



OSTIS-2012

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

МОДЕЛЬ ИНТЕРНЕТ-ПОРТАЛА НАЦИОНАЛЬНОГО ЦЕНТРА АНТАРКТИЧЕСКИХ ДАННЫХ

Глоба Л.С., д.т.н., проф., Кузин И.А., Мочалкина К.С., Новогрудская Р.Л.

*Национальный Технический Университет Украины «Киевский Политехнический Институт»,
г. Киев, Украина*

lgloba@its.kpi.ua

ksu_mochalkina@yahoo.com

ihkuzin@gmail.com

rinan@ukr.net

В статье представлен подход к построению интернет-портала знаний в области антарктических исследований. В качестве информационной модели портала используются онтологии.

Ключевые слова: интернет-портал, представление знаний, модель, онтология.

Введение

В настоящее время осуществляется большое количество исследований в Антарктике в различных научных направлениях. Однако существует проблема доступа к полученным результатам и их дальнейшей обработке. Важные результаты, как теоретические, так и практические, которые получают исследователи, остаются сосредоточенными лишь на уровне конкретных исследовательских институтов. Происходит процесс сосредоточения знаний, они становятся доступными узкому кругу ученых и специалистов. Поэтому актуальной задачей становится предоставление возможности использования этих знаний большому количеству исследователей, для которых они предназначены.

Согласно требованиям Резолюции 4.1 XXII АТСМ 1998 года все Стороны Договора об Антарктике, осуществляющих антарктические исследования, обязаны создать Национальные центры Антарктических Данных (НЦАД). Исходя из этих требований, пока актуальна задача по созданию такого Центра в Украине. Эффективное функционирование НЦАД поможет решить ряд проблем, связанных со сбором, хранением, обработкой и передачей, как первичных данных так и обработанных результатов антарктических исследований широкому кругу ученых и научным учреждениям в Украине, а также заинтересованным зарубежным научным организациям. Кроме этого, создание такого информационного пространства позволит Украине

полноценно войти в мировой антарктического научного сообщества, что повысит эффективность работы наших ученых и поднимет международный авторитет нашего государства. [Национальный Центр Антарктических Данных и конструкций, 2005]

Из всего вышесказанного следует необходимость построения качественной модели представления знаний на портале. В качестве модели представления знаний на портале использовались онтологии.

1. Онтология портала

1.1. Представление знаний в области антарктических исследований

Рассмотрим построение Интернетпортала в области антарктических исследований. Информационную основу портала составляют онтологии – концептуальная модель, с помощью которой осуществляется формализация некоторой области знаний. Вводя формальные описания понятий предметной области в виде классов объектов и отношений между ними, онтология портала задает структуры для представления реальных данных и связей между ними. Использование онтологий для построения информационной основы портала позволяет не только целостно представить такие трудно формализуемые предметные области, как технические, но и автоматизировать процесс

сбора и накопления информации по выбранной тематике. Такая концептуальная модель позволяет обеспечить единообразное представление знаний и данных по выбранной тематике, их семантическую связанность.

Онтологии состоят из следующих элементов: экземпляров, классов, атрибутов и отношений.

- Атрибуты – описывают классы по определенным характеристикам. Каждый атрибут имеет, по крайней мере, имя и значение, и используется для хранения информации, которая специфична для объекта и привязана к нему.

- Отношения – устанавливают связи (описывающие зависимости) между элементами онтологии. Обычно отношением является атрибут, значением которого является другой объект.

- Экземпляры – это основные, нижнеуровневые компоненты онтологии. Экземпляры могут представлять собой как физические объекты (люди, дома, планеты), так и абстрактные (числа, слова).

- Онтология представляет собой иерархию классов, связанных отношениями. Для построения онтологии обычно используются следующие 4 вида отношений:

- Ассоциативное отношение – это отношения задаваемые пользователем. Наличие таких отношений позволяет осуществлять содержательный поиск.

- Отношение часть-целое – это отношение позволяет устанавливать связи между классами на уровне иерархии. При поиске информации, связанной отношениями такого типа, осуществляется транзитивное замыкание.

- Отношение наследования – это отношение позволяющее осуществить передачу всех отношений и атрибутов от родительского класса дочернему.

- Отношение вида класс-данные – это отношение позволяющее связывать конкретные экземпляры понятий с классом.

Базируясь на описанных выше понятиях, построена онтология портала знаний в области антарктических исследований. Формально онтология может быть задана как: $O = \{C, A, R, T, F, D\}$, где C – множество классов, описывающих понятия предметной области; A – множество атрибутов, описывающих свойства понятий и отношений; R – множество отношений, заданных на классах: $R = \{R_{AS}, R_{IA}, R_n, R_{CD}\}$ - отношения следующих типов: R_{AS} - ассоциативное, R_{IA} - «часть-целое», R_n - наследования, R_{CD} - «класс-данные»; T – множество стандартных типов значений атрибутов; F – множество ограничений на значения атрибутов понятий и отношений; D – множество экземпляров классов [Загоруйко Ю.А. и др., 2007].

Определенная таким образом онтология может служить для представления понятий, необходимых

как для описания знаний в области антарктических исследований, так и для выполняемой в ее рамках обработки данных.

1.2. Онтология портала «Национального Центра Антарктических Данных»

Онтология портала в области антарктических исследований состоит из четырех онтологий: онтология внешнего пользователя, онтология антарктического знания, онтология исследователя и онтология предметной области (рис. 1). Онтология внешнего пользователя (ОВП) включает классы, относящиеся к организации доступа внешнего пользователя к материалам портала. Онтология антарктического знания (ОАЗ) содержит следующие метапонятия, задающие структуры для описания рассматриваемой предметной области. Онтология исследователя (ОИ) описывает возможности, предоставляемые исследователям на портале по доступу и обработке размещенных на портале материалов. Онтология предметной области (ОПО) отражает общие знания о предметной области, такие как иерархия классов понятий, семантические отношения на этих классах. [Глобал.С., и др., 2011]



Рисунок 1 – Онтология портала «Национальный Центр Антарктических Данных»

На (рис. 2) приведены классы ОВП, ОАЗ и ОИ, а также отношения заданные на этих классах. Классы отображены заштрихованной овальной областью которая расположена в определенном прямоугольнике, который соответствует той онтологии, к которой принадлежит данный класс.

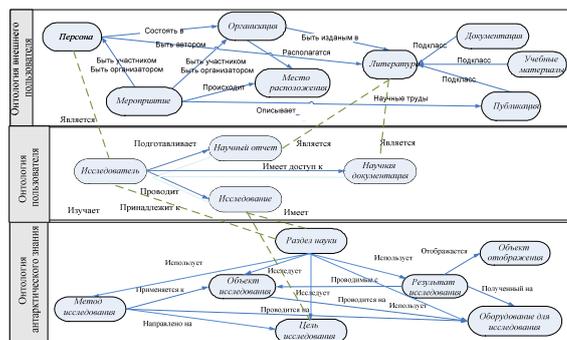


Рисунок 2 – Элементы онтологий портала

Для примера рассмотрим формальное описание онтологии внешнего пользователя, которая включает восемь общих классов понятий, относящиеся к организации научной деятельности, которые связаны отношениями различных типов. Различные свойства каждого понятия описываются на основе атрибутов понятий и ограничений, наложенных на область их значений. Исходя из этого ОВП может быть описана следующим образом:

$$O_1 = \{C_{o_1}, A_{o_1}, R_{o_1}, T_{o_1}, F_{o_1}, D_{o_1}\}, \text{ где} \quad (1)$$

$$C_{o_1} = \{C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9\},$$

$$A_q = (A_{C_1}, A_{C_2}, A_{C_3}, A_{C_4}, A_{C_5}, A_{C_6}, A_{C_7}, A_{C_8}, A_{C_9}),$$

$$D_q = (D_{C_1}, D_{C_2}, D_{C_3}, D_{C_4}, D_{C_5}, D_{C_6}, D_{C_7}, D_{C_8}, D_{C_9}),$$

$$R_{o_1} = (R_{AS_1}(O_1), \dots, R_{AS_{10}}(O_1), R_{L_1}(O_1), \dots, R_{L_6}(O_1),$$

$$R_{n_1}(O_1), \dots, R_{n_6}(O_1), R_{CD_1}(O_1), \dots, R_{CD_3}(O_1)).$$

ОВП включает следующие классы понятий:

Персона - $C_1(O_1)$. К этому классу относятся понятия, связанные с субъектами научной деятельности: исследователями, сотрудниками и членами организаций. Атрибутами персоны являются: персональные данные, ученая степень, звание, направления научной деятельности, место проживания.

$$C_1(O_1) = (A_{C_1}, D_{C_1}) \quad (2)$$

Организация - C_2 . Понятия этого класса описывают различные организации, научные сообщества и ассоциации, институты, исследовательские группы и другие объединения. Атрибутами организации являются: название и место расположения.

$$C_2(O_1) = (A_{C_2}, D_{C_2}) \quad (3)$$

Мероприятие - C_3 . В этот класс входят понятия, описывающие научно-организационную или научно-исследовательскую деятельность – научные мероприятия, конференции, исследовательские поездки, проекты, программы и т.п. К атрибутам события относятся: название, место проведения, дата начала, дата окончания, степень завершенности.

$$C_3(O_1) = (A_{C_3}, D_{C_3}) \quad (4)$$

Публикация - C_4 . Этот класс служит для описания различного рода публикаций, (в периодических изданиях и тех, которые издаются в результате проведения конференций, научных мероприятий и т. д.). К атрибутам публикации относятся: название, описание, дата публикации и язык публикации.

$$C_4(O_1) = (A_{C_4}, D_{C_4}) \quad (5)$$

Местоположение - C_5 . Этот класс понятий позволяет описывать географическую и административно-территориальную локализацию объектов исследования, организаций и т.п. Атрибутами этого класса являются название местоположения и географический тип.

$$C_5(O_1) = (A_{C_5}, D_{C_5}) \quad (6)$$

Литература - C_6 . Этот класс служит для описания литературы (представленных в печатном или электронном форматах), которая используется в деятельности исследователей (монографии, статьи, отчеты, труды конференций, периодические издания, фото- и видеоматериалы и др.). К атрибутам публикации относятся: название, описание, дата публикации и язык публикации.

$$C_6(O_1) = (A_{C_6}, D_{C_6}) \quad (7)$$

Документация - C_7 . Этот класс служит для описания различного рода документации, такой как ГОСТЫ, ISO и т. д. К атрибутам документация относятся: название, описание, дата публикации и язык.

$$C_7(O_1) = (A_{C_7}, D_{C_7}) \quad (8)$$

Учебные материалы - C_8 . Этот класс содержит существующие учебники, справочники, учебные пособия, мануалы... К атрибутам класса относятся: название, описание, дата публикации и язык.

$$C_8(O_1) = (A_{C_8}, D_{C_8}) \quad (9)$$

1.3. Онтология предметной области

Онтология предметной области отражает общие знания о предметной области, такие как иерархия классов понятий, семантические отношения на этих классах. Основой онтологии предметной области для портала антарктических знаний в области антарктических исследований послужило постановление Кабинета Министров Украины.

Онтология предметной области (ОПО) описывает классификацию направлений исследований, которые проводятся на антарктической станции (рис. 3), в различных разделах науки и включает формальное и неформальное описание понятий и отношений между ними.

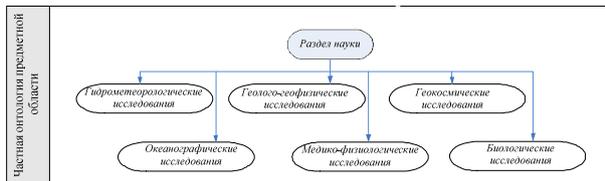


Рисунок 3 – Классификация направлений исследований НЦАД.

Эти понятия являются реализациями метапонятий онтологии Антарктического знания и могут быть упорядочены в иерархию общее-частное и часть-целое. Так, например, *Объектами* исследования соответствуют природные явления Антарктики такие как: рельеф дна и суши, наземные и подводные экосистемы, электромагнитное поле и др., а в качестве *Методов* исследования выступают такие понятия, как: сейсмографические исследования, зондирование почвы, прогнозирование погоды и др. Пример предметной области для геолого-геофизических исследований приведен на (рис. 4). [ГлобалС., и др., 2010]

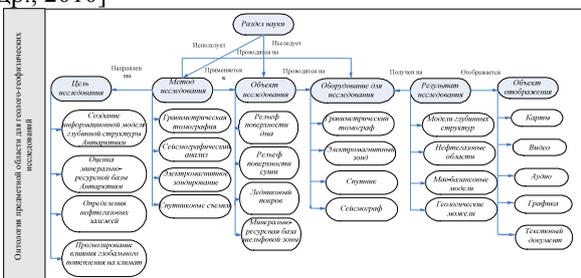


Рисунок 4 – Фрагмент модели предметной области.

2. Функции портала

Согласно рассмотренной выше онтологии портал НЦАД должен решать ряд функциональных задач:

- Качественное представления знаний на портале
- Систематизация и структурирование информации
- Формализация антарктических знаний
- Эффективный и содержательный поиск

Портал обеспечит доступ к базам данных, справочной литературе, мануалам, эксперсинформации, ресурсам сети и т. пр. На портале возможна реализация просмотра блоков новостей, информации о конференциях (предстоящих и прошедших), конкурсах и грантах, а также другая информация о событиях, имеющих отношение к данной области знаний. Важным этапом построения портала является структуризация и систематизация информации и знаний в области антарктических исследований, что позволит конечному пользователю осуществлять просмотр и поиск конкретных сведений в рассматриваемой сфере. Структуризация и размещение на портале организуется удобным для конечного пользователя образом, реализуя проблемно-ориентированные

средства навигации и поиска по информационному пространству портала. При этом поиск информации организован так, что пользователь имеет возможность задавать запрос не только и не столько по ключевым словам, сколько в знакомых ему терминах предметной области портала (рис. 5).

Функционально Национальный Центр Антарктических Данных:

- создает Единое информационное пространство для представления результатов антарктических исследований украинских ученых в среде Интернет;
- обеспечивает единую точку доступа ко всем видам научных ресурсов;
- использует такие базовые подходы, как:
 - одновременность системной разработки, реализации и использования данных при обеспечении одновременной работы всех ученых;
 - стандартизацию всех типовых компонентов информационных процессов;
 - интеграцию как метод организации взаимодействия отдельных компонентов;
 - интеллектуализацию как метод вычислений необходимых аналитических и статистических данных.

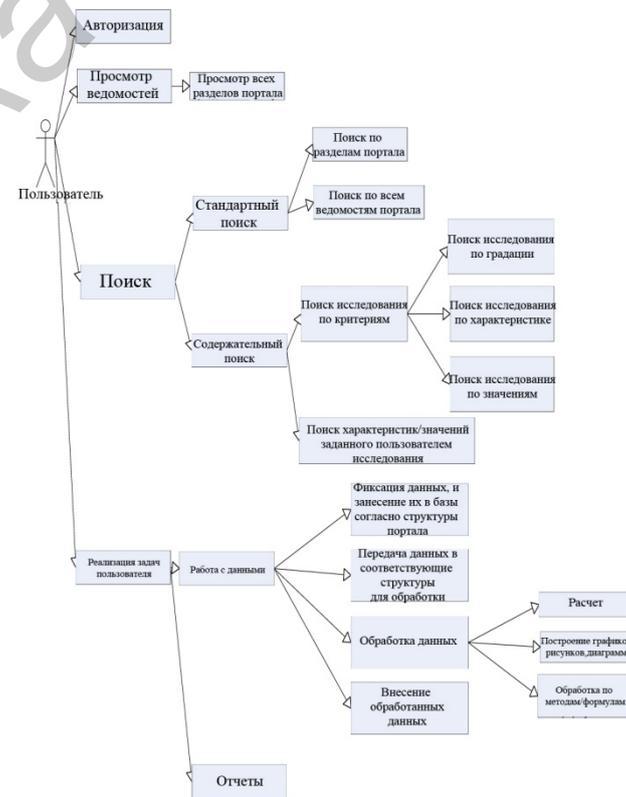


Рисунок 5 – Общая функциональная схема

Для поддержки процессов сбора, хранения, обработки и передачи данных необходимо учесть наличие таких составных частей единой информационной среды:

- необходимо учесть наличие таких составных частей единой информационной среды:
- портала, который используется как единая точка входа;
- единой сервисной шины для объединения всех сервисов хранения, обработки и представления информационных ресурсов антарктических исследований;
- базы метаданных и онтологий, которые задают структуры для представления реальных данных и связей между ними, позволяют не только целостно представить всю необходимую информацию о данных антарктических исследований, но и обеспечить их семантическую связанность;
- базы актуальных данных антарктических исследований; технической документации; мониторинга состояния оборудования; сохранения и анализа статистических данных;
- базы геоданных для отображения информации в картографическом виде;
- программы для сохранения, резервного копирования и восстановления информационных ресурсов;
- хранилища архивных ресурсов для сохранения исторических информационных ресурсов антарктических исследований в разрезе времени;
- средств и технологий интерактивного взаимодействия.

В процессе разработки программного и информационного обеспечения необходимо предусмотреть работу системы в следующих основных режимах:

- Режим обычного пользователя;
- Режим научного исследователя;
- Режим администрирования банка данных.

Режим обычного пользователя предусматривает доступ к уровню портала включающий общедоступную информацию, такую как описание деятельности НЦАД, направлений исследований, проводимых на станции Академик Вернадский и другие. Главное требование к режиму обычного пользователя - возможность доступа к portalу с любой рабочей станции.

Режим научного исследователя предполагает доступ к рабочим корпоративным материалам, научным результатам и другим необходимым данным только специалистам, работающим с данной информацией, в том числе и на Украинской антарктической станции Академик Вернадский. Главное требование к режимам научного исследователя в процессе функционирования системы - асинхронность, то есть, независимость одного от другого. Этот режим должен предоставлять пользователю возможность выполнять необходимые работы с учетом распределения доступа к базе данных и ее отдельных элементов. Вместе с тем предусматривается возможность доступа к

конкретным элементам базы данных, к более детализированной информации.

Режим администрирования банка данных должен включать в себя все виды работ, выполняемых обслуживающим персоналом для поддержки всех составных частей системы в работоспособном состоянии. Функционирование режима администрирования не должно мешать пользователям портала. Если этого избежать невозможно, необходимо предусмотреть переход пользователей в автономный режим работы или предупредить о временном прекращении работы пользователей с системой. Как правило, режим администрирования должен выполняться в нерабочее для большинства пользователей время. Режим администрирования системы должен обеспечивать выполнение необходимых действий связанных с хранением и защитой информации от возможного разрушения по техническим причинам, подключение новых пользователей, установка паролей и уровня доступа пользователей к базам данных и их составных частей и отдельных элементов и другие функции.

Система должна обеспечивать защиту информации от несанкционированного доступа с помощью системы паролей, определяющих уровень доступа к базе данных и ее составных частей

ЗАКЛЮЧЕНИЕ

Рассмотрена возможность построения портала «Национальный Центр Антарктических Данных». Информационную основу портала составляют онтологии, с помощью которой осуществляется систематизация и структуризация информации, организовывается эффективный поиск и навигация по информационному пространству портала Национального Центра Антарктических Данных. Также рассмотрены функции портала реализуемые на основании созданной онтологии.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Национальный Центр Антарктических Данных и конструкций, 2005] Национальный Центр Антарктических Данных и конструкций / Редкол.: В. Т. Троценко (отв. ред.) и др. – К.: Академперіодика, 2005. – 1088с.
- [Загорюлько Ю.А. и др., 2007] Загорюлько Ю.А., Боровикова О.И. Технология построения онтологий для порталов научных знаний / Загорюлько Ю.А., Боровикова О.И. // Весник НГУ. Серия: Информационные технологии. - Том 5, выпуск 2. – С.12–15.
- [Глоба Л.С., и др., 2011] Глоба Л.С., д.т.н., проф., Новогрудская Р.Л. Модель интернет-портала «Прочность материалов» / Глоба Л.С., д.т.н., проф., Новогрудская Р.Л. // Международный научно-технический сборник «Надежность и долговечность машин и сооружений». - Київ; - 2011, стр. 183-189
- [Глоба Л.С., 2010] Глоба Л.С., д.т.н., проф., Новогрудская Р.Л. Подход к построению портала инженерных знаний / Глоба Л.С., д.т.н., проф., Новогрудская Р.Л. // ИАИ – 2010. Труды конференции. С53-62.

NATIONAL ANTARCTIC DATA CENTRE INTERNET-PORTAL MODEL

Globa L.S., Kuzin I.A., Mochalkina K.S.,
Novogradskaya R.L.

*National Technical University of Ukraine "Kiev
Polytechnic Institute", c. Kiev, Ukraine*

lgloba@its.kpi.ua

ksu_mochalkina@yahoo.com

ihkuzin@gmail.com

rinan@ukr.net

The article presents the approach for the designing of knowledge Internet portal in the field Antarctic researches. Ontology is used as portal informational model.

INTRODUCTION

There are currently a large number of studies in the Antarctic in various scientific fields. However, there is a problem of access to the results obtained and their further processing. Important results, both theoretical and practical remain focused only on the level of individual research institutes. Therefore, an urgent task is to provide opportunities to use this knowledge to more researchers, for whom they are intended.

According to the requirements of Resolution 4.1 XXII ATCM 1998 all Antarctic Treaty Parties engaged in Antarctic research have to establish National Antarctic Data Centre (NTSAD). From all of these points there is a need to build a qualitative model of knowledge representation on the portal. As a model of knowledge representation portal ontology is used.

MAIN PART

Recently, for the construction of scientific knowledge portals ontological model are being used. Ontology is a conceptual model, which helps in a certain area's formalization. Introducing a formal description of domain concepts as classes of objects and their relationships, the ontology defines the structure of the portal to provide real data and relationships between them.

Ontologies consist of the following elements: instances, classes, attributes and relationships. Classes - these are abstract groups, sets or collections of objects, which include instances other classes, or the same combination and order of the two. Attributes - describe the classes of certain characteristics. Each attribute has at least the name and value, and is used to store information that is specific to the object and attached to it. Relationships - establish a connection (describing the dependence) between the elements of the ontology.

Ontology of National Antarctic Data Centre (NADC)portal consists of four ontologies: the ontology of the external user the ontology of Antarctic knowledge, the ontology of a researcher and the ontology of a subject area. The ontology of the external user (OEU) includes classes related to providing access to materials for an external user. Ontology of the Antarctic knowledge (OAK) contains the following metadata defining a framework for the description of the subject area. Ontology of the researcher (OR) describes the possibilities provided to researchers at the portal to access and processing of materials posted on the portal. Ontology of a subject area (OSA) reflects the general domain knowledge, such as class hierarchy of concepts, semantic relationships for these classes.

According to the discussed above ontology portal NADC must solve a number of functional problems:

- Quality of knowledge representation in the portal
- Organizing and structuring of information
- The formalization of Antarctic knowledge
- Effective and meaningful search

The portal will provide access to databases, reference books, manuals, express information, resources, networks, etc., etc. In the portal can be implemented blocks viewing of news, information about conferences (upcoming and past), competitions and grants, as well as other information about events relevant to this area of expertise. An important step in the construction of the portal is the structuring and systematization of information and knowledge in the field of Antarctic research, which will allow end users to browse and search for specific information in this field. Structuring and placement on the site is organized for the end user friendly way by implementing problem-oriented navigation tools and search the information space portal. The search information is organized so that the user can specify the request is not only and not by words, but in terms familiar to him the domain portal.

CONCLUSION

The possibility of building a portal "National Centre for Antarctic Data" was discussed. Portal's information base are ontologies, which helps in systematizing and structuring of information, helps to organize an effective search and navigation in information space of the National Antarctic Data Centre portal. The portal's functions implemented on the basis of described ontology are presented.