Ontology
Machine Learning
Data Mining





# 49. Ontologies and Machine Learning Systems

Shoba Tegginmath, [illegible]el Pears, Nikola Kasabov

In this chapter we review the uses of ontologies within bio[illegible]matics and the various attempts to combine M[illegible]d ontologies, and the uses of data mining ontologies. This is a diverse field and there is enormous potential for wider use of ontologies in bioinformatics. A systems biology approach comprising of experimental and computational research using biological, medical, and clinical data is needed to understand complex biological processes and help scientists draw meaningful inferences and to answer questions scientists have not even attempted so far.[TS0]



## 49.1 ~~Introduction~~

[TS1] In modern computer science, ontology is a data model that represents knowledge within a domain (a part of the world), providing a common understanding about the type of objects and concepts that exist in the domain. The biological sciences are replete with descriptive terms, and ontologies are useful here to reach a common understanding of these terms. The explicit definition of concepts in an ontology supports the sharing and reuse of formally represented knowledge among systems [49.1] and helps people and machines to communicate semantically, and not just syntactically [49.2].

Data mining (DM) is a part of the larger overall process of knowledge discovery; it is the process of discovering meaningful patterns in data [49.3] and as such is wholly relevant to the field of biology with the massive quantities and types of data available. Ontologies and DM work in parallel to identify and formalize knowledge – ontologies help in expressing the knowledge in a meaningful way while DM and machine learning (ML) help in extracting useful knowledge from data.

The current use of ontologies in biomedical sciences, however, is limited. The term ontology is used loosely in biomedicine and refers to a number of artifacts – such as controlled vocabularies, terminologies, and ontologies [49.4]. Ontologies have largely been used to facilitate interoperability among the various databases that contain datasets of biomedical experiments by indexing databases with standard terms to help locate and retrieve information. Ontologies have also been used for storing microarray experiment results and data; the microarray gene expression data ontology provides the common terminology and structure used for microarray experiments.

## 49.2 Ontologies in Bioinformatics

Gene ontology (GO) [49.4] is a preeminent example of the most common usage of ontologies in bioinformatics. GO provides a set of controlled, structured vocabularies to describe key domains of molecular and cellular biology and biological processes, including gene product attributes and biological sequences.

[TS0] Please provide index entries.
[TS1] The name “Introduction” is not allowed for the first section. Please supply a different name.

Editor’s or typesetter’s annotations (will be removed before the final TeX run)



SPIN: 12742608 (Springer Handbook of Springer Handbook of Nanomaterials)
MS ID: hb23-049 Proof 1
Created on: 30 January 2013 17:12 CET

Brain-Gene ontology
Ontology!Biomedical
Protege





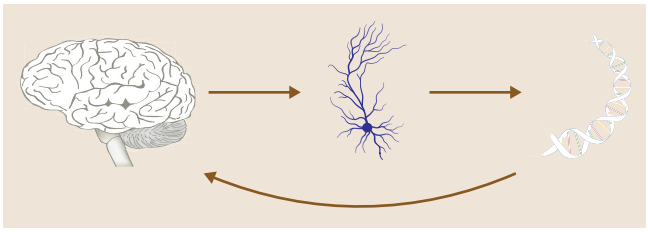

**Fig. 49.1** BGO is concerned with the relationships between the brain and genes TS2

Each ontology is structured in a classification that supports *is-a* and *part-of* relationships. The control in this controlled vocabulary arises from the commitment to use that ontology delivered vocabulary to describe attributes of classes of gene products in community-wide resources [49.5]. Collaborating databases provide data sets with links between database objects and GO terms – these are called annotations. A GO annotation is a link between a gene product type and a molecular function, biological process, or cellular component type. These annotations are what make the ontology useful and able to support computational reasoning about the instances and GO terms. Observations from experiments and inferences drawn from such experiments are used to create annotations.

The brain-gene ontology (BGO) [49.6] is a biomedical ontology that integrates information from different disciplines such as neuroscience, bioinformatics, genetics, and computer and information sciences. BGO is focused on the gene-disease relationship and includes various concepts, facts, data, software simulators, graphs, videos, animations, and other information forms related to brain functions, in diseases, their genetic basis, and the relationship between all of them.

BGO [49.7] has been implemented in the Protégé ontology building environment [49.8]. BGO is based on GO [49.4] and the Unified Medical Language System [49.9]. In addition, knowledge acquired from biology domain experts, from other biological data sources such as Entrez Gene, Swissprot, and Interpro, and from literature databases such as PubMed has also been incorporated.

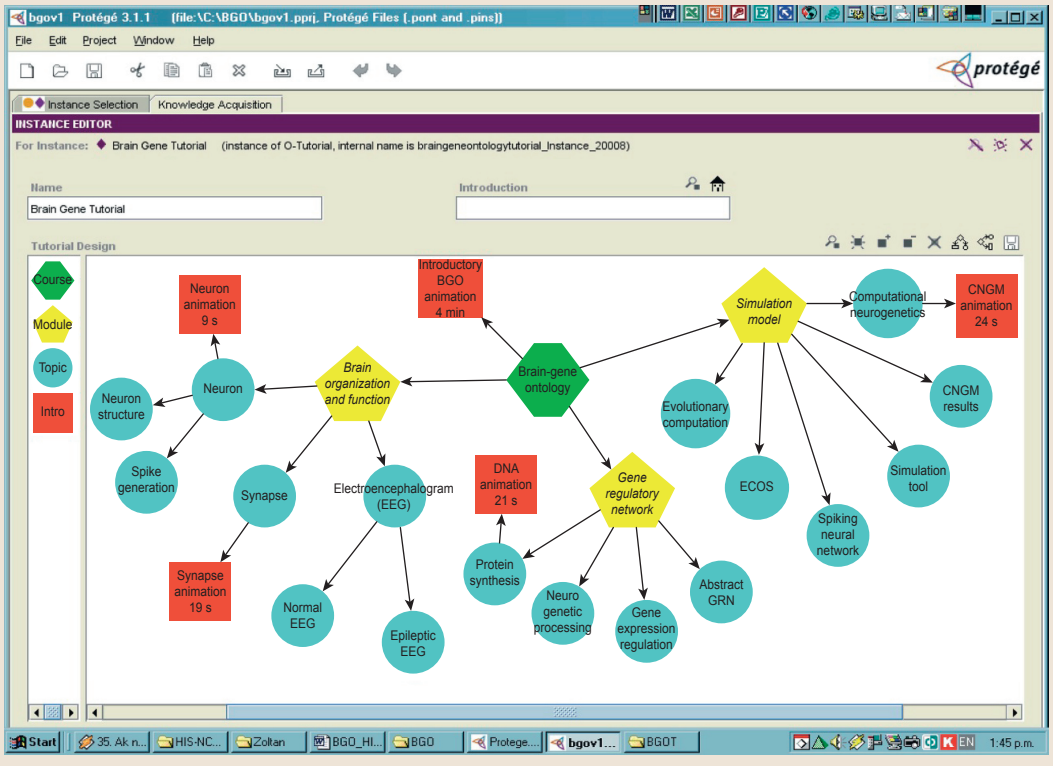


**Fig. 49.2** BGO information structure

TS2 Please refer to this figure in the text.

Editor's or typesetter's annotations (will be removed before the final TeX run)



SPIN: 1274260
MS ID: hb23-0
Created on: 30

**Table 49.1** General structure of CDO domains

| Organism domain | Molecular domain | Medical domain | Nutritional domain | Biomedical informatics map |
|---|---|---|---|---|
| Human | Gene | Disease | Nutrients | Disease gene map |
| Group | Mutation | Clinical findings | Source | |
| Population group | Protein | Signs | Function | |
| Patient group | | Symptoms | | |
| | | Laboratory tests | | |

The overall system as shown in Fig. 49.2 comprises three main parts:

1. Brain organization and functions, which contains information about neurons, their structure, the process of spike generation, and processes in synapses
2. Genes and gene regulatory networks (GRN) is divided into sections on neurogenetic processing, gene expression regulation, protein synthesis, and abstract GRNs
3. A simulation module that has sections on computational neurogenetic modeling, evolutionary computation, and evolving connectionist systems (ECOS).

ECOS [49.10] are modular connectionist-based systems that evolve their structure and functionality in a continuous, self-organized, adaptive way from incoming information, and are capable of processing both data and knowledge in a supervised or unsupervised manner.

One of the main applications of BGO is the integration between ontology and ML tools in relation to feature selection, classification, and prognostic modeling. Software machine learning environments such as NeuCom [49.11], WEKA [49.12], and Siftware [49.13] can be used to aid novel discoveries. By integrating results from ML with genetic information in BGO, a more complete understanding of the pathogenesis of brain diseases is facilitated [49.14]. However, the discoveries from ML environments are currently entered back to the BGO manually, which is untenable.

Chronic disease ontology [49.15] (CDO) is a Protégé-based ontology [49.8], which contains information about genes involved in three diseases and their mutations, health and nutrition information, and life history data. The diseases are the top three common chronic diseases in developed countries–cardiovascular disease, type2 diabetes, and obesity. These diseases are thought to be mainly caused by interactions of common factors such as genes, nutrition, and life-style. Five domains – organism domain, molecular domain, medical domain, nutritional domain, and a biomedical informatics map exist in CDO. These five classes contain further subclasses and instances as shown in Table 49.1. Each subclass has a set of slots which provides information about each instance and have relationships among other slots, instances, and concepts [49.16]. There are about 76 genes in the ontology. Each gene instance has diverse information associated with the gene and has relationships with other domains. The population group of the organism domain contains information on 50 different population groups; the patient group contains individual patient data.

CDO makes it possible to input information on individual patients such as symptoms, gene maps, diet, and life history details, and generate a personalized

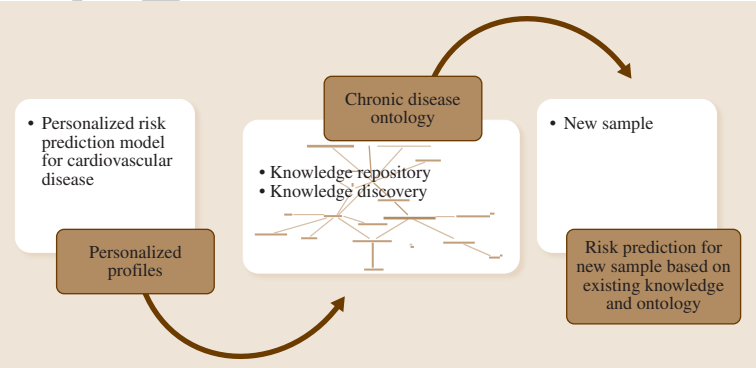


**Fig. 49.3** Utilization of knowledge within the CDO



SPIN: 12742608 (Springer Handbook of Springer Handbook of Nanomaterials)
MS ID: hb23-049 Proof 1
Created on: 30 January 2013 17:12 CET

model of risks, profiles, and recommendations, based on the genes of interest and relevant diet components identified, as shown in Fig. 49.3. Large datasets can be imported into the ontology using the data-master plug-in while individual patient information, such as medical, genetic, clinical and nutritional, can be added manually. CDO has helped in identifying interrelationships in personalized risk evaluation for chronic diseases for individuals as well as for groups of individuals [49.1[illegible]]. A personalized model is created for a single person, based on their personal data and the information in the ontology. A transductive neuro-fuzzy inference system with weighted data normalization [49.18] is used to evaluate personalized risk.

Ontologies in the biological domain have also been used supporting text mining and information retrieval. *Khelif*, *Dieng-Kuntz*, and *Barbry* [49.19] discuss the knowledge available in textual documents and the efficient detection and use of this knowledge, which is a challenging task. Biologists need tools to support them in the interpretation and validation of their experiments to facilitate future experiments. The authors propose an ontology-based approach for generation of semantic annotations and information retrieval and suggest that the proposed approach can probably be extended to other massive analysis of biological events. However, the authors do not present any experimental work. The proposed work is only suitable for microarray DNA experiment data. The proposed approach uses an annotation base and relations, and in certain cases these relations can be used interchangeably by the biologists and there is a gap between system-based ontology and user-based ontology integration. To bridge the gap, biologists are required to strengthen the system-based ontologies by providing regular feedback of their ontology usage. The contextual information in the text documents is also very important in designing the ontologies in biological domain. The major problem, however, is that there are a number of text mining techniques consisting of different operators, but as yet there is no generic ontology framework for text mining in the biological domain.

*Kuo*, *Lonie*, *Sonenberg*, and *Paizis* [49.20] report on using a domain ontology-driven approach to DM using clinical data of patients undergoing treatment for chronic kidney disease. The authors acknowledge the challenges in mining such a multiple attribute dataset; the attributes in the dataset useful for DM are not apparent without domain knowledge. The reported work explores the use of medical domain ontology as a source of domain knowledge in both extracting and expressing knowledge in a useful format. Domain ontology has been used to categorize attributes in preparation for mining association rules in the data. The authors report that domain ontology driven DM can obtain more meaningful results than naive mining. However, determining the meaningfulness of the results is not an easy task – a strong association rule does not guarantee a useful rule. Also, although domain ontology has been used to provide domain expert knowledge to guide the mining process, a domain expert is still required to gauge the usefulness of the rules/results.

## 49.3 Data Mining Specific Ontologies

In this section, we present some recent work in ontologies for the DM domain.

*Diamantini*, *Potena*, and *Storti* [49.21] discuss the highly complex, iterative, and interactive, goal-driven and domain-dependent nature of the knowledge discovery in databases (KDD) and the DM process. The effective design of KDD process, the continuous development of new algorithms, their many characteristics, and the applicability to different kinds of data are all considered and the authors discuss issues in designing a successful knowledge discovery process. The main contribution of the work is the design of KDDONTO, an ontology for supporting both the discovery of suitable KDD algorithms and the composition of the KDD process. An ontology building methodology has been proposed with the justification that previously proposed methodologies for ontology design were mainly borrowed from the software engineering field. There are a number of similarities between software engineering and ontology fields but the goals are totally different. For instance, the goal of software engineering is the development of implementable classes while the knowledge engineer looks for the formal representation of the domain. The proposed methodology satisfies the quality requirements of the formal ontology such as coherence, clarity, extensibility, minimization of encoding bias, and minimization of ontological commitment. For KDDONTO implementation, the authors have chosen the OWL-DL language. However, there are some issues yet to be resolved due to the expressive semantics re-



SPIN: 1274260 ook of Springer Handbook of Nanomaterials)
MS ID: hb23-0
Created on: 30 2 CET

strictions in the properties of data and model classes. At present the KDDONTO implementation is formed of 95 classes, 31 relations, and more than 140 instances. Each step of the proposed methodology returns as output a valid ontology represented in a different language. The existing limitation of the work is that the KDD process composition using the proposed methodology is not automatic. However, authors plan to report the semi-automatic process composition in the future.

In order to help data miners consider the vast design space of possible processes thoroughly, *Bernstein*, *Provost*, and *Hill* [49.22] present an intelligent discovery assistant (IDA) that works with an explicit ontology of DM techniques to navigate space and present the user with a systematic enumeration of valid DM processes. IDA considers the characteristics of the data, the desired results, and works with the ontology to search for and enumerate valid plans to achieve the desired results. The processes may be ranked on criteria such as speed, accuracy, model comprehensibility, etc. The strength of the work is that the authors discuss how the proposed IDA tool can play an important role in knowledge sharing in the team of data miners. However, there are a few limitations in the work such as the fact that IDA provides a large number of valid plans to choose from but does not provide assistance in recommending a combination of processes to assist the user at this stage. Furthermore, there is no indication of which particular induction algorithm will be better to choose based on heuristic knowledge. IDA also lacks the incorporation of the characteristics of a dataset, which is a very important aspect before applying any DM technique. The ontology designed in this work is lightweight and does not contain the internal structure of DM operators [49.23]. The major gap is in the identification of interesting DM processes, which we think should be the primary objective of applying any DM technique.

Based on the need for a unifying framework for DM and an as yet unfulfilled proposal to describe the hierarchies of DM operations in terms of their signatures [49.24], *Panov*, *Dzeroski*, and *Soldatova* [49.23] put forward the idea of a DM ontology called OntoDM for the description of the DM domain. The authors' aim is to design OntoDM with sound theoretical foundations. They argue that ontologies proposed thus far are light-weight ontologies which can be easily developed and meet the specific purpose for which they were created but do not follow good practices of ontology development. They tend to be shallow, with no rigid relations between defined entities. The DM domain, unlike other domains, requires detailed inference over data and, therefor, a rigid, heavy weight ontology is needed. In contrast to the work of *Bernstein* et al. [49.22], the proposed OntoDM covers the details of basic DM entities such as data types, datasets, DM tasks, and ~~the DM algorithm and components~~ CE[3]. OntoDM follows the philosophy of OBI (ontology for biomedical investigations) and EXPO ontology of scientific experiments. OBI's top-level ontology BFO (basic formal ontology) is used to define upper level classes. OBO RO (relational ontology) is also used to define the semantics of the relationships between the DM entities. The methodology is developed keeping in mind the complex entities and more popular research areas such as constraint based mining. The proposed ontology consists of three main components: classes, a hierarchical structure of classes, and relations between classes. The basic entities considered are based on a framework proposed by one of the authors in a prior work. The entities describe the different orthogonal dimensions of DM, and it is the authors' intention that different combinations of these can be used to describe ~~most of the present~~ DM approaches CE[4]. The entities are: dataset, data type, DM tasks, generalizations, DM algorithm, and components of DM algorithms such as distance functions, kernel functions, and features. Furthermore, it provides for defining more complex entities such as constraints and DM tasks. We find that the claim that OntoDM covers all the basic entities of DM and can be used for complex DM tasks needs proof of concept and further evaluation. The rapidly growing field of DM is incorporating intense nature inspired algorithms that are different from the traditional DM approaches and algorithms. We believe that such advancement should also be covered in the proposed ontology. Furthermore, the use of previous knowledge and the sharing of information have not been covered in the proposed ontology, which is an important aspect of KDD.

In *Panov*, *Soldatova*, and *Dzeroski* [49.25] TS[5] the authors extend the work presented in [49.23] and present an updated version of OntoDM. The version described is updated in a number of ways. Alignment of the structure of the ontology with the top level structure of the OBI ontology introduced new entities in the ontology. For instance, the entity DM algorithm was split into three entities, each capturing a different aspect of the algorithm such as algorithm specification, algorithm implementation, and algorithm description. The set of relations used in the initial version is extended with relations defined in OBI ontology in order to express the relations between informational entities or entities

CE[3] Please check that this is the intended meaning.
CE[4] Please check that this is the intended meaning.
TS[5] Please check reference.

Editor's or typesetter's annotations (will be removed before the final TeX run)



SPIN: 12742608 (Springer Handbook of Springer Handbook of Nanomaterials)
MS ID: hb23-049 Proof 1
Created on: 30 January 2013 17:12 CET

that are realized in a process and processes. The authors have also extended the OBI classes with DM specific classes for defining complex entities such as DM scenarios and queries. However, the ontology presented is still in the early stages of development. There is a need to populate the proposed classes of DM entities with individuals and to refine the structure of OntoDM as needed in order to cover the various aspects of the DM domain.

An illustrative example of the use of ontology for DM has been provided in [49.26], which reports on an ontology-based DM system for discovering knowledge from incomplete data and the effectiveness of ontology in knowledge management. The results of applying ontologies for DM with incomplete data in a classroom environment are reported. The limitation is that a very simple example has been used for the demonstration of ontology development. As mentioned previously, there is a strong need for ontology development for complex DM tasks and procedures.

*Hilario*, *Kalousis*, *Nguyen*, and *Woznica* [49.27] present their vision of a DM ontology designed to support meta-learning for algorithm and model selection. The authors argue that previous research has focused on aligning experiments and performance metrics and not much work has been done on explaining observations in terms of the internal logic and mechanism of learning algorithms. The authors extended the previously proposed Rice model, which related dataset descriptions to performance of algorithms, by adding algorithm features to dataset features as a parameter of algorithm selection. The key components of an algorithm include the structure and parameters of the models produced, the cost function used to quantify the appropriateness of a model, and the optimization strategy adopted to find the model parameter values that minimize the cost function. The authors discuss the next steps for the continuation of the proposed work. The first step is to gather interested data miners and ontology engineers to consolidate the core concepts of the DM ontology proposed, and the next step is to show how DM ontology can be used to improve algorithm selection through meta learning. However, there are still some ongoing issues in the proposed work. DM research has identified other components of bias for learning algorithms in addition to those described in this paper.

## 49.4 Discussion and Future Work

It is generally recognized that a standardized ontology framework makes data easily available for advanced methods of analysis and artificial intelligence algorithms. In the biomedical domain, GO and the subsequent OBO consortium have been instrumental in allowing the community of scientists to speak the same language, add to the knowledge by creating annotations, and use this in drawing inferences. The interested reader is referred to a detailed review [49.28] of trends in biomedical ontologies. The potential for wider use of ontologies in bioinformatics will be realized with greater use of ML methods.

Our review also found also that the use of ontologies with DM is either in a specific domain of interest, such as bioinformatics, or in the DM domain itself. There are ongoing research efforts both in the construction of light-weight DM ontologies and in the construction of top-level DM ontologies. Neither meet the needs of the KDD community entirely. Light-weight DM ontologies meet the specific purpose they were created for but do not follow good practices of ontology development. General purpose, top-level DM ontologies are not conceived for achieving specific support requirements, like discovery of algorithms and process composition [49.21]. While such a general purpose top-level ontology is meant to be useful in supporting different activities it ends up providing inefficient support ~~in each~~ CE ontologies, with their high level of abstraction, suffer as their construction and usage have been decoupled [49.29]. We believe that a coupling or integration of these two types of ontologies will open up new grounds for potentially rich areas of research into DM ontologies. Such integration is capable of providing excellent solutions to treat complex data and sophisticated DM methods and algorithms.

In the DM domain, ontologies are used to detect patterns in data, and further to retrieve facts or information. Knowledge in the ontology may be used to deduce features from the ontology, which help modify classical feature representations. We believe more work is required in enhancing an ML model with knowledge from ontology, resulting in a richer model.

We propose to develop a wholly-integrated ontology system capable of evaluating newly discovered knowledge and evolving in a recurring, automatic manner. The integrated ontology system will be composed of

CE6 Please check: each what?

Editor's or typesetter's annotations (will be removed before the final TeX run)



SPIN: 1274260 ook of Springer Handbook of Nanomaterials)
MS ID: hb23-0
Created on: 30 2 CET

ML methods, ontology, and associated knowledge base. To the best of our knowledge, there is no methodology that has dealt with these two types of ontologies in a single framework. This integration will ensure that usage will continuously add to the data and semantics of the ontology. For this to be realized a major challenge pertains to resolving the process of ontology augmentation – the knowledge, or additional facts, need to be confirmed as improving the accuracy of predictions on new data before they can be acknowledged as conceptual changes (new concepts or relationships), or as explication changes (changes to existing concepts or relationships) to the ontology [49.14]. We propose to use BGO in this exploration.

## 49.5 Summary

This chapter has reviewed current research in bioinformatics ontologies and found that, in general, biomedical ontologies are prolific and current ML efforts reuse existing ontologies in order to support text mining and/or other ML efforts. Research has concentrated on the solution of domain-specific problems or DM issues. We emphasize a need for a methodology that deals with the two types of ontologies in a single framework. Such an integrated solution can improve the analytical powers of modern DM systems. Our proposal for an integrated environment will also aid the coupling of construction and usage of ontologies and continued refinement, where usage will continuously add to the data and semantics of the ontology for significant advancement in computing research. Enhancement the decision-making process based on both ontology and ML is a rich and exciting area with substantial potential.

### References


49.1 B. Chandrasekaran, J.R. Josephson, V.R. Benjamins: What are ontologies, and why do we need them?, Intell. Syst. Appl. **14**, 20–26 (1999)

49.2 A. Maedche, B. Motik, L. Stojanovic, R. Studer, R. Volz: Ontologies for enterprise knowledge management, Intell. Syst. IEEE **18**(2), 26–33 (2003)

49.3 I.H. Witten, E. Frank, M.A. Hall: *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, Burlington 2011)

49.4 M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock: Gene ontology: Tool for the unification of biology, Nat. Genet. **25**(1), 25–29 (2000)

49.5 Stevens and Lord (2009) TS

49.6 www.kedri.org TS

49.7 Brain-gene ontology, current version can be downloaded from http://www.kedri.info TS

49.8 Protégé http://protege.stanford.edu TS

49.9 Unified Medical Language System http://www.nlm.nih.gov/research/umls/ TS

49.10 N. Kasabov: *Evolving Connectionist Systems The Knowledge Engineering Approach*, 2nd edn. (Springer, Berlin, Heidelberg 2007) p. 451

49.11 NeuCom http://www.theNeuCom.com TS

49.12 WEKA http://www.cs.waikato.ac.nz/ml/weka TS

~~49.13 Siftware http://www.peblnz.com~~ TS

49.14 ~~N. Kasabov, V. Jain, L. Benuskova: Integrating evolving brain-gene ontology and connectionist-based system for modelling and knowledge discovery, Neural Netw. **2**, 266–275 (2008)~~

~~49.15 www.kedri.info~~ TS

49.16 A. Verma, N. Kasabov, E. Rush, Q. Song: Ontology based personalized modeling for chronic disease risk analysis: An integrated approach. In: TS LNCS, Vol. 5506 (Springer, Heidelberg 2008) pp. 1204–1210

~~49.17 A. Verma, M. Fiasche, M. Cuzzola, P. Iacopino, F.C. Morabito, N. Kasabov: Ontology based personalized modeling for type 2 diabetes risk analysis: An integrated approach. In:~~ TS ~~, LNCS, Vol. 5864~~ (Springer, Heidelberg 2009) pp. 360–366

49.18 Q. Song, N. Kasabov: TWNFI – a transductive neuro-fuzzy inference system with weighted data normalization for personalized modeling, Neural Netw. **19**, 1556–1591 (2006)

49.19 K. Khelif, R. Dieng-Kuntz, P. Barbry: An ontology-based approach to support text mining and information retrieval in the biological domain, J. Univers. Comput. Sci. **13**(12), 1881–1907 (2007)

49.20 Y. Kuo, A. Lonie, L. Sonenberg, K. Paizis: *Domain Ontology Driven Data Mining: A Medical Case Study*, ACM SIGKDD Workshop on Domain Driven DATA MINING (DDDM2007) (ACM, San Jose 2007)

49.21 C. Diamantini, D. Potena, E. Storti: KDDONTO: An ontology for discovery and composition of KDD algorithms, Third Generation Data Mining: To-


TS7 *Please supply more information.*
TS9 *Please supply the chapter/article/book title.*
TS9 *Please supply the chapter/article/book title.*

Editor's or typesetter's annotations (will be removed before the final TeX run)

**SPIN:** 12742608 (Springer Handbook of Springer Handbook of Nanomaterials)
**MS ID:** hb23-049 Proof 1
**Created on:** 30 January 2013 17:12 CET

wards Service-Oriented Knowledge Discovery 19–24 (2009)

49.22 A. Bernstein, F. Provost, S. Hill: Toward intelligent assistance for a DATA MINING process: An ontology-based approach for cost-sensitive classification, IEEE Trans. Knowl. Data Eng. **17**(14), 503–518 (2005)

49.23 P. Panov, S. Dzeroski, L. Soldatova: OntoDM: An ontology of Data Mining, IEEE Int. Conf. DATA MINING Workshops (IEEE, Washington 2008) pp. 752–760

49.24 R. Ramakrishnan, R. Agrawal, J.-C. Freytag, T. Bollinger, C.W. Clifton, S. Dzeroski, J. Hipp, D. Keim, S. Kramer, H.-P. Kriegel, U. Leser, B. Liu, H. Mannila, R. Meo, S. Morishita, R. Ng, J. Pei, P. Raghavan, M. Spiliopoulou, J. Srivastava, V. Torra: Data mining: The next generation, Perspectives Workshop: Data Mining: The Next Generation, number 04292, Dagstuhl Seminar Proc., ed. by R. Agrawal, J.C. Freytag, R. Ramakrishnan (Internationales Begegnungs- and Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl 2005)

49.25 TS10 P[illegible]: Towards an Ontology of Data Mining Investigations. In: *Discovery Science*, ed. by J. Gama (Springer Berlin, Heidelberg 2009) pp. 257–271

49.26 H. Wang, S. Wang: Ontology for data mining and its application to mining incomplete data, J. Database Manag. **19**(4), 81–90 (2008)

49.27 M. Hilario, A. Kalousis, P. Nguyen, W. Woznica: A DATA MINING ontology for algorithm selection and meta-mining, Third Generation Data Mining: Towards Service Oriented Towards Service-Oriented Knowledge Discovery (SoKD) (2009) p. 76

49.28 O. Bodenreider, R. Stevens: Bio-ontologies: Current trends and future directions, Brief. Bioinform. **7**(3), 256–274 (2006)

49.29 M. Hepp: Ontologies: State of the art, business potential, and grand challenges. In: *Data Management*, ed. by M. Hepp, P. De Leenheer, A. de Moor, Y. Sure (Springer, Berlin, Heidelberg 2007) pp. 3–24



SPIN: 1274260 ook of Springer Handbook of Nanomaterials)
MS ID: hb23-0
Created on: 30 2 CET



TS10 Please provide author's initials.

Editor's or typesetter's annotations (will be removed before the final TeX run)