International Journal of Geo-Information

MDPI

Article

Point-of-Interest (POI) Data Validation Methods: An Urban Case Study

Lih Wei Yeow [1,2,†], Raymond Low [1,†], Yu Xiang Tan [3] and Lynette Cheah [1,*]

[1] Engineering Systems and Design, Singapore University of Technology and Design, Singapore 487372, Singapore; lihwei.yeow@mail.utoronto.ca (L.W.Y.); raymond_low@mymail.sutd.edu.sg (R.L.)
[2] Department of Civil & Mineral Engineering, University of Toronto, Toronto, ON M5S 1A4, Canada
[3] Information Systems Technology and Design, Singapore University of Technology and Design, Singapore 487372, Singapore; yuxiang_tan@mymail.sutd.edu.sg
* Correspondence: lynette@sutd.edu.sg
† These authors contributed equally to this work.

Abstract: Point-of-interest (POI) data from map sources are increasingly used in a wide range of applications, including real estate, land use, and transport planning. However, uncertainties in data quality arise from the fact that some of this data are crowdsourced and proprietary validation workflows lack transparency. Comparing data quality between POI sources without standardized validation metrics is a challenge. This study reviews and implements the available POI validation methods, working towards identifying a set of metrics that is applicable across datasets. Twenty-three validation methods were found and categorized. Most methods evaluated positional accuracy, while logical consistency and usability were the least represented. A subset of nine methods was implemented to assess four real-world POI datasets extracted for a highly urbanized neighborhood in Singapore. The datasets were found to have poor completeness with errors of commission and omission, although spatial errors were reasonably low (<60 m). Thematic accuracy in names and place types varied. The move towards standardized validation metrics depends on factors such as data availability for intrinsic or extrinsic methods, varying levels of detail across POI datasets, the influence of matching procedures, and the intended application of POI data.

Keywords: point of interest; volunteered geographic information (VGI); data quality

check for updates

Citation: Yeow, L.W.; Low, R.; Tan, Y.X.; Cheah, L. Point-of-Interest (POI) Data Validation Methods: An Urban Case Study. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 735. https://doi.org/10.3390/ijgi10110735

Academic Editors: Nicholas Hamm, Qian (Chayn) Sun, Keone Kelobonye and Wolfgang Kainz

Received: 30 June 2021
Accepted: 23 October 2021
Published: 29 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Points of interest or POI refers to places of interest frequently visited by human traffic throughout the day, including restaurants, supermarkets, transportation hubs, parks, cafes, and tourist attractions. Given the ubiquitous use of mobile devices and advancements in various location-aware technologies [1,2], it is now possible for mobile service providers and technology companies to analyze users' mobility data at increasing geospatial-temporal resolutions to identify neighboring POIs [3]. Location-based social networks (LBSNs), such as Google Maps and Swarm by Foursquare, also emerged. LBSNs rely on their community of end-users to maintain their geospatial database. This is accomplished by soliciting reviews and different semantic information about a recently visited location, benefiting other users in the process [4]. Government agencies and commercial data providers also maintain their own proprietary databases of business establishments and critical facilities, which support various purposes such as market research, policy-making, and urban planning.

Given the availability of geospatial information from different sources, POI data were used in a wide range of applications, from real estate valuations [5], disaggregated employment size estimation [6] and land use classification [7], to the transportation domain, with trip purpose inference [8], stop activity prediction [9] and travel demand modeling [10].

However, there are concerns about data quality of the POI information stored in these data sources, especially those that rely heavily on crowdsourced data or voluntary contributions from their end-users [11]. While different POI data providers implemented their own set of data validation workflows, the detailed documentation of these workflows is often kept proprietary and/or lacking. Furthermore, it is often challenging to compare between different POI sources as they do not follow a standard set of validation metrics that comprehensively and objectively evaluate different aspects of their databases' data quality.

This study first provides a review of POI validation methods that evaluate different aspects of data quality, with the further aim of identifying a set of validation metrics that can be applied across POI datasets. To demonstrate the application of the validation methods identified, they will be implemented and applied to four different POI datasets within a study area in Singapore. The POI datasets are obtained from OpenStreetMap (OSM), Google Maps, HERE Maps, and OneMap, while POI data from the Singapore Land Authority (SLA) is used as reference data. With the variety of commercial and governmental data sources available, users of POI data are likely to encounter the mentioned challenges when comparing different data sources. As a secondary contribution to the literature, a thorough analysis of the validation results will be conducted between the different datasets evaluated in this study to provide readers with the means of interpreting the results.

2. Review of Approaches for Validating POI Data Quality

A literature review of approaches for validating POI data quality was conducted with the search terms "POI", "points-of-interest", "data quality", "assessment", "validation", and "methods" on the search engine Google Scholar and Web of Science, seeking publications published from year 2010 onwards. The search returned 40 studies, including 7 review papers or editorials [11–17]. In addition, to ensure that the review is as representative as possible, further validation approaches were obtained from previous review articles [16,17] and the most recent review by Fonte et al. in 2017 [11].

As shown in Figure 1, the total number of articles found in the literature review fluctuates across the years. The highest number of articles (seven) were found in 2014, with about one review or editorial per year on average from 2014 onwards. Articles from the peak in 2014 generally focus on laying the frameworks for assessing volunteered geographic information (VGI) quality [18,19], along with the development of methods for assessment [20–23]. In recent years, despite the lower number of articles, the field developed in several new directions, such as in more sophisticated statistical machine learning methods for quality assessment [24,25] and POI matching [26,27]. Recent studies also explore novel data sources beyond OSM, such as LBSNs and review websites [27,28], and broader applications of POI data [10,29,30].

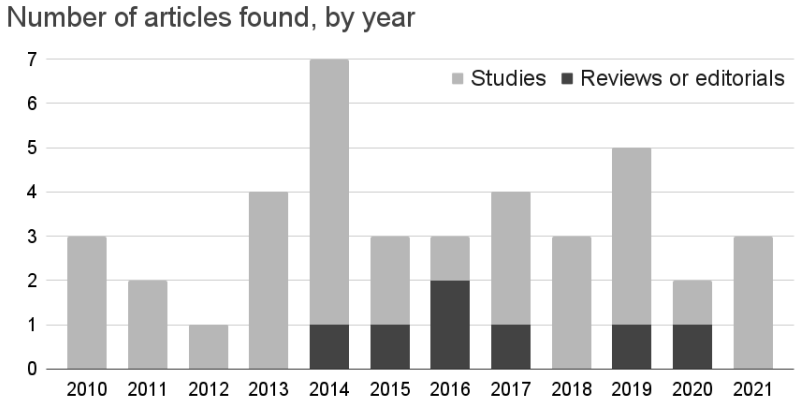

Figure 1. Number of articles found in literature review, by year and type.

The relevant approaches applicable to POI data sources were extracted from the surveyed literature and categorized into the six elements of data quality defined by ISO

19157:2013 [31], which are completeness, logical consistency, positional accuracy, temporal quality, thematic accuracy, and usability. Each element will be elaborated upon in the following subsections.

The approaches were also categorized into intrinsic or extrinsic approaches. Intrinsic approaches do not require the use of external knowledge to evaluate a POI source against—such as a ground truth reference dataset—while extrinsic approaches do require the use of such data [17,18]. There was a focus on intrinsic approaches in the context of research assessing crowdsourced POI data or VGI, as such sources are increasingly "more complete and accurate than authoritative datasets" [16]. However, extrinsic approaches using reference data still have value as intrinsic approaches are limited in their power to make definitive evaluations of data quality [32]. In this paper, D_{eval} refers to the dataset under evaluation and D_{ref} refers to the reference dataset.

This review focuses on point-based POI data, which have location information represented with a pair of latitudinal and longitudinal coordinates. Evaluation methods that rely on a building's or area's footprint data are excluded from the review as such data are often not available in POI datasets. As such, two studies [22,23] did not have relevant approaches extracted, as they used the geometric properties (e.g., land area of parks and gardens) of entities in validating POI data quality. The focus on point-based data allows the approaches evaluated in this study to be easily transferable and applicable to datasets that are commonly encountered.

2.1. Completeness

Completeness refers to the presence or absence of POIs in a data source compared to reality. Errors can arise from commission, where excess data exists in a dataset, or omission, where data are missing from a dataset. The simplest approach for measuring completeness is to compare the number of points in D_{eval} and in D_{ref} [10,33–36]. If a correspondence between points in D_{eval} and in D_{ref} was established, the proportion of points in D_{eval} that are found in D_{ref}, and vice versa, serves as a measure of completeness. This measure is superior to "simple feature counts" as errors of commission and omission can be detected [36].

When D_{ref} is unavailable, the intrinsic approach of analyzing community activity and growth rates over time can give an indication of completeness [10,18,21,30,35].

2.2. Logical Consistency

Logical consistency is defined as the "degree of adherence to logical rules of data structure, attribution, and relationships" [31] and can be further broken down into conceptual consistency, domain consistency, format consistency, and topological consistency. Logical consistency was assessed by analyzing the spatial relations between POIs and other map features within the same data source [4,37]. These can be done through a series of topo-semantic checks, where different relational rules are applied according to the type of POI. For example, 'bus_stop' POIs must lie outside of roads, buildings and nature polygons [37]. This approach can also evaluate the positional accuracy of POIs as elaborated in the subsequent subsection.

Logical consistency can also be evaluated by assessing the positional plausibility of POIs using coexistence patterns [24], which was proposed within the context of evaluating OSM POIs at the data entry stage. The positional plausibility of a candidate POI is evaluated with spatial association rules, which are determined by the frequency of occurrence of other POI types within varying distance bands. In OSM, POI types are expressed with tags, which consist of key-value pairs. The authors [24] demonstrated the concept of positional plausibility with Automated Teller Machines (ATMs) in Paris, where adding an ATM POI (amenity: ATM) to the middle of a park would be less plausible as supposed to the Paris downtown area. As its authors note, the success of the method depends on the inherent frequency of POI types and the strength of the underlying patterns in the POI data, if they exist. While some generic patterns in the coexistence of POIs were found across cities,

the method was recommended to be trained and implemented within a single city, thus limiting its ability to be easily applied to new urban areas.

2.3. Positional Accuracy

The positional accuracy of POIs refers to the accuracy of a POIs spatial location, expressed in geographic coordinates, with reference to its true position. When a correspondence between points in D_{eval} and in D_{ref} was established, positional accuracy can be directly measured by calculating the distance between the locations of the corresponding points in D_{eval} and in D_{ref}. The distance metric can be in the form of Euclidean distances or in terms of distances in the x- or y-dimensions [4,28,29,33,34,36,38–41], after transforming geographic coordinates into a planar coordinate system. While the Euclidean distance provides a single value of the distance between points, the x and y distances also indicate the direction of biases, which might arise from systematic data errors that could cause points in D_{eval} to be shifted by a fixed amount in a particular direction.

If other features are available in D_{ref} or D_{eval}, the spatial relation to other features such as roads, rivers, or buildings could provide indication of the positional accuracy of POIs [4,28,37]. Similar to the approach of identifying coexistence patterns for evaluating logical consistency, certain POIs are more likely to be colocated with other POI types, and deviations from the pattern could indicate errors in position. However, interpreting these spatial relations is still necessary to make judgments on positional accuracy [4]. Additionally, large stacks of POIs with the same location coordinates could occur. If these stacks occur in unlikely areas, such as in a sparse rural neighborhood, then the positional accuracy of the POI stack is doubtful and manual verification is necessary [28].

Spatial analysis tools can also be used to evaluate positional accuracy, such as by comparing the nearest neighbor index (NNi) or the Cross-K function between D_{eval} and D_{ref} [28]. The NNi expresses the extent of clustering (NNi < 1) or dispersion (NNi > 1) of a point pattern relative to complete spatial randomness. While the magnitude of the NNi is not a quality criterion, it can reveal differences in the spatial distributions of POI between data sources. On the other hand, the Cross-K function depicts the relative clustering between points from different data sources. When executing the Cross-K function between D_{eval} and D_{ref}, the relative clustering can give an indication of positional accuracy in terms of the similarity in their spatial coverage.

When D_{ref} is unavailable, intrinsic approaches focusing on the POI contributions and contributors can provide an indication of positional accuracy. Machine learning methods can estimate positional accuracy based on the spatial, temporal and user characteristics of contributions [42]. Some external data will first be required to train the estimation model, although unsupervised learning methods can be used to glean insights into the characteristics of contributors [25].

Otherwise, the map scale and screen resolution at which POI is entered by the user can quantify positional errors due to POI placement [43], if such data are available. Finally, the number of contributors in an area has also been shown to be an indicator of positional accuracy [44] by the concept of Linus' Law, which states that quality is assured when sufficient parties are involved.

2.4. Temporal Quality

The temporal quality refers to the time-related aspects of POIs. This could be related to the accuracy of temporal details of POI, such as the opening hours of a restaurant, or currency, which refers to the time when the POI was created or updated. Most of the temporal-related approaches reviewed are applicable to OSM POI data, which allows for temporal analysis because POI modifications are tracked and published. The most basic approach is to observe the object capture dates and object versions in OSM [34], which will indicate the currency of the POI data. Analyzing the number of annual contributions, the number of features edited over time [45], or the frequency and magnitude of changes in POI position and attributes [4] can also be indicators of temporal quality. Going further,

the temporal evolution of trustworthiness can be analyzed based on the changes in the thematic, geometric, and qualitative spatial aspect of POIs [46].

2.5. Thematic Accuracy

Thematic accuracy refers to the correctness of the thematic tags of each POI. Thematic tags could refer to any further information associated with a specific POI, but the most common tags encountered in this review are related to the POI attributes of name and place type.

When a correspondence between points in D_{eval} and in D_{ref} was established, the string similarity in POI names in D_{eval} and in D_{ref} can be quantified with several string distance measures like the Levenshtein distance, Longest Common Subsequence (LCS), or the mean of the Token Sort Ratio and Token Set Ratio [4,19,29,40]. For POI place types, similar comparisons can be made, ranging from a simple calculation of the proportion of POIs in D_{eval} with the same place type as the corresponding point in D_{ref} [20,34,41,47,48], or by comparing semantic similarity (such as with the WordNet Similarity Metric) between the place types in D_{eval} and D_{ref} [40,49–51]. When D_{ref} is unavailable, calculating the proportion of POIs in D_{eval} with missing attributes provides an indication of thematic accuracy [20,41,52].

2.6. Usability

Usability is a broad element of data quality which refers to the extent to which the data fits the user requirements and could involve multiple data quality elements previously mentioned. The fitness-for-use of POIs can be defined by splitting the data into different use cases, each having different areas of focus. Examples of use cases could be object-referencing and geo-referencing [53]. POIs can also be incorporated into models to evaluate their ability to explain urban phenomena, such as in the relationship between coffee shop density and housing prices [29].

2.7. Summary of Reviewed Approaches

Table 1 shows 23 distinct approaches for evaluating POI data quality that were extracted from the literature. These approaches were categorized according to the element of data quality it applies to, its intrinsic or extrinsic nature, and if POIs in D_{eval} and D_{ref} need to be matched (matching procedure described in Section 3.4). A subset of these approaches will be chosen to evaluate four datasets labeled A to D. More details on the datasets are provided in Section 3.2, while the rationale for implementing a subset of approaches can be found in Section 3.3.

In general, of the six elements of data quality, positional accuracy had the most number of possible approaches (8), followed by thematic accuracy (4) and temporal quality (4). Logical consistency and usability were the least represented elements with only two approaches found. While there was an even spread of intrinsic and extrinsic approaches, some elements of data quality appeared to be better represented with either an intrinsic or extrinsic approach. For example, completeness, thematic accuracy, and usability had more extrinsic approaches, while all of the reviewed approaches for temporal quality were intrinsic.

Some approaches appeared to be more frequently used in the literature compared to others. Analyzing the distribution of spatial error between corresponding points in D_{eval} and D_{ref} was the most popular approach, with 10 studies using this approach for evaluating positional accuracy. Another common approach was the proportion of POIs in D_{eval} with the same classification as the corresponding point in D_{ref}, which was used in five studies.

Table 1. Reviewed approaches for evaluating POI data quality. D_{eval} refers to POI dataset being evaluated (A to D), while D_{ref} refers to reference POI dataset. Checks and crosses indicate if approach was chosen for implementation in case study. Elaboration on why an approach was not chosen is provided in remarks column.

Measure	Approach		Reference(s)	A	B	C	D	D_{ref}	Remarks
Completeness	Ex	Comparison of number of points in D_{eval} and D_{ref}	[10,33–36]	✓	✓	✓	✓	✓	
	Ex, M	Proportion of points in D_{eval} found in D_{ref}, and vice versa	[4,29,30,39]	✓	✓	✓	✓	✓	
	In	Community activity, growth rates over time	[10,18,21,30, 35]	✗	✗	✗	○	✗	
Logical Consistency	In	Evaluates the positional plausibility of incoming POI using coexistence patterns	[24]	✗	✗	✗	✗	✗	Specific to validation of new POI, not generalized for application on verifying entire dataset.
	Ex/In	Spatial relation to other features (e.g., roads, rivers, buildings) (Also: positional accuracy)	[4,28,37]	✗	✗	✗	✗	✗	Manual verification not scalable to large datasets.
Positional Accuracy	Ex, M	Distribution of spatial error (euclidean distance, x-y distance) between corresponding points in D_{eval} and D_{ref}	[4,28,29,33, 34,36,38–41]	✓	✓	✓	✓	✓	
	Ex/In	Spatial relation to other features (e.g., roads, rivers, buildings) (Also: logical consistency)	[4,28,37]	✗	✗	✗	✗	✗	Manual verification not scalable to large datasets.
	Ex	Manual verification of large POI stacks (POIs with the same location)	[28]	✗	✗	✗	✗	✗	Manual verification not scalable to large datasets.
	Ex	Comparing Nearest Neighbour Index of D_{eval} and D_{ref}	[28]	✓	✓	✓	✓	✓	
	Ex	Comparing relative clustering with Cross-K function	[28]	✓	✓	✓	✓	✓	
	In	Machine learning methods on spatial, temporal and user characteristics of contributions	[25,42]	✗	✗	✗	○	✗	
	In	Map scale and screen resolution at which POI is entered by the user	[43]	✗	✗	✗	✗	✗	Metadata at the point of data entry is not available.
	In	Number of contributors in an area (Linus' Law)	[44]	✗	✗	✗	○	✗	
Temporal Quality	In	Frequency and magnitude of changes in POI position and attributes	[4]	✗	✗	✗	○	✗	
	In	Object capture date, object version	[34]	✗	✗	✗	○	✗	
	In	Number of nodes, way, and relations contributed annually in OSM, Number of features edited over time	[45]	✗	✗	✗	○	✗	
	In	Measures trustworthiness based on the changes in the thematic, geometric, and qualitative spatial aspect of each feature in any map that supports versioning	[46]	✗	✗	✗	○	✗	
Thematic Accuracy	Ex, M	Distribution of string similarity of POI names between corresponding points in D_{eval} and D_{ref} using Lowest Common Subsequence (LCS), Levenshtein distance, and the mean of Token Sort and Token Set Ratio.	[4,19,29,40]	✓	✓	✓	✓	✓	
	Ex, M	Proportion of POIs in D_{eval} with the same classification as corresponding point in D_{ref}	[20,34,41,47, 48]	✓	✓	✓	✓	✓	
	Ex, M	Comparing semantic similarity between category name from D_{eval} and D_{ref}(e.g., WordNet Similarity Metric)	[40,49–51]	✓	✓	✓	✓	✓	

Table 1. *Cont.*

Measure	Approach		Reference(s)	A	B	C	D	D_{ref}	Remarks
	In	Proportion of POIs in D_{eval} with missing attributes	[20,41,52]	✓	✓	✓	✓	✓	
Usability	Ex	Modeling POIs in D_{eval} with urban phenomena (e.g., coffee shop density and housing prices)	[29]	✗	✗	✗	✗	✗	Method's success depends on data availability and quality of chosen urban phenomena.
	Ex	Defining Fitness for use for POI by splitting into 2 different use cases (Object-referencing and geo-referencing) each having different focuses	[53]	✗	✗	✗	✗	✗	Dependent on use case, not applicable for generic validation of data.

Legend: Ex—Extrinsic. In—Intrinsic. M—Matching required. ✓—Implemented. ○ —Possible to implement. ✗—Not implemented.

3. Case Study: Implementation of Reviewed Approaches in Singapore

As a case study, a selection of the reviewed approaches was implemented on four POI data sources in the study area of Singapore to demonstrate their applicability. The source code for all of the POI validation methods implemented in this study is shared in a public code repository (Appendix A) to ensure reproducibility and enable the adoption of these methods in other study areas of interest.

3.1. Study Area

The area of interest in this study is the town of Tampines located in the east of Singapore (Figure 2). The Tampines Planning Area has a total land area of 20.97 km^2 [54] housing a population of 261,230 people [55] in 2015, giving an average population density of 12,457 people per km^2. Tampines has several large public transit nodes surrounded by multiple large retail malls. The town also houses several industrial estates and business parks, in addition to community services like schools, places of worship and medical facilities. The high population density and the wide representation of diverse land uses and activities make Tampines a good study area for implementing and testing the approaches for validating POI data quality.

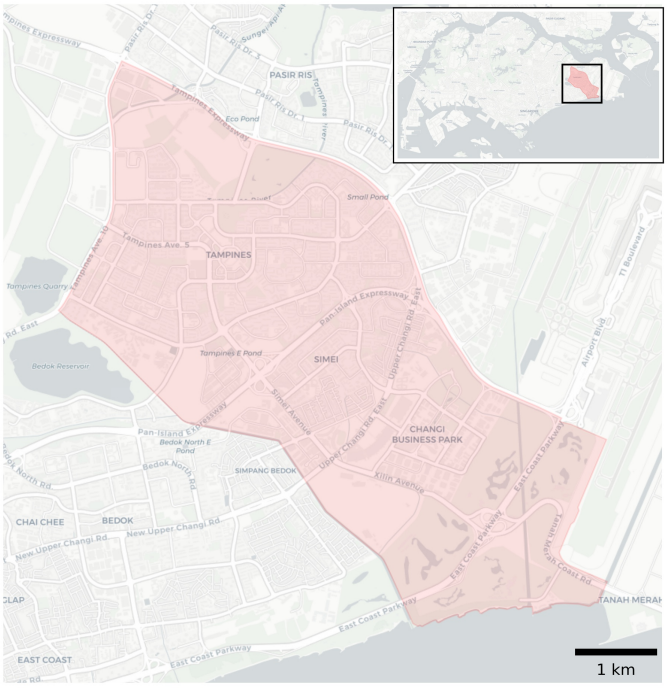

Figure 2. Overview of study area of Tampines (red polygon), located in east of Singapore (inset). Map tiles by Carto, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

3.2. POI Dataset Description

This section provides a thorough description of the four POI data sources selected for evaluation in this case study. These data sources are OpenStreetMap (OSM), Google Places, HERE Maps, and OneMap. These datasets are selected based on various factors, namely, their popular use in the literature, popularity among users, and availability. Heatmaps showing the geographic distribution of POIs from each data source within the study area are available in the Appendix A (Figure A1).

OSM is the result of a community-driven initiative to develop a global geospatial data source maintained by a community of volunteers. Due to the initiative's dedication towards the growth, development, and distribution of free geospatial data, the database is made freely available online for use by any parties and for any purposes. Users are provided with a wide range of alternatives to download the data through various online channels and at different geospatial scales, from the entire planet down to individual continents, countries, metropolitan areas, and specific user-defined areas using the Overpass API [56]. Given that the data procurement process for OSM is heavily reliant on a crowdsourcing approach in maintaining the database, this led to some inconsistencies in how new locations are added or updated in the database by different contributors [57]. While several guidelines are generally in place to reduce such inconsistencies between different contributors, it is challenging to enforce.

Google Maps is a commercial mapping service developed by Google, providing users with different features such as satellite imagery, aerial photography, street maps, panoramic views of streets, real-time traffic conditions and route planning. Users of the platform can also obtain detailed POI information about a specific region using the Places API by including its geospatial coordinates and the search radius [58]. To maintain the relevancy of its geospatial database, the platform currently relies on various approaches such as vetting the current database against authoritative sources and encouraging its end-users to provide timely feedback on any discrepancies. Recently, the Google Maps team also began leveraging satellite imagery and computer vision advancements to automate the mapping process of building outlines [59].

HERE Maps is another example of a commercial mapping service provider that provides a rich set of geospatial POI data to its end-users. While the company advertises state-of-the-art technology and leading mapping processes to compile its database of geospatial data [60], specific details about the platform's data validation process were not available from their documentation. Users of the mapping service can obtain different attributes about a POI, such as the name and address information, by using the Places (Search) API [61] as well as report any mapping discrepancies by using the Map Feedback API [62].

Apart from commercial data providers, another authoritative source of geospatial data are government agencies that rely on such datasets to support land development, housing, employment, and transportation policies. An example is the authoritative national map of Singapore named OneMap. OneMap allows the general public to access a wide variety of up-to-date information and services supported by various government agencies through a mobile or web application. Some examples of such services include allowing users to query for different information on a particular location, such as land information and nearby amenities, as well as traffic conditions and travel directions using various modes of transportation [63]. Users can also utilize the OneMap RESTful API to query for different thematic information from various government agencies, such as parking lots, healthcare centers, food establishments, parks, historic sites, and museums [64].

In addition to the four POI data sources considered, Singapore Land Authority's (SLA) POI dataset will be used as the authoritative reference dataset for extrinsic validation approaches. SLA is the national mapping agency of Singapore and maintains the POI dataset. This dataset was chosen as the reference dataset due to its authoritative source which is assumed to provide reliable and accurate data. The SLA POI dataset was chosen over OneMap as SLA, being the national mapping agency, is deemed to be more reliable

than the combination of sources that makes up OneMap. The licensed SLA dataset contains detailed attributes of 27 different place types within Singapore, such as education institutions, transportation hubs, religious buildings, local government offices, and key medical facilities.

3.3. Choice of Approaches for Implementation

From the range of approaches reviewed in Section 2, a subset of approaches were selected for implementation in the case study. Table 1 shows the applicability and suitability of implementing each approach on the different POI sources. Except for D_{ref}, each evaluated data source (D_{eval}) is relabeled from 'A' to 'D' to bring attention to the interpretation of the results and avoid comparing the superiority of data sources. The implemented approaches were selected for their fit with inherent features of the POI data sources identified in this study. More specifically, the selected approaches utilized particular POI attributes that were present in the data sources considered, such as the POI's name, address, or place type information. Furthermore, identifying commonly found attributes across datasets can be used to inform the future development of standardized POI validation metrics.

The remaining approaches were not implemented for one of the following reasons: (1) the approach required highly specific data that was not present across all POI datasets, (2) the approach was not scalable to large datasets, or (3) the approach fell outside the study scope of large, existing, generalized POI sources. In the first case, some approaches were only applicable to POI sources that contain historical data of POI edits, contributions, and details of the contributors (such as with data source D). Examples of these approaches are the use of community activity and growth rates over time to evaluate completeness [18,21,35], or analyzing the capture date and version of POIs to evaluate temporal quality [34]. In the second case, approaches such as observing spatial relations [4,28] or the verification of large POI stacks [28] relied on human input and thus are not scalable to large datasets.

The reasons for excluding certain approaches from further implementation are provided in Table 1. In total, the list of reviewed approaches was narrowed down to nine approaches, most of which are extrinsic, and cover the elements of completeness, positional accuracy, and thematic accuracy.

3.4. Matching Procedure

Given that several of the extrinsic approaches first require the POIs in D_{eval} to be matched with their corresponding counterparts in D_{ref} before evaluation, a POI matching procedure was implemented in this study and applied to each of the four POI data sources (i.e., OSM, Google Places, HERE Maps, and OneMap—D_{eval}) to identify the matching POIs in the SLA reference dataset (D_{ref}).

The POI matching procedure, adapted from previous work [65], considered three similarity measures (i.e., spatial similarity, name similarity, and address similarity) in a two-stage matching process to identify POI matches between D_{eval} and D_{ref}. The first stage begins by considering the spatial similarity between the POIs in D_{eval} and D_{ref} by filtering out all neighboring POIs that fall within 100 m around the POI to be matched. The second stage of the POI matching process subsequently compares each neighboring POI against the POI to be matched based on their name and address similarity measures to identify matches.

The name similarity measure between each POI pair is calculated by first tokenizing the POIs' name information and alphabetically sorting the resulting tokens before calculating the Levenshtein distance between the two formatted name strings.

However, given that the address information between two neighboring POIs is likely to be very similar with slight differences in terms of their street or block numbers, a string comparison approach, which places an equal weight on each matching string, would likely be unsuitable. Therefore, a weighted approach was adopted when calculating the address similarity measure between each POI pair by placing a smaller weight on

string matches that occur more frequently (i.e., street name and country), while placing a larger weight on string matches that occur less frequently (i.e., block number and street number). Based on these requirements, the address similarity measure is calculated by first vectorizing the address information for all neighboring POIs and the POI to be matched using Term Frequency-inverse Document Frequency (TF-IDF) before calculating the cosine similarity score between each pair of address vectors. TF-IDF is a numerical statistic that can represent the significance of a word t within the document and other documents in the same collection by increasing proportionally based on the number of times it appears in document d but is offset when it repeatedly appears in the entire collection D (refer to Equation (1)).

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) * \text{IDF}(t,d) \tag{1}$$

where

$$\text{TF}(t,d) = \frac{\sum_{j \in d} 1_{j=t}}{|d|} \tag{2}$$

$$\text{IDF}(t,d) = \log \frac{|D|}{\sum_{k \in D} 1_{t \in k}} \tag{3}$$

In the context of this study, d corresponds to the address string of a neighboring POI, while D represents the address list of all neighboring POIs that fall within 100 m from the POI to be matched. After calculating the address similarity metric based on the cosine similarity score between each POI pair's address vectors, the result is combined with their name similarity scores and passed into a machine learning classifier to infer if they are a match.

A detailed evaluation of the POI matching procedure described above was conducted in a previous study [65] using a ground truth POI dataset that was collected in the eastern region of Singapore (Southwest: 1.331747, 103.961258; Northwest: 1.339397, 103.961258; Northeast: 1.339397, 103.969027; Southeast: 1.331747, 103.969027) and manually labeled to indicate all 200 POI matches and 8498 POI non-matches in the region of interest. By performing a 75/25 train-test split on the ground truth dataset, the machine learning classifier was trained using a combination of hybrid sampling techniques and ensemble learning algorithms such as Gradient Boosting [66] and Bootstrap Aggregation [67] to overcome the imbalance between the POI matches and non-matches in the training dataset. By evaluating the resulting model on the test dataset, the model was able to report matching accuracies up to 97.6% for balanced accuracy and 97.2% for overall accuracy, outperforming all baseline approaches considered in the study.

4. Validation Approaches and Results

The following results will be presented for the four evaluated POI data sources mentioned in Section 3.2. As previously mentioned in Section 3.3, the evaluated data sources (D_{eval}) are labelled A to D. In the approaches that require a correspondence between POIs in D_{eval} and D_{ref}, the matching procedure described in Section 3.4 was used.

4.1. Completeness

Table 2 shows the comparison of the number of points in D_{eval} and D_{ref} and the proportion of points in D_{eval} found in D_{ref}, and vice versa. As some data sources might be more abundant in POIs of a particular place type compared to the others, the comparison was also conducted for a subset of POIs of a particular place type (school-related POIs), such as 'school', 'university', 'secondary_school', 'primary_school' and 'college'. School-related POIs were chosen because the place type was common across most of the data sources, except data source A, and was relatively abundant in D_{ref}.

Considering all POIs, data source C has the highest number of POIs, with about 8-times as many points as the next largest data sources (data source A and B). Data source D and D_{ref} have the least number of points. The larger number of points for data source

C could indicate a true wider coverage of points or errors of commission in the form of duplicate or excess POIs that do not truly exist.

Table 2. Results for completeness across all POI place types and a subset of school-related POIs.

Data Source	(1) Number of Points	(2) Number (Proportion) of Points in D_{eval} Found in D_{ref}	(3) Number (Proportion) of Points in D_{ref} Found in D_{eval}
All POIs			
A	1206	34 (0.028)	30 (0.054)
B	1210	493 (0.407)	245 (0.44)
C	9618	702 (0.073)	315 (0.566)
D	631	167 (0.265)	193 (0.346)
D_{ref}	557	—	—
School-Related POIs			
A	—	—	—
B	143	118 (0.825)	100 (0.99)
C	248	55 (0.222)	63 (0.624)
D	38	22 (0.579)	31 (0.307)
D_{ref}	101	—	—

Overall, the proportion of points in D_{eval} found in D_{ref} and vice versa (columns 2 and 3) is low (less than 0.566). In terms of the proportion of points in D_{eval} found in D_{ref} (column 2), data source A has the lowest proportion of points found in D_{ref} (0.028), indicating that the level of completeness is lacking relative to D_{ref}. This is followed by data source C with a slightly higher proportion of 0.073. Data source D has the next highest proportion (0.265), followed by data source B with the highest proportion of 0.407.

In terms of the proportion of points in D_{ref} found in D_{eval} (column 3), data source A still has the lowest proportions of 0.054, followed by data sources D (0.346), B (0.44) and C (0.566). A low proportion of POIs in data source C could be found in D_{ref}, but a much higher proportion of points in D_{ref} can be found in data source C. Other sources have a more balanced proportion (e.g., 0.407 vs. 0.44 for data source B).

A low proportion of points in D_{eval} found in D_{ref} could indicate errors of commission, where there are excess points in D_{eval} which do not exist in D_{ref} or in reality. On the other hand, a low proportion of points in D_{ref} found in D_{eval} could indicate errors of omission, where D_{eval} is lacking in points that exist in D_{ref}. In both cases, a low proportion could also be due to differences in coverage for specific place types between D_{eval} and D_{ref}. This is evidenced by the considerably higher proportion of points in D_{eval} found in D_{ref} and vice versa when considering the subset of school-related POIs. Additionally, it was observed that data source A has a higher proportion of POIs relating to restaurants, food outlets and shops. On the other hand, being a government source, a majority of POIs in D_{ref} are related to ATMs, schools, political venues, and doctors.

By comparing the number of points in D_{eval} that were found in D_{ref} and vice versa (columns 2 and 3), it is observed that for data sources B and C, about twice as many points in D_{eval} were found in D_{ref} (column 2) compared to the number of points in D_{ref} that were found in D_{eval} (column 3). This is an indication of the extent to which many-to-one matches in the data sources are present.

4.2. Positional Accuracy

4.2.1. Distribution of Spatial Error (Euclidean Distance, x-y Distance) between Corresponding Points in D_{eval} and D_{ref}

Figures 3 and 4 show the distribution of spatial error between corresponding points in D_{eval} and D_{ref}, where spatial error is measured with both the euclidean distance and the x-y distance. All points were projected to the planar SVY21 coordinate reference system (EPSG: 3414) before distances were calculated. In general, the points in D_{eval} are within 70 m of the corresponding point in D_{ref}. In terms of mean euclidean distances, Data sources B and D have lower spatial error compared to data sources A and C. The lower spatial error for data sources B and D might be a result of possible bulk POI imports from governmental sources.

The spatial error of data source A has a comparable mean and distribution compared to data source C, but the low number of points in data source A limits any further conclusions on its positional accuracy. Visually, the distribution of x-y errors (Figure 4) do not exhibit bias in any direction, and the errors are scattered radially about the origin (0, 0).

The wider spread of spatial errors for data sources A and C could be due to the data sources having more detailed POIs of individual establishments. When located within a larger complex, these detailed POIs are matched to the aggregated version of the POI in D_{ref}, which might be located at the centroid of the complex. This phenomenon will be elaborated in the subsequent section on many-to-one matches (Section 5.1).

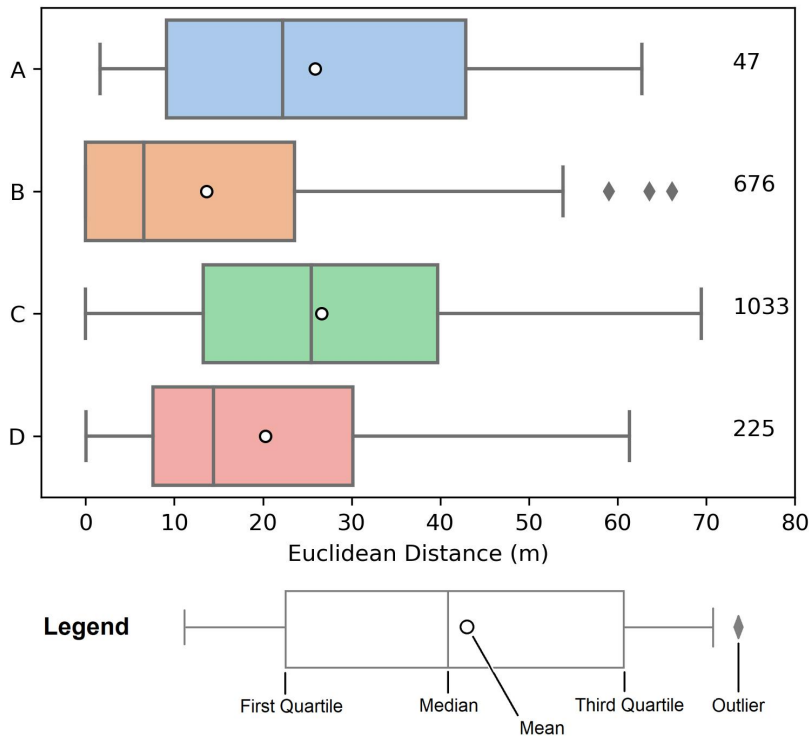

Figure 3. Distribution of spatial error based on euclidean distance between corresponding points in four data sources A to D (D_{eval}) and D_{ref}. Numbers on right of distribution indicate number of corresponding points in each data source.

4.2.2. Comparing Nearest Neighbor Index of D_{eval} and D_{ref}

The Nearest Neighbor Index (NNi) provides an indication of how a point pattern is clustered relative to a point pattern with complete spatial randomness (CSR). NNi is calculated with the equation:

$$NNi = \frac{r_A}{r_E}, \tag{4}$$

where r_A is the average distance to the nearest neighbor in the analyzed point pattern, and r_E is the expected nearest neighbor distance in a CSR point pattern. For a CSR point pattern with intensity λ (points per unit area), r_E is given by:

$$r_E = \frac{1}{2\sqrt{\lambda}} \tag{5}$$

When the NNi of a point pattern is less than one, the point pattern is interpreted to have a greater degree of clustering relative to a CSR point pattern, as the average distance to the nearest neighbor in the point pattern is less than the expected average distance in a CSR point pattern ($r_A < r_E$). In the extreme case where all the points in a point pattern are in exactly the same location, the average distance to the nearest neighbor (r_A) will be zero,

giving an NNi of zero. On the other hand, a point pattern with an NNi greater than one indicates a tendency towards evenly spaced points, or dispersion.

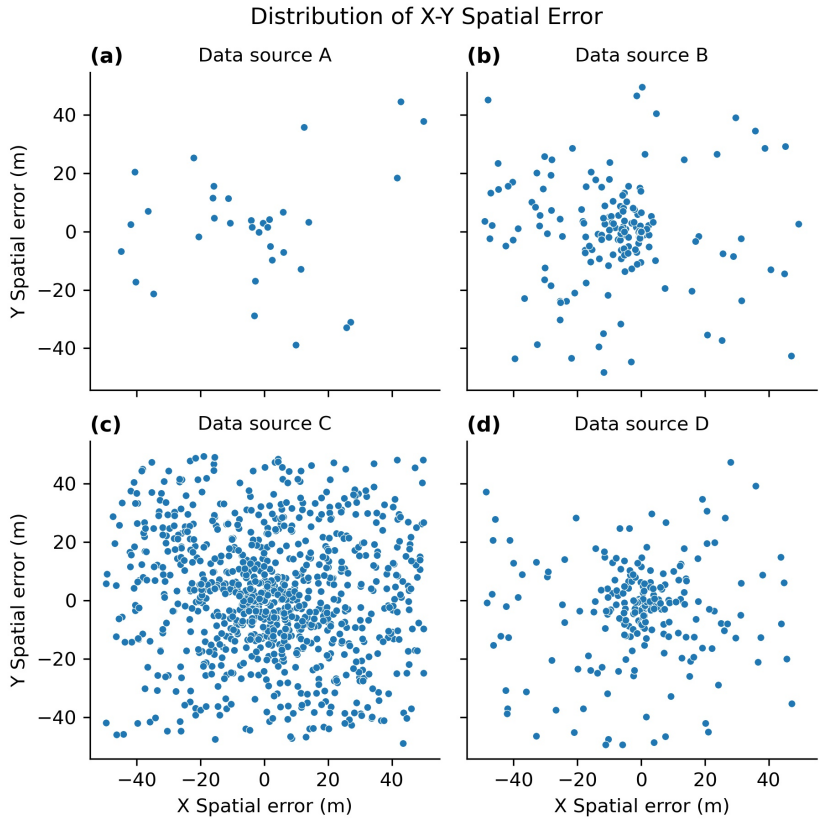

Figure 4. Distribution of spatial error in X (longitude) and Y (latitude) axes between corresponding points in D_{eval} and D_{ref} for the four data sources A to D (**a**–**d**).

The NNi values for each of the data sources, along with the observed mean and expected average distance between points are shown in Table 3. Similar to the findings of Hochmair et al. (2018) [28], the NNi values across maps are less than one, indicating clustering. Data sources A and B have the highest degree of clustering (lowest NNis) while data source D and D_{ref} have the lowest degree of clustering (highest NNis). Data source D is clustered in the manner that is closest to D_{ref} (0.51 vs. 0.47), while data source A is much more clustered compared to D_{ref} (0.23 vs. 0.45).

The differences in NNi could be explained by the different types of POIs being covered with different spatial characteristics. For example, as noted by Hochmair et al. (2018) [28], restaurants and shops tend to be clustered along the same commercial areas, which might have a higher representation in data source A and C. More investigation is needed to examine clustering patterns within each data source to explain the different NNi values observed.

Table 3. Nearest Neighbor Index (NNi) values for four evaluated data sources (D_{eval}) and reference data source (D_{ref}).

Data Source	Mean Distance, r_A (m)	Expected Distance, r_E (m)	Nearest Neighbour Index ($r_A \div r_E$)
A	20.69	88.23	0.23
B	22.94	92.86	0.25
C	10.11	30.02	0.34
D	63.39	124.72	0.51
$\boldsymbol{D_{\text{ref}}}$	60.90	128.42	0.47

4.2.3. Comparing Relative Clustering with Cross-K Function

The cross-K function is a method for analyzing how two point patterns cluster relative to each other. This is the bivariate version of Ripley's K function [68]. The value of $K_{ij}(r)$ of points from sources i and j as a function of distance r is given by:

$$K_{ij}(r) = \frac{\mathbb{E}(\#\text{ of points from } j \text{ within distance } r \text{ of a randomly chosen point in } i)}{\lambda_j}, \quad (6)$$

where λ_j is the density of the points in j, or the number of points per unit area.

This is followed by a test for statistical significance using a Monte Carlo simulation in which the points from i and j are randomly relabeled. The observed cross-K function is then compared with the simulation envelope of cross-K functions with random relabeling. If the observed cross-K function falls within the simulation envelope, then the point patterns cluster similarly to each other. If the observed cross-K function falls below the simulation envelope, then the point patterns are spatially segregated from each other. In this case, it is suggested that conflating POIs from different spatially segregated sources can be beneficial in improving data coverage [28].

Figure 5 shows the cross-K functions for the various POI data sources (D_{eval}) when compared with D_{ref}. The vertical axis shows the value of $K_{ij}(r)$ against various values of distance r. The observed cross-K function, simulated mean, and simulation envelopes are plotted. As observed in the plots, for data sources A, C, and D, the observed cross-K function falls within the simulation envelope, indicating that the point patterns cluster similarly to each other. However, the observed cross-K function for data source B falls below the simulation envelope to a small degree, indicating that the POIs from data source B and D_{ref} are spatially segregated from each other.

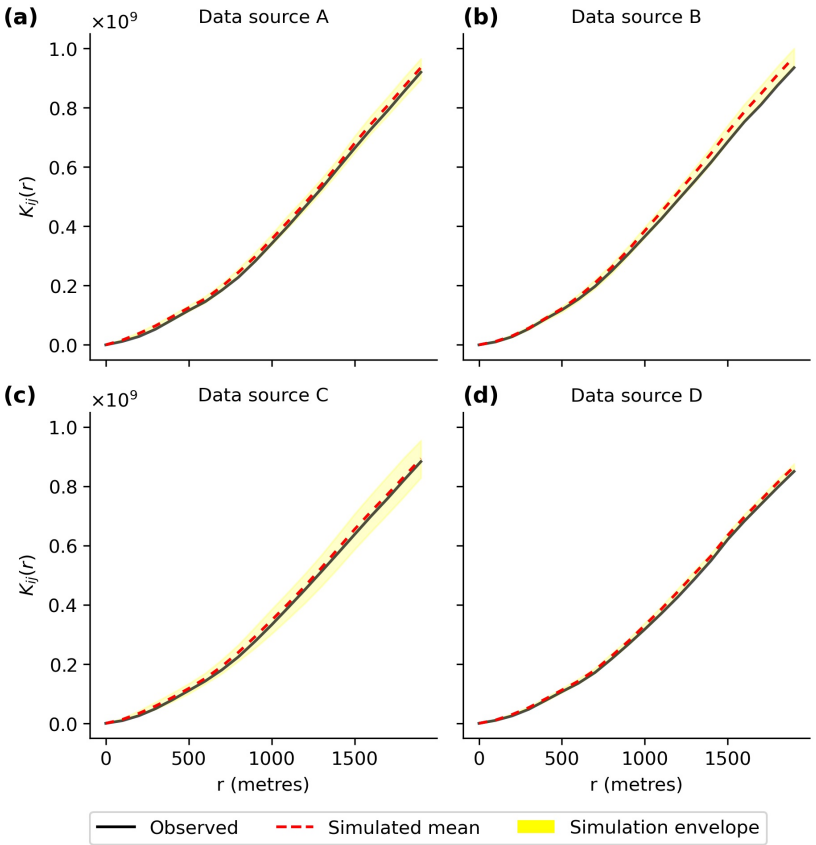

Figure 5. Cross-K functions of D_{eval} against D_{ref} for the four data sources A to D (**a**–**d**).

4.3. Thematic Accuracy

4.3.1. Distribution of String Similarity of POI Names between Corresponding Points in D_{eval} and D_{ref}

String similarity scores are normalized to range between zero and one, where a score of one corresponds to an exact match. The equations below detail the calculation of string similarity scores, $S(i,j)$, between two POI names i and j of length L_i and L_j for the three string distance measures: the Longest Common Subsequence (LCS), Levenshtein distance, and the mean of token sort and token set ratios.

The Longest Common Subsequence (LCS) measure, $L_{\text{LCS}}(i,j)$, refers to the subsequence with the most number of characters that appears in both strings that are being evaluated. If two strings are exactly equal, then the LCS will be equal to the length of the strings. The equation is as follows:

$$S_{\text{LCS}}(i,j) = \frac{2 \times L_{\text{LCS}}(i,j)}{L_i + L_j} \tag{7}$$

The Levenshtein distance, $L_{\text{Lev}}(i,j)$, is the minimum number of edits (insertions, substitutions, and deletions) required to transform one string to another. The similarity score based on the Levenshtein distance is given by:

$$S_{\text{Lev}}(i,j) = 1 - \frac{L_{\text{Lev}}(i,j)}{\max(L_i, L_j)} \tag{8}$$

The mean of the Token Sort Ratio and Token Set Ratio uses the corresponding functions from the 'fuzzywuzzy' Python package of the same name. Taking the mean of the two ratios was found to have higher accuracy scores when identifying matching POIs compared to that of using the ratios individually [69]. The two functions tokenize the words in the POI names and rearrange the tokens in alphabetical order. The Token Sort Ratio "computes the similarity of the two re-ordered strings" while the Token Set ratio "computes the similarity between the intersection and the shorter of the two strings (i.e., the string with the least amount of characters)" [69]. The mean of the two ratios is taken as the string similarity measure as follows:

$$S_{\text{mean token sort, token set}} = \frac{L_{\text{token sort ratio}}(i,j) + L_{\text{token set ratio}}(i,j)}{2} \tag{9}$$

Figure 6 shows the distribution of string similarity of POI names for the three measures and data sources. Across the three string measures, data sources B and D have a higher proportion of exact name matches and higher mean similarity scores compared to data sources A and C. Regardless of the string similarity measure, data source A has lower string similarity scores compared to that of the other data sources. Similar to the case of distribution of spatial error, data sources B and D might have high similarity scores due to bulk imports from governmental sources.

Comparing between the string similarity measures, the Levenshtein measure has the lowest string similarity scores, followed by LCS and finally the mean of Token Sort and Token Set ratios, which has the highest string similarity scores across the data sources. This could be due to the matching procedure used in this study, which evaluates similarity in POI names with the Token Sort Ratio. The relative differences between the string similarity scores of the data sources is maintained regardless of the string similarity measure used. While this might suggest that the choice of string similarity measure is not critical, each string similarity measure has its advantages when dealing with spelling errors, differences in word order and partial name matches.

The LCS and Levenshtein measures differ in that the LCS uses insertions and deletions to convert one string to another, while the Levenshtein distance allows for substitutions, in addition to insertions and deletions. Apart from this, LCS and Levenshtein measures are highly similar, as observed by the similar distributions in string similarity scores. On

the other hand, the Token Sort and Token Set Ratios differ from the LCS and Levenshtein measures due to the additional step of tokenization. As a result, compared to the LCS and Levenshtein measures, the Token Sort and Token Set Ratios are more robust to differences in word order and partial name matches, but less robust to spelling errors. Between the two ratios, the Token Set Ratio provides some independence from name differences arising from varying levels of detail. For example "Noodle Restaurant" and "Chicken Noodle Restaurant" gives a Token Set Ratio of 1. The Token Sort Ratio provides some independence from changes in naming order. For example "Chicken Noodle Restaurant" and "Restaurant Chicken Noodle" gives a Token Sort Ratio of 1.

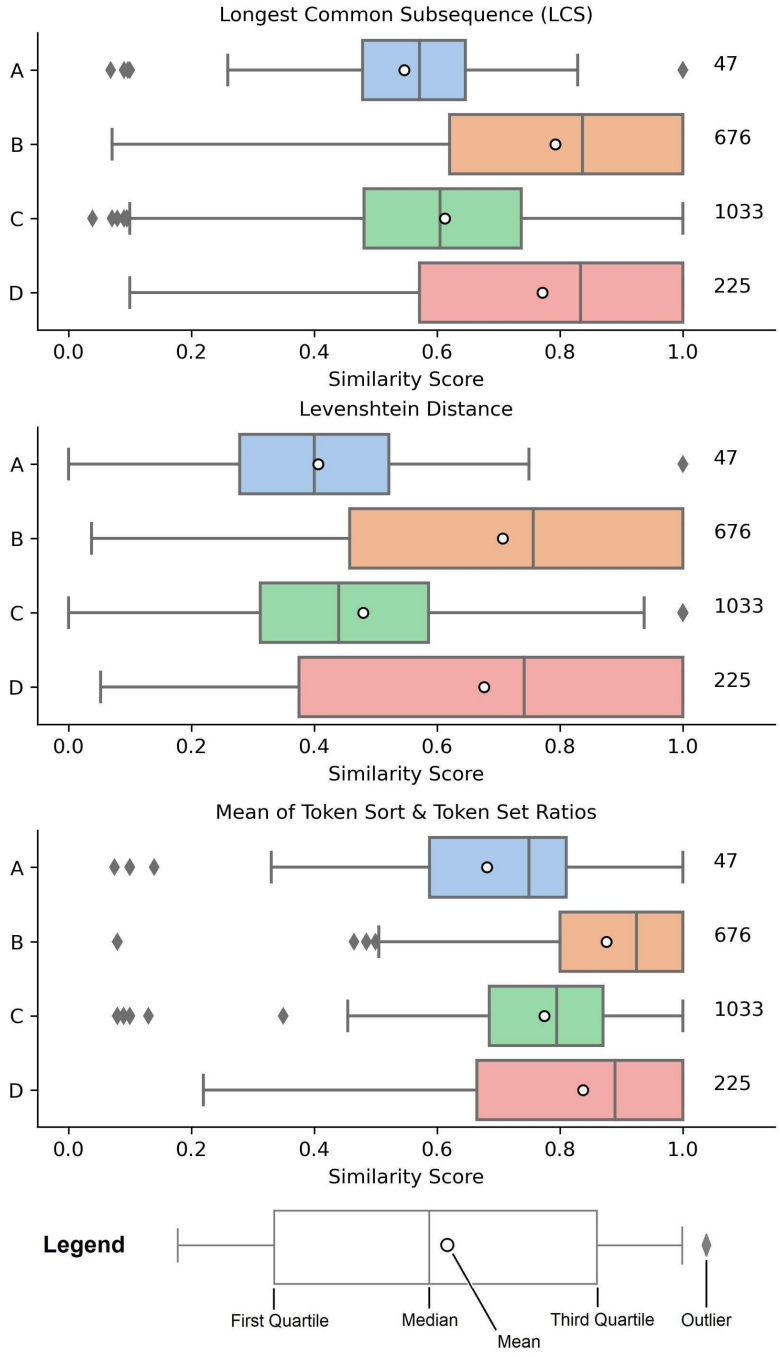

Figure 6. Distribution of string similarity scores of POI names between corresponding points in four data sources (D_{eval}) and D_{ref}. String similarity scores were calculated according to three string similarity measures. Numbers on right of the distribution indicate number of corresponding points in each data source.

4.3.2. Proportion of POIs in D_{eval} with the Same Place Type as Corresponding Point in D_{ref}

For each pair of corresponding points in D_{eval} and D_{ref}, the place types from each data source were compared. Then, the proportion of POIs in D_{eval} which shared at least one place type with the corresponding point in D_{ref} was calculated. The results are shown in Table 4. Using this evaluation criteria of having at least one exact match, the proportions of POIs with the same classification are generally low (<0.29). Data sources B and D have the highest proportion, while data source A has none of the matched POIs sharing the same classification as D_{ref}.

Table 4. Proportion of POIs with same place type using exact match, manual place type mapping, and WordNet similarity metric.

Data Source	Proportion of POIs in D_{eval} with Same Place Type		(3) WordNet Similarity Metric
	(1) Exact Match	(2) With Manual Place Type Mapping	
A	0.000	0.043	0.163
B	0.217	0.235	0.366
C	0.126	0.138	0.529
D	0.293	0.329	0.479

A major reason for the low scores can be attributed to different categorization schema (or taxonomy) of place types. For example, under an exact match, a POI with the place type 'school' would not be considered as having the same classification as another POI with the place type 'primary school'. To consider the case of similar place types, the place types of corresponding points in D_{eval} and D_{ref} were manually mapped, and the proportion was recalculated. An example of the manually mapped place types for the case of data source D is shown in Figure 7.

As shown in column 2 of Table 4, the proportions increased slightly when using the manually mapped place types as compared to using exact matches. This increase is expected as place types without an exact string match but are highly similar in semantic meaning are now considered as having the same classification under the manual mapping process, as previously mentioned with the example of the 'school' and 'primary school' place types. However, manually mapping place types is a labor-intensive process and is dependent on how the place types are interpreted or understood by the individuals doing the mapping. This is especially evident when broad or ambiguous place types are involved. For example, a POI with a 'service' place type might be interpreted as locations where physical services are offered, such as a repair shop or a dry-cleaning shop. The place type could also be interpreted more broadly as a location offering any type of service, such as a bank (financial service), restaurant (food service) or transit station (transportation service). In this study, place types were mapped as conservatively as possible with minimal interpretation.

In addition to using manually mapped place types, the WordNet Similarity Metric was also used to compare semantic similarity between category names from D_{eval} and D_{ref}. For each pair of category names, the names are first tokenized before using WordNet's lin's similarity to compute the similarity score. It is done so using the following equation [70]:

$$N_{\text{sim}}(N_1, N_2) = \frac{\sum_{l_1 \in L_1}[\max_{l_2 \in L_2} \text{sim}(l_1, l_2) + \sum_{l_2 \in L_2}[\max_{l_1 \in L_1} \text{sim}(l_1, l_2)]}{|L_1| + |L_2|}, \quad (10)$$

where $|L_1|$ and $|L_2|$ are the length of the token set and $\text{sim}(l_1, l_2)$ is the lin's similarity metric in WordNet. For example, when evaluating the category pair N_1 = "Secondary School" and N_2 = "school", the pair will be tokenized to the token sets L_1 = ["Secondary", "School"] and L_2 = ["school"]. The similarity of the sets will then be computed as stated in Equation (10). In the scenario where each pair of matched POI has multiple place types, the maximum score out of every combination will be taken as the final score. Taking the maximum score results in higher scores for datasets with multiple place types for each POI,

as there is a higher probability to find a place type that has a similar semantic meaning to the reference place type. This is intended as we are looking to measure the semantic similarity of the two closest place types. For a matched pair of POIs, it is unlikely that all the place types are semantically similar to all the reference place types.

The results of the WordNet similarity metric are shown in column 3 of Table 4. Using the WordNet similarity metric results in higher scores overall compared to columns 1 and 2. Similar to the case of manual place type mappings, the WordNet similarity metric introduces new links between place types which were absent in the manual mapping. As observed in the place type mapping between D_{ref} and data source D (Figure 7), the WordNet similarity metric assigns a relatively high score (>0.5) for place type pairs that did not appear in the manual mapping. For example, 'police' in D_{ref} was only mapped to 'police' in D_{eval} under manual mapping, but WordNet also assigned scores greater than 0.5 for 'office' and 'service'. This would lead to the overall inflation in WordNet similarity scores observed in Table 4. As such, users seeking to evaluate place type similarities with either manual mapping or WordNet should take into account the strengths and limitations of each method, and strike a balance between scalability and accuracy.

Unlike the other data sources, the WordNet similarity score (column 3) for data source C is much higher than that of the exact match and manual place type mapping (columns 1 and 2). This is likely due to the broad naming used in data source C's taxonomy (e.g., establishment). As observed in the case of 'establishment', the broad category is semantically similar to most place types, with the WordNet similarity metric assigning relatively high scores (>0.5) with most place types in D_{ref} (Figure 8). Combined with the fact that 'establishment' is one of the most common place types in the data source C, these factors lead to the much higher WordNet similarity scores for data source C in Table 4. In addition, 'establishment' was not manually mapped to any of the place types in D_{ref} due to its broad nature.

By matching the semantic information of place types, the WordNet similarity metric is able to accommodate the different place type taxonomies used by different datasets in an automatic manner, which is more scalable than mapping place types manually. However, it is possible that WordNet might not be able to account for all place type taxonomies as it is dependent on the data that WordNet was trained on.

4.3.3. Proportion of POIs in D_{eval} and D_{ref} with Missing Attributes

Table 5 shows the proportion of POIs in D_{eval} and D_{ref} that have missing attributes. While some map services might offer additional POI attributes like opening hours, websites, contact numbers, and user reviews, this study will only evaluate the POI attributes that are common between the data sources considered. These attributes relate to the POI's name, coordinates, place type, and address string.

Across all POI sources, coordinates and place types are fully present. However, a significant proportion (0.67) of points in data source D do not have name information. Most of the points are parking spaces, and swimming pools and sports pitches that are mostly located within private property. It can be argued that these place types are ancillary in nature and thus name information would be less critical. Furthermore, it may be possible to automatically generate logical place names, such as by combining place type and address (e.g., "Swimming pool at 1 Example Street").

A small proportion of points in data sources B, C, and D have missing address strings (<0.107). In the case of data source B, most of the points with missing address strings are park facilities within parks and bicycle racks, where addresses are less applicable. While missing addresses can be recovered by reverse geocoding services, examining the prevalence of missing attributes is still a key step, especially when using POI sources for human-facing applications.

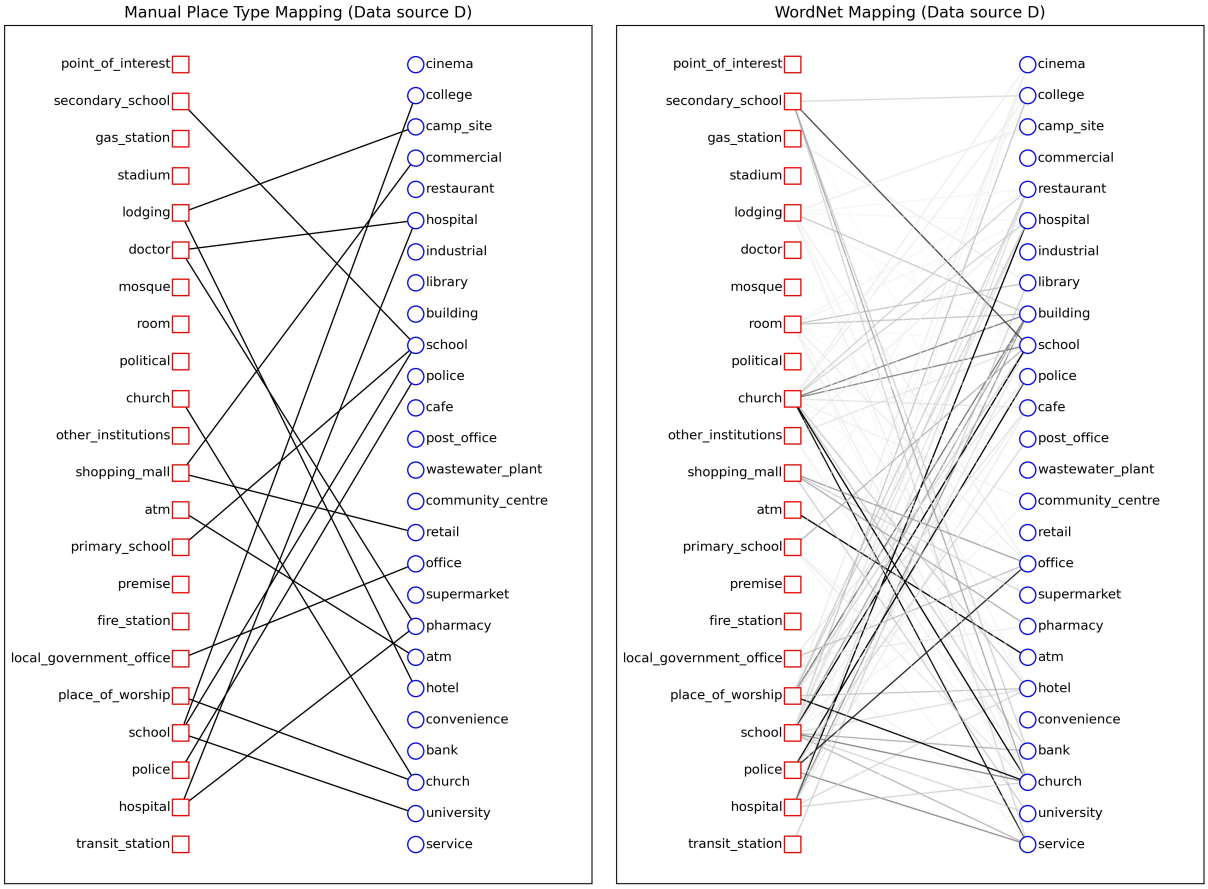

Figure 7. Example of mapping place types between D_{ref} and data source D (D_{eval}) manually (**left**) and with WordNet similarity metric (**right**). Square red nodes on left are place types from D_{ref}, while circular blue nodes on right are place types from data source D (D_{eval}). Darker lines indicate a higher WordNet similarity metric. Only pairs with a WordNet similarity score greater than 0.5 are displayed.

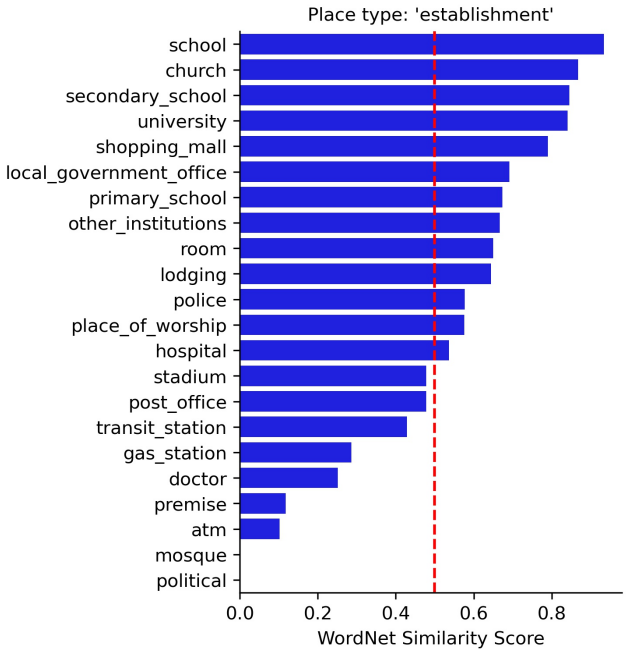

Figure 8. WordNet similarity scores between 'establishment' place type in data source C and various place types in D_{ref}.

Table 5. Proportion of POIs with missing attributes. None of POIs have missing coordinates and place types. These attributes were excluded from table.

Data Source	POI Name	Address String
A	0.0	0.000
B	0.0	**0.107**
C	0.0	**0.010**
D	**0.67**	**0.016**
D_{ref}	0.0	0.000

5. Discussion: Towards Standardized POI Data Quality Metrics

While an industry standard currently exists for the elements of geographic data quality, these standards can be further enhanced to include standardized validation methods for POI data. Having a set of standardized quality indicators would allow potential data users to have a quick, transparent overview of data quality to pick the most suitable data source(s) for their needs. Data administrators adopting these standards can use them to track changes in data quality as POI data are enhanced and augmented with diverse data sources and processed with more sophisticated algorithms. The development of a standardized set of metrics for POI data quality should account for the following issues: (1) differences in the levels of detail between POI datasets, (2) differences between extrinsic and intrinsic methods and the corresponding data and algorithmic requirements, and (3) the intended application(s) of POI data.

5.1. Varying Levels of POI Detail: Presence of Many-to-One Matches

As previously mentioned in Section 4.1, comparing the number of points in D_{eval} that were found in D_{ref} and vice versa suggested the presence of many-to-one matches between D_{eval} and D_{ref}. Closer examination of the matches indeed highlighted several instances where multiple POIs in D_{eval} are matched with a single POI in D_{ref}. More specifically, 27.5%, 71.2%, 40.6% and 13.3% of the matches in data sources A to D respectively, involved many-to-one matches, while the rest were one-to-one matches. An example of many-to-one matching is provided in Figure 9, where multiple POIs in data source C were matched to a single POI in D_{ref}. In this example, 68 POIs in data source C matched to a single POI in D_{ref}, which represented a large mall in the study area. All POIs fall within the building outline in grey, with the exception of one point which was related to a nearby road intersection. The POIs in data source C contained shops, restaurants, and other establishments located within and around the mall. The higher level of detail in POI data sources like data source C, as compared to that of D_{ref}, is a contributing factor to the many-to-one matches observed.

While duplicated POIs in data source C were removed based on their ID information, duplicates could still remain if they were assigned different IDs. Removing such duplicates could be addressed with more sophisticated data cleaning procedures, which are beyond the scope of this study.

With the wide range of POI data sources in many cities around the world, and increasing urbanization and density of activities, there will be a need to match points between data sources with differences in coverage and levels of detail when implementing extrinsic validation methods. The many-to-one matches results in an imbalance in the number of points in D_{eval} found in D_{ref} and vice versa. The distribution of spatial error and string similarity scores would also be artificially increased, since the disaggregated, fine-grained POIs would be matched to the aggregated, coarser version of the POI, which would typically be spatially located at the centroid of the cluster of disaggregated points, and have different POI names and place types (for example, "IKEA Restaurant" being matched to "IKEA").

The differences in levels of detail can also manifest semantically in the form of (A) differences in word order and length of POI names, or (B) differences in the level of specificity of POI place types. In the former (A), measuring differences in POI names with

the mean set sort ratio is superior to the LCS and Levenshtein measures as it is more robust to differences in word order and length. In the latter (B), comparing place types using an exact match might have an edge over the WordNet similarity metric, as the WordNet similarity metric tends to overcompensate for generic place types. Therefore, the selection and interpretation of POI validation metrics should account for differences in the levels of detail between POI datasets, especially when extrinsic methods are used to evaluate datasets with reference data of different levels of detail.

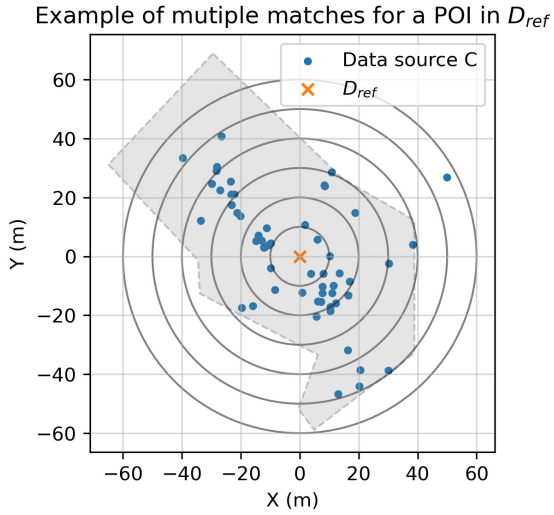

Figure 9. An example where multiple POIs from data source C are matched to a single POI in D_{ref}. Building outline is shown in grey.

5.2. Extrinsic and Intrinsic Methods

Extrinsic and intrinsic methods have their unique characteristics and applications, and these differences should be taken into account. Extrinsic methods focus on measuring the similarities between the evaluation and reference datasets, and a high degree of similarity confers the reliability and accuracy of the reference dataset to the evaluated dataset.

However, there are many limitations with regards to extrinsic methods. Firstly, it might be difficult to obtain reference data if the data are not open-sourced. The validity of the reference data might also be unknown if the method of data collection and verification is not published. As such, extrinsic methods can only evaluate the quality of a dataset up to the extent of the reference dataset's quality. Lastly, the matched data are a subset of the evaluation dataset, which leads to the assumption that the unmatched data points would have a similar data quality compared to the matched data points.

Apart from its inherent limitations, extrinsic methods also have additional issues that have to be considered. These are the potential effects of bulk imports into datasets and the influence of the methods used to match points between the evaluation and reference datasets. Bulk imports refer to additions of geographic data from external sources into a dataset at a large scale. This was reported to occur in OSM and its effect on completeness was assessed for the case of road data in the U.S. [45]. Bulk imports were also raised as a possible reason for similar distributions in tags for OSM relations between cities [71]. However, the impact of POI bulk imports on positional and thematic accuracy is less understood and could be an area of future work. This would be especially relevant if the reference data chosen for extrinsic validation methods is the source of the bulk imports.

The results from extrinsic approaches that require matching corresponding points in D_{eval} and D_{ref} will be subject to the accuracy of the matching method used. Furthermore, matching procedures could make use of POI attributes like geographical coordinates, POI names and/or place types [26,27,40]. This could confound the interpretation of results when these attributes are used in approaches for validating POI data, making it difficult to ascertain if the observed results are due to artefacts of the matching procedure, or truly

reflect POI data quality. As it is not currently known if matching procedures definitively cause POI data sources to appear more or less accurate, isolating the influence of the matching procedure on the results of validating POI data via extrinsic means can be an area of future work.

Intrinsic methods, on the other hand, have the ability to reveal internal statistics of the dataset and allows for data quality analyses without the regional or financial limitations arising from obtaining reference data. It gives us a better understanding of what the dataset has, allowing us to infer the usability of the dataset. For example, by checking the proportion of missing attributes, a dataset with a large proportion of missing postal code information is less usable for postal services.

However, compared to extrinsic methods using reference datasets, intrinsic methods are less able to provide definitive evaluations of data quality [32]. The review of current POI validation approaches revealed that the majority of intrinsic methods currently require temporal data in the form of edit histories to analyze changes in POI data over time. However, with the exception of OSM, change logs for POI data are not easily accessible to researchers or external users. Therefore, evaluations of POI data using intrinsic methods are mostly restricted to industry players who manage or produce the datasets. Sharing these change logs would allow the wider geospatial research community and future users to conduct independent assessments of POI data quality with intrinsic methods. Further developing intrinsic methods that do not require temporal data or edit histories would also reduce reliance on such data.

5.3. Use-Cases

As previously mentioned in Section 2.6, the usability of a dataset relates to how well it fits user requirements, and evaluating this can involve a combination of approaches which are deemed to be suitable for the specific use case. At its minimum, geographical information should contain enough information to allow users to know what objects are located where. As such, it can be argued that positional and thematic accuracy should be adequately validated before POI data are used, regardless of the data's eventual use case.

Depending on the specific use cases of the data, attention should be paid to the respective aspects of POI data that users wish to gain insight from. For example, a mobility application running on POI data might require a high level of positional accuracy so that passengers and goods are transported to their intended destinations. In this case, analyzing the distribution of spatial error between corresponding points would be useful in validating spatial accuracy. On the other hand, researchers seeking to understand urban form using big data [72] might be more interested in the spatial distribution of points in an urban area. In this case, the tools of spatial statistics would be more relevant, such as by calculating NNis and analyzing the Cross-K function.

If POI data are used for urban planning and policy decisions, then more attention should be given to how representative the dataset is. For example, it was found that OSM data tends to be more developed in wealthier communities [44]. Basing policy decisions solely on POI data without an understanding of the inherent biases and the local context could exacerbate existing social inequalities. In such situations, more attention should be paid to the completeness of the dataset, where the proportion of points found in datasets under consideration can provide an indication of errors of commission or omission.

POI data has also been used for real estate valuations [5], real estate search websites and calculating walkability scores [73]. In these contexts, thematic accuracy in the POI's place type would be most critical, especially if there are errors in the place types of key amenities such as supermarkets or transit nodes. The discrete nature of place types might also cause the resulting model outputs (such as the real estate values or walkability scores) to be more sensitive to place type misclassifications, as compared to deviations in POI location.

In addition to the established use cases of POI data, the development of standardized quality metrics will have to keep pace with upcoming advancements in the field, such as

(1) the use of POI data in dense, high-rise urban environments and (2) POI unification. In the urban context, activities are increasingly conducted in vertical environments with multiple floors. For example, in a large retail mall, shoppers and delivery personnel will have to navigate within a 3D indoor environment to accomplish their objectives. Given that the vast majority of POI locations are currently represented in 2D space, it is important that vertical information—such as the floor and unit number, which are typically found in the address attribute of POI data—are present and accurate. This implies that until POI data are able to meaningfully represent points in 3-dimensional space, accuracy in address strings will have an increased importance in very dense, high-rise urban environments.

The field has begun to look towards the unification of POI data from various sources with the goal of obtaining a POI dataset that is more comprehensive and of a higher quality than the individual sources. Data validation approaches can play a key role in POI unification and data fusion, such as in developing a unification process that is sensitive to the varying qualities of each POI data source. A novel set of validation approaches might also be required to evaluate the quality of unified POI datasets to quantify improvements relative to the computational or algorithmic requirements of the unification process.

6. Conclusions

A review of POI validation methods led to the identification of 23 methods which were then categorized into the various elements of data quality. Methods that evaluated the positional accuracy of a dataset were the most common, while very few methods were found to evaluate logical consistency and usability. A selection of nine validation methods covering the aspects of completeness, positional and thematic accuracy was implemented to assess four real-world POI datasets (Google Maps, HERE Maps, OSM, and OneMap) within the study area of a town in Singapore.

The low proportion of points found in both the evaluated (D_{eval}) and reference (D_{ref}) datasets suggests errors of commission and omission. Positional accuracy was good across the evaluated datasets, with reasonably low levels of spatial error (<60 m), and no significant difference in relative clustering. Finally, thematic accuracy in POI names and place types were highly dependent on the types of metrics used, and key attributes such as place type and coordinates were present in all datasets.

The key issues in the development of standardized POI validation metrics can be broadly categorized into three areas. The first relates to the characteristics of the POI datasets, such as the availability of datasets for intrinsic and extrinsic methods, and the varying levels of detail across POI datasets. This is followed by the methods and processes employed to conduct the validation, such as the inherent differences between extrinsic and intrinsic methods, and the influence of matching procedures. Lastly, the selection of metrics should be tailored to the intended application of POI data and be adapted to suit specific needs.

This work demonstrates the real-world application of POI validation methods and guides the development of standardized POI validation metrics. As the applications of POI data continues to grow and become more tightly integrated with society, assuring POI data quality will be increasingly relevant. Therefore, improving the methods for assessing and improving POI data quality is essential to instill confidence in the use of POI data, ultimately harnessing its full potential.

Author Contributions: Conceptualization, Lynette Cheah; methodology, Lih Wei Yeow, Raymond Low, Lynette Cheah; software, Lih Wei Yeow, Raymond Low and Yu Xiang Tan; validation, all; formal analysis, all; investigation, all; resources, Lynette Cheah and Raymond Low; data curation, Lih Wei Yeow, Raymond Low and Yu Xiang Tan; writing—original draft preparation, Lih Wei Yeow, Raymond Low and Yu Xiang Tan; writing—review and editing, all; visualization, Lih Wei Yeow; supervision, Lynette Cheah; project administration, Lynette Cheah; funding acquisition, Lynette Cheah. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: OpenStreetMap data analyzed in this study are publicly available and can be found here: https://wiki.openstreetmap.org/wiki/Downloading_data (accessed on 19 May 2021) under the Open Database License. OpenStreetMap data © OpenStreetMap contributors. OneMap data was accessed on 2 September 2020 from the Singapore Land Authority which are made available under the terms of the Singapore Open Data Licence version 1.0 (accessed on 22 June 2021). Restrictions apply to the availability of data from Google and HERE Maps. Data are available at https://developers.google.com/maps/documentation/places/web-service/overview (accessed on 22 June 2021) and https://developer.here.com/develop/rest-apis (accessed on 19 May 2021) with the permission of Google and HERE Maps. Singapore Land Authority (SLA) Street Directory Premium Data and Points of Interest (POI) Data are licensed from SLA. © Singapore Land Authority. All Rights Reserved.

Acknowledgments: We thank Google and HERE for permitting the use of their POI data in the case study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

All of the methods applied in Section 3 were implemented with the Python programming language in Jupyter notebooks, which also contain further details of the methods. The notebooks can be found in a public repository at this link: https://doi.org/10.6084/m9.figshare.14839644 (accessed 25 June 2021).

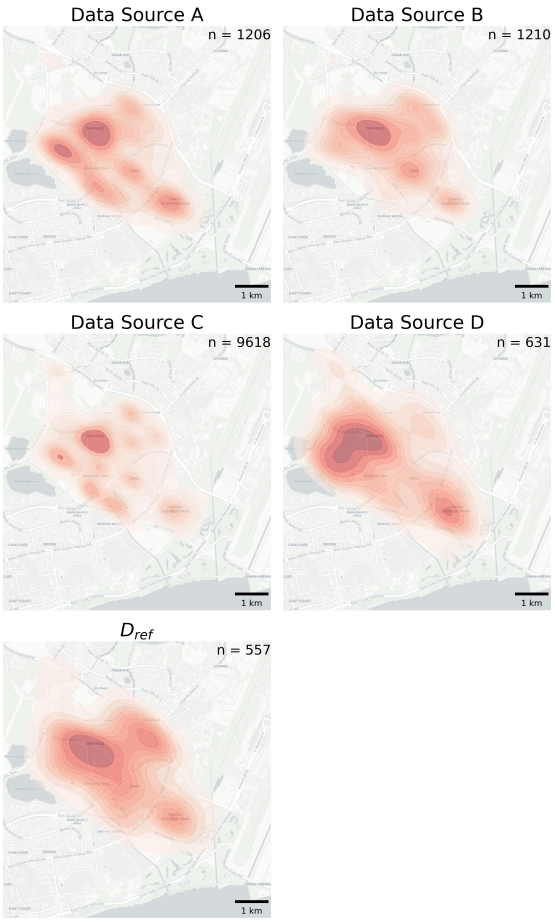

Figure A1. Heatmaps of POIs from each data source within study area. Numbers on top-right corner are number of POIs in data source. Map tiles by Carto, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

References

1. Miller, H.J.; Shaw, S.L. Geographic Information Systems for Transportation in the 21st Century. *Geogr. Compass* **2015**, *9*, 180–189. [CrossRef]
2. Tekler, Z.D.; Low, R.; Gunay, B.; Andersen, R.K.; Blessing, L. A scalable Bluetooth Low Energy approach to identify occupancy patterns and profiles in office spaces. *Build. Environ.* **2020**, *171*, 106681. [CrossRef]
3. Vhaduri, S.; Poellabauer, C.; Striegel, A.; Lizardo, O.; Hachen, D. Discovering places of interest using sensor data from smartphones and wearables. In Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; pp. 1–8. [CrossRef]
4. Touya, G.; Antoniou, V.; Olteanu-Raimond, A.M.; Van Damme, M.D. Assessing Crowdsourced POI Quality: Combining Methods Based on Reference Data, History, and Spatial Relations. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 80. [CrossRef]
5. Fu, Y.; Xiong, H.; Ge, Y.; Yao, Z.; Zheng, Y.; Zhou, Z.H. Exploiting geographic dependencies for real estate appraisal: A mutual perspective of ranking and clustering. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 1047–1056. [CrossRef]
6. Rodrigues, F.; Alves, A.; Polisciuc, E.; Jiang, S.; Ferreira, J.; Pereira, F. Estimating disaggregated employment size from points-of-interest and census data: From mining the web to model implementation and visualization. *Int. J. Adv. Intell. Syst.* **2013**, *6*, 41–52.
7. Liu, X.; Tian, Y.; Zhang, X.; Wan, Z. Identification of Urban Functional Regions in Chengdu Based on Taxi Trajectory Time Series Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 158. [CrossRef]
8. Gong, L.; Liu, X.; Wu, L.; Liu, Y. Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartogr. Geogr. Inf. Sci.* **2016**, *43*, 103–114. [CrossRef]
9. Low, R.; Cheah, L.; You, L. Commercial Vehicle Activity Prediction With Imbalanced Class Distribution Using a Hybrid Sampling and Gradient Boosting Approach. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 1401–1410. [CrossRef]
10. Klinkhardt, C.; Woerle, T.; Briem, L.; Heilig, M.; Kagerbauer, M.; Vortisch, P. Using OpenStreetMap as a Data Source for Attractiveness in Travel Demand Models. *Transp. Res. Rec.* **2021**, *2021*, 0361198121997415 [CrossRef]
11. Fonte, C.C.; Antoniou, V.; Bastin, L.; Estima, J.; Arsanjani, J.J.; Bayas, J.C.L.; See, L.; Vatseva, R. Assessing VGI data quality. In *Mapping and the Citizen Sensor*; Ubiquity Press: London, UK, 2017; pp. 137–163.
12. Bordogna, G.; Carrara, P.; Criscuolo, L.; Pepe, M.; Rampini, A. On predicting and improving the quality of Volunteer Geographic Information projects. *Int. J. Digit. Earth* **2016**, *9*, 134–155. [CrossRef]
13. See, L.; Mooney, P.; Foody, G.; Bastin, L.; Comber, A.; Estima, J.; Fritz, S.; Kerle, N.; Jiang, B.; Laakso, M.; et al. Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 55. [CrossRef]
14. Basiri, A.; Haklay, M.; Foody, G.; Mooney, P. Crowdsourced Geospatial Data Quality: Challenges and Future Directions. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 1588–1593. [CrossRef]
15. Fernandes, V.; Elias, E.; Zipf, A. Integration of Authoritative and Volunteered Geographic Information for Updating Urban Mapping: Challenges and Potentials. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 261–268. [CrossRef]
16. Antoniou, V.; Skopeliti, A. Measures and indicators of VGI quality: An overview. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2* 345–351. [CrossRef]
17. Senaratne, H.; Mobasheri, A.; Ali, A.L.; Capineri, C.; Haklay, M. A review of volunteered geographic information quality assessment methods. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 139–167. [CrossRef]
18. Barron, C.; Neis, P.; Zipf, A. A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Trans. GIS* **2014**, *18*, 877–895. [CrossRef]
19. Meek, S.; Jackson, M.J.; Leibovici, D.G. A flexible framework for assessing the quality of crowdsourced data. In Proceedings of the 17th AGILE International Conference on Geographic Information Science, Castellón, Spain, 3–6 June 2014.
20. Fan, H.; Zipf, A.; Fu, Q.; Neis, P. Quality assessment for building footprints data on OpenStreetMap. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 700–719. [CrossRef]
21. Gröchenig, S.; Brunauer, R.; Rehrl, K. Estimating Completeness of VGI Datasets by Analyzing Community Activity Over Time Periods. In *Connecting a Digital Europe Through Location and Place*; Huerta, J., Schade, S., Granell, C., Eds.; Lecture Notes in Geoinformation and Cartography; Springer: Cham, Switzerland, 2014; pp. 3–18. [CrossRef]
22. Ali, A.L.; Schmid, F. Data quality assurance for volunteered geographic information. In *International Conference on Geographic Information Science*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 126–141.
23. Ali, A.L.; Schmid, F.; Al-Salman, R.; Kauppinen, T. Ambiguity and plausibility: Managing classification quality in volunteered geographic information. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Dallas, TX, USA, 4–7 November 2014, pp. 143–152.
24. Kashian, A.; Rajabifard, A.; Richter, K.F.; Chen, Y. Automatic analysis of positional plausibility for points of interest in OpenStreetMap using coexistence patterns. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 1420–1443. [CrossRef]
25. Madubedube, A.; Coetzee, S.; Rautenbach, V. A Contributor-Focused Intrinsic Quality Assessment of OpenStreetMap in Mozambique Using Unsupervised Machine Learning. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 156. [CrossRef]

26. Deng, Y.; Luo, A.; Liu, J.; Wang, Y. Point of Interest Matching between Different Geospatial Datasets. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 435. [CrossRef]
27. Piech, M.; Smywinski-Pohl, A.; Marcjan, R.; Siwik, L. Towards Automatic Points of Interest Matching. *ISPRS Int. J. -Geo-Inf.* **2020**, *9*, 291. [CrossRef]
28. Hochmair, H.H.; Juhász, L.; Cvetojevic, S. Data Quality of Points of Interest in Selected Mapping and Social Media Platforms. In *Progress in Location Based Services 2018*; Lecture Notes in Geoinformation and Cartography; Kiefer, P., Huang, H., Van de Weghe, N., Raubal, M., Eds.; Springer: Cham, Switzerland, 2018; pp. 293–313. [CrossRef]
29. Zhang, L.; Pfoser, D. Using OpenStreetMap point-of-interest data to model urban change—A feasibility study. *PLoS ONE* **2019**, *14*, e0212606. [CrossRef]
30. Briem, L.; Heilig, M.; Klinkhardt, C.; Vortisch, P. Analyzing OpenStreetMap as data source for travel demand models A case study in Karlsruhe. *Transp. Res. Procedia* **2019**, *41*, 104–112. [CrossRef]
31. International Organization for Standardization. Geographic Information—Data Quality. Technical Report ISO 19157. 2013. Available online: https://www.iso.org/standard/32575.html (accessed on 23 May 2021).
32. Degrossi, L.C.; Porto de Albuquerque, J.; Santos Rocha, R.D.; Zipf, A. A taxonomy of quality assessment methods for volunteered and crowdsourced geographic information. *Trans. GIS* **2018**, *22*, 542–560. [CrossRef]
33. Ciepłuch, B.; Jacob, R.; Mooney, P.; Winstanley, A.C. Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. In Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resuorces and Enviromental Sciences, Leicester, UK, 20–23 July 2010; p. 337.
34. Girres, J.F.; Touya, G. Quality Assessment of the French OpenStreetMap Dataset. *Trans. GIS* **2010**, *14*, 435–459. [CrossRef]
35. Hochmair, H.; Zielstra, D. *Development and Completeness of Points Of Interest in Free and Proprietary Data Sets: A Florida Case Study*; Verlag der Österreichischen Akademie der Wissenschaften: Viena, Austria, 2013; Volume 1, pp. 39–48. [CrossRef]
36. Jackson, S.P.; Mullen, W.; Agouris, P.; Crooks, A.; Croitoru, A.; Stefanidis, A. Assessing Completeness and Spatial Error of Features in Volunteered Geographic Information. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 507–530. [CrossRef]
37. Zacharopoulou, D.; Skopeliti, A.; Nakos, B. Assessment and Visualization of OSM Consistency for European Cities. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 361. [CrossRef]
38. Ciepłuch, B.; Mooney, P.; Winstanley, A.C. Building Generic Quality Indicators for OpenStreetMap. In Proceeding of the 19th Annual GIS Research UK (GISRUK), Portsmouth, UK, 27–29 April 2011.
39. Mashhadi, A.; Quattrone, G.; Capra, L. The Impact of Society on Volunteered Geographic Information: The Case of OpenStreetMap. In *OpenStreetMap in GIScience: Experiences, Research and Applications*; Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., Eds.; Lecture Notes in Geoinformation and Cartography; Springer: Cham, Switzerland, 2015; pp. 125–141. [CrossRef]
40. Novack, T.; Peters, R.; Zipf, A. Graph-Based Matching of Points-of-Interest from Collaborative Geo-Datasets. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 117. [CrossRef]
41. Stark, H.J. Quality assessment of volunteered geographic information using open Web map services within OpenAddresses. In *Proceedings of Geospatial Crossroads@ GI_Forum*; Wichmann: Berlin/Heidelberg, Germany, 2011.
42. Mohammadi, N.; Malek, M. Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. *J. Spat. Sci.* **2015**, *60*, 119–135. [CrossRef]
43. De Tré, G.; Bronselaer, A.; Matthé, T.; Van de Weghe, N.; De Maeyer, P. Consistently Handling Geographical User Data. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*; Communications in Computer and Information Science; Hüllermeier, E., Kruse, R., Hoffmann, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 85–94. [CrossRef]
44. Haklay, M.M.; Basiouka, S.; Antoniou, V.; Ather, A. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *Cartogr. J.* **2010**, *47*, 315–322. [CrossRef]
45. Zielstra, D.; Hochmair, H.H.; Neis, P. Assessing the Effect of Data Imports on the Completeness of OpenStreetMap—A United States Case Study. *Trans. GIS* **2013**, *17*, 315–334. [CrossRef]
46. Fogliaroni, P.; D'Antonio, F.; Clementini, E. Data trustworthiness and user reputation as indicators of VGI quality. *Geo-Spat. Inf. Sci.* **2018**, *21*, 213–233. [CrossRef]
47. Kounadi, O. Assessing the quality of OpenStreetMap data. In *MSc Geographical Information Science*; Department of Civil, Environmental and Geomatic Engineering, University College of London: London, UK, 2009; p. 19.
48. Antoniou, V. User Generated Spatial Content: An Analysis of the Phenomenon and Its Challenges for Mapping Agencies. Ph.D. Thesis, University College London, London, UK, 2011.
49. Vandecasteele, A.; Devillers, R. Improving volunteered geographic data quality using semantic similarity measurements. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *2*, 143–148. [CrossRef]
50. Vandecasteele, A.; Devillers, R. Improving Volunteered Geographic Information Quality Using a Tag Recommender System: The Case of OpenStreetMap. In *OpenStreetMap in GIScience: Experiences, Research, and Applications*; Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., Eds.; Lecture Notes in Geoinformation and Cartography; Springer: Cham, Switzerland, 2015; pp. 59–80. [CrossRef]
51. Al-Bakri, M.; Fairbairn, D. Assessing similarity matching for possible integration of feature classifications of geospatial data from official and informal sources. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1437–1456. [CrossRef]
52. Sehra, S.S.; Singh, J.; Rai, H.S. Assessing OpenStreetMap Data Using Intrinsic Quality Indicators: An Extension to the QGIS Processing Toolbox. *Future Internet* **2017**, *9*, 15. [CrossRef]

53. Jonietz, D.; Zipf, A. Defining Fitness-for-Use for Crowdsourced Points of Interest (POI). *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 149. [CrossRef]
54. Master Plan 2014 Planning Area Boundary (No Sea). Available online: https://data.gov.sg/dataset/master-plan-2014-planning-area-boundary-no-sea (accessed on 13 May 2021).
55. Singstat. General Household Survey (GHS) 2015—Basic Demographic Characteristics. Available online: https://www.singstat.gov.sg/publications/ghs/ghs2015content (accessed on 13 May 2021).
56. OpenStreetMap Wiki. Downloading Data—OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/wiki/Downloading_data (accessed on 19 May 2021).
57. Haklay, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. Plan. Des.* **2010**, *37*, 682–703. [CrossRef]
58. Google Places. Available online: https://developers.google.com/maps/documentation/places/web-service/overview (accessed on 22 June 2021).
59. Google Maps 101: How We Map the World. Available online: https://www.blog.google/products/maps/google-maps-101-how-we-map-world/ (accessed on 19 May 2021).
60. Here Map Data. Available online: https://www.here.com/products/mapping/map-data (accessed on 19 May 2021).
61. HERE Map Rest APIs. Available online: https://developer.here.com/develop/rest-apis (accessed on 19 May 2021).
62. HERE Map Submit Feedback. Available online: https://developer.here.com/documentation/map-feedback/dev_guide/topics/quick-start-submit-feedback.html (accessed on 19 May 2021).
63. OneMap. Available online: https://www.onemap.gov.sg/docs (accessed on 22 June 2021).
64. OneMap Themes. Available online: https://www.onemap.gov.sg/docs/#themes (accessed on 22 June 2021).
65. Low, R.; Tekler, Z.D.; Cheah, L. An End-to-end Point of Interest (POI) Conflation Framework. *arXiv* **2021**, arXiv:cs.LG/2109.06073.
66. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
67. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
68. Dixon, P.M. Ripley's K function. In *Encyclopedia of Environmetrics*; El-Shaarawi, A.H., Walter, W.P., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2002; Volume 3, pp. 1796–1803.
69. Novack, T.; Peters, R.; Zipf, A. Graph-based strategies for matching points-of-interests from different VGI sources. In Proceedings of the 20th AGILE Conference, Wageningen, The Netherlands, 9–11 May 2017, pp. 9–12.
70. Tansalarak, N.; Claypool, K.T. QMatch—Using paths to match XML schemas. *Data Knowl. Eng.* **2007**, *60*, 260–282. [CrossRef]
71. Pruvost, H.; Mooney, P. Exploring Data Model Relations in OpenStreetMap. *Future Internet* **2017**, *9*, 70. [CrossRef]
72. Boeing, G. Spatial information and the legibility of urban form: Big data in urban morphology. *Int. J. Inf. Manag.* **2021**, *56*, 102013. [CrossRef]
73. Walk Score. Find Apartments for Rent and Rentals—Get Your Walk Score. Available online: https://www.walkscore.com/ (accessed on 13 April 2021).