Review

Semantic Data Mining in Ubiquitous Sensing: A Survey

Grzegorz J. Nalepa [1,2,*] , Szymon Bobek [1,2] , Krzysztof Kutt [1] and Martin Atzmueller [3,*]

[1] Institute of Applied Computer Science and Jagiellonian Human-Centered Artificial Intelligence Laboratory (JAHCAI), ul. Prof. Stanislawa Lojasiewicza 11, Jagiellonian University, 30-348 Krakow, Poland; szymon.bobek@uj.edu.pl (S.B.); krzysztof.kutt@uj.edu.pl (K.K.)

[2] Department of Applied Computer Science, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland

[3] Semantic Information Systems Group, Osnabrück University, 49074 Osnabrück, Germany

* Correspondence: gjn@gjn.re (G.J.N.); martin.atzmueller@uni-osnabrueck.de (M.A.)

Abstract: Mining ubiquitous sensing data is important but also challenging, due to many factors, such as heterogeneous large-scale data that is often at various levels of abstraction. This also relates particularly to the important aspects of the explainability and interpretability of the applied models and their results, and thus ultimately to the outcome of the data mining process. With this, in general, the inclusion of domain knowledge leading towards semantic data mining approaches is an emerging and important research direction. This article aims to survey relevant works in these areas, focusing on semantic data mining approaches and methods, but also on selected applications of ubiquitous sensing in some of the most prominent current application areas. Here, we consider in particular: (1) environmental sensing; (2) ubiquitous sensing in industrial applications of artificial intelligence; and (3) social sensing relating to human interactions and the respective individual and collective behaviors. We discuss these in detail and conclude with a summary of this emerging field of research. In addition, we provide an outlook on future directions for semantic data mining in ubiquitous sensing contexts.

Keywords: semantics; data mining; declarative methods; explainability; industrial sensors

1. Introduction

With the advent of ubiquitous sensing driven by, for example, mobile computing, the Internet of Things and Industry 4.0, many novel research directions and interesting applications have emerged through the use of large-scale sensor data as well as advanced analysis and processing methods. While there is a multitude of such powerful processing and analytics methods in data mining (DM) and machine learning (ML), there are also specific challenges relating to the characteristics of the data in ubiquitous sensing. These include, for example, some of the common challenges of Big Data [1,2] such as volume, velocity and variety of the data. However, most importantly, aspects such as the interpretability and explainability of the applied data mining models and their results, respectively, motivate, require, or even enforce the application of domain/background knowledge to data mining and machine learning approaches.

In the context of ubiquitous sensing and sensor data processing, Big Data, for example, requires not only the proper selection and curation of potentially relevant data, but also the use of dimensionality reduction and feature construction and engineering. This also relates to making *Big Data* smart, that is, transforming it into *Smart Data* [3]. In particular, a number of important questions regarding the understandability and interpretability of machine learning models used in sensitive applications of DM and artificial intelligence (AI) (e.g., medicine) have recently been raised [4]. When applying the models for knowledge discovery and/or decision support, their transparency and explainability is often crucial, otherwise limiting acceptance and trust in their adoption in such sensitive contexts. Using background/domain knowledge enables semantic enrichment and ultimately semantic

interpretation, leading to a knowledge-based analysis approach, which we call *semantic data mining*.

This article aims to survey relevant works in these areas, specifically focusing on different semantic data mining approaches and methods, but also on specific selected—current, prominent and emerging—application areas of ubiquitous sensing. Here, we consider in particular: (1) environmental sensing; (2) ubiquitous sensing in industrial applications of artificial intelligence; and, finally, (3) social sensing relating to human interactions—observing and capturing the respective individual and collective behaviors. We discuss these in detail and conclude with a broad outlook on future directions for semantic data mining approaches in ubiquitous sensing contexts.

There exist several comprehensive surveys on using semantic knowledge in data mining [5–7] or exploring the possibilities of combining data mining with background knowledge [8], and others that tackle the issue of moving from raw data to smart data [9–12]. However, they do not approach the subject holistically, instead focusing on particular aspects of semantic data mining or narrowing the view on how semantics can be defined and introduced to the data mining pipeline. We present a broader perspective on the issue and compare existing practical tools and frameworks that aid data scientists in building ubiquitous sensing systems.

Our contributions are summarized as follows:

1. We provide a comprehensive perspective on semantic data mining, including different methods and techniques from related areas to be captured under this common topic.
2. We discuss relevant applications in the context of ubiquitous sensing, exemplifying the specific implementation of semantic techniques in context.
3. We outline interesting future directions for the development and application of approaches and methods for semantic data mining in ubiquitous sensing.

The rest of the article is structured as follows: Section 2 provides an overview of semantic data mining, and in Section 3 we discuss specific application areas in ubiquitous sensing. Finally, Section 4 concludes with a summary and discusses interesting future challenges and perspectives for semantic data mining in ubiquitous sensing.

2. Overview of Semantic Data Mining Approaches

In the following, we structure the discussed relevant literature around the topics of data mining process models, semantic and declarative approaches for data mining, and the role of interpretability and explainability in DM.

2.1. Data Mining Process

The general goal of data mining is to uncover novel, interesting and ultimately understandable patterns [13]—that is, relating to valuable, useful and implicit knowledge. It is an iterative and often incremental process, such that a partial solution is often refined in order to arrive at the final one. There exist several approaches for data mining (see [14]), the most prominent of which is provided by the CRISP-DM process. It can be roughly divided into three sub-processes: domain focusing (understanding and data preparation); pattern modeling (the mining step); and model implementation (evaluation and deployment). CRISP-DM—consisting of six phases in total—is thus split into five iterative phases: Business Understanding (defining the goals of data mining); Data Understanding (making sure that data is applicable and clarifies semantics); Data Preparation (which usually needs about 80% of the total effort of the process for transforming and cleaning the data, including feature engineering, e.g., [15,16]); Modeling (the central phase: regularities and patterns are extracted from the data for constructing the data mining model); Evaluation (where the quality of the mined model needs to be assessed); and finally Deployment (where the model is applied, e.g., for pattern understanding, prediction, classification or clustering).

It is worth noting that CRISP-DM was proposed in the context of a long-time tradition of Knowledge Discovery from Databases (KDD) [13]. An overview [14] discusses the evolution of such approaches for KDD as well as for DM processes. Some of them include

aspects that CRISP-DM omits, for example, Domain Knowledge Elicitation and Knowledge Post-processing. While a decade ago, the term "explanation" was rarely explicitly used, the explanatory aspect of the DM or KDD process was somewhat considered in the latter phases of related approaches, but still to a limited extent. Later, we will discuss how explanation is especially important in relation to our scope of much more complex DM approaches and processes today. This is, in particular, specifically relevant for semantic data mining in ubiquitous sensing due to the complex, heterogeneous, and typically uncertain and noisy, data.

Connected to this, some additional deficiencies were already identified a long time ago. First of all, it is a common opinion among DM experts that data understanding and preparation are typically the most costly and time-consuming (about 80%) phases in CRISP-DM, before even mentioning proper Business Understanding, for example, [15,16]. Moreover, the lack of feedback loops is emphasized [14].

To improve CRISP-DM, or to offer alternative solutions, more recently several new approaches have been proposed. The SAS Institute proposed its own SEMMA (Sample, Explore, Modify, Model and Assess) sequential approach for DM [17]; for a comparison with CRISP-DM see [18]. IBM proposed its own extension to the original CRISP-DM process, called ASUM-DM, to focus more on the operations side of implementing DM projects [19]. However, these two approaches remain sequential and do not consider the role of the domain knowledge, nor the explanative aspect. Most recently, the so-called *Model Development Process* was proposed [20]. It extends the ideas of Rational Unified Process, and partially considers the need to introduce explanations. However, it does not support the explicit elicitation of knowledge in any of the phases. In addition, there have been several proposals for including domain knowledge regarding the data mining process in general, for example, [21,22]. However, those mainly relate to the (declarative) specification of the process itself, or the "fine-tuning" of the process, but not to cross-links between the different steps of the process.

As such, in the next subsection we discuss a variety of approaches that aim to enhance data mining from a knowledge-oriented perspective.

2.2. Semantic, Knowledge-Based and Declarative Data Mining

Using background knowledge in DM has been proposed in the area of *semantic data mining*, where the knowledge is typically represented in a knowledge repository, such as an ontology or a knowledge base. The main aspect of semantic DM is the explicit integration of this knowledge into the DM and KDD modeling step, where the algorithms for data mining/modeling or post-processing make use of the formalized knowledge to improve the overall results. There has been growing interest in this issue (e.g., [23–28]) in various domains, for example in the medical domain [24,29–31] but also for industrial applications [26]. Here, [32] present a collaborative approach for specifying task-configurations of specific DM methods. Further examples include using ontologies in specific DM tasks (i.e., subgroup discovery and network analysis) [33,34].

However, in those approaches, domain knowledge is only used in a very specific setting, that is, modeling, so is not generalized to the whole DM process. The same observation holds for approaches that stress the importance of contextual knowledge for data mining, for example, [35,36], which applied context-aware approaches to the process.

Several toolkits allow for embedding *declarative knowledge* into the learning process. In [37], the authors proposed a neural network mechanism that allows for the representation of structured knowledge in the form of n-dimensional vectors. This is an approach equivalent to word2vec [38]. In [39], the usage of variational autoencoders for graph structure embedding was presented. Due to the variational nature of an encoder, it not only allows for embedding graphs, but also supports the generation of such. In particular, there have been various approaches in the fields of Semantic Web and Linked Open Data for DM, although their full potential is still to be unlocked [7]. Traditional DM processes still face major challenges in terms of massive data [40]. In addition, the application of data

mining still faces serious challenges, one of which is reproducing already known knowledge. At the same time, DM systems typically make very little use of existing corporate knowledge [41]. Here, existing DM methodologies only provide general directions and directives, while users ultimately require explanations and recommendations on how to effectively perform the steps of the DM methodology. This is currently not provided or enabled by standard DM approaches, cf. [42].

An idea for *declarative data analysis* is presented in [22], which specifically targets declarative problem formulation; however, they do not tackle a specific data mining process, nor augment the specific methods or connect between them explicitly. In the area of constraint programming there have been approaches (e.g., [43,44]) that re-frame a DM method using constraint-based programming. These approaches actually specify what the ML or DM task is about rather than utilizing contextual domain knowledge in a declarative way. So, this mainly relates to the interpretability of the specification of such approaches, not to the ultimate understandability/explainability of the process and/or its outcomes, since the proposed declarative systems mainly transform the declarative specifications in a kind of black-box manner [44].

2.3. Explainability and Interpretability in Data Mining

According to [45,46], the term 'explanation' has been widely investigated in different disciplines. Explanations are in some sense always answers to questions, supporting humans in their decision-making [47]. In particular, explanations are a central component for advanced data mining approaches. This becomes especially relevant when considering complicated black-box models that provide recommendations and predictions in sensitive application contexts like medicine, Industry 4.0 and so forth. Here, nontransparent methods and models make it more difficult to spot errors and can thus lead to biased decisions. For example, this can be based on incorrect training data, or training data that is actually not suitable for application—for example, relating to its contained data quality. In general, nontransparent and non-explainable methods stretch the trust humans have (and should rightfully have) in the respective predictions. Then, the potential competitive advantage through better predictions for humans, for businesses, and for society as a whole comes at the cost of reduced explanatory power—which is specifically problematic for sensitive application areas like those for ubiquitous sensing. This is particularly important in the light of the European Union's new General Data Protection Regulation and the "right to explanation" (providing users the right to obtain an explanation for any algorithmic decisions that were made about them), cf. [48].

Recently, with these developments and more and more complex models, there has been growing interest in the development of so-called *eXplainable AI (XAI)* systems. One of the triggers was the NASA XAI Challenge [49]. From [50], XAI is described as "one that produces details or reasons to make its functioning clear or easy to understand." That paper also outlines the differences between key concepts such as comprehensibility, interpretability, explainability, transparency, and the most important one: understandability (intelligibility).

As the challenges of XAI are mostly related to ML models and their use in the DM process, two main cases are considered: different levels of transparent ML models and post-hoc explainers for black-box ML models. Furthermore, we are interested in the hybrid approaches combining these two, for example, see [51]. For some authors, it is clear that the role of knowledge in the process of using proper ML models, and their use in the DM process, is paramount [52]. Furthermore, it is worth emphasizing that interpretability goes far beyond the model itself, and needs to be considered in the scope of the whole process of designing a system [53].

An overview of so-called interpretable ML techniques can be found in [4]. The interpretation of the results of ML explanation models, such as LIME [54], SHAP [55] or Anchor [56], highly depends on expert knowledge and domain knowledge. However, these frameworks do not provide any means for encoding such knowledge, relying purely on the *manual* examination and interpretation of their results. Many attempts have been made to aid domain experts or data scientists in the interpretation and incorporation of explanation results into the DM process. Most of them focus on the visual presentation of such results. This includes saliency maps for Deep Neural Networks [57], task specific visualisations [58] and more general frameworks [59], which are still narrowed to only one phase of the DM process and hardly use any domain knowledge to enhance explanations nor interpretability of the models. For some examples, in [60], the authors demonstrate how the combination of deep tensor and knowledge graph embedding methods can be used for generating explanations for a model in intrusion detection and genomic medicine. In [61], an approach aiming at predicting and explaining interactions between nodes in a knowledge graph is presented. In [62], an approach for exploiting knowledge graphs for the purpose of explanation is sketched. A medical ontology and temporal domain knowledge was successfully incorporated into the prediction model described in [63], for explaining decisions to the end-user.

In general, nontransparent methods and models make it more difficult to comprehend the decisions of the methods and models in general; also, it becomes very difficult to perform validation and to, for example, spot mistakes, since algorithmic methods can learn "bad habits" from the data. For example, if their training data contains misleading/wrongly classified examples, then it is highly likely that the resulting model incorporates specific biases induced by this training data. This can then simply lead to wrong conclusions and decisions due to, for example, incorrect or biased data capture, insufficient data preprocessing [64,65], and so forth. All these aspects are also particularly important and relevant in a ubiquitous sensing context, since we need to provide and ensure representative training and testing procedures, which can also be supported by the inclusion of semantic information and domain knowledge. In such cases, in general, the inability to provide an explanation as justification is a significant drawback of such methods, limiting acceptance and trust in their adoption in sensitive applications of DM—like those we discuss in the context of ubiquitous sensing. Therefore, interpretability and explainability are crucial for a successful DM process. This is enabled by developing computational methods—to "make sense" of complex information and knowledge processes—in a knowledge augmented DM approach. Here, the explanation has to be pushed through all the steps of the DM process.

There have been several attempts to provide methodological approaches for the evaluation and verification of given explanation results [66,67]. Among many qualitative approaches, there are also those that allow for quantitative evaluation. In [68], measures such as fidelity, consistency and stability were coined, which can be used for a numerical comparison of methods. In [69], the aforementioned measures were used to improve overall explanations. In [70], a measure that allows the capture of the stability or robustness of explanations was introduced. Context in terms of explanations is mostly considered in terms of the similarity of training instances within its vector space, not in the broader context of the domain. In [71], the authors exploit the context of features within a training instance to improve explanations generated with LIME. In [72], the context of an instance that is being explained is generated for the purpose of up-sampling and generating explanations. A more advanced approach was discussed in [73], where an interactive explanation architecture was presented that allows for interactive verification and ad-hoc personalization of the explanations.

The overview provided in this section emphasizes the role of knowledge and explanation in the DM process, and different approaches for introducing them. In the next section, we discuss selected illustrative examples of the applications of semantic data mining in ubiquitous sensing.

3. Applications in Ubiquitous Sensing

We distinguish different sensing areas and contexts that are mostly relevant to semantic DM approaches: environmental sensing, sensing in industrial artificial intelligence and social sensing. Here, we observe large scale and/or complex data, motivating the application of a semantic approach. Below, for data analysis, data mining and machine learning, we survey approaches and methods for the inclusion of domain knowledge in those specific contexts.

3.1. Environmental Sensing

Due to the fact that environmental and industrial (see Section 3.2) data originate from the sensors of multiple manufacturers, use different measurement methods and return data in a variety of measurement units, their use in DM requires appropriate semantization [12]. This is understood here as an appropriate formatting of data, enriching it with tags or labels and combining it with contextual knowledge to create a unified description that can be easily processed by automatic systems (e.g., DM tools) [10].

The most popular markups for sensory data semantization are [10] the following:

- Resource Description Framework (RDF) [74] and the Web Ontology Language (OWL) [75]—two Semantic Web standards developed by the World Wide Web Consortium (W3C). On their basis, many detailed models have been developed to describe a certain type of data or measurement context.
- Sensor Measurement Lists (SenML; in draft versions it was also called Sensor Markup Language) [76]—a standard aimed at small packets with simple sensor measurements that are easy to use in constrained networks, proposed by the Internet Engineering Task Force (IETF).
- Entity Notation (EN) [77]—another standard aimed at providing semantics for low-resource sensors. It provides the definition of short packets, which are transferred via communication links, and complete packets, derived from the short ones, useful for connection with ontologies.
- Observations and Measurements (O&M; https://www.ogc.org/standards/om; accessed on 13 April 2021) and Sensor Model Language (SensorML; https://www.ogc.org/standards/sensorml; accessed on 13 April 2021)—two complementary specifications proposed by the Open Geospatial Consortium (OGC) for observations and sensors description.

They differ not only in their expressivity, but also in their corresponding processing-related energy consumption [78]; this is important, as environmental and physiological sensing often takes place in real-time under resource-constraints in edge computing [79].

What these standards have in common is that they combine sensory data from various Internet of Things (IoT) devices into some kind of (knowledge) graph. The W3C stack of standards is the foundation of the Semantic Web. The SenML and EN notations do not constitute the semantic graph itself, but they can be easily translated into RDF (see [80] and [77], respectively). The O&M and SensorML standards are part of a broader set of services and languages developed by OGC for the Semantic Sensor Web [81]. They can also be integrated with the Semantic Web stack of technologies [82]. As a result, regardless of the notation used to collect and transmit data, measurements can be described using Semantic Web methods, creating a so-called Semantic Web of Things (SWoT) [83,84].

Proper semantization requires not only choosing the right markup, but also the right vocabulary. From both the Semantic Web and Data Mining perspectives, it is important to ensure that multiple datasets use the same set of vocabulary. In the case of weather data processing, this can allow for, for example, an unambiguous statement that both the X value from set S_1 and the Y value from set S_2 represent a measurement of air temperature 2 m above the ground. Appropriate metadata also allow us to determine that the X value is expressed in K and the Y value in °C, which will facilitate their conversion to a common unit.

To address the latter issue, in [85], the authors introduce Custom Datatypes (CDT) [86]—a vocabulary based on The Unified Code for Units of Measure (UCUM; https://ucum.org/; accessed on 14 April 2021) for representing measurements along with their units in data semantized according to the RDF standard. It includes a general type of measurement (`cdt:ucum`) as well as more specific ones (e.g., `cdt:temperature`, `cdt:pressure`). The authors provide a working Java implementation that allows on-the-fly conversions performed during query execution to return results in the desired unit, regardless of the unit in which the measurement was stored (https://ci.mines-stetienne.fr/lindt; accessed on 14 April 2021).

To address the need for a standardized vocabulary for measurement representation, the W3C and OGC joined forces in the Spatial Data on the Web Working Group (for an overview of other approaches to sensory data semantization see [12]), which led to the development of the Semantic Sensor Network (SSN) ontology [87]. Its core concepts were further refined, leading to the creation of a lightweight self-contained Sensor, Observation, Sample and Actuator (SOSA) ontology [88]. From the ubiquitous sensing point of view, the key concept is the Observation. It has a Result of a Procedure performed at a specific time by a Sensor (e.g., thermometer) that observes some object (Observable Property, e.g., the air at the top of the Eiffel Tower) and measures a particular Feature of Interest (e.g., temperature), as summarized in Figure 1. The SSN and SOSA ontologies contain generic high-level terminology and thus can be further refined in more detailed application- and domain-specific ontologies, for example, expanded for the whole IoT area in an IoT-Lite ontology [89] or adapted to describe energy consumption related data [90].
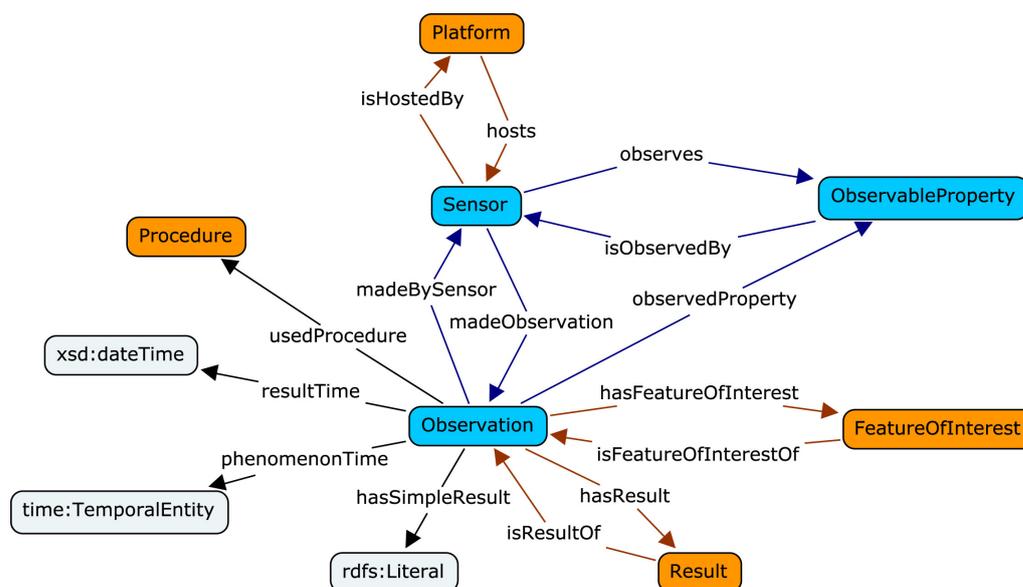

Figure 1. Observation and related concepts in SOSA ontology. Reprinted from [88], with permission from Elsevier.

The practical application and usefulness of data semantization in environmental data preprocessing may be illustrated using weather data. In this scenario, temperature, humidity, or rainfall data may come from multiple heterogeneous sources that need to be properly integrated [91,92]: private sensors connected via networks like Weather Underground (https://www.wunderground.com/; accessed on 14 April 2021), sensors managed by local authorities (e.g., [93]), data provided by (commercial) internet services (e.g., OpenWeatherMap) and data published as (Linked) Open Data.

In order to combine them into a single knowledge graph, it is necessary to first define an appropriate vocabulary and scheme. Then, a dedicated wrapper should be developed for each source to tag the data using the pre-defined scheme. In Listing 1, an example of a semantized data sample adapted from [94] is shown. It uses previously

mentioned vocabularies: SOSA (`sosa:`) to represent observations and CDT (`cdt:`) to describe measured values. The custom namespace (`weather:`) is also used for other pieces of information. Finally, one can see the use of the GeoSPARQL (`geo:`) ontology [95]. The specification of the geographic coordinates according to this standard (see line 15 in Listing 1) makes it possible to easily search the dataset and find particular points, for example, those closest to a given location or all the points in a given area. Listing 2 shows a sample query that extracts all temperature data for the area of Kraków, PL. The filter in line 16 is responsible for selecting points from the indicated area (defined by a set of coordinates). In line 15, it is specified that the result should be given in Kelvin—automatic unit conversion will be done where necessary. Data generated in this way are therefore standardized and ready for further analysis. Depending on the needs, the semantic description can be more detailed, including, for example, sensor specifications and more metadata about the measurement site [94].

Listing 1. An example of a semantized data sample taken from OpenWeatherMap reporting air temperature at Jagiellonian University.

```
1 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
2 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3 @prefix cdt: <http://w3id.org/lindt/custom_datatypes#> .
4 @prefix sosa: <http://www.w3.org/ns/sosa/> .
5 @prefix geo: <http://www.opengis.net/ont/geosparql#> .
6 @prefix weather: <https://geist.re/sdm/weather/> .
7
8 weather:obs_1618483279_19.933107698819033_50.06088286523941_temp
9   a sosa:Observation ;
10  sosa:madeBySensor weather:openweathermap ;
11  sosa:resultTime "2021-04-15T12:41:19+02:00"^^xsd:dateTime ;
12  sosa:hasFeatureOfInterest weather:feat_air ;
13  sosa:observedProperty weather:prop_air_temp ;
14  sosa:hasResult [ a sosa:Result ; weather:value "20.45 Cel"^^cdt:
      temperature ] ;
15  geo:hasGeometry [ a geo:Point; geo:asWKT "POINT(19.933107698819033
      50.06088286523941)"^^geo:wktLiteral ] .
```

Listing 2. A sample query to extract all temperature measurements for the Kraków, PL area.

```
1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX cdt: <http://w3id.org/lindt/custom_datatypes#>
3 PREFIX sosa: <http://www.w3.org/ns/sosa/>
4 PREFIX geo: <http://www.opengis.net/ont/geosparql#>
5 PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
6 PREFIX weather: <https://geist.re/sdm/weather/>
7
8 SELECT ?time ?temperature
9 WHERE {
10  ?observation a sosa:Observation ;
11               sosa:resultTime ?time ;
12               sosa:observedProperty weather:prop_air_temp ;
13               sosa:hasResult ?raw_temp ;
14               geo:hasGeometry ?position
15  BIND(( "0 K"^^cdt:temperature + ?raw_temp ) AS ?temperature)
16  FILTER geof:sfWithin(?position, "POLYGON((19.79120 49.99440, 20.07877
      49.99899, 20.11130 50.12376, 19.80786 50.10799))"^^geo:wktLiteral)
17 }
```

There are state-of-the-art frameworks that facilitate the entire process of sensory data semantization. The Linked Sensor Middleware (LSM) [96] provides both an automatic tagging mechanism by applying custom wrappers to data sources, and a web interface for manual annotation. The whole is complemented by a web service that allows for data extraction using the SPARQL query language. SWoT4CPS (Semantic Web of Things for Cyber-Physical Systems) [84], in turn, provides a more sophisticated ontology, which includes, for example, cause–effect relationships, that can be used not only for data seman-

tization, but also for further processing and mining on a "semantic" layer of the system. For an overview of other sensing semantization frameworks, see [84].

Environmental sensing is not limited to the observation of changing weather conditions. Other applications include, but are not limited to, automatic temperature and humidity control with anomaly detection in smart buildings [84], early detection of collisions between pedestrians, cyclists and drivers to generate timely alerts on mobile devices [97], data collection from in-car sensors to predict generated noise, travel time and fuel consumption [98], and overseeing the food production process in the agri-food sector to reduce CO_2 emission levels and energy consumption levels [99]. The applications discussed are summarised in Table 1.

Table 1. Semantic Web of Things (SWoT) for Environmental Sensing [10,12,83,84].

Ref.	Sensing Framework	Semantic Formalism	Explainability	Domain
[91]	Custom	Simple taxonomy	Visual interface	Smart City
[92]	Custom	Ontology	No	Smart City
[94]	Custom	Ontology (SOSA, SSN)	Visual interface	Smart City
[97]	Custom	Relational database	Visual interface	Traffic
[98]	Custom	Ontology (O&M)	Visual interface	Traffic
[99]	None	*Out of the scope*	Visual interface	Agri-food
[100]	None	*Out of the scope*	No	Emotions
[101]	None	*Out of the scope*	No	Education
[102]	Custom	Relational database	Visual interface	e-Health
[103]	None	*Out of the scope*	No	Fatigue detection
[96]	LSM	Ontology (SSN)	Visual interface	cross-domain
[84]	SWoT4CPS	Ontology (SSN), rules	No	cross-domain

An interesting area of environmental sensing is physiological sensing, which aims to measure various characteristics originating from the body. A diverse set of sensors, including smart bands, smartwatches [102], sensors embedded in phones [104], portable EEGs [103] and smart textiles [105,106], are used to measure temperature, heart action, respiration, galvanic-skin reaction, electroencephalography and other signals. Among the most common applications, there are stress levels and changes in emotions detection [100], assessment of involvement in education [101], health monitoring [102] and various cognitive enhancement tasks, including driver fatigue detection and the assessment of air traffic controllers' mental fatigue [103].

3.2. Sensing in Industrial AI

Environmental sensing is nowadays most extensively developed in the area of Industry 4.0 applications. Industry 4.0 (I4.0) defines an ongoing transformation of traditional business processes by the adaptation of new technologies and automation systems. Although the term originally referred only to manufacturing, currently it can be extended to almost every sector where technology plays an important role. Figure 2 depicts the advanced technology uptake in different sectors as of 2019. This can be considered as the expansion of I4.0 among enterprises, which is growing rapidly in almost all sectors, for example (discrete) manufacturing [107,108], especially in the context of the adaptation of Artificial intelligence as presented in Figure 3.
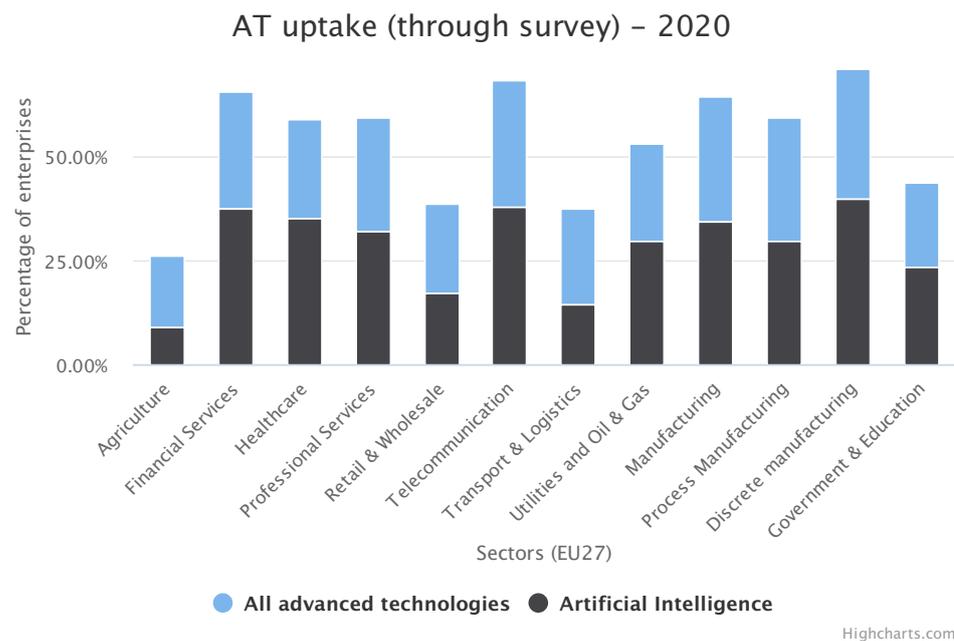
AT uptake (through survey) – 2020

Figure 2. Percentage of advanced technology uptake in Industry 4.0 and AI in this uptake. Generated with https://ati.ec.europa.eu/data-dashboard (accessed on 8 April 2021).

AT uptake (through survey) – Artificial Intelligence
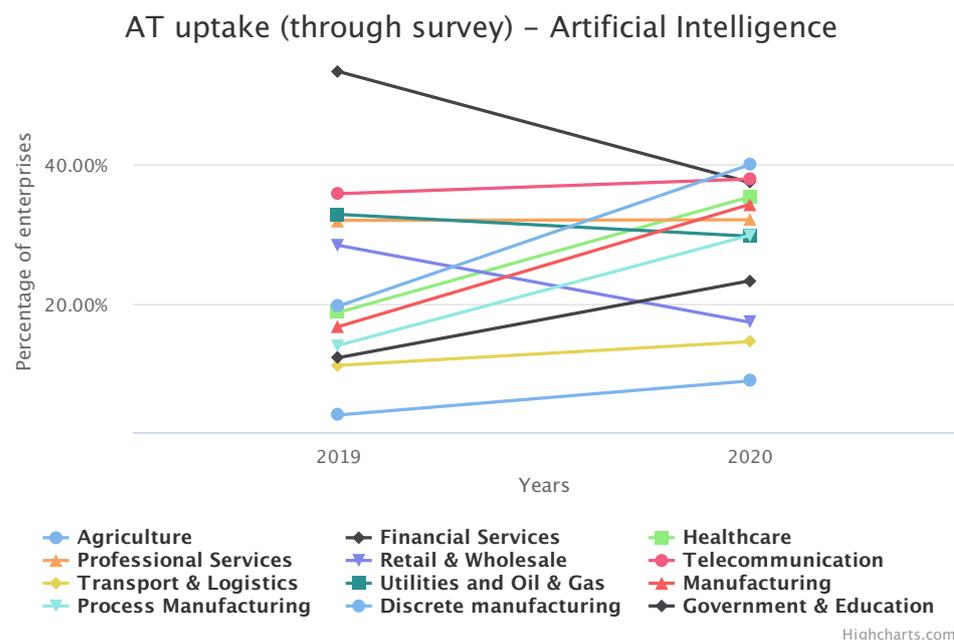
Figure 3. Trend in percentage uptake of advanced technology in 2019 and 2020. Generated with https://ati.ec.europa.eu/data-dashboard (accessed on 8 April 2021).

This allows us to extend the definition of I4.0, after [109], to a complex technological system that embraces a plethora of technologies, the implementation of which allows the development of intelligent manufacturing processes, composed by devices that are able to exchange information, perform actions and control each other. These technologies include, but are not limited to, Cyber–Physical Systems, Internet of Things, Robotics, Big Data, Artificial intelligence and so forth. The adaptation of such technologies in all of the cases is performed on many different levels of abstraction. Such levels can be generalised to three stages of process/data/control flow, as shown in Figure 4.
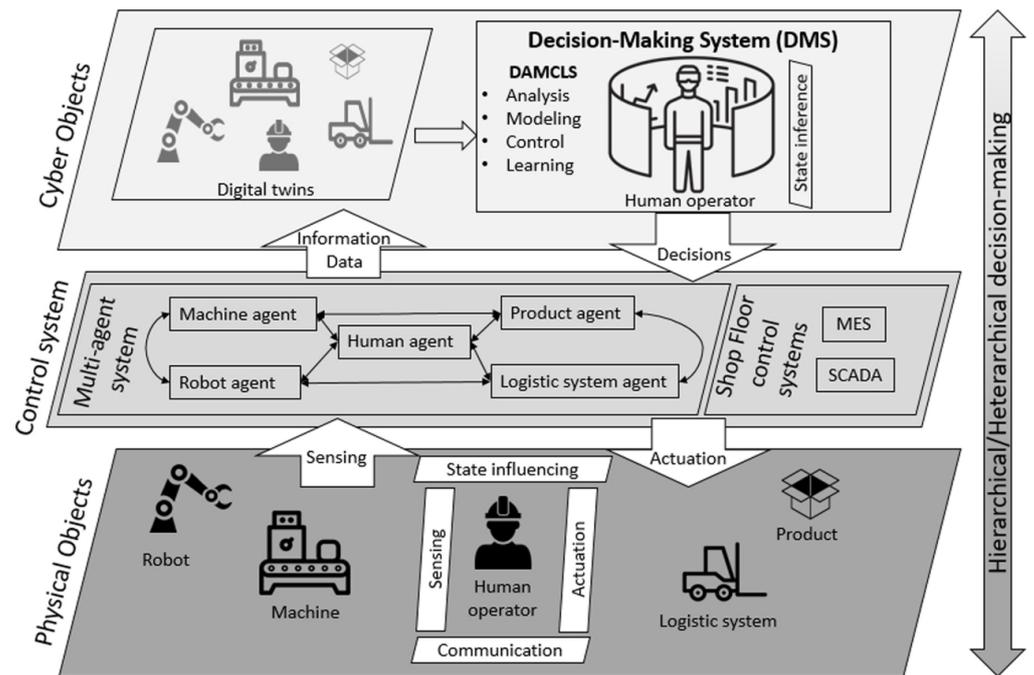
Figure 4. Three levels of the Cyber–Physical system in Industry 4.0 [110].

These three layers represent different levels of interaction of humans with the system, and hence different levels of knowledge and semantics' exploitation by automated algorithms (including data mining and machine learning systems). At the Physical layer, humans directly interact with machines and other equipment. This is usually supported by built-in interfaces and does not require any additional layers to be fully operational. The control layer serves as a middleware between the Physical layer and the Cyber layer. It can be considered as a technical layer for exchanging and storing data from the machines and other system components, but also extends the control over the larger parts of the system, such as SCADA. The Cyber layer mirrors the physical environment that is formed by the concatenation of the Control system and Physical objects. This layer is mostly responsible for tasks related to the analysis of the data, modelling, learning, decision support, predictive maintenance and other high-level tasks [111].

Despite the automation of the system, humans still play an important role in each of the aforementioned levels and are mostly present in Physical and Cyber layers presented in Figure 4. Such a system, where a human operator is actively involved in the automated Cyber–Physical process is called a Human Cyber–Physical System [112]. In this work, we focus on the level where data acquisition, processing and utilization takes place; therefore, we reach in our discussion the moment where such systems are designed and built up to the deployment phase. Due to the broad nature of the problem, we narrowed the discussion to the areas in the process that are data-driven and, yet, the interaction with humans exists. This coexistence of the semantic world and the data driven world is especially difficult to embrace in the industrial process. Due to the fact that the knowledge that is possessed by human operators is rich in semantics but most often lacks formalism, the integration of such knowledge into the system is a challenging task.

This coexistence is achieved with many different methods, depicted in Figure 5. In the following sections, a description of particular applications of selected methods in the area of I4.0 will be presented. They are also summarized in Table 2.
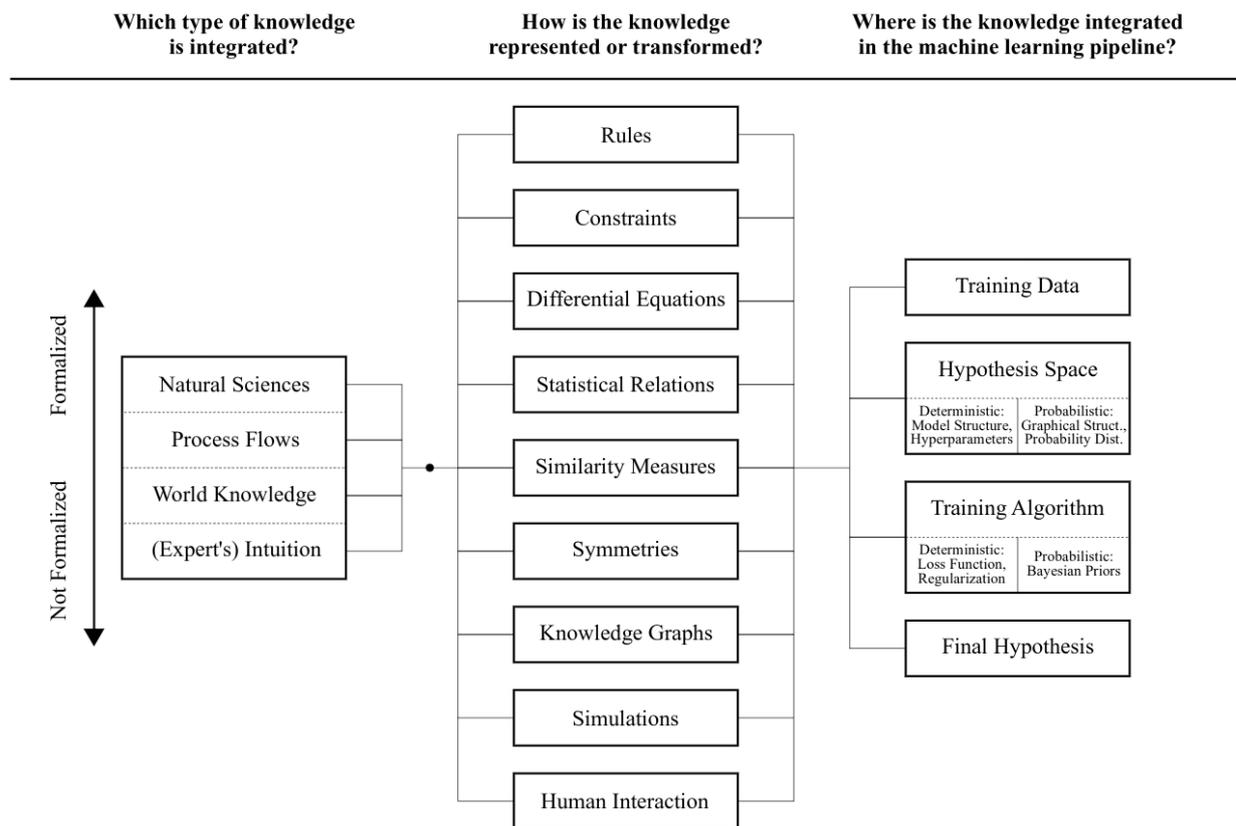
Figure 5. Knowledge source, its formalization and application to different ML/DM stages [8].

Table 2. Sensing platforms in Industrial AI.

Reference	Sensing Framework	Semantic Formalism	Explainability	Domain
[113]	SANSA stack	Semantic Web	Visual interface	domain-specific (electronic mounting)
[114,115]	I40KG framework	Ontologies	No	cross-domain
[116]	SWeTI framework	Semantic Web	No	cross-domain
[117–119]	None	Physics equations	No	domain-specific
[120]	None	Physics approximation model	No	domain-specific
[121–123]	Custom	Constraints	visualization dashboard, knowledge mediation	cross-domain
[124–126]	Custom	knowledge graph	No	cross-domain
[127–130]	None	knowledge graph	knowledge-graph extensions	cross-domain
[131,132]	None	Rules	Shapely values	domain-specific
[133–135]	None	knowledge graph	visual, symbolic, statistical	cross-domain

3.2.1. Formalization of Semantics for Industrial AI Sensing

In Figure 5, different levels of formalization of knowledge are presented. Depending on the level of formalization, different methods for the knowledge transfer to machine learning pipeline can be chosen. In this section, we present how such knowledge can be used in practical applications. We will focus mostly on ontologies as one of the most expressive mechanisms for encoding and processing domain knowledge in many fields.

In [113], the integration of Semantic Web techniques in a large Industry 4.0 context was presented. The authors deployed the SANSA Stack to enable uniform access to Surface-Mount Technology (SMT) data. An ergonomic visual user interface was proposed to help non-technical users coping with the various concepts underlying the process and conveniently interacting with the data. In [114], the authors propose a framework for constructing a semantically annotated knowledge graph for Industry 4.0 related standards called I40KG. The graph provides a Linked Data-conform collection of annotated, classified reference guidelines, supporting newcomers and experts alike in understanding how to implement Industry 4.0 systems. The authors illustrate the suitability of the graph for various use cases, its already existing applications, and present the maintenance process

and evaluate its quality. In [115], the authors provide a practical example and evaluation of the Bosh implementation of I40KG. In [116], the authors present a Semantic Web of Things for Industry 4.0 (SWeTI) platform, which is a cross domain, cross platform solution. In their approach, they extend the hierarchy of I4.0 layers by adding on top of the cyber layer the data analytic and application layers, which are supposed to be the main components for knowledge exchange. It serves as an intelligent interface with all of the knowledge hidden within the bottom layers. As this is only a framework, it does not provide any specific ontology but allows integration with domain knowledge formalized in such a way at the level of the cyber layer.

The remaining components for knowledge transformation and encoding are presented in Figure 5 and do not require complex frameworks such as knowledge graphs; however, they may lack the expressive power of the former in many cases. In the following section, we discuss the application of selected knowledge embedding methods to different phases in the ML/DM pipeline.

3.2.2. Knowledge Embedding Methods

Knowledge embedding within the ML/DM pipeline can be conducted at many different stages, as depicted in Figure 5. Depending on the stage of the ML/DM workflow, different methods are more applicable. This natural consequence yields serious difficulties in communication and knowledge transfer between ML/DM phases. Below, we discuss methods that allow the integration of specific knowledge into the ML/DM pipeline.

Simulations, equations, statistical relations similarity measures or symmetries are mostly used in the first two phases of the DM/ML process. They usually do not use any formal knowledge representation that is utilised by data mining algorithms, as they are implemented within the pipeline or algorithm by data analysts and experts. In [117], the authors present a differentiable physics engine that can be integrated as a module into a deep neural network as a layer in its architecture to improve overall performance. Similar exploitation of physics equations is presented in [118], where the authors embed Gaussian process regression with stochastic equations that model the well-defined physics of the power grid dynamics. In [120], the authors approach the problem from a different perspective and embed domain knowledge that can be generated with simulations within a surrogate neural network that produces the approximation of simulations, which can later be used at further stages of the ML/DM pipeline.

In [121], the authors present a framework for ML/DM tasks configuration, which is based on the profiling that is generated through interactions with users. The knowledge representation and inference was not specified, while more emphasis was placed on the aspect of knowledge mediation involving the user/data scientist. They extended their work in [122] by showing how the profiles and user needs and expectations can be modelled and used in the ML/DM pipeline, for example, in a form of constraints. The same knowledge representation technique was presented in [123], where constraints were applied to the output space of classification neural network architecture to minimize the need to label the data, which might be costly. In [119], the authors approached a similar problem of using machine learning models and statistical learning on datasets that are relatively small. They demonstrate how embedding domain knowledge for the machine learning of complex material systems can improve its performance in the case of small datasets.

A comprehensive review on usage knowledge graphs in Industry 4.0 was presented in [136]. In [124,125], the authors present an approach for knowledge fusion in manufacturing operations with the use of knowledge graph embedding methods. Such knowledge can later be used by an arbitrary ML model for further training and decision making. A similar approach was presented in [126], where the system for predicting the geographic centers of fuel cells is enhanced with knowledge gathered form heterogeneous sources and unified in a form of RDF-based knowledge graph. The knowledge is used to generate training data for machine learning models that implement predictive maintenance tasks.

Human interactions are considered informal knowledge transfer methods that can be used to infer new knowledge by an algorithm, or to translate them into more formalized queries that can be processed by a machine. Such systems are usually built on top of one of the aforementioned formalisms. A comprehensive review of methods that allow the combination of virtual reality and augmented reality visualization techniques with machine learning and knowledge graphs was provided in [137].

3.2.3. Decision Explanation Methods

An emerging field of science in recent years is eXplainable Artificial Intelligence (XAI). As stated in Section 2.3, one of the most important roles of these methods is to reverse the process of knowledge transfer. While the majority of the methods described in the previous section focus on incorporating domain knowledge into the ML/DM tasks, the goal of XAI methods is to translate the decision making process performed by an ML algorithm in a way that can be understood by the data scientist or domain expert. This can either be accomplished by augmenting and/or contextualizing the provided decision with additional information, such as background knowledge or similar/historic applied cases, or with domain knowledge providing an extended context, for example including declarative knowledge [138]. Ultimately, this bridges the gap between knowledge discovery and data mining to decision support, and its contextualization, ultimately enabling computational sensemaking, for example, see [138].

In many cases, this requires the embedding of the raw decision with additional semantics that will be aligned with the expectations of the addressee. In many task-specific solutions, such explanations are delivered partially with the use of formalism for the ML model, for example, knowledge graphs. In [127], the utilisation of the knowledge graph was expanded. It not only serves as an input for the ML model, but can also be extended by statistical learning methods, enriching knowledge about the domain or ML decisions. Similar approaches for application in the Cyber–Physical system were also provided in [128–130]. In [131,132], the authors provide a method that aims to bring more semantics to clusters discovered by automated methods in an Industry 4.0 setting. Such semantic information can then be represented, for example in the form of rules, and can be used to extend the knowledge about the machinery states in the Cyber–Physical setting. In addition, [139] presented a method for visualizing interesting parts of the decision space of a model in order to make the respective modeling and ultimate decisions interpretable for humans. Further directions include the application of interpretable methods for obtaining explanatory patterns—for contextualization and explanation, for example, see [133–135].

3.3. Social Sensing

Social sensing [140–143] relates to observing human interactions and capturing the respective individual and/or collective behaviors by way of sensors, which can, for example, relate to both offline as well as online sensors, for example, [144–151] In the following, we first outline the general area of social sensing, before discussing two specific subareas: First, we consider *semantic social sensing* relating to the use of strongly formalized knowledge structures for integrating semantic information, such as ontologies, into the respective mining and analysis approaches; second, we discuss *semantic social network analysis*, focusing on the analysis methodology of social network analysis applied to rich social sensing data. Here, we also revisit semantics given light-weight knowledge structures, that is, collaborative tagging leading to folksonomies for integrating (semantic) information. Table 3 summarizes the discussed semantic social sensing and social network-based frameworks/platforms.

Table 3. Semantic Social Sensing and Social Network-Based Frameworks/Platforms.

Ref.	Sensing Framework	Semantic Formalism	Domain/Sensing
[91]	Custom	Ontology	Smart City
[152,153]	Custom	Ontology	Smart City
[154]	Custom	Ontology (SOSA, SSN)	IoT/Heterogeneous Sensors
[155]	Custom	Ontology	IoT/Heterogeneous Sensors
[156,157]	None	Ontology	Social Networks
[150]	Custom	Folksonomy-Based	Social Network/Human Sensors
[151]	Custom	Folksonomy-Based	Social/Human Sensors/IoT
[144,158]	Custom	Folksonomy-Based	Social/Human Sensors
[159]	None	Network-Based	Social/Textual/User-Generated Content
[160]	Custom	Ontology	Healthcare

3.3.1. Social Sensing in Ubiquitous and Social Environments

For social sensing in ubiquitous and social environments, a variety of heterogeneous sensor data can be observed and analyzed, for example, considering specific sensors, social media and the ubiquitous social web, and so forth, cf. [147,161–163]. This specifically relates to observing human interactions, that is, social and physical activities [140,144, 145,148,151,164]. While [161,162], for example, describe social sensing on the ubiquitous social web, [165] discusses social sensing in the context of social media and human face-to-face interactions, using the OpenBeacon badges of the SocioPatterns consortium [144]. Other prominent sensors for social sensing in similar contexts include the Sociometric badges [145] and successors such as the Rhythm badge [164]. Besides data mining on social interaction networks [166], social sensing is also relevant for (computational) social science and digital epidemiology [149,167], as well as for applications in human sensing in industrial contexts, such as manufacturing [168].

Regarding the issues of data modeling, mining and analysis in social sensing, we follow the presentation in [166] and focus on *social interaction networks* [169–172]. These enable a wide range of modeling and analysis options, cf. [166,171,173], that is, user-related social networks capturing social relations inherent in social interactions, social activities and other social phenomena which either directly connect users or act as proxies for social user relatedness. This then also includes interaction data from sensors and mobile devices, as long as the data are created by real users. In this way, social sensing transcends offline scenarios using hardware sensors (only) such as in mobile and ubiquitous computing, and can be linked to virtual sensor data as well—that is, data captured in online contexts. For social sensing in such broadly defined contexts, we can thus consider, for example, users who connect their mobile phones via Bluetooth, interact similarly with online applications such as Flickr, communicate in a similar way (or about similar topics) on Twitter or Facebook, or explicitly establish "contacts" within certain social applications, for example, [173,174]. Furthermore, we consider real-world contacts as determined by other ubiquitous computing applications [175,176], the ubiquitous web [161,177,178], and the principle of object-centric sociality [179], where objects of a specific actor—for example, resources—mediate connections to other actors.

3.3.2. Semantic Social Sensing

Depending on the respective types of social sensing data, in particular when different representations are integrated in a multi-modal strategy, different representations can be derived. This is possible, for example, when including sequential/time series sensor data together with unstructured information from web pages and structured information from ontologies. For example, [180] proposes such an approach called *semantic social sensing*, making use of ontologies and semantic augmentation together with textual analysis on user generated information such as comments. Such data representations are necessary in order to prepare the analysis, drive explanations, or guide exploratory approaches. Tabular (structured) data can usually be mapped and normalized in a straight-forward way, whereas unstructured (e.g., text) and semi-structured data need further processing. Here, informa-

tion first needs to be extracted or abstracted. In this respect, including background and/or domain knowledge in such approaches becomes more and more important; to improve the models, to drive explanations and, ultimately, to allow *computational sensemaking*.

Regarding formalized knowledge, providing this in the context of sensing with respect to smart cities—ontological requirements and useful semantic information as an ontological representation of urban data—has been discussed in [152,153]. Furthermore, [154] presents the modeling of smart sensors on top of the SOSA/SSN ontologies [87,88] (as already introduced above), also with the semantic smart sensor network (S3N) modular ontology. In [181], the authors present semantic social sensing applications in the form of a semantic sensing middleware for the Web of Things. In [155], the authors introduce an ontology for hybrid semantic sensor networks (HSSN) which extends the Semantic Sensor Network (SSN) ontology described above for, for example, more heterogeneous sensors and platforms in order to enable extended analysis and mining. Furthermore, [156,157,182] specifically discuss semantic social network analysis for modeling social interaction networks in order to create richer (semantic) models given the social sensing data, which can also be used for detailed analysis. Basically, a network between humans, ontologies and their interlinks is provided and analyzed, also taking semantic information on the respective graph structure into account.

Applications of such approaches include the analysis of organizational social networks [159] or data mining for recommender systems, for example, see [183]. Specifically, in [184], a data mining method for generating recommendations in the context of software development is provided. Given social sensing data, as well as additional domain knowledge and information extracted from CVS logs, the performance of the recommendations could be significantly improved, combining both sources of information and knowledge.

Altogether, using semantic social sensing, for example, unusual social activities can be detected using geo-tagged microblogs [185], or human activity and interesting patterns can be analyzed using ubiquitous social data in social [186,187] or urban contexts [188,189]. Here, as mentioned in [142], it is particularly important to include rich information, such as place and space semantics, about the respective social and spatial interactions. Semantic signatures/labels can also be applied [189]; example applications include the health care domain considering the detection/monitoring activities of daily living (ADL) [160,190–192]. Here, rule-based approaches are also relevant [193,194].

3.3.3. Semantic Social Network Analysis

Overall, the analysis of online social network data has received significant attention for analyzing large and complex systems, such as large-scale social network systems or the internet infrastructure, and so forth cf. [195–197]. While there has been foundational work on social network analysis and mining on social sensing data, such as the analysis of face-to-face contact networks, for example in [144], semantic data mining on those networks is still a rather new field of research. When integrating semantic information into social network analysis approaches, this leads to *semantic social network analysis* [198].

Regarding the semantic information to be included in data mining, we distinguish different types, for example, formalized in ontologies, taxonomies, or folksonomies. Semantic structures emerge from collaborative tagging, which can be used at the level of the respective tagged objects, such as images or locations on a map. These structures are called folksonsomies [199,200]. In the scope of social sensing, a folksonomy is also called a *sensonomy* [201,202], for example, for urban sensing. Then, we can apply this on the level of maps and data mining to the interactions, spatial structures and so forth. An example is the analysis of social interaction networks integrating multi-modal sensing information, for example from the WideNoise and Airprobe systems, implemented using the Ubicon system [151,203], for observing and analyzing social and physical activities. Here, semantics are applied in the form of subjective information, in addition to folksonomies, as collaborative tag vocabularies.

4. Summary, Challenges and Future Directions

Today's sensing technology often utilizes diverse hardware sensors that generate huge volumes of heterogeneous data—Big Data. Therefore, the processing of these data is mostly performed by the use of data mining methods and tools. However, proper interpretation of these data often requires the inclusion of certain knowledge regarding the operational context of data acquisition or requirements of a specific domain in which the sensory system is used. The objective of this survey was to provide a concise overview of a range of approaches that aim to extend the typical data mining process with the use of semantic information, the introduction of knowledge and declarative representations. We refer to these approaches using the general term *semantic data mining*. We also emphasized the role of interpretability and explainability in data mining, which can be achieved with the use of semantic data mining.

Furthermore, we discussed selected applications of data mining for ubiquitous sensing, where—in our opinion—semantic interpretation of data can be particularly useful. We selected three specific areas of interest for which we described a number of approaches categorized based on the sensing framework used, semantic formalism, explainability and domain. The first area is environmental sensing, for which we mentioned standardized protocols and discussed their use for weather data and physiological sensing. The second important area is industrial AI, for which we considered how different knowledge sources can be formalized and applied for data mining. The third area is social sensing, including semantic social sensing, and semantic social network analysis.

It is worth noting that the adoption of the semantic data mining approaches we discussed may face different *challenges*, which we identified—from the surveyed papers, as well as from our discussion. Thereby, we deduce the following challenges, which we present in order to inspire future research in these areas:

1. The first challenge is related to the availability of domain knowledge, its form and representation. Semantic data mining approaches differ with respect to the knowledge representation used, for example, from simple annotations to formalized knowledge models. This selection also has an impact on the possible cognitive load of human experts participating in the knowledge acquisition process. Furthermore, in certain domains, formalized knowledge is in fact present in the form of rules, constraints, structures and vocabularies. The introduction of such knowledge into the DM process—if successful—can allow for the alignment of the results of the process with the domain requirements.

2. The second important challenge is the proper selection of the phase of the DM where the knowledge is introduced. As we discussed, it is often the case that preliminary stages of the process are very time consuming, so a proper understanding of the data can be achieved. This is why the use of domain knowledge in this stage could be beneficial, for example, as a part of the feature engineering activity. However, in practice, such an approach—while possible—is often overlooked.

3. The third challenge is related to the provision of explainability methods. The use of complex black-box machine learning models that offer superior accuracy can result in certain risks in terms of their interpretability. The need to formulate explanations instrumental for understanding the results of the DM process and for putting it in the context of specific domains, is an important requirement. As such, the use of semantic data mining methods can be of particular interest and value for interpretability and explainability, as we have also discussed throughout the methods and application sections.

To summarize, we expect in the near future a growing interest in semantic data mining approaches, especially in ubiquitous sensing. Specific future directions of the introduction of semantic methods in data mining could include, for example: semantic interoperability in sensing for limiting data pre-processing; semi-automated methods for data annotation in the early phases of the data mining process; domain knowledge modelling during

data acquisition (possibly as part of the feature engineering); emphasis on explanations regarding both the data and the output of the data mining process, and so forth.

In addition, since both symbolic and sub-symbolic data mining and machine learning methods can be applied in the context of ubiquitous sensing, domain knowledge—enabling semantic data mining—can be added as a third dimension in order to allow for a fruitful combination of those different methods, and also potentially to serve as a strong promoter of their combination and application in sensing scenarios. With the above mentioned observations in mind, we envisage the increasing adoption of semantic data mining in research as well as in the wide range of fields of application.

References

1. Cuzzocrea, A.; Song, I.Y.; Davis, K.C. Analytics Over Large-Scale Multidimensional Data: The Big Data Revolution! In Proceedings of the 14th International Workshop on Data Warehousing and OLAP at 20th International Conference on Information and Knowledge Management, CIKM 2011, Glasgow, UK, 24–28 October 2011; ACM: New York, NY, USA, 2011; pp. 101–104. [CrossRef]
2. Klein, D.; Tran-Gia, P.; Hartmann, M. Big Data. *Inform. Spektrum* **2013**, *36*, 319–323. [CrossRef]
3. Luengo, J.; García-Gil, D.; Ramírez-Gallego, S.; García, S.; Herrera, F. Smart Data. In *Big Data Preprocessing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 45–51.
4. Molnar, C. Interpretable Machine Learning. Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 28 Feburary 2020).
5. Dou, D.; Wang, H.; Liu, H. Semantic Data mining: A Survey of Ontology-Based Approaches. In Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015), Anaheim, CA, USA, 7–9 Feburary 2015; pp. 244–251.
6. Sirichanya, C.; Kraisak, K. Semantic Data Mining in the Information Age: A Systematic Review. *Int. J. Intell. Syst.* **2021**. [CrossRef]
7. Ristoski, P.; Paulheim, H. Semantic Web in Data Mining and Knowledge Discovery: A Comprehensive Survey. *Web Semant.* **2016**, *36*, 1–22. [CrossRef]
8. von Rueden, L.; Mayer, S.; Beckh, K.; Georgiev, B.; Giesselbach, S.; Heese, R.; Kirsch, B.; Pfrommer, J.; Pick, A.; Ramamurthy, R.; et al. Informed Machine Learning—A Taxonomy and Survey of Integrating Knowledge Into Learning Systems. *arXiv* **2020**, arXiv:1903.12394.
9. Kim, J.; Wang, G.; Bae, S.T. A Survey of Big Data Technologies and How Semantic Computing Can Help. *Int. J. Semant. Comput.* **2014**, *8*, 99–117. [CrossRef]
10. Shi, F.; Li, Q.; Zhu, T.; Ning, H. A Survey of Data Semantization in Internet of Things. *Sensors* **2018**, *18*, 313. [CrossRef] [PubMed]
11. Kacfah Emani, C.; Cullot, N.; Nicolle, C. Understandable Big Data: A survey. *Comput. Sci. Rev.* **2015**, *17*, 70–81. [CrossRef]
12. Rhayem, A.; Mhiri, M.B.A.; Gargouri, F. Semantic Web Technologies for the Internet of Things: Systematic Literature Review. *Internet Things* **2020**, *11*, 100206. [CrossRef]
13. Fayyad, U.M. Data Mining and Knowledge Discovery: Making Sense Out of Data. *IEEE Expert* **1996**, *11*, 20–25. [CrossRef]
14. Kurgan, L.A.; Musilek, P. A Survey of Knowledge Discovery and Data Mining Process Models. *Knowl. Eng. Rev.* **2006**, *21*, 1–24. [CrossRef]
15. Zhang, S.; Zhang, C.; Yang, Q. Data Preparation for Data Mining. *Appl. Artif. Intell.* **2003**, *17*, 375–381. [CrossRef]
16. Morik, K.; Scholz, M. The Miningmart Approach to Knowledge Discovery in Databases. In *Intelligent Technologies for Information Analysis*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 47–65.
17. SAS Institute. *SAS Enterprise Miner: Introduction to SEMMA*; SAS Institute: Cary, NC, USA, 2017.
18. Azevedo, A.; Santos, M. KDD, SEMMA and CRISP-DM: A Parallel Overview. In Proceedings of the IADIS European Conference on Data Mining , Amsterdam, The Netherlands, 22–27 July 2008.
19. IBM. *Have You Seen ASUM-DM*; IBM: Armonk, New York, NY, USA, 2015.

20. Biecek, P. Model Development Process. *arXiv* **2019**, arXiv:1907.04461.
21. Blockeel, H. Data Mining: From Procedural to Declarative Approaches. *New Gener. Comput.* **2015**, *33*, 115–135. [CrossRef]
22. Blockeel, H. Declarative Data Analysis. *Int. J. Data Sci. Anal.* **2017**, *6*, 217–223. [CrossRef]
23. Atzmueller, M.; Puppe, F.; Buscher, H.P. Exploiting Background Knowledge for Knowledge-Intensive Subgroup Discovery. In Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05), Edinburgh, Scotland, 30 July–5 August 2005; pp. 647–652.
24. Atzmueller, M.; Seipel, D. Declarative Specification of Ontological Domain Knowledge for Descriptive Data Mining (Extended Version). In 17th International Conference on Applications of Declarative Programming and Knowledge Management (INAP 2007) and 21st Workshop on (Constraint), Würzburg, Germany, 4–6 October 2007; Spriner: Berlin/Heidelberg, Germany, 2008.
25. Svátek, V.; Rauch, J.; Ralbovský, M. Ontology-Enhanced Association Mining. In *Semantics, Web and Mining*; Spriner: Berlin/Heidelberg, Germany, 2005; Volume 4289, pp. 163–179.
26. Atzmueller, M.; Sternberg, E. Mixed-Initiative Feature Engineering Using Knowledge Graphs. In Proceedings of the 9th International Conference on Knowledge Capture (K-Cap), Austin, TX, USA, 4–6 December 2017; ACM Press: New York, NY, USA, 2017.
27. Ławrynowicz, A. Semantic Data Mining—An Ontology-Based Approach. In *Studies on the Semantic Web*; IOS Press: Amsterdam, The Netherlands, 2017; Volume 29. [CrossRef]
28. Nalepa, G.J. Modeling with Rules Using Semantic Knowledge Engineering. In *Intelligent Systems Reference Library*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 130. [CrossRef]
29. Cespivova, H.; Rauch, J.; Svatek, V.; Kejkula, M. Roles of Medical Ontology in Association Mining CRISP-DM Cycle. In Proceedings of the ECML/PKDD 2004 Workshop on Knowledge Discovery and Ontologies, Pisa, Italy, 20–24 September 2004.
30. Kuo, Y.T.; Lonie, A.; Sonenberg, L.; Paizis, K. Domain Ontology Driven Data Mining: A Medical Case Study. In Proceedings of the DDDM '07: 2007 International Workshop on Domain Driven Data Mining, San Jose, CA, USA, 12 August 2007; ACM: New York, NY, USA; pp. 11–17. [CrossRef]
31. Sikora, M.; Wróbel, Ł.; Gudyś, A. GuideR: A guided separate-and-conquer rule learning in classification, regression, and survival settings. *Knowl. Based Syst.* **2019**, *173*, 1–14. [CrossRef]
32. Atzmueller, M.; Lemmerich, F.; Reutelshoefer, J.; Puppe, F. Wiki-Enabled Semantic Data Mining—Task Design, Evaluation and Refinement. In Proceedings of the 2nd International Workshop on Design, Evaluation and Refinement of Intelligent Systems (DERIS2009), Kraków, Poland, 28 November 2009; Volume 545.
33. Lavrač, N.; Vavpetič, A.; Soldatova, L.; Trajkovski, I.; Novak, P.K. Using Ontologies in Semantic Data Mining With Segs and G-Segs. In Proceedings of the International Conference on Discovery Science, Espoo, Finland, 5–7 October 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 165–178.
34. Kralj, J.; Robnik-Sikonja, M.; Lavrac, N. NetSDM: Semantic Data Mining With Network Analysis. *J. Mach. Learn. Res.* **2019**, *20*, 1–50.
35. Singh, S.; Vajirkar, P.; Lee, Y. Context-Based Data Mining Using Ontologies. In Proceedings of the International Conference on Conceptual Modeling, Chicago, IL, USA, 13–16 October 2003; Springer: Berlin/Heidelberg, Germany, 2003; pp. 405–418.
36. Matei, O.; Rusu, T.; Bozga, A.; Pop-Sitar, P.; Anton, C. Context-Aware Data Mining: Embedding External Data Sources in a Machine Learning Process. In Proceedings of the International Conference on Hybrid Artificial Intelligence Systems, Oviedo, Spain, 20–22 June 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 415–426.
37. Narayanan, A.; Chandramohan, M.; Venkatesan, R.; Chen, L.; Liu, Y.; Jaiswal, S. Graph2vec: Learning Distributed Representations of Graphs. *arXiv* **2017**, arXiv:1707.05005.
38. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781.
39. Kipf, T.N.; Welling, M. Variational Graph Auto-Encoders. *arXiv* **2016**, arXiv:1611.07308.
40. Grady, N.W. KDD Meets Big Data. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 5–8 December 2016; pp. 1603–1608.
41. Charest, M.; Delisle, S.; Cervantes, O.; Shen, Y. Bridging the Gap Between Data Mining and Decision Support: A Case-Based Reasoning and Ontology Approach. *Intell. Data Anal.* **2008**, *12*, 211–236. [CrossRef]
42. Charest, M.; Delisle, S.; Cervantes, O.; Shen, Y. Intelligent Data Mining Assistance via CBR and Ontologies. In Proceedings of the 17th International Workshop on Database and Expert Systems Applications (DEXA'06), Krakow, Poland, 4–8 September 2006; pp. 593–597.
43. Raedt, L.D. Declarative Modeling for Machine Learning and Data Mining. In Proceedings of the Formal Concept Analysis—10th International Conference ICFCA, Leuven, Belgium, 7–10 May 2012; Domenach, F., Ignatov, D.I., Poelmans, J., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7278, p. 2-2. [CrossRef]
44. Guns, T.; Dries, A.; Nijssen, S.; Tack, G.; De Raedt, L. MiningZinc: A Declarative Framework for Constraint-Based Mining. *Artif. Intell.* **2017**, *244*, 6–29. [CrossRef]
45. Roth-Berghofer, T.R.; Richter, M.M. On Explanation. *Künstl. Intell.* **2008**, *22*, 5–7.
46. Atzmueller, M.; Roth-Berghofer, T. The Mining and Analysis Continuum of Explaining Uncovered. In Proceedings of the 30th SGAI International Conference on Artificial Intelligence (AI-2010), Cambridge, UK, 14–16 December 2010.

47. Schank, R.C. Explanation: A First Pass. In *Experience, Memory, and Reasoning*; Kolodner, J.L., Riesbeck, C.K., Eds.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1986; pp. 139–165.

48. Goodman, B.; Flaxman, S. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". *AI Mag.* **2017**, *38*, 50–57. [CrossRef]

49. Gunning, D. *Explainable Artificial Intelligence (XAI).* Defense Advanced Research Projects Agency (DARPA): Arlington, VA, USA, 2017; Volume 2.

50. Barredo Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]

51. Donadello, I.; Serafini, L.; d'Avila Garcez, A.S. Logic Tensor Networks for Semantic Image Interpretation. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 1596–1602. [CrossRef]

52. Rudin, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [CrossRef]

53. Costabello, L.; Lécué, F.; Giannotti, F.; Guidotti, R.; Minervini, P.; Hitzler, P.; Sarker, K. AAAI 2019 Tutorial on on Explainable AI: From Theory to Motivation, Applications and Limitations. In Proceedings of the 33rd AAAI Conference on Artificial Intelligence Technical report, AAAI, Honolulu, HI, USA, 27 January–1 February 2019.

54. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why Should I Trust You: Explaining the Predictions of Any Classifier. In Proceedings of the ACM SIGKDD, San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA, 2016; pp. 1135–1144.

55. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.

56. Ribeiro, M.T.; Singh, S.; Guestrin, C. Anchors: High-Precision Model-Agnostic Explanations. In Proceedings of the AAAI Publications, Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.

57. Selvaraju, R.R.; Das, A.; Vedantam, R.; Cogswell, M.; Parikh, D.; Batra, D. Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization. *arXiv* **2016**, arXiv:1610.02391.

58. Mujkanovic, F.; Doskoc, V.; Schirneck, M.; Schäfer, P.; Friedrich, T. timeXplain—A Framework for Explaining the Predictions of Time Series Classifiers. *arXiv* **2020**, arXiv:2007.07606.

59. Collaris, D.; van Wijk, J.J. ExplainExplore: Visual Exploration of Machine Learning Explanations. In Proceedings of the 2020 IEEE Pacific Visualization Symposium (PacificVis), Tianjin, China, 3–5 June 2020; pp. 26–35. [CrossRef]

60. Fuji, M.; Morita, H.; Goto, K.; Maruhashi, K.; Anai, H.; Igata, N. Explainable AI Through Combination of Deep Tensor and Knowledge Graph. *Fujitsu Sci. Tech. J.* **2019**, *55*, 58–64.

61. Zhang, W.; Paudel, B.; Zhang, W.; Bernstein, A.; Chen, H. Interaction Embeddings for Prediction and Explanation in Knowledge Graphs. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, VIC, Australia, 11–15 February 2019.

62. Lecue, F. On the Role of Knowledge Graphs in Explainable AI. *Semant. Web* **2019**, *11*, 1–11. [CrossRef]

63. Panigutti, C.; Perotti, A.; Pedreschi, D. Doctor XAI: An Ontology-Based Approach to Black-Box Sequential Data Classification Explanations. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT), Barcelona, Spain, 27–30 January 2020; ACM: New York, NY, USA, 2020; pp. 629–639.

64. Schmidt, A.; Atzmueller, M.; Hollender, M. Data Preparation for Big Data Analytics: Methods & Experiences. In *Enterprise Big Data Engineering, Analytics, and Management*; IGI Global: Hershey, PA, USA, 2016.

65. García, S.; Ramírez-Gallego, S.; Luengo, J.; Benítez, J.M.; Herrera, F. Big Data Preprocessing: Methods and Prospects. *Big Data Anal.* **2016**, *1*, 1–22. [CrossRef]

66. Mohseni, S.; Zarei, N.; Ragan, E.D. A Survey of Evaluation Methods and Measures for Interpretable Machine Learning. *arXiv* **2018**, arXiv:1811.11839.

67. Sokol, K.; Flach, P.A. Explainability Fact Sheets: A Framework for Systematic Assessment of Explainable Approaches. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 27–30 January 2020.

68. Robnik-Šikonja, M.; Bohanec, M. Perturbation-Based Explanations of Prediction Models. In *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent*; Zhou, J., Chen, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 159–175. [CrossRef]

69. Yeh, C.; Hsieh, C.; Suggala, A.S.; Inouye, D.I.; Ravikumar, P. How Sensitive are Sensitivity-Based Explanations? *arXiv* **2019**, arXiv:1901.09392.

70. Alvarez-Melis, D.; Jaakkola, T.S. On the Robustness of Interpretability Methods. *arXiv* **2018**, arXiv:1806.08049.

71. Zhang, Z.; Yang, F.; Wang, H.; Hu, X. Contextual Local Explanation for Black Box Classifiers. *arXiv* **2019**, arXiv:1910.00768.

72. Liu, N.; Shin, D.; Hu, X. Contextual Outlier Interpretation. *arXiv* **2017**, arXiv:abs/1711.10589.

73. Sokol, K.; Flach, P. One Explanation Does Not Fit All. *KI Künstl. Intell.* **2020**, *34*, 235–250. [CrossRef]

74. Schreiber, G.; Raimond, Y. RDF 1.1 Primer. W3C Working Group Note, W3C. 2014. Available online: https://www.w3.org/TR/rdf11-primer/ (accessed on 13 April 2021).

75. Hitzler, P.; Krötzsch, M.; Parsia, B.; Patel-Schneider, P.F.; Rudolph, S. OWL 2 Primer. W3C Recommendation, W3C. 2012. Available online: https://www.w3.org/TR/owl2-primer/ (accessed on 13 April 2021).

76. Jennings, C.; Shelby, Z.; Arkko, J.; Keranen, A.; Bormann, C. Sensor Measurement Lists (SenML). Technical Report, IETF. 2018. Available online: https://tools.ietf.org/html/rfc8428 (accessed on 13 April 2021).

77. Su, X.; Riekki, J.; Haverinen, J. Entity Notation: Enabling Knowledge Representations for Resource-Constrained Sensors. *Pers. Ubiquitous Comput.* **2012**, *16*, 819–834. [CrossRef]

78. Su, X.; Riekki, J.; Nurminen, J.K.; Nieminen, J.; Koskimies, M. Adding Semantics to Internet of Things. *Concurr. Comput. Pract. Exp.* **2015**, *27*, 1844–1860. [CrossRef]

79. Al-Osta, M.; Ahmed, B.; Gherbi, A. A Lightweight Semantic Web-Based Approach for Data Annotation on IoT Gateways. In Proceedings of the 8th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2017)/the 7th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2017)/Affiliated Workshops, Lund, Sweden, 18–20 September 2017; Shakshuki, E.M., Ed.; Procedia Computer Science; Elsevier: Amsterdam, The Netherlands, 2017; Volume 113, pp. 186–193. [CrossRef]

80. Su, X.; Zhang, H.; Riekki, J.; Keränen, A.; Nurminen, J.K.; Du, L. Connecting IoT Sensors to Knowledge-Based Systems by Transforming SenML to RDF. In Proceedings of the 5th International Conference on Ambient Systems, Networks and Technologies (ANT 2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014), Hasselt, Belgium, 2–5 June 2014; Shakshuki, E.M., Yasar, A., Eds.; Procedia Computer Science; Elsevier: Amsterdam, The Netherlands, 2014; Volume 32, pp. 215–222. [CrossRef]

81. Aggarwal, C.C.; Ashish, N.; Sheth, A.P. The Internet of Things: A Survey From the Data-Centric Perspective. In *Managing and Mining Sensor Data*; Aggarwal, C.C., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 383–428. [CrossRef]

82. Tagliolato, P.; Fugazza, C.; Oggioni, A.; Carrara, P. Semantic Profiles for Easing SensorML Description: Review and Proposal. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 340. [CrossRef]

83. Jara, A.J.; Olivieri, A.C.; Bocchi, Y.; Jung, M.; Kastner, W.; Skarmeta, A.F. Semantic Web of Things: An Analysis of the Application Semantics for the IoT Moving Towards the IoT Convergence. *Int. J. Web Grid Serv.* **2014**, *10*, 244–272. [CrossRef]

84. Wu, Z.; Xu, Y.; Yang, Y.; Zhang, C.; Zhu, X.; Ji, Y. Towards a Semantic Web of Things: A Hybrid Semantic Annotation, Extraction, and Reasoning Framework for Cyber-Physical System. *Sensors* **2017**, *17*, 403. [CrossRef] [PubMed]

85. Lefrançois, M.; Zimmermann, A. The Unified Code for Units of Measure in RDF: Cdt: Ucum and Other UCUM Datatypes. In *ESWC (Satellite Events)*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2018; Volume 11155, pp. 196–201. [CrossRef]

86. Lefrançois, M.; Zimmermann, A. Custom Datatypes. Unofficial draft, École Nationale Supérieure des Mines de Saint-Étienne. 2020. Available online: https://ci.mines-stetienne.fr/lindt/v3/custom_datatypes (accessed on 14 April 2021).

87. Compton, M.; Barnaghi, P.M.; Bermudez, L.; Garcia-Castro, R.; Corcho, Ó.; Cox, S.J.D.; Graybeal, J.; Hauswirth, M.; Henson, C.A.; Herzog, A.; et al. The SSN Ontology of the W3C Semantic Sensor Network Incubator Group. *J. Web Semant.* **2012**, *17*, 25–32. [CrossRef]

88. Janowicz, K.; Haller, A.; Cox, S.J.D.; Phuoc, D.L.; Lefrançois, M. SOSA: A Lightweight Ontology for Sensors, Observations, Samples, and Actuators. *J. Web Semant.* **2019**, *56*, 1–10. [CrossRef]

89. Bermúdez-Edo, M.; Elsaleh, T.; Barnaghi, P.M.; Taylor, K. IoT-Lite: A Lightweight Semantic Model for the Internet of Things and Its Use With Dynamic Semantics. *Pers. Ubiquitous Comput.* **2017**, *21*, 475–487. [CrossRef]

90. Radulovic, F.; Poveda-Villálon, M.; Vila-Suero, D.; Rodríguez-Doncel, V.; García-Castro, R.; Gómez-Pérez, A. Guidelines for Linked Data Generation and Publication: An Example in Building Energy Consumption. *Autom. Constr.* **2015**, *57*, 178–187. [CrossRef]

91. Quarati, A.; Clematis, A.; Roverelli, L.; Zereik, G.; D'Agostino, D.; Mosca, G.; Masnata, M. Integrating Heterogeneous Weather-Sensors Data Into a Smart-City App. In Proceedings of the 2017 International Conference on High Performance Computing & Simulation, HPCS 2017, Genoa, Italy, 17–21 July 2017; pp. 152–159. [CrossRef]

92. Ramar, K.; Mohan, G.; Narayanasamy, P. Ontological Based Interoperability and Integration Framework for Heterogeneous Weather Systems. *Rev. Téc. Ing. Univ. Zulia* **2016**, *39*, 185–192.

93. Shah, J.; Mishra, B. IoT Enabled Environmental Monitoring System for Smart Cities. In Proceedings of the 2016 International Conference on Internet of Things and Applications (IOTA), Pune, India, 22–24 January 2016; pp. 383–388. [CrossRef]

94. Roussey, C.; Bernard, S.; André, G.; Boffety, D. Weather Data Publication on the LOD Using SOSA/SSN Ontology. *Semant. Web* **2020**, *11*, 581–591. [CrossRef]

95. Battle, R.; Kolas, D. Enabling the Geospatial Semantic Web With Parliament and GeoSPARQL. *Semant. Web* **2012**, *3*, 355–370. [CrossRef]

96. Phuoc, D.L.; Nguyen-Mau, H.Q.; Parreira, J.X.; Hauswirth, M. A Middleware Framework for Scalable Management of Linked Streams. *J. Web Semant.* **2012**, *16*, 42–51. [CrossRef]

97. Boronat, P.; Pérez-Francisco, M.; Calafate, C.T.; Cano, J. Towards a Sustainable City for Cyclists: Promoting Safety Through a Mobile Sensing Application. *Sensors* **2021**, *21*, 2116. [CrossRef]

98. Bröring, A.; Remke, A.; Stasch, C.; Autermann, C.; Rieke, M.; Möllers, J. EnviroCar: A Citizen Science Platform for Analyzing and Mapping Crowd-Sourced Car Sensor Data. *Trans. GIS* **2015**, *19*, 362–376. [CrossRef]

99. Miranda, J.; Ponce, P.; Molina, A.; Wright, P.K. Sensing, Smart and Sustainable Technologies for Agri-Food 4.0. *Comput. Ind.* **2019**, *108*, 21–36. [CrossRef]

100. Dzedzickis, A.; Kaklauskas, A.; Bucinskas, V. Human Emotion Recognition: Review of Sensors and Methods. *Sensors* **2020**, *20*, 592. [CrossRef]

101. Yadegaridehkordi, E.; Noor, N.F.B.M.; Ayub, M.N.B.; Affal, H.B.; Hussin, N.B. Affective Computing in Education: A Systematic Review and Future Research. *Comput. Educ.* **2019**, *142*, 103649. [CrossRef]

102. Hosseini, A.; Buonocore, C.M.; Hashemzadeh, S.; Hojaiji, H.; Kalantarian, H.; Sideris, C.; Bui, A.A.T.; King, C.E.; Sarrafzadeh, M. Feasibility of a Secure Wireless Sensing Smartwatch Application for the Self-Management of Pediatric Asthma. *Sensors* **2017**, *17*, 1780. [CrossRef]

103. Aricò, P.; Borghini, G.; Di Flumeri, G.; Sciaraffa, N.; Babiloni, F. Passive BCI Beyond the Lab: Current Trends and Future Directions. *Physiol. Meas.* **2018**, *39*, 08TR02. [CrossRef]

104. Politou, E.A.; Alepis, E.; Patsakis, C. A Survey on Mobile Affective Computing. *Comput. Sci. Rev.* **2017**, *25*, 79–100. [CrossRef]

105. Kiaghadi, A.; Homayounfar, S.Z.; Gummeson, J.; Andrew, T.; Ganesan, D. Phyjama: Physiological Sensing via Fiber-Enhanced Pyjamas. In *ACM Interactactive Mobile Wearable Ubiquitous Technologies*; Association for Computing Machinery: New York, NY, USA, 2019; Volume 3, pp. 1–29. [CrossRef]

106. Wicaksono, I.; Tucker, C.I.; Sun, T.; Guerrero, C.A.; Liu, C.; Woo, W.M.; Pence, E.J.; Dagdeviren, C. A Tailored, Electronic Textile Conformable Suit for Large-Scale Spatiotemporal Physiological Sensing in Vivo. *NPJ Flex. Electron.* **2020**, *4*, 5. [CrossRef]

107. Zhong, R.Y.; Xu, X.; Klotz, E.; Newman, S.T. Intelligent Manufacturing in the Context of Industry 4.0: A Review. *Engineering* **2017**, *3*, 616–630. [CrossRef]

108. van den Hoogen, J.; Bloemheuvel, S.; Atzmueller, M. The Di-Plast Data Science Toolkit – Enabling a Smart Data-Driven Digital Circular Economy for the Plastics Industry. In Proceedings of the Dutch-Belgian Database Day. Jheronimus Academy of Data Science, 's-Hertogenbosch, The Netherlands, 11 December 2019.

109. Liu, Q.; Liu, Z.; Xu, W.; Tang, Q.; Zhou, Z.; Pham, D.T. Human-Robot Collaboration in Disassembly for Sustainable Manufacturing. *Int. J. Prod. Res.* **2019**, *57*, 4027–4044. [CrossRef]

110. Cimini, C.; Pirola, F.; Pinto, R.; Cavalieri, S. A Human-in-the-Loop Manufacturing Control Architecture for the Next Generation of Production Systems. *J. Manuf. Syst.* **2020**, *54*, 258–271. [CrossRef]

111. Angelopoulos, A.; Michailidis, E.T.; Nomikos, N.; Trakadas, P.; Hatziefremidis, A.; Voliotis, S.; Zahariadis, T. Tackling Faults in the Industry 4.0 Era—A Survey of Machine-Learning Solutions and Key Aspects. *Sensors* **2020**, *20*, 109. [CrossRef]

112. Wang, B.; Li, X.; Freiheit, T.; Epureanu, B.I. Learning and Intelligence in Human-Cyber-Physical Systems: Framework and Perspective. In Proceedings of the 2020 Second International Conference on Transdisciplinary AI (TransAI), Irvine, CA, USA, 21–23 September 2020; pp. 142–145. [CrossRef]

113. Mami, M.N.; Grangel-González, I.; Graux, D.; Elezi, E.; Lösch, F. Semantic Data Integration for the SMT Manufacturing Process Using SANSA Stack. In *The Semantic Web: ESWC 2020 Satellite Events*; Harth, A., Presutti, V., Troncy, R., Acosta, M., Polleres, A., Fernández, J.D., Xavier Parreira, J., Hartig, O., Hose, K., Cochez, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; pp. 307–311.

114. Bader, S.R.; Grangel-Gonzalez, I.; Nanjappa, P.; Vidal, M.E.; Maleshkova, M. *A Knowledge Graph for Industry 4.0. The Semantic Web*; Harth, A., Kirrane, S., Ngonga Ngomo, A.C., Paulheim, H., Rula, A., Gentile, A.L., Haase, P., Cochez, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; pp. 465–480.

115. Grangel-González, I.; Lösch, F.; ul Mehdi, A. Knowledge Graphs for Efficient Integration and Access of Manufacturing Data. In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 8–11 September 2020, Volume 1; pp. 93–100. [CrossRef]

116. Patel, P.; Ali, M.I.; Sheth, A. From Raw Data to Smart Manufacturing: AI and Semantic Web of Things for Industry 4.0. *IEEE Intell. Syst.* **2018**, *33*, 79–86. [CrossRef]

117. de Avila Belbute-Peres, F.; Smith, K.; Allen, K.; Tenenbaum, J.; Kolter, J.Z. End-to-End Differentiable Physics for Learning and Control. In *Advances in Neural Information Processing Systems*; Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R., Eds. Curran Associates, Inc.: Red Hook, NY, USA, 2018; Volume 31.

118. Ma, T.; Barajas-Solano, D.A.; Tipireddy, R.; Tartakovsky, A.M. Physics-Informed Gaussian Process Regression for Probabilistic States Estimation and Forecasting in Power Grids. *arXiv* **2020**, arXiv:2010.04591.

119. Childs, C.M.; Washburn, N.R. Embedding Domain Knowledge for Machine Learning of Complex Material Systems. *MRS Commun.* **2019**, *9*, 806–820. [CrossRef]

120. Pfrommer, J.; Zimmerling, C.; Liu, J.; Kärger, L.; Henning, F.; Beyerer, J. Optimisation of Manufacturing Process Parameters Using Deep Neural Networks as Surrogate Models. *Procedia CiRP* **2018**, *72*, 426–431. [CrossRef]

121. Zacarias, A.G.V.; Reimann, P.; Mitschang, B. A Framework to Guide the Selection and Configuration of Machine-Learning-Based Data Analytics Solutions in Manufacturing. *Procedia CIRP* **2018**, *72*, 153–158. [CrossRef]

122. Zacarias, A.G.V.; Ghabri, R.; Reimann, P. AD4ML: Axiomatic Design to Specify Machine Learning Solutions for Manufacturing. In Proceedings of the 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), Las Vegas, NV, USA, 11–13 August 2020; pp. 148–155.

123. Stewart, R.; Ermon, S. Label-Free Supervision of Neural Networks With Physics and Domain Knowledge. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.

124. Ringsquandl, M.; Lamparter, S.; Lepratti, R.; Kröger, P. Knowledge Fusion of Manufacturing Operations Data Using Representation Learning. In Proceedings of the IFIP International Conference on Advances in Production Management Systems, APMS 2017: Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing, Hamburg, Germany, 3–7 September 2017; Lödding, H., Riedel, R., Thoben, K.D., Kiritsis, D., von Cieminski, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 302–310. [CrossRef]
125. Atzmueller, M.; Kloepper, B.; Mawla, H.A.; Jäschke, B.; Hollender, M.; Graube, M.; Arnu, D.; Schmidt, A.; Heinze, S.; Schorer, L.; et al. Big Data Analytics for Proactive Industrial Decision Support: Approaches & First Experiences in the Context of the FEE Project. *Atp Ed.* **2016**, *58*, 62–74.
126. Duan, W.; Chiang, Y.Y. Building Knowledge Graph From Public Data for Predictive Analysis: A Case Study on Predicting Technology Future in Space and Time. In Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, San Francisco, WA, USA, 31 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 7–13. [CrossRef]
127. Ringsquandl, M.; Kharlamov, E.; Stepanova, D.; Lamparter, S.; Lepratti, R.; Horrocks, I.; Kröger, P. On Event-Driven Knowledge Graph Completion in Digital Factories. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1676–1681. [CrossRef]
128. Banerjee, A.; Dalal, R.; Mittal, S.; Joshi, K.P. Generating Digital Twin Models Using Knowledge Graphs for Industrial Production Lines. In Proceedings of the 2017 ACM on Web Science Conference, WebSci '17, Troy, NY, USA, 25–28 June 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 425–430. [CrossRef]
129. Grangel-González, I.; Halilaj, L.; Vidal, M.E.; Lohmann, S.; Auer, S.; Müller, A.W. Seamless Integration of Cyber-Physical Systems in Knowledge Graphs. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, Pau, France, 9–13 April 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 2000–2003. [CrossRef]
130. Grangel-González, I.; Halilaj, L.; Vidal, M.E.; Rana, O.; Lohmann, S.; Auer, S.; Müller, A.W. Knowledge Graphs for Semantically Integrating Cyber-Physical Systems. In *Database and Expert Systems Applications*; Hartmann, S., Ma, H., Hameurlain, A., Pernul, G., Wagner, R.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 184–199.
131. Szelążek, M.; Bobek, S.; Gonzalez-Pardo, A.; Nalepa, G.J. Towards the Modeling of the Hot Rolling Industrial Process. Preliminary Results. In Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning, Guimaraes, Portugal, 4–6 November 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 385–396.
132. Cao, Q.; Samet, A.; Zanni-Merk, C.; de Bertrand de Beuvron, F.; Reich, C. Combining Chronicle Mining and Semantics for Predictive Maintenance in Manufacturing Processes. *Semant. Web* **2020**, *11*, 927–948. [CrossRef]
133. Atzmueller, M.; Mollenhauer, D.; Schmidt, A. Big Data Analytics Using Local Exceptionality Detection. In *Enterprise Big Data Engineering, Analytics, and Management*; IGI Global: Hershey, PA, USA, 2016.
134. Atzmueller, M.; Hayat, N.; Schmidt, A.; Klöpper, B. Explanation-Aware Feature Selection Using Symbolic Time Series Abstraction: Approaches and Experiences in a Petro-Chemical Production Context. In Proceedings of the IEEE International Conference on Industrial Informatics (INDIN), Emden, Germany, 24–26 July 2017.
135. Atzmueller, M.; Kloepper, B. Mining Attributed Interaction Networks on Industrial Event Logs. In Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning, Workshop on Methods for Interpretation of Industrial Event Logs, Madrid, Spain, 21–23 November 2018; Spriner: Berlin/Heidelberg, Germany, 2018.
136. Buchgeher, G.; Gabauer, D.; Martinez-Gil, J.; Ehrlinger, L. Knowledge Graphs in Manufacturing and Production: A Systematic Literature Review. *arXiv* **2020**, arXiv:1903.12394.
137. Lampropoulos, G.; Keramopoulos, E.; Diamantaras, K. Enhancing the Functionality of Augmented Reality Using Deep Learning, Semantic Web and Knowledge Graphs: A Review. *Vis. Inform.* **2020**, *4*, 32–42. [CrossRef]
138. Atzmueller, M. Declarative Aspects in Explicative Data Mining for Computational Sensemaking. In Proceedings of the International Conference on Declarative Programming (DECLARE), Würzburg, Germany, 19–22 September 2017; Springer: Berlin/Heidelberg, Germany, 2018.
139. Schwenke, L.; Atzmueller, M. Show Me What You're Looking For: Visualizing Abstracted Transformer Attention for Enhancing Their Local Interpretability on Time Series Data. In Proceedings of the 34th International Florida Artificial Intelligence Research Society Conference (FLAIRS-2021), Miami, FL, USA, 16–19 May 2021; FLAIRS/Florida Online Journals: North Miami Beach, FL, USA, 2021.
140. Eagle, N.; Pentland, A.S. Reality Mining: Sensing Complex Social Systems. *Pers. Ubiquitous Comput.* **2006**, *10*, 255–268. [CrossRef]
141. Aggarwal, C.C.; Abdelzaher, T. Social Sensing. In *Managing and Mining Sensor Data*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 237–297.
142. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social Sensing: A New Approach to Understanding Our Socioeconomic Environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [CrossRef]
143. Wang, D.; Szymanski, B.K.; Abdelzaher, T.; Ji, H.; Kaplan, L. The Age of Social Sensing. *Computer* **2019**, *52*, 36–45. [CrossRef]
144. Cattuto, C.; Benz, D.; Hotho, A.; Stumme, G. Semantic Grounding of Tag Relatedness in Social Bookmarking Systems. In Proceedings of the International Semantic Web Conference, Karlsruhe, Germany, 26–30 October 2008; Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T.W., Thirunarayan, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5318, pp. 615–631.

145. Kim, T.; McFee, E.; Olguin, D.O.; Waber, B.; Pentland, A.S. Sociometric Badges: Using Sensor Technology to Capture New Forms of Collaboration. *J. Organ. Behav.* **2012**, *33*, 412–427. [CrossRef]

146. Atzmueller, M. Mining Social Media. *Inform. Spektrum* **2012**, *35*, 132–135. [CrossRef]

147. Atzmueller, M. Mining Social Media: Key Players, Sentiments, and Communities. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2012**, *2*, 411–419. [CrossRef]

148. Atzmueller, M.; Becker, M.; Doerfel, S.; Kibanov, M.; Hotho, A.; Macek, B.E.; Mitzlaff, F.; Mueller, J.; Scholz, C.; Stumme, G. Ubicon: Observing Social and Physical Activities. In Proceedings of the IEEE International Conference On Cyber, Physical and Social Computing (CPSCom 2012), Besancon, France, 20–23 November 2012; IEEE Computer Society: Washington, DC, USA, 2012; pp. 317–324.

149. Salathé, M.; Bengtsson, L.; Bodnar, T.J.; Brewer, D.D.; Brownstein, J.S.; Buckee, C.; Campbell, E.M.; Cattuto, C.; Khandelwal, S.; Mabry, P.L.; et al. Digital Epidemiology. *PLoS Comput. Biol.* **2012**, *8*, e1002616. [CrossRef]

150. Atzmueller, M.; Benz, D.; Doerfel, S.; Hotho, A.; Jäschke, R.; Macek, B.E.; Mitzlaff, F.; Scholz, C.; Stumme, G. Enhancing Social Interactions at Conferences. *IT Inf. Technol.* **2011**, *53*, 101–107. [CrossRef]

151. Atzmueller, M.; Becker, M.; Kibanov, M.; Scholz, C.; Doerfel, S.; Hotho, A.; Macek, B.E.; Mitzlaff, F.; Mueller, J.; Stumme, G. Ubicon and Its Applications for Ubiquitous Social Computing. *New Rev. Hypermedia Multimed.* **2014**, *20*, 53–77. [CrossRef]

152. Gao, X.; Yu, W.; Rong, Y.; Zhang, S. Ontology-Based Social Media Analysis for Urban Planning. In Proceedings of the IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Torino, Italy, 4–8 July 2017; Volume 1, pp. 888–896.

153. Espinoza-Arias, P.; Poveda-Villalón, M.; García-Castro, R.; Corcho, O. Ontological Representation of Smart City Data: From Devices to Cities. *Appl. Sci.* **2019**, *9*, 32. [CrossRef]

154. Sagar, S.; Lefrançois, M.; Rebaï, I.; Khemaja, M.; Garlatti, S.; Feki, J.; Médini, L. Modeling Smart Sensors on Top of SOSA/SSN and WoT TD With the Semantic Smart Sensor Network (S3N) Modular Ontology. In Proceedings of the Internal Semantic Web Conference, Crete, Greece, 3–7 June 2018, pp. 163–177.

155. Mansour, E.; Chbeir, R.; Arnould, P. HSSN: An Ontology for Hybrid Semantic Sensor Networks. In Proceedings of the 23rd International Database Applications & Engineering Symposium, Athens, Greece, 10–12 June 2019; pp. 1–10.

156. Jung, J.J.; Euzenat, J. Towards Semantic Social Networks. In Proceedings of the European Semantic Web Conference, Portorož, Slovenia, 28 May–1 June 2007; Springer: Berlin/Heidelberg, Germany; pp. 267–280.

157. Erétéo, G.; Buffa, M.; Gandon, F.; Corby, O. Analysis of a Real Online Social Network Using Semantic Web Frameworks. In Proceedings of the International Semantic Web Conference, Chantilly, VA, USA, 25–29 October 2009; Springer: Berlin/Heidelberg, Germany; pp. 180–195.

158. Van den Broeck, W.; Cattuto, C.; Barrat, A.; Szomszor, M.; Correndo, G.; Alani, H. The Live Social Semantics application: A platform for integrating face-to-face presence with on-line social networking. In Proceedings of the 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), Lucca, Italy, 12–15 April 2010; pp. 226–231.

159. Gloor, P.A.; Zhao, Y. Analyzing Actors and Their Discussion Topics by Semantic Social Network Analysis. In Proceedings of the Tenth International Conference on Information Visualisation (IV'06), London, UK, 5–7 July 2006; pp. 130–135.

160. Woznowski, P.; King, R.; Harwin, W.; Craddock, I. A Human Activity Recognition Framework for Healthcare Applications: Ontology, Labelling Strategies, and Best Practice. In Proceedings of the International Conference on Internet of Things and Big Data, SCITEPRESS, Rome, Italy, 23–25 April 2016; Volume 2, pp. 369–377.

161. Sheth, A. Computing for Human Experience: Semantics-Empowered Sensors, Services, and Social Computing on the Ubiquitous Web. *IEEE Internet Comput.* **2009**, *14*, 88–91. [CrossRef]

162. Sheth, A.; Anantharam, P.; Henson, C. Physical-Cyber-Social Computing: An Early 21st Century Approach. *IEEE Intell. Syst.* **2013**, *28*, 78–82. [CrossRef]

163. Zempo, K.; Arai, T.; Aoki, T.; Okada, Y. Sensing Framework for the Internet of Actors in the Value Co-Creation Process with a Beacon-Attachable Indoor Positioning System. *Sensors* **2021**, *21*, 83. [CrossRef] [PubMed]

164. Lederman, O.; Mohan, A.; Calacci, D.; Pentland, A.S. Rhythm: A Unified Measurement Platform for Human Organizations. *IEEE MultiMedia* **2018**, *25*, 26–38. [CrossRef]

165. Atzmueller, M. Onto Collective Intelligence in Social Media: Exemplary Applications and Perspectives. In Proceedings of the International Workshop on Modeling Social Media (MSM 2012), Milwaukee, WI, USA, 25 June 2012; ACM Press: New York, NY, USA, 2012.

166. Atzmueller, M. Data Mining on Social Interaction Networks. *arXiv* **2014**, arXiv:1312.6675.

167. Rashid, M.T.; Wang, D. CovidSens: A Vision on Reliable Social Sensing for COVID-19. *Artif. Intell. Rev.* **2021**, *54*, 1–25. [CrossRef]

168. Zhou, J.; Zhou, Y.; Wang, B.; Zang, J. Human–Cyber–Physical Systems (HCPSs) in the Context of New-Generation Intelligent Manufacturing. *Engineering* **2019**, *5*, 624–636. [CrossRef]

169. Mitzlaff, F.; Benz, D.; Stumme, G.; Hotho, A. Visit Me, Click Me, Be My Friend: An Analysis of Evidence Networks of User Relationships in Bibsonomy. In Proceedings of the 21st ACM Conference on Hypertext and Hypermedia, Toronto, ON, Canada, 13–16 June 2010; ACM: New York, NY, USA, 2010.

170. Mitzlaff, F.; Atzmueller, M.; Benz, D.; Hotho, A.; Stumme, G. Community Assessment Using Evidence Networks. In *Analysis of Social Media and Ubiquitous Data*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6904, pp. 79–98.

171. Mitzlaff, F.; Atzmueller, M.; Stumme, G.; Hotho, A. Semantics of User Interaction in Social Media. In *Complex Networks IV*; Studies in Computational Intelligence; Ghoshal, G., Poncela-Casasnovas, J., Tolksdorf, R., Eds.; Spriner: Berlin/Heidelberg, Germany, 2013; Volume 476. [CrossRef]

172. Mitzlaff, F.; Atzmueller, M.; Benz, D.; Hotho, A.; Stumme, G. User-Relatedness and Community Structure in Social Interaction Networks. *arXiv* **2013**, arXiv:1309.3888.

173. Mitzlaff, F.; Atzmueller, M.; Hotho, A.; Stumme, G. The social distributional hypothesis: A pragmatic proxy for homophily in online social networks. *Soc. Netw. Anal. Min.* **2014**, *4*, 216. [CrossRef]

174. Mitzlaff, F.; Atzmueller, M.; Benz, D.; Hotho, A.; Stumme, G. Community Assessment Using Evidence Networks. In Proceedings of the International Workshop on Mining Ubiquitous and Social Environments (MUSE2010), Toronto, ON, Canada, 13 June 2010.

175. Weiser, M. The Computer for the 21st Century. *Sci. Am.* **1991**, *265*, 66–75. [CrossRef]

176. Weiser, M. Some Computer Science Issues in Ubiquitous Computing. *Commun. ACM* **1993**, *36*, 74–84. [CrossRef]

177. Zhong, N.; Liu, J.; Yao, Y. In Search of the Wisdom Web. *Computer* **2002**, *35*, 27–31. [CrossRef]

178. Hotho, A.; Stumme, G. Towards the Ubiquitous Web. *Semant. Web* **2010**, *1*, 117–119. [CrossRef]

179. Knorr-Cetina, K. Sociality With Objects: Social Relations in Postsocial Knowledge Societies. *Theory Cult. Soc.* **1997**, *14*, 1–43. [CrossRef]

180. Dimitrova, V.; Steiner, C.M.; Despotakis, D.; Brna, P.; Ascolese, A.; Pannese, L.; Albert, D. Semantic Social Sensing for Improving Simulation Environments for Learning. In Proceedings of the European Conference on Technology Enhanced Learning, Paphos, Cyprus, 17–21 September 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 601–602.

181. Calbimonte, J.P.; Sarni, S.; Eberle, J.; Aberer, K. XGSN: An Open-Source Semantic Sensing Middleware for the Web of Things. In Proceedings of the International Semantic Web Conference, Trentino, Italy, 20 October 2014; pp. 51–66.

182. Erétéo, G.; Limpens, F.; Gandon, F.; Corby, O.; Buffa, M.; Leitzelman, M.; Sander, P. Semantic social network analysis: A concrete case. In *Handbook of Research on Methods and Techniques for Studying Virtual Communities: Paradigms and Phenomena*; IGI Global: Hershey, PA, USA, 2011; pp. 122–156.

183. Davoodi, E.; Kianmehr, K.; Afsharchi, M. A Semantic Social Network-Based Expert Recommender System. *Appl. Intell.* **2013**, *39*, 1–13. [CrossRef]

184. Macek, B.E.; Atzmueller, M.; Stumme, G. Profile Mining in CVS-Logs and Face-to-Face Contacts for Recommending Software Developers. In Proceedings of the IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9–11 October 2011; IEEE Computer Society: Boston, MA, USA, 2011; pp. 250–257.

185. Lee, R.; Wakamiya, S.; Sumiya, K. Discovery of Unusual Regional Social Activities Using Geo-Tagged Microblogs. *World Wide Web* **2011**, *14*, 321–349. [CrossRef]

186. Scholz, C.; Atzmueller, M.; Barrat, A.; Cattuto, C.; Stumme, G. New Insights and Methods for Predicting Face-to-Face Contacts. In Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, Cambridge, MA, USA, 8–11 July 2013; Kiciman, E., Ellison, N.B., Hogan, B., Resnick, P., Soboroff, I., Eds.; AAAI Press: Palo Alto, CA, USA, 2013.

187. Kibanov, M.; Atzmueller, M.; Illig, J.; Scholz, C.; Barrat, A.; Cattuto, C.; Stumme, G. Is Web Content a Good Proxy for Real-Life Interaction? A Case Study Considering Online and Offline Interactions of Computer Scientists. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Paris, France, 25–28 August 2015; IEEE Press: Boston, MA, USA, 2015.

188. Hasan, S.; Zhan, X.; Ukkusuri, S.V. Understanding Urban Human Activity and Mobility Patterns Using Large-Scale Location-Based Data From Online Social Media. In Proceedings of the International Workshop on Urban Computing, UrbComp '13, Chicago, IL, USA, 11 August 2013; ACM: New York, NY, USA, 2013.

189. Janowicz, K.; McKenzie, G.; Hu, Y.; Zhu, R.; Gao, S. Using Semantic Signatures for Social Sensing in Urban Environments. In *Mobility Patterns, Big Data and Transport Analytics*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 31–54.

190. Tonkin, E.L.; Woznowski, P.R. Activities of Daily Living Ontology for Ubiquitous Systems. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece, 19–23 March 2018; pp. 573–578.

191. Woznowski, P.R.; Tonkin, E.L.; Flach, P.A. Activities of Daily Living Ontology for Ubiquitous Systems: Development and Evaluation. *Sensors* **2018**, *18*, 2361. [CrossRef]

192. Bennasar, M.; Price, B.A.; Stuart, A.; Gooch, D.; McCormick, C.; Mehta, V.; Clare, L.; Bennaceur, A.; Cohen, J.; Bandara, A.K.; et al. Knowledge-Based Architecture for Recognising Activities of Older People. *Procedia Comput. Sci.* **2019**, *159*, 590–599. [CrossRef]

193. Baryannis, G.; Woznowski, P.; Antoniou, G. Rule-Based Real-Time Adl Recognition in a Smart Home Environment. In *International Symposium on Rules and Rule Markup Languages for the Semantic Web*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 325–340.

194. Atzmueller, M.; Hayat, N.; Trojahn, M.; Kroll, D. Explicative Human Activity Recognition Using Adaptive Association Rule-Based Classification. In Proceedings of the IEEE International Conference on Future IoT Technologies, Eger, Hungary, 18–19 January 2018; IEEE Press: Boston, MA, USA, 2018.

195. Ahn, Y.Y.; Han, S.; Kwak, H.; Moon, S.; Jeong, H. Analysis of Topological Characteristics of Huge Online Social Networking Services. In Proceedings of the 16th International Conference On the World Wide Web (WWW), Banff, AB, Canada, 8–12 May 2007; ACM: New York, NY, USA, 2007; pp. 835–844.

196. Mislove, A.; Marcon, M.; Gummadi, K.P.; Druschel, P.; Bhattacharjee, B. Measurement and Analysis of Online Social Networks. In Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, San Diego, CA, USA, 24–26 October 2007; ACM: New York, NY, USA, 2007; pp. 29–42.

197. Kwak, H.; Lee, C.; Park, H.; Moon, S. What Is Twitter, a Social Network or a News Media? In Proceedings of the 19th International Conference On World Wide Web (WWW), Raleigh, NC, USA, 26–30 April 2010; ACM: New York, NY, USA, 2010; pp. 591–600.

198. Chelmis, C.; Wu, H.; Sorathia, V.; Prasanna, V.K. Semantic Social Network Analysis for the Enterprise. *Comput. Inform.* **2014**, *33*, 479–502.

199. Gruber, T. Ontology of Folksonomy: A Mash-Up of Apples and Oranges. *Int. J. Semant. Web Inf. Syst. (IJSWIS)* **2007**, *3*, 1–11. [CrossRef]

200. Tang, J.; Leung, H.f.; Luo, Q.; Chen, D.; Gong, J. Towards Ontology Learning From Folksonomies. In Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, CA, USA, 14–17 July 2009; Volume 9, pp. 2089–2094.

201. Miyaki, T.; Rekimoto, J. Sensonomy: EnvisioningFolksonomic Urban Sensing. In *UbiComp 2008 Workshop Programs*; Citeseer: State College, PA, USA, 2008; pp. 187–190.

202. Rekimoto, J. Sensonomy: Intelligence Penetrating Into the Real Space. In Proceedings of the 14th International Conference on Intelligent User Interfaces, Sanibel Island, FL, USA, 8–11 February 2009; pp. 3–4.

203. Atzmueller, M.; Mueller, J.; Becker, M. Exploratory Subgroup Analytics on Ubiquitous Data. In *Mining, Modeling and Recommending 'Things' in Social Media*; Number 8940 in LNAI; Springer: Berlin/Heidelberg, Germany, 2015.