



r-markers>

character name="evelyn_ellis" total-hits="259" distinct-

#prominent-eyelashes #uniform

<character name="frankie" total-hits="239" distinct-

#long-hair #high-heeled-shoes #skirt #wingtips #dress-gloves #full-cup_bra

<character name="evelyn_ellis" total-hits="225" distinct-markers="12">#prominent-breasts #dress #long-hair #earring #dress-gloves #high-heeled-shoes #lipstick

</character>

<character name="frankie" total-hits="215" distinct-markers="15">#suit #wingtips #fedora #train-agent-hat #bib-overall

<gender-markers>

<female total-count="14" number_of_characters="14">

#prominent-breasts #uniform #long-hair #high-heeled-shoes #wingtips #dress-gloves #full-cup_bra #uniform-jockey #mink-coat

TE1/CBML ANALYSIS

XQUERY



A TEI/CBML FILE IS AN XML DOCUMENT, AND CAN BE MANIPULATED LIKE ANY OTHER COMPUTER FILE USING SCRIPTING AND XML EDITORS. FOR EXAMPLE, A SOPHISTICATED XML EDITOR LIKE OXYGEN, A COMMERCIAL TOOL USED BY PROGRAMMING PROFESSIONALS, CAN REVEAL QUITE A LOT ABOUT THE CONTENTS USING BUILT-IN SEARCH AFFORDANCES. BY SEARCHING FOR "#PHANTOM_LADY" AS AN ATTRIBUTE VALUE, FOR EXAMPLE, OXYGEN FINDS 95 HITS, IN BOTH THE CBML: PANEL "CHARACTERS" AND CBML: BALLOON "WHO" ATTRIBUTES.

The screenshot shows the Oxygen XML Editor interface. On the left, the 'Open/Find Resource' panel is open, showing 'In file paths' selected. The main workspace displays two files: 'TEI_Data-TEIHeader-csv.xml' and 'PL17_TEI.xml'. The 'PL17_TEI.xml' file is currently active, showing XML code with several lines highlighted in yellow. To the right of the code, there's a search and replace dialog. The 'Replace with:' field contains the placeholder text 'PH...'. Below it, the 'XPath:' field is set to 'Type XPath expression'. Under 'Direction', 'Forward' is selected. Under 'Scope', 'All' is selected. In the 'Options' section, 'Case sensitive', 'Wrap around', and 'Enable XML search options <<' are checked. The 'Search only in' section includes checkboxes for 'Element names', 'Element contents', 'Attribute names', and 'Attribute values', with 'Attribute values' being checked. Buttons for 'Find All', 'Replace All', and 'Replace to End' are visible at the top right of the dialog.

Open/Find Resource

In file paths In content In reviews

You should consider re-indexing.

Dexed: 0 Reindex

Project Open/Find Re... Archive Brow...

Data Source Explorer

- analysis-tools
- backgroundReading
- curriculum
- data
 - gender-markers.xml
 - PL17_TEI.xml

description - 42 items

```

oxygen:/eXist-db localhost (3)$eXist-db localhost/db/apps/teicomix/data/PL17_TEI.xml (42 items)
<cbml:panel n="1.1.1" type="title" characters="#phantom_lady #anonymous_criminal">
<cbml:panel n="1.5.1" ana="#subject-to-subject" characters="#phantom_lady">
<cbml:panel n="1.5.2" ana="#action-to-action" characters="#phantom_lady #betty #don_borden">
<cbml:panel n="1.5.3" ana="#action-to-action" characters="#phantom_lady">
<cbml:panel n="1.7.1" ana="#scene-to-scene" characters="#phantom_lady">
<cbml:panel n="1.7.3" ana="#action-to-action" characters="#phantom_lady">
<cbml:panel n="1.7.4" ana="#action-to-action" characters="#phantom_lady">

```

Replace with:

Find All Replace All Replace to End

XPath: Type XPath expression

Direction: Forward Scope: All

Options

Case sensitive Whole words only

Incremental Regular expression

Wrap around Dot matches all

Canonical equivalence

Enable XML search options >>

42 matches found Close

USING A SIMPLE REGULAR EXPRESSION [A PROGRAMMER'S TOOL FOR SEARCHING CODE], WE DETERMINE THAT "#PHANTOM_LADY" OCCURS 42 TIMES AS CBML: PANEL "CHARACTERS" ATTRIBUTE VALUES, EITHER ALONE OR WITH OTHER CHARACTERS.



WHAT OXYGEN AND SIMILAR TOOLS CANNOT CREATE THROUGH QUERY STRINGS ARE SUCH THINGS AS WORD CONCORDANCES WITH USAGE FREQUENCY, OR BREAKDOWNS OF GENDER-MARKERS BY CHARACTER, OR TYPES OF SPEECH BY CHARACTER. QUERIES OF THIS COMPLEXITY CAN BE PERFORMED USING A VARIETY OF SCRIPTING LANGUAGES. SINCE TEI/CBML FILES ARE IN XML, ONE STRAIGHTFORWARD METHOD IS TO USE THE QUERY LANGUAGE CREATED BY THE W3C SPECIFICALLY FOR XML, XQUERY. IF THE TEI/CBML FILE IS SAVED IN AN XML DATABASE LIKE EXIST-DB, AN OPEN-SOURCE PROJECT THAT CAN BE INSTALLED ON YOUR COMPUTER, WHATEVER THE OPERATING SYSTEM, IT CAN BE QUERIED USING A COMMERCIAL XML EDITOR THAT SUPPORTS XQUERY, LIKE OXYGEN, OR AN OPEN-SOURCE TEXT EDITOR LIKE ATOM.

ATOM SHIPS WITH THE ABILITY TO LINK SEAMLESSLY WITH EXIST-DB, MAKING IT EXTREMELY EASY TO PROGRAM AND RUN XQUERY SCRIPTS USING A TEI/CBML FILE. ONE OF THE GREAT ADVANTAGES OF USING THIS APPROACH IS THE ABILITY TO RUN QUERIES AGAINST MULTIPLE TEI/CBML FILE -- HUNDREDS, OR THOUSANDS. SO, FOR INSTANCE, IF WE HAD ALL OF THE PHANTOM LADY COMIC BOOKS MARKED UP AS TEI/CBML FILES, WE COULD PERFORM "BIG-DATA" ANALYSES ON THEM, REVEALING CHANGES IN CHARACTERIZATION OVER THE DECADES.

TEI/CBML

```
<?xml-model
href="http://localhost:8080/exist/apps/teicomix/resources/schema/cbml.rnc"
type="application/relax-ng-compact-syntax" xmlns:cbml="http://www.cbml.org/ns/1.0"
<tei:TEI xmlns:tei="http://www.tei-c.org/ns/1.0">
<tei:teiHeader>
  <!--tei:teiHeader is where the object of analysis is described: the comic book or graphic novel being analyzed, the parties responsible for the encoding, and so on. -->
  <tei:fileDesc>
    <tei:titleStmt>
      <!-- tei:title: name of the comic book or graphic novel -->
      <tei:title>Phantom Lady</tei:title>
      <!-- tei:author: repeatable tag for listing all individuals who created or contributed to the content, whether writer, colorist, editor, etc. Use roles such as "writer" or "colorist" if applicable, and create them using a "role" attribute -->
      <tei:author role="writer">[Enter the name of the writer here]</tei:author>
      <tei:author role="colorist">[Enter the name of the colorist here] Matthew Baker</tei:author>
      <tei:editor>[If the comic or graphic novel has an editor, use this tag to record the name]</tei:editor>
      <tei:respStmt>
        <!-- party responsible for the creation and/or encoding of the TEI file -->
        <tei:responsible>[Enter the name of the person or organization responsible for the creation and/or encoding of the TEI file]</tei:responsible>
    </tei:titleStmt>
    <tei:content>
      <!-- The content of the comic book or graphic novel -->
      <tei:body>
        <!-- A single page of the comic book or graphic novel -->
        <tei:page>
          <!-- A single panel on a page -->
          <tei:panel>
            <!-- The visual content of a panel -->
            <tei:img alt="A panel from the comic book 'Phantom Lady' showing a woman in a mask." data-bbox="106 178 904 880"/>
          </tei:panel>
        </tei:page>
      </tei:body>
    </tei:content>
  </tei:fileDesc>
</tei:teiHeader>
<tei:body>
  <!-- The content of the comic book or graphic novel -->
  <tei:page>
    <!-- A single panel on a page -->
    <tei:panel>
      <!-- The visual content of a panel -->
      <tei:img alt="A panel from the comic book 'Phantom Lady' showing a woman in a mask." data-bbox="106 178 904 880"/>
    </tei:panel>
  </tei:page>
</tei:body>
</tei:TEI>
```

```
<word w="better" frequency="20">
    <character name="anonymous_criminal"/>
    <character name="dishonest_jockey"/>
    <character name="evelyn_ellis"/>
    <character name="frankie"/>
    <character name="joe"/>
    <character name="phantom_lady"/>
    <character name="sandra_knight"/>
</word>
<word w="betty" frequency="10">
    <character name="don_borden"/>
    <character name="phantom_lady"/>
    <character name="sandra_knight"/>
```

EVEN WITH ONLY
A SINGLE
COMIC BOOK,
PL#17, WE CAN
SCRIPT
XQUERIES THAT
CAN GENERATE
WORD
CONCORDANCE
S FOR EACH
OR ALL
CHARACTERS:

WE CAN SCRIPT XQUERIES THAT WILL GIVE US THE STATS ON THE GENDER-MARKERS AND OCCURRENCE BY CHARACTER:

XQUERY

```
<character name="sandra_knight" total-hits="239" distinct-markers="20">#dress  
#long-hair #bare-legs #lipstick #prominent-eyelashes #prominent-breasts #cleavage  
#high-heeled-shoes #apron #bare-arms #stockings #context-kitchen #necklace  
#blouse #skirt #dress-hat #dress-gloves #purse #full-cup_bra  
#uniform-jockey</character>  
<character name="evelyn_ellis" total-hits="225"  
distinct-markers="12">#prominent-breasts #dress #long-hair #mink-coat #necklace  
#earring #dress-gloves #high-heeled-shoes #lipstick #bare-legs #prominent-eyelashes  
</character>  
<character name="frankie" total-hits="115" distinct-markers="8">#short-hair  
#necktie #suit #wingtips #fedora #train-agent-hat #bib-overalls #shirt</character>  
<character name="anonymous_crimal" total-hits="103"  
distinct-markers="9">#short-hair #fedora #suit #shirt #necktie #bowtie #facial-hair  
#pants #wingtips</character>
```

WE CAN SCRIPT XQUERIES THAT PROFILE THE STATISTICS AND TYPES OF SPEECHES MADE BY EACH CHARACTER, AN IMPORTANT YARDSTICK FOR DEVELOPING A MEANINGFUL PERSONALITY PROFILE, USED IN CONJUNCTION WITH IBM WATSON PERSONALITY INSIGHTS:

```
<speaker>#sandra_knight</speaker>  
<count>46</count>  
<type-of-speech>  
    <dialogue>31</dialogue>  
    <monologue>1</monologue>  
    <interjection>0</interjection>  
    <thought>14</thought>  
    <command>0</command>  
    <narration>0</narration>  
</type-of-speech>  
</speeches>  
<speeches>  
    <speaker>#phantom_lady</speaker>  
    <count>44</count>  
    <type-of-speech>  
        <dialogue>15</dialogue>  
        <monologue>0</monologue>  
        <interjection>0</interjection>  
        <thought>20</thought>  
        <command>0</command>
```

WE CAN SEE AT A GLANCE THE STATISTICAL SPREAD OF THE PANEL TRANSITION TYPES IN PL#17:

```
<panel-transitions>
  <stats>
    <transitions>#subject-to-subject #action-to-action #scene-to-scen
    #aspect-to-aspect #moment-to-moment</transitions>
    <total>200</total>
    <moment-to-moment total="1"/>
    <action-to-action total="116"/>
    <subject-to-subject total="58"/>
    <scene-to-scene total="24"/>
    <aspect-to-aspect total="1"/>
    <non-sequitur total="0"/>
  </stats>
```

IT IS TRIVIALLY EASY TO GENERATE A FILE OF ALL OF THE SPEECHES FOR A CHARACTER, WHICH CAN BE COPIED INTO OTHER TOOLS, LIKE IBM WATSON PERSONALITY INSIGHTS OR TAPORWARE TOOLS FOR CORPUS LINGUISTIC ANALYSIS:

IF YOU ARE INTERESTED IN USING THESE TECHNOLOGIES FOR YOUR FINAL PROJECT, CONTACT THE INSTRUCTORS.

every-results

```
<results>SIT DOWN CHOWDER HEAD BEFORE I SCRATCH THOSE SQUINTY EYES OUT OF THAT THICK HEAD! SHUT UP, FRANKIE. I DARE YOU TO PULL THAT TRIGGER! YOU HAVEN'T THE GUTS, YA CHEAP PHONEY! GO ON SHOOT! OKAY, NOW THAT'S SETTLED. SIT DOWN AND I'LL TELL YA WHAT I GOT ON MY MIND. I GOTTA KNOW ONE THING FIRST. ARE YA WITH ME? OKAY, THEN LISTEN. THE JOB I GOT IN MIND IS WORTH SEVENTY FIVE GRAND! WE'RE GONNA KNOCK OFF THE UNION PACIFIC EXPRESS. THEY CARRY MORE REGISTERED MAIL THAN THE POST OFFICE AND BEFORE YA OPEN YA YAPS, IT AIN'T AS HARD AS YA THINK .... GET OUT OF THE CHAIR POP BEFORE IT BECOMES YA PERMANENT RESTING PLACE! SORRY POP, WE AIN'T AIMIN' TO BE INTERRUPTED! SHUT THAT TRAP, WISE GUY, OR I'LL DO IT FOR YA! I BETTER LET HIM HAVE A LUMP TOO, WE CAN'T TAKE ANY CHANCES. I HOPE YOU GOT A HARD HEAD, STUPID. I DON'T WANT TO KNOCK YOUR BRAINS OUT. YOU GOT LITTLE ENOUGH AS IT IS! THROW THEM INTO THE TICKET BOOTH WATCH THEM EVERY SECOND, JOE,
```

LINK TO EXIST-DB DOWNLOAD: [HTTPS://BINTRAY.COM/EXISTDB/RELEASES/EXIST/3.4.1/VIEW](https://bintray.com/existdb/releases/EXIST/3.4.1/view)

LINK TO ATOM DOWNLOAD: [HTTPS://ATOM.IO](https://atom.io)