Declarative Aspects in Explicative Data Mining for Computational Sensemaking

Martin Atzmueller

Tilburg University,
Department of Cognitive Science and Artificial Intelligence
Warandelaan 2, 5037 AB Tilburg, Netherlands

m.atzmuller@uvt.nl

Abstract. Computational sensemaking aims to develop methods and systems to "make sense" of complex data and information. The ultimate goal is then to provide insights and enhance understanding for supporting subsequent intelligent actions. Understandability and interpretability are key elements of that process as well as models and patterns captured therein. Here, *declarativity* helps to include guiding knowledge structures into the process, while *explication* provides interpretability, transparency, and explainability. This paper provides an overview of the key points and important developments in these areas, and outlines future potential and challenges.

Keywords: computational sensemaking, data mining, declarative modeling, domain knowledge, explicative data analysis, knowledge graph, statistical relational learning

1 Introduction

Computational sensemaking aims to "make sense" in the context of complex information and knowledge processes. This is enabled using computational methods for *analysis*, *interpretation*, and *intelligent decision-support*. While the latter is mostly supported by human-computer interaction techniques, the former two are supported by data mining approaches, in particular, *explicative data mining* methods.

Overall, data mining systems are commonly applied to obtain a set of *novel*, *potentially useful*, and ultimately *interesting* patterns from (large) data sets [27]. While the resulting patterns are typically interpretable, e. g., in pattern mining, the large result sets of potentially interesting patterns that the user needs to assess, require further *exploration* and *interpretation* techniques. In general, facilitating the understandability and interpretability of the process as well as its "products" (e. g., in the form of patterns) need to consider two important aspects: *declarativity*, in order to include guiding knowledge structures into the process, as well as *explication* in order to provide interpretability, transparency, and explainability. Both *declarative* as well as *explicative* approaches work together in that context, complementing each other. This paper provides an overview of the key points and important developments in these areas, and outlines future potential and challenges.

Preprint of: Martin Atzmueller. Declarative Aspects in Explicative Data Mining for Computational Sensemaking.
Proc. International Conference on Declarative Programming (DECLARE), Springer, Heidelberg, Germany, 2018.

2 Declarative Aspects in Explicative Data Mining

While declarative approaches allow for the incorporation of background knowledge and the guidance of the data mining process, explicative data mining [3] focuses specifically on obtaining interpretable models and patterns, on transparency on the data mining process, and on explanainable or explanation-aware mining. In that way, both complement each other quite well, such that declarative aspects can be incorporated into explicative data mining for enhancing interpretability, transparency, and explainability.

Below, we briefly introduce declarative aspects on data mining, especially focussing on the modeling of background knowledge and the specification of knowledge to be incorporated into the data mining process. For that, we first introduce explicative data mining methods, including exploratory and explanation-aware approaches. Here, we discuss examples in the context of pattern mining methods [1, 2, 4, 13, 45, 94], since pattern mining is a prominent research direction for obtaining interpretable patterns, enabling a transparent data mining process. In particular, we discuss the relation to incorporating prior knowledge, e. g., in the form of knowledge patterns [12] and knowledge graphs [16, 36, 92] into the data mining process. This enables hybrid approaches that incorporate semantic knowledge into the process, e. g., supporting modeling and explanation methods.

2.1 Explicative Data Mining

Data mining methods are commonly applied to obtain a set of *novel*, *potentially useful*, and ultimately *interesting* patterns from (large) data sets cf., [27]. This can be achieved e. g., utilizing exploratory data mining techniques like association rule mining or subgroup discovery, as sketched above.

However, most common data mining methods and approaches lack important aspects, i. e., *interpretability*, *transparency* and *explainability* in order to be *explicative* towards its users. Especially considering complicated black-box models this becomes relevant, e. g., when providing recommendations and filtering. Prominent application examples include, for example, large online social networks, e. g., when providing posts or news to users, but also in predictive settings such as user scoring or classification in e-commerce. Here, intransparent methods and models make it more difficult to spot mistakes and can lead to biased decisions, e. g., based on incorrect training data; in general, they 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.

This is particularly important in the light of the European Union's new General Data Protection Regulation, which will as of this year enforce a "right to explanation" (providing users the right to obtain an explanation for any algorithmic decisions that were made about them). Overall, there will be a major impact on business, technology, and society. In particular, in the area of data mining, these developments give rise to major research challenges and a major impact on the interaction of humans with such algorithms and according technology in itself, cf., [31].

Explicative data mining [3] is a comprehensive paradigm for interpretable, transparent and explainable data analysis. Similar to the philosopical process of *explication* cf., [19, 54] which aims to make the implicit explicit, explicative data mining aims to model, describe and explain the underlying structure in the data.

Explicative data mining targets interpretable (and transparent) models utilizing exploratory and explanation-aware methods. These can be constructed and inspected on different layers and levels. This ranges from pure data summarization to pattern-based exploratory data mining. Furthermore, these features also provide for different options for including the human in the loop, e. g., using visualization methods embedded into interactive and semi-automatic approaches and methods.

Below, we discuss how to include declarative aspects into explicative data mining, focussing on interpretable models as well as explainable or explanation-aware approaches. For the former, we focus on how to model and provide explication knowledge that is then integrated into the data mining process. We take a pragmatic view, and consider the typical data mining process, e. g., structured according to the CRISP-DM [20, 93] cycle, "as is" – and thus incorporate important elements of purely declarative approaches, e. g., [17, 18]. We first focus on exploratory approaches, before we discuss explanation-aware methods.

2.2 Exploratory Data Mining

In the scope of explication, exploratory data mining techniques can provide a first view on the data in order to detect interesting patterns. Exploratory techniques span from statistical approaches for characterizing a dataset or determining key (influence) factors, e. g., [29, 89] to more refined (semi-)automatic approaches, e. g., for *local pattern detection* [43, 58, 59]. Local pattern detection aims to discover *local models* chacterizing (or describing) parts of the data given an interestingness measure, e. g., [43]. In addition, interactive visualization methods, e. g., [28, 30, 39, 42, 76, 78, 79, 82, 86] can be applied (or combined with automatic methods) for further supporting data exploration.

Pattern Mining Overall, a broadly applicable and powerful set of methods is provided by the area of pattern mining. Common methods include those for association rule mining [1] or subgroup discovery, e. g., [2, 40, 94]. The latter is at the intersection of descriptive and predictive data mining [45] and can be applied for a variety of different analytical tasks. Essentially, *subgroup discovery* [2, 40, 45, 94] is an exploratory approach for discovering interesting subgroups – as an instance of local pattern detection [2, 43, 48, 58, 59]. The interestingness is usually defined by a certain property of interesting formalized by a quality function. Essentially, subgroup discovery is a flexible method for detecting relations between dependent (characterizing) variables and a dependent target concept, e. g., comparing the share or the mean of a nominal/numeric target variable in the subgroup vs. the share or mean in the total population, respectively. The interestingness of a pattern can then be flexibly defined, e. g., by a significant deviation from a model that is derived from the total populatioctn. In the simplest case, (see the example above) a binary target variable is considered, where the share in a subgroup can be compared to the share in the dataset in order to detect (exceptional) deviations.

More complex target concepts consider sets of target variables. Here, *exceptional model mining* [2, 25, 47] focuses on more complex quality functions, considering complex *target models*, like comparing regression models or graph structures. Essentially, exceptional model mining tries to identify interesting patterns with respect to a local model derived from a set of attributes, cf., [23, 24]. This can be extended, e. g., for network analysis and (exceptional) graph mining, e. g., [38]. Below, we introduce and define subgroup models (and patterns), as well as association rules more formally.

Domain Knowledge for Semantic Data Mining In many domains, a lot of (semantic) domain knowledge is available in order to support reasoning processes. However, in data mining, semantic knowledge is scarcely exploited so far. Domain knowledge is a natural resource for knowledge-intensive data mining methods, e.g., [37,70], and can be exploited for improving the quality of the data mining results significantly. Appropriate domain knowledge can increase the representational expressiveness and also focus the algorithm on the relevant patterns. Furthermore, for increasing the efficiency of the search method, the search space can often be constrained, e.g., [9].

There are several approaches, which show how to effectively provide and include domain knowledge into data mining approaches, e. g., [9, 22, 56, 57, 60, 64, 88] thus supporting explicative data mining by providing semantic specifications. For modeling expected relations for pattern discovery, for example, according methods are presented in [7] utilizing Bayesian network formalizations. For relational data analysis approach, also the comparison of hypotheses with a semantic data model using Bayesian techniques (first order Markov chains) has been targeted in [11]. Furthermore, statistical relational learning, e. g., [21, 65, 71] combines both probabilistic and complex relational learning approaches, in particular also enabling complex logic-based methods. Such methods then enable powerful declarative approaches in order to provide domain knowledge for data mining.

Easing Knowledge Acquisition Costs However, knowledge acquisition is often challenging and costly, a fact that is known as the so-called *knowledge acquisition bottleneck*. Thus, an important idea is to ease the knowledge acquisition by reusing existing domain knowledge, i.e., already formalized knowledge that is contained in existing ontologies or knowledge bases. Furthermore, we aim to simplify the knowledge acquisition process itself by providing knowledge concepts that are easy to model and to comprehend.

We propose high-level knowledge, such as properties of ontological objects for deriving simpler constraint knowledge that can be directly included in the data mining step, as discussed, e.g., in [8, 9]. Modeling such higher-level ontological knowledge, i.e., properties and relations between domain concepts, is often easier for the domain specialist, since it often corresponds to the mental model of the concrete domain. Below, we outline simple specifications of domain knowledge, before we discuss more complex modeling approaches including integrated knowledge graphs and statistical relational learning approaches.

Subgroups and Association Rules Subgroup models [9, 41], often provided by conjunctive rules, describe 'interesting' subgroups of cases, e. g., "the subgroup of 16-20 year old men that own a sports car are more likely to pay high insurance rates than the people in the reference population." Subgroup discovery [2, 40, 46, 94] is a powerful method, e. g., for (data) exploration and descriptive induction, i. e., to obtain an overview of the relations between a so-called target concept and a set of explaining features. These features are represented by attribute/value assignments, i. e., they correspond to binary features such as items known from association rule mining [1]. As discussed below, in its simplest case the target concept is often represented by a binary variable, but can also extend to more complex target concepts, e. g., considering sets of variables, and their relations.

Formally, a *database* $DB = (I, A)$ is given by a set of individuals I and a set of attributes A. For each attribute $a \in A$, a range $dom(a)$ of values is defined. An attribute/value assignment $a = v$, where $a \in A, v \in dom(a)$, is called a *feature*. We define the feature space V to be the (universal) set of all features.

Basic elements used in subgroup discovery are patterns and subgroups. Intuitively, a *pattern* describes a *subgroup*, i. e., the subgroup consists of instances that are covered by the respective pattern. It is easy to see, that a pattern describes a fixed set of instances (subgroup), while a subgroup can also be described by different patterns, if there are different options for covering the subgroup' instances. In the following, we define these concepts more formally, following an adapted notation of [12].

Definition 1. *A (subgroup)* pattern P *is defined as a conjunction*

$$P = s_1 \wedge s_2 \wedge \cdots \wedge s_n$$

of (extended) features $s_i \subseteq V$ *, which are then called selection expressions, where each* s_i *selects a subset of the range* $dom(a)$ *of an attribute* $a \in A$*.*

A selection expression s is thus a Boolean function $I \to \{0, 1\}$ that is true if the value of the corresponding attribute is contained in the respective subset of V for the respective individual.

Definition 2. *A* subgroup (extension) $I_P := ext(P) := \{i \in I | P(i) = true\}$ *is the set of all individuals which are covered by the pattern* P*.*

The subgroup mentioned above, for example, is described by the relation between the independent (explaining) variables (Sex = male, Age ≤ 20, Car = sports car). Furthermore, there is a dependent (target) variable, i. e., Insurance Rate = high; this target variable relates to the *concept of interest* used in subgroup discovery, which is utilized to estimate the interestingness of a subgroup using a quality measure. This is captured by the notion of a subgroup model described below.

In general, the applied quality measure can also be defined using a set of target variables, or more complex models such as Bayesian networks or topological graph structures which relates to the area of *exceptional model mining* [2, 25]. In the scope of this paper and our simple example, we focus on simple *binary* target variables given by (simple) features as defined above.

An *association rule*, e.g., [1], is given by a rule of the form $P_B \rightarrow P_H$, where P_B and P_H are patterns; the rule body P_B and the rule head P_H specify sets of items. For the insurance domain, for example, we can consider an association rule showing a combination of potential risk factors for high insurance rates and accidents:

$$\text{Sex} = \text{male} \wedge \text{Age} \leq 20 \wedge \text{Car} = \text{sports car}$$
$$\rightarrow \text{Insurance Rate} = \text{high} \wedge \text{Accident Rate} = \text{high}$$

A *subgroup model* is a special association rule, namely a horn clause $P \rightarrow e$, where P is a pattern and the feature $e \in V$ is called the target variable. For subgroup discovery, a fixed rule head is considered.

In general, the quality of an association rule is measured by its support and confidence, and the data mining process searches for association rules with arbitray rule heads and bodies, e. g., using the apriori algorithm [1]. For subgroup models there exist various (more refined) quality measures, e.g., [2,40]: Since an arbitrary quality function can be applied, the anti-monotony property of support used in association rule mining cannot be utilized in the general case.

The applied quality function can also combine the difference of the confidence and the apriori probability of the rule head with the size of the subgroup. Since mining for interesting subgroup patterns is more complicated, usually a fixed, atomic rule head is given as input to the knowledge discovery process.

Declarative Specifications of Domain Knowledge As we have presented in [12], a prerequisite for the successful application and exploitation of domain knowledge is given by a concise *declarative* specification of the domain knowledge. A concise specification also provides for better documentation, extendability, and standardization. Below, we summarize the approaches proposed in [12] and provide examples of its instantiation in the field of pattern mining as outlined above.

In contrast to existing approaches, e.g., [70,95] we focus on domain knowledge that can be easily declared in symbolic form. Furthermore, the presented approach features the ability of deriving *simpler* low-level knowledge (constraints) from high-level ontological knowledge. In general, the search space considered by the data mining methods can be significantly reduced by *shrinking* the value ranges of the attributes. Furthermore, the search can often be focused if only *meaningful* values are taken into account. This usually depends on the considered ontological domain.

The considered classes of domain knowledge include ontological knowledge and (derived) constraint knowledge, as a subset of the domain knowledge described in [8,9]. Figure 1 shows the knowledge hierarchy proposed in [12], from the two knowledge classes to the specific types, and the objects they apply to.

Prolog-based Specifications For specifying the properties and relations of the concepts contained in the domain ontology, we utilize Prolog rules as a compact and versatile representation, cf., [12]. Using these rules, we obtain a suitable representation formalism for ontological knowledge. Using these, we can automatically derive ad-hoc relations between ontological concepts using (simple) rules.

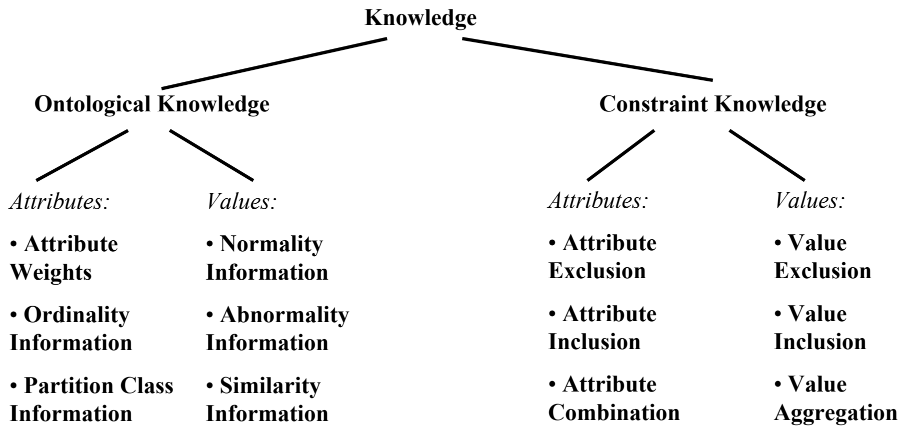

Fig. 1. Hierarchy of (abstract) knowledge classes and specific types, cf., [12].

Essentially, the declarative features of Prolog allow simple and transparent knowledge specification, integration and advancement: Depending on the experience of the domain specialist, new knowledge can be added extending the existing knowledge, new relations can be introduced, and furthermore additional advanced features like derivation rules can be directly implemented using Prolog. In addition, using domain specific languages built on top, e. g., [75, 85] the declarativity can even be further enhanced, while also providing an even simpler interface to the domain specialist. Then, this both provides for a concise specification and also comprehensive overview, documentation and summary for the domain specialist, which is typically easy to comprehend, to interpret and to extend.

Below, we focus on selected examples proposed in [12]. For brevity, we focus on simple examples considering attributes, e. g., using attribute weights, and attribute inclusion/exclusion constraints, cf., Figure 1. With these, attributes can be selected (or excluded) such that they do not occur in patterns constructed by the applied pattern mining method. In that way, for example, exclusion constraints restrict the pattern space. Also, combination constraints inhibit the examination of specified sets of concepts. In that way, they help to find more understandable results. For increasing the representational expressiveness and thus the interpretability of patterns, modifications of the considered attributes (and their combinations) can be utilized to make the discovered patterns and models more meaningful for the user. It is easy to see, that the specifications regarding combinations of attributes, and their explicit exlusion and inclusion directly map to attribute values, and features, respectively. Then, regarding the presented pattern mining methods the format of the considered patterns, and their "building blocks" can be conveniently

Examples. As outlined above, we summarize some simple examples regarding the declarative specifications presented in [12], for which we refer to for an in-depth description and detailed discussion.

As a first example, *partition class information*, provides semantically distinct groups of attributes. These disjoint subsets usually correspond to certain problem areas of the application domain. E.g., in the medical domain such partitions are representing different organ systems like *liver*, *kidney*, *pancreas*, *stomach*, *stomach*, and *intestine*. For each organ system a list of attributes is given:

```
attribute_partition(inner_organs, [
   [fatty_liver, liver_cirrhosis, ...],
   [renal_failure, nephritis, ...], ... ]).
```

Furthermore, *attribute weights* denote the relative importance of attributes, and are a common extension for knowledge-based systems [14]. In the car insurance domain, for example, we can state that the attribute *Age* is more important than the attribute *Car Color*, since its assigned weight is higher:

```
weight(age, 4).
weight(car_color, 1).
```

Deriving Constraints. We can construct attribute exclusion constraints using attribute weights to filter the set of relevant attributes by a weight threshold or by subsets of the weight space.

```
dsdk_constraint(exclude(attribute), A) :-
   weight(A, N), N =< 1.
dsdk_constraint(include(attribute), A) :-
   weight(A, N), N > 1.
```

Partition class information can be used to infer attribute combination constraints in order to prevent the combination of individual attributes that are contained in separate partition classes. Alternatively, inverse constraints can also be derived, e.g., to specifically investigate inter-organ relations in the medical domain.

```
dsdk_constraint(exclude(attribute_pair), [A1, A2]) :-
   attribute_partition(_, P),
   member(As1, P), member(As2, P), As1 \= As2,
   member(A1, As1), member(A2, As2).
```

Finally, we can use a generic Prolog rule for detecting conflicts w.r.t. these rules and the derived knowledge (*automatic verification*):

```
dsdk_constraint(error(include_exclude(X), Y) :-
   dsdk_constraint(include(X), Y),
   dsdk_constraint(exclude(X), Y).
```

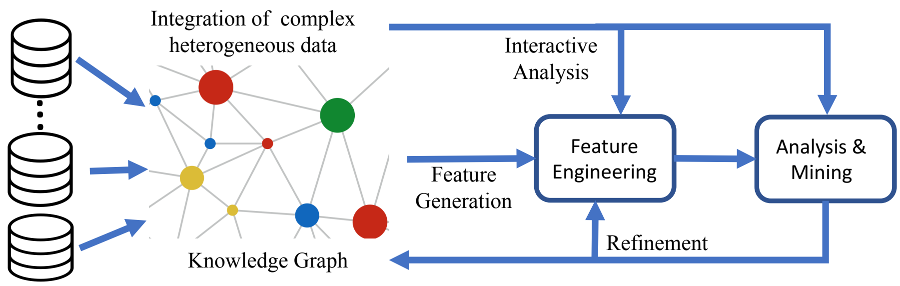

Fig. 2. Overview on a framework for mixed-initiative feature engineering and data mining using knowledge graphs, cf., [13] for a detailed discussion.

Knowledge Graphs A further effective approach for modeling explication knowledge is given by constructing a knowledge graph cf., e. g., [16,36]: Here, the data is integrated into a comprehensive knowledge structure capturing the relations between concepts and their properties in an explicit way, cf., [16, 36, 72, 92]. Then, this structure can be exploited in order to facilitate data mining, e. g., by applying ontologies in the data mining step. However, so far the approaches only apply a "shallow" coupling, that is, typically there is no deep integration of knowledge graph and mining approach.

First approaches for integrating knowledge graphs, i. e., based on ontologies and a set of instance data has been proposed in the area of semantic data mining [66, 87, 88]. In [87, 88] an ontology is used for instantiating pattern elements. Compared to the approach making use of declarative specifications discussed above, this makes of the relations modeled in the ontology, in order to connect different concepts. However, compared to the logic-based approach using the versatile Prolog representation, no simple specification/declaration of further processing knowledge like inference and derivation rules is possible. Likewise, in [66] mainly the instantiation of the knowledge elements is utilized in the mining process.

Typically, the knowledge graph mainly focuses on the structuring of the concepts and their relations, while specific modeling tasks, as well as data characteristics (e. g., distributions, correlations) are typically not captured. [13] presents a mixed-initiative approach, for semantic feature engineering using a knowledge graph. In a semi-automatic process, the knowledge graph is engineered and refined. Finally, the engineered features are provided for data mining. A similar approach is applied in [6]. Here, data from heterogeneous data sources is integrated into a knowledge graph, which then provides the basis for data mining by supporting feature selection, pattern mining, and interpretation in an integrated way. In particular, the constructed knowledge graph serves as a data integration and exploration mechanism, such that the modeled relations and additional information about the contained entities can be utilized by advanced graph mining methods, that work on such attributed graphs, e. g., [11].

Also, the obtained knowledge graph itself can be applied for providing additional context regarding the results of the data mining step, e. g., in order to provide explanations [10, 83] as discussed below in more detail.

2.3 Explicative and Explanation-Aware Data Mining

The term explanation has been widely investigated in different disciplines. In this context, *explanation-aware* approaches have been a prominent research direction in artificial intelligence and data science, e. g., [10, 44, 73, 91].

On Explanation Knowing about kinds of explanations helps with structuring available knowledge and deciding which knowledge further is required for exhibiting certain explanation capabilities. In [74], Roth-Berghofer and Cassens outline the combination of goals and kinds of explanations, in the context of case-based reasoning. In [80], several useful kinds of explanations are discussed in the context of knowledge-based systems, referring to *concept explanations*, *purpose explanations*, *why explanations*, *action explanations*, and *how explanations*, cf., [10, 80] for a detailed discussion.

In the data mining context, *concept*, *why* and *how* explanations are then particularly useful, since they provide insights into knowledge elements utilized in modeling, and also in the model itself by explicating model mechanisms and outcomes. Explanation goals, on the other hand, help to focus on user needs and expectations towards explanations. They aim at addressing to understand what and when the system has to be able to explain (something). Sørmo et al. [77] suggest a set of explanation goals addressing transparency, justification, relevance, conceptualisation, and learning.

Explicative Modeling Recently, the concept of transparent and explainable models has also gained a strong focus and momentum in the data mining and machine learning community, e. g., [15,50,68], also see [32] for a survey on explaining black box models. Several methods focus on specific model types, e. g., tree-based models [84] or pattern-based approaches [26] for getting a better understanding of where a classifier does not work using local pattern mining techniques. Here, also methods integrating associative classification, i. e., based utilizing a set of (class) association rules [5, 49, 53, 81] can be applied for obtaining interpretable models for explicative data mining. While the methods sketched above focus on specific modeling methods, there are several approaches for model agnostic explanation methods, e. g., [67, 69]. In particular, general directions are given by methods considering counterfactual explanation, e. g., [55, 90]. Furthermore, other general methods consider data perturbation and randomization techniques as well as interaction analysis methods, e. g., [33–35].

In general, for explicative data mining, the transparency of the respective patterns and models and their explanation-awareness is an important factor for supporting the user. In particular, if explanations for the complete models, or parts thereof can be provided, then the acceptance of the patterns and models, as well as their assessment and evaluation can often be significantly improved, e. g., [10].

Explanation-Aware Data Mining The generation of explanations in the general data mining process is described in [10], the *mining and analysis continuum of explanation*. In particular, if explanations for the complete models, or parts thereof can be provided, then the acceptance can often be significantly improved, e. g., [10]. As put forward and described in the *Mining and Analysis Continuum of Explaining* [10] appropriate data representation and abstraction can facilitate explanation-awareness, also supporting and featuring different analysis and presentation levels. Then, data and models can be inspected at different levels of detail, from aggregated representations to the original ones in drill-down fashion combined with appropriate explanation capabilities. Typically, the user starts on an aggregated view that can be refined subsequently, for getting insights into the relations in the data and the constructed model, respectively. Here, different dimensions provide distinct view on the explanation space. Figure 3 provides an overview on the explanation framework and its dimensions. For a detailed discussion, we refer to [10].

In [4], for example, we consider symbolic representations, i. e., decision tree models and sequence representations of time series given by the symbolic aggregate approximation (SAX) [51, 52] as a convenient data abstraction. In a general process model for explanation-aware data analytics, we investigate this abstraction together with a decision tree model in the context of feature selection and assessment, and present a case study in a petro-chemical production context.

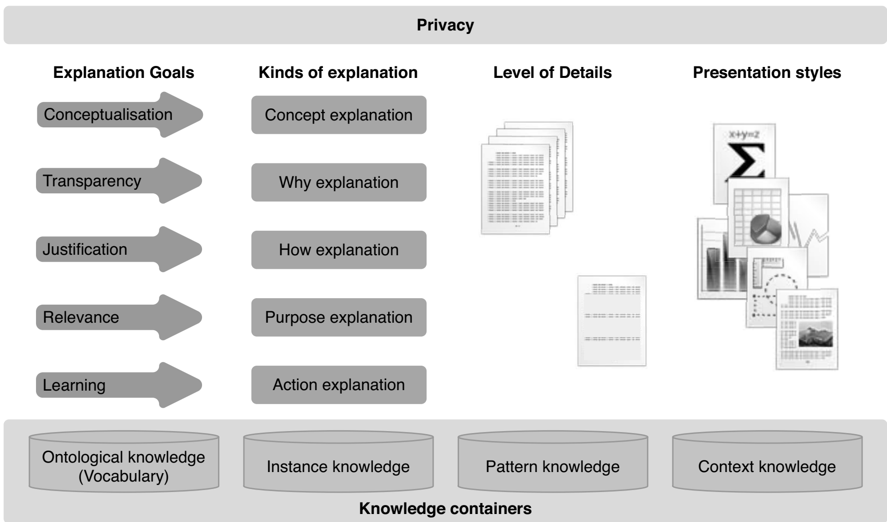

Fig. 3. Overview on the explanation dimensions of the mining and analysis continuum of explanation cf., [10].

Some recent approaches for introducing declarativity in explanation-aware approaches, include the knowledge-graph-based data mining approach outlined in [6] which is detailed in [11] regarding the applied pattern mining techniques. Furthermore, linked open data inspired approaches for interpreting pattern-based models, e. g., [62, 63] and also explanations using linked open data for recommender systems [61] are first promising starting points in that context.

3 Conclusions

In this paper, we have provided an overview on declarative aspects in explicative data mining, targeting the overall goal of computational sensemaking. We have discussed the modeling of domain knowledge as well as extended knowledge structuring using knowledge graphs. Furthermore, we have summarized the paradigm of explicative data mining providing interpretable, transparent and explainable approaches.

We have introduced *explicative data mining* as a comprehensive paradigm. Similar to the philosopical process of *explication* cf., [19, 54] which aims to make the implicit explicit, explicative data mining aims to model, describe and explain the underlying structure in the data. In that way, this paves the way to *computational sensemaking* which focuses on computational methods and models for "making sense" of complex data and information. Here, the goal is to understand structures and processes and to provide intelligent decision support through analysis and (semantic) interpretation. Therefore, explicative data mining coupled with declarative approaches is crucial since this can then both provide the necessary means for comprehensive analysis and giving meaning to models and results, respectively.

References

1. Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules. In: Proc. VLDB. pp. 487–499. Morgan Kaufmann (1994)
2. Atzmueller, M.: Subgroup Discovery. WIREs Data Mining and Knowledge Discovery 5(1), 35–49 (2015)
3. Atzmueller, M.: Onto Explicative Data Mining: Exploratory, Interpretable and Explainable Analysis. In: Proc. Dutch-Belgian Database Day. TU Eindhoven, Netherlands (2017)
4. 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: Proc. IEEE INDIN. IEEE Press, Boston, MA, USA (2017)
5. Atzmueller, M., Hayat, N., Trojahn, M., Kroll, D.: Explicative Human Activity Recognition using Adaptive Association Rule-Based Classification. In: Proc. IEEE International Conference on Future IoT Technologies. IEEE Press, Boston, MA, USA, 2018)
6. Atzmueller, M., Kloepper, B., Mawla, H.A., Jäschke, B., Hollender, M., Graube, M., Arnu, D., Schmidt, A., Heinze, S., Schorer, L., Kroll, A., Stumme, G., Urbas, L.: Big Data Analytics for Proactive Industrial Decision Support: Approaches & First Experiences in the Context of the FEE Project. atp edition 58(9) (2016)
7. Atzmueller, M., Lemmerich, F.: A Methodological Approach for the Effective Modeling of Bayesian Networks. In: Proc. Annual German Conference on Artificial Intelligence (KI). pp. 160–168. Springer, Heidelberg, Germany (2008)
8. Atzmueller, M., Puppe, F.: A Methodological View on Knowledge-Intensive Subgroup Discovery. In: Proc. EKAW. pp. 318–325. Springer, Heidelberg, Germany (2006)
9. Atzmueller, M., Puppe, F., Buscher, H.P.: Exploiting Background Knowledge for Knowledge-Intensive Subgroup Discovery. In: Proc. 19th International Joint Conference on Artificial Intelligence (IJCAI-05). pp. 647–652. Edinburgh, Scotland (2005)
10. Atzmueller, M., Roth-Berghofer, T.: The Mining and Analysis Continuum of Explaining Uncovered. In: Proc. AI (2010)
11. Atzmueller, M., Schmidt, A., Kloepper, B., Arnu, D.: HypGraphs: An Approach for Analysis and Assessment of Graph-Based and Sequential Hypotheses. In: Postproceedings NFMCP 2016. Springer (2017)

12. Atzmueller, M., Seipel, D.: Declarative Specification of Ontological Domain Knowledge for Descriptive Data Mining. In: Proc. International Conference on Applications of Declarative Programming and Knowledge Management (INAP). Springer, Heidelberg, Germany (2008)
13. Atzmueller, M., Sternberg, E.: Mixed-Initiative Feature Engineering Using Knowledge Graphs. In: Proc. K-CAP. ACM (2017)
14. Baumeister, J., Atzmueller, M., Puppe, F.: Inductive Learning for Case-Based Diagnosis with Multiple Faults. In: Advances in Case-Based Reasoning. LNAI, vol. 2416, pp. 28–42. Springer, Heidelberg, Germany (2002)
15. Biran, O., Cotton, C.: Explanation and Justification in Machine Learning: A Survey. In: IJCAI-17 Workshop on Explainable AI (2017)
16. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia - A Crystallization Point for the Web of Data. Web Semantics 7(3), 154–165 (2009)
17. Blockeel, H.: Data Mining: From Procedural to Declarative Approaches. New Generation Computing 33(2), 115–135 (2015)
18. Blockeel, H.: Declarative Data Analysis. International Journal of Data Science and Analytics pp. 1–7 (2017)
19. Carnap, R.: Logical foundations of probability (1962)
20. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: CRISP-DM 1.0: Step-by-Step Data Mining Guide. CRISP-DM consortium: NCR Systems Engineering, DaimlerChrysler AG, SPSS Inc. and OHRA Verzekeringen en Bank Groep B.V (2000)
21. De Raedt, L., Kersting, K.: Statistical Relational Learning. In: Encyclopedia of Machine Learning, pp. 916–924. Springer (2011)
22. Dou, D., Wang, H., Liu, H.: Semantic Data Mining: A Survey of Ontology-based Approaches. In: IEEE ICSC. pp. 244–251. IEEE (2015)
23. Duivesteijn, W., Knobbe, A., Feelders, A., van Leeuwen, M.: Subgroup Discovery Meets Bayesian Networks–An Exceptional Model Mining Approach. In: Proc. International Conference on Data Mining (ICDM). pp. 158–167. IEEE, Washington, DC, USA (2010)
24. Duivesteijn, W., Feelders, A., Knobbe, A.J.: Different Slopes for Different Folks: Mining for Exceptional Regression Models with Cook's Distance. In: Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 868–876. ACM, New York, NY, USA (2012)
25. Duivesteijn, W., Feelders, A.J., Knobbe, A.: Exceptional Model Mining. Data Min. Knowl. Disc. 30(1), 47–98 (2016)
26. Duivesteijn, W., Thaele, J.: Understanding Where Your Classifier Does (Not) Work – The SCaPE Model Class for EMM. In: Proc. ICDM. pp. 809–814. IEEE (2014)
27. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery: An Overview. In: Advances in Knowledge Discovery and Data Mining, pp. 1–34. AAAI Press (1996)
28. Gamberger, D., Lavrac, N., Wettschereck, D.: Subgroup Visualization: A Method and Application in Population Screening. In: Proc. IDAMAP (2002)
29. Gaskin, C.J., Happell, B.: On Exploratory Factor Analysis: A Review of Recent Evidence, an Assessment of Current Practice, and Recommendations for Future Use. International Journal of Nursing Studies 51(3), 511–521 (2014)
30. Goethals, B., Moens, S., Vreeken, J.: MIME: A Framework for Interactive Visual Pattern Mining. In: Proc. ACM SIGKDD. pp. 757–760. ACM (2011)
31. Goodman, B., Flaxman, S.: European union regulations on algorithmic decision-making and a" right to explanation". arXiv preprint arXiv:1606.08813 (2016)
32. Guidotti, R., Monreale, A., Turini, F., Pedreschi, D., Giannotti, F.: A Survey of Methods for Explaining Black Box Models. arXiv preprint arXiv:1802.01933 (2018)

33. Henelius, A., Puolamäki, K., Boström, H., Asker, L., Papapetrou, P.: A Peek Into the Black Box: Exploring Classifiers by Randomization. Data Mining and Knowledge Discovery 28(5-6), 1503–1529 (2014)
34. Henelius, A., Puolamäki, K., Karlsson, I., Zhao, J., Asker, L., Boström, H., Papapetrou, P.: Goldeneye++: A Closer Look Into The Black Box. In: International Symposium on Statistical Learning and Data Sciences. pp. 96–105. Springer (2015)
35. Henelius, A., Puolamäki, K., Ukkonen, A.: Interpreting Classifiers through Attribute Interactions in Datasets. Proc. 2017 ICML Workshop on Human Interpretability in Machine Learning (WHI 2017) (2017)
36. Hoffart, J., Suchanek, F.M., Berberich, K., Weikum, G.: YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia. Artificial Intelligence 194, 28–61 (2013)
37. Jaroszewicz, S., Simovici, D.A.: Interestingness of Frequent Itemsets using Bayesian Networks as Background Knowledge. In: Proc. SIGKDD. pp. 178–186. ACM (2004)
38. Kaytoue, M., Plantevit, M., Zimmermann, A., Bendimerad, A., Robardet, C.: Exceptional Contextual Subgraph Mining. Machine Learning 106(8), 1171–1211 (2017)
39. Keim, D., Ward, M.: Intelligent Data Analysis, chap. 11: Visualization. Springer, Heidelberg, Germany (2003)
40. Klösgen, W.: Explora: A Multipattern and Multistrategy Discovery Assistant. In: Advances in Knowledge Discovery and Data Mining, pp. 249–271. AAAI Press (1996)
41. Klösgen, W.: Handbook of Data Mining and Knowledge Discovery, chap. 16.3: Subgroup Discovery. Oxford University Press, New York (2002)
42. Klösgen, W., Lauer, S.R.W.: Handbook of Data Mining and Knowledge Discovery, chap. 20.1: Visualization of Data Mining Results. Oxford University Press, New York (2002)
43. Knobbe, A.J., Cremilleux, B., Fürnkranz, J., Scholz, M.: From Local Patterns to Global Models: The LeGo Approach to Data Mining. In: From Local Patterns to Global Models: Proceedings of the ECML/PKDD-08 Workshop (LeGo-08). pp. 1 – 16 (2008)
44. Kolodner, J.L.: Making the Implicit Explicit: Clarifying the Principles of Case-based Reasoning. Case-based Reasoning: Experiences, Lessons & Future Directions pp. 349–370 (1996)
45. Lavrac, N.: Subgroup Discovery Techniques and Applications. In: Ho, T.B., Cheung, D.W., Liu, H. (eds.) Proc. 9th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD05), LNAI, vol. 3518. Springer, Heidelberg, Germany (2005)
46. Lavrac, N., Kavsek, B., Flach, P., Todorovski, L.: Subgroup Discovery with CN2-SD. Journal of Machine Learning Research 5, 153–188 (2004)
47. Leman, D., Feelders, A., Knobbe, A.: Exceptional Model Mining. Machine Learning and Knowledge Discovery in Databases pp. 1–16 (2008)
48. Lemmerich, F., Becker, M., Atzmueller, M.: Generic Pattern Trees for Exhaustive Exceptional Model Mining. In: Proc. ECML/PKDD. Springer, Heidelberg, Germany (2012)
49. Li, W., Han, J., Pei, J.: CMAR: Accurate and Efficient Classification Based on Multiple Class-Association Rules. In: Cercone, N., Lin, T.Y., Wu, X. (eds.) Proc. International Conference on Data Mining (ICDM). pp. 369–376. IEEE Computer Society (2001)
50. Li, X., Huan, J.: Constructivism Learning: A Learning Paradigm for Transparent Predictive Analytics. In: Proc. SIGKDD. pp. 285–294. ACM (2017)
51. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A Symbolic Representation of Time Series, with Implications for Streaming Algorithms. In: Proc. 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. pp. 2–11. ACM (2003)
52. Lin, J., Keogh, E., Wei, L., Lonardi, S.: Experiencing SAX: A Novel Symbolic Representation of Time Series. Data Mining and Knowledge Discovery 15(2), 107–144 (2007)
53. Liu, B., Hsu, W., Ma, Y.: Integrating Classification and Association Rule Mining. In: Proc. SIGKDD. pp. 80–86. AAAI Press (August 1998)
54. Maher, P.: Explication Defended. Studia Logica 86(2), 331–341 (2007)

55. Mandel, D.R.: Counterfactual and Causal Explanation: From Early Theoretical Views To New Frontiers. In: The Psychology of Counterfactual Thinking, pp. 23–39. Routledge (2007)
56. Mitzlaff, F., Atzmueller, M., Hotho, A., Stumme, G.: The Social Distributional Hypothesis. Journal of Social Network Analysis and Mining 4(216), 1–14 (2014)
57. Mitzlaff, F., Atzmueller, M., Stumme, G., Hotho, A.: Semantics of User Interaction in Social Media. In: Ghoshal, G., Poncela-Casasnovas, J., Tolksdorf, R. (eds.) Complex Networks IV, Studies in Computational Intelligence, vol. 476. Springer, Heidelberg, Germany (2013)
58. Morik, K.: Detecting Interesting Instances. In: Hand, D., Adams, N., Bolton, R. (eds.) Pattern Detection and Discovery, LNCS, vol. 2447, pp. 13–23. Springer Berlin Heidelberg (2002)
59. Morik, K., Boulicaut, J., Siebes, A. (eds.): Local Pattern Detection, International Seminar, Dagstuhl Castle, Germany, April 12-16, 2004, Revised Selected Papers, Lecture Notes in Computer Science, vol. 3539. Springer (2005)
60. Morshed, A., Dutta, R., Aryal, J.: Recommending Environmental Knowledge as Linked Open Data Cloud using Semantic Machine Learning. In: Proc. IEEE ICDEW. pp. 27–28. IEEE (2013)
61. Musto, C., Narducci, F., Lops, P., de Gemmis, M., Semeraro, G.: Linked Open Data-Based Explanations for Transparent Recommender Systems. International Journal of Human-Computer Studies (2018)
62. Paulheim, H.: Explain-a-LOD: Using Linked Open Data for Interpreting Statistics. In: Proc. ACM IUI. pp. 313–314. ACM (2012)
63. Paulheim, H.: Generating Possible Interpretations for Statistics from Linked Open Data. In: Extended Semantic Web Conference. pp. 560–574. Springer (2012)
64. Paulheim, H., Fümkranz, J.: Unsupervised Generation of Data Mining Features from Linked Open Data. In: Proc. WIMS. p. 31. ACM (2012)
65. Pujara, J., Miao, H., Getoor, L., Cohen, W.: Large-scale Knowledge Graph Identification using PSL. In: AAAI Fall Symposium on Semantics for Big Data (2013)
66. Rauch, J., Simunek, M.: Learning Association Rules from Data through Domain Knowledge and Automation. In: International Workshop on Rules and Rule Markup Languages for the Semantic Web. pp. 266–280. Springer (2014)
67. Ribeiro, M.T., Singh, S., Guestrin, C.: Model-Agnostic Interpretability of Machine Learning. Proc. 2016 ICML Workshop on Human Interpretability in Machine Learning (2016)
68. Ribeiro, M.T., Singh, S., Guestrin, C.: Why Should I Trust You?: Explaining the Predictions of Any Classifier. In: Proc. ACM SIGKDD. pp. 1135–1144. ACM (2016)
69. Ribeiro, M.T., Singh, S., Guestrin, C.: Anchors: High-Precision Model-Agnostic Explanations. AAAI (2018)
70. Richardson, M., Domingos, P.: Learning with Knowledge from Multiple Experts. In: Proc. ICML. pp. 624–631. AAAI Press (2003)
71. Richardson, M., Domingos, P.: Markov Logic Networks. Machine learning 62(1-2), 107–136 (2006)
72. Ristoski, P., Paulheim, H.: Semantic Web in Data Mining and Knowledge Discovery: A Comprehensive Survey. Web Semantics 36, 1–22 (2016)
73. Roth-Berghofer, T., Schulz, S., Leake, D., Bahls, D.: Explanation-Aware Computing. AI Magazine 28(4) (2007)
74. Roth-Berghofer, T.R., Cassens, J.: Mapping Goals and Kinds of Explanations to the Knowledge Containers of Case-Based Reasoning Systems. In: Case-Based Reasoning Research and Development, Proc. ICCBR. pp. 451–464. No. 3620 in LNAI, Springer, Heidelberg (2005)
75. Seipel, D., Nogatz, F., Abreu, S.: Domain-Specific Languages in Prolog for Declarative Expert Knowledge in Rules and Ontologies. Computer Languages, Systems & Structures 51, 102–117 (2018)

76. Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: Proc. IEEE Symposium on Visual Languages. pp. 336–343. Boulder, Colorado (1996)
77. Sørmo, F., Cassens, J., Aamodt, A.: Explanation in case-based reasoning – perspectives and goals. Artificial Intelligence Review 24(2), 109–143 (2005)
78. Spenke, M.: Visualization and Interactive Analysis of Blood Parameters with InfoZoom. Artificial Intelligence In Medicine 22(2), 159–172 (2001)
79. Spenke, M., Beilken, C.: Visual, Interactive Data Mining with InfoZoom – The Financial Dataset. In: Workshop Notes on Discovery Challenge at the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases. pp. 15–18 (1999)
80. Spieker, P.: Natürlichsprachliche Erklärungen in technischen Expertensystemen. Dissertation, University of Kaiserslautern (1991)
81. Thabtah, F.: A Review of Associative Classification Mining. Knowl. Eng. Rev. 22(1), 37–65 (Mar 2007)
82. Theus, M.: Interactive Data Visualization using Mondrian. Journal of Statistical Software 7(11), 1–9 (2003)
83. Tiddi, I., d'Aquin, M., Motta, E.: An Ontology Design Pattern to Define Explanations. In: Proc. K-Cap. ACM, New York, NY, USA (2015)
84. Tolomei, G., Silvestri, F., Haines, A., Lalmas, M.: Interpretable predictions of tree-based ensembles via actionable feature tweaking. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 465–474. ACM (2017)
85. Van Deursen, A., Klint, P., Visser, J.: Domain-Specific Languages: An Annotated Bibliography. ACM Sigplan Notices 35(6), 26–36 (2000)
86. Van Leeuwen, M.: Interactive Data Exploration Using Pattern Mining. In: Interactive Knowledge Discovery and Data Mining in Biomedical Informatics, pp. 169–182. Springer (2014)
87. Vavpetič, A., Lavrač, N.: Semantic Subgroup Discovery Systems and Workflows in the SDM-Toolkit. The Computer Journal 56(3), 304–320 (2013)
88. Vavpetic, A., Podpecan, V., Lavrac, N.: Semantic Subgroup Explanations. J. Intell. Inf. Syst. 42(2), 233–254 (2014)
89. Velicer, W.F., Eaton, C.A., Fava, J.L.: Construct Explication Through Factor or Component Analysis: A Review and Evaluation of Alternative Procedures for Determining the Number of Factors or Components. Problems and Solutions in Human Assessment: Honoring Douglas N. Jackson at Seventy pp. 41–71 (2000)
90. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR (2017)
91. Wick, M.R., Thompson, W.B.: Reconstructive Expert System Explanation. Artif. Intell. 54(1-2), 33–70 (1992)
92. Wilcke, X., Bloem, P., de Boer, V.: The Knowledge Graph as the Default Data Model for Learning on Heterogeneous Knowledge. Data Science (Preprint), 1–19 (2017)
93. Wirth, R., Hipp, J.: CRISP-DM: Towards a Standard Process Model for Data Mining. In: Proc. 4th Intl. Conference on the Practical Application of Knowledge Discovery and Data Mining. pp. 29–39. Morgan Kaufmann (2000)
94. Wrobel, S.: An Algorithm for Multi-Relational Discovery of Subgroups. In: Proc. PKDD. pp. 78–87. Springer (1997)
95. Zelezny, F., Lavrac, N., Dzeroski, S.: Using Constraints in Relational Subgroup Discovery. In: Proc. International Conference on Methodology and Statistics. pp. 78–81. University of Ljubljana (2003)