

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ СОЦИОКУЛЬТУРНЫХ КОММУНИКАЦИЙ
Кафедра компьютерной лингвистики и лингводидактики

Реферат дипломной работы

РАЗРАБОТКА ЛИНГВИСТИЧЕСКОЙ БАЗЫ ДАННЫХ
ЮРИДИЧЕСКОГО ДОМЕНА В ФОРМАТЕ NOOJ

Кисель Ксения Сергеевна

Руководитель Воронович Валерий Викторович

Минск, 2024

РЕФЕРАТ

дипломной работы Кисель К.С.

РАЗРАБОТКА ЛИНГВИСТИЧЕСКОЙ БАЗЫ ДАННЫХ ЮРИДИЧЕСКОГО ДОМЕНА В ФОРМАТЕ NOOJ

Дипломная работа включает 77 страниц, 28 рисунков, 2 таблицы, 4 приложения и основана на 57 источниках.

Ключевые слова: ЛИНГВИСТИЧЕСКАЯ БАЗА ДАННЫХ, NOOJ, КОРПУС ТЕКСТОВ, МАШИННЫЙ СЛОВАРЬ, ЮРИДИЧЕСКАЯ ТЕРМИНОЛОГИЯ.

Цель исследования – разработка лингвистической базы данных юридического домена в формате NooJ.

Объект исследования – лингвистические базы данных в юридическом домене.

Предмет исследования – методы и подходы к созданию и применению лингвистической базы данных в юридическом домене.

Актуальность исследования заключается в повышении эффективности и качества перевода юридических текстов при помощи лингвистических баз данных юридического домена и словарей юридической терминологии в системах машинного перевода.

Методы, использованные при написании дипломной работы: анализ, описание, синтез, сравнение, структурно-функциональный метод.

Научная новизна состоит в том, что в данной работе на основании практического исследования были созданы многоязычный словарь юридических терминов и корпус текстов, содержащий Кодексы Республики Беларусь на белорусском языке.

Практическая значимость работы состоит в улучшении качества перевода юридических текстов с помощью разработанной лингвистической базы данных и многоязычного словаря юридических терминов.

Результаты исследования были апробированы на 80-ой научной конференции студентов и аспирантов БГУ, Минск, 15.05.2023 (выступление с докладом «Словари юридической терминологии в системах машинного перевода»); 81-ой научной конференции студентов и аспирантов БГУ, Минск, 20.05.2024 (выступление с докладом «Разработка лингвистической базы данных юридического домена в формате NooJ»).

Автор работы подтверждает, что приведенный в работе материал объективно отражает состояние исследуемого процесса, а все заимствования из источников сопровождаются ссылками на их авторов.

РЭФЕРАТ

дыпломнай працы Кісель К. С.

РАСПРАЦОЎКА ЛІНГВІСТЫЧНАЙ БАЗЫ ДАНЫХ ЮРЫДЫЧНАГА ДАМЕНА Ў ФАРМАЦЕ NOOJ

Дыпломная праца ўключае 77 старонак, 28 малюнкаў, 2 табліцы, 4 дадаткі і заснавана на 57 крыніцах.

Ключавыя словы: ЛІНГВІСТЫЧНАЯ БАЗА ДАНЫХ, NOOJ, КОРПУС ТЭКСТАЎ, МАШЫННЫ СЛОЎНІК, ЮРЫДЫЧНАЯ ТЭРМІНАЛОГІЯ.

Мэта даследавання – распрацоўка лінгвістычнай базы даных юрыдычнага дамена ў фармаце NooJ.

Аб'ект даследавання – лінгвістычныя базы даных у юрыдычным дамене.

Прадмет даследавання – метады і падыходы да стварэння і прымянення лінгвістычнай базы даных у юрыдычным дамене.

Аktуальнасць даследавання заключаецца ў павышэнні эфектыўнасці і якасці перакладу юрыдычных тэкстаў пры дапамозе лінгвістычных баз даных юрыдычнага дамена і слоўнікаў юрыдычнай тэрміналогіі ў сістэмах машыннага перакладу.

Метады, выкарыстаныя пры напісанні дыпломнай працы: аналіз, апісанне, сінтэз, параўнанне, структурна-функцыянальны метад.

Навуковая навізна складаецца ў тым, што ў дадзенай працы на падставе практычнага даследавання былі створаны шматмоўны слоўнік юрыдычных тэрмінаў і корпус тэкстаў, які змяшчае кодэксы Рэспублікі Беларусь на беларускай мове.

Практычная значнасць працы складаецца ў паляпшэнні якасці перакладу юрыдычных тэкстаў з дапамогай распрацаванай лінгвістычнай базы даных і шматмоўнага слоўніка юрыдычных тэрмінаў.

Вынікі даследавання былі апрабаваныя на 80-ай навуковай канферэнцыі студэнтаў і аспірантаў БДУ, Мінск, 15.05.2023 (выступленне з дакладам «Слоўнікі юрыдычнай тэрміналогіі ў сістэмах машыннага перакладу»); 81-ай навуковай канферэнцыі студэнтаў і аспірантаў БДУ, Мінск, 20.05.2024 (выступленне з дакладам «Распрацоўка лінгвістычнай базы даных юрыдычнага дамена ў фармаце NooJ»).

Аўтар працы пацвярджае, што прыведзены ў працы матэрыял аб'ектыўна адлюстроўвае стан даследаванага працэсу, а ўсе запазычаныя з крыніц суправаджаюцца спасылкамі на іх аўтараў.

SUMMARY

of diploma paper by Kisel K.S.

DEVELOPMENT OF A LINGUISTIC DATABASE OF A LEGAL DOMAIN IN THE NOOJ FORMAT

The diploma paper includes 77 pages, 28 drawings, 2 table, 4 appendixes and is based on 57 sources.

Key words: LINGUISTIC DATABASE, NOOJ, CORPUS OF TEXTS, MACHINE DICTIONARY, LEGAL TERMINOLOGY.

The objective is to development of a linguistic database of the legal domain in the NooJ format.

The research object is linguistic databases in the legal domain.

The research subject is methods and approaches to the creation and application of a linguistic database in a legal domain.

The relevance of the study is improving the efficiency and quality of translation of legal texts using linguistic databases of the legal domain and dictionaries of legal terminology in machine translation systems.

Methods used in writing the diploma paper: analysis, description, synthesis, comparison, structural and functional method.

The scientific novelty lies in the fact that in this work, on the basis of practical research, a multilingual dictionary of legal terms and a corpus of texts containing the Codes of the Republic of Belarus in the Belarusian language were created.

The practical significance of the work is to improve the quality of translation of legal texts using a developed linguistic database and a multilingual dictionary of legal terms.

The research results were approved at the 80th Scientific Conference of Students and Postgraduates of BSU, Minsk, 15.05.2023 (presentation of the report "Dictionaries of legal terminology in machine translation systems"); 81st Scientific Conference of Students and Postgraduates of BSU, Minsk, 20.05.2024 (presentation of the report "Development of a linguistic database of a legal domain in the NooJ format").

The author of the paper confirms that the material correctly and objectively reflects the state of the investigated process and all borrowed theoretical, methodological and methodical concepts from literature and other sources are accompanied by links to their authors.