Journal of the Brazilian Computer Society, 2023, 29:1, doi: 10.5753/jbcs.2023.2458

Context Elements Taxonomy for Intelligent Transportation Systems

Amirton Chagas [**Federal University of Pernambuco** | ***abc@cin.ufpe.br***]
Carlos Ferraz [**Federal University of Pernambuco** | ***cagf@cin.ufpe.br***]

Federal University of Pernambuco - Informatics Center - Av. Jorn. Aníbal Fernandes, s/n - Cidade Universitária, Recife - PE, 50740-560, Brazil.

Received: 21 January 2022 • **Accepted:** 21 April 2023 • **Published:** 02 May 2023

Abstract Design and development of context-aware Intelligent Transportation Systems (ITS) are not trivial due to the large number of possible context elements that may be relevant to the application and the lack of structured information to guide system designers in this task. This paper proposes that context elements with common characteristics can be grouped into categories, and these categories can be organized in a taxonomy. This taxonomy could help system designers with the task of modeling and developing new context-aware ITS. We performed a literature review of 68 articles describing 70 ITS applications with context-aware features to identify context elements used in this type of application. Furthermore, we also analyzed three commercial ITS applications. We used data collected from the analysis of these 73 projects to define the categories and identify their relationships. We propose a taxonomy with 79 categories, with 57 leaf categories (a category without children subcategories). We also performed two experiments to validate whether the exposure to this taxonomy could improve the quality of an ITS application during its design, with favorable results showing a 2.7 times increase in the average amount of relevant context elements used in the application. Finally, we compiled a knowledge base of which context element categories are used in the 73 analyzed projects. It is another companion information that can be used to help system designers. The proposed taxonomy of context element categories organizes the information of the context-aware ITS domain in a way that can ease the task of designing such systems and improve the usage of context-aware features. The overall methodology used in this work to create the taxonomy for the ITS domain could be applied to other popular domains of context-aware applications.

Keywords: Software engineering, Intelligent Transportation Systems, Context awareness, Taxonomy, Knowledge management

1 Introduction

Vehicles are central pieces of modern life, thus, many cities have grown around and had their current shape defined due to motorized transportation. In the past, cars were mechanical machines, but the electronic revolution has embedded microcontrollers even in the simplest vehicles. Now, the driving activity is assisted by dozens of onboard computers and millions of lines of code [Dibaei *et al.*, 2020].

Until recently, in-car computers operated individually, merely exchanging information with each other, with these computers being mainly used for critical issues. Now, sensor networks deployed in cars collect machine-related data to provide feedback to drivers or even actuate when life-threatening events for vehicle occupants happen.

The trend became connected vehicles, where computers support from driving activities to passenger's well-being and entertainment. Atzori *et al.* [2010] predicted that different types of vehicles, including non-motorized ones (e.g., bicycles), would become part of the Internet of Things (IoT), enhanced with the ability to communicate with other devices. This vision is now a reality, with affordable new vehicles providing features supported by an Internet connection.

As fully autonomous driving has yet to be regulated in many places throughout the world, most of the features brought by IoT to vehicles must consider that drivers must stay focused on driving and avoid distractions. Context-aware computing techniques [Dey, 2001] fit in this scenario. Vehicles already have many sensors, and Internet connectivity brings more power to collect data that can be used to infer the vehicle's context and to act or adapt accordingly. Given the strength of such techniques, investigation, both academic and commercial, focusing on the contextual characteristics of vehicular network applications is growing in numbers [Vahdat-Nejad *et al.*, 2016].

Modeling information that can be sensed in or retrieved from a vehicle environment helps to improve the development process, and consequently, the quality of Intelligent Transportation Systems (ITS) applications. Understanding which context elements categories could be useful in this environment allows one to create applications adaptable to the current context to offer the best possible user experience.

Applications with few context elements (or little contextual data) often provide a poorer user experience. More context-related data processed by the system during its execution could enrich the context information to enable a better user experience [Hu *et al.*, 2014]. For example, an application that utilizes not only location information (e.g., home, work) but also context information related to activities (e.g., mobility pattern) and environments (e.g., temperature) tends to offer a better service to its users. An example is an application that reacts to a user's location to turn on an air conditioner upon arrival. If the system only uses the user's location to achieve this goal, it can still work adequately. Nonetheless, with access to information such as the current room temperature and traffic information on the user's path so that the time of arrival can be calculated, it can optimize the moment to turn the air conditioner on so that the room is at a target temperature when they arrive. However, simply increasing the number of context elements in an application development does not solve the identified problem. The various context elements must be well-articulated, and it is not trivial to integrate a large number of possible context elements that may be relevant to the application. This way, a common vocabulary must be structured to describe and classify context information to guide software engineers in this task.

This paper focuses on defining a taxonomy for classifying context elements for connected-vehicle applications, with 79 categories, 4 of them being supra-categories and the other 75 distributed as subcategories of those four top-level ones. It should support decision-making on the use of context elements in the development of context-aware ITS. The overall goal is to enable context and situational awareness in the development and use of ITS applications, allowing system designers to have a better understanding of the possibilities of context awareness when designing their ITS applications. The validation process of the taxonomy is described, with the three approaches used being explained in detail: a blind experiment, the development of an ITS application, and the compilation of a knowledge base using the taxonomy categories as its guide. The taxonomy proved to be useful, with a significant increase in the number of context elements used in applications designed with its aid.

The remainder of the paper is organized as follows: Section 2 builds upon the literature to define the computational context and the vehicular applications scenario. Section 3 describes the proposed taxonomy. Section 4 describes an application designed using the proposed taxonomy as a validation of the model. Section 5 discusses the results obtained, and section 6 focuses on concluding remarks and identifies potential topics for further research.

2 Background and Related Work

This section discusses context modeling in the vehicular domain. Initially, it provides a background on Computational Context, and then on ITS and Context-Aware Vehicle Applications. It ends with a discussion of related works that categorize and model context in applications of the ITS domain.

2.1 Computational Context

Temdee and Prasad [2018] have performed an extensive review of definitions for context awareness. Their work analyzed the most commonly used definitions for context awareness and concluded that no agreement on a general definition of context exists. However, some points in common were identified, in particular, all definitions agree that any information used to characterize the situation of entities relevant to a system is part of the context.

Another point of convergence in the works of Dey [2001]; Vieira *et al*. [2009]; Abowd and Mynatt [2000] and Zimmermann *et al*. [2007] is the need for classification of useful data for context reasoning. The five Ws of Abowd and Mynatt [2000] (Who, What, Where, When, and Why) are a basic and general way of categorizing contextual information. Dey [2001] and Zimmermann *et al*. [2007] *dimensions* are a step further in modeling context regardless of domain. Vieira *et al*. [2009] go even further and propose a general ontology of context elements, representing not only categories but also their relationships. Being generic, their model has a coarse granularity and cannot represent particularities in specific domains. In [Vieira *et al*., 2009] an important concept is defined to distinguish *context elements* from *Context*. *Context elements* are unique pieces of information that together can define the *Context*, which is "the set of instantiated *context elements* that are needed to support the task at hand".

Modeling context has since become a specific research topic. The domain knowledge of a situation can be described by a hierarchy of *context elements*, where some of them have been evaluated, either to a specific value or to another context element [Brezillon and Brezillon, 2007].

Bettini *et al*. [2010] discuss requirements of context modeling and reasoning techniques, delineating what they defined as contextual information types and their relationships. Among the requirements, Bettini *et al*. state that contextual models must be heterogeneous, being able to deal with information gathered from different sources in distinct update rates, and have a wide range of semantic levels. Models should also be able to describe clearly the relationships and dependencies between different types of information. Timeliness, using past states and possible predictions of the future state, is another feature that a context model should have, as well as representing the potential imperfections of information used and generated.

Bettini *et al*. [2010] also elicit that the model's formalisms have an acceptable degree of usability, so application designers can correctly use the proposed model to translate the real-world concepts to the constructs used in the model. Another requirement related to this desired usability is the provisioning of contextual information, defining paths to reach each contextual information. Bettini *et al*. affirm that the dimensions proposed by Dey [2001] are commonly used as the root of such paths, thus named as *primary context*.

2.2 ITS and Context-Aware Vehicular Applications

Based on the definitions of Figueiredo *et al*. [2001] and Guerrero-Ibanez *et al*. [2015], it is possible to establish that an ITS is the application of communication, information, and electronics to minimize pollutant emissions, vehicular wear, and time spent on commuting while maximizing fuel efficiency, road usage, and safety. Guerrero-Ibanez *et al*. [2015] state that emerging technologies, such as the trend of connected vehicles, Cloud Computing and the Internet of Things (IoT), will shape the future development of ITS.

According to Vahdat-Nejad *et al*. [2016], some of the possible services provided by the so-called *Vehicular Network Applications* include collision warnings, road hazards, traffic conditions, and points of interest notifications, and overtaking assistance.

A large number of vehicular network application projects already make use of context awareness. Vahdat-Nejad *et al.* [2016] provide an extensive survey with the state-of-the-art of *Context-aware Vehicular Network Applications*, categorizing them according to different criteria, such as service type, context type, context gathering methods, environment type, system architecture, and others. Diverse vehicular applications try to solve a wide range of problems in this domain, ranging from safety issues, traffic congestion avoidance, environment protection, information, entertainment, and driving comfort [Vahdat-Nejad *et al.*, 2016]. The same study shows that applications can focus on urban, freeway, or both environments and that while some applications were still in the design phase, there are already plenty of context-aware vehicular applications deployed. Overall, Vahdat-Nejad *et al.* [2016] support that applications in this domain are very diverse but share some common characteristics.

More recent research such as those by Swarnamugi and Chinnaiyan [2020]; Chavhan *et al.* [2021], and Dzemydienė and Burinskienė [2021] show that context-awareness in ITS is indeed a relevant and effective approach to enrich this kind of application.

It is worth noting that it is common to find research similar to Zheng *et al.* [2016], where the usage of Big Data on Social Transportation is reviewed, mentioning several topics clearly related to context-awareness, but without even mentioning the word "context". A careful researcher in this area must be aware of this fact because most of the projects related to ITS have some use of computational context. One example of such a phenomenon is the work of Li *et al.* [2020], where ontologies are used to generate test cases for automated and autonomous driving scenarios. Such ontologies are certain to represent several context awareness-related features, however, the work does not mention context awareness.

Finally, Laña *et al.* [2021] explores the relationship of ITS applications with context awareness, citing several examples where context-aware systems are used in this domain. It also explores how information such as demographic characteristics, and information about the road, network, mode, and travel are used in context-aware ITS models. Furthermore, they also list cases where social information originated from mobile and wearable devices, as well as from social media, and are used in features such as traffic analysis and forecasting and the generation of recommendations in the context of transportation. The impacts of regulations such as data privacy controls on context-aware ITS applications are also briefly discussed by the authors.

2.3 Related Works

We performed a literature review looking for other taxonomies and models focused on context elements for generic vehicular applications. We defined the following query that encompassed the terms that we found most probable to identify works with similar goals in relation to our research:

```
("taxonomy" OR "ontology") AND ((("context
aware" OR "context awareness") AND ("v2x"
OR "vehicle to everything")) OR (("context
aware" OR "context awareness") AND ("smart
cars" OR "intelligent transportation
systems")) OR (("v2x" OR "vehicle to
everything") AND ("smart cars" OR
"intelligent transportation systems")))
```

The search was performed using Google Scholar, and since this tool has no support for such a query, we generated the 24 different queries that are possible given the logic operators in the previously shown query and performed the searches individually. We collected all results from each query, joined and removed duplicated results, and ordered them by their number of citations. We analyzed the works from the most cited to the least, discarding those that are not related works, until we collected the seven works cited in this section.

We did not find any work that summarizes such information in this domain. However, some alternatives specialized in sub-domains of context-aware vehicular applications were found. Vieira *et al.* [2011] list context elements useful for public transportation systems but mention no hierarchy or relationship among them.

Vahdat-Nejad *et al.* [2016] provide a categorization of context types into four groups. Each of these groups contains context information that shares similar characteristics. One group is local context, for information that describes local entities, such as the vehicle where the system is running or its driver. Another group is the external context, which describes the same information but is related to other vehicles or drivers in the scenario. Vahdat-Nejad *et al.* [2016] define other two groups: General-related to transportation, for information that is directly connected to the context of transportation but is not related to a vehicle or its driver (such as parking information); and General-unrelated to transportation, for any other information relevant to an ITS but not related to transportation, such as weather.

The coarse granularity of the categorization defined by Vahdat-Nejad *et al.* [2016] is interesting for their objective of identifying types of projects that usually need each kind of information. However, for the aim of helping software professionals to understand better the possible uses of context awareness in their ITS applications, it is too coarse. We would need a categorization in finer granularity based on context elements.

On the opposite direction of the coarse-grained categorization provided by Vahdat-Nejad *et al.* [2016], we have found the ontology defined by Klotz *et al.* [2018]. The ontology provides a very fine-grained definition of signals available in the vehicle, going as low as to identify specific doors or seats in the vehicle, for instance. Our research does not need such a low level of detail, and yet the ontology is limited to vehicle information, but it is a good source of inspiration and validation of our model.

Kannan *et al.* [2010] used an OWL-based ontology to model context regarding vehicle safety, defining context elements and their relationships for safety-driven context-aware vehicular applications. Xiong *et al.* [2016] also propose an

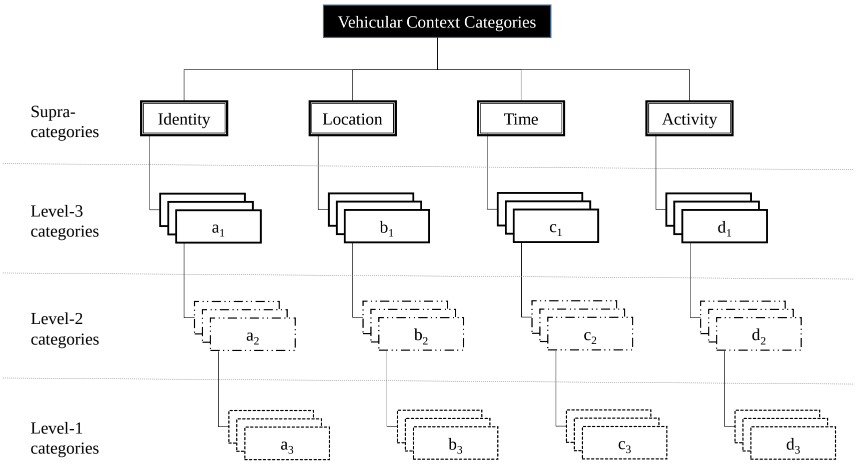

Figure 1. Proposed taxonomy model

ontology focused on context elements related to the driving activity, and most of their use scenarios are around Advanced Driver-Assistance Systems (ADAS), sometimes extrapolating the ontology's use to self-driving vehicles. Like other already discussed ontologies, this one focuses on a particular sub-domain of ITS (driving). Driving, for instance, is only part of what the taxonomy proposed in our research covers.

Arooj *et al*. [2022] proposes a taxonomy of Big Data in the domain of the Internet of Vehicles (IoV), which is very similar to the domain of ITS and can be considered a subset of it. Different from our proposal, the taxonomy defined by Arooj *et al*. [2022] is a classification of everything related to Big Data in IoV, with its topmost categories representing the phases required to use Big Data, such as Data Acquisition, Data Storage, Data Processing and Data Analysis. While the overall data contained in their taxonomy is very different from our proposal, both of them share a common goal of helping people who will work on the design of systems to understand the domain better. In Arooj *et al*. [2022] case, researchers of IoV are explicitly declared as their audience, while in our work, the focus is on ITS software engineers.

Finally, Sobral *et al*. [2020] designed an ontology-based approach to integrate and visualize data in the Urban Mobility domain, which is also very correlated to the ITS domain. They propose the VUMO Ontology, with four upper classes: Urban Mobility Concept (UMC), Data Concept, Visualization Concept, and Domain Expert Concept. UMC upper class is particularly related to our research: it classifies some concepts that exist in the Urban Mobility domain, using the classes Agent, Infrastructure Component, and Event to further subclass it. The specializations of these classes, such as Passenger, Vehicle, Route, SocialMediaPost, and TravelEvent are also present as categories of our taxonomy.

3 Context Elements Taxonomy for Vehicular Applications

We consider the definition of context elements given by Vieira *et al*. [2009] as a starting point for our proposal. Understanding the concept of a context element as a single piece of information that can carry enough value to be processed and infer the current context is a central factor in the definition of our model. Our initial proposal presents a hierarchical model where each subcategory has only one parent category, and each context element is part of a single category. While we understand the limitations of favoring such a restricted model instead of a richer ontology-based model where elements could belong to multiple categories, as well as categories could relate to others, we also perceive that such a simple model has advantages.

Understanding a hierarchical model is straightforward since the relations are simpler and linear. While a full ontology formalizes the relations among different elements in the modeled world, such relations can visually pollute graphic representations of the model [Lohmann *et al*., 2016] and may not be relevant to most users. A hierarchical categorization focuses on establishing only the strongest links among its elements, freeing readers to define less relevant relations that can be useful to their scenario. The concept of representing useful contextual information using this kind of model has already been defined for other domains such as Digital Television, Web Navigation, Medical Monitoring applications and health software defects in [Chagas and Ferraz, 2012; Villegas *et al*., 2011; Mitchell *et al*., 2011; Rajaram *et al*., 2019].

The methodology to create the taxonomy is similar to the one defined by Papatheocharous *et al*. [2018]. The taxonomy was built in three main phases: In the first phase, we identified the problem through a literature review on ITS and Context-Sensitive Systems, in general, to understand their development process and identify how a taxonomy could help improve this process. The second phase was the design of the taxonomy, starting with the identification of context elements used in the ITS applications found in the projects reviewed in the literature. After collecting the context elements, they were classified according to the similarity of their characteristics into various categories. The third phase was the validation, which is further detailed in Section 4 of this work. Similar to the process followed by Papatheocharous *et al*. [2018], there were several iterations between the second and third phases, which served to improve the taxonomy design.

Our proposed taxonomy model (Figure 1) is based on the four basic context types defined by Dey *et al*. [2001]: Iden-

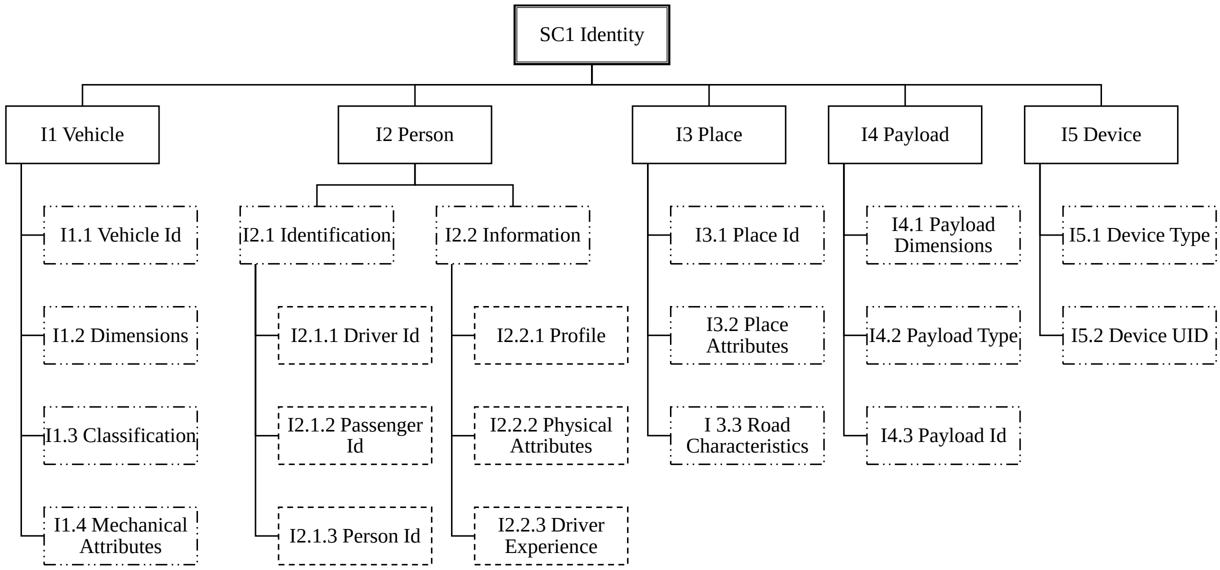

Figure 2. Identity supra-category of the proposed taxonomy

tity, Location, Time, and Activity (or status). Other models include other top-level categories, such as the ontology for context representation in groupware proposed by Vieira *et al.* [2005] or the categories proposed by Kaltz *et al.* [2005] for web applications, but these four categories are always present either as a distinct or as part of another category. Due to size constraints, it will be presented in parts according to the four basic context types. These four basic context types are called *supra-categories* in this work, as they are the topmost categories in our taxonomy.

The categories represented in the model contain information that can define the context for many different actors in ITS. When the term Vehicle is used, it encompasses both motorized vehicles such as cars, buses, motorbikes, or trucks, and also human-powered vehicles such as bicycles. When categories using terms like "Driver" are used, we expect the reader to be able to extrapolate the term to "the entity in control of the vehicle", such as a pilot for motorbikes, the cyclist for bicycles, or software for an autonomous vehicle.

3.1 Identity

Identity (SC1) is the most straightforward category and is depicted in Figure 2, consisting of information that helps to identify the main elements in the scenario, as well as to characterize them with their immutable attributes. It can be further sub-categorized by defining the main stakeholders in a vehicular application: the vehicle itself, the driver, the passengers, and the cargo being transported.

Vehicle (I1) identity is the first category we will describe. It has information whose purpose is to portray the vehicle so that contextual applications can have enough data to uniquely identify and reason about the vehicle's characteristics. Vehicle identity information can be specialized into four categories:

Vehicle Id (I1.1): composed of context elements like Vehicle Identification Number (VIN), License Plate, national registration numbers, or any other information that can uniquely identify the vehicle. The vehicle's brand and model are also information that fits in this category.

Dimensions (I1.2): Information about the size of the vehicle. Length, height, and width, as well as weight specs (unloaded and maximum weights, for instance), are context elements in this category. It comprises also the dimensions and weights of possible trailers towed by the vehicle. It is important to emphasize that the vehicle's current weight does not fit in this category, due to its high mutability.

Classification (I1.3): information about vehicle categorization, such as its type (Hatchback, Sedan, SUV, etc.). While initially, this seems to be not a category but a single context element, vehicle classification is complex, and many alternative schemes exist, grouping vehicles according to different attributes.

Mechanical attributes (I1.4): Context elements that carry invariable features of car mechanics. Engine displacement, suspension and gear characteristics, and any other information that helps to define the vehicle based on its mechanics.

Context elements in the vehicle identity category must be immutable, or at least stable enough to rarely change (such as license plates or registration numbers, depending on the jurisdiction). This allows them to be known in advance, requiring limited integration with sensors. Values for these context elements can be hard-coded in the vehicle's onboard computers, inputted by users on the first use of applications, or fetched from online services covering vehicle identification.

Person (I2) identity is equivalent to the former category but contains elements that define people related to the application, such as the driver, passengers, or pedestrians. This category is divided into the identification and information subcategories.

The *Identification* (I2.1) category holds three subcategories: *Driver Id* (I.2.1.1), *Passenger Id* (I.2.1.2), and *Person Id* (I2.1.3). Mostly, the same information, such as a Driver's License or any official identification document, can be part of any of these subcategories. What defines to which of them it is part in a specific system is the semantics: The driver's license, when used to identify the driver, is part of the Driver Id. However, when it is used to uniquely identify a pedestrian, would be part of the Person Id, since that person is not a driver in the context of the application.

Elements from either of these categories can come from knowledge, physical, and possession characteristics, as defined in terms of authentication factors. Names and official document numbers are examples of knowledge identification factors. These elements will generally be informed by users during setup, sign-up, or log in to applications. Given that other driver identification elements are present, they could be fetched from local or online databases.

Biometric data are physical identification elements. Considering that facial recognition technology is currently mature enough even to differentiate identical twins [Leyvand *et al.*, 2011], we can use user pictures or 3D mapping data as physical identification of the driver.

Finally, possession-based identification elements can consist of tokens, cards (both contact and contactless), or any other hardware (e.g. smartphone) that can provide data to identify the user and is supposed to be in his exclusive custody. Devices like these are already in use for security-related applications, such as anti-theft systems and an increasing trend in remote keyless entry systems. This kind of information can be obtained from card readers, but most commonly from wireless sensors, to avoid the hassle of fitting keys, cards, or other devices in specific places. When wireless technology is used, drivers can just carry the identification hardware in their pockets and the system will still be able to retrieve the required information.

As already mentioned, most types of identification can be part of the Driver, Passenger, or Person Id categories. However, there are some unique elements of Driver Id, such as a professional driver's registration number (within a company, for example), or of the Passenger Id subcategory, such as a train ticket number.

The *Information* (I2.2) category contains identity elements that cannot be used to uniquely identify a user. This information is either static or at least expected not to vary frequently. Its three subcategories are Profile, Physical Attributes, and Driver Experience.

Profile (I2.2.1): Elements representing general characteristics of persons involved in the application are part of the *Profile* category. Information such as the person's name, address, phone number, birthday, or social media links is part of this category. Infotainment-related preferences, such as music style or preferred radio station are also part of the profile. It is mostly obtained via manual input, but some of the information can be collected through the use of web services. It is important to remember that user profile information can be subject to privacy laws and must be protected accordingly.

Physical attributes (I2.2.2): Applications designed to improve ergonomics must have access to information such as height, weight, and other more specific physical characteristics. Accessibility-based applications also need to gather information regarding the person's physical abilities. Applications with features of safety and emergency-handling situations can be improved by having access to elements of this category. Acquiring context elements in this category is possible via user input and sensors.

Driver experience (I2.2.3): Another category of context elements that can be used by applications. The word *experience* can be considered here encompassing both the length of the driving experience as well as the driving skills. How long this driver is licensed is relevant. Other drivers can be warned that nearby drivers are under training or are newly licensed.

Regarding driving skills, we would include categories of vehicles that the driver is licensed to drive and eventual training programs he has completed (hazardous cargo training or any other courses focused on professional drivers).

Information about elements in the driving experience category can be collected in several ways. User input can be used but is unreliable since drivers can lie or be too optimistic about their skills. Obtaining data by fetching services based on the driver's other Identity information can be useful and provide better results. A novel approach would be using AI to infer some of the driver skills based on the driving activities performed, something already viable [Johnson and Trivedi, 2011].

The previously mentioned categories would have all the required information to characterize drivers and vehicles. While these are arguably the most important components from the point of view of vehicular applications, other components can also play marginal or central roles depending on the requirements and objectives of the application. We will now describe the categories and subcategories defined to contain context elements related to places, passengers, and cargo.

Place (I3) holds information that can be used to identify or characterize a place, in this sense, examples of elements in the *Place Id* (I3.1) subcategory are Addresses or business names (e.g., Museum of Science and Technology).

Another subcategory for Place is *Place Attributes* (I3.2). This category encompasses elements such as the dimensions of the place, opening hours, and also restrictions such as the maximum allowed weight or height for vehicles to come inside the location, or the minimum age for people to be admitted to the place.

Road Characteristics (I3.3) contain static information about a road section. Its pavement, the number of lanes, length, and other relevant road information fall into this category.

Payload (I4), in its turn, is very important when we consider commercial vehicles.

Payload Dimensions (I4.1): these are relevant to applications related to freight, and encompass information such as width, height, and weight.

Payload Type (I4.2): It is defined as a category and not a single context element for the same reasons as the vehicle classification. There are several different methods to classify cargo and applications may need to use more than one classification system at the same time. Particularly, cargo hazard classifications are very useful in a wide range of applications, from inspection to handling emergencies. Examples are the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) [Winder *et al.*, 2005] or the United States Environmental Protection Agency Toxicity Category [Environmental Protection Agency, 2021].

Payload Identification (I4.3): this category holds context elements used to uniquely identify the payload. Barcodes, parcel tracking numbers, and any other information that can identify the cargo are part of this category.

Cargo information can be collected by the input of cargo manifest into the system, preferably using integration with other systems to avoid human error. Weight sensors and video or infrared cameras can also be useful to gather information about cargo currently loaded in vehicles.

Device (I5) identity is important to handle data from other equipment, such as traffic lights, sensors, and network infrastructure devices. Its first subcategory is *Device Type* (I5.1), for context elements containing information regarding the type and capabilities of the device. Another subcategory is the *Device UID* (I5.2), for information that can uniquely identify a device, such as a MAC address or a traffic light identification number.

3.2 Location

Mobility is a key factor in vehicular applications. Such software will run most of the time while cars are in motion, some of them only being meaningful in this situation. Context elements regarding the vehicle position are for some applications the most important to its proper working. Location may not seem like a category with a great diversity of elements, but this is not correct. While Geographic coordinates are important and well-known information to define the location, other ways of defining position also exist and might be relevant in our environment. Figure 3 shows the categories and subcategories in the Location (SC2) supra-category.

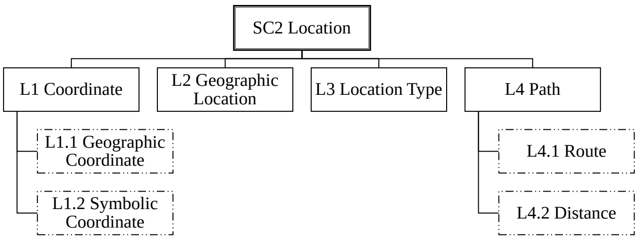

Figure 3. Location supra-category of the proposed taxonomy

Two types of **Coordinates** (L1) can be used, each with its own possible context elements. The first holds context elements related to *Geographic Coordinates* (L1.1). Information like Latitude, Longitude, and Altitude would be classified in this subcategory. The ready availability of sensors that gather data related to the geographic coordinates from GPS and GLONASS satellites makes their use widespread. Map-based applications are very common and have changed the way drivers get prepared to travel to unknown places. Along with other contextual information described in some other categories here, applications like Waze, which can help deal with traffic and other commuting issues, have been developed and are in use, improving the driving experience for millions of people.

Symbolic Coordinates (L1.2) is the other type of coordinates able to define a location in a vehicular environment. This category is mentioned by Bettini *et al*. [2010], and context elements that provide coordinates and identifiers not related to the physical world can be fit here. Examples would be the Cell Id of a cellular network base station, identifiers for other wireless networks, or special-purpose beacons placed in strategic locations.

Geographic Location (L2), or Semantic Location, provides more semantics to the location information. Addresses, Road names, floor numbers, and any other information that can be used to identify places without being connected to Geographic features of the location are part of this category. Generally, the most reliable way to obtain such information is based on geographic or symbolic coordinates and geographic information systems.

Location type (L3): in automotive applications, it can be very useful to know whether the vehicle is currently on an urban street, on the road, or in a parking lot, for instance. This subcategory serves to identify not the specific and unique place where the vehicle is, but the type of this place. Different rules can apply according to such information. Like the address subcategory, this information is also generally dependent upon geographic or symbolic coordinate values, but depending on the environment, symbolic coordinates can be more important and even be used independently, such as when vehicles are indoors, in multi-story car parks.

Another location subcategory is **Path** (L4). It is further subdivided into *Route* and *Distance*. The context elements that are part of *Route* (L4.1) are locations that define a way, from the starting point of the journey until its destination, including both the start and the destination. Manual and automatic alternatives exist to obtain route information. Manual methods include user input to define its route. Automatic methods for obtaining routes are based on user history or connected to web services that contain the user's agenda. *Distance* (L4.2) holds context elements that represent distances between two points, such as the distance between a vehicle and a destination, or between vehicles.

These categories are useful for characterizing the current location of the vehicle as well as its route. In general, location data can be associated with time information to help make the current context of the vehicle clearer.

3.3 Time

Time-related context elements can be used to refine the identification of the current context. Similar to location, Time (SC3) can be thought of as a simple, indivisible category, but different types of information can be collected and used based on time values, and we categorize these elements into three categories: Local Time, Schedule, and Travel Time. The organization of the elements in this category is illustrated in Figure 4.

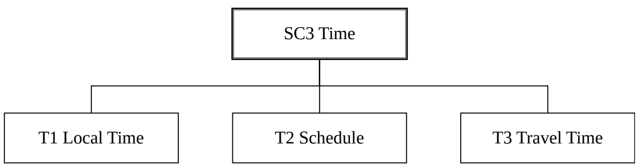

Figure 4. Time supra-category of the proposed taxonomy

Local Time (T1) is related to time information of the current vehicle location. This includes the date, time, day of the week, and more subtle or subjective information, such as whether the current day is a holiday, workday, or weekend day. Applications dealing with traffic information can use it to predict traffic conditions and suggest better alternatives.

Timestamps can be collected from local devices' time settings, or more accurately from time servers online. Holiday information can also be consumed from web services.

Schedule (T2) is designed to contain context elements that represent information on scheduled appointments of drivers or passengers, or due dates and times of arrival of the transported cargo. This information can be collected from integration with user's agenda systems (like Google Calendar or smartphone applications), integration with enterprise systems (in case of cargo due dates), historical data, or ultimately but not ideally, user input.

Travel Time (T3) is another subcategory, which aims at collecting time information regarding the travel itself. Information like the *time a journey has started* and the *last rest stop* by the driver are in this category. Such data can be used to measure tiredness probability and recommend drivers to make unplanned stops, for instance. In countries with regulations for maximum continuous driving journeys for professional drivers, such information can be used in the inspection of such rules. Together with information from other categories, applications could also suggest rest stops before the driver reaches the legal limits.

3.4 Activity

Context-aware applications, in general, have to deal with user and device activity, in the meaning of their set of tasks, both current and background. Vehicular context-aware applications have multiple users and devices contributing to the processing of the current context. Thus, we have defined subcategories for each of the involved components whose activities are useful to such applications. These components include the driver, the vehicle itself, the network connecting the vehicle to other vehicles and devices, passengers, and the surroundings of the vehicle and its route. Figure 5 demonstrates the organization of the Activity (SC4) category and its subcategories.

The first subcategory we have defined holds context elements regarding **Driver**'s activity (A1), as follows. *Driver Status* (A1.1) and context elements in this category can be used, for instance, to measure a driver's attention, tiredness, and other important information related to the tasks a driver is executing while driving. Biometrics data such as pulse rate, temperature, or blood alcohol content are also part of this category. When combined with context elements from other categories, driver activity data enables applications to deliver information on adequate media, affecting driver's concentration as low as reasonably achievable.

Driving tasks (A1.2) are expected to be the most common type of tasks handled by drivers. Driver's actions on pedals, wheel, gears and every other car interface used to control the vehicle are low-level data that can be considered as context elements for this category. Such data can be retrieved through a vehicle bus like OBD-II (On-Board Diagnostics) or specific sensors in each of the interfaces. More advanced low-level data can be collected using cameras, which could perform eye-tracking to identify the direction where the driver is looking at. Using inference techniques and combining with other data, applications could generate high-level context elements which fit into this category, such as the identification that a driver is performing an evasive maneuver, emergency braking, or parking. Driver concentration level can be considered as another high-level context element that can be computed using other data as input and fits in this category.

External tasks (A1.3) is a category to hold any activity a driver is performing that is not related to driving or handling any other device in the car that controls information or entertainment systems. Their interaction with passengers or equipment that are not part of the car is in this category.

Infotainment tasks (A1.4) are related to the driver actions related to devices providing information and entertainment. While not their main task, drivers commonly have to deal with equipment like vehicle radios or GPS devices. Nonetheless, there is also high-level information in this category, such as knowing which station is tuned or which route on the GPS device is being followed.

Passenger's activity (A2) is another subcategory of the Activity category. Context elements representing the current activity and status of passengers would fit in here. It would have context elements to define the number of passengers on board, their current seats, and other information that identify not a particular passenger, but the group of passengers currently in the vehicle. Their actions, objectives, and tasks are useful information to infer the current context. Also, dynamic information about the passenger, such as biometric data like temperature and pulse rate, are useful to some applications. Gathering values for these elements is probably the most difficult of all the categories in this model since passengers usually have very low interaction with vehicle controls or interfaces. Using cameras, image and motion recognition, presence or weight sensors, and combining other information is required to obtain valid and useful values for the elements in this subcategory.

The **Pedestrian** (A3) category is divided into two subcategories: *Pedestrian Movement* (A3.1), which holds elements such as the speed, direction, and acceleration of pedestrians nearby the vehicle, and *Pedestrian Role* (A3.2), that holds elements that define the role of that pedestrian in the ITS, such as whether they are potential future passengers or if their activity is relevant to the system (such as a traffic agent or a first responder).

After describing the person-related activity categories, we define the **Vehicle** (A4) activity category. The category is subdivided into vehicle *activity*, *movement*, and *mechanical status*.

Information concerning the tasks the vehicle itself is performing, as well as the metrics for some variable information from the vehicle, according to the vehicle or driver activities, are context elements related to the *Vehicle Activity* (A4.1) subcategory. Another important part of *Vehicle activity* is context elements that provide information regarding what the vehicle is being used for. While these elements might seem the same as those in the Identity/Vehicle/Classification category, their context elements are different. Vehicle classification is static, while the type of service the vehicle is providing is dynamic. A pickup truck does not change its classification as a light-duty vehicle, but the same vehicle might be used, at different moments, as a passenger vehicle, for emergency handling, or cargo hauling, for instance. Collecting data that

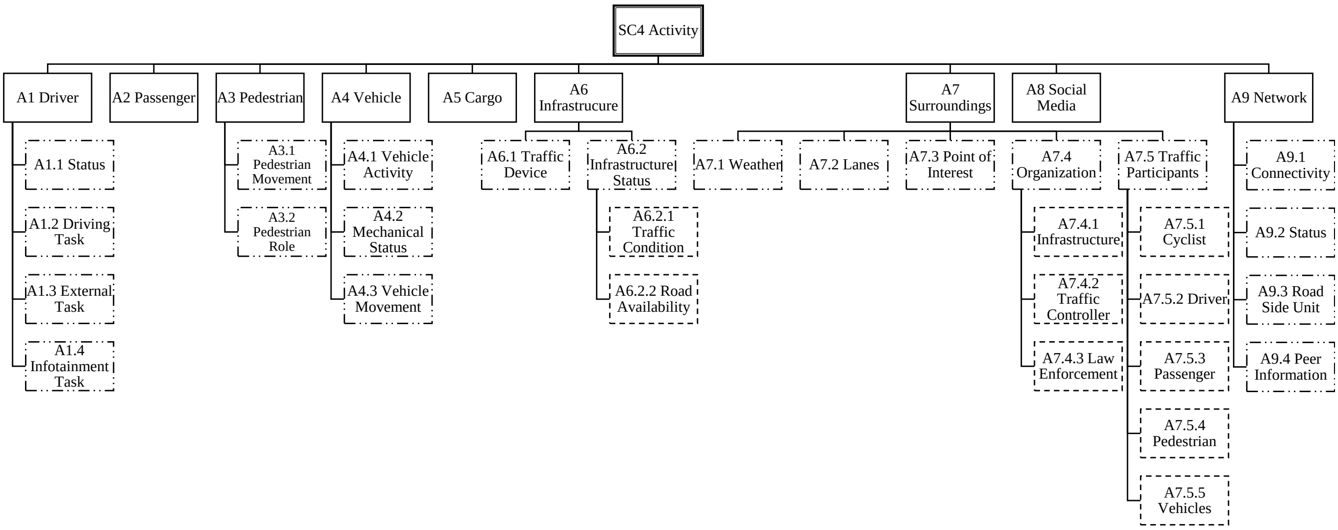

Figure 5. Activity supra-category of the proposed taxonomy

can be used to infer the current service can be challenging. The inference from other context elements can be a valid approach, as values for some possibly present sensors could be used to identify the service.

Mechanical Status (A4.2) is the subcategory to hold most of the information that can be collected from vehicle data buses as OBD-II or sensors on car parts. Interfacing with default buses to obtain data is not hard, but collecting information from sensors in parts not connected to such infrastructure can be challenging. Status messages that show that maintenance is required or the extent to which the vehicle has been handled can be higher-level information about the mechanical status.

Vehicle Movement (A4.3) subcategory holds information that can describe the motion attributes of a vehicle. Context elements that might fit in this category are speed, direction, and acceleration. Speed can be easily obtained using the vehicle's OBD-II interface. Acceleration (not to be confused with the pressure on the throttle pedal) can be calculated using distance, speed, and timers. Direction can be obtained from modern GPS receivers.

The **Cargo** (A5) Activity category holds context elements that characterize the payload interaction with the vehicle, either when already loaded or while waiting to be loaded. The temperature of the cargo, in the case of perishable materials, and pressure for the transport of gases are some of the elements in this category. Monitoring data of livestock being transported is another example.

The **Infrastructure** (A6) category is meant to hold information regarding the status of road-related equipment and the road itself. It is further divided into two subcategories, *Traffic Device* and *Infrastructure Status*.

The *Traffic Device* (A6.1) subcategory holds information about road equipment such as traffic lights, messaging boards, traffic signs, toll plazas, and other road devices. Such equipment has data that can be very important to many vehicular applications. Traffic lights can share their current color and how long it would take for it to change, message boards could broadcast their current message or more detailed information that would otherwise be not feasible to be displayed due to its size restrictions, traffic signs could share their enforcing rules or warnings to vehicles without needing to rely on online databases which can be not updated, making sure that vehicular applications receive the same data as the driver can see. V2I communication, with the devices "broadcasting" their status, is a viable solution to obtain road device information.

Infrastructure Status (A6.2) contains two subcategories. *Traffic Condition* (A6.2.1) is a category with context elements to ITS applications, such as the level of traffic jams and their causes (such as cars stopped on the road, potholes, and other events that can impact the traffic flow). ITS applications can use this kind of information regarding traffic and road conditions to infer current context and predict future situations. The other subcategory of Infrastructure Status is *Road Availability* (A6.2.2), which contains elements regarding the possible blocks (either total or partial) on the road. Indeed, applications such as Waze make heavy use of information from both *Traffic Condition* and *Road Availability* to provide driver assistance. Some data related to road availability and traffic conditions can be collected from devices broadcasting the road status, inferred through cameras and other sensors, or obtained through web services. Such services can be kept updated by using crowdsourcing techniques, as is used in the already mentioned Waze application.

Another subcategory in the Activity category is related to the **Surroundings** (A7) of the vehicle. This subcategory is a parent to several other subcategories that will be further described. Information can be collected by using wireless networks or the integration of GPS data with online services. Cameras, RADARs, and LIDARs can also be used to gather data about the surroundings of the vehicle. This category not only knows which vehicles, people, or points of interest are in the surroundings but mainly knows of their current activities. For example, understanding that a pedestrian close to the vehicle is on a cell phone, potentially distracted, is crucial to avoiding an accident; having information on the identity and activity of nearby vehicles can be very useful if there are emergency vehicles in the vicinity; knowing that a nearby restaurant is currently open is essential to the completion and usefulness of the information.

The first subcategory in the Surroundings subcategory is *Weather* (A7.1). Information regarding temperature, wind, air humidity, and rain or snow forecasts is obviously useful for many ITS applications. Also important in the domain of Surroundings is the *Lanes* (A7.2) subcategory. It holds information about the number, availability, and current way of lanes (in case of reversible lanes) in the vicinity of the vehicle. Another Surroundings subcategory is *Point of Interest* (A7.3). It holds information about any location in the vicinity of the vehicle or its route that might be useful to the context of the application. Gas station fuel prices, tourist attraction information, or stores in the route are possible information of this category that are used in ITS projects.

Two of the Surroundings subcategories are further subcategorized as shown in Figure 5: *Organization* and *Traffic Participants*. *Organization* (A7.4) has the *Infrastructure* (A7.4.1) subcategory, which holds information about road infrastructure near the vehicle, i.e., this subcategory focuses on the presence and status of infrastructure items only in the vicinity of the vehicle, in contrast to the *Activity/Infrastructure/Status (A6.2)* subcategory. *Traffic Controller* (A7.4.2) holds information about entities that have the power to control the traffic flow, such as traffic agents or traffic signs near the vehicle or on its route. Location is a key differentiating factor when such information is in this category or in *Activity/Infrastructure/Traffic Device (A6.1)*. The state of a traffic sign is always part of the *Activity/Infrastructure/Traffic Device*, but if that traffic sign can directly affect the vehicle, it is also part of the *Traffic Controller* subcategory. *Law Enforcement* (A7.4.3) regards the presence and role of police, traffic agents, speed cameras, and other entities involved in traffic law enforcement. While a common (albeit controversial) use of elements in this category is to warn drivers of the presence of these entities on their route, less disputed uses of elements in this category exist, such as automated first-responder allocation systems and other security and safety applications.

The *Traffic Participants* (A7.5) subcategory includes *Cyclist* (A7.5.1), *Driver* (A7.5.2), *Passenger* (A7.5.3), *Pedestrian* (A7.5.4) and *Vehicle* (A7.5.5). Their elements are both the presence of any of these participants in the vicinity as well as any other activity information relevant to the system regarding one of those participants.

Social Media Activity (A8) is a category to hold context elements regarding data coming from social media. The usefulness of data obtained from social networks, such as Twitter, to predict traffic jams and other transit-related issues is well proven, and various research has already been performed in such direction [Wongcharoen and Senivongse, 2016; Essien *et al*., 2020]. Information gathered from social media regarding friends and acquaintances nearby is also useful in some ITS applications.

The **Network** (A9) Activity category contains context elements that represent the state of the network that a vehicle is using to communicate with other vehicles (Vehicle-to-Vehicle - V2V), road infrastructure (Vehicle-to-Infrastructure - V2I), or the Internet, and as a consequence, with any other connected device (Vehicle-to-Everything - V2X). In this category, the *Connectivity* (A9.1) subcategory holds information about the network, such as bandwidth, type of network, and level of connectivity. The *Status* (A9.2) subcategory contains information about network statistics that are not part of the Connectivity subcategory. *Road Side Unit* (A9.3) holds context elements representing information collected from Road Side Units in the same network of the vehicle, such as traffic flow on a road segment [Woodard *et al*., 2016]. *Peer information* (A9.4) would hold context elements about the peers, the kind of device they are, and their interfaces to obtain more information.

3.5 History

Historical data can have a multitude of uses, being useful in predicting next context situations based on previous ones, such as knowing the traffic intensity information for a long period can help to predict future traffic. Previously captured data about fuel consumption, location, and several other information are useful and are already used in applications. Every previously mentioned context element from the aforementioned categories could be stored if useful in some context to an application. Such accumulated data can also be used to infer why some previous activities and outcomes happened, obtaining reasoning over which the application can rely on to better identify and adapt to future context situations.

Major issues related to context elements in this category are not related to gathering data, but to their storage: Depending on the granularity, a large amount of data can be generated, making local storage alternatives unfeasible, bringing us to cloud storage as a viable alternative.

Still, we have to be concerned with privacy because historical information is dangerous in the wrong hands, so data security is essential if a system would require storing such data online. When correctly used, it can have multiple good outcomes besides the ones already mentioned.

Historical use of context elements was not specified in terms of categories in this work, since we have observed that all history-related information is also part of some subcategory of the other supra-categories.

Those are the categories defined in our model. While we strongly believe that these categories reflect most of the existing useful context elements of the vehicular application domain, we reinforce that this model is not exhaustive, so it can be extended.

3.6 Using the taxonomy

Our proposed approach for using the taxonomy to help on designing an application is depicted in Figure 6 and is made up of five steps, where only the first one does not use our taxonomy. The first step in our approach is the initial requirements elicitation, where stakeholders' needs and expectations are collected through any technique or process that the analyst desires to use.

The second step in our approach involves analyzing the taxonomy to identify the categories that have already been covered during the initial requirement gathering. This step helps in identifying the existing categories that are relevant to the application under design. This can be done by just analyzing Figures 2, 3, 4 and 5, but it is recommended that the

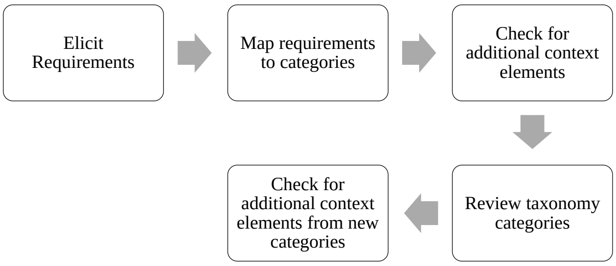

Figure 6. Process for using the taxonomy on the design of an application

person has also read sections 3.1, 3.2, 3.3 and 3.4. By marking the identified categories, the development team can have a clear understanding of the requirements that have already been captured and can focus on identifying new categories or requirements that may not have been initially identified.

The third step in our approach is the verification of the identified categories in the table found at https://amirton-chagas.github.io/CETITS-AnnexB.pdf. This verification can be done by the analysts themselves or by consulting with domain experts and stakeholders who can provide insights into whether the other context elements in the same already-used category could generate new features or requirements, or enhance any of the existing ones. This step helps in ensuring that most of the relevant requirements are captured in the elicitation process.

The fourth step in our approach involves analyzing the taxonomy again to identify other categories that could be useful for the application. These categories may not have been initially identified during the initial requirement gathering but could provide valuable inputs for enhancing the application's functionality or usability. This step involves a thorough review of the taxonomy to identify any missed categories that could be relevant to the application. Once the new categories are identified, the final step is the verification of these categories in the aforementioned table in a similar way that was performed in step 3, but now for the new categories that were identified as potentially useful.

4 Evolution and Validation

Three approaches were used to validate and evolve the taxonomy from its initial forms until the current proposed taxonomy: The first approach was a blind experiment where software development professionals were assigned the task of designing a context-aware vehicular application to a particular scenario. The second approach was to proceed with the complete process of designing and developing a vehicular application using this taxonomy in the process. Finally, with a mature iteration of the taxonomy, we used it to build a knowledge base of existing ITS projects in the literature.

4.1 Blind experiment

In this experiment, we had the participation of 21 subjects, in two rounds. In each round, the concepts of ITS, vehicular applications, context awareness, and context elements were introduced to all subjects and an explanation of the scenario was given. The chosen scenario was a vehicular application to improve the effectiveness of service provided by emergency vehicles, such as ambulances, fire engines, or police cars.

All subjects were briefed on what they were expected to accomplish, the overall designing of a coherent application in the described scenario, with the description of features that could be desirable, and mentioning which information the application would require to properly work. Participants were advised not to be constrained by what they think is currently possible with existing or deployed technology, with the sole restriction being the application being viable and coherent.

Then, they were divided into two groups, each in a separate room. The first was a control group, where participants had no contact with the proposed taxonomy (initial version), being allowed to start immediately designing the application. The second group was presented with that taxonomy and a table with an abstract of the description of each category, with similar content to what was presented in section 3. A very brief explanation was given to this second group about the hierarchical organization of the taxonomy, and then they were allowed to start designing the application.

Each participant was given one and a half hours to design the application according to informed rules. Subjects were free to format the output in the way they preferred, but it was recommended that a brief description of the objectives of their application was first given, and then a list of features. For each feature, the participant gave short descriptions of why it would be useful, who would benefit from such a feature, input data that would be required, and output data that the feature would generate.

Subjects were instructed not to identify the group they took part in the responses form, so an unbiased evaluation could be possible, with each response being identified only by a code. The list of participants in each group was hidden from the researchers until the end of the evaluation.

Two different researchers evaluated the proposals. Each researcher needed to assess a 0-10 score to five aspects of the application: *Coherence*, *Usefulness*, *Number of distinct context elements used*, *Number of distinct context elements incorrectly used*, and *Viability under current technology*.

The overall application *usefulness* was evaluated as 16.3% higher in the group that used the taxonomy when compared to the applications idealized by those in the control group. No significant difference was found in the *viability under current technology* and *coherence* of the application aspects. The *number of context elements used*, however, was much higher in the applications of the experimental groups. It is important to observe that the total number of context elements used could reflect an overuse of context elements too, and that is why we also evaluated the number of context elements that were elicited incorrectly by analyzing the application proposal and identifying which of the context elements mentioned by the application designer were not proper to the proposed application. This was performed by reading and understanding the goal of the application that was defined by the designer and assessing any possible use of each of the context elements elicited in the application which could provide a benefit to its performance, usability, or final result.

The context element was counted as *incorrectly used* if none of such benefits were found by any of the researchers.

We used the *number of distinct context elements used* and the *number of distinct context elements incorrectly used* to calculate another metric, the *correctly-used context elements*, as a simple subtraction of the number of incorrectly-used context elements from the total number of context elements used. While applications of the control groups averaged 4.5 *correctly used context elements*, applications designed by subjects in the experimental groups averaged a much higher average of 12.2 *correctly used context elements* per application. This indicates that access to the taxonomy of context element categories increased software engineers' awareness of the possibility of using different context elements.

This validation also served to improve the taxonomy, with categories regarding traffic devices, conditions, and surrounding traffic participants being suggested by the subjects in the experimental groups and added to the taxonomy.

4.2 Design and development of a vehicular application using the proposed model

We designed an ITS application using the taxonomy as input to aid in defining the application's context-related features. The following sections will describe the process followed to design the application, its results, and its limitations.

4.2.1 Procedure

The process to design a context-aware vehicular application was divided into two parts:

(a) A software engineer with previous knowledge in vehicular applications and context awareness designed the application accompanied by one of this work's researchers. The software engineer was a member of the Vehicular Innovation Laboratory of the Federal University of Pernambuco, Brazil (`http://live.cin.ufpe.br/`). Both of them worked together, under the core requirement of designing a system to assist drivers taking part in car groups to follow the same route. The researcher presented the taxonomy to the engineer, answering questions raised by the engineer on demand. Apart from the core requirement, they were free to suggest any other feature that they considered useful given the scenario.

(b) The same conditions were also given to another software engineer, with similar experience, but depriving him of the exposure to the taxonomy. Both of them were given one week to elicit the requirements for the application, with the researcher available at any moment to clarify issues.

4.2.2 Designed Application

The application designed by the engineer exposed to the taxonomy was called Convoy (Context-Oriented Navigation of Vehicles On the waY), a system intended to assist drivers taking part in car groups to follow the same route, navigating through roads and traffic, and reacting accordingly to unforeseen events which might happen. The description that follows includes the context elements that are used in the application and that form part of the taxonomy.

One of the drivers is designated the leader who must define the route that the other vehicles in the group must follow. Two options to define the leader can be made: Driver (using elements of the category *Identity/Person/Identification/-Driver Id* - I2.1.1) or Vehicle (using elements of the category *Identity/Vehicle/Vehicle Id* - I1.1). In addition, different versions of the Convoy can suggest the most suitable driver to be the leader on a journey, using information from context elements from the categories *Identity/Person/Information/-Driver Experience - I2.2.3*, *Identity/Person/Information/Profile - I2.2.1*, and *History*.

The leader must also choose when the convoy should stop to rest, eat, or sightseeing (*Activity/Surroundings/Points of Interest* - A7.3). Convoy should assist group leaders by notifying optimal times to rest or stop to eat, by assessing the values of context elements from several categories. Values from *Location/Path/Route* (L4.1), *Location/Coordinate/-Geographic Coordinates* (L1.1) and *Time /Schedule* (T2) can be used to define the distance to the destination, relate it to the passengers' schedules, and infer whether it would be worth the while to stop. Information from the *Time/-Travel Time* (T3) and *Identity/Person/ Information/Physical Attributes* (I2.2.2) categories can help prevent tired drivers from being kept on the road. The use of *Identity/Person-/Identification/Passenger Id* (I2.1.2) and *Identity/Person/Information/Profile* (I2.2.1) can also help if children, pregnant women, people with disabilities, or any type of passenger with special needs are present in the group, so that stops can be scheduled accordingly with their needs. Not least, the use of *Activity/Surroundings/Points of interest* (A7.3) and historical data can help identify safe places to rest. An ideal application could merge information from all the aforementioned context elements to decide the best time to suggest a stop for resting or sightseeing.

Convoy also focuses on sharing the following information among the vehicles in a group:

- Position of each vehicle, with an adaptive map that zooms in to fit the vehicles in the viewing window according to contextual information (*Location/Coordinate/Geographic Coordinates* - L1.1, *Activity/Vehicle/-Movement* - A4.3, and *Location/Path/Route* - L4.1);
- Vehicle status, such as whether it is moving, stuck in slow traffic, or in an emergency (using elements in the *Activity/Vehicle/Movement* - A4.3, *Activity/Vehicle/Mechanical Status* - A4.2, *Activity/Driver/Status* - A1.1 and *Activity/Passenger* - A2 categories);
- Route change notifications (*Location/Coordinate/Geographic Coordinates* - L1.1 and *Location/Path/Route* L4.1);
- Warning when a vehicle is too slow compared to the leader (*Activity/Vehicle/Movement* - A4.3).
- Notification that the leader is too fast when compared to other vehicles in the group (*Activity/Vehicle/Movement* - A4.3).

Every participant can declare an emergency, which would be due to context elements in the *Activity/Vehicle/Mechani-*

cal Status (A4.2), *Activity/Driver/ Status* (A1.1) or *Activity/-Passenger* (A2) categories.

In the design process – section 4.2.1 (b), the resulting application designed by the engineer not exposed to the taxonomy presented a much smaller context-enabled feature set. This developer elicited the use of maps in a similar way thought by the other professional who had the support of our taxonomy (1 feature), but none of the other features appeared in his design (11 other features).

The application with the one core feature could still achieve its goal, but at a very rudimentary level, whereas the application designed with the help of the taxonomy and its supplementary material was much richer and would be more useful for the described scenario.

4.2.3 Results

Overall, the application designed without the support of the taxonomy had the core feature of identifying the position of each member of the group on the map but had none of the other supporting features that Convoy had to provide services to this kind of group travel, such as automatic vehicle status, notifications on route changes, and suggestions of rest stop strategies.

An initial version of Convoy implementing some of the above features was developed and could be used as an application to be embedded in automotive systems in the near future. Concerning non-functional requirements such as performance, being aware of context elements in the *Activity/Network/Connectivity* (A9.1) category can improve application performance and resource usage.

4.2.4 Limitation of this Evaluation

It is not possible to affirm that the taxonomy was the sole responsible for the different results. However, we attempted to make the environment the most similar as possible so that the major contributor to any difference in the result was the exposure to the taxonomy.

4.3 Creation of a knowledge base of context element categories used in projects available in the literature

The final taxonomy validation step was the creation of a knowledge base (KB) from ITS projects available in the literature, listing the context element categories of the taxonomy which are used in the projects in question. Therefore, we would validate the practicality of the categories defined in the taxonomy since there must be at least one project in the current literature that uses elements from each of the defined categories.

First, we performed a literature review to find articles in scientific journals and conference proceedings that describe existing ITS. There was no specific filter regarding context awareness, as systems commonly have context-aware features without explicitly mentioning them. Articles describing systems developed since 1998 were found, and even if the older projects are out of date, the set of context elements used by them can still be useful and represent valid information to assess the categories defined in this taxonomy. Five surveys were identified [Vahdat-Nejad *et al*., 2016; Baras *et al*., 2018; Gomes *et al*., 2020; Khekare and Sakhare, 2012; Soyturk *et al*., 2016], with their subject being varied, but connected to vehicular applications and context awareness. The projects cited in those surveys were also analyzed to check which of them were appropriate to this research's objectives, since the subjects of the surveys always encompassed our domain, but were also broader than it.

The review resulted in a total of 70 projects collected from 68 papers. Two of the papers describe two projects each: Bifulco *et al*. [2014] reports two cases of ITS usage, one in Singapore and the other in Amsterdam, while David *et al*. [2013] presents four smart city-related systems, two of them (Loading Zone Management and Communicating Bus Stop) being useful for our knowledge base. 41 of the papers were identified through the previously mentioned surveys: 19 papers were collected from [Vahdat-Nejad *et al*., 2016], 7 from [Baras *et al*., 2018], 9 from [Gomes *et al*., 2020], 2 from [Khekare and Sakhare, 2012], and 4 from [Soyturk *et al*., 2016]. Some of the papers appear in more than one survey, so the numbers cited mention only the first occurrence of the paper in a survey.

The 70 projects use context elements from at least one of the context element categories defined in our taxonomy. Furthermore, three worldwide used applications, Waze, Uber, and Moovit, were also analyzed and included in the knowledge base, leading to 73 projects in total. The authors analyzed Waze and Uber based on their personal experiences as users, supported by full navigation through the applications' features. In the case of Moovit, the authors are not frequent users of the application, so alongside the navigation of the application, an article that contains a section describing Moovit helped on mapping the categories used in the application [Santos and Nikolaev, 2021]. All categories in our proposed taxonomy appear in at least one project in this knowledge base.

During the development of the knowledge base, we did not map whether the usage of elements from a category is for current or historical data because while this information could be valuable for some potential uses of this knowledge base, our objective is to validate whether the categories in the taxonomy are indeed used in real projects. Thus, we marked the category as used in a project either when its data usage was current, historical, or both.

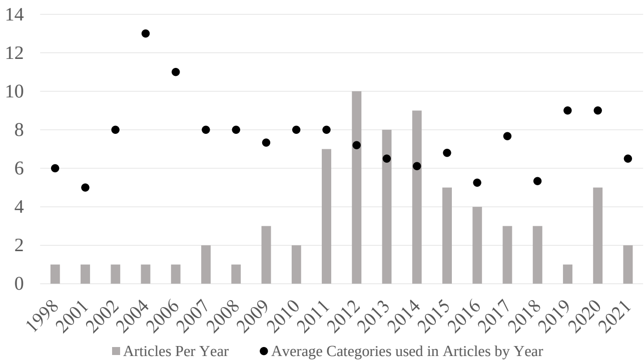

Figure 7. Distribution of projects per year and the average number of categories used in the projects of the articles of each year

Table 1. Top-10 categories used in the 73 analyzed projects.

Rank	Category	% of projects that use the category
1	Location/Coordinate/Geographic (L1.1)	53.4%
2	Activity/Vehicle/Movement (A4.3)	47.9%
3	Time/Local Time (T1)	41.1%
4	Activity/Infrastructure/Status/Traffic Condition (A6.2.1)	32.9%
5	Activity/Surroundings/Traffic Participants/Vehicle (A7.5.5)	31.5%
5	Identity/Vehicle/Vehicle Id (I1.1)	31.5%
7	Identity/Person/Information/Profile (I2.2.1)	30.1%
8	Activity/Surroundings/Weather (A7.1)	28.8%
8	Identity/Vehicle/Classification (I1.3)	28.8%
10	Time/Schedule (T2)	26%

Figure 7 shows the distribution of projects per year and the average number of categories used by the projects each year. It does not take Waze and Uber apps into account as it is not possible to map when the current feature set of each of these projects has been defined. We highlight the period between 2011 and 2017, which concentrates most projects in the knowledge base. During this period, the average number of categories in a project ranges from 5.25 to 8. The average number of categories used per project for the whole knowledge base is 7.2. In addition, Table 1 shows the top-10 categories used by the projects in the knowledge base (KB) – as expected, geographic coordinates representing the location of an object ranks number 1, being used by 53.4% of the projects in the KB.

The aggregation of how many context elements categories each of the analyzed projects use, grouped by the supra-categories *activity*, *identity*, *location*, and *time*, can be observed in Annex A: https://amirton-chagas.github.io/CETITS-AnnexA.pdf. It is important to note that Waze and Uber use context elements from far more categories (30 and 25, respectively) than the average of the projects found in the reviewed articles, i.e., about four times more than in the other projects. Two hypotheses have been put forward: 1. Mature commercial products evolve, get richer in features, and naturally use more context elements, hence, more categories; 2. Research projects might use more context elements than it has been possible to capture from their publications. Further research could check whether any of these hypotheses are valid.

5 Discussion

Our model has been defined to summarize and create a comprehensive taxonomy of Context elements contributing to help the development of vehicular applications. Using Dey *et al*. [2001] as a starting point, we used four basic context types to create a structure that can categorize such elements, calling them *supra-categories* in this work. This effort resulted in a hierarchical model with 57 leaf categories (those with no child subcategories), from a total of 79 categories in the model, including the four supra-categories and all intermediary categories between them and the leaf categories. As far as we could trace, this is the broadest number of context element categories documented in one single place, specially designed for the vehicular application domain.

Designing an application using the proposed model was helpful, and some of the features emerged by observing the model and identifying possibilities related to the application's core idea. When comparing our model to the work of Kannan *et al*. [2010], we could check that our proposed taxonomy is indeed more general, and some of the features of the application designed as validation of this model would not be able to be modeled using the more specific domain existent in Kannan *et al*. [2010].

However, we have identified a limitation of the hierarchical format. The choice of Identity, Location, Time, and Activity was consistent with other contextual models aiming at a more generic model, but, in this scenario, another valid alternative would be to root the model using the categories Vehicle, Driver, Passengers, and Environment. A model with similar expressiveness can be created if this is considered, and this observation will be taken into account when we evolve this model into an ontology, which is capable of representing both relationships.

Before the ontology, we believe there is room for more formal qualitative research to improve the validation of the semantic value of this taxonomy. We would need to elicit which measurements could be used as evidence of the effectiveness and afterward collect and analyze them.

In summary, this model is efficient and complete enough to help design a context-aware ITS application. Although only one application was developed for validation, we believe the model is generic enough to be helpful in most applications of this vast domain of ITS. This belief is based on the robust literature we used to identify the context elements and categorize them into this taxonomy.

6 Conclusions

This paper showed that this work resulted in two products, a taxonomy and a knowledge base, and a context-aware vehicular application (an initial implemented version of the Convoy application, as mentioned in Section 4.2).

The main contribution is the taxonomy consisting of 79 context elements categories commonly used in ITS, and this taxonomy can help software developers design richer context-aware applications in this domain by using more context elements to define a context. The validation experiment showed that engineers designing an ITS project using the proposed taxonomy during application design correctly used 2.7 times more context elements in their applications when compared to engineers who designed the project without the help of our taxonomy.

Through the knowledge base, created based on the literature review from 1998 to 2021, the developer knows which and how many ITS projects use which categories of context elements. That can further help engineers in the design process of a new ITS project, as they can browse the knowledge base, find projects similar to the one they are designing, and verify that the context elements used in those projects are adequate for their new ones. We currently have work in progress using this knowledge base to create a recommendation system to further ease the task of designing ITS applications.

In another validation effort, a context-aware ITS application was specially developed with the support of the taxonomy to guide the application designer in identifying the context elements that might be useful in it. Following the other validation experiment, 16 categories of context elements that meet the application requirements were elicited with the support of the taxonomy. This is almost 2.3 times or 128% of the average number of categories used in the projects of the knowledge base, i.e. 7.2, which is also in line with the results of the validation experiment.

The process used to design, evaluate and evolve the taxonomy proposed in this work for the ITS domain could be followed to create similar structures for other domains. More research is needed to design and validate such taxonomies in other domains and to assess their usefulness in these environments. Another direction for future work is to use the proposed model as a basis to define and build a formal ontology that can also be used at runtime.

Declarations

Authors' Contributions

Amirton Chagas: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Carlos Ferraz**: Writing - Review Editing, Supervision, Methodology, Validation.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Materials are available on the links provided in the article. In case the links become broken, the materials can be made available upon request.

References

Abowd, G. D. and Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction*, 7(1):29–58. DOI: 10.1145/344949.344988.

Arooj, A., Farooq, M. S., Akram, A., Iqbal, R., Sharma, A., and Dhiman, G. (2022). Big Data Processing and Analysis in Internet of Vehicles: Architecture, Taxonomy, and Open Research Challenges. *Archives of Computational Methods in Engineering*, 29(2):793–829. DOI: 10.1007/s11831-021-09590-x.

Atzori, L., Iera, A., and Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805. DOI: 10.1016/j.comnet.2010.05.010.

Baras, S., Saeed, I., Tabaza, H. A., and Elhadef, M. (2018). VANETs-based intelligent transportation systems: An overview. In *Lecture Notes in Electrical Engineering*, volume 474, pages 265–273. Springer Verlag. DOI: 10.1007/978-981-10-7605-3_44.

Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., and Riboni, D. (2010). A survey of context modelling and reasoning techniques. *Pervasive and Mobile Computing*, 6(2):161–180.

Bifulco, F., Amitrano, C. C., and Tregua, M. (2014). Driving smartization through intelligent transport. *Chinese Business Review*, 13(4).

Brezillon, J. and Brezillon, P. (2007). Context modeling: Context as a dressing of a focus. *Proceedings of the 5 th International and Interdisciplinary Conference on Modeling and Using Context*.

Chagas, A. and Ferraz, C. (2012). ConnecTV: An environment to develop and run component-based customized iDTV applications. In *WebMedia'12 - Proceedings of the 2012 Brazilian Symposium on Multimedia and the Web*. DOI: 10.1145/2382636.2382657.

Chavhan, S., Gupta, D., Nagaraju, C., A, R., Khanna, A., and Rodrigues, J. J. (2021). An Efficient Context-Aware Vehicle Incidents Route Service Management for Intelligent Transport System. *IEEE Systems Journal*. DOI: 10.1109/JSYST.2021.3066776.

David, B., Xu, T., Jin, H., Zhou, Y., Chalon, R., Zhang, B., Yin, C., and Wang, C. (2013). User-oriented system for smart city approaches. In *12th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, volume 46, pages 333–340. Elsevier. DOI: 10.3182/20130811-5-US-2037.00072.

Dey, A., Abowd, G., and Salber, D. (2001). A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. *Human-Computer Interaction*, 16(2):97–166. DOI: 10.1207/S15327051HCI16234_02.

Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5(1):4–7. DOI: 10.1007/s007790170019.

Dibaei, M., Zheng, X., Jiang, K., Abbas, R., Liu, S., Zhang, Y., Xiang, Y., and Yu, S. (2020). Attacks and defences on intelligent connected vehicles: a survey. *Digital Communications and Networks*, 6(4):399–421. DOI: 10.1016/J.DCAN.2020.04.007.

Dzemydienė, D. and Burinskienė, A. (2021). Integration of Context Awareness in Smart Service Provision System Based on Wireless Sensor Networks for Sustainable

Cargo Transportation. *Sensors 2021, Vol. 21, Page 5140*, 21(15):5140. DOI: 10.3390/S21155140.

Environmental Protection Agency (2021). Code of Federal Regulations. 40 CFR 156.62 - Toxicity Category.

Essien, A., Petrounias, I., Sampaio, P., and Sampaio, S. (2020). A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. *World Wide Web 2020 24:4*, 24(4):1345–1368. DOI: 10.1007/S11280-020-00800-3.

Figueiredo, L., Jesus, I., Machado, J. A. T., Ferreira, J. R., and de Carvalho, J. L. M. (2001). Towards the development of intelligent transportation systems. In *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)*, pages 1206–1211, Oakland, CA, USA. IEEE. DOI: 10.1109/itsc.2001.948835.

Gomes, J. Z., Victória Barbosa, J. L., Geyer, C. F. R., Anjos, J. C. S. D., Canto, J. V., and Pessin, G. (2020). Ubiquitous intelligent services for vehicular users: A systematic mapping. *Interacting with Computers*, 31(5):465–479. DOI: 10.1093/iwcomp/iwz030.

Guerrero-Ibanez, J. A., Zeadally, S., and Contreras-Castillo, J. (2015). Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and internet of things technologies. *IEEE Wireless Communications*, 22(6):122–128. DOI: 10.1109/mwc.2015.7368833.

Hu, X., Li, X., Ngai, E. C., Leung, V. C., and Kruchten, P. (2014). Multidimensional context-aware social network architecture for mobile crowdsensing. *IEEE Communications Magazine*, 52(6):78–87. DOI: 10.1109/MCOM.2014.6829948.

Johnson, D. A. and Trivedi, M. M. (2011). Driving style recognition using a smartphone as a sensor platform. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1609–1615. IEEE. DOI: 10.1109/ITSC.2011.6083078.

Kaltz, W. J., Ziegler, J., and Lohmann, S. (2005). Context-aware Web engineering: Modeling and applications. *Revue d'intelligence artificielle*, 19(3):439–458.

Kannan, S., Thangavelu, A., and Kalivaradhan, R. (2010). An intelligent Driver Assistance System (I-DAS) for vehicle safety modelling using ontology approach. *International Journal of UbiComp*, 1(3):15–29.

Khekare, G. S. and Sakhare, A. V. (2012). Intelligent traffic system for VANET: a survey. *International Journal of Advanced Computer Research*, 2(6):99–102.

Klotz, B., Troncy, R., Wilms, D., and Bonnet, C. (2018). VSSo-A Vehicle Signal and Attribute Ontology. In *SSN Workshop at ISWC. CEUR Workshop Proceedings*.

Laña, I., Sanchez-Medina, J. J., Vlahogianni, E. I., and Ser, J. D. (2021). From Data to Actions in Intelligent Transportation Systems: A Prescription of Functional Requirements for Model Actionability. *Sensors 2021, Vol. 21, Page 1121*, 21(4):1121. DOI: 10.3390/S21041121.

Leyvand, T., Meekhof, C., Yi-Chen Wei, Jian Sun, and Baining Guo (2011). Kinect Identity: Technology and Experience. *Computer*, 44(4):94–96. DOI: 10.1109/MC.2011.114.

Li, Y., Tao, J., and Wotawa, F. (2020). Ontology-based test generation for automated and autonomous driving functions. *Information and Software Technology*, 117:106200. DOI: 10.1016/J.INFSOF.2019.106200.

Lohmann, S., Negru, S., Haag, F., and Ertl, T. (2016). Visualizing ontologies with VOWL. *Semantic Web*, 7(4):399–419. DOI: 10.3233/SW-150200.

Mitchell, M., Meyers, C., Wang, A.-I. A., and Tyson, G. (2011). Contextprovider: Context awareness for medical monitoring applications. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5244–5247.

Papatheocharous, E., Wnuk, K., Petersen, K., Sentilles, S., Cicchetti, A., Gorschek, T., and Shah, S. M. A. (2018). The GRADE taxonomy for supporting decision-making of asset selection in software-intensive system development. *Information and Software Technology*, 100:1–17. DOI: 10.1016/J.INFSOF.2018.02.007.

Rajaram, H. K., Loane, J., MacMahon, S. T., and Caffery, F. M. (2019). Taxonomy-based testing and validation of a new defect classification for health software. *Journal of Software: Evolution and Process*, 31(1):e1985. DOI: 10.1002/SMR.1985.

Santos, G. and Nikolaev, N. (2021). Mobility as a Service and Public Transport: A Rapid Literature Review and the Case of Moovit. *Sustainability 2021, Vol. 13, Page 3666*, 13(7):3666. DOI: 10.3390/SU13073666.

Sobral, T., Galvão, T., and Borges, J. (2020). An Ontology-based approach to Knowledge-assisted Integration and Visualization of Urban Mobility Data. *Expert Systems with Applications*, 150:113260. DOI: 10.1016/J.ESWA.2020.113260.

Soyturk, M., Muhammad, K. N., Avcil, M. N., Kantarci, B., and Matthews, J. (2016). Chapter 8 - From vehicular networks to vehicular clouds in smart cities. In Obaidat, M. S. and Nicopolitidis, P., editors, *Smart Cities and Homes*, pages 149 – 171. Morgan Kaufmann, Boston. DOI: https://doi.org/10.1016/B978-0-12-803454-5.00008-0.

Swarnamugi, M. and Chinnaiyan, R. (2020). Context—Aware Smart Reliable Service Model for Intelligent Transportation System Based on Ontology. *Lecture Notes in Electrical Engineering*, 597:23–30. DOI: 10.1007/978-3-030-29407-6_3.

Temdee, P. and Prasad, R. (2018). Context and Its Awareness. In *Context-Aware Communication and Computing: Applications for Smart Environment*, pages 15–31. Springer, Cham, Switzerland. DOI: 10.1007/978-3-319-59035-6_2.

Vahdat-Nejad, H., Ramazani, A., Mohammadi, T., and Mansoor, W. (2016). A survey on context-aware vehicular network applications. *Vehicular Communications*, 3(C):43–57. DOI: 10.1016/j.vehcom.2016.01.002.

Vieira, V., Caldas, L. R., and Salgado, A. C. (2011). Towards an Ubiquitous and Context Sensitive Public Transportation System. In *2011 Fourth International Conference on Ubi-Media Computing*, pages 174–179, São Paulo, Brazil. IEEE. DOI: 10.1109/U-MEDIA.2011.19.

Vieira, V., Tedesco, P., and Salgado, A. C. (2005). Towards an ontology for context representation in groupware. In *11th International Workshop, CRIWG*, volume 3706, pages 367–375, Porto de Galinhas, Brazil.

Vieira, V., Tedesco, P., and Salgado, A. C. (2009). A process for the design of Context-Sensitive Systems. In *2009 13th International Conference on Computer Supported Cooperative Work in Design*, pages 143–148, Santiago, Chile. IEEE. DOI: 10.1109/cscwd.2009.4968049.

Villegas, N. M., Müller, H. A., Muñoz, J. C., Lau, A., Ng, J., and Brealey, C. (2011). A dynamic context management infrastructure for supporting user-driven web integration in the personal web. In *Proceedings of the 2011 Conference of the Center for Advanced Studies on Collaborative Research*, pages 200–214, Toronto, Canada. IBM Corp. Riverton, NJ, USA.

Winder, C., Azzi, R., and Wagner, D. (2005). The development of the globally harmonized system (GHS) of classification and labelling of hazardous chemicals. *Journal of Hazardous Materials*, 125(1-3):29–44. DOI: 10.1016/J.JHAZMAT.2005.05.035.

Wongcharoen, S. and Senivongse, T. (2016). Twitter analysis of road traffic congestion severity estimation. *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*. DOI: 10.1109/JCSSE.2016.7748850.

Woodard, M., Wisely, M., and Sarvestani, S. S. (2016). A Survey of Data Cleansing Techniques for Cyber-Physical Critical Infrastructure Systems. *Advances in Computers*, 102:63–110. DOI: 10.1016/BS.ADCOM.2016.05.002.

Xiong, Z., Dixit, V. V., and Waller, S. T. (2016). The development of an Ontology for driving Context Modelling and reasoning. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 13–18. DOI: 10.1109/ITSC.2016.7795524.

Zheng, X., Chen, W., Wang, P., Shen, D., Chen, S., Wang, X., Zhang, Q., and Yang, L. (2016). Big Data for Social Transportation. *IEEE Transactions on Intelligent Transportation Systems*, 17(3):620–630. DOI: 10.1109/tits.2015.2480157.

Zimmermann, A., Lorenz, A., and Oppermann, R. (2007). An Operational Definition of Context. In *Modeling and Using Context*, pages 558–571. Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-74255-5_42.