# ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO

LABORATORIO No 6

MODELOS Y BASES DE DATOS (MBDA)

DANIEL FELIPE ROJAS

MARIA VALENTINA TORRES

#### **MODELOS Y BASES DE DATOS**

XML y SQL

2023-02

Laboratorio 6/6

#### **OBJETIVOS**

Evaluar el logro de las competencias adquiridas para:

- 1. Definir un archivo XML que cumplan con las especificaciones de ejemplares específicos.
- 2. Validar que un archivo XML está bien formado.
- 3. Visualizar la estructura de un archivo XML en forma de árbol.
- 4. Proponer e implementar consultas en XPath
- 5. Definir el esquema emergente de sus datos usando el estandar DTD
- 6. Verificar que los datos de un archivo XML cumplen lo definido en un esquema.
- 7. Integrar información en XML a una base de datos relacional

PUNTO UNO. Sólo XML. CATALOGO DE CURSOS (courses.xml)

El caso es uno de los propuestos en el curso de bases de datos de Stanford.

#### A. Explorando

1. Explore el archivo usando un navegador y un editor. ¿Cuáles son las ventajas de uno u otro?

En un editor de texto como visual podemos mediante extensiones visualizar mejor la estructura jerárquica y en el navegador es que se puede realizar la extracción de información.

2. Describa la información que encuentra en el archivo.

El archivo contiene una etiqueta principal que describe el catálogo de los cursos, los cuales están compuestos por las etiquetas de departamento. Cada etiqueta de departamento tiene las etiquetas de títle, chair y las etiquetas de coursers.

La etiqueta de título describe el nombre del departamento.

La etiqueta de chair contiene la etiqueta de Professor, la cual está compuesta por las etiquetas de First\_Name y Last\_Name.

La etiqueta de

- B. Consultas iniciales Implemente las siguientes consultas:
  - 1. ¿Cuántos departamentos tenemos? ¿Cuáles son? [Nombre]

```
count(/Course_Catalog/Department/@Code)
```

```
1 3.0
```

/Course Catalog/Department/@Code

```
1 Code="CS"
2 Code="EE"
3 Code="LING"
```

- 2. ¿De cuáles cursos se tiene información? [a. Nombre con etiqueta] [b. Nombre sin etiqueta] [c. Número con atributo] [d. Número sin atributo] [e. Número total de cursos]
  - a. Nombre con etiqueta

# /Course\_Catalog/Department/Course/Title

b. Nombre sin etiqueta

/Course\_Catalog/Department/Course/Title/text()

Programming Methodology
Programming Abstractions
Computer Organization and Systems
Introduction to Probability for Computer Scientists
From Languages to Information
Compilers
Introduction to Databases
Artificial Intelligence: Principles and Techniques
Structured Probabilistic Models: Principles and Techniques
Machine Learning
Digital Systems I
Digital Systems II
From Languages to Information

#### c. Número con atributo

/Course\_Catalog/Department/Course/@Number

Number=CS106A

Number=CS106B

Number=CS107

Number=CS109

Number=CS124

Number=C5124

Number=CS143

Number=CS145

Number=CS221

Number=CS228

Number=CS229

Number=EE108A

Number=EE108B

Number=LING180

# d. Número sin atributos

### e. Número total de cursos

count(/Course\_Catalog/Department/Course/@Number)

1 13.0

3. ¿Cuál es la información del curso Introduction to Databases? [a. Toda la información] [b. Todos los atributos] [c. Todas las etiquetas]

/Course\_Catalog/Department/Course[Title='Introduction to Databases']

R /Course\_Catalog/Department/Course[Title='Introduction to Databases']/@\*

# Number="CS145" Enrollment="130"

/Course\_Catalog/Department/Course[Title='Introduction to Databases']/\*

4. ¿A qué departamento pertenece el curso Introduction to Databases? [Código]

/Course\_Catalog/Department[Course/Title = 'Introduction to Databases']/@Code

```
1 Code="CS"
```

5. ¿Cuáles directores de departamento no colaboran en cursos? [Toda la información del profesor] (Los profesores se pueden identificar por su apellido)

 $/ Course\_Catalog/Department/Chair/Professor[not(/Course\_Catalog/Department/Course/instructors/Professor)]/* \\$ 

6. ¿Cuál es el total de inscripciones de los cursos?

```
sum(/Course Catalog/Department/Course/@Enrollment)
```

7. ¿De cuáles cursos no se conoce el número de inscritos? [Nombre]

8. ¿Cuál curso tiene el mayor número de inscritos? [Nombre]

/Course\_Catalog/Department/Course[@Enrollment=max(/Course\_Catalog/Department/Course/@Enrollment)]/Title

```
<Title>Programming Methodology</Title>
```

9. ¿Cuáles cursos son iniciales (no tienen prerrequisitos)? [Nombre]

/Course\_Catalog/Department/Course[not(Prerequisites)]/Title/text()

```
1 Text = Programming Methodology
2 Text = Artificial Intelligence- Principles and Techniques
3 Text = Structured Probabilistic Models- Principles and Techniques
4 Text = Machine Learning
5 Text = Digital Systems I
```

10. ¿Cuáles cursos son finales (no son prerrequisitos de ninguno)? [Nombre]

 $/ Course\_Catalog/Department/Course[not(@Number=/Course\_Catalog/Department/Course/Prerequisites/Prereq)]/ Title$ 

## B. Consultas quiz Standford

1. Return all Title elements (of both departments and courses). (16)

```
//Title/text()
```

```
Computer Science
Programming Methodology
Programming Abstractions
Computer Organization and Systems
Introduction to Probability for Computer Scientists
From Languages to Information
Compilers
Introduction to Databases
Artificial Intelligence: Principles and Techniques
Structured Probabilistic Models: Principles and Techniques
Machine Learning
Electrical Engineering
Digital Systems I
Digital Systems II
Linguistics
From Languages to Information
```

2. Return last names of all department chairs. (3)

```
/Course_Catalog/Department/Chair/Professor/Last_Name/text()
```

Widom Horowitz Levin

3. Return titles of courses with enrollment greater than 500. (2)

/Course\_Catalog/Department/Course[@Enrollment > 500]/Title

```
<Title>Programming Methodology</Title>
<Title>Programming Abstractions</Title>
```

4. Return titles of departments that have some course that takes "CS106B" as a prerequisite. (2)

/Course Catalog/Department[Course/Prerequisites/Prereq = 'CS106B']/Title

```
<Title>Computer Science</Title>
<Title>Electrical Engineering</Title>
```

5. Return last names of all professors or lecturers who use a middle initial. Don't worry about eliminating duplicates. (5)

//Professor[Middle\_Initial]/Last\_Name | //Lecturer[Middle\_Initial]/Last\_Name

```
<Last_Name>Cain</Last_Name>
<Last_Name>Cain</Last_Name>
<Last_Name>Aiken</Last_Name>
<Last_Name>Horowitz</Last_Name>
<Last_Name>Dally</Last_Name>
```

6. Return the title of courses that have a cross-listed course (i.e., that have "Cross-listed" in their description). (2)

/Course\_Catalog/Department/Course[contains(Description, 'Cross-listed')]/Title

```
<Title>From Languages to Information</Title>
<Title>From Languages to Information</Title>
```

7. Return the enrollment of all courses in the CS department. (10)

/Course\_Catalog/Department[@Code='CS']/Course/@Enrollment

```
Enrollment=1070
Enrollment=620
Enrollment=500
Enrollment=280
Enrollment=60
Enrollment=90
Enrollment=130
Enrollment=110
Enrollment=320
```

8. Return last names of instructors teaching at least one course that has "system" in its description and enrollment greater than 100. (2)

/Course\_Catalog/Department/Course[contains(Description, 'system') and @Enrollment > 100]//(Professor|Lecturer)/Last\_Name

```
<Last_Name>Zelenski</Last_Name>
<Last_Name>Widom</Last_Name>
```

9. Return the title of the course with more than 2 instructors (\*)(1)

```
/Course_Catalog/Department/Course[<u>count(</u>Instructors/*) > 2]/Title
```

```
<Title>Programming Methodology</Title>
```

10. Return course numbers of courses that have the same title as some other course. (Hint: You might want to use the "preceding" and "following" navigation axes for this query, which were not covered in the video or our demo script; they match any preceding or following node, not just siblings.) (2)

//Course[Title = following::Course/Title or Title=preceding::Course/Title]/@Number

Number=CS124 Number=LING180

11. Return the courses numbers of courses that have no lecturers as instructors. (\*)(10)

/Course\_Catalog/Department/Course[not(/Instructors/Lecturer)]/@Number

Number=CS106A

Number=CS106B

Number=CS107

Number=CS109

Number=CS124

Number=CS143

Number=CS145

Number=CS221

Number=CS228

Number=CS229

Number=EE108A

Number=EE108B

Number=LING180

12. Return titles of courses taught by the chair of a department. For this question, you may assume that all professors have distinct last names. (1)

- 13. Return titles of courses taught by a professor with the last name "Ng" and by a professor with the last name "Thrun". (\*)(1)
- 14. Return course numbers of courses that have a course taught by Eric Roberts as a prerequisite. (4) Return last names of instructors teaching at least one course that has system in its description and enrollment greater than 100.(2)
- C. Consultas propias

Propongan e implementen cinco consultas propias

1. El número de los cursos que tienen más de 100 inscritos y no tiene prerrequisitos /Course Catalog/Department/Course[@Enrollment > 100 and not(Prerequisites)]/@Number

Number=CS106A Number=CS221 Number=CS228 Number=CS229

Code=LING

2. El código del departamento que solo tiene un curso

```
/Course_Catalog/Department[count(Course) = 1]/@Code
```

3. Los cursos que tiene el mismo número de instructores que el máximo de todos los cursos

/Course\_Catalog/Department/Course[max(count(Instructors))]/@Number

Number=CS106A Number=EE108A Number=LING180

4. El número de estudiantes inscritos en el curso que tenga prerrequisito CS106A /Course\_Catalog/Department/Course[Prerequisites/Prereq = 'CS106A']/@Enrollment

Enrollment=620

5. La suma de los estudiantes inscritos en cursos que no tengan prerrequisito /Course\_Catalog/Department/sum(Course[not(Prerequisites)]/@Enrollment)

## D. Esquema

Proponga un DTD para estos datos. Para verificar use la herramienta que aparece en moodle.

```
<!DOCTYPE Course_Catalog [</pre>
   <!ELEMENT Course_Catalog (Department+)>
   <!ELEMENT Department (Title, Chair, Course+)>
   <!ATTLIST Department Code CDATA #REQUIRED>
   <!ELEMENT Title (#PCDATA)>
   <!ELEMENT Chair (Professor)>
   <!ELEMENT Professor (First_Name, (Middle_Initial?, Last_Name))>
   <!ELEMENT First_Name (#PCDATA)>
   <!ELEMENT Middle_Initial (#PCDATA)>
   <!ELEMENT Last_Name (#PCDATA)>
   <!ELEMENT Course (Title, Description?, Instructors, Prerequisites?)
   <!ATTLIST Course Number CDATA #REQUIRED
                    Enrollment CDATA #REQUIRED>
   <!ELEMENT Description (#PCDATA)>
   <!ELEMENT Instructors (Professor+ | Lecturer+)>
   <!ELEMENT Lecturer (First_Name, (Middle_Initial?, Last_Name))>
   <!ELEMENT Prerequisites (Prereq+)>
   <!ELEMENT Prereq (#PCDATA)>
```

E. Nuevos datos Incluya en el archivo la información de tres cursos que no existan en el archivo. Verifique que cumple el esquema definido anteriormente.

PUNTO DOS. (XML - Oracle) Uber PQRS. Atributo: TDescripcion

Este punto vamos a perfeccionar uno de los atributos de la base de datos.

1) Proponga la estructura XML necesaria para tener la información del atributo. DTD y ejemplos XML OK y XML NoOK. Explique.

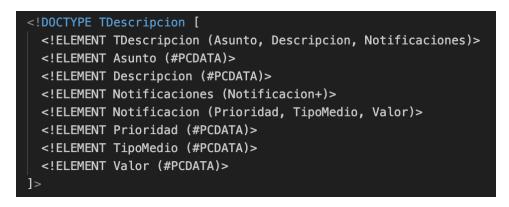
Descripcion del Astah:

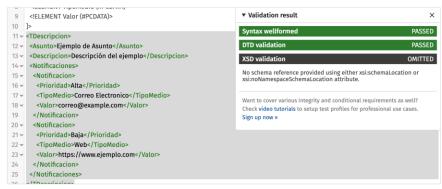
Incluye: asunto, texto descriptivo del caso, medios de notificación de respuesta.

Cada notificación tiene definida una prioridad, un tipo de medio y un valor

Los tipos de medio pueden ser teléfono, correo electrónico y dirección.

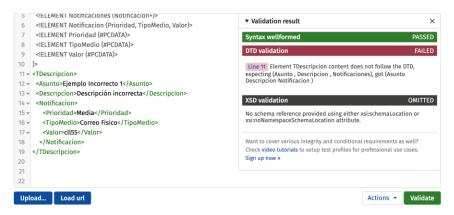
(Ejemplo: La prioridad número uno es para el teléfono con valor 3133456756).



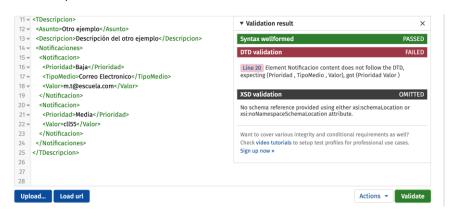




### Falta la rama de <Notificaciones>



#### Falta el tipoMedio en una de las notificaciones del PQRS



2) Actualice la tabla y los datos adicionados a la base de datos. (PoblandoOK, ProblandoNoOK)

```
CREATE TABLE PORS(
    TICKED
                      VARCHAR(13)
                                     NOT NULL.
     radicacion
                      DATE
                                      NOT NULL.
     cierre
                      DATE.
    descripcion
                      XMLTYPE.
                      CHAR(1)
                                      NOT NULL,
    tipo
    estado
                      CHAR(1)
                                      NOT NULL.
     solicitud
                      NUMBER(9)
                                      NOT NULL
);
```

```
VALUES ('P180812345078', TO_DATE('2023-10-05', 'YYYY-MM-DD'), TO_DATE('2023-10-10', 'YYYY-MM-DD'), '
     <TDescripcion>
       <Asunto>No lo se</Asunto>
       <Descripcion>Hola
       <Notificaciones>
            Notificacion>
                <Prioridad>1</Prioridad>
<TipoMedio>Correo Electronico</TipoMedio>
                 <Valor>1456</Valor>
                 Antificacion>
                 <Prioridad>3</Prioridad>
                <TipoMedio>Web</TipoMedio>
<Valor>10</Valor>
            </Notificacion>
            </Notificacion>
       </Notificaciones>
     </TDescripcion>', 'P', 'P', 1808);
INSERT INTO PQRS (ticked, radicacion, cierre, descripcion, tipo, estado, solicitud) VALUES ('P280816345678', TO_DATE('2023-10-06', 'YYYY-MM-DD'), TO_DATE('2023-10-12', 'YYYY-MM-DD'), '
     <TDescripcion>
       <Asunto>Conductor grosero</Asunto>
       <Descripcion>Hola/Descripcion>
<Notificaciones>
            ◇Notificacion>
                <Prioridad>2</Prioridad>
                 <TipoMedio>Web</TipoMedio>
                <Valor>10</Valor>
            </Notificacion>
     </Notificaciones>
    </TDescripcion>', 'P', 'P', 2808);
```

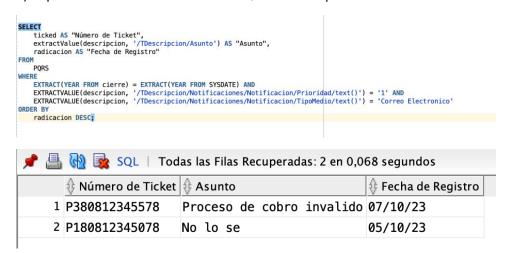
```
INSERT INTO PQRS (ticked, radicacion, cierre, descripcion, tipo, estado, solicitud)
VALUES ('P380812345578', T0_DATE('2023-10-07', 'YYYY-MM-DD'), T0_DATE('2023-10-15', 'YYYY-MM-DD'), '

*TDescripcions
- Asunto-Proceso de cobro invalido</Asunto>
- Oescripcion-Hola-Vorere Electronico-
- Abotificaciones-
- Abotificaciones-
- Prioridad-1c/Prioridad-
- < TipoMedio-Ocrree Electronico-/TipoMedio>
- < Notificaciones-
- *Notificaciones-
- *Notificaciones-
- *Notificaciones-
- *TDEScripcion-', 'P', 'P', 3808);

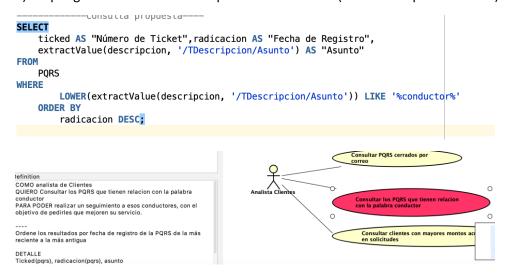
- PoblarNOOK
- INSERT INTO PQRS (ticked, radicacion, cierre, descripcion, tipo, estado, solicitud)
VALUES ('P480812345678', T0_DATE('2023-10-08', 'YYYY-MM-DD'), T0_DATE('2023-10-18', 'YYYY-MM-DD'), 'Descripciā'n 4', 'P', 'P', 4808);

INSERT INTO PQRS (ticked, radicacion, cierre, descripcion, tipo, estado, solicitud)
VALUES ('P580812345678', T0_DATE('2023-10-09', 'YYYY-MM-DD'), T0_DATE('2023-10-20', 'YYYY-MM-DD'), 'Descripciā'n 5', 'P', 'P', 5808);
```

3) Implemente la consulta Consultar PQRS cerrados por correo



4) Proponga otra nueva consulta que use ese atributo (Diseño e implementación)



5) Extienda la información de este atributo (DTD) y proponga una nueva consulta que ilustre la pertinencia de la nueva información registrada en XML. (Diseño e implementación)

#### **RETROSPECTIVA**

1. ¿Cuál fue el tiempo total invertido en el laboratorio por cada uno de ustedes? (Horas/Hombre)

17 /Rojas

17/ Torres

2. ¿Cuál es el estado actual del laboratorio? ¿Por qué?

Esta parcialmente completo ya que del punto de consultas nos hizo falta algunas consultas

3. ¿Cuál consideran fue el mayor logro? ¿Por qué?

Poder lograr realizar consultas que estén relacionada con el atributo de tipo xml.

4. ¿Cuál consideran que fue el mayor problema técnico? ¿Qué hicieron para resolverlo?

Nuestro mayor problema técnico es que no sabíamos donde ubicar los archivos .dtd y ejecutarlos

5. ¿Qué hicieron bien como equipo? ¿Qué se comprometen a hacer para mejorar los resultados?

Trabajar en equipo, escuchar al otro y dialogar para poder desarrollar el otro pedazo