to describe and analyze data structures and algorithms presuming that geometry and semantics of objects can be added to topology. As basis for BIM and GIS applications, 3D/4D geospatial data management has been identified as a key concept to support the geometric, topological, thematic, and temporal modeling of buildings and cities and subsurface models. 3D/4D geospatial data management also has started to serve as an integration platform for different geospatial and spatio-temporal data representations. As geospatial data management is closely connected to and dependent on data modelling and data visualization, it is important that in the last decade, progress in modelling and visualization of massive geospatial features on web platforms has been made. In particular, the application of discrete global grid systems as an alternative geospatial reference framework and the development of visualization clusters on top of parallel data stores have provided significant research progress for the geospatial community. Finally, the extensive use of geo-sensor data sources has been identified as a milestone in research, not least because Michael Goodchild´s paradigm of "citizens as sensors" has revolutionized the handling of geospatial data, leading to new requirements including geospatial data management for unstructured and structured data to be solved in the next decade.

Geospatial data science is expected to become a key research field for the next decade to solve the integration of unstructured and structured data as well as the extraction of geospatial knowledge, i.e., data and patterns, in interdisciplinary applications. Furthermore, geospatial data science will bridge the gap between modern information technology concepts and the geo-related sciences. In this context, topology will continue to be a powerful and general concept to analyze GIS and BIM data structures and spatial relations that will be of great importance in emerging applications such as smart cities and digital twins. Streaming of big geospatial data and "in-situ" geo-computing will play a central role in the next decade. Thus, the bridging of geospatial data management with data streaming libraries and "in-situ" geo-computing on objects executed directly on the sensors has been identified as a future direction of research. This will revolutionize geo-information science and bridge geo-computing with geospatial data management. In future, visualization will be more integrated into web applications. Thus, the development of advanced geospatial data visualization on web platforms has been presented as a future direction of research. This will enable the representation of dynamically changing geospatial features or moving objects trajectories directly loaded from geospatial databases. Finally, geospatial data management will increasingly support big geospatial data analysis. This future direction of research will include the management of unstructured and structured geospatial data as well as the extraction of knowledge and patterns in geospatial data. Graph databases are expected to experience a revival on top of parallel and distributed data stores supporting big geospatial data analysis. Finally, raster data gained from satellites will be managed efficiently in geospatial data management systems. This includes the realization of seamless geospatial and temporal access to satellite data.

References

1. Tan, Z.; Yue, P.; Gong, J. An Array Database Approach for Earth Observation Data Management and Processing. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 220. [CrossRef]

2. Kim, H.-S.; Sun, C.-G.; Cho, H.-I. Geospatial Big Data-Based Geostatistical Zonation of Seismic Site Effects in Seoul Metropolitan Area. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 174. [CrossRef]

3. Arroyo Ohori, K.; Diakité, A.; Krijnen, T.; Ledoux, H.; Stoter, J. Processing BIM and GIS Models in Practice: Experiences and Recommendations from a GeoBIM Project in The Netherlands. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 311. [CrossRef]

4. Hor, A.-H.; Jadidi, A.; Sohn, G. Bim-Gis Integrated Geospatial Information Model Using Semantic Web and RDF Graphs. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 73–79. [CrossRef]

5. Wu, C.; Zhu, Q.; Zhang, Y.; Du, Z.; Ye, X.; Qin, H.; Zhou, Y. A NoSQL–SQL Hybrid Organization and Management Approach for Real-Time Geospatial Data: A Case Study of Public Security Video Surveillance. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 21. [CrossRef]

6. Ge, M.; Dohnal, V. Quality Management in Big Data. *Informatics* **2018**, *5*, 19. [CrossRef]

7. Hu, F.; Xu, M.; Yang, J.; Liang, Y.; Cui, K.; Little, M.M.; Lynnes, C.S.; Duffy, D.Q.; Yang, C. Evaluating the Open Source Data Containers for Handling Big Geospatial Raster Data. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 144. [CrossRef]

8. Zhu, J.; Wright, G.; Wang, J.; Wang, X. A Critical Review of the Integration of Geographic Information System and Building Information Modelling at the Data Level. *IJGI* **2018**, *7*, 66. [CrossRef]

9. Sacks, R.; Ma, L.; Yosef, R.; Borrmann, A.; Daum, S.; Kattel, U. Semantic Enrichment for Building Information Modeling: Procedure for Compiling Inference Rules and Operators for Complex Geometry. *J. Comput. Civ. Eng.* **2017**, *31*, 04017062. [CrossRef]

10. Borrmann, A.; König, M.; Koch, C.; Beetz, J. (Eds.) *Building Information Modeling*; Springer International Publishing: Berlin, Germany, 2015.

11. Borrmann, A. From GIS to BIM and back again—A spatial query language for 3D building models and 3D city models. In Proceedings of the 5th International 3D Geoinfo Conference, Berlin, Germany, 3–4 November 2010.

12. Reshetyuk, Y. *Terrestrial Laser Scanning: Error Sources, Self Calibration and Direct Georeferencing*; VDM: Stockholm, Sweden, 2009.

13. El-Mekawy, M.; Östman, A.; Hijazi, I. A Unified Building Model for 3D Urban GIS. *IJGI* **2012**, *1*, 120–145. [CrossRef]

14. Dave, B.; Buda, A.; Nurminen, A.; Främling, K. A framework for integrating BIM and IoT through open standards. *Autom. Constr.* **2018**, *95*, 35–45. [CrossRef]

15. Hor, A.-H.; Sohn, G.; Claudio, P.; Jadidi, M.; Afnan, A. A Semantic Graph Database For Bim-Gis Integrated Information Model for an Intelligent Urban Mobility Web Application. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *4*, 89–96. [CrossRef]

16. Irizarry, J.; Karan, E.P.; Jalaei, F. Integrating BIM and GIS to improve the visual monitoring of construction supply chain management. *Autom. Constr.* **2013**, *31*, 241–254. [CrossRef]

17. Amirebrahimi, S.; Rajabifard, A.; Mendis, P.; Ngo, T. A BIM-GIS integration method in support of the assessment and 3D visualisation of flood damage to a building. *J. Spat. Sci.* **2016**, *61*, 317–350. [CrossRef]

18. Kang, T.W.; Hong, C.H. A study on software architecture for effective BIM/GIS-based facility management data integration. *Autom. Constr.* **2015**, *54*, 25–38. [CrossRef]

19. Fosu, R.; Suprabhas, K.; Rathore, Z.; Cory, C. Integration of Building Information Modeling (BIM) and Geographic Information Systems (GIS)—A Literature Review and Future Needs. In Proceedings of the 32nd CIB W78 Conference 2015, Eindhoven, The Netherlands, 27–29 October 2015; p. 9.

20. Geiger, A.; Benner, J.; Haefele, K.-H. Generalization of 3D IFC Building Models. In *3D Geoinformation Science*; Springer: Cham, Switzerland, 2015; pp. 19–35.

21. Alexandrov, P.S. Diskrete Räume. *Matematicheskii Sbornik* **1937**, 501–518.

22. Egenhofer, M.J. Reasoning about binary topological relations. In *Advances in Spatial Databases*; Günther, O., Schek, H.-J., Eds.; Springer: Berlin/Heidelberg, Germany, 1991; Volume 525, pp. 141–160, ISBN 978-3-540-54414-2.

23. Winter, S.; Frank, A.U. Topology in Raster and Vector Representation. *GeoInformatica* **2000**, *4*, 35–65. [CrossRef]

24. Ozel, F. Spatial databases and the analysis of dynamic processes in buildings. In Proceedings of the Fifth Conference on Computer Aided Architectural Design Research in Asia, Singapore, 18–19 May 2000; Volume 2, pp. 97–106.

25. Kohler, N.; Lockemann, P. *Arbeits- und Entwicklungsbericht: Planungsplattform für Dynamische Gebäude*; Institut für industrielle Bauproduktion (IFIB): Karlsruhe, Germany, 2003.

26. Boguslawski, P.; Mahdjoubi, L.; Zverovich, V.; Fadli, F.; Barki, H. BIM-GIS Modelling in Support of Emergency Response Applications. *WIT Trans. Built Environ.* **2015**, *149*, 381–391.

27. Liu, L.; Li, B.; Zlatanova, S.; Liu, H. The Path from Bim to a 3d Indoor Framework—A Requirement Analysis. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII–4*, 373–378.

28. Breunig, M. *On the Way to Component-Based 3D/4D Geoinformation Systems*; Lecture Notes in Earth Sciences; Springer International Publishing: Berlin, Germany, 2001; Volume 94.

29. Bradley, P.E.; Paul, N. Using the Relational Model to Capture Topological Information of Spaces. *Comput. J.* **2010**, *53*, 69–89. [CrossRef]

30. Lienhardt, P. N-Dimensional Generalized Combinatorial Maps and Cellular Quasi-Manifolds. *Int. J. Comput. Geom. Appl.* **1994**, *4*, 275–324. [CrossRef]

31. Bradley, P.E.; Paul, N. Comparing G-maps with other topological data structures. *Geoinformatica* **2014**, *18*, 595–620. [CrossRef]

32. Paul, N. Topologische Datenbanken für Architektonische Räume. Ph.D. Thesis, Universität Karlsruhe, Karlsruhe, Germany, 2008.

33. Vitalis, S.; Ohori, K.; Stoter, J. Incorporating Topological Representation in 3D City Models. *IJGI* **2019**, *8*, 347. [CrossRef]

34. Sagan, H. *Space Filling Curves*; Springer: New York, NY, USA, 1994. [CrossRef]

35. Guan, X.; van Oosterom, P.; Cheng, B. A Parallel N-Dimensional Space-Filling Curve Library and Its Application in Massive Point Cloud Management. *IJGI* **2018**, *7*, 327. [CrossRef]

36. Wang, J.; Shan, J. Space filling curve based point clouds index. In Proceedings of the 8th International Conference on GeoComputation, Ann Arbor, MI, USA, 31 July–3 August 2005; pp. 551–562.

37. Van Oosterom, P.; Martinez-Rubi, O.; Ivanova, M.; Horhammer, M.; Geringer, D.; Ravada, S.; Tijssen, T.; Kodde, M.; Gonçalves, R. Massive point cloud data management: Design, implementation and execution of a point cloud benchmark. *Comput. Graph.* **2015**, *49*, 92–125. [CrossRef]

38. Sirdeshmukh, N.; Verbree, E.; Oosterom, P.V.; Psomadaki, S.; Kodde, M. Utilizing a Discrete Global Grid System for Handling Point Clouds with Varying Locations, Times, and Levels of Detail. *Cartographica: Int. J. Geogr. Inf. Geovis.* **2019**, *54*, 4–15. [CrossRef]

39. Langran, G.; Stuart, N. Time in Geographic Information Systems. *Int. J. Geograph. Inf. Syst.* **1992**, *6*, 547. [CrossRef]

40. Le, H.H.; Gabriel, P.; Gietzel, J.; Schaeben, H. An object-relational spatio-temporal geoscience data model. *Comput. Geosci.* **2013**, *57*, 104–115. [CrossRef]

41. Le, H.H.; Schaeben, H.; Jasper, H.; Görz, I. Database Versioning and its Implementation in Geoscience Information Systems. *Comput. Geosci.* **2014**, *70*, 44–54. [CrossRef]

42. Le, H.H. Spatio-temporal Information System for the Geosciences: Concepts, Data models, Software, and Applications. Ph.D. Thesis, Technische Universität Bergakademie Freiberg, Freiberg, Germany, 2014.

43. Gabriel, P.; Gietzel, J.; Le, H.H.; Schaeben, H. GST: A Network Based Datastore for Geoscience Data and Geomodels and its Implementation—ProMine's Contribution Towards Interoperability. In *3D, 4D and Predictive Modelling of Major Mineral Belts in Europe*; Weihed, P., Ed.; Springer International Publishing: Cham, Switzerland, 2015; pp. 59–71, ISBN 978-3-319-17427-3.

44. Breunig, M.; Schilberg, B.; Thomsen, A.; Kuper, P.V.; Jahn, M.; Butwilowski, E. DB4GeO, a 3D/4D Geodatabase and Its Application for the Analysis of Landslides. In *Geographic Information and Cartography for Risk and Crisis Management*; Konecny, M., Zlatanova, S., Bandrova, T.L., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 83–101, ISBN 978-3-642-03441-1.

45. Breunig, M.; Kuper, P.V.; Butwilowski, E.; Thomsen, A.; Jahn, M.; Dittrich, A.; Al-Doori, M.; Golovko, D.; Menninghaus, M. The story of DB4GeO—A service-based geo-database architecture to support multi-dimensional data analysis and visualization. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 187–205. [CrossRef]

46. Hopf, H.; Alexandroff, P. *Topologie*; Springer International Publishing: Berlin, Germany, 1935; p. 158.

47. Kuper, P.V. Efficient and Practical Handling of Spatio-Temporal Data Based on Time-Dependent Net Components. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII–4*, 321–327. [CrossRef]

48. Dutton, G.H. *A Hierarchical Coordinate System for Geoprocessing and Cartography, Lecture Notes in Earth Sciences*; Springer: New York, NY, USA, 1999.

49. Peuquet, D.J. A Conceptual Framework and Comparison of Spatial Data Models. *Cartogr. Int. J. Geogr. Inf. Geovis.* **1984**, *21*, 66–113. [CrossRef]

50. Mallet, J.L. *Geomodeling*; Oxford University Press: New York, NY, USA, 2002.

51. Langran, G.; Chrisman, N.R. A Framework for Temporal Geographic Information. *Cartographica* **1988**, *25*, 1–14. [CrossRef]

52. Worboys, M. A generic model for planar geographical objects. *Int. J. Geogr. Inf. Syst.* **1992**, *6*, 353–372. [CrossRef]

53. Mallet, J.L. GOCAD: A computer aided design program for geological applications. In *Three-Dimensional Modeling with Geoscientific Information Systems*; Turner, A.K., Ed.; NATO ASI; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1992; Volume 354, pp. 123–142.

54. Thomsen, A.; Breunig, M.; Butwilowski, E. Towards the Unified Modelling and Management of Topology in Multiple Representation Databases. *Schweizerbart'sche Verlagsbuchhandlung* **2008**, *3*, 176–185.

55. Yao, Z.; Nagel, C.; Kunde, F.; Hudra, G.; Willkomm, P.; Donaubauer, A.; Adolphi, T.; Kolbe, T.H. 3DCityDB—A 3D geodatabase solution for the management, analysis, and visualization of semantic 3D city models based on CityGML. *Open Geospat. Data Softw. Stand.* **2018**, *3*, 5. [CrossRef]

56. Yao, Z.; Kolbe, T.H. A new Approach to Model Transformation using Graph Transformation System. In *PFGK18-Photogrammetrie-Fernerkundung-Geoinformatik-Kartographie, 37. Jahrestagung in München 2018*; Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.: Munich, Germany, 2018; pp. 831–834.

57. 3dcitydb. Available online: https://github.com/3dcitydb/3dcitydb (accessed on 12 November 2019).

58. Peterson, M.P. (Ed.) *Online Maps with APIs and WebServices*; Lecture Notes in Geoinformation and Cartography; Springer: Berlin/Heidelberg, Germany, 2012; ISBN 978-3-642-27484-8.

59. Stefanakis, E. Web mercator and raster tile maps: Two cornerstones of online map service providers. *Geomatica* **2017**, *71*, 100–109. [CrossRef]

60. Keates, J.S. *Cartographic Design and Production*, 2nd ed.; Longman Scientific & Technical: Harlow, Essex, UK; Wiley: New York, NY, USA, 1989; ISBN 978-0-582-30133-7.

61. Robinson, A.H.; Morrison, J.L.; Muehrcke, P.C.; Kimerling, A.J.; Guptill, S.C. *Elements of cartography*, 6th ed.; John Wiley & Sons: New York, NY, USA, 1995; ISBN 978-0-471-55579-7.

62. Goodchild, M.F. Reimagining the history of GIS. *Ann. GIS* **2018**, *24*, 1–8. [CrossRef]

63. Purss, M. Topic 21: Discrete Global Grid Systems Abstract Specification. Available online: http://docs.opengeospatial.org/as/15-104r5/15-104r5.html (accessed on 12 November 2019).

64. OGC. OGC Announces a New Standard That Improves the Way Information Is Referenced to the Earth. Available online: https://www.opengeospatial.org/pressroom/pressreleases/2656 (accessed on 12 November 2019).

65. OGC. Vector Tiles Pilot. Available online: https://www.opengeospatial.org/projects/initiatives/vt-pilot-2018 (accessed on 13 November 2019).

66. ICA Commission on Generalisation and Multiple Representation. Available online: https://generalisation.icaci.org/ (accessed on 30 December 2019).

67. Annapurna, S.; Teja, P.K.V.S.; Murty, S. A Comparative Study on Mobile Platforms (Android vs. IOS). *Int. J. Adv. Res. Comput. Eng. Technol. IJARCET* **2016**, *5*, 547–553.

68. Nittel, S. A Survey of Geosensor Networks: Advances in Dynamic Environmental Monitoring. *Sensors* **2009**, *9*, 5664–5678. [CrossRef]

69. Huber, D. *Background Positioning for Mobile Devices—Android vs. iPhone*; Technical Report; School of Electrical and Computer Engineering: Berlin, Germany, 2011; p. 7.

70. Söderholm, S.; Bhuiyan, M.Z.H.; Thombre, S.; Ruotsalainen, L.; Kuusniemi, H. A Multi-GNSS Software-Defined Receiver: Design, Implementation, and Performance Benefits. *Ann. Telecommun.* **2016**, *71*, 399–410. [CrossRef]

71. Sauerländer-Biebl, A.; Brockfeld, E.; Suske, D.; Melde, E. Evaluation of a transport mode detection using fuzzy rules. *Transp. Res. Procedia* **2017**, *25*, 591–602. [CrossRef]

72. Li, S.; Xu, L.D.; Zhao, S. The internet of things: A survey. *Inf. Syst. Front.* **2015**, *17*, 243–259. [CrossRef]

73. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805. [CrossRef]

74. Jara, A.J.; Olivieri, A.C.; Bocchi, Y.; Jung, M.; Kastner, W.; Skarmeta, A.F. Semantic Web of Things: An analysis of the application semantics for the IoT moving towar the IoT convergence. *IJWGS* **2014**, *10*, 244–272. [CrossRef]

75. Chen, F.; Deng, P.; Wan, J. Data Mining for the Internet of Things: Literature Review and Challenges. *Int. J. Distrib. Sens. Netw.* **2015**. [CrossRef]

76. Tao, F.; Cheng, Y.; Da Xu, L.; Zhang, L.; Li, B.H. CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System. *IEEE Trans. Ind. Inf.* **2014**, *10*, 1435–1442.

77. Bahrepour, M.; Meratnia, N.; Havinga, P. *Automatic Fire Detection: A Survey from Wireless Network Perspective*; Technical Report TR-CTIT-08-73; Centre for Telematics and Information Technology, University of Twente: Enschede, The Netherlands, 2008; ISSN 1381-362513p.

78. Muller, C.; Chapman, L.; Grimmond, C.S.B.; Young, D.; Cai, X. Sensors and the City: A Review of Urban Meteorological Networks. *Int. J. Climatol.* **2013**, *33*, 1585–1600. [CrossRef]

79. Moeller, A.K.; Lukacs, P.M.; Horne, J.S. Three Novel Methods to Estimate Abundance of Unmarked Animals Using Remote Cameras. *Ecosphere* **2018**, *9*, e02331. [CrossRef]

80. Eldawy, A.; Mokbel, M.F. SpatialHadoop: A mapreduce framework for spatial data. In Proceedings of the 2015 IEEE 31st International Conference on Data Engineering, Seoul, Korea, 13–17 April 2015.

81. Alarabi, L.; Mokbel, M.F.; Musleh, M. ST-Hadoop: A mapreduce framework for spatio-temporal data. *Geoinformatica* **2018**, *22*, 785–813. [CrossRef]

82. Aji, A.; Wang, F.; Vo, H.; Lee, R.; Liu, Q.; Zhang, X.; Saltz, J. Hadoop GIS: A high performance spatial data warehousing system over mapreduce. *Proc. VLDB Endow.* **2013**, *6*, 1009–1020. [CrossRef]

83. Günther, W.A.; Mehrizi, M.H.R.; Huysman, M.; Feldberg, F. Debating Big Data: A Literature Review on Realizing Value From Big Data. *J. Strateg. Inf. Syst.* **2017**, *26*, 191–209. [CrossRef]

84. Raifer, M.; Troilo, R.; Kowatsch, F.; Auer, M.; Loos, L.; Marx, S.; Przybill, K.; Fendrich, S.; Mocnik, F.B.; Zipf, A. OSHDB: A Framework for Spatio-Temporal Analysis of Openstreetmap History Data. *Open Geospat. Data Softw. Stand.* **2019**, *4*, 3. [CrossRef]

85. Arsanjani, J.J.; Zipf, A.; Mooney, P.; Helbich, M. (Eds.) *OpenStreetMap in GIScience—Experiences, Research, and Applications, Lecture Notes in Geoinformation and Cartography*; Springer: Berlin/Heidelberg, Germany, 2015; p. 325.

86. Berners-Lee, T. Linked Open Data. 2006. Available online: https://www.w3.org/DesignIssues/LinkedData.html (accessed on 3 February 2020).

87. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]

88. Elwood, S. Volunteered geographic information: Key questions, concepts and methods to guide emerging research and practice. *GeoJournal* **2008**, *72*, 133–135. [CrossRef]

89. Elwood, S. Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal* **2008**, *72*, 173–183. [CrossRef]

90. Chilton, S. CrowdSourcing is radically changing the geodata landscape: Case study of OpenStreetMap. In Proceedings of the 24th International Cartographic Conference, Santiago, Chile, 15–21 November 2009.

91. Steiger, E.; Resch, B.; Zipf, A. Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1694–1716. [CrossRef]

92. Goodchild, M.F.; Glennon, J.A. Crowdsourcing Geographic Information for Disaster Response: A Research Frontier. *Int. J. Digit. Earth* **2010**, *3*, 231–241. [CrossRef]
93. Dhar, V. Data science and prediction. *Commun. ACM* **2013**, *56*, 64–73. [CrossRef]
94. Agarwal, R.; Dhar, V. Big data, data science, and analytics: The opportunity and challenge for IS research. *IS Res. J.* **2014**, *25*, 443–448. [CrossRef]
95. Omoniwa, B.; Hussain, R.; Javed, M.A.; Bouk, S.H. Fog/Edge Computing-based IoT (FECiOT): Architecture, Applications, and Research Issues. *IEEE Internet Things J.* **2018**. [CrossRef]
96. Meijer, A.; Bolivar, M.P.R. Governing the Smart City: A Review of the Literature on Smart Urban Governance. *Int. Rev. Adm. Sci.* **2016**, *82*, 392–408. [CrossRef]
97. Enders, M.R.; Hoßbach, N. Dimensions of Digital Twin Applications—A Literature Review. In Proceedings of the 25th Americas Conference on Information Systems, Cancún, Mexico, 15–17 August 2019; p. 10.
98. Bradley, P.E.; Paul, N. Topologically consistent space, time, version, and scale using Alexandrov topologies. In *Contemporary Strategies and Approaches in 3-D Information Modeling*; IGI Global: Pershey (PA), USA, 2018; pp. 52–82.
99. Ji, Q.; Barr, S.; James, P.; Fairbairn, D. A Geospatial Analysis Framework for Fine Scale Urban Infrastructure Networks. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII–4*, 291–296. [CrossRef]
100. Giovanella, A.; Bradley, P.E.; Wursthorn, S. Evaluation of Topological Consistency in CityGML. *IJGI* **2019**, *8*, 278. [CrossRef]
101. Arioli, M.; Benzi, M. A finite element method for quantum graphs. *IMA J. Numer. Anal.* **2018**, *38*, 1119–1163. [CrossRef]
102. Gelinas, U.J.; Sutton, S.G.; Fedorowicz, J. Online Transaction Processing. In *Open Textbooks for Hong Kong*; Open University of Hong Kong: Hong Kong, China, 2019.
103. Conn, S.S. OLTP and OLAP data integration: A Review of Feasible Implementation Methods and Architectures for Real Time Data Analysis. In Proceedings of the IEEE Southeast Conference, Ft. Lauderdale, FL, USA, 8–10 April 2005; p. 6. [CrossRef]
104. Zhang, C.; Huang, Y.; Griffin, T. Querying Geospatial Data Streams in SECONDO. In Proceedings of the 17th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems (ACM-GIS 2009), Seattle, WA, USA, 4–6 November 2009. [CrossRef]
105. Bauer, A.; Wissink, A.; Potsdam, M.; Jayaraman, B. ParaView Catalyst Computes Particle Paths In Situ. *Kitware Blog* **2016**, *39*.
106. Karimi, H.A. (Ed.) *BIG DATA: Techniques and Technologies in Geoinformatics*; CRC Press: Boca Raton, FL, USA, 2014; 312p, ISBN 978-1-138-07319-7.
107. VoPham, T.; Hart, J.E.; Laden, F.; Chiang, Y.-Y. Emerging Trends in Geospatial Artificial Intelligence (geoAI): Potential Applications for Environmental Epidemiology. *Environ. Health* **2018**, *17*, 40. [CrossRef]
108. Breunig, M.; Zlatanova, S. 3D geo-database research: Retrospective and future directions. *Comput. Geosci.* **2011**, *37*, 791–803. [CrossRef]
109. Balovnev, O.; Breunig, M.; Cremers, A.B. Geotoolkit: Opening Access to Object-Oriented Geodata Stores. In *Interoperating Geographic Information Systems*; Goodchild, M., Egenhofer, M., Fegeas, R., Kottman, C., Eds.; Springer: Boston, MA, USA, 1999; pp. 235–247, ISBN 978-1-4613-7363-6.
110. Güting, R.H. GraphDB: Modeling and Querying Graphs in Databases. In Proceedings of the 20th International Conference on Very Large Data Bases, Santiago de Chile, Chile, 12–25 September 1994; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1994; pp. 297–308.
111. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Prentice Hall Series in Artificial Intelligence; Prentice Hall: Englewood Cliffs, NJ, USA, 1995; ISBN 978-0-13-103805-9.
112. Baader, F.; Nutt, W. Basic description logics. In *The Description Logic Handbook. Theory, Implementation and Applications*; Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., Eds.; Cambridge University Press: Cambridge, UK, 2003; pp. 43–95.
113. Rudas, I.; Fodor, J. Intelligent systems. *Int. J. Comput. Commun. Control* **2008**, *3*, 132–138.
114. Li, L.; Li, S.; Zhao, S. QoS-Aware Scheduling of Services-Oriented Internet of Things. *IEEE Trans. Ind. Inf.* **2014**, *10*, 1497–1505.
115. Viriyasitavat, W.; Xu, L.D. Compliance Checking for Requirement-Oriented Service Workflow Interoperations. *IEEE Trans. Ind. Inf.* **2014**, *10*, 1469–1477. [CrossRef]

116. He, L.; Yue, P.; Jiang, L.; Zhang, M. Fuzzy spatial relation ontology driven detection of complex geospatial features in a web service environment. *Earth Sci. Inf.* **2015**, *8*, 63–76. [CrossRef]

117. Yuan, J.; Yue, P.; Gong, J.; Zhang, M. A Linked Data Approach for Geospatial Data Provenance. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 5105–5112. [CrossRef]