

Предлагается способ организации обмена электронными документами в высшем учебном заведении и рассматривается создание электронной системы свободного обмена информацией. В ходе работы проанализированы возможные способы решения данной задачи, заданы критерии оценки результата, разработана архитектура системы и анализируются результаты внедрения.

© С.Н. Горбач, Л.А. Макаревич,
К.Ю. Талецкий, В.А. Петрухин,
В.И. Шоломова, 2013

УДК 004.4

С.Н. ГОРБАЧ, Л.А. МАКАРЕВИЧ, К.Ю.
ТАЛЕЦКИЙ,
В.А. ПЕТРУХИН, В.И. ШОЛОМОВА

ОРГАНИЗАЦИЯ ОБМЕНА ЭЛЕКТРОННЫМИ ДОКУМЕНТАМИ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

Введение. За последние двадцать лет способы обмена информацией радикально изменились. Мало кто мог подумать, что на смену многотомным энциклопедиям придут сайты, а их авторами будут уже не сотни человек, а миллионы. Есть мнение, что само существование печатных изданий под угрозой. Газеты и журналы со столетней историей борются за влияние и аудиторию с миллионами он-лайн-дневников, чьи авторы излагают и оценивают любое событие по-своему, создавая доселе невообразимую широту взглядов. Телевидение начинает проигрывать свои позиции сайтам, на которых ежедневно выкладываются сотни тысяч коротких роликов. Файлообменные сети дают возможность за считанные минуты передать в любую точку мира любимый фильм или музыкальное произведение.

Очевидно, что те проекты, которые основывались на принципах общедоступности и открытости, пользуются огромной популярностью. Примерами таких поразительно успешных социальных проектов могут служить Википедия или торренты. Опен-сорс сообщество – ярчайшая демонстрация того, как сложнейшие системы могут быть созданы людьми, объединившимися для решения общей проблемы, а не

для извлечения быстрой | ресурсов приходит на смену персональным
выгоды. Совместное | приложениям и небольшим коллективам
использование и развитие | авторов
и разработчиков. Образовательный процесс с этой точки зрения не исключение. Помогая учащимся и преподавателям интенсифицировать обмен информацией, можно повысить скорость и качество обучения, а также удовлетворенность учеников самим процессом.

Настоящая работа посвящена попытке предоставления студентам простого, а, главное, естественного способа обмена и совместного пользования документами.

Созданная система способна существовать как отдельное решение на коммерческой основе, так и быть внедренной как часть инфраструктуры образовательного учреждения, органично вписываясь в его процессы и учитывая интересы руководства и студентов. Благодаря такому подходу система обладает большой гибкостью в распространении.

1. Свободный обмен информацией в образовании

1.1. Критерии реализации Свободного хранилища документов

Основная цель Свободного хранилища документов – предоставить студентам и, возможно, преподавателям простой и удобный способ обмена, хранения и структуризации любых документов. Это позволит значительно стимулировать образовательный процесс за счет значительного ускорения обмена знаниями между студентами и консолидации самых актуальных учебных материалов, отбираемых самими студентами.

Перед построением любой системы необходимо учитывать опыт разработок, решающих схожие задачи, схожих по основным структурным решениям, направленных на схожую целевую аудиторию. Одним из ключевых моментов предварительного анализа является выработка критериев удобства использования («юзабилити») и определение наиболее важных свойств системы с точки зрения пользователя.

С наиболее общей точки зрения нашу Систему можно охарактеризовать как свободное хранилище документов, ориентированное на активное использование в учебном процессе для хранения и обмена информацией между его участниками и формируемое самими участниками.

По общей концепции такая система близка к современным социальным ресурсам вроде Википедии или YouTube. На примере этих систем можно попытаться проследить общие законы таких ресурсов, без которых, возможно, они бы не стали сверхпопулярными.

Одним из основных факторов, повышающих интерес аудитории к этим ресурсам, является сама аудитория. Она постоянно генерирует содержимое. Механизмы саморегуляции сообщества обеспечивают естественный отбор самой интересной информации и самых популярных авторов. Это привлекает новых посетителей и создает отличную саморекламу.

Для первоначального набора необходимой «критической массы» аудитории всегда принимаются дополнительные меры, но все они малоэффективны, если процесс создания нового содержимого не упрощен до предела, а доступ к информации производится без каких-либо ограничений.

Хотелось бы выделить общие черты этих систем, которые в значительной мере обусловили их популярность.

1. Свободный просмотр любого содержимого.
2. Свободное добавление любого содержимого.
3. Все время пополняющийся и актуальный контент.
4. Простота просмотра и добавления новых материалов.
5. Покрытие спроса пользователей в широкой области.
6. Начальная «критическая масса» содержимого.

На основе данных критериев можно сформулировать общие требования к Свободному хранилищу документов для образовательного процесса. Они повторяют шесть перечисленных пунктов.

Пп. 1 и 2 могут быть удовлетворены за счет создания простых удобных интерфейсов и выбора бизнес-модели, не подразумевающей непосредственной монетизации контента. П. 5 – это требование легко удовлетворить, предоставив пользователям самим управлять содержимым и поощрять добавление информации, интересной сообществу.

П. 6, но хронологически самый первый критерий, требует нахождения способа быстрого набора «критической массы» документов, после чего ресурс начнет развиваться очень активно уже сам по себе. Наиболее быстрый и малозатратный способ это сделать – найти место, где уже пересекаются основные информационные потоки, откуда пользователи смогут наиболее просто и удобно добавлять свой контент в Хранилище.

Если от пользователя будет требоваться самое минимально возможное количество действий, чтобы добавить свой документ в Хранилище, он иногда будет этой функцией пользоваться. Чем больший поток данных и количество пользователей удастся охватить этим процессом, тем быстрее будет происходить это «тунеллирование» документов в Хранилище через «потенциальный барьер» лени пользователей.

При реализации пп. 3, 4 также возникают сложности, требующие анализа потоков информации в образовательном процессе, предпочтений пользователей и специфики задачи.

1.2. Информационные потоки в образовании

В процессе обучения студенты постоянно используют информацию из книг, лекций или семинаров и обмениваются ею друг с другом. Причем многие из них считают, что обмен знаниями друг с другом не только помогает в освоении нового материала, но и является отличным способом структурирования знаний и необходим для глубокого понимания.

На примере образовательного процесса в Московском физико-техническом институте (МФТИ) выделим основные источники знаний, используемые студентами:

- конспекты лекционных курсов и семинарских занятий;
- книжные библиотеки;
- электронные книги и пособия;
- самостоятельные письменные работы.

Все эти разнородные источники информации встречаются в одном месте – при переводе бумажных источников в электронный вид, а электронных – в бумажный.

За последние пять лет распространенность персональных компьютеров в студенческом городке МФТИ выросла в несколько раз: от одного компьютера на 4 – 5 человек до одного-двух ПК на одного человека. Интенсивность использования любых электронных источников информации выросла многократно с ростом доступности различных Интернет-изданий и онлайн-библиотек.

С увеличением интенсивности использования электронных документов необходимость в переводе из бумажного в электронный вид и наоборот только возрастает – бумажные документы значительно удобнее читать, некоторые документы создаются именно в бумажном виде, при сдаче лабораторных, курсовых и дипломных работ используются печатные варианты.

Так мы можем концептуально обозначить «грань» между «мирами» реальных и виртуальных документов. Студенты постоянно переводят нужные им документы через эту «грань». Поэтому можно утверждать, что эти документы:

- непосредственно участвуют в образовательном процессе;
- заведомо интересны учащимся;
- постоянно обновляются.

Впрочем их поток со временем только увеличивается. А если оптимизировать и процесс перевода через «грань», это только еще больше стимулирует процессы перехода.

Становится понятным, что, возможно, самое удобное место, где документы смогут «туннелировать» в Свободное хранилище и есть эта «граница» между электронными и бумажными документами. Благодаря интенсивному потоку документов даже при малой начальной вероятности «туннелирования» документа, Хранилище начнет пополняться и через некоторое время выйдет на режим самоподдержки и перейдет в стадию экспоненциального роста.

1.3. Основная концепция построения Свободного хранилища документов

Основываясь на приведенной модели первоначального пополнения Хранилища, разработана система, логически состоящая из двух компонент.

- **Распределенная система общественной печати** – та самая «грань», через которую пользователи будут централизованно переводить свои документы.
- **Свободное хранилище документов** – объединено с системой печати, имеет с ней общий пользовательский интерфейс и реализует принцип минимальности действий при добавлении и просмотре содержимого.

Такая архитектура системы решает ряд внедренческих и экономических проблем.

Во-первых, система печати сама по себе оказывается очень востребованной, что уже делает всю систему привлекательной как для обучающихся, так и для руководства образовательных учреждений.

Во-вторых, система может быть легко коммерциализована за счет платного оказания услуг печати. Это легко принимают пользователи, а система становится менее зависимой от конкретных условий в образовательном учреждении, что делает ее привлекательной с точки зрения инвестиций и позволяет ей быть как минимум самоокупающейся.

Рассмотрим архитектуру и основные концепции, инструменты, программные и организационные решения, примененные в распределенной системе общественной печати и свободном хранилище документов, которые функционируют уже третий год на территории студенческого городка МФТИ.

2. Архитектура

Распределенная система общественной печати – это информационная система объединяющая программно аппаратный комплекс (рис. 1) общедоступных принтеров и систему обработки, хранения и анализа информации. С его помощью пользователи могут печатать документы, находить интересующую для себя информацию, максимально простым и удобным образом делиться своими документами.

Для того чтобы система, основной акцент которой делается на пользователей, была эффективна, должна быть в первую очередь удобна в использовании. Именно критерии удобства определили архитектуру системы в том виде, в котором она есть сейчас.

Логически система состоит из трех типов компонентов.

- Веб-сервер обеспечивает работу пользователей с сайтом системы.
- Сервер БД содержит описание документов системы и пользовательские настройки.
- Обслуживающие сервера служат для анализа, хранения и отправки их на печать документов.

Сервера объединены между собой в виртуальную сеть с применением технологии VPN. В качестве VPN-сервера и клиентов используется OpenVPN [3].

2.1. Веб-сайт

С точки зрения пользователя центральный узел системы – сайт. Такое архитектурное решение является не случайным. Именно применение Веб-технологий позволяет решить множество проблем и сделать систему существенно более простой и удобной в использовании. Пользователь не должен устанавливать дополнительное программное обеспечение. Веб-браузер имеет любая операционная система. Отсутствие какого бы то ни было дополнительного ПО и требований к клиентскому компьютеру, обеспечивает Системе платформенную независимость.

Основные направления использования сайта – поиск и печать на принтерах системы разнообразных учебных материалов, документов, статей. Процедура

добавления нового документа также является ключевой и не требует никаких сложных дополнительных процедур.

Основной Веб-сервер системы – Apache2, для генерации динамического содержимого и обработки запросов применяется PHP версия 5.3 в виде mod_php. Веб-сервер настроен с учетом характера возникающих нагрузок.

Структура системы

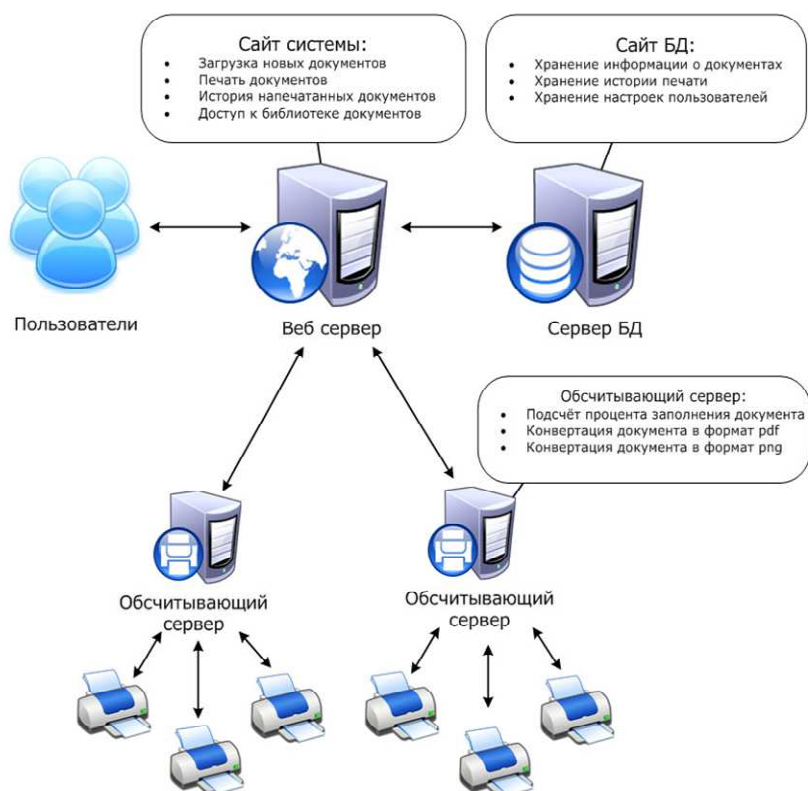


РИС. 1. Схема организации компонентов распределенной системы печати

2.2. Обсчитывающие сервера системы

Задача обработки документов – одна из ключевых для последующего анализа и организации совместного использования. Кроме того, при отправке исходные документы могут иметь достаточно большие размеры, что значительно затрудняет их передачу через Интернет. Для преодоления этих трудностей используются обсчитывающие сервера, располагающиеся в непо-

средственной близости как к принтерам, так и к пользователям. Основные функции обчитывающих серверов – анализ документов, конвертирование их в необходимые форматы, хранение, а также передача информации о них на центральный сервер.

Обчитывающие сервера также предоставляют доступ пользователям к Веб-сайту системы, выступая в роли прокси-серверов. В качестве прокси используется `mod_proxy` Веб-сервера Apache2 с установленным `mod_rewrite` для обеспечения HRU [2]. Связь с Веб-сервером и сервером БД осуществляется через VPN.

Отправка документов возможна либо с помощью сайта системы, либо с помощью драйвера принтера посредством сетевого протокола `ipp`. Загрузка документов вторым способом предпочтительней, так как в этом случае снимается ограничение на изначальный формат документа, и напечатан может быть любой документ, который пользователь может открыть у себя.

При отправке документов через сайт существует значительное количество технических сложностей, связанных с обработкой документов. Для того чтобы аналитический сервер имел возможность правильно совершить процедуры конверсий ему необходимо точно знать с каким типом документа он работает. Анализ типа документа и конверсия документов из разнообразных форматов, является задачей, которая не решается в общем случае. Отправка документов посредством драйвера принтера не имеет таких ограничений.

Система осуществляет конвертирование в документы трех форматов:

PNG – растровый графический формат, использующий сжатие без потерь [3]. Данный формат необходим системе для генерации собственного предпросмотра, для отображения в Интернет-браузере, оценки процента заполнения для учета тонера при печати заказа;

PDF – кроссплатформенный формат электронных документов, созданный фирмой Adobe Software [4]. Предназначен, в первую очередь, для электронного представления полиграфической продукции. В нашем случае используется как основной формат документов в библиотеке;

PCL – язык управления принтером [5]. Инструкции данного языка позволяют управлять режимами печати принтера, задавать различные опциональные возможности, получать статусные сообщения от принтера.

Указанные преобразования осуществляются из исходного формата **Postscript** [6]. Это высокоуровневый аппаратно-независимый язык описания страниц, применяется как язык разметки полиграфической продукции. Он позволяет получать от пользователей документы в легко интерпретируемом формате вне зависимости от вида ОС, ее версии и прикладного ПО, которое сгенерировало заказ на печать. Делает систему независимой от моделей принтеров и их возможностей.

Для приема заказов по протоколу IPP [7] используется CUPS-сервер [8]. Протокол IPP был выбран как замена протоколу SMB/CIFS [9] из-за плохой переносимости последнего и невозможности обеспечить полную поддержку с помощью свободного ПО. IPP обеспечивает возможность передачи заданий печати через любые сети, поскольку является надстройкой над HTTP/1.1 [10], имеет отличную поддержку в CUPS во всех ОС, где он используется, и в ОС

семейства Windows в числе стандартных, включенных по умолчанию компонент, кроме Windows Vista/7, где его требуется включать [11].

Для управления подключенными принтерами, обеспечения обратной связи с принтерами и передачи заданий печати, используются разработанные нами backend'ы для CUPS.

3. Результаты

В качестве экспериментального внедрения для определения востребованности подобной системы выбран студенческий городок МФТИ (Московского физико-технического института).

Испытание системы [12] естественным образом разбито на два этапа. Сначала, при малом количестве пользователей и соответственно объемах документов поступающих в систему, основное развитие происходило как развитие сервиса печати. В этот период наблюдается монотонный рост количества пользователей и количества документов, проведенных через систему (рис. 2). Отдельно отметим корреляцию объемов печати с периодами сессии. Данный факт явно указывает на то, что система более востребована в период интенсивного обучения.

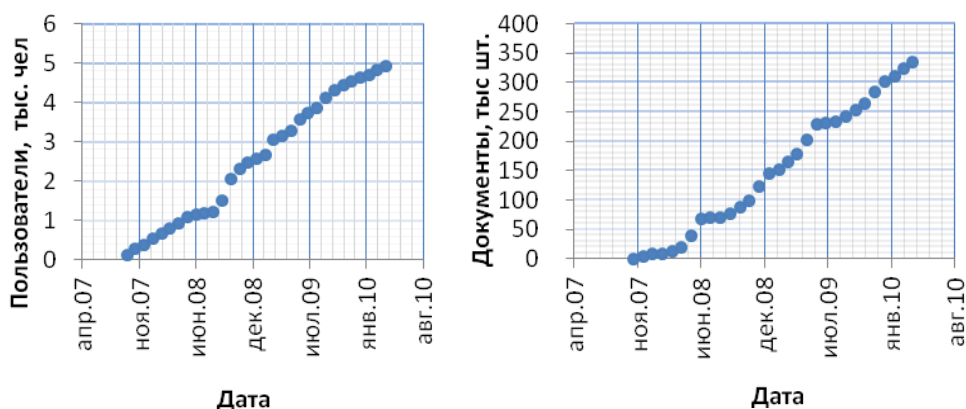


РИС. 2. Рост числа пользователей системы (слева) и числа напечатанных документов (справа)

Через два года после начала использования системы, когда накопилось достаточное количество активных пользователей, появилась возможность развития открытого хранилища. Сразу после своего создания открытая библиотека продемонстрировала быстрый рост количества документов. Частота добавления документов в хранилище сначала резко выросла, затем постепенно стала снижаться, а частота извлечения документов остается на примерно постоянном уровне (рис. 3). Из этого можно сделать вывод, что пользователи

быстро наполнили хранилище самыми часто используемыми учебными материалами. Анализ наиболее популярных документов подтвердил данную гипотезу: преимущественно, это учебники и конспекты лекций по основным учебным дисциплинам.

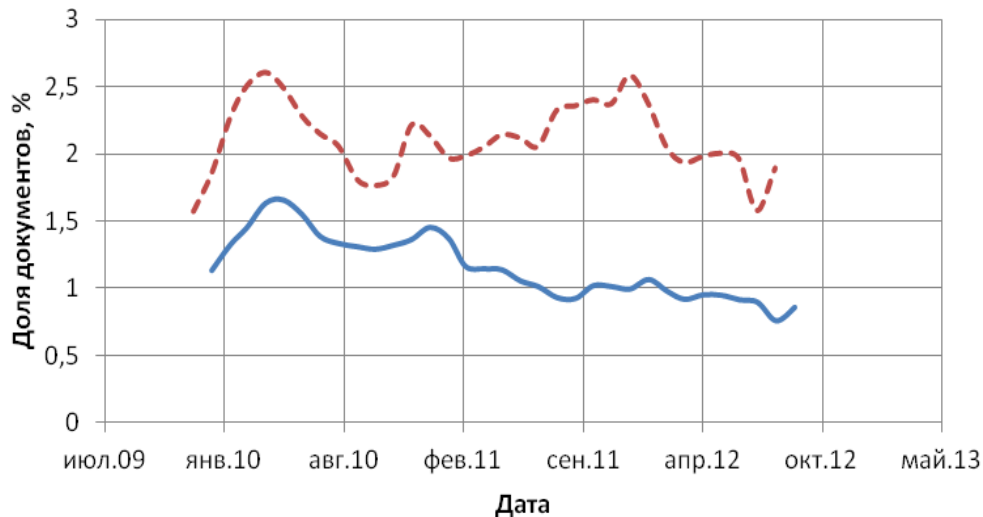


РИС. 3. Доля документов от общего числа напечатанных в системе за день, попавших в библиотеку (сплошная линия) и напечатанных из библиотеки (пунктирная линия)

С.М. Горбач, Л.О. Макаревич, К.Ю. Талецкий, В.О. Петрухин, В.И. Шоломова

ОРГАНІЗАЦІЯ ОБМІНУ ЕЛЕКТРОННИМИ ДОКУМЕНТАМИ ДЛЯ ПІДВИЩЕННЯ ЯКОСТІ ОСВІТНЬОГО ПРОЦЕСУ

Пропонується спосіб організації обміну електронними документами у вищому навчальному закладі та розглядається створення електронної системи вільного обміну інформацією. У ході роботи проаналізовано можливі способи розв'язання даної задачі, задано критерії оцінки результату, розроблено архітектуру системи та аналізуються результати впровадження.

S.N. Gorbach, L.A. Makarevich, K.Yu. Taletsky, V.A. Petrukhin, V.I. Sholomova

ORGANIZING ELECTRONIC DOCUMENTS EXCHANGE IN ORDER TO IMPROVE THE QUALITY OF EDUCATIONAL PROCESS

A method of organization of electronic documents exchange for educational institutions is proposed. Development of a system aimed to simplify this process is considered. In this study, various solutions to the problem are analyzed, quantitative criteria for the result are proposed. Also, architecture of such system is developed and efficiency of deployment of the system is analyzed.

1. *Brumbaugh, scott*. Deploying a VPN with PKI. Technology books, tech conferences, it courses, news – o'reilly media. [Электронный ресурс] 10 21, 2004.
http://www.oreillynet.com/pub/a/security/2004/10/21/vpns_and_pki.html.

2. *HRU*. Wikipedia. [Электронный ресурс] <http://en.wikipedia.org/wiki/HRU>.
3. *W3C members*. Portable Network Graphics (PNG) Specification (Second Edition). World Wide Web Consortium (W3C). [Электронный ресурс] 10 November 2003. <http://www.w3.org/TR/PNG/>.
4. *Adobe Systems Incorporated*. PDF Reference and Adobe Extensions to the PDF Specification. Adobe. [Электронный ресурс] June 2008. http://www.adobe.com/devnet/pdf/pdf_reference.html.
5. *Corp., HP*. PCL 5 Technical Reference Manual. б.м. : HP Corp., 1992.
6. *Adobe Systems Incorporated*. PostScript Language Reference Third Edition. – 1999.
7. *Xerox Corporation, Novell, Inc., Astart Technologies, Utah Valley State College* // Internet Printing Protocol/1.1: Model and Semantics. Internet Engineering Task Force. [Электронный ресурс] September 2000 г. <http://www.ietf.org/rfc/rfc2911.txt>.
8. *Apple Inc.* CUPS. Common Unix Printing System. [Электронный ресурс] 2010. <http://www.cups.org/>.
9. *Server Message Block*. Wikipedia.org. [Электронный ресурс] http://en.wikipedia.org/wiki/Server_Message_Block.
10. *World Wide Web Consortium (W3C)*. Hypertext Transfer Protocol. — HTTP/1.1. World Wide Web Consortium (W3C). [Электронный ресурс] <http://www.w3.org/Protocols/rfc2616/rfc2616.html>.
11. *TMG Ltd*. Включение Internet Printing Client в ОС Windows Vista / Server 2008. Физтех-Печать. [Электронный ресурс] 2009 г. http://old.print.mipt.ru/docs/vista_ipp/.
12. *TMG Ltd*. Физтех печать. [Электронный ресурс] 2007–2010. <http://print.mipt.ru>.

Получено 12.02.2013

Об авторах:

Горбач Сергей Николаевич,

аспирант Московского физико-технического института (государственного университета),
E-mail: sergey.gorbach@tmg-co.ru

Макаревич Леонид Александрович,

аспирант Московского физико-технического института (государственного университета),
E-mail: leo.makarevich@tmg-co.ru

Талецкий Константин Юрьевич,

аспирант Московского физико-технического института (государственного университета),
E-mail: konstantin.taletsky@tmg-co.ru

Петрухин Владимир Алексеевич,

ведущий научный сотрудник Института кибернетики имени В.М. Глушкова НАН Украины,
профессор Московского физико-технического института (государственного университета),
E-mail: vapetr@gmail.com

Шоломова Виктория Игоревна,

аспирантка Московского государственного университета имени М.В. Ломоносова.
E-mail: vsholomova@gmail.com