

АНОТАЦІЯ

Жучий Л. І. Інтеграція та узгодження даних інформаційних систем залізничного транспорту онтологічними засобами. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» – Український державний університет науки і технологій, Дніпро, 2022.

Дисертація присвячена питанням обґрунтування та розробки методів інтеграції роз'єднаних інформаційних систем залізниць та відповідної нормативної документації, підвищення контролю їх узгодженості.

У дисертаційній роботі отримані нові науково обґрунтовані теоретичні та експериментальні результати, що у сукупності є суттєвими для рішення актуальної науково-технічної задачі обробки даних та перевірки їх узгодженості на залізничному транспорті.

Наукова новизна отриманих результатів.

В роботі вперше:

- 1) виконано концептуалізацію і формалізацію онтології різних типів джерел, що дозволяє об'єднувати та узгоджувати таблиці, представлені в різних інформаційних та програмних середовищах. На відміну від інших, вона враховує специфічну табличну структуру даних та заснована на поступовому узагальненні зв'язків між елементами даних;
- 2) виконано концептуалізацію прототипу онтологічного забезпечення залізничної колії та його інтеграцію з поїзною, вагонною і відправочною моделями АСК ВП УЗ-Є. На відміну від інших, він враховує формалізовані положення нормативно-правових актів;
- 3) формалізовано процедуру формування онтологій залізничного домену засобами конструктивно-продукційного моделювання. На

відміну від інших, вона враховує інтеграцію роз'єднаних джерел даних.

Поліпшено:

- 4) технології обробки даних і процесів інформаційного супроводження: паспорту під'їзної колії; допустимих швидкостей руху поїздів на залізничних коліях загального користування на основі відповідного онтологічного забезпечення;
- 5) методи конструктивно-продукційного моделювання у частині розробки ланцюгів взаємопов'язаних конструкторів та їх застосування у залізничному домені.

Отримали подальший розвиток:

- б) засоби семантичного анотування нормативних документів. Для цього представлена схема анотування на основі онтології залізничного домену.

Практичне значення отриманих результатів полягає в розробці прототипу на основі онтологій, який дає формалізоване представлення технологічних процесів залізниці, що дозволяє перевірити узгодженість даних інформаційних систем між собою та з нормативною документацією.

Зараз на українських залізницях використовується централізована система «Єдина Автоматизована Система Управління Вантажними Перевезенням» (АСК ВП УЗ-Є), яка об'єднує основні операційні підсистеми Укрзалізниці. У зв'язку з тим, що АСК ВП УЗ-Є створена шляхом об'єднання і доопрацювання окремих підсистем, вона до сих пір страждає від фрагментації, що є певною проблемою. Крім того, виникла потреба в певній децентралізації у зв'язку з реорганізацією Укрзалізниці та створенням на її базі кількох компаній та співпрацею з інформаційними системами зарубіжних країн, іншими перевізниками та деякими замовниками.

У першому розділі виконано аналіз онтологічних розробок в залізничному та інших доменах. Встановлено що онтології на залізничному

транспорті Євросоюзу використовуються для інтеграції даних опису інфраструктури, розкладу руху поїздів та інших. При цьому недостатньо уваги приділяється нормативному забезпеченню перевізного процесу. Існують програмні засоби для анотування текстів, видобування знань з таблиць і розробки онтологій, але вони не використовуються для підтримки перевізного процесу на залізничному транспорті України. Визначено, що актуальним завданням є анотація нормативної документації для встановлення зв'язку між онтологією та текстами інструкцій.

Науково обґрунтовано можливості використання онтологічних засобів на залізничному транспорті для: формалізації нормативного забезпечення; перетворення та інтеграції даних; перевірки узгодженості даних та інструкцій інформаційної системи.

У другому розділі представлені розроблені методи: концептуалізації табличного представлення знань; формування онтології залізничної інфраструктури; інтеграції прототипу онтології залізничної інфраструктури з: онтологією джерел; онтологією поїзної, вагонної і відправочної моделей залізниці; конструктивно-продукційного моделювання онтологічного забезпечення документообігу.

Здійснена формалізація онтологічних елементів табличного представлення знань та їх відношень у вигляді графічної нотації формальної мови. Запропоновано метод формування складних понять. Концептуалізація здійснюється на основі глибокої деталізації відношень між різними поняттями.

В основі концепту лежить застосування відношення порядку, варіативність якого дає можливість моделювати різні реальні об'єкти різної природи.

Розроблено метод семантичного контролю заснований на моделі багаторівневої конкретизації та інтеграції онтологій джерел та залізничної інфраструктури. Механізми реалізації складових онтологій та їх інтеграції продемонстровано на прикладі.

У третьому розділі розроблено засоби онтологічної підтримки для паспорту під'їзної колії та допустимих швидкостей залізничних колій загального користування.

Виконано формалізацію табличного представлення даних та правил інструкцій залізничного транспорту для паспорту під'їзної колії та обмежень швидкості руху онтологічними засобами та інструментами «перетворення даних» (data wrangling) та «видобування даних» (data extraction). Це виключає частину трудомісткого процесу введення концептів в онтологію. Розроблено засоби семантичного контролю даних з різних джерел про допустимі швидкості на елементах справних залізничних колій загального користування. Для формалізації обмежень «Правил технічної експлуатації залізниць України» використовується метод семантичного анотування. Формування онтології здійснюється на основі композиції відношень.

Розроблено модульну онтологію для інтеграції даних відомостей стрілочних переводів і колій, наказів, які встановлюють допустимі швидкості на елементах залізничної інфраструктури, «Правил технічної експлуатації залізниць України» та «Будівельних норм». Цей підхід забезпечує зв'язок між природно-мовними інструкціями, інформаційними системами та онтологіями обмежень швидкості поїздів.

Онтологія несправності залізничної колії об'єднує інформацію з наступних джерел: баз даних, станційних креслень та природно-мовних інструкцій. Семантична перевірка здійснюється за допомогою формалізованих інструкцій з технічного обслуговування колії, колійних робіт та автоматизованої видачі попереджень про обмеження швидкості. Застосування розробленої онтології надає можливість підвищення безпеки руху поїздів шляхом перевірки дотримання швидкісного режиму і несправності колії, своєчасного зняття обмежень і перевірки достовірності ручного введення даних. Досліджено проблему уніфікації та інтелектуалізації інформаційних систем залізничного транспорту в Україні засобами онтологічного забезпечення. Розроблено модульну онтологію, яка дозволяє інтегрувати

інформаційні системи залізничних підсистем: станційну, вагонну, поїзну, відправочну.

У четвертому розділі доцільність запропонованого підходу експериментально підтверджена на прикладі ділянок залізниці «Маріуполь Порт-Волноваха» України, «Plovdiv-Burgas» Болгарії та «Coswig-ESig G» Німеччини та відповідного нормативного забезпечення. Онтологія може бути застосована як до високошвидкісних і звичайних залізниць (наприклад до 160 км/год), а також в подальшому інтегрована з поїзною моделлю. Оцінка онтології була проведена експертним шляхом і за допомогою застосування Ontology Pitfall Scanner.

Ключові слова: онтологія, залізниця, OWL, база знань, база даних, конструктивно-продукційне моделювання, інформаційна система, нормативно-правове забезпечення, інструкція, семантичний аналіз, анотування текстів, таблиця.

ABSTRACT

Zhuchyi L. I. Integration and harmonization of railway transport information systems data by ontological means – Qualifying scientific work as a manuscript.

Thesis submitted for obtaining the Doctor of Philosophy degree in Information technology, speciality 122 – Computer Science – Ukrainian State University of Science and Technologies, Dnipro, 2022.

The dissertation is devoted to issues of feasibility demonstration and development of methods of integration of siloed railway information systems and relevant regulations, improvement of checking their consistency.

In the dissertation work, new scientifically based theoretical and experimental results were obtained, which in the aggregate are essential for solving the actual scientific and technical problem of railway transport data processing and checking their consistency.

The scientific novelty of the obtained results.

At work for the first time:

- 7) the ontology of various types of sources was conceptualized and formalized, which allows for combining and harmonizing the tables presented in various information and software environments. Unlike others, it takes into account the specific tabular structure of data and is based on the gradual generalization of relationships between data elements;
- 8) the conceptualization of the prototype of the ontological support of the railway track and its integration with the train, wagon and dispatch models of the United Automated Management System of Freight Transportation of Ukrzalisnytsya (AMS FT UZ-U). Unlike others, it takes into account the formalized regulations;
- 9) the procedure for forming railway domain ontologies is formalized through constructive-synthesizing modelling. Unlike others, it takes into account the integration of siloed data sources.

Improved:

- 10) data processing technologies and information support processes: connection track passport; permitted railway train speeds on public railway tracks based on the corresponding ontological support;
- 11) methods of constructive-synthesizing modelling in the part of developing chains of interconnected constructors and their application in the railway domain.

Received further development:

- 12) means of semantic annotation of regulations. For this, an annotation scheme based on railway ontology is presented.

The practical value of the obtained results is in the development of a prototype based on ontologies, which provides a formalized representation of the railway technological processes, which allows for checking the consistency of the data of information systems with each other and with regulations.

Now on Ukrainian railways, the centralized system AMS FT UZ-U is used, which unites the main operational subsystems of Ukrzaliznytsya. Since the AMS FT UZ-U was created by combining and augmenting individual subsystems, it still suffers from fragmentation, which is a certain problem. In addition, there was a need for a certain decentralization because of the reorganization of Ukrzaliznytsya and the creation of several companies on its basis and cooperation with information systems of foreign countries, other carriers and some customers.

In the first section, ontological developments in the railway and other domains are analyzed. It has been established that the European Union railway transport ontologies are used for the integration of infrastructure description data, train timetables and others. At the same time, not enough attention is paid to the transportation process regulations. There are software tools for text annotation, table knowledge extraction and ontology development, but they are not used to support the Ukrainian railway transport transportation process. It was determined that the

actual task is the annotation of regulatory documentation to establish a connection between ontology and the regulation texts.

The feasibility of using ontological tools in railway transport is justified for: formalization of regulatory support; data transformation and integration; checking the consistency of information systems data and regulations.

In the second section, author presents the developed methods for:

- 13) conceptualization and of knowledge tabular representation;
- 14) formation of the railway infrastructure ontology;
- 15) integration of railway infrastructure ontology prototype with:
 - a) ontology of sources;
 - b) ontology of railway train, wagon and dispatch models;
- 16) constructive-synthesizing modelling of ontological support of document flow.

The ontological elements of tabular knowledge representation and their relations are formalized in the form of a formal language graphic notation. A method of forming complex concepts is suggested. The conceptualization basis is deep detailing of relations between different concepts.

The basis of the concept is the application of the order relation, the variability of which makes it possible to model different real objects of different natures.

A method of semantic checking based on a model of multi-level concretization and integration of ontologies of sources and railway infrastructure has been developed. Mechanisms of implementation of ontologies and their integration are demonstrated in examples

In the third section, the means of ontological support for the connection track passport and permitted speeds of public railway tracks have been developed.

The tabular presentation of data and railway transport regulation rules for the connection track passport and speed restrictions has been formalized by ontological means and tools of «data wrangling» and «data extraction». This eliminates part of the time-consuming process of ontology population.

Means of semantic control of data from various sources on permitted speeds on operational public railway track elements have been developed. To formalize the restrictions of «Technical exploitation Rules of Ukrainian railways» the method of semantic annotation is used. Ontology formation is based on the composition of relations.

A modular ontology has been developed for the integration of data on railway switches and tracks, orders that establish permitted speeds on the railway infrastructure elements, «Technical exploitation Rules of Ukrainian railways» and «Building codes». This approach provides a connection between natural language instructions, information systems and ontologies of the railway train speed restrictions.

The railway track defect ontology combines information from the following sources: databases, station drawings and natural language regulations. Semantic checking is carried out with the help of the railway track maintenance formalized, track work and automated issuance of speed restriction warnings regulations. Application of the developed ontology makes it possible to increase the safety of railway train traffic by checking compliance with the speed regime and railway track defects, timely removal of restrictions and checking the validity of manual data entry.

The problem of unification and intellectualization of information systems of railway transport in Ukraine utilizing ontological support has been investigated. A modular ontology has been developed that allows for the integration of information systems of railway subsystems: station, wagon, train, and dispatch.

In the fourth section, the feasibility of the proposed approach is experimentally confirmed on the example of the sections of the railway «Mariupol Port-Volnovakha» of Ukraine, «Plovdiv-Burgas» of Bulgaria and «Coswig-ESig G» of Germany and the corresponding regulations. The ontology can be applied to both high-speed and conventional railways (for example, up to 160 km/h), as well as further integrated with the railway train model. The evaluation of the ontology was carried out by a domain expert and by the Ontology Pitfall Scanner.

Keywords: ontology, railway, OWL, knowledge base, database, constructive-synthesizing modelling, information system, regulations, instruction, semantic analysis, annotation of texts, table.