

# *Інформаційна технологія маніпулювання вмістом XML-файлів*

---

**ВИКОНАВ**

*СІМОНЕНКО ДМИТРО  
ІВАНОВИЧ*

**КЕРІВНИК:**

*МЕСЮРА ВОЛОДИМИР*

**Мета дослідження** – зменшення часу пошуку необхідного фрагменту та виконання необхідних маніпуляцій із вмістом XML-файлу.

**Об'єкт дослідження** – процеси документообігу на підприємстві із використанням сучасних інформаційних технологій.

**Предмет дослідження** – методи та засоби аналізу XML-файлів.

## Задачі дослідження:

- провести аналіз задачі використання даних аналізу XML-файлів на сучасному етапі розвитку інформаційних технологій;
- розглянути існуючі методи вирішення задачі аналізу XML-файлів та обрати й обґрунтувати вибір методу, який задовольняє меті даної магістерської кваліфікаційної роботи;
- вдосконалити метод розбору структури XML-файлів та удосконалити його згідно мети магістерської кваліфікаційної роботи;
- розробити структуру та алгоритм роботи програмного засобу маніпулювання вмістом XML-файлів;
- провести тестування розробленого програмного продукту.

## **Наукова новизна одержаних результатів полягає в такому:**

- подальший розвиток інформаційної технології роботи із структурою XML-файлів за рахунок введення етапів перевірки XML-файлу вимогам на коректність;
- удосконалено метод розбору структури XML-файлів із можливістю маніпулювання їх вмістом за рахунок комбінованого способу маніпулювання вмістом XML-файлів на основі XSLT та DTD;
- удосконалено алгоритм програми аналізу структури XML-файлів із можливістю маніпулювання їх вмістом.

## Практичне значення одержаних результатів:

- розроблено комбінований спосіб аналізу структури XML-файлів із можливістю маніпулювання їх вмістом.
- розроблено алгоритм функціонування програмного забезпечення для виконання процесів структуризації XML-файлів.
- розроблено програмний засіб для аналізу структури XML-файлів із можливістю маніпулювання його вмістом.
- Результати роботи впроваджені ФОП Швець Е. І.

## АКТУАЛЬНІСТЬ РОБОТИ

Бурхливий розвиток інформаційного суспільства призводить до стрімкого зростання інформаційних потоків, поширення безпаперових технологій, що збільшує важливість підвищення ефективності ведення електронного документообігу, продуктивності процесів пошуку необхідних фрагментів у великих XML-документах.

Саме ці фактори й обумовлюють актуальність розробленої у роботі інформаційної технологія маніпулювання вмістом XML-файлів, яка забезпечує підвищення продуктивності процесів пошуку необхідних фрагментів у великих XML-документах.

# РЕЗУЛЬТАТИ АНАЛІЗУ СТРУКТУРИ XML-ДОКУМЕНТІВ

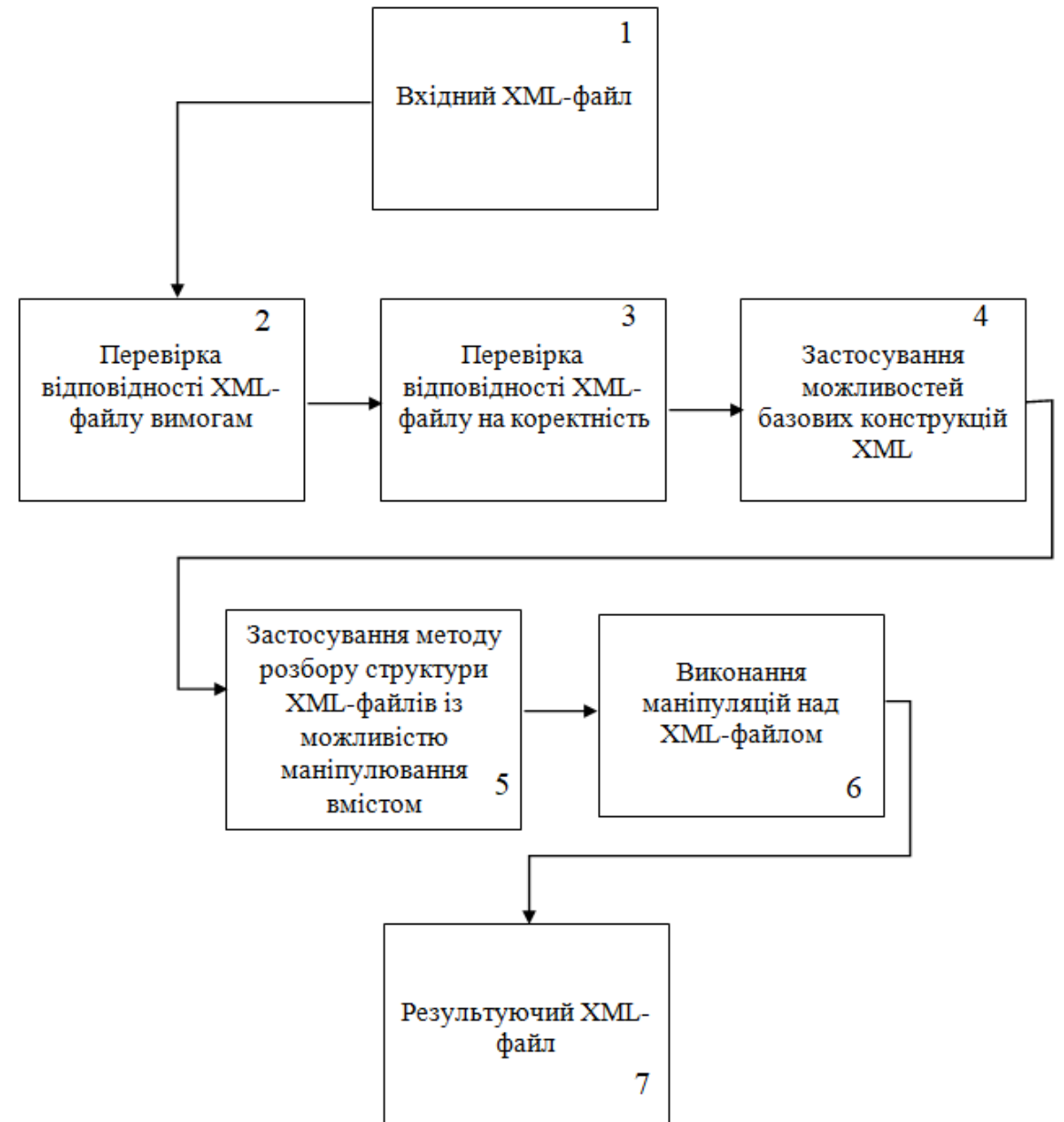
Найбільш поширеними підзадачами процесів побудові XML-подібного формату даних є:

- визначення структури документів;
- перевірка вивантажених даних на правильність.

Основними способами розв'язку таких підзадач є:

- визначення типу документу (Document Type Declaration);
- формування схеми документа (XML Schema), що є більш складним, а у деяких реалізаціях навіть надто громіздким, але й надає значно більше можливостей

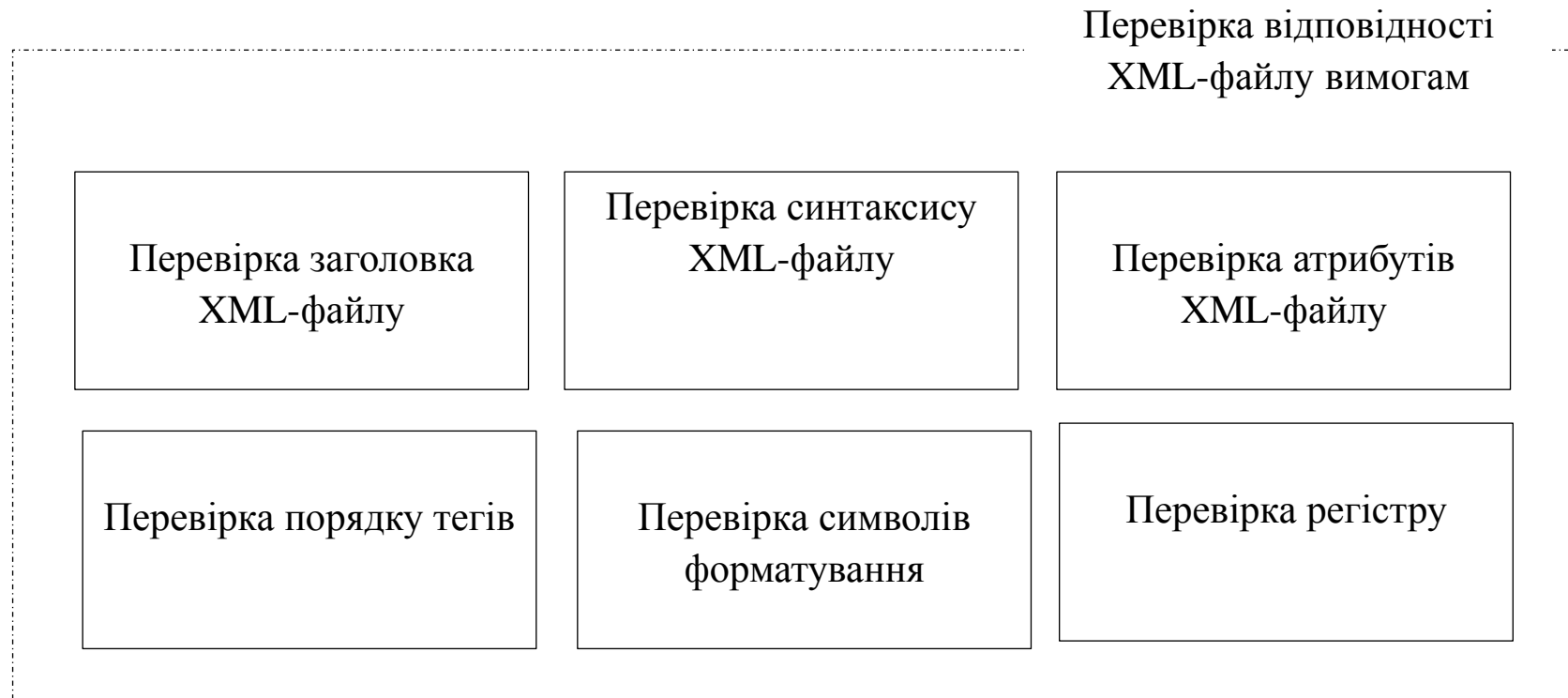
# ланцюжок маніпулювання вмістом XML- файлів





# Схема перевірки відповідності XML-файлу вимогам

---



# Схема перевірки XML-файлу на коректність

---

Перевірка відповідності XML-файлу на коректність

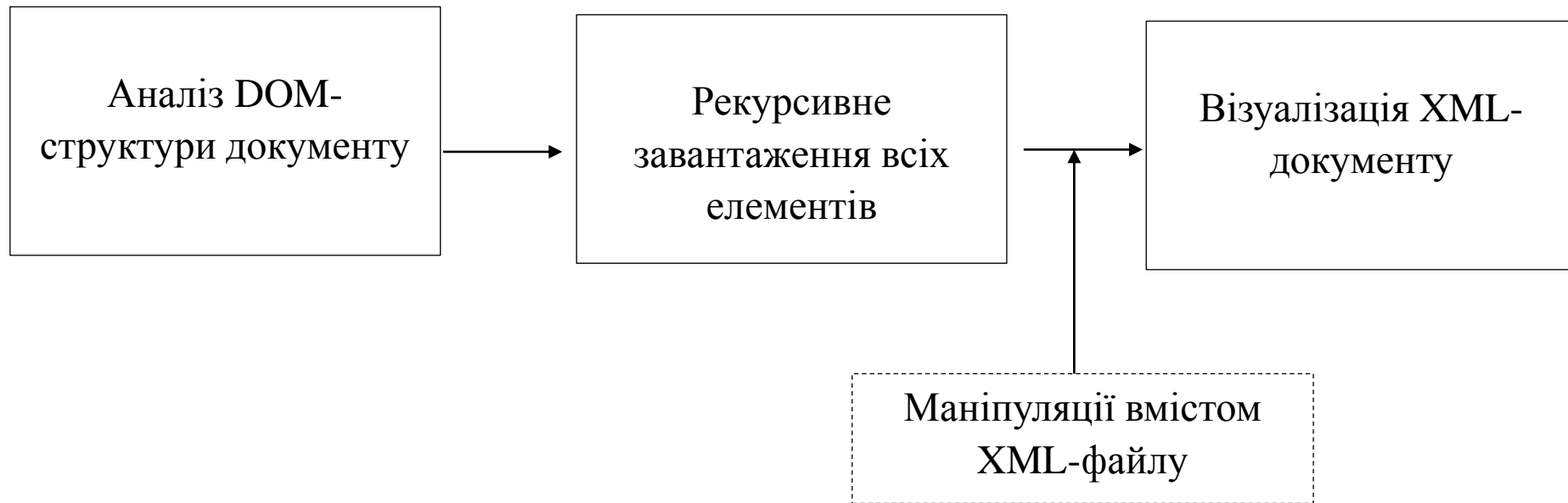
Перевірка  
структури

Відповідність  
Document Type  
Definition

Логічна структуризація

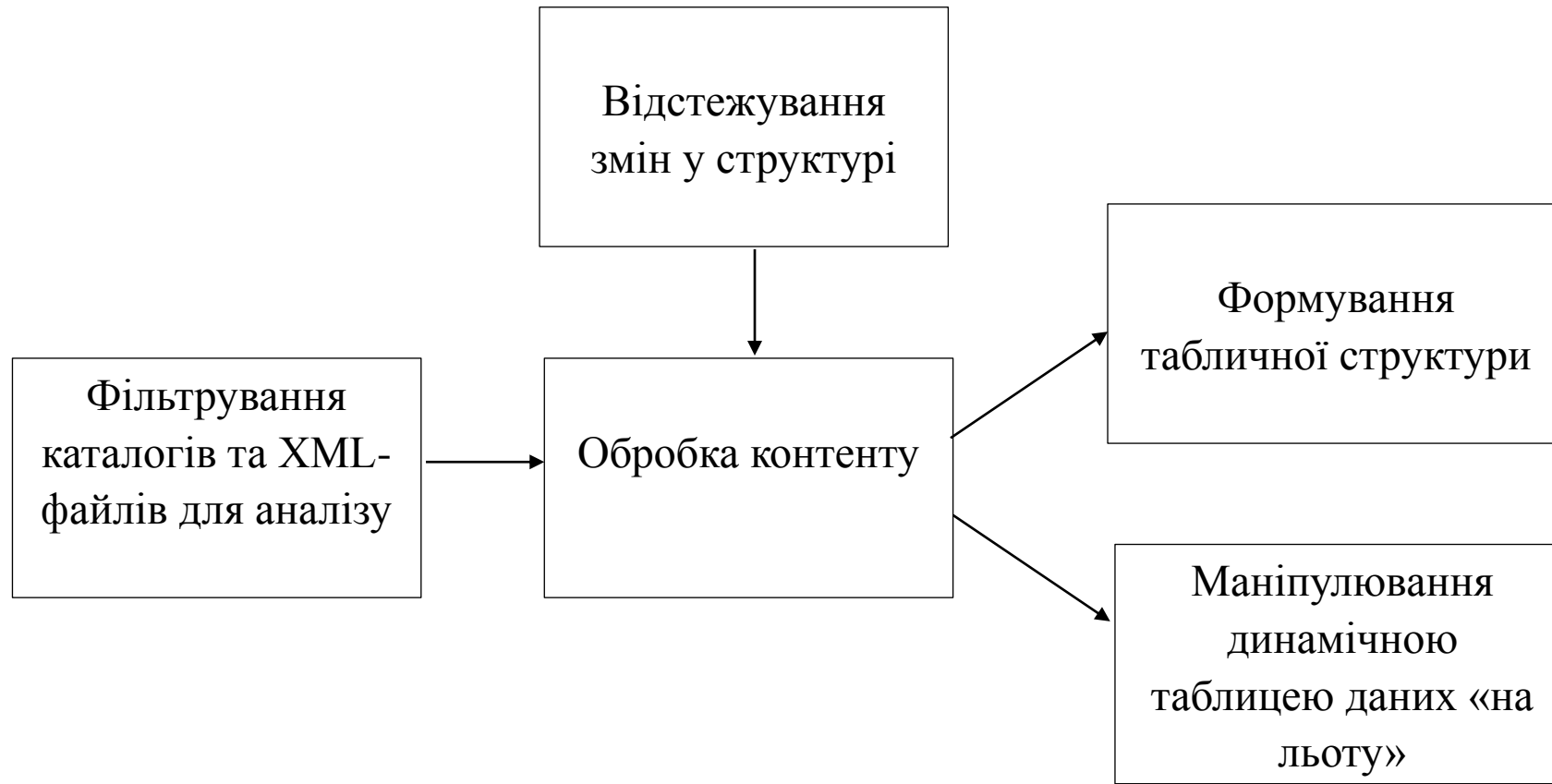
# Загальна схема методу розбору структури XML-файлів із можливістю маніпулювання їх вмістом

---



# Обробка контенту DOM-структури XML-документу

---



# Візуалізація схеми XML-документу

---

Візуалізація XML-  
документу

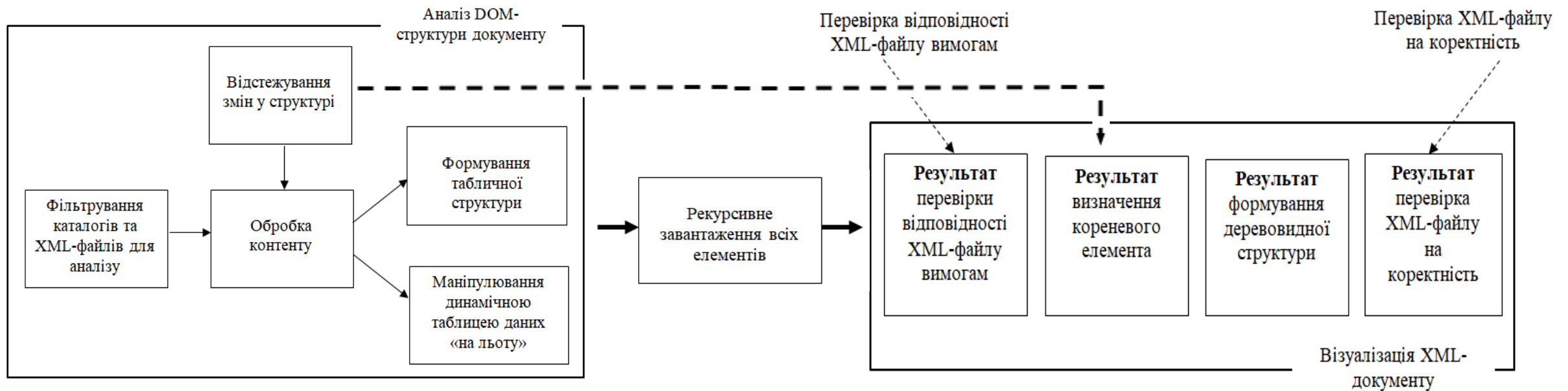
**Результат**  
перевірки  
відповідності  
XML-файлу  
вимогам

**Результат**  
визначення  
кореневого  
елемента

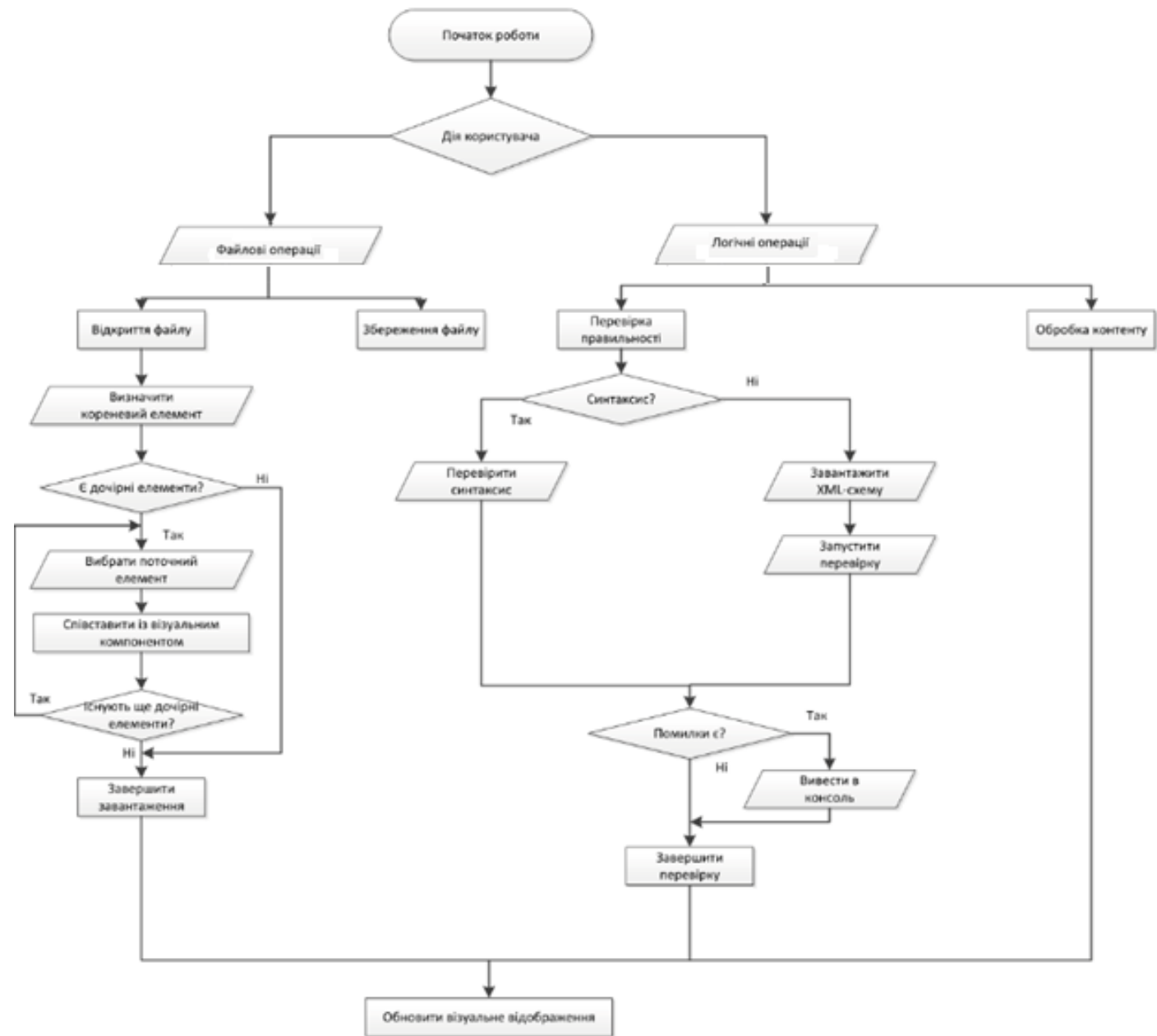
**Результат**  
формування  
деревовидної  
структури

**Результат**  
перевірка  
XML-файлу  
на коректність

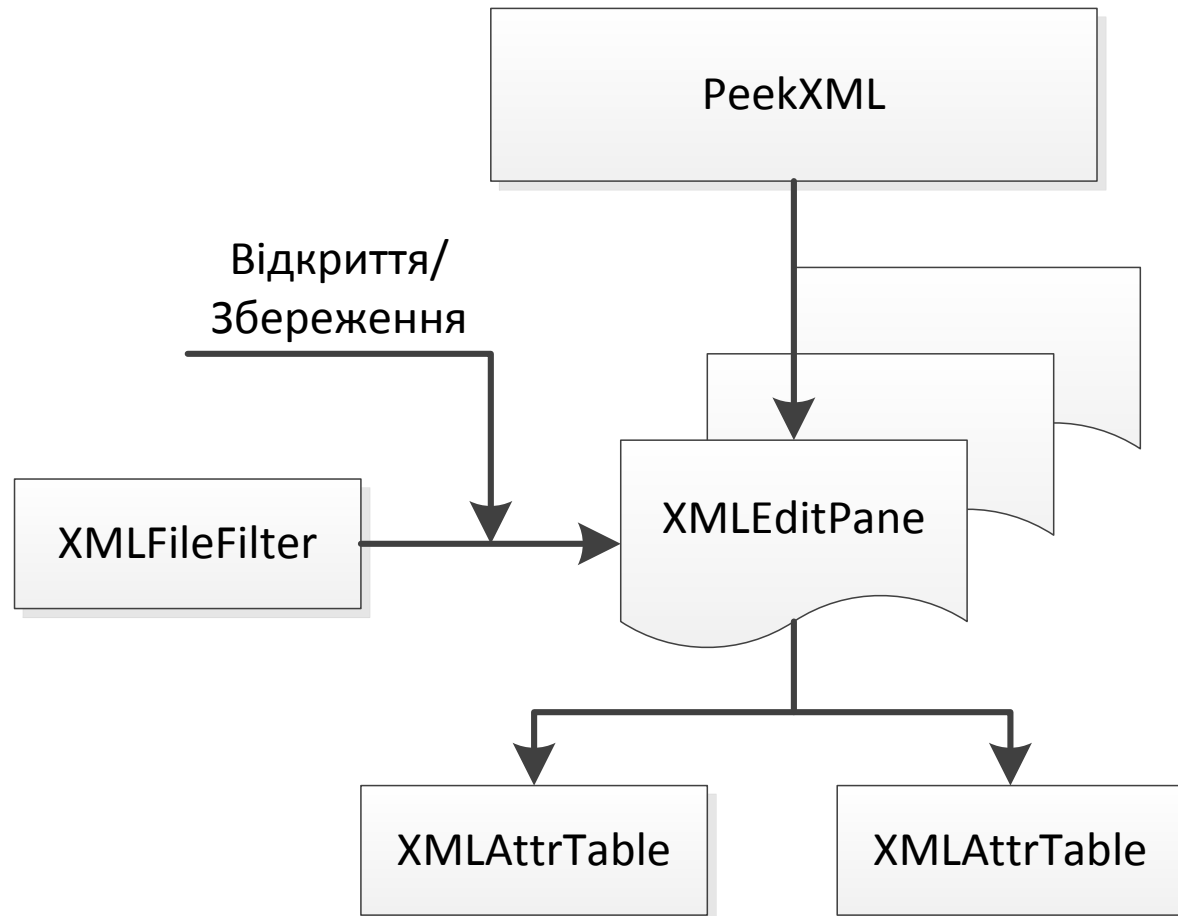
# Метод розбору структури XML-файлів із можливістю маніпулювання їх вмістом



# Алгоритм роботи програми

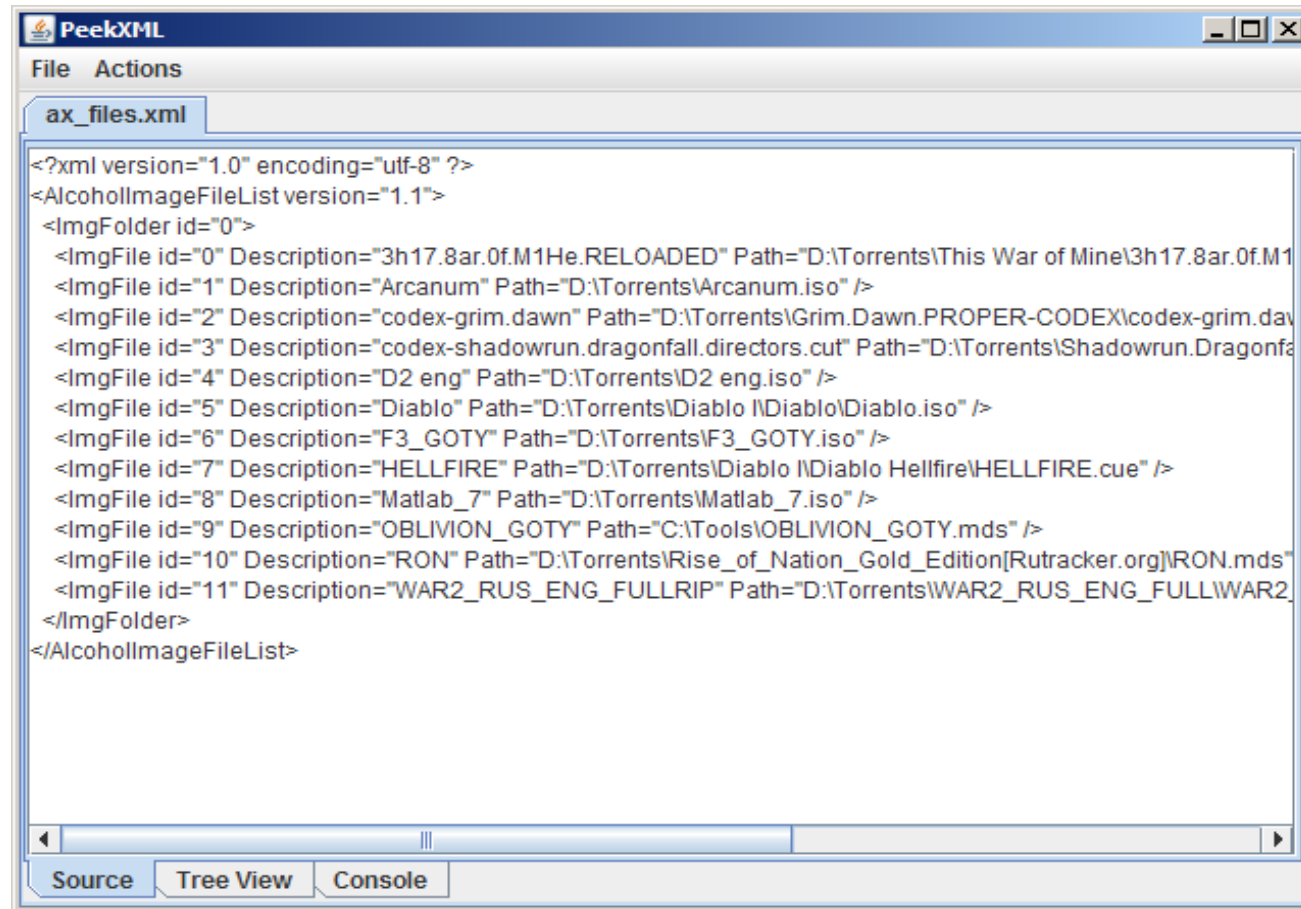


# Діаграма класів програми





# Відкриття XML-документу в режимі редагування

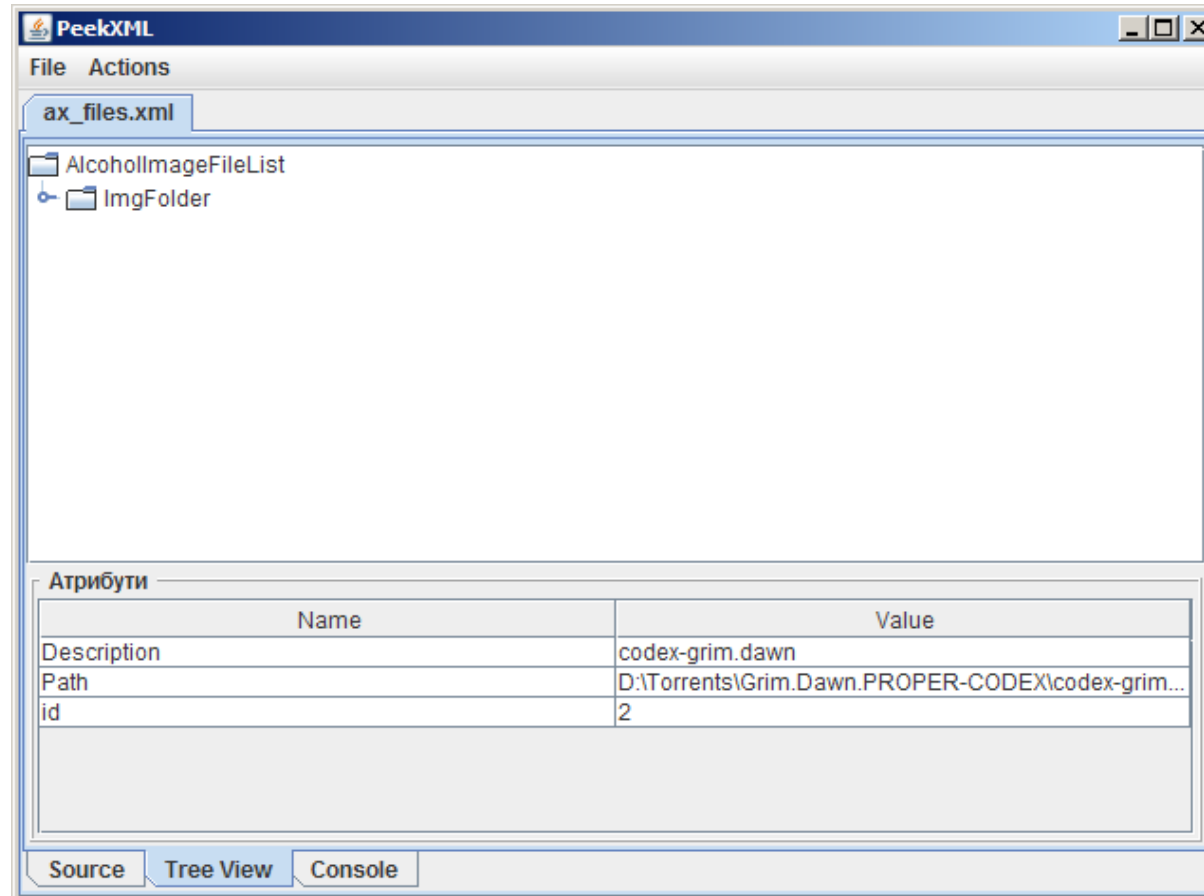


The screenshot shows the PeekXML application window with the title bar 'PeekXML'. The menu bar contains 'File' and 'Actions'. The active tab is 'ax\_files.xml'. The main text area displays the following XML code:

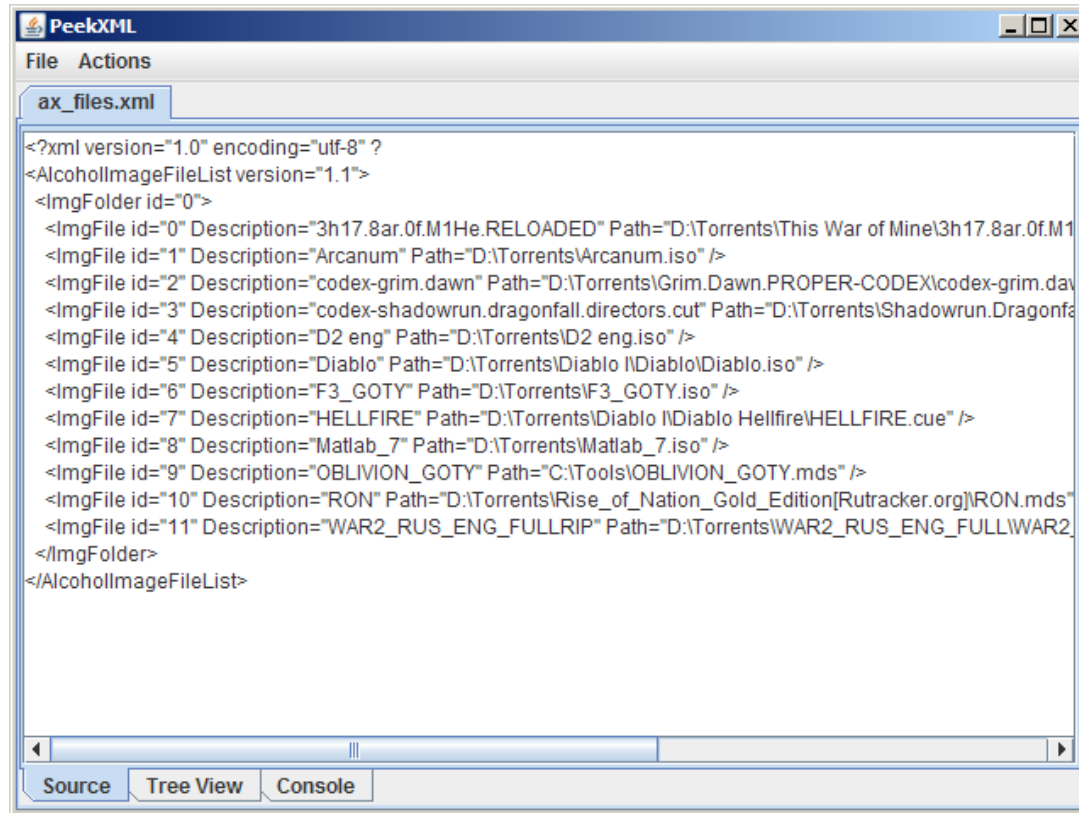
```
<?xml version="1.0" encoding="utf-8" ?>
<AlcoholImageFileList version="1.1">
  <ImgFolder id="0">
    <ImgFile id="0" Description="3h17.8ar.0f.M1He.RELOADED" Path="D:\Torrents\This War of Mine\3h17.8ar.0f.M1
    <ImgFile id="1" Description="Arcanum" Path="D:\Torrents\Arcanum.iso" />
    <ImgFile id="2" Description="codex-grim.dawn" Path="D:\Torrents\Grim.Dawn.PROPER-CODEX\codex-grim.daw
    <ImgFile id="3" Description="codex-shadowrun.dragonfall.directors.cut" Path="D:\Torrents\Shadowrun.Dragonfa
    <ImgFile id="4" Description="D2 eng" Path="D:\Torrents\D2 eng.iso" />
    <ImgFile id="5" Description="Diablo" Path="D:\Torrents\Diablo \Diablo\Diablo.iso" />
    <ImgFile id="6" Description="F3_GOTY" Path="D:\Torrents\F3_GOTY.iso" />
    <ImgFile id="7" Description="HELLFIRE" Path="D:\Torrents\Diablo \Diablo Hellfire\HELLFIRE.cue" />
    <ImgFile id="8" Description="Matlab_7" Path="D:\Torrents\Matlab_7.iso" />
    <ImgFile id="9" Description="OBLIVION_GOTY" Path="C:\Tools\OBLIVION_GOTY.mds" />
    <ImgFile id="10" Description="RON" Path="D:\Torrents\Rise_of_Nation_Gold_Edition[Rutracker.org]\RON.mds"
    <ImgFile id="11" Description="WAR2_RUS_ENG_FULLRIP" Path="D:\Torrents\WAR2_RUS_ENG_FULL\WAR2
  </ImgFolder>
</AlcoholImageFileList>
```

At the bottom of the window, there are three tabs: 'Source' (selected), 'Tree View', and 'Console'. A scrollbar is visible on the right side of the text area.

# Перегляд XML-файлу в режимі структурної візуалізації

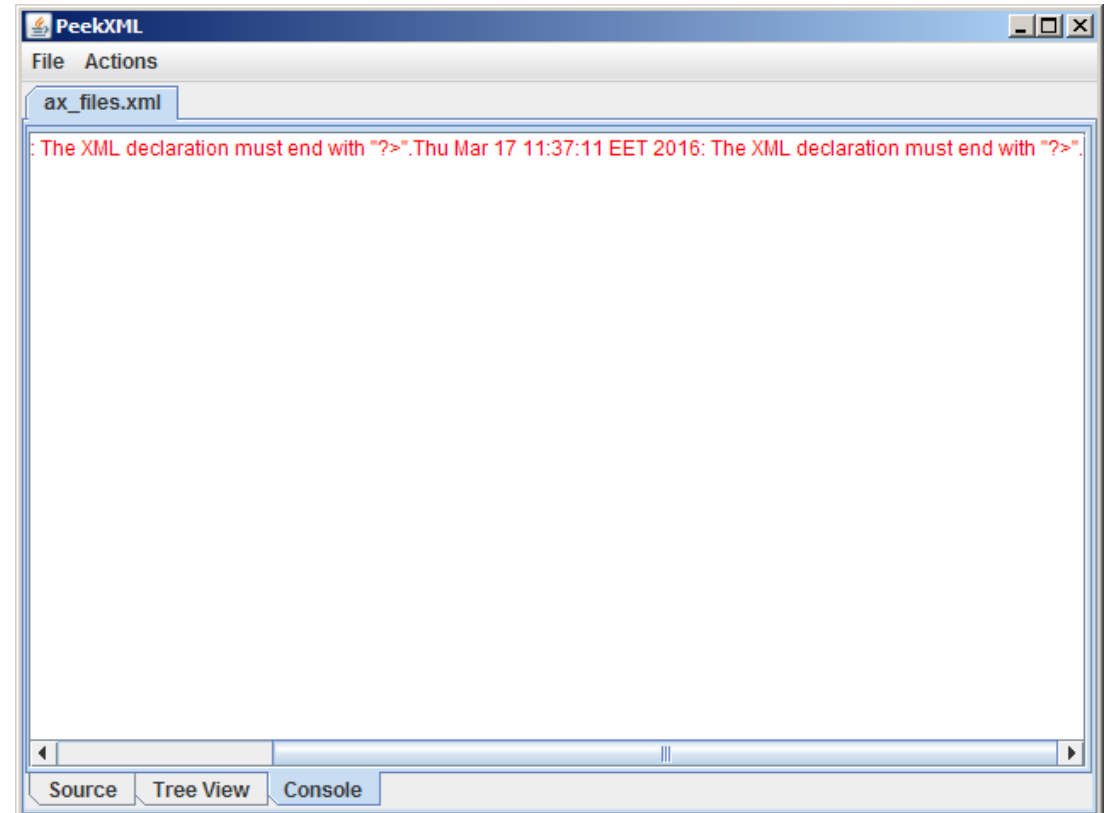


# Відображення помилки синтаксису



The screenshot shows the PeekXML application window with the file 'ax\_files.xml' open. The XML content is displayed in the main pane, and the 'Source' tab is selected at the bottom. The XML is well-formed and valid.

```
<?xml version="1.0" encoding="utf-8" ?>
<AlcoholImageFileList version="1.1">
  <ImgFolder id="0">
    <ImgFile id="0" Description="3h17.8ar.0f.M1He.RELOADED" Path="D:\Torrents\This War of Mine\3h17.8ar.0f.M1" />
    <ImgFile id="1" Description="Arcanum" Path="D:\Torrents\Arcanum.iso" />
    <ImgFile id="2" Description="codex-grim.dawn" Path="D:\Torrents\Grim.Dawn.PROPER-CODEX\codex-grim.daw" />
    <ImgFile id="3" Description="codex-shadowrun.dragonfall.directors.cut" Path="D:\Torrents\Shadowrun.Dragonfa" />
    <ImgFile id="4" Description="D2 eng" Path="D:\Torrents\D2 eng.iso" />
    <ImgFile id="5" Description="Diablo" Path="D:\Torrents\Diablo I\Diablo\Diablo.iso" />
    <ImgFile id="6" Description="F3_GOTY" Path="D:\Torrents\F3_GOTY.iso" />
    <ImgFile id="7" Description="HELLFIRE" Path="D:\Torrents\Diablo II\Diablo Hellfire\HELLFIRE.cue" />
    <ImgFile id="8" Description="Matlab_7" Path="D:\Torrents\Matlab_7.iso" />
    <ImgFile id="9" Description="OBLIVION_GOTY" Path="C:\Tools\OBLIVION_GOTY.mds" />
    <ImgFile id="10" Description="RON" Path="D:\Torrents\Rise_of_Nation_Gold_Edition[Rutracker.org]\RON.mds" />
    <ImgFile id="11" Description="WAR2_RUS_ENG_FULLLRIP" Path="D:\Torrents\WAR2_RUS_ENG_FULL\WAR2" />
  </ImgFolder>
</AlcoholImageFileList>
```



The screenshot shows the PeekXML application window with the file 'ax\_files.xml' open. A red error message is displayed at the top of the main pane, indicating a syntax error in the XML declaration. The 'Source' tab is selected at the bottom.

```
The XML declaration must end with "?>". Thu Mar 17 11:37:11 EET 2016: The XML declaration must end with "?>".
```

## ПУБЛІКАЦІЇ ПО ТЕМІ РОБОТИ

Сімоненко Д. В. Інформаційна технологія маніпулювання вмістом XML-файлів / Д. В. Сімоненко // «Наукові підсумки 2017 року» : матеріали XV Міжнародної інтернет-конференції «Наукові підсумки 2017 року», 15 грудня 2017 р. м. Вінниця. – С. 63-68

Сімоненко Д. В. Комбінований спосіб маніпулювання вмістом XML-файлів на основі XSLT та DTD / Д.В. Сімоненко, В.І. Месюра // Регіональна науково-практична конференція студентів, аспірантів та молодих науковців "Молодь в науці: дослідження, проблеми, перспективи" (МН-2018). - Вінниця: ВНТУ, 2017. - Режим доступу:

<https://conferences.vntu.edu.ua/index.php/mn/mn/paper/viewFile/3597/3049>

# ОСНОВНІ РЕЗУЛЬТАТИ РОБОТИ

Розроблено та програмно реалізовано вдосконалену інформаційну технологію маніпулювання вмістом XML-файлів, яка дозволила збільшити швидкість пошуку фрагменту XML-файлу завдяки використанню запропонованого методу розбору структури XML-файлів із можливістю маніпулювання їх вмістом, покращеному за рахунок розробки комбінованого способу маніпулювання вмістом XML-файлів на основі XSLT та DTD.

Здійснено аналіз задачі визначення структури документів, результати якого дозволили вдосконалити інформаційну технологію і методи розбору структури XML-файлів за рахунок врахування процесів перевірки вивантажених даних на правильність та коректність структури документу, що є ключовим під час аналізу XML-файлів.

Запропоновано комбінований підхід, при якому застосування механізмів аналізу структури XML-файлів на основі XSLT-перетворення та формату DTD, дозволило збільшити швидкість пошуку фрагменту XML-файлу.

Результати роботи апробовані на двох конференціях та впроваджені в процесі організації документообігу на підприємстві ФОП Швець Е. І.

Дякую за увагу!