



«КОНТЕКСТУМ»: легальный контент малотиражных изданий

Несмотря на трудности понятийного аппарата, проблемы открытого доступа: «копирайт» – «копилефт», изменяющееся законодательство, серьёзные финансовые вложения, проекты, связанные с внедрением ЭБС получают всё большее распространение. "Контекстум" – тому подтверждение.

О самой системе, её возможностях и перспективах развития – в интервью с председателем Совета директоров ОАО ЦКБ «БИБКОМ» **Виталием ИВАНОВЫМ.**

— **Виталий Сергеевич, к известным проектам «БИБКОМа» добавляется ещё один, связанный с электронными ресурсами. Рынок давно об этом догадывался... Каков в целом, на Ваш взгляд, потенциал этого сегмента?**

— Электронное книгоиздание выходит сегодня на перспективные позиции. Свидетельство тому – множество профессиональных конференций, которые проходят на разных площадках – от книжных ярмарок до Интернет-сообществ, общественных организаций и т.д., и бурные дискуссии в специализированной прессе. Естественно, что мы не остались в стороне от тенденций отрасли.

Я бы сегодня воздержался от количественных прогнозов, связанных с потенциалом и объёмом рынка. Однако уже сегодня существуют данные о том, что объёмы продаж издательской продукции в цифровой форме на некоторых известных торговых площадках сопоставимы с объёмом продаж той же продукции на бумажных носителях.

Около трёх лет назад мы задумались о возможности создания современной IT-технологии для комплексного решения проблем, связанных с вопросами легального использования цифровых версий изданий. Работа проводилась в рамках консорциума «Контекстум» в составе нескольких организаций, имеющих многолетний опыт создания масштабных информационных технологий: Агентство «Книга – Сервис», ЦКБ «БИБКОМ», Агентство по распространению зарубежных изданий.

— **Что это за проект – «Контекстум», расскажите подробнее.**

— Мы задумались: а в чём, собственно, суть проблемы, суть противоречий между «копирайтом» и «копилефтом», которые лежат в основе многих профессиональных «баталий»? Ведь к интересам «копилефта» в полной мере можно отнести такие значимые сферы деятельности, как библиотечное дело и образование, которые отстаивают необходимость свободного использования информации в цифровой век.

Мы пришли к выводу, что эти противоречия во многом условны. Когда идут дискуссии на философском уровне, сложно понять, кто прав, кто виноват. Но если конфликт свести к технологической проблеме, в частности, к проблеме документооборота, то многое проясняется: с точки зрения «копилефта», невозможно вести лицензионную работу с правообладателями в объёме всего российского авторского потенциала.

Мы же посчитали, что теоретически это возможно. В последующем мы, правда, в полной мере ощутили и оценили мнение некоторых известных специалистов, что данную задачу невозможно решить в принципе в силу её многоаспектности, сложности и большой информационной размерности.

И всё же – если представить, сколько у нас объектов авторского права формируется в стране, то что надо лицензировать? По нашим подсчётам, это 600 – 700 тыс. объектов авторского права, включая и мультимедиа и софт. Цифра большая, но конечная. Таким образом, данную социально-правовую проблему («копирайт» – «копилефт») можно попытаться сформулировать в качестве сложной информационно-технологической задачи и определить требования к её решению. Предполагается, что, если речь идёт о легальном использовании до

Собственные электронно-библиотечные ресурсы ВУЗов

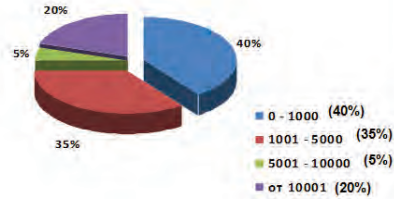
(март 2010 года)

Статистика по 343 ВУЗам из прямого подчинения по федеральным округам

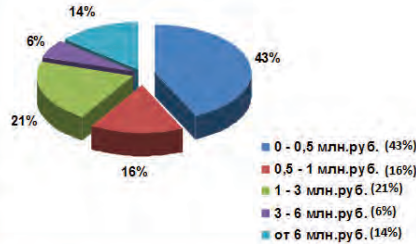
Дальневосточный федеральный округ	20	СНГ	4
Приволжский федеральный округ	64	Уральский федеральный округ	24
Северо-Западный федеральный округ	41	Центральный федеральный округ	105
Сибирский федеральный округ	47	Южный федеральный округ	38

Отсутствует 36 в 10 ВУЗах (3%)

Распределение количества ЭПД в ВУЗах



Финансирование создания ЭБС



- Всего в библиотеках данных ВУЗов содержится порядка **1,4 млн.** электронных документов.
- В среднем ВУЗовская библиотека содержит **4,2 тысячи** ЭПД.

- Всего на создание электронных библиотек ВУЗами затрачено порядка **1,14 млрд. руб.**
- В среднем ВУЗовская библиотека затратила **3,3 млн. руб.** на создание электронной библиотеки.

Приведённая стоимость одного ЭПД в ЭБС ВУЗа — 810 руб.

Рис. 1

700 тыс. объектов авторского права в год, то соответствующие технологии должны обеспечивать заключение по всей России одну-две тысячи лицензионных договоров в день и ведение соответствующей базы данных. Не обязательно весь объём работ осваивать одной фирме, мы ни в коем случае не претендуем на монополию.

Проект «Контекстум» — это прежде всего информационная технология, позволяющая вести массовую работу по оформлению и ведению лицензионных договоров в соответствии с IV частью ГК РФ, т.е. с наиболее острой и актуальной проблемой в области ведения и использования легальных цифровых ресурсов.

По сути дела, это некая «система закупок». Сегодня большинство проектов, связанных с агрегацией цифрового контента, за редким исключением (например, проект создания ФИПРа), если говорить с точки зрения бизнеса, построены как «системы продаж», или системы использования контента — в терминах 4-й части ГК РФ. Уточню: мы говорим только о цифровом контенте в скачиваемой форме.

Нами же создана технология, которая позволяет обеспечивать массовую легальную «закупку» этого контента и его «складирование» в некий базовый массив. Этот базовый массив предназначен для всех участников рынка использования контента. И для агрегаторов в том числе, которые могут брать из него контент, который им интересен. Мы гарантируем, что этот контент легальный: каждый файл связан с лицензионным договором в письменной форме.

— О каком контенте мы говорим?

— Сфера наших интересов лежит, прежде всего, в области малотиражной специализированной литературы, которая издается в российских регионах, в том числе вузами.

К примеру, монография в вузе печатается тиражом до 500 экземпляров. Считайте, что она вообще на рынок не выходит. А по результатам исследований, которые мы проводили, только российские вузы производят издательской продукции в области науки и образования не менее 150–170 тыс. наименований в год (без учёта диссертационных работ). Мы занимались этой проблемой, исследовали

взаимодействие издательской деятельности вузов с библиотечной. В своё время мы проанкетировали 17% российских вузов и получили достаточно репрезентативные данные. Фактически можно говорить об отсутствии в России важного и необходимого сегмента рынка авторской и издательской продукции.

— Ну, у ЦКБ «БИБКОМ» уже был опыт работы в этом сегменте... Не очень эффективный, как мне кажется.

— Да, мы уже лет восемь занимаемся практическими работами и исследованиями в области издательской деятельности вузов, издавали бумажный каталог изданий российских вузов, занимались распространением этой литературы в бумажном виде. Коммерчески он не был сверхуспешным, но мы этой цели и не преследовали. А вот опыт накопили немалый. ▶



Рис. 2

ОРГАНИЗАЦИОННО-ПРАВОВАЯ СХЕМА СОЗДАНИЯ ЭБС ВУЗа НА БАЗЕ ИТ «Контекстум»

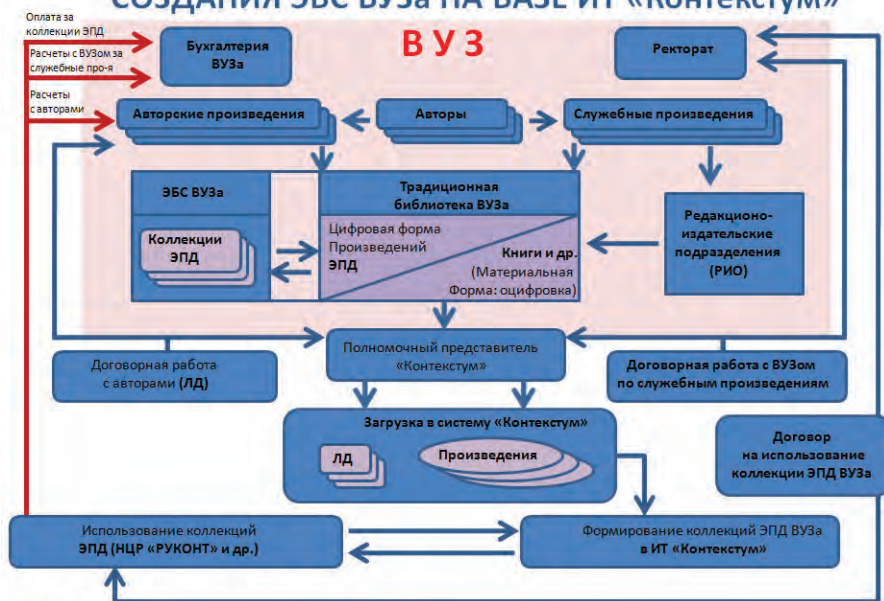


Рис. 3

Что же касается тех же самых изданий в электронной форме, здесь возникает совершенно иная экономическая ситуация. На наш взгляд, важно создание рыночных условий, механизмов и мотиваций по переводу этого типа издательской продукции в цифровую форму и агрегирование в современные высокотехнологичные национальные информационные ресурсы, использующие полнотекстовые документы. Важно отметить, что данное направление основывается на «Стратегии развития информационного общества Российской Федерации», утверждённой Президентом РФ ещё в 2008 г. (от 07.02.2008 № Пр. – 212).

Мы проанализировали состояние электронных библиотек в 343 вузах прямого подчинения Минобрнауки России (рис. 1). Данные предоставляли сами вузовские библиотеки. Получается, что вузы истратили за последние несколько лет на создание электронных полнотекстовых документов 1,4 млрд рублей. Сегодня стоимость одного полнотекстового документа в вузовских библиотеках – 810 рублей. Основная масса полнотекстовых библиотек – 40% – содержат до 1000 документов. Это очень небольшие коллекции. К тому же большинство из них являются не полностью легитимными с точки зрения ГК РФ. Значительную часть контента вузы не решаются вводить в гражданский оборот, потому что не уверены в соответствии данного контента требованиям законодательства в области авторского права.

— В чём суть проекта и технологии «Контекстум», как выстроена система сбора контента?

— Информационная технология «Контекстум» (рис. 2) состоит из трёх взаимосвязанных си-

стем. Первая – подсистема «Лицензионный договор». Она обрабатывает 7 групп и 16 типов произведений. Эта классификация – наша собственная. По сути дела, это классификация правовых отношений при заключении лицензионных договоров. Технологически это экспертная система («искусственный интеллект»), которая позволяет неспециалисту в области авторского права сгенерировать текст лицензионного договора. Договор в письменной форме – это следствие технологического процесса. Когда ведётся работа, а она ведётся через веб-интерфейс непосредственно в базе данных «Контекстума», по её результатам формируется запись,

эквивалентная лицензионному договору. Остаётся только нажать кнопку, и текст договора будет сгенерирован в pdf-формате. А дальше про бумагу можно забыть. Документ в письменной форме может понадобиться, только если возникает конфликтная ситуация.

Система непростая. Мы привлекали к её разработке ведущих юристов, которые участвовали в подготовке 4-й части ГК РФ.

Дальше возникает вопрос: а кто будет работать с подсистемой «Лицензионный договор», чтобы охватить достаточно большой ассортимент, сосредоточенный, в основном, в регионах? За эту часть технологии «отвечает» подсистема «Полномочный представитель». По нашим оценкам, необходимо иметь около тысячи полномочных представителей, имеющих доступ к нашей технологии. Доступ осуществляется через веб-интерфейс. Представителей необходимо просто наделить соответствующими полномочиями. Всё это заложено в технологии. К слову, ведение большой представительской сети – очень сложная технологическая задача.

Единицей хранения в базе является электронный полнотекстовый документ (ЭПД), который включает в себя как сам файл или несколько файлов, так и метаописание. Для каждого ЭПД заполняются метаданные, иначе он не загрузится в массив... Поиск ЭПД осуществляется на основе распознавания и индексирования файлов, там где это возможно.

Третья подсистема – это «Контент-менеджмент», служебная. Здесь и администрирование информационных массивов, и контроль качества ЭПД, и многое другое.

При всей технологической сложности проекта работа в системе, как показывает реальный опыт,

достаточно проста. Чтобы обрести первичные навыки полномочного представителя в системе требуется максимум 2–3 часа.

— Кто может стать полномочным представителем в Вашем проекте?

— Как правило, откликаются на наше предложение представители библиотек. Нас это устраивает, поскольку данная категория сотрудников и контент хорошо знает, и может профессионально подготовить метаданные.

— Сколько представителей у Вас сегодня? Какова динамика сбора контента?

— С учётом того, что проект стартовал в июне текущего года, на сегодняшний день в системе работают более 40 полномочных представителей, причём большинство из них не в Москве, а в регионах. В базе данных собрано около 16000 ЭПД. За 2 года мы хотели бы построить сеть из примерно 300 представителей, которые обеспечили бы поступление 100–150 тысяч объектов авторского права в год. Наша технология позволяет выдержать соответствующую нагрузку.

— А как стать Вашим представителем и как выстраивается система их мотивации? Как осуществляется обучение?

— Система разработана именно с тем расчётом, чтобы с нею легко мог работать любой пользователь, имеющий опыт работы с компьютером. Интерфейс разработан по подобию стандартной линейки всем известных майкрософтовских продуктов, включая веб-интерфейсы. Я уже говорил, что на обучение представителя уходит не более 2–3 часов. А стать нашим представителем очень просто: достаточно направить в наш адрес заявку в произвольной форме. Мотивация нашего представителя осуществляется в виде роялти по результатам дальнейшего использования произведений.

— Со сбором информации понятно: ваши представители заключают лицензионные договоры с правообладателями на издания, статьи, произведения и дистанционно агрегируют их в вашей базе. Как выстраивается организационно-правовая система использования этого контента?

— Ключевыми моментами при использовании издательской продукции в цифровой форме являются задачи, связанные с необходимостью работы с большим количеством правообладателей и различными типами их произведений, защита права на результаты интеллектуальной деятельности в со-

ответствии с требованиями 4 части ГК РФ. Эти требования сводятся к заключению прямых лицензионных договоров с правообладателями. Каждый лицензионный договор имеет сложную информационную структуру (необходимо контролировать и учитывать множество различных параметров — способы использования, сроки, возмездность и т.д.). Всё это заложено в нашей технологии.

Например, при работе с вузами — правообладателями контента в любом случае становится сам вуз. По сути дела, создаётся новый объект авторского права — коллекция, электронная библиотека полнотекстовых документов. Вуз как правообладатель может её использовать без каких-либо ограничений в учебном процессе. Если мы эту коллекцию используем в коммерческих целях на какой-либо торговой площадке, например в ФИПР, — на неё кто-то подписывается, мы получаем деньги и отчисляем роялти правообладателю. Причем, биллинг тоже заложен в технологию системы и проводится автоматически. Организационно-правовая схема отношений с вузом представлена на рис. 3.

Таким образом, с помощью нашей технологии вузы могут создать ЭБС, отвечающую требованиям, установленным федеральными государственными образовательными стандартами высшего профессионального образования.

— Осуществляется ли каким-то образом правовая экспертиза контента?

— Да, осуществляется: с помощью нашей технологии — полномочными представителями, а также контролем на входе в базовый массив документов.

— Какие дополнительные опции, сервисы предлагаются вашим клиентам в рамках проекта?

— Поскольку значительная часть контента поступает в нашу базу на основе безвозмездных договоров с правообладателями, то и пользователям он предоставляется бесплатно. Это один из наших дополнительных бонусов тем организациям, которые размещают свой контент на нашем ресурсе. ■

