

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУАЗИЯ
ҰЛТТЫҚ УНИВЕРСИТЕТІ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН
ЕВРАЗИЙСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ Л.Н. ГУМИЛЕВА

MINISTRY OF EDUCATION AND SCIENCE
OF THE REPUBLIC OF KAZAKHSTAN
L.N. GUMILYOV EURASIAN NATIONAL UNIVERSITY



16-18 маусым
Нұр-Сұлтан, 2022

«TURKLANG 2022»

«Түркі тілдерін компьютерлік өңдеу»
атты X халықаралық конференция
ЕҢБЕКТЕРІ

ТРУДЫ

X Международной конференции
«Компьютерная обработка тюркских языков»

«TURKLANG 2022»

PROCEEDINGS

of the X International Conference
on Computer processing of Turkic Languages

«TURKLANG 2022»

**ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ**

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН
ЕВРАЗИЙСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ Л.Н. ГУМИЛЕВА**

**MINISTRY OF EDUCATION AND SCIENCE OF
THE REPUBLIC OF KAZAKHSTAN
L.N. GUMILYOV EURASIAN NATIONAL UNIVERSITY**

**«TURKLANG 2022»
«Түркі тілдерін компьютерлік өңдеу»
атты X халықаралық конференция
ЕҢБЕКТЕРІ
16-18 маусым 2022 ж.**

**ТРУДЫ
X Международной конференции
«Компьютерная обработка тюркских языков»
«TURKLANG 2022»
16-18 июня 2022 г.**

**PROCEEDINGS
of the X International Conference
on Computer processing of Turkic Languages
«TURKLANG 2022»
16-18 June 2022**

Нұр-Сұлтан, 2022

УДК 80/81:004
ББК 81.2:32-973
Т 90

Техникалық редакция:

Ергеш Б.Ж.
Елибаева Г.К.
Турсынова Н.А.

Т 90 ТҮРКІ ТІЛДЕРІН КОМПЬЮТЕРЛІК ӨНДЕУ. X халықаралық конференция: Еңбектері = КОМПЬЮТЕРНАЯ ОБРАБОТКА ТЮРКСКИХ ЯЗЫКОВ. X международная конференция: Труды. / - Нұр-Сұлтан: «Булатов А.Ж.» ЖК, 2022.= Нур-Султан: ИП «Булатов А.Ж.»

ISBN 978-601-326-645-9

Жинақта «Түркі тілдерін компьютерлік өңдеу» атты X халықаралық конференция қатысушыларының баяндамалары енген.

Компьютерлік лингвистика бағыты бойынша оқитын студенттерге, магистранттарға, докторанттарға және мамандарға арналған.

Жинақ «BR11765535» Қазақ тілі мәдениетін арттыру және функцияларды кеңейту бойынша ғылыми-лингвистикалық негіздер мен IT-ресурстарды әзірлеу» бағдарламасы есебінен жарияланды.

В сборнике представлены доклады участников X международной конференции «Компьютерная обработка тюркских языков».

Предназначен для студентов, магистрантов, докторантов и специалистов специализирующихся в областях компьютерной лингвистика.

Сборник издан за счет средств программы BR11765535 «Разработка научно-лингвистических основ и IT-ресурсов по расширению функций и повышению культуры казахского языка».

УДК 80/81:004
ББК 81.2:32-973

ISBN 978-601-326-645-9

© Л.Н.Гумилев атындағы Еуразия ұлттық университеті, 2022

© Евразийский национальный университет им. Л.Н. Гумилева, 2022

ЭОК 004.89

¹Жұмаиш Б., ²Муканова А. С.¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті²Астана Халықаралық университеті

Нұр-Сұлтан, Қазақстан

¹zh.batyrzhan@bk.ru, ²asiserikovna@gmail.com

ОНТОЛОГИЯЛЫҚ МОДЕЛЬ НЕГІЗІНДЕ КӨЛІК САЛЫҒЫ БОЙЫНША АҚПАРАТТЫҚ ЖҮЙЕНІ ӘЗІРЛЕУ

Андатпа: Мақалада «Көлік салығы» пәндік облысы бойынша білім базасы негізінде ақпараттық жүйені жобалауға байланысты мәселелер көрсетілген. Жүйені әзірлеудің әдіснамалық тәсілін қарастыра отырып, зерттеу саласы анықталып, пәндік облыстың онтологиясын құрудың теориялық аспектілері көрсетілген. «Көлік салығы» пәндік облысы бойынша білімдерді құрылымдау және формализациялау жүргізілген. Онтологиялық моделдің негізгі класстар мен ішкі класстар жіктеліп, оларды бейнелейтін сипаттамалар анықталған. «Көлік салығы» пәндік облысының онтологиялық моделі құрылып, білім базасы негізінде ақпараттық жүйе әзірленген.

Түйін сөздер: онтологиялық модел, білімдер базасы, пәндік облыс, SPARQL.

УДК 004.89

¹Жұмаиш Б., ²Муканова А. С.¹Евразийский национальный университет имени Л.Н. Гумилева,²Международный университет Астана

Нур-Султан, Казахстан

¹zh.batyrzhan@bk.ru, ²asiserikovna@gmail.com

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ПО ТРАНСПОРТНОМУ НАЛОГУ НА ОСНОВЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ

Аннотация: В статье отражены вопросы, связанные с проектированием информационной системы на основе базы знаний по предметной области «Транспортный налог». рассмотрен методологический подход к разработке системы, определена область исследования, выявлены теоретические аспекты построения онтологии предметной области. Произведена структуризация и формализация

знаний по предметной области «Транспортный налог». Определены характеристики элементов онтологической модели и описаны их значения. Были классифицированы основные классы, подклассы и выявлены характеристики, описывающие данные понятия. Создана онтологическая модель предметной области «Транспортный налог» и разработана информационная система на основе базы знаний.

Ключевые слова: онтологическая модель, база знаний, предметная область, SPARQL

UDC 004.89

¹Zhumash B., ²Mukanova A.

¹L. N. Gumilyov Eurasian National University

²Astana International University

Nur-Sultan, Kazakhstan

¹zh.batyrzhan@bk.ru, ²asiserikovna@gmail.com

DEVELOPMENT OF A TRANSPORT TAX INFORMATION SYSTEM BASED ON AN ONTOLOGICAL MODEL

Abstract: The article presents issues related to the design of an information system based on the knowledge base in the subject area "Transport tax". Considering the methodological approach to the development of the system, the field of research is determined and the theoretical aspects of building the ontology of the subject area are indicated. Structuring and formalization of knowledge in the subject area "transport tax" was carried out. In the ontological model, the main classes and subclasses are classified and the characteristics that represent them are determined. An ontological model of the subject area "transport tax" has been created, and an information system based on the knowledge base has been developed.

Keywords: ontological model, knowledge base, subject area, SPARQL

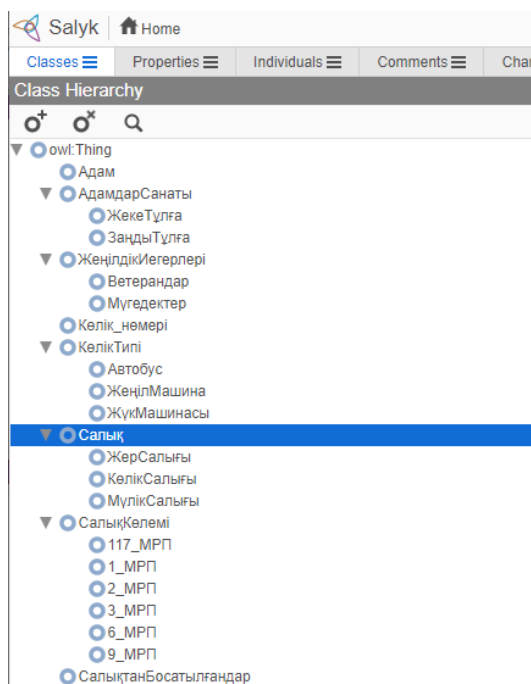
Онтология кез келген пәндік облыс бойынша білімдерді сипаттау үшін қолданылатын ұғымдарды жинақтайды. Пәндік облыс туралы ақпаратты бірлесіп пайдалануы қажет қосымшалар, мәліметтер базасы және адамдар онтологияны пайдалана алады. Онтологияға пәндік аймақтың негізгі ұғымдарының машинамен оқылатын, яғни формальды түрде бейнеленген анықтамалары және олардың арасындағы қатынастар кіреді. Олар пәндік облыстан және онымен байланысты басқа салалардан берілген білімдерді қайта пайдалана отырып жаңа білімдер алуға мүмкіндік береді [1].

Әдетте онтология егжей-тегжейлі, дәл, дәйекті, негізделген және класстарды, қасиеттер мен қатынастарды нақты бөле алатын логикаға негізделген тілде жасалады. Онтологиямен жұмыс істеудің кейбір құралдары онтологияны қолдана отырып, автоматты ойлауды жүзеге асыра алады, осылайша тұжырымдамалық, семантикалық іздеу, программалық қамтама агенттері, шешімдерді қолдау, сөйлеу және табиғи тілді тану, білімді басқару, смарт мәліметтер базасы және электрондық коммерция сияқты смарт қосымшалар үшін жетілдірілген қызметтерді ұсынады [2].

Онтология семантикалық желінің пайда болуында маңызды рөл атқарады, өйткені құжаттар семантикасының көрінісі және осы семантиканы желілік қосымшалар мен интеллектуалды агенттер қолдана алады. Онтология қазіргі уақытта жинақталып, стандартталатын метадеректер терминдерінің мағынасын құрылымдау және сипаттау тәсілі ретінде қоғам үшін пайдалы болуы мүмкін. Онтологияны қолдана отырып, болашақтың қосымшалары "интеллектуалды" бола алады, яғни олар адамның ойлау деңгейінде және тілінде дәлірек жұмыс істей алады. Сондықтан да онтология табиғи тілде деректерді өңдеумен айналысатын барлық салаларда кеңінен қолданылады. Онтологияны әртүрлі қосымшаларда қолдануға байланысты оларды ұсынудың стандартталған тәсілдерін құру қажеттілігі туындады. Барлық жүйелерде қолдануға болатын әртүрлі тілдердің дамуы басталды, олардың ішіндегі ең танымалдары -RDF және OWL.

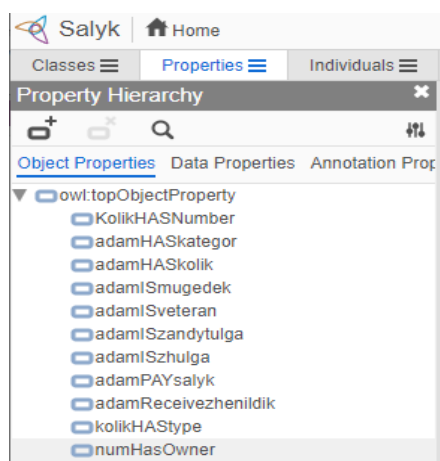
Веб-онтология тілі (OWL) – веб-онтологияларды анықтауға және жасауға арналған тіл. OWL онтологиясы класстардың, қасиеттердің және олардың даналарының сипаттамасын қамтиды. OWL сөздіктердегі терминдердің мағынасын және сол терминдер арасындағы қатынастарды айқын көрсету үшін қолданылады [3].

Бұл жұмыста «Көлік салығы» саласы бойынша онтологиялық модель құрылып, білімдер базасы жасалды, SparQL сұратымдары арқылы табиғи тілде берілген сұрауларға жауап беретін жүйе құрылды. 1 суретте «Көлік салығы» онтологиясындағы негізгі класстар көрсетілген.

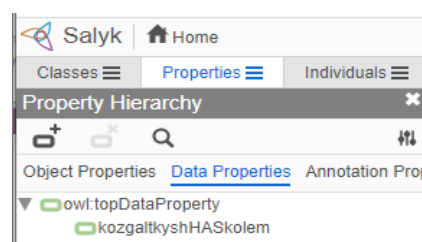


Сурет 1. Класстар иерархиясы

Суретте көрсетілгендей, көлік салығын төлеуші адамдарды «Жеке тұлға» және «Заңды тұлға» деп қарастырылған. Сонымен қатар онтологиялық модельде салық түрлері, транспорт түрлері мен салық төлеуден босатылған немесе жеңілдік түрлері көрсетілген. Пәндік саланың негізгі түсініктерінен басқа, олардың қасиеттері мен қатынастары анықталған. Сурет 2 және Сурет 3 те объектілер қасиеттері мен деректер қасиеттері көрсетілген.

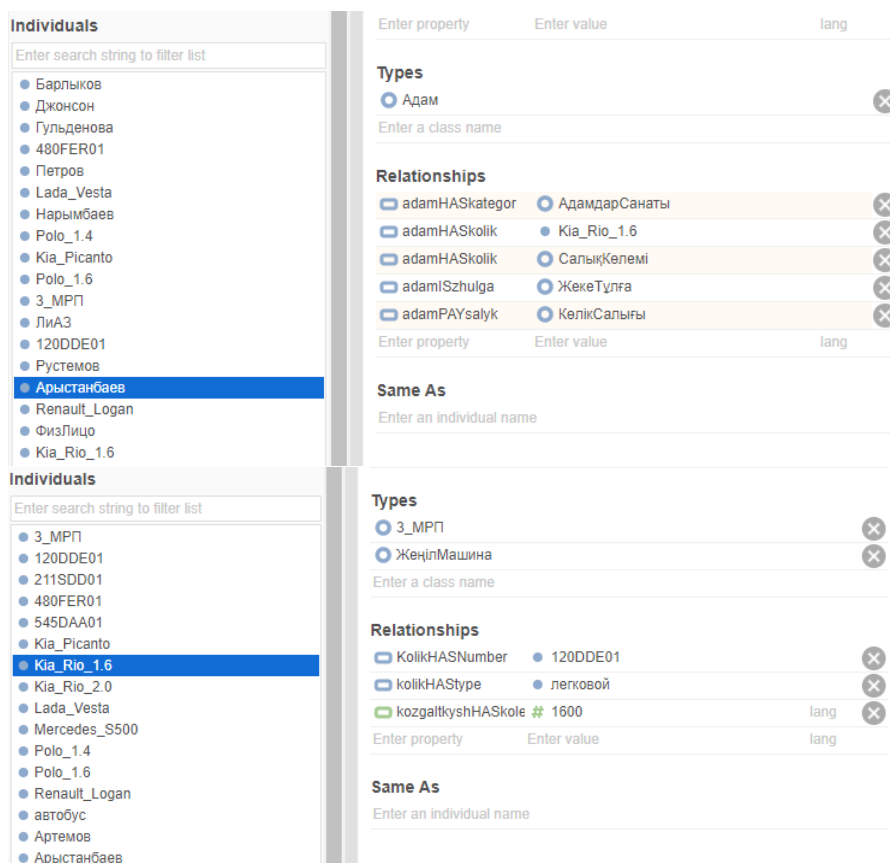


Сурет 2. Объектілер қасиеттері



Сурет 3. Деректер қасиеттері

Пәндік облыс класстарына сәйкес индивидтер қосылып, олардың арасындағы қатынастар көрсетілді (Сурет 4).



Сурет 4. Индивидтер беті

«Көлік салығы» пәндік облысы бойынша құрылған онтология негізінде табиғи тілде қажетті ақпарат алуға мүмкіндік беретін жүйе жасалды. Ол үшін Java программалау тілі қолданылды. Java – тілінде онтологиялармен жұмыс істеуге арналған «Jena ontology API» таңдалды. Ақпараттық жүйе Apache Netbeans IDE программалау ортасында жазылды. Онтологиядан деректерді алу үшін SPARQL сұралым тілі қолданылды.

SPARQL (ағылшын тілінен рекурсивті акроним. SPARQL Protocol and RDF Query Language) - RDF моделі бойынша ұсынылған деректерге сұрау салу тілі, сондай-ақ осы сұрау салулар мен оларға жауап беруге арналған хаттама. SPARQL-бүкіләлемдік ғаламтор консорциумының (W3C) ұсынысы және семантикалық веб технологиясының бірі[4]. SPARQL семантикалық вебті көрудің маңызды элементтерінің бірі болып табылады - жобаның бастамашысы және бастамашысы Тим Бернерс-Ли 2006 жылдың мамыр айында сұхбатында "SPARQL үлкен өзгеріс әкеледі"деп атап өтті [5].

Жұмыста көрсетілген онтология үшін сұратылым префикстері төмендегідей жазылды:

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX owl: <http://www.w3.org/2002/07/owl#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
 PREFIX onto:
<http://www.semanticweb.org/batyr/ontologies/2022/2/untitled-ontology-15#>

Мысалы, белгілі бір адамға тиесілі көліктер тізімін алу үшін келесі сұратылым жасалынады (Сурет 5).

The screenshot shows a SPARQL query interface with the following query:

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX onto: <http://www.semanticweb.org/batyr/ontologies/2022/2/untitled-ontology-15#>
SELECT ?s ?o {
  ?s onto:человекИмеетТранспорт ?o;
}

```

The results table shows the following data:

s	o
Барлықов	Kia_Rio_1.6
Джонсон	Kia_Rio_1.6
Арыстанбаев	Kia_Rio_1.6
Гульденова	Kia_Rio_1.6
Рустемов	Kia_Rio_1.6
Петров	Lada_Vesta
Нарымбаев	Kia_Rio_2.0
Петров	Renault_Logan
Артемов	Polo_1.6
Гульденова	Polo_1.4

Сурет 5. SPARQL сұратымдарын қолдану мысалы

Netbeans ортасында java тілінде сұратылымдар жасалып, онтологиядан деректер алу үшін функция құрылды. Пәндік онтология негізінде жасалған «Көлік салығы» бойынша табиғи тілде сұратымдар жасауға мүмкіндік беретін жүйенің фрагменттері (Сурет 6) да көрсетілген.

Автокөлік маркалары

иесі кім?

Renault_Logan
Polo_1.6
Polo_1.4
Mercedes_S500
Lada_Vesta
Kia_Rio_2.0
Kia_Rio_1.6

ардагер ме?

Жауабы: Белгісіз

Қозғалтқыш көлемі (см³): 0

Автокөлік маркалары

иесі кім?

Renault_Logan
Polo_1.6
Polo_1.4
Mercedes_S500
Lada_Vesta
Kia_Rio_2.0
Kia_Rio_1.6

Рустемов
Джонсон
Гульденова
Барлыков
Арыстанбаев

ардагер ме?

Жауабы: Белгісіз

Рустемов
Джонсон
Гульденова
Барлыков
Арыстанбаев

ардагер ме?

Жауабы: Иа, Соғыс ардагері. Көліктерінің саны : 1

Ұлы Отан соғысының ардагерлері,
басқа мемлекеттер аумағындағы ұрыс қимылдарының ардагерлері
Егер автокөліктер саны 1 ден аспаса көлік салығынан босатылады.

600
тг

Автокөлік маркалары

иесі кім?

Renault_Logan
Polo_1.6
Polo_1.4
Mercedes_S500
Lada_Vesta
Kia_Rio_2.0
Kia_Rio_1.6

Рустемов
Джонсон
Гульденова
Барлыков
Арыстанбаев

Қозғалтқыш көлемі (см³): 1600

Сумма налога 3 мрп = 9189 тг

Сурет 6. «Көлік салығы» бойынша табиғи тілде сұратымдар жасауға мүмкіндік беретін жүйенің фрагменттері

Көрсетілген жұмыста «Көлік салығы» пәндік саласы бойынша негізгі түсініктер мен олардың арасындағы қарым-қатынас және соларға сәйкес ережелер анықталып, солардың негізінде пәндік облыстың

онтологиялық моделі құрылды, білімдер базасы жасалды. Білімдер базасына сұратымдар жасау арқылы табиғи тілде машинаның түрі, көлемі, адамдардың санатына қарай көлік салығы бойынша ақпарат алуға мүмкіндік беретін жүйе жасалды.

Әдебиеттер тізімі

1 OWL Язык Сетевых онтологий: Варианты использования и требования. — Текст : электронный // https://www.w3.org/2006/04/OWL_UseCases-ru.html : [сайт]. — URL: (дата обращения: 23.05.2022).

2 Date C.J. Deit K.D. Vvedenie v sistemy baz dannyx // М.: Izd. Dom «Viliams», 2001. – 72 s.

3 OWL Web Ontology Language. – URL: <https://www.w3.org/TR/2004/REC-owlfeatures-20040210/>

4 W3C Semantic Web Activity Publications - www.w3.org/2001/sw/Specs.html (англ.). W3C. — Перечень публикаций W3C по проекту семантической паутины.

5 Berners-Lee looks for Web's big leap web.archive.org/web/20070930221904/http://news.zdnet.co.uk/internet/0,1000000097,39270671,00.htm (англ.). — Интервью Тима Бернерса-Ли.