

ОНТОЛОГИЯ «ОБРАЗЫ МЕДВЕЖЬИХ ПЕСЕН ХАНТОВ» ДЛЯ ПОРТАЛА «ФОЛЬКЛОР НАРОДОВ СИБИРИ»¹

В настоящее время особенно актуальным является вопрос доступа к специализированной информации, которая зачастую слабо структурирована, рассредоточена по разным сайтам, архивам и библиотекам, что не позволяет осуществлять эффективный поиск по ней. Чтобы удовлетворить потребность в содержательном доступе к специализированной информации, Ю. А. Загоруйко и его коллегами предложена концепция и архитектура Интернет-порталов знаний, которые обеспечивают систематизацию знаний и информационных ресурсов по заданной научной тематике, их интеграцию в единое информационное пространство, а также содержательный доступ к ним [3].

Электронный портал знаний «Фольклор народов Сибири» (olk.philology.nsc.ru), описанию которого посвящена данная статья, разрабатывается на базе технологий для порталов знаний, а также на основе академического издания «Памятники фольклора народов Сибири и Дальнего Востока». Портал «Фольклор народов Сибири» представляет собой информационную систему, посвященную фольклору народов, исконно населявших Сибирь. Со временем портал объединит материалы по фольклору всех народов, публикуемых в серии «Памятники фольклора народов Сибири и Дальнего Востока». На данный же момент к публикации готовится материал по фольклору бурят (героический эпос) и хантов (обрядовый фольклор).

© А. А. Гриневич, 2016

Информационная система портала объединяет как текстовые, так и мультимедийные материалы (фото, видео, аудио, ноты). Информационная система разрабатывается в русле Semantic Web, согласно которому гипертекстовые страницы снабжаются дополнительной разметкой, несущей сведения о семантике включаемых в страницы элементов, поэтому система способна без участия человека извлечь какие-либо сведения из текста. Таким образом, текст на портале представлен в виде электронного корпуса текстов. Особенностью корпуса на портале является его разметка, основанная на элементах поэтической системы. В отличие от лингвистических корпусов, имеющих в настоящее время достаточно широкое употребление, фольклорный корпус текстов создается впервые.

Неотъемлемым компонентом Semantic Web является понятие онтологии, описывающее смысл семантической разметки. Термин «онтология» имеет долгую историю в философии, где со времен Аристотеля обозначал науку о бытии, в настоящее же время его активно использует компьютерная инженерия, где под онтологией понимается система понятий некоторой предметной области (в данном случае фольклористики, а точнее, концептуальная картина мира фольклорных произведений), которая представляется как набор сущностей, соединенных различными отношениями. Новое понимание онтологии рождает много дискуссий, но общепризнанным в этой области знания является определение, данное

¹ Статья подготовлена при поддержке гранта РГНФ № 14-04-12010 «Электронный ресурс «Фольклор народов Сибири»: текстовое, семантическое и мультимедийное представление».

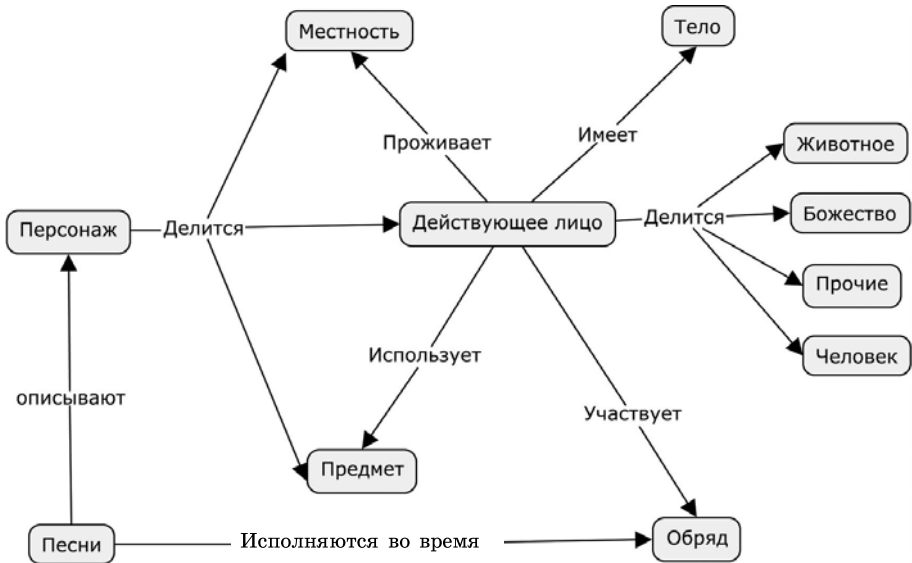


Рис. 1. Базовая онтология

Н. Грубером, согласно которому онтология «is an explicit specification of a conceptualization “это формальная спецификация концептуализации”» [6]. Концептуализация – это упрощенное представление о мире, на ней основывается любая репрезентация знаний. Набор объектов, а также отношений между ними отражаются в специальном словаре, при помощи которого программа, ориентированная на определенную область знаний, представляет то или иное знание.

В построении онтологии необходим строгий формализм, который в то же время обеспечивает гибкие средства описания предметной области, а также установления между ними семантических связей. Понятия предметной области выстраиваются в иерархию «общее-частное» с наследованием свойств по этой иерархии.

Базовая онтология «Образы медвежьих песен хантов» для портала знаний «Фольклор народов Сибири» разрабатывалась на основе анализа обрядового фольклора хантов, а именно песен медвежьего праздника. В песнях были выделены устойчивые элементы, формирующие

поэтическую систему, – поэтические строки устойчивой структуры и лексики, формулы, – которые, в свою очередь, были тематически разделены на классы, между которыми установлены связи (подробнее о формулах см.: [4, с. 3–10; 5, 109–120]. Эта онтология включает в себя следующий набор понятий, связанных отношениями «класс-подкласс» или «часть-целое» (рис. 1).

Базовая онтология была представлена средствами портала знаний. Формально онтология порталов знаний включает в себя следующие взаимосвязанные сущности.

$$O = \langle C, R, T, D, A, F, Ax \rangle,$$

где O – Онтология портала знаний;

C – Конечное непустое множество классов, описывающих ПО;

R – Конечное непустое множество бинарных отношений, заданных на классах (понятиях);

T – Конечное непустое множество стандартных типов;

D – Множество доменов, отношений заданного типа;

A – Конечное множество атрибутов, описывающее свойства понятий C;

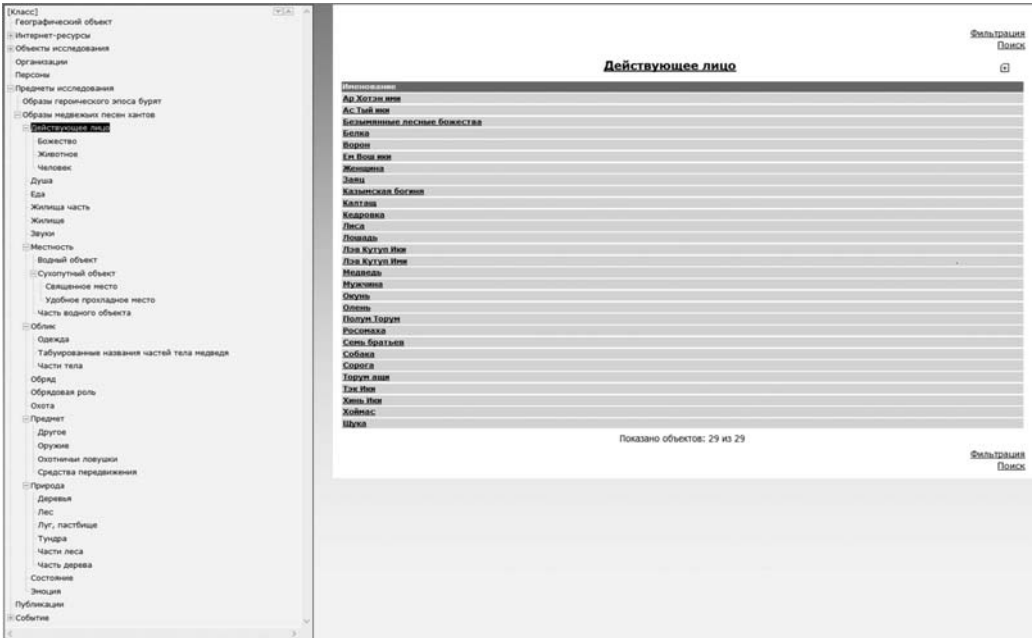


Рис. 2. Редактор онтологии

F – Множество ограничений на значения атрибутов понятий и отношений;

Ax – Множество аксиом, определяющих семантику классов и отношений онтологии [1, с. 162] (подробнее о структуре онтологии см.: [2]).

В редакторе портала знаний онтология представлена в виде дерева понятий (рис. 2).

Отношения между объектами устанавливаются на уровне классов. Эти свойства наследуют все объекты, входящие в один класс. Так, для класса «Действующее лицо» насчитывается 13 отношений (рис. 3).

Онтология «Образы медвежьих песен хантов» может стать образцом для других онтологий, которые будут разрабатываться для разметки фольклорных

Имя атрибута	Тип	Имя домена	Множеств.	Обязат.	Уник.	Ссылка	Кр.ссылка	Использований
1 Именованное	строка		нет	нет	нет	1	1	0+29
2 Возраст	в домене	Возраст	нет	нет	нет			0+3

Отношения класса Действующее лицо								
Название	Тип отношения	Правый аргумент	Обязат.	Уник.	Ссылка	Кр.ссылка	Использований	
1 участвуетВ	Класс-Класс	Медвежий праздник хантов	нет	нет			0+1	
2 живётВ	Класс-Класс	Жилеа	нет	нет			0+1	
3 испытывает	Класс-Класс	Эмоция	нет	нет			0+4	
4 обладает	Класс-Класс	Облик	нет	нет			0+6	
5 ест	Класс-Класс	Еда	нет	нет			0+1	
ситуация	в домене	ситуация	нет	нет			0+1	
6 слышит	Класс-Класс	Звуки	нет	нет			0+1	
7 использует	Класс-Класс	Предмет	нет	нет			0	
ситуация	в домене	ситуация	нет	нет			0	
8 ходитНа	Класс-Класс	Охота	нет	нет			0+4	
9 исполняет	Класс-Класс	Обрядовая роль	нет	нет			0+6	
возраст исполнителя	в домене	Возраст	нет	нет			0+3	
10 состоитИз	Класс-Класс	Образы медвежьих песен хантов	нет	нет			0	
11 означает	Класс-Класс	Образы медвежьих песен хантов	нет	нет			0	
12 переживает	Класс-Класс	Состояние	нет	нет			0+5	
13 становится	Объект-Класс	Действующее лицо	нет	нет			0+1	

Рис. 3. Список отношений класса «Действующее лицо»

текстов, поскольку она содержит основные элементы, отражающие взаимосвязь человека и окружающего его универсума. Действительно, фольклор сохраняет главные ценностные установки человека,

его представление о себе и своей роли в окружающем мире. Поэтому такая онтология может стать основой для других онтологий, посвященных не только фольклору, но и человеку в целом.

Литература

1. Загоруйко Ю. А. Подход к обеспечению многоязычного доступа к систематизированным знаниям и информационным ресурсам заданной предметной области // Извест. Томск. политех. ун. – 2009. – Т. 314. – № 5. – С. 161–165.
2. Загоруйко Ю. А., Боровикова О. И. Информационная модель портала научных знаний // Информационные технологии. – 2009. – № 12. – С. 2–7.
3. Загоруйко Ю. А., Боровикова О. И. Подход к построению порталов научных знаний // Автометрия. – 2008. – Т. 44. № 1. – С. 100–110.
4. Кузьмина Е. Н. Указатель типических мест героического эпоса народов Сибири (алтайцев, бурят, тувинцев, хакасов, шорцев, якутов): Эксп. изд.; Рос. акад. наук, Сиб. отделение, Ин-т филологии. – Новосибирск: Изд-во СО РАН, 2005. – 1383 с.
5. Гриневич А. А. Поэтика обрядовых песен медвежьего праздника казымских хантов: Дис. ... канд. филол. наук. – Новосибирск. 2012. – 256 с.
6. Gruber T. R., A Translation Approach to Portable Ontology Specifications. // Knowledge Acquisition. – 1993, 5(2), pp. 199–220. // Ссылка доступа <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>

References

1. Zagorul'ko Iu. A. Podkhod k obespecheniiu mnogoiazychnogo dostupa k sistematizirovannym znaniiam i informatsionnym resursam zadannoi predmetnoi oblasti [Approach to providing a multilanguage access to systematized knowledge and information resources in the specified subject area] Izvestiia Tomskogo politekhnicheskogo universiteta [News of Tomsk Polytechnic University]. 2009, Vol. 314, no 5, pp. 161–165.
2. Zagorul'ko Iu. A., Borovikova O.I. Informatsionnaia model' portala nauchnykh znanii [Information model of the scientific knowledge portal] Informatsionnye tekhnologii [Information technologies]. 2009, no 12, pp. 2–7.
3. Zagorul'ko Iu. A., Borovikova O.I. Podkhod k postroeniiu portalov nauchnykh znanii [Approach to building the scientific knowledge portals] Avtometriia [Autometria]. 2008, Vol. 44, no 1, pp. 100–110.
4. Kuz'mina E. N. Ukazatel' tipicheskikh mest geroicheskogo eposa narodov Sibiri (altaitsev, buriat, tuvintsev, khakasov, shortsev, iakutov) [Index of typical locations of a heroic epic of peoples of Siberia (Altaians, Buryats, Tuvinians, Khakass, Shors, Yakuts)]. Siberian branch of the Russian Academy of Sciences, Institute of Philology, Novosibirsk: Izd-vo SO RAN, 2005, 1383 p.
5. Grinevich A. A. Poetika obriadovykh pesen medvezh'ego prazdnika kazymskikh khandov [Poetics of the Kazym Khants' bear holiday ritual songs]. Abstract of PhD thesis, Novosibirsk. 2012. 256 s.
6. Gruber T. R., A Translation Approach to Portable Ontology Specifications. // Knowledge Acquisition. 1993, 5(2), pp. 199–220. URL: <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>

Онтология «Образы медвежьих песен хантов» для портала «Фольклор народов Сибири»

В статье описана онтология, посвященная образам обрядового фольклора хантов. Онтология понимается в русле инженерного знания и представляет собой систему взаимосвя-

Ontology “Images of Ostyak bear songs” for the portal “Folklore of Siberia”

The article is devoted to the description of ontology of ritual Ostyak folklore. Ontology is understood in line with the engineering knowledge as a system of interrelated concepts.

занных понятий. В фольклорном материале автором выделены устойчивые элементы – поэтические формулы, которые представлены в виде системы понятий, разделенные на классы и наследующие свойства по иерархии.

Ключевые слова: онтология, обрядовый фольклор, портал знаний, корпус текстов.

The author marked stable elements (poetic formulas) in the folklore material, which are presented in the form of hierarchically systemized concepts divided into classes.

Keywords: ontology, ritual folklore, knowledge portal, text corpus.

Гриневиц Анна Александровна, научный сотрудник Института филологии СО РАН, кандидат филологических наук.

E-mail: annazor@mail.ru

Anna Grinevich, Candidate of Folkloristics, researcher at the Institute of philology of the SB of RAS.

E-mail: annazor@mail.ru