Лабораторна робота № 1.

Вивчення базових операцій обробки ХМС-документів

Метою роботи ϵ здобуття практичних навичок створення програм, орієнтованих на обробку XML-документів засобами мови Python.

Завдання роботи полягає у наступному:

- 1. Виконати збір інформації зі сторінок Web-сайту за варіантом.
- 2. Виконати аналіз сторінок Web-сайту для подальшої обробки текстової та графічної інформації, розміщеної на ньому.
- 3. Реалізувати функціональні можливості згідно вимог, наведених нижче.

Функціональні вимоги

1. На основі базової адреси Web-сайту виконати обхід наявних сторінок сайту, відокремлюючи текстову та графічну інформацію від тегів HTML. Пошук вузлів виконувати засобами XPath. Наступну сторінку для аналізу **цього ж сайту** обрати як одне із гіперпосилань на даній сторінці (тег). Обмежитись аналізом 20 сторінок сайту. Зберегти XML у вигляді файлу. Формат XML-документу:

```
<data>
  <page url="wwww.server.com/index.hml">
 <fragment type="text">
.... знайдений текст
 </fragment>
 <fragment type="image">
.... url зображення
 </fragment>
 </page>
 <page url="wwww.server.com/index1.hml">
 <fragment type="text">
.... знайдений текст
</fragment>
 <fragment type="image">
.... url зображення
</fragment>
 </page>
</data>
```

- 2. Виконати аналіз отриманих даних засобами XML згідно варіанту та вивести результати у консольне вікно. Відбір вузлів виконувати засобами XPath.
- 3. Проаналізувати вміст Web-сторінок інтернет-магазину (див. варіант). Отримати ціну, опис та зображення для 20 товарів з нього за допомогою

DOM-парсеру та мови XPath для пошуку відповідних вузлів. Результат записати в XML-файл.

4. Перетворити отриманий XML-файл у XHTML-сторінку за допомогою мови XSLT. Дані подати у вигляді XHTML-таблиці та записати його у файл.

Вимоги до інтерфейсу користувача

Використовувати консольний (текстовий) інтерфейс користувача.

Вимоги до інструментарію

- 1. Мова програмування Python 2.7
- 2. Бібліотека для обробки XML libxml2 або (xml.dom, xml.xpath.. вбудовані пакети) або lxml на вибір студента.
- 3. Середовище розробки програмного забезпечення PyCharm Community Edition 2016 (*опціонально*)

Вибір варіанту

Робота виконується індивідуально. Варіант обирається шляхом взяття останніх двох цифр номеру залікової книжки студента.

Варіанти

№ вар.	Базова сторінка (завдання 1)	Зміст завдання 2	Адреса інтернет- магазину (завдання 3)
1.	www.kpi.ua	Максимальна кількість	www.rozetka.ua
		текстових фрагментів	
2.	www.ukr.net	Середня кількість	www.repka.ua
		текстових фрагментів	
3.	www.bigmir.net	Мінімальна кількість	www.sokol.ua
		графічних фрагментів	
4.	www.korrespondent.net	Кількість текстових	www.hotline.ua
		фрагментів по кожному	
		документу	
5.	www.football.ua	Кількість графічних	www.moyo.ua
		фрагментів по кожному	

документу Вивести список гіперпосилань www.portativ.ua 7. www.tsn.ua Мінімальна кількість графічних фрагментів www.www.potmarket.ua 8. www.golos.ua Середня кількість текстових фрагментів www.petmarket.ua 9. www.ridna.ua Вивести список гіперпосилань www.meblium.com.ua 10. www.ridna.ua фрагментів по кожному документу www.nestiki.com.ua 11. www.museum- ukraine.org.ua Середня кількість текстових фрагментів www.hozmart.com.ua 12. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 13. www.doroga.ua Вивести список гіперпосилань www.mebli-lviv.com.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Гередня кількість графічних фрагментів www.odissey.kiev.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.zvetsad.com.ua 17. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 18. www.shkola.ua Кількість текстових фрагментів				
6. www.isport.ua гіперпосилань www.wallet.ua 7. www.tsn.ua Мінімальна кількість графічних фрагментів www.wallet.ua 8. www.golos.ua Середня кількість текстових фрагментів www.petmarket.ua 9. www.sport.ua Вивести список гіперпосилань www.meblium.com.ua 10. kiлькість графічних фрагментів по кожному документу www.veliki.com.ua 11. www.museum- ukraine.org.ua вивести список гіперпосилань www.hozmart.com.ua 12. www.doroga.ua Середня кількість текстових фрагментів www.freedelivery.in.ua 13. www.stejka.com Вивести список гіперпосилань www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.freedelivery.in.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.zvetsad.com.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.auto-store.kiev.ua 18. www.shkola.ua Кількість текстових фрагментів www.tennismag.com.ua 19. www.shkola.ua Вив			документу	
7. www.tsn.ua Мінімальна кількість графічних фрагментів www.wallet.ua 8. www.golos.ua Середня кількість текстових фрагментів www.petmarket.ua 9. www.xsport.ua Вивести список гіперпосилань www.meblium.com.ua 10. www.ridna.ua фрагментів по кожному документу www.neblium.com.ua 11. www.museum- ukraine.org.ua Вивести список гіперпосилань www.instrument.in.ua 12. www.doroga.ua Середня кількість текстових фрагментів www.hozmart.com.ua 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.posolstva.org.ua Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.zvetsad.com.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.shkola.ua Мінімальна кількість гекстових фрагментів www.tennismag.com.ua 19. www.shkola.ua фрагментів	6.	www.isport.ua	Вивести список	<u>www.portativ.ua</u>
7. www.tsn.ua графічних фрагментів 8. www.golos.ua Середня кількість текстових фрагментів 9. www.xsport.ua Вивести список гіперпосилань www.meblium.com.ua 10. www.ridna.ua Кількість графічних фрагментів по кожному документу www.veliki.com.ua 11. www.museum-ukraine.org.ua Вивести список гіперпосилань www.instrument.in.ua 12. www.doroga.ua Середня кількість текстових фрагментів www.freedelivery.in.ua 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.ua.igotoworld.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.posolstva.org.ua Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.zvetsad.com.ua 17. www.osvita.ua Середня кількість текстових фрагментів www.auto-store.kiev.ua 18. www.shkola.ua Кількість текстових фрагментів www.tennismag.com.ua 19. www.shkola.ua фрагментів по кожному документу www.fishing- 20. www.ostriv.in.ua Вивести список			_	
Second Part of the composition of the compositio	7.	www.tsn.ua	Мінімальна кількість	www.wallet.ua
8. www.golos.ua текстових фрагментів 9. www.xsport.ua Вивести список гіперпосилань www.meblium.com.ua 10. кількість графічних фрагментів по кожному документу www.veliki.com.ua 11. www.museum- ukraine.org.ua Вивести список гіперпосилань www.hozmart.com.ua 12. www.doroga.ua Мінімальна кількість графічних фрагментів www.hozmart.com.ua 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу 20. www.ostriv.in.ua Вивести список www.fishing-				
текстових фрагментів 9. www.xsport.ua Вивести список гіперпосилань www.meblium.com.ua 10. www.ridna.ua Кількість графічних фрагментів по кожному документу www.veliki.com.ua 11. www.museum-ukraine.org.ua Вивести список гіперпосилань www.instrument.in.ua 12. www.doroga.ua Середня кількість текстових фрагментів www.hozmart.com.ua 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу www.fishing- 20. www.ostriv.in.ua Вивести список www.fishing-	8.	www.golos.ua	Середня кількість	www.petmarket.ua
9. www.xsport.ua гіперпосилань 10. www.ridna.ua Кількість графічних фрагментів по кожному документу 11. www.museum- ukraine.org.ua Вивести список гіперпосилань www.instrument.in.ua 12. www.doroga.ua Середня кількість текстових фрагментів www.hozmart.com.ua 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua Кількість текстових фрагментів по кожному документу www.fishing- 20. www.ostriv.in.ua Вивести список www.fishing-			текстових фрагментів	
10. мим. гідпа. ца Кількість графічних фрагментів по кожному документу мим. лайна праводи п	Q	www.xsport.ua	Вивести список	www.meblium.com.ua
10.www.ridna.uaфрагментів по кожному документу11.www.museum- ukraine.org.uaВивести список гіперпосиланьwww.instrument.in.ua12.www.uartlib.orgСередня кількість текстових фрагментівwww.hozmart.com.ua13.www.doroga.uaМінімальна кількість графічних фрагментівwww.freedelivery.in.ua14.www.stejka.comВивести список гіперпосиланьwww.mebli-lviv.com.ua15.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість графічних фрагментівwww.auto-store.kiev.ua19.www.shkola.uaКількість текстових фрагментів по кожному документуwww.fishing-20.www.ostriv.in.uaВивести списокwww.fishing-) ,		гіперпосилань	
документуВивести список гіперпосиланьwww.instrument.in.ua11.www.museum- ukraine.org.uaВивести список текстових фрагментівwww.hozmart.com.ua12.www.doroga.uaМінімальна кількість графічних фрагментівwww.freedelivery.in.ua13.www.doroga.uaВивести список гіперпосиланьwww.mebli-lviv.com.ua14.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість графічних фрагментівwww.auto-store.kiev.ua19.www.shkola.uaКількість текстових фрагментів по кожному документуwww.fishing-20.www.ostriv.in.uaВивести списокwww.fishing-		www.ridna.ua	• •	www.veliki.com.ua
11.www.museum-ukraine.org.uaВивести список гіперпосиланьwww.instrument.in.ua12.www.uartlib.orgСередня кількість текстових фрагментів13.www.doroga.uaМінімальна кількість графічних фрагментів14.www.stejka.comВивести список гіперпосиланьwww.mebli-lviv.com.ua15.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість графічних фрагментівwww.auto-store.kiev.ua19.www.shkola.uaКількість текстових фрагментівwww.tennismag.com.ua19.www.shkola.uaфрагментів по кожному документу20.www.ostriv.in.uaВивести списокwww.fishing-	10.		фрагментів по кожному	
11.ukraine.org.uaгіперпосиланьwww.hozmart.com.ua12.www.uartlib.orgСередня кількість текстових фрагментівwww.hozmart.com.ua13.www.doroga.uaМінімальна кількість графічних фрагментівwww.freedelivery.in.ua14.www.stejka.comВивести список гіперпосиланьwww.mebli-lviv.com.ua15.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість графічних фрагментівwww.auto-store.kiev.ua19.www.shkola.uaфрагментів по кожному документуwww.tennismag.com.ua20.www.ostriv.in.uaВивести списокwww.fishing-			документу	
ukraine.org.uaгіперпосилань12.www.uartlib.orgСередня кількість текстових фрагментівwww.hozmart.com.ua13.www.doroga.uaМінімальна кількість графічних фрагментівwww.freedelivery.in.ua14.www.stejka.comВивести список гіперпосиланьwww.mebli-lviv.com.ua15.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість графічних фрагментівwww.auto-store.kiev.ua19.www.shkola.uaКількість текстових фрагментів по кожному документуwww.fishing-20.www.ostriv.in.uaВивести списокwww.fishing-	11	www.museum-	Вивести список	www.instrument.in.ua
12.www.uartib.orgтекстових фрагментів13.www.doroga.uaМінімальна кількість графічних фрагментів14.www.stejka.comВивести список гіперпосиланьwww.mebli-lviv.com.ua15.www.ua.igotoworld.comСередня кількість графічних фрагментівwww.zvetsad.com.ua16.www.posolstva.org.uaМаксимальна кількість текстових фрагментівwww.odissey.kiev.ua17.www.uahotels.infoСередня кількість текстових фрагментівwww.zvetsad.com.ua18.www.osvita.uaМінімальна кількість гекстових фрагментівwww.auto-store.kiev.ua19.www.shkola.uaКількість текстових фрагментів по кожному документуwww.tennismag.com.ua20.www.ostriv.in.uaВивести списокwww.fishing-	11.	ukraine.org.ua	гіперпосилань	
текстових фрагментів 13. www.doroga.ua Мінімальна кількість графічних фрагментів www.freedelivery.in.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua Кількість текстових фрагментів по кожному документу www.fishing- 20. www.ostriv.in.ua Вивести список www.fishing-	12.	www.uartlib.org	Середня кількість	www.hozmart.com.ua
13. www.doroga.ua графічних фрагментів www.mebli-lviv.com.ua 14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-				
14. www.stejka.com Вивести список гіперпосилань www.mebli-lviv.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-	13.	www.doroga.ua	Мінімальна кількість	www.freedelivery.in.ua
14. www.stejka.com гіперпосилань www.zvetsad.com.ua 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-			графічних фрагментів	
Гіперпосилань 15. www.ua.igotoworld.com Середня кількість графічних фрагментів www.zvetsad.com.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. www.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. www.shkola.ua фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-	1.4	www.stejka.com	Вивести список	www.mebli-lviv.com.ua
15. www.ua.igotoworld.com графічних фрагментів www.odissey.kiev.ua 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. мww.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-	14.		гіперпосилань	
графічних фрагментів 16. www.posolstva.org.ua Максимальна кількість текстових фрагментів www.odissey.kiev.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. муж.овуіта.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriy.in.ua Вивести список www.fishing-	15.	www.ua.igotoworld.com	Середня кількість	www.zvetsad.com.ua
16. www.posolstva.org.ua текстових фрагментів www.zvetsad.com.ua 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. мww.osvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. мww.ostriv.in.ua Вивести список www.fishing-			графічних фрагментів	
текстових фрагментів 17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. мим.овуіта.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. мим.овутіу.in.ua вивести список www.fishing-	16	www.posolstva.org.ua	Максимальна кількість	www.odissey.kiev.ua
17. www.uahotels.info Середня кількість текстових фрагментів www.zvetsad.com.ua 18. мим.оsvita.ua Мінімальна кількість графічних фрагментів www.auto-store.kiev.ua 19. кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. мим.оstriv.in.ua Вивести список www.fishing-	10.		текстових фрагментів	
18. текстових фрагментів www.auto-store.kiev.ua 19.	17	www.uahotels.info		www.zvetsad.com.ua
18. www.osvita.ua графічних фрагментів 19. Кількість текстових фрагментів по кожному документу www.tennismag.com.ua 20. www.ostriv.in.ua Вивести список www.fishing-	1/.		текстових фрагментів	
трафічних фрагментів кількість текстових www.tennismag.com.ua 19. фрагментів по кожному документу 20. www.ostriv.in.ua Вивести список www.fishing-	18.	www.osvita.ua	Мінімальна кількість	www.auto-store.kiev.ua
19. www.shkola.ua фрагментів по кожному документу 20. www.ostriv.in.ua Вивести список www.fishing-			графічних фрагментів	
документу 20. www.ostriv.in.ua Вивести список www.fishing-	19.	www.shkola.ua	Кількість текстових	www.tennismag.com.ua
20. www.ostriv.in.ua Вивести список www.fishing-			фрагментів по кожному	
20. www.ostriv.in.ua			документу	
гіперпосилань <u>mart.com.ua</u>	20.	www.ostriv.in.ua	Вивести список	www.fishing-
			гіперпосилань	mart.com.ua

Протокол лабораторної роботи має містити: титульний аркуш, варіант студента, вміст файлу XSL, текст програми, 2-3 копії екранних форм (screenshots) у інверсному поданні (білий фон), а також вміст XHTML-файлу.

Контрольні запитання

- 1. Дати визначення well-formed та коректному XML-документу.
- 2. Назвати способи обробки ХМL-документів.
- 3. Охарактеризувати мову XPath.
- 4. Охарактеризувати мову XSLT.