

# Построение модели онтологии предметной области нефтедобычи

Т.Г. Дидык  
Институт экономики и управления  
Уфимский государственный авиационный  
технический университет  
Уфа, Россия  
e-mail: tanayg@mail.ru

Е.И. Филосова  
Институт экономики и управления  
Уфимский государственный авиационный  
технический университет  
Уфа, Россия  
e-mail: filsova@yandex.ru

Ю.В. Шаронова  
Институт экономики и управления  
Уфимский государственный авиационный технический университет  
Уфа, Россия  
e-mail: hedviga@mail.ru

## Аннотация<sup>1</sup>

В работе представлена технология моделирования онтологии для описания свойств скважин при нефтедобыче и применения полученных результатов для решения задачи эффективного поиска научной информации по указанному направлению в Интернет источниках.

## 1. Введение

Целью данных исследований являлась разработка технологии публичного, отраслевого или корпоративного оперирования информационными ресурсами с большим количеством пространственно привязанных данных. Для решения данной задачи была применена технология онтологического моделирования. Онтология понятий лежит в основе развития современных компьютерных технологий. Данные исследования проводятся как в нашей стране, так и за рубежом. Теме разработки информационных систем на основе онтологий посвящены работы Лапшина В.А. (объемный обзор инженерных приложений онтологий дан в его монографии [1]), Нахайновой Л.В. [2], Доброva Б.В., Иванова В.В., Лукашевич Н.В., Соловьева Н.В. [3] и др.

Главной задачей моделей, построенных на основе онтологий, является развитие процедуры навигации по информационным ресурсам. Настоящее исследование посвящено онтологиям в сфере недропользования на примере нефтедобычи. Анализ

---

Труды Шестой всероссийской научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 28-31 мая, Уфа-Ставрополь, Россия, 2018

научных работ по данной тематике, опубликованных за последние несколько лет, позволяет сделать вывод о недостаточной активности авторов в этой предметной области, хотя в других сферах приложения онтологий широко применяются.

## 2. Моделирование предметной области нефтедобычи

Для представления пространственных данных широко используются геоинформационные системы (ГИС). Применение ГИС в области нефтедобычи играет особую роль для информационного обеспечения предприятий. При освоении месторождения, добыче и транспортировке накапливаются и обрабатываются большие объемы пространственной геолого-физической информации. Затем накопленные данные используются при эксплуатации месторождения. Выбор метода организации данных в геоинформационной системе, и, в первую очередь, модели данных, т. е. способа цифрового описания пространственных объектов, зачастую значительно важнее, чем выбор программного обеспечения для его реализации.

Для моделирования данных предметной области нефтедобычи был использован объектный подход, основанный на принципах прикладного системного анализа. Язык UML представляет собой общецелевой язык визуального моделирования, который разработан для спецификации, визуализации, проектирования и документирования компонентов программного обеспечения, бизнес-процессов и других систем. А оптимальной методикой формирования алгоритмов формализации и представления знаний является использование технологии моделирования онтологии, основанное на представлении концептов и зависимостей между ними в формате OWL на базе D-логики. Онтология

определяет общий словарь для ученых, которым нужно совместно использовать информацию в предметной области. Она включает машинно-интерпретируемые формулировки основных понятий предметной области и отношения между ними. Указанная методология используется совместно с объектным подходом к моделированию процессов и данных области нефтедобычи в терминах языка UML. Предлагаемая методика обеспечивает комплексный иерархический подход к процессу разработки алгоритмов онтологического моделирования знаний в указанной предметной области.

Применяемый метод осуществления процесса сбора и интеграции распределенных данных базировался на трех технологиях:

1. Объектные репозитории данных.
2. Механизм Web-сервисов, как средство построения внешних интерфейсов к таким репозиториям.
3. Аппарат рабочих процессов, как средство управления обработкой и интеграции информационных потоков.

Объектно-ориентированный подход позволяет выстроить информационные иерархии, организовать стройную систему методов-сервисов. Методами, работающими с этими информационными объектами, являются сервисы репозитория: ассоциативный поиск по хранилищу, генерирование динамических каталогов, профилирование и т. д. При работе с репозиторием пользователи должны получить эффективные интеллектуальные средства поиска информации на основе семантической индексации, автоматической классификации и каталогизации найденных онтологий с построением семантических связей между ними и автоматического их реферирования с использованием знаний. Построение объектно-ориентированной модели репозитория информационных объектов обеспечит быстрый и эффективный путь к ее реализации.

Для описания объектно-ориентированных моделей данных применяется ряд языков описания объектных схем данных, например:

- ODL – стандарта ODMG объектно-ориентированных БД;
- RDFS (Resource Definition Framework Schema) – W3C стандарт позволяет описывать схемы классов и их свойств, с учетом их наследования, ограничений;
- OWL (Web Ontology Language) – специализация RDFS, ориентированная на описание предметных онтологий.

Процесс сбора и интеграции данных может представлять собой логически сложную композицию обращений к хранилищам информационных

сущностей посредством интерфейсов Web-сервисов. Они представляют собой технологию интеграции приложений, которая может использоваться в Internet. В данном контексте Web-сервис выступает в роли автономного приложения, которое предоставляет средства доступа к информации внешним клиентам через набор предоставляемых им услуг.

В основе методов и средств поиска информации для определения концептов в области недропользования было положено описание системы в терминах сущностей, отношений между ними и преобразование сущностей, которое выполняется в процессе решения определенной задачи. Основной характерной чертой этого подхода является разделение реального мира на составляющие - классы объектов и совокупность фундаментальных свойств этих объектов. Необходимо решить следующие задачи:

- провести анализ понятийного аппарата конкретной предметной области, например, гидрохимические способы разведки нефти в целом или их разделы (газовая съемка, радиоактивная съемка и т.д.) и определить критерии отбора терминов;
- распределить отобранные термины по категориям и построить графические схемы понятий;
- по возможности устранить синонимы и многозначности;
- создать модели полученных концептов, определить их взаимосвязи и отношения;
- сформировать классификатор отобранных терминов, распределив их по отношениям класс – надкласс, класс - подкласс.

При разработке любой онтологии стоит задача внести ясность в толкование терминов заданной предметной области и визуально представить взаимосвязи между этими терминами. Кроме этого ещё необходимо создание семантического элемента, имеющего перспективу многократного использования при поиске информации интеллектуальными системами и системами поддержки принятия решений в заданной предметной области.

При выделении концептов предметной области мы воспользовались следующей последовательностью действий. Во-первых, изучили открытые источники по тематике выполняемых научных работ. Во-вторых, выделили, на наш взгляд, наиболее значимые. К таким источникам мы отнесли фундаментальную литературу, наиболее часто используемую при обучении в данной предметной области. в результате для разработки онтологии по предлагаемой предметной области была создана следующая иерархия классов и подклассов (табл. 1).

Таблица 1  
Иерархия классов онтологии

Уровень иерархии	Название класса (подкласса)
-	Нефтедобыча
--	Бурение скважин
---	Виды бурения
---	Проведение исследования
---	Ремонт скважин
---	Углубление скважин
---	Удаление породы
--	Подготовка нефти
--	Поиск нефти
--	Скважина
---	Исследование скважин
---	Классификация скважин
----	Категории по проекту разработки
----	Категории при заложении
---	Типы скважин
----	Наблюдательные
----	Нагнетательные
----	Оценочные
---	Фонд скважин
----	Неэксплуатационные скважины
----	Законсервированные скважины
----	Ликвидированные скважины
----	Эксплуатационные скважины
---	Способы добычи

В результате получилась иерархия, представленная на рис. 1.

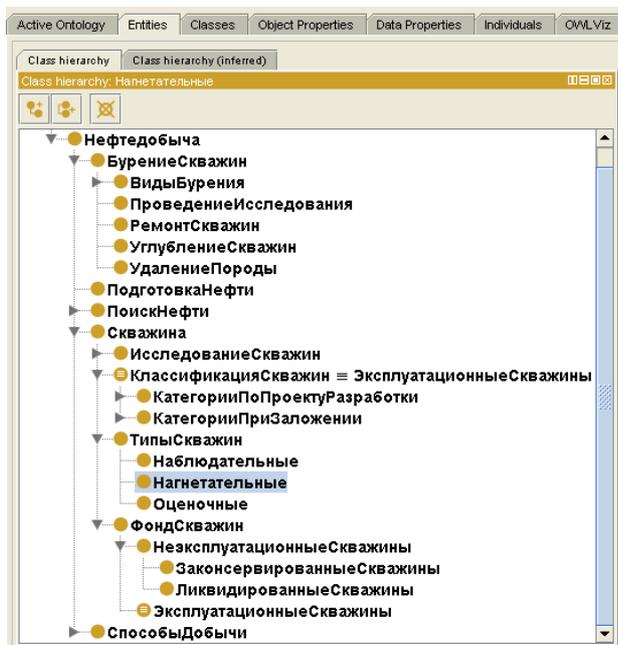


Рис. 1. Таксономия

Построение модели онтологии предметной области нефтедобычи

### 3. Построение онтологии нефтедобычи

Рассмотрим методы онтологического описания конкретной предметной области. Под онтологией можно понимать:

- надежный семантический базис в определении содержания;
- общую логическую теорию, которая состоит из словаря и набора утверждений на некотором языке логики;
- базис для взаимопонимания людей, специалистов в разных предметных областях.

В настоящее время для создания и поддержки онтологий существует целый ряд инструментов, которые помимо общих функций редактирования и просмотра выполняют поддержку документирования онтологий, импорт и экспорт онтологий разных форматов и языков, поддержку графического редактирования, управление библиотеками онтологий и т.д. В данном исследовании был использован редактор Protégé – локальная, свободно распространяемая Java-программа, разработанная в Стенфордском университете.

Разработка онтологий для Protege состоит из 5 шагов:

1. Выделение области (масштаба) онтологии, иначе определение границ онтологии.
2. Определение классов;
3. Организация иерархии классов;
4. Формирование описания классов, подклассов, экземпляров через определение слотов, т. е. свойств;
5. Определение значений.

Разработка онтологии предполагает поэтапную работу над контентом предметной области, включающим термины, понятия и сведения, выявленные в разных источниках. Процедура спецификации определяет цели создания, предполагаемое использование и потенциальных пользователей онтологии. Обозначенные цели позволяют перейти к этапу концептуализации, который обеспечивает структурирование предметных знаний области нефтедобычи. Дальнейшая формализация трансформирует структуру знаний области нефтедобычи, представленную в концептуальном виде, в формальную или вычислительную модель. В процессе реализации онтологии происходит генерация моделей в виде программного кода на языке представления знаний, например, OWL.

Процедура проверки включает идентификацию и редактирование неопределенных терминов; выявление конфликтов и несогласованностей; проверку онтологии на полноту; идентификацию семантических различий между двумя терминами в различных онтологиях; синтаксический анализ;

анализ таксономии; семантическую оценку. Результатом данного этапа является разработанная онтология предметной области.

Для примера рассмотрим формирование контента для способов добычи нефти. Согласно принятой классификации способы нефтедобычи делятся на:

- Фонтан, где выход флюида происходит за счет разности давления в продуктивном пласте и давления на устье скважины;
- Газлифт, являющийся методом добычи нефти, основанным на эффекте газлифта.
- Насосно-компрессорную добычу, способ подъема нефти по скважине за счет энергии сжатого газа, нагнетаемого в скважину с поверхности.

В рамках построенной онтологии эти концепты становятся подклассами первого уровня (рис. 2).

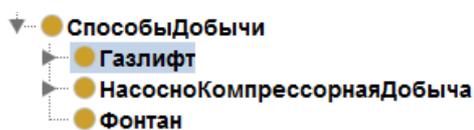


Рис. 2. Подклассы первого уровня «Способы добычи»

Рассмотрим подробно ветку второго уровня Газлифты.

Газлифты подразделяются по:

- Видам источников рабочего агента;
- Конструкциям газлифтных подъемников;
- Схеме действий.

В рамках построенной онтологии эти концепты становятся подклассами третьего уровня (рис. 3).

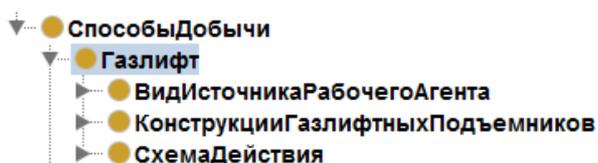


Рис. 3. Подклассы второго уровня «Газлифты»

По схеме подачи от вида источника рабочего агента – газа (воздуха) различают:

- Компрессорный газлифт, где для сжатия попутного газа применяются компрессоры;
- Безкомпрессорный газлифт, где используется газ газового месторождения, находящийся под давлением, или из других источников.

В рамках построенной онтологии эти концепты становятся подклассами четвертого уровня (рис. 4).

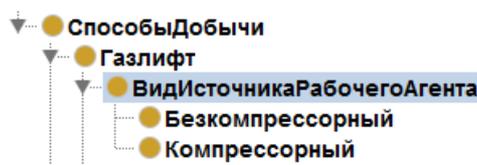


Рис. 4. Подклассы четвертого уровня «Вид источника рабочего агента»

В процессе создания онтологий основная идея состоит в том, что имея описание некоторых понятий, можно давать определения, которые, в конечном итоге и будут задавать наши понятия. Для этого используют аксиомы, которые задают односложные утвердительные высказывания, позволяющие описать тот или иной объект. Чем больше аксиом мы используем для описания понятия, тем «полнее» это понятие будет определено, что в конечном итоге позволяет получать «неявные» знания о рассматриваемом объекте. Аксиомы задаются по правилам дискреционных логик первого порядка.

Существует пять видов задания ограничений:

- Экзистенциальные ограничения описывают класс индивидов имеющих, по крайней мере, одно (some) отношение указанного свойства для индивида, являющегося членом заданного класса. Экзистенциальные ограничения указывают наличие, по крайней мере, одного отношения по данному свойству для индивида, являющегося членом определенного класса (указывается диапазон);
- Универсальные ограничения (only) обозначаются символом  $\forall$ . Они ограничивают отношения по данному свойству для отдельных индивидов, которые являются членами определенного класса;
- Ограничения мощности (exactly) указывает точное количество Ограничения, которые так описывают классы, известны как ограничения мощности;
- Минимальное ограничение мощности (min) определяет минимальное количество отношений, которые индивид может иметь;
- Максимальное ограничение мощности (max) определяет максимальное количество отношений, которые индивид может иметь.

Для данных концептов характерны следующие свойства: видом источника рабочего агента может быть либо безкомпрессорный, либо компрессорный, поэтому введем функциональное свойство «Имеет вид» и опишем с его помощью следующих ограничений (рис. 5).

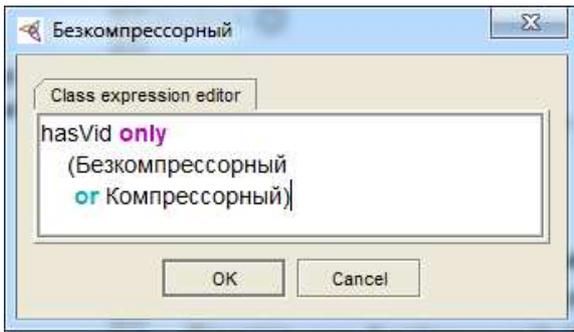


Рис. 5. Ограничение

Остальные свойства были сформированы из свойств вышележащих классов (рис. 6).

По такому принципу рассматриваются все концепты в области добычи нефти.

В виде графа концепты для способа добычи нефти с помощью газлифтового оборудования будут выглядеть следующим образом (рис. 7).

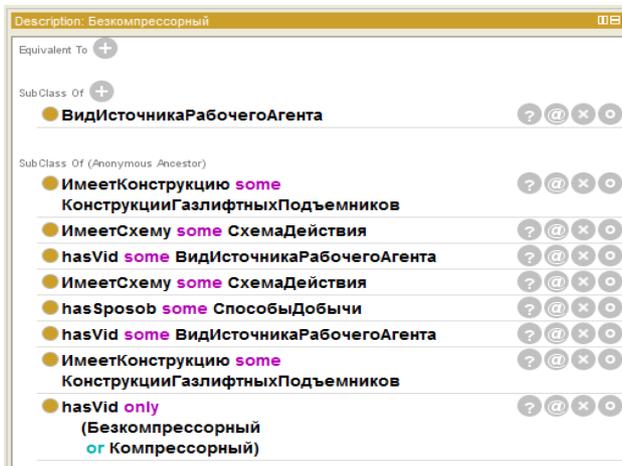


Рис. 6. Остальные ограничения

После формирования контента предметной области была применена технология разработки онтологии, представленная на рисунке 8 [4].

Разработка онтологии предполагает поэтапную работу над контентом предметной области, включающим термины, понятия и сведения, выявленные в разных источниках. Дальнейшая формализация трансформирует структуру знаний области нефтедобычи, представленную в концептуальном виде, в формальную или вычислительную модель. В процессе реализации онтологии происходит генерация моделей в виде программного кода на языке представления знаний, например, OWL.

Процедура проверки включает идентификацию и редактирование неопределенных терминов; выявление конфликтов и несогласованностей; проверку онтологии на полноту; идентификацию семантических различий между двумя терминами в различных онтологиях; синтаксический анализ; анализ таксономии; семантическую оценку. Результатом данного этапа является разработанная онтология предметной области [5].

#### 4. Заключение

Сформированная онтология объектов нефтедобычи позволяет средствами логических запросов получать новые знания в указанной области. Результатом могут быть оформлены в виде новых объектов как следствие визуального анализа структуры предметной области, либо как результаты выполнения семантических запросов.

Также были решены инженерные, производственные и образовательные задачи в результате сетевой каталогизации и обеспечение доступа к понятийным ресурсам пространственных данных в области недропользования на примере нефтедобычи.

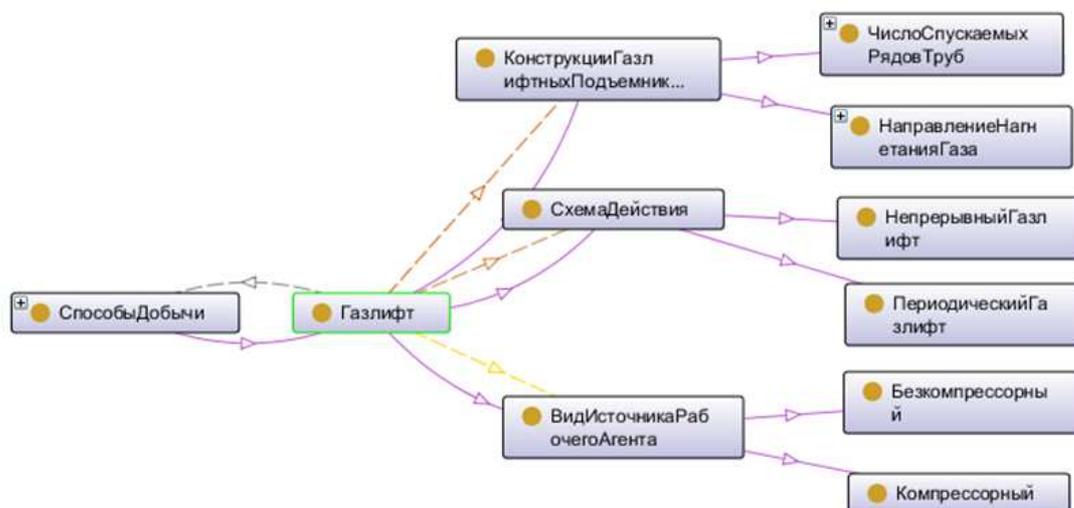


Рис. 7. Онтологический граф для способа добычи нефти с помощью газлифтового оборудования

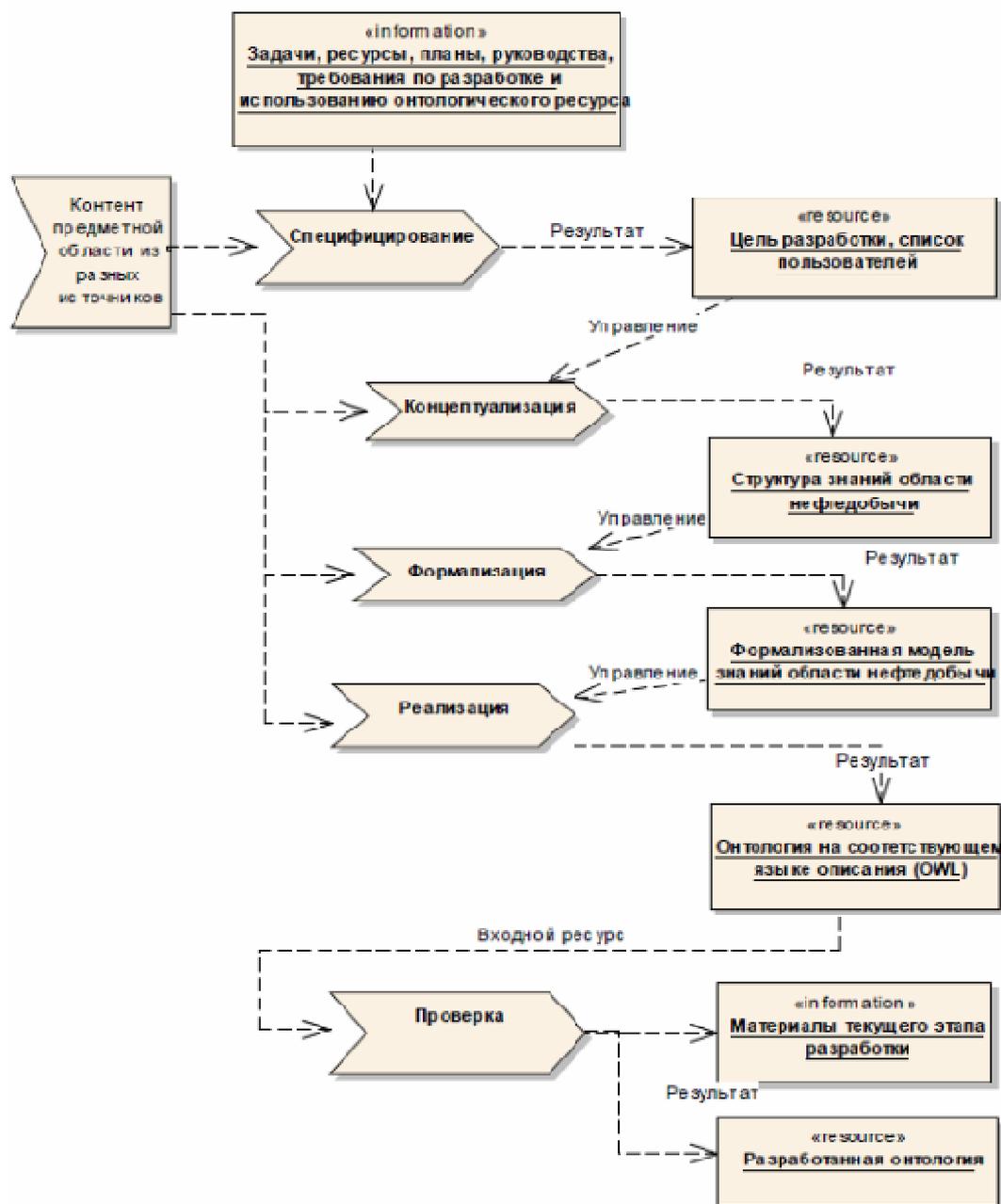


Рис. 8. Основные процедуры процесса «Разработка онтологии»

### Список используемых источников

1. Лапшин В.А. Онтология в компьютерных системах. — М.: Научный мир, 2010. — 224с.
2. Найханова Л.В. Технология создания методов автоматического построения онтологий с применением генетического и автоматного программирования: монография. — Улан-Удэ: Изд-во БНЦ СО РАН, 2008. — 244с.
3. Онтологии и тезаурусы: модели, инструменты, приложения / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, Н.В. Соловьев. — М.: БИНОМ, 2009. — 173с.
4. Дидык Т.Г., Рыков В.И., Шаронова Ю.В. Алгоритмы и средства формирования онтологии заданной предметной области // Современные проблемы науки и образования. 2013. № 6. С. 64.
5. Мартынов В.В., Филосова Е.И., Фандрова Л.П., Шаронова Ю.В., Дидык Т.Г. // Использование операций и технологий управления онтологиями в предметной области «Нефтедобыча». Proceedings of the 19th International Workshop on Computer Science and Information Technologies. Germany, Baden-Baden, October 08 - 10, 2017. Volume 1. Ufa State Aviation Technical University, 2017. P. 77-83.