

# Структуризация мультимедиа ресурсов для семантических web-сайтов

Мошенко С.Г.; Колб Д.Г.

Кафедра ИИТ, ФИТиУ

БГУИР

Минск, Беларусь

e-mail: moshenko@bsuir.by

**Аннотация** — В данной статье рассматривается проблема, связанная с поиском нужной информации среди мультимедиа ресурсов в сети Internet. Также в этой статье приводится пример структуризации мультимедиа ресурсов на сайте [www.conf.ostis.net](http://www.conf.ostis.net).

**Ключевые слова:** семантическая структуризация; медиа ресурсы; web-сайт

## I. ВВЕДЕНИЕ

Актуальность работ в данном направлении обусловлена все возрастающей потребностью предоставления пользователю информации в виде различных мультимедиа данных и проблемами, возникающими у пользователя при работе с такими медиа ресурсами посредством сети Internet.

В настоящее время в сети Internet размещено множество различных web-сайтов, использующих мультимедиа данные. Алгоритмы поиска, используемые в настоящее время, среди мультимедийных ресурсов либо не работают, либо затрачивают много времени на поиск, либо дают неверный результат [1]. Эффективным способом облегчения поиска информации для таких web-сайтов, является использование семантической структуризации мультимедийных ресурсов, размещенных на сайте.

Один из подходов к решению этой задачи использован при реализации сайта международной конференции Open Semantic Technologies for Intelligent Systems (OSTIS). В основе подхода лежит структуризация мультимедиа ресурсов, с использованием семантически структурированного гипертекста [2]. Основным способом кодирования таких гипертекстов является SCn-код – способ псевдоестественного кодирования семантически структурированных гипертекстов[2].

Все данные размещенные на сайте касаются информации о конференциях прошедших или планирующихся в рамках проекта OSTIS. Мультимедийные ресурсы, имеющиеся на сайте, включают в себя: фотографии форматов png и jpg, гипермедийные ссылки, документы в формате pdf, презентации, географические данные.

В рамках сайта конференций OSTIS все имеющиеся на данный момент мультимедийные ресурсы можно классифицировать по способу предоставления информации на графические (фотографии, географические данные) и текстовые (гипермедийные ссылки, текстовые документы).

Также мультимедийные ресурсы на сайте OSTIS классифицируются по интерактивности на интерактивные (гипермедийные ссылки, географические данные) и не интерактивные (фотографии, текстовые документы).

## II. СТРУКТУРИЗАЦИЯ ИЗОБРАЖЕНИЙ

Изображения, расположенные на сайте разбиваются на несколько групп. В первую группу входят изображения, которые являются логотипом

организации, такие изображения входят в состав информации, специфицирующей конкретную организацию, и связываются отношением “синонимичные идентификаторы”, и уточняются атрибутивным отношением эмблема.

Вторая группа изображений – это изображения, которые используются для обозначения гиперссылок. Семантика таких изображений совпадает с семантикой использования гипермедийных ссылок.

И, наконец, третью наиболее многочисленную группу составляют изображения людей или фотографии. Фотографии, представленные на сайте, отображают историю проведения конференции. Фотографии структурируются по дате (на какой конференции был произведен снимок), персонам (кто изображен на снимке) и мероприятиям (какое событие изображено на снимке).

По дате фотографии делятся на:

- OSTIS-2011;
- OSTIS-2012.

По персонам фотографии делятся на:

- персональные;
- коллективные.

По мероприятиям фотографии делятся на:

- регистрация;
- открытие конференции;
- интервью;
- секции;
- доклады;
- перерыв;
- награждение;
- закрытие конференции.

Фотография связывается отношениями:

- “кто изображен” с персонами;
- “что изображено” с мероприятиями.

## III. СТРУКТУРИЗАЦИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ

Текстовые документы, размещенные на web-сайте, представлены файлами формата pdf, содержащими тексты докладов участвующих в конференции.

Текстовый документ связывается с докладом отношением “Опубликованные материалы”. Также текстовый документ связывается отношениями:

- “ссылка по стандарту” с названием доклада оформленным по ГОСТу;

- “аннотация” с аннотацией данного доклада;
- “комментарий” с пояснениями к данному докладу.

Текстовые документы структурируются по дате (на какой конференции был заслушан доклад) и секциям (на какой секции был заслушан доклад).

По дате текстовые документы делятся на:

- OSTIS-2011;
- OSTIS-2012.

Также текстовые документы содержащие доклады участников делятся на:

- состоявшиеся доклады;
- не состоявшиеся доклады.

#### IV. СТРУКТУРИЗАЦИЯ МУЛЬТИМЕДИА ПРЕЗЕНТАЦИЙ

Переходя к вопросу структуризации мультимедиа презентаций, отметим, что с точки зрения пользователя презентация связана с докладом отношением “Иллюстрационные материалы”.

Так как презентации, размещенные на сайте, представляют собой вспомогательные иллюстрационные материалы для докладов, то их структуризация будет подобна структуризации текстовых документов.

Презентации структурируются по дате (на какой конференции был произведен снимок) и секциям (на какой секции был заслушан доклад).

Презентации не делятся на состоявшиеся и не состоявшиеся.

#### V. СТРУКТУРИЗАЦИЯ ГЕОДАНЫХ

Географические данные структурируются по типу учреждения (какое учреждение находится по данному адресу):

- гостиница;
- пункт питания;
- организатор конференции;
- спонсор конференции.

Географические данные связываются с объектом, расположение на который указывают, отношением “месторасположение”. Месторасположение для географических объектов задается с помощью использования данных почтового адреса и одновременно с использованием методов геокодирования отображается на интерактивной географической карте, API для работы с которыми предоставляется компанией Yandex.

#### VI. СТРУКТУРИЗАЦИЯ ГИПЕРМЕДИЙНЫХ ССЫЛОК

Использование семантически-структурированных гипертекстов позволяет устанавливать семантический тип перехода по гипермедийной ссылке. Такая возможность появляется за счет использования SCn-кода как основного способа представления информации на сайте.

С точки зрения пользователя гипермедийные ссылки, представленные на сайте, являются частью sc.n-предложения, описывающего связь описываемого объекта с некоторым объектом, указанным в предложении. Любое sc.n-предложение состоит из sc-идентификатора описываемого объекта и одного sc.n-поля описания связи.

Гипермедийные ссылки структурируются по месту расположения в sc.n-предложении:

- sc.n-поле описания связи;
- объект, с которым связывается описываемый объект;
- смешанные.

Также гипермедийные ссылки структурируются по виду sc.n-предложения, в котором они находятся:

- sc.n-предложение об авторах;
- sc.n-предложение об адресе;
- sc.n-предложение об изданных материалах;
- sc.n-предложение об организаторах;
- sc.n-предложение об организационном комитете;
- sc.n-предложение об официальном приглашении;
- sc.n-предложение об официальной программе;
- sc.n-предложение об участниках конференции;
- sc.n-предложение об электронном адресе;

- sc.n-предложение об электронном почтовом адресе;

- sc.n-предложение о времени завершения;
- sc.n-предложение о времени начала;
- sc.n-предложение о времени проведения;
- sc.n-предложение о включении;
- sc.n-предложение о декомпозиции;
- sc.n-предложение о докладе;
- sc.n-предложение о ключевых объектах;
- sc.n-предложение о названии;
- sc.n-предложение о местах размещения участников;
- sc.n-предложение о месте проведения;
- sc.n-предложение о местоположении;
- sc.n-предложение о персональной фотографии;
- sc.n-предложение о поддержке;
- sc.n-предложение о пояснении;
- sc.n-предложение о правилах оформления публикуемых материалов;
- sc.n-предложение о примере;
- sc.n-предложение о принадлежности;
- sc.n-предложение о программном комитете;
- sc.n-предложение о публикациях;
- sc.n-предложение о пунктах питания участников;
- sc.n-предложение о телефонном номере;
- sc.n-предложение о тематике;
- sc.n-предложение о семантической эквивалентности;
- sc.n-предложение о семантической близости;
- sc.n-предложение о синонимии;
- sc.n-предложение о строгом включении;
- sc.n-предложение о фотоматериалах.

Используя такой подход к структуризации, пользователь может легко визуально заметить связь, описываемого понятия с понятием к которому ведет гипермедийная ссылка.

#### VII. ЗАКЛЮЧЕНИЕ

Отметим, что рассмотренные классы мультимедиа ресурсов соответствуют сущностям некоторой предметной области. И хотя роли таких мультимедиа ресурсов для разных предметных области разные, ряд отношений, выделенных в данной работе, и подход в целом, можно использовать для семантической структуризации и систематизации любых web-ресурсов, обладающих мультимедийным содержанием.

Подход апробирован в рамках работы над сайтом конференции OSTIS, который доступен по адресу <http://www.conf.ostis.net>.

[1] P. Hobson, I. Kompatsiaris, “Semantic Multimedia and Ontologies: Theory and Applications”. Berlin.: Springer, 2007.

[2] Д. Г. Колб. Направления, методы и средства применения семантических сетей в internet-технологиях. Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2011). Мн.: БГУИР, 2011, с. 443-462.