

Наука и инновации в информационном обществе

ЕДИНОЕ ЦИФРОВОЕ ПРОСТРАНСТВО НАУЧНЫХ ЗНАНИЙ В МИРОВОМ ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ

Статья рекомендована к публикации членом редакционного совета А.М. Елизаровым 30.09.2020.

Каленов Николай Евгеньевич

Д.т.н., профессор

Межведомственный суперкомпьютерный центр Российской академии наук – филиал Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», главный научный сотрудник

Москва, Россия

nekalenov@mail.ru

Соболевская Ирина Николаевна

К.ф.-м.н.

Межведомственный суперкомпьютерный центр Российской академии наук – филиал Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», старший научный сотрудник

Москва, Россия

ins@jicc.ru

Сотников Александр Николаевич

Д.ф.-м.н.

Межведомственный суперкомпьютерный центр Российской академии наук – филиал Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», заместитель директора

Москва, Россия

asotnikovov@mail.ru

Аннотация

Научное пространство является подпространством существующего информационного пространства, в котором содержится вся научная информация, представленная в различной форме (в виде печатных и электронных документов, аудио- и видеоматериалов, изображений, моделей пространственных объектов т.п.). В свою очередь, цифровое научное пространство также является подмножеством всего информационного пространства, в котором содержится научная информация в цифровой форме. Цифровое пространство знаний (ЦПЗ) – подмножество цифрового пространства, содержащее достоверную информацию. Единое цифровое пространство научных знаний (ЕЦПНЗ) – подмножество ЦПЗ, содержащее проверенную научную информацию, представленную в соответствии с единой онтологией. Под онтологией понимается свод правил построения и предоставления пользователям элементов ЕЦПНЗ. Поэтому, для построения ЕЦПНЗ, необходимо определить его элементы и онтологию их представления. В статье рассматривается структура ЕЦПНЗ как совокупности подпространств знаний по различным научным направлениям, виды и типы данных, составляющих контент ЕЦПНЗ, операции над ними, принципы управления пространством.

Ключевые слова

цифровое научное пространство, электронная библиотека, единое цифровое пространство, предметная онтология, научная сеть

© Каленов Н.Е., Соболевская И.Н., Сотников А.Н., 2021. Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «Атрибуция — Некоммерческое использование — На тех же условиях» Всемирная 4.0 (Creative Commons Attribution – NonCommercial - ShareAlike 4.0 International; CC BY-NC-SA 4.0). См. <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.ru>

Введение

Единое цифровое пространство научных знаний (ЕЦПНЗ) является одним из важнейших объектов современного информационного общества. Пространство в его математическом (формализованном) понимании есть множество некоторых объектов с определенными правилами работы с ними и набор аксиом, которым эти правила должны подчиняться, т.е. это множество с введенной на нем структурой. Определяя множество и структуру по-разному, можно получать разные пространства. Мировое информационное пространство содержит все сведения, накопленные человечеством в процессе его эволюции, представленные на материальных носителях. Оно включает различного рода документы, представленные в печатной, рукописной или электронной формах (публикации, архивные материалы, научно-техническую документацию и т.п.), фотографии, кино- видео- и аудиоматериалы, мультимедиа и 3D-модели объектов реального мира.

Цифровое пространство знаний (ЦПЗ), являющееся частью мирового информационного пространства, представляет собой компьютерную среду, содержащую формализованные характеристики разнородных знаний (это и законы механики, и химические формулы, и расписание движения транспорта, и рецепты приготовления различных блюд, и пр.). Цифровое пространство научных знаний (ЦПНЗ) – часть ЦПЗ, содержащая достоверную фундаментальную, научно-образовательную и научно-популярную информацию по различным областям науки, представленную в разнообразных формах. Единое цифровое пространство научных знаний (ЕЦПНЗ) – компьютерная среда, содержащая информацию, отраженную в ЦПНЗ, организованную и предоставляемую пользователям по единым правилам для всех областей науки (рис. 1).



Рисунок 1 – Информационные пространства

Иными словами, ЕЦПНЗ состоит из совокупности подпространств, относящихся к отдельным направлениям науки, связанных между собой на основе онтологии, общей для всего пространства [1, 2]. Общая онтология включает ряд предметных онтологий, описывающих с помощью тезаурусов и классификационных систем отдельные научные направления (рис. 2).

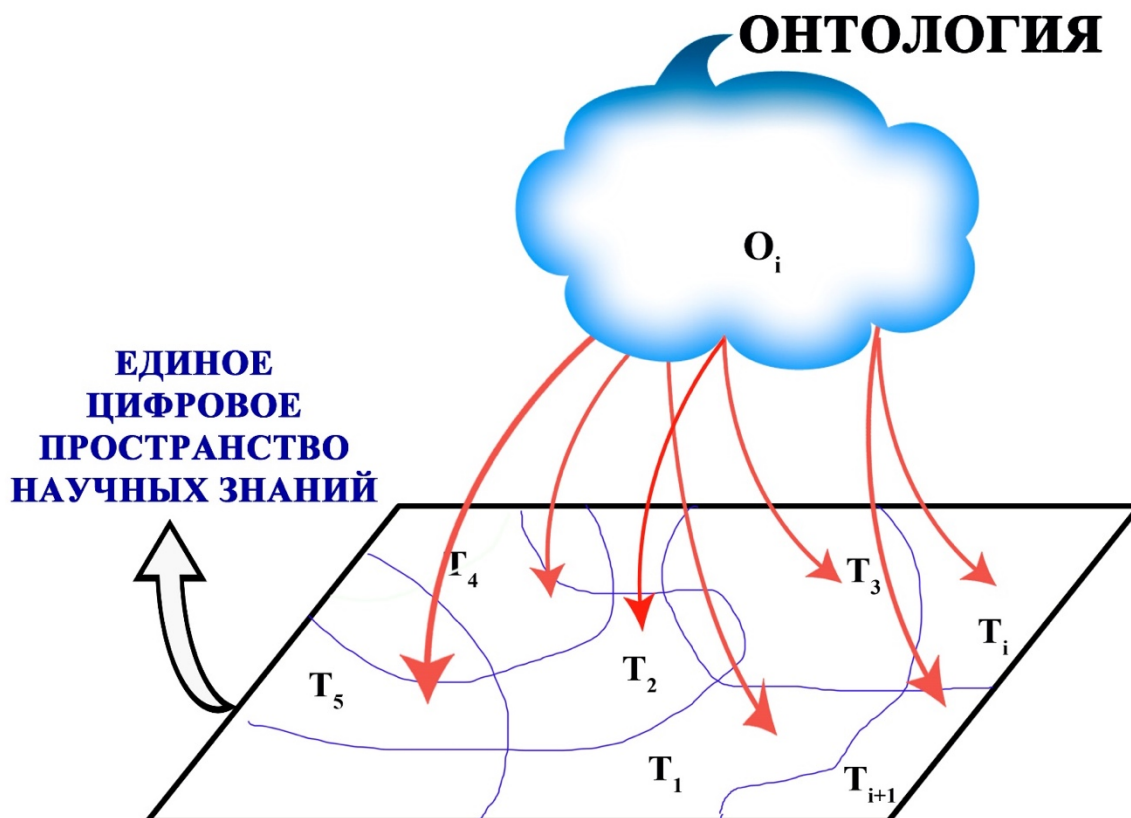


Рисунок 2. - Единое цифровое пространство научных знаний с общей онтологией.

Здесь

$T_1, \dots, T_i, T_{i+1}, \dots$ - тематические подмножества ЕЦПНЗ, представляющие собой ресурсы по различным областям науки (физике, математике, биологии и т.п.). Эти подмножества могут быть независимыми, а могут (в большинстве случаев) пересекаться.

O_i - множество предметных онтологий каждого тематического подпространства T_i

Каждое подпространство ЕЦПНЗ включает аксиомы и фундаментальные результаты, лежащие в основе данного научного направления, а также динамическую часть, содержащую информацию о новейших достижениях в данной научной области.

Для каждой отдельной области науки определены конкретные научные знания, но почти в любой области можно выделить два класса знаний - теоретические и экспериментальные [3].

Теоретические знания - совокупность аксиом, методов исследования и базовых результатов, полученных на основе применения этих методов. Аксиомы и методы исследований являются формальными и признаны научным сообществом со всеми устоявшимися элементами и понятиями [1]. При этом базовые результаты теоретических исследований - понятие, которое определяется экспертными оценками научного сообщества.

Экспериментальные знания - совокупность знаний и умений, приобретённых человеком в процессе взаимодействия с внешним по отношению к нему миром [1]. Эти знания основаны на результатах применения теоретических знаний, подтверждённых практикой. Очевидно, что такое деление относительно, поскольку множество теоретических знаний возникло в результате экспериментов (законы физики, механики и других естественных наук).

Источником наполнения ЕЦПНЗ являются ресурсы цифрового научного пространства, которые должны быть проанализированы на предмет достоверности, важности и актуальности.

В последнее время цифровое научное пространство активно пополняется ресурсами различного вида - формируются репозитории открытого доступа (<https://arxiv.org>) <https://ru.wikipedia.org/wiki/ArXiv.org> развиваются научные и образовательные социальные сети. Научные социальные сети предоставляют многочисленные сервисы по обмену разнообразной

информацией, размещению результатов исследований, их рецензированию и комментированию, поиску вакансий и т. п. [3].

Примерами научных социальных сетей могут служить такие сервисы, как:

Academia (www.academia.edu);

ResearchGate (www.researchgate.net);

Social Science Research Network (www.ssrn.com);

Mendeley (www.mendeley.com);

Citeulike (www.citeulike.org);

Myexperimen (www.myexperiment.org);

Соционет (<https://socionet.ru>);

Учёные России (www.russian-scientists.ru);

Научная сеть (www.scipeople.ru);

Социальная научная сеть (<https://www.science-community.org/>).

В настоящее время активно развиваются научные электронные библиотеки, предоставляющие широкому кругу пользователей такие сервисы, как:

Поиск метаданных документов по различным научным дисциплинам и источникам;

Полнотекстовый поиск в цифровых документах;

Предоставление доступа к найденным информационным ресурсам;

Обеспечение сохранности коллекций цифровых документов;

Примером такой научной электронной библиотеки, предоставляющей пользователю перечисленные выше сервисы, может служить Национальная цифровая библиотека по науке США (The National Science Digital Library – NSDL, <https://nsdl.oercommons.org>) [4, 5]. Этот ресурс основан Национальным научным фондом (NSF) в 2000 г. NSDL агрегирует ресурсы из цифровых библиотек и проектов, поддерживаемых NSF, а также формирует собственные обзоры вебсайтов по своим направлениям.

Среди достаточно развитых российских научных электронных библиотек необходимо выделить электронную библиотеку «Научное наследие России» (ЭБ ННР) [2, 3, 6]. Эта библиотека представляет собой объектно-ориентированную информационную систему, интегрирующую разнородные научные ресурсы институтов памяти (библиотек, архивов, музеев). Пользовательский блок ЭБ ННР (<http://e-heritage.ru>) предоставляет возможность многоаспектного поиска таких объектов как «персона», «публикация», «музейный предмет», «архивный документ», и навигации по найденным ресурсам.

Одним из недостатков электронных библиотек является поисковый сервис, не обеспечивающий достаточную точность и полноту поиска. В частности, в некоторых из них не предусмотрено использование при формировании запросов булевой логики.

Кроме электронных библиотек в интернете представлены различные коммерческие и некоммерческие научные сервисно-поисковые системы [7, 8]. Можно выделить следующие наиболее распространенные сервисно-поисковые системы:

Google Scholar (<https://scholar.google.ru/>) – поисковая система научной литературы;

Scholar.ru (<http://www.scholar.ru>) – поисковая система научных публикаций. Этот проект создан для упрощения поиска научных документов на русском языке. Основная цель этой поисковой системы – сбор информации о свободно скачиваемых научных публикациях;

Science Research Portal (<http://www.science-research.com/scienceresearch>) – научная поисковая система, осуществляющая полнотекстовый поиск в журналах многих крупных научных издательств, таких как Elsevier, Highwire, IEEE, Nature и др.;

CiteSeer Publications Research Index (<http://citeseer.ist.psu.edu/>) – научная поисковая система, индексирующая статьи в PostScript и PDF формате с научных веб-сайтов;

Ingenta (<http://www.ingentaconnect.com/>) – библиографическая база данных универсального типа, включающая описания статей из англоязычных журналов и сборников. Объем базы составляет более 13 млн записей [9].

Наиболее популярными источниками справочной информации являются, в частности, такие ресурсы, как Wikipedia и Wikidata. Эти ресурсы не верифицируются научным сообществом, поэтому не могут быть источником достоверных научных знаний.

Описанные выше примеры демонстрируют, что несмотря на то, что каждый из ресурсов мирового информационного пространства играет определённую роль, ни один из них не может претендовать на роль основы единого цифрового пространства научных знаний. Это связано, в том числе, с тем, что:

В большинстве случаев, имеющиеся ресурсы ориентированы на предоставление сведений о публикациях, которые содержат запрашиваемую информацию. В ряде случаев информационный сервис предоставляет полные тексты таких публикаций. Исключение составляют узкотематические фактографические системы, которые ориентированы на выдачу конкретных научных знаний. Они могут являться элементами соответствующих тематических подпространств, но не охватывают все виды объектов и не обеспечивают связи между подпространствами.

Контент политематических ресурсов (Wikipedia, Wikidata, Europeana) является случайным, и часто содержит неточную, непроверенную или неверную информацию [10].

Во всех системах отсутствует разделение контента по уровням на фундаментальные знания, научно-популярные и учебные материалы.

Формирование ЕЦПНЗ предполагает разработку специальных подходов и алгоритмов, которые базируются на новых принципах.

Структура ЕЦПНЗ

Пространство научных знаний должно включать две составляющие – статическую и динамическую. Статическая составляющая – это основополагающие, проверенные временем и практикой теоретические и экспериментальные данные. Динамическая составляющая – часть ЕЦПНЗ, охватывающая появляющиеся новые знания.

Эти составляющие можно рассматривать как две части пространства знаний, одна из которых (базис) отражает зафиксированные научные знания, другая (надстройка) – новую научную информацию. При этом, после прохождения экспертного фильтра, вторая часть переходит в первую.

Связи между базисом и надстройкой ЕЦПНЗ могут быть реализованы, в том числе, на уровне междисциплинарной общенаучной онтологии [1]. При этом, базис и надстройка представляют собой совокупность подпространств (фасетов) по различным научным направлениям.

Базис включает три взаимосвязанных уровня:

Образовательный, рассчитанный, в частности, на учащихся школ. Этот уровень статический, формируется на основе информации, полученной из академической литературы, например, учебников, научно-популярной литературы. На этом же уровне могут быть представлены верифицированные мультимедийные ресурсы, например, оцифрованные музейные объекты.

Научно-популярный, рассчитанный на более «продвинутых» пользователей, но не специалистов в данной области. Этот уровень квазистатический. Он может изменяться при появлении новых существенных для данной области знания открытий или результатов. Формирование этого уровня строится на основе энциклопедической и справочной информации, а также монографий и учебников для высшей школы. На этом уровне наряду с извлечёнными метаданными, пользователь может получить ссылки на источники и их цифровые копии в соответствии с законом об охране авторских прав.

Фундаментальный, предназначенный для специалистов в данной области науки. Этот уровень включает более глубокую по сравнению с предыдущими двумя уровнями информацию по данному научному направлению. Которая, также является статической. Контент на этом уровне может быть дополнен и подтвержден экспертными оценками новой информацией.

На рисунке 3 показано распределение новой информации по трем базисным уровням.



Рисунок 3 – Распределение новой информации по трем базисным уровням

Надстройка, ориентированная на специалистов в данной области науки, является источником текущего контента, необходимого для проведения научных исследований. Она содержит новые результаты в той или иной научной области. Контент надстройки формируется за счёт новой информации, прошедшей экспертизу, например, информации из статей, опубликованных в верифицированных журналах, из сборников научных статей с подтверждённым научным уровнем, патентов, материалов, ведущих мировых научных конференций и пр. Наполнение надстройки, динамически меняется. Это происходит по причине исключения информации, ценность которой не подтвердилась временем, или благодаря новой, проверенной информации.

Контент ЕЦПНЗ

Информационные ресурсы являются источниками научных знаний. Они содержат описания постулатов, теорий, экспериментов, полученных на их основе результатов в той или иной области науки, представленные на материальном носителе. Материальные носители эволюционировали от глиняных табличек и папирусов до печатных, фото-, кино-, видеоматериалов и цифровых музейных моделей [11-15]. Современные цифровые носители обладают качественно иными возможностями с точки зрения компактного хранения, извлечения, поиска и предоставления знаний пользователям.

Базис ЕЦПНЗ формируется из имеющихся библиотечных, архивных, музейных, энциклопедических, фактографических, словарных и других ресурсов. Этот контент создается научно-образовательным сообществом на основе существующих систем без нарушения их функционирования. Как правило, информация, содержащаяся в этих ресурсах, является достоверной и проверенной временем. Тем не менее, для принятия решения о том, что будет загружено в ЕЦПНЗ, необходима экспертиза. Экспертами должны выступать квалифицированные представители научного сообщества в области, соответствующей данному подпространству.

Динамическая часть (надстройка) ЕЦПНЗ формируется информационно-библиотечными специалистами совместно с учёными на основе существующих мировых баз данных научной информации и коллекций полнотекстовых цифровых документов.

Базис и надстройка ЕЦПНЗ состоят из ядра и оболочки.

Элементы ядра представляют собой информационные объекты, которые содержат метаданные, полученные из перечисленных выше ресурсов, определяющих основные постулаты, методы и результаты исследований в конкретной области науки для базиса и надстройки.

Каждый информационный объект должен:

Содержать элементы типа «Что (кто), где, когда»;

Указывать на связи с другими объектами как данного подпространства, так и, в общем случае, других подпространств;

Иметь ссылку на источник информации.

Оцифрованные публикации, архивные материалы, образы музейных экспонатов, мультимедийные материалы, тематические базы данных, поддерживаемые в научных организациях, образуют *оболочку ЕЦПНЗ*.

Предметная онтология ЕЦПНЗ

Предметные онтологии различных научных областей широко применяются в задачах построения поисковых систем, систем представления знаний, а также при решении различного рода семантических задач [16-18]. Процедура введения онтологических представлений в накопленный массив данных представляет собой определение предметной области через описание множества используемых в ней понятий и связей между ними [19]. Одной из задач формирования ЕЦПНЗ является разработка платформы, которая позволит автоматизировать процесс построения связей между объектами предметной области. Эта платформа может выступать в роли независимой информационной среды, а может быть внедрена в качестве одного из элементов в более масштабную информационную среду, создаваемую, в том числе, для образовательного процесса и научных исследований. Одной из важнейших задач при построении ЕЦПНЗ является формирование обобщенной предметной онтологии, включающей единые тезаурусы общенаучных терминов, а также словари номенклатурных наименований, географических названий, имён персоналий, событий, обозначения временных интервалов, тематики и другие категории научной лексики. Обобщенная онтология должна строиться на базе предметных онтологий, которые создаются внутри каждого подпространства, отражающего конкретное научное направление. Эти онтологии включают ключевые слова и корреляции между ними, а также средства, обеспечивающие взаимодействие различных систем классификации научных знаний в той или иной научной области.

Унифицированное представление связей между объектами, единые принципы описания информационных объектов, единые терминологические словари и тезаурусы, единые принципы оцифровки источников информации – поддержка всего этого обеспечивает единство пространства научных знаний.

Операции над объектами ЕЦПНЗ

Для ЕЦПНЗ определены следующие операции над информационными объектами:

Поиск объектов по любым элементам метаданных и их комбинациям в каждом из подпространств;

Поиск объектов по любым элементам метаданных и их комбинациям в цифровом пространстве в целом;

Выдача списков найденных объектов;

Сортировка найденных объектов по заданному элементу;

Навигация от найденных объектов к связанным с ними объектам, находящимися, как внутри подпространства, так и в пространстве в целом с использованием возможностей общей онтологии ЕЦПНЗ;

Формирование различных коллекций объектов, отвечающих заданному параметру.

Для числовых значений метаданных определены операции:

«равно»;

«не равно»;

«меньше»;

«больше».

Для текстовых объектов определены операции:

«равно» – это значение предполагает точное совпадение элемента метаданных с заданным фрагментом текста с возможностью выбора значения из перечислимого множества);

«начинается с...»;

«содержит»;

«не содержит»;

правое и левое усечения от начала слова.

Для данных типа «дата» определены операции:

«равно»;

«до заданной даты»;

«после заданной даты»;

«в заданном интервале».

Поисковый запрос может содержать множество элементов, связанных операторами:

«и»;

«или»;

«и не».

Развитие ЕЦПНЗ предполагает расширение наполнения и увеличение номенклатуры входящих в него объектов, что, в свою очередь, повлечет появление новых запросов и потребует введения новых операций на множестве объектов ЕЦПНЗ.

Управление ЕЦПНЗ

Единое пространство знаний формируется на основе распределённой (децентрализованной) подготовки метаданных. Эти метаданные должны создаваться по единым правилам с централизованным хранением при единой централизованной системе контроля качества. Материалы, относящиеся к частям ЕЦПНЗ (оцифрованные публикации, архивные документы, музейные предметы и т. п.), могут храниться у владельцев этих данных или в центральном блоке пространства.

С функциональной точки зрения управление ЕЦПНЗ включает три уровня [20]:

Верхний уровень управления обеспечивает центральный блок, связанный через обобщенную онтологию с блоками управления ядрами тематических подпространств ЕЦПНЗ, который, в свою очередь связан с блоками управления данными оболочек подпространств (рис. 4).

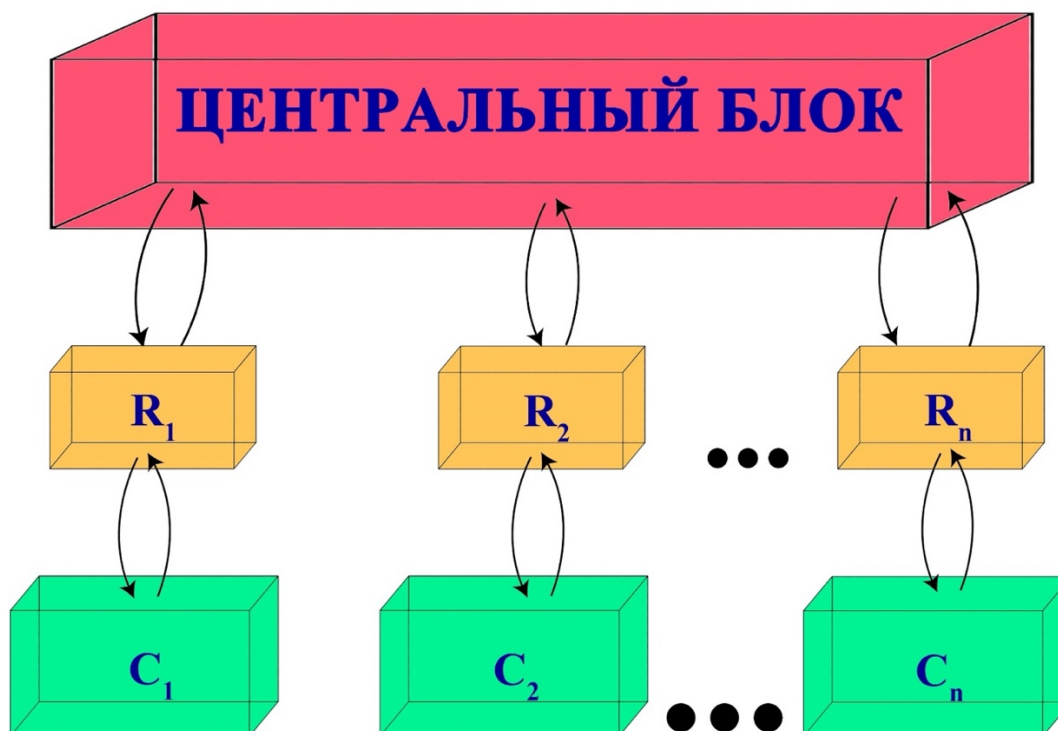


Рисунок 4. Структура управления ЕЦПНЗ

Здесь

$R_1, R_2 \dots R_n$ – блоки управления ядрами тематических подпространств ЕЦПНЗ;

$C_1, C_2 \dots C_n$ – блоки управления оболочками тематических подпространств ЕЦПНЗ.

Блоки управления ядрами обеспечивают загрузку и поддержку содержимого ядра каждого подпространства. Блоки управления оболочками – обеспечивают загрузку цифровых объектов и их связи с ядрами.

Заключение

Единое цифровое пространство научных знаний должно создаваться усилиями научных организаций, библиотек, музеев и архивов. При этом необходим мониторинг существующих ресурсов; определение принципов отбора наиболее значимых объектов, разработка соответствующей программной оболочки и технологии наполнения контента. Необходимо разработать общесистемную онтологию и начать поэтапную реализацию отдельных тематических подпространств. Работы по перечисленным выше направлениям ведутся в настоящее время в МСЦ РАН в кооперации с другими академическими институтами и научными библиотеками в рамках государственного задания, а также при поддержке РФФИ.

Работа выполнена в МСЦ РАН – филиале ФГУ ФНЦ НИИСИ РАН рамках государственного задания 065-2019-0014.

Литература

1. Антопольский А.Б., Каленов Н.Е., Серебряков В.А., Сотников А.Н., 2019. О едином цифровом пространстве научных знаний. Вестник Российской Академии Наук. Т. 89(7), сс. 728-735.
2. Антопольский А.Б., 2019. О разработке Единого российского электронного пространства знаний. Информации и инновации. Т. 14(1), сс. 7-17.
3. Антопольский А.Б., Ефременко Д.В., 2018. К вопросу о едином электронном пространстве знаний. Вестник Российской Академии Наук. Т. 88(1), сс. 89-95.

4. Andrews-Todd, J., Forsyth Carol M., 2018. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. In *Computers in human behaviour*. Vol. 104, UNSP 105759.
5. Arshad H., Jantan A., Hoon G.K., Abiodun I.O., 2020. Formal knowledge model for online social network forensics. In *Computers & security*. Vol. 89, UNSP 101675.
6. Погорелко К.П., 2015. Анализ спроса на электронную библиотеку «научное наследие России». Информационное обеспечение науки: новые технологии. Сборник научных трудов, сс. 191-199.
7. Zhmailo S.V., Ulyanin O.V.. Sci-tech libraries within the knowledge management system: from information specialist's viewpoint. 2020. In *Nauchnye i tekhnicheskie biblioteki-scientific and technical libraries*. vol. 2, pages 9–23.
8. Sanchez A.S., 2017. Knowledge Organization Systems: Definition and Historical Development. In *E-ciencias de la information*. vol. 7(2).
9. Corrales D.C . Ledezma, A., Corrales, J.C., 2020. A case-based reasoning system for recommendation of data cleaning algorithms in classification and regression tasks. In *Applied soft computing*. Vol. 90, UNSP 106180.
10. Deena G., Raja K., 2019. Designing an Automated Intelligent e-Learning System to Enhance the Knowledge using Machine Learning Techniques. I n *International journal of advanced computer science and applications*. Vol. 10(12), pages 112-119.
11. Dresel R., Henkel M., Scheibe K., Zimmer F., Stock WG., 2020. A Nationwide Library System and Its Place in Knowledge Society and Smart Nation: The Case of Singapore. In *LIBRI*. vol. 70(1), pages 81-94.
12. Frandsen T.F., Tibyampansha D., Ibrahim G.R., von Isenburg M., 2017. Library training to promote electronic resource usage: A case study in information literacy assessment. In *Information and Learning Science*. vol. 118(11-12), pages 618-628.
13. Gabrys J., 2020. Smart forests and data practices: From the Internet of Trees to planetary governance. In *Big data & society*. Vol. 7(1), UNSP 2053951720904871.
14. Gushul Y.V., Teslya E.V., 2020. Information and analytical support: current goals and pathways. In *Nauchnye i tekhnicheskie biblioteki-scientific and technical libraries*, pages 24-44.
15. Maggio A., Kuffer J., Lazzari M., 2017. Advances and trends in bibliographic research: Examples of new technological applications for the cataloging of the georeferenced library heritage. In *Journal of librarianship and information science*. vol. 49(3), pages 299-312.
16. Yeh J.H., Huang X.M., 2018. BKOntoVR: A Virtual Reality Exhibition System for Biographic Ontology-Based Semantic Structure. In *Proceedings of 2018 2nd international conference on software and e-business*, pages 69-73.
17. Lou W., Pi RF., Wang H., Ju Y., 2019. Low-cost similarity calculation on ontology fusion in knowledge bases. In *Journal of information science*. SCITEPRESS.
18. Bonacchi C., Krzyzanska M., 2019. Digital heritage research re-theorised: ontologies and epistemologies in a world of big data. In *International journal of heritage studies*. Vol. 25(12), 1235-1247.
19. Shahzad F., Alwosaibi F.M. 2017. Development of an e-Library Web application. IMSCI. In *11th International Multi-Conference on Society; Orlando; the United States, Cybernetics and Informatics*. Pages 153-158.
20. Kalenov N.E., Sobolevskaya I.N., Sotnikov A.N., 2019. Hierarchical representation of information objects in a digital library environment. In *17th Russian Conference, RCAI 2019, Ulyanovsk, Russia*. SCITEPRESS.

COMMON DIGITAL SPACE OF SCIENTIFIC KNOWLEDGE IN THE WORLD INFORMATION SPACE

Kalenov, Nikolai Evgenievich

Doctor of technical sciences, professor

*Joint Supercomputer Center of the Russian Academy of Sciences – Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, chief researcher
Moscow, Russia
nekalenov@mail.ru*

Sobolevskaya, Irina Nikolaevna

Candidate of physical and mathematical sciences

*Joint Supercomputer Center of the Russian Academy of Sciences – Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, senior researcher
Moscow, Russia
ins@jsc.ru*

Sotnikov, Alexander Nikolaevich

Doctor of physical and mathematical sciences

*Joint Supercomputer Center of the Russian Academy of Sciences – Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, deputy director
Moscow, Russia
asotnikovov@mail.ru*

Abstract

The scientific space is a subspace of the existing information space. The scientific space contains all the scientific information presented in various forms (printed sources, digital documents, audio and video materials, images, 3D-models of various objects, etc.). In turn, the digital scientific space is also a subset of the entire information space, which contains scientific information in digital form. Digital Knowledge Space (DKS) is a subset of the digital space that contains reliable information. The Common digital space of scientific knowledge (CDSSK) is a subset of the DKS containing verified scientific information presented in accordance with a single ontology. Ontology is understood as a set of rules for constructing and providing users with the elements of the CDSSK. Therefore, in order to build the CDSSK, it is necessary to determine its elements and the ontology of their presentation. The article discusses the structure of the CDSSK as a set of knowledge subspaces in various scientific areas, the types and types of data that make up the content of the CDSSK, operations with them, the principles of space management.

Keywords

digital knowledge space, information space, ontology, digital library, subject ontologies

References

1. Antopol'skij A.B., Kalenov N.E., Serebryakov V.A., Sotnikov A.N., 2019. O edinom cifrovom prostranstve nauchnyh znaniy. Vestnik Rossijskoj Akademii Nauk. T. 89(7), ss. 728-735.
2. Antopol'skij A.B., 2019. O razrabotke Edinogo rossijskogo elektronnoho prostranstva znaniy. Informacii i innovacii. T. 14(1), ss. 7-17
3. Antopol'skij A.B., Efremenko D.V., 2018. K voprosu o edinom elektronnom prostranstve znaniy. Vestnik Rossijskoj Akademii Nauk. T. 88(1), ss. 89-95.
4. Andrews-Todd, J., Forsyth Carol M., 2018. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. In Computers in human behaviour. Vol. 104, UNSP 105759.
5. Arshad H., Jantan A., Hoon G.K., Abiodun I.O., 2020. Formal knowledge model for online social network forensics. In Computers & security. Vol. 89, UNSP 101675.
6. Pogorelko K.P., 2015. Analiz sprosa na elektronnyuyu biblioteku «nauchnoe nasledie Rossii». Informacionnoe obespechenie nauki: novye tekhnologii. Sbornik nauchnyh trudov, ss. 191-199.

7. Zhmailo S.V., Ulyanin O.V.. Sci-tech libraries within the knowledge management system: from information specialist's viewpoint. 2020. In *Nauchnye i tekhnicheskie biblioteki-scientific and technical libraries*. vol. 2, pages 9–23.
8. Sanchez A.S., 2017. Knowledge Organization Systems: Definition and Historical Development. In *E-ciencias de la information*. vol. 7(2).
9. Corrales D.C . Ledezma, A., Corrales, J.C., 2020. A case-based reasoning system for recommendation of data cleaning algorithms in classification and regression tasks. In *Applied soft computing*. Vol. 90, UNSP 106180.
10. Deena G., Raja K., 2019. Designing an Automated Intelligent e-Learning System to Enhance the Knowledge using Machine Learning Techniques. In *International journal of advanced computer science and applications*. Vol. 10(12), pages 112-119.
11. Dresel R., Henkel M., Scheibe K., Zimmer F., Stock WG., 2020. A Nationwide Library System and Its Place in Knowledge Society and Smart Nation: The Case of Singapore. In *LIBRI*. vol. 70(1), pages 81-94.
12. Frandsen T.F., Tibyampansa D., Ibrahim G.R., von Isenburg M., 2017. Library training to promote electronic resource usage: A case study in information literacy assessment. In *Information and Learning Science*. vol. 118(11-12), pages 618-628.
13. Gabrys J., 2020. Smart forests and data practices: From the Internet of Trees to planetary governance. In *Big data & society*. Vol. 7(1), UNSP 2053951720904871.
14. Gushul Y.V., Teslya E.V., 2020. Information and analytical support: current goals and pathways. In *Nauchnye i tekhnicheskie biblioteki-scientific and technical libraries*, pages 24-44.
15. Maggio A., Kuffer J., Lazzari M., 2017. Advances and trends in bibliographic research: Examples of new technological applications for the cataloging of the georeferenced library heritage. In *Journal of librarianship and information science*. vol. 49(3), pages 299-312.
16. Yeh J.H., Huang X.M., 2018. BKOntoVR: A Virtual Reality Exhibition System for Biographic Ontology-Based Semantic Structure. In *Proceedings of 2018 2nd international conference on software and e-business*, pages 69-73.
17. Lou W., Pi RF., Wang H., Ju Y., 2019. Low-cost similarity calculation on ontology fusion in knowledge bases. In *Journal of information science*. SCITEPRESS.
18. Bonacchi C., Krzyzanska M., 2019. Digital heritage research re-theorised: ontologies and epistemologies in a world of big data. In *International journal of heritage studies*. Vol. 25(12), 1235-1247.
19. Shahzad F., Alwosaibi F.M. 2017. Development of an e-Library Web application. IMSCI. In *11th International Multi-Conference on Society; Orlando; the United States, Cybernetics and Informatics*. Pages 153-158.
20. Kalenov N.E., Sobolevskaya I.N., Sotnikov A.N., 2019. Hierarchical representation of information objects in a digital library environment. In *17th Russian Conference, RCAI 2019, Ulyanovsk, Russia*. SCITEPRESS.