



UNIVERSIDAD DE MÁLAGA



Graduado en Ingeniería de la Salud

Resistencia a antibióticos en Europa

Antimicrobial Resistance in Europe

Realizado por

Alejandro Pascual Mellado
David Cubillos del Toro
David Ramírez Arco
Nerea Martín Serrano

Tutorizado por

Laura Panizo Jaime



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA
GRADUADO EN INGENIERÍA DE LA SALUD

Resistencia a antibióticos en Europa

Antimicrobial Resistance in Europe

Realizado por
Alejandro Pascual Mellado
David Cubillos del Toro
David Ramírez Arco
Nerea Martín Serrano

Tutorizado por
Laura Panizo Jaime

Resumen

En este proyecto se va a desarrollar una página web que incluya información sobre bacterias con resistencia a los antibióticos en Europa.

Se va a comenzar por extraer información de un dataset (en formato XML). A este archivo se le aplicarán transformaciones XSLT para poder integrar la información en una página web. Esta página web será diseñada para que los distintos usuarios puedan visualizar de forma sencillos los datos recogidos. Para ello se van a crear dos interfaces, una dirigida a investigadores y otra dirigida a usuarios que simplemente quieran informarse sobre el tema.

Además, se realizarán posibles consultas que puedan realizar los usuarios sobre el dataset en lenguajes como XQuery y SPARQL.

El proyecto se aloja en el siguiente repositorio de gitHub:

github.com/Antimicrobial-Resistance-in-Europe.

Y la implementacion de la página web final está hecha en GitHub Pages, y se puede acceder con el siguiente enlace:

[Archerd6.github.io/Proyecto Estandares de Datos Web](https://Archerd6.github.io/Proyecto_Estandares_de_Datos_Web).

[a desarrollar conforme se vaya realizando el proyecto]

Palabras clave: Bacteria, Antibiótico, Base de datos, Página web, HTML, XSLT, XML

Índice

| | |
|--|-----------|
| 1. Introducción | 4 |
| 1.1. Motivación | 4 |
| 1.2. Objetivos | 4 |
| 1.3. Tecnologías usadas | 4 |
| 2. Descripción del Dataset | 5 |
| 3. Diseño XQuery | 7 |
| 3.1. Query 1 | 7 |
| 3.2. Query 2 | 7 |
| 4. Página Web | 8 |
| 4.1. Interfaz inicio | 8 |
| 4.2. Interfaz investigadores | 9 |
| 4.3. Interfaz usuarios | 10 |
| 5. Integración de datos | 11 |
| 5.1. Transformación tabla bacterias | 11 |
| 5.2. Transformación tabla grupos de estudio | 12 |
| 5.3. Transformaciones de las tablas del mapa | 13 |
| 6. Grafo RDF | 14 |
| 7. Conclusiones y Líneas Futuras | 16 |
| 7.1. Conclusiones | 16 |
| 7.2. Líneas Futuras | 16 |
| Apéndice A. Manual de Instalación | 17 |

1

Introducción

1.1. Motivación

Un problema emergente es la resistencia de los antibióticos que han desarrollado algunas bacterias. Esta resistencia ocurre cuando las bacterias ya no se eliminan con los antibióticos creados con ese objetivo. Esto es un grave problema pues las bacterias pueden causar infecciones con graves consecuencias, y la manera de curar estas infecciones es eliminar la bacteria que lo causa.

Se quiere crear una página web para que tanto el personal médico, como investigadores y otros usuarios puedan tener un listado de las bacterias que hayan desarrollado esta resistencia a antibióticos. Así los usuarios podrán intentar evitar el contacto con estas bacterias, el personal médico podrá saber si existe tratamiento para la infección bacteriana que estén tratando...

1.2. Objetivos

El objetivo es tener una página web donde los usuarios puedan mantenerse informados acerca de este tipo de bacterias. Por un lado, informando a los usuarios sobre en que zonas de Europa se encuentran este tipo de bacterias, y por otro lado, ampliando la información de los investigadores que estén llevando a cabo un estudio relacionado con el tema.

1.3. Tecnologías usadas

Las tecnologías que se van a usar a lo largo de este proyecto son: Visual Studio, \LaTeX , Protégé, eXist-db

2

Descripción del Dataset

El dataset elegido guarda información recogida de la base de datos European Anti-microbial Resistance Surveillance (EARS), en concreto la información relativa a aquellas bacterias que han desarrollado resistencia a los antibióticos. El dataset se encuentra en el siguiente enlace: kaggle.com/datasets/euro-resistance.

Cada una de las filas del dataset representa un estudio realizado sobre un grupo de personas. En las columnas se nos presenta la siguiente información sobre cada estudio:

- Distribución: se refiere a si los datos se extrajeron de un género o grupo de edad en particular.
- RegionName: país de la institución informante.
- Tiempo: el año.
- Categoría: grupo de edad o el género (dependiendo de la distribución).
- Valor: porcentaje de bacterias que fueron resistentes al grupo de antibióticos.
- Bacterias: se refiere a las bacterias en las que se estudió la resistencia.
- Antibiótico: grupo de antibióticos que se usó para matar la bacteria.

En la imagen 1 se encuentra el dataset en formato .XMLSpy (tal cual se descarga de Kaggle).

Se necesita que los datos estén en formato XML para poder realizar las distintas consultas y transformaciones. XML es un metalenguaje de marcado. Un lenguaje de marcado o lenguaje de marcas es una forma de codificar un documento que, junto con el texto,

```
,Distribution,Unit,Time,RegionCode,RegionName,Category,Value,Bacteria,Antibiotic
0,"R - resistant isolates proportion, by age",%,2012,BG,Bulgaria,0-4,0.0,Acinetobacter spp.,Aminoglycosides
1,"R - resistant isolates proportion, by age",%,2012,BG,Bulgaria,5-18,100.0,Acinetobacter spp.,Aminoglycosides
2,"R - resistant isolates proportion, by age",%,2012,BG,Bulgaria,19-64,62.857142857,Acinetobacter spp.,Aminoglycosides
3,"R - resistant isolates proportion, by age",%,2012,BG,Bulgaria,65+,56.52173913,Acinetobacter spp.,Aminoglycosides
4,"R - resistant isolates proportion, by age",%,2012,CY,Cyprus,0-4,33.333333333,Acinetobacter spp.,Aminoglycosides
5,"R - resistant isolates proportion, by age",%,2012,CY,Cyprus,5-18,0.0,Acinetobacter spp.,Aminoglycosides
6,"R - resistant isolates proportion, by age",%,2012,CY,Cyprus,19-64,57.142857143,Acinetobacter spp.,Aminoglycosides
7,"R - resistant isolates proportion, by age",%,2012,CY,Cyprus,65+,62.5,Acinetobacter spp.,Aminoglycosides
8,"R - resistant isolates proportion, by age",%,2012,DE,Germany,0-4,0.0,Acinetobacter spp.,Aminoglycosides
9,"R - resistant isolates proportion, by age",%,2012,DE,Germany,5-18,0.0,Acinetobacter spp.,Aminoglycosides
10,"R - resistant isolates proportion, by age",%,2012,DE,Germany,19-64,9.090909091,Acinetobacter spp.,Aminoglycosides
11,"R - resistant isolates proportion, by age",%,2012,DE,Germany,65+,5.479452055,Acinetobacter spp.,Aminoglycosides
12,"R - resistant isolates proportion, by age",%,2012,DK,Denmark,0-4,0.0,Acinetobacter spp.,Aminoglycosides
13,"R - resistant isolates proportion, by age",%,2012,DK,Denmark,5-18,16.666666667,Acinetobacter spp.,Aminoglycosides
14,"R - resistant isolates proportion, by age",%,2012,DK,Denmark,19-64,15.384615385,Acinetobacter spp.,Aminoglycosides
15,"R - resistant isolates proportion, by age",%,2012,DK,Denmark,65+,7.142857143,Acinetobacter spp.,Aminoglycosides
16,"R - resistant isolates proportion, by age",%,2012,FR,France,0-4,9.375,Acinetobacter spp.,Aminoglycosides
```

Figura 1: Dataset formato .XMLSpy

incorpora etiquetas o marcas que contienen información adicional acerca de la estructura del texto o su presentación Para ello se ha usado la siguiente herramienta online: www.convertcsv.com. Esta herramienta da como salida el archivo .XML de la imagen 2.

Como se puede observar en la imagen, se ha generado un nodo raíz, al que se ha llamado *root*, y por cada fila del dataset se ha generado un nodo *row*, el cual tiene como nodos hijos las columnas del dataset.

```
<?xml version="1.0" encoding="UTF-8"?>
<root>
  <row>
    <FIELD1>0</FIELD1>
    <Distribution>R - resistant isolates proportion, by age</Distribution>
    <Unit>%</Unit>
    <Time>2012</Time>
    <RegionCode>BG</RegionCode>
    <RegionName>Bulgaria</RegionName>
    <Category>0-4</Category>
    <Value>0.0</Value>
    <Bacteria>Acinetobacter spp.</Bacteria>
    <Antibiotic>Aminoglycosides</Antibiotic>
  </row>
  <row>
    <FIELD1>1</FIELD1>
    <Distribution>R - resistant isolates proportion, by age</Distribution>
    <Unit>%</Unit>
    <Time>2012</Time>
    <RegionCode>BG</RegionCode>
    <RegionName>Bulgaria</RegionName>
    <Category>5-18</Category>
    <Value>100.0</Value>
    <Bacteria>Acinetobacter spp.</Bacteria>
    <Antibiotic>Aminoglycosides</Antibiotic>
  </row>

```

Figura 2: Dataset formato .XML

3

Diseño XQuery

En esta sección se van a hacer dos posibles consultas que harían distintos usuarios a la base de datos, para lo que se usan las queries o consultas. Una consulta sirve encontrar y extraer elementos y atributos de una base de datos.

Para realizar estas consultas se va a usar XQuery, un lenguaje para realizar consultas en documentos XML. Se basa en funciones, expresiones XPath y predicados.

Las operaciones más comunes para hacer queries son SELECT, que selecciona los datos que serán visibles en la tabla resultado; FROM, para señalar la o las tablas de las cuales quiere obtener la información, WHERE, utilizada para delimitar la información que queremos obtener.

3.1. Query 1

3.2. Query 2

4

Página Web

La página Web va a permitir a los usuarios acceder de forma concisa a la información que queremos presentar. Se van a hacer dos interfaces, una para investigadores y otra para usuarios que simplemente quieran informarse. Se va añadir una interfaz que se muestre al inicio de la aplicación donde los usuarios elijan que interfaz desean ver.

4.1. Interfaz inicio

Esta interfaz va a constar de dos botones, uno que redirija a la interfaz de usuarios y otro a la de investigadores, como se puede comprobar en la imagen 3.



Figura 3: Interfaz Inicio

4.2. Interfaz investigadores

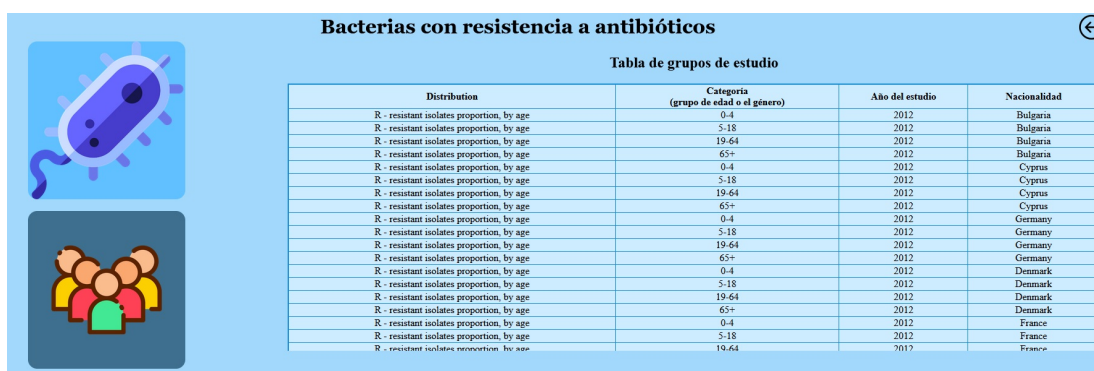
Esta interfaz esta enfocada para personas dedicadas a la investigación. A la izquierda de la interfaz van a encontrarse dos botones y a la derecha una tabla. Al seleccionar uno de los botones de la izquierda (el de la imagen de una bacteria), en la tabla de la derecha se presentará información relacionada con las bacterias, los antibióticos a los que son resistentes y el porcentaje de resistencia de la bacteria en ese estudio (es decir, se filtran de las columnas del dataset aquellas columnas de interés). Esto se refleja en la imagen 4.



| Nombre Bacteria | Antibiotico testado | Porcentaje de resistencia |
|----------------------|---------------------|---------------------------|
| Acinetobacter spp. | Aminoglicosidos | 100.0% |
| Enterococcus faecium | Aminopenicilinas | 99.236641221% |
| Enterococcus faecium | Aminopenicilinas | 99.206349206% |
| Enterococcus faecium | Aminopenicilinas | 99.152542173% |
| Acinetobacter spp. | Fluoroquinolones | 99.130454783% |
| Enterococcus faecium | Aminopenicilinas | 99.038461538% |
| Enterococcus faecium | Aminopenicilinas | 98.98989899% |
| Enterococcus faecium | Aminopenicilinas | 98.969072165% |
| Enterococcus faecium | Aminopenicilinas | 98.901090901% |
| Enterococcus faecium | Aminopenicilinas | 98.88888889% |
| Enterococcus faecium | Aminopenicilinas | 98.863636364% |
| Enterococcus faecium | Aminopenicilinas | 98.795180723% |
| Enterococcus faecium | Aminopenicilinas | 98.666666667% |
| Enterococcus faecium | Aminopenicilinas | 98.564593301% |
| Enterococcus faecium | Aminopenicilinas | 98.550724638% |
| Enterococcus faecium | Aminopenicilinas | 98.507462687% |
| Enterococcus faecium | Aminopenicilinas | 98.484848485% |
| Enterococcus faecium | Aminopenicilinas | 98.461538462% |
| Enterococcus faecium | Aminopenicilinas | 98.412698413% |
| Enterococcus faecium | Aminopenicilinas | 98.4% |

Figura 4: Tabla bacterias

En el caso de seleccionar el otro botón en la tabla de la derecha se presentará la información relacionada con los grupos a los que se le realizó el estudio (imagen 5).



| Distribucion | Categoria (grupo de edad o el género) | Año del estudio | Nacionalidad |
|---|---------------------------------------|-----------------|--------------|
| R - resistant isolates proportion, by age | 0-4 | 2012 | Bulgaria |
| R - resistant isolates proportion, by age | 5-18 | 2012 | Bulgaria |
| R - resistant isolates proportion, by age | 19-64 | 2012 | Bulgaria |
| R - resistant isolates proportion, by age | 65+ | 2012 | Bulgaria |
| R - resistant isolates proportion, by age | 0-4 | 2012 | Cyprus |
| R - resistant isolates proportion, by age | 5-18 | 2012 | Cyprus |
| R - resistant isolates proportion, by age | 19-64 | 2012 | Cyprus |
| R - resistant isolates proportion, by age | 65+ | 2012 | Cyprus |
| R - resistant isolates proportion, by age | 0-4 | 2012 | Germany |
| R - resistant isolates proportion, by age | 5-18 | 2012 | Germany |
| R - resistant isolates proportion, by age | 19-64 | 2012 | Germany |
| R - resistant isolates proportion, by age | 65+ | 2012 | Germany |
| R - resistant isolates proportion, by age | 0-4 | 2012 | Denmark |
| R - resistant isolates proportion, by age | 5-18 | 2012 | Denmark |
| R - resistant isolates proportion, by age | 19-64 | 2012 | Denmark |
| R - resistant isolates proportion, by age | 65+ | 2012 | Denmark |
| R - resistant isolates proportion, by age | 0-4 | 2012 | France |
| R - resistant isolates proportion, by age | 5-18 | 2012 | France |
| R - resistant isolates proportion, by age | 19-64 | 2012 | France |
| R - resistant isolates proportion, by age | 65+ | 2012 | France |

Figura 5: Tabla estudios

4.3. Interfaz usuarios

En este interfaz se presenta un mapa de Europa. Delante de este mapa se van a poner una serie de botones por zonas (que incluyan más de un país) para que el usuario pueda hacer click y obtener información sobre aquellas bacterias con resistencia a antibióticos en esa zona de Europa.

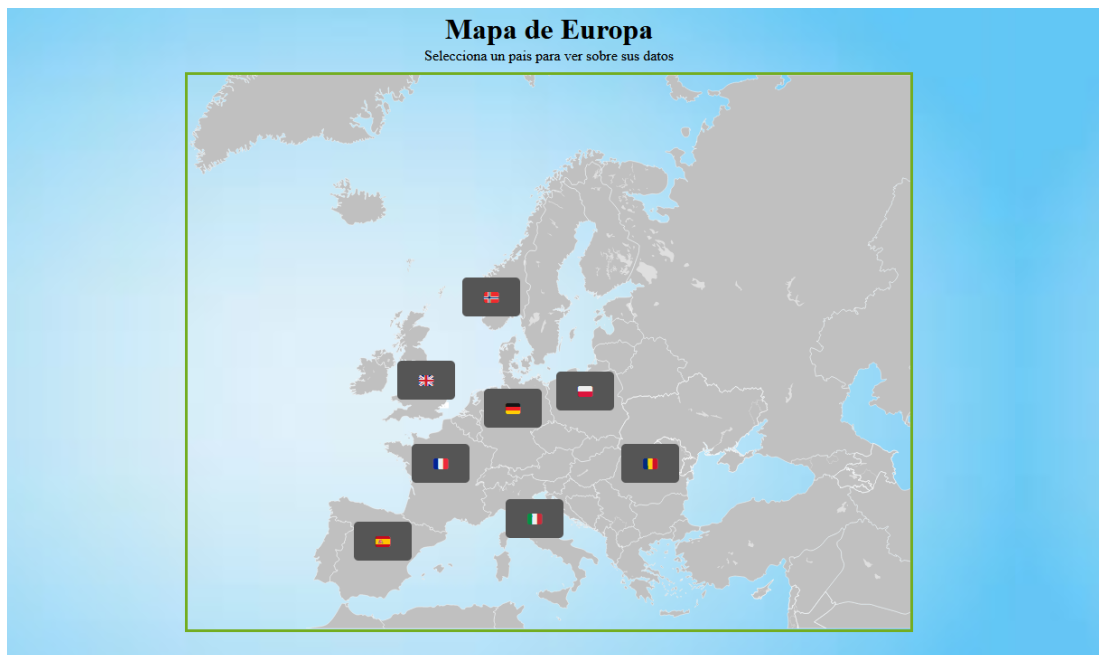


Figura 6: Interfaz usuarios

5

Integración de datos

En esta sección se va a describir como se van a integrar los datos en la aplicación. Se van a usar transformaciones XSLT, las cuales permiten definir plantillas para transformar un documento XML en otro tipo de documento (HTML en este caso) utilizando sintaxis XML.

5.1. Transformación tabla bacterias

En la interfaz de investigadores, para generar la tabla que aparece al pulsar en la bacteria se ha creado una transformación XSLT.

Al principio de la transformación se ha creado una etiqueta `<style>` para poder dar formato a la tabla. A continuación, se ha creado la tabla y se han definido el nombre de las columnas. Para poder rellenar la tabla, dentro de cada fila (*row*) del archivo XSL y se han seleccionado los nodos hijos mencionados anteriormente.

En la imagen 7 se puede ver como se ha usado un `for-each` para ir recorriendo las filas cuyas bacterias tenga más de un 20 % de resistencia a los antibióticos. Después, se han ordenado las bacterias por mayor resistencia a antibióticos. Por último, se usa un `xsl:if` para comprobar que los datos no se repiten, pues hay varios estudios donde coinciden los resultados. Si la condición se cumple, se crea una nueva fila en la tabla con los datos seleccionados en los `xsl:value-of`.

```

<xsl:for-each select="//row[Value>20]">
  <xsl:sort select="Value" data-type="number" order="descending"/>
  <xsl:if test="Bacteria[generate-id()=generate-id(key('idBacteria',.))[1]]
    or Value[generate-id()= generate-id(key('idValue',.))[1]]">
    <tr>
      <td><xsl:value-of select="Bacteria"/></td>
      <td><xsl:value-of select="Antibiotic"/></td>
      <td><xsl:value-of select="Value"/>%</td>
    </tr>
  </xsl:if>
</xsl:for-each>

```

Figura 7: Transformación tabla bacterias

5.2. Transformación tabla grupos de estudio

Se ha realizado otra transformación para el caso en el que el investigador pulse el botón de grupos de estudio. El principio es igual que en el caso anterior, se empieza dándole formato a la tabla y definiendo las columnas. En la imagen 8 se ve como se crean las columnas. Se vuelven a recorrer las distintas filas, esta vez sin condición, y se seleccionan los valores que necesitamos para rellenar la tabla.

```

21      <center>
22        <h2>Tabla de grupos de estudio</h2>
23        <table>
24          <thead>
25            <tr>
26              <th>Distribution</th>
27              <th>Categoría <br> (grupo de edad o el género) </br></th>
28              <th>Año del estudio</th>
29              <th>Nacionalidad</th>
30            </tr>
31          </thead>
32          <tbody>
33            <xsl:for-each select="//row">
34              <tr>
35                <td><xsl:value-of select="Distribution"/></td>
36                <td><xsl:value-of select="Category"/></td>
37                <td><xsl:value-of select="Time"/></td>
38                <td><xsl:value-of select="RegionName"/></td>
39              </tr>
40            </xsl:for-each>
41          </tbody>
42        </table>
43      </center>
44

```

Figura 8: Transformación tabla grupos

5.3. Transformaciones de las tablas del mapa

Se ha realizado transformaciones para obtener información sobre aquellas bacterias con resistencia a antibióticos en ese país de Europa

```
<h2>Tabla de bacterias</h2></center>
<table>
  <thead>
    <tr>
      <th>Field</th>
      <th>Distribution</th>
      <th>Unit</th>
      <th>Time</th>
      <th>RegionName</th>
      <th>Category</th>
      <th>Value</th>
      <th>Nombre Bacteria</th>
      <th>Antibiotico testado</th>
    </tr>
  </thead>
  <tbody>
    <xsl:for-each select="//row">
      <xsl:if test="contains(RegionCode,'ES')">
        <tr>
          <td>
            <xsl:value-of select="FIELD1" />
          </td>
          <td>
            <xsl:value-of select="Distribution" />
          </td>
          <td>
            <xsl:value-of select="Unit" />
          </td>
          <td>
            <xsl:value-of select="Time" />
          </td>
          <td>
            <xsl:value-of select="RegionName" />
          </td>
          <td>
            <xsl:value-of select="Category" />
          </td>
          <td>
            <xsl:value-of select="Value" />
          </td>
          <td>
            <xsl:value-of select="Bacteria" />
          </td>
          <td>
            <xsl:value-of select="Antibiotic" />
          </td>
        </tr>
      </xsl:if>
    </xsl:for-each>
  </tbody>
</table>
```

Figura 9: Transformación tabla grupos

6

Grafo RDF

RDF es un lenguaje de modelado para describir recursos en la web. RDF permite modelar recursos (datos) y sus relaciones mediante un grafo dirigido.

Se ha tenido que desarrollar un mapa de clases que están relacionados mediante propiedades de datos. Definiendo las clases, operaciones:

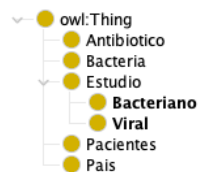


Figura 10: Clases



Figura 11: Relaciones

El mapa se ha jerarquizado en torno al estudio que se esta realizando, de el parten las relaciones.

Para expresar atributos de las clases se han creado propiedades:

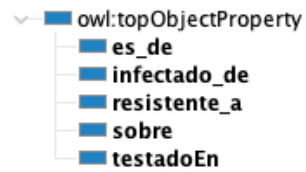


Figura 12: Clases

Estas propiedades complementan a ciertas clases, aportando información sobre ellas. La mayoría se trata de propiedades de una clase a un tipo de dato int o string peor rango de edad se ha declarado como un enumerado.



Figura 13:

7

Conclusiones y Líneas Futuras

7.1. Conclusiones

7.2. Líneas Futuras

Apéndice A

Manual de Instalación



UNIVERSIDAD
DE MÁLAGA

| **uma.es**

E.T.S. DE INGENIERÍA INFORMÁTICA

E.T.S de Ingeniería Informática
Bulevar Louis Pasteur, 35
Campus de Teatinos
29071 Málaga