УДК 004.041+004.045

В.А. СТАРЫХ
V.A. STARYKH

ОТКРЫТАЯ ТЕРРИТОРИАЛЬНО-РАСПРЕДЕЛЁННАЯ СИСТЕМА УПРАВЛЕНИЯ ИНФОРМАЦИОННЫМИ РЕСУРСАМИ. THE OPEN TERRITORIALLY-DISTRIBUTED CONTROL SYSTEM OF INFORMATION RESOURCES.

В статье рассматривается территориально-распределённая модель управления системой распределенного доступа, электронной публикации и хранения информационных ресурсов в форме мегабиблиотеки со встроенной системой поиска федеративного уровня. Информационный ресурс хранится в локальном репозитории, как глобально и уникально именуемый набор структурированных данных, содержащих сведения о ресурсе, его свойствах, атрибутах и связях. Репозитории взаимодействуют между собой и с компонентами вышележащего тематического уровня, содержащего один компонент – реестр сред, выполняющего функции систематизации репозиториев в соответствии с тематическим классификатором. Высокий уровень абстракции позволяет концептуально объединить информацию из разных систем, с учётом разнообразия используемых ими моделей более низкого уровня, описывая их в единой терминологии атрибутированных ресурсов и связей между ними.

Ключевые слова: информационный ресурс, модель, управление, реестр, репозиторий, хранилище, xml-документ, метаданные, спецификации.

In article the territorially-distributed model of management by the system of distributed access, the electronic publication and storage of information resources in the form of megalibrary with the built in system of search of federal level is considered. The information resource is stored in a local repository as it is global and is unique a called set of the structured data containing data on a resource, its properties, attributes and communications. Repositories co-operate among themselves and with components of the overlying thematic level containing one component – the register of the environments, carrying out functions of ordering of repositories according to the thematic qualifier. High level of abstraction allows to unite conceptually the information from different systems, with the account of a variety of models of lower level used by them, describing them in uniform terminology.

Keywords: an information resource, model, management, the register, a repository, storehouse, the xml-document, metadata, specifications.

В настоящее время существует множество образовательных, научных и других специализированных учреждений (библиотеки, информационные центры) имеющих потребность в оптимальном управлении собственными образовательными, научными и другими электронными информационными ресурсами (ЭИР), а также во взаимодействии с другими владельцами информационных ресурсов.

Тем не менее, типовые бизнес-процессы управления электронными информационными ресурсами не доведены до уровня формального описания в стандартной нотации и опубликования для открытого обсуждения. Существует реальная потребность в формировании технологической инфраструктуры управления ЭИР сферы образования и науки путем создания условий для создания и апробации ее ключевых технологий представляемых в открытых технических спецификациях и реализуемых в типовых инструментальных средствах.

Традиционно, проблема управления информационными ресурсами решается на базе подхода основанного на централизации сервисов хранения и обработки ЭИР. Наряду с достоинствами данного подхода, заключающимися в простоте реализации и обслуживания системы, организации единой точки доступа ко всем ресурсам, существует ряд серьезных недостатков, сдерживающих развитие систем с такой архитектурой. Основные из них - это необходимость в дорогостоящем аппаратно-программном обеспечении для организации обработки огромных объемов данных и поступающих запросов пользователей, жесткие управленческие схемы администрирования, отсутствие мотивации владельца передавать

свои ЭИР в централизованные хранилища и др. Актуальное значение при разработке подобных систем приобретает идея децентрализации управления ЭИР.

При реализации предлагаемого подхода владелец электронного информационного ресурса избавлен от необходимости перемещения информации в какое-либо централизованное хранилище. Информационный ресурс размещается там, где он впервые был опубликован автором, при условии, что ему предлагается web-инструментарий с простым дружественным пользовательским интерфейсом, позволяющий создавать стандартные метаописания ЭИР[1], ориентированные на общепринятые специфицированные модели. Идея децентрализации заключается в том, что распределенная система будет заинтересована как в простой поставке новых информационных ресурсов так и в долгосрочном сотрудничестве с профессиональными сообществами, локально осуществляющими самостоятельное управление собственным подмножеством ЭИР. Кроме того, при распределенном подходе автоматически решается проблема обеспечения авторских прав на контент, поскольку ЭИР находится в системе владельца.

Распределенный подход к обработке информации, находящейся в различных местах, в локальных хранилищах, или наследуемых СУБД является основой предложенной программной архитектуры. При таком подходе осуществляется непосредственная обработка и консолидированный федеративный поиск ресурсов в распределенном хранилище, состоящем из доступных через Интернет, территориально-распределённых ресурсных локальных репозиториев, взаимодействующих между собой на основе открытых стандартов. Необходимо принять во внимание, что при данном подходе не предлагается полный отказ от реализации идеи централизованного управления ЭИР, так как для обеспечения взаимодействия компонентов территориально-распределенной системы (локальных репозиториев ЭИР) необходимо наличие управляющего звена, обеспечивающего структуризацию множества репозиториев, распределенный поиск, сбор статистики.

Модель территориально-распределенной системы управления информационными ресурсами, представленная на рисунке 1, состоит из двух уровней:

1. Содержательный уровень.
2. Тематический уровень.

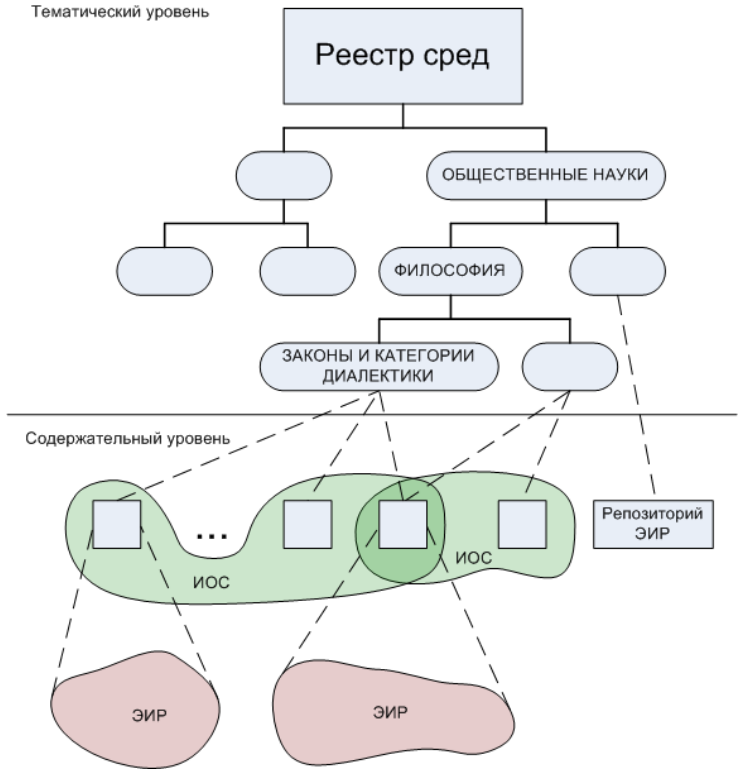


Рисунок 1- Структура системы управления информационными ресурсами

Содержательный уровень содержит множество локальных репозиториев, выполняющих функции управления метаописаниями ЭИР. Репозитории на данном уровне взаимодействуют между собой и с компонентами тематического уровня.

Тематический уровень содержит один компонент – реестр сред, выполняющий функции систематизации репозиториев в соответствии с тематическим классификатором. Основная цель реестра сред состоит в обеспечении компонентов содержательного уровня (репозиториев) сведениями о репозиториях, принадлежащих выбранному тематическому направлению для организации распределенного поиска. Такой подход призван обеспечить наиболее полную релевантность запросов пользователей в выбранной области.

Тематическое направление определяется элементами классификатора. На основе данного направления происходит формирование областей для навигационных механизмов информационно-образовательных сред (ИОС) [2]. Среда служит областью взаимодействия всех заинтересованных участников в рамках формируемого таким образом сообщества. ИОС призвана обеспечить:

- качество информационных ресурсов за счет взаимодействия всех участников процессов создания, распространения и применения ресурса;
- доступность информационных ресурсов за счет предоставления возможностей оперативного поиска, выбора, получения ресурса, соответствующего конкретным потребностям и условиям применения;
- использование информационных ресурсов для осуществления предметно-ориентированной деятельности.

1 Организация распределенного поиска ЭИР

Одной из основных функций системы управления информационными ресурсами сферы образования является обеспечение эффективного поиска ЭИР. В условиях распределенной архитектуры самым простым вариантом является поиск среди всех репозиториев (плоская модель поиска). Такой подход может быть эффективным при сравнительно небольшом количестве репозиториев. В противном случае возникает ситуация, при которой поиск производится в репозиториях, не соответствующих запросу пользователя. В результате релевантность найденной информации резко снижается. В дополнение к этому возникают проблемы производительности при «широковещательном» поисковом запросе.

Решение проблемы может заключаться в ограничении множества репозиториев, в которых производится поиск информации. Наиболее логичным вариантом выделения областей поиска является тематическая фильтрация, которая предполагает, что запрос пользователя снабжается метками принадлежности к той или иной предметной области. Реестр сред, как компонент, обладающий полной информацией о привязке репозиториев к тематическому классификатору, осуществляет выделение тематических областей и предоставляет эту информацию компоненту, выполняющему непосредственный запрос на поиск ЭИР. При этом сохраняется возможность гибкого поиска по репозиториям путем задания принадлежности к любому уровню классификатора реестра сред, что позволяет искать сразу в нескольких тематических областях.

2 Реестр сред

Реестр сред, приведённый на рисунке 2, представляет собой web-приложение, обеспечивающее единую точку доступа к интересующему пользователя информационному образовательному сообществу, в соответствии с выбранной тематикой.

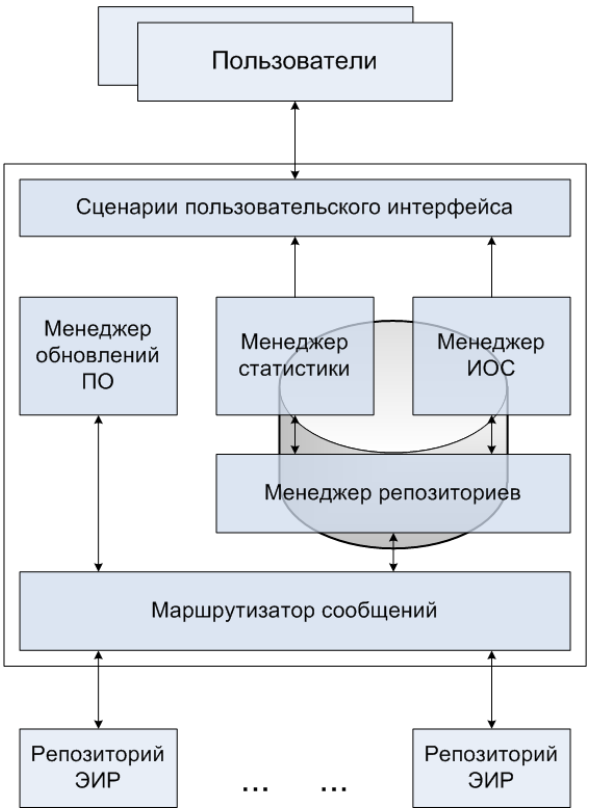


Рисунок 2 - Структура реестра сред

Реестр сред выполняет следующие функции:

1. Поддержка иерархического тематического классификатора
2. Ведение реестра репозиториев с привязкой их к узлам тематического классификатора (формирование областей для навигационных механизмов информационных образовательных сред)
3. Поддержка механизма статистики. Сбор, хранение и представление информации по количеству репозиториев в рамках тематики, а также зарегистрированных в них ЭИР, посещаемости и т. д.
4. Взаимодействие с репозиториями на основе механизма сообщений.
5. Поддержка контроля версий ПО зарегистрированных репозиториев.
6. Предоставление интерфейса пользователя с возможностью поиска информации по распределенному хранилищу ЭИР.

В состав реестра сред входят следующие компоненты:

1. База данных реестра. Содержит следующие информационные объекты:
   - Тематический классификатор, представляющий собой иерархический справочник.
   - Список репозиториев с привязкой к тематическому классификатору.
   - Статистическая витрина.
2. Менеджер ИОС. Управляет списком информационных образовательных сообществ в соответствии с тематическим классификатором. В его функции входят добавление, удаление репозиториев в состав ИОС.
3. Менеджер репозиториев. Управляет общим списком репозиториев, зарегистрированных в системе. В его функции входят добавление, удаление, изменение регистрационных сведений репозитория, а также актуализация общесистемных классификаторов в нем.
4. Менеджер статистики. Управляет сбором и обработкой статистических сведений по различным показателям. Информация собирается со всех зарегистрированных репозиториев.

5. Менеджер обновлений ПО. Осуществляет поддержку версий ПО репозиториев. В его функции входит хранение и предоставление информации по текущей версии ПО репозитория.
6. Маршрутизатор сообщений. Перенаправляет запросы репозиториев к соответствующим менеджерам и подсистемам.
7. Сценарии пользовательского интерфейса. Предоставляют пользователю web - интерфейс для взаимодействия с реестром сред.

3 Репозиторий ЭИР

Репозиторий ЭИР представляет собой web-приложение, предоставляющее пользователю возможность создания метаописания ЭИР и его публикации, поиска в локальном и распределенном хранилище метаописаний. Компонентная структура репозитория ЭИР представлена на рисунке 3.

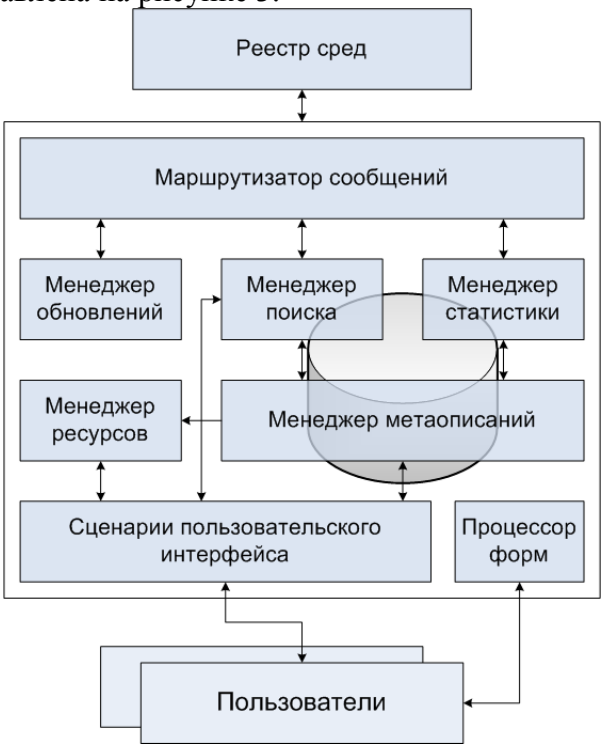


Рисунок 3 - Структура репозитория ЭИР

Репозиторий ЭИР выполняет следующие функции:

1. Регистрация создателя метаданных ЭИР.
2. Поддержание групп создателей метаописаний ЭИР, принадлежащих к структурным единицам организации.
3. Поддержание профиля пользователя, отражающего уровень детализации метаописаний ЭИР.
4. Создание, редактирование, удаление метаописаний ЭИР.
5. Публикация ЭИР в рамках репозитория с возможностью открытого доступа к нему.
6. Навигация по хранилищу метаописаний ЭИР с возможностью задания группировки по различным атрибутам метаданных.
7. Многокритериальный поиск по распределенному хранилищу ЭИР.
8. Регистрация репозитория в реестре сред.

В состав репозитория ЭИР входят следующие компоненты:

1. База данных метаописаний ЭИР. Представляет собой хранилище XML-документов – метаописаний ЭИР. Содержит следующие информационные объекты:
    - Тематический классификатор (копия тематического классификатора реестра сред)

- Каталог метаописаний ЭИР
- Системный каталог

2. Менеджер обновлений. Обеспечивает периодическую проверку появления новых версий ПО репозитория и тематического классификатора в реестре.
3. Менеджер поиска. При организации поиска по распределенному хранилищу в его функции входят: получение из реестра сред списка репозиториев, соответствующих тематике запроса пользователя; рассылку поисковых запросов по полученному множеству репозиториев; группировка результатов поиска по временному признаку. В режиме локального использования обеспечивает: поиск по локальному хранилищу метаданных ЭИР в соответствии с запросом пользователя, группировку метаданных в по заданным критериям.
4. Менеджер статистики. Обеспечивает периодический сбор следующих статистических сведений: количество поисковых запросов к репозиторию за период времени, количественный состав документов в хранилище репозитория с учетом набора фиксированных критериев.
5. Менеджер ресурсов. Обеспечивает публикацию ЭИР пользователя в рамках домена репозитория, с возможностью описания структуры модульного ресурса.
6. Менеджер метаописаний ЭИР. Обеспечивает управление метаописаниями ЭИР. Реализует операции чтения, сохранения и удаления документов в хранилище метаописаний репозитория.
7. Сценарии пользовательского интерфейса. Предоставляют пользователю web - интерфейс для взаимодействия с репозиторием.
8. Процессор форм. Предоставляет web-интерфейс с возможность заполнения реквизитов формы метаописания ЭИР в соответствии с профилем пользователя. Реализует операции создания и изменения документа метаописания.
9. Маршрутизатор сообщений. Перенаправляет запросы компонентов репозитория к реестру сред и другим репозиториям.

4 Организация взаимодействия компонентов системы

Компоненты данной информационной системы взаимодействует друг с другом посредством реализации механизма обмена сообщениями. Данный механизм включает в себя взаимодействие репозитория с реестром сред, а также с другими репозиториями. Сообщения представляют собой XML-документы определенного формата.

В системе определены следующие сообщения:

- Запрос на регистрацию репозитория в реестре сред.
- Запрос на удаление регистрационной информации о репозитории из реестра сред.
- Запрос на получение списка репозиториев заданной тематики.
- Поисковый запрос к репозиториям на получение ссылок на метаописания ЭИР, которые соответствуют заданным критериям пользователя.
- Запрос на получение информации об актуальной версии ПО репозитория.
- Запрос на обновление тематического классификатора с целью его синхронизации с тематическим классификатором реестра сред.
- Запрос на обновление статистики репозитория к реестру сред.

5 Реализация реестра сред

В настоящей версии программной реализации реестра сред поддерживаются такие наиболее часто используемые классификаторы, как ГРНТИ и УДК. Тем не менее, реестр сред не ограничен применением двух вышеперечисленных классификаторов. Возможно использование любого иерархического классификатора. Однако, использование ГРНТИ представляется наиболее целесообразным, так как ГРНТИ, в отличие от УДК, содержит три

уровня вложенности, что позволяет организовать несложный пользовательский интерфейс для построения запроса на поиск ЭИР.

Структура информационной модели базы данных реестра сред представлена на рисунках 4а и 4б.

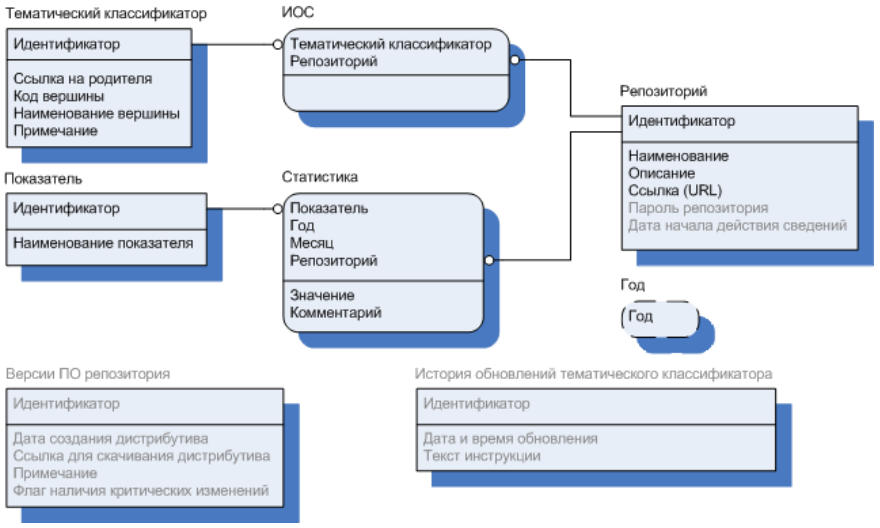


Рисунок 4а - Логическая модель базы данных реестра сред

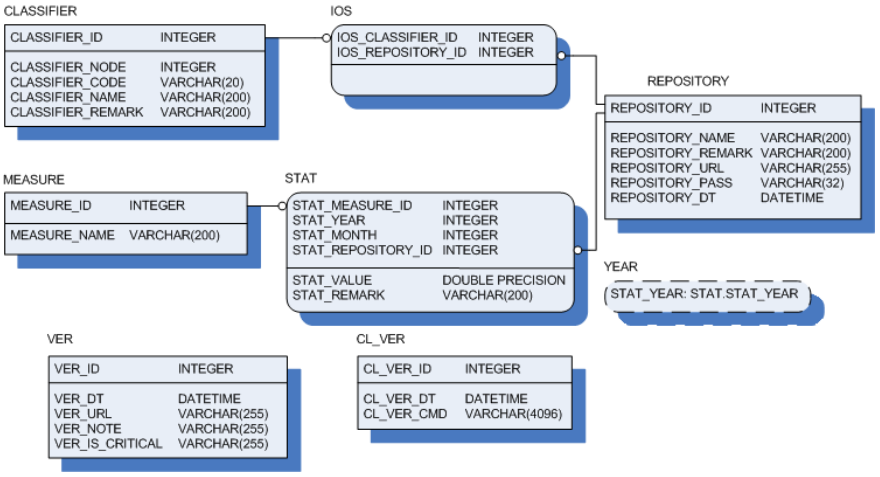


Рисунок 4б - Физическая модель базы данных реестра сред

6 Реализация репозитория ЭИР в составе участников виртуального сообщества.

Так как метаописание ЭИР представляется в виде XML-документа, возникает проблема выбора механизма для хранения и обеспечения эффективного доступа к коллекциям документов. В настоящее время имеется ряд коммерческих и свободно распространяемых проектов и реализаций XML-Хранилищ, ориентированных на спецификацию XML DB:API. Целью этой спецификации является разработка общего для всех стандарта API для доступа к XML базам данных. Этот API нейтрален по отношению к любому множеству баз данных и языковым средствам. Выбор СУБД с полной поддержкой данной спецификации, таким образом гарантирует, что разработанные средства могут быть впоследствии легко перенесены на любые технологические платформы.

В соответствии со спецификацией XML DB:API XML-хранилище представляет собой набор коллекций документов, образующих иерархическую структуру. На рисунке 5 представлена структура коллекции репозитория.

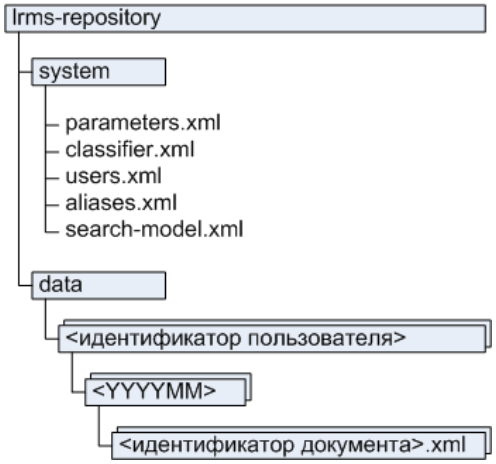


Рисунок 5 - Схема базы документов репозитория

Прототип информационной системы, разработанной на основе данного подхода, базируется на использовании свободно распространяемого ПО и может функционировать на любой программно-аппаратной платформе. Система легко настраивается и конфигурируется, и сразу предоставляет как сервисы для публикации информационного ресурса (книги, статьи, учебники, монографии, учебно-методическое обеспечение и т.д), так и средства для создания библиографических XML-карточек, основанных на спецификации базовой информационно модели RUS-LOM [3].

ЛИТЕРАТУРА

1. Башмаков А.И., Старых В.А., Принципы построения и описания профилей стандартов и спецификаций информационно-образовательных сред. Метаданные для информационно-образовательных ресурсов сферы образования. Серия: Нормативно-техническое обеспечение информационных технологий в образовании. Выпуск 1.. Научное издание. – М.:ФГУ ГНИИ ИТТ «ИНФОРМИКА», «Европейский центр по качеству», 2009. – 376 с., ил.
2. Башмаков А.И., Старых В.А. Принципы и технологические основы создания открытых информационно-образовательных сред / ФГУ ГНИИ ИТТ "Информика". – М.: "БИНОМ. Лаборатория знаний", 2010. – 719 с. – ISBN 978-5-9963-0285-7.
3.http://spec.edu.ru/sights/spec.nsf/(All)/A397FD2AF61045BCC32572B40040A706/$File/Спецификация%20информационой%20модели.pdf

Старых Владимир Александрович
ФГУ ГНИИ ИТТ «Информика» г. Москва
Кандидат технических наук, доцент, заместитель директора
Тел. (499) 155-87-28
E-mail: vstar@informika.ru