

**Использование формализованных методов при подготовке текста
для web-сайтов учреждений культуры**

Formal Text-Preparation Methods for Cultural Web-Sites

**Використання формалізованих методів при підготовці тексту
для web-сайтів установ культури**

Н. И. Гендина

*НИИ информационных технологий социальной сферы
Кемеровского государственного университета культуры и искусств, Кемерово, Россия*

Н. И. Колкова, О. И. Алдохина

Кемеровский государственный университет культуры и искусств, Кемерово, Россия

Nataliya Gendina

*Research Institute of Information Technologies in Social Sphere,
Kemerovo State University of Culture and Arts, Kemerovo, Russia*

Nadezhda Kolkova and Olga Aldokhina

Kemerovo State University of Culture and Arts, Kemerovo, Russia

Н. І. Гендіна

*НДІ інформаційних технологій соціальної сфери
Кемеровського державного університету культури і мистецтв, Кемерово, Росія*

Н. І. Колкова, О. І. Алдохіна

Кемеровський державний університет культури і мистецтв, Кемерово, Росія

Рассматриваются типичные недостатки контента web-сайтов учреждений культуры. Предлагается использование формализованных методов анализа и синтеза информации для построения качественного контента web-сайта учреждений культуры. Раскрывается суть методики формализованного составления текстов для отражения рубрик web-сайтов учреждений культуры. Описываются подходы к созданию унифицированных аспектно – маркерных описаний рубрик контента web-сайта.

Typical drawbacks of cultural web-sites are considered, and formal methods are suggested for the analysis and synthesis of information that will serve as a basis of quality web-site content. Described are the formal method used for writing texts in compliance with the headings of cultural web-sites and approaches that must be to create uniform aspect-marker descriptions of the headings of the web-site content.

Розглядаються типові недоліки контенту web-сайтів установ культури. Пропонується використання формалізованих методів аналізу і синтезу інформації для побудови якісного контенту web-сайту установ культури. Розкривається суть методики формалізованого складання текстів для відображення рубрик web-сайтів установ культури. Описуються підходи до створення уніфікованих аспектно-маркерних описів рубрик контенту web-сайту.

В современных условиях web-сайты учреждений культуры становятся мощным средством обеспечения доступа к культурному наследию и информационным ресурсам широкого круга удаленных пользователей, на практике претворяя в жизнь важнейший принцип информационного общества – всеобщий доступ к информации, образованию и культурным ценностям.

Полноценное выполнение web-сайтом его сущностных функций, прежде всего, определяется качеством его информационного наполнения (контента). Проблемы обеспечения высокого качества контента web-сайтов учреждений культуры в настоящее время являются предметом пристального внимания профессиональной общественности не только на российском, но и международном уровнях. Одним из важнейших результатов работы в этом направлении является выработанный в рамках проекта Minerva «Определение потребностей пользователей, содержания и критериев качества web-сайтов по культуре» документ «Принципы качества веб-сайтов по культуре. Руководство» (1). В нем представлено десять принципов качества web-сайтов музеев, библиотек и других

учреждений культуры. Важнейшими из них провозглашаются прозрачность, эффективность, ориентация на пользователя и др. Их реализация предполагает, что информационное наполнение высококачественного web-сайта должно быть релевантным, простым и доступным для пользователя. Однако ни в отечественной, ни в зарубежной литературе пока нет ответов на вопросы, о том каким образом можно обеспечить достижение таких потребительских качеств web-сайта, какова технология подготовки качественного контента сайта.

Создание контента web-сайта неразрывно связано с подготовкой текстов для наполнения его рубрик и подразбук. Требования, предъявляемые к текстам для web-сайтов, могут быть подразделены на общие (относящиеся ко всем информационным ресурсам) и специфические (характерные для сетевых информационных ресурсов). К числу общих требований, предъявляемых ко всем к текстам, включая тексты для web-сайтов, относятся: полнота, точность, логичность, последовательность, структурированность, четкость и ясность изложения материала, соответствие литературным нормам русского языка. Специфическими требованиями, обусловленными сетевой формой представления информации являются: лаконичность («телеграфный стиль»), порционность и обязательная рубрицированность; диверсификация различных стилей представления информации (официального, научного, публицистического, разговорного) в зависимости от типа сайта и выполняемых им функций; гипертекстовая форма представления информации (преодоление линейности обычных текстов).

Проведенный в НИИ информационных технологий социальной сферы Кемеровского государственного университета культуры и искусств (КемГУКИ) анализ показывает, что недостатки контента существующих web-сайтов различных учреждений культуры связаны с нарушением обеих групп требований. Типичными недостатками являются неполнота либо избыточность информации; разнородность и неструктурированность информации; несоответствие языка и стиля представления информации статусу, типу сайта и типу рубрик контента сайта; погрешности языка и стиля; затрудненность восприятия информации; монотонность подачи информации преобладание текста в ущерб иллюстративному представлению информации; перенос в состав контента web-сайта типичной для традиционных документов, но не характерной для web-пространства, линейной формы представления информации. Безусловно, что web-сайты, контент которых обладает такими недостатками, даже при релевантности его содержания запросам пользователей, не будут оценены ими как качественные.

Эти недостатки во многом объясняются тем, что созданием контента сайтов зачастую занимаются специалисты, не обладающие профессиональной подготовкой в области работы с текстами. Однако и для тех, кто обладает базовыми филологическими, библиотечно-информационными, редакторскими знаниями и умениями, работа над текстами для web-сайтов представляет собой весьма сложную и трудоемкую задачу, решение которой не обеспечено необходимым теоретическим, технологическим и методическим инструментарием. Ситуация усугубляется хаотичностью лексики сайтостроения, диффузионностью классификации web-сайтов, несовершенством системы оценки качества web-сайтов, фрагментарностью описания технологии создания сайта, отсутствием типовых схем контента web-сайтов, неразработанностью справочной и инструктивно-методической базы создания web-сайтов. Не случайно в подготовке текстов для web-сайтов доминируют эмпирически – описательные методы.

Одним из путей повышения качества текстов для web-сайтов, на наш взгляд, является использование методов формализации, в частности, формализованных методов информационного анализа и синтеза информации (2-5). Формализованные методы вошли в информационную практику как реакция на сверхбольшие объемы информации, явившиеся следствием информационного взрыва XX столетия, и невозможность в массовом технологическом производстве решить проблему аналитико-синтетической переработки информации за счет интеллектуальных методов.

Повышение качества подготовки вторичных документов за счет снижения субъективизма библиотечно-информационных специалистов, призванных каждодневно решать сложную интеллектуальную задачу выявления в тексте первичного документа основного смыслового содержания и отделения существенной информации от не существенной; обеспечение возможности четкой ориентации пользователей в текстах вторичных документов, оперативной узнаваемости их сущности; массовая воспроизводимость результатов формализованного свертывания; уменьшение

временных, интеллектуальных и иных трудозатрат на аналитико-синтетическую переработку информации – все эти причины обусловили популярность и востребованность в библиотечно-информационной практике формализованных методов.

Формализованные методы работы с информацией, как правило, ассоциируются в библиотечно-информационной деятельности с процессами свертывания информации и созданием различных вторичных документов. Традиционными областями применения формализованных методов, в которых убедительно доказана их эффективность, являются технологии свертывания информации: индексирования, аннотирования и реферирования, составления обзоров, успешно реализуемые как на массивах научно-технических, так и публицистических текстов.

Вместе с тем формализованные методы способны обеспечить существенное повышение эффективности работы с информацией и при использовании их в технологиях развертывания информации, связанных с подготовкой научных статей, докладов, диссертаций и других первичных документов. Использование метода формализации олицетворяет осознанное стремление подняться над эмпирически-описательными методами конструирования текстов и на этой основе обеспечить повышение их качества. Суть методики формализованного составления текста заключается в ориентации на заранее выделенный перечень относительно самостоятельных смысловых фрагментов – аспектов содержания. При этом для каждого аспекта устанавливается определенная совокупность только ему присущих формальных признаков (маркеров), однозначно определяющих его содержание. Маркеры могут быть выражены устойчивыми словами или словосочетаниями, фрагментами фраз.

Термин «формализация» в контексте рассматриваемых задач может трактоваться как аспектно – маркерный метод представления содержания определенной тематической рубрики или подрубрики контента web-сайта с использованием перечня заранее заданных для нее аспектов и соответствующих им маркеров. Использование аспектно – маркерного метода представления содержания рубрики или подрубрики контента web-сайта предполагает предварительное определение вида рубрики. Как правило, для описания контента сайта используются рубрики двух видов: рубрики справочно-информационного характера («Новости», «Мероприятия», «Проекты», «Персоны» и т. п.) и рубрики коммуникативного характера («Вопрос – ответ», «Наиболее часто задаваемые вопросы», «Заказ» и др.). Следует заметить, что рубрики справочно-информационного характера в составе контента web-сайта могут отражать объекты как единичного («История музея», «История библиотеки», «История филармонии», «История театра» и т. п.), так и собирательного характера («Новости», «Персоны», «Проекты» и т. д.). При этом отражение объектов как единичного, так и собирательного характера в составе контента web-сайта с помощью аспектно – маркерного метода сводится к выполнению несложного алгоритмического предписания, предусматривающего ознакомление с унифицированным описанием конкретной рубрики (или подрубрики), содержащим максимально полный перечень всех аспектов ее содержания; выбор из имеющегося перечня аспектов содержания, подлежащих отражению в составе текста той или иной рубрики (подрубрики), конструирование в соответствии с заданным перечнем аспектов содержания текста рубрики (подрубрики) и соответствующих им маркеров текста рубрики (подрубрики) web-сайта.

Составление унифицированных описаний может быть обеспечено с помощью полного заимствования готовых схем описаний объектов; разработки собственных схем описаний объектов; сочетания заимствованных и самостоятельно разработанных схем описания объектов (комбинированный способ). Использование комбинированных способов структурирования описаний объектов осуществляется в тех случаях, когда набор представленных в заимствованных описаниях характеристик объекта оказывается недостаточным для его отражения на web-сайте.

Основой для составления такой унифицированной схемы с помощью метода заимствования могут стать методические рекомендации открытой энциклопедии «Рубрикана», размещенной на web-сайте «Рубикон» (www.rubicon.com). Данный сетевой информационный ресурс включает значительный массив энциклопедических описаний различных объектов культуры и искусства: театров, библиотек, цирков, кинотеатров, архивов и т. п. Наряду с энциклопедическими описаниями открытой энциклопедии «Рубрикана» в качестве источников для выявления аспектов, раскрывающих содержание тех или иных рубрик, могут также использоваться нормативные, справочные, учебные и научные документы и ИПЯ (УДК, ББК, тезаурусы и т. п.).

Разработка собственных схем описаний объектов предполагает формирование массива web-сайтов, содержащих однотипные рубрики, например, рубрику «Новости». Затем тексты, отражающие данную рубрику на всех анализируемых web-сайтах, подлежат формализованному анализу с целью выявления максимально полного перечня аспектов содержания. Выявленные аспекты упорядочиваются в ранжированный ряд в зависимости от частоты встречаемости. Наиболее часто встречаемые аспекты упорядочиваются в соответствии с семантической формулой. Так, например, для получения унифицированной схемы описания рубрики «Новости», нами было проанализировано по десять новостных сообщений, представленных на тридцати web-сайтах учреждений культуры (театры, музеи, библиотеки). Таким образом, всего формализованному анализу подлежало 300 новостных сообщений. В ходе такого анализа в тексте каждого новостного сообщения выделялись все представленные в нем аспекты содержания. В результате было выявлено семь аспектов содержания, которые использовались в большинстве (более 50%) анализируемых сообщений. Эти аспекты в обобщенном виде могут быть представлены семантической формулой: ««Когда? «+ «Где? «+ «О Чем? «+ «Кто? «+ «Что сделано (Что представлено)? «+ «Как? «+ «Поэтому». В соответствии с представленной семантической формулой унифицированная схема текста рубрики «Новости» любого web-сайта учреждения культуры должна включать следующие аспекты: предмет рассмотрения (событие, тема, проблема); место происходящего события; время происходящего события; характеристика события; лица, принимающие участие в событии; результаты, выводы; оценка события).

В качестве примера ниже приведена аспектная структура новостного сообщения, построенного в соответствии с представленной выше семантической формулой. Основой для ее построения явился один из фрагментов рубрики «Новости», представленный на сайте одного из объектов нашего исследования.

Аспект содержания	Пример
Время происходящего события	27 апреля 2007 г.
Место происходящего события	Новгородская областная научная библиотека
Предмет рассмотрения (событие, тема, проблема)	Презентация электронного альбома «Особо охраняемые природные территории Новгородской области»
Лица, принимающие участие в событии	В презентации приняли участие студенты третьего курса Новгородского государственного университета
Характеристика события	После презентации прошло обсуждение вопросов, связанных с возможностью использования электронного альбома для развития туристической деятельности в Новгороде
Результаты, выводы	Известные в Новгороде туристические фирмы предложили новые маршруты, связанные с охраняемыми территориями
Оценка события	Презентация получила высокую оценку экспертов в области туризма и охраны природы, а также специалистов Комитета культуры и туризма Новгородской области

В виде связного текста эти сведения могут быть представлены следующим образом:

«27 апреля 2007 г. в Новгородской областной научной библиотеке состоялась презентация электронного альбома «Особо охраняемые природные территории Новгородской области». В презентации приняли участие студенты третьего курса Новгородского государственного университета. После презентации прошло обсуждение вопросов, связанных с возможностью использования

электронного альбома для развития туристической деятельности в Новгороде. Известные в г. Новгороде туристические фирмы предложили новые маршруты, связанные с охраняемыми территориями. Презентация получила высокую оценку экспертов в области туризма и охраны природы, а также специалистов Комитета культуры и туризма Новгородской области».

Использование аспектно – маркерного метода как средства формализации подготовки текстов для web-сайтов позволяет обеспечить построение качественного контента web-сайта, удовлетворяющего требованиям целостности и полноты его содержания. Вместе с тем, наличие унифицированных в пределах web-сайта схем описания объектов позволяет четко структурировать материал, делает его «прозрачным», тем самым значительно облегчает восприятие информации для пользователя за счет логичности, последовательности, четкости и ясности изложения описания объектов сайтостроения. Аспекты содержания, явно выраженные в тексте той или иной рубрики web-сайта, могут служить основой для формирования системы гиперссылок и, тем самым, для преодоления линейности исходных текстов. Наряду с использованием аспектно- маркерного метода для построения рубрик контента, он также может использоваться и для совершенствования разработки схемы поиска информации, представленной в составе контента web-сайта. При этом создание контента web-сайта становится не только менее сложным, но и менее дорогостоящим делом.

Литература

1. Принципы качества веб-сайтов по культуре. Руководство [Текст]/Под. Ред. Темпера М. Т. Н., Темпера А. -М.,2006. -61 с.
2. |Гендина, Н. И. Выявление аспектов содержания документов с помощью формальных текстовых признаков [Текст] /Н. И. Гендина // Науч. -техн. информ. Сер. 2. Информационные процессы и системы. -1979. – №4. – С. 7-12.
3. Гендина, Н. И. Реферирование [Текст] /Н. И. Гендина // Справочник библиографа / науч. ред. А. Н. Ванеев, В. А. Минкина. – 2-е изд., перераб и доп. – СПб., Профессия, 2003. – С. 375- 387.
4. Гендина, Н. И. Аннотирование [Текст] /Н. И. Гендина, В. П. Леонов // Справочник библиографа / науч. ред. А. Н. Ванеев, В. А. Минкина. – 2-е изд., перераб и доп. – СПб., Профессия, 2003. – С. 362- 374.
5. Гендина, Н. И. Формализация свертывания краеведческой информации как способ повышения качества информационных продуктов и услуг: результаты исследований научно-исследовательского института информационных технологий социальной сферы Кемеровского государственного университета культуры и искусств [Электронный ресурс]/ Н. И. Гендина, О. И. Алдохина // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: материалы конф. – Электрон. дан. – М.: ГПНТБ России, 2004. – 1 электрон. опт. диск (CD-ROM). – Систем. требования: IBM PC, Windows 95 и выше. – ISBN 5-85638-091-6. – № гос. регистрации 0320400576,1500 экз.