Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Кафедра Штучного інтелекту

Дисципліна: «Інтелектуальні технології в Internet та Semantic Web»

Лабораторна робота № 3

на тему: «РОЗРОБКА ЗАПИТІВ XQUERY ІЗ ВИКОРИСТАННЯМ XPATH-ВИРАЗІВ»

|  |  |  |
| --- | --- | --- |
| Виконали:  ст. гр. ІТКН-17-7  Савенков Д.В. |  | Перевірив:  Бібічков І. Є. |
|  |  |  |

Харків 2020

3 РОЗРОБКА ЗАПИТІВ XQUERY ІЗ ВИКОРИСТАННЯМ XPATH-ВИРАЗІВ

**3.1 Мета роботи**

Навчитися використовувати XPath-вирази для XQuery-запитів. Створити XQuery запити для обробки XML – документів

**3.2 Виконання завдань**

**3.2.1 Отримати загальну кількість описаних елементів (товарів, послуг або ін.)**

**XQuery**: count(/candidates/candidate)

**Python code**:

from lxml import etree

tree = etree.parse('djinni.xml')

print(tree.xpath('count(/candidates/candidate)'))

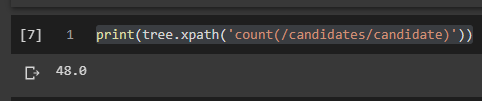


Рис. 1 – Результати виконання завдання 1

**3.2.2 вивести інформацію за однією із характеристик для всіх описаних складних елементів.**

**XQuery**: //\*[@quantity]

**Python** **code**:

for node in tree.xpath('//\*[@quantity]'):

print(etree.tostring(node, pretty\_print=True, encoding='unicode'))

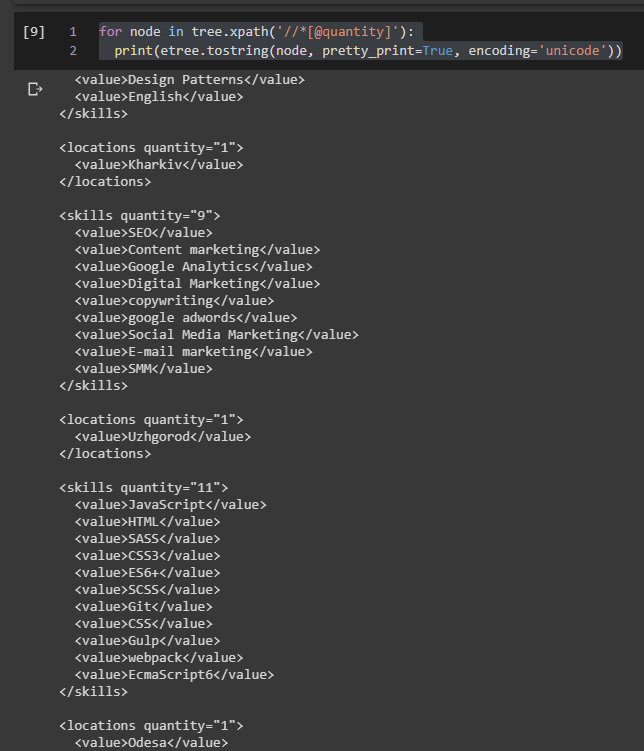


Рис. 2 – Результати виконання завдання 2

**3.2.3 отримати інформацію1 за однією із характеристик2 для одного складного елементу.**

**XQuery**: //locations[@quantity]

**Python code**:

for node in tree.xpath('//locations[@quantity]'):

print(etree.tostring(node, pretty\_print=True, encoding='unicode'))

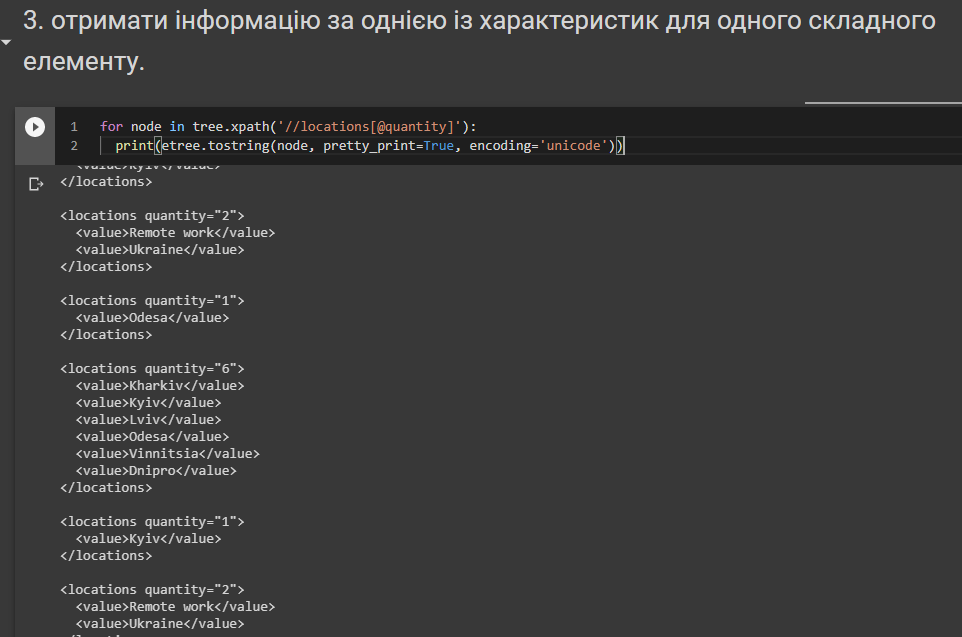


Рис. 3 – Результати виконання завдання 3

**3.2.4 вивести номер елемента, назва якого складається більше, ніж із одного слова**

**XQuery**: //candidate[position[contains(., ' ')]]/@id

**Python code**:

# Work in tests, but not in Python's frameworks/libs (atleast, I haven't found one)

# http://www.xpathtester.com/xquery

# Problems with if and for-loop statements

query = "for $x in //candidate return if($x/position[contains(., ' ')]) then <index>{data($x/@id)}</index> else ''"

query = "//candidate[position[contains(., ' ')]]/@id" # Every item has atleast 2 words (fucntion can be checked on different fields)

print(tree.xpath(query))

# Works in Python (but not pure XPath/XQuery)

for candidate in tree.xpath('//candidate'):

if candidate.xpath("position[contains(., ' ')]"):

print(candidate.xpath("@id")[0]) # Every item has atleast 2 words (fucntion can be checked on different fields)

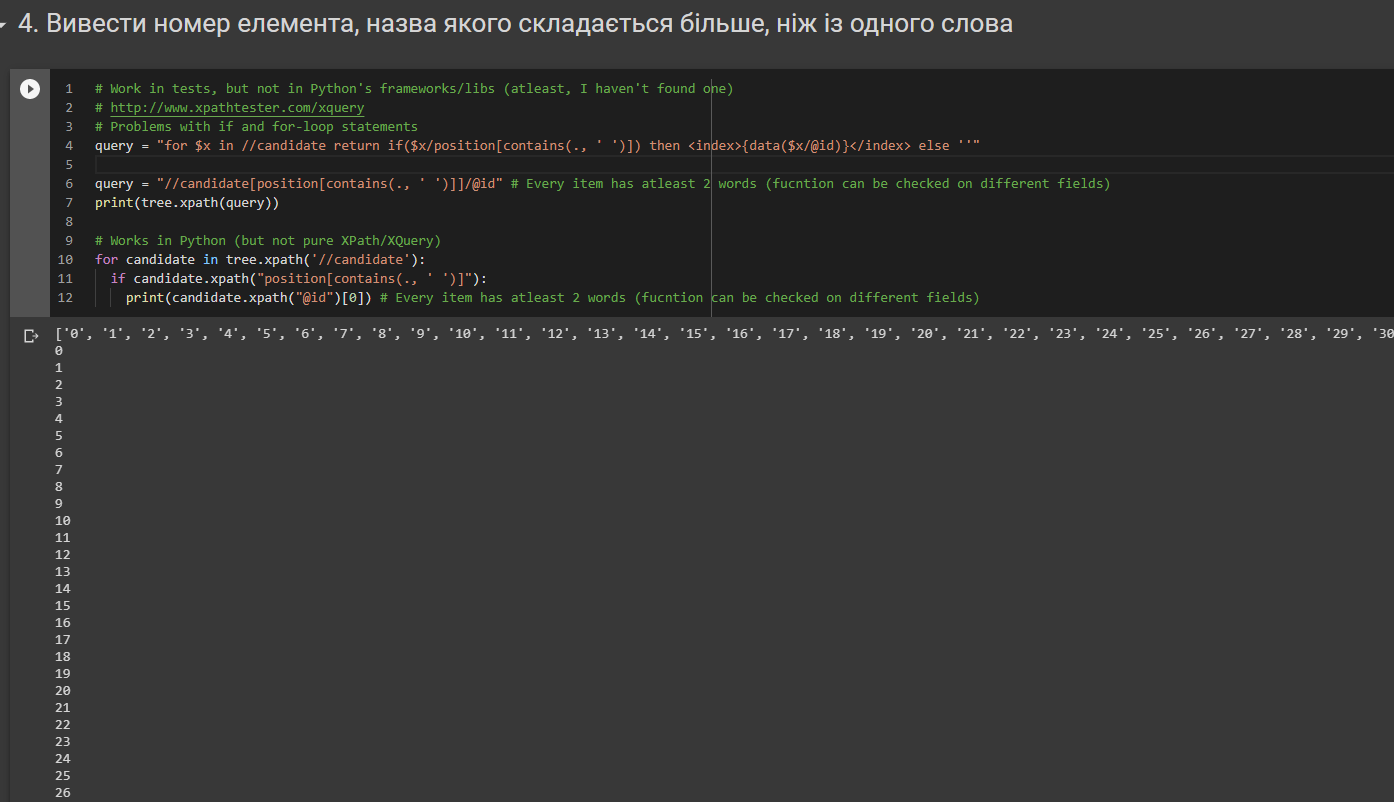


Рис. 4 – Результати виконання завдання 4

**3.2.5 вивести тільки 1, 2, 3 параметр одного складного елемента.**

**XQuery**: //skills[1]/value[i], де i = [1, 2, 3]

**Python code**:

for i in range(1, 4):

print(f"============{i}============")

for node in tree.xpath(f'//skills[1]/value[{i}]'):

print(etree.tostring(node, pretty\_print=True, encoding='unicode'))

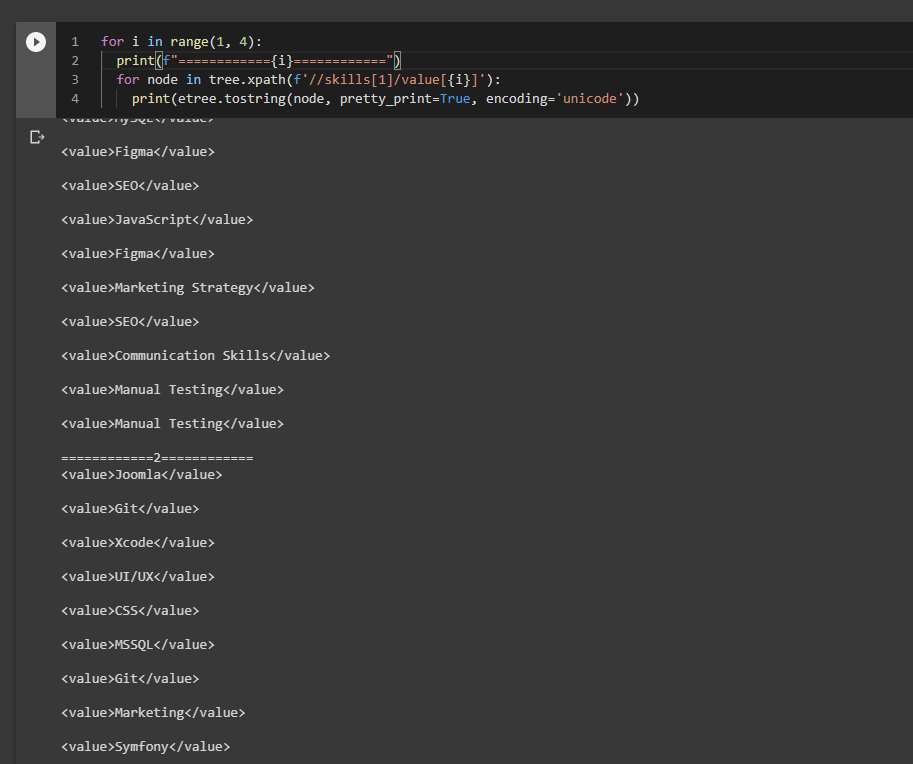


Рис. 5 – Результати виконання завдання 5, 6, 7

**3.2.6 вивести назву елемента, якщо його ціна (або кількість відгуків, зірочок, тощо) знаходиться в заданому діапазоні4, а також містить задану послідовність літер у назві**

**XQuery**: /candidates/\*[skills/@quantity > 10 and skills/@quantity < 15 and starts-with(position, 'Senior')]/position

**Python code**:

# Work in tests, but not in Python's frameworks/libs (atleast, I haven't found one)

# http://www.xpathtester.com/xquery

# Problems with if and for-loop statements

query = "for $x in //candidate return if((count($x/skills/value) > 10) and (count($x/skills/value) < 15) and (starts-with($x/position, 'Senior'))) then $x/position else ''"

# Works in Python, XPath, not XQuery

query = "/candidates/\*[skills/@quantity > 10 and skills/@quantity < 15 and starts-with(position, 'Senior')]/position"

print(etree.tostring(tree.xpath(query)[0], pretty\_print=True, encoding='unicode'))

# Works in Python but not pure XQuery or Xpath

for candidate in tree.xpath('//candidate'):

if candidate.xpath('skills/@quantity > 10 and (count(skills/value) < 15)') and candidate.xpath("starts-with(position, 'Senior')"):

print(etree.tostring(candidate.xpath('position')[0], pretty\_print=True, encoding='unicode'))

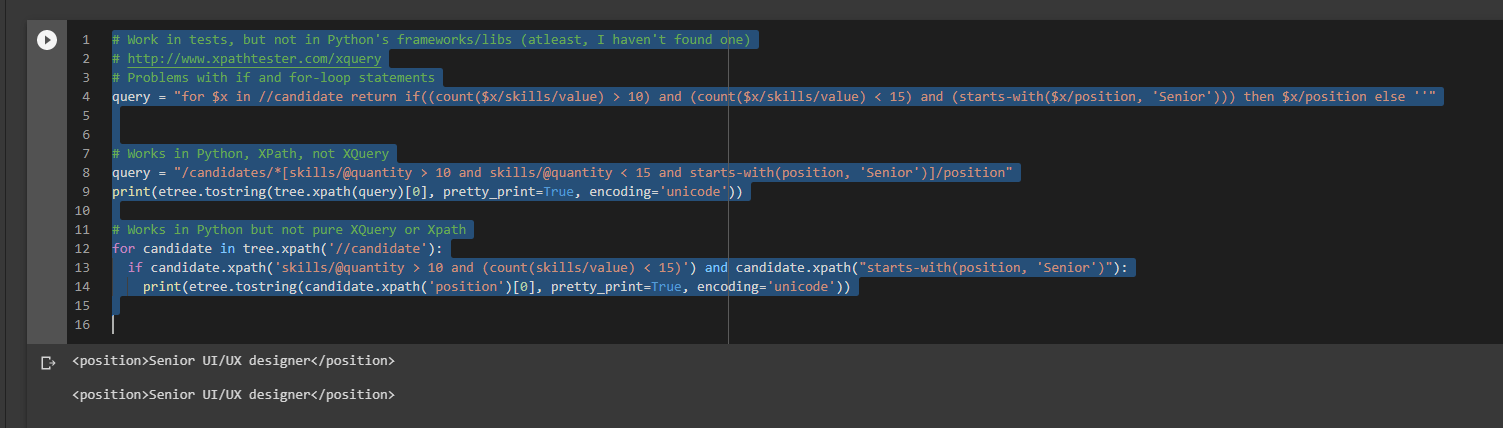


Рис. 6 – Результати виконання завдання 8

**3.2.7 отримати інформацію1 за однією із характеристик2 для одного складного елементу.**

**XQuery**: //candidate[position() mod 5 = 0]/position | //candidate[position() mod 5 = 0]/salary

**Python code**:

for candidate in tree.xpath('//candidate[position() mod 5 = 0]/position | //candidate[position() mod 5 = 0]/salary'):

print(etree.tostring(candidate, pretty\_print=True, encoding='unicode'))

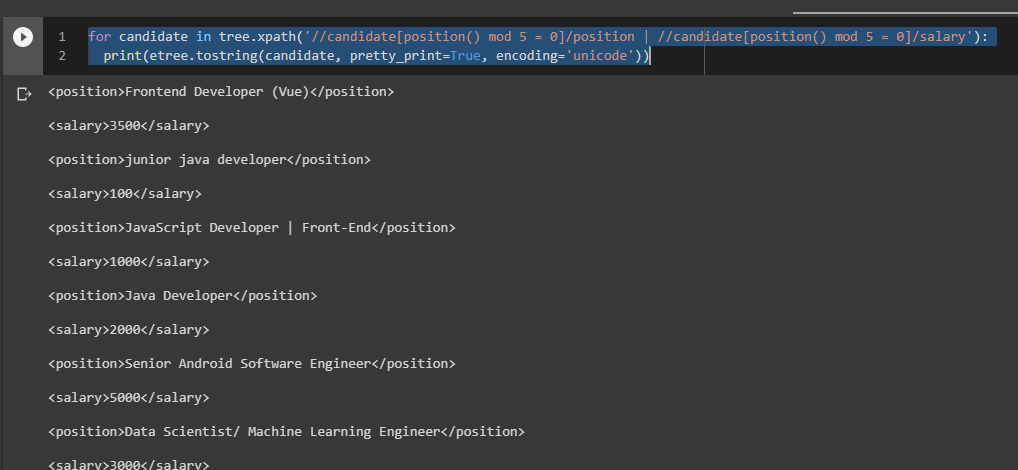


Рис. 7 – Результати виконання завдання 9

**3.2.8 Вивести номер елементу (товару, послуги тощо) та додатково, ще один параметр (напр. ціна) для кожного другого складного елемента в документі.**

**XQuery**: //candidate[/\*[count(./\*) > 1]]/@id | //candidate/\*[count(./\*) > 1]/@quantity

**Python code**:

res = tree.xpath("//candidate[/\*[count(./\*) > 1]]/@id | //candidate/\*[count(./\*) > 1]/@quantity")

print(res)

for i in range(0, len(res)-1, 2):

print(res[i], res[i+1])

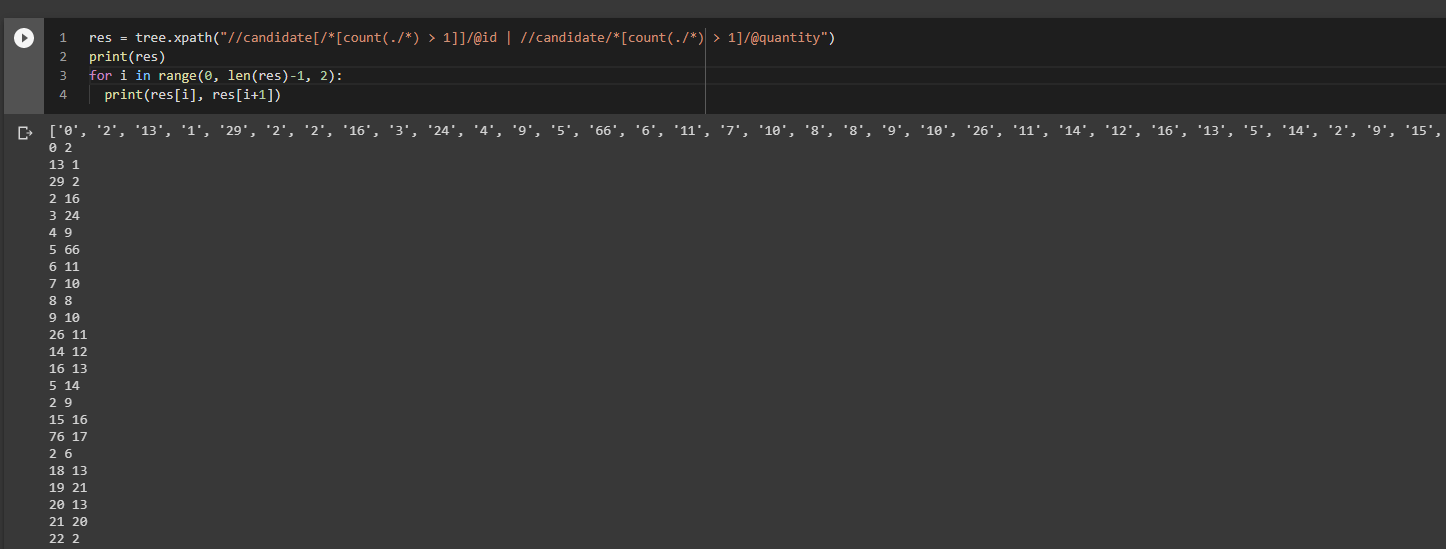


Рис. 8 – Результати виконання завдання 10

**3.3 Висновки**

У результаті виконання даної практичної роботи були засвоєні основні принципи роботи із XPath та XQuery для запитів для отримання даних елементів з XML-документу, що дозволяє програмним шляхом (наприклад, використовуючи мову програмування Python) працювати з вмістом файлу.

Були розглянуті принципи взаємодії з елементами та атрибутами різних рівнів складності та вкладеності, а також були розглянуті основні функції XQuery для складних запитів.

Отриманні знання та навички можна буде використати при виконанні наступних практичних робот або під час взаємодії з xml-документами для виконання поставленої задачі в майбутніх та існуючих проектах.