|  | **Софийски университет „Св. Кл. Охридски”**  Факултет по математика и информатика  *Бакалавърска програма „Софтуерно инженерство”* |  |
| --- | --- | --- |

**Предмет: XML технологии за семантичен Уеб**

***Зимен семестър, 2022/2023 год.***

**Тема № 03: “Каталог на манастири - 1”**

**Курсов проект**

*Автори:*

*Мария Симеонова, фак. номер 855361*

*Валерия Стояновска, фак. номер 855365*

януари, 2023 г.

София

**Съдържание**

1 Въведение……………………………………………………………………………………………………………… 3

2 Анализ на решението……………………………………………………………………………………………. 3

2.1 Работен процес……………………………………………………………………………………..…… 3

2.2 Структура на съдържанието………………………………………………………………………. 3

2.3 Тип и представяне на съдържанието………………………………………………………… 5

3 Дизайн……………………………………………………………………………………………………………………. 6

4 Тестване…………………………………………………………………………………………………………………. 8

5 Заключение и възможно бъдещо развитие………………………………………………………. 10

6 Разпределение на работата………………………………………………………………………………… 10

7 Използвани литературни източници и Уеб сайтове…………………………………………… 10

# **1** **Въведение**

Настоящият документ представлява шаблон за оформление на документация на курсов проект за предмета „XML технологии за семантичен Уеб“. Изискванията за изработване и представяне на проекта са описани на сайта на този предмет в Мудъл.

Насоки за разработка на секцията:

* Информацията в настоящия документ е напълно актуално
* Проектът ни позволява да могат бързо да се прелистват профилите на манастирите с техните църкви и свещеници.
* Проблемът се разрешава чрез XML базиран каталог, който съхранява основната информация за регионите, манастирите, църквите и свещениците и ги обединява логически.

Надолу в документа се съдържа информация за нашите анализ на решението, дизайн, тестване и заключение за бъдещо развитие на проекта.

# **2** **Анализ на решението**

## **2.1** **Работен процес**

Работният процес е следният:

* Данните са взети от wikipedia.org
* Те са ръчно въведени в документа, като е спазена йерархична наредба, която е описана по-долу.
* Изходът е структурираната информация, която потребителят може да използва за своите цели.

## **2.2** **Структура на съдържанието**

В проекта се съдържат четири основни типа данни:

1. Региони:

<Regions>

<Region region\_id="">

<direction></direction>

</Region>

</Regions>

1. Манастири:

<Monasteries>

<Monastery monastery\_id="" region\_ref="" church1\_ref="" church2\_ref="" abbot1\_ref="" abbot2\_ref="">

<name></name>

<image src=""/>

<construction\_year></construction\_year>

<area></area>

<description></description>

</Monastery>

</Monasteries>

1. Църкви:

<Churches>

<Church church\_id="" abbot\_ref="">

<church\_name></church\_name>

<church\_area></church\_area>

</Church>

</Churches>

1. Свещеници:

<Abbots>

<Abbot abbot\_id="" church\_ref="">

<abbot\_name></abbot\_name>

<age></age>

</Abbot>  
</Abbots>

Както може да се забележи, манастирите имат уникален идентификационен номер, както и референции към регион, две църкви и два свещеникa. Църквите имат уникален идентификационен номер и референция към съответния свещеник. Свещениците имат уникален идентификационен номер и референция към църквата в която работят, а регионите имат само уникален идентификационен номер.

## **2.3** **Тип и представяне на съдържанието**

Използвали сме 4 текстови файла:

* monastery.xml – Size: 9 KB
* monastery.xsl - Size: 9 KB
* monastery.dtd - Size: 2 KB
* monastery.pdf - Size: 4.43 MB

Използвали сме 7 файла с изображение:

* BachkovoMonastery.jpg - Size: 153 KB
* CherepishMonastery.jpg - Size: 73.5 KB
* Dryanovo.jpg - Size: 138 KB
* Kremikovtsi Monastery.jpg - Size: 225 KB
* RilaMonastery.jpg - Size: 3.32 MB
* RozhenMonastery.jpg - Size: 457 KB
* Shipka Monastery.jpg - Size: 78.4 KB

# **3** **Дизайн**

**3.1 Свързаност на данните**

Както вече казахме в секция 2.2, връзките в проекта са следните:

* Манастир -> Регион
* Манастир -> Църква 1
* Манастир -> Църква 2
* Манастир -> Свещеник 1
* Манастир -> Свещеник 2
* Църква -> Свещеник
* Свещеник -> Църква

За постигане на тези връзки използваме следните атрибути:

* Манастир -> Регион -region\_ref=““
* Манастир -> Църква 1 - church1\_ref=““
* Манастир -> Църква 2 - church2\_ref=““
* Манастир ->Свещеник 1 - abbot1\_ref=““
* Манастир -> Свещеник 2 - abbot2\_ref=““
* Църква -> Свещеник - abbot\_ref=““
* Свещеник -> Църква - church\_ref=““

**3.2 Валидация**

Валидацията чрез DTD се изразява в това, че всеки елемент, атрибут на елемент или  
поделемент е експлицитно деклариран:

* Елемент може ли да има стойност, ще се парсва ли от процесора или не?
* Елемент какви поделементи/атрибути има.
* Поделементите в каква последователност трябва да са, кои от тях са не/задължителни, кои от тях могат повече от веднъж да се повтарят.
* Атрибутите на даден елемент какви стойности могат да приемат, дали са задължителни или не.
* Надграждане над основните типове стойности позволени от DTD.
* Какви стойности могат да приемат(за атрибутите).
* Какви под елементи или атрибути съдържат, в каква последователност трябва да бъдат.
* Кои от тях са задължителни и кои не.

Например, за валидирането на елемента Monastery сме използвали следното:

<!ELEMENT Monastery (name, image, construction\_year, area, description)>

<!ATTLIST Monastery monastery\_id ID #REQUIRED

region\_ref IDREF #REQUIRED

church1\_ref IDREF #REQUIRED

church2\_ref IDREF #REQUIRED

abbot1\_ref IDREF #REQUIRED

abbot2\_ref IDREF #REQUIRED>

<!ELEMENT name (#PCDATA)>

<!ELEMENT image EMPTY>

<!ATTLIST image src ENTITY #REQUIRED>

<!ELEMENT construction\_year (#PCDATA)>

<!ELEMENT area (#PCDATA)>

<!ELEMENT description (#PCDATA)>

За валидирането на нашия xml файл сме използвали външно и вътрешно DTD, като вътрешното само добавя допълнителни елементи, конкретни за стойностите на атрибутите на елементите в xml файла – например custom ENTITY, чрез които се реферира към съответната снимка :

<!ENTITY Bachkovo SYSTEM "images/BachkovoMonastery.jpg" NDATA JPG>

Понеже използваме JPEG => не е сред основните типове стойности във вътрешното DTD трябва експлицитно да го добавим. Това става чрез :

<!NOTATION JPG SYSTEM "image/jpg">

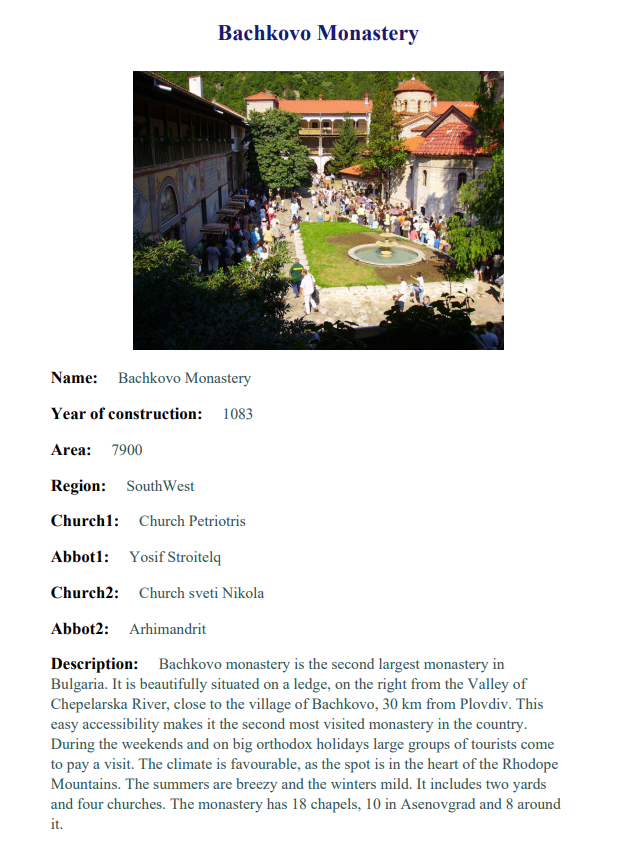
Така ако решим може да преизползваме външното DTD за друг проект, без да съдържа  
референции локални за самия проект.

# **4** **Тестване**

Тестване на предложеното решение:

* Направихме DTD валидация, за да сме сигурни, че XML документът е валиден спрямо DTD граматиката и проектът е добре форматиран
* Крайният резултат е представен в pdf формат, който е универсален за различните платформи за гледане на pdf файлове
* Примери от крайния резултат:





**5** **Заключение и възможно бъдещо развитие**

Резултатът от проекта е опростено портфолио в pdf формат. Заради липсата на всякакви визуални ефекти, това е лесен за използване каталог, който се фокусира върху предаването на информация. В съвременния свят съществуват множество формати за пренасяне на информация, които биха свършили същата работа. Пример за такъв формат е JSON, който до голяма степен може да прави същото, което прави XML.

Този проект е подходящ за бързи справки както от хора, които се интересуват от манастири и църкви, така и от такива, които нищо не разбират. Той съдържа само най-важните атрибути на обектите и може да се използва за запознаване с даден манастир, техните църкви и свещеници.

# **6** **Разпределение на работата**

Всички части от проекта са изготвени взаимно, чрез взаимна комуникация и консултации.

Например: Мария Симеонова е започнала с писането на xml файлът, след което Валерия Стояновска е написала съответната dtd граматика. Това се отнася за целия проект.

* Валидиране на xml файл чрез dtd
* Трансформация от xml към pdf чрез използване на xslt файл
* Създаване на дизайна
* Събиране и внедряване на информацията ( Мария -> информация за манастири и църкви, Валерия -> информация за свещеници и намиране на снимки )
* Изготвяне на документацията ( Валерия -> точки 1,2,5, Мария -> точки 3,4,6,7 )

# **7** **Използвани литературни източници и Уеб сайтове**

1. <https://en.wikipedia.org/wiki/Main_Page> -> част от информацията за обектите в проекта, включително изображенията
2. <https://www.xmlvalidation.com/> -> DTD валидация, за проверка на валидността на XML документът спрямо DTD граматиката
3. <https://www.bulgariamonasteries.com/en/> -> част от информацията за църквите в проекта