

# ID2SBVR: Semantics of Business Vocabulary and Rules

Subjects: Computer Science, Interdisciplinary Applications | Computer Science, Software Engineering

Contributor: Irene Tangkawarow, Riyanarto Sarno, Daniel Siahaan

Semantics of Business Vocabulary and Rules (SBVR) is a standard that is applied in describing business knowledge in the form of controlled natural language. Business process designers develop SBVR from formal documents and later translate it into business process models. In many immature companies, these documents are often unavailable and could hinder resource efficiency efforts. ID2SBVR mines fact type candidates using word patterns or extracting triplets (actor, action, and object) from sentences.

Keywords: SBVR ; resource efficiency ; fact type ; operational rules ; informal document to SBVR ; ID2SBVR

---

## 1. Introduction

Business knowledge is an essential aspect of the early stages of systems development and evaluation in information systems engineering. The features of model-driven development and transformation are essential <sup>[1]</sup>. Determining business vocabulary and business rules is laborious because it requires much time and many resources. A more straightforward method to define business vocabulary (BV) and business rules (BR) is to conduct interviews and then extract them automatically using a natural language processing approach. The document resulting from the interview is an informal document.

The analysts who build a process model by gathering information require different techniques, such as document review and interviews <sup>[2]</sup>. Determining semantic business vocabulary and rules is challenging because of problems that result from the interview process. The processes described are not sequential, there are missing processes, and the interview document contains statements that are not relevant to the subject matter (noise). This documentation is not always well-structured and can be challenging to solve.

Informal documents usually refer to information that lacks a data model and that computer programs cannot easily use <sup>[3]</sup>. According to Praveen and Chandra <sup>[4]</sup> and Baig, Shuib, and Yadegaridehkordi <sup>[5]</sup>, unstructured documents include files such as word processing documents, spreadsheets, PDFs, social media and message system content, graphical content, videos, and multimedia. The unstructured document does not have structured data information and precise data types and rules to apply to its stored data.

The difference between formal and informal documents is that formal documents are written following specific standards. In contrast, informal documents are more casual, conversational, and do not have a writing standard <sup>[6]</sup>. Examples of formal documents are documents containing standard operating procedures (SOP), laws and regulations, official script procedures, and policy documents. Examples of informal documents are news documents, documented interview results, memos, personal letters, and software requirements specifications (SRS).

The fact type defines the relationship between different concepts in BR and business process models: the noun concept indicates the name of the actor and the action verb indicates the process <sup>[7]</sup>. Informal documents that pass through the preprocessing stage are the basis for determining Semantics of Business Vocabulary and Business Rules (SBVR). SBVR is a standard to describe business knowledge in the form of controlled or structured natural language. Research to determine the transformation rules from SBVR to Business Process Modeling Notation (BPMN) in terms of structural rules and operational rules has been carried out. The research illustrated the transformation of data input using data already in the form of SBVR <sup>[8][9]</sup>. The enhanced Natural Language Processing (NLP) SBVR extraction provides recognition of entities, noun and verb phrases, and multiple associations <sup>[10]</sup>. They presented NLP-enhanced and pattern-based algorithms for SBVR automatic extraction from UML case diagrams. Previous research on NLP-enhanced algorithms was extended with a model-to-model (M2M) transformation approach <sup>[11]</sup>. According to Mishra and Sureka <sup>[12]</sup>, there are inconsistencies between BPMN and SBVR. They generated Extensible Markup Language (XML) from a BPMN diagram,

extracted triplets (actor, action, and object) using grammatical relations, searched node-induced sub-graphs, and applied algorithms to detect instances of semantic inconsistency. These indicate that recent research developments in natural language aim to deliver automatic model transformation.

## **2. Informal Document to Semantics of Business Vocabulary and Rules**

The NLP research that has focused on business process modeling and SBVR has had various proposed methodologies. Several works that concern NLP, SBVR, and BPMN can be separated into six groups: works that discuss business process improvement and business process re-engineering to optimize the process and increase efficiency [13][14][15][16]; works that discuss SBVR transformation related to Software Requirements Specification (SRS) into XML [17][18][19]; works that discuss generating Unified Modeling Language (UML) class models from SRD using NLP [20]; works that discuss transformation from SBVR to BPMN where SBVR structured English (SE) specification is consistent and complete [12][21][22][23][24]; works that discuss producing SBVR from UML (use case diagram) [10][11]; and works that discuss generating natural language from business processes [2][25][26][27]. Further explanation regarding the grouping of related works is discussed below.

The Business Process Management (BPM) is an approach for advancing workflow in order to align processes with customer needs in an organization [13]. BPM covers both business process improvement and business process re-engineering [14]. Business Process (BP) focuses on re-engineering of processes and constant process improvement to achieve optimized procedures and increase efficiency and effectiveness [15].

Aiello et al. [17] investigated a mapping methodology and SBVR transformation grammar to produce rules that are ready to process in a rule engine. The main objective of their research is to overcome some weaknesses in the software development process that can result in inconsistencies between the identification of domain requirements and the functionality of the software implemented. Arshad, Bajwa, and Kazmi [18] provided an approach for translating SBVR specifications of software requirements into an XML schema. The translation mapped verb concept, noun concept, characteristic, and quantification. Akhtar et al. [19] generated a knowledge graph based on Resource Description Framework SBVR (RDFS) from SBVR. They used SBVR rules and created a triplet (actor, action, and object), then generated the RDF and the RDFS [28].

Mohanar and Samuel [20] generated UML class models instantly from software requirement specifications (SRS) using a modern approach. Their approach used OpenNLP for lexical analysis and generated required POS tags from the requirement specification. In their further research, they developed a prototype tool that can generate accurate models in a shorter time [29]. It reduces the cost and budget for both the designers and the users.

BP modeling has a long-standing tradition in several domains. This discipline persists in the constant improvement of process and issue solving [21]. They examined the basic principle and the disparity between the specifications of BV and BR modeling and BP modeling. Another research transformed BR in SBVR into BPMN to assist the business expert in the requirement validation phase [22]. The focus was on the model transformation where the SBVR Structured English (SE) specification is consistent and complete. Kluza and Honkisz [24] presented an interoperability solution for transforming a subset of the SBVR rules into the BPMN and Decision Model and Notation (DMN) models. They combined process and decision models with translation algorithms to translate the SBVR vocabulary and structural and operational rules. Bazhenova, et al., [30] succeeded to identify a group of patterns that grab potential data representations in BPMN processes and it can be used to conduct the derivation of de-cision models related to current process models. Purificação and da Silva [31] succeeded in validating SBVR business rules that deliver content to assist users writing SBVR rules. This method supplied the functionality to update parts of the defined grammar with runtime and to locate and extract verb concepts that can be validated from the BR. Mishra and Sureka [12] investigated automatic techniques to detect inconsistencies between BPMN and SBVR. The research transformed rules to graphics and applied subgraph-isomorphism to detect instances of inconsistencies between BPMN and SBVR models.

Danenas et al. [10] succeeded in producing the SBVR from UML (use case diagrams) by automatic extraction. This research enhanced recognition of entities, entire nouns and verb phrases, improved various associations extraction capabilities, and produced better quality extraction results than their previous solution [11]. Their main contributions were pre- and post-processing algorithms and extraction algorithms using a custom-trained POS tagger.

Rodrigues, Azevedo, and Revoredo [25] investigated a language-independent framework for automatically generating natural language texts from business process models. They found empirical support that, in terms of knowledge representation, the textual work instructions can be considered equivalent to process models represented in BPMN. The

research investigating the natural language structure showed that mapping rules and correlations between words representing the grammatical classes indicate a process element through keywords and/or verb tenses [2]. Furthermore, a semi-automatic approach successfully identified process elements from the natural language with a process description [26][27]. There were 32 mapping rules to recognize business process text elements using natural language processing techniques. This was discovered through an empirical study of texts comprising explanations of a process [2].

The ID2SBVR presents a new approach for extracting fact types from informal documents. The ID2SBVR allows a business process designer to translate natural language from an interview document into operational rules in SBVR, which in turn can be transformed into BPMN. The novelty of ID2SBVR is that it uses informal documents as a substitute for the formal documents that usually are required by BPMN. The informal documents are the result of an open-ended interview. The data are formed from irregularly structured natural language.

The ID2SBVR succeeds in SBVR operational rule extraction from informal documents on the basis of sentence extraction relevant to SBVR and its sequence. The unstructured data is successfully converted into semi-structured data for use in the pre-processing. The ID2SBVR method translates informal documents that are unstructured into structured ones with a high accuracy value of 0.91. The standard deviation of the ID2SBVR accuracy value in each process is 0.17. The ID2SBVR accuracy value does not show any large data deviations. The ID2SBVR method succeeded in extracting the types of facts including compound, complex, and complex-compound sentences, with an average value of 0.91 for precision and recall, and an almost perfect accuracy of 0.97.

---

## References

1. Skersys, T.; Danenas, P.; Butleris, R. Model-Based M2M Transformations Based on Drag-and-Drop Actions: Approach and Implementation. *J. Syst. Softw.* 2016, 122, 327–341.
2. Lopez, H.A.; Marquard, M.; Muttenthaler, L.; Stromsted, R. Assisted Declarative Process Creation from Natural Language Descriptions. In *Proceedings of the IEEE International Enterprise Distributed Object Computing Workshop, EDOC W*, Paris, France, 28–31 October 2019; pp. 96–99.
3. Deng, J.; Deng, X. Research on the Application of Data Mining Method Based on Decision Tree in CRM. *J. Adv. Oxid. Technol.* 2018, 21, 20–28.
4. Praveen, S.; Chandra, U. Influence of Structured, Semi-Structured, Unstructured Data on Various Data Models. *Int. J. I sc. Eng. Res.* 2017, 8, 67–69.
5. Baig, M.I.; Shuib, L.; Yadegaridehkordi, E. Big Data Adoption: State of the Art and Research Challenges. *Inf. Process. Manag.* 2019, 56, 102095.
6. McConnell, C.R.; Fallon, F.L. *Human Resource Management in Health Care*; Jones & Bartlett Publishers: Burlington, MA, USA, 2013.
7. Tangkawarow, I.; Sarno, R.; Siahaan, D. Modeling Business Rule Parallelism by Introducing Inclusive and Complex Gateways in Semantics of Business Vocabulary and Rules. *Int. J. Intell. Eng. Syst.* 2021, 14, 281–295.
8. Mickeviciute, E.; Butleris, R.; Gudas, S.; Karciauskas, E. Transforming BPMN 2.0 Business Process Model into SBVR Business Vocabulary and Rules. *Inf. Technol. Control* 2017, 46, 360–371.
9. Mickeviciute, E.; Skersys, T.; Nemuraite, L.; Butleris, R. SBVR Business Vocabulary and Rules Extraction from BPMN Business Process Models. In *Proceedings of the 8th IADIS International Conference Information Systems 2015, IS 2015*, Funchal, Portugal, 4–16 March 2015; pp. 211–215.
10. Danenas, P.; Skersys, T.; Butleris, R. Natural Language Processing-Enhanced Extraction of SBVR Business Vocabularies and Business Rules from UML Use Case Diagrams. *Data Knowl. Eng.* 2020, 128, 101822.
11. Skersys, T.; Danenas, P.; Butleris, R. Extracting SBVR Business Vocabularies and Business Rules from UML Use Case Diagrams. *J. Syst. Softw.* 2018, 141, 111–130.
12. Mishra, A.; Sureka, A. A Graph Processing Based Approach for Automatic Detection of Semantic Inconsistency between BPMN Process Model and SBVR Rules. In *Mining Intelligence and Knowledge Exploration; Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9468, pp. 115–129.
13. Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H.A. *Fundamentals of Business Process Management*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 8–10.

14. Leon, R.D.; Rodríguez-Rodríguez, R.; Gómez-Gasquet, P.; Mula, J. Business Process Improvement and the Knowledge Flows That Cross a Private Online Social Network: An Insurance Supply Chain Case. *Inf. Process. Manag.* 2020, 57, 102237.
15. Kluza, K.; Nalepa, G.J. Formal Model of Business Processes Integrated with Business Rules. *Inf. Syst. Front.* 2019, 21, 1167–1185.
16. Kluza, K.; Nalepa, G.J. A Method for Generation and Design of Business Processes with Business Rules. *Inf. Softw. Technol.* 2017, 91, 123–141.
17. Aiello, G.; Di Bernardo, R.; Maggio, M.; Di Bona, D.; Re, G.L. Inferring Business Rules from Natural Language Expressions. In *Proceedings of the IEEE 7th International Conference on Service-Oriented Computing and Applications, SOCA 2014, Matsue, Japan, 17–19 November 2014*; pp. 131–136.
18. Arshad, S.; Bajwa, I.S.; Kazmi, R. Generating SBVR-XML Representation of a Controlled Natural Language. In *Communications in Computer and Information Science*; Springer: Singapore, 2019; Volume 932, pp. 379–390.
19. Akhtar, B.; Mehmood, A.A.; Mehmood, A.A.; Noor, W. Generating RDFS Based Knowledge Graph from SBVR. In *Intelligent Technologies and Applications. INTAP 2018. Communications in Computer and Information Science*; Springer: Singapore, 2019; Volume 932, pp. 618–629.
20. Mohanan, M.; Samuel, P. Natural Language Processing Approach for UML Class Model Generation from Software Requirement Specifications via SBVR. *Int. J. Artif. Intell. Tools* 2018, 27, 6.
21. Skersys, T.; Kapocius, K.; Butleris, R.; Danikauskas, T. Extracting Business Vocabularies from Business Process Models: SBVR and BPMN Standards-Based Approach. *Comput. Sci. Inf. Syst.* 2014, 11, 1515–1536.
22. Tantan, O.; Akoka, J. Automated Transformation of Business Rules into Business Processes From SBVR to BPMN. In *Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE, Vancouver, BC, Canada, 1–3 July 2014*; Volume 25, pp. 684–687.
23. Kluza, K.; Kutt, K.; Wozniak, M. SBVRwiki (Tool Presentation). *CEUR Workshop Proc.* 2014, 1289, 703–713.
24. Kluza, K.; Honkisz, K. From SBVR to BPMN and DMN Models. Proposal of Translation from Rules to Process and Decision Models. In *Artificial Intelligence and Soft Computing: 15th International Conference, ICAISC 2016; Lecture Notes in Computer Science*; Springer: Zakopane, Poland, 2016; Volume 9693, pp. 453–462.
25. Rodrigues, R.D.A.; Azevedo, L.G.; Revoredo, K.C. BPM2Text: A Language Independent Framework for Business Process Models to Natural Language Text. *ISys-Braz. J. Inf. Syst.* 2016, 10, 38–56.
26. Ferreira, R.C.B.; Thom, L.H.; Fantinato, M. A Semi-Automatic Approach to Identify Business Process Elements in Natural Language Texts. In *Proceedings of the ICEIS 2017: 19th International Conference on Enterprise Information Systems, Porto, Portugal, 26–29 April 2017*; Volume 3, pp. 250–261.
27. Delicado, L.; Sánchez-Ferreres, J.; Carmona, J.; Padró, L. NLP4BPM—Natural Language Processing Tools for Business Process Management. *CEUR Workshop Proc.* 2017, 1920, 1–5.
28. Iqbal, U.; Bajwa, I.S. Generating UML Activity Diagram from SBVR Rules. In *2016 6th International Conference on Innovative Computing Technology, INTECH 2016; IEEE Xplore: Dublin, Ireland, 2017*; pp. 216–219.
29. Mohanan, M.; Bajwa, I.S. Requirements to Class Model via SBVR. *Int. J. Open Source Softw. Process.* 2019, 10, 70–87.
30. Bazhenova, E.; Zerbato, F.; Oliboni, B.; Weske, M. From BPMN Process Models to DMN Decision Models. *Inf. Syst.* 2019, 83, 69–88.
31. da Purificação, C.E.P.; da Silva, P.C. A Controlled Natural Language Editor for Semantic of Business Vocabulary and Rules. In *Advances in Intelligent Systems and Computing*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 738, pp. 499–508.