ГОДИШНИК НА СОФИЙСКИЯ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА
Том 110

ANNUAL OF SOFIA UNIVERSITY “ST. KLIMENT OHRIDSKI”
FACULTY OF MATHEMATICS AND INFORMATICS
Volume 110

ENTERPRISE DATA AND SEMANTIC MODELLING: CONCEPTUAL MODEL OF INFORMATION TECHNOLOGY INCIDENT MANAGEMENT

KRISTINA ARNAOUDOVA and MARIA NISHEVA-PAVLOVA

Knowledge management methods and their efficient implementation across the organization determine sound and resilient management of processes. This paper studies the semantic integration of enterprise data sources essential to service management processes. Implementing a semantic layer within the enterprise architecture uses various tools, methods, and techniques. The semantic conceptual model unifies and implements intelligent integration of multiple data sources across the enterprise, achieving consistency and more accessible interpretation. Specifically, we draw our attention to incident and problem management within enterprises. We propose an ontology – a conceptual model for the incident management process. The incident ontology presented as an intelligent data integration layer component aims to achieve operational excellence. Besides, this ontology is a fundamental part of the proactive process in problem management. An ontology as a logic-based system supports integrity validation. It infers new, no explicitly modeled facts in the problem domain, thus helping experts better analyze and understand the problem. We discuss the conducted experiment results with the proposed in this article conceptual model using the enterprise knowledge graph platform. It can be perceived as a framework for a query-answering system with components, including ontology schema, data mapping, and classification methods for data graph enrichment.

Keywords: knowledge management, knowledge representation, ontology, semantic model, hybrid classification model, enterprise semantic layer, enterprise knowledge graph, incident management, ITIL, problem management

CCS Concepts:

- Information Systems~Information retrieval~Document representation~Ontologies;
- Information Systems~Information retrieval~Retrieval models and ranking;
- Information Systems~Information retrieval~Evaluation of retrieval results

DOI: 10.60063/GSU.FMI.110.25-36

1. Introduction

The business processes within an organization are of great variety. Their automation includes many applications. The enterprises are targeting improvement in many aspects, including better customer experience and manageable processes, where data integration is one of the main challenges. There are many approaches, platforms, and tools. With the significant increase in data volume, the complexity of the analysis and insight discoveries grows considerably, becoming an ultimate challenge. The centralized data store is one of the often-used components in the overall enterprise architecture. In many aspects, it is a costly approach that does not give results as expected. Vast methods and platforms are addressing this topic. Some ensure centralized storage and consistent reporting within the organization but are designed to solve specific tasks and reflect various aspects of applications, but not the business domain and the underlying relationships. Thus, it does not represent the problem domain and the concept's relations. A complementary approach is an intelligent integration using a virtual unified data layer.

The type of data assets in organizations, structured or unstructured, increases the difficulties in various tasks of the integration processes. The text data assets are a source of intensive knowledge. They are the primary source of additional information for the methods that might add other steps in the pipeline, like necessary preprocessing steps or concept mapping, and affirm challenges before information extraction. The lack of unified metadata or catalog leads to ambiguity and inconsistent automation. This paper presents a conceptual model essential for semantic data integration and apprehends the metadata for the data sources across one organization, allowing an ontology-based data access approach [18].

The particular domain is Information Technology Service Management (ITSM). The IT Infrastructure Library (ITIL) is an IT service management framework that outlines best practices for delivering IT services [3,17]. Implementing a good ITSM process as ITIL has numerous benefits for an organization [6]. The complementary component to the processes is the semantic layer – a technique to implement intelligent integration and interoperability. The specific process we are reviewing is Information Technology service incident management, a part of the processes and practices of Information Technology Infrastructure Library (ITIL) methodology, widely adopted in many enterprises. An incident and its resolution are critical for service reliability and organization credibility. Extracting and enriching organizational knowledge is a fundamental part of the pipeline of construction of our holistic view of corporate IT assets. Our model, an essential component in the semantic data layer, may contribute to operational excellence proactively.

2. Conceptual model

IT service management encompasses many practices, including the operational one of resolving incidents. According to ITIL Operation management, the incident is [3]: "An unplanned interruption to an IT Service or reduction in the quality of an IT Service. Failure of a configuration item that has not yet impacted the service configuration item is also an incident, for example, failure of one disk from the mirror. Incident management deals with all incidents; this can include failures, questions, or queries reported by the user by technical staff or automatically reported by the event".

Solving an incident that has interrupted service is critical and includes determining the failure, taking the right action, and restoring the service's regular work. Implementing a fast process of resolving incidents is an indispensable asset of an organization where the critical aspect is to include all possible sources of information. Gathering expertise about the incident is the expert responsibility, defining the incident, including essential attributes and a text description. The enrichment using other, typically from unstructured data, information extraction techniques could be valuable for more accurate first-level incident analysis.

2.1. Conceptual schema. Knowledge and data

The incident management process can benefit from semantically modeled data in finding new information, streamlining the resolution in a resilient and predictive way. The classical approach for enrichment is to infer new implicit facts by applying logical axioms embedded in the schema. Automatically deduce a taxonomy with subsumed classes with our categories or infer new relations among them. The solution can be found using expert knowledge or knowledge from already solved incidents stored in the database.

The priority assessment is among many steps in incident management, essential for the escalation process. An integral part of the process is the estimation of impact and urgency. The impact reflects how many configuration items such as storage, servers, or applications are affected, and the urgency reproduces how fast other configuration items will be affected.

Our proposed conceptual schema aims to categorize the incident according to different characteristics automatically. For example, the incident priority, impact, and urgency can be defined with appropriate axioms as a part of the conceptual schema. We can categorize according to business rules and components and consequently have better ideas that might complement determining possible causes.

A fundamental component of the data integration approach in the semantic data layer is the conceptual model, realized further as ontology. The proposed article may help implement a non-incident-centric strategy for operational excellence following the proactive understanding of the incident management process based on a semantic data graph. We have planned various experiments with the proposed model. We have used a knowledge graph platform as a design schema integration data tool, query answering, and database for them. The knowledge platforms are integration tools that combine different approaches for finding insights and have functionalities for semantical enrichments. They can link various data sources that could be extremely helpful in resolving incidents, including regulations and organization procedures.

Semantic network

Different descriptive formalisms for knowledge representation are organized as structural and formal languages. The Semantic Web is perceived as a semantic network of immense scale. The semantic networks as a knowledge representation structured formalism are at the heart of the described model. The semantic network is a structured formalism, and it is many variants John Sowa defines [2]: “A semantic network or net is a graph structure representing knowledge in patterns of interconnected nodes and arcs. Computer implementations of semantic networks were first developed for artificial intelligence and machine translation, but earlier versions have long been used in philosophy, psychology, and linguistics.” All semantic networks are graphically represented knowledge that supports automated systems for reasoning. Some versions are informal, but others are formally defined logic systems [2].

Ontology

The concepts and the relations among them should be unambiguously defined, and that is the main target of ontology as metadata about domain entities and relationships among them. Tom Gruber establishes the idea of ontology as [13] “an explicit specification of a conceptualization”, creating machine-readable explicit formalization of a domain. An ontology consists of concepts and relationships constrained by domain-specific business rules.

The knowledge is represented with formal languages; those based on first-order predicate calculus are RDF Schema (RDFS) [10] for concept, relations, and restriction, and the Ontology Web Language (OWL) [8] as an RDFS extension. The fundamental standard Resource Description Framework (RDF) [11] widely adopted a graph data model of objects and its presentation in subject-predicate-object triples. Applying the Shape Conceptual Language (SHACL) [12] enables distributed validation over data silos and rules described by OWL 2 RL [7] and SWRL [16].

OWL is probably the most used knowledge representation language. The OWL sublanguages OWL Lite, DL, and Full enable the knowledge engineer to balance expressiveness and semantic capabilities while ensuring efficient inference.

SHACL is an RDF-based language that can check for integrity constraints not globally as OWL 2 axioms. Instead, it can be performed for each data set. SHACL is appropriate for expressing data integrity, while the traditional data integrity constraints are global and thus not readily adopted to distributed data.

Simple Protocol and RDF (Resource Description Framework) Query Language SPARQL [14] can define queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF. The result of a SPARQL query can be a data set or an RDF graph.

The above techniques are used in various realizations for enterprise-wide usage, ensuring interoperability and state-of-the-art enterprise management processes.

2.2. Conceptual model

The ontology concepts, relations, and constraints are central to the proposed semantic model. As such, the conceptual schema design results from concentrated domain knowledge analysis. Our model follows the top-down approach based on methodology and human expertise. The high-level concepts and their relations are presented in Figure 1.

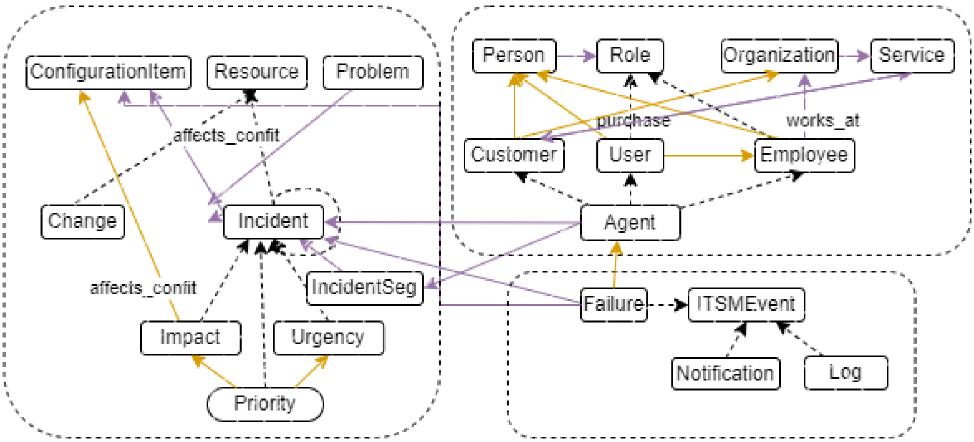

Figure 1. Incident conceptual model

2.2.1. Concepts

The main concepts, designed as owl classes, are *Request*, *Incident*, *Agent*, *Resource*, and *Configuration Item* examples: server, storage, application, service, and agents. The request for Resources could be for implementing a change, incident resolution, or service desk request. An example of an *Application as a subclass of Resources* is associated with its modules and architecture. An Agent is specialized as a person, and a person has different roles. The important roles in the incident process are client, employee, and user.

2.2.2. Ontology specification

OWL Class: inc:Incident
The class *Incident* includes different types of service interruption with the attributes describing the object.
Sub-class-of: inc:Request
Restriction: domain of object property comprise_seg

OWL Class: inc:Agent
The class *Agent* groups the agents or system who notifies of the service interruption.
Sub-class-of: owl:Thing
Restriction: domain of object property ann_incident

OWL Class: inc:Customer
The class *customer* includes those person or organization entities who have acquired a service or product from an organization.
Sub-class-of: owl:Role
Restriction: domain of object property *purchases*

OWL Class: inc:User
The user denotes persons who work at an organization or customers of an organization and are authorized to use the services of the organization.
Sub-class-of: owl:Role
Restriction: domain of object property register

OWL Class: inc:ISMEvent
The class *ITSMEvent* includes the system operation messages
sub-class-of: owl:Thing

OWL Class: inc:Role
The class *Role* includes different roles that a person or organization may have
sub-class-of: owl:Thing

OWL Class: inc:ConfigurationItem
The class *Configuration item* includes a comprehensive type of IT assets
sub-class-of: owl:Thing
Restriction: domain of object property subject_of

3. Enterprise knowledge graph

"Knowledge graph" is a term proposed by Google in 2012 as RDF graph data model, without details about its realization [18]. The knowledge graph may be described as a framework for realizing a data layer for querying information across various data sources, thus unifying different silos and data access across many sources. It may be implemented as a database or virtually with a transformed query. Its main component as a knowledge-based system is the data schema or ontology. It has embedded logic axioms, which could be formalized differently, performing constraints and inferring additional, not explicitly modeled concepts or relations. Statistical reasoning is another approach for enriching the knowledge graph in many knowledge platforms.

The pipeline is not fully automated but could be partially realized and enriched using semantic parsing, natural language understanding, machine learning, and logic inference. Different profiles ensuring different levels of expressiveness are adopted to manage the performance limitation of the reasoning engines. The semantic layer data's graph unifies distributed silos. The construction of a knowledge graph is supposed to have good scalability and performance, which is achieved with an on-demand inference or approximate one for knowledge graph storage virtualization and querying.

To validate the proposed conceptual mode for incident management, we have defined the ontology using OWL and the mapping procedure for conceptualizing the data. The processed experiments with the knowledge platform Stardog [15] based on a graph model that supports various formats and reasoning, including RDFS, OWL, SHACL, and SWRL. Following are some excerpts from the schema and screen with reasoning.

3.1. Graph data

We use data from an open data repository – the Incident management process enriched event log Data Set[1]. The dataset reflects information about incidents and the workflow for incident resolution, related agents, affected configuration items, and the different stages of the process. It is an event log of an incident management process extracted from data gathered from an instance of the ServiceNow™ platform. The event log is enriched with data loaded from a relational database underlying a corresponding process-aware information system. Information is anonymized for privacy. The instances are 141 712 events, but 24 918 incidents. The concept mapping is created by file with platform-specific formats associating the data row to the classes and relationships. The result is confirmed by the created incident nodes with attributes, configuration items, and other concepts.

3.2. Incident data graph

The primary type of objects are concepts, designed as owl classes, and are *Request*, *Incident*, *Agent*, *Resource*, and *Configuration Item.* The incident is instantiated from the customer, user, or system, and an incident announcement is an event using a specific component. We model the incidents as sequences of segments denoting the different steps in the incident resolution ontology, following the design patterns for ordering relationships [1]. For example, the introduction of the incident relates to opening the next step could be the investigation performed by the assigned expert, then closing the incident.

3.2.1. Priority and impact

The incident priority is formulated as a defined class, categorizing the incident in a subcategory, i.e., high-priority-incident. The incident is classified according to other parameters like business service affected, urgency, or impact. The priority formulation is in terms of impact and urgency, and the incident impact is defined in terms of affected configuration items. For example, the classification axioms infer a medium impact if the affected items are more than substantial value, which is specific to the organization.

[1] https://archive.ics.uci.edu/ml/datasets/Incident+management+process+enriched+event+log

3.3. Configuration item data graph

The location is identified as a physical place where a person or organization may reside or a virtual address. The configuration items in Figure 2 and Figure 3 are the operations assets and the configuration management database (CMDB) entities with the corresponding hardware, application, and services subclasses. The configuration items are the range of the affected configuration item relation.

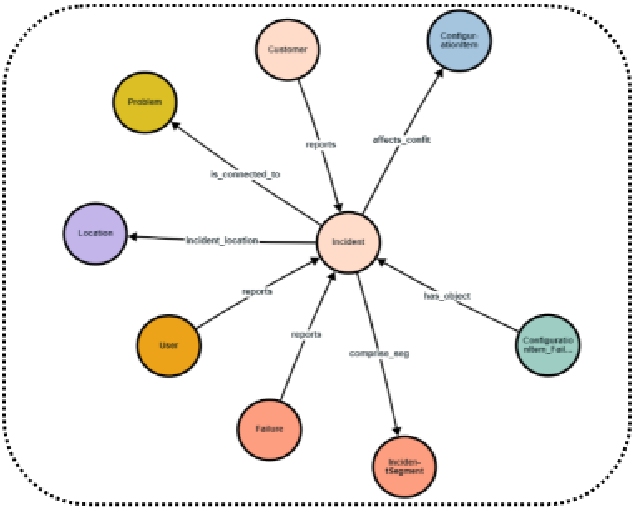

Figure 2. Incident schema

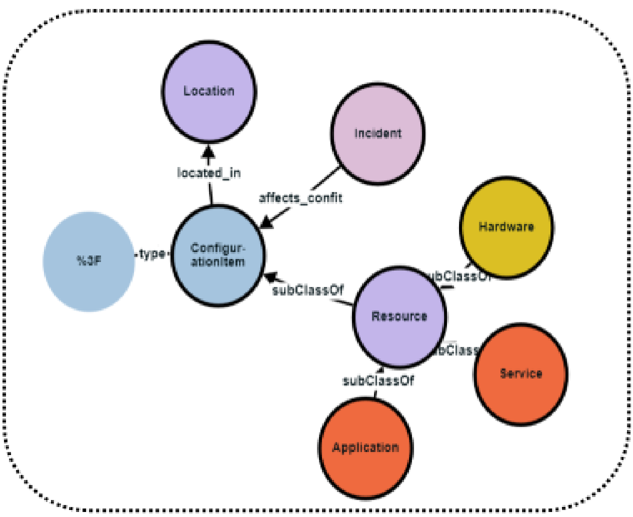

Figure 3. Configuration item type and instances

3.4. Agent data graph

The *Agent* class is specified in the subclasses *person*, *customer*, and *system notification*. The concept underpins the incident logging step of the incident management process. The incident management process follows several steps, among them incident categorization, prioritization, assignment, and resolution.

Organization

The *organization* class in Figure 3 relates to the *owner* and *representative*, which are the specification of class roles. An organization instance has employees and an internal structure defined in an organizational hierarchy in employee groups. The *organization* produces *products* that clients purchase.

3.5. Role schema

A person or organization may have different temporal roles, presented in Figure 4 and Figure 5, like a company's customer, user, owner, or representative. The constraint for a customer is to have at least one product. In Figure 5, we can follow the customer who has instantiated the presented incidents object.

ISTM event

Figure 6 illustrates the system's event interpretation – the notification of the status code of the systems, i.e., warning or failure. The failure of some components may automatically create an incident in the workflow. The typical channel is service support, where a customer reports the incident while users open an incident due to a malfunctioning component of the IT environment.

4. Classification

Our planned experiments include inferences based on the OWL 2 Lite and the OWL 2 DL [9] description logic reasoner, and the OWL 2 RL semantics with SWRL support. The instances are classified with eager reasoning based on axioms. In Figure 7, the experiments with the prioritization process assign a member to the related incident category – medium priority, according to its impact and urgency. It is realized with appropriate ontology axioms and rules. Such classification step could be included in the incident resolution pipeline as part of the solution for automatic inference or validation of an existing manual priority assessment process. Thus, we can derive new, non-explicitly defined facts.

Another approach to classification is reasoning, performed at query time. In the performed experiments, the data set was relatively limited. We plan various experiments over different data sources to reveal the model's capability further. Among them, virtual and materialized unification are planned. As a machine learning aspect, we plan to apply the inductive approach to predict values or search for

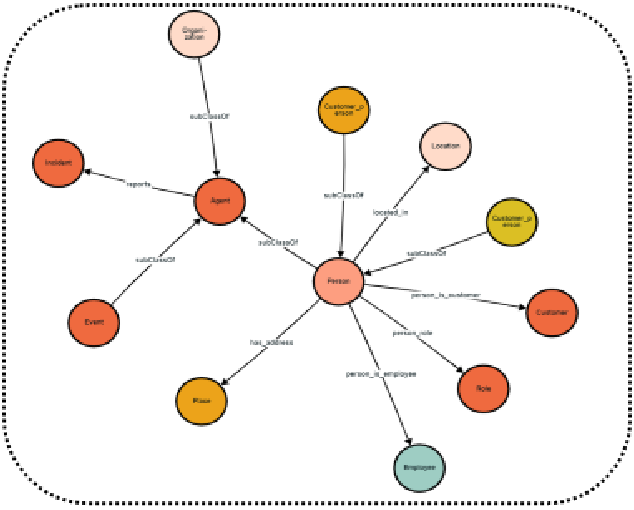

Figure 4. Agent, Person, and Role class

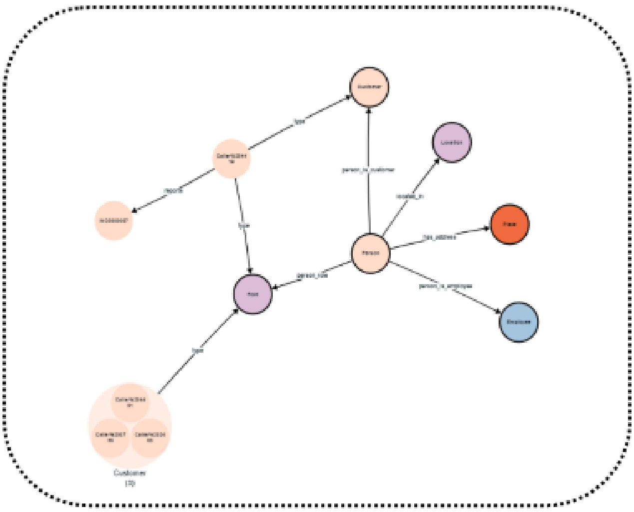

Figure 5. Role class and instances and relations

similarities in the data embedded in the graph [5]. Another intended experiment in the future is to use suitable NLP techniques to extract additional information about the incident from text data sources.

5. Conclusion

The semantic approach to search engines within an organization can significantly improve the processes within the organization, and the ontology approach

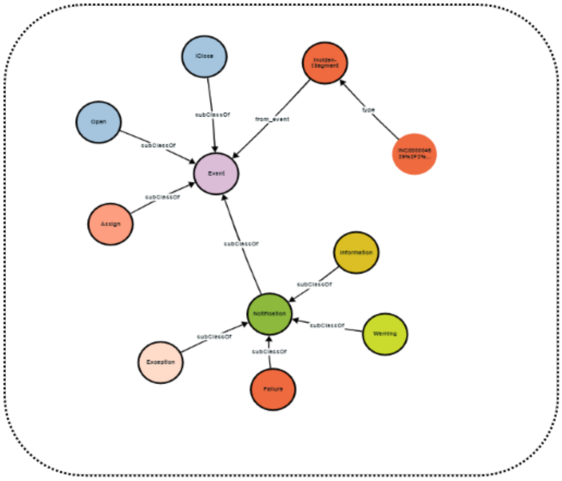

Figure 6. ITSM Event

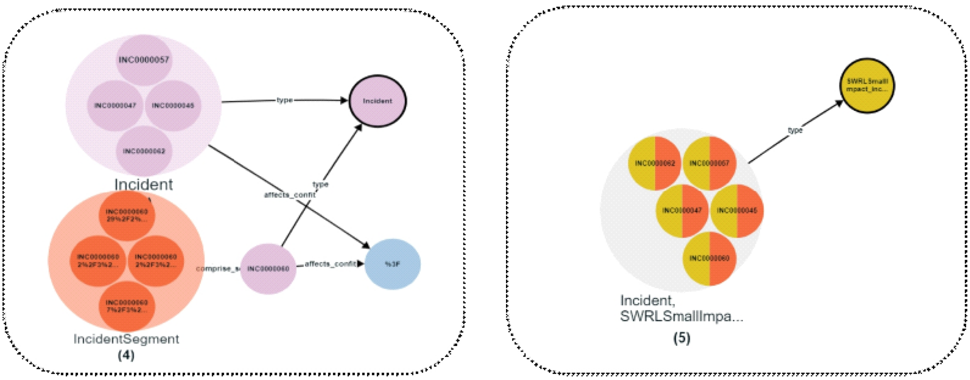

Figure 7. Incident classification

may solve different generic tasks within the organization [4]. It may rely on graph data, comprehended holistically, including various data silos. Constructing a semantic data model depends on multiple methods and techniques to build, support and enrich the enterprise's knowledge assets. Our experiments on the operational process for IT incident management with semantically modeled proactive incident discovery show promising results. Further development of such an approach can potentially leverage the organization's operational excellence.

References

[1] Defining N-ary Relations on the Semantic Web, https://www.w3.org/TR/swbp-n-aryRelations.

[2] T. Gruber, Principles for the design of ontologies used for knowledge sharing, Int. J. Hum. Comput. Stud. 43 (1993) 907–928.

[3] ITIL – IT Infrastructure Library, https://www.ibm.com/topics/it-infrastructure-library [Accessed on 11 April 2023].
[4] K. Kaloyanova, I. Patias and B. Robev, Prototyping data management model and tasks, in: CEUR Workshop Proc., ISSN 1613-0073, Vol. 3061 (2021) 256–264.
[5] K. Liang, L. Meng, M. Liu et al., A survey of knowledge graph reasoning on graph types: Static, dynamic, and multimodal, arXiv:2212.05767v6 [cs.AI].
[6] M. Marrone, L. Kolbe and M. Lutz, Impact of IT Service Management Frameworks on the IT Organization, Bus. Inf. Syst. Eng. **3**(1) (2011) 5–18, https://doi.org/10.1007/s12599-010-0141-5.
[7] OWL 2 RL in RIF (2nd Edition), https://www.w3.org/TR/rif-owl-rl/.
[8] OWL 2 Web Ontology Language. Structural Specification and Functional-Style Syntax (2nd Edition), W3C Recommendation, 11 December 2012, https://www.w3.org/TR/owl2-syntax/.
[9] OWL 2 Web Ontology Language Profiles, https://www.w3.org/TR/owl2-profiles/#Profile_Specification_3.
[10] RDF Schema 1.1, https://www.w3.org/TR/rdf-schema/.
[11] RDF 1.1 Concepts and Abstract Syntax, https://www.w3.org/TR/rdf11-concepts/.
[12] Shapes Constraint Language (SHACL), W3C Recommendation, 20 July 2017, https://www.w3.org/TR/shacl/#terminology.
[13] J. Sowa, Semantic Networks, in: Encyclopedia of Artificial Intelligence, ed. by S. Shapiro, 1987.
[14] SPARQL 1.1 Query Language, W3C Recommendation, 21 March 2013, https://www.w3.org/TR/sparql11-query/.
[15] Stardog Tutorial, https://docs.stardog.com/tutorials/.
[16] SWRL: A Semantic Web Rule Language Combining OWL and RuleML, https://www.w3.org/Submission/SWRL/.
[17] What is ITIL? Your guide to the IT Infrastructure Library | CIO, https://www.cio.com/article/272361/infrastructure-it-infrastructure-library-itil-definition-and-solutions.html.
[18] G. Xiao, D. Calvanese, R. Kontchakov et al., Ontology-based data access: A survey, in: Proc. 27th Int. Joint Conf. on Artificial Intelligence IJCAI 2018, Survey Track, 5511–5519, https://doi.org/10.24963/ijcai.2018/777.

Received on March 31, 2023
Accepted on May 7, 2023

Kristina Arnaoudova
Faculty of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
5 James Bourchier Blvd.
1164 Sofia
BULGARIA
E-mail: k.arnaoudova@fmi.uni-sofia.bg

Maria Nisheva-Pavlova
Faculty of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
5 James Bourchier Blvd.
1164 Sofia
BULGARIA

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 8
1113 Sofia
BULGARIA

E-mail: marian@fmi.uni-sofia.bg