Location Data Analytics in the Business Value Chain: A Systematic Literature Review

LUIS E. FERRO-DÍEZ[1], NORHA M. VILLEGAS[1,3], (Senior Member, IEEE), AND JAVIER DÍAZ-CELY[1,2]

[1]Department of Information and Telecommunication Technologies, Universidad Icesi, Cali 760031, Colombia
[2]MO Tecnologias, Bogotá 110221, Colombia
[3]Department of Computer Science, University of Victoria, Victoria, BC V8P 5C2, Canada

Corresponding author: Norha M. Villegas (nvillega@icesi.edu.co)

ABSTRACT Context information has become a significant asset to optimize the value obtained from information systems. Location is an important type of context information that refers to the place in which an event occurs. In business environments, the implementation of location-based analytics systems to aid decision making processes is of paramount importance for business development. However, after an exhaustive literature review, we found that researchers and practitioners still lack a comprehensive characterization of location-based data analytics systems that have been effectively applied to business processes. This paper presents the results of a systematic literature review (SLR), in which we characterized a total of 168 location-based and business oriented analytics solutions that were published between 2014 and 2019. To conduct this SLR we defined three characterization dimensions: *business aspects*, through which we identified value chain business processes or activities that may be benefited with the proposed solution; *data source*, which allowed us to report on the data used in each of the studies; and *data analytics*, through which we report on the analytics techniques and validation strategies implemented by the studied approaches. The contribution of our SLR is twofold. First, it provides business and data analytics practitioners with a comprehensive catalog of location-based data analytics approaches that could be applied to improve value generation, at different levels, along their businesses' value chains. And second, it provides researchers with a complete landscape of recent advancements and open challenges in the field.

INDEX TERMS Business data processing, business value chain, location analysis, location context, location intelligence, machine learning.

I. INTRODUCTION

Nowadays, because of the proliferation of new technologies and services that facilitate the collection of location data [1], there is plenty of availability of location information. Furthermore, the evolution of technology has facilitated the realization of operational work in several aspects of our lives. In particular, data is very useful to drive decision making in operations [2]. In the business context, researchers have demonstrated that location information can be exploited in favor of the achievement of operational goals [3], [4].

The value of location information, as an asset to improve decision support mechanisms, has been studied thoroughly in several problem types [5]. However, we found that

researchers, specially practitioners interested in exploiting location context in the business domain, still lack comprehensive and domain-independent surveys, particularly systematic literature reviews, that connect common business needs with data analytics approaches, in which the usage of location information plays a major role. We found other eight surveys associated with business processes and location context, nevertheless, their results could be considered narrow since they focus on specific industries or analysis techniques.

With the goal of providing researchers and practitioners with a more comprehensive characterization of approaches that exploit location context and that are applicable to business processes, we surveyed 168 papers by following the guidelines proposed by Kitchenham *et al.* [6]. Our survey was conducted to answer five research questions that are

relevant to researchers and practitioners interested in this research field: *i)* what are common use cases using location information to generate business value? *ii)* what analytics techniques have been used to generate business value using location context? *iii)* what have been the primary sources for data extraction, and what kind of information was used? *iv)* how location information has been exploited to generate business value? And *v)* what metrics have been used to evaluate and validate the models?

The findings obtained from the papers characterized in our SLR are classified into three groups: *i) business process findings*, which refer to business process aspects that were addressed using the studied approaches; *ii) data source findings*, which refer to the primary data sources that were exploited in the studied solutions; and *iii) data analysis findings*, which report on the techniques, metrics, and validation strategies used or proposed by the analyzed studies. To characterize business process aspects, we revisited the value chain concept proposed by Porter [7].

Findings suggest that there is an increasing interest in this field. Regarding business process aspects, there are applications in all processes of the value chain, particularly in operational processes. Data source findings show that most of the studies use proprietary and surveyed data sets, although there are several studies using location-based social networks (LBSN). Finally, we collected and organized the data analysis findings to easily observe the problem types, data sources, and analysis techniques used by the authors of the surveyed studies. Location analysis, in particular applied to the facility location problem, was the most common approach among the studies.

This paper is structured as follows. Section II explains the background of this SLR. Section III discusses related work by analyzing the contributions of our SLR with respect to other similar surveys in the area. Section IV explains the methodology we followed to conduct this review. Sections V-VII present the findings and contributions of this work. Section VIII summarizes limitations and research opportunities we identified during the elaboration of this SLR. Finally, Section IX concludes the paper.

II. BACKGROUND
This section presents foundational concepts that are relevant to the application of location analytics to business value chain challenges.

A. THE BUSINESS VALUE CHAIN
The concept of *business value chain* was introduced by Porter in 1985 [7]. In his book, Porter describes the business value chain as the collection of activities a company performs in order to design, produce, market, deliver and support products. He also defines a set of primary activities that can be categorized into five groups:

- **Inbound logistics:** Activities related to receiving, sorting, handling and managing the inputs or raw material the company needs in order to create its products.

- **Operations:** Activities associated with the transformation of the inputs into the final product, including design, prototyping, testing, and fabrication.
- **Outbound logistics:** Activities associated with the collection, storage, management and distribution of the final product to the customers.
- **Marketing & sales:** Activities related to the promotion, communication, sales force and everything a company does to enable customers to purchase products.
- **Service:** Activities that are related to service providing, maintenance, adjustments, and support necessary to enhance or sustain the life of the product or acquisition.

The business value chain concept has matured well in the business domain, and it is actively used as a reference to elaborate business strategies to attain competitive advantage. It is also compared with other definitions, such as *core process*, and it is used as the basis for other business strategy concepts [8]. In this SLR, the analysis of the business dimension is based on Porter's business value chain concept.

B. LOCATION CONTEXT
According to Abowd *et al.*, context is "*any information useful to characterize the situation of an entity (e.g., a user or an item) that can affect the way users interact with systems*" [9]. In general, context information has the potential to improve the quality of results of data-oriented solutions.

There are several categories of context information, Villegas and Müller classify them as *Individual*, *Location*, *Time*, *Activity*, and *Relational*. Furthermore, they define location context as "*all the information about the place of settlement or activity of an object*" [10]. Location context is the subject of study in this paper.

In recent years, location context has been widely used in different industry domains such as biofuel [11], [12], transportation [13], [14], tourism [15], [16], and retail [17], [18], among others. This wide interest in location context demonstrates that the positional relationship that may exist among entities is relevant to imply other information that affects businesses and customers. Hence, investigating new methods and techniques to exploit location context in the business realm is not only relevant, but also generates a significant impact [19].

Traditional statistical analysis can be applied accordingly to spatial analysis, e.g., a researcher would want to look at the spatial distribution of the studied instances, and perform experiments. Moreover, location context can play an additional role. For example, two frequency distributions, numerically speaking, can be equivalent, however, the way they are located in the space can change how the data is interpreted. Fig. 1 shows this phenomenon graphically. In both images, we observe the same number of instances in the same number of squares, however, they have different spatial distributions. The spatial relationship between instances can significantly contribute to the discovery of latent information that might reveal additional insights of significant value for the use case.
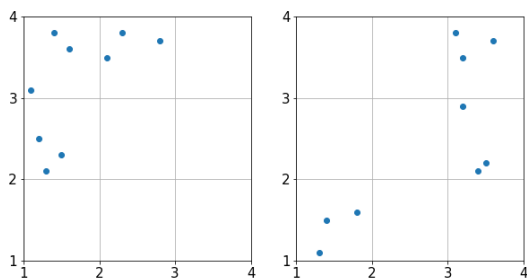
FIGURE 1. Spatial distributions with an equivalent number of cells and instances but different spatial patterns.

C. LOCATION CONTEXT PROBLEM TYPES

Based on the findings of our SLR, the following can be considered the most common location context problem types:

- **Location analysis:** Location analysis focuses on selecting the optimal location for a particular goal [20]. It includes the facility location and coverage problems. Actual applications may vary depending on the industry domain, e.g., selecting optimal locations for a factory or designing transportation routes to optimize service coverage [4].
- **Spatial analysis:** Spatial analysis extends traditional statistical analysis by including the spatial components of the data, such as the spatial relationship between instances [21]. There are several sub-types of analysis depending on the needs and context. For example, pattern identification, clustering, and geographic relationship analysis.
- **Recommendation systems:** Recommendation Systems emerged as a response to a question that is relevant to any industry sector: What are the most appropriate items for a particular user? There are several approaches to solve this problem, and some include context information. Recommender systems that rely on context information to improve their recommendation models are known as context-aware recommendation systems [22].
- **Point-of-interest (POI) recommendation:** This refers to a specific type of recommendation system where the item to recommend is a location itself. A POI is a specific place a user might find useful or interesting. Several applications have emerged from the need of recommending places (POIs) to users, for example Foursquare, TripAdvisor, Yelp, and Google Maps. As a result, in recent years, POI Recommendation has gained considerable popularity as a research field. We consider this as a separate problem type given its significance to the findings of our SLR.
- **Behavioral pattern detection:** Location context is exploited to detect or infer behavioral patterns among elements of a system. One common use case is exploiting the movement behavior of users to improve mobile services [23].
- **Predictive analysis:** This problem type is analogous to predictive analytics or supervised learning. The idea is to model data, entities or events, and predict values for new instances. The use cases are significantly diverse. In what respects to location context, some examples are predicting the best business type for a particular location [17], predicting business performance according to location and social context [24], and predicting the quality of mobile services according to user locations [25].
- **Trajectory tracking:** Trajectory data from entities is collected to be used as input to analyze movement and behavioral patterns. The goal is to learn movement patterns in conjunction with other variables, to build a model capable of predicting the next location for an entity or model routing patterns. Imai et al. [26] suggest that early destination prediction helps improve personalized services.
- **Location based relationships:** Using location data, it is possible to model similarities between two entities, as there might exist a latent relationship between them depending on how they move or the places they frequent [27], [28]. User similarity can be used for POI, users, or other recommendation types.
- **Location-aware advertising:** Location-aware advertising is a marketing technique aimed to do better-targeted advertising, i.e., to send ads related to a particular product or service to locations where there seems to be people more likely to engage with it. Ubiquitous devices are a crucial tool given their location sensing capabilities [29].
- **Matching:** Matching is an ephemeral association established between entities based on their current context and state, e.g., pairing two different entities where the first one has a request that must be satisfied by the second one. There must be a criteria that can match them together, and this relationship can last until the request is satisfied. Matching is common, for example, among services that depend on a geographical location like ride-sharing or transportation [14].
- **Location estimation:** Location estimation is the action of inferring the position of an entity based on other information about it. In the context of location-based social networks (LBSN), this is useful because, despite the importance of location context, a significant amount of data still does not include it [30]. Some possible reasons are: users choose not to share their location, or the samples are not collected with this attribute. The lack of geotagged samples create the need to identify the latent location from the available data [31].

D. DATA ANALYTICS

Data analytics is a common approach to address the challenges of the location context problem types presented before. The following are the types of data analytics identified in the papers studied in this SLR:

- **Non-predictive approaches:** Non-predictive approaches are commonly based on descriptive or unsupervised learning. Descriptive analysis approaches, based on the concept of *descriptive statistics* [32], seek to describe and summarize the main characteristics in a data set,

TABLE 1. Related work - Comparing our SLR with other surveys about location context.

Author/Year	SLR	Not focused on a specific industry domain	Not focused on a specific problem type	Business process findings	Data source findings	Data analysis findings
Lim et al. [37] - 2018				✓		✓
Murray [38] - 2019		✓				✓
Nielsen et al. [39] - 2017	✓		✓	✓	✓	
Stock [40] - 2018	✓		✓		✓	✓
Yang et al. [41] - 2014			✓	✓		✓
Pelka and Hellbrück [42] - 2016		✓	✓		✓	✓
Keenan and Jankowski [5] - 2019		✓	✓			
Toch et al. [43] - 2019		✓	✓		✓	✓
Our literature review	✓	✓	✓	✓	✓	✓

with the aid of statistical and mathematical modeling, and present the results in the form of tables, charts, and graphs, among other visualization options. In machine learning approaches, other *unsupervised learning* techniques can be used to perform numerical analysis and discover insights from the data [33].

- **Predictive approaches:** Predictive approaches are associated with predictive analytics or supervised learning. In these cases, data sets contain the target values to predict, and an algorithm is used to "learn" numeric relationships between the attributes and the expected value [33]. The prediction of continuous numeric values is called regression [2], [34], and the prediction of a class or category is called classification [2].

- **Inference:** Whereas machine learning is generally more focused on prediction, statistics is more concerned with inference, i.e., it focuses on how data is generated as a function of the input data [35], and on the relationships between the independent and the dependent features.

- **Optimization:** Mathematical or numerical optimization, according to Nocedal and Wright [36], is the process of numerically analyzing a physical system to formulate an objective $f(X_n)$ that depends on certain characteristics of the system. The objective should be something to improve or *optimize* e.g., *maximize* profit or *minimize* costs. The system characteristics are *variables* for which their optimal values are yet unknown, and the objective is often subject to *constraints*, i.e., the optimal solution should also meet a set of particular numeric constrains [36]. Findings suggest that most solutions to the facility location problem are addressed as mathematical optimization problems.

III. RELATED WORK

We found eight surveys published in relevant journals since 2010. These surveys consider location context as a primary aspect to analyze together with at least a particular aspect of a business or system. We did not consider reviews in which location context was a secondary attribute, since our main focus is precisely the exploitation of this information in favor of the results.

Aiming at a comprehensive understanding of the use cases, applications and techniques, our SLR is an industry & problem type independent characterization of location context usage in studies or solutions focused on the business value chain. We characterized the surveyed papers in light of three dimensions: *business process*, *data sources*, and *data analysis*.

Table 1 compares our SLR (last row) with the eight relevant location-related surveys found in the state of the art. This comparison is based on six criteria that we define as follows: *i) SLR*, the survey follows a systematic literature review [6] approach as methodology; *ii) Not focused on a specific industry domain*, the survey scope is not tied to a specific industry or sector, e.g., the study does not focus solely on the hotel industry; *iii) Not focused on a specific problem type*, the survey scope is not tied to a specific problem type, e.g., the study is not focused solely on optimization problems; *iv) Business process findings*, the survey includes business level characterization dimensions based on accepted business or operation strategy models; *v) Data source findings*, the survey contains data sources or data characterization dimensions; *vi) Data analysis findings*, the study elaborates on the solution approaches followed by the surveyed papers. The check mark (✓) in the corresponding cell indicates the study is compliant with that given criterion.

According to Table 1, four surveys focus on particular industry domains: Hotels & Tourism [37], [41]; multinational enterprises [39]; and social media [40]. Two reviews work on specific problem types: Recommendation systems [37], and Optimization [38]. Only three studies dedicated a section to elaborate on business details: Lim *et al.* [37] include dimensions to study operation aspects of tourism businesses; Nielsen *et al.* [39] elaborate on business processes of multinational enterprises (MNEs); and Yang *et al.* [41] elaborate on hotel industry business details. Finally, Keenan and Jankowski [5] survey over a considerable period of 30 years of *spatial decision support systems*, but they focus on building networks and publication relationships among the authors of the surveyed studies. None of the studies considered general and domain-independent business processes or activity aspects nor general technique usage.

IV. METHODOLOGICAL ASPECTS

This SLR followed the methodology suggested by Kitchenham *et al.* [6]. Our goal was to understand how location context is being used to enhance analytics in business areas. For this, we defined the following research questions:

- **RQ1:** What are common use cases using location information to generate business value?
- **RQ2:** What analytics techniques have been used to generate business value using location context?
- **RQ3:** What have been the primary sources for data extraction, and what kind of information was used?
- **RQ4:** How location information has been used to generate business value?
- **RQ5:** What metrics have been used to evaluate and validate the models?

We conducted a bibliographic search on journal papers published in ACM, IEEE, Scopus, ScienceDirect, and SpringerLink. These were selected because of the affiliation with the topic and the acceptable quality of their publications. The search string used was: *("location analytics" OR "location analysis" OR "location context" OR "location intelligence") AND (consumer OR customer OR client OR retail OR sales OR marketing OR logistics OR operation OR service)*.

The inclusion criteria consisted of: *i) publication date*, we selected papers published between 2014 and 2019; *ii) publication type*, we considered only Journal articles; *iii) the number of pages and language*, we excluded publications with less than eight pages and written in non-english languages; *iv) relevance*, we studied article titles and abstracts and verified the relevance to our research questions. Additionally, when the database offered the option to filter by area, we considered the following: *engineering, computer science, environmental sciences, decision sciences, business and management, earth sciences, economics and finances*. When the database did not offered this option, we inspected the results manually and considered only articles within the mentioned areas.

Fig. 2 shows the look-up and filtering process. The first search resulted in 3,223 articles. Due to query limitations with the logic connectors in ScienceDirect, we had to split the query into two parts, which generated some duplicates that we had to exclude manually. While inspecting the content of the articles, we excluded 2,848 during the first screening, then another 207 due to lack of relevance, or because it was not possible to identify the information required by our characterization dimensions.

Since we strictly focused on location context in business processes, we defined an additional set of inclusion and exclusion criteria regarding the content and purpose of each study. In short, we looked up for studies where location context played a mayor role in solving business process challenges using analytics techniques. Furthermore, we excluded studies focused on environmental settings such as weather, traffic, natural disasters, or waste control since we do not consider these cases to be easily traceable as
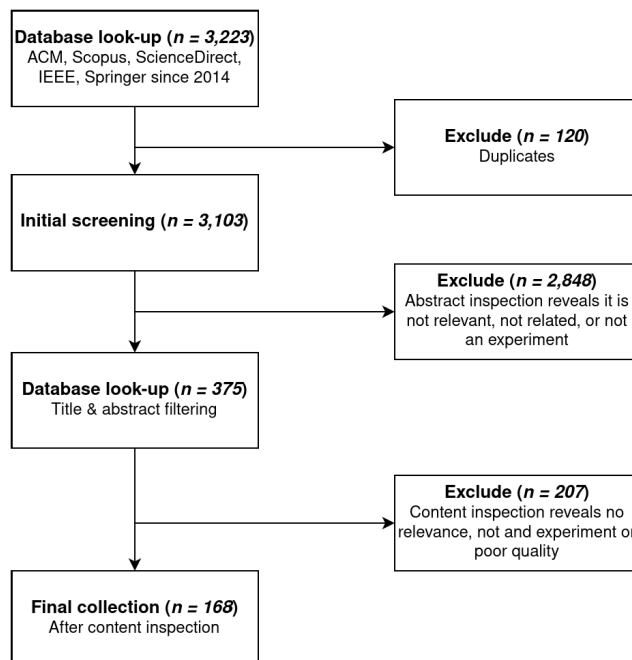

FIGURE 2. Article look-up process.

business processes. Similarly, we excluded studies related to human nature aspects such as culture, crime, and society since we consider these cases beyond of our scope. Finally, we excluded pure theoretical research work, particularly those focused on new mathematical methods and proofs without demonstrating their applicability. At the end, we got a total of 168 relevant articles that we characterized comprehensively.

V. BUSINESS PROCESS FINDINGS

This section summarizes the findings related to the contributions of the studied approaches to the business processes introduced in Section II-A, along with others aspects that we considered relevant to answer **RQ1**.

We observed a diverse set of industries exploiting location information in different problem types. Fig. 3 presents a cross between the business value chain processes and the problem types described in Section II. We observe the prevalence of *location analysis* and *spatial analysis* in different industries and segments. Also, we observe that with respect to *logistics*, there are only location and spatial analysis applications. Finally, the set of addressed problem types was diverse along different industries and value chain processes.

Next, we discuss our findings in light of three subdimensions: *i) Industry domain*, by identifying to what industry domain the study was being targeted; *ii) Business value chain*, by identifying and classifying the studied approaches with respect to the impact on the value chain processes; and *iii) Decision impact*, by identifying, at a high level, in terms of time and cost, the impact of the model or study output, and whether it was used to support strategic decisions or as an end user functionality.
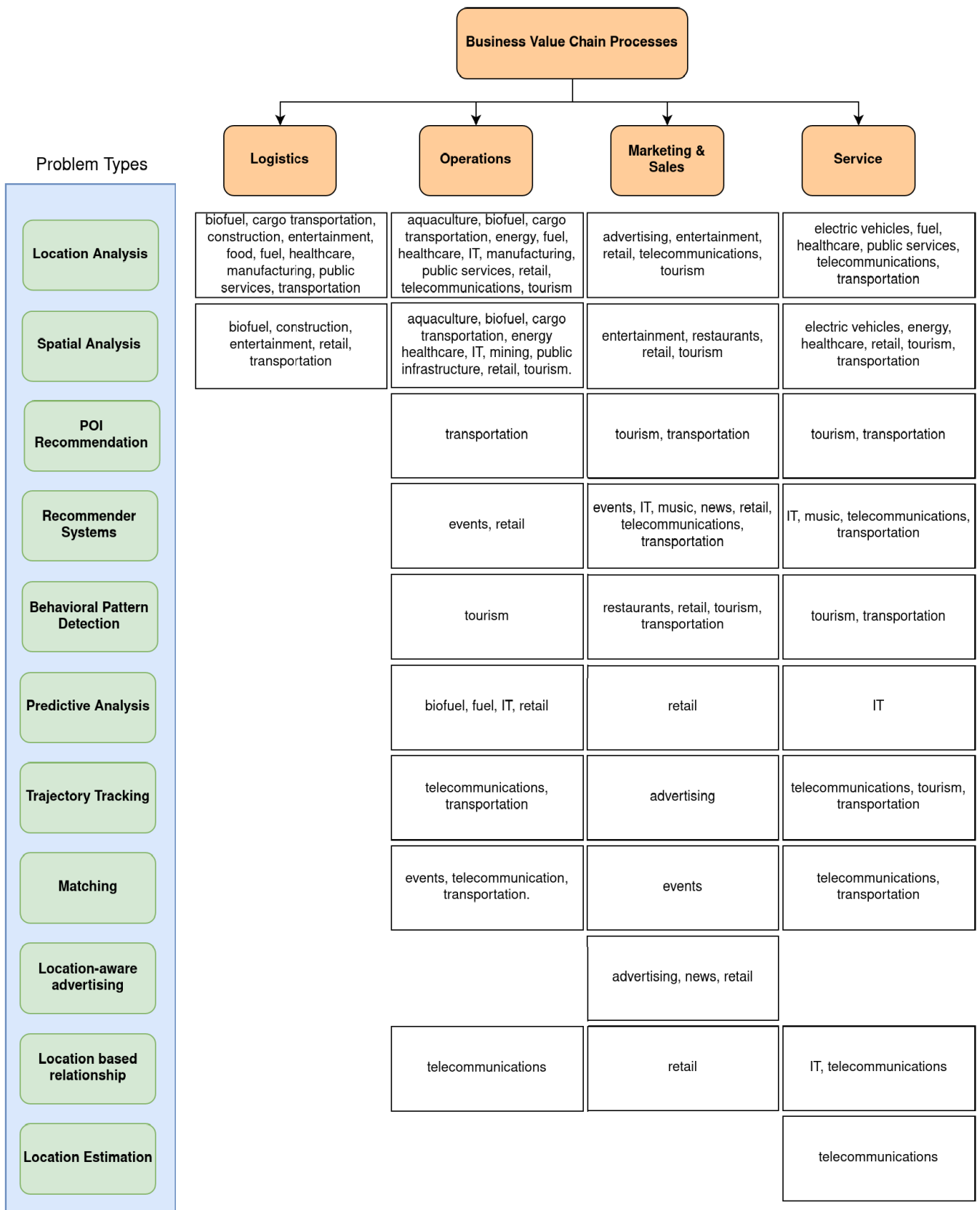
FIGURE 3. Location context problem types commonly addressed in different industry sectors accross business value chain processes.

TABLE 2. Publications per industry domain.

Industry Domain	Total Papers	%	References
Not specific	40	23.81%	[23]–[25], [28], [44]–[79]
Biofuel	17	10.12%	[11], [12], [80]–[94]
Transportation	14	8.33%	[13], [14], [95]–[106]
Health-care	13	7.74%	[107]–[119]
Retail	14	8.33%	[17], [18], [120]–[131]
Tourism	13	7.74%	[15], [16], [132]–[142]
Cargo transportation	10	5.95%	[143]–[152]
Energy	9	5.36%	[153]–[161]
Telecommunications	5	2.98%	[162]–[166]
Public services	4	2.38%	[167]–[170]
It	4	2.38%	[171]–[174]
Fuel	4	2.38%	[175]–[178]
Manufacturing	3	1.79%	[179]–[181]
Electric vehicles	3	1.79%	[182]–[184]
Advertising	3	1.79%	[185]–[187]
Other	12	7.19%	[188]–[199]
Total	**168**	**100%**	

A. INDUSTRY DOMAIN FINDINGS

We define *industry domain* as the particular economic activities or business types for which a study was applied to. We found 23 different industry domains (cf. Table 2). Within the element named *Other*, we included domains for which only one study was identified. These include: *aquaculture, mining, music, news, food, public infrastructure, entertainment,* and *construction.*

Table 2 also shows the proportion of papers per industry domain. The predominant category is ''*Not specific*'' which represents 23.81% of the studied papers. This group corresponds to those that were not specific in solving a unique business case, but their setting, findings and results are connected to real business problems and were validated with realistic data. One example of this category is the work of Önden [44], which elaborates on the single facility location problem in the context of logistics activities in Istanbul, Turkey. The study focuses on the inclusion of Geographic Information Systems (GIS) to location selection processes. Another interesting case is the work by Xin *et al.* [23], which elaborates on recommendation systems that exploit user mobile traces.

Biofuel, transportation, and health-care industries occasionally share interests in using location information in their respective use cases. Biofuel generally focuses on facility location problems, i.e., to find the best location to install refineries and processing facilities. Transportation, including cargo, also focuses on locating logistics stations such as bike sharing stations [13], [95]. Other studies focus on urban transportation networks [97], [100] and ride sharing [14], [103]. Health-care focuses mostly on designing attention center networks to optimize user accessibility [107], [108].

In retail, we found a diverse set of approaches, several of them focused on analyzing market behavior, prices and

economic conurbations [121], [125], [127]. Yu *et al.* [17] exploit location information for shop type recommendation, i.e., given a particular location, recommend the best shop type to open there. In Tourism, we found several approaches in hospitality that address both the company and market performance standpoints. For example, Gutiérrez *et al.* [134] apply spatial analysis to assess the impact of Airbnb in Barcelona.

B. BUSINESS VALUE CHAIN FINDINGS

Certain approaches were applied to more than one of the business value chain processes presented in Section II-A, hence, several studies were assigned to more than one process. The categorization was a manual inspection task that implied understanding the situation and analyzing it in light of the processes of the value chain.

As shown in Table 3, we found that the majority of studies relate to *operation* processes. For instance, Guaita-Pradas *et al.* [161] studied the sustainable development of solar power in Valencia, Spain, with the purpose of analyzing geographical regions in terms of their solar radiation impact and the viability of locating photo voltaic panels in that geographic region. Another example is the study by Huang *et al.* [105], which focuses on solutions for taxi-related services.

Service was the second most addressed value chain process. We found diverse applications on public transportation services [97], [106], cargo transportation and delivery [102], as well as IT services, in particular to predict the Quality of Service for Web services [173].

In *logistics*, the problems tended to be more related to facility (station) location [13], [95] and distribution network design [55], [152], [196].

Marketing & sales approaches were more focused on customers, i.e., exploiting user information to improve communication and advertising campaigns. For example, by exploiting user generated content that includes geographical information to improve traveling marketing decisions [15] and recommend POI [79].

C. DECISION IMPACT FINDINGS

We found a total of 138 studies from which we could extract the impact of the proposed solutions on the decision making process for the corresponding business. We characterized this impact using two sub-dimensions: *i) Decision impact*, to assess, at a high level, the impact of the decision in terms of time and cost; *ii) Decision type*, to assess whether the solution was used to support the business strategy, or to implement a service for end users.

To characterize the impact of the decisions made with the surveyed approaches, we defined three categories:

- **Long term:** The decision has a high economical and/or timely impact, i.e., the decision implies a significant investment, and/or is expected to last from months to years. For example, building a new facility, making a significant investment, or defining a time consuming operation.

TABLE 3. Publications per value chain process. (Percentages calculated against the total surveyed papers).

Value Chain Process	Total Papers	%	References
Logistics	40	23.81	[11], [13], [44], [45], [48]–[50], [53], [55], [56], [65], [78], [82], [87], [88], [95], [96], [101], [108], [109], [113], [123], [143]–[148], [150]–[152], [170], [175], [177], [179]–[181], [196], [198], [199]
Marketing & sales	51	30.36	[15]–[17], [23], [24], [46], [56], [57], [59], [61]–[63], [66]–[70], [72]–[75], [79], [120]–[122], [124]–[131], [135]–[138], [140]–[142], [162], [163], [185]–[188], [190], [191], [194], [195], [198]
Operation	80	47.62	[11], [12], [14], [17], [18], [44], [51], [52], [54], [58], [64], [65], [71], [76], [78], [80]–[87], [89]–[94], [97], [99], [100], [102], [104], [105], [111]–[113], [117], [119], [121], [122], [124], [132]–[134], [136]–[138], [145], [148]–[161], [165]–[171], [174], [176], [178], [179], [181], [189], [192], [193], [197]
Service	58	34.52	[13]–[16], [18], [23], [25], [28], [47]–[49], [57], [60], [62], [67], [71], [77], [95]–[112], [114]–[119], [133]–[135], [139], [153], [162]–[165], [168], [172], [173], [175], [182]–[184], [194]

TABLE 4. Publications per decision impact. (Percentages calculated against the decision informed papers).

Decision Impact	Total Papers	%	References
Long term	75	40.76	[11], [12], [17], [24], [44], [45], [47]–[50], [52], [54], [55], [60], [64], [76], [78], [80]–[85], [87]–[93], [95]–[97], [107], [108], [111]–[114], [117], [118], [121], [126], [127], [129], [132], [137], [138], [143], [147], [149]–[154], [156], [157], [159], [168], [169], [171], [175]–[184], [192], [197], [199]
Mid term	53	28.80	[12]–[15], [24], [45], [48], [50], [53], [55], [58], [60], [65], [70], [78], [82], [87], [88], [92], [95]–[97], [102], [107]–[111], [113], [121], [125], [127], [141]–[143], [145], [146], [148], [150], [168], [170], [174], [175], [178], [181], [182], [184]–[186], [189], [196], [198], [199]
Short term	56	30.43	[14], [15], [23], [25], [28], [46], [57]–[59], [61], [62], [65]–[75], [77], [79], [101]–[106], [109], [115], [131], [139]–[141], [144]–[146], [148], [150], [162], [164], [165], [165], [167], [170], [172]–[174], [178], [187], [188], [194]–[196]

- **Mid term:** The decision has a moderate economical and/or timely impact, i.e., a moderate investment decision that is expected to last from days to months. For example, changing distribution routes, or adjusting seasonal parameters.
- **Short term:** The decision has a small economical and/or timely impact, i.e., the decision is made immediately, has low individual cost, and is expected to last from minutes to days. For example, a user accepting an item recommendation.

Decisions made with the surveyed approaches were characterized through the following values:

- **Strategic:** It refers to a business strategic decision that is taken internally.
- **End user:** It refers to a functionality provided to end users.

The majority of studies (40.8%) were related to long term decisions. For example, all facility location use cases correspond to this category since there are high costs and expectations associated with new facilities. Mid term decisions constitute 28.8% of the papers, which includes studies associated with travel route planning [15], [70], [141], [142]; prediction of event participants [189]; as well as resource optimization and/or calibration during operation [65], [167], [174], [178]. Finally, short term decisions constitute 30.4% of the studies, where we found recommendation systems

applied to telecommunications [163] and logistics [148], [170], among others.

With respect to decision type, the majority of the surveyed studies focus on strategic business decisions (81.5%), such as facility location problems or distribution chain planing. Industries such as biofuel, fuel, health-care, retail, energy, and tourism show great interest in this area. For example, in retail, selecting a shop type given a particular location is an internal strategic decision [17]. Regarding end user oriented functionalities (18.5%), we found diverse applications such as carpooling recommendation systems [103], event organization planning [188], and travel marketing decision planning [15].

Tables 4 and 5 classify the surveyed approaches according to their decision impact and decision type, respectively.

VI. DATA COMPONENT FINDINGS

To answer **RQ3**, we analyzed the data information disclosed in the surveyed studies, and documented which were the most common data sources.

Table 6 summarizes data source findings. We found that 48.81% of the studies reported usage of *proprietary data sources*, whereas 9.52% of the studied papers used *surveyed data sources*, which were collected by researchers through surveys or computing programs. Another category is *synthetic data sources*, which correspond to data generated by

TABLE 5. Publications per decision type. (Percentages calculated against the decision informed papers).

Decision Type	Total Papers	%	References
End user	34	18.48	[15], [15], [23], [28], [46], [57], [59], [61], [66]–[70], [70], [72]–[75], [77], [79], [103], [104], [131], [139]–[141], [141], [142], [162], [165], [188], [189], [194], [195]
Strategic	150	81.52	[11], [12], [12]–[14], [14], [17], [24], [24], [25], [44], [45], [45], [47], [48], [48]–[50], [50], [52]–[55], [55], [58], [58], [60], [60], [62], [64], [65], [65], [71], [76], [78], [78], [80]–[82], [82]–[85], [87], [87], [88], [88]–[92], [92], [93], [95], [95], [96], [96], [97], [97], [101], [102], [102], [105]–[107], [107], [108], [108], [109], [109]–[111], [111]–[113], [113]–[115], [117], [118], [121], [121], [125]–[127], [127], [129], [132], [137], [138], [143], [143]–[145], [145], [146], [146]–[148], [148]–[150], [150], [150]–[154], [156], [157], [159], [164], [165], [167], [168], [168]–[170], [170]–[174], [174], [175], [175]–[178], [178], [178]–[181], [181], [182], [182]–[184], [184]–[187], [192], [196], [196]–[199], [199]

the researchers. *Undisclosed data sources*, corresponding to 7.74% of the surveyed papers, are those that were not explicitly mentioned by the authors.

Several other studies declared to use LBSN data such as Twitter, Gowalla, Yelp, Dianping, Baidu, Flickr, Foursquare and Tripadvisor. The studies using LBSN data cite the pervasiveness, volume, and information quality as significant factors for their usage.

Other data sources correspond to dedicated location information providers such as the Environmental Systems Research Institute (ESRI[1]), Openstreetmap,[2] and Google maps/earth. ESRI, for example, has a strong presence in business systems. Platforms, such as ArcGIS,[3] facilitate spatial analysis.

Finally, under the *other* category, we grouped all the data sources that appeared less than three times among the surveyed studies. This category includes Facebook, Panoramio, Doubanevent, Geodata, WS-Dream, Movielens, YouTube, Wikipedia, Wind nrel, T-drive, Plancast, 17life, Meetup, Last.fm, Lashou, Airbnb, Groupon, Gomaji, Factual, Ditu, Coin-or-foundation, Callcab, Brightkite, Beijing taxi passenger, Amazon, Alibaba, and Kaggle.

The nature of the variables included in the data sets is very diverse. In facility location problems, particularly in fuel refineries, the environmental aspects are crucial to choose where to extract the natural resources from, and the implications of the logistics costs of the selected location [80], [81], [83], [84]. In Tourism, environmental features can affect business dynamics and user preferences, e.g., prices, distances, and landmarks. [16], [135], [137], [138].

Use cases involving business and operational data range from exploiting city dynamics for car-sharing solutions [14] to planning distribution networks and hub-port locations [96], [144], [145]. Retail is a similar case where business data is useful for market and customer analysis [121]–[123], [127].

Demographic data is widely used in solutions dealing with customers from the market analysis perspective, which is expected. However, it was also used in other applications such as health-care facility and equipment location

problems [110], [111], [113], as well as urban and transportation analysis [13], [95], [98], [101]. User and item information, usually accompanied with social data, are used as the natural definition of entities in recommendation systems in different applications such as coupon recommendations [131], and location-aware news feeds [195].

As expected, economy data was mostly used in business performance related approaches. Some studies focused on external factors, such is the case of Teng *et al.* [51], who analyzed the performance of multinational enterprise (MNE) subsidiaries using country wide economic indicators [171].

In terms of geographic scope, only two studies reported usage of data originated from worldwide. An *et al.* [63] studied audiences and customer segmentation for online platforms using data generated from YouTube. The other case was the approach by Liu *et al.* [173], which used the WS-Dream[4] data set that contains web service log requests labeled with different locations for a quality of service prediction and service recommendation system. Both cases use a discrete representation of location information, i.e., the location was just a name or label referenced from where the event occurred.

Similarly, only two studies reported to work within a single venue. East *et al.* [133] combined GPS traces with survey data from customers visiting the Marwell Zoo in Hampshire, UK, in an interesting approach to discover the behavioral patterns of customers within the installations of the Zoo. The other case was Cheng and Shen [194], who worked on the concept of venue aware recommender systems for music, and used a venue-labeled music data set in which songs were tagged with a venue type such as *gym, restaurant, and library*.

VII. DATA ANALYSIS FINDINGS

This section summarizes the findings related to the data analysis approaches applied by the authors in the surveyed papers. In order to answer research questions **RQ2**, **RQ4**, and **RQ5**, we characterized these findings in four sub-dimensions, *i) problem type*, a categorization of the studies by the problem types defined in Section II-C; *ii) techniques*, a summary of the techniques used by the researchers to solve the addressed

[1]https://www.esri.com/

[2]https://www.openstreetmap.org/

[3]https://www.arcgis.com

[4]http://wsdream.github.io/

TABLE 6. Publications per data source.

Data sources	Total Papers	%	References
Proprietary	82	48.81 %	[11]–[13], [44], [45], [47], [49]–[51], [53]–[56], [58], [60], [64], [76], [78], [82]–[84], [86]–[90], [93]–[98], [101], [102], [106]–[111], [113], [114], [116]–[120], [122]–[124], [127], [128], [132], [134]–[138], [143], [148], [149], [153], [155], [156], [158], [160], [161], [163], [167], [171], [175], [178]–[183], [192], [193], [196]–[198]
Surveyed	16	9.52 %	[16], [18], [28], [57], [121], [123], [133], [147], [148], [164], [167], [170], [176], [190], [194], [196]
Esri	15	8.93 %	[45], [81], [95], [108], [113], [128], [133], [135], [149], [154], [168], [170], [191], [198], [199]
Undisclosed	13	7.74 %	[48], [81], [99], [100], [105], [115], [129], [130], [157], [166], [174], [185], [187]
Foursquare	9	5.36 %	[57], [61], [66]–[68], [72], [74], [104], [195]
Twitter	7	4.17 %	[59], [68], [71], [74], [125], [186], [189]
Gowalla	6	3.57 %	[66], [67], [69], [72], [75], [77]
Dianping	6	3.57 %	[17], [23], [73], [74], [129], [162]
Openstreetmap	4	2.38 %	[47], [121], [171], [183]
Synthetic	4	2.38 %	[24], [64], [104], [123]
Flickr	4	2.38 %	[15], [70], [139], [141]
Google maps/earth	3	1.79 %	[15], [113], [132]
Tripadvisor	3	1.79 %	[79], [135], [140]
Baidu	3	1.79 %	[17], [124], [138]
Yelp	3	1.79 %	[24], [62], [79]
Other	36	21.43 %	[25], [46], [58], [61], [63], [66], [70], [71], [75], [75], [79], [84], [103], [129], [131], [131], [131], [131], [131], [131], [134], [134], [139], [142], [142], [142], [142], [144], [151], [172], [173], [188], [189], [189], [191], [194], [199]

problems; *iii) metrics and validation strategies*, a summary of the metrics and validation strategies reported in the studies to validate their findings; and *iv) location information usage*, a classification of the studies according to the role played by location context.

Fig. 4 presents a taxonomy of the data analysis findings. Columns refer to the four data analysis types, and rows correspond to the problem types. This information should be read as the frequently used techniques and data sources for a given cross between a data analysis type and an problem type. For example, we observed that approaches that were applied to *location analysis* (problem type) were based on *optimization* (analysis type), and the frequently used techniques were *mathematical optimization, genetic algorithms, matrix factorization*, and *regression*.

Findings show that predictive analysis has implementations in almost all of the problem types with a wide variety of techniques and data sources. This suggests that the interest in exploiting location context in the business realm is an interesting research topic and it is attracting researchers with diverse backgrounds.

Another fascinating aspect is the usage of LBSN data in several problems and analysis types. Researchers have been showing interest in exploiting LBSN information given their rich geotagged data options that they can offer. Further details about data analysis findings are presented in the following subsections.

A. PROBLEM TYPE FINDINGS

Table 7 summarizes problem type findings. We found that 50.6% of the studies fall into the location analysis

problem type. Within location analysis, *the facility location problem* stood out as an important problem in different contexts. There is high interest in researching about the optimal location to install new facilities, according to industry and business needs, by considering the sustainability and economic viability of the locations as decisive investment factors [11], [12], [80], [87], [91].

For example, in the public transport subdomain, the bicycle share station location problem was commonly studied [13], [95]. Another commonly studied approach was the optimization of transportation networks [96], [97], [144], [150].

The second most addressed problem type was *spatial analysis*, which represents 32.74% of the papers. All the studies that performed some level of descriptive analysis, inference, or observation of spatial properties of events used spatial statistics techniques to study particular phenomena by exploiting location context. Some cases focused purely on analyzing variables related to location context. For example, analyzing hotel room price variations according to location context [135], as well as identifying retail centers and customer interactions according to their intra-market movement patterns [18], [125], [127].

Spatial analysis was also a complementary tool in other cases where a clear business objective was the target of the study. For instance, North and Miller [198] first studied the geographical characteristics of a particular region in Bavaria, Germany, and then formulated a location analysis problem to deploy an entertainment troop seeking at maximizing audience. Another example is Yang *et al.* [191] who studied the geographical
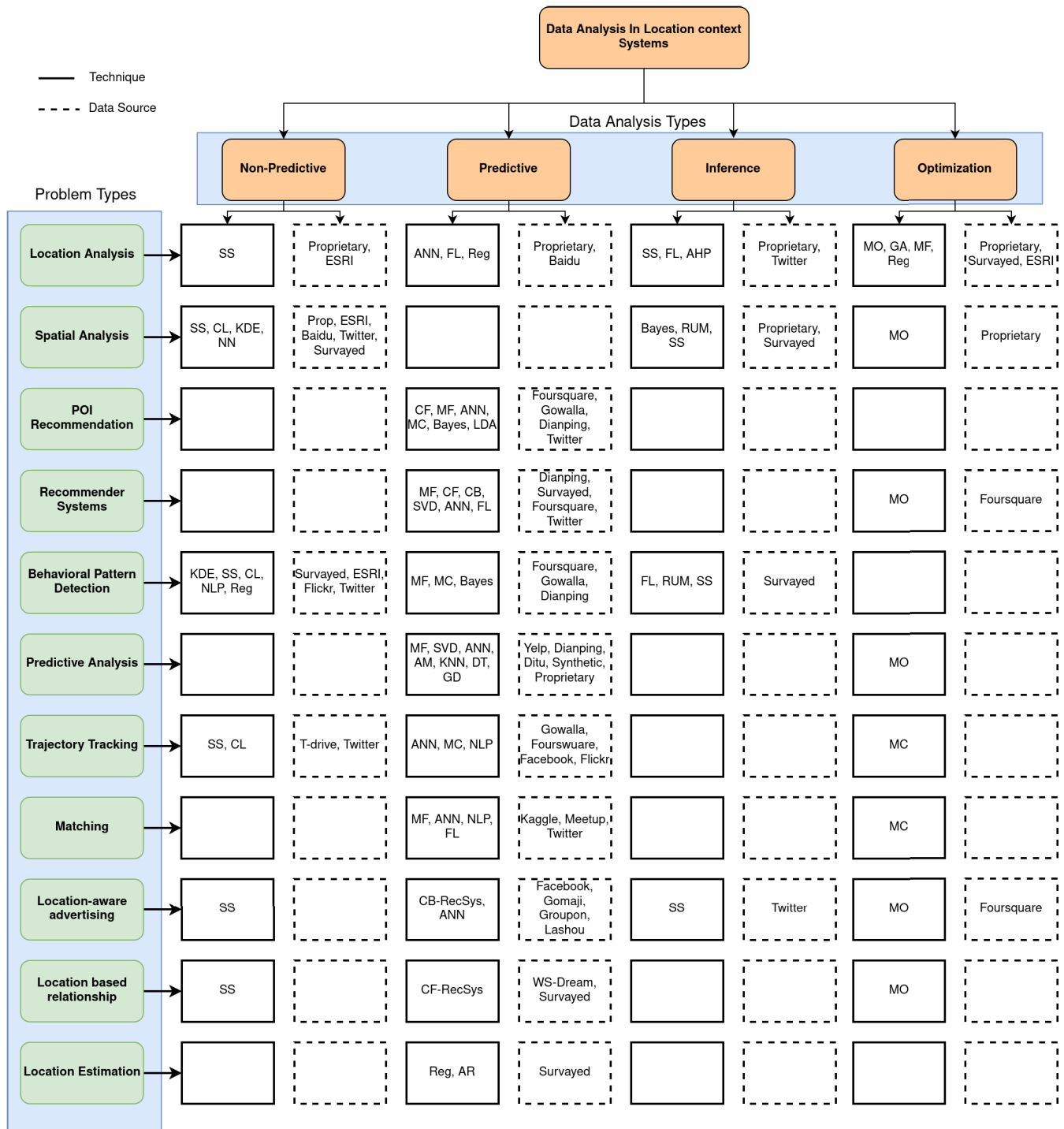
FIGURE 4. Taxonomy of data analysis in location context systems. AF: Affinity Models, AHP: Analytic Hierarchy Process, ANN: Artificial Neural Networks, AR: Association Rules, CB: Content Based, CF: Collaborative Filtering, CL: Clustering, DT: Decision Trees, FL: Fuzzy Logic, GA: Genetic Algorithms, GD: Gradient Descent, KDE: Kernel Density Estimation, KNN: K Nearest Neighbors, LDA: Latent Dirichlet Allocation, MC: Markov Chain, MF: Matrix Factorization, MO: Mathematical Optimization, NLP: Natural Language Processing, Reg: Regression, RUM: Random Utility Maximization, SS: Spatial Statistics, SVD: Singular Value Decompositio.

planing of restaurants in US, based on socio-demographic characteristics.

POI recommendation has received an increasing attention in recent years because of the availability of fine grained geotagged data, and the adoption of GPS enabled devices and LBSN [66], [72], along with opportunities derived from activities such as tourism, leisure, restaurants, retail, among other location related ones. POI recommendation was the

TABLE 7. Publications per problem type.

Problem Type	Total Papers	%	References
Location analysis	85	50.6%	[11]–[13], [44], [45], [47]–[50], [52]–[55], [60], [64], [65], [76], [78], [80]–[91], [95]–[97], [101], [102], [104], [107]–[109], [111]–[115], [117], [118], [121], [126], [138], [143]–[148], [150]–[152], [155], [157], [159], [161], [163], [166]–[171], [174]–[184], [186], [192], [196], [198], [199]
Spatial analysis	55	32.74%	[11], [12], [14], [16], [18], [51], [56], [71], [84]–[87], [89], [91]–[95], [98]–[100], [106], [107], [110], [116], [119]–[125], [127], [128], [130], [132], [134]–[137], [149], [153], [154], [156], [158], [160], [161], [171], [182], [191]–[193], [197]–[199]
POI recommendation	17	10.12%	[23], [46], [61], [66]–[70], [72]–[74], [79], [105], [139]–[142]
Recommendation systems	17	10.12%	[17], [25], [28], [57], [59], [61], [75], [77], [103], [129], [131], [162], [173], [188], [189], [194], [195]
Behavioral pattern detection	13	7.74%	[15], [16], [23], [25], [63], [71], [72], [98], [101], [106], [125], [133], [190]
Predictive analysis	10	5.95%	[17], [24], [25], [58], [62], [92], [129], [172], [174], [178]
Trajectory tracking	9	5.36%	[67], [71], [103]–[105], [139], [165], [185], [187]
Matching	6	3.57%	[14], [103], [105], [165], [188], [189]
Location-aware advertising	5	2.98%	[131], [185]–[187], [195]
Location based relationship	4	2.38%	[28], [130], [165], [173]
Location estimation	2	0.60%	[164]

third most frequent problem type, representing 10.12% of the surveyed studies.

As for context-aware recommendation systems, we found 16 papers, and most of them focused on the technical aspects of the problem type. However, several studies focused on specific business applications for recommending, for example, shop-types [17], [129], web or mobile services [162], [173], news feeds [195], social events and gatherings [188], carpooling [103], and music [194].

Within behavioral pattern detection, we found tourism applications focused on studying customer patterns and the hotel market [15], [16], [133]. In transportation, researchers studied the dynamics of networks, carriers, and passengers [98], [101], [106].

Under predictive analysis we classified studies that exploit location context to improve particular tasks such as business or service performance prediction [24], [62], and to detect anomalies in services [58], [178], among other applications.

Trajectory tracking applications include some level of observation of the routes and movements of the entities of the system, either to be used as an attribute of the study or the actual study subject. Trajectory tracking was common in transportation [101], [103]–[105], tourism [139], and advertising [185], [187].

Under the matching problem type, we found use cases in transportation, more specifically, applications to optimize the match between passengers and transporters [14], [105]. Also, in telecommunications, we found approaches exploiting location context to optimize the association between mobile users and cellular networks [165]. Similarly, we found approaches that exploit location for matching participants in social events [188], [189].

In location-aware advertising, aside from the two cases that also involve trajectory tracking, and location-aware news feeds, we found applications exploiting Twitter data to enable targeted advertising in cities [186], and coupon recommendations, which also exploit social context [131].

With respect to the location-based relationship problem type, we found applications in the construction of communities [28], long lived associations between users consuming services [165], [173], and discovering of retail areas according to spatial relationships [130].

Finally, we found only one instance of location estimation: Ahmed et al. [164] studied the estimation of cellular network degradation based on service performance.

B. TECHNIQUE FINDINGS

We found a rich variety of techniques among the surveyed papers, because each study elaborated thoroughly in implementation details. As a result, we only focused on characterizing the most common ones.
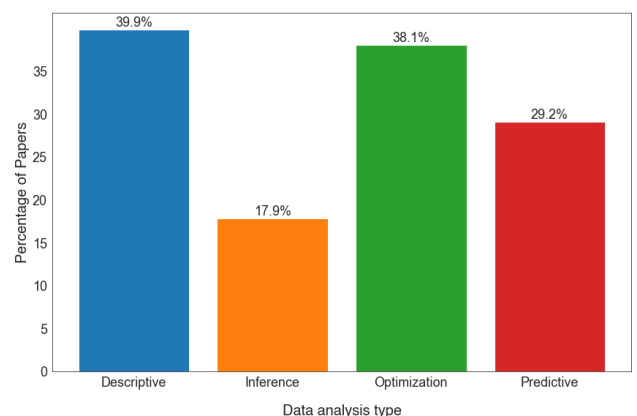
FIGURE 5. Publications per analysis type.

Fig. 5 presents the data analysis types used in the surveyed papers. We found that most of the studies used descriptive

analysis or unsupervised learning representing 39.88% of the approaches. The second most used approach was optimization, representing 38.10% of the studies. The third one was predictive analysis with 19.71%. Finally, inference analysis with 17.86%.

The two most frequent techniques were *mathematical optimization*, and *spatial analysis*, representing 36.9% and 34.52% respectively. Mathematical optimization is commonly used in location analysis problems such as facility location and spatial coverage problems. Spatial analysis is used to analyze and describe scenarios with geographical data.

Recommendation systems was the third most frequent technique representing 14.88% of the papers. Similar to location analysis, recommender systems is actually a complete research field, and not necessarily a single technique. However for simplicity, we grouped all the studies that cited the use of recommendation system techniques, such as collaborative-filtering, content-based, or hybrid techniques.

Matrix factorization is also a common tool to solve recommendation system problems. Nevertheless, we considered matrix factorization separately because it can be applied in other contexts such as predictive analysis [62], [189] or pattern detection [63]. Also, there are different methods to factorize a matrix, hence we group all of them together under this category.

Clustering techniques are a conventional tool to describe or discover patterns in geotagged entities or events. The fact that geographic information is commonly encoded in cartesian coordinates makes it a suitable candidate for clustering analysis. For example, Arbia *et al.* [116] implemented spatio-temporal clustering techniques to explore geographical aspects of medical device manufacturing industries. And Lloyd and Cheshire [125] employed clustering and KDE techniques to discover geographical center locations for retailing.

Artificial neural networks (ANN) have gained considerable attention in the research community since in several cases have demonstrated superior performance than traditional machine learning methods. According to our findings, there are few studies using ANN compared with the rest of the techniques, (only 4.76%). Nevertheless, all the studies that implemented an ANN-based solution focused on different industry and problem types, demonstrating valid solution approaches to their respective settings. In particular, ANN was used for location analysis problems in the hotel [138] and solar farm industries [159], as well as for context-aware recommendation systems [57], location-wise anomaly detection [58], social event recommendations [188], and location-aware advertising [187].

Under *Other* techniques we grouped the ones that appeared three or less times in the surveyed papers. Among these cases we found Latent Dirichlet Allocation [74], [75], [194], Natural Language Processing [15], [105], and Bayesian models [72], [74], [197].

Table 8 presents the characterization of the studies according to the techniques identified in our SLR.

C. METRIC AND VALIDATION FINDINGS

Depending on the problem type and techniques, there are several suitable metrics and validation protocols. In this section we summarize some aspects that we considered relevant with respect to the validation performed for the surveyed approaches.

1) METRIC FINDINGS

We were able to identify metrics in 69 of the surveyed papers. Among the different metrics reported, we categorized and documented the most frequent ones, i.e., metrics that appeared seven or more times among the surveyed studies. The findings are summarized in Table 9. The majority of the reported metrics are associated with predictive analysis approaches, whereas descriptive analysis and inference related studies use statistical measurements such as standard error, p-value, or chi-square tests. Mathematical optimization approaches seldom report metrics usage since these methods work differently than statistical or machine learning-based methods, e.g., maximizing an objective function subject to some constraints yields a single value.

In predictive analysis or supervised learning, the usual approach consists in, as the name implies, predicting a target value using other features that are associated with it. Depending on the model objective, there are some metrics that are more suitable than others. We identified the following metrics:

- **Accuracy:** It is the fraction of predictions correctly classified by the model among all the number of samples.
- **Precision:** Also known as true positive rate, it measures the proportion of positive classes against the total positives predicted by the model.
- **Recall:** Also known as sensitivity, it measures the proportion of positive classes against the total samples that should have been predicted as positive by the model.
- **F1-score:** Also known as the harmonic mean between precision and recall, it measures the model accuracy by considering both precision and recall.
- **RMSE:** The root mean squared error is a loss function consisting in calculating the squared root of the averaged squared loss per sample.
- **MAE:** The mean absolute error is a loss function calculated by averaging the absolute errors between the observed value and the predicted by the model.
- **Coverage:** It represents the amount of area covered given a certain objective and parameters, e.g., how much area a delivery service can cover using a certain configuration. In certain sub-categories of location analysis problems, it is defined as an objective to maximize.
- **Standard error:** It is a statistical measure of the difference between the mean of a sample and the total population.

TABLE 8. Publications per technique.

Techniques	Total Papers	%	References
Mathematical optimization	62	36.90%	[12], [13], [45], [47]–[50], [52]–[55], [60], [65], [76], [78], [80]–[83], [88], [90], [96], [97], [101], [102], [107], [109], [111]–[115], [117], [118], [126], [141], [143]–[148], [150], [152], [155], [163], [166]–[168], [174]–[177], [179]–[184], [195], [196], [198]
Spatial statistics	58	34.52%	[11], [12], [51], [56], [64], [81], [84]–[87], [89], [91]–[95], [98]–[101], [104], [106], [108], [110], [116], [119]–[125], [128], [130], [132]–[137], [149], [153], [154], [156], [158], [160], [161], [166], [170], [171], [178], [185], [186], [191]–[193], [198], [199]
Recommendation systems	25	14.88%	[23], [28], [57], [59], [61], [62], [66], [67], [69], [70], [73]–[75], [77], [79], [103], [129], [131], [139]–[142], [162], [173], [194]
Matrix factorization	14	8.33%	[17], [25], [46], [57], [59], [62], [63], [68], [72], [74], [77], [129], [151], [189]
Clustering	13	7.74%	[15], [56], [64], [66], [70], [71], [106], [116], [124], [125], [127], [140], [153]
Linear regression	10	5.95%	[14], [121], [133]–[136], [138], [158], [164], [192]
Fuzzy logic	9	5.36%	[44], [65], [112], [156], [157], [159], [162], [169], [188]
Artificial Neural Networks	8	4.76%	[57], [58], [79], [105], [138], [159], [187], [188]
Kernel Density Estimation	6	3.57%	[14], [15], [81], [110], [125], [133]
Markov models	4	2.38%	[23], [67], [72], [165]
Other	49	29.17%	[14]–[18], [18], [24], [44], [57], [58], [58], [59], [70], [72], [74], [74], [75], [92], [95], [98], [104]–[107], [120], [121], [121], [124], [126], [126], [128], [129], [132], [137], [138], [138], [138], [140], [141], [164], [171], [172], [180], [182], [186], [190], [191], [194], [197]

TABLE 9. Publications per metrics.

Metrics	Total Papers	%	References
Precision	18	18.00%	[23], [24], [28], [46], [58], [66]–[69], [71], [74], [129], [139], [141], [186], [188], [189], [194]
Recall	18	18.00%	[14], [23], [24], [28], [46], [58], [66]–[69], [71], [72], [74], [129], [139], [141], [186], [188]
RMSE	11	11.00%	[17], [25], [57]–[59], [62], [141], [158], [164], [171], [172]
F1-score	10	10.00%	[24], [28], [58], [66], [71], [77], [129], [139], [141], [188]
MAE	9	9.00%	[17], [25], [58], [62], [73], [79], [138], [172], [173]
Accuracy	8	8.00%	[24], [66], [71], [73], [75], [105], [131], [188]
Coverage	8	8.00%	[13], [68], [108], [118], [186], [187], [190], [198]
Standard error	7	7.00%	[51], [131], [132], [134], [137], [148], [191]

TABLE 10. Publications per validation strategies.

Validation strategies	Total Papers	%	References
AdHoc	86	57.33%	[13], [15], [44], [45], [47]–[50], [52]–[55], [60], [64], [65], [76], [78], [80]–[86], [88], [90], [91], [93]–[97], [99], [101]–[103], [106]–[109], [111]–[115], [117], [118], [127], [132], [136], [137], [142]–[151], [154]–[156], [158], [165], [165]–[169], [171], [174]–[184], [191], [196], [198]
Adhoc-desc	32	21.33%	[18], [51], [56], [64], [81], [85]–[87], [89], [91], [93]–[95], [98], [100], [101], [106], [107], [116], [119], [120], [130], [149], [153], [156], [160], [161], [166], [170], [178], [185], [193]
Holdout	30	20.00%	[17], [23]–[25], [28], [46], [58], [59], [61], [63], [66]–[69], [72]–[75], [79], [105], [129], [162], [164], [172], [173], [186], [187], [189], [194], [195]
Cross-validation	14	9.33%	[17], [23], [57], [62], [70], [77], [125], [133], [138], [139], [141], [186], [188], [192]
Other	6	4.00%	[25], [46], [58], [61], [63], [66], [70], [71], [75], [75], [79], [84], [103], [129], [131], [131], [131], [131], [131], [131], [134], [134], [139], [142], [142], [142], [142], [144], [151], [152], [172], [173], [188], [189], [189], [191], [194], [199]

2) VALIDATION STRATEGY FINDINGS

Validation strategies define the procedures or protocols followed to evaluate a data analysis solution or experiment. Findings are presented in Table 10. In supervised learning approaches, it is common to partition the original data set into training and testing subsets. The former is used for model training, whereas the later is used to verify how the model behaves with unobserved data. However, the validation strategy to follow often depends on the data analysis type and the techniques used. For example, in mathematical optimization, the optimal solution depends on the set of variables and values used, and partitioning the data could

TABLE 11. Publications per usage of location information.

Usage of Location info	Total Papers	%	References
Model output	94	55.95%	[12]–[14], [23], [44]–[50], [52]–[55], [57], [60], [61], [66]–[70], [72]–[74], [76], [78]–[81], [85], [86], [88], [90], [91], [96], [97], [101], [102], [104], [105], [107]–[109], [111]–[115], [117], [118], [121], [125], [126], [138]–[152], [155]–[157], [159], [164], [167]–[169], [175]–[184], [186], [192], [195], [196], [198], [199]
Model parameter	76	45.24%	[11], [15]–[18], [24], [25], [28], [51], [56]–[59], [62]–[65], [71], [75], [77], [82]–[84], [87], [89], [92]–[95], [98]–[100], [103], [106], [110], [116], [119], [120], [122]–[124], [127]–[137], [153], [154], [158], [160]–[162], [165], [165], [166], [170]–[174], [185], [187]–[191], [193], [194], [197]

lead to different results that are indeed valid for each partition, but not necessarily valid in different settings. For unsupervised learning approaches, experiments and ad-hoc testing can be more suitable, i.e., validating a model definition using real data or testing the definition in different settings.

We were able to identify a validation strategy in 150 of the surveyed papers. The validation strategies identified in the studied papers are characterized as follows:

- **Adhoc:** This strategy consists in experimenting with the designed and implemented solution by testing it against a real data set or setting, to evaluate how it behaves and whether it delivers the expected results. For example, optimization problems are often evaluated against a real setting.
- **Adhoc-descriptive:** This strategy follows the same concept as Adhoc, however the application is in descriptive settings like spatial analysis. We separated them because the analysis types and techniques are different.
- **Holdout:** This strategy consists in splitting the data set in two sets: training and testing, usually 70% and 30% respectively. However, the proportions may vary depending on the setting.
- **K-fold cross-validation:** This is a more sophisticated strategy that consists in partitioning the data set in *k* equally sized groups called folds, then iterating over each fold to perform training and testing phases, using a different test subset in every iteration.
- **Other:** We grouped under this category validation strategies with usage frequency of two or less. We included here: Student's t-test, Hypothesis testing, Binomial test, and Pearson's chi-squared tests.

D. LOCATION INFORMATION USAGE FINDINGS
This section summarizes the findings related to the usage of location information. We categorized the studies into two classes: *i) Model output*, which implies that the expected result of the model is a location or place; and *ii) Model parameter*, which means that the location information was used as an additional parameter to analyze or improve the performance of a model with other purposes.

Table 11 presents the findings about the usage of location information. We found that 55.36% of the approaches used location context data as the expected result, whereas the other 45.24% used it as a model parameter. Some studies used

location information as both model parameter and output. For example, Xu and Chow [195] exploited location information to recommend news feeds to users based on their current locations, and included a mechanism to predict the next location.

With respect to location information as model output, location analysis and POI recommendation are clear examples of problem types that focus on delivering a location as a result. Another problem type that used location information as result or subject of study was spatial analysis.

Among the studies that used location information as model parameter, we found problem types such as predictive analysis. For instance, Bojesen *et al.* [92] studied approaches to forecast the biogas production in Denmark; and Zhao *et al.* [62] exploited location context in LBSN to predict user ratings on certain items.

VIII. LIMITATIONS & RESEARCH OPPORTUNITIES
This section summarizes the limitations we faced during the elaboration of this SLR, as well as relevant research opportunities we identified.

A. LIMITATIONS
We encountered some limitations while elaborating this review, particularly with the depth and scope of the information we were able to cover. Most of the characterization was the product of manual inspection and interpretation of the study content. For some studies, it was difficult to derive certain characterization dimensions due to the lack of clearness in the scenarios the researchers described. Furthermore, given the significant variety of approaches and settings, elaborating on more in-depth details would have made this survey excessively extensive. As for depth, the most significant limitations were as follows:

- Given that it is not common to explicitly state business value chain aspects in research papers, the business component dimensions were mostly inferred according to our understanding of the individual cases. Hence, some elements might be limited.
- Given the plethora of solutions presented by the researchers, we had to limit the amount of information to be included in the final report, otherwise the survey would have been too extensive. This also limits the amount of detail certain researchers and practitioners interested in particular solutions can gather from our survey. However, the characterization tables we obtained

from our SLR are comprehensive enough to guide readers interested in getting further details from the surveyed papers.

With respect to the scope of this SLR:

- We did not include indoor-location studies mainly because the settings differ considerably from outdoor approaches. We recognize the contributions made in this area, and similar surveys could be conducted specifically for indoor locations.
- We acknowledge the importance of privacy, particularly with respect to location context in settings where information about individuals is collected. However, we did not consider privacy implications since we were focused on understanding location analytics.

B. RESEARCH OPPORTUNITIES

We observed an increasing interest in this topic in the research community. This might be the product of the surge and availability of geotagged information from different sources that enable researchers to explore scenarios that were not possible or too difficult before.

Borrowing from some of the surveyed papers' future work and research opportunities, in the context of exploiting LBSN for context-aware recommendation systems, Korakakis *et al.* [70] mention the inclusion of semantic analysis in the geo-clustering process, e.g., natural language processing techniques. Similarly, in the context of targeted advertising, Anagnostopoulos *et al.* [186] suggest exploring the trajectories of users when designing physical ad campaigns.

In the context of location analysis problems, particularly those that are based on mathematical optimization approaches, since the cases and scenarios are specific, some researchers suggest that their solutions can be tried in other similar industries, and/or can include additional features such as time, and other location-based properties [143], [152], [175], [198].

Other studies reported satisfactory results with their presented solutions and pointed in the direction of validating those approaches with other real-world data sets [24], [172], and exploring deeply other attributes associated with geotagged events [23], [62]. Furthermore, other authors plan to optimize the training of their models, and/or explore other advanced techniques such as deep learning [67], [140]. Others show interest in combining additional information associated with the events that are being studied besides the location data [17], [74].

Additionally, aside from the future directions reported by the authors of the surveyed papers, we observe opportunities in the location-aware advertising field. For example, exploring user generated content from LBSN to perform geographic market segmentation. Considering that a recurrent problem in marketing is the low effectiveness of advertising campaigns compared to the invested budget, solutions tailored to optimize the effort and resources needed to run marketing campaigns can be of utmost interest for that industry. Similarly, researchers can pursue the elaboration of other surveys that focus on particular aspects such as a single industry or a single problem type. The focus on a single element would allow researchers to expand more on internal details that can improve the visibility of the state of the art.

IX. CONCLUSION

This paper presented a comprehensive characterization of the application of location data analytics in the business value chain, based on the findings of a SLR that we conducted on papers published between 2014 and 2019. This SLR characterizes recent state of the art for several industries and problem types, in which location context is exploited, along with the techniques and data sources that are commonly employed. Findings were organized according to business concepts that are easy to understand, which should help practitioners identify paths to solve existing real life challenges, or researchers discover new paths that lead to investigate unexplored areas.

This study was conducted with the goal of mapping the different settings in which location context and data analytics can be effectively used, and exploited in favor of concrete business value chain processes, additionally, to help researchers and practitioners identify recent approaches and research opportunities. The main results provide an understandable and systematic characterization of the industries in which location data analytics is being exploited in specific business segments, as well as general understanding of the data sources, techniques, and experiment protocols used to implement the solutions. This way, this SLR allowed us to systematically answer the research questions that motivated it.

In spite of the scope of this study, and given the significant diversity of approaches and applications, it is unfeasible to derive definitive conclusions about the approaches, effectiveness, and peculiarities of every aspect of the settings we have found. One aspect that hampers a comprehensive generalization is the differential nature among the approaches that makes them difficult to compare. Nevertheless, it is possible to conclude that measures and validation strategies can improve such that a comprehensive benchmark of the studies can be identified, particularly with location and spatial analysis.

REFERENCES

[1] H. Huang, G. Gartner, J. M. Krisp, M. Raubal, and N. Van de Weghe, "Location based services: Ongoing evolution and research agenda," *J. Location Based Services*, vol. 12, no. 2, pp. 63–93, Apr. 2018.

[2] E. Alpaydin, *Machine Learning New AI*, vol. 91. Cambridge, MA, USA: MIT Press, 2017.

[3] H. A. Eiselt and V. Marianov, *Applications of Location Analysis*. Cham, Switzerland: Springer, 2015.

[4] H. A. Eiselt, V. Marianov, and J. Bhadury, *Location Analysis in Practice*, vol. 232. Cham, Switzerland: Springer, 2015.

[5] P. B. Keenan and P. Jankowski, "Spatial decision support systems: Three decades on," *Decis. Support Syst.*, vol. 116, pp. 64–76, Jan. 2019.

[6] B. Kitchenham, B. Kitchenham, and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Durham Univ., Durham, U.K., Joint Rep., 2007.

[7] M. E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*. New York, NY, USA: Free Press (Simon & Schuster), 1985.

[8] P. Harmon, "Strategy, value chains, business initiatives, and competitive advantage," in *Business Process Change*. Amsterdam, The Netherlands: Elsevier, Jan. 2019, pp. 27–50.

[9] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, *Towards a Better Understanding of Context and Context-Awareness*, (Lecture Notes in Computer Science), vol. 1707. Berlin, Germany: Springer, 1999, pp. 304–307.

[10] N. M. Villegas and H. A. Müller, *Managing Dynamic Context to Optimize Smart Interactions and Services* (Lecture Notes in Computer Science), vol. 6400. Berlin, Germany: Springer, 2010, pp. 289–318.

[11] N. Martinkus, S. A. M. Rijkhoff, S. A. Hoard, W. Shi, P. Smith, M. Gaffney, and M. Wolcott, "Biorefinery site selection using a stepwise biogeophysical and social analysis approach," *Biomass Bioenergy*, vol. 97, pp. 139–148, Feb. 2017.

[12] M. Bojesen, M. Birkin, and G. Clarke, "Spatial competition for biogas production using insights from retail location models," *Energy*, vol. 68, pp. 617–628, Apr. 2014.

[13] L. Conrow, A. T. Murray, and H. A. Fischer, "An optimization approach for equitable bicycle share station siting," *J. Transp. Geography*, vol. 69, pp. 163–170, May 2018.

[14] C. Willing, K. Klemmer, T. Brandt, and D. Neumann, "Moving in time and space—Location intelligence for carsharing decision support," *Decis. Support Syst.*, vol. 99, pp. 75–85, Jul. 2017.

[15] S. J. Miah, H. Q. Vu, and J. G. Gammack, "A location analytics method for the utilisation of geotagged photos in travel marketing decision-making," *J. Inf. Knowl. Manage.*, vol. 18, no. 1, Mar. 2019, Art. no. 1950004.

[16] L. Masiero, Y. Yang, and R. T. R. Qiu, "Understanding hotel location preference of customers: Comparing random utility and random regret decision rules," *Tourism Manage.*, vol. 73, pp. 83–93, Aug. 2019.

[17] Z. Yu, M. Tian, Z. Wang, B. Guo, and T. Mei, "Shop-type recommendation leveraging the data from social media and location-based services," *ACM Trans. Knowl. Discovery Data*, vol. 11, no. 1, pp. 1–21, Aug. 2016.

[18] L. Kowalski, "Comparing spatial-interaction and hybrid agent-based modelling approaches: An application to location analysis of services," *J. Artif. Soc. Social Simul.*, vol. 22, no. 1, 2019. [Online]. Available: http://jasss.soc.surrey.ac.uk/index_by_issue.html#v22

[19] Z. Dashdorj and S. Sobolevsky, "Characterization of behavioral patterns exploiting description of geographical areas," in *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXVII* (Lecture Notes in Computer Science), vol. 9860. Berlin, Germany: Springer, 2016, pp. 159–176.

[20] H. A. Eiselt and V. Marianov, *Foundations of Location Analysis*. New York, NY, USA: Springer, 2011.

[21] A. Mitchell, *The ESRI Guide to GIS Analysis: Spatial Measurements and Statistics*, vol. 2. Redlands, CA, USA: ESRI, 2015.

[22] N. M. Villegas, C. Sánchez, J. Díaz-Cely, and G. Tamura, "Characterizing context-aware recommender systems: A systematic literature review," *Knowl.-Based Syst.*, vol. 140, pp. 173–200, Jan. 2018.

[23] M. Xin, Y. Zhang, S. Li, L. Zhou, and W. Li, "A location-context awareness mobile services collaborative recommendation algorithm based on user behavior prediction," *Int. J. Web Services Res.*, vol. 14, no. 2, pp. 45–66, Apr. 2017.

[24] X. Chang and J. Li, "Business performance prediction in location-based social commerce," *Expert Syst. Appl.*, vol. 126, pp. 112–123, Jul. 2019.

[25] S. Wang, Y. Ma, B. Cheng, F. Yang, and R. N. Chang, "Multi-dimensional QoS prediction for service recommendations," *IEEE Trans. Services Comput.*, vol. 12, no. 1, pp. 47–57, Jan. 2019.

[26] R. Imai, K. Tsubouchi, T. Konishi, and M. Shimosaka, "Early destination prediction with spatio-temporal user behavior patterns," in *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, 2017, vol. 1, no. 4, Art. no. 142.

[27] H. Pham, C. Shahabi, and Y. Liu, "Inferring social strength from spatiotemporal data," *ACM Trans. Database Syst.*, vol. 41, no. 1, pp. 1–47, Apr. 2016.

[28] N. D. A. B. Navarro, C. A. da Costa, J. L. V. Barbosa, and R. D. R. Righi, "Spontaneous social network: Toward dynamic virtual communities based on context-aware computing," *Expert Syst. Appl.*, vol. 95, pp. 72–87, Apr. 2018.

[29] K. Liu and X. Li, "Hiding media data via shaders: Enabling private sharing in the clouds," in *Mobile SmartLife via Sensing, Localization, and Cloud Ecosystems*. Boca Raton, FL, USA: CRC Press, Dec. 2017, pp. 223–242.

[30] D. Jurgens, T. Finnethy, J. McCorriston, Y. T. Xu, and D. Ruths, "Geolocation prediction in twitter using social networks: A critical analysis and review of current practice," in *Handbook of Psychology*. Palo Alto, CA, USA: AAAI Press, 2003, pp. 1–10.

[31] G. Skoumas, D. Pfoser, A. Kyrillidis, and T. Sellis, "Location estimation using crowdsourced spatial relations," *ACM Trans. Spatial Algorithms Syst.*, vol. 2, no. 2, pp. 1–23, Jul. 2016.

[32] L. P. Fávero and P. Belfiore, "Univariate descriptive statistics," in *Data Science for Business and Decision Making*. Amsterdam, The Netherlands: Elsevier, Jan. 2019, pp. 21–91.

[33] A. Géron, *Hands-On Machine Learning With Scikit-Learn and Tensor-Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, cA, USA: O'Reilly Media, Oct. 2017.

[34] C. Nyce, "Predictive analytics white paper," in *American Institute for Chartered Property Casualty Underwriters*. Malvern, PA, USA: AICPCU, 2007, p. 16.

[35] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to Statistical Learning*, vol. 7, no. 10. New York, NY, USA: Springer, 2000.

[36] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer, 2006.

[37] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, "Tour recommendation and trip planning using location-based social media: A survey," *Knowl. Inf. Syst.*, vol. 60, no. 3, pp. 1247–1275, Sep. 2019.

[38] A. T. Murray, "Contemporary optimization application through geographic information systems," *Omega*, vol. 99, Dec. 2019, Art. no. 102176.

[39] B. B. Nielsen, C. G. Asmussen, and C. D. Weatherall, "The location choice of foreign direct investments: Empirical evidence and methodological challenges," *J. World Bus.*, vol. 52, no. 1, pp. 62–82, Jan. 2017.

[40] K. Stock, "Mining location from social media: A systematic review," *Comput., Environ. Urban Syst.*, vol. 71, pp. 209–240, Sep. 2018.

[41] Y. Yang, H. Luo, and R. Law, "Theoretical, empirical, and operational models in hotel location research," *Int. J. Hospitality Manage.*, vol. 36, pp. 209–220, Jan. 2014.

[42] M. Pelka and H. Hellbrück, "Survey of challenges and towards a unified architecture for location systems," *J. Netw. Comput. Appl.*, vol. 67, pp. 75–85, May 2016.

[43] E. Toch, B. Lerner, E. Ben-Zion, and I. Ben-Gal, "Analyzing large-scale human mobility data: A survey of machine learning methods and applications," *Knowl. Inf. Syst.*, vol. 58, no. 3, pp. 501–523, Mar. 2019.

[44] I. Önden, "Integrating GIS with F-AHP for locating a single facility," *Transport*, vol. 33, no. 5, pp. 1173–1183, Dec. 2018.

[45] A. Z. Acar, I. Onden, and K. Kara, "Evaluating the location of regional return centers in reverse logistics through integration of GIS, AHP in integer programming," *Int. J. Ind. Eng., Theory Appl. Pract.*, vol. 22, no. 4, pp. 399–411, Aug. 2015.

[46] B. Zou, B. Zou, C. Li, C. Li, L. Tan, L. Tan, H. Chen, and H. Chen, "Location-based recommendation using incremental tensor factorization model," in *Advanced Data Mining and Applications* (Lecture Notes in Computer Science), vol. 8933. Cham, Switzerland: Springer, Dec. 2014, pp. 227–238.

[47] M. Cebecauer and L. Buzna, "A versatile adaptive aggregation framework for spatially large discrete location-allocation problems," *Comput. Ind. Eng.*, vol. 111, pp. 364–380, Sep. 2017.

[48] P. Venkateshan, R. H. Ballou, K. Mathur, and A. P. P. Maruthasalam, "A two-echelon joint continuous-discrete location model," *Eur. J. Oper. Res.*, vol. 262, no. 3, pp. 1028–1039, Nov. 2017.

[49] O. Berman, I. Hajizadeh, D. Krass, and A. Rahimi-Vahed, "Reconfiguring a set of coverage-providing facilities under travel time uncertainty," *Socio-Econ. Planning Sci.*, vol. 62, pp. 1–12, Jun. 2018.

[50] C. S. Guazzelli and C. B. Cunha, "Exploring K-best solutions to enrich network design decision-making," *Omega (United Kingdom)*, vol. 78, pp. 139–164, Jul. 2018.

[51] L. Teng, D. Huang, and Y. Pan, "The performance of MNE subsidiaries in China: Does it matter to be close to the political or business hub?" *J. Int. Manage.*, vol. 23, no. 3, pp. 292–305, Sep. 2017.

[52] H. Fan, J. Ma, and X. Li, "A reliable location model for heterogeneous systems under partial capacity losses," *Transp. Res. C, Emerg. Technol.*, vol. 97, pp. 235–257, Dec. 2018.

[53] X. Zheng, M. Yin, and Y. Zhang, "Integrated optimization of location, inventory and routing in supply chain network design," *Transp. Res. B, Methodol.*, vol. 121, pp. 1–20, Mar. 2019.

[54] I. Hernandez, J. Emmanuel Ramirez-Marquez, C. Rainwater, E. Pohl, and H. Medal, "Robust facility location: Hedging against failures," *Rel. Eng. Syst. Saf.*, vol. 123, pp. 73–80, Mar. 2014.

[55] M. B. C. Menezes, D. Ruiz-Hernández, and V. Verter, "A rough-cut approach for evaluating location-routing decisions via approximation algorithms," *Transp. Res. B, Methodol.*, vol. 87, pp. 89–106, May 2016.

[56] J. Beckers, M. Vanhoof, and A. Verhetsel, "Returning the particular: Understanding hierarchies in the Belgian logistics system," *J. Transp. Geography*, vol. 76, pp. 315–324, Apr. 2019.

[57] M. Unger, A. Bar, B. Shapira, and L. Rokach, "Towards latent context-aware recommendation systems," *Knowl.-Based Syst.*, vol. 104, pp. 165–178, Jul. 2016.

[58] R. Corizzo, M. Ceci, and N. Japkowicz, "Anomaly detection and repair for accurate predictions in geo-distributed big data," *Big Data Res.*, vol. 16, pp. 18–35, Jul. 2019.

[59] R. Pálovics, P. Szalai, J. Pap, E. Frigó, L. Kocsis, and A. A. Benczúr, "Location-aware online learning for top-K recommendation," *Pervas. Mobile Comput.*, vol. 38, pp. 490–504, Jul. 2017.

[60] M. Meraklı and H. Yaman, "Robust intermodal hub location under polyhedral demand uncertainty," *Transp. Res. B, Methodol.*, vol. 86, pp. 66–85, Apr. 2016.

[61] M. Sarwat, J. J. Levandoski, A. Eldawy, and M. F. Mokbel, "LARS*: An efficient and scalable location-aware recommender system," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 6, pp. 1384–1399, Jun. 2014.

[62] G. Zhao, X. Qian, and C. Kang, "Service rating prediction by exploring social mobile user's geographical locations," *IEEE Trans. Big Data*, vol. 3, no. 1, Mar. 2017, pp. 67–78.

[63] J. An, H. Kwak, S.-G. Jung, J. Salminen, and B. J. Jansen, "Customer segmentation using online platforms: Isolating behavioral and demographic segments for persona creation via aggregated user data," *Social Netw. Anal. Mining*, vol. 8, no. 1, p. 54, Dec. 2018.

[64] F. Chen, J. Qi, H. Lin, Y. Gao, and D. Lu, "GOAL: A clustering-based method for the group optimal location problem," *Knowl. Inf. Syst.*, vol. 61, no. 2, pp. 873–903, Nov. 2019.

[65] A. Alizadeh and S. Yousefi, "An integrated Taguchi loss function–fuzzy cognitive map–MCGP with utility function approach for supplier selection problem," *Neural Comput. Appl.*, vol. 31, no. 11, pp. 7595–7614, Jun. 2019.

[66] H. Bagci and P. Karagoz, "Context-aware location recommendation by using a random walk-based approach," *Knowl. Inf. Syst.*, vol. 47, no. 2, pp. 241–260, May 2016.

[67] W. X. Zhao, N. Zhou, A. Sun, J.-R. Wen, J. Han, and E. Y. Chang, "A time-aware trajectory embedding model for next-location recommendation," *Knowl. Inf. Syst.*, vol. 56, no. 3, pp. 559–579, Sep. 2018.

[68] B. Xia, T. Li, Q. Li, and H. Zhang, "Noise-tolerance matrix completion for location recommendation," *Data Mining Knowl. Discovery*, vol. 32, no. 1, pp. 1–24, Jan. 2018.

[69] D. Zhou, S. M. Rahimi, and X. Wang, "Similarity-based probabilistic category-based location recommendation utilizing temporal and geographical influence," *Int. J. Data Sci. Anal.*, vol. 1, no. 2, pp. 111–121, Jul. 2016.

[70] M. Korakakis, E. Spyrou, P. Mylonas, and S. J. Perantonis, "Exploiting social media information toward a context-aware recommendation system," *Social Netw. Anal. Mining*, vol. 7, no. 1, p. 42, Dec. 2017.

[71] Z. Yang and N. Japkowicz, "Anomaly behaviour detection based on the meta-morisita index for large scale spatio-temporal data set," *J. Big Data*, vol. 5, no. 1, p. 23, Dec. 2018.

[72] X. Li, D. Han, J. He, J. Liao, and M. Wang, "Next and next new POI recommendation via latent behavior pattern inference," *ACM Trans. Inf. Syst.*, vol. 37, no. 4, pp. 1–28, Dec. 2019.

[73] J. Sang, T. Mei, and C. Xu, "Activity sensor: Check-in usage mining for local recommendation," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 3, pp. 1–24, May 2015.

[74] Q. Fang, C. Xu, M. S. Hossain, and G. Muhammad, "STCAPLRS: A spatial-temporal context-aware personalized location recommendation system," *ACM Trans. Intell. Syst. Technol.*, vol. 7, no. 4, pp. 1–30, Jul. 2016.

[75] H. Yin, B. Cui, L. Chen, Z. Hu, and C. Zhang, "Modeling location-based user rating profiles for personalized recommendation," *ACM Trans. Knowl. Discovery Data*, vol. 9, no. 3, pp. 1–41, Apr. 2015.

[76] G. Erdoğan, N. Stylianou, and C. Vasilakis, "An open source decision support system for facility location analysis," *Decis. Support Syst.*, vol. 125, Oct. 2019, Art. no. 113116.

[77] E. Khazaei and A. Alimohammadi, "Context-aware group-oriented location recommendation in location-based social networks," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 9, p. 406, Sep. 2019.

[78] E. Fernández and A. Sgalambro, "On carriers collaboration in hub location problems," *Eur. J. Oper. Res.*, vol. 283, no. 2, pp. 476–490, Jun. 2020.

[79] D. Margaris, C. Vassilakis, and D. Spiliotopoulos, "Handling uncertainty in social media textual information for improving venue recommendation formulation quality in social networks," *Social Netw. Anal. Mining*, vol. 9, no. 1, p. 64, Dec. 2019.

[80] Y. Costa, A. Duarte, and W. Sarache, "A decisional simulation-optimization framework for sustainable facility location of a biodiesel plant in colombia," *J. Cleaner Prod.*, vol. 167, pp. 174–191, Nov. 2017.

[81] J. Höhn, E. Lehtonen, S. Rasi, and J. Rintala, "A geographical information system (GIS) based methodology for determination of potential biomasses and sites for biogas plants in southern finland," *Appl. Energy*, vol. 113, pp. 1–10, Jan. 2014.

[82] L. He-Lambert, B. C. English, D. M. Lambert, O. Shylo, J. A. Larson, T. E. Yu, and B. Wilson, "Determining a geographic high resolution supply chain network for a large scale biofuel industry," *Appl. Energy*, vol. 218, pp. 266–281, May 2018.

[83] C. Miret, P. Chazara, L. Montastruc, S. Negny, and S. Domenech, "Design of bioethanol green supply chain: Comparison between first and second generation biomass concerning economic, environmental and social criteria," *Comput. Chem. Eng.*, vol. 85, pp. 16–35, Feb. 2016.

[84] E. Nikas, A. Sotiropoulos, and G. A. Xydis, "Spatial planning of biogas processing facilities in greece: The sunflower's capabilities and the waste-to-bioproducts approach," *Chem. Eng. Res. Des.*, vol. 131, pp. 234–244, Mar. 2018.

[85] F. Wang and X. Shi, "Geospatial analysis for utilizing the marginal land in regional biofuel industry: A case study in Guangdong Province, China," *Biomass Bioenergy*, vol. 83, pp. 302–310, Dec. 2015.

[86] R. Singh, F. Kemausuor, and M. Wooldridge, "Locational analysis of cellulosic ethanol production and distribution infrastructure for the transportation sector in ghana," *Renew. Sustain. Energy Rev.*, vol. 98, pp. 393–406, Dec. 2018.

[87] T. Morato, M. Vaezi, and A. Kumar, "Developing a framework to optimally locate biomass collection points to improve the biomass-based energy facilities locating procedure—A case study for Bolivia," *Renew. Sustain. Energy Rev.*, vol. 107, pp. 183–199, Jun. 2019.

[88] F. Zhang, D. Johnson, M. Johnson, D. Watkins, R. Froese, and J. Wang, "Decision support system integrating GIS with simulation and optimisation for a biofuel supply chain," *Renew. Energy*, vol. 85, pp. 740–748, Jan. 2016.

[89] B. Sharma, S. Birrell, and F. E. Miguez, "Spatial modeling framework for bioethanol plant siting and biofuel production potential in the U.S.," *Appl. Energy*, vol. 191, pp. 75–86, Apr. 2017.

[90] F. Zhang, J. Wang, S. Liu, S. Zhang, and J. W. Sutherland, "Integrating GIS with optimization method for a biofuel feedstock supply chain," *Biomass Bioenergy*, vol. 98, pp. 194–205, Mar. 2017.

[91] S. Silva, L. Alçada-Almeida, and L. C. Dias, "Biogas plants site selection integrating multicriteria decision aid methods and GIS techniques: A case study in a Portuguese region," *Biomass Bioenergy*, vol. 71, pp. 58–68, Dec. 2014.

[92] M. Bojesen, H. Skov-Petersen, and M. Gylling, "Forecasting the potential of Danish biogas production—Spatial representation of Markov chains," *Biomass Bioenergy*, vol. 81, pp. 462–472, Oct. 2015.

[93] M. Bojesen, L. Boerboom, and H. Skov-Petersen, "Towards a sustainable capacity expansion of the Danish biogas sector," *Land Use Policy*, vol. 42, pp. 264–277, Jan. 2015.

[94] P. Patrizio and D. Chinese, "The impact of regional factors and new bio-methane incentive schemes on the structure, profitability and CO_2 balance of biogas plants in Italy," *Renew. Energy*, vol. 99, pp. 573–583, Dec. 2016.

[95] T. Wuerzer and S. G. Mason, "Retail gravitation and economic impact: A market-driven analytical framework for bike-share station location analysis in the united states," *Int. J. Sustain. Transp.*, vol. 10, no. 3, pp. 247–259, Mar. 2016.

[96] B. Soylu and H. Katip, "A multiobjective hub-airport location problem for an airline network design," *Eur. J. Oper. Res.*, vol. 277, no. 2, pp. 412–425, Sep. 2019.

[97] K.-I. Tanaka, T. Furuta, and S. Toriumi, "Railway flow interception location model: Model development and case study of tokyo metropolitan railway network," *Oper. Res. Perspect.*, vol. 6, Jan. 2019, Art. no. 100085.

[98] M. S. Hasnine, A. Graovac, F. Camargo, and K. N. Habib, "A random utility maximization (RUM) based measure of accessibility to transit: Accurate capturing of the first-mile issue in urban transit," *J. Transp. Geography*, vol. 74, pp. 313–320, Jan. 2019.

[99] X. Chen and L. Lin, "The node-place analysis on the 'hubtropolis' urban form: The case of Shanghai Hongqiao air-rail hub," *Habitat Int.*, vol. 49, pp. 445–453, Oct. 2015.

[100] A. Moyano, H. S. Martínez, and J. M. Coronado, "From network to services: A comparative accessibility analysis of the Spanish high-speed rail system," *Transp. Policy*, vol. 63, pp. 51–60, Apr. 2018.

[101] R. Li and D. Tong, "Incorporating activity space and trip chaining into facility siting for accessibility maximization," *Socio-Econ. Planning Sci.*, vol. 60, pp. 1–14, Dec. 2017.

[102] B. F. Santos, S. Limbourg, and J. S. Carreira, "The impact of transport policies on railroad intermodal freight competitiveness—The case of Belgium," *Transp. Res. D, Transp. Environ.*, vol. 34, pp. 230–244, Jan. 2015.

[103] D. Zhang, T. He, Y. Liu, S. Lin, and J. A. Stankovic, "A carpooling recommendation system for taxicab services," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 254–266, Sep. 2014.

[104] S. Wang, Z. Bao, J. Shane Culpepper, T. Sellis, and G. Cong, "Reverse *k* nearest neighbor search over trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 4, pp. 757–771, Apr. 2018.

[105] Z. Huang, G. Shan, J. Cheng, and J. Sun, "TRec: An efficient recommendation system for hunting passengers with deep neural networks," *Neural Comput. Appl.*, vol. 31, no. S1, pp. 209–222, Jan. 2019.

[106] A.-S. Briand, E. Côme, M. K. El Mahrsi, and L. Oukhellou, "A mixture model clustering approach for temporal passenger pattern characterization in public transport," *Int. J. Data Sci. Analytics*, vol. 1, no. 1, pp. 37–50, Apr. 2016.

[107] P. V. Amaral and A. T. Murray, "Equity in regional access to renal dialysis in Brazil," *Regional Sci. Policy Pract.*, vol. 8, nos. 1–2, pp. 27–44, Mar. 2016.

[108] O. Kotavaara, T. Pohjosenperä, J. Juga, and J. Rusanen, "Accessibility in designing centralised warehousing: Case of health care logistics in Northern Finland," *Appl. Geography*, vol. 84, pp. 83–92, Jul. 2017.

[109] S. X. Zhu and E. Ursavas, "Design and analysis of a satellite network with direct delivery in the pharmaceutical industry," *Transp. Res. E, Logistics Transp. Rev.*, vol. 116, pp. 190–207, Aug. 2018.

[110] S. Moon, T. F. Vadeboncoeur, W. Kortuem, M. Kisakye, M. Karamooz, B. White, P. Brazil, D. W. Spaite, and B. J. Bobrow, "Analysis of out-of-hospital cardiac arrest location and public access defibrillator placement in metropolitan Phoenix, Arizona," *Resuscitation*, vol. 89, pp. 43–49, Apr. 2015.

[111] S. Khodaparasti, M. E. Bruni, P. Beraldi, H. R. Maleki, and S. Jahedi, "A multi-period location-allocation model for nursing home network planning under uncertainty," *Oper. Res. Health Care*, vol. 18, pp. 4–15, Sep. 2018.

[112] M. Mohammadi, S. Dehbari, and B. Vahdani, "Design of a bi-objective reliable healthcare network with finite capacity queue under service covering uncertainty," *Transp. Res. E, Logistics Transp. Rev.*, vol. 72, pp. 15–41, Dec. 2014.

[113] P. Chaiwuttisak, H. Smith, Y. Wu, C. Potts, T. Sakuldamrongpanich, and S. Pathomsiri, "Location of low-cost blood collection and distribution centres in thailand," *Oper. Res. Health Care*, vol. 9, pp. 7–15, Jun. 2016.

[114] J. N. Ares, H. de Vries, and D. Huisman, "A column generation approach for locating roadside clinics in Africa based on effectiveness and equity," *Eur. J. Oper. Res.*, vol. 254, no. 3, pp. 1002–1016, Nov. 2016.

[115] P. L. van den Berg, G. J. Kommer, and B. Zuzáková, "Linear formulation for the maximum expected coverage location model with fractional coverage," *Operations Res. Health Care*, vol. 8, pp. 33–41, Mar. 2016.

[116] G. Arbia, G. Espa, D. Giuliani, and M. M. Dickson, "Spatio-temporal clustering in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis," *Regional Sci. Urban Econ.*, vol. 49, pp. 298–304, Nov. 2014.

[117] R. Batta, M. Lejeune, and S. Prasad, "Public facility location using dispersion, population, and equity criteria," *Eur. J. Oper. Res.*, vol. 234, no. 3, pp. 819–829, May 2014.

[118] S. Taymaz, C. Iyigun, Z. P. Bayindir, and N. P. Dellaert, "A healthcare facility location problem for a multi-disease, multi-service environment under risk aversion," *Socio-Econ. Planning Sci.*, vol. 71, Sep. 2020, Art. no. 100755.

[119] I. Barbarisi, G. Bruno, A. Diglio, J. Elizalde, and C. Piccolo, "A spatial analysis to evaluate the impact of deregulation policies in the pharmacy sector: Evidence from the case of Navarre," *Health Policy*, vol. 123, no. 11, pp. 1108–1115, Nov. 2019.

[120] D. T. Robinson and A. M. Balulescu, "Comparison of methods for quantifying consumer spending on retail using publicly available data," *Int. J. Geographical Inf. Sci.*, vol. 32, no. 6, pp. 1061–1086, Jun. 2018.

[121] T. Wieland, "Market area analysis for retail and service locations with MCI," *R J.*, vol. 9, no. 1, p. 298, 2017.

[122] M. D. Rice, A. Ostrander, and C. Tiwari, "Decoding the development strategy of a major retailer: Wal-Mart's expansion in the United States," *Prof. Geographer*, vol. 68, no. 4, pp. 640–649, Oct. 2016.

[123] L. Chen and T. Notteboom, "A cost perspective on the location of value-added logistics services in supply chains," *Int. J. Logistics Syst. Manage.*, vol. 18, no. 1, pp. 24–48, 2014.

[124] F. Wang, C. Chen, C. Xiu, and P. Zhang, "Location analysis of retail stores in Changchun, China: A street centrality perspective," *Cities*, vol. 41, pp. 54–63, Dec. 2014.

[125] A. Lloyd and J. Cheshire, "Deriving retail centre locations and catchments from geo-tagged Twitter data," *Comput., Environ. Urban Syst.*, vol. 61, pp. 108–118, Jan. 2017.

[126] N. Yıldız and F. Tüysüz, "A hybrid multi-criteria decision making approach for strategic retail location investment: Application to Turkish food retailing," *Socio-Econ. Planning Sci.*, vol. 68, Dec. 2019, Art. no. 100619.

[127] M. D. Pratt, J. A. Wright, S. Cockings, and I. Sterland, "Delineating retail conurbations: A rules-based algorithmic approach," *J. Retailing Consum. Services*, vol. 21, no. 5, pp. 667–675, Sep. 2014.

[128] Y.-S. Shi, J. Wu, and S.-Y. Wang, "Spatio-temporal features and the dynamic mechanism of shopping center expansion in Shanghai," *Appl. Geography*, vol. 65, pp. 93–108, Dec. 2015.

[129] X. Mao, X. Zhao, J. Lin, and E. Herrera-Viedma, "Utilizing multi-source data in popularity prediction for shop-type recommendation," *Knowl.-Based Syst.*, vol. 165, pp. 253–267, Feb. 2019.

[130] S. Mishra, V. Leroy, and S. Amer-Yahia, "Colloquial region discovery for retail products: Discovery and application," *Int. J. Data Sci. Anal.*, vol. 4, no. 1, pp. 17–34, Aug. 2017.

[131] Y.-M. Li, C.-L. Chou, and L.-F. Lin, "A social recommender mechanism for location-based group commerce," *Inf. Sci.*, vol. 274, pp. 125–142, Aug. 2014.

[132] L. Fang, H. Li, and M. Li, "Does hotel location tell a true story? Evidence from geographically weighted regression analysis of hotels in Hong Kong," *Tourism Manage.*, vol. 72, pp. 78–91, Jun. 2019.

[133] D. East, P. Osborne, S. Kemp, and T. Woodfine, "Combining GPS and survey data improves understanding of visitor behaviour," *Tourism Manage.*, vol. 61, pp. 307–320, Aug. 2017.

[134] J. Gutiérrez, J. C. García-Palomares, G. Romanillos, and M. H. Salas-Olmedo, "The eruption of Airbnb in tourist cities: Comparing spatial patterns of hotels and peer-to-peer accommodation in barcelona," *Tourism Manage.*, vol. 62, pp. 278–291, Oct. 2017.

[135] J. Kim, S. Jang, S. Kang, and S. H. J. Kim, "Why are hotel room prices different? Exploring spatially varying relationships between room price and hotel attributes," *J. Bus. Res.*, vol. 107, pp. 118–129, Feb. 2020.

[136] S. K. Lee, "Quality differentiation and conditional spatial price competition among hotels," *Tourism Manage.*, vol. 46, pp. 114–122, Feb. 2015.

[137] H. Luo and Y. Yang, "Intra-metropolitan location choice of star-rated and non-rated budget hotels: The role of agglomeration economies," *Int. J. Hospitality Manage.*, vol. 59, pp. 72–83, Oct. 2016.

[138] Y. Yang, J. Tang, H. Luo, and R. Law, "Hotel location evaluation: A combination of machine learning tools and Web GIS," *Int. J. Hospitality Manage.*, vol. 47, pp. 14–24, May 2015.

[139] Y.-T. Wen, J. Yeo, W.-C. Peng, and S.-W. Hwang, "Efficient keyword-aware representative travel route recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 8, pp. 1639–1652, Aug. 2017.

[140] R. Logesh, V. Subramaniyaswamy, V. Vijayakumar, X. Z. Gao, and G. G. Wang, "Hybrid bio-inspired user clustering for the generation of diversified recommendations," *Neural Comput. Appl.*, vol. 32, no. 7, pp. 2487–2506, Mar. 2019.

[141] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," *Knowl. Inf. Syst.*, vol. 54, no. 2, pp. 375–406, Feb. 2018.

[142] A. V. Smirnov, A. M. Kashevnik, and A. Ponomarev, "Context-based infomobility system for cultural heritage recommendation: Tourist assistant–TAIS," *Pers. Ubiquitous Comput.*, vol. 21, no. 2, pp. 297–311, Apr. 2017.

[143] Y. Zhang, L. V. Snyder, M. Qi, and L. Miao, "A heterogeneous reliable location model with risk pooling under supply disruptions," *Transp. Res. B, Methodol.*, vol. 83, pp. 151–178, Jan. 2016.

[144] Z. Dong and M. A. Turnquist, "Combining service frequency and vehicle routing for managing supplier shipments," *Transp. Res. E, Logistics Transp. Rev.*, vol. 79, pp. 231–243, Jul. 2015.

[145] T. G. Crainic, P. Dell'Olmo, N. Ricciardi, and A. Sgalambro, "Modeling dry-port-based freight distribution planning," *Transp. Res. C, Emerg. Technol.*, vol. 55, pp. 518–534, Jun. 2015.

[146] Z. Kartal, S. Hasgul, and A. T. Ernst, "Single allocation P-hub median location and routing problem with simultaneous pick-up and delivery," *Transp. Res. E, Logistics Transp. Rev.*, vol. 108, pp. 141–159, Dec. 2017.

[147] K. Komchornrit, "The selection of dry port location by a hybrid CFA-MACBETH-PROMETHEE method: A case study of Southern Thailand," *Asian J. Shipping Logistics*, vol. 33, no. 3, pp. 141–153, Sep. 2017.

[148] M. Catalano and M. Migliore, "A stackelberg-game approach to support the design of logistic terminals," *J. Transp. Geography*, vol. 41, pp. 63–73, Dec. 2014.

[149] E. Özceylan, M. Erbaş, M. Tolon, M. Kabak, and T. Durğut, "Evaluation of freight villages: A GIS-based multi-criteria decision analysis," *Comput. Ind.*, vol. 76, pp. 38–52, Feb. 2016.

[150] N. Romero, L. K. Nozick, and N. Xu, "Hazmat facility location and routing analysis with explicit consideration of equity using the gini coefficient," *Transp. Res. E, Logistics Transp. Rev.*, vol. 89, pp. 165–181, May 2016.

[151] T. Van Nguyen, J. Zhang, L. Zhou, M. Meng, and Y. He, "A data-driven optimization of large-scale dry port location using the hybrid approach of data mining and complex network theory," *Transp. Res. E, Logistics Transp. Rev.*, vol. 134, Feb. 2020, Art. no. 101816.

[152] M. Basirati, M. R. Akbari Jokar, and E. Hassannayebi, "Bi-objective optimization approaches to many-to-many hub location routing with distance balancing and hard time window," *Neural Comput. Appl.*, vol. 32, no. 17, pp. 13267–13288, Sep. 2020.

[153] W. Arowolo, P. Blechinger, C. Cader, and Y. Perez, "Seeking workable solutions to the electrification challenge in Nigeria: Minigrid, reverse auctions and institutional adaptation," *Energy Strategy Rev.*, vol. 23, pp. 114–141, Jan. 2019.

[154] A. C. Chiang, G. A. Keoleian, M. R. Moore, and J. C. Kelly, "Investment cost and view damage cost of siting an offshore wind farm: A spatial analysis of Lake Michigan," *Renew. Energy*, vol. 96, pp. 966–976, Oct. 2016.

[155] H. Shiraki, S. Ashina, Y. Kameyama, S. Hashimoto, and T. Fujita, "Analysis of optimal locations for power stations and their impact on industrial symbiosis planning under transition toward low-carbon power sector in Japan," *J. Cleaner Prod.*, vol. 114, pp. 81–94, Feb. 2016.

[156] M. K. Firozjaei, O. Nematollahi, N. Mijani, S. N. Shorabeh, H. K. Firozjaei, and A. Toomanian, "An integrated GIS-based ordered weighted averaging analysis for solar energy evaluation in iran: Current conditions and future planning," *Renew. Energy*, vol. 136, pp. 1130–1146, Jun. 2019.

[157] İ. Erol, S. Sencer, A. Özmen, and C. Searcy, "Fuzzy MCDM framework for locating a nuclear power plant in Turkey," *Energy Policy*, vol. 67, pp. 186–197, Apr. 2014.

[158] V. Katinas, G. Gecevicius, and M. Marciukaitis, "An investigation of wind power density distribution at location with low and high wind speeds using statistical model," *Appl. Energy*, vol. 218, pp. 442–451, May 2018.

[159] M. Tavana, F. J. Santos Arteaga, S. Mohammadi, and M. Alimohammadi, "A fuzzy multi-criteria spatial decision support system for solar farm location planning," *Energy Strategy Rev.*, vol. 18, pp. 93–105, Dec. 2017.

[160] C. Calderón, P. James, J. Urquizo, and A. McLoughlin, "A GIS domestic building framework to estimate energy end-use demand in UK sub-city areas," *Energy Buildings*, vol. 96, pp. 236–250, Jun. 2015.

[161] I. Guaita-Pradas, I. Marques-Perez, A. Gallego, and B. Segura, "Analyzing territory for the sustainable development of solar photovoltaic power using GIS databases," *Environ. Monitor. Assessment*, vol. 191, no. 12, Dec. 2019.

[162] M. Xin, M. Lu, and W. Li, "An adaptive collaboration evaluation model and its algorithm oriented to multi-domain location-based services," *Expert Syst. Appl.*, vol. 42, no. 5, pp. 2798–2807, Apr. 2015.

[163] Q. Ma, Y.-F. Liu, and J. Huang, "Time and location aware mobile data pricing," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2599–2613, Oct. 2016.

[164] F. Ahmed, J. Erman, Z. Ge, A. X. Liu, J. Wang, and H. Yan, "Detecting and localizing End-to-End performance degradation for cellular data services based on TCP loss ratio and round trip time," *IEEE/ACM Trans. Netw.*, vol. 25, no. 6, pp. 3709–3722, Dec. 2017.

[165] A. Chattopadhyay, B. Blaszczyszyn, and E. Altman, "Location aware opportunistic bandwidth sharing between static and mobile users with stochastic learning in cellular networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 8, pp. 1802–1815, Aug. 2019.

[166] F. Xu, Y. Li, H. Wang, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 1147–1161, Apr. 2017.

[167] J. Meseguer, J. M. Mirats-Tur, G. Cembrano, V. Puig, J. Quevedo, R. Pérez, G. Sanz, and D. Ibarra, "A decision support system for on-line leakage localization," *Environ. Model. Softw.*, vol. 60, pp. 331–345, Oct. 2014.

[168] S. Fernandes, M. E. Captivo, and J. Clímaco, "A DSS for bicriteria location problems," *Decis. Support Syst.*, vol. 57, pp. 224–244, Jan. 2014.

[169] M. B. Secron, M. Montaño, M. G. Miguez, A. Jonoski, J. P. S. D. Azevedo, I. Popescu, and P. C. C. Rosman, "Proposal of a hydric index to support industrial site location decision-making applying a fuzzy multi-attribute methodology," *Ecol. Indicators*, vol. 83, pp. 427–440, Dec. 2017.

[170] J. R. Kinobe, T. Bosona, G. Gebresenbet, C. B. Niwagaba, and B. Vinnerås, "Optimization of waste collection and disposal in Kampala city," *Habitat Int.*, vol. 49, pp. 126–137, Oct. 2015.

[171] J. Kinne and B. Resch, "Analysing and predicting micro-location patterns of software firms," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 1, p. 1, Dec. 2018.

[172] H. Wu, K. Yue, C.-H. Hsu, Y. Zhao, B. Zhang, and G. Zhang, "Deviation-based neighborhood model for context-aware QoS prediction of cloud and IoT services," *Future Gener. Comput. Syst.*, vol. 76, pp. 550–560, Nov. 2017.

[173] J. Liu, M. Tang, Z. Zheng, X. Liu, and S. Lyu, "Location-aware and personalized collaborative filtering for Web service recommendation," *IEEE Trans. Services Comput.*, vol. 9, no. 5, pp. 686–699, Sep. 2016.

[174] M. A. Islam, K. Ahmed, H. Xu, N. H. Tran, G. Quan, and S. Ren, "Exploiting spatio-temporal diversity for water saving in geo-distributed data centers," *IEEE Trans. Cloud Comput.*, vol. 6, no. 3, pp. 734–746, Jul. 2018.

[175] J. E. Kang and W. Recker, "Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles," *Transp. Sci.*, vol. 49, no. 4, pp. 767–783, Nov. 2015.

[176] A. Tajbakhsh and A. Shamsi, "A facility location problem for sustainability-conscious power generation decision makers," *J. Environ. Manage.*, vol. 230, pp. 319–334, Jan. 2019.

[177] N. D. S. S. Hermeto, V. J. M. Ferreira Filho, and L. Bahiense, "Logistics network planning for offshore air transport of oil rig crews," *Comput. Ind. Eng.*, vol. 75, pp. 41–54, Sep. 2014.

[178] P. Amir-Heidari and M. Raie, "Response planning for accidental oil spills in persian gulf: A decision support system (DSS) based on consequence modeling," *Mar. Pollut. Bull.*, vol. 140, pp. 116–128, Mar. 2019.

[179] D. Strong, M. Kay, B. Conner, T. Wakefield, and G. Manogharan, "Hybrid manufacturing—Integrating traditional manufacturers with additive manufacturing (AM) supply chain," *Additive Manuf.*, vol. 21, pp. 159–173, May 2018.

[180] S. Wu and Z. Yang, "Optimizing location of manufacturing industries in the context of economic globalization: A bi-level model based approach," *Phys. A, Stat. Mech. Appl.*, vol. 501, pp. 327–337, Jul. 2018.

[181] J. Moellmann and V. M. Thomas, "Social enterprise factory location and allocation model: Small scale manufacturing for East Africa," *Socio-Econ. Planning Sci.*, vol. 68, Mar. 2019, Art. no. 100694.

[182] G. Dong, J. Ma, R. Wei, and J. Haycox, "Electric vehicle charging point placement optimisation by exploiting spatial statistics and maximal coverage location models," *Transp. Res. D, Transp. Environ.*, vol. 67, pp. 77–88, Feb. 2019.

[183] G. Brandstätter, M. Kahr, and M. Leitner, "Determining optimal locations for charging stations of electric car-sharing systems under stochastic demand," *Transp. Res. B, Methodol.*, vol. 104, pp. 17–35, Oct. 2017.

[184] A. Karasan, I. Kaya, and M. Erdogan, "Location selection of electric vehicles charging stations by using a fuzzy MCDM method: A case study in Turkey," *Neural Comput. Appl.*, vol. 32, no. 9, pp. 4553–4574, 2020.

[185] D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, and Y. Wu, "SmartAdP: Visual analytics of large-scale taxi trajectories for selecting billboard locations," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 1–10, Jan. 2017.

[186] A. Anagnostopoulos, F. Petroni, and M. Sorella, "Targeted interest-driven advertising in cities using Twitter," *Data Mining Knowl. Discovery*, vol. 32, no. 3, pp. 737–763, May 2018.

[187] X. Li, L. Yang, Y. Li, and C. Zhou, "Deep trajectory: A deep learning approach for mobile advertising in vehicular networks," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2813–2825, Jul. 2019.

[188] M. Badami, F. Tafazzoli, and O. Nasraoui, "A case study for intelligent event recommendation," *Int. J. Data Sci. Anal.*, vol. 5, no. 4, pp. 249–268, Jun. 2018.

[189] J.-Y. Jiang and C.-T. Li, "Who should I invite: Predicting event participants for a host user," *Knowl. Inf. Syst.*, vol. 59, no. 3, pp. 629–650, Jun. 2019.

[190] Y.-H. Hsiao, L.-F. Chen, C.-C. Chang, and F.-H. Chiu, "Configurational path to customer satisfaction and stickiness for a restaurant chain using fuzzy set qualitative comparative analysis," *J. Bus. Res.*, vol. 69, no. 8, pp. 2939–2949, Aug. 2016.

[191] Y. Yang, W. S. Roehl, and J.-H. Huang, "Understanding and projecting the restaurantscape: The influence of neighborhood sociodemographic characteristics on restaurant location," *Int. J. Hospitality Manage.*, vol. 67, pp. 33–45, Oct. 2017.

[192] A. Gimpel, V. Stelzenmüller, B. Grote, B. H. Buck, J. Floeter, I. Nú nez-Riboni, B. Pogoda, and A. Temming, "A GIS modelling framework to evaluate marine spatial planning scenarios: Co-location of offshore wind farms and aquaculture in the German EEZ," *Mar. Policy*, vol. 55, pp. 102–115, May 2015.

[193] K. C. Malpeli and P. G. Chirico, "A sub-national scale geospatial analysis of diamond deposit lootability: The case of the central African republic," *Extractive Ind. Soc.*, vol. 1, no. 2, pp. 249–259, Nov. 2014.

[194] Z. Cheng and J. Shen, "On effective location-aware music recommendation," *ACM Trans. Inf. Syst.*, vol. 34, no. 2, pp. 1–32, Apr. 2016.

[195] W. Xu and C.-Y. Chow, "A location-and diversity-aware news feed system for mobile users," *IEEE Trans. Services Comput.*, vol. 9, no. 6, pp. 846–861, Nov. 2016.

[196] C.-I. Hsu and W.-T. Chen, "Optimizing fleet size and delivery scheduling for multi-temperature food distribution," *Appl. Math. Model.*, vol. 38, no. 3, pp. 1077–1091, Feb. 2014.

[197] L. A. Sierra, V. Yepes, T. García-Segura, and E. Pellicer, "Bayesian network method for decision-making about the social sustainability of infrastructure projects," *J. Cleaner Prod.*, vol. 176, pp. 521–534, Mar. 2018.

[198] J. North and F. L. Miller, "Facility location using GIS enriched demographic and lifestyle data for a traveling entertainment troupe in Bavaria, Germany," *Decision Support Syst.*, vol. 99, pp. 30–36, Jul. 2017.

[199] K. Mommens and C. Macharis, "Location analysis for the modal shift of palletized building materials," *J. Transp. Geography*, vol. 34, pp. 44–53, Jan. 2014.

NORHA M. VILLEGAS (Senior Member, IEEE) received the B.Sc. degree in software systems engineering and a Specialist Diploma in management of information systems from Universidad Icesi, Cali, Colombia, and the Ph.D. degree in computer science (with focus on software engineering) from the University of Victoria, Canada. She is currently an Associate Professor with the Department of Information and Communication Technologies, Universidad Icesi, Colombia, the Director of the Software Systems Engineering Bachelor program, and an Adjunct Assistant Professor with the University of Victoria. She has 20 years of professional experience in academia and industry. Her research has been recognized internationally by her peers for its quality and relevance. She has published an important number of book chapters and refereed journal and conference proceedings articles that have been co-authored with top researchers worldwide. She investigates the application of software engineering models, techniques, and architectures to the development of smart software systems, i.e., systems that are context-aware and self-adaptive. In particular, she is interested in the application of feedback loops, dynamic context management, and autonomic computing mechanisms to interdisciplinary areas that are crucial for the advancement of society, i.e., smart cyber-physical systems, cognitive computing, and digital twins. She is also highly engaged in Engineering Education research. Over the last nine years, she has co-chaired several international workshops, and served as a program committee member for many IEEE international conferences and symposia in the field of software engineering.

LUIS E. FERRO-DÍEZ received the B.Sc. degree in software systems engineering from the Institución Universitaria Antonio José Camacho, Cali, Colombia, and the M.Sc. degree in informatics and telecommunications (with focus on machine learning engineering) from Universidad Icesi, Cali. He is a Software and Machine Learning Research Engineer with over 13 years of professional experience in several technology sectors. He has been a team lead for different top-tier location and technology projects, helping with design, implementation, validation, and delivery endeavors. His area of expertise includes software engineering and development; system and architecture design; big data platforms; cloud environments; and machine learning. He has worked with important clients in the technology industry, including HERE Technologies, and IPG Mediabrands (Kinesso) under PSL - Perficient, Colombia, designing and implementing data-oriented solutions to support critical business processes. He is currently working as a Researcher with the Institute of Advanced Research in Artificial Intelligence (IARAI), Vienna, Austria.

JAVIER DÍAZ-CELY received the B.Sc. degree in computer engineering from the Pontificia Universidad Javeriana, Cali, Colombia, the master's degree in artificial intelligence, pattern recognition, and applications and the Ph.D. degree in informatics from Sorbonne Université, and the master's degree in corporate finance from the Conservatoire National des Arts et Métiers, Paris, France. He is currently the Chief Analytics Officer of MO Technologies, a Fin-tech that aims at revolutionizing the credit scoring industry and promoting financial inclusion, by giving everyone, independently from their credit history, the possibility to access nano and microloans in a simple and effective way, leveraging innovative credit scoring through the use of alternative data sources. He believes that business analytics offers an incredible set of tools that enable organizations to achieve strategic goals by optimizing processes, solving problems, and discovering new and disruptive opportunities for creating value from data. With more than 15 years of experience in the field of artificial intelligence in France and Colombia, he has gathered considerable experience applying machine learning to real-world problems in several sectors, i.e., banking and finance, telecommunications, retail, health, and education; either working directly for a company (Orange France Télécom, Sociéte Générale, Banco Falabella) or as a Consultant (Altran, Universidad Icesi). He also had the chance to work as a University Professor and Applied Analytics Researcher, first as a part of the LIP6 Laboratory, Paris, and currently as a Faculty Member of the School of Engineering, Universidad Icesi, where he was also the Director of the Master's in Data Science program.

- - -