**Софийски университет „Св. Климент Охридски”**



Факултет по математика и информатика

*Бакалавърска програма*

*„Софтуерно инженерство”*

**Предмет: XML технологии за семантичен Уеб**

***Зимен семестър, 2021/2022 год.***

**Тема №8: „Каталог на университети - 2“**

Курсов проект

*Автори:*

*Мария Велева, фак. номер 62445*

*Виктория Месова, фак. номер 62446*

януари, 2022 г.

София

**Съдържание**

[1. Въведение 3](#_Toc92568674)

[1.1. Задание на проекта 3](#_Toc92568675)

[1.2. Актуалност на избраната тема за проект 3](#_Toc92568676)

[1.5. Структура на останалата част от документа 4](#_Toc92568677)

[2. Анализ на решението 4](#_Toc92568678)

[2.1 Работен процес 4](#_Toc92568679)

[2.2 Структура на съдържанието 5](#_Toc92568680)

[2.3 Тип и представяне на съдържанието 7](#_Toc92568681)

[3. Дизайн 7](#_Toc92568682)

[4. Тестване 10](#_Toc92568683)

[5. Заключение и възможно бъдещо развитие 11](#_Toc92568684)

[6. Разпределение на работата 11](#_Toc92568685)

[7. Използвани литературни източници и уеб сайтове 11](#_Toc92568686)

## 1. Въведение

### 1.1. Задание на проекта

Да се създаде каталог на университетите в България по региони, базиран на XML документи с текстово и графично съдържание, описващо възможните характеристики на всеки един университет в страната. Каталогът да представя графичното съдържание чрез XML Entities. Връзките в каталога – напр. между региони и университети, между университети и техните филиали и др. – да се описват чрез (съставни) ключове и референции към тях. Да се състави описанието на 7 – 8 университета и да се валидира чрез подходящо изграден за целта XML Schema документ. Създаденото XML съдържание да се представи в HTML браузър чрез CSS и XSLT, с възможност за сортиране на представените обекти по няколко от техните признаци.

### 1.2. Актуалност на избраната тема за проект

При избор на университет е важно човек да има ясна и подробна информация относно наличните възможности, местоположение и факултети. Вместо да търси информацията за факултетите в сайта на всеки един университет по отделно, би било хубаво тази информация да бъде достъпна добре структурирана на едно място, и съответно да търси подробности само за отговарящите на изискванията университети. Представеният проект се фокусира точно върху този проблем.

**1.3. Проблем, който решава проектът**

Създаденият проект „Каталог на университети - 2“, съдържа списък с всички университети в България. За всеки университет е представена главната информация– име, местоположение, регион, в който се намира, телефонен номер, ако е наличен – втори телефонен номер, година на създаване, сайт на университета, брой студенти, както и списък на факултетите. Каталогът предлага разглеждане на университетите, групирани по региони, както и разделени по филиали. Проектът съдържа информацията за всички университети в България, намиращи се във всичките 6 региона на страната, както и за техните филиали, ако има такива. Всички университети са сортирани в низходящ ред спрямо броя на студентите в тях.

**1.4. Технологии, използвани за решаването на зададения проблем**

Проектът използва XML Entities, за да представи нужната информация за всеки един университет. Връзките между университети и региони, както и между университети и техните филиали, се осъществяват чрез (съставни) ключове и референции към тях. Съдържанието на XML документа е валидирано чрез изградената XML Schema и DTD граматика. За графичното представяне съдържанието в уеб браузър, се използва XSLT документ, който преобразува XML данните. Използвано е CSS стилизиране на съдържанието в уеб браузъра. За сортиранията в низходящ ред по брой студенти, по региони и филиали се използва JavaScript.

### 1.5. Структура на останалата част от документа

В останалата част от документацията на настоящия проект са включени следните точки:

1. Анализ на решението - Описание на входните документи, както и тяхното съдържание.
2. Дизайн - Какви са връзките, чрез които е изграден каталога, както и как е валидиран.
3. Тестване - Тестване на проекта чрез различни браузъри, както и резултатът от тестването.
4. Заключение и възможно бъдещо развитие
5. Разпределение а работата
6. Използвани литературни източници и уеб сайтове

## 2. Анализ на решението

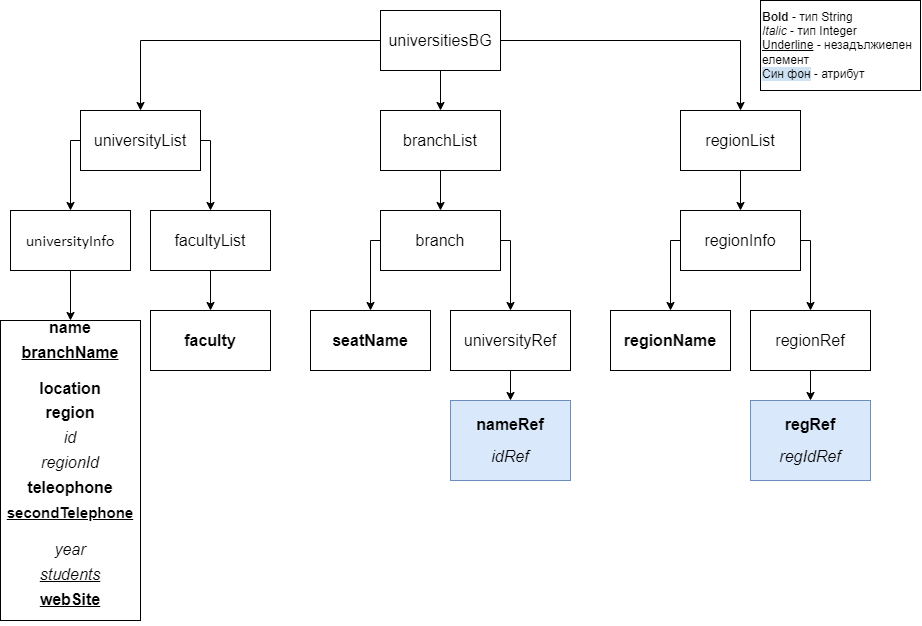
### 2.1 Работен процес

* Входните данни на каталога се съдържа в документа universityCatalogue.xml. Информацията, както текстовото, така и графичното съдържание, са получени от [www.rsvu.mon.bg/](http://www.rsvu.mon.bg/), както и от уеб сайтовете на съответните университети.
* Входното XML съдържание се валидира чрез XML Schema и DTD граматика. За визуализацията в уеб браузър се използва XSLT трансформация, както и CSS стилизиране.
* Създаденият „Каталог на университети - 2“ е лесно достъпен за потребителите в уеб браузър - Mozilla Firefox, Internet Explorer и други.



Фигура 1: Извадка от съдържанието на XML документа

### 2.2 Структура на съдържанието

Каталогът има следната обща структура:

Описание на елементите и атрибутите:

* **universitiesBG** – коренов елемент.
* **universityList** – елемент, който съдържа списък на университетите; може да се съдържа най – много веднъж в universitiesBG.
* **university** – поделемент на universityList, университет, може да се съдържа неограничен брой пъти в universityList.
* **universityInfo** – поделемент на university, информация за конкретния университет.
* **name** – поделемент на universityInfo, име на университета, задължителен елемент от тип string.
* **location** – поделемент на universityInfo, име на населеното място, в което се намира университета, задължителен елемент от тип string.
* **region** – поделемент на universityInfo, име на региона, в който се намира университета, задължителен елемент от тип string.
* **id** – поделемент на universityInfo уникален идентификатор на университета, задължителен елемент от тип integer.
* **regionId** – поделемент на universityInfo, id на региона, задължителен елемент от тип integer.
* **telephone** – поделемент на universityInfo, телефонен номер на университета, задължителен елемент от тип string.
* **secondTelephone** – поделемент на universityInfo, допълнителен телефонен номер на университета, незадължителен елемент от тип string.
* **year** – поделемент на universityInfo, година на основаване на университета, задължителен елемент от тип integer.
* **students** – поделемент на universityInfo, брой студенти в университета, ако не е филиал, незадължителен елемент от тип integer.
* **webSite** – поделемент на universityInfo, сайт на университета, ако има такъв, незадължителен елемент от комплексен тип, който е string и представлява референция към сайта на съответния университет.
* **facultyList** – поделемент на university, списък на всички факултети.
* **faculty** – поделемент на facultyList, име на факултет; съдържа се неограничен брой пъти в facultyList.
* **branchList** – елемент, списък на университетите, които имат филиали; може да се съдържа най-много веднъж.
* **branch** – поделемент на branchList, университет, който има филиали; съдържа се неограничен брой пъти в branchList.
* **seatName** – поделемент на branch, име на седалището на университета, елемент от тип string.
* **universityRef** – празен поделемент на branch, който се съдържа точно веднъж в елемента branch.
* **nameRef** – атрибут на universityRef, който реферира към името на седалището на университета от тип string.
* **idRef** – атрибут на universityRef, който реферира id-то на седалището на университета от тип integer.
* **regionList** – елемент, списък на регионите в България; може да се съдържа най – много веднъж.
* **regionInfo** – поделемент на regionList, име на регион; съдържа се неограничен брой пъти в regionList.
* **regionName** – поделемент на regionList, име на регион, елемент от тип string.
* **regionRef** – празен поделемент на regionList, който се съдържа точно веднъж в елемента regionInfo.
* **regRef** –атрибут на regionRef, който реферира към името на региона от тип string.
* **regIdRef** –атрибут на regionRef, който реферира id-то на дадения регион от тип integer.

### 2.3 Тип и представяне на съдържанието

Съдържанието на проекта се намира в един текстов и 61 графични файла:

* 1 текстов файл – universityCatalogue.xml, който съдържа информацията за университетите, валидиран от XML Schema. Файлът използва UTF-8 кодиране и xml version 1.0. Размер на текстовия файл: 55 KB.
* 61 графични файла – изображения на университетите, получени от техните уеб сайтове. Форматът на изображенията е JPG и всички са преформатирани, за да бъдат с един и същи размер, когато се визуализират в браузър. Общ размер на графичните файлове: 25,7 MG.

## 3. Дизайн

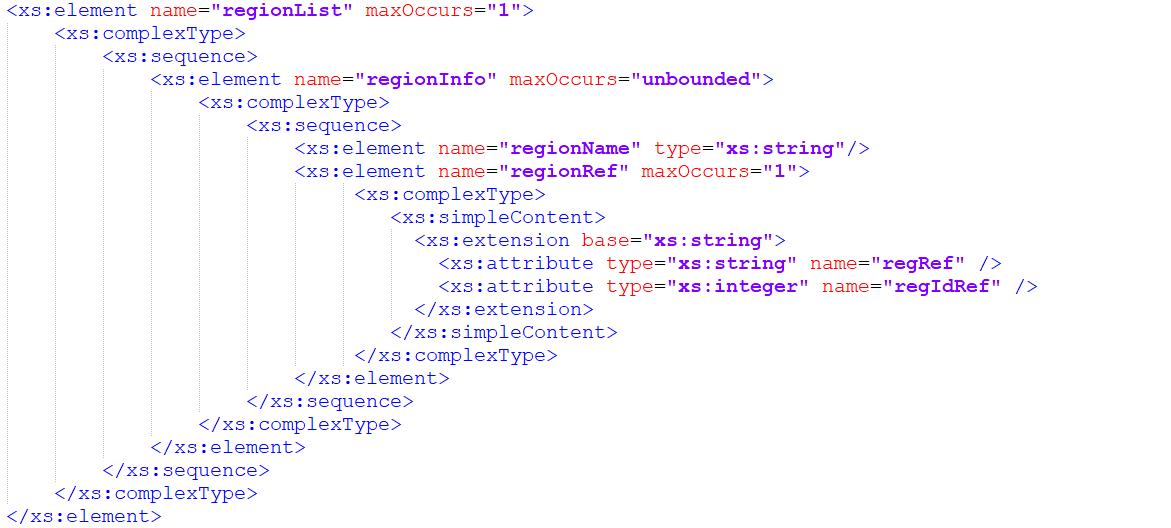
За цялостното решение на проекта са разработени следните файлове:

* **universityCatalogue.xml** – основен файл на проекта. Съдържа информация за университетите като използва XML Entities, както и ключове, чрез които се осъществяват връзките между университети и региони, и университети и техните филиали. Документът е валидиран чрез XML Schema.
* **universityCatalogueScheme.xsd** – XML Schema, която съдържа описание на елементите, както и връзките между тях. Примери с описанието на елемента universityInfo, branchList и regionList, както и референциите по филиали и региони:



Фигура 2: Валидация на елемента universityInfo чрез XML Schema

 Фигура 3: Валидация на елемента branchList чрез XML Schema

Фигура 4: Валидация на елемента regionList чрез XML Schema

Картина, която съдържа текст

Описанието е генерирано автоматично

Фигура 5: Валидация на ключа regionKey и референцията regionRef чрез XML Schema

* **universityCatalogueXslt.xsl** – XSLT файл, чрез който се осъществява трансформацията към HTML, за представяне на съдържанието в уеб браузър.
* **style.css** – css файл, който стилизира съдържанието на проекта.
* **unisInBG.dtd** – dtd граматика, която съдържа Entities за изображенията от каталога.
* **/images** – директория, която съдържа графичните файлове, използвани в проекта.

## 4. Тестване

За тестване на решението на проекта са използвани следните браузъри:

* + Mozilla Firefox
  + Internet Explorer

Получена е следната визуализация в гореспоменатите браузъри:

Graphical user interface

Description automatically generated with low confidence

Фигура 6: Изглед на проекта, тестван с Internet Explorer. Университети

Graphical user interface

Description automatically generated with low confidence

Фигура 7: Изглед на проекта, тестван с Internet Explorer. Региони

Graphical user interface

Description automatically generated with medium confidence

Фигура 8: Изглед на проекта, тестван с Internet Explorer. Филиали

## 5. Заключение и възможно бъдещо развитие

С използването на езика XML, проектът е добре структуриран и връзките между елементите са представени ясно и разбираемо. Чрез XSLT се осъществява лесно преобразуването, за да бъде достъпно съдържанието през уеб браузър. При наличие на промяна, в зависимост от това каква е тя, промени трябва да бъдат извършени на по-малко места, тъй като съдържанието и оформлението са разделени в отделни файлове.

В посока на бъдещото развитие на проекта, могат да бъдат включени всички специалности, които факултетите и филиалите на университетите предлагат. Могат да бъдат включени и рейтинги на университетите и дали те са държавни или частни, както и да се добавят още възможности за сортиране по различни признаци.

## 6. Разпределение на работата

Както документите по проекта – .xml, .xslt, .xsd, .css, .dtd така и документацията, са разработени съвместно от членовете на екипа.

## 7. Използвани литературни източници и уеб сайтове

1. [www.rsvu.mon.bg](http://www.rsvu.mon.bg)
2. [www.uni-sofia.bg](http://www.uni-sofia.bg) - уеб сайт на Софийски университет "Св. Климент Охридски"
3. [www.tu-sofia.bg](http://www.tu-sofia.bg) - уеб сайт на Технически университет - София
4. [www.natfiz.bg](http://www.natfiz.bg) - уеб сайт на Национална академия за театрално и филмово изкуство

"Кръстьо Сарафов"

1. [www.mvr.bg/academy](http://www.mvr.bg/academy) - уеб сайт на Академия на Министерството на вътрешните работи – София
2. [www.vsu.bg](http://www.vsu.bg) – уеб сайт на Висше строително училище "Любен Каравелов" – София
3. [www.vtu.bg](http://www.vtu.bg) – уеб сайт на Висше транспортно училище "Тодор Каблешков" – София
4. [www.vuzf.bg](http://www.vuzf.bg) – уеб сайт на Висше училище по застраховане и финанси – София
5. [www.utp.bg](http://www.utp.bg) – уеб сайт на Висше училище по телекомуникации и пощи – София
6. [www.rndc.bg/](http://www.rndc.bg/) - уеб сайт на Военна академия "Георги Стойков Раковски" – София
7. [www.ltu.bg](http://www.ltu.bg) – уеб сайт на Лесотехнически университет – София
8. [www.mu-sofia.bg](http://www.mu-sofia.bg) – уеб сайт на Медицински университет – София
9. [www.mgu.bg](http://www.mgu.bg) – уеб сайт на Минно-геоложки университет "Св. Иван Рилски"- София
10. [www.nma.bg](http://www.nma.bg) – уеб сайт на Национална музикална академия "Проф. Панчо Владигеров"- София
11. [www.nsa.bg](http://www.nsa.bg) – уеб сайт на Национална спортна академия "Васил Левски"- София
12. [www.nha.bg](http://www.nha.bg) – уеб сайт на Национална художествена академия – София
13. [www.nbu.bg](http://www.nbu.bg) – уеб сайт на Нов български университет – София
14. [www.unwe.bg](http://www.unwe.bg) – уеб сайт на Университет по национално и световно стопанство – София
15. [www.uacg.bg](http://www.uacg.bg) – уеб сайт на Университет по архитектура, строителство и геодезия – София
16. [www.unibit.bg](http://www.unibit.bg) – уеб сайт на Университет по библиотекознание и информационни технологии – София
17. [www.uctm.edu](http://www.uctm.edu) – уеб сайт на Химикотехнологичен и металургичен университет – София
18. [www.aubg.edu](http://www.aubg.edu) – уеб сайт на Американски университет в България
19. [www.swu.edu](http://www.swu.edu) – уеб сайт на Югозападен университет "Неофит Рилски" – Благоевград
20. [www.bfu.bg](http://www.bfu.bg) – уеб сайт на Бургаски свободен университет
21. [www.vfu.bg](http://www.vfu.bg) – уеб сайт на Варненски свободен университет "Черноризец Храбър"
22. [www.naval-acad.bg](http://www.naval-acad.bg) – уеб сайт на Висше военноморско училище "Никола Вапцаров"– Варна
23. [www.vum.bg](http://www.vum.bg) – уеб сайт на Висше училище по мениджмънт – Варна
24. [www.ue-varna.bg](http://www.ue-varna.bg) – уеб сайт на Икономически университет – Варна
25. [www.mu-varna.bg](http://www.mu-varna.bg) – уеб сайт на Медицински университет "Проф. д-р Параскев Стоянов" – Варна
26. [www.tu-varna.bg](http://www.tu-varna.bg) – уеб сайт на Технически университет – Варна
27. [www.uni-vt.bg](http://www.uni-vt.bg) – уеб сайт на Великотърновски университет "Св. св. Кирил и Методий"
28. [www.nvu.bg](http://www.nvu.bg) – уеб сайт на Национален военен университет "Васил Левски" - Велико Търново
29. [www.togap.bg](http://www.togap.bg) – уеб сайт на Технически университет – Габрово
30. [www.af-acad.bg](http://www.af-acad.bg) – уеб сайт на Висше военновъздушно училище "Георги Бенковски"- Долна Митрополия
31. [www.epubg.eu](http://www.epubg.eu) – уеб сайт на Европейски политехнически университет – Перник
32. [www.mu-pleven.bg](http://www.mu-pleven.bg) – уеб сайт на Медицински университет – Плевен
33. [www.uni-plovdiv.bg](http://www.uni-plovdiv.bg) – уеб сайт на Пловдивски университет "Паисий Хилендарски"
34. [www.au-plovdiv.bg](http://www.au-plovdiv.bg) – уеб сайт на Аграрен университет – Пловдив
35. [www.artacademyplovdiv.com](http://www.artacademyplovdiv.com) – уеб сайт на Академия за музикално, танцово и изобразително изкуство – Пловдив
36. [www.uard.bg](http://www.uard.bg) – уеб сайт на Висше училище по агробизнес и развитие на регионите – Пловдив
37. [www.vusi.bg](http://www.vusi.bg) – уеб сайт на Висше училище по сигурност и икономика – Пловдив
38. [www.mu-plovdiv.bg](http://www.mu-plovdiv.bg)– уеб сайт на Медицински университет – Пловдив
39. [www.uft-plovidiv.bg](http://www.uft-plovidiv.bg) - уеб сайт на Университет по хранителни технологии – Пловдив
40. [www.uni-ruse.bg](http://www.uni-ruse.bg) - уеб сайт на Русенски университет "Ангел Кънчев"
41. [www.uni-svishtov.bg](http://www.uni-svishtov.bg) - уеб сайт на Стопанска академия "Димитър Ценов" – Свищов
42. [www.uni-sz.bg](http://www.uni-sz.bg) - уеб сайт на Тракийски университет - Стара Загора
43. [www.shu.bg](http://www.shu.bg) – уеб сайт на Шуменски университет "Епископ Константин Преславски"
44. [www.w3school.com](http://www.w3school.com)
45. Лекции по “XML технологии за семантичен Уеб“, зимен семестър 2021/2022,

проф. д-р Боян Бончев