**ЛАБОРАТОРНАЯ РАБОТА**

XML: ОСНОВЫ СОЗДАНИЯ

И ОБРАБОТКИ СТРУКТУРИРОВАННЫХ ДАННЫХ

**Цель работы**: освоить принципы создания, валидации и обработки XML-

документов, развить навыки анализа данных и работы с инструментами для

обработки структурированных форматов.

**Описание выполненных задач**

**1 - Создание XML-документа**: Был создан файл library.xml, содержащий структуру библиотеки, включающую 3 книги. Каждая книга описана с использованием таких элементов, как title, author, year, genre и price. Дополнительно для каждой книги предусмотрен атрибут id, обеспечивающий её уникальность. Файл строго соответствует синтаксису XML.

**2 - Создание XSD-схемы**: Разработан файл library.xsd, задающий правила для валидации XML-документа. Схема проверяет:

-Наличие всех обязательных элементов (title, author, year, genre, price).

-Правильный формат и диапазон значений для элементов year (от 1800 до текущего года) и price (положительное число с двумя знаками после запятой).

-Наличие уникального атрибута id.

**3 - Валидация XML-документа**: Написана программа на языке Java, которая с помощью библиотеки javax.xml.validation проверяет, соответствует ли library.xml разработанной XSD-схеме. Программа успешно валидирует документ или выводит ошибки в случае несоответствия.

**4 - Обработка XML-документа**: Создана Java-программа для чтения, анализа и вывода данных из файла library.xml. Программа выполняет следующие функции:

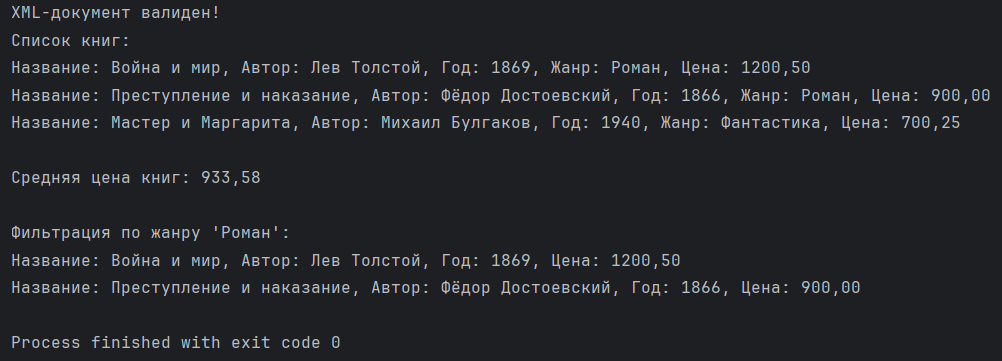
-Выводит список всех книг с их подробной информацией.

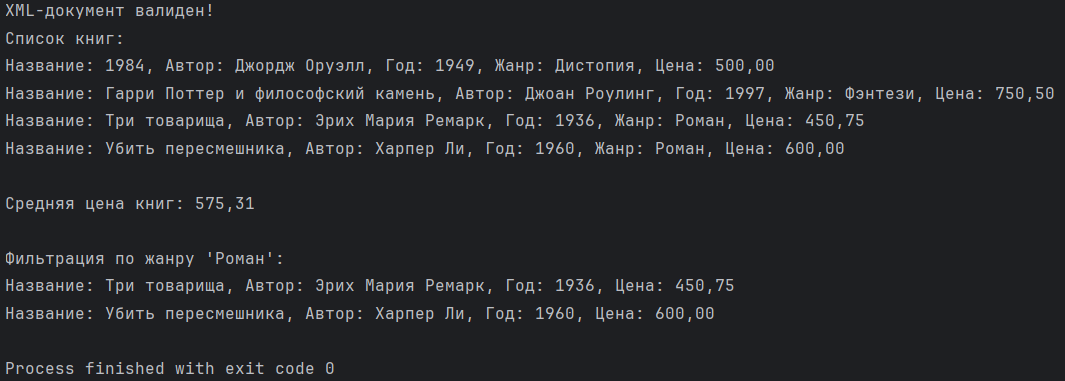
-Рассчитывает среднюю цену книг.

-Фильтрует книги по заданному критерию, например, по жанру "Роман".

**5 - Исследование и анализ реального XML-документа**: Был найден пример реального XML-документа (RSS-лента). Выполнен анализ его структуры: корневым элементом является <rss>, внутри находятся <channel> и <item>. Указаны элементы <title>, <link> и <description> для каждого обновления. Документ используется для автоматического предоставления новостей и обновлений.

**Примеры вывода программы**.





**Исследование и сравнение**.

**1 - Основные теги**: XML-документ состоит из тегов <rss>, <channel>, <item>, <title>, <link> и <description>.

**2 - Вложенность**: <rss> — корневой элемент, содержащий <channel>, который, в свою очередь, включает <title>, <link>, <description> и несколько <item>. Каждый <item> описывает одну запись (новость).

**3 - Атрибуты**: У тега <rss> есть обязательный атрибут version, указывающий на версию RSS-протокола (в данном случае — 2.0).

**4 - Назначение**: Документ представляет собой RSS-ленту, предназначенную для распространения обновлений сайта, таких как новости или статьи.

**5 - Простота и читаемость**: Структура документа проста и стандартизирована, что облегчает его обработку как людьми, так и программами (например, агрегаторами новостей).

**6 - Ценность данных**: Элементы <item> содержат заголовки, ссылки и описания новостей, делая их доступными для быстрого просмотра.

7 - Данный документ применяется для автоматического уведомления пользователей о новых материалах на сайте, предоставляя ссылки и краткое описание каждой новости. Это упрощает процесс информирования аудитории.