

УДК 013, 004.65

ОНТОЛОГИЯ ВСПОМОГАТЕЛЬНЫХ И ПОЛИТЕМАТИЧЕСКИХ ПРЕДМЕТНЫХ КЛАССОВ ЕДИНОГО ЦИФРОВОГО ПРОСТРАНСТВА НАУЧНЫХ ЗНАНИЙ

С. А. Власова¹ [0000-0003-1533-5850], Н. Е. Каленов² [0000-0001-5269-0988],

А. Н. Сотников³ [0000-0002-0137-1255]

¹⁻³Межведомственный суперкомпьютерный центр (МСЦ) РАН – филиал ФГУ
ФНЦ Научно-исследовательский институт системных исследований (НИИСИ)
РАН

¹vlav.svetlana2013@yandex.ru, ²nekalenov@mail.ru, ³asotnikov@jssc.ru

Аннотация

Одними из основных компонентов Единого Цифрового Пространства Научных Знаний (ЕЦПНЗ) являются предметные онтологии отдельных тематических подпространств, включающие в себя основные понятия, относящиеся к данному научному направлению. Задача построения предметных онтологий на первом этапе требует формирования массива ключевых терминов в заданной области науки с последующим установлением связей между ними. Настоящая работа является развитием исследований, проводимых авторами в области создания ЕЦПНЗ. В рамках предыдущих исследований была предложена унифицированная структура представления онтологии элементов ЕЦПНЗ (подпространств, классов и атрибутов объектов, связей между объектами или атрибутами). В процессе моделирования онтологии на примере универсального и ряда тематических подпространств ЕЦПНЗ выявилась необходимость некоторой корректировки структуры онтологии, касающейся справочников ЕЦПНЗ, для обеспечения возможности описания вложенных атрибутов данных. Кроме того, в онтологию введено понятие «тип словаря значений атрибутов данных», определены два типа словарей – «статические» и «динамические». Эта информация позволяет упростить алгоритмы формально-логического контроля при формировании контента ЕЦПНЗ. Указание

на тип словаря введено в структуру справочников атрибутов объектов. В представленной работе описана модифицированная структура онтологии на примере 11-ти вспомогательных и 10-ти предметных классов универсального подпространства (УПП) ЕЦПНЗ. Приведены примеры справочников каждого класса, построенные в соответствии с моделью структуры онтологии, перечень атрибутов объектов и примеры статических словарей.

Ключевые слова: *цифровое пространство научных знаний, онтология, классы объектов, атрибуты, структуризация, связанные данные.*

ВВЕДЕНИЕ

Настоящая работа является продолжением исследований, проводимых в МСЦ РАН и связанных с созданием Единого цифрового пространства научных знаний (ЕЦПНЗ) как структурированной интегрированной информационной среды, отражающей достижения в различных областях науки [1, 2]. На предыдущих этапах работы была определена архитектура построения ЕЦПНЗ [3], рассмотрены вопросы сетевого обеспечения ЕЦПНЗ [4] и отражения в нем 3D-моделей мультимедийных объектов [5, 6].

В материалах конференции «Научный сервис в сети Интернет» в 2022 году опубликован доклад, описывающий предложенную нами модель структуры онтологии ЕЦПНЗ [7]. Модифицированная версия модели представлена в [8, 9]. Согласно этой модели ЕЦПНЗ представляется в виде 5-ти уровневой иерархической структуры (ЕЦПНЗ – подпространства – классы объектов – атрибуты объектов класса – значения атрибутов), дополненной связями трех типов (универсальные, квазиуниверсальные и специфические). Информация обо всех элементах ЕЦПНЗ хранится в справочниках и словарях. Элементы каждого иерархического уровня описаны справочниками фиксированной для каждого уровня структуры. Справочниками представлены также связи каждого из трех типов. Значения атрибутов и связей хранятся в словарях, информация о которых содержится в соответствующих справочниках. Словари подразделяются на две группы – статические и динамические. Первые содержат значения «стандартизованных» атрибутов (таких как перечень ученых степеней, званий и должностей, рубрик классификационных систем и т. п.). Такие словари заполняются при первоначальной инсталляции ЕЦПНЗ

или при формировании его подпространств и корректируются администраторами, наделенными соответствующими правами.

Словари второго типа наполняются по мере формирования контента ЕЦПНЗ конкретными данными (фамилии персон, наименования публикаций, ссылки на сетевые ресурсы и т. п.).

КЛАССЫ УНИВЕРСАЛЬНОГО ПОДПРОСТРАНСТВА И ИХ АТТРИБУТЫ

Реализация предложенной структуры моделировалась на примере формирования элементов универсального подпространства (УПП) ЕЦПНЗ, содержащего вспомогательные классы объектов, а также объекты мультидисциплинарного характера и связи между ними. В рамках модели выделено 11 вспомогательных и 10 предметных классов объектов.

В качестве вспомогательных выделены следующие классы объектов:

- Форматы данных;
- Универсальные классификационные системы;
- Местоположение (географические характеристики);
- Временные характеристики, даты;
- Количественные характеристики;
- Мировые константы;
- Единицы измерения;
- Языки;
- Группы персон;
- Коллекции;
- Числовые значения.

К предметным классам УПП отнесены следующие:

- Персоны;
- Публикации;
- Квалификационные работы;
- Документы;
- Музейные предметы;
- Изображения и мультимедийные объекты;
- События / мероприятия;
- Организации;

- Политематические базы данных, каталоги ресурсов;
- Награды, гранты.

Для каждого из предметных и вспомогательных классов составлены перечни атрибутов. При определении набора атрибутов объектов учитывался многолетний положительный опыт эксплуатации электронной библиотеки «Научное наследие России» [10], современная версия которой построена на совокупности связанных разнородных данных и развивается как модель составляющей ЕЦПНЗ [11, 12].

Ниже приведен перечень атрибутов объектов УПП. В скобках после наименования атрибута указано, является ли его значение обязательным (о) или факультативным (ф).

Атрибуты вспомогательных классов

Важнейшую роль в структуре онтологии ЕЦПНЗ играет класс «Форматы данных». Его объекты представляют собой структурированный набор правил описания атрибутов других классов. Указание на формат представления данных присутствует во всех справочниках атрибутов объектов и связей. Информация о форматах используется для формально-логического контроля вводимых данных при потоковой загрузке, а также в процессе диалога с оператором «ручного» ввода данных. Кроме того, этот класс объектов может использоваться в качестве источника информации о тех или иных форматах данных, с которыми встречаются пользователи ЕЦПНЗ.

Атрибуты объектов класса «Форматы представления данных»:

- тип представления данных (о); атрибут может принимать значения «текст», «целое число», «дата в формате гггг[.мм[.дд]]», «связи» и т. д.);
- вид формата (ф); атрибут конкретизирует формат, например, «pdf», «jpeg», «URL», ряд значений его словаря описывает структуру представления различных видов связей ЕЦПНЗ;
- обязательное (r) или факультативное (f) значение атрибута (о);
- уникальное (u) или множественное (m) значение атрибута (о);
- ограничения по структуре (ф) (атрибут содержит наименование конкретной структуры, которой должен соответствовать тот или иной атри-

бут, например, «ГОСТ 7-1.2003: Библиографическое описание¹» или «алгоритм контроля ISBN²», или «необходимые требования к структуре адреса электронной почты³» и т. п.);

- ссылка на подробное описание формата (ф).

Атрибуты объектов класса «Универсальные классификационные системы»:

- наименование классификации (о);
- наименование раздела (о);
- индекс (о).

Между объектами этого класса могут устанавливаться различные простые связи универсального типа.

- Атрибуты объектов класса «Местоположение»:
- название объекта (о);
- ид объекта (о);
- дополнительная информация (ф).

Объекты этого класса могут быть связаны с объектами подпространства «География», перечень атрибутов которых существенно шире.

Атрибуты объектов класса «Временные характеристики, даты»:

- название характеристики (о);
- словесное значение характеристики или URN числового значения (ф).

Атрибуты объектов класса «Единицы измерения»:

- наименование единицы измерения (о);
- предмет измерения (о);
- обозначения (аббревиатура) (о);
- дополнительная информация (ф).

Атрибуты объектов класса «Языки»:

- наименование языка (о);
- двухсимвольный код языка (в соответствии со стандартом группы ISO 639) (о);

¹ со ссылкой на официальное описание ГОСТа

² с приведением формулы контроля

³ с формулировкой требований

- трехсимвольный код языка (в соответствии со стандартом группы ISO 639) (о).

Атрибуты объектов класса «Количественные характеристики»:

- наименование характеристики (о);
- единица измерения (о);
- числовое значение (о).
- Атрибуты объектов класса «Мировые константы»:
- наименование (о);
- описание (о);
- обозначение (о);
- значение (о).

Атрибуты объектов класса «Группы персон»:

- наименование группы (о);
- описание (ф).

Атрибуты объектов класса «Коллекции»:

- наименование коллекции (о);
- классы включаемых объектов (о);
- условия включения объекта в коллекцию (о).

Атрибутами объектов вспомогательного класса «Числовые значения» являются любые числа.

Атрибуты предметных классов УПП

- Атрибуты объектов класса «Персоны»:
- фамилия (о);
- имя (о);
- отчество (ф);
- псевдоним (ф);
- дата рождения точная (ф);
- дата рождения приблизительная (ф);
- место рождения (о);
- дата смерти (ф);
- место смерти (ф);
- ученая степень (ф);

- ученое звание (ф);
- биография (о);
- библиография персоны как автора (ф);
- библиография о персоне (ф).

Дополнительная информация (область научных интересов, трудовая деятельность, научные открытия, идентификаторы в базах данных и т. д.) оформляется как именованные связи с соответствующими объектами.

Атрибуты объектов класса «Публикации»:

- наименование (о);
- год издания (о);
- вид (о);
- библиографическое описание (о);
- ISBN (ф);
- ISSN (ф);
- аннотация (ф);
- URL полного текста (ф);
- полный текст (если он представлен в ЕЦПНЗ) (ф).

Дополнительная информация оформляется в виде связей с персонами (может принимать значения «автор», «переводчик», «редактор», «о персоне» и т. д.), организациями, классификационными системами и другими классами объектов.

Атрибуты объектов класса «Квалификационные работы»:

- наименование (о);
- вид работы (о);
- дата выпуска (защиты) (ф);
- URL полного текста (ф);
- Полный текст, если он включен в ЕЦПНЗ (ф).

Дополнительная информация оформляется в виде связей с персонами (автор, научный руководитель, оппонент и т. д.), организациями (место выполнения, место защиты, ведущая организация и т. п.), классификационными системами и пр.

Атрибуты объектов класса «Музейные предметы»:

- наименование объекта (о);
- наименование музея, где хранится оригинал (именованная связь с объектом класса «организации») (о);
- вид источника поступления (о);
- дата поступления в музей (о);
- дата обнаружения (создания) предмета (о);
- описание предмета (о).

Музейные объекты могут быть связаны специфическими связями с персонай (изготовитель объекта, владелец, «автор сбора» – для естественнонаучных коллекций), географическим объектом (место сбора), публикациями и документами, касающимися данного объекта, и т. д.

Атрибуты объектов класса «События / мероприятия»:

- наименование (о);
- тип события (о);
- дата начала события (мероприятия) (о);
- дата окончания события (мероприятия) (о),
- описание события (мероприятия) (ф).

Эти данные дополняются связями с объектами других классов по аналогии с предыдущими. В частности, связь события с персонай (если речь идет о конференции) может принимать значения «председатель оргкомитета», «член программного комитета», «спонсор» и т. п.

Атрибуты объектов класса «Документы»:

- наименование (о);
- вид документа; (о);
- текст документа (ф).

Атрибуты объектов класса «Изображения и мультимедийные материалы»:

- наименование (о);
- вид объекта (о);
- адрес объекта (о);
- дата создания объекта (ф);
- развернутая информация об объекте (ф).

Атрибуты объектов класса «Организации»:

- наименование (о);
- тип организации – научная, образовательная, производственная и т.п. (о);
- вид организации – бюджетная, ООО, ЗАО и т. п. (о).;
- местонахождение (о);
- контактные данные (ф);
- дата создания (ф);
- дата ликвидации (ф).

Атрибуты объектов класса «Политематические базы данных, каталоги ресурсов»:

- наименование (о);
- вид объекта (о);
- условия доступа (ф);
- URL ресурса (о).

Атрибуты объектов класса «Награды, гранты»:

- наименование (о);
- вид награды (о);
- раздел науки (о);
- дата награждения (получения гранта) (ф);
- дата окончания гранта (ф).

Для формирования справочников атрибутов и словарей их значений разработана диалоговая программа, с помощью которой сформированы справочники вышеперечисленных классов объектов и атрибутов, а также элементы статических словарей их значений.

Ниже приведен ряд справочников классов, фрагментов справочников атрибутов и статических словарей.

Class.1: Форматы представления данных; UN; UNFT; A_UNFT; форматы представления атрибутов объектов (число, время, дата, текст и т. п.).

Справочник атрибутов объектов класса «Форматы»:

A_UNFT.1: тип представления данных; ; N_A_UNFT.1; S ;

A_UNFT.2: вид формата; ; N_A_UNFT.2; S

A_UNFT.3: обязательное (r) или факультативное (f) значение атрибута; ;
N_A_UNFT.3; S ;

A_UNFT.4: уникальное (u) или множественное (m) значение атрибута; ;
N_A_UNFT.4; S

A_UNFT.5: ограничения по структуре; ; N_A_UNFT.5; D

A_UNFT.6: ссылка на описание формата; ; N_A_UNFT; D

Фрагменты словарей значений атрибутов объектов класса «Форматы»:

N_A_UNFT.1.1: текст

N_A_UNFT.1.2: изображение

N_A_UNFT.1.3: видео

N_A_UNFT.1.4: звук

N_A_UNFT.1.5: время в формате чч[.мин[.сек]]

N_A_UNFT.1.6: дата в формате гггг[.мм[дд]]

N_A_UNFT.1.7: формула

N_A_UNFT.1.8: связь

N_A_UNFT.1.9: ссылка на сетевой ресурс

И т. д.

N_A_UNFT.2.1: MP3

N_A_UNFT.2.2: PDF

N_A_UNFT.2.3: Excel, csv

N_A_UNFT.2.4: JPG

N_A_UNFT.2.5: MP4

N_A_UNFT.2.6: простая связь URNc первого типа между объектами, атрибутами или значениями O1 и O2 вида <URNc>:<URNO1><URNO2>, где URNc – URN конкретной связи. Пример: фамилия «Крылов» эквивалентна «Krylov» и «Krylow»; отдел входит в состав института и т. д.

N_A_UNFT.2.7: простая связь второго типа, указывающая на субъект, объект, URN связи и URN значения связи. Формат представления связи имеет вид:

<URNc>:<URN субъекта><URN объекта>=<URN элемента словаря значений соответствующего атрибута связи>. Пример: персона P1 является сотрудником организации O1 в должности инженера (значение атрибута).

И т. д.

N_A_UNFT.3.1: r

N_A_UNFT.3.2: f

N_A_UNFT.4.1: u

N_A_UNFT.4.2: m

N_A_UNFT.5.1: арабские цифры

N_A_UNFT.5.2: библиографическое описание по ГОСТ
<https://docs.cntd.ru/document/1200034383>

N_A_UNFT.5.3: структура адреса электронной почты – буквенно-цифровая строка, содержащая внутри символ «@», по крайней мере, одну точку не в начале и не в конце строки и не содержащая спецсимволов.

N_A_UNFT.5.4: только буквы

N_A_UNFT.5.5: структура URL – строка символов, начинающаяся с http:// или
<https://>

И т. д.

N_A_UNFT.6.1: <https://ru.wikipedia.org/wiki/JPEG>,
<https://open-file.ru/types/jpeg> [описание формата JPEG]

N_A_UNFT.6.2: <https://ru.wikipedia.org/wiki/MPEG-4>,
<https://open-file.ru/types/mp4> [описание формата mp4]

И т. д.

Фрагмент словаря форматов:

UNFT.1: N_A_UNFT.1.1; ; N_A_UNFT3.1; N_A_UNFT.4.1 [текст, атрибут обязательный, значение уникальное]

UNFT.2: N_A_UNFT.1.7; ; N_A_UNFT3.1; N_A_UNFT.4.1 [дата в формате гггг[.мм[.дд]], атрибут обязательный, значение уникальное]

UNFT.3: N_A_UNFT.1.15; ; N_A_UNFT.3.2; N_A_UNFT.4.2; N_A_UNFT.5.5 [ссылка на внешний ресурс, атрибут необязательный, повторяющийся]

UNFT.4: N_A_UNFT.1.1; ; N_A_UNFT3.1; N_A_UNFT.4.2 [текст, только буквы, атрибут обязательный, значение повторяющееся]

UNFT.5: N_A_UNFT.1.1; ; N_A_UNFT3.2; N_A_UNFT.4.1 [текст, атрибут факультативный, значение уникальное]

UNFT.6: N_A_UNFT.1.1; ; N_A_UNFT3.2; N_A_UNFT.4.2 [текст, атрибут факультативный, значение повторяющееся]

UNFT.7: N_A_UNFT.1.1; ; N_A_UNFT3.1; N_A_UNFT.4.1; N_A_UNFT.5.2 [библиографическое описание элемент обязательный, значение уникальное]

UNFT.8: N_A_UNFT.1.9; ; N_A_UNFT3.2; N_A_UNFT.4.2; N_A_UNFT.5.5 [ссылка на интернет-ресурс, элемент факультативный, значение повторяющееся].

Class.2: Единицы измерения; UN; UNMU; A_UNMU; стандартные единицы измерения различных физических величин.

Справочник атрибутов объектов класса «Единицы измерения»:

A_UNMU.1: наименование единицы измерения; UNFT.1; N_A_UNMU.1; S

A_UNMU.2: предмет измерения; UNFT.4; N_A_UNMU.2; S

A_UNMU.3: обозначения (аббревиатура); UNFT.4; N_A_UNMU.3; S

A_UNMU.4: дополнительная информация; UNFT.18; N_A_UNMU.5; S

Фрагменты словарей значений атрибутов класса «Единицы измерения»:

N_A_UNMU.1.1: секунда

N_A_UNMU.1.2: метр

N_A_UNMU.1.3: грамм

N_A_UNMU.1.4: ньютон

И т. д.

N_A_UNMU.2.1: время

N_A_UNMU.2.2: длина

N_A_UNMU.2.3: масса

N_A_UNMU.2.4: сила

И т. д.

N_A_UNMU.3.1: с.
N_A_UNMU.3.2: сек.
N_A_UNMU.3.3: м.
N_A_UNMU.3.4: г.
N_A_UNMU.3.5: дж.
N_A_UNMU.3.6: н.
И т. д.

N_A_UNMU.4.1: Определение секунды в Википедии
<https://ru.wikipedia.org/wiki/Секунда>

N_A_UNMU.4.2: Определение секунды в Большой российской энциклопедии <https://bigenc.ru/physics/text/3546123>
И т. д.

Элемент словаря объектов класса «Единицы измерения», относящийся к секунде, будет иметь вид:

UNMU.1: N_A_UNMU.1.1; N_A_UNMU.2.1; N_A_UNMU.3.1; N_A_UNMU.3.2;
N_A_UNMU.4.1; N_A_UNMU.4.2.

Class.3: Персоны; UN; UNPS; A_UNPS; ; информация о персонах, в той или иной мере связанных с научными исследованиями.

Фрагмент справочника атрибутов объектов класса «Персоны»:

A_UNPS.1: фамилия; UNFT.1; N_A_UNPS.1; D; фамилия выбирается из словаря, при отсутствии она вводится и проверяется на эквивалентность с другими написаниями

A_UNPS.4: дата рождения; UNFT.2; N_A_UNTC.2; D; дата выбирается из словаря временных характеристик, при отсутствии вводится в словарь в соответствии с указанным форматом

A_UNPS.8: квалификация (ученая степень); UNFT.6; N_A_UNPS.8; S; выбирается из словаря;

Фрагмент статического словаря значений атрибута «квалификация (ученая степень)»:

N_A_UNPS.8.1: доктор физ.-мат. наук

N_A_UNPS.8.2: кандидат техн. наук

N_A_UNPS.8.3: PhD в области химии

Class.4: Публикации; UN; UNPB; A_UNPB; книги, сборники, журналы, статьи и т. п.

A_UNPB.1: наименование; UNFT.1; N_A_UNPB.1; D

A_UNPB.2: год издания; UNFT.1; N_UNND; D

A_UNPB.3: вид публикации; UNFT.1; N_A_UNPB.3; S

A_UNPB.4: библиографическое описание; UNFT.7; N_A_UNPB.4; D

Если издание имеет несколько названий, это указывается в виде универсальной связи «эквивалентно» между элементами словаря N_A_UNPB.1. Это позволит поисковому модулю обрабатывать запросы, включающие элементы всех названий данного издания.

Указание на то, что статья опубликована в данном журнале (сборнике), формируется как связь «входит в состав».

Роли персон и организации в данной публикации указываются в виде специфических связей с соответствующими значениями.

Фрагмент статического словаря значений атрибута «вид публикации»:

N_A_UNPB.3.1: монография

N_A_UNPB.3.2: сборник

N_A_UNPB.3.3: статья

N_A_UNPB.3.4: журнал

N_A_UNPB.3.5: выпуск журнала

Class.5: Квалификационные работы UN; UNDS; A_UNDS; диссертации, авторефераты и т. п.

A_UNDS.1: наименование; UNFT.1; N_A_UNDS.1; D

A_UNDS.2: дата выпуска (защиты) работы; UNFT.2; N_A_UNTC.2; D

A_UNDS.3: вид работы; UNFT.1; N_A_UNDS.3; S

A_UNDS.4: URL полного текста; UNFT.9; N_A_UNDS.4; D заполняется в случае отсутствия текста в ЕЦПНЗ, при наличии полного текста в ЕЦПНЗ формируется специфическая связь «квалификационная работа – документ» или «квалификационная работа – изображение»;

A_UNDS.5: Дополнительная информация; UNFT.5; N_A_UNDS.5; D
Фрагмент статического словаря значений атрибута «Вид работы»:

N_A_UNDS.3.1: диссертация докторская

N_A_UNDS.3.2: диссертация кандидатская

N_A_UNDS.3.3: диссертация PhD

Тема диссертации (номер специальности ВАК, индексы УДК и ГРНТИ и др.) указывается как элементы квазиуниверсальной связи с объектами класса «Универсальные классификационные системы». Роли персон и организаций в подготовке и защите диссертации указываются как элементы именованных специфических связей.

Class.6: Музейные предметы; UN; UNMS; A_UNMS; цифровые копии предметов, хранящихся в музеях

A_UNMS.1: наименование предмета; UNFT.1; N_A_UNMS.1; D; ; ;

A_UNMS.2: наименование музея, где хранится в настоящее время оригинал; UNFT.1; N_A_UNOR.1; D; выбирается из словаря организаций, при отсутствии вводится в словарь;

A_UNMS.3: вид источника поступления; UNFT.1; N_A_UNMS.3; S; ;

A_UNMS.4: дата поступления в музей; UNFT.2; N_A_UNTC.2; D; проверяется по словарю временных характеристик. при отсутствии вводится в словарь

A_UNMS.5: дата обнаружения (создания) объекта; UNFT.2; N_A_UNTC.2; D; проверяется по словарю временных характеристик. при отсутствии вводится в словарь

A_UNMS.6: описание музейного предмета; UNFT.5; N_A_UNMS.6; D

Фрагменты статического словаря значений атрибута «Вид источника поступления»

N_A_UNMS.3.1: приобретен музеем

N_A_UNMS.3.2: получен в дар

N_A_UNMS.3.3: изготовлен для музея

Именованные связи музейных предметов с персонами, организациями, событиями и другими классами объектов отражаются в справочниках и словарях соответствующих специфических связей.

ЗАКЛЮЧЕНИЕ

Использование предложенной модели онтологии ЕЦПНЗ позволяет унифицировать алгоритмы создания контента пространства, разработать типовой интерфейс добавления новых элементов вне зависимости от подпространства и конкретного вида данных, упростить и ускорить алгоритмы поиска и навигации по связанным данным. В настоящее время в МСЦ РАН ведутся исследования по развитию и конкретизации предложенной модели в части алгоритмизации формирования вложенных связей, а также моделирования формирования фрагментов ЕЦПНЗ на примере реальных данных, в том числе, составляющих контент электронной библиотеки «Научное наследие России».

Работа выполнена в Межведомственном суперкомпьютерном центре РАН в рамках государственного задания по теме FNEF-2024-0014.

СПИСОК ЛИТЕРАТУРЫ

1. Антопольский А.Б., Каленов Н.Е., Серебряков В.А., Сотников А.Н. О едином цифровом пространстве научных знаний // Вестник Российской академии наук. 2019. Т. 89. № 7. С. 728–735.
<https://doi.org/10.31857/S0869-5873897728-735>.
2. Савин Г.И. Единое цифровое пространство научных знаний: цели и задачи // Информационные ресурсы России. 2020. № 5. С. 3–5.
<https://doi.org/10.51218/0204-3653-2020-5-3-5>.
3. Каленов Н.Е., Сотников А.Н. Архитектура единого цифрового пространства научных знаний // Информационные ресурсы России. 2020. № 5. С. 5–8. <https://doi.org/10.51218/0204-3653-2020-5-5-8>.

4. *Абрамов А.Г., Гончар А.А., Евсеев А.В.* Национальная исследовательская компьютерная сеть нового поколения как инфраструктурно-сервисная платформа Единого цифрового пространства научных знаний // Информационные ресурсы России. 2020. № 5. С. 43–46.

<https://doi.org/10.51218/0204-3653-2020-5-43-46>.

5. *Соболевская И.Н.* Об особенностях представления мультимедийных объектов в едином цифровом пространстве научных знаний // Информационные ресурсы России. 2020. № 5. С. 31–34.

<https://doi.org/10.51218/0204-3653-2020-5-31-34>.

6. *Sobolevskaya I.* Some Aspects of 3D-objects Presentation in a Common Digital Space of Scientific Knowledge // CEUR Workshop Proceedings (CEUR-WS.org). 2021. Vol. 2990. P. 117–124.

<https://doi.org/10.51218/1613-0073-2990-117-124>.

7. *Каленов Н.Е., Сотников А.Н.* О структуре онтологии Единого цифрового пространства научных знаний // Научный сервис в сети Интернет: труды XXIV Всероссийской научной конференции. 2022. С. 203–221.

<https://doi.org/10.20948/abrau-2022-23>.

8. *Каленов Н.Е., Сотников А.Н.* Унифицированное представление онтологии единого цифрового пространства научных знаний // Электронные библиотеки. 2023. Т. 26. № 1. С. 80–103.

<https://doi.org/10.26907/1562-5419-2023-26-1-80-103>.

9. *Kalenov N.E., Sobolevskaya I.N., Sotnikov A.N.* Common Digital Space of Scientific Knowledge Ontology Structurization // Lobachevskii Journal of Mathematics. 2023. Vol. 44. No. 7. P. 2733–2743. <https://doi.org/10.1134/S1995080223070235>

10. *Погорелко К.П.* Динамика использования электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии: Сб. науч. тр. 2017. С. 192–200.

11. *Pogorelko K.* A New Version of the Software for the Information System “Scientific Heritage of Russia” // CEUR Workshop Proceedings (CEUR-WS.org). 2021. Vol. 2990. P. 110–116. <https://doi.org/10.51218/1613-0073-2990-110-116>.

12. *Каленов Н.Е., Погорелко К.П., Сотников А.Н.* О развитии электронной библиотеки «Научное наследие России» как составляющей Единого цифрового

пространства научных знаний // Информационные процессы. 2022. Т. 22. № 3. С. 155–166. https://doi.org/10.53921/18195822_2022_22_3_155.

ONTOLOGY OF THE UNIVERSAL SUBSPACE OF COMMON DIGITAL SPACE OF SCIENTIFIC KNOWLEDGE

S. A. Vlasova¹ [0000-0003-1533-5850], **N. E. Kalenov**² [0000-0001-5269-0988],

A. N. Sotnikov³ [0000-0002-0137-1255]

¹⁻³ Joint Supercomputer Center of the Russian Academy of Sciences – JSCC

¹*vlas.svetlana2013@yandex.ru*, ²*nekalenov@mail.ru*, ³*asotnikov@jssc.ru*

Abstract

The work is a development of research conducted by the authors in the field of creating a Common Digital Space for Scientific Knowledge (CDSSK). Previous research has proposed a unified structure for representing the ontology of CDSSK elements (subspaces, classes and attributes of objects, relationships between objects or attributes). In the process of modeling the ontology using the example of the universal and a number of thematic subspaces of the CDSSK, the need for some adjustments to the structure of the ontology regarding CDSSK directories was revealed to ensure the possibility of describing nested data attributes. In addition, the concept of “data attribute value dictionary type” was introduced into the ontology; two types of dictionaries were defined – “static” and “dynamic”. This information makes it possible to simplify formal-logical control algorithms when generating CDSSK content. An indication of the dictionary type has been introduced into the structure of object attribute directories. The presented work describes the modified structure of the ontology using the example of 11 auxiliaries and 10 subject classes of the CDSSK universal subspace (USS). Examples of directories of each class, built in accordance with the ontology structure model, a list of object attributes and examples of static dictionaries are given.

Keywords: digital space of scientific knowledge, ontologies, structuring, linked data.

REFERENCES

1. *Antopol'skiy A.B., Kalenov N.Ye., Serebryakov V.A., Sotnikov A.N.* // O yed-inom tsifrovom prostranstve nauchnykh znaniy // Vestnik Rossiyskoy akademii nauk. 2019. T. 89. № 7. S. 728–735.
2. *Savin G.I.* Edinoe cifrovoe prostranstvo nauchny`x znaniy: celi i zadachi // Informacionny`e resursy` Rossii. 2020. № 5. S. 3–5. <https://doi.org/10.51218/0204-3653-2020-5-3-5>.
3. *Kalenov N.Ye., Sotnikov A.N.* Arkhitektura yedinogo tsifrovogo pros-transtva nauchnykh znaniy // Informatsionnyye resursy Rossii. 2020. № 5. S. 5–8. <https://doi.org/10.51218/0204-3653-2020-5-5-8>.
4. *Abramov A.G., Gonchar A.A., Yevseyev A.V.* Natsional'naya issle-dovatel'skaya komp'yuternaya set' novogo pokoleniya kak infrastruktarno-servisnaya platforma Yedinogo tsifrovogo prostranstva nauchnykh znaniy // Informatsionnyye resursy Rossii. 2020. № 5. S. 43–46.
5. *Sobolevskaya I.N.* Ob osobennostyakh predstavleniya mul'timediynykh ob"yektov v yedinom tsifrovom prostranstve nauchnykh znaniy // Informatsionnyye resursy Rossii. 2020. № 5. S. 31–34. <https://doi.org/10.51218/0204-3653-2020-5-31-34>.
6. *Sobolevskaya I.* Some Aspects of 3D-objects Presentation in a Common Digital Space of Scientific Knowledge // CEUR Workshop Proceedings (CEUR-WS.org). 2021. Vol. 2990. P. 117–124. <https://doi.org/10.51218/1613-0073-2990-117-124>.
7. *Kalenov N.Ye., Sotnikov A.N.* O strukture ontologii Yedinogo tsifrovogo prostranstva nauchnykh znaniy // Nauchnyy servis v seti Internet: trudy XXIV Vse-rossiyskoy nauchnoy konferentsii. 2022. S. 203–221. <https://doi.org/10.20948/abrau-2022-23>.
8. *Kalenov N.Ye., Sotnikov A.N.* Unifitsirovannoye predstavleniye ontologii yedinogo tsifrovogo prostranstva nauchnykh znaniy // Elektronnyye biblioteki. 2023. T. 26, № 1. S. 80–103. <https://doi.org/10.26907/1562-5419-2023-26-1-80-103>.
9. *Kalenov N.E., Sobolevskaya I.N., Sotnikov A.N.* Common Digital Space of Scientific Knowledge Ontology Structurization // Lobachevskii Journal of Mathematics. 2023. Vol. 44. No. 7. P. 2733–2743. <https://doi.org/10.1134/S1995080223070235>.

10. *Pogorelko K.P.* Dinamika ispol'zovaniya elektronnoy biblioteki "Nauchnoye naslediyе Rossiі" // Informatsionnoye obespecheniye nauki: novyye tekhnologii: Sb. nauch. tr. 2017. S. 192–200.

11. *Pogorelko K.* A New Version of the Software for the Information System "Scientific Heritage of Russia" // CEUR Workshop Proceedings (CEUR-WS.org). 2021. Vol. 2990. P. 110–116. <https://doi.org/10.51218/1613-0073-2990-110-116>.

12. *Kalenov N.Ye., Pogorelko K.P., Sotnikov A.N.* O razvitii elektronnoy biblioteki "Nauchnoye naslediyе Rossiі" kak sostavlyayushchey Yedinogo tsifrovogo prostanstva nauchnykh znaniy // Informatsionnyye protsessy. 2022. T. 22, № 3. S. 155–166. https://doi.org/10.53921/18195822_2022_22_3_155.

СВЕДЕНИЯ ОБ АВТОРАХ



ВЛАСОВА Светлана Александровна – ведущий научный сотрудник Межведомственного Суперкомпьютерного Центра РАН – филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», кандидат технических наук.

Svetlana Aleksandrovna VLASOVA – Leading Researcher of Joint Super Computer Center of the Russian Academy of Sciences – Branch of Federal State Institution «Scientific Research Institute for System Analysis of the Russian Academy of Sciences», Candidate of Technical Sciences.

email: vlas.svetlana2013@yandex.ru

ORCID: 0000-0003-1533-5850



КАЛЕНОВ Николай Евгеньевич – главный научный сотрудник Межведомственного Суперкомпьютерного Центра РАН – филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», доктор технических наук, профессор.

Nikolay Evgenievich KALENOV – Chief Researcher of Joint Super Computer Center of the Russian Academy of Sciences – Branch of Federal State Institution «Scientific Research Institute for System Analysis of the Russian Academy of Sciences», Doctor of Technical Sciences, Professor.

email: nekalenov@mail.ru

ORCID: 0000-0001-5269-0988



СОТНИКОВ Александр Николаевич – заместитель директора Межведомственного Суперкомпьютерного Центра РАН – филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», доктор физико-математических наук, профессор.

Alexander Nikolaevich SOTNIKOV – Deputy Director of the Joint SuperComputer Center of the Russian Academy of Sciences – Branch of the Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, Doctor of Sciences (Math), Professor.

email: asotnikov@jssc.ru

ORCID: 0000-0002-0137-1255

Материал поступил в редакцию 3 декабря 2023 года