

**Ефимов А.А.**

## **ОПИСАНИЕ ДАННЫХ МЕТАДААННЫМИ НА ПРИМЕРЕ СОЗДАНИЯ ЭЛЕКТРОННЫХ УЧЕБНЫХ МАТЕРИАЛОВ**

*yefimov@mail.ru*

*Уральский государственный педагогический университет*

*г. Екатеринбург*

В настоящее время мы ежедневно сталкиваемся в сети с информацией о новом проекте или среде разработки, соответствующих концепции WEB 2.0. При этом, мало кто из авторов осмеливается дать точное, честное и внятное определение этому термину.

По каким-то причинам, принято связывать WEB 2.0 с инструментами разработки, использованными при воплощении проекта, или с нишевой функциональностью, но, ни в коем случае с его структурой.

Так, сплетение языка разметки XML (eXtended Markup Language) и языка программирования JavaScript, получившее название AJAX (Asynchronous Javascript and XML) стало чуть ли не основным идейным знаменем WEB 2.0. а другие инструменты. Например язык программирования PHP вдруг стал не достаточно гибким и вообще не удобным для реализации модных сервисов. Данному языку сразу припомнили всё, как техническое несовершенство, так и идейные особенности. Например, на полном серьезе говорится что язык PHP (PHP: Hypertext Preprocessor, рекурсивная аббревиатура) провоцирует автора писать неаккуратный или нелогичный код.

Мы придерживаемся более общей точки зрения относительно определения WEB 2.0 в целом и семантической паутины в частности. Мы не привязываем WEB 2.0 к конкретным инструментам или областям применения, лишь считаем, что использование некоторых инструментов желательным, но ни как не достаточным, для создания сервиса, удовлетворяющего концепции WEB 2.0.

Что же до самого определения WEB 2.0 - оно не сформировано. Его формализация - задача недалёкого будущего, но на данном этапе в определённой практической и структурной плоскости, WEB 2.0 можно отождествить с семантической паутиной - структурой, наполненной смысловым содержанием о предоставляемой информации, пригодным для машинного анализа. Для сравнения можно привести определение Тима О'рейли, согласно которому, WEB 2.0 есть методика проектирования систем, которые путем учета сетевых взаимодействий, становятся тем лучше, чем больше людей ими пользуются.

Выбор инструмента для иллюстрации работы концепции одновременно и предсказуем и неоднозначен.

Для иллюстрации снабжения метаданными электронных учебных материалов, нами был выбран язык разметки текста HTML (Hypertext Markup Language) версии 4.

В наши дни IT индустрия использует формализованные языки разметки, подобные XML или даже SGML (Standard Generalized Markup Language), но проанализировав круг стоящих перед нами задач, по своей сути не далеко ушедших от типографской вёрстки текста, было принято решение об использовании более простого языка HTML. Для разработки реальных электронных материалов нами используется язык, XHTML (eXtensible Hypertext Markup Language), но лишь для иллюстрации семантической даже его возможности избыточны.

Чтобы конечный продукт, в данном случае - текст лекции, был максимально удобен для приведения его в соответствие с концепцией, он должен иметь "верную" структуру: текст необходимо разбить на составные части, озаглавить каждую часть, ввести перекрёстные ссылки, сноски, указатели, вводить определения, в то же время не злоупотребляя ими и датами.

Для иллюстрации использования метаданных возьмем фрагмент текста третьей главы книги "Основы права интеллектуальной собственности", который приведён ниже:

### *Смежные права*

Ранее мы выяснили, что объектами авторского права являются произведения науки, литературы и искусства, являющиеся результатом творческой деятельности, независимо от назначения и достоинства произведения, а вот объектами смежных прав являются исполнения, постановки, фонограммы, передачи организаций эфирного или кабельного вещания. Примечательно, что постановки в международных соглашениях не включаются в число объектов смежных прав.

Согласно "Глоссарию терминов по авторскому праву и смежным правам Всемирной организации интеллектуальной собственности", смежные права - это права, предоставляемые для охраны интересов исполнителей, производителей фонограмм и организаций вещания в связи с публичным использованием произведений авторов, всевозможных выступлений артистов или доведением до всеобщего сведения событий, информации и каких-либо звуков или изображений.

Субъектами смежных прав являются исполнители, производители фонограмм и организации эфирного или кабельного вещания.

В ст. 4 ЗоАП дается следующее определение исполнителя: это актер, певец, музыкант, танцор или иное лицо, которое играет роль, читает, декламирует, поет, играет на музыкальном инструменте или иным образом исполняет произведения литературы или искусства, а также режиссер-постановщик спектакля и дирижер.

В этой же статье дается определение изготовителя фонограммы, хотя далее по тексту ЗоАП употребляется термин "производитель фонограммы". Итак, "изготовитель фонограммы - физическое или

юридическое лицо, взявшее на себя инициативу и ответственность за первую звуковую запись исполнения или иных звуков; при отсутствии доказательств иного изготовителем фонограммы признается физическое или юридическое лицо, имя или наименование которого обозначено на этой фонограмме и (или) на содержащем ее футляре обычным образом".

Важно отметить, что производитель фонограммы - это лицо, которое ответственно за первую звуковую запись. На практике производитель фонограммы заказывает тиражирование записей сторонней организации. Запись - это фиксация звуков и (или) изображений с помощью технических средств в какой-либо материальной форме, позволяющей осуществлять их неоднократное восприятие, воспроизведение или сообщение.

Исполнитель осуществляет свои права при условии соблюдения прав автора исполняемого произведения. Производитель фонограммы, организация эфирного или кабельного вещания осуществляют свои права в пределах прав, полученных по договору с исполнителем и автором записанного на фонограмме или передаваемого в эфир или по кабелю произведения. Следует заметить, что организациям вещания могут принадлежать права на собственные авторским правом произведения.

Проанализировав данный фрагмент, легко заметить, что в нём вводится как минимум четыре определения: авторского права; субъектов смежных прав; исполнителя; изготовителя фонограммы. Уже не так очевидно, но исходя из тематики материала, логично дать ссылки на юридическое определение таких терминов как исполнение, постановка фонограммы, организация эфирного или кабельного вещания и некоторых других.

Так же, имея перед глазами весь текст главы, можно запросто создать аннотацию, которая в данном случае может выглядеть следующим образом: "Условия признания прав субъектов смежных прав. Исключительные права на объекты смежных прав. Судебная практика - Интернет и смежные права", и разбить главу на параграфы. Название первого параграфа очевидно даже из его содержания и текста аннотации - "Смежные права, понятие, сущность. Субъекты и объекты смежных прав".

Многие авторы, вооружившись разноцветными карандашами, размечают текст на бумаге - выделяют определения, делят на главы, параграфы, помечают термины, которые стоит пояснить и пр. Другие работают непосредственно с электронной формой. По нашему мнению оба подхода являются состоятельными и приводят практически к одному и тому же результату - "верной" структуре, наличию нужного и отсутствию лишнего.

Для создания правильной структуры лекции в электронном виде надо отметить базовые средства разметки документа средствами языка HTML где помимо инструментов создания абзацев, заголовков, ссылок и списков, могут быть описаны цитаты, определения, переменные, сноски, листинги кода и пр. Так же, одинаково выглядящий полужирный текст может быть как просто полужирным, как и типографским усилением (выглядит зачастую одинаково, но несёт разную смысловую нагрузку) и таких примеров можно привести великое множество.

Язык HTML создавался именно как виртуальный аналог типографии. Его корни уходят к задаче, поставленной учёными, нуждающимися в инструменте для обмена своими научными трудами, а это значит, что на нём удобно верстать классические тексты, аналогичные книжным, совсем не интерактивные, но структурированные и "верные".

Используя теги разметки H1 и H2 для названия главы и параграфа, DFN - для определений, EM - для определяемых терминов, SUP и A - для введения сносок, CITE - для выдержек из законов и нескольких других, мы получаем предсказуемый результат.

При использовании соответствующего программного обеспечения, которым в данном случае могут являться практически любой HTML браузер и сервис, аналогичный сервису валидации от W3C, мы можем выбрать все цитаты из документа, все определения, получить список всех терминов, трактовка которых однозначно прописана в законе, а также построить структуру документа (список глав с разделением на параграфы) и список аннотаций к главам и т.п..

Здесь важно учесть, что для достижения результата не требуется ни специальное программное обеспечение, ни значительные трудовые затраты.

В заключении хотелось бы отметить, что практически любые учебные материалы, при переработке под новые семантические стандарты, ничего не теряют, а лишь приобретают. Материал попутно может быть актуализирован, может быть подкорректирована его структура и содержание, ведь дисбаланс количества определений или размера параграфов сразу бросится в глаза.

С философской точки зрения, работа ведётся над формой, но при этом, улучшается качество содержания. Говоря о форме и содержании, стоит отметить, что концепция семантической паутины всё же подразумевает машинный анализ информации, и те метаданные, которыми мы описываем полезную информацию для пользователя являются формой, а для анализатора, которым может быть как пользователь, так и программа - содержанием. Но данный вопрос выходит за рамки этой статьи.

#### *Литература*

1. Тим О'Рейли. Что такое Веб 2.0 // Компьютерра. - 2005. - № 37.
2. Серго А. Г. Интернет и право. - М. : Бестселлер, 2003. - 272 с.