OSTIS-2016
(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ ДЛЯ СОЗДАНИЯ СЦЕНАРИЕВ ОБРАБОТКИ ДАННЫХ

Глоба Л.С.

Национальный технический университет Украины «Киевский политехнический институт», г. Киев, Украина

lgloba@its.kpi.ua

Предложен метод формирования сценариев обработки данных, основанный на онтологическом описании как информационных, так и вычислительных ресурсов (web-сервисов). Поиск и принятие решения о возможности включения web-сервиса в состав сценария выполняется на основе оценки семантической близости, которая учитывает входы и выходы операций, множество типов параметров и «шаблоны сервисов». Установление порядка выполнения сервисов в рамках одного сценария выполняется с помощью упорядоченного дерева, порядок обхода вершин которого соответствует порядку вычислений. Такой метод позволяет автоматизировать процесс формирования сценариев обработки больших объемов данных, расположенных в среде Интернет.
Ключевые слова: сценарий обработки данных, web-сервис, метаграфы, онтологические модели.

Введение

Проблематика обработки значительных объемов информации, накапливаемых в различных сферах деятельности: производстве, здравоохранении, торговле, государственном управлении и т.д., приобретает все большую остроту и определяется сегодня термином Big Data (большие данные). Ожидается, что развитие и начало широкого использования всех источников больших данных инициирует проникновение технологий больших данных как в научно-исследовательскую деятельность, так и в коммерческий сектор, а также сферу государственного управления.

В качестве определяющих характеристик для больших данных отмечают «четыре V»: объём, скорость (в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов), многообразие (в смысле возможности одновременной обработки различных типов структурированных, полуструктурированных и неструктурированных данных с точки зрения двух аспектов: синтаксиса и семантики), стоимость, которая особенно связана с их коммерческой ценностью [Business Information Systems, 2014].

Современные требования к огромным объемам уже накопленных данных состоят в переходе к поддержке принципа:

ДАННЫЕ→ ИНФОРМАЦИЯ→ЗНАНИЯ.

Данные рассматриваются как массивы несвязных разнородных территориально-распределенных данных, по ним можно лишь выполнить сбор статистики или осуществить элементарный поиск.

Информация рассматривается как набор данных агрегированных по определенным характеристикам или требованиям и представляет собой предварительно обработанные и обобщенные данные, классифицированные по определенному признаку или объединённые по определённой тематике. Получение информации из данных выполняется с использованием методов аналитической обработки данных, обобщения, группирования, математических моделей и методов, методов классификации, структуризации и систематизации, аналитической обработки данных, обобщения, системного и сценарного анализа и т.д.

Знания рассматриваются как набор данных, сгенерированных на основе имеющихся данных. Получение знаний из данных и информации осуществляется на основании методов экспертных оценок, интеллектуальной обработки информации и интеллектуального анализа данных, прогнозирования, моделей представления знаний, подсистем устойчивого развития и других.

В качестве научного базиса при создании многих интеллектуальных информационных систем, призванных обрабатывать значительные объемы информации (транзакционные системы, системы поддержки принятия решений, порталы знаний и т.п.) выступают следующие математические и интеллектуальные модели и соответствующие им методы: семантические сети, онтологическая

модель, нечеткие базы знаний, метаграфы, метаописания, сценарии динамической обработки заданий пользователя. Создание таких моделей и методов является непростой задачей и требует, как знаний в определенных областях деятельности, так и понимания общесистемных взаимосвязей. Инженерная и научная сферы деятельности имеют особенно острую потребность в средствах извлечения знаний, важнейшими среди которых есть сценарии динамической обработки заданий пользователя. С точки зрения задач инженерных областей или научных исследований очевидно, что выполнение конкретных расчетов в процессе анализа требует решения ряда определённых подзадач, для чего необходимо не только описать доступные пользователю расчетные задачи, но и установить связи между конкретными сервисами обработки (набором сервисов) для реализации конкретных рабочих процессов (workflow (WF)) в рамках общих рабочих процессов и связей сервисов с информационными ресурсами, которые необходимы для выполнения расчетов. Важным фактором является оптимизация таких связей. Кроме того, в зависимости от параметров, заданных пользователем, тематикой и особенностями выполняемых расчетов, последние могут выполняться с помощью разных алгоритмов. В связи с этим при построении реальных научных и инженерных задач используют их декомпозицию на подзадачи в зависимости от направленности общей задачи и параметров, необходимых для выполнения расчета [Новогрудська, 2015]. Подрасчеты динамически компонуют в общую задачу, при этом одна подзадача может использоваться в ряде разных общих задач. Это обуславливает потребность учета и описания логики связности расчетных задач и их корректной интеграции в информационное пространство портала. Таким образом, при построении единой информационной среды в инженерных и научных сферах деятельности недостаточно использования семантических или онтологических моделей, в этом случае возникает ряд дополнительных проблем:

- определение механизма связывания информационных элементов с функциональными, а также формирование сложного расчета «на лету» на момент его выполнения в зависимости от текущих условий;
- описание характеристик функциональных и информационных элементов в соответствии с предметной областью;
- описание последовательности взаимодействия функциональных элементов для реализации глобального расчета.

Объективными трудностями при создании эффективных моделей и методов обработки больших объемов информации являются недостаточные систематизация и структуризация информации, неоднозначность метаописаний в онтологиях многих научных и инженерных порталов в случае, если информация разбросана в Интернет, невозможность автоматизированного создания workflow для извлечения знаний, хранящихся в распределенных информационных хранилищах.

В статье предлагается с целью систематизации больших объемов данных и повышения производительности их обработки объединить три важные группы математических моделей и методов: онтологические модели как средство описания наборов данных в сети Интернет, теорию метаграфов и нечеткую логику как средства обработки больших объемов информации и извлечения знаний из них.

1. Модели и методы представления и обработки данных

Ключевая тенденция развития Интернета заключается в переходе от документов машиночитаемых к документам, которые машина понимает, а это требует наличия средств понимания смысла при обработке больших объемов данных. Процесс такой обработки требует автоматизированного выполнения ряда этапов: создания онтологий, связывающих ресурсы под конкретные задачи на основании онтологических моделей; формирования поисковых запросов «на лету» с учетом возможной неточности метаописаний и их эффективного исполнения на основании автоматизированных рабочих процессов (сценариев обработки информации, учитывающих альтернативы). Структура программных средств, построенных с использованием онтологических моделей состоит из следующих компонент:

- онтология,
- коллекционер онтологической информации о ресурсах,
- конструктор запросов,
- формирователь ответов.

Для более четкого понимания и анализа связей между элементами онтологической модели удобно использовать графовые структуры, чтобы визуализировать связи между отдельными компонентами. Это облегчает процесс структурирования данных в процессе их обработки. Кроме того, графы используют не только в качестве основы для визуализации в процессе получения вывода, но они также могут служить базисом алгебраических операций, которые позволяют более строго проводить обработку больших объемов информации. Однако, отображение нечеткости метаописаний свойств отдельных объектов сложно отобразить в виде простой графовой структуры, поэтому для этих целей удобно использовать специальный вид графов, который называют метаграфами [Amit Basu, 2008].

1.1. Онтологические модели

Онтологические модели можно использовать, для:

- моделирования доменов программных сред;
- формирования автоматизированных рассуждений и получения выводов;

– автоматического программирования и удовлетворения спецификаций;

– построения программных систем, основанных на моделях;

– технологий семантики (онтологий и семантического web).

Особенностью онтологических систем является способность поддерживать систематизацию знаний и информационных ресурсов моделируемых областей знаний, обеспечивать содержательный доступ к ним. В рамках общей онтологии выделяют специфические онтологии и проектирование общей онтологической модели, которую создают путем доопределения и развития базовых онтологий. Онтологии включают как описание предметной области (ПО), так и описание релевантных ей ресурсов. Описание предметной области включает совокупность терминов и отношений, семантически значимых для данной предметной области, а также правил, согласно которым можно строить утверждения об элементах ПО. Описание ресурса сети включает список ссылок на страницы и сайты, а также описание связей между ними.

Основные преимущества использования онтологических моделей в качество одной из составляющих научного базиса:

– Онтологии позволяют сделать систему предметно-ориентированной, благодаря чему пользователь может задавать поисковый запрос в хорошо известных ему терминах предметной области.

– Онтологии являются новым интеллектуальным средством для поиска ресурсов в сети Интернет, новыми методами представления и обработки знаний и запросов.

– Онтологии способны точно и эффективно описать семантику данных для определенной предметной области и решить проблему несовместимости и противоречивости данных.

– Онтологии владеют собственными средствами обработки (логического вывода) и соответствующими моделями и методами семантической обработки данных.

1.2. Метаописания

Основным элементом онтологической модели предметной области знаний являются метаописания ее объектов, именно с помощью метаописаний и устанавливаются связи между объектами.

Метаописание – объект рассматриваемой ПО, который служит для определения элементов системы, информационных и вычислительных ресурсов, учитывает их смысловое значение. Метаописания позволяют однозначно идентифицировать элементы и характеризуются двойкой: $M_i = \langle A, V \rangle$, где *A* – атрибут, *V* – значение атрибута.

Синтаксис обозначения метаописаний: порядковый номер метаописания - номер элемента, который характеризируется метаописанием, название атрибута, отражающее его смысловое значение и собственно значение атрибута. Формат записи:

$$\textit{Атрибут} = \textit{значение}\ M_i(\mathrm{E}): Atr = "Value".$$

Для порталов знаний выделено несколько видов метаописаний:

– Стандартные метаописания – определяются\строятся согласно стандарту дублинского ядра, и описывают основные характеристики элементов.

– Специфические метаописания – метаописания связные с предметной областью.

– Метаописания, устанавливающие связь, – используются для определения связи элемента с другими элементами, например, информационными ресурсами.

Для установления связей между метаописаниями используют правило объедения, которое формулируется: «Если два элемента имеют одинаковые метаописания, то они связаны». Существует два вида связности: 1) элементы компонуются для объединения в общий элемент либо 2) один из элементов включается как подъэлемент во второй.

Сравнение элементов происходит по значениям атрибутов, т.е. метаописаний. Вначале происходит сравнение значений, потом проверка на равенство, далее проверка самих элементов для организации их в более общий элемент. Однако, такое сравнение не всегда является точным или хотя бы числовым, поэтому для нахождения объектов по их метаописаниям и принятия решения о принадлежности элементов к тем или иным классам требуется решение задачи классификации на нечеткой области значений.

1.3. Обнаружение web-сервисов в мульти-онтологической среде

С целью извлечения знаний используют сценарии обработки заданий пользователя, называемые workflow, которые представляются как объединение web-сервисов, хранящихся в распределенных информационных хранилищах порталов знаний, в цепочку последовательно-параллельных вычислений. Для такой реализации workflow и последующей обработки полученных данных формируют семантические отношения между процессами (web-сервисами), чтобы отразить сходства и различия между запросами конечных пользователей и моделями, описывающими web-сервисы. Обнаружению web-сервисов в среде Интернет посвящены работы [Larysa Globa, 2015]. Однако, в предлагаемых подходах имеется ряд минусов. Прежде всего, это проблема неоднозначности языковых конструкций, поэтому web-службы поиска используют метаописания онтологических понятий, написанные исключительно с помощью инструментария, такого как: язык DAML (DARPA Agent Markup Language), описания в виде RDF-графов (Resource Description Framework), язык LARKS (Language for Advertisement and Request for Knowledge Sharing),

описания для целей машинного обучения LSD (Learning Source Descriptions), основанные на распределении вероятностей экземпляров различных заранее определенных онтологических моделей. Несмотря на значительное число имеющихся решений, остается существенным влияние человеческого фактора и, кроме того, не поддерживаются описания для входов и выходов web-сервисов. Для улучшения релевантности обнаружения web-сервисов в среде Интернет предлагается использовать семантическую близость, которая учитывает входы и выходы операций, например, множество типов параметров и «шаблоны сервисов».

Подход к поиску web-сервисов в нескольких онтологических средах состоит из трех последовательных этапов:

1. Создание «шаблона сервиса», базируясь на запросе пользователя;
2. Сравнение «шаблона сервиса» с несколькими web-сервисами, которые были определены как сервисы-кандидаты;
3. Возвращение web-сервисов, удовлетворяющих минимально приемлемой оценке сходства с запрошенными пользователем в упорядоченном списке.

Семантический шаблон сервиса описывает запрос пользователя. Это позволяет конечному пользователю определить набор необходимых операций, их свойств, входов и выходов. Шаблон сервиса не имеет конкретной реализации, потому что рассматривается как промежуточная абстракция - прокси web-сервис.

Шаблон сервиса может быть определен как:

$$ST = \langle N_{ST}, D_{ST}, OPs_{ST} \langle N_{OP}, D_{OP}, O_{OP}, I_{OP} \rangle \rangle, \quad (1)$$

где N_{ST} - имя web-сервиса,
D_{ST} - текстовое описание web-сервиса,
OPs_{ST} - множество операций web-сервиса.

Каждая из операций web-сервиса в свою очередь определяется с помощью: N_{OP} - имени операции, D_{OP}- текстовое описания операции, O_{OP} и I_{OP} - входных и выходных параметров операции.

Шаблон сервиса сравнивается с набором сервисов-кандидатов (CS) - шаблонов сервисов, полученных с помощью анализа web-сервисов из предварительно определенного набора web-сервисов. При сравнении используют для оценки синтаксическое и семантическое сходство, определяемое как:

$$\Theta(ST, CS) = \frac{w_H \cdot H(ST, CS) + w_\Phi \cdot \Phi(ST, CS)}{w_H + w_\Phi}, \quad (2)$$

где $\Theta(ST, CS)$ - обобщенная оценка, $H(ST, CS)$ - синтаксическое сходство, $\Phi(ST, CS)$ - семантическое (функциональное) сходство, w_i - весовой коэффициент, соответствующий каждому типу сходства, предназначенной для боле е гибкого управления критерием сравнения. Подробное описание предложенного алгоритма обнаружения web-сервисов в среде Интернет приведено в работе [Larysa Globa, 2015].

1.4. Динамическое формирование и выполнение сценариев обработки заданий пользователя

Основной задачей при проектировании последовательности элементов сценариев обработки заданий пользователя является упорядоченное объединение соответствующих частичных элементов сценария для включения их в сложный сценарий. Сложный сценарий пользователя можно идентифицировать со сложным инженерным или научным расчетом, который реализуется в программной среде в виде конкретного workflow, а каждый частичный элемент сценария идентифицируется с подрасчетом решения конкретной задачи в ПО, реализуемым в виде web-сервиса. Множество частичных элементов сценария – чСэ хранится в независимом хранилище, которое не имеет связей с хранилищем сложных сценариев – сСэ. Структура каждого из сложных и частичных сценариев Сэ включает название и параметры Сэ. Поиск частичных сценариев чСэ, которые необходимо включить в сложный, выполняется в соответствии с подходом, представленным в разделе 1.3, и базируется на сравнении параметров чСэ, поскольку в сложный сСэ включаются только те частичные, у которых есть общие параметры, или области значений соответствующих параметров пересекаются. Формальное описание сценариев Сэ:

1) множество сложных сСэ:

$$R^{з} \ni r_k^{з}, r_k^{з} = \langle T_k^{з}, p_{kj}^{з} \rangle,$$

где $R^{з}$ - множество сложных сценариев сСэ,
$r_k^{з}$ - *k*-й сценарий Сэ из множества сложных сценариев сСэ,
$T_k^{з}$- название *k*-ого сценария Сэ из множества сложных сценариев сСэ,
$p_{kj}^{з}$- *p*-й параметр *k*-ого сценария Сэ из множества сложных сценариев сСэ.

2) множество частичных сценариев чСэ:

$$R^{ч} \ni r_l^{ч}, r_l^{ч} = \langle T_l^{ч}, p_{lq}^{ч} \rangle,$$

где $R^{ч}$ - множество частичных сценариев чСэ,
$r_l^{ч}$- *l*-й сценарий Сэ из множества частичных сценариев чСэ,
$T_l^{ч}$- название *l*-ого сценария Сэ из множества частичных сценариев чСэ,
$p_{lq}^{ч}$- *q*-й параметр *l*-ого сценария Сэ из множества частичных сценариев Сэ.

Множества $P_{kj}^{з}$ и $P_{lq}^{з}$ могут пересекаться или не пересекаться, что отображает наличие у обоих Сэ одинаковых параметров или их отсутствие соответственно. Сложный сценарий Сэ формируется из частичных сценариев чСэ на базе сравнения значений параметров, которые используются в сложном сценарии Сэ. Это позволяет автоматизировать процесс формирования

сложного сценария, устанавливать порядок его формирования, выполнять формирование на «лету». Предлагается сложный сценарий Сэ представить в виде упорядоченного дерева, порядок обхода вершин которого соответствует порядку вычислений с помощью найденных web-сервисов, что упростит процедуру его формирования и модификации. Исходя из того, что рассматриваемое дерево сценария является упорядоченным, деревом с корневым узлом и заданным порядком прохождения дочерних узлов, следует, что последовательность выполнения частичных сценариев Сэ будет соответствовать введенному порядку на дереве и может выполняться в процессе построения дерева, т.е. на «лету». Этапы метода динамического формированию сложного сценария:

Этап 1. На первом этапе происходит отбрасывание из множества $R^{ч}$сценариев, в которых ни один $p_{kj}^{з}$ не равняется никакому $p_{lq}^{ч}$.

Этап 2. На этом этапе происходит сравнение значений $p_{lq}^{ч}$ сценариев, отобранных на этапе 1, для отсечения подмножества сценариев, которые имеют общие параметры, но множества их значений не пересекаются.

Этап 3. На этапе 3 проводится упрощение формулы частичного сценария Сэ.

Подробное описание метода формирования сложного сценария представлено в работе [Новогрудська, 2015]. Полученный сложный сценарий для определенного инженерного, научного расчета или для процесса анализа больших объемов данных может быть представлен как шаблон для дальнейшего использования.

Для визуального анализа полученного на «лету» сценария предлагается использовать метаграфы [Globa, 2015]. Метаграф - это двойка: порождающее множество вершин и метавершин и множество ребер. В определении метаграфа важным является рассмотрение двух составляющих: множество метавершин и множество метаребер. Так как для решения задачи визуализации метаграфа важны типы узлов и их соотношения, целесообразно выделить отдельно множество вершин и множество метавершин метаграфа. Выделять и отдельно рассматривать метаребра нет необходимости, так как при визуализации метаграфа важным является наличие ребер, но не их тип. В связи с этим множество ребер будет содержать все ребра метаграфа, независимо от того, какие типы узлов метаграфа они соединяют. Цель применения метаграфов заключается в возможности моделирования бизнес-процессов с использованием диаграмм выполнения работ (workflow), визуальном представлении и анализе аномалий в нечетких наборах данных. С точки зрения задачи моделирования бизнес-процессов (workflow) при нечеткости метаописаний web-сервисов, вершины можно представить в виде метаописаний тех web-сервисов, при поиске которых все критерии поиска выполнены, а метавершины – это метаописания таких web-сервисов, критерии поиска которых удовлетворены частично или когда для решения одной задачи найдено множество сервисов-претендентов. В таком случае визуальный анализ аномалий в метаграфе полученного workflow может позволить избежать ошибочных решений. Подробное описание метода визуализации workflow средствами метаграфов представлено в работах [Amit Basu, 2008, Globa, 2015]. Таким образом, использование онтологий в комплексе с методами семантического поиска позволяет сформировать целостную методологию для создания сценариев (workflow) решения сложных задач в среде Интернет, а визуализация сценариев обработки информации в виде метаграфов повысить релевантность полученных решений, что упростит выполнение сложных наукоемких расчетов широкому кругу пользователей в различных сферах деятельности.

2. Комплекс инструментальных средств построения сценариев обработки данных

Комплекс инструментальных средств построения сценариев обработки данных является частью общесистемной структуры программных средств порталов знаний и включает подсистему ведения онтологий (коллекционер онтологической информации), конструктор запросов и формирователь ответов. Коллекционер онтологической информации о ресурсах использует стандартные механизмы поиска, последовательно обрабатывает ссылки на документы, страницы, сайты, определяя их тематику, структуру, степень релевантности между тематикой и информацией, которая в них содержится, сохраняет их в виде онтологии. Коллекционер анализирует новые источники и проверяет старые. Формирователь ответов обрабатывает источники, отслеживает степень их релевантности понятиям и выдает список наиболее подходящих для данного понятия ресурсов. Портал знаний имеет язык запросов, близкий к естественному языку. Это требует наличия в нем соответствующего лингвистического обеспечения, одним из основных компонентов которого является словарь терминов, органически связанный с онтологией. На основе лингвистического обеспечения решаются следующие важные задачи: задание лингвистической информации о данной предметной области, для чего используют словарь, организованный с учетом проблемно-ориентированной лексики, морфологии, словообразования, синонимии.

В виде онтологической модели выполняется описание системы знаний о предметной области как комплекса понятий, связанных между собой отношениями.

Один из возможных вариантов реализации подсистемы ведения онтологий на портале знаний представлен на рисунке 1. Такая подсистема ориентирована на участие экспертов из конкретных областей знаний. Она позволяет им собрать и

систематизировать в рамках единого информационного пространства обширные знания и данные из этих областей знаний. Для построения онтологий и управления ими служит редактор онтологий, реализованный как web-сервис. Формы для ввода конкретных информационных объектов и связей между ними автоматически генерируются на основе онтологии. Управление информационным контентом осуществляется с помощью редактора данных, который позволяет создавать, редактировать и удалять информационные объекты и связи между ними.

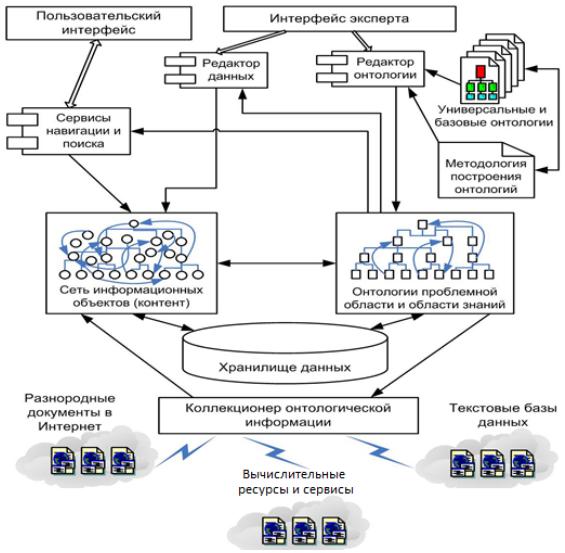

Рисунок 1 – Подсистема ведения онтологий на портале знаний

Комплекс инструментальных средств построения бизнес-процессов и обработки больших объемов данных включает набор интегрированных программных технологий и продуктов для обработки и хранения информации и позволяет:

– предоставлять пользователям доступ к данным, которые расположенным в Интернет-среде;
– предоставлять данные группе конечных пользователей в форме, которая соответствует их коллективному представлению о данных;
– сокращать времени ответа на запрос;
– предоставлять данные, структурированные в соответствии с требованиями доступа;
– упрощать выполнение задач очистки, загрузки, преобразования, интеграции и анализа данных;
– формировать сценарии обработки данных «на лету».

Таким образом, при обращении к порталу знаний пользователь имеет возможность получить в ответ ресурсы семантически релевантные его запросу и обработать их с помощью соответствующего запросу сценария.

Заключение

Рассмотренный в статье подход к автоматизации построения сценариев обработки данных отличается использованием онтологических моделей для описания информационных и вычислительных ресурсов, теории метаграфов для визуализации и анализа полученных сценариев, является особенно эффективным в таких сложных сферах деятельности как инженерия и научные исследования.

Использование онтологий в комплексе с методами семантического поиска позволяет сформировать сценарии (workflow) решения сложных наукоемких задач в среде Интернет.

Визуализация сценариев обработки информации в виде метаграфов позволяет повысить релевантность полученных «на лету» сценариев, использование шаблонов упрощает выполнение сложных наукоемких расчетов широкому кругу пользователей в различных сферах деятельности.

Среди вопросов, которые необходимо дополнительно исследовать, можно выделить исследование неоднозначности построенных сценариев в случае смысловой нечеткости описания онтологий, многозначности описаний одних и тех же ресурсов в среде Интернет, а также развитие подхода к созданию сценариев обработки больших объемов данных с целью повышения производительности их обработки.

Библиографический список

[Business Information Systems, 2014] Lecture Notes in Business Information Processing / Witold Abramowicz, Angelika Kokkinaki (Eds.) - ©Springer International Publishing Switzerland, 2014, 265 .

[Новогрудська, 2015] Новогрудська Р.Л. Інформаційна технологія створення та підтримки порталів інженерних знань / Р.Л. Новогрудська // диссерт. на соиск. ученой степени к.т.н.; – Киев, 2015, 186 с.

[Amit Basu, 2008] Amit Basu, RobertW. Blanning Metagraphs and their applications/ A. Basu, R. Blanning // A. Basu, R. Blanning Library of Congress Control Number: 2006930395, Springer Science+Business Media, LLC, 2008, 173 p.p.

[Larysa Globa, 2015] L. Globa, M. Kovalskyi, A. Stryzhak Increasing web services discovery relevancy in the multi-ontological environment/A. Wiliński, Jerzy Pejaś Advances in Intelligent Systems and Computing// ISBN 978-3-319-15147-2 (eBook): Springer International Publishing Switzerland, 2015, p. 335-344

[Globa, 2015] L.Globa, M.Ternovoy, O.Shtogrina, O. Kryvenko Based on Force-Directed Algorithms Method for Metagraph Visualization /A. Wiliński, Jerzy Pejaś Advances in Intelligent Systems and Computing// ISBN 978-3-319-15147-2 (eBook): Springer International Publishing Switzerland, 2015 – p. .359-370.

ONTOLOGIES USING FOR DATA PROCESSING SCENARIOS DESIGNING

Globa L.S.

NTUU «KPI », Kiev, Ukraine

lgloba@its.kpi.ua

The paper deals with method of the data processing scenarios (workflow) designing based on the ontological description of the information and computing resources in conditions of the fuzzy metadata presence, evaluation of semantic similarity, the data processing scenario tree visualization as matagraph for the scenario problem control.