Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Кафедра штучного інтелекту

Звіт з лабораторної роботи №1

з курсу "Інтелектуальні технології в Internet та Semantic Web"

Тема: "Розробка XML-документів та відповідних XML-схем в редакторі XML"

|  |  |  |
| --- | --- | --- |
| Виконав:  ст. гр. ІТКН-17-7  Гур’єв О.Р. |  | Перевірив викладач:  Бібічков І.Є. |

Харків 2020

**Ссылка на гит-репозиторий:** <https://github.com/HurievA/IT-ISW>

**1. Цель работы.**

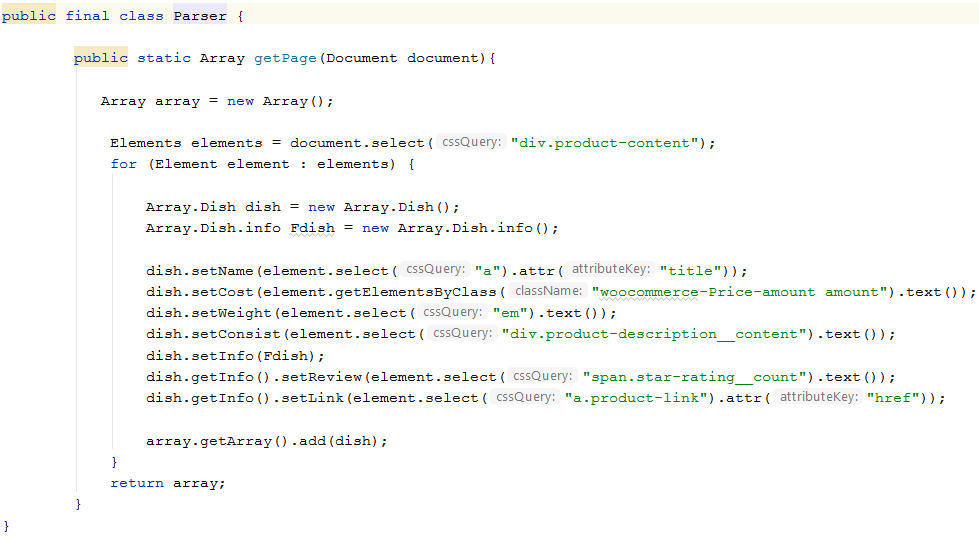
Отримання даних із віддаленого WEB-ресурсу та їх обробка. Ознайомлення із роботою у редакторі XMLSpy – середовищі для роботи з кількома XML-технологіми. Розробка XML-схеми, XML-документа, його валідація.

**2. Ход выполнения.**

В ходе работы была подключена библиотека Jsoup, библиотека предназначенная для анализа, извлечения и обработки данных, хранящихся в документах HTML. Был создан класс Array с XML аннотациями, что позволит правильно сгенерировать XML-документ. Класс Array представлен в виде списка, а также хранит в себе класса Dish, который содержит конструкторы и хранит параметры объекта после парсинга.



Далее создается класс Parser, в котором мы получаем все необходимые данные с веб-ресурса для дальнейшей работы с ними.



Полученные данные выводим в отдельный XML-документ.

public static void convertObjectToXml(Array dish, String filePath) throws FileNotFoundException,ParserConfigurationException, JAXBException {

try {

JAXBContext context = JAXBContext.*newInstance*(Array.class);

Marshaller marshaller = context.createMarshaller();

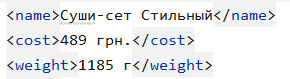
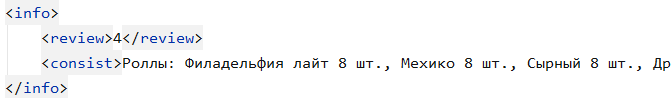
marshaller.setProperty(Marshaller.*JAXB\_FORMATTED\_OUTPUT*, Boolean.*TRUE*);

marshaller.marshal(dish, new File(filePath));

} catch (JAXBException e) {

e.printStackTrace(); }}

Пример сложноного и простого элементов:

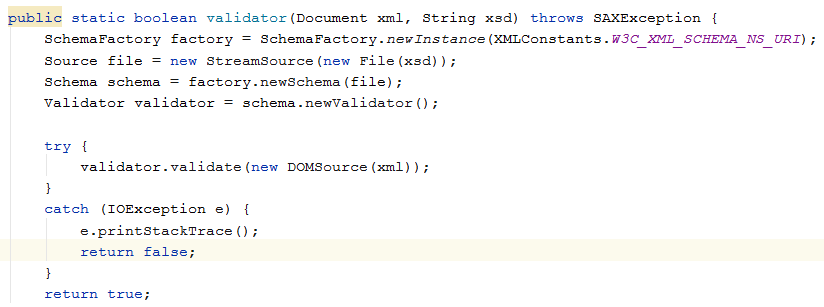




На основе XML файла была сгенерирована XSD схема



После этого был создан валидатор, который проверяет XML-файл на достоверность.



**3. Выводы**

В результате работы, были получены навыки работы с библиотекой Jsoup, архитектурой JAXB, а также приобретены навыки получение информации (парсинг) с веб-ресурсов. Полученные данные были обработаны, выведены в XML-документ и провалидированы. Валидация XML-схемы необходима для ее сравнение с XSD-схемой. XML документ должен соответствовать определенному типу документов и придерживаться всех синтаксических правил.

4. Контрольні запитання

1. Синтаксис XML.

XML документ должен иметь корневой элемент

XML элемент должен иметь закрывающий тег

XML теги регистрозависимы

XML элементы должны соблюдать последовательность вложенности

Значения XML атрибутов должны заключаться в кавычки

1. В чому різниця між поняттям правильно сформованого (well formed) та валідного (valid) XML-документа?

XML документ с корректным синтаксисом является "синтаксически верным".

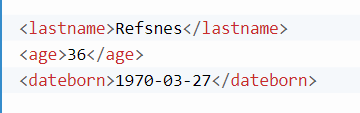
Первое правило для валидного XML документа то, что он должен быть синтаксически верным. Второе правило — валидный XML документ должен соответствовать определенному типу документов.

1. Що таке XML-схема?

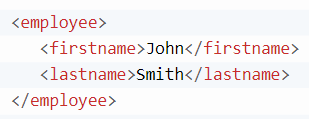
XML схема:

* определяет элементы, которые могут появляться в XML документе
* определяет атрибуты, которые могут появляться в XML документе
* определяет, какие элементы являются дочерними
* определяет порядок дочерних элементов
* определяет количество дочерних элементов
* определяет, пустой ли элемент или может содержать текст
* определяет типы данных элементов и атрибутов
* определяет фиксированные значения и значения по умолчанию элементов и атрибутов

1. Які елементи XML-документа мають простий тип? Як вказується/створюється простий тип?



1. Які елементи мають комплексний тип? Як створюється комплексний тип?



6. Якого типу можуть бути атрибути в XML-документі?

-string

-decimal

-integer

-boolean

-date

-time