

УДК 004.5, 004.8, 004.9.

## Поддержка интеллектуальности пользовательского интерфейса системы распределенного семантического поиска: проблемы и решения

П.А. Ломов, В.А. Путилов, А.В. Маслобоев

*Институт информатики и математического моделирования технологических процессов Кольского научного центра РАН, лаборатория региональных информационных систем*

**Аннотация.** В статье рассматривается задача поддержки интеллектуальности пользовательского интерфейса системы распределенного семантического поиска. В качестве решения предложено использование общесистемного расширяемого тезауруса как основы интеллектуального интерфейса информационной системы распределенного семантического поиска. Представлено расширенное элементами онтологии DOLCE и метасвойствами определение разделяемого тезауруса. Разработана технология формирования запроса на основе правил определения возможных вариантов его усложнения, что позволяет использовать тезаурус в качестве основы для разработки интеллектуального пользовательского интерфейса.

**Abstract.** In this paper the distributed semantic retrieval information system user interface intellectuality support problem has been considered. As a solution the application of expanded systems-wide thesaurus as the basis of intelligent interface of the distributed semantic retrieval information systems has been proposed. The definition of shared thesaurus expansion by units of the top-level ontology and meta-properties has been represented. The inquiry formation technology based on identification procedures of its meshing possible variants that allows the thesaurus application as the basis for intelligent user interface design has been developed.

**Ключевые слова:** система информационного поиска, онтология, тезаурус, семантическая интеграция информации, формирование запроса, интеллектуальный пользовательский интерфейс

**Key words:** retrieval information system, ontology, thesaurus, semantic information integration, query manipulation, intelligent user interface

### 1. Введение

Важным аспектом создания информационных систем распределенного семантического поиска является разработка интеллектуального пользовательского интерфейса. Интеллектуальность пользовательского интерфейса в данном контексте предполагает наличие у него возможности оперировать при взаимодействии с пользователем языком, близким к естественному, а также интерпретировать пользовательское сообщение в виде некоторого плана подзапросов на специализированном языке. Отличительной особенностью таких интерфейсов является их тесная интеграция с системой понятий предметной области, выраженной, как правило, в виде онтологии, разработка которой, а также различных методов и средств ее использования и развития (логический вывод, добавление новых фактов, интеграция с другими онтологиями и т.п.) может рассматриваться как поддержка интеллектуальности. Поддержка интеллектуальности также предполагает наличие специальных механизмов согласования и реализацию технологий интеграции разнородных онтологий предметных областей, разрабатываемых независимыми специалистами, что позволяет обеспечить высокий уровень "дружественности" взаимодействия пользователя с системой в процессе формирования поисковых запросов.

На сегодняшний день существует несколько подходов к созданию интеллектуального пользовательского интерфейса поисковых систем, основанных в большей степени на использовании реляционных баз данных (*Benzi et al.*, 1999; *Pichat, Saker*, 1995; *Keim*, 1996). Среди недостатков таких интерфейсов, согласно (*Catarci et al.*, 1997), выделяют: отсутствие правил логического вывода, позволяющих получить новое знание, вытекающее из хранимых фактов, невозможность использования метаинформации, семантику которой нельзя выразить с использованием реляционной модели данных, и запросов, дающих приближенный результат, подстраивание системы под конкретного пользователя на основе анализа семантики некоторой совокупности прошлых запросов.

Среди современных концепций создания интеллектуальных пользовательских интерфейсов можно также выделить подходы, использующие в своей основе динамическую фасетную классификацию понятий для поддержки пользователя в процессе формулировки запроса (*Обухова и др.*, 2009). Это

позволяет использовать простые правила уточнения запроса на каждом шаге его формирования, но так как сама фасетная классификация достаточно слабо отражает все разнообразие связей между объектами предметной области, то возможности такого уточнения достаточно ограничены.

Наиболее перспективными являются подходы, направленные на создание пользовательского интерфейса на основе формальной онтологической модели, содержащей понятия и отношения, в которые отображаются элементы данных, хранящихся в различных гетерогенных источниках (Catarci et al., 2004; Bechhofer et al., 1999; Feikje et al., 2008). Под формальной онтологией понимается онтология, понятия которой описываются аксиомами и определениями в логике или на языке, автоматически транслируемом в логику (Sowa, 2009). В этом случае онтология используется как система связанных между собой понятий предметной области, в которой, как правило, пользователь ориентируется. Это дает возможность пользователю оперировать известными ему терминами, осуществлять экстенциональную и интенциональную навигацию между ними и итеративно формировать сложный запрос, внося в него дополнительные объекты поиска и определяя ограничения. В итоге это приводит к формулировке запроса, наиболее точно отражающего семантику сложного объекта поиска. Также пользователи имеют возможность ознакомиться с предметной областью без совершения запросов, для более ясного видения контекста того или иного термина, определяемого его иерархическим положением и взаимосвязями с другими терминами. Однако такие подходы часто опираются на единую онтологию, задающую общую семантику понятий, описываемых в различных информационных источниках, что соответствует централизованному подходу к интеграции информации с использованием онтологий, имеющему следующие недостатки: усреднение семантики понятий исходных онтологий при их отождествлении с понятиями централизованной, невозможность последующего добавления новых информационных ресурсов, описывающих понятия, которые не могут быть отражены в общей онтологии. Для их устранения было предложено использовать разновидность гибридного подхода (Ломов, Шишаев, 2008) к семантической интеграции с использованием расширяемого общесистемного тезауруса. Однако тезаурус в меньшей степени отражает формальную семантику понятий, в сравнении с единой онтологией, что снижает его возможности в отношении поддержки пользователя при поиске информации. На решение данной проблемы и направлена настоящая работа.

В работе рассматриваются вопросы, связанные с использованием общесистемного расширяемого тезауруса как основы интеллектуального интерфейса информационной системы распределенного семантического поиска. Предлагается расширенное элементами онтологии верхнего уровня DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*) и метасвойствами, являющимися инвариантными относительно онтологических спецификаций предметных областей, определение разделяемого тезауруса, позволяющее сохранить понятийную семантику и обеспечивающее повышение уровня его выразительности. Описывается процесс формирования запроса на основе правил определения возможных вариантов его усложнения, что позволяет использовать тезаурус в качестве основы для разработки интеллектуального пользовательского интерфейса.

Методологическую основу работы составляют исследования отечественных и зарубежных ученых в следующих областях: онтологическое моделирование – Н. Гуарино, К. Велти (Guarino, Welty, 2000), Т. Груббер (Grubber, 1993), А.С. Клещев (2007), И.Л. Артемьева (2008); семантическая интеграция данных и информационный поиск – Л.А. Калиниченко (1983), М.Р. Коголовский (2005), В.А. Витгих (Витгих, Минаков, 2007), А.В. Смирнов (Смирнов и др., 2002); машинное обучение и математическая логика – Д.А. Поспелов (1989), Г.С. Осипов (2007); семантическая паутина (Semantic Web) – Т. Бернерс-Ли (Berners-Lee et al., 2001), Д. МакГиннесс (McGuinness, 2004), Ф. Баадер (Baader, Narendran, 2001) и др.

## 2. Определение общесистемного расширяемого тезауруса

В общем случае тезаурус можно определить как словарь терминов, связанных определенным набором семантических отношений (синонимия, гипонимия, родовая связь). Выделяется также особый класс информационно-поисковых тезаурусов, в которых термины обязательно связываются со своими определениями. Вследствие наличия сходных с формальной онтологией элементов, таких как понятия и отношения, тезаурус может быть отнесен к ресурсам онтологического типа (Добров и др., 2008). Можно также рассматривать тезаурус как обогащенную семантикой, выраженной в дополнительных типах отношений, таксономию (Garshol, 2004).

Основным отличием формальной онтологии от тезауруса является формальное определение понятий и отношений с помощью, как правило, дескриптивных логик, представляющих разрешимый фрагмент логики предикатов первого порядка и являющихся формальной основой языка онтологий OWL (*Web Ontology Language*). Это и определяет направленность использования тезауруса для описания лексики в проекции на семантику, а онтологии – для моделирования семантики и прагматики в проекции на язык представления (Гладун, Рогошина, 2007).

Однако в данной работе понятие тезауруса не в полной мере соответствует общепринятому и расширяет его определением дополнительных ограничений и взвешенный отношений. Тезаурус позиционируется как каноническая модель для осуществления семантической интеграции, представленная в работе (Ломов, Шишаев, 2009). Основной задачей тезауруса являются:

- централизованное хранение элементов онтологий (концептов и атрибутов) исходных информационных источников;
- установления и поддержка различного рода связей между терминами, отражающих их семантическую близость;
- определение и хранение общезначимых атрибутов и определенных на их основе объектов (Ломов, Шишаев, 2009).

Для формального определения тезауруса введем следующие обозначения:

$C$  – множество всех понятий, содержащихся в онтологиях исходных интегрируемых ресурсов;

$A$  – множество всех свойств концептов онтологий.

Каждый концепт из множества  $C$  представляется в тезаурусе соответствующим ему элементом тезауруса типа "Объект":

$$O_i = \langle N_i, A_i \rangle,$$

где  $N_i$  – описание объекта, соответствующее какому-либо описанию представляемого им концепта, которым может быть название или определение на естественном языке,  $A_i$  – множество онтологий, в которых представлен концепт, соответствующий объекту  $O_i$ .

$O$  – множество всех объектов тезауруса.

Каждый атрибут из множества  $A_i$  будет представлен в тезаурусе соответствующим элементом типа "Свойство", которое представим в виде:

$$P_i = \langle N_i, A_i \rangle,$$

где  $N_i$  – символьное имя свойства или описание  $P_i$ , соответствующее наименованию или описанию атрибута из онтологии,  $A_i$  – множество онтологий, в которых представлен атрибут, соответствующий свойству  $P_i$ .

$P$  – множество всех свойств тезауруса.

Для обозначения элементов тезауруса, соответствующих концепту или атрибуту из онтологии (и наоборот) будем использовать обозначения:

$n(t)$  – элемент онтологии  $n$ , соответствующий элементу тезауруса  $t$ ,  $n \in C \cup A$ ,  $t \in O \cup P$ .

$t(n)$  – элемент тезауруса  $t$ , соответствующий элементу онтологии  $n$ .

В тезаурусе определяется также элемент "Значение свойства", который используется для описания множества допустимых значений атрибутов в виде множества номиналов определенного типа. Значения атрибутов могут быть разных типов как элементарные – строковый, числовой, логический, так и характерные для предметных областей – номер свидетельства, идентификатор документа и т.д.:

$$R_i = \langle N_i, A_i \rangle,$$

где  $N_i$  – описание множества значений,  $A_i$  – множество онтологий, в которых используется данное множество значений.

Атомарное значение – конкретный номинал определяется элементом:

$$V_i = \langle VAL_i, A_i \rangle,$$

где  $VAL_i$  – литеральное значение свойства,  $A_i$  – множество онтологий, в которых используется данное множество значений.

Далее будем обозначать множество онтологий, в которых присутствует элемент  $t$ , как  $A(t)$ .

На множествах различных элементов тезауруса заданы следующие отношения:

$HP \subseteq (O \times O) \cup (P \times P) \cup (V \times V)$  – отношение гипонимии;

$PR \subseteq (O \times P)$  – отношение принадлежности свойства объекту тезауруса;

$FD \subseteq (O \times O) \cup (P \times P)$  – отношение формального различия.

Принадлежность свойства объекту и формальное различие объектов и свойств наследуется по связи гипонимии:

$$\forall o, b \in O, p \in P, oHPb \wedge bPRp \rightarrow oPRp,$$

$$\forall p, n, k \in O, P, pHPn \wedge nFDk \rightarrow pFDk.$$

Отношение принадлежности свойства объекту может быть определено, если они оба принадлежат, хотя бы к одной, общей для них, онтологии:

$$\forall b \in O, p \in P, bPRp \rightarrow (A_b \cap A_p) \neq \emptyset$$

$PQ(b) = \{p_i | p_i \in P \vee (p_i PRb)\}$  – множество свойств объекта  $b$ .

$VL \in (P \times V) \cup (P \times O) \cup (P \times R)$  – отношение принадлежности свойству множества значений, конкретного значения или объекта, для отображения литеральных и объектных свойств онтологий.

$IN \in (R \times V) \cup (R \times R)$  – отношения вхождения атомарного значения или множества в определенное множество значений.

Отношения гипонимии и формального различия обладает транзитивностью:

$$\begin{aligned} \forall o, b, n \in O, oHPb \wedge bHPn &\rightarrow oHPn, \\ \forall p, n, k \in O, P, pFDn \wedge nFDk &\rightarrow pFDk. \end{aligned}$$

На множестве элементов тезауруса также заданы отношение синонимии и отношение ассоциации, имеющее определенный вес. Определим эти отношения следующим образом:

$SYS \subseteq (O \times O) \cup (P \times P)$  – отношение синонимии;

$ASC \subseteq (O \times O) \cup (P \times P)$  – отношение ассоциации.

$WS:ASC \rightarrow [0,1]$  – функция, возвращающая вес ассоциативного отношения. Вес отношения характеризует степень семантической схожести понятий, соответствующих элементам тезауруса, входящим в отношение. Нулевой вес означает отсутствие сходства, единичный – полную семантическую эквивалентность понятий.

Определим также предельные значения веса отношения ассоциации:

$UWS$  – верхнее предельное значение: в дальнейших процедурах термины-аргументы расцениваются как семантически эквивалентные;

$DWS$  – нижнее предельное значение: в дальнейших процедурах термины-аргументы не расцениваются как ассоциации.

Необходимыми условиями для установления синонимии между семантически эквивалентными объектами и свойствами тезауруса являются:

$\forall p, n \in O, pSYNn \rightarrow \neg(pFDk)$  – объекты-синонимы не могут формально различаться;

$\forall b, m \in P, bSYNm \rightarrow \neg(pFDk) \wedge (bVLr \wedge mVLg \wedge \exists v:(rINv \wedge mINv))$  – свойства-синонимы не являются формально различными и их множества значений имеют хотя бы один общий элемент.

### 3. Расширения тезауруса концептами DOLCE и метасвойствами OntoClean

Одной из основополагающих задач тезауруса является определение отношений между понятиями различных онтологий, однако для этого часто необходима дополнительная семантика, которую нельзя отразить, используя существующее определение тезауруса. Малое количество основных видов связей между элементами также снижает выразительные возможности тезауруса как основы для интеллектуальной системы формирования запросов. Эти недостатки связаны с тем, что тезаурус можно в большей степени отнести к онтологическим моделям представления (Добров и др., 2008), ориентированным на представление знаний, а не того смыслового содержания понятий предметных областей и отношений между ними, которые можно использовать при установлении семантической близости понятий.

Данную проблему можно решить путем определения некоего общего основания, способного отразить фундаментальные смысловые особенности понятий любой предметной области или задачи, и в то же время не накладывающего ограничений на смысловое содержание понятий предметной области. Одним из способов реализации такого подхода является использование онтологий верхнего уровня, включающих понятия и отношения, которые можно выделить в каждой онтологии. Поэтому онтологии верхнего уровня можно рассматривать как некое общее основание, обеспечивающее возможность анализа любой предметной области для определения ее понятий и ее последующего моделирования. Включение же элементов таких онтологий в определение разделяемого тезауруса позволит в большей степени сохранить исходную семантику терминов при их добавлении в тезаурус, с большей точностью определить и обосновать смысловую близость или ее отсутствие между понятиями, формально определить базовые понятия предметной области и их основные атрибуты, задать правила вывода, используемые для выявления скрытых отношений, а также позволит упростить разработку и обеспечить высокую степень правильности создания онтологических спецификаций исходных информационных ресурсов.

В качестве онтологии верхнего уровня, или мета-онтологии, авторами настоящей работы было предложено использовать онтологию DOLCE (Masolo et al., 2006). Использование именно этой онтологии объясняется ее ориентацией на различные социальные субъекты, объекты и процессы, такие как организации, коллективы, планы и нормы.

#### 3.1. Методология OntoClean как основа создания онтологии DOLCE

Онтология DOLCE является первой из онтологий в библиотеке базовых онтологий проекта WonderWeb. Онтологию DOLCE предполагается применять в Semantic Web для согласования между

интеллектуальными агентами, использующими разную терминологию (Добров и др., 2008). DOLCE имеет когнитивный уклон, поскольку фиксирует онтологические категории естественного языка и знания "здорового смысла" и, не претендуя на звание некой универсальной онтологии, определяет семантический базис для описания понятий, присущих отдельным предметным областям.

При разработке онтологии DOLCE применялась методика OntoClean (Guarino, Welty, 2004), которая позволяет:

- выявить ошибки в иерархии классов;
- провести валидацию таксономий, показав неверные решения при моделировании сущностей;
- увидеть явные пути интерпретации атомарных свойств или классов;
- показать точку зрения на онтологические модели специалиста, осуществляющего моделирование.

Заметим, что данная методология трактует понятие "свойство" шире, чем это принято, например, в OWL онтологиях. Свойством в данной методологии также является, то что обычно называется классом или концептом онтологии. Например, концепт "Персона" может быть рассмотрен как набор объектов реального мира, обладающих свойством "являться персоной".

Основой данной методологии являются так называемые метасвойства, независимые от предметной области и присущие любому объекту или понятию реального мира. К данным метасвойствам относят:

*Существенность (Essential)* – неотъемлемость свойства сущности, является, своего рода, признаком, используя который отдельный экземпляр можно причислить к определенному классу. Можно рассматривать существенное свойство как основу интерпретации концепта или отношения онтологической модели;

*Стойкость (Rigid)* – неизменность существенных свойств во времени и пространстве, определяет постоянство членства экземпляра в определенном классе;

*Единство (Unity)* – определяет экземпляры как целые сущности, составные части которых связаны только с ними. Например, понятие "Собака", обозначающее конкретных животных, обладает единством, так как в данном контексте нет возможности разделить экземпляр данного класса на части и соотнести их с другим экземпляром. Характерным примером сущностей, не обладающих данным свойством, могут служить различные наборы, множества, количества. Например, класс "Группа людей" не обладает свойством единства, так как если его экземпляры разбивать по определенному числу людей или, наоборот, объединять будут получаться также экземпляры этого класса. Отмечается также, что если определенное свойство обладает единством, проявляющемся через определенное отношение, то это свойство должно также обладать существенностью (Guarino, Welty, 2002). Это позволяет говорить о понятиях, экземпляры которых являются целыми в любом подразумеваемом мире и в любой промежуток времени. Так, например, отдельный кусочек глины, можно считать целым до тех пор, пока нет другого куска, к которому он может быть прикреплен, что приведет к потере им свойства единства, так как в этом случае он уже является частью большого куска;

*Зависимость (Dependence)* – наличие зависимости у определенного класса указывает на необходимость существования других сущностей для существования его экземпляров, например, экземпляр класса "Комната" требует существования экземпляров класса "Стена";

*Идентифицируемость (Identity)* – наличие возможности дифференцирования экземпляров одного класса по определенному критерию.

В работе (Guarino, Welty, 2004) подчеркивается возможность отождествления свойств единства и идентифицируемости, поскольку они оба указывают на наличие некоего разделяющего критерия у понятия. Фундаментальное их отличие друг от друга состоит в том, что наличие идентифицируемости говорит о некой возможности разделить экземпляры концепта, а единство, в свою очередь, указывает на наличие критерия, дающего возможность отделить части одного экземпляра от частей других, то есть определить принадлежность части к только одному конкретному экземпляру.

### 3.2. Анализ состава концептов и отношений онтологии DOLCE

Общая иерархия классов DOLCE задает разделение понятий, определенных согласно OntoClean как *стойкие (rigid)* свойства, что следует из теоремы о правильной онтологии (Guarino, Welty, 2002): "В правильно построенной онтологии типы являются стойкими свойствами". Выделяются следующие основные классы:

- Абстрактные (Abstract);
- Длющиеся (Endurant);
- Постоянные (Perdurant);
- Качества (Quality).

Основным признаком абстрактных сущностей служит отсутствие у них пространственно-временных характеристик, то есть они существуют вне времени и пространства. Примером абстрактных

сущностей могут служить математические объекты, тогда как различные объекты реального мира или события таковыми не являются.

Длящиеся сущности характеризуются, как представляемые полностью в каждый момент их существования. Примером может являться предмет "Книга", так как в любой момент, пока она существует, она имеет свое содержание, то есть, представлена полностью.

Постоянные сущности "размазаны" по некоторому промежутку времени и представляются в каждый момент времени своими временными частями, которые упрощенно можно рассматривать как наборы значений свойств сущности в определенный момент времени. Примером может служить процесс кипения воды, так как он не существует полностью в единый момент времени, а может определяться как некая комбинация состояний воды при ее нагревании до температуры кипения.

Качества можно рассматривать как сущности, которые характеризуют другие сущности. Однако необходимо отметить, что в рассматриваемой онтологии качество не является синонимом свойства класса. Качество в данном случае является *частностью* (*particular*), а свойство, в свою очередь, является универсалией, то есть в отличие от качества имеет экземпляры. Качества присущи объектам, пока те существуют. Также в контексте данной онтологии производится разделение между качеством, характеризующим конкретный объект реального мира, и его *значением* (*quale*), которое можно определить как точку в *пространстве значений* (*quality space*). Качества, в свою очередь, принадлежат определенному *типу качества* (*quality type*), например, цвет, размер, вес. Для каждого типа качества определяется свое пространство значений. Внутри пространства состояний можно выделить *области* (*regions*). Области позволяют в ряде случаев более точно отразить семантику высказываний. Например, если рассмотреть высказывание "цвет изменился с серого на белый", то здесь речь идет не о конкретных точках в пространстве значений, а скорее об областях серого и белого цветов.

Области выделяются в отдельный класс абстрактных сущностей, экземплярами которого являются как конкретные значения в пространстве или времени – *атомарные области* (*atomic region*), так и множества значений, близких в пространственном или временном смысле (собственно *области*). Данный класс имеет несколько подклассов, среди которых можно выделить:

- *Абстрактная область* (*Abstract region*) – область, не определенная ни во времени, ни в пространстве;
- *Область во времени* (*Temporal region*);
- *Область в пространстве* (*Physical region*).

Рассмотрим основные идеи реализации отношений, связывающих концепты онтологии DOLCE. Все отношения, как и качества, разделяются на *непосредственные* (*immediate*) и *опосредованные* (*mediate*). Такое решение было принято для того, чтобы достаточно полно отразить семантику взаимосвязи между постоянными и длящимися сущностями, проявляющуюся в их пространственно-временных положениях:

*Постоянные* (*Perdurant*) имеют непосредственно заданное временное расположение, а их пространственное расположение определяется опосредованно через задание расположения их временных частей. Например, процесс кипения воды происходит во времени с 12:00 до 12:03, а его пространственное расположение определяется расположением емкости, в которой он происходит;

*Длящиеся* (*Endurant*) имеют непосредственно заданное пространственное расположение, тогда как их временное задается опосредовано, через постоянные сущности, в которых они принимают участие.

Надо отметить, что не следует отождествлять классические транзитивные отношения в дескриптивных логиках и опосредованные отношения DOLCE, которые в силу того, что несут более общий смысл, не могут быть выражены строго формально.

Так как в основе DOLCE лежит методология OntoClean, то ее основные отношения являются реализацией метасвойств методологии. Эти отношения называются *базовыми примитивными отношениями* (*basic primitive relation*), основным назначением которых является отражение различных онтологических утверждений таким образом, чтобы они оставались верными для любой предметной области. К ним относятся:

*Отношение "часть"* (*parthood*) / *"временная часть"* (*temporary parthood*) – применяется для обозначения меронимии между сущностями. Временная разновидность данного свойства применяется к длящимся сущностям, часто материальным объектам реального мира, для отражения того факта, что во время своего существования они могут приобретать или терять части;

*Отношение "зависимость"* (*dependence*) / *"пространственная зависимость"* (*spatial dependence*) – определяется следующим образом: если X зависит от Y, то Y представлен всякий раз, когда X представлен. Необходимо отметить, что "быть представленным" в данном случае рассматривается только во временном аспекте, для принятия во внимания еще и пространственной зависимости используется отношение "пространственная зависимость";

*Отношение "составлять"* (*to constitute*) – определяется следующим образом: если X составляет Y, то X имеет такие несущественные свойства, что для Y эти свойства являются существенными. Для

иллюстрации приведем такой пример, если для куска глины свойство "форма" является несущественным, так как при изменении его значения сущность глины не изменится, то для статуи, сделанной из этой глины, свойство "форма" уже является существенным. Данный пример также отражает отличие отношения "составлять" от отношения меронимии. Можно также заметить, что отношение "быть подклассом" также отличается, так как при его использовании множество существенных свойств наследуются подклассом;

*Отношение партиципации (participation)* – применяется для указания "принадлежности" длящихся сущностей к постоянным. Необходимо отметить, что для указания данного факта в контексте данной онтологии нельзя использовать отношения меронимии, так как длящиеся сущности не являются частями постоянных – только постоянные могут быть частями других постоянных сущностей;

*Отношения принадлежности (inherence) / временной принадлежности (temporal inherence)* качества сущности – связывает сущность с качеством, которое ее характеризует;

*Отношения принадлежности (has quale) / временной принадлежности (has temporal quale)* значения качеству – определяет значение, постоянное или меняющееся, качества некоторой сущности.

### 3.3. Отображение элементов онтологии DOLCE и метасвойств OntoClean в разделяемом тезаурусе

Для отображения основных понятий онтологии DOLCE необходимо определить их принадлежность к типам объектов тезауруса, при этом необходимо учитывать семантику, заложенную в них, для обеспечения непротиворечивости при отображении.

Необходимо также отметить, что, несмотря на то, что данная онтология верхнего уровня получена в результате исследования метасвойств, общих для понятий различных предметных областей, сами метасвойства непосредственно в ней не представлены. В данном случае ее основные концепты задают некоторую совокупность значений метасвойств. Например, из определения концепта "Состояние" (state) следует, что он обладает свойствами "существенность" (essential) и "стойкость" (rigidy). Это позволяет проверять корректность построения отображения между понятиями, то есть определяет необходимые условия для установления факта эквивалентности понятий. С этой целью, наряду с концептами DOLCE, можно включить в тезаурус новые элементы, соответствующие метасвойствам методологии OntoClean (Скворцов, 2009), и впоследствии использовать их для аннотирования понятий или отношений между ними.

Рассмотрим понятия *абстрактный (abstract), длящийся (endurant), постоянный (perdurant) объект*, они являются подклассами концепта "Частность" (particular), обозначающим любую сущность. Исходя из этого, их можно представить объектами тезауруса, являющимися прямыми гипонимами объекта "Сущность" – о, гипернима всех объектов тезауруса:

$$\begin{aligned} prd &= \langle \text{"Perdurant"}, A_{prd} \rangle, prdHPo, \\ abs &= \langle \text{"Abstract"}, A_{abs} \rangle, absHPo, \\ end &= \langle \text{"Endurant"}, A_{end} \rangle, endHPo. \end{aligned}$$

Для определения метасвойств предлагается включить в определение тезауруса функции, отображающие множества его элементов на множества значений. Для этого определим следующие функции, соответствующие метасвойствам методологии OntoClean:

$RG: O \cup C \rightarrow \{\text{"rigid"}, \text{"non-rigid"}, \text{"antirigid"}\}$  – функция, соответствующая свойству "Стойкость" (Rigid).

Данные значения применимы к понятиям онтологии, а не к свойствам, так как в OntoClean они определяют неизменность *значимого свойства (essential property)* класса, которое, как правило, непосредственно не указывается, и считается указанием на неизменность членства экземпляров в данном классе. Исходя из этого, значение "antirigid" можно использовать для определения концептов-ролей какого-либо объекта, например, классы "Несовершеннолетний" и "Совершеннолетний" могут иметь значение свойства "antirigid", так как их экземпляры, имеющие существенное свойство – определенный возраст, теряют его со временем, что в последствии может привести к разрыву отношений членства с данными классами.

$UN: O \cup C \rightarrow \{\text{"unity"}, \text{"non-unity"}\}$  – функция, определяющая значение свойства "Единство" (Unity).

Аналогично свойству "Стойкость", применяется к понятиям онтологии, так как атрибут понятия, через который оно проявляется, как правило, явно не указывается.

$ID: O \cup C \cup A \cup P \rightarrow \{\text{"id"}, \text{"non-id"}\}$  – функция, определяющая значение свойства "Идентифицируемость" (Identity).

$DP: O \cup C \cup A \cup P \rightarrow \{\text{"dep"}, \text{"non-dep"}\}$  – функция, определяющая значение свойства "Зависимость" (Dependence).

Наличие свойств идентифицируемости и/или зависимости у классов онтологии или объектов тезауруса может определяться вследствие наличия аналогичных значений метасвойств у одного или

нескольких атрибутов класса или свойства объекта, характеризующих данный класс или объект. Например, если в тезаурусе объект "Автомобиль" определяется как идентифицируемый, если существует свойство, сопряженное с ним, являющееся идентифицирующим, например, "имеет номерной знак".

$ES: A \cup P \rightarrow \{\text{"ess"}, \text{"non-ess"}\}$  – функция, определяющая значение свойства "Существенность" (*Essential*), определена на множестве атрибутов понятий из онтологий.

Рассмотрим далее вопрос включения понятия *качества* (*quality*) онтологии DOLCE. Заметим, что в тезаурусе изначально задан элемент типа "Свойство" – *c*, который выступает как гиперним для всех атрибутов, использованных в различных онтологиях, включаемых в тезаурус, поэтому данный элемент можно считать эквивалентным качеству, так как они оба являются *частностями* (*particular*) и имеют сходный смысл. Соответственно, подклассы *качества* (*quality*): *абстрактное* (*abstract*), *физическое* (*physical*) и *временное* (*temporal*) *качество* – будут представлены гипонимами свойства тезауруса:

$$\begin{aligned} pabs &= \langle \text{"abstract-quality"}, A_{prd} \rangle, pabsHPc, \\ pphs &= \langle \text{"physical-quality"}, A_{pphs} \rangle, pphsHPc, \\ ptmp &= \langle \text{"temporal-quality"}, A_{ptmp} \rangle, ptmpHPc. \end{aligned}$$

Концепт "*Область*" (*region*) семантически соответствует элементу тезауруса "Значение свойства", поэтому каким-либо образом дополнительно не определяется. Элемент "Атомарное значение", в свою очередь, соответствует концепту *значение* (*quale*).

Отображение в тезаурусе отношений между концептами онтологии DOLCE осуществляется определением дополнительных отношений на множествах элементов тезауруса, соответствующих отношениям *партисипации* (*participateIn*), *меронимии* (*partOf*), *зависимости* (*dependOn*), *составления* (*consituentOf*), а также их *временными* (*temporal*) и *пространственными* (*spatial*) разновидностями. Это обусловлено тем, что данные отношения применимы к понятиям любой предметной области и могут быть использованы для связывания элементов, представляющих их в тезаурусе.

Введем следующие обозначения:

$Perdurant = \{b_i | b_i HPprd\} \subseteq O$  – множество гипонимов объекта "Perdurant".

$Endurant = \{b_i | b_i HPend\} \subseteq O$  – множество гипонимов объекта "Endurant".

$Abstract = \{b_i | b_i HPabs\} \subseteq O$  – множество гипонимов объекта "Abstract".

$$Perdurant \cap Endurant \cap Abstract = \emptyset$$

Дополнительные отношения между элементами тезауруса определим следующим образом:

$PRC \subseteq (Endurant \times Perdurant)$  – отношение партисипации (*participate in*).

$PRT \subseteq (O \times O) \cup (P \times P)$  – отношение меронимии (*part of*).

$TPRT \subseteq (O \times Endurant)$  – отношение временной принадлежности (*temporal part of*).

$TDPN \subseteq (O \times O)$  – отношение зависимости во временном аспекте (*temporal dependence*).

$SDPN \subseteq (O \times O)$  – отношение зависимости в пространственном аспекте (*spatial dependence*).

$CST \subseteq (O \times O)$  – отношение составления (*to constitute*).

$RLC \subseteq (R \times R)$  – отношение вхождения одного множества значений свойств в другое.

Оставшиеся концепты онтологии DOLCE могут быть включены в тезаурус как гипонимы его объектов и свойств в соответствии со своим иерархическим положением.

#### 4. Использование тезауруса в процессе формирования запроса

Формирование, планирование и выполнение запросов являются важными аспектами функционирования системы интеграции разнородных информационных ресурсов, от которых во многом зависит быстрота, полнота и точность результатов поиска информации об определенном объекте предметной области. В данном разделе статьи представлены основные этапы процедуры формирования запроса, а также правила, на основании которых может осуществляться поддержка пользователя, путем предоставления различных способов расширения запроса, семантически соответствующих исходному.

##### 4.1. Общий вид процедуры формирования запроса

Основной задачей формулировки пользовательского запроса является как можно более полное и точное определение объекта поиска с помощью элементов тезауруса. Впоследствии это будет играть важную роль для получения результатов, в достаточной степени релевантных той информации, которую намеревался найти пользователь.

Процедура формулировки пользовательских запросов начинается с определения начального или предполагаемого объекта поиска путем его выбора из множества объектов тезауруса. Далее пользователю предлагаются как объекты, связанные с начальным различными явными отношениями, так



и близкие в отношении семантики объекты, тем самым представляя различные контексты проведения поиска, выбор которых позволит более точно передать семантику запроса для получения в итоге более релевантных результатов. В процессе формирования запроса пользователь также накладывает различные ограничения на литеральные свойства объекта. На следующем этапе сложный запрос разбивается на некоторый набор простых запросов, содержащих объекты, свойства и ограничения и адресованных к тем онтологиям, в которых определены соответствующие концепты и отношения.

Простой запрос является атомарной составляющей более сложных запросов. Он задается к онтологии  $N$ , в которой определены его составляющие, и характеризуется указанием одного объекта  $t$  с определением ограничений на его литеральные свойства в виде множества пар "свойство" – "ограничение":

$$SQ(t) = \langle t, \{ \langle p_j, v_j \rangle \} \rangle,$$

где  $tPRp_j, p_jVLv_j, N \in (A(t) \cap A(p_j) \cap A(v_j))$ .

Граф простого запроса представлен на рис. 1.

Результатом выполнения запроса будет множество экземпляров из онтологии исходного информационного ресурса, значения свойств которых удовлетворяют ограничениям, заданным в запросе на объект поиска:

$$Ans(SQ(t)) = \{ d_i : d_i ISA c(t) \vee L(d_i, r(p_j)) \in v_j \},$$

где  $ISA \subseteq (E \times C)$  – родовое отношение онтологии между элементами множества понятий  $C$  и элементами множества экземпляров  $E$ ,  $c(t)$  – понятие онтологии, соответствующее объекту тезауруса,  $r(p_j)$  – атрибут в онтологии, соответствующий  $j$ -му свойству объекта  $t$  в тезаурусе,  $L(d_i, r_j)$  – значение свойства  $r$  экземпляра  $d_i$ .

Каждый отдельный экземпляр онтологии является парой, включающей некий внутренний идентификатор и набор значений свойств класса, потомком которого экземпляр является:

$$d = \langle ident, \{ \langle r_i(p_i), L(r_i) \rangle \} \rangle,$$

где  $ident$  – внутренний идентификатор, например, [http://www.owl-ontologies.com/ont\\_zags\\_1.3.owl#prsn01](http://www.owl-ontologies.com/ont_zags_1.3.owl#prsn01).

На этапе формирования простой запрос может быть расширен дополнительными объектами, связями, свойствами и ограничениями как в глубину, так и в ширину, в зависимости от необходимости более точного определения или от расширения контекста объекта поиска.

#### 4.2. Расширение запроса

Расширение в глубину простого запроса можно также назвать запросом с учетом объектных свойств, который характеризуется поиском объекта, определяемого другими объектами, вступающими с ним в определенные отношения. Данный вид запросов имеет вид:

$$DQ(t) = \langle t, \{ \langle p_i, SQ_i(b_i) \rangle \} \rangle$$

где  $tPRp_i, b_iVLp_i, (A(t) \cap A(b) \cap A(p_i)) \neq \emptyset$ .

Граф запроса представлен на рис. 2.

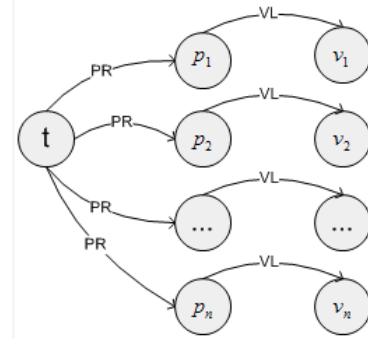


Рис. 1. Граф простого запроса

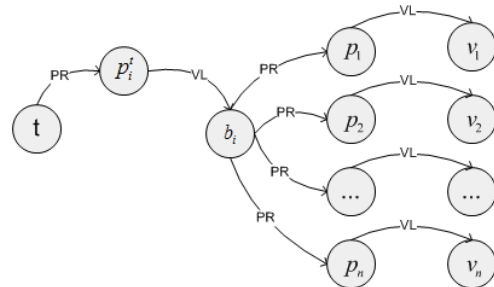
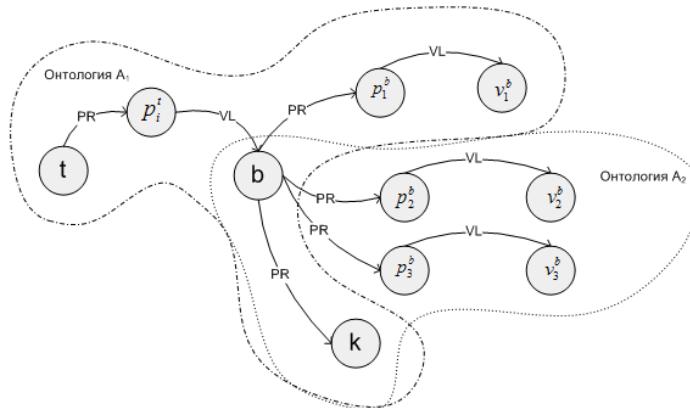


Рис. 2. Расширение в глубину простого запроса

Рис. 3. Условие наличия объединяющего свойства  $k$  при расширении в глубину запроса



Однако в этом случае при формировании запроса необходимо учесть возможность наличия у объекта  $b_i$  набора свойств из тех онтологий, в которых не определен объект  $t$ . В этом случае задавать ограничения на них при формировании запроса имеет смысл только тогда, когда их онтологии включают хотя бы одно отличительное свойство –  $k$  или его синоним из онтологии  $t$ , полученное значение которого позволит использовать результаты выполнения вложенного простого запроса для поиска объекта  $b_i$  и далее связанного с ним объекта  $t$  (рис. 3).

**Условие связности запросов.** Для выполнимости запроса  $DQ(t) = \langle t, \{ \langle p_i^t, SQ_i(b_i) \rangle \} \rangle$ , где  $SQ_i(b_i) = \langle b_i, \{ \langle p_j^b, v_j \rangle \} \rangle$ , по поиску объекта  $t$  по заданному набору ограничений на свойства  $p_j$  объекта  $b_i$  необходимо выполнение следующего условия:

$$\{Pb_j | (b_i PR p_j) \wedge (A(p_j) \cap A(t) \neq \emptyset)\} = \emptyset \rightarrow \exists k | (b_i PR k) \wedge ((A(k) \cap A(t) \neq \emptyset) \vee ((sSYNk) \wedge (A(s) \cap A(t) \neq \emptyset))).$$

Помимо этого условия, которое необходимо соблюдать, для обеспечения выполнимости запроса можно также сформулировать правила вывода, позволяющие формировать динамические связи на этапе построения и выполнения запроса.

**Правило транзитивной идентификации.** Если у некоторого объекта  $t$  имеется идентифицирующее свойство  $p_{id}^t$ , значением которого является другой объект –  $b$ , также имеющий идентифицирующее свойство  $p_{id}^b$ , то в этом случае между свойством  $p_{id}^b$  и объектом  $b$  формируется динамическая связь принадлежности свойства объекту (рис. 4). Таким образом, пользователь имеет возможность сразу задать ограничение на свойство  $p_{id}^b$  без перехода к обзору свойств объекта  $b$ . Разумеется, в случае наличия более длинной цепочки объектов, идентифицирующих друг друга, все их идентифицирующие свойства динамически прикрепляются к первому.

Формально правило транзитивной идентификации имеет вид:

$$(tPRp_t) \wedge (p_tVLb) \wedge (bPRp_b) \wedge (ID(p_t) = "id") \wedge (ID(p_b) = "id") \rightarrow tPRp_b.$$

**Правило транзитивности свойств синонима.** В случае наличия у объекта  $t$  связи с объектом  $b$ , имеющим в свою очередь синоним – объект  $f$ , между  $b$  и свойствами  $f$  устанавливаются динамические связи принадлежности (рис. 5).

Формально правило транзитивности свойств синонима можно записать в виде:

Рис. 5. Правило транзитивности свойств  $f$  – синонима  $t$

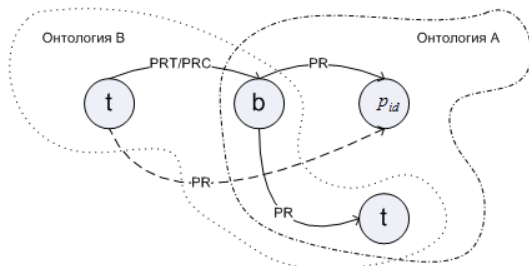
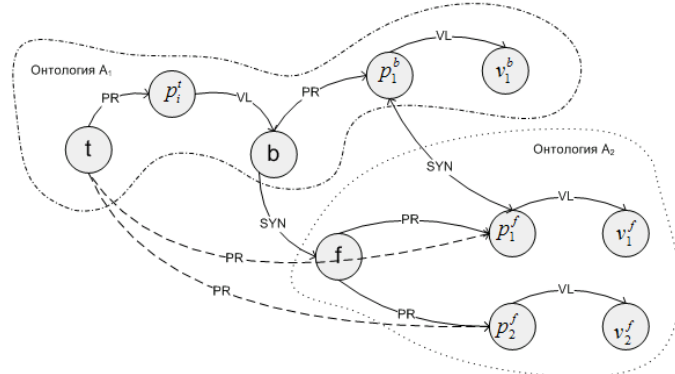


Рис. 6. Правило партисипативной / партномической идентификации через объект-часть или объект-участник

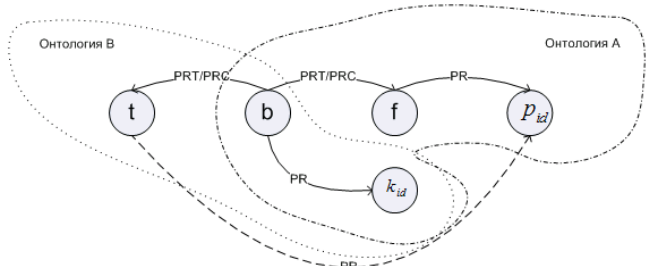


Рис. 7. Правило партисипативной/партномической идентификации объекта-части или объекта-участника посредством других частей или участников

$$(tPRp_i^t) \wedge (p_i^t VLb) \wedge (bSYNf) \wedge (fPRp_i^f) \rightarrow tPRp_i^f,$$

где для объектов  $t$  и  $f$  выполняется условие связности запросов.

**Правило партисипативной/партономической идентификации.** Данное правило может быть использовано в следующих случаях:

1. Объект может быть идентифицирован посредством указания значений идентификационных значений его объектов-частей или объектов-участников (рис. 6).

Правило партономической идентификации объекта  $t$  через его часть – объект  $b$ :

$$(tPRTb) \wedge (bPRp) \wedge (ID(p) = "id") \wedge (bPRk) \wedge (ID(k) = "id") \wedge (A(k) \cap A(t) \cap A(b) \neq \emptyset) \rightarrow tPRp.$$

Правило партисипативной идентификации объекта  $t$  через его участника – объект  $b$ :

$$(tPRCb) \wedge (bPRp) \wedge (ID(p) = "id") \wedge (bPRk) \wedge (ID(k) = "id") \wedge (A(k) \cap A(t) \cap A(b) \neq \emptyset) \rightarrow tPRp.$$

2. Объект-часть или объект-участник может быть идентифицирован через объект, частью которого он является или участвует, посредством указания значений идентификационных свойств других частей или участников (рис. 7).

Формально данное правило можно записать следующим образом:

$$((bPRTt) \wedge (bPRTf)) \vee ((bPRCt) \wedge (bPRCf)) \wedge (fPRp) \wedge (ID(p) = "id") \wedge (bPRk) \wedge (ID(k) = "id") \wedge (A(k) \cap A(t) \cap A(b) \neq \emptyset) \rightarrow tPRp.$$

Необходимо заметить, что применение данного правила требует введения дополнительного ограничения, наряду с условием связности запросов по свойству –  $k$ , в отношении того, что оно должно также быть идентифицирующим. Однако такое ограничение определяется в случае нахождения объектов  $b$  и  $f$  в разных онтологиях. Если же объекты  $b$  и  $f$  находятся в одной онтологии  $A$ , то роль связующего свойства выполняет внутренний идентификатор найденного экземпляра концепта  $b$ .

Перечисленные правила приводят к расширению запроса в глубину, но в данном случае это производится неявно, что при практическом использовании позволяет повысить интеллектуальность интерактивного конструктора запросов путем использования метаинформации и логического вывода, что, в свою очередь, облегчает работу пользователя с системой.

Расширение запроса в ширину происходит путем определения дополнительных объектов поиска из числа тех, которые входят во множество синонимов и ассоциаций. При этом пользователю, помимо самих объектов отображаются также их гиперонимы, что позволяет представить контекст использования того или иного синонима или ассоциации и помогает пользователю расширить множество объектов поиска более осознанно (рис. 8).

Далее по каждому отображенному объекту проводится при необходимости процесс расширения запроса в глубину, правила которого были рассмотрены ранее.

Таким образом, расширение запроса по поиску начального объекта  $t$  в ширину можно определить как множество простых запросов, объекты поиска которых имеют связи синонимии и ассоциации между собой:

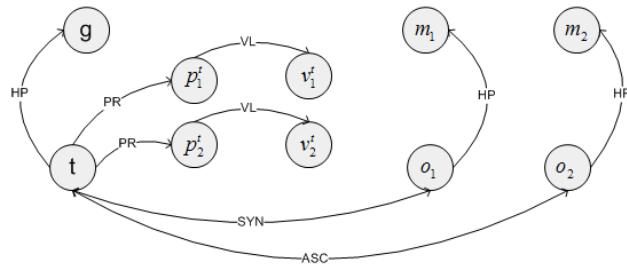


Рис. 8. Расширение запроса в ширину синонимами и ассоциациями

$$EXQ(t) = SQ(t) \cup \{ \langle o_i | (o_i ASC t \wedge WS(o_i, t) > DWS) \vee (o_i SYN t), \{ \langle p_j^i | o_i PR p_j^i, v_j^i | p_j^i VL o_i \rangle \} \},$$

где  $SQ(t)$  – простой запрос.

## 5. Практическая реализация

Для практической реализации представленных в работе результатов использованы технологии Semantic Web. В частности, общесистемный тезаурус был разработан с использованием модели данных SKOS (Simple Knowledge Organization System, простая система организации знаний) (World Wide Web Consortium, 2009), который является подмножеством RDF (Resource Description Framework) (World Wide Web Consortium, 2004) и используется для создания моделей, выражающих базовую структуру и содержимое концептуальных схем, таких как тезаурусы, классификационные схемы, списки именованных объектов, таксономий и других подобных типов словарей.

Так как SKOS не может в полной мере отобразить все элементы тезауруса в связи с тем, что предоставляет легковесный язык концептуального моделирования, то некоторые элементы

отображаются с помощью языка веб-онтологий OWL. В результате расширенный тезаурус представляется как OWL-Full онтология. В качестве низкоуровневого языка запросов, в которые транслируются запросы в терминах тезауруса, был использован SPARQL, являющийся языком запросов к RDF-графам. Использование технологий Semantic Web, продвигаемых консорциумом W3C и ориентированных на оперирование семантическими метаданными, позволит избежать, вследствие их широкого применения в информационных системах, проблем на синтаксическом и системном уровне и даст возможность использовать разработанные модели и технологию в других системах.

## 6. Заключение

В работе рассмотрена проблема использования общесистемного расширяемого тезауруса в качестве основы интеллектуального интерфейса информационной системы распределенного семантического поиска. В ходе проведенных исследований были получены следующие основные результаты:

1) Предложено расширенное элементами онтологии DOLCE и метасвойствами определение разделяемого тезауруса, позволяющее сохранить понятийную семантику, тем самым повысить уровень выразительности тезауруса.

2) Разработана технология формирования запроса на основе правил определения возможных вариантов его усложнения, что позволяет использовать тезаурус в качестве основы для разработки интеллектуального пользовательского интерфейса информационных систем распределенного семантического поиска.

Новизна полученных результатов заключается в том, что при формировании запроса используются не только онтологические отношения между объектами предметной области, но также и их отношения, определяемые в контексте онтологии верхнего уровня и наличия у них определенных метасвойств, что позволяет вводить в запрос понятия и атрибуты разных онтологий. Это является существенным результатом в направлении решения проблемы семантической интеграции информации. Расширение тезауруса концептами и отношениями онтологии верхнего уровня и метасвойствами обеспечило также возможность дальнейшего совершенствование методики оценки близости терминов в отношении введения критерия сходства на основе их понятийной семантики (Скворцов, 2009).

Полученные результаты могут найти широкое применение при решении практических задач, связанных с семантической интеграции информации, построении онтологических моделей предметных областей и задач, семантическим поиском и разработкой интеллектуального пользовательского интерфейса.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 08-07-00301-а "Разработка информационной технологии и распределенной информационно-аналитической среды поддержки инновационной деятельности").

## Литература

- Baader F., Narendran D.** Unification of concepts terms in description logics. *Journ. of Symbolic Computation*, v.31(3), p.277-305, 2001.
- Bechhofer S., Stevens R., Ng G., Jacoby A., Goble C.** Guiding the user: An ontology driven interface. *UIDIS*, p.158-161, 1999.
- Benzi F., Maio D., Rizzi S.** VISIONARY: A viewpoint-based visual language for query in relational databases. *J. Vis. Lang. Comput.*, v.10(2), p.117-145, 1999.
- Berners-Lee T., Hendler J., Lassila O.** The Semantic Web. *Scientific American*, v.284(5), p.34-43, 2001.
- Catarci T., Dongilli P., Dimascio T., Franconi E., Santucci G., Tessaris S.** An Ontology-based Visual Tool for Query Formulation Support. *In Proceedings of the 16<sup>th</sup> European Conference on Artificial Intelligence*, 2004.
- Catarci T., Francesca M., Levialdi S., Batini C.** Visual query systems for databases: A survey. *Journ. of Visual Languages and Computing*, v.8(2), p.215-260, 1997.
- Feikje H., Chris M., Peter E.** Evaluating an Ontology-Driven WYSIWYM Interface. *Proceedings of the 5<sup>th</sup> International Conference on Natural Language Generation*, p.138-146, 2008.
- Garshol L.** Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all. *Journ. of Information Science*, v.30, N 4, p.378-391, 2004.
- Grubber T.** A translation approach to portable ontology specifications. *Knowledge Acquisition*, v.5(2), p.199-220, 1993.
- Guarino N., Welty C.** An overview of OntoClean. *Handbook on Ontologies*, Springer, Berlin, p.151-172, 2004.
- Guarino N., Welty C.** Identity and subsumption. *The Semantics of Relationships: An Interdisciplinary Perspective*, p.111-125, 2002.

- Guarino N., Welty C.** Ontological analysis of taxonomic relationships. *Proceedings of ER-2000: The 19<sup>th</sup> International Conference on Conceptual Modeling*. Springer-Verlag LNCS, October, 2000.
- Keim D.** Databases and visualization, tutorial. *Int. Conf. on Management of Data Montreal, Canada*, p.543, 1996.
- Masolo C., Borgo S., Gangemi A., Guarino N., Oltramari A., Schneider L.** DOLCE: A descriptive ontology for linguistic and cognitive engineering. *DOLCE documentation*, URL: <http://www.loa-cnr.it/DOLCE.html>, 2006.
- McGuinness D.** Question answering on the semantic Web. *In IEEE Intelligent Systems*, v.19(1), p.82-85, 2004.
- Pichat E., Saker D.** An Automatic and Cooperative Visual Database Interface. *In: Proc. of the IFIP 2.6 Third Working Conference on Visual Database Systems (VDB-3), Lausanne, Switzerland*, p.333-348, 1995.
- Sowa J.** Building, sharing, and merging ontologies. URL: <http://www.jfsowa.com/ontology/ontoshar.htm>, 2009.
- World Wide Web Consortium. SKOS Simple Knowledge Organization System Reference W3C Recommendation. URL: <http://www.w3.org/TR/skos-reference>, 2009.
- World Wide Web Consortium. RDF Primer W3C Recommendation. URL: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210>, 2004.
- Артемьева Л.** Специализированные оболочки интеллектуальных систем для сложно-структурированных предметных областей. *Труды XI национальной конференции по искусственному интеллекту с международным участием КНИИ-2008*, т.1, с.95-103, 2008.
- Витгих В., Минаков И.** Интеграция профессиональных знаний: основные положения подхода. *Проблемы управления и моделирования в сложных системах: Тр. IX Междунар. конф., Самара, 22 июня – 29 июня 2007. Самара, СНЦ РАН*, с.191-197, 2007.
- Гладун А.Я., Рогошина Ю.В.** Формирование тезауруса предметной области как средства моделирования информационных потребностей пользователя при поиске в Интернете. *Вестник компьютерных и информационных технологий*, № 1, с.56-68, 2007.
- Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д.** Онтологии и тезаурусы: модели, инструменты, приложения. Учебное пособие. М., *Интернет-Университет Информационных Технологий, БИНОМ. Лаборатория знаний*, 172 с., 2008.
- Калнинченко Л.** Методы и средства интеграции неоднородных баз данных. М., *Наука*, 420 с., 1983.
- Клещев А.** Использование онтологий в разработке программного обеспечения. *Мат. Всерос. конф. с междунар. участием "Знания – Онтологии – Теории" (ЗОНТ-07) 14-16 сентября 2007г., Новосибирск, Изд-во ООО "Омега Принт"*, т.1, с.122-130, 2007.
- Когаловский М.** Энциклопедия технологий баз данных. М., *Финансы и статистика*, 800 с., 2005.
- Ломов П.А., Шишаев М.Г.** Интеграция данных на основе онтологий для обеспечения информационной поддержки управленческих решений. *Труды Института системного анализа РАН, М., Книжный дом "ЛИБРОКОМ"*, т.39, с.159-173, 2008.
- Ломов П.А., Шишаев М.Г.** Разработка метода семантической интеграции информации в сфере государственного и муниципального управления. *Труды XI Всерос. науч. конф. "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Петрозаводск, КарНЦ РАН*, с.78-86, 2009.
- Ломов П.А., Шишаев М.Г.** Использование базовых классов для установления смысловой эквивалентности в семантически гетерогенных информационных ресурсах. *Информационные технологии в региональном развитии: Сб. науч. тр. ИИММ КНЦ РАН, вып. VIII, Апатиты, КНЦ РАН*, с.62-68, 2008.
- Обухова О.Л., Бирюкова Т.К., Гершкович М.М., Соловьев И.В., Чочиа А.П.** Метод динамического создания связей между информационными объектами базы знаний. *Труды XI Всерос. науч. конф. "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Петрозаводск, КарНЦ РАН*, с.39-45, 2009.
- Осипов Г.С., Завьялова О.С., Смирнов И.В., Тихомиров И.А.** Интеллектуальный семантический поиск с привлечением средств метапоиска. *Труды междунар. конференции ИАИТ'2005, Киев*, 2005.
- Поспелов Д.** Моделирование рассуждений. Опыт анализа мыслительных актов. М., *Радио и связь*, 184 с., 1989.
- Скворцов Н.** Специфика подходов к отображению онтологий. *Труды семинара "Знания и Онтологии \*ELSEWHERE\* 2009", М.*, с.91-103, 2009.
- Смирнов А., Пашкин М., Левашова Т., Шипов Н.** Основные принципы организации систем быстрой интеграции знаний. *Труды СПИИРАН, СПб., СПИИРАН*, вып. 11, т.2, 2002.