

Contextual Modeling and Intelligent Decision-Making for IoT Systems: A Combined Ontology and Machine Learning Approach [†]

Sanaa Mouhim 

Interdisciplinary Applied Research Laboratory-LIDRA, International University of Agadir-Universiapolis, Agadir 08000, Morocco; sanaa.mouhim@e-polytechnique.ma

[†] Presented at the 7th edition of the International Conference on Advanced Technologies for Humanity (ICATH 2025), Kenitra, Morocco, 9–11 July 2025.

Abstract

In the context of the Internet of Things (IoT), this article proposes an innovative approach combining ontologies and the Apache Spark MLlib library to design an intelligent system capable of dynamically adapting to its environment. The aim is to model the context including users, devices, events, and environmental conditions, and exploit massive sensor data to generate intelligent, contextualized predictions. The architecture relies on two pillars: an ontology as a formal way to structure and semantically annotate knowledge and Spark MLlib in order to execute big data machine learning algorithms and notably random forest regression. The solution is targeted to real-time applications such as energy or air quality management in smart homes. The results demonstrate the value of combining ontology and machine learning in order to improve contextual knowledge and automatic decision-making.

Keywords: context; context awareness; smart home; ontologies; Spark MLlib; stream processing



Academic Editors: Brahim El Bhiri and Ayman S. Mosallam

Published: 18 November 2025

Citation: Mouhim, S. Contextual Modeling and Intelligent Decision-Making for IoT Systems: A Combined Ontology and Machine Learning Approach. *Eng. Proc.* **2025**, *112*, 71. <https://doi.org/10.3390/engproc2025112071>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Smart homes based on the Internet of Things (IoT) are transforming how we interact with our living environment [1]. Combining sensors, actuators, and high-level communication networks, these systems enable the efficient control of comfort, security, and energy conservation. However, the design and deployment of such systems pose certain modeling and architectural challenges.

One of the main concerns is the complexity of handling heterogeneous devices as part of a unified and interoperable environment. Smart homes need to handle data streams from multiple sensors (temperature, humidity, air quality, presence, etc.) and synchronize the actions in real time while making the system scalable and robust.

Another strong limitation is modeling context and interactions. The IoT generates vast amounts of data that must be analyzed and processed very well. Models constructed need to be able to encapsulate the dynamic context of a smart home, including user preference, environmental condition, and usage context [2].

Moreover, security and privacy requirements place stringent architectural choices. Open user exposure of their sensitive data to attacks compels them to employ stringent authentication, encryption, and access control within their architectural design choices [3].

Finally, the scalability and maintainability of smart home Smart Home architectures pose another challenge. The systems must be designed so that they are scalable by incor-

porating new devices and services without compromising their performance. Modularity and ontology-based architecture can help achieve better interoperability and effective management of complexity.

The research methodology adopted in this study is based on an interdisciplinary approach combining artificial intelligence, the Internet of Things (IoT), and contextual modeling techniques. To propose an effective and innovative solution, we first carried out an exhaustive review of different contextual modeling and reasoning methods to identify the best practices and their limitations in IoT environments. Then, a hybrid model was developed, integrating ontologies for semantic data management and machine learning techniques, notably via the Apache Spark 4.0.1 MLlib library, to perform large-scale contextual predictions.

2. Toward a Better Understanding of Context and Context Awareness

Context refers to all relevant information that characterizes the situation of a system, user, or environment at a given moment. In intelligent systems, context enables a system's behavior to be adapted to current conditions and interactions.

Schilit and Theimer [4] define context as “any information that can be used to characterize the situation of an entity”, where an entity can be a person, a place, or an object that interacts with a computer system.

Context awareness is the ability of a system to automatically sense, understand, and respond to the environment using contextual information present [5]. Context awareness is a central concept to IoT, artificial intelligence, and intelligent systems because it enables devices and applications to respond pre-actively and intelligently without direct human intervention.

A context-aware system usually follows three primary steps:

- Context perception: collect data from sensors, networked devices, databases, etc.
- Context interpretation: use of artificial intelligence algorithms, logical rules, or ontologies to infer situations from contextual information and analysis.
- Action and adaptation: system behavior adjustment according to the determined context.

Of the modeling approaches, key-value pair models are simple and sufficient for the manipulation of basic contexts such as the status of a device [6]. Object-oriented models, grounded in object-oriented design principles, allow for the manipulation of complex objects with characteristics, associations, and inheritance rules [7]. Finally, ontology-based models allow for a formal shared representation of the knowledge, enabling interoperability as well as machine-based reasoning and enabling implicit knowledge to be inferred through axioms [8].

Based on contextual reasoning, supervised learning such as decision trees and support vector machines is utilized to train the models against specific data sets [9]. Unsupervised learning, instead, is used to identify patterns and anomalies of data without labels. Simple yet efficient IF-THEN rules are used to link contextual variables to one another, while ontological reasoning is utilized to deduce new information by applying logical rules as well as axioms [9].

However, these approaches have their limitations. Key-value pair models are poor at uncertainty and complex relationships, and object-oriented models are rigid for evolving systems. Ontologies, while excellent at knowledge representation, are computationally expensive and difficult to extend. Supervised learning methods require large quantities of labeled data, and unsupervised learning is difficult to quantify without well-defined criteria. To add to that, rule-based systems become more complicated and difficult to maintain with additional rules.

3. Related Works

IoT-based smart home research has been the subject of extensive study, dealing with various facets of methodologies and technological solutions to improve their efficiency and reliability. In this section, we review the main contributions in the field, reporting on the recent advances, challenges raised, and solutions adopted to overcome current limitations. In particular, we report on the modeling approaches, architectures deployed, and security and data management mechanisms that have been reported in the scientific literature.

The article “Taxonomy and Software Architecture for Real-time Context-aware Collaborative Smart Environments” [10] proposes a taxonomy and a JSON structure for the homogeneous definition of domain and contextual data in IoT environments. It also presents a software architecture that combines the processing of this data using complex event processing (CEP) technologies. The objective is to improve the contextualization of situations by integrating data from different domains, which facilitates decision-making. The study demonstrates the feasibility and effectiveness of this approach through performance tests on various system configurations.

The work presented in [11] describes a contextual middleware for modeling and reasoning in smart cultural environments that is based on a hybrid of contextual categorization schemas. It defines a model based on five major classes: reason, activity, time, location, and thing. It aims to improve user interactions and facilitate preventive conservation of artifacts.

Authors of [12] present SEArch as an architecture for intelligent environments such as smart homes. They give great importance to context information to understand and anticipate user behaviors, which is crucial for the system’s operation. The contextual elements mentioned in the document are mainly environmental information, the activity understanding module, user profiles, and APIs for external data. Several classes or elements are used to structure the system and ensure its efficient operation in smart environments. The main classes are sensors, actuators, profile, context, and reasoning. These classes interact dynamically to offer a contextually relevant service environment, allowing users to benefit from personalized assistance in their daily lives.

The article [13] presents Context-Oriented Behavioral Programming (COBP), a new approach to the development of contextual systems, which combines behavioral programming with context-oriented idioms. COBP allows for the natural and incremental specification of context-dependent behaviors, thus facilitating the management of complex requirements. The author provides formal semantics for COBP and demonstrates its effectiveness through real-world case studies in robotics and IoT.

The work presented in [14] explores how to model environmental data in smart environment systems to improve their efficiency and adaptability. It highlights the use of ontologies to structure knowledge, as well as data models that define the relationships between various types of information. Processing algorithms, like machine learning, are employed to analyze contextual data and trigger corresponding action. The systems also create user profiles, updating preferences and behavior dynamically to provide personalized services. Finally, dynamic interaction strategies are employed to make the system’s communication adaptive to the gathered data.

The article [15] mentions that the proposed smart home system is based on the use of sensors and actuators allowing remote control of appliances using the Internet of Things (IoT). For the context, techniques such as facial detection and motions are integrated to enhance authentication security. Data are structured through a modular architecture whose main components are Security, Web services, data exchange, and notifications.

The article [16] mentions that human activity recognition is essential for environments such as smart homes or health centers allowing the adaptation of the systems according to the residents' activities and monitoring patients' routines. The document does not specifically detail the environmental data used for context awareness, but it mentions that smart environments are equipped with embedded systems and sensors that can capture and communicate information.

In the paper by Bouroumi et al. [17], contextual data is modeled using several approaches, mainly through metamodels that are used to structure and organize this information. The authors propose a context metamodel to represent contextual information in a structured way. This metamodel consists of several elements: the global context, the sub-context which includes specific properties, and the contextual relationships which define the interactions between these elements. A business process model for integrating this contextual data includes an ontology and a modular architecture that facilitates the reuse and adaptation of the models in different systems or applications.

Table 1 presents a comparative summary of several papers [10–17] on context-aware systems, focusing on contextual data, system components, data structures, event processing, and the advanced technologies used.

Table 1. A comparative study of context-aware systems focused on contextual data, system components, data structures, event processing, and the advanced technologies used.

Article	Contextual Data	System Components	Data Structure	Event Processing	Advanced Technologies
[10]	Person, Place, Environment, Technology, Activity	Event Types, CEP Engine, Semi-Static Databases, REST Services, Actions and Persistent Context	JSON format	Complex event processing engine	Provides for integration of ML
[11]	User Data, Environmental Data, Temporality, Activity Data, Derived Data	Context Acquisition, Pre-processing, Reasoning Module, Incertitude Module, Machine Learning Module, Knowledge Base, Applicative Interface	NoSQL database, specifically, MongoDB	Rule-based, if-then type	Hybrid Reasoning, Machine Learning, Fuzzy Logic and Probabilistic Logic
[12]	x	Sensors and Actuators, Reasoning Components, User Interfaces, Data Management	x	Rule-based language, a case-based raiser	Provides for ML integration
[13]	Physical Entities, Device Status, Users, Contextual Conditions, Requiring and Updating	Contextual Behaviors, Context Model, Architectural Layers	A relational database, managed by a data access layer	Event Arbiter, Effect Functions, Compartmental Threads, Conditional Transitions	Artificial Intelligence, Big data, Ontology, Distributed systems, IoT
[14]	User personal data, Environmental data, Specific contextual information, Temporal data, Relationships between entities	Compartmental Threads, Data Access Layer, Context Schema, Events, Effect Functions	Ontologies, data models, NoSQL database, graphs	Reactive event processing, Temporal event scheduling, Use of decision algorithms	Artificial Intelligence, Big data, Ontology, Multi-agent systems
[15]	Sensor data, Electronic device data, User data, Safety data Environmental conditions	User Interface, Transmission Mode, Central Controller, Electronic Devices Connection Manager, Notification Module, Web Service, Object Management Module	x	Automatization rules, Order Management Module, Artificial Intelligence and Machine Learning	Artificial Intelligence, Big data, Ontology, Multi-agent systems Facial Recognition and Liveness Detection, Chatbots and Natural Language Processing (NLP), Machine Ap-learning and Artificial Intelligence

Table 1. Cont.

Article	Contextual Data	System Components	Data Structure	Event Processing	Advanced Technologies
[16]	Sensor data, Temporal data, Interaction data, Activity history, Personal context, Location data	Sensors, Data collection, Data pre-processing, Machine learning algorithms, User interface, Data storage and management, Feedback system	x	Temporal analysis, Contextual interpretation, Re-action, and adaptive learning	Computer Vision, Artificial Intelligence and Machine Learning, Cloud Computing and Big Data
[17]	Location, identity, localization, sensor data, event	Sensors, Context, Contextualized Business Process Model, Ontology of Contextualized Business Process Model, Support Infrastructure	x	Events, Integration of Contextual Elements, Task and Gateway Management, Notification/Alert Systems	Ubiquitous Computing, Iot, Data Analysis

A detailed analysis of this work has enabled us to highlight the following points:

- Contextual Data: The majority of articles cover a variety of data types, such as user, environmental, temporal, and activity data. Some works [10,13,14] include relationships between entities or contextual conditions, enabling richer modeling.
- System Components: Architectures are well detailed, with specialized modules such as reasoning [11], event management [10,13], or user interfaces [15]. Some works [11,14], integrate machine learning (ML) modules and knowledge bases. However, few articles deal with load management or component distribution.
- Data structure: Diverse formats: JSON [10], NoSQL [11,14], ontologies [13,14], and relational databases [9]. Some works [10] use graphs to model relationships between entities. A hybrid combination of several data structures would be an ideal solution.
- Event processing: Various approaches: CEP engines [10], if-then rules [11,12], and temporal algorithms [14]. Some works, [13,14], integrate effect functions or reactive threads. This work can be enhanced by real-time processing tools such as Flink, KAFKA, or Spark.
- Advanced technologies: Wide adoption of AI/ML [10,11,14,15], ontologies [13,14], and big data [16] as well as emerging technologies such as multi-agent systems [14], [15] or computer vision [16]. However, no details are provided on the implementation of the models deployed, or on the frameworks or tools used.

To fill these gaps, we present in the following section a complete description of smart home context awareness system components.

4. Context Awareness System Components

4.1. Contextual Data

- User-related data: This data enables the home's functionalities to be tailored to the needs and preferences of its inhabitants:
 - Identity: facial recognition, RFID badges.
 - Presence: motion detection.
 - Personal preferences: temperature, lighting, music, daily routines.
 - Current activities: watching TV, cooking, sleeping, working.
- Usage history: frequent settings, devices used at certain times of the day. Environmental data: These enable the home to be adapted to suit climatic conditions and comfort:
 - Indoor and outdoor temperature;
 - Air humidity;
 - Air quality (CO₂, COV, fine particles PM2.5);

- Ambient brightness;
 - Rain, wind, external weather conditions;
 - Smoke and toxic gas detection;
 - State of windows and doors (open/closed).
- Data from connected devices: They can be used to monitor and control intelligent objects in the home.
 - Appliance status (on/off, operating mode);
 - Energy consumed by each appliance;
 - Light status (intensity).
- Time data:
 - Day/night cycle (automatic activation of shutters, lights, heating);
 - Frequency of equipment use (habit detection).
- Security data: This information guarantees the safety of the home and its occupants.
 - Intrusion detection (cameras, motion sensors, unauthorized door opening);
 - User authentication (facial recognition, NFC badges);
 - Fire and gas leak detection;
 - Alert and notification system (SMS, email, alarm siren).
 - Data from interactions with the home: This information concerns the occupants' interactions with the system.
 - Voice commands;
 - Interactions via a mobile app or web interface;
 - Management of pre-programmed scenarios.

4.2. Data Structure

For a context-aware system, the choice of the database depends on the specific requirements of the system, including the nature of the data, performance needs, scalability, and flexibility. While our system integrates Artificial Intelligence (AI) for prediction, it is important to choose databases that not only efficiently manage historical data, but also enable fluid integration with machine learning (ML) algorithms and predictive systems.

For critical data requiring ACID transactions, we opted for a relational database such as PostgreSQL. This mainly concerns data on users, access rights, parts, devices, and their status. NoSQL databases are ideal for unstructured data, real-time data streams, and sensors. Table 2 shows the databases used for the different types of contextual data presented in Section 4.1.

Table 2. Recommended database for different data types used in the context of smart home system.

Data Type	Recommended Database	Reason
User-related data	PostgreSQL	Ideal for structured data with complex relations
Environmental data	PostgreSQL	Can be used for temporal data (temperature, humidity, air quality) with efficient time-series queries.
Data from connected devices	MongoDB	Flexible for semi-structured device data (status, energy consumption).
Time data	PostgreSQL	Use time functions to manage day/night cycles and frequency of use.
Security data	PostgreSQL	Guarantees secure, structured management of sensitive data (intrusion, authentication)
Data from interactions with the home	MongoDB	Adapted to handle a variety of interactions (voice commands, mobile or web interfaces).

4.3. Event Processing

For event processing, we opt to use advanced technologies such as Kafka and Spark, enabling us to efficiently process large quantities of data in real time, while providing large-scale analysis and prediction capabilities.

As shown in Figure 1, the integration of Kafka and Spark for event processing works as follows:

- Event collection: events from sensors or smart devices are sent to Kafka topics.
- Event processing: Spark Streaming consumes these events from Kafka and processes them in real time to perform specific analyses, filtering, and calculations. For example, determining whether the temperature exceeds a certain threshold, or whether a person is detected by the camera.
- Prediction and decision-making: thanks to its MLlib library, Spark applies artificial intelligence models to predict actions to be taken, such as adjusting the temperature or sending a security alert.
- Results storage: The results of processing, i.e., predictions or decisions, are sent to Kafka Topics to trigger actions in the smart home. Thanks to Kafka Connect, this data is stored in a database such as PostgreSQL or MongoDB for future access.

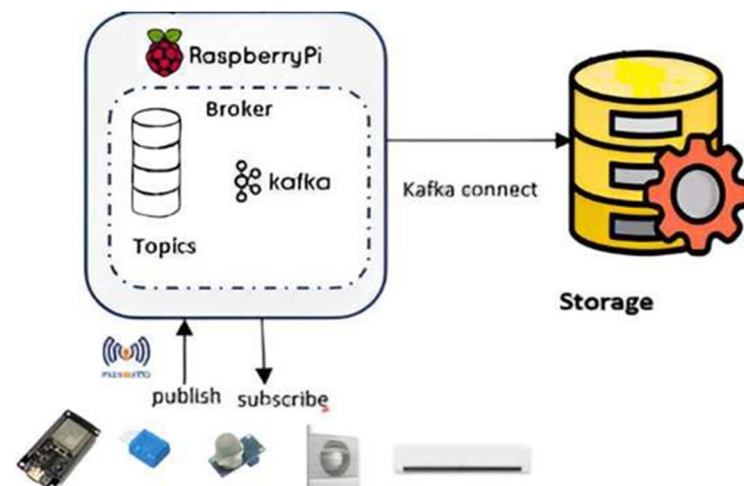


Figure 1. The use of Kafka topics and Kafka connect for events handling and database storage.

4.4. Advanced Technologies

As shown in Figure 2, our approach is based on the collaborative use of an ontology and the Apache Spark MLlib library to design an intelligent system able to understand and react to its environment. The ontology is used to model the various components of the context, such as entities (users, devices, environment), actions, events, as well as spatial and temporal relationships. Using formal languages such as RDF or OWL provides a rich semantic representation of domain knowledge. In parallel, Apache Spark MLlib is an open-source machine learning library designed to handle very large volumes of data, making it particularly well suited to IoT environments where data is produced continuously. Its integration with Spark Streaming facilitates the processing of real-time data streams, useful, in particular, for continuous monitoring of air quality or energy consumption. MLlib offers a wide range of learning algorithms (classification, regression, clustering, recommendation), enabling the generation of predictive models based on the analysis of historical data.



Figure 2. The reasoning layer combining Ontologies for context modeling and Spark MLlib for result interpretation and decision-making.

By integrating an ontology with MLlib, data is given a semantic meaning that helps guide the modeling process and makes it easier to interpret the results. The concepts in the ontology are used to label the data, making the models easier to understand. This technique might, for instance, forecast a user's preferred temperature in a smart home setting based on weather and other variables like the time of day. By connecting predictions and suggestions to domain-specific knowledge, ontology helps in providing them with meaning.

5. Discussion and Conclusions

In this paper, we propose an innovative approach that combines the use of an ontology with the massive processing and machine learning capabilities offered by Apache Spark MLlib. The aim is to build an intelligent system able to dynamically understand and adapt to its contextual environment, particularly in smart home applications. The ontology enables fine semantic modeling of context, while Spark MLlib ensures efficient processing of massive real-time data streams from IoT sensors.

The integration of these two technologies not only improves contextual understanding but also optimizes prediction and automatic decision-making based on historical data. These techniques provide a solid framework for developing intelligent, adaptive, and scalable systems.

However, some improvements can be envisioned to enhance the performance of the architecture shown. These mainly involve the refinement of the ontology by enriching the knowledge base with more sophisticated logical rules in order to enable more contextualized and more detailed reasoning. In addition, the use of deep learning methods, as a complement to Spark ML, seems to be an interesting research path to enhance predictions and provide a better description of situations. Finally, the addition of a security layer dedicated to the protection of sensitive data collected. This layer must guarantee data confidentiality, integrity, and access control, in particular through encryption mechanisms, strong authentication, and fine-tuned context-dependent authorization management. By integrating security right from the design stage, we can reinforce user confidence in these intelligent systems.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Strang, T.; Linnhoff-Popien, C. A Context Modeling Survey. In Proceedings of the Workshop on Advanced Context Modeling, Reasoning, and Management as Part of UbiComp 2004—6th International Conference on Ubiquitous Computing, Nottingham, UK, 7 September 2004.
2. Colace, F.; Lorusso, A.; Marongiu, F.; Santaniello, D.; Troiano, A.; Valentino, C. An Internet of Things based approach for Smart Home Management. *Res. Briefs Inf. Commun. Technol. Evol.* **2022**, *8*, 182–192. [\[CrossRef\]](#)
3. Li, M.; Wu, Y. Intelligent control system of smart home for context awareness. *Int. J. Distrib. Sens. Netw.* **2022**, *18*, 155013292210820. [\[CrossRef\]](#)
4. Schilit, B.; Adams, N.; Want, R. Context-aware computing applications. In Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications, Santa Cruz, CA, USA, 8–9 December 1994.
5. Perera, C.; Zaslavsky, A.; Christen, P.; Georgakopoulos, D. Context Aware Computing for The Internet of Things: A Survey. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 414–454. [\[CrossRef\]](#)
6. Robles, R.J.; Kim, T.-H. Context Aware Systems, Methods and Trends in Smart Home Technology. In Proceedings of the International Conference on Security-Enriched Urban Computing and Smart Grid, Hualien, Taiwan, 21–23 September 2011.
7. Cristea, V.; Dobre, C.; Pop, F. Context-Aware Environments for the Internet of Things. In *Internet of Things and Inter-Cooperative Computational Technologies for Collective Intelligence*; Bessis, N., Xhafa, F., Varvarigou, D., Hill, R., Li, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 460.
8. Guermah, H.; Fissaa, T.; Hafiddi, H.; Nassar, M.; Kriouile, A. An Ontology Oriented Architecture for Context Awar.pdf. Available online: <https://arxiv.org/pdf/1404.3280> (accessed on 24 May 2024).
9. Prasetya Dwi Wibawa, I.; Kallista, M.; Phaijoo, G.R. A Literature Survey of Human Activity Recognition Using Deep Learning and Nonparametric Model with Some Exchanges in Karl Popper’s Viewpoint and Kuhn’s Paradigm in Philosophy of Science. *J. Meas. Electron. Commun. Syst.* **2022**, *9*, 18–28. [\[CrossRef\]](#)
10. Bazan-Muñoz, A.; Ortiz, G.; Augusto, J.; Garcia-de-Prado, A. Taxonomy and software architecture for real-time context-aware collaborative smart environments. *Internet Things* **2024**, *26*, 101160. [\[CrossRef\]](#)
11. Michalakakis, K.; Christodoulou, Y.; Caridakis, G.; Voutos, Y.; Mylonas, P. A Context-Aware Middleware for Context Modeling and Reasoning: A Case-Study in Smart Cultural Spaces. *Appl. Sci.* **2021**, *11*, 5570. [\[CrossRef\]](#)
12. Augusto, J.; Giménez-Manuel, J.; Quinde, M.; Oguego, C.; Ali, M.; James-Reynolds, C. A Smart Environments Architecture (Search). *Appl. Artif. Intell.* **2020**, *34*, 155–186. [\[CrossRef\]](#)
13. Elyasaf, A. Context-Oriented Behavioral Programming. *Inf. Softw. Technol.* **2020**, *133*, 106504. [\[CrossRef\]](#)
14. Babli, M.; Onaindia, E. A context-aware knowledge acquisition for planning applications using ontologies. *arXiv* **2019**, arXiv:1904.09845. [\[CrossRef\]](#)
15. Fahd Al-Mutawa, R.; Albouraey Eassa, F. A Smart Home System based on Internet of Things. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2020**, *11*, 260–267. [\[CrossRef\]](#)
16. Hussain, Z.; Sheng, Q.Z.; Zhang, W.E. A review and categorization of techniques on device-free human activity recognition. *J. Netw. Comput. Appl.* **2020**, *167*, 102738. [\[CrossRef\]](#)
17. Bouroumi, J.E.; Guermah, H.; Nassar, M. Enhancing Business Process Modeling with Context and Ontology. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2021**, *12*, 373–380. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.