

Part A:

Url:

<http://lyle.smu.edu/~rmandhare/cse7347/movieland/A3.php>

<http://lyle.smu.edu/~rmandhare/cse7347/movieland/movies.xml>



← → ↻ file.smu.edu/~rmandhare/cse7347/movieland/movies2.xml 🔍 ☆ 0

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
▼ <movies>
  <title rm1="Good Will Hunting"></title>
  <year>1997</year>
  <country>USA</country>
  <genre>Drama</genre>
  ▼ <director>
    <firstname>Gus</firstname>
    <lastname>sant</lastname>
    <born>1952</born>
  </director>
  ▼ <actor>
    <firstname>Matt</firstname>
    <lastname>Damon</lastname>
    <role>Main character</role>
  </actor>
  ▼ <summary xml:lang="en">
    A touching tale of a wayward young man who struggles to find his identity, living in a world where he can solve any problem, except the one brewing deep within himself, until one day he meets his soul mate who opens his mind and his heart.
  </summary>
  <length>126 min</length>
  <location>Massachusetts Turnpike</location>
  <title rm2="The Truman Show"></title>
  <year>1998</year>
  <country>USA</country>
  <genre>Drama</genre>
  ▼ <director>
    <firstname>Peter</firstname>
    <lastname>Weir</lastname>
    <born>1944</born>
  </director>
  ▼ <actor>
    <firstname>Jim</firstname>
    <lastname>Carrey</lastname>
    <role>Main character</role>
  </actor>
  ▼ <summary xml:lang="en">
```

.rng File

```
<?xml version="1.0" encoding="UTF-8"?>
<grammar ns="" xmlns="http://relaxng.org/ns/structure/1.0"
datatypeLibrary="http://www.w3.org/2001/XMLSchema-datatypes">
  <start>
    <element name="movies">
      <oneOrMore>
        <element name="movie">
          <attribute name="id">
            <text/>
          </attribute>
          <element name="title">
            <text/>
          </element>
          <element name="year">
            <data type="gYear">
              <param name="maxInclusive">2020</param>
              <param name="minInclusive">1800</param>
            </data>
          </element>
          <element name="country">
            <text/>
          </element>
          <element name="genre">
            <text/>
          </element>
          <element name="director">
```

```

        <ref name="lastname"/>
        <ref name="firstname"/>
        <element name="born">
            <data type="gYear" />
        </element>
    </element>
    <zeroOrMore>
        <element name="actor">
            <ref name="firstname"/>
            <ref name="lastname"/>
            <element name="role">
                <text/>
            </element>
        </element>
    </zeroOrMore>
    <element name="summary">
        <optional>
            <attribute>
                <name>xml:lang</name>
                <text/>
            </attribute>
        </optional>
        <text/>
    </element>
    <element name="length">
        <data type="integer"/>
    </element>
    <element name="location">
        <text/>
    </element>
</element>
</oneOrMore>
</element>
</start>
<define name="lastname">
    <element name="lastname">
        <text/>
    </element>
</define>
<define name="firstname">
    <element name="firstname">
        <text/>
    </element>
</define>
</grammar>

```

Part B:

PHP code:

```
<?php
```

```
$now = new DateTime();
$fh = fopen('alias.txt','r');

$doc1 = new DOMDocument();
$doc1->load('A3Merged.xml');

while($line=fgets($fh)) {
    $url = "http://lyle.smu.edu/$line";

    $ch = curl_init($url);
    curl_setopt_array($ch, array(
        CURLOPT_FOLLOWLOCATION => TRUE,
        CURLOPT_RETURNTRANSFER => TRUE,
    ));
    curl_exec($ch);

    $realUrl = curl_getinfo($ch, CURLINFO_EFFECTIVE_URL);

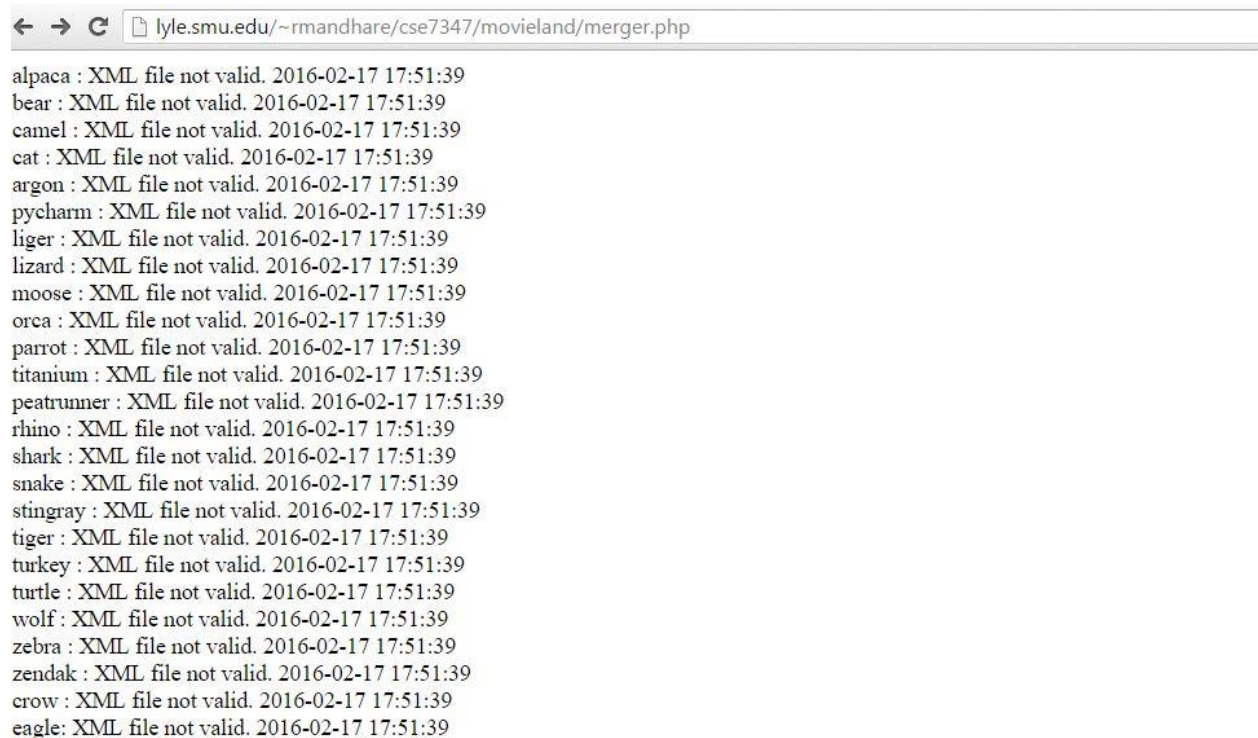
    $doc2 = new DOMDocument();
    $doc2->load($realUrl);

    $handle = curl_init($realUrl);
    curl_setopt($handle, CURLOPT_RETURNTRANSFER, TRUE);
    $response = curl_exec($handle);
    $statusCode = curl_getinfo($handle, CURLINFO_HTTP_CODE);
    if($statusCode==404 || $statusCode==301 || $statusCode==403){
        $b = false;
    }
    curl_close($handle);
    $b = true;

    if($b){
        if($doc2->relaxNGValidate("A3.rng")){
            $root = $doc1->documentElement;
            $movie = $doc2->getElementsByTagName('movie');
            for ($i = 0; $i < $movie->length; $i++) {
                $item = $movie->item($i);
                $item1 = $doc1->importNode($item, true);
                $root->appendChild($item1);
            }
            $doc1->save('A3Merged.xml');
        }
        else{
            echo "$line: XML file not valid. ".$now->format('Y-m-d H:i:s')."<br>";
        }
    }
    else{
        echo "$line: Could not find file. ".$now->format('Y-m-d H:i:s')."<br>";
    }
}
fclose($fh);?>
```

All Merge File - <http://lyle.smu.edu/~rmandhare/movieland/A3Merged.xml>

Output:



The screenshot shows a web browser window with the address bar displaying `lyle.smu.edu/~rmandhare/cse7347/movieland/merger.php`. The main content area lists 25 items, each followed by the message "XML file not valid. 2016-02-17 17:51:39". The items are: alpaca, bear, camel, cat, argon, pycharm, liger, lizard, moose, orca, parrot, titanium, peatranner, rhino, shark, snake, stingray, tiger, turkey, turtle, wolf, zebra, zendak, crow, and eagle.

```
alpaca : XML file not valid. 2016-02-17 17:51:39
bear : XML file not valid. 2016-02-17 17:51:39
camel : XML file not valid. 2016-02-17 17:51:39
cat : XML file not valid. 2016-02-17 17:51:39
argon : XML file not valid. 2016-02-17 17:51:39
pycharm : XML file not valid. 2016-02-17 17:51:39
liger : XML file not valid. 2016-02-17 17:51:39
lizard : XML file not valid. 2016-02-17 17:51:39
moose : XML file not valid. 2016-02-17 17:51:39
orca : XML file not valid. 2016-02-17 17:51:39
parrot : XML file not valid. 2016-02-17 17:51:39
titanium : XML file not valid. 2016-02-17 17:51:39
peatrunner : XML file not valid. 2016-02-17 17:51:39
rhino : XML file not valid. 2016-02-17 17:51:39
shark : XML file not valid. 2016-02-17 17:51:39
snake : XML file not valid. 2016-02-17 17:51:39
stingray : XML file not valid. 2016-02-17 17:51:39
tiger : XML file not valid. 2016-02-17 17:51:39
turkey : XML file not valid. 2016-02-17 17:51:39
turtle : XML file not valid. 2016-02-17 17:51:39
wolf : XML file not valid. 2016-02-17 17:51:39
zebra : XML file not valid. 2016-02-17 17:51:39
zendak : XML file not valid. 2016-02-17 17:51:39
crow : XML file not valid. 2016-02-17 17:51:39
eagle : XML file not valid. 2016-02-17 17:51:39
```

Part C:

PHP code:

```
<?php
ini_set('display_errors',1);

$query = $_GET['query'];
$values = explode(':', $query);

$xml = simplexml_load_file("A3Merged.xml");

if($values[0] == 'director'){
//values[1]
```

```

$resultArr1 = $xml-
>xpath("/movies/movie/director[lastname='$values[1]']/preceding-
sibling::title");
$resultArr2 = $xml-
>xpath("/movies/movie/director[lastname='$values[1]']/preceding-
sibling::year");
foreach($resultArr1 as $result){
    print "Title: $result <br/> ";
}

foreach($resultArr2 as $result) {
    print "Year: $result <br/> ";
}

}

else if ($values[0] == 'actor'){
$resultArr1 = $xml-
>xpath("/movies/movie/actor[lastname='$values[1]']/preceding-
sibling::title");
$resultArr2 = $xml-
>xpath("/movies/movie/actor[lastname='$values[1]']/preceding-sibling::year");
foreach($resultArr1 as $result){
    print "Title: $result <br/> ";
}

foreach($resultArr2 as $result) {
    print "Year: $result <br/> ";
}

}

else if ($values[0] == 'time'){
$resultArr1 = $xml->xpath("/movies/movie[year>=$values[1] and
year<=$values[2]]/title");
$resultArr2 = $xml->xpath("/movies/movie[year>=$values[1] and
year<=$values[2]]/year");
foreach($resultArr1 as $result){
    print "Title: $result <br/> ";
}

foreach($resultArr2 as $result) {
    print "Year: $result <br/> ";
}

}

else{
echo " Not valid query" ;
}

?>

```

Output:

← → ↻ lyle.smu.edu/~rmandhare/movieland/partc.php?query=time:2000:2008

Title: Miss Congeniality
Title: Iron Man
Title: The Notebook
Title: The Pursuit of Happyness
Title: I Am Legend
Title: Black Hawk Down
Title: The Alamo
Title: Gladiator
Title: Star Wars: The Force Awakens
Title: enemy at the gates
Title: Crouching Tiger, Hidden Dragon
Title: Harry Potter and the Sorcerer's Stone
Title: Slumdog Millionaire
Title: The Dark Knight
Title: Mulholland Dr.
Title: Stardust
Title: Keeping the Faith
Title: Finding Nemo
Title: The Notebook
Title: Gladiator
Year: 2005
Year: 2004
Year: 2006
Year: 2001
Year: 2006
Year: 2005
Year: 2005
Year: 2003
Year: 2006

← → ↻ lyle.smu.edu/~rmandhare/movieland/partc.php?query=director:Nolan

Title: Inception
Title: Inception
Title: The Dark Knight
Title: The Dark Knight Rises
Year: 2010
Year: 2010
Year: 2008
Year: 2012

← → ↻ lyle.smu.edu/~rmandhare/movieland/partc.php?query=actor:DiCaprio

Title: The Revenant
Title: Inception
Title: Inception
Title: Inception
Title: Titanic
Title: Titanic
Title: The Revenant
Year: 2015
Year: 2010
Year: 2010
Year: 2010
Year: 1997
Year: 1997
Year: 2015