International Journal of
Geo-Information

Article

Towards the Semantic Enrichment of Trajectories Using Spatial Data Infrastructures

Jarbas Nunes Vidal-Filho [1,2,*], Valéria Cesário Times [1], Jugurta Lisboa-Filho [3] and Chiara Renso [4]

1 Informatics Center (CIn), Federal University of Pernambuco (UFPE), Recife 50740-560, Pernambuco, Brazil; vct@cin.ufpe.br
2 Research, Innovation and Software Laboratory (LAPIS)—Federal Institute of Ceara (IFCE), Tabuleiro do Norte 62960-000, Ceará, Brazil
3 Informatics Department (DPI), Federal University of Viçosa (UFV), Viçosa 36570-900, Minas Gerais, Brazil; jugurta@ufv.br
4 Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Consiglio Nazionale delle Ricerche (CNR), 56127 Pisa, Italy; renso.chiara@gmail.com
* Correspondence: jnvf@cin.ufpe.br or jarbas.vidal@ifce.edu.br; Tel.: +55-88-99958-4035

Abstract: The term Semantic Trajectories of Moving Objects (STMO) corresponds to a sequence of spatial-temporal points with associated semantic information (for example, annotations about locations visited by the user or types of transportation used). However, the growth of Big Data generated by users, such as data produced by social networks or collected by an electronic equipment with embedded sensors, causes the STMO to require services and standards for enabling data documentation and ensuring the quality of STMOs. Spatial Data Infrastructures (SDI), on the other hand, provide a shared interoperable and integrated environment for data documentation. The main challenge is how to lead traditional SDIs to evolve to an STMO document due to the lack of specific metadata standards and services for semantic annotation. This paper presents a new concept of SDI for STMO, named SDI4Trajectory, which supports the documentation of different types of STMO—holistic trajectories, for example. The SDI4Trajectory allows us to propose semi-automatic and manual semantic enrichment processes, which are efficient in supporting semantic annotations and STMO documentation as well. These processes are hardly found in traditional SDIs and have been developed through Web and semantic micro-services. To validate the SDI4Trajectory, we used a dataset collected by voluntary users through the MyTracks application for the following purposes: (i) comparing the semi-automatic and manual semantic enrichment processes in the SDI4Trajectory; (ii) investigating the viability of the documentation processes carried out by the SDI4Trajectory, which was able to document all the collected trajectories.

Keywords: spatial data infrastructure; semantic annotation; semantic trajectories of moving objects

Citation: Vidal-Filho, J.N.; Times, V.C.; Lisboa-Filho, J.; Renso, C. Towards the Semantic Enrichment of Trajectories Using Spatial Data Infrastructures. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 825. https://doi.org/10.3390/ijgi10120825

Academic Editors: Michail Vaitis and Wolfgang Kainz

Received: 1 October 2021
Accepted: 3 December 2021
Published: 6 December 2021

1. Introduction

For many years, studies addressing Spatial Data Infrastructures (SDI) focused on storing, recovering, and sharing raw spatial data to assist user groups of various public and private organizations [1,2]. In contrast, the literature reports only a few discussions on the use of SDI to document and share Semantic Trajectories of Moving Objects (STMO) and its possibility to meet the demands of external applications. In addition, the analysis of some examples of the traditional and consolidated SDIs available, such as NSDI (NSDI: https://fgdc.gov/nsdi, accessed on 21 March 2021), INSPIRE (INSPIRE: https://inspire.ec.europa.eu, accessed on 21 March 2021), and MSDI (MSDI: https://msdi.data.gov.mt, accessed on 21 March 2021), revealed a lack of specific metadata standards and services to support STMO documentation. The architectural models of these SDIs focus on serving groups of users through data and services. For Alam et al. [3], due to the data heterogeneity and the spatial and temporal dependencies of the spatio-temporal data, extracting and analyzing semantic information is rather challenging, also due to the limitations of SDIs and web interface standards, among others. Therefore, traditional SDIs do not support the documentation of STMO for various reasons, including the lack of services specialized in the semantic enrichment of trajectories.

Currently, several projects provide raw data of trajectories to the international community —for example, the NYC Taxi and Limousine Commission (NYC Taxi & Limousine Commission: https://www1.nyc.gov/site/tlc/index.page, accessed on 21 March 2021) and Geolife (Geolife: https://www.microsoft.com/en-us/download/details.aspx?id=52367, accessed on 21 March 2021). These raw data of trajectories are designated as Raw Trajectories of Moving Objects (RTMO) and do not contain semantic annotations and nor metadata. The growing availability of RTMO imposes the need of analysis, semantical enrichment, and sharing with communities that are interested in their application [4–6]. These semantically enriched trajectories are named STMO and correspond to a sequence of spatial-temporal points with associated semantic information (e.g., information on the place visited, type of transport of the moving object, the aim of the movement etc.). Thus, it is important that SDI used for documenting STMO satisfy the specific technical requirements to support STMO documentation, such as the provision of semantic enrichment processes and metadata standards definition. In addition, the use of metadata standards to document STMO increases data quality [7,8].

The Geolife project dataset, for example, holds 17,621 raw trajectories collected by 78 users while using different modes of transportation. Such data have no semantic information, which hampers the extraction of unknown insights. Conversely, the enrichment, documentation, and storage of this dataset in SDI enable users and external applications to request services and STMO data, for example, to make inferences about places of traffic congestion or identify patterns in taxi drivers' routes. SDI services are essential for allowing users to document, search, analyze, retrieve, integrate, and share data. Semantic enrichment processes enable us to query the activities performed by the moving objects or places visited during the corresponding data collection. Using a metadata standard, an user specialized in documentation can catalog the STMO according to the technical and non-technical aspects of data collection [9]. Metadata include semantic information, algorithms used, adopted reference systems, among other components, favoring the availability of the documented STMO.

Another point that motivates the definition and construction of STMO-based SDIs is related to data heterogeneity generated by Big Data, social networks, and sensors. The growth of these areas allowed the collection of real-time data with associated semantics from moving objects. According to Petry et al. [10], the possibility of raw trajectories receiving semantic enrichment from multiple information sources in real-time promoted the new concept of Holistic Trajectories. This type of trajectory uses data analysis for real-time decision-making. For example, suppose that a user is taking a walk in an ecological park and decides to use their smartwatch to collect the trajectory and measure their heart rate and blood pressure. If the collected user's health parameters detect a cardiac problem in real time, the smartwatch can interface with SDI services to consult data and metadata on the location of nearby health units specialized in cardiovascular diseases. Therefore, SDI containing different data sources and location-based services can promote the fulfillment of the sharing needs of services and data of external users and applications.

SDIs are known as interoperable environments constituted of services and standards aimed at documenting and providing data through their metadata [11]. The goal of SDI development is to provide the storage, sharing, and access to data by means of a metadata standard [12]. These metadata allow the documentation of a set of raw trajectories [7]. SDI services are defined by the Open Geospatial Consortium (OGC), which has standardized the services for accessing maps, obtaining information from maps, among other functions [13,14]. STMO documented and provided in SDIs can be used to support applications of urban mobility analysis, benefit the identification of animal migration patterns, provide location-based services, assist recommendation services, among others. For example, SDI can contain a dataset of STMO referring to "Birds of Prey" and identify patterns of movement and feeding of birds from algorithms. Finally, this dataset containing avian STMOs can be documented with technical information from the data collection and semantic annotation process; therefore, users or applications can reuse it for further analysis and decision making regarding the bird movement pattern.

This paper introduces the SDI concept to the scope of STMO and demonstrate the evolution of SDI proposals to document different types of STMO. Thus, SDI for STMO can be defined as a set of users, professionals, technologies, standards, policies, data, metadata, Web services for semantic annotation, Maps Application Programming Interfaces (APIs), and different microservices. The use of a microservice-based architecture facilitates the scalability of Web services, communication between applications, and interface with APIs and Representational State Transfer (REST). The use of modern Web service interface standards facilitates the use of algorithms for complex data analysis and allow external applications and SDI services to interact [12]. Therefore, enabling semantic enrichment processes and documentation of STMO in SDIs is the main motivation of this research.

This study introduces the following main contributions: (i) definition of a new concept of SDI for STMO due to the lack of processes to perform semantic annotation of trajectories in SDI; (ii) creation of semantic enrichment processes for the documentation of STMO in SDIs; (iii) adaptation of a metadata standard to document STMO based on the ISO 19115-1:2014; (iv) use of a microservice-based architecture to facilitate the expansion of SDI concepts to meet the demands of external applications and provide a means of documenting STMO; (v) definition of an API REST for the semantic enrichment of trajectories, and (vi) assessment of the proposed SDI through experiments with data collected by voluntary users of the MyTracks app. Initially, this assessment was aimed at comparing semi-automatic and manual processes of semantic enrichment for the proposed STMO-based SDI to document a quantitative analysis of the trajectories.

All the remaining content of this paper is organized as follows. Section 2 discusses the basic concepts and some related works. Section 3 brings new definitions and a new SDI architecture for STMO data. Next, Section 4 presents SDI4Trajectory (SDI4Trajectory: http://sdi4trajectory.lapis.ifce.edu.br, accessed on 21 March 2021) which is an instance of the SDI architecture for the documentation of the STMO here proposed. Subsequently, Section 5 introduces a comparative study between the processes of semantic annotations provided by the SDI proposed. Section 6 presents discussions on the results and main problem. Finally, Section 7 reports our final remarks.

2. Basic Concepts and Related Works

Before outlining the research goals of this paper, we present a brief background on micro-services and SDI, providing a comparison between the proposed ideas and traditional SDIs, and indicating relevant related literature on semantic annotation and semantic SDIs.

2.1. Spatial Data Infrastructure

New policies for data sharing started to be created as the production of spatial information became an essential process for large organizations. An SDI consists of a set of people, technologies, standards, and policies whose goal is to share data, metadata, and geospatial services for the end users [15,16]. Figure 1 shows a simplified architectural model of traditional SDIs. A catalog of metadata is used to search for spatial data in a traditional SDI. The definition of metadata is data about data [17], for which many standards have been established: *Content Standards for Digital Geospatial Metadata* (CGSDM) from the Federal Geographic Data Committee (FGDC), regulations of ISO/TC 211:19115, ISO 19115-1:2014 or INSPIRE (ISO/TS 19139:2007). In Brazil, the metadata standard of MGB profile was created based on the norm ISO 19115-1:2014, described in a document from the Brazilian Institute of Geography and Statistics (IBGE) and the Geographic Service Directorate (DSG) of the Brazilian Army [18]. The importance of the metadata standard is to register and catalog datasets from different organizations, aiming at standardization, retrieval, and exploitation.

The definitions of [14,16,19] concerning traditional SDIs demonstrate their suitability to the context of raw trajectories since the data are collected directly from a piece of equipment and then shared and documented. In these SDIs, users access the services provided by the OGC and the available metadata standards to document the set of collected raw trajectories. Figure 2 illustrates a diagram of processes for documenting and providing raw trajectories together with their metadata. For example, if the raw trajectories of Geolife had been documented in traditional SDIs, initially the data would be selected and then documented based on a metadata standard. Finally, using a metadata catalog, the dataset can be made available with the corresponding metadata. However, the metadata refers to the technical information of the data production process, without the semantic information of the trajectories.

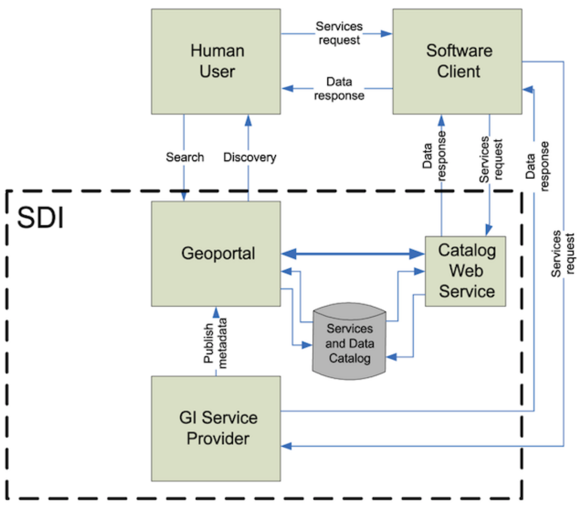

Figure 1. Architectural model of a traditional SDI. (Source: adapted Davis-Júnior and Alves [14]).

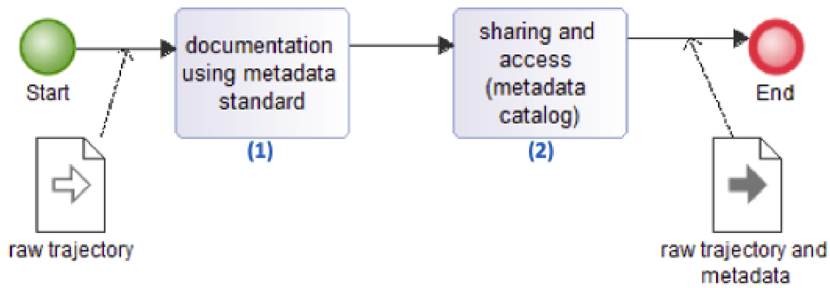

Figure 2. Documentation of trajectories in traditional SDI.

Crompvoets et al. [20] showed the benefits of national SDI developed in Europe for better control, performance, and decision making in the government scope. The SDI implementation within the INSPIRE project is the base of this analysis in an effort to provide data to the European community, in addition to strengthening SDI development. The focus of this European project is on addressing policies and activities that impact the environment. References [21–28] focus on the development and study of SDI in local, regional, and national contexts for control and management of natural disaster issues, land control, police emergencies or health-related applications, and sharing of road information that require spatial data for decision making.

The architectural models of traditional SDIs are now obsolete [29], mainly due to the use of a monolithic structure. With the emergence of Big Data and the possibility of analyzing large datasets, the Web service interface standards for the exchange of spatial data have been faced difficulties in using predefined libraries. As a result, OGC has been recently standardizing new services using the concept of REST API, which tends to improve the exchange of spatial data and contribute to the development of decentralized SDI architectures using microservices.

It is worth pointing out from the aforementioned definitions and works that current SDIs have not advanced towards a vision of SDI focused on communicating with external applications. For example, a mobile application for data collection may require the use of SDIs service to identify stop points or allow the own user to manually perform the trajectory semantic annotations. Such application enables to record the places visited by the user, which is named Point of Interest (POI) [30], Lines of Interest (LOI) [31], or Region of Interest (ROI) [31].

Most traditional SDIs are based purely on the services defined by the OGC, such as the INDE Brazil and MSDI-Malta, which are useful for sharing good quality data using consolidated metadata standards. INDE Brazil uses the ISO 19115:2014 standard, while MSDI-Malta follows the INSPIRE guidelines. Both projects tend to benefit from the new features of the OGC-API, allowing external applications to make requests to these SDIs. For example, both OGC API-Process and OGC API-Maps work as a service capable of retrieving the Minimum Bounding Boxes (MBR) of spatial objects. Finally, the application can overlay the MBR information with trajectory data and identify if the user has been at the locations represented by the MBR of the spatial objects. Therefore, it is convenient to adopt Web services standards based on REST for SDIs due to the facilitated insertion of new services for complex data analysis.

Other traditional SDIs use more modern interface standards, such as the INSPIRE that encourages the use of the OGC API and is based on the specification "Invoke Spatial Data Service" and SOAP protocol. However, since it is more advantageous for enabling the use of several data types, many enterprises have implemented Web Services in SDIs through REST API and micro-services; it is easy to apply, it provides good performance, and prevents points of failure in the system—for example, a failure in data visualization service will not hamper the use of other services by external users or applications.

2.2. Micro-Services and Micro-Service-Based SDI

The emergence of heterogeneous data has generated discussions about Web services used by SDI [12,29,32,33]. These services need to be inter operable, reusable, and light in order to facilitate the execution of complex data analysis required by SDI applications. Miranda [34] and Borba et al. [35] proposed adaptations to traditional SDI architectures for the insertion of Web services with voluntary collaboration, corresponding to the use of Volunteered Geographic Information (VGI) in SDI. VGI provides a means of collecting data produced by non-specialized users [36]. Bordogna et al. [37] present a SDI that follows OGC standards completely aiming to integrate time-series data derived from multiple information sources with VGI data to support applications in the agricultural sector. These SDIs use monolithic structures and Web interface standards that make the exchange of data among applications difficult.

In contrast, we have identified an innovative approach to SDI based on micro-services. Micro-services are small parts of software that work independently and execute a set of specific activities [38]. In turn, a set of micro-services is known as micro-service architecture. Assis et al. [12] propose and validate a micro-service-based SDI architecture aimed at analyzing and documenting data to assist in decision making in the context of deforestation areas. Nascimento et al. [29] propose resilient SDI architecture containing micro-services to improve the availability and reliability of resources and provide data using the Circuit Breaker pattern to improve service recovery. Figure 3 displays a micro-service-based architecture model for STMO documenting illustrating the processes of semi-automatic and manual semantic enrichment, trajectory catalog, and trajectory visualization as the minimum set of requirements for building an STMO-based SDI.

The use of micro-services as a Web interface standard in SDI semantic enrichment processes is important for strengthening discussions about SDIs focused on applications and allowing the use of APIs to provide interface with external applications. For example, a real-time data collection application about athletes' health may need to analyze geospatial data to direct the individuals to the nearest sports activity environment. Thus, SDI can provide data and services via a Web interface to assist applications in decision making. In addition, the micro-services available with API resources can facilitate STMO documentation and access to metadata. Since data documentation is usually performed manually, it is possible to use micro-services to automatically document STMO based on a metadata standard. Micro-services can also be implemented to verify the applicability of a dataset metadata, consequently reusing the proper data only in decision-making. This may reduce computational costs related to accesses to inappropriate data by an application. Therefore, highly available services that are light and automated favor integration with metadata standards and external applications.

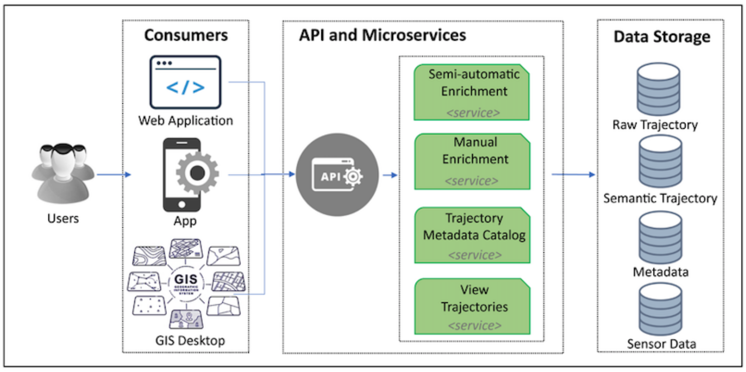

Figure 3. A System Architecture for STMO-enabled SDI.

The TerraBrasilis project (TerraBrasilis project: http://terrabrasilis.dpi.inpe.br, accessed on 25 June 2021) is the most recent example of the use of micro-services in SDI. Based on a hexagonal architecture model and using different micro-services, this SDI was developed to share, access, and analyze data on the deforestation of the Legal Amazon and Cerrado. This project is part of the Amazon Program, which collaborates with the Brazilian environmental monitoring. The advantages of this SDI are related to the use of a decentralized architecture and an inter operation provision between services by allowing less application failure points, analyzing data with complex algorithms, and improving the availability of Web services. Therefore, using micro-services in SDI applications has been proved promising and regarded as a technical contribution considering the disadvantage of the lack of automated data documentation.

Even with recent OGC efforts, only a few studies are reported in the literature that address SDIs and micro-services exclusively. Typically, the currently available SDIs data and services are not reused by applications. For example, location-based apps can use APIs to request the closest restaurant to the user, which occurs because the SDIs Web service interface standards are based on OGC definitions. According to Assis et al. [12], OGC standards have become obsolete for the new era of Big Data, which requires lighter and more flexible services for complex data analysis. Therefore, the use of micro-services strengthens the concept of STMO-based SDIs, since applications that use trajectory data need quality data and services available for spatial analysis and semantic inferences.

2.3. Semantic SDI and Semantic Annotation

Other works use a semantic approach to SDI for metadata specification, data integration, and information retrieval. Kalantari et al. [39] established a conceptual metadata project to document VGI, in which Linked Data are used to explore information from voluntary user's comments on the Web and assign semantics to metadata. Huang et al. [1]

discussed the problem related to the spatial data integration and semantic heterogeneity in traditional SDI environments, concluding that using Linked Data approaches in SDIs improves spatial data integration. This study uses Resource Description Framework (RDF) repositories applied to the SDI context to support the discussion. Based on an ontology, Janowicz et al. [40] and Fugazza et al. [41] investigated how to incorporate semantic information into SDI search engines to improve data retrieval. Henriques [42] defined a method of automatic semantic annotation of geographical data based on open data for information recovery in SDI. Wiemann and Bernard [2] used spatial data fusion techniques to establish new guidelines for data integration in SDI, in which OGC services are integrated with semantic Web standards.

Conversely, studies on raw trajectories have advanced towards the labeling of stop points with semantic information. Since traditional SDIs already have metadata standards available, metadata editors, metadata catalogs and services for enabling accesses to the documented data, Figure 2 shows a documentation process for raw trajectories. However, traditional SDIs lack a formal process for STMO documenting. In this context, we have identified the absence of new SDI definitions to assist STMO documentation. As a result, Figure 4 exhibits a diagram of processes containing steps for STMO documenting in SDI. As shown, before using metadata standards to document STMO, semantic annotation processes should be integrated with SDI to enable STMO documentation. Thus, we added step (1) to Figure 4 to represent the semantic enrichment processes in SDIs.

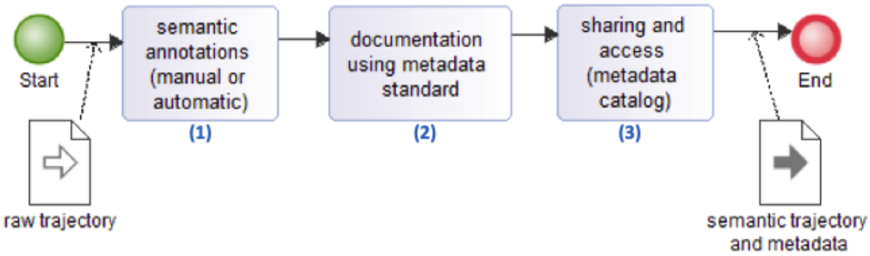

Figure 4. Diagram representing STMO documentation in SDI.

In general, a sequence of points $TR = < p_1, \ldots, p_n >$ is known as a raw trajectory [43], where $p_i = (x_i, y_i, t_i)$, $p_i \in TR$, x and y are coordinates to the object position in space and t_i is the timestamp of p_i collected from the moving object. STMO corresponds to a raw trajectory with semantic information associated with its stop points/moves [44]. According to Spaccapietra et al. [5], stop points are places where the object has remained for some time, and a move represent the displacement between two stop points. In the semantic annotation process (see example in Figure 5), initially, different computational techniques are used to identify slow parts (i.e., stop points) or regions with an increase in the speed (i.e., movements) of a raw trajectory. Such identification allows methods to be used to determine the semantic information to be used in the semantic annotation—for example, the set of POIs closest to a given stop point may be the chosen semantic information. Finally, the selected semantic attributes are incorporated into the parts of the raw trajectory, resulting in the corresponding STMO. Figure 5 presents an example of a semantic annotation process.

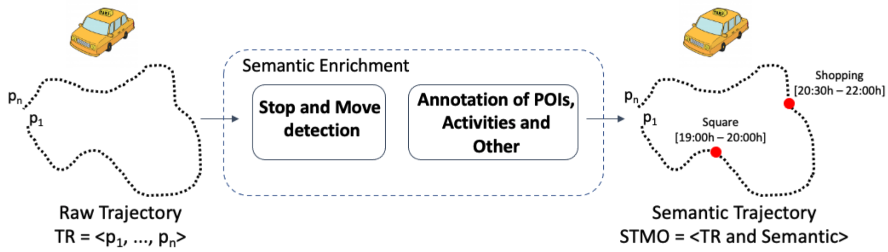

Figure 5. Example of semantic annotation process.

Several research works propose methods or algorithms for the implementation of semantic enrichment processes based on different approaches [31,42,45]. Alvares et al. [46] propose the semantic enrichment of trajectories using geographic information through the SMoT algorithm. Yan et al. [31] developed a framework capable of performing semantic annotations in raw trajectories. These trajectories are enriched with heterogeneous data, such as POIs, ROIs, and Satellite Images. Furletti et al. [30] present a method to infer human activities from raw trajectories. From the analyses of POIs associated with chosen stop points, activity inference is performed based on the Newton's gravitational model. Ruback et al. [47] propose a framework for the semantic enrichment of trajectories with Linked Open Data. Soares-Júnior et al. [48] developed a platform named VISTA for the manual annotation of trajectories applying methods to segment trajectories and assigning manual annotations to the segmented trajectories.

Considering the studies previously discussed, we understand that the efforts reported in the literature for the enrichment of raw trajectories are enormous [31,44,46,47]. Several data models for representing semantically enriched raw trajectories have been discussed as well [5,43]. However, we have identified a lack of approaches for STMO documentation using metadata standards, services, technical descriptions, and semantic information. There is also the lack of shared environments with Web micro-services, semantic enrichment processes, data, or metadata, which would enable the sharing and analysis of STMO data. Finally, no specific metadata standards are provided for semantic trajectories documentation, an issue to which studies/researchers have paid little attention.

The main motivation behind the use of an STMO-based SDI is that applications that collect trajectories can be integrated with sensor data analysis services of this SDI. These data collection applications can be based on the Internet of Things (IoT) technologies, which have equipment capable of controlling and monitoring environments [49]. OGC provides the Sensor Web Enablement (SWE) standard that allows developers to create applications for collecting and storing data derived from the Web by using sensors. This OGC standard provides services for communication with sensors, data collection, and the configuration of sensor parameters, in addition to allowing the creation of SDI solutions having an interface with IoT to provide documented datasets (e.g., temperature and humidity sensors) and use of SDI services for weather forecasting. Thus, for example, raw trajectories collected in real-time can reuse SDI data and services to infer the nearest lunch location.

Studies addressing STMO-based SDIs can benefit from research on a new type of STMO as well, designated as holistic trajectories, which are defined as a new type of STMO data with different semantic aspects and dimensions—for example, data extracted from sensors and social networks [32]. According to Ferrero et al. [33], a holistic trajectory is a sequence of elements $<e_1, e_2, \ldots, e_n>$ in which each element has the dimension x, y, t, Asp, where x and y represent the location of the moving object in space and time t (e.g., raw trajectory), and $Asp = (Asp_1, Asp_2, \ldots, Asp_m)$ is the set of semantic annotations, also called semantic aspects. This type of trajectory aims to represent the real movement of a moving object in displacement with different semantic annotations. For example, consider the trajectory of Paul, who moves from home to work. Paul may have an equipment at home that collects information on sleep quality, room temperature, climate, voice recognition patterns for feeling analysis. Consider also that, at Paul's working place, there are sensors for collecting data on noise, air pollution, humidity, and other types of data. For a given displacement performed by Paul, suppose that sensor data and raw trajectory are collected and the collected raw trajectory is enriched with data from sensors, activities and POIs, producing the corresponding holistic trajectory. Therefore, SDI can use OGC standards to document sensor data, make it available to be reused by external applications in real time and improve STMO semantic analyses.

3. A Novel Approach for Building STMO-Based SDIs: SDI4Trajectory

This section presents our approach to enable STMO documentation in SDI4Trajectory. Previously, we defined the concept of SDI4Trajectory, including new definitions for semantic annotation processes in SDI, and subsequently presented the metadata catalog and a set of fields defined from the ISO 19115-1:2014 standard, aimed at the sharing and standardization of STMO documentation. Finally, we detailed the semi-automatic and manual enrichment processes using diagrams and algorithms.

The construction of the SDI4Trajectory is based on the basic requirements of traditional SDI, in addition to contributions from the Semantic Web, standards of inter operable Web services used for data exchange, and the concepts of REST API and micro-services. Therefore, we highlight our four initial requirements: (1) semantic annotation processes; (2) definition of a set of metadata to document STMO; (3) definition of a metadata catalog; and (4) data visualization and sharing services. It is important to highlight that this paper focuses on presenting contributions to enable STMO documentation and offer convenience for new types of STMO documentation through the customization of the SDI4Trajectory's available services.

We identified that, before performing STMO documentation, SDIs need to provide semantic annotation processes to enable the semantic enrichment of STMOs. These processes can be conducted manually and semi-automatically by users or external applications. Definitions 1 and 2 address the new concepts of semantic enrichment in SDI. Metadata standard and catalog are essential to facilitate retrieval and attribute quality to data. Finally, visualization and sharing services favor the access to documented STMOs.

Definition 1. *(Manual semantic enrichment): The semantic enrichment performed by a user (manual) is a process that allows the manual creation of semantic annotations derived from the users' previous knowledge, using micro-services of the STMO-based architecture that is proposed for SDIs.*

The manual enrichment process allows users to interact with the SDI4Trajectory to manually record the visited POIs and the activities performed during the collection of the users' trajectory. In Sections 3.1.1 and 3.1.2, the Algorithms 1–3 present a proposal for validating the Definition 1.

Definition 2. *(Semi-automatic semantic enrichment): The semantic enrichment performed by the SDI (semi-automatic) is a micro-service of STMO-based architecture SDI that generates semantic annotations in an automatized way from the interaction of applications or users.*

In the semi-automatic enrichment process, the annotation of visited POIs and possible activities is performed semi-automatically as it requires the calibration of similarity parameters by users or external applications. In Sections 3.1.1 and 3.1.2, the Algorithms 1–4 aim to validate the Definition 2. By instantiating the semantic enrichment processes of Definitions 1 and 2, it is possible to document different types of STMO. The emergence of new types of STMO depends on the network interface used by equipment that collects raw trajectory and the web standards used by the applications. For example, an user who uses a smartwatch to collect raw trajectory and request information from sensors about rainy weather conditions is constructing a trajectory based on multiple heterogeneous aspects, that is, a so-called holistic trajectory. This user would require an equipment connected to the internet to provide an interface with Web standards for data exchange. Therefore, Definitions 1 and 2 can be used to instantiate the documentation of new STMO types, but these new semantic processes tend to require simpler, lighter, and more interoperable standard Web interfaces.

Algorithm 1 Compute Clusters

```
Input  : Raw Trajectory, TR = (p_1,...p_n), Clustering method (Cm), set of parameters for clustering (Sp = (s_1,...s_n))
Output : clusters, C = (c_1,...c_n)
clusters = new List < TrackPoint >;
TrackPoint = null;
for each point p IN TR do
    /* findStopPoints is the function that computes the clusters */
    cluster = findStopPoints (p, Cm, Sp);
    if p ∈ cluster then
        /* clusters is a list of points that belong to clusters */
        clusters.addAll(p);
    else
        /* clusters are lists of points that belong to clusters */
        moves.addAll(p);
return clusters;
```

Algorithm 2 Identification of Semantic Point

```
Input  : ClusterList, C = (c_1,...c_n) and SemanticPointMethod
Output : StopList, S = (s_1,...s_n)
StopList = new List < semanticpoint >;
semanticpoint = null;
for each cluster c IN clusters do
    if stopMethod == WEIGHTED then
        /* StopsDiscovering method to identify semantic point in clusters */
        semanticpoint = StopsDiscoveringWeightedAverage(c);
    else
        semanticpoint = StopsDiscoveringCentralPoint(c);
    StopList.addAll(semanticpoint);
return StopList;
```

The previous definitions demonstrate that processes of semantic enrichment are the main elements when constructing SDI for STMO for corresponding to the first task in the STMO documenting process (see step (1) of Figure 4). In contrast, we consider that when building SDI for STMO, besides the services for identifying stop points, semantic annotation, inference of activities, or inference of behaviors for moving objects, it is also necessary to provide other services such as documentation, sharing, and retrieval of STMO. These other services arise from the need to establish a quality reuse of STMO. For example, a public policy manager in the area of public safety may need to reuse STMO collected by police vehicles to implement improvements in the police patrol. Therefore, the reuse of STMOs by users or applications requires reliability, quality, and documentation with metadata standards for STMOs.

The API4Trajectory defined in Definition 3 is an API developed as a service provider of SDI4trajectory and holds many features. For example, it provides access to many APIs containing semantic information on climate, tourism, restaurants, and POIs. Additionally, it is accessible and used by other applications, accesses external bases for manipulation, allows to incorporate new algorithms for data mining, trajectory segmentation, machine learning, and big data processing, among other important characteristics.

Algorithm 3 Manual Semantic Enrichment

```
Input  : Coordinates of semantic point in the cluster (lat/lng), VisitedPOI[] and
         Maximum Walking Distance (MWD)
Output : STMO (semantic points and POIs)
RelatePOIsToSemanticPoint ();
poisToSemanticPoint = new Map ();
for each semanticpoint do
    /* maximum time taken to walk after leaving the car to consider a
       semantic point */
    SemanticPointMaxTimeInMillis = radius/this.WALK-AVG*1000;
    for each visitedPOI do
        distance = haversine(semanticpoint.coordinate, visitedPOI.location);
        visitedPOI.distanceToSemanticPoint = distance;
        if MWD >= distance then
            if visitedPOI.horary.start == undefined || visitedPOI.horary.end ==
             undefined then
                mostLikelyStop == poisToSemanticPoint.get(visitedPOI);
            if mostLikelyStop == null || distance < haversine
             (mostLikelySemanticPoint.coordinate, visitedPOI.location) then
                poisToSemantic.set(visitedPOI.semanticpoint);
        else
            startDateInMillis = new Date (visitedPOI.horary.start).getTime();
            endDateInMillis = new Data (visitedPOI.horary.end).getTime();
            startSemantiPointInstantMillis = new Date
             (semanticpoint.instant).getTime() + 3600 * 3 * 1000;
            if (startDateInMillis >= startSemanticPointInstantMillis) &
             (endDateInMillis <= endSemanticPointInstantInMillis) then
                SemanticPoint.visitedPOI.push (visitedPOI);
return STMO;
```

Definition 3. *(API4Trajectory): The API4Trajectory is an API that allows the visualization of STMO, determination of the stop points for moving objects, identification of semantic points in clusters, clustering using similarity parameters, extraction of data from social networks and sensors, data mining, and other methods for trajectory analyses.*

The SDI4Trajectory creates a great precedent for the new SDI to be developed. Naturally, the SDI4Trajectory can provide access to diverse applications for data analysis or collection, adopt new standards for documentation, and improve the inter operation between services and data. The definition of SDI4Trajectory proposed in this paper is given by Definition 4.

Definition 4. *(SDI4Trajectory): The SDI4Trajectory is an SDI that allows the combination of technologies, people, metadata, policies, and standards, in addition to integrating API4Trajectory with processes of trajectory semantic enrichment.*

With the use of micro-services and REST, SDI4Trajectory services can be customized or reused to document multi-aspect trajectories. REST also allows the provision of Web resources by URL and offers better communication with external applications, being able to communicate with IoT equipment. For example, consider an IoT environment that employs the Message Queuing Telemetry Transport (MQTT) protocol. The MQTT is the protocol that manages the publication and receive of messages from IoT applications. Thus, a REST

API may be integrated into this type of environment [50] easily. The OGC SWE standard already uses the MQTT protocol for publishing and receiving messages over its services.

Algorithm 4 Search POIs semi-automatic

```
Input  : Coordinates of semantic point in the cluster (lat/lng) and Maximum
         Walking Distance (MWD)
Output : Set of POIs and Activity
POIList = new List < POI >;
POI = null;
/* nearbySearchRequest is a function to search for POIs near the
   point semantic                                                    */
response = nearbySearchRequest.POI(SemanticPoint.getCoordinate());
results [] = addAll.response;
for each result IN results do
    resultCoord = new Coordinate (geometry.lat,geometry.lng);
    distanceSemanticPoint = Haversine.distance(semanticpoint, resultCoord);
  if distanceToSemantiPoint > MWD then
      BREAK;
  else
      POI = new Semi-AutomaticPOI (Id, name, coordinate, viewport, address);
      period = details.openingHours.POI;
      /* verifies if the POI is open at the time of the semantic
         point collection                                            */
      if OpenPOI.isOpen(period, semanticpoint.getInstant()) then
          POIList.add(POI);
/* GravityNewton is the function that returns the probable
   activity of the POIs                                              */
Activity = GravityNewton (POIList);
return Activity;
return POIList;
```

3.1. Semantic Enrichment of SDI4Trajectory

The semantic enrichment described in this section represents step (1) shown in Figure 4 (i.e., corresponding to the semantic annotation service of STMO documentation process). However, from the works described in Section 2, we understand that researchers have made little effort to incorporate semantic annotation services in SDIs. Our study attempts to fill such research gap by developing two proposals of semantic annotation processes for SDI4Trajectory—one is semi-automatic and the other manual. For validation purposes, these two processes use clustering algorithms and semantic point identification methods for spatial data. Initially, candidate stop points are identified and grouped into clusters. For each cluster, a semantic point is identified. Then, it becomes necessary to define the form of semantic enrichment (manual or semi-automatic). Finally, the STMO being documented is registered in the SDI.

Figure 6 illustrates the functioning of the semantic enrichment considered in this research. For example, assume that Smith uses the shared bike system from the city of Alpha to know touristic routes. Using his own smartphone, Smith collects raw trajectories of his ride and makes them available for semantic annotation. For the semantic enrichment of Smith's trajectories, this work considered four steps: (i) collection of raw trajectory data (Figure 6a); (ii) computation of clusters for the acquired trajectory (Figure 6b); (iii) identification of semantic points (Figure 6c); and (iv) execution of semantic annotations (Figure 6d). Each step represents part of the semantic enrichment proposed in this article. Finally, the steps are detailed below.

(a) Acquisition of raw trajectories
It refers to the trajectory acquisition process (Figure 6a). Smith has recorded his movement in Alpha city and made the trajectory available for documentation in SDI.

(b) Computation of clusters
The goal is to identify similarity patterns in the trajectory, possibly corresponding to the regions where Smith has performed stops. These regions are often characterized by places where the moving object slows down, which requires the use of a spatial clustering approach to classify the trajectory points based on the object speed value of each collected point, and subsequently to group these points into clusters. Next, the classification by similarity may be performed after calculating the instantaneous speed value of each point and verifying if it does not exceed the value of a predefined input speed parameter. Then, the user defines this speed parameter as a similarity criterion to be given as input to the clustering algorithm. Following this, the formation of candidate clusters for stop points is performed based on the similarity values between the calculated speed of each point and the input parameter value defined by the user that is seen as a speed threshold. After grouping the points by similarity, the algorithm returns the clusters $<C_1, \ldots, C_n>$ for the Smith's trajectory (Figure 6b). Each cluster has a set of points grouped by a pattern of similarity and is identified by circles in Figure 6b.

(c) Identification of semantic points
The main motivation behind the adoption of semantic points is given as follows: (1) Easy acquisition of geographic coordinates to retrieve POIs, activities, distances, and other semantic information; (2) The possibility of retrieving a larger set of POIs from APIs that provide the necessary place information; (3) The APIs that return place information require a pair of geographic coordinates as input; (4) The possibility of overlapping the semantic points with the bounding rectangle of the returned POIs to validate the stop point; (5) The possibility of adopting different semantic point identification methods for different means of transport because the determination of a semantic point may be affected by the chosen means of transport. However, the use of semantic points may not exclude other approaches (e.g., convex hull or bounding rectangle), since they may be seen as complementary solutions for adding semantics.

In this proposal, after computing the clusters, we introduce the concept of semantic points [51]. Instead of considering the cluster as a stop point, we use two methods of semantic point identification. This point is used to represent the stop point and is assigned to semantic annotations. The central point method identifies the cluster center as the semantic point, while the weighted average method, defined in [51], uses the weighted average concept to calculate the average of the geographic coordinates of the points that belong to a cluster, resulting in the definition of the semantic point. This point, in turn, is represented by new geographic coordinates generated from the weighted average calculation. Figure 6c shows the semantic points $<SP_1, ..., SP_n>$ generated for Smith's trajectory. Finally, a set of POIs is associated with each semantic point using the minimum walking distance informed by the user as follows. The distances between each semantic point and all probable POIs are computed. However, only POIs whose distances are less than or equal to the minimum walking distance are associated with a semantic point.

(d) Making semantic annotations
It aims to annotate semantic information (e.g., visited place or performed activities by a moving object) for the stop points of Smith's trajectory (see Figure 6d). For each semantic point, all POIs located within the buffer region created from the distance given as an input parameter are identified. Only the POIs closest to the semantic point are considered, out of which those containing the opening hours belonging to the cluster formation time interval are selected. The cluster formation time is calculated by the difference of the timestamp of two points belonging to the formed

cluster, considered as the last and first in this research. In addition, semantic points can receive annotations of activities likely to be performed by the user. The activity inference method used depends on the type of semantic enrichment process chosen by the user or application. If a manual enrichment is conducted, the user manually notes his activities by associating them to each stop point, while in semi-automatic enrichment, an adaptation of [30] may be used.

For inference of activities, a list is initially created containing different types of activities and respective POI categories. For example, the “food” activity may contain POIs from the “supermarket” and “restaurant” categories. Thus, the POIs previously associated with each semantic point are categorized by type of activity, according to the list previously created. The activity inference is based on the set of categorized POIs, as the distances between the semantic point and the selected POIs are obtained. For each set of categorized POI, we calculate the amount of existing POIs divided by the shortest distance between the POIs, then the result is multiplied by a normalization factor. Finally, the activity of the set of POIs categorized with the highest result after the normalization calculation is associated with the semantic point.

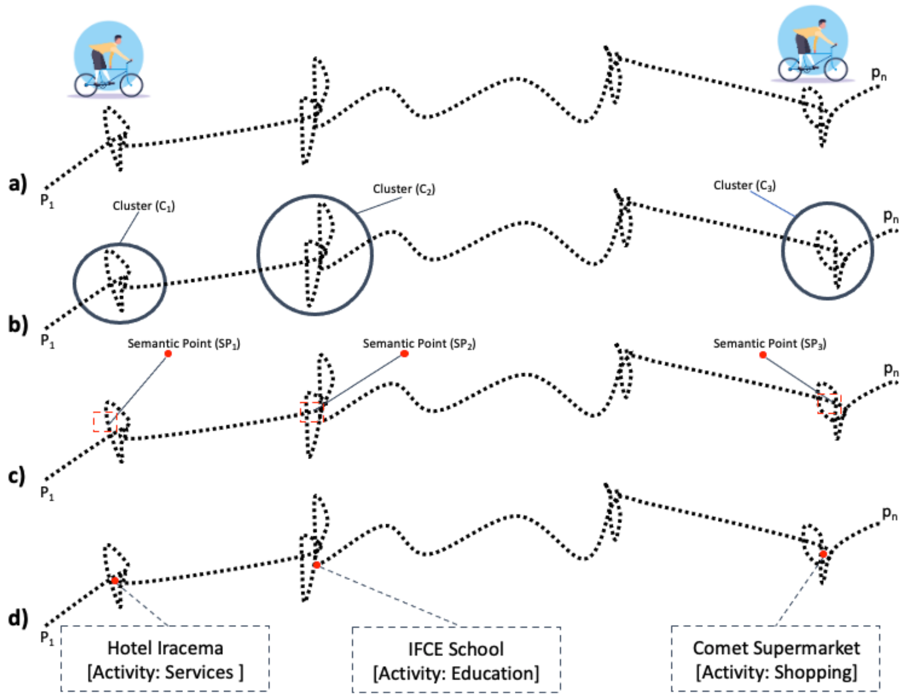

Figure 6. Semantic enrichment: (**a**) Raw trajectory example. (**b**) Computer clusters. (**c**) Semantic point identification. (**d**) Annotation of probable POI visited and activity.

The execution of the steps in Figure 6 ensures the STMO generation. The instantiating of the semantic enrichment process in SDI is described by Algorithm 5. Firstly, the clusters are computed as shown in line 2 of Algorithm 5, regarded as clusters of candidate stop points in a raw trajectory. Subsequently, a single semantic point is identified by Vidal-Filho et al. [51] in the cluster (line 3) using Algorithm 2, further detailed in this paper. The semantic point is useful for associating semantic information (e.g., a POI or ROI) with the clusters candidate stop points. Then, we propose two approaches for the semantic enrichment of trajectories through SDI (from line 4 to 7), namely semi-automatic and manual, described by Algorithms 3 and 4, respectively, which, in turn, are explained next. Finally, Algorithm 5 generates STMO containing metadata on the chosen semantic enrichment process (line 9) and carries out the STMO documentation process based on a metadata standard (line 11), such as the ISO 19115-1:2014, available in the Trajectory Metadata Catalog (TMC)—see steps 2 and 3 of Figure 4.

Algorithm 5 Semantic Enrichment of SDI4Trajectory

```
Input  : Raw Trajectory, TR = (tr_1,...tr_n), process and metadata_standard
Output : STMO and metadata
for each trajectory tr IN TR do
    ComputerCluster (trajectory);
    IdentificationSemanticPoint (clusters);
    if process == semi-automatic then
        SearchPOIsSemi-Automatic ();
    else
        AnnotationManual ();
    /* method of STMO and metadata generation                     */
    Generation (stmo_file);
    /* method of STMO documentation with metadata standard        */
    Documentation (stmo_file, metadata_standard);
return documented_stmo;
```

The objective is to supply both services (manual and semi-automatic) for users of the SDI and applications who seek to communicate with the API4Trajectory services. For example, a user specialized in data documentation can manually annotate the POIs to the stop points of the raw trajectory, later documenting the STMO. In contrast, the same user may need to document a dataset without semantic annotations, requiring semi-automatic services to automatically infer semantic information. Finally, it is valid to have both semantic annotation methods, since the manual process will possibly compensate for the lack of precision for some cases of the semi-automatic process.

In Sections 3.1.1 and 3.1.2, we present the proposals for the semantic annotation processes of Definitions 1 and 2 in this section. The proposals are grounded on concepts of data mining, APIs, parameters of time and distance, among other computational resources to generate processes of semantic enrichment in SDIs. These processes ensure the documentation of STMOs and are described and presented through algorithmic proposals/approaches at each step of the process. Therefore, the need for having semantic annotation processes in SDIs is the main motivation behind the proposal of new definitions for STMO-based SDIs.

3.1.1. Semi-Automatic Semantic Enrichment

The proposal of Semantic Enrichment without user support has the function of consuming raw trajectory data, subsequently performing the four steps shown in Figure 7. Finally, it aims to generate STMO to be submitted to documentation process.

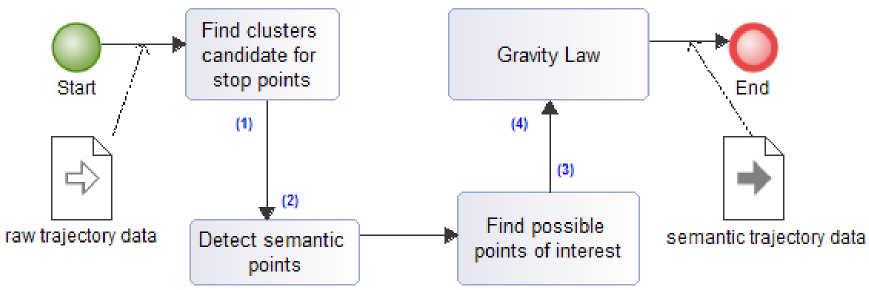

Figure 7. Process of semantic enrichment performed by the SDI.

Initially, the clusters of candidate stop points are detected in the raw trajectory (see step (1)—Figure 7). Cluster is a term used in statistics to classify elements that have similar characteristics [52]. A spatial clustering algorithm is used at this step (see Algorithm 1), since the objective is to group probable stop points with similar characteristics (e.g., velocity), subsequently choosing a single stop point in the cluster to be associated with semantic information (see Algorithm 2), called semantic point [51].

Algorithm 1 expresses that SDI4Trajectory provides a service to compute clusters, namely "clusters candidate stop points". Before using this service, the user is required to define a set of input parameters for Algorithm 1, which includes the chosen clustering method and its corresponding parameter values (e.g., known as thresholds) (line 4). These thresholds are necessary to find clusters with similar characteristics in the trajectory and are chosen based on the selected clustering method and user's expertise. They can be defined according to the transportation type used by the moving object.

Step (2) in Figure 7 involves the identification of a semantic point chosen from the cluster. After computing the clusters, it is necessary to define a single point in the cluster as a semantic point, since setting out from this point the closest POIs is searched using the functionalities of an API to search for localities and extract location data. Algorithm 2 is used by SDI4Trajectory to determine the semantic point. Initially, efforts have primarily focused only on defining a central point in the region of the cluster. However, the choice of the central point hardly selects a semantic point belonging to the cluster. Therefore, a second method proposed by Vidal-Filho et al. [51] based on the weighted average to identify a semantic point in the cluster was defined for SDI4Trajectory. Such a method consists of the ordering of points based on certain values of a relevant feature (e.g., points of low velocity are more weighted within a cluster of the trajectory). In any case, the user will be able to choose which method to use to identify the semantic point in the cluster, thus providing flexibility to the enrichment.

In step (3) of Figure 7, POIs are associated with the semantic point of each cluster, based on a minimum distance traveled (defined by the user). For each POI associated with the semantic point of the cluster, there is some information to be retrieved about the POI, such as name, coordinates, surrounding rectangle, category, opening hours, and the distance to the semantic point. This information is determined using an API to search for places and extract location data.

Algorithm 4 performs the semantic enrichment by inferring the probable POIs and possible activities for each possible semantic point. Algorithm 4 receives the geographic coordinate of each semantic point of the cluster as input parameter. Then, the location data extraction API returns the closest POIs to the semantic point (line 4).

The location data extraction API returns a set of POIs to be associated with the semantic point (lines 4 and 5). However, Algorithm 4 will only associate POIs with the semantic point: (1) if the distance between the semantic point and the POI is less than or equal to the maximum distance traveled—informed by the user (from line 9 to 16); (2) if the time at which the semantic point is collected belongs to the POI's functioning time interval (lines 15 and 16). After the execution of Algorithm 4, it is possible to have more than one POI associated with a semantic point. Initially, this is carried out by considering the minimum distance of a POI to the semantic point and verifying the opening time of the POI. Afterwards, the probable POI of a semantic point is identified also taking into consideration the probable activity of the POI in relation to the semantic point, as explained as follows.

Step (4) in Figure 7 seeks to infer activities based on Newton's gravity model, used by Furletti et al. [30]. The gravitational law defines how two bodies with differing masses stand concerning each other. In this paper, we adopt the solution proposed in [30] by adapting Newton's model to represent the relationship between a semantic point and a POI, thus determining the probable activity of the semantic point. Such adaptation is implemented by the method *GravityNewton*() of Algorithm 4 to infer the trajectory activities (line 18).

3.1.2. Manual Semantic Enrichment

The Manual Semantic Enrichment allows the inclusion of semantic annotations manually through a service available in SDI4Trajectory (see Figure 8). That is to say, the user can manually annotate POIs or activities by associating them directly with possible clusters of candidate stop points formed from the clustering algorithm adopted.

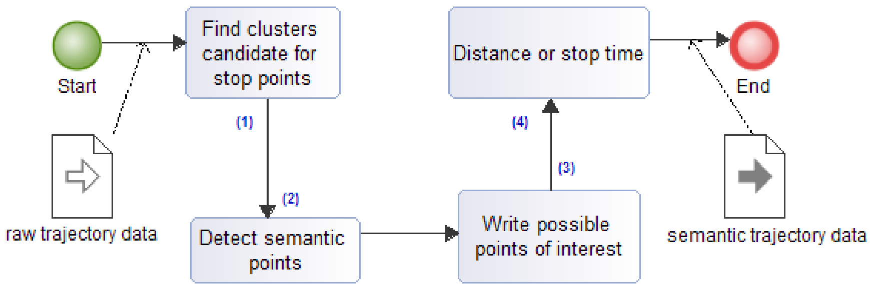

Figure 8. Semantic Enrichment performed by the user.

The approaches of steps 1 and 2 in Figure 8 are the same as steps 1 and 2 in Figure 7, respectively. Algorithm 3 represents steps 3 and 4 of the process described in Figure 8. In Step 3, the user registers the POI visited using a GUI provided by the SDI4Trajectory manual annotation service. In step 4, for each POI manually inserted, Algorithm 3 computes the shortest distance between the POI and the semantic point (line 6) and checks whether the instant of the semantic point collection belongs to the opening and closing time interval of the POI manually annotated (from line 8 to 10). In case it occurs, the POI will be assigned to the semantic point. Otherwise, if there is no record of the POI visiting hours, Algorithm 3 selects the POI with the shortest distance to the semantic point (from line 11 to 12) using the Haversine calculation formula by Alam et al. [53].

3.2. Trajectory Metadata Catalog

Trajectory Metadata Catalog (TMC) is part of the architectural components of the proposed SDI4Trajectory. The existence of the TMC results from the need to define a metadata standard and make STMOs available, ending the process described in steps 2 and 3 of Figure 4. These data need to be documented according to a metadata standard and provided through a metadata catalog. Many are the metadata standards available (e.g., MGB profile, ISO 19115-1:2014) that can be used by a specialist when documenting datasets. Therefore, after performing the semantic enrichment process, it is the expert's responsibility to choose the appropriate metadata standard for data documentation.

The data documenting process, that is, describing metadata from a set of data, still represents a bottleneck since the process is essentially manual. The central proposal of the TMC is to describe the STMO through a consolidated metadata standard. Then, the metadata is imported into an inter operable standard and cataloged in a metadata manager. Thus, we understand that STMO must follow the definitions of an ISO metadata standard, as it will follow international definitions for data inter operation. Finally, the use of semantic enrichment definitions in SDIs together with metadata standards and catalogs favor the construction of other SDI4Trajectory at local, regional, and national levels.

STMO documentation using SDIs requires a metadata standard with a reduced number of fields, based on ISO standards, easy to use, and support the quality information recording with respect to user trajectories. In this context, we have identified that the ISO 19115-1:2014 standard meets the requirements. Another important point for using the ISO 19115-1: 2014 standard is the elimination of mandatory fields when filling in the metadata. Table 1 shows a minimum set of fields (18 attributes) that are seen as important for documenting STMO.

Table 1 is organized as packages, attributes, and attribute descriptions. The attributes are organized by Packages: (1) Identification information (six attributes); (2) Dataset identification (three attributes); (3) Distribution information (two attributes); (4) Reference system information (two attributes); (5) Data quality (one attribute); and (6) Metadata information (four attributes). Each attribute has a filling description, which can later be performed manually or automatically—depending on the metadata field.

Table 1. Set of fields for STMO document.

Packages	Attributes	Descriptions
Identification information	Title	Name assigned to the shared resource
	Data creation date	Date of creation, publication or update of the resource
	Summary	Brief description about the data or dataset
	Metadata Status	Report the metadata status, whether it has been completed or is in progress
	Responsible for the resource	Information about the producer of the resource to be provided
	Keywords	Describe words that represent the resource
Dataset identification	Spatial representation	Inform the type of representation of geographic information
	Geographical extent	Inform on the spatial extent (MBR)
	Trajectory category	Inform whether it is a social network trajectory, traffic, etc
Distribution information	Distribution format	Inform the name of the data distribution format
	Resources for online access	Inform the address to access the resource
Reference system information	Reference systems	Information about projection and datum
	Coordinate systems	Inform the type of coordinates
Data quality	Lineage	Describes in detail the semantic enrichment process
Metadata information	Responsible for metadata	Information about the responsible for metadata
	Metadata standard	Insert the standard used
	Metadata Date	Insert metadata creation date
	Metadata profile	Inform the profile created for the trajectory

4. An Instance of the SDI4Trajectory Architecture

The SDI4Trajectory aims to propose a new approach for building SDIs capable of documenting and sharing different types of STMO by seeking to meet the functional requirements of first, second, and (mainly) third-generation SDIs. For this latest generation, SDI4Trajectory provides new standards for Web services to validate the new semantic enrichment processes (Definitions 1 and 2). The Java language (*Spring Boot based solution*) was applied to develop the available resources in the SDI4Trajectory. JavaScript Object Notation (JSON) was chosen as the file format that can contain the raw trajectory data and the semantic annotations (e.g., STMO, holistic trajectories).

4.1. Instrumenting the Proposed Semantic Enrichment Process

This section is dedicated to demonstrate the use of semantic enrichment processes and the metadata catalog to document and share STMO. To show how to utilize the services available in the SDI4Trajectory, we set the process instances of semantic enrichment available in the proposed SDI. In addition to introducing an example of STMO documentation through metadata, we describe some of the technologies used. The raw trajectory data used for the instrumentation of the proposed enrichment processes were collected from the MyTracks app.

The clustering algorithm used for grouping points of trajectories is the CB-SMoT proposed by Palma et al. [52], which was chosen due to being easy to understand and apply, in addition to being well defined in the literature. However, SDI can provide other clustering algorithms—for example, the use of medoid clustering Steinbach et al. [54]. The methods used to identify the semantic point were instantiated through the definitions of the weighted average and central point methods previously discussed.

In both semantic enrichment processes, users need to submit each trajectory to the SDI4Trajectory, and then define the parameters for the initialization of the semantic annotation process. The following parameters are required: choice of the semantic enrichment process to be used, CB-SMoT algorithm thresholds (e.g., *stop time*, *speed limit*, *average speed* and *maximum walking distance*), and the selection of identification method of the semantic point using the front-end of SDI4Trajectory (e.g., the central point method or weighted average method).

4.1.1. An Example of Semi-Automatic Semantic Enrichment

In this section, we exemplify the use of the SDI4Trajectory semi-automatic semantic enrichment (i.e., Definition 2). For this purpose, the weighted average method was used to identify the semantic point and the values of input parameters chosen for the CB-SMoT algorithm are: *stop time = 300 s*, *speed limit = 4 m/s*, *average speed = 3 m/s* and *maximum walking distance = 100 m*. Figure 9 presents an example of the STMO generated by executing this semi-automatic semantic enrichment. The STMO of Figure 9 contains three semantic points, at which two of them indicate the activity of shopping, represented by an icon shopping bag, while the activity of the third semantic point is related to food, represented by two cutleries.

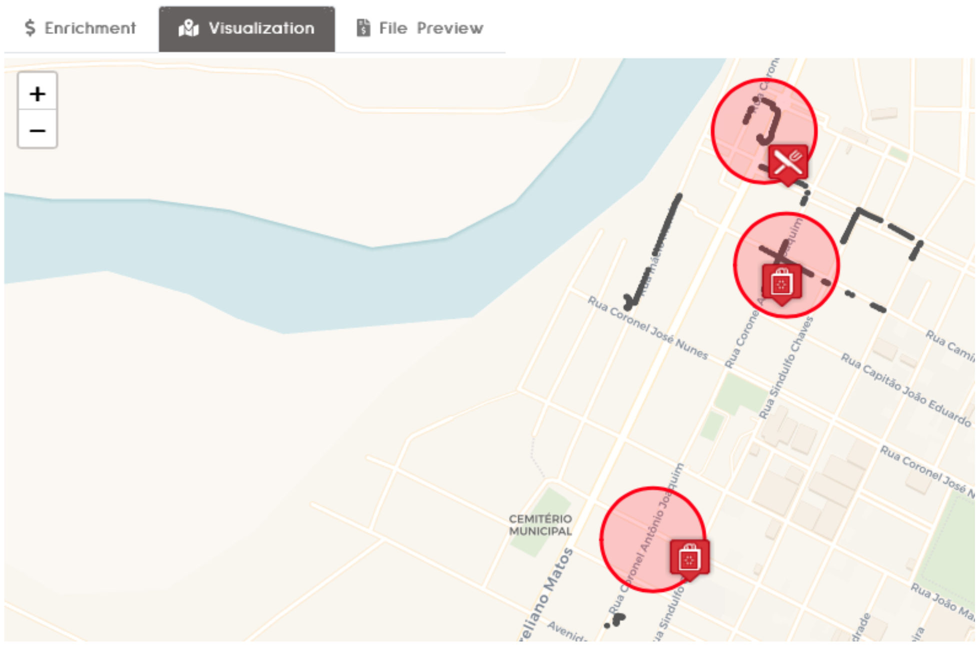

Figure 9. Interface after executing the semi-automatic semantic enrichment.

During this documentation process, the user may query the map and STMO annotations using the "Visualization" and "File Preview" functionalities, respectively (see Figure 9). If necessary, the user can adjust the parameters defined for semantic enrichment and, later, finalize with the STMO documentation in the SDI4Trajectory.

4.1.2. Example of Manual Semantic Enrichment

In this section, we provide an example of usage of the manual semantic enrichment (i.e., Definition 1). The parameters used to compute the clusters and identify the semantic points are the same as those used in the previous section. Then, the user manually registers in the buffer area created for clusters of candidate stop points (exhibited on the system interface), the possible visited POIs, and probable performed activities (Figure 10). The user can also manually inform the time interval elapsed during his visit to the POI. Such time information is used by Algorithm 3 to infer the probable POI. However, the user is not obliged to provide the exact time (optional) of the visit, as they may not recollect it.

This example presented in this section is based on the VISTA platform proposed by [48]. The authors used user-generated semantic annotation to identify trajectory segments. For this, the user performs manual semantic annotation (i.e., labels with semantic information) to the points of the raw trajectory referring to activities and means of transport. Finally, the proposed model performs inference of the trajectory segments.

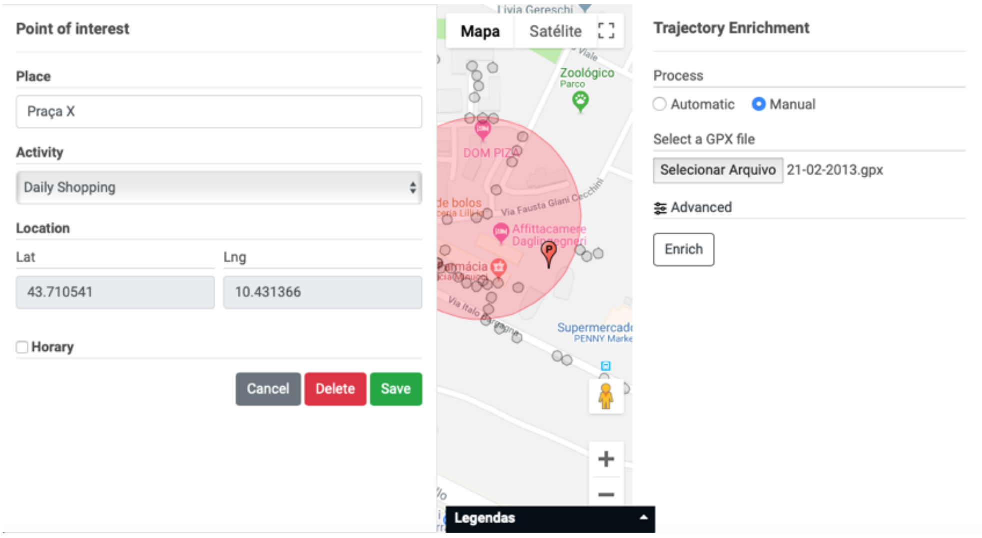

Figure 10. Interface of the Manual Semantic Enrichment.

4.2. Usefulness of the SDI4Trajectory

We believe that setting out from the creation and availability of the services of the SDI4Trajectory, SDI becomes a new environment to enrich, document and provide STMO data. To view and download a documented STMO, the user searches in the metadata catalog using a set of keywords. SDI4Trajectory provides services for the visualization of STMO and queries to obtain metadata—such as values of the input parameters of the adopted clustering algorithm. Other services can be found in the proposed SDI4Trajectory, like download, upload, and exportation of STMO.

Figure 11 illustrates parts of the STMO of Figure 9 that are now displayed in JSON format. Figure 11 contains data of the raw trajectory, identification of stop points, list of possible visited POIs, probable POI and activity of a STOP, in addition to metadata on the chosen clustering algorithms and details on the available methods for semantic enrichment. Every POI contains registration of name, category, opening hours, surrounding triangle, and other semantic information.

The results shown in Figure 11 were obtained from Algorithms 1–4 presented in this paper. Definition 2 was used to formalize the semantic enrichment process. These results can be improved by redefining the parameters of the enrichment used or by including new processes. Therefore, the formalization of semantic enrichment processes of trajectories for SDI allows SDI4Trajectory to be used as a shared repository to make STMO available. SDI4Trajectory can provide STMO and respective metadata, which can be combined and used by other systems to improve decision-making processes on the movement of moving objects.

After executing the semantic enrichment process, we used a set of fields extracted from the ISO 19115-1:2014 standard to document the STMO in Figure 11. The ISO 19115-1:2014 standard was adapted for two reasons: (1) it is a consolidated standard that enables the selection of fields required for dataset documentation; (2) it is simple to use as it does not contain mandatory fields. Figure 12 exhibits an example of a STMO documented by Semi-Automatic Semantic Enrichment of the proposed SDI4Trajectory, as stated by Definition 2. By way of the TMC, one can share STMO and metadata with information that adds quality to the data of moving objects.

```
{
  "metadata": {
      "process": "semi-automatic",
      "file": "2019-03-12 10_01_07.gpx",
      "semanticPointMethod": "Weighted",
      "stopMinimumTime": 300,
      "maxAverageSpeed": 3,
      "maxSpeedLimit": 4,
      "maxWalkingDistance": 100
  }
}
```


(**a**)


```
"stops":
{
  "estimatedPosition": {
    "lat": -5.153035590621075,
    "lng": -38.10183131385969
  },
  "estimatedInstant": "2019-03-12T13:06:29.951",
    "pointsOfInterest": [
        {
        "name": "Somar Calçados",
        "position": {
              "lat": -5.15296,
              "lng": -38.10162
        }
}
}
```


(**b**)


```
"activitiesProbabilities": {
    "SHOPPING": 79.24427466757128,
    "SERVICES": 16.3398413839807,
    "HEALTH_CARE_SERVICES": 4.4158839484480366
},
    "probablePOI": {
      "name": "Somar Calçados",
        "position": {
            "lat": -5.15296,
            "lng": -38.10162
        }
    }
```


(**c**)

Figure 11. Semantic Trajectory Example: (**a**) Metadata on the methods used for enrichment. (**b**) Set of identified stops. (**c**) Semantic information on activities and the POI likely to be visited.

Figure 12. Example of STMO documentation in the metadata catalog.

Initially, we had adopted the metadata catalog customization by Geonetwork; however, we have two types of metadata: (1) semantic metadata generated through semantic enrichment process; and (2) metadata related to the attributes of Table 1. The semantic enrichment metadata (see Figure 11) is stored together with the raw trajectory data in the

SDI4Trajectory database. These metadata refer to stops, semantic information, and descriptions of semantic enrichment methods, while the metadata of Table 1 are stored and managed by Geonetwork, which provides support for STMO data retrieval. The developed micro-services provide an interface for analyzing raw trajectories and querying the corresponding semantic metadata, as well as requesting available datasets via OGC API.

5. Experimental Usability Assessment

Two approaches were employed to assess the proposed use of the SDI4Trajectory: (i) a comparative study between the processes of semi-automatic and manual semantic enrichment, to validate the new definitions presented in this work; (ii) assessment to verify if the shared data could be documented by the proposed SDI. The testing scenario contains a dataset collected from the MyTracks app by a group of voluntary participants. The group consisted of teachers, students, lawyers and small business owners, among others. Figure 13 presents the profile of users and the quantitative data collected by car or by walking.

(**a**) (**b**)

Figure 13. (**a**) Exhibits the quantity of collected data and users' profiles. (**b**) Users' profiles.

The dataset applied to validate this proposal contains data from 15 trajectories (Figure 13a). The data were collected over three weeks. As illustrated in Figure 13b, ten users were invited to collaborate (academics, professionals, and others). However, only five users decided to participate, and performed work or leisure activities for this purpose. Thus, we established a comparison between the enrichment processes described in this paper. Users make the trajectories available together with manual annotations of the places likely to be visited. Based on this, it is important to recognize that users may have mistakenly annotated places. For example, the lawyer user manually noted that he visited a square, but in fact, he was at the city hall.

5.1. Comparing Manual and Semi-Automatic Semantic Enrichment Processes

Table 2 exhibits a set of annotated trajectories by semi-automatic and manual processes performed by volunteers. The idea is to compare the annotations of both processes to validate the micro-services available in the SDI4Trajectory. The trajectories were uploaded into the system and were enriched using semi-automatic and manual processes of semantic enrichment. The parameter values defined for the algorithm CB-SMoT are the same for all trajectories (e.g., *stop time = 300 s*, *speed limit = 4 m/s*, *average speed = 3 m/s*, and *maximum walking distance = 100 m*). After completing the semi-automatic enrichment, users were asked to register their manual notes using the SDI4Trajectory service. From this, the results of both enrichment processes were compared.

On each line, Table 2 exhibits trajectory identification, the number of clusters candidate stop points per trajectory, the quantity of possible POIs returned for each annotation process, the number of hits of semi-automatic annotation comparing with manual annotations, and the percentage of correct hits in carrying out the comparison. This is performed by comparing the quantity of POIs that were noted down in a semi-automatic and manual way for each trajectory.

Table 2. A Comparison between the Semantic Enrichment Processes.

Tracks	Clusters	Semi-Automatic	Manual	Number of Hits	Hits (%)
Track 01	3	3	2	2	100
Track 02	3	3	3	3	100
Track 03	2	2	2	2	100
Track 04	1	1	2	1	50
Track 05	3	3	3	2	67
Track 06	8	8	3	3	100
Track 07	4	4	2	2	100
Track 08	3	3	3	3	100
Track 09	4	4	2	2	100
Track 10	1	1	1	1	100
Track 11	3	3	3	3	100
Track 12	4	4	4	4	100
Track 13	1	1	1	1	100
Track 14	2	2	2	2	100
Track 15	2	2	2	2	100

The results from Table 2 correspond to a comparison of the two forms of semantic annotations, identifying the number of correct hits between each of them. After performing the semi-automatic semantic annotation, each trajectory contains several clusters where a probable POI is associated with each one. Afterwards, the voluntary users were asked to note down the POIs of each stop point for manual semantic annotation. Subsequently, we carried out comparative experiments from the semi-automatic and manual annotations.

In both the semantic annotation approaches considered, each cluster of candidate stop points receives a probable POI. Accordingly, a comparison was necessary to consider how many of the POIs, which were manually informed, identifiable in the semi-automatic process. For example, in Track 05 of Table 2, the semi-automatic process computed three stop-points and noted three POIs to the clusters of candidate stop points. However, the user noted down (manually) only three POIs for these stop points. Thus, we can conclude that, within a universe of three annotations, there was a correct hit of 67% (two out of three) in the semi-automatic annotations concerning the manual annotations. This represents that the semi-automatic annotation got two manually annotated POIs right. The values of the column Hits (%) of Table 2 are calculated using the Formula (1):

$$Hits(\%) = (Number of Hits/Manual) * 100. \quad (1)$$

The results in Table 2 demonstrate that the semi-automatic enrichment process is performing satisfactorily, which can be questioned, though, since the number of hits for some trajectories is below 100%. Nonetheless, this could arise due to the number of clusters formed, uncertainties in the identification of semantic points, or because of data provided by Google Places API for some trajectories.

We found that the difference in the amount of POIs annotated by the two semantic enrichment processes or the appearance of a POI identified in the semi-automatic process that is not described in the manual enrichment occurred for three reasons. The first may be due to the Google Places API being out of date, for example, Track 13 was manually annotated as "IFCE campus Jaguaruana", while the semi-automatic annotation returned "Multi-sport Gym Gerardo Correia Lima". However, by consulting Google Maps API, we see that the "Multi-sport Gym Gerardo Correia Lima" belongs to the "IFCE campus Jaguaruana". The second reason may be that the point of semantic location in the cluster is close to a larger quantity of POIs; consequently, the algorithm that returns a set of POIs with their respective activities (e.g., Algorithm 4) uses the gravitational model to infer the closest POI. Finally, the third refers to inconsistencies in manual annotations.

Therefore, none of the semantic enrichment processes should be discarded since the semantic annotations produced by both processes are associated with stop points physically located in the same region, which is demonstrated in Table 2, since from the POIs manually annotated (e.g., 35 probable POIs) in the set of STMOs analyzed, the semi-automatic enrichment correctly hit 89% (e.g., 31 probable POIs) of the stop locations of the voluntary user. A better calibration in the semantic enrichment algorithms is required to inform the probable POI more precisely.

5.2. Evaluation of STMO Documentation

In the STMO documentation process, all trajectories of Table 2 were submitted to the documentation process of SDI4Trajectory. For each trajectory, the metadata describes a set of information (e.g., algorithms used, semantic information, or technical information on the dataset). It is worth noting that even though we have used a small dataset, all collected trajectories have been documented correctly by the SDI4Trajectory.

Analyzing the ISO 19115-1:2014 standard, we realized that it is interesting to generate metadata on semantic enrichment (see Figure 11). These correspond to a large amount of data and can be filled in automatically, facilitating documentation. In addition, the metadata on the semantic enrichment attribute quality to the STMO, allowing the validation of the semantic annotation process used, as well as the carrying out of comparative studies with other methods. From the ISO 19115-1:2014 standard, we have also identified that 18 attributes are quite enough to document STMO (see Table 1) and that such standard does not require mandatory attributes. Finally, we observe that, from the 18 attributes used for STMO documentation (see Table 1), 14 of them (~80%) are filled automatically, while four attributes (~20%) are completed manually, which are: *Title, Summary, Keywords and Trajectory Category*.

From the results obtained with the documentation, we identified that a metadata proposal for STMO should prioritize semantic enrichment information and identification of the STMO dataset. However, we propose to use a reduced number of fields to reduce effort and minimize documentation problems. The use of the ISO standard is internationally recognized [7,8]. Therefore, the main findings reported in this paper follow the standards of ISO 19115:2014 for a spatial data set.

6. Discussions

This section presents discussions of this paper from the perspective of STMO documentation, semantic enrichment of trajectories using SDI, and benefits of micro-services for SDIs.

6.1. Providing SDI with Documented STMO

Investigating and evolving the concepts of traditional SDI to enable STMO documentation in SDI is the main driver behind this research. Documented STMOs have great value for different types of applications, for example, urban traffic planning and social networking location services. Through a set of STMO with annotations of POIs, it is possible to identify the places most visited by users, in addition to recommending them to other users. Traffic-related STMOs with semantic annotations can be useful to identify "hot spots" in urban traffic and search for alternatives to minimize flow in a given region.

The main motivation behind the adoption of a new concept of SDI for STMO documentation is the lack of processes to perform semantic annotation. We realized that the SDI scientific community has paid little attention to the adoption of SDI for the semantic enrichment of trajectories. After analyzing several SDI examples, we understood that most of them focus on being a data repository without associated semantic information. Also, the minimum requirements for STMO documentation in SDI are not discussed in the current SDI literature. Thus, semantic enrichment processes are part of the SDI4Trajectory to support STMO documentation. The use of these processes make semantically enriched trajectories available in SDI4Trajectory.

We have studied different metadata standards. However, we found a lack of specific standard for STMO documentation. In this work, we have adapted the ISO 19115-1:2014 standard for assisting in the documentation of STMO. This standard has been chosen for the following reasons: (1) it follows the ISO international standard; (2) it is up to date; (3) it does not require mandatory fields; and (4) it allows specific fields to be chosen to document different types of datasets. In this context, we have defined a minimum set of fields from this ISO standard to document STMOs (see Table 1). The proposal of this minimal set aims to make the documentation process easier and faster since it usually occurs manually. Furthermore, the aforementioned standard has a large set of attributes that can be adapted for documenting new STMO types, if necessary.

Metadata fields have been selected based on criteria such as the existence of the attribute in other standards, ease of automatic completion, quality attribution to the dataset, and support for applications that require data and metadata. This last criterion is interesting, as these specific attributes can be accessed via micro-services to retrieve data. For example, imagine that an application based on location services needs to retrieve STMO from taxis to analyze traffic at a certain time of day. Instead of the application performing cloud processing to access and process the data, it can directly request the metadata catalog (e.g., using OGC API). Searching for metadata on a dataset avoids application overhead, especially in the Big Data era, where large amounts of data are processed for decision making. Therefore, it is possible to recover the necessary data only for the final processing.

6.2. Semantic Enrichment in SDI

Concerning the state-of-the-art on trajectory semantic enrichment, we have identified a lack of the use of SDI for STMO enrichment or documentation. We have attempted to fill this research gap by developing two semantic annotation processes for SDI4Trajectory—one is semi-automatic and the other manual. These two proposals use the concepts of clusters and semantic points to identify stop points in raw trajectories. The formation of clusters is based on similarity parameters related to speed and time, while the semantic point is the spatial representation of a single point in the cluster. The technical implementation of the semantic point approach is considered in this paper to identify the closest point of the trajectory to the exact location visited by the moving object. Finally, the semantic point can be assigned to semantic annotations either manually or semi-automatically.

The choice of the semantic point is one of the challenges of this research. The difficulties in collecting voluntary data hampered a further exploration of the semantic enrichment processes. During the experiments, we realized that the chosen semantic point identification method can significantly change the most probable POI or affect the inference of activities. We observed that the same semi-automatically enriched trajectory may have notes of different activities, depending on the semantic point identification method used. This is due to the change in the physical location of the semantic point, as each method tends to return different geographic coordinates for the semantic points. For example, the weighted average method indicated 90% accuracy for "food" activity in Track 01 of Table 2, while the center point method indicated 60% accuracy for "services" activity of the same trajectory, and using the same thresholds of clustering. Therefore, we identified the need to explore the methods using other datasets, for example, Geolife. These new analyses will allow us to infer which semantic point identification method should be used based on the means of transport (e.g., bike, bus, and car).

Although some trajectories have different semantic annotations between the semantic enrichment processes, this does not imply rejecting the proposed methods. These differences occur due to various factors, such as the quantity of POIs returned by the Google

Places API and the method used to identification of semantic points, among many others. The available place in APIs have many limitations for returning semantic information. Therefore, alternatives should be considered such as using the services of the OGC API to reuse semantic information from other datasets or SDI, exploring new services and data sources for semantic enrichment.

6.3. Benefits of Adopting Microservices in SDI

The use of micro-service-based architectures is regarded as a technical contribution to SDI based on STMO. We highlight the ease of retrieving data and metadata through the use of micro-services, REST, and other standards. These new technologies allow external applications to make requests directly to the SDI, allowing SDI to move from being a data server for users to provide services and data to external applications and their users. Traditional SDI architectures offer centralized services that are difficult to reuse because the Web standards used for data exchange are not updated. In traditional SDI, the environment becomes prone to failures since, if the server fails, the entire system will be offline. This does not occur with the use of micro-services, as micro-service-based SDIs use a decentralized architecture. For example, if the semantic annotation services fails, the TMC becomes available. A recent work on the OGC API has employed the concept of decentralized architecture in SDI. Therefore, this will lead to a software layer in SDI with simpler and more reusable Web interface standards.

This paper proposes the adoption of STMOs that require little complexity in Web interface standards for semantic annotation. However, it is worth noting that other STMOs can be documented based on the customization of the services implemented in the SDI4Trajectory. For example, the multi-aspect trajectory is a type of raw trajectory enriched with semantic information from different data sources and formats (i.e., data heterogeneity). This type of trajectory requires medium and high-complexity Web interface standards. Thus, the acquisition of the raw trajectory will continue to exist, but the form of semantic enrichment defines new types of STMO. Therefore, exploring the use of micro-services in SDI is important to enable the documentation of STMOs that require heterogeneous data for analysis and decision making.

We believe that moving objects using Internet-connected equipment to collect raw trajectories and sensors data may require new Web services standards from SDI to document new types of STMO. Currently, data documentation is practically manual, therefore, a micro-service to allow automatic documentation would be interesting. This micro-service can be an important contribution to SDI as it automates metadata documentation by requesting attributes from different data sources. Therefore, we realized that the SDI4Trajectory is in the right direction and can be used to document new STMO types or be expanded with new OGC API standards.

7. Conclusions and Further Studies

This paper presents a new definition for documenting and enriching trajectories using SDI. The system architecture proposal presented is based on micro-services and favors the scaling and integration of the SDI, stimulating the development of SDI for STMO. We define a new approach to enable the use of SDIs to enrich trajectories, both automatically and manually. The architectural proposal was instantiated to generate solutions and validations. The results indicated that STMO documentation in the SDI4Trajectory is enabled by the incorporation of semantic processes. In addition, the metadata fields defined from the ISO standard are also able to document STMO.

We presented a comparison between the manual and semi-automatic processes of enrichment, which performs a satisfactory validation concerning the semi-automatic enrichment implemented for SDI. However, it is worth noting that it is necessary to improve the algorithm calibration or perform queries using new APIs for the extraction of location data. The results presented are satisfactory for the validation of the semantic enrichment processes, since almost all trajectories had 100% of correct answers of the probable POIs manually noted by the user. These results can be further explored and improved to previously indicate in the metadata which is the best semantic point identification method to be used for a set of STMO collected using GPS and car as a means of transport, for example.

Documentation of STMOs was carried out manually. All STMOs were documented through the ISO 19115-1:2014 standard, but not all metadata fields were filled out due to excessive information or lack of knowledge in filling out the field. This is a limitation for the SDI environment and has generated many efforts by the person in charge of providing the STMO. This context generates new demands for the SDI4Trajectory as it requires new definitions for the creation of micro-services specialized in STMO documentation, either automatically or semi-automatically. In addition, it highlights the importance of specifying a metadata core focused on STMO documentation, which is simple, easy to use, based on ISO 19115: 2014, and integrated with the micro-service.

This paper also encourages new changes in the use of SDI to document and analyze different types of STMO. In addition, it proposes an environment that can avoid major processing, as the use of micro-services favors prior queries of metadata about data before reusing it properly in the cloud. This new concept allows the integration with OGC and communication with technologies of Industry 4.0 aimed at performing complex analyses of spatial data.

For further studies, we intend to analyze methods for the detection of semantic points in clusters and expand the comparative analysis of semantic enrichment processes using the consolidated dataset of the Geolife project. This dataset will also be used to validate STMO documentation processes in SDIs and validate the algorithms proposed in this research. It is also intended to investigate the use of ontology to describe metadata of semantic trajectories aiming the selection of datasets. Additionally, it would be interesting to propose new methods for data collection in holistic or real-time trajectories, which are integrated with the SDI4Trajectory. Another interesting research goal would be to define a specific metadata core to document different STMOs and validation using automated micro-services, as well as to investigate new approaches to document different sets of STMO in SDI4Trajectory, such as the use of the OGC API to perform semantic annotations on semantic trajectories.

Author Contributions: Conceptualization, Jarbas Nunes Vidal-Filho; Methodology, Jarbas Nunes Vidal-Filho, Jugurta Lisboa-Filho and Valéria Cesário Times; Software, Jarbas Nunes Vidal-Filho; Writing—original draft, Jarbas Nunes Vidal-Filho; Writing—review and editing, Jugurta Lisboa-Filho and Valéria Cesário Times; Supervision, Chiara Renso. All authors have read and agreed to the published version of the manuscript.

Funding: This research received external funding from the Informatics Center (CIn) of the Federal University of Pernambuco (UFPE).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this research were collected by volunteer users using the MyTracks application. This data can be extracted from: https://github.com/lapislabifce/voluntarydata.git, accessed on 28 September 2020.

Acknowledgments: SDI4Trajectory is an institutional project aimed at enriching, analyzing and documenting Semantic Trajectories of Mobile Objects. The authors would like to thank the National Research Council (Cnr) in Pisa/Italy and the Foundation for the Support of Science and Technology of Pernambuco (FACEPE) in Recife/Brazil for supporting the initial development of this research.

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

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Programming Interfaces
CONCAR	National Commission of Cartography
CSDGM	Content Standards for Digital Geospatial Metadata
FGDC	Federal Geographic Data Committee
INDE	National Spatial Data Infrastructure of Brazil
INSPIRE	Infrastructures for Spatial Information in Europe
IoT	Internet of Things
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
LOI	Lines of Interest
MBR	Minimum Bounding Boxes
MGB	Brazil Geospatial Metadata Profile
MQTT	Message Queuing Telemetry Transport
MSDI	Malta Spatial Data Infrastructure
NSDI	National Spatial Data Infrastructure of EUA
OGC	Open Geospatial Consortium
POI	Point of Interest
RDF	Resource Description Framework
REST	Representational State Transfer
ROI	Region of Interest
RTMO	Raw Trajectories of Moving Objects
SDI	Spatial Data Infrastructures
STMO	Semantic Trajectories of Moving Objects
SWE	Sensor Web Enablement
TMC	Trajectory Metadata Catalog
VGI	Volunteered Geographic Information

References

1. Huang, W.; Raza, S.A.; Mirzov, O.; Harrie, L. Assessment and benchmarking of spatially enabled RDF stores for the next generation of spatial data infrastructure. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 310. [CrossRef]
2. Wiemann, S.; Bernard, L. Spatial data fusion in spatial data infrastructures using linked data. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 613–636. [CrossRef]
3. Alam, M.M.; Torgo, L.; Bifet, A. A Survey on Spatio-temporal Data Analytics Systems. *ACM Comput. Surv.* **2021**, *1*, 44.
4. Wu, T.; Qin, J.; Wan, Y. TOST: A topological semantic model for GPS trajectories inside road networks. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 410. [CrossRef]
5. Spaccapietra, S.; Parent, C.; Damiani, M.L.; Macedo, J.A.; Porto, F. Vangenot, C. A conceptual view on trajectories. *Data Knowl. Eng.* **2008**, *65*, 126–146. [CrossRef]
6. Ribeiro de Almeida, D.; de Souza Baptista, C.; Gomes de Andrade, F.; Soares, A. A survey on big data for trajectory analytics. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 88. [CrossRef]
7. Kalantari, M.; Syahrudin, S.; Rajabifard, A.; Subagyo, H.; Hubbard, H. Spatial Metadata Usability Evaluation. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 463. [CrossRef]
8. Brodeur, J.; Coetzee, S.; Danko, D.; Garcia, S.; Hjelmager, J. Geographic information metadata—An outlook from the international standardization perspective. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 280. [CrossRef]
9. Quarati, A.; De Martino, M.; Rosim, S. Geospatial Open Data Usage and Metadata Quality. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 30. [CrossRef]
10. Petry, L.M.; Ferrero, C.A.; Alvares, L.O.; Renso, C.; Bogorny, V. Towards semantic-aware multiple-aspect trajectory similarity measuring. *Trans. GIS* **2019**, *23*, 960–975. [CrossRef]
11. Rajabifard, A.; Feeney, M.E.F.; Williamson, I.P. Future directions for SDI development. *Int. J. Appl. Earth Obs. Geoinf.* **2002**, *4*, 11–22. [CrossRef]
12. Assis, F.G.L.F.; Ferreira, K.R.; Vinhas, L.; Maurano, L.; Almeida, C.; Carvalho, A.; Rodrigues, J.; Maciel, A.; Camargo, C. TerraBrasilis: A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 513. [CrossRef]
13. Lisboa-Filho, J.; Vegi, L.; Souza, W.; Lamas, J. Uma Infraestrutura de Dados Espaciais para o projeto Geominas com metadados definidos no perfil MGB da INDE. *Rev. Bras. De Cartogr.* **2013**, *65*, 123–138.
14. Davis-Júnior, C.A.; Alves, L.L. Infraestrutura de Dados Espaciais: Potencial para uso local. *Rev. Informática Pública* **2006**, *8*, 65–80.

15. Crompvoets, J.; Bregt, A.; Rajabifard, A.; Williamson, I. Assessing the worldwide developments of national spatial data clearinghouses. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 665–689. [CrossRef]
16. Rajabifard, A. A Spatial Data Infrastructure for a Spatially Enabled Government and Society. In *A Multi-View Framework to Assess Spatial Data Infrastructures*; Crompvoets, J., Fernandez, T., Eds.; Digital Print Centre: Melbourne, Australia, 2008; pp. 71–90.
17. Nogueras-Iso, J.; Zarazaga-Soria, F.J.; Muro-Medrano, P.R. Spatial Data Infrastructures and related concepts. In *Geographic Information Metadata for Spatial Data Infrastructures*; Nogueras-Iso, J., Zarazaga-Soria, F.J., Muro-Medrano, P.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1–29.
18. IBGE.; DSG. *Perfil de Metadados Geoespaciais do Brasil—Perfil MGB 2.0*; IBGE: Rio de Janeiro, Brazil, 2021; pp. 1–106.
19. Hjemager, J.; Moellering, H.; Cooper, A.; Delgado, T.; Rajabifard, A.; Rapant, P.; Martynenki, A. An initial formal model for spatial data infrastructures. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 1295–1309. [CrossRef]
20. Crompvoets, J.; Vancauwenberghe, G.; Ho, S.; Masser, I.; Vries, W.T. Governance of national spatial data infrastructures in Europe. *Int. J. Spat. Data Infrastructures Res.* **2018**, *13*, 253–285.
21. Putra, T.Y.D.; Sekimoto, Y.; Shibasaki, R. Toward the Evolution of National Spatial Data Infrastructure Development in Indonesia. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 1–27. [CrossRef]
22. Scholten, H.; Fruijter, S.; Dilo, A.; Borkulo, V.E. Spatial Data Infrastructure for emergency response in Netherlands. In *Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters*; Nayak, S., Zlatanova, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 179–197.
23. Asante, K.O.; Verdin, J.P.; Crane, M.P.; Tokar, S.A.; Rowland, J. The Role of spatial data infrastructure in the management of natural disasters. In Proceedings of the 9th International Conference for the Global Spatial Data Infrastructure, Santiago, Chile, 6–10 November 2006; ACM: Santiago, Chile, 2006.
24. Ajmar, A.; Perez, F.; Terzo, O. WFP spatial data infrastructure (SDI) implementation in support of emergency management. In *XXI the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; ISPRS—International Society of Photogrammetry and Remote Sensing: Beijing, China, 2008; pp. 1097–1103.
25. Giuliani, G.; Peduzzi, P. The PREVIEW Global Risk Data Platform: A geoportal to serve and share global data on risk to natural hazards. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 53–66. [CrossRef]
26. Snoeren, G.; Zlatanova, S.; Crompvoets, J.; Scholten, H. Spatial Data Infrastructure for Emergency Management: The View of the Users. In Proceedings of the 3rd International Symposium on Gi4DM, Toronto, ON, Canada, 23–25 May 2007; pp. 1–12.
27. Manfré, L.A.; Hirata, E.; Silva, J.; Shinohara, E.J.; Giannotti, M.A. An analysis of geospatial technologies for risk and natural disaster management. *ISPRS Int. J. Geo-Inf.* **2012**, *1*, 166–185. [CrossRef]
28. Zwirowicz-Rutkowska, A.; Michalik, A. The use of spatial data infrastructure in environmental management: an example from the spatial planning practice in Poland. *Environ. Manag.* **2016**, *58*, 619–635. [CrossRef]
29. Nascimento, H.L.; de Souza Baptista, C.; de Andrade, F.G.; dos Santos, L.C. Towards a Resilient Spatial Data Infrastructure. In Proceedings of the XXI GEOINFO, São José dos Campos, Brazil, 30 November–3 December 2020; INPE: São José dos Campos, Brazil, 2020.
30. Furletti, B.; Cintia, P.; Renso, C.; Spinsanti, L. Inferring human activities from GPS tracks. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, Chicago, IL, USA, 11 August 2013; ACM: Chicago, IL, USA, 2013; pp. 1–8.
31. Yan, Z.; Chakraborty, D.; Parent, C.; Spaccapietra, S.; Aberer, K. SeMiTri: A framework for semantic annotation of heterogeneous trajectories. In Proceedings of the 14th International Conference on Extending Database Technology, Uppsala, Sweden, 21–24 March 2011; ACM: Uppsala, Sweden, 2011; pp. 259–270.
32. Mello, R.D.S.; Bogorny, V.; Alvares, L.O.; Santana, L.H.Z.; Ferrero, A.A.; Renso, C. MASTER: A multiple aspect view on trajectories. *Trans. GIS* **2019**, *23*, 804–822. [CrossRef]
33. Ferrero, C.A.; Petry, L.M.; Alvares, L.O.; da Silva, C.L.; Zalewski, W.; Bogorny, V. MasterMovelets: Discovering heterogeneous movelets for multiple aspect trajectory classification. *Data Min. Knowl. Discov.* **2020**, *34*, 652–680. [CrossRef]
34. Miranda, T.S. Uma Arquitetura Para Contribuição Geográfica Voluntária em Infraestruturas de Dados Espaciais: Um Estudo de Caso no Munícipio de viçOsa. Ph.D. Thesis, Universidade Federal de Viçosa (UFV), Viçosa, Brazil, 2010.
35. Borba, R.L.R.; Strauch, J.C.M.; Souza, J.M.; Coleman, D.J. Uma proposta para a nova geração de infraestrutura de dados espaciais. *Rev. Bras. De Cartogr.* **2015**, *67*, 1145–1166.
36. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]
37. Bordogna, G.; Kliment, T.; Frigerio, L.; Brivio, P.A.; Crema, A.; Stroppiana, D.; Boschetti, M.; Sterlacchini, S. A Spatial Data Infrastructure Integrating Multisource Heterogeneous Geospatial Data and Time Series: A Study Case in Agriculture. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 73. [CrossRef]
38. Jamshidi, P.; Pahl, C.; Mendonça, N.C.; Lewis, J.; Tilkov, S. Microservices: The journey so far and challenges ahead. *IEEE Softw.* **2018**, *35*, 24–35. [CrossRef]
39. Kalantari, M.; Rajabifard, A.; Olfat, H.; Williamson, I. Geospatial metadata 2.0—An approach for volunteered geographic information. *Comput. Environ. Urban Syst.* **2014**, *48*, 35–48. [CrossRef]
40. Janowicz, K.; Schade, S.; Broring, A.; Kleber, C.; Maué, P.; Stasch, C. Semantic Enablement for Spatial Data Infrastructures. *Trans. GIS* **2010**, *14*, 111–129. [CrossRef]

41. Fugazza, C.; Pepe, M.; Pavesi, F.; Carrara, P. A Holistic, Semantics-aware Approach to Spatial Data Infrastructures. In Proceedings of the 3rd International Conference on Data Management Technologies and Applications, Vienna, Austria, 29–31 August 2014; SCITEPRESS: Vienna, Austria, 2014; pp. 349–356.
42. Henriques, H.B. Anotação Automática de Dados Geográficos Baseada em Bancos de Dados Abertos e Interligados. Ph.D. Thesis, Universidade Federal de Campina Grande (UFPB), Campina Grande, Brazil, 2015.
43. Bogorny, V.; Renso, C.; Aquino, A.R.; Lucas, S. F.and Alvares, L.O. Constant—A conceptual data model for semantic trajectories of moving objects. *Trans. GIS* **2014**, *18*, 66–88. [CrossRef]
44. Albanna, B.H.; Moawad, I.F.; Moussa, S.M.; Sakr, M.A. Semantic Trajectories: A Survey from Modeling to Application. In Proceedings of the Information Fusion and Geographic Information Systems (IFGIS' 2015), Grenoble, France, 18–20 May 2015; Springer: Cham, Switzerland, 2015; pp. 56–76.
45. Parent, C.; Spaccapietra, S.; Renso, C.; Andrienko, G.; Andrienko, N.; Bogorny, V.; Yan, Z. Semantic trajectories modeling and analysis. *ACM Comput. Surv.* **2013**, *45*, 1–42. [CrossRef]
46. Alvares, L.O.; Bogorny, V.; Kuijpers, B.; Macedo, J.A.F.; Moelans, B.; Vaisman, A. A model for enriching trajectories with semantic geographical information. In Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems, Washington, DC, USA, 7–11 November 2007; Samet, H., Ed.; ACM: New York, NY, USA, 2007; pp. 1–8.
47. Ruback, L.; Casanova, M.A.; Raffaeta, A.; Renso, C.; Vidal, V. Enriching mobility data with linked open data. In Proceedings of the 20th International Database Engineering & Applications Symposium, Montreal, QC, Canada, 11–13 July 2016; ACM: Montreal, QC, Canada, 2016; pp. 173–182.
48. Soares-Júnior, A.; Rose, J.; Etemad, M.; Renso, C.; Matwin, S. VISTA: A visual analytics platform for semantic annotation of trajectories. In Proceedings of the International Conference on Extended DataBase Technology (EDBT), Lisbon, Portugal, 26–29 March 2019; OpenProceedings.org: Lisbon, Portugal, 2019; pp. 570–573.
49. Mena, M.; Corral, A.; Iribarne, L.; Criado, J. A progressive web application based on microservices combining geospatial data and the internet of things. *IEEE Access* **2019**, *7*, 104577–104590. [CrossRef]
50. Collina, M.; Corazza, G.E.; Vanelli-Coralli, A. Introducing the QEST broker: Scaling the IoT by bridging MQTT and REST. In Proceedings of the 23rd International Symposium on Personal, Indoor and Mobile Radio Communications-(PIMRC), Sydney, Australia, 9–12 September 2012; IEEE: Sydney, Australia, 2012; pp. 36–41.
51. Vidal-Filho, J.N.; Times, V.C.; Lisboa-Filho, J. Towards the Identification of Semantic Points in Trajectories of Moving Objects with Weighted Averages. In Proceedings of the XXI GEOINFO, São José dos Campos, Brazil, 30 November–3 December 2020; INPE: São José dos Campos, Brazil, 2020.
52. Palma, A.T.; Bogorny, V.; Kuijpers, B.; Alvares, L.O. A clustering-based approach for discovering interesting places in trajectories. In Proceedings of the 23rd Symposium on Applied Computing, Ceará, Brazil, 16–20 March 2008; ACM: Fortaleza, Brazil, 2008; pp. 863–868.
53. Alam, C.N.; Manaf, K.; Atmadja, A.R.; Aurum, D.K. Implementation of haversine formula for counting event visitor in the radius based on Android application. In Proceedings of the 4th International Conference on Cyber and IT Service Management, Bandung, Indonesi, 26–27 April 2016; IEEE: Bandung, Indonesi, 2016; pp. 1–6.
54. Steinbach, M.; Kumar, V.; Tan, P. Cluster analysis: Basic concepts and algorithms. In *Introduction to Data Mining*; Tan, P.-N., Michael Steinbach, V.K., Eds.; Pearson Addison Wesley: Boston, MA, USA, 2005; pp. 487–559.