

# Taller de Lenguajes y Tecnologías de la Web Semántica

## Introducción

Concepción de Sistemas de Información

2021



# Motivación

- Consideremos las siguientes personas y la información que tiene Wikipedia:

### John Lennon



John Lennon en marzo de 1969

#### Información personal

<b>Nombre de nacimiento</b>	John Winston Lennon
<b>Otros nombres</b>	<i>John Winston Ono Lennon</i>
<b>Nacimiento</b>	9 de octubre de 1940 Liverpool, Reino Unido
<b>Fallecimiento</b>	8 de diciembre de 1980 (40 años) Nueva York, Estados Unidos
<b>Causa de la muerte</b>	Asesinato por arma de fuego (Ver más detalles)
<b>Sepultura</b>	Central Park <span>✎</span>
<b>Residencia</b>	Liverpool, Nueva York y 251 Menlove Avenue <span>✎</span>
<b>Nacionalidad</b>	Británica
<b>Religión</b>	Ateísmo <span>✎</span>

#### Características físicas

### Albert Einstein



Albert Einstein en 1947

#### Información personal

<b>Nacimiento</b>	14 de marzo de 1879 Ulm (Alemania)
<b>Fallecimiento</b>	18 de abril de 1955 (76 años) Princeton (Nueva Jersey) (Estados Unidos)
<b>Causa de la muerte</b>	Aneurisma de aorta abdominal
<b>Nacionalidad</b>	Estadounidense (desde 1940) Alemana (1918-1933) Austrohúngara (1911-1912) Suiza (desde 1901) Sin nacionalidad (1896-1901) Alemana (hasta 1896)
<b>Religión</b>	Agnóstico
<b>Familia</b>	
<b>Padres</b>	Hermann Einstein Pauline Koch

### Eduardo VIII del Reino Unido

Rey del Reino Unido y de sus Dominios de Ultramar y emperador de la India



Eduardo, duque de Windsor, en 1945

#### Reinado

20 de enero de 1936-11 de diciembre de 1936  
(326 días)

<b>Predecesor</b>	Jorge V
<b>Sucesor</b>	Jorge VI

#### Información personal

<b>Nombre secular</b>	Edward Albert Christian George Andrew Patrick David
<b>Otros títulos</b>	Duque de Windsor (1937-1972)
<b>Nacimiento</b>	23 de junio de 1894 White Lodge, Richmond Park, Inglaterra
<b>Fallecimiento</b>	28 de mayo de 1972

### Alan Turing



Foto de pasaporte de Alan Turing a los 16 años

#### Información personal

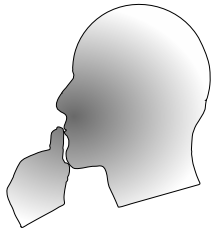
<b>Nombre de nacimiento</b>	Alan Mathison Turing <span>✎</span>
<b>Nacimiento</b>	23 de junio de 1912 <span>✎</span> Maida Vale (Reino Unido de Gran Bretaña e Irlanda) o Warrington Lodge (Reino Unido de Gran Bretaña e Irlanda) <span>✎</span>
<b>Fallecimiento</b>	7 de junio de 1954 <span>✎</span> (41 años) Wilmslow (Reino Unido) <span>✎</span>
<b>Causa de la muerte</b>	Intoxicación cianhídrica <span>✎</span>
<b>Sepultura</b>	Woking Crematorium <span>✎</span>
<b>Residencia</b>	Maida Vale y Guildford <span>✎</span>
<b>Nacionalidad</b>	Británica

# Motivación: Variedad en la Información

- No todos tienen la misma información:
- Con información compartida entre varios:
  - Todos: Nacimiento, Fallecimiento
  - Turing, Lennon, Einstein: Nacionalidad, Ocupación, Educado en, Padres
  - Turing, Einstein: Supervisor Doctoral, Area
  - Lennon, Turing: Alumno de, Sepultura
  - Eduardo VIII: Predecesor, Sucesor, Entierro, Madre, Casa Real
  - Turing: Estudiantes Doctorales
  - Lennon: Discograficas, Tipo de Voz, Instrumentos

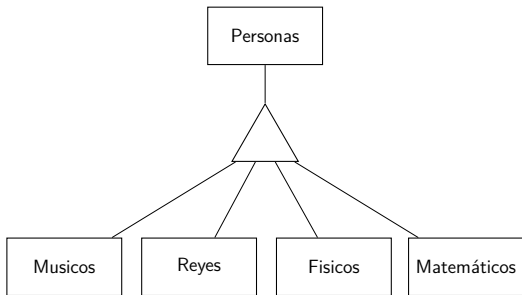
# Motivación: Problema de la Variedad

Cómo modelar para poner a todos en una sola estructura?



# Motivación: Posibles Soluciones?

- Clases y subclasses UML (categorizaciones de MER)



- Si vienen nuevas personas con datos diferentes hay que agregar una nueva clase.

# Motivación: Posibles Soluciones?

- Una tabla relacional  
**Personas( nombre\_nacimiento,  
nacimiento,  
fallecimiento,  
sepultura,  
discograficas,  
...  
)**
  - Si vienen nuevas personas con datos diferentes hay que agregar un nuevo atributo.
  - Muchísimos valores nulos en los atributos que no corresponden.

# Motivación: Otros Problemas

- Hay atributos distintos que representan lo mismo:
  - Lennon: Sepultura
  - Eduardo VIII: Entierro
- Hay atributos iguales pero, que aparentemente, tienen distinto “tipo”.
  - Biden: Residencia - Casa Blanca
  - Turing: Residencia - Maida Vale y Guilford

# Datos Abiertos

*“...pueden ser usados, reusados y redistribuidos por cualquier persona o entidad con la única condición de reconocimiento de las fuentes.”*

*D. Dietrich. **Open Data Handbook**. Open Knowledge Foundation.*

- Estos datos son publicados por organizaciones, usualmente gubernamentales, como un medio para promover la transparencia en la gestión.



# Datos Abiertos: Principios

- Originalmente, se proponía que estos datos fueran:

**Completos** Están todos los datos públicos del organismo.

**Primarios** Los datos deben ser obtenidos en la fuente, con el mayor nivel de granularidad posible, sin ser modificados ni agrupados.

**Oportunos** Los datos deben quedar disponibles tan pronto como sea necesario para preservar su valor.

**Accesibles** Los datos están disponibles para la mayor cantidad de usuarios y propósitos.

**Procesables** *Los datos están razonablemente estructurados, de forma de permitir su procesamiento automático.*

# Datos Abiertos: Principios

- No Discriminatorios** Los datos están disponibles para todos, sin requisito de registro.
- No Propietarios** Los datos están disponibles en un formato sobre el que ninguna entidad tiene control exclusivo.
- Sin Licencia** Los datos no están sujetos a ningún tipo de regulación de derechos, patentes o registros de marca. Se podrán permitir restricciones razonables de privacidad, seguridad o privilegio.

# Publicaciones de Datos Abiertos

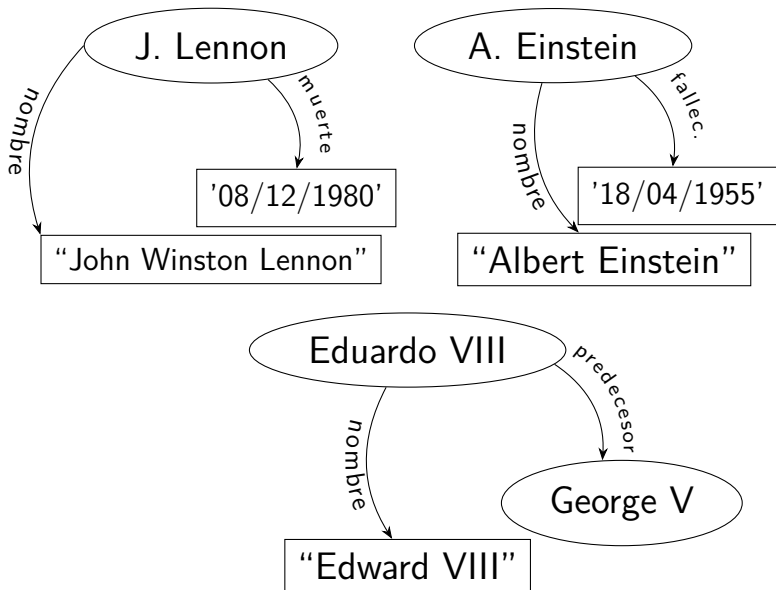
- Gobiernos Nacionales:
  - EEUU: [www.data.gov](http://www.data.gov)
  - UK: [www.data.gov.uk](http://www.data.gov.uk)
  - Uruguay: [catalogodatos.gub.uy](http://catalogodatos.gub.uy)
- Gobiernos Locales:
  - New York: [nycopendata.socrata.com](http://nycopendata.socrata.com)
  - Londres: [data.london.gov.uk](http://data.london.gov.uk)
  - Montevideo: IM - [catalogodatos.gub.uy](http://catalogodatos.gub.uy)
- Diferentes Organizaciones:
  - Banco Mundial: [databank.worldbank.org](http://databank.worldbank.org)
  - Eurostat: [ec.europa.eu/eurostat/data](http://ec.europa.eu/eurostat/data)
  - Naciones Unidas: [data.un.org](http://data.un.org)

# Qué se necesita?

- Un modelo de datos:
  - Flexible
  - Simple
  - General
  - Con un lenguaje de consultas razonable
  - Distribuido

G R A F O S

# Modelo de Datos = Grafos



# Proyectos

- Lod: lod-cloud
- Dbpedia: dbpedia.org

# Contenido del Curso

- Qué es la Web Semántica?
  - Comparación con la Web “Normal”.
  - Arquitectura General.
- Modelos de Datos y Semántica.
  - Repaso de Modelos de Datos.
  - Semántica de los Modelos de Datos: Extensión e Intensión
- Introducción a los Lenguajes y Tecnologías de la Web Semántica.
  - RDF
  - RDFS
  - SPARQL
  - Herramientas

# Contenido del Curso

- RDF y RDFS: Sintaxis y Semántica.
- SPARQL: Consultas y Acceso.
- Razonamiento en la Web Semántica.
  - Reglas para RDF-RDFS.
  - Entailment Regimes.
- OWL
  - OWL : Sintaxis y Semántica.
  - OWL y Modelado de Datos: Perfiles.

Laboratorio



# Evaluación del Curso

- Laboratorio en grupos.
  - Implementación de caso de estudio.
  - Presentación y documento.
- Prueba Final.