



Факултет по математика и информатика

Бакалавърска програма
„Софтуерно инженерство“

Предмет: XML технологии за семантичен Уеб
Зимен семестър, 2022/2023 год.

Тема №09: „КАТАЛОГ НА IT ФИРМИ - 1“

КУРСОВ ПРОЕКТ

Автори:

Моника Красиминова Митрева, фак. номер 62522
Кристиан Филипов Иванов, фак. номер 0MI0600137

9 януари 2023 г.
гр. София

Съдържание

1. Въведение	3
1.1. Актуалност на проекта	3
1.2. Проблем, който решава проектът	3
1.3. Технологии, използвани за реализация на проекта	4
1.4. Структура на останалата част от документа.....	4
2. Анализ на решението	4
2.1. Работен процес	4
2.2. Структура на съдържанието	9
2.3. Тип и представяне на съдържанието	12
3. Дизайн	13
4. Тестване	20
5. Заключение и възможно бъдещо развитие.....	23
6. Разпределение на работата	24
7. Използвани литературни източници и Уеб сайтове	24

1. Въведение

Настоящият документ представлява подробна документация, описваща курсов проект на тема „Каталог на IT фирми - 1“ за предмета „XML технологии за семантичен Уеб“.

Реализираното съвместно решение е изградено възможно най-оптимално спрямо изискванията на представеното задание на проекта. Важни уточнения, породени от условията на заданието, които предопределиха избора на конкретни IT фирми за разработката на текущия проект, са:

- › Разгледаните IT фирми са разположени в четири региона на територията на Република България – София, Пловдив, Варна, Бургас.
- › За всяка отделна област са представени по две IT фирми.
- › Подбрани са само софтуерни компании, които са разделени в две категории, базирани на филтър ‚дейност‘ – продуктови и service. Основната причина за така направената селекция на фирмите е оскъдната и/или изцяло липсваща информация за други производители в IT сектора.

1.1. Актуалност на проекта

През последните години се наблюдава бурно развитие на IT сектора както в световен мащаб, така и у нас. Въпреки кризисната обстановка, породена от пандемията COVID-19, поради високото ниво на технологичност и естество на работа, компаниите от сектора бяха едни от най-бързо адаптивните към новата ситуация в страната. Данни от проучване на Българската асоциация на софтуерните компании, сочат, че противно на пандемичните обстоятелства, IT секторът в България нараства¹. Именно поради тази причина, чрез настоящия проект, разработен под формата на каталог, заинтересованите от темата лица ще могат да получат бърз и централизиран достъп до необходимата им информация, както и условията, предлагани от разгледаните фирми.

1.2. Проблем, който решава проектът

Основната цел на проекта е да се създаде систематизирано съдържание на информацията за някои от IT фирмите в страната под формата на каталог. Всяка една компания се характеризира с основна прилежаща информация – име, актуална снимка, година на основаване, дейност, кратка информация под формата на визитка, използвани технологии и контакти. Предимство на така представените данни е, че каталогът може да служи за справка, както на бъдещи и настоящи студенти, търсещи стаж и/или редовна работа в IT сектора, така и за популяризиране на дейността/продуктите на избраните фирми, посредством наличните PDF визуализации, които да се разпространяват под формата на флаери, достигайки таргетираната аудитория.

¹ [BACOM: IT секторът в България расте с 10% през 2020 г. въпреки коронавируса \(capital.bg\)](https://capital.bg/)

1.3. Технологии, използвани за реализацията на проекта

Реализацията на проекта се базира на езикът XML. За представяне на графичното съдържание в XML документа са използвани ENTITIES (единици). Връзките между регионите и фирмите, и тези между фирмите и дейността, която извършват, се осъществяват посредством използването на ID/IDREF ((съставни) ключове и референции към тях). Валидацията на създаденият XML документ се контролира от уникален външен DTD документ. Цялото генерирано XML съдържание може да се представи в PDF формат чрез XSL документ.

1.4. Структура на останалата част от документа

В останалата част от настоящия документ последователно се разглеждат:

- › Анализ на решението
- › Дизайн
- › Тестване
- › Заключение и възможно бъдещо развитие
- › Разпределение на работата по курсовия проект
- › Използвани литературни източници и уеб сайтове

2. Анализ на решението

2.1. Работен процес

- › Документът **IT_companies.xml** съхранява входните данни на каталога. Използваните данни под формата на текстово и графично съдържание е извлечена от сайта **DEV.BG**², който предоставя широкоспектърна информация за IT работодателите в България.
- › DTD документът, именуван **IT_companies.dtd**, контролира валидността на предоставеното входно съдържание.
- › XSL документът **IT_companies.xsl** предоставя възможността за генерация на текстов файл във формат pdf. Благодарение на него, крайният потребител получава достъп до информацията в лесно четим и систематизиран за него вид, стилизирана в един PDF документ. Изходният текстови файл е лесен за печат и разпространение, което го превръща в универсално средство с цел справка и/или реклама.

² [Компании – ИТ обяви за работа, придобивки, екип и мнения \(dev.bg\)](http://dev.bg)

Фиг.1. Извадка от документа **IT_companies.xml**

```

1  <?xml version = "1.0" encoding = "UTF-8"?>
2  <?xml-stylesheet type = "text/xsl" href = "IT_companies.xsl"?>
3
4  <!DOCTYPE catalogue SYSTEM "IT_companies.dtd">
5
6  <catalogue>
7      <!-- Regions -->
8      <regions>
9          <region regionID = "B" companyRef = "PaysafeGroup Adastra">
10             <regionName> Варна </regionName>
11          </region>
12          <region regionID = "C" companyRef = "AsteaSolutions Melon">
13             <regionName> София </regionName>
14          </region>
15          <region regionID = "PB" companyRef = "DraftKings Strypes">
16             <regionName> Пловдив </regionName>
17          </region>
18          <region regionID = "A" companyRef = "SigmaSoftware SoftServe">
19             <regionName> Бургас </regionName>
20          </region>
21      </regions>
22
23      <!-- Activities -->
24      <activities>
25          <activity activityID = "product" companyRef = "PaysafeGroup AsteaSolutions DraftKings SigmaSoftware">
26             <activityName> Продуктови компании </activityName>
27          </activity>
28          <activity activityID = "service" companyRef = "Adastra Melon Strypes SoftServe">
29             <activityName> Service компании </activityName>
30          </activity>
31      </activities>
32
33      <!-- Companies -->
34      <companies>
35          <!-- PaysafeGroup -->
36          <company companyID = "PaysafeGroup" regionRef = "B" activityRef = "product">
37             <companyName> Paysafe Group </companyName>
38             <image href = "paysafeGroup"/>
39             <yearOfEstablishment> 1996 г. </yearOfEstablishment>
40             <description>
41                 Paysafe Limited ("Paysafe") (NYSE: PSFE) (PSFE.WS) е водещ глобален доставчик на цялостни плат
42                 позволи на бизнеса и потребителите да се свързват и осъществяват трансакции безпроблемно, като
43                 на разплащания, дигитални портфейли и онлайн парични решения. Наетите в България над 1400 служ
44                 в глобален мащаб. Екип от софтуерни инженери в България развиват едни от ключовите продуктови
45                 ползвани от милиони крайни потребители и бизнеси, с което компанията следва ангажмента си да

```

Фиг.2. Извадка от документа **IT_companies.dtd**

```

1 <!ELEMENT catalogue (regions, activities, companies)>
2
3 <!ELEMENT regions (region+)>
4 <!ELEMENT region (regionName)>
5 <!--ATTLIST region regionID ID #REQUIRED-->
6 <!--ATTLIST region companyRef IDREFS #IMPLIED-->
7 <!--ELEMENT regionName (#PCDATA)-->
8
9 <!--ELEMENT activities (activity+)>
10 <!--ELEMENT activity (activityName)>
11 <!--ATTLIST activity activityID ID #REQUIRED-->
12 <!--ATTLIST activity companyRef IDREFS #IMPLIED-->
13 <!--ELEMENT activityName (#PCDATA)-->
14
15 <!--ELEMENT companies (company+)>
16 <!--ELEMENT company (companyName, image, yearOfEstablishment, description, technologies, contacts)>
17 <!--ATTLIST company companyID ID #REQUIRED-->
18 <!--ATTLIST company regionRef IDREF #IMPLIED-->
19 <!--ATTLIST company activityRef IDREF #IMPLIED-->
20 <!--ELEMENT companyName (#PCDATA)-->
21 <!--ELEMENT image EMPTY-->
22 <!--ATTLIST image href ENTITY #REQUIRED-->
23 <!--ELEMENT yearOfEstablishment (#PCDATA)-->
24 <!--ELEMENT description (#PCDATA)-->
25 <!--ELEMENT technologies (#PCDATA)-->
26
27 <!--ELEMENT contacts (address, email?, site?, facebook?, linkedin?)>
28 <!--ELEMENT address (city, street?)>
29 <!--ELEMENT city (#PCDATA)-->
30 <!--ELEMENT street (#PCDATA)-->
31 <!--ELEMENT email (#PCDATA)-->
32 <!--ELEMENT site (#PCDATA)-->
33 <!--ELEMENT facebook (#PCDATA)-->
34 <!--ELEMENT linkedin (#PCDATA)-->
35
36 <!--NOTATION JPEG SYSTEM "IMAGES/jpg"-->
37
38 <!--ENTITY adastra SYSTEM "IMAGES/adastra.jpg" NDATA JPEG-->
39 <!--ENTITY asteaSolutions SYSTEM "IMAGES/asteaSolutions.jpg" NDATA JPEG-->
40 <!--ENTITY draftKings SYSTEM "IMAGES/draftKings.jpg" NDATA JPEG-->
41 <!--ENTITY melon SYSTEM "IMAGES/melon.jpg" NDATA JPEG-->
42 <!--ENTITY paysafeGroup SYSTEM "IMAGES/paysafeGroup.jpg" NDATA JPEG-->
43 <!--ENTITY sigmaSoftware SYSTEM "IMAGES/sigmaSoftware.jpg" NDATA JPEG-->
44 <!--ENTITY softServe SYSTEM "IMAGES/softServe.jpg" NDATA JPEG-->
45 <!--ENTITY strypes SYSTEM "IMAGES/strypes.jpg" NDATA JPEG-->

```

Фиг.3. Извадка от документа **IT_companies.xsl**

```

1 <?xml version = "1.0" encoding = "UTF-8"?>
2 <xsl:stylesheet version = "1.0"
3   xmlns:xsl = "http://www.w3.org/1999/XSL/Transform"
4   xmlns:fo = "http://www.w3.org/1999/XSL/Format">
5   <xsl:output method = "xml" indent = "yes"/>
6
7   <xsl:template match = "/">
8     <fo:root>
9       <fo:layout-master-set>
10        <fo:simple-page-master page-height = "297mm" page-width = "210mm" margin = "5mm 20mm 5mm 20mm" master-name = "template">
11          <fo:region-body margin = "20mm 0mm 20mm 0mm"/>
12        </fo:simple-page-master>
13      </fo:layout-master-set>
14
15      <!-- Project title page -->
16      <fo:page-sequence master-reference = "template">
17        <fo:flow flow-name = "xsl-region-body">
18          <fo:block position = "absolute" text-align = "right" font-size = "40pt"
19            margin-bottom = "30mm" font-family = "Impact" padding-before = "30mm">
20            Тема №09:<fo:block/>КАТАЛОГ НА ИТ #ИПМИ - 1
21          </fo:block>
22          <fo:block position = "absolute" font-family = "Franklin Gothic Book">
23            <fo:block margin-bottom = "5mm" text-align = "center" font-size = "24pt" font-style = "italic">Курсов проект</fo:block>
24            <fo:block text-align = "center" font-size = "16pt">XML технологии за семантичен уеб, зимен семестър 2022/2023 г.</fo:block>
25            <fo:block text-align = "justify" padding-before = "30mm" font-size = "14pt">
26              <fo:table>
27                <fo:table-column column-width="35mm"/>
28                <fo:table-column column-width="150mm"/>
29                <fo:table-body>
30                  <fo:table-row>
31                    <fo:table-cell>
32                      <fo:block font-style = "italic">Изготвен от:</fo:block>
33                    </fo:table-cell>
34                    <fo:table-cell>
35                      <fo:block>Моника Красивярова Митрева, фак. номер 62522</fo:block>
36                      <fo:block margin-bottom = "10mm">Кристиан Филипов Иванов, фак. номер 0MI0600137</fo:block>
37                    </fo:table-cell>
38                  </fo:table-row>
39                  <fo:table-row>
40                    <fo:table-cell>
41                      <fo:block font-style = "italic">Проверен от:</fo:block>
42                    </fo:table-cell>
43                    <fo:table-cell>
44                      <fo:block margin-bottom = "30mm">гл. ас. д-р Явор Данков</fo:block>
45                    </fo:table-cell>
46                  </fo:table-row>
47                </fo:table-body>
48              </fo:table>
49            </fo:block>
50            <fo:block text-align = "center">09 януари 2023 г.</fo:block>

```

Фиг.4. Извадка от генерираният от инструмента OXYGEN XML EDITOR³ чрез XSL-FO трансформация документ **IT_companies.pdf**

Paysafe Group

Plug into **Paysafe**

Визитна картичка:

Paysafe Limited ("Paysafe") (NYSE: PSFE) (PSFE.WS) е водещ глобален доставчик на цялостни платежни решения. Основната цел на компанията е да позволи на бизнеса и потребителите да се свързват и осъществяват трансакции безпроблемно, като използват водещи за индустрията възможности в обработването на разплащания, дигитални портфейли и онлайн парични решения. Наетите в България над 1400 служители представляват над една трета от всички работещи в Paysafe Group. Това превръща софийския офис в най-големия за компанията в глобален мащаб. Екип от софтуерни инженери в България развиват едни от ключовите продуктови линии на компанията – дигиталните портфейли Skrill и NETELLER – ползвани от милиони крайни потребители и бизнеси, с което компанията следва ангажимента си да предоставя решения, ориентирани към клиента.

Технологии / програмни езици, които се използват в компанията:

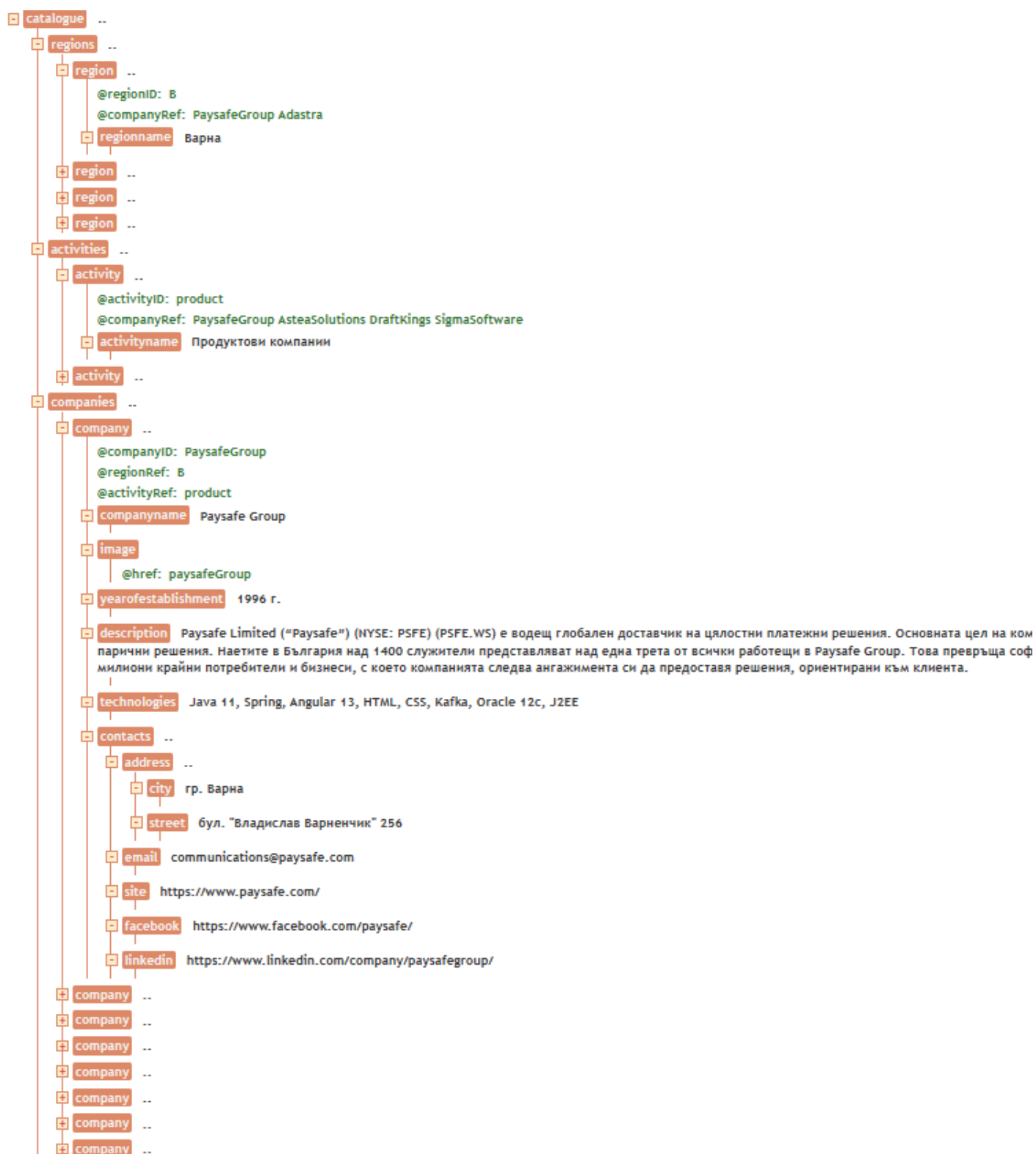
Java 11, Spring, Angular 13, HTML, CSS, Kafka, Oracle 12c, J2EE

Година на основаване:	1996 г.
Адрес:	гр. Варна
Имейл:	communications@paysafe.com
Официален уебсайт:	https://www.paysafe.com/
Facebook:	https://www.facebook.com/paysafe/
LinkedIn:	https://www.linkedin.com/company/paysafegroup/

³ [Oxygen XML Editor](#)

2.2. Структура на съдържанието

Фиг.5. Дървовидно представяне на входното XML съдържание посредством Code Beautify⁴



⁴ [Best Online XML Viewer, XML Formatter, XML Editor, Analyser, Beautify-Beautifier, Minify, Tree structure \(codebeautify.org\)](https://codebeautify.org/)

На база йерархията, представена на **Фиг.5.**, каталогът има следната структура на съдържанието:

- › **catalogue** – коренов елемент на документа **IT_companies.xml**. Съдържа три поделементи – региони (regions), дейности (activities) и фирми (companies).
 - › **regions** – списък от регионите в България. Съдържа се точно веднъж в документа **IT_companies.xml**. Има единствен поделемент – регион (region). Състои от един или повече елемента region.
 - › **region** – индивидуален регион. Съдържа се един или повече пъти в документа **IT_companies.xml**. Има единствен поделемент - име на региона (regionName) и два атрибута – regionID и companyRef.
 - *regionID* – атрибут с регионален идентификатор. Представлява задължителен атрибут от тип ID, който ще се реферира от фирмите, намиращи се в конкретната територия.
 - *companyRef* – атрибут с имена на фирмите. Представлява незадължителен атрибут от тип IDREF, свързващ фирмите с регионите.
 - › **regionName** – име на региона. Представлява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент region.
 - › **activities** – списък от дейностите, извършвани от фирмите. Съдържа се точно веднъж в документа **IT_companies.xml**. Има единствен поделемент – дейност (activity). Състои от един или повече елемента activity.
 - › **activity** – индивидуална дейност. Съдържа се един или повече пъти в документа **IT_companies.xml**. Има единствен поделемент - име на дейността (activityName) и два атрибута – activityID и companyRef.
 - *activityID* – атрибут с идентификатор на дейността. Представлява задължителен атрибут от тип ID, който ще се реферира от фирмите, извършващи конкретната дейност.
 - *companyRef* – атрибут с имена на фирмите. Представлява незадължителен атрибут от тип IDREF, свързващ фирмите с дейността, която извършват.
 - › **activityName** – име на дейността. Представлява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент activity.
 - › **companies** – списък от фирмите. Съдържа се точно веднъж в документа **IT_companies.xml**. Има единствен поделемент – фирма (company). Състои от един или повече елемента company.
 - › **company** – индивидуална фирма. Съдържа се един или повече пъти в документа **IT_companies.xml**. Притежава няколко поделементи - име на компанията (companyName), снимка (image), година на основаване (yearOfEstablishment), кратко описание (description), използвани от фирмата езици за програмиране/технологии (technologies), контакти (contacts) и три атрибута – companyID, regionRef и activityRef.

- *companyID* - атрибут с идентификатор на фирмата. Представява задължителен атрибут от тип ID, който ще се реферира от регионите, в които се намира компанията, както и от дейността, която тя извършва.
- *regionRef* – представлява незадължителен атрибут от тип IDREF, рефериращ към идентификатора на региона, в който се намира фирмата.
- *activityRef* – представлява незадължителен атрибут от тип IDREF, рефериращ към идентификатора на дейността, който се извършва от фирмата.
 - › **companyName** - име на фирмата. Представява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент company.
 - › **image** – снимка (лого/мото на фирмата). Съдържа се точно веднъж във всеки елемент company. Представява елемент от тип EMPTY, добавен в документа посредством XML entities. Има единствен атрибут – href.
 - › **href** - задължителен атрибут, осъществяващ връзка с елемент, който не е в XML формат.
 - › **yearOfEstablishment** – година на основаване. Представява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент company.
 - › **description** – кратко описание на фирмата под формата на визитка. Представява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент company.
 - › **technologies** - използвани от фирмата езици за програмиране/технологии. Представява елемент от тип #PCDATA. Съдържа се точно веднъж във всеки елемент company.
 - › **contacts** – списък с контакти на фирмите. Съдържа пет поделементи – адрес (address), имейл (email), официален уебсайт (site), Facebook (facebook), Linkedin (linkedin).
 - › **address** – адрес на фирмата. Представява елемент от тип #PCDATA. Съдържа се най-много веднъж за всеки елемент contacts. Има два поделементи – city и street.
 - › **city** – градът, в който е разположен офис на фирмата. Представява елемент от тип #PCDATA.
 - › **street** – улицата, на която се намира офиса. Представява елемент от тип #PCDATA.

- › **email** – имейл на фирмата. Представява елемент от тип #PCDATA. Съдържа се най-много веднъж за всеки елемент contacts.
- › **site** – официален уебсайт на фирмата. Представява елемент от тип #PCDATA. Съдържа се веднъж за всеки елемент contacts.
- › **facebook** – официална Facebook страница на фирмата. Представява елемент от тип #PCDATA. Съдържа се най-много веднъж за всеки елемент contacts.
- › **linkedin** – официален LinkedIn профил на фирмата. Представява елемент от тип #PCDATA. Съдържа се най-много веднъж за всеки елемент contacts.

2.3. Тип и представяне на съдържанието

Съдържанието на проекта е разположено в:

- › **2 текстови файла:**
 - › **IT_comapnies.xml** – основният файл, съдържащ цялата входна информация за фирмите, валидиран чрез IT_companies.dtd документа. IT_comapnies.xml използва xml version 1.0 и UTF-8 кодиране.
Размер: 17 KB
 - › **IT_comapnies.xsl** – XSLT документ, съдържащ в себе си подробно описание на XSL трансформациите, които трябва да се приложат върху файла IT_comapnies.xml, за да се генерира XSL-FO документ, посредством който да се възпроизведе крайния PDF файл. IT_comapnies.xsl използва xml version 1.0 и UTF-8 кодиране.
Размер: 41 KB
- › **8 графични файла** – изображения (лого/мото) на фирмите във формат .jpg:

› adastra.jpg – размер: 113 KB	› paysafeGroup.jpg – размер: 24,8 KB
› asteaSolutions.jpg – размер: 42,2 KB	› sigmaSoftware.jpg – размер: 133 KB
› draftKings.jpg – размер: 410 KB	› softServe.jpg – размер: 5,67 KB
› melon.jpg – размер: 109 KB	› strypes.jpg – размер: 119 KB
- › **1 мултимедиен файл:**
 - › **Generate_IT_companies_pdf.mp4** – файл, в който е демонстрирана генерацията на документа IT_companies.pdf посредством инструмента OXYGEN XML EDITOR.
Размер: 20 099 KB

3. Дизайн

За цялостното решение на поставеното задание са разработени файловете:

- › **IT_companies.xml** – основният файл, съдържащ входните данни на проекта – подробна информация за 8 IT фирми, заедно с тяхното местоположение и извършваната от тях дейност. Чрез ID/IDREFS са представени връзките между всяка една фирма (company) и локацията на нейният офис (т.е. region), както и тези между всяка фирма (company) и дейността, която извършва (т.е. activity).
- › **IT_companies.dtd** – уникална DTD граматика, която контролира валидацията на IT_companies.xml. Съдържа описание на елементите и атрибутите, връзките между тях и типа на тяхното съдържание. В този DTD документ са дефинирани XML entities, посредством които се осъществява добавянето на графичната информация.
- › **IT_companies.xsl** – XSLT документ, съдържащ XSL трансформации които трябва да се приложат върху файла IT_companies.xml, за да се генерира XSL-FO документ, посредством който да се възпроизведе крайния PDF файл. За създаването на този PDF документ е използван инструментът OXYGEN XML EDITOR.
- › **IT_companies.pdf** – PDF документ, генериран от инструмента OXYGEN XML EDITOR в резултат на успешна XSL-FO трансформация върху документа IT_companies.xml.
- › **/IMAGES** – допълнителна папка в директорията с гореспоменатите файлове на проекта, която съхранява само графичната информация, използвана за разработката на заданието.

Фиг.6. Извадка от документа **IT_companies.xml**, на която се виждат релациите между регионите, фирмите и дейностите

```
1 <?xml version = "1.0" encoding = "UTF-8"?>
2 <?xml-stylesheet type = "text/xsl" href = "IT_companies.xsl"?>
3
4 <!DOCTYPE catalogue SYSTEM "IT_companies.dtd">
5
6 <catalogue>
7   <!-- Regions -->
8   <regions>
9     <region regionID = "B" companyRef = "PaysafeGroup Adastra">
10       <regionName> Бърна </regionName>
11     </region>
12     <region regionID = "C" companyRef = "AsteaSolutions Melon">
13       <regionName> София </regionName>
14     </region>
15     <region regionID = "PB" companyRef = "DraftKings Strypes">
16       <regionName> Пловдив </regionName>
17     </region>
18     <region regionID = "A" companyRef = "SigmaSoftware SoftServe">
19       <regionName> Бургас </regionName>
20     </region>
21   </regions>
22
23   <!-- Activities -->
24   <activities>
25     <activity activityID = "product" companyRef = "PaysafeGroup AsteaSolutions DraftKings SigmaSoftware">
26       <activityName> Продуктови компании </activityName>
27     </activity>
28     <activity activityID = "service" companyRef = "Adastra Melon Strypes SoftServe">
29       <activityName> Service компании </activityName>
30     </activity>
31   </activities>
32
33   <!-- Companies -->
34   <companies>
35     <!-- PaysafeGroup -->
36     <company companyID = "PaysafeGroup" regionRef = "B" activityRef = "product">
```

Фиг.7. Извадка от документа **IT_companies.dtd**, на която се виждат прилежащите на елемента **companies** елементи и атрибути, типа на тяхното съдържание, честотата им на срещане и с какъв характер са – задължителен (#REQUIRED) или не (#IMPLIED)

```

15 <!ELEMENT companies (company+)>
16 <!ELEMENT company (companyName, image, yearOfEstablishment, description, technologies, contacts)>
17 <!ATTLIST company companyID ID #REQUIRED>
18 <!ATTLIST company regionRef IDREF #IMPLIED>
19 <!ATTLIST company activityRef IDREF #IMPLIED>
20 <!ELEMENT companyName (#PCDATA)>
21 <!ELEMENT image EMPTY>
22 <!ATTLIST image href ENTITY #REQUIRED>
23 <!ELEMENT yearOfEstablishment (#PCDATA)>
24 <!ELEMENT description (#PCDATA)>
25 <!ELEMENT technologies (#PCDATA)>
26 <!ELEMENT contacts (address?, email?, site, facebook?, linkedin?)>
27 <!ELEMENT address (city, street)>
28 <!ELEMENT city (#PCDATA)>
29 <!ELEMENT street (#PCDATA)>
30 <!ELEMENT email (#PCDATA)>
31 <!ELEMENT site (#PCDATA)>
32 <!ELEMENT facebook (#PCDATA)>
33 <!ELEMENT linkedin (#PCDATA)>
..

```

Фиг.8. Извадка от документа **IT_companies.dtd**, на която се вижда как са използвани XML ENTITIES за добавяне на графично съдържание

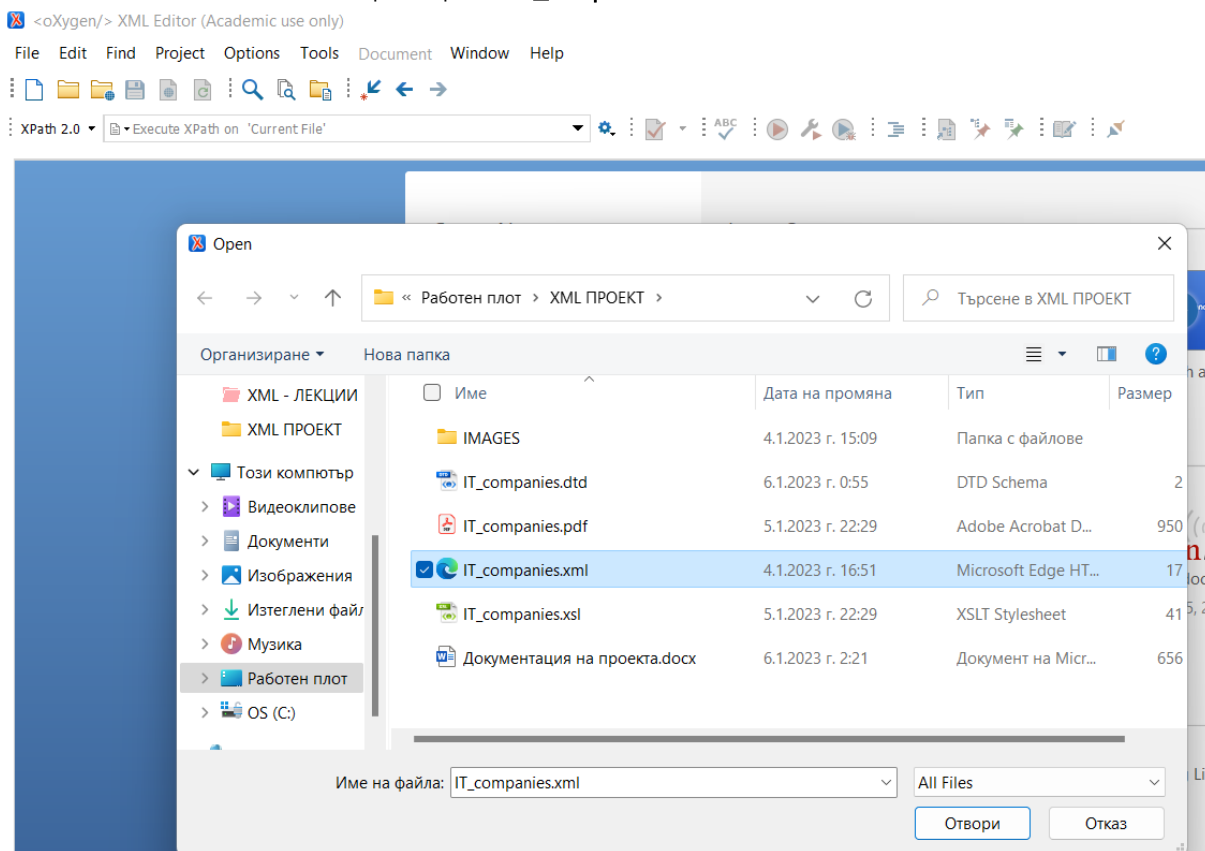
```

34
35 <!NOTATION JPEG SYSTEM "IMAGES/jpg">
36
37 <!ENTITY adastra SYSTEM "IMAGES/adastra.jpg" NDATA JPEG>
38 <!ENTITY asteaSolutions SYSTEM "IMAGES/asteaSolutions.jpg" NDATA JPEG>
39 <!ENTITY draftKings SYSTEM "IMAGES/draftKings.jpg" NDATA JPEG>
40 <!ENTITY melon SYSTEM "IMAGES/melon.jpg" NDATA JPEG>
41 <!ENTITY paysafeGroup SYSTEM "IMAGES/paysafeGroup.jpg" NDATA JPEG>
42 <!ENTITY sigmaSoftware SYSTEM "IMAGES/sigmaSoftware.jpg" NDATA JPEG>
43 <!ENTITY softServe SYSTEM "IMAGES/softServe.jpg" NDATA JPEG>
44 <!ENTITY strypes SYSTEM "IMAGES/strypes.jpg" NDATA JPEG>
45

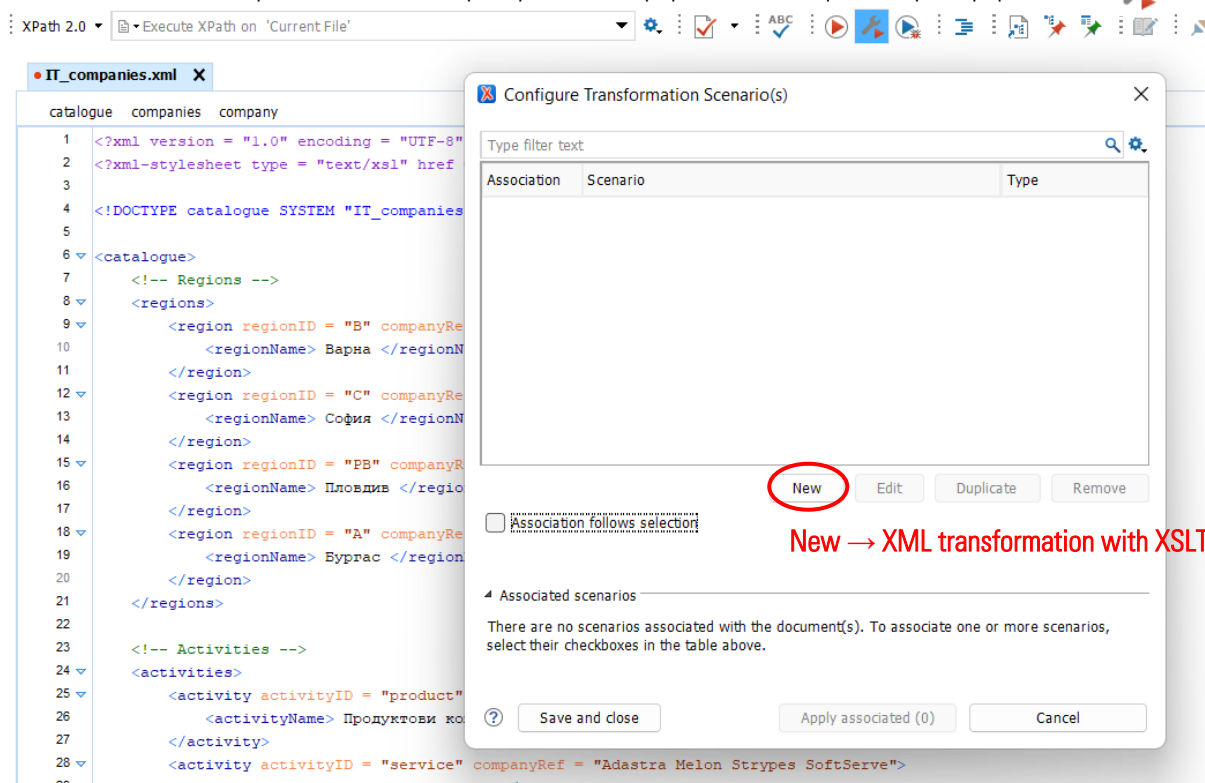
```

Фиг.9. Представяне на процеса по генериране на файла IT_companies.pdf чрез OXYGEN XML EDITOR

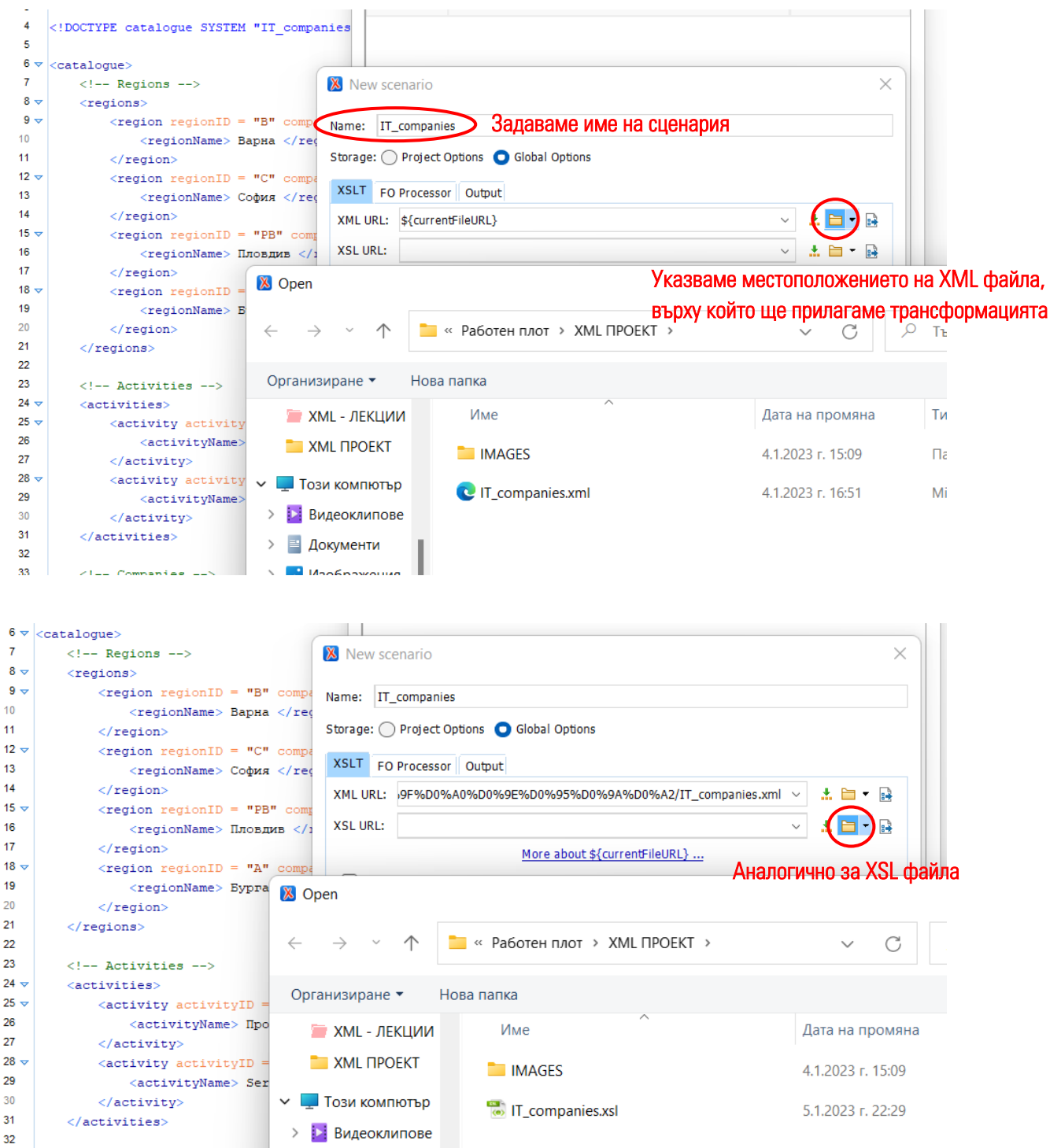
Фиг.9.1. Отваряме файла IT_companies.xml във OXYGEN XML EDITOR

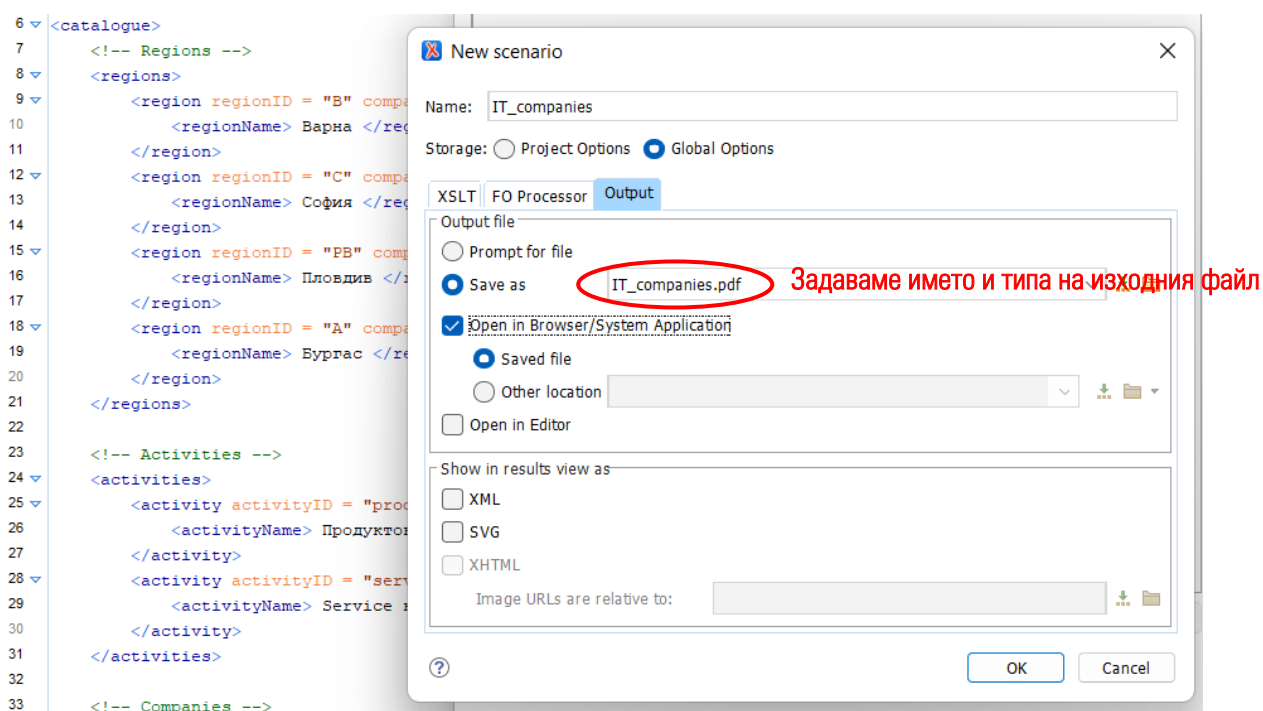
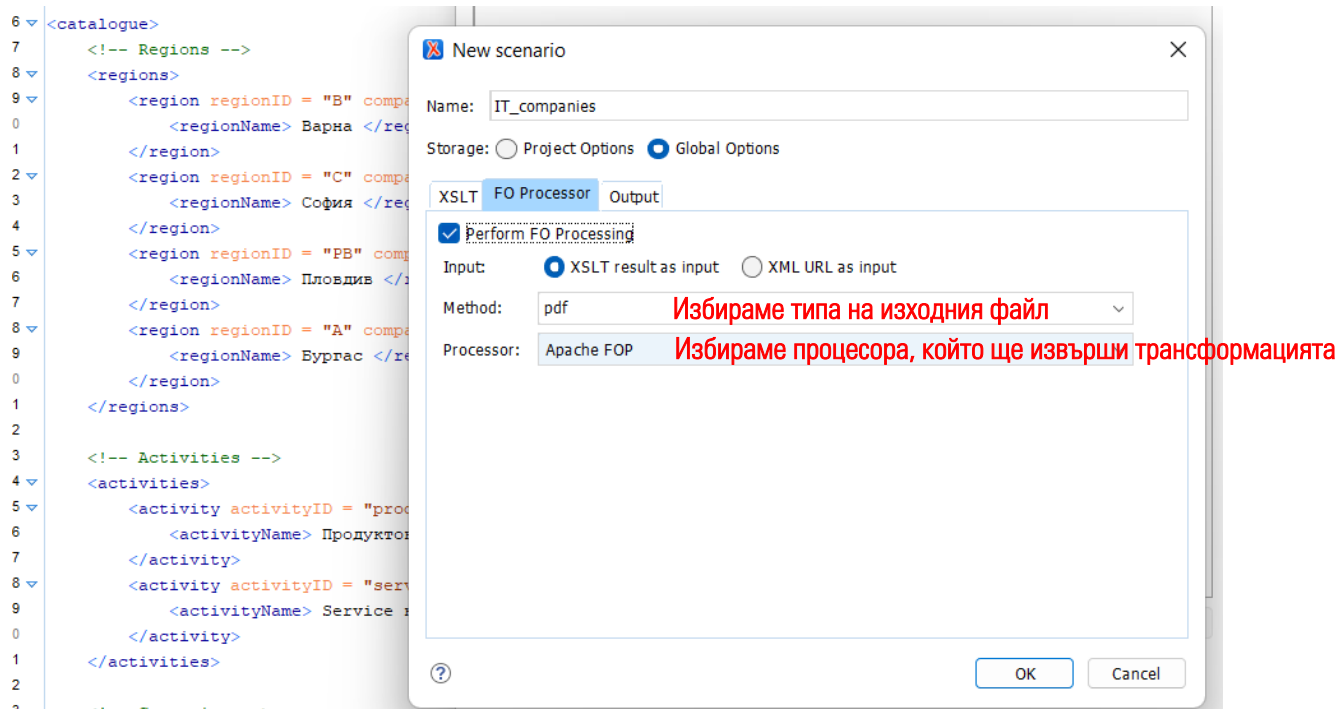


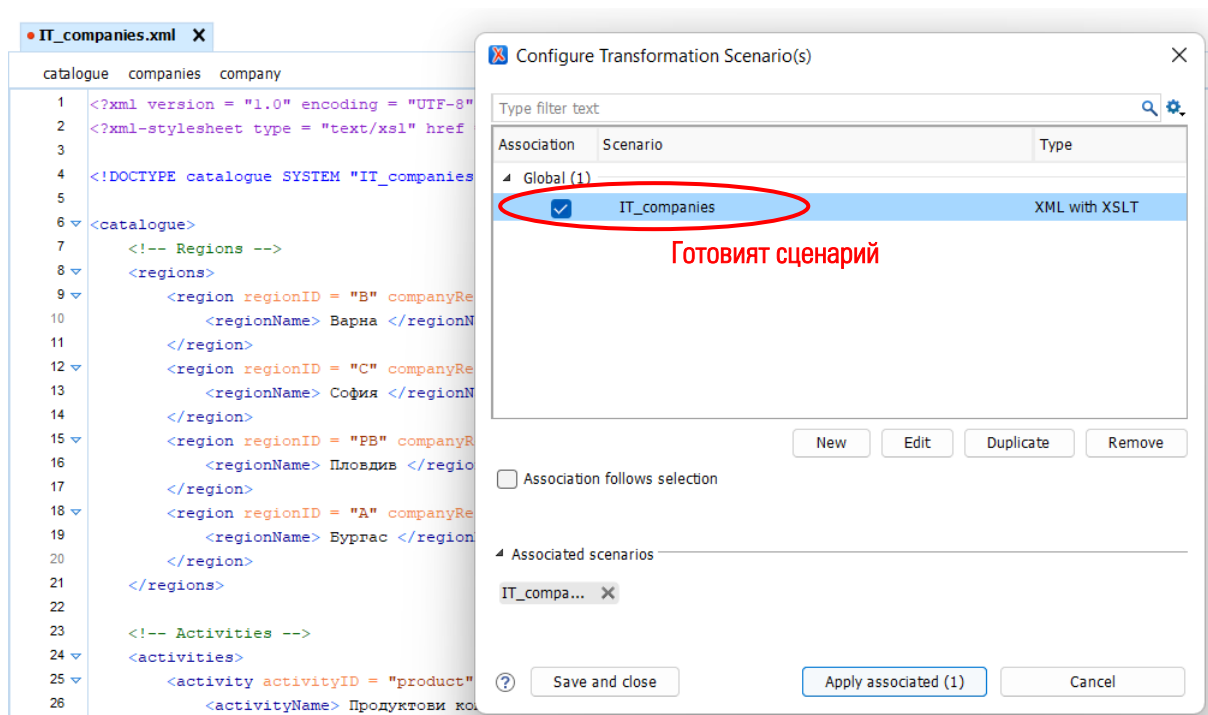
Фиг.9.2. Отваряме диалоговия прозорец за генериране на сценарии за трансформация от



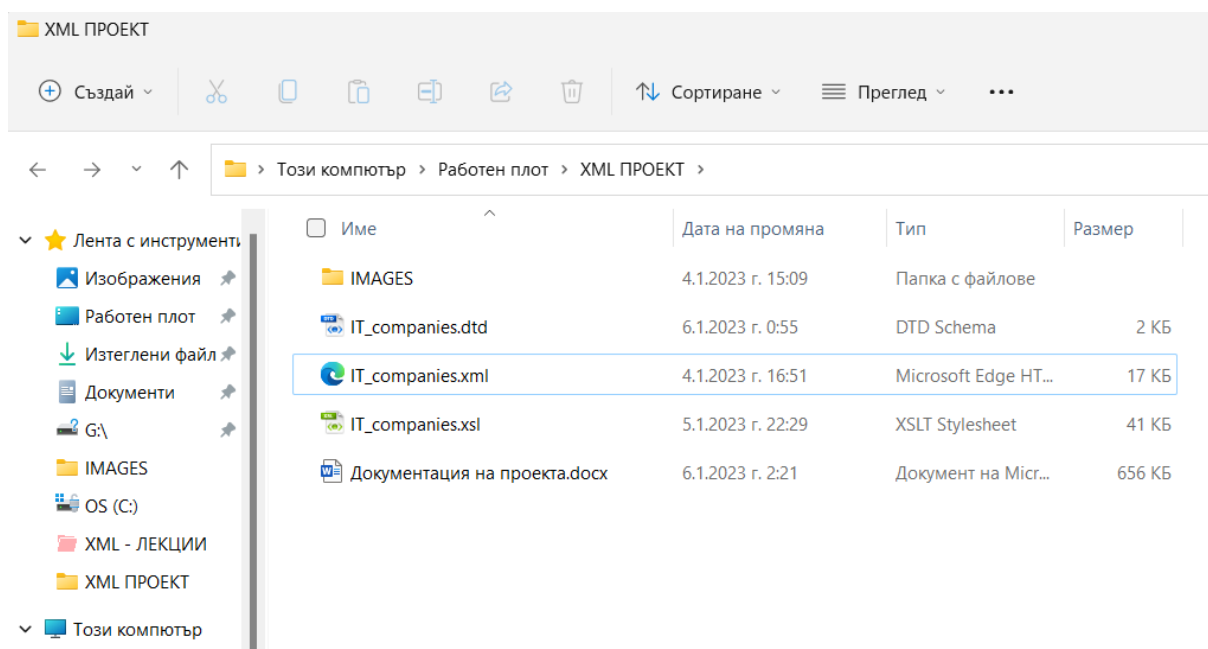
Фиг.9.3. Генериране на сценарий за трансформацията



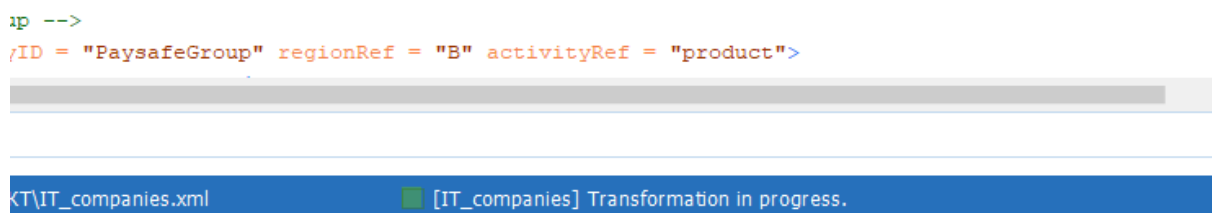




Фиг.9.4. Папката със съдържанието на проекта преди да сме натиснали **Apply associated (1)**



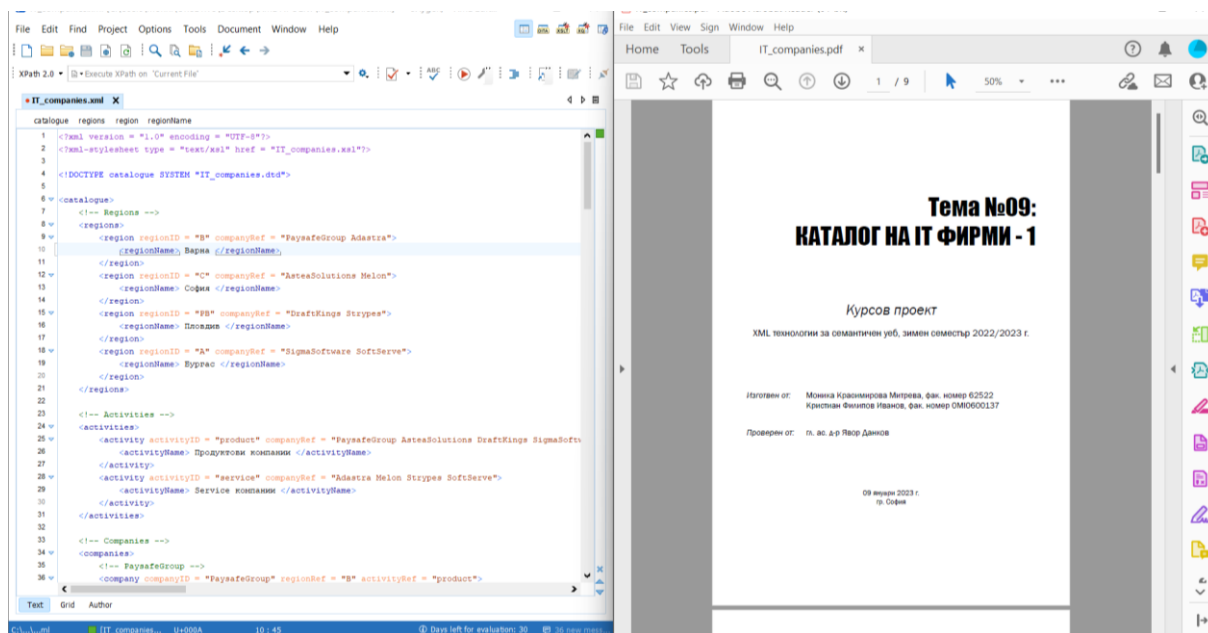
Фиг.9.5. Натискаме **Apply associated (1)**



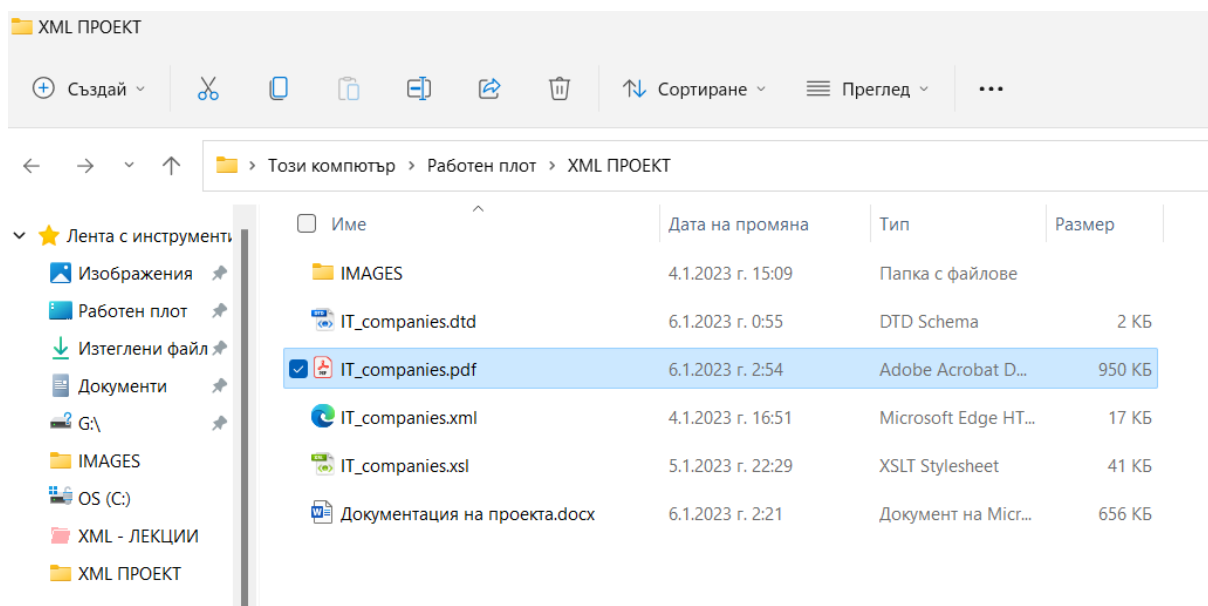
```
regionRef = "B" activityRef = "product">
```

[IT_companies] Transformation successful (3.7s)

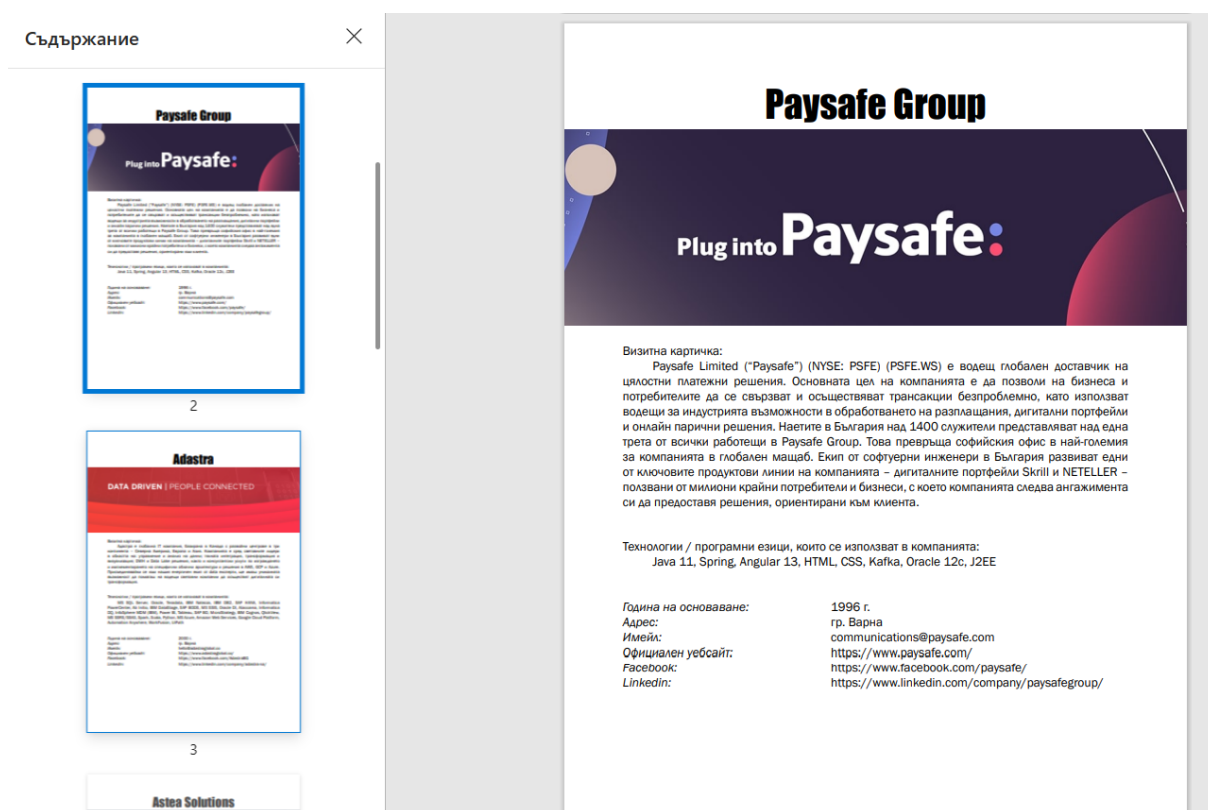
Фиг.9.6. Документът **IT_companies.pdf** е успешно генериран според XSL инструкции от файла **IT_companies.xml**



Фиг.9.7. Папката със съдържанието на проекта след като сме натиснали **Apply associated (1)**



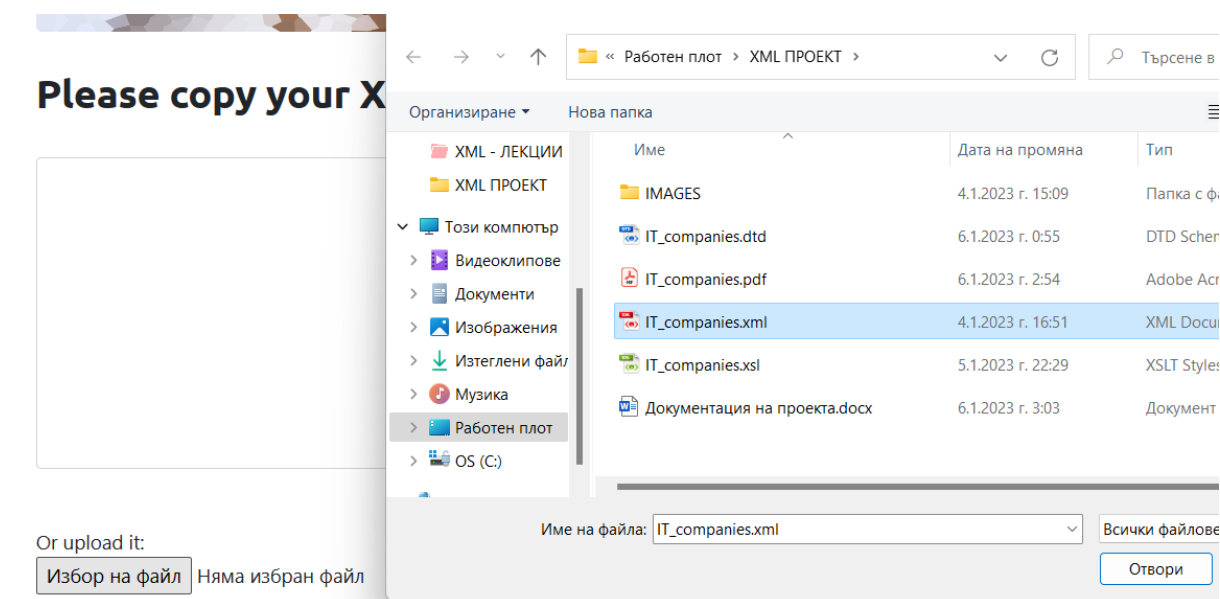
Фиг.9.8. Извадка от документа IT_companies.pdf



4. Тестване

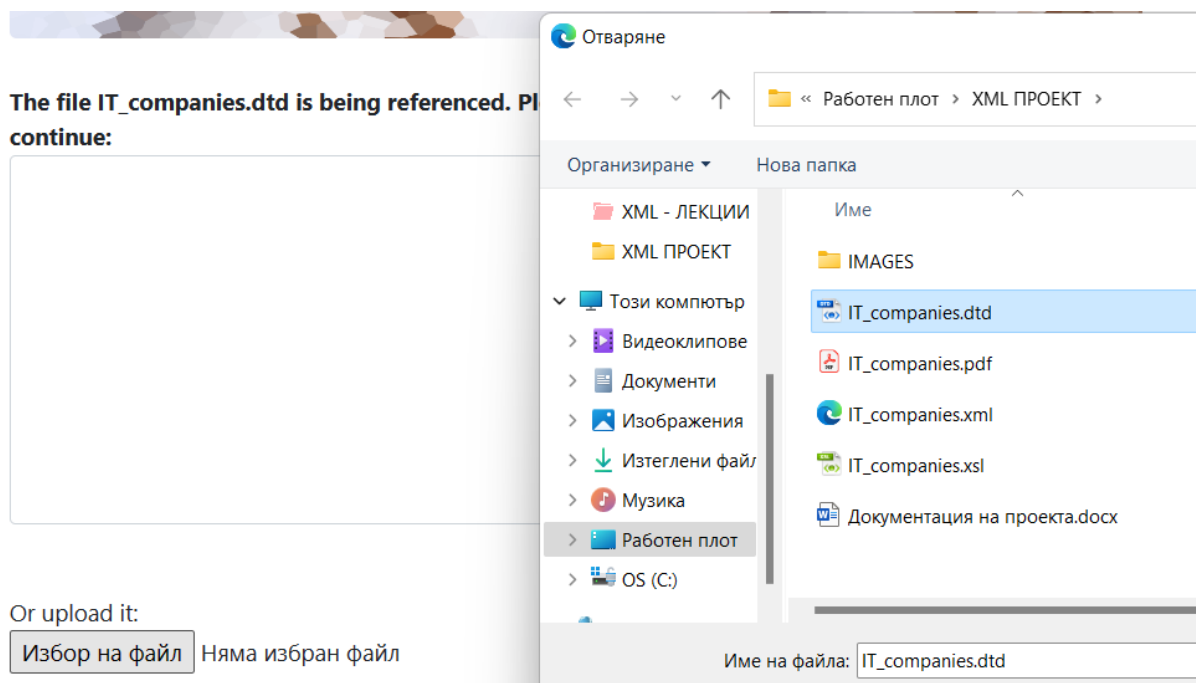
Проверката на DTD валидацията и входните данни от файла **IT_companies.xml** беше направена в сайта XML validation⁵.

Фиг.10. Зареждаме файла IT_companies.xml

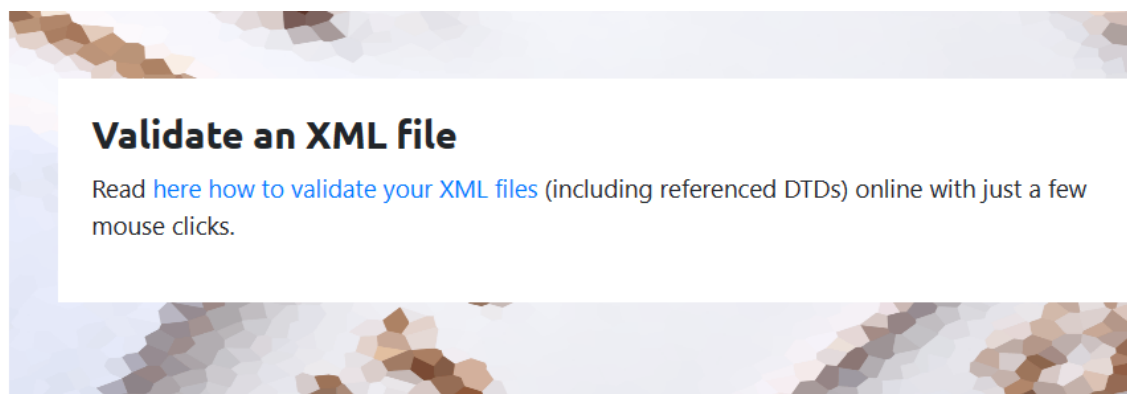


⁵ [Validate XML files \(xmlvalidation.com\)](https://www.xmlvalidation.com/)

Фиг.11. Прикачваме файла **IT_companies.dtd** (Сайтът изисква да го качим допълнително, тъй като **IT_companies.dtd** е външна граматика за **IT_companies.xml**)



Фиг.12. Резултат



No errors were found

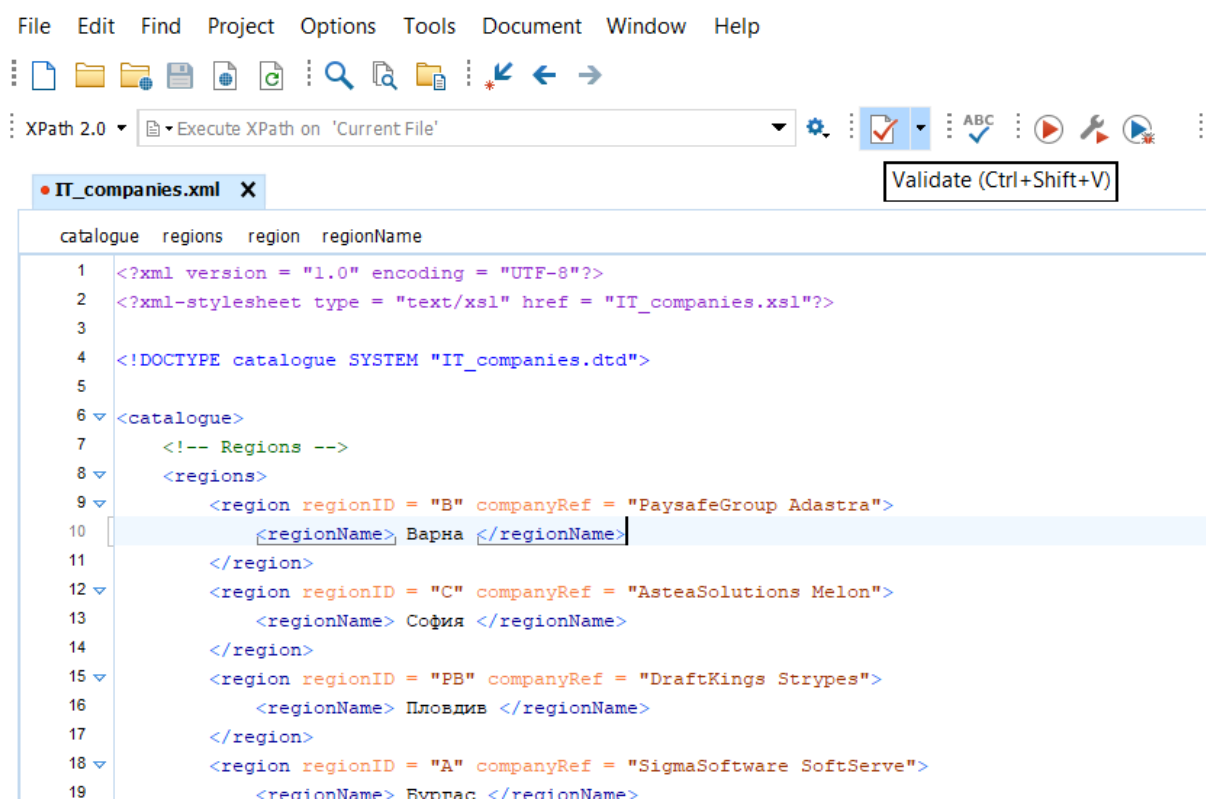
The following files have been uploaded so far:

[XML document:](#)

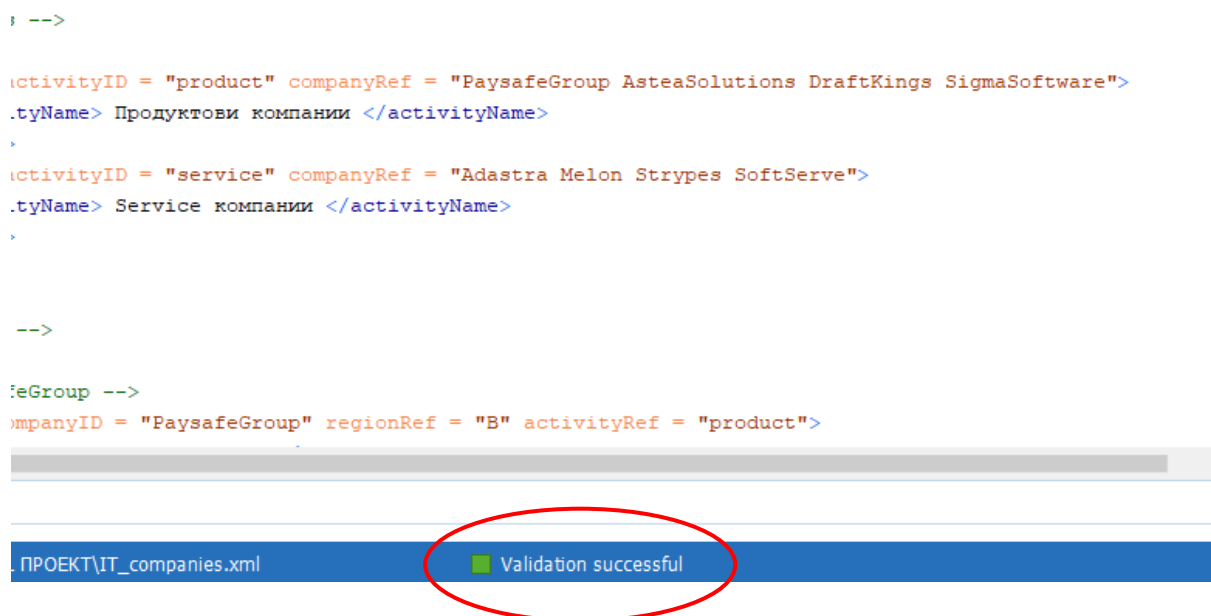
[IT_companies.dtd](#)

Click on any file name if you want to edit the file.

Фиг.13. Валидация посредством OXYGEN XML EDITOR



Фиг.14. Резултат



Създаване на IT_companies.pdf

[Generate IT_companies_pdf.mp4](#) - кратка демонстрация на създаване и прилагане на трансформиращ сценарий

5. Заключение и очаквано бъдещо развитие

5.1. Oxygen XML Editor

Създадената към момента реализация на каталог за IT фирми в България е базова интерпретация на заданието за курсовия проект. В процеса на разработка се запознахме както с XSL-FO, за да изградим изискващата се според условието трансформация, така и с инструмента Oxygen XML Editor, който да я реализира и да генерира крайният продукт – PDF файл.

Oxygen XML Editor е широкоспектърен редактор за качествена обработка на XML съдържание. Освен възможността да се генерират различни изходни файлове посредством XSL трансформации, платформата предлага и възможност за debugging, моментна валидация на входните XML данни, както и проверка дали подаденият .xml документ е добре структуриран, която характеристика е основополагаща при работа с подобен тип съдържание.

От съществена полза за нас по време на разработката на проекта беше начина, по който се извършва трансформацията на XML съдържанието към PDF такова. Трансформиращите сценарии се дефинират лесно и доста интуитивно – задава се име на сценария, посочва се източника на входни данни, както и .xsl файлът и се задават желаните име и формат на изходните данни. Веднъж генериран сценарият става преизползваем и при всяка следваща модификация в .xsl данните, не е необходимо да се генерира отново, достатъчно е само да се стартира и промените се прилагат, което прави инструмента Охуген изключително удобен за работа.

5.2. Нашият проект

„Каталог за IT фирми - 1“ към момента успешно се валидира от персонално създадената за него DTD граматика, което му осигурява запазването на цялостната структура и консистентност на информацията.

Прилежащият му XSL файл, с помощта на Oxygen XML Editor, лесно и бързо генерира изходен PDF файл, но не имплементиран максимално оптимално и за бъдещи реализации на решението може да бъде оптимизиран като се използват цикли и условни проверки. Освен това, стилового оформление на данните може да се изнесе в отделен CSS файл, което допълнително ще повиши четимостта и лесната поддръжка на текущия код.

„Каталог за IT фирми - 1“ поставя само основата. Текущият продукт, който създадохме, лесно се модифицира и според нас би намерил широко приложение в сферата на IT индустрията. Предимство на представените данни във лесно четим формат е, че каталогът може да служи за справка, както на бъдещи и настоящи студенти, търсещи стаж и/или редовна работа в IT сектора, така и за популяризиране на дейността/продуктите на избраните фирми, посредством наличните PDF визуализации, които да се разпространяват под формата на флаери или текстово съдържание в социалните медии, достигайки таргетираната аудитория.

6. Разпределение на работата

Структурата и съдържанието на настоящия курсов проект, както и равномерното разпределение на работата по търсене и анализ на необходимата информация бяха обсъдени от двамата членове на екипа. Проектът е разработен съвместно като следното разпределение се базира основно върху това на коя от отделните части е наблегнал повече всеки един от членовете на екипа.

- › **Моника Митрева:**
 - XML документ (IT_companies.xml)
 - DTD граматика (IT_companies.dtd)
 - PDF файл (IT_companies.pdf)
 - Документация на проекта
- › **Кристиан Иванов:**
 - XML документ (IT_companies.xml)
 - XSLT документ (IT_companies.xsl)
 - Документация на проекта

7. Използвани литературни източници

В настоящата точка може да се намери информация за използваните литературни източници и такива от Интернет, които послужиха за изготвянето на настоящия проект.

1. БАСКОМ: ИТ секторът в България расте с 10% през 2020 г. въпреки коронавируса (capital.bg) – сайт на CAPITAL.bg
2. Компании – ИТ обяви за работа, придобивки, екип и мнения (dev.bg) – сайт на DEV.BG
3. Oxygen XML Editor – официален сайт на OXYGEN XML EDITOR - инструментът, използван за генериране на изходния PDF файл
4. Best Online XML Viewer, XML Formatter, XML Editor, Analyser, Beautify-Beautifier, Minify, Tree structure (codebeautify.org) – инструментът, използван за генериране на XML дървото – Фиг.5.
5. Validate XML files (xmlvalidation.com) – онлайн инструмент, използван за проверка за входните данни в XML файла и DTD граматиката
6. XSL-FO Tutorial (sinsixx.com) – сайт с пълно описание на употребата на XSL-FO
7. Бончев, Б. XSLT в детайли. Лекционен курс: XML ТЕХНОЛОГИИ ЗА СЕМАНТИЧЕН УЕБ 2022/2023 г.