



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CONSTRUCCIÓN AUTOMÁTICA DE CAJAS DE INFORMACIÓN PARA WIKIPEDIA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL EN COMPUTACIÓN

TOMÁS ANDRÉS SÁEZ BINELLI

PROFESOR GUÍA:
AIDAN HOGAN

MIEMBROS DE LA COMISIÓN:
GONZALO NAVARRO BADINO
ANDRÉS MUÑOZ ORDENES

SANTIAGO DE CHILE
ENERO 2018

Resumen

Las Infobox son tablas de resumen, que pretenden describir brevemente una entidad mediante la presentación de sus principales características de forma clara y en un formato establecido. Lamentablemente estas Infoboxes son construidas de forma manual por editores de Wikipedia, lo que se traduce en que muchos artículos en idiomas poco frecuentes no cuentan con Infoboxes o éstas son de baja calidad.

Utilizando Wikidata como fuente de información, el desafío de este trabajo es ordenar y seleccionar las propiedades y valores según importancia, para lograr una Infobox concisa con la información ordenada según pertenencia. Con este objetivo en mente, este trabajo propone una estrategia de control y 4 estrategias experimentales para la construcción de Infoboxes en forma automática.

Durante el desarrollo de este trabajo se implementa una API en Django, que se recibe una petición indicando la entidad, el lenguaje y la estrategia a utilizar para generar la Infobox. Como respuesta se obtiene un JSON que representa la Infobox generada. Se construye adicionalmente una interfaz gráfica que permita una rápida utilización de dicha API y opere como facilitador de un proceso de evaluación comparativo entre las diversas estrategias.

La evaluación comparativa se realiza enfrentando a encuestados a un listado de 15 entidades cuyas 5 Infoboxes (una por estrategia) han sido previamente calculadas y dispuestas en forma paralela. Asignando una nota de 1 (menor valoración) a 7, 12 usuarios proceden a evaluar cada Infobox; obteniendo un total de 728 valoraciones. Los resultados indican que la estrategia mejor evaluada combina la frecuencia de una propiedad y el PageRank de su valor como indicadores de orden de importancia.

Agradecimientos

A mi familia por acompañarme y apoyarme durante estos largos ocho años de universidad.

A Alejandra Corvalán, por sus constantes palabras de ánimo y la ayuda con la corrección ortográfica de este informe.

A Gabriel Dintrans, Pedro Rivas y Jaime Castillo por su constante preocupación.

A Francisco Madrid y al equipo de Epistemonikos, por su colaboración y sus consejos en el desarrollo de este trabajo.

A Gonzalo Tartari por proveer los cálculos del algoritmo PageRank.

A Aidan Hogan por su ayuda y su buena disposición durante la totalidad de esta memoria.

Tabla de contenido

1. Introducción	1
1.1. Justificación del trabajo de memoria	3
1.2. Objetivos de la memoria	3
1.3. Metodología de trabajo	4
1.4. Estructura del documento	4
2. Marco teórico	6
2.1. Wikipedia	6
2.2. Wikidata	6
2.3. Web Semántica	7
2.3.1. Resource Description Framework (RDF)	7
2.3.2. SPARQL Protocol and RDF Query Language (SPARQL)	8
2.4. Sistemas de Ranking	9
2.4.1. Relevancia	10
2.4.2. Importancia	10
3. Creación de Infoboxes	12
3.1. Extracción desde Wikidata	12
3.2. Orden y Selección	14
3.2.1. Estrategia Aleatoria	14
3.2.2. Conteo de Frecuencia	15
3.2.3. PageRank	16
3.2.4. Estrategias compuestas	16
4. Arquitectura	18
4.1. Application Programming Interface (API)	19
4.2. Graphical User Interface (GUI)	20
5. Evaluación	23
5.1. Descripción	23
5.2. Evaluación comparativa	26
5.2.1. Grupo Control	26
5.2.2. Frecuencia y PageRank	26
5.2.3. Sistemas híbridos	27
5.3. Evaluación por entidad	27
6. Conclusiones	31

Bibliografía	32
A.	34
A.1. Tablas de promedios por entidad	34
A.2. Tabla total de promedios	35
A.3. Infoboxes de evaluación	36
A.4. Fichas de evaluación	52

Índice de tablas

1.1. Número de artículos por idioma	1
1.2. Ejemplo de relaciones propiedad-valor	2
1.3. Plan de trabajo por semana	4
2.1. Resultados de consulta SPARQL	9
3.1. Valores de X_{max} y X_{min}	17
4.1. Valores de parámetro <i>strategy</i>	19
A.1. Tabla de promedios de estrategia por entidad	34
A.2. Desviación estándar de estrategia por entidad	35
A.3. Tabla de promedios y desviación estándar total por estrategia	35

Índice de ilustraciones

1.1. Ejemplo de Infobox ¹	2
4.1. Arquitectura del proyecto	18
4.2. Modelo de datos	19
4.3. Interfaz de usuario	20
4.4. Selector de idioma	20
4.5. Selector de estrategia	21
4.6. Mensaje de error	21
4.7. Vista de Infobox en la interfaz	22
5.1. Ejemplo de documento de evaluación	25
5.2. Propiedades en Infobox de frecuencia	26
5.3. Promedios de puntaje por entidad	29
5.4. Promedios de puntaje por entidad	30
A.1. Infoboxes automáticas para Douglas Adams	37
A.2. Infoboxes automáticas para Michelle Bachelet	38
A.3. Infoboxes automáticas para David Lynch	39
A.4. Infoboxes automáticas para Chile	40
A.5. Infoboxes automáticas para Zimbabwe	41
A.6. Infoboxes automáticas para Oro	42
A.7. Infoboxes automáticas para Agua	43
A.8. Infoboxes automáticas para Perro	44
A.9. Infoboxes automáticas para Ornitorrinco	45
A.10. Infoboxes automáticas para la Biblia	46
A.11. Infoboxes automáticas para 12 Angry Men	47
A.12. Infoboxes automáticas para Marte	48
A.13. Infoboxes automáticas para Betelgeuse	49
A.14. Infoboxes automáticas para la Torre Eiffel	50
A.15. Infoboxes automáticas para el Museo Guggenheim	51

Capítulo 1

Introducción

Wikipedia es una enciclopedia libre, multilinguaje y colaborativa que cuenta con más de 40 millones de artículos en más de 250 idiomas. Se constituye actualmente como una de las mayores recopilaciones de información en internet, en conjunto con el resto de iniciativas mantenidas por la fundación Wikimedia, dueña y administradora del proyecto.

La construcción y el desarrollo de la enciclopedia se lleva a cabo mediante editores, cuya tarea es crear, editar y completar los artículos. Pero esta estructura de trabajo conlleva una de las grandes desventajas existentes en Wikipedia, la cantidad y calidad de los artículos no aumenta de forma uniforme para todos los idiomas, debido al desequilibrio de los editores por cada uno de estos. Esto se evidencia claramente al analizar la cantidad de artículos por idioma (Tabla 1.1):

Idioma	Nro Artículos
Inglés	5.440.611
Cebuano	4.865.009
Sueco	3.789.676
Alemán	2.076.701
Neerlandés	1.906.225
Francés	1.888.270
Ruso	1.407.358
Italiano	1.369.984
Español	1.334.753
Samareño	1.262.773

Tabla 1.1: Número de artículos por idioma¹

Con respecto a la cantidad de artículos, el problema de crear artículos desde cero es demasiado grande para solucionar de forma rápida, y por tanto dentro del marco de este trabajo. Pero sí se puede realizar un aporte que mejore la calidad de éstos utilizando estrategias alternativas.

¹List of Wikipedias, https://en.wikipedia.org/wiki/List_of_Wikipedias

De forma paralela Wikimedia mantiene el proyecto Wikidata [7], una base de datos similar a Wikipedia, cuya descripción es la siguiente:

*“Wikidata es una base de datos secundaria, gratuita, colaborativa y multilinguaje que recopila información de forma estructurada para apoyar el trabajo de Wikipedia, Wikimedia Commons, otras iniciativas de Wikimedia y a cualquier otra iniciativa en el mundo”*²

Entidad	Propiedad		Valor	
	Id	Etiqueta	Id	Etiqueta
Q310657	-	label	-	Matías Fernández @en Matías Fernández @es
	-	description	-	Chilean footballer @en futbolista chileno @es
	P31	instance of @en instancia de @es	Q5	human @en ser humano @es
	P21	sex or gender @en sexo o género @es	Q6581097	male @en masculino @es
	P27	country of citizenship @en país de nacionalidad @es	Q298	Chile @en Chile @es

Tabla 1.2: Ejemplo de relaciones propiedad-valor

La información en Wikidata se encuentra organizada en entidades, que cuentan con un identificador único construido mediante la letra Q seguida de un número (Ej: Q310657 es el identificador correspondiente al futbolista Matías Fernández). A esta entidad se le asocian datos mediante pares propiedad-valor (Tabla 1.2); las propiedades cuentan, en su mayoría, con un identificador similar al de las entidades (antecedido por la letra P) y se encuentran asociados a un valor, que puede corresponder a un número, una entrada de texto, una entidad, etc.

El factor de interés para este trabajo es el componente multilingüe de Wikidata, pues en esencia ésta carece de un idioma. Toda su información se encuentra disponible mediante identificadores, y son estos los que cuentan con etiquetas para diversos lenguajes de forma simultánea. Esto la coloca como una posible solución para generar Infoboxes de forma automática. Pero diseñar e implementar una aplicación que utilice la información de Wikidata para escribir un artículo en Wikipedia es una tarea de gran magnitud, por lo que se debe elegir una estrategia alternativa.

Una Infobox, como podemos observar en la Figura 1.1, corresponde a un pequeño compendio de las principales características del sujeto del artículo. Se destaca por representar un

Personal information			
Full name	Matías Ariel Fernández Fernández		
Date of birth	15 May 1986 (age 31)		
Place of birth	Caballito, Argentina		
Height	1.78 m (5 ft 10 in)		
Playing position	Attacking midfielder		
Club information			
Current team	Necaxa		
Number	3		
Youth career			
1996–1998	Unión La Calera		
1998–2004	Colo-Colo		
Senior career*			
Years	Team	Apps	(Gls)
2004–2006	Colo-Colo	82	(37)
2006–2009	Villarreal	71	(7)
2009–2012	Sporting CP	69	(12)
2012–2017	Fiorentina	96	(7)
2016–2017	– Milan (loan)	13	(1)
2017–	Necaxa	6	(0)

Figura 1.1: Ejemplo de Infobox³

²Wikidata: Introduction, <https://www.wikidata.org/wiki/Wikidata:Introduction>

³Matías Fernández, https://en.wikipedia.org/wiki/Mat%C3%ADas_Fern%C3%A1ndez

resumen gráfico, accesible y fácil de leer para el lector. Su elaboración se encuentra descrita y normada por los mismos editores de Wikipedia⁴.

Las similitudes entre los datos presentes en un Infobox con el modelo propiedad-valor de Wikidata es lo que justifica este trabajo. El desafío es crear un sistema que, dada una entidad y un idioma, extraiga la información desde Wikidata y la procese de tal forma que seleccione solo las propiedades relevantes y sus valores correspondientes, entregando como resultado un Infobox en el idioma correspondiente.

1.1. Justificación del trabajo de memoria

Wikipedia es una fuente de información muy importante, usada frecuentemente en todo el mundo. Lamentablemente la cantidad de editores en idiomas menos populares es deficiente y las horas-hombre destinadas a crear artículos de calidad no son las necesarias. Esto repercute directamente en la cantidad y la calidad de la información presente en las Infoboxes⁵, debido a que son revisados con menor frecuencia.

Debido a que la creación de un artículo de forma actualizada constituye una tarea de gran dificultad, la propuesta consiste en generar Infoboxes para entidades y lenguajes donde ésta no existe. Además, la creación de una herramienta de esta naturaleza también se presenta como una oportunidad de utilizar la fuente para mantener las Infoboxes al día, pues al ingresar una actualización en Wikidata la información podría propagarse a varios lenguajes de forma instantánea (a diferencia del proceso actual, donde los editores deben actualizar la información en cada idioma manualmente).

La tarea propuesta presenta diversos desafíos desde el punto de vista de la ingeniería de software y las ciencias de la computación. Uno de los principales problemas es procesar la gran cantidad de pares propiedad-valor existentes para cada entidad de Wikidata y seleccionar solo aquellos relevantes para la creación de la Infobox, pues el consenso en la comunidad de editores de Wikipedia es evitar una longitud excesiva. A modo de sortear esta dificultad se espera seleccionar una estrategia de priorización como base (Ej: selección aleatoria) y enfrentarla contra 4 estrategias adicionales. Se espera identificar, mediante un estudio comparativo, qué estrategia resulta en una mejor Infobox y cómo mejorar sobre ella.

1.2. Objetivos de la memoria

El objetivo principal de este trabajo es llevar a cabo el diseño y construcción de una aplicación que permita la generación automática de Infoboxes para artículos de Wikipedia en múltiples idiomas, utilizando Wikidata como fuente de información. De forma adicional se espera alcanzar los siguientes objetivos específicos:

⁴Help: Infobox <https://en.wikipedia.org/wiki/Help:Infobox>

⁵DBPedia Statistics <http://wiki.dbpedia.org/dbpedia-2016-04-statistics>

- Crear un prototipo que dada una entidad en Wikidata realice una consulta SPARQL, con el fin de obtener toda la información necesaria para crear la Infobox correspondiente.
- Implementar diversas estrategias de Ranking que permitan identificar qué propiedades son importantes de incluir dentro de un Infobox y realizar un análisis comparativo entre ellas, a modo de comprender de mejor forma como construirlas mediante la información disponible.
- Construir Infoboxes de forma estandarizada y de similar calidad, sin importar el idioma de salida. Con esto se espera poder aportar significativamente al desarrollo de artículos de Wikipedia en idiomas cuya cantidad de editores es deficiente.

1.3. Metodología de trabajo

Para llevar a cabo esta memoria, se diseñó un plan de trabajo de 16 semanas que permitiera el avance natural del proyecto. Este plan se describe a continuación:

- **T1 - Implementación de prototipo para interfaz gráfica y creación de Infobox:** Etapa inicial, cuyo objetivo principal es realizar una prueba de concepto y entrenar el uso de la API de consulta de Wikidata.
- **T2 - Creación de API e implementación ranking de base:** En este paso se busca consolidar la arquitectura final del proyecto, e implementar la rutina base del trabajo.
- **T3 - Obtención de datos:** Esta fase tiene como objetivo principal recolectar la información necesaria para la realización de los rankings por conteo de frecuencia y PageRank.
- **T4 - Implementación de los sistemas de ranking:** Etapa principal del trabajo, en la cual se implementan las rutinas de ranking en la API.
- **T5 - Evaluación final:** Durante este proceso se crean las herramientas de evaluación y e procede a llevar ésta a cabo.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
T1	X	X	X													
T2			X	X	X											
T3						X	X									
T4								X	X	X	X	X	X			
T5														X	X	X

Tabla 1.3: Plan de trabajo por semana

1.4. Estructura del documento

El siguiente capítulo describe el marco teórico del trabajo, describiendo principalmente los principios técnicos utilizados durante este trabajo. Posteriormente en el capítulo 3 se

describen los procesos de obtención y manejo de información, mientras que el capítulo 4 describe el sistema computacional construido y las aplicaciones implementadas. En el capítulo 5 se describe el proceso de evaluación comparativa que se llevó a cabo, para posteriormente realizar una interpretación de ésta, basada en indicadores estadísticos. Finalmente, el capítulo 6 presenta las conclusiones y el trabajo a futuro.

Capítulo 2

Marco teórico

En este capítulo se describen los conceptos y servicios involucrados en el trabajo de memoria. Además, se describe brevemente la arquitectura y tecnologías utilizadas.

2.1. Wikipedia

Wikipedia es una enciclopedia en línea, libre, políglota y colaborativa creada en enero del año 2001. Es administrada por la Fundación Wikimedia, una organización sin fines de lucro con sede en San Francisco, California.

En la actualidad el proyecto consta de 46 millones de artículos en 288 idiomas, redactados colaborativamente por un conjunto de voluntarios de todo el mundo. La dimensión de este proyecto lo ha transformado en uno de los principales compendios del conocimiento humano y en uno de los diez sitios más visitados de la web.

El éxito y la influencia alcanzada por Wikipedia ha sido un catalizador importante para la creación de proyectos paralelos, dentro de los que se incluyen Wikimedia, Wikiquote, Wikisource, Wikiversity, Wikidata y muchos más; siendo este último el de mayor interés para el desarrollo de este trabajo.

2.2. Wikidata

Wikidata es una base de datos de edición colaborativa basada en documentos, fundada en octubre de 2012, cuya principal función es mantener una fuente común de datos para el resto de los proyectos de la fundación Wikimedia. Su característica más importante es la mantención de un formato de información que facilita la lectura y escritura de datos, tanto para humanos como para computadores.

En la actualidad el proyecto contiene casi 40 millones de entidades descritas¹, soportando etiquetas en cientos de lenguajes². Wikidata basa su estructura de datos en los modelos y estándares propuestos por la iniciativa de Web Semántica.

2.3. Web Semántica

La web que conocemos en la actualidad fue diseñada para el intercambio de documentos, como páginas web. Este diseño es muy útil al extender los usos para los cuales se pensó la estructura inicial, pero presenta variadas dificultades al momento de responder requerimientos más complejos, sobre todo cuando esto involucra múltiples fuentes de información. Esto se debe a que las máquinas no pueden extraer e integrar datos de forma automática si no cuentan con una estructura de base [3].

La Web Semántica es un conjunto de estándares y tecnologías impulsadas por el World Wide Web Consortium (desde ahora en adelante W3C ³) para impulsar páginas Web cuyo contenido sea de fácil lectura y escritura para máquinas. Sus principales objetivos son:

- Impulsar la utilización de formatos comunes para las estructuras de datos en la web.
- Integrar las diversas fuentes de datos en la web, para su acceso universal.

Dentro de la batería de estándares pertenecientes a esta iniciativa se pueden destacar RDF y SPARQL, los cuales son utilizados en este trabajo.

2.3.1. Resource Description Framework (RDF)

RDF [6] es un formato en estructura de grafo usado frecuentemente en la Web. Basa su estructura en la declaración de propiedades para un recurso mediante la estructura sujeto-predicado-objeto, permitiendo establecer relaciones entre dos recursos (sujeto y objeto) mediante el uso de una propiedad (predicado).

- **Sujeto:** recurso que se busca describir.
- **Predicado:** propiedad o relación del sujeto que se busca establecer.
- **Objeto:** puede representar el valor de la propiedad o el otro recurso con el que se establece la relación.

Una de las ventajas de utilizar este formato es que no impone una nomenclatura predeterminada, lo que permite a los usuarios crear recursos, propiedades y valores personalizados. Adicionalmente permite integrar datos de diversas fuentes de forma rápida, pues solo se requiere realizar una unión de grafos. Un buen caso para ejemplificar el uso de RDF, basado en la estructura de información de Wikidata, es el siguiente:

¹Wikidata: Statistics <https://www.wikidata.org/wiki/Wikidata:Statistics>

²Wikipedia Language codes https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

³W3C <https://www.w3.org/>

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
```

Iniciamos nuestro documento declarando la documentación de los prefijos a utilizar. En este caso *rdfs* es parte de las librerías estándar provistas por W3C, *schema* pertenece a una iniciativa colaborativa y *wd* junto con *wdt* pertenecen a Wikidata.

```
wd:Q310657 rdfs:label 'Matías Fernández'@es.
wd:Q310657 schema:description 'futbolista chileno'@es.
```

La primera declaración establece que la etiqueta del elemento *wd:Q310657* es *Matías Fernández*, mientras la segunda indica que su descripción es *futbolista chileno*. De esta forma podemos establecer la existencia de múltiples entidades en nuestro sistema de datos:

```
wd:Q5 rdfs:label 'ser humano'@es.
wd:Q476028 rdfs:label 'club de fútbol'@es.
wd:Q1280895 rdfs:label 'Club Necaxa'@es.
```

Pero la declaración de estas entidades cobra poco sentido si no existe la capacidad de establecer relaciones entre ellas, a modo de aportar un mayor volumen de información. En búsqueda de este objetivo, el siguiente paso es establecer propiedades y utilizarlas para caracterizar objetos entre sí.

```
wdt:P31 rdfs:label 'instancia de'@es.
wdt:P54 rdfs:label 'miembro del equipo deportivo'@es.
```

```
wd:Q310657 wdt:P31 wd:Q5.           #(I)
wd:Q1280895 wdt:P31 wd:Q476028.    #(II)
wd:Q310657 wdt:P54 wd:Q1280895.    #(III)
```

En este paso se establece *wdt:31* como *instance of* y *wdt:54* como *member of sports team*, para posteriormente utilizarlas como propiedades. En este caso se declara a Matías Fernández como un ser humano (I), a Club Necaxa como un club de fútbol (II) y se establece al primero como un miembro del segundo (III).

Esta serie de ejemplos evidencian que la simple estructura de RDF, sumado a la flexibilidad otorgada para crear de entidades y propiedades, otorgan una poderosa herramienta descriptiva a la Web Semántica.

2.3.2. SPARQL Protocol and RDF Query Language (SPARQL)

Todo modelo de datos requiere, para dar utilidad a la información almacenada, un sistema que permita realizar búsquedas de forma eficiente. RDF no es la excepción a esta regla y SPARQL [2] es el estándar impulsado por W3C, que busca satisfacer esa necesidad.

Una de las virtudes de SPARQL es que cuenta con una batería de herramientas muy similares a las utilizadas por su equivalente para bases de datos relacionales, Structured Query Language (SQL). Uniones, consultas anidadas y joins son parte de las herramientas de consulta comunes entre los dos lenguajes, pero SPARQL adiciona nuevas funcionalidades específicas para la búsqueda en RDF, como la posibilidad de retornar la información en forma de grafos.

Una consulta SPARQL básica tiene la siguiente estructura:

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
SELECT ?x
WHERE {
    ?x wdt:P31 wd:Q5 .
    ?x rdfs:label ?label .
}
```

- **PREFIX:** se utiliza para definir la sintaxis de las entradas RDF en la búsqueda.
- **SELECT:** corresponde a los datos que componen los resultados de búsqueda.
- **?:** este carácter se utiliza para declarar variables dentro del contexto de la búsqueda
- **WHERE:** contiene los patrones RDF que deben cumplir los datos, para ser seleccionados en el resultado final.

En el caso de la consulta ejemplificada, al correrla sobre los datos definidos en la estructura RDF de ejemplo, la tabla resultante sería la siguiente:

?x	?label
wd:Q310657	Matias Fernandez

Tabla 2.1: Resultados de consulta SPARQL

2.4. Sistemas de Ranking

Uno de los problemas más frecuentes a la hora de consultar bases de datos de gran magnitud, es diferenciar qué resultados de una búsqueda son más relevantes según el dato de entrada utilizado. Este es el principal caso de estudio de la recuperación de información y es de vital importancia para este trabajo, pues Wikidata cuenta con una gran cantidad de propiedades por cada entidad y para poder crear Infobox de calidad se necesita priorizar la información obtenida.

Un sistema de ranking se utiliza para ordenar los resultados de una búsqueda según su importancia y/o relevancia. En su mayoría estos sistemas logran este objetivo mediante la utilización de una métrica que establece un puntaje a cada resultado, permitiendo establecer un orden de prioridad sobre estos.

La utilización de estos sistemas es amplia, pero casi imperceptible. Google y otros buscadores web utilizan dentro de sus algoritmos de búsqueda sistemas de ranking para garantizar a sus usuarios que los resultados retornados sean relevantes. Dentro de los sistemas de ranking se identifican dos conceptos primordiales a evaluar:

2.4.1. Relevancia

En búsquedas de información, el concepto de relevancia representa la correlación existente entre cada uno de los elementos encontrados y lo que se desea buscar. Al realizar una búsqueda de texto, los documentos relevantes para ese proceso son aquellos que contienen las palabras ingresadas; mientras que aquellos que no las contengan no presentan relación alguna con lo pedido. Un buen ejemplo sería buscar “países de América”, donde se esperaría obtener dentro de los resultados a Chile, Argentina, Venezuela, etcétera; pero cualquier país que no pertenezca al continente americano debiera ser desechado de los resultados por falta de relevancia.

Existen diversos algoritmos para recuperación de información cuyo principal objetivo es evaluar la relevancia de cada documento (como TF-IDF [1]); pero debido a que se necesita priorizar un conjunto de propiedades de una entidad en Wikidata (y, por tanto, de relevancia asegurada) para la creación de una Infobox, este trabajo no requiere de sus usos.

2.4.2. Importancia

El concepto de importancia hace alusión a la necesidad de ordenar los resultados relevantes, a modo de establecer cuáles son de mayor interés. Al realizar la búsqueda “Obama” en Wikipedia, la página redirige de forma automática al perfil del 44to presidente de los Estados Unidos, Barack Obama. En este caso el algoritmo de búsqueda no retorna múltiples artículos como respuesta, pues asume que de la totalidad de resultados éste es el documento más importante. De esta forma se establece un orden de prioridad de este documento por sobre el monte Obama (en Antigua y Barbuda) o la ciudad de Obama (en Japón), pues estos son resultados menos conocidos, y por tanto, menos probables de ser el real objetivo de la búsqueda.

Dentro de los algoritmos de importancia podemos identificar PageRank [5], que corresponde al utilizado por Google para ordenar los resultados de una consulta en su buscador. PageRank se basa en los enlaces existentes entre páginas web para determinar la importancia de cada una, siendo aquellas con mayor cantidad de referencias las cuales destacan. Cabe destacar que PageRank no solo considera el destino de la referencia, sino también el origen; de esta forma los enlaces provenientes de páginas importantes tienen mayor relevancia en la cuantificación final de la escala. Si bien PageRank es el caso más connotado existen otros ejemplos, como HITS [4] que asigna dos formas de importancia. Preferimos PageRank pues solo necesitamos asociar un puntaje a los valores incluidos en una Infobox. Por ejemplo, esperamos que un *Premio Nobel* sea de mayor importancia al compararlo con un *Premio Nacional de Ciencias*, y por tanto debe ser priorizado como un valor a incluir en la Infobox.

Este será el concepto primordial en este trabajo, pues nos permitirá ordenar los valores obtenidos de una entidad en Wikidata, y seleccionar aquellos de mayor importancia para colocarlos en la Infobox correspondiente.

PageRank se aplica recursivamente sobre un grafo dirigido, aplicando la siguiente formula:

$$P_x = \frac{1 - d}{m} + d * \sum_{i=1}^n \frac{P_i}{C_i}$$

- m es la cantidad totales de nodos existentes en el grafo.
- P_i para $(1 \leq i \leq n, i \neq x)$, son todos los nodos con un enlace que apunta hacia el nodo x .
- P_x y P_i corresponden a los valores de PageRank para los nodos x e i respectivamente.
- La constante d es el factor de amortización. Este busca representar la posibilidad de que un usuario pase de un nodo a otro, utilizando las conexiones del grafo como referencia. Si bien se define $d \in [0, 1]$, el consenso de los expertos es que su valor es 0,85.
- C_i es el total de enlaces salientes desde el nodo i .

Esta fórmula se aplica sobre Wikidata modelando cada entidad como un nodo del grafo dirigido y cada triple RDF, cuyo objeto no sea un literal, como una arista dirigida desde el sujeto al objeto.

Capítulo 3

Creación de Infoboxes

Dada una entidad en Wikidata, el objetivo es crear una Infobox que la describa de forma clara, informativa y concisa. Para lograr este proceso se realizan dos pasos de alto nivel:

1. Extraer los datos base correspondientes a la entidad desde Wikidata, limitándonos al lenguaje seleccionado.
2. Ordenar los pares propiedad-valor obtenidos y seleccionar aquellos que conformarán finalmente la Infobox.

En este capítulo se describe con mayor detalle estos dos pasos.

3.1. Extracción desde Wikidata

El primer paso consiste en identificar e implementar la estrategia de extracción de datos desde Wikidata. Para realizar este trabajo se utilizó la interfaz de consulta SPARQL¹ que Wikidata mantiene de forma permanente.

Dada una entidad y un lenguaje, la idea es que al consultar el servicio de Wikidata este nos retorne todas las tuplas cuya propiedad y objeto cuenten con una traducción en el idioma seleccionado.

La consulta a realizar, utilizando como ejemplo la entidad Q1 y el lenguaje español, es la siguiente:

```
SELECT ?pLabel ?prop ?val ?valLabel
WHERE {
  wd:Q1 ?prop ?val .
  ?ps wikibase:directClaim ?prop .
  ?ps rdfs:label ?pLabel .
  SERVICE wikibase:label { bd:serviceParam wikibase:language 'es'. }
```

¹Wikidata Query Service <https://query.wikidata.org/>

```

FILTER((LANG(?pLabel)) = 'es' && (?prop != wdt:P18))
}

```

La cláusula WHERE define en la primera línea la selección de todos los triples que tengan como objeto al elemento seleccionado. Esto corresponde a la totalidad de los descriptores existentes en la base de datos.

Las variables *?prop* y *?val* corresponden solo a los identificadores, por lo que se deben agregar nuevas cláusulas para obtener las etiquetas pertenecientes al idioma seleccionado. La etiqueta de *?val* se obtiene utilizando la cláusula SERVICE, que refiere a un plugin del servicio de Wikidata. Esto permite extraer el nombre del identificador en el idioma seleccionado como el tercer valor del triple descrito (en este caso “es”, que corresponde a español).

En el caso de *?prop* la solución no es tan simple, pues las etiquetas de propiedades se encuentran asociadas al identificador mediante un nodo intermediario. A modo de solución se incluye el segundo y tercer triple presente en la cláusula WHERE, donde la variable *?ps* representa el intermediario y *?pLabel* la etiqueta en cuestión. Para finalizar se procede a filtrar, mediante la cláusula FILTER, todas aquellas etiquetas de propiedades que no correspondan al idioma seleccionado aprovechando de desechar las propiedades que correspondan a imágenes (*wdt:P18*).

El output de esta consulta se obtiene en el siguiente formato JSON:

```

{"head": { ... },
"results": {
  "bindings": [
    {
      "pLabel": [ ],
      "prop": [ ],
      "val": [ ],
      "valLabel": [ ]
    },
    { ... },
    ...
  ]
}}

```

La llave *head* contiene el nombre de las variables declaradas en SELECT, mientras que la llave *bindings* corresponde a la lista de resultados.

Con la finalidad de facilitar el procesamiento de las propiedades en los siguientes pasos, se decide obtener el nombre, la descripción y la imagen de la entidad deseada de forma paralela, pues estas tres características son vitales en la construcción de una Infobox. Por tanto, se crea una segunda consulta al servicio de Wikidata, que nos permite sustraer de forma independiente estos valores.

```

SELECT ?label ?description ?picture
WHERE {

```

```

wd:Q1 rdfs:label ?label .
wd:Q1 schema:description ?description.
OPTIONAL { wd:Q1 wdt:P18 ?picture }
OPTIONAL { wd:Q1 wdt:P41 ?picture }
FILTER((LANG(?label)) = 'es' && (LANG(?description) = 'es'))
}
LIMIT 1

```

Las propiedades *rdfs:label* y *schema:description* contienen la etiqueta y la descripción respectivamente, mientras la cláusula `FILTER` nos permite seleccionar exclusivamente aquellas que se encuentren escritas en el idioma seleccionado.

Para el caso de las imágenes se utiliza la cláusula `OPTIONAL` pues existe la posibilidad de que ésta no esté presente, o en el caso de entidades de tipo país este contenida en otra propiedad (*wdt:P41*, que corresponde a *flag of*).

3.2. Orden y Selección

Los JSON de respuesta a las consultas de extracción no son útiles por si solos; se deben procesar los resultados para seleccionar aquellas características que conformarán las Infobox. Los aspectos clave para crear una Infobox es seleccionar un conjunto acotado de las propiedades más sobresalientes y presentarlas en orden de importancia. Para efectos de este trabajo, se realizó una investigación preliminar de artículos de Wikipedia que estimó que la cantidad de pares propiedad-valor adecuada para la confección de una Infobox es aproximadamente 25. De esta forma se fija 25 como la cantidad de valores a incluir por cada Infobox generada en el sistema.

A continuación, se describen las distintas estrategias para procesar los pares propiedad-valor.

3.2.1. Estrategia Aleatoria

La primera estrategia para procesar los datos corresponde a una selección aleatoria de pares propiedad-valor. Esto a modo de generar una Infobox que cumpla el rol de grupo control al momento de comparar sistemas de ranking. El algoritmo consulta a Wikidata por la lista de pares propiedad-valor y al recibir la respuesta procede a reordenar el registro aleatoriamente.

Para finalizar se procede a recuperar los primeros 25 elementos de la lista, para conformar el conjunto de atributos que conforman la Infobox resultante. A estos se les añade de forma automática la etiqueta de la entidad objetivo, su descripción y su imagen; cortesía de la segunda consulta que se realiza a Wikidata.

El resultado de este algoritmo es un objeto JSON, que cuenta con el siguiente formato:

```
{
  "image": ...,
  "label": ...,
  "description": ...,
  "properties": [{...}, ...]
}
```

- **image:** contiene la dirección web donde se encuentra almacenada la imagen del objeto.
- **label:** corresponde a la etiqueta del objeto.
- **description:** almacena la descripción obtenida.
- **properties:** es la lista de propiedades que conforman el cuerpo de la Infobox resultante, en este caso, aquellas seleccionadas aleatoriamente.

3.2.2. Conteo de Frecuencia

Una forma de atribuir la importancia de un objeto en una muestra es comparar el número de apariciones totales. El sentido común nos dice que, a mayor número de ocurrencias, mayor relevancia adquiere el objeto. La ocupación de una persona es una propiedad que los editores ingresan con mayor frecuencia en Wikidata en comparación con el grupo sanguíneo, por tanto, tiene un mayor conteo absoluto; esto posicionaría la ocupación como una propiedad más sobresaliente (al menos para la mayoría de los casos y usuarios).

Utilizando de base esta hipótesis, la primera estrategia experimental prioriza las propiedades en función de su frecuencia de aparición absoluta en la base de datos de Wikidata.

Para llevar a cabo esta estrategia se realiza una consulta a Wikidata que nos entregue la frecuencia absoluta de ocurrencia para cada propiedad en el sistema de datos.

```
SELECT ?p (count(*) AS ?count)
WHERE { ?s ?p ?o. }
GROUP BY ?p
```

El resultado de esta consulta es una lista de tuplas, en que cada una de ellas contiene:

- El identificador de Wikidata correspondiente a la propiedad representada.
- El número de ocurrencias totales de esa propiedad en la base de datos RDF de Wikidata.

Almacenando esta información de forma previa, el objetivo del algoritmo es consultar por la información correspondiente a Wikidata. A continuación, se procede a obtener, propiedad por propiedad, el valor de la frecuencia correspondiente según el identificador único.

A diferencia de la estrategia aleatoria, la secuencia de propiedades se distribuye utilizando el valor de la frecuencia absoluta como referencia, en orden decreciente. De esta forma el resultado es una secuencia en que las primeras entradas corresponden a las propiedades de la entidad consultada cuya frecuencia absoluta es mayor, sobre la cual se realiza la selección final de los 25 primeros valores para generar la Infobox.

3.2.3. PageRank

El conteo de frecuencia ordena la información de la entidad, basándose en la importancia estimada de la propiedad que la describe, pero no toma en consideración el objeto del triple. Si deseamos construir una Infobox para un actor famoso y debemos incluir dentro de ésta los premios con los que fue galardonado, es de mayor importancia seleccionar un premio Oscar que un premio Donostia (Festival de San Sebastián, España).

Siguiendo esta lógica, se lleva a cabo el cálculo de PageRank para todas las entidades de Wikidata, obteniendo como resultado una lista de tuplas que contiene:

- El identificador de Wikidata correspondiente al nodo representado.
- El valor asignado por el algoritmo PageRank al respectivo nodo.

Almacenando esta información de forma previa, el algoritmo de selección procede a obtener, para cada par propiedad-valor, el valor PageRank correspondiente al sujeto. En el caso de que el sujeto no corresponda a un identificador Wikidata, se procede a asignar 0 como valor.

A continuación, se procede a distribuir la lista de tuplas según el puntaje PageRank calculado, en orden decreciente (a mayor PageRank, más importante es la tupla). Para finalizar, se realiza la selección de los 25 primeros valores para generar la Infobox.

3.2.4. Estrategias compuestas

El conteo de frecuencia y PageRank atacan el problema de orden y selección desde distintas miradas. Aparte de asumir hipótesis diferentes, se basan en distintos componentes de la estructura propiedad-valor, lo que deriva en la aparición de claras desventajas a la hora de generar Infobox.

En el caso del conteo de frecuencia, como fue descrito anteriormente, no es capaz de distinguir diferencias de importancia entre dos valores para una misma propiedad. Mientras que PageRank no es capaz de distinguir si un objeto importante se encuentra asociado mediante una propiedad que no lo es.

Con la finalidad de resolver estos problemas, se diseñan dos estrategias que buscan aprovechar las ventajas de cada una, y minimizar el efecto de sus puntos débiles. La primera consiste en asignar un puntaje al par propiedad-valor basado en la multiplicación de los puntajes de frecuencia y PageRank; mientras la segunda es homóloga, pero aplicando un puntaje aditivo. Para ambos casos se lleva a cabo previamente una normalización de las escalas, evitando magnificar la dispersión de datos.

El algoritmo a seguir es el siguiente:

1. Obtener las tuplas propiedad-valor desde Wikidata.
2. Consultar por el puntaje asociado a cada propiedad, según el conteo de frecuencia.

3. Consultar por el puntaje PageRank asociado a cada sujeto.
4. Realizar la normalización basada en la unidad para ambos valores. Esto se hace utilizando la fórmula:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Siendo X_{max} el máximo valor del conjunto de valores, X_{min} el mínimo, X el valor original y X' el valor normalizado.

	PageRank	Frecuencia
X_{max}	0.041	941678168
X_{min}	0	1

Tabla 3.1: Valores de X_{max} y X_{min}

5. Calcular el puntaje multiplicativo o aditivo, según corresponda.
6. Ordenar el conjunto de forma decreciente en función del valor calculado y seleccionar los primeros 25 pares propiedad-valor para confeccionar la Infobox.

Capítulo 4

Arquitectura

Diseñadas las estrategias de construcción, es necesario construir el software que sea capaz de llevar a cabo todas las tareas requeridas para obtener una Infobox en función de los parámetros requeridos. Para este fin, se diseñan dos servicios:

- El primero es una interfaz de programación de aplicaciones (API), que es la encargada de realizar los procesos de obtención, ordenamiento y selección de los datos. Como tarea final, esta retorna un JSON que contiene la información correspondiente a la Infobox. Su implementación se realiza en Python, utilizando el framework web Django.
- El segundo es una interfaz gráfica de usuario (GUI), que actúa como intermediario entre la API y el consumidor que desea obtener una Infobox para un determinado par entidad-idioma. El principal objetivo es mostrar la información contenida en el JSON en un formato similar a una Infobox de Wikipedia. Su implementación se realiza utilizando Javascript y AngularJS.

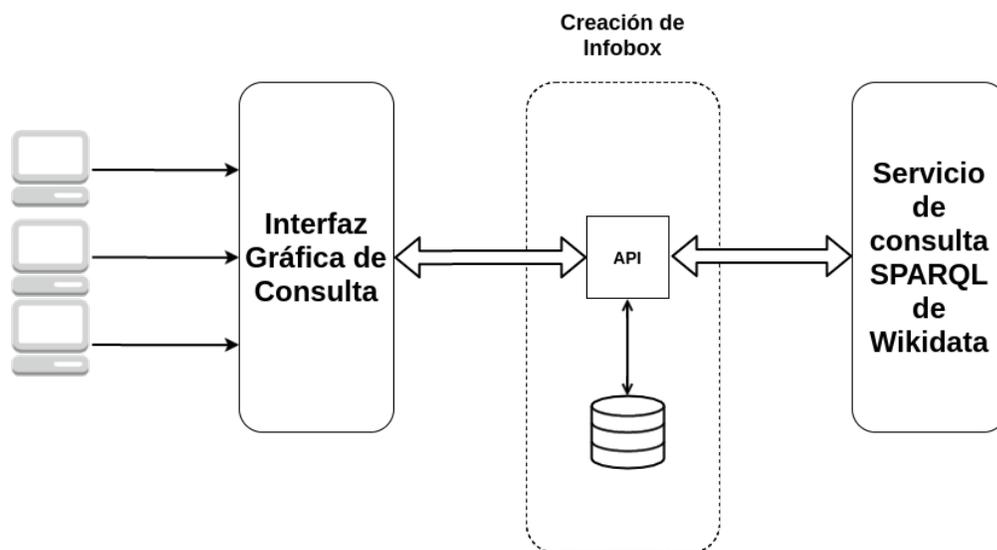


Figura 4.1: Arquitectura del proyecto

4.1. Application Programming Interface (API)

La API es la encargada de obtener la información desde Wikidata y procesarla para obtener la Infobox correspondiente. Para llevar a cabo esta meta cuenta con una dirección de consulta, la cual puede ser accedida mediante una petición HTTP GET a la ruta “/entity”. La consulta a esta ruta debe contar con los siguientes parámetros:

- **id:** corresponde al número del identificador Wikidata para la entidad requerida. Es decir, no se debe incluir la *Q* inicial.
- **lang:** corresponde al código Wikimedia del lenguaje deseado para la Infobox. Ej: “es” para Español, “en” para Inglés y “fr” para Francés.
- **strategy:** esta variable indica a la API que estrategia quiere aplicar el usuario para la creación de la Infobox. Los valores son los siguientes:

Estrategia	Valor
Aleatoria	<i>baseline</i>
Frecuencia	<i>frequency</i>
PageRank	<i>pagerank</i>
Multiplicativa	<i>product</i>
Aditiva	<i>sum</i>

Tabla 4.1: Valores de parámetro *strategy*

Para realizar la confección de las Infobox, la API necesita almacenar en la base de datos los puntajes asociados al conteo de frecuencia para las propiedades y el *score* PageRank para las entidades. Para esto se crea un modelo de datos (Figura 4.2) utilizando las librerías de Django, que nos permite disponer de la información recopilada según lo requerido.

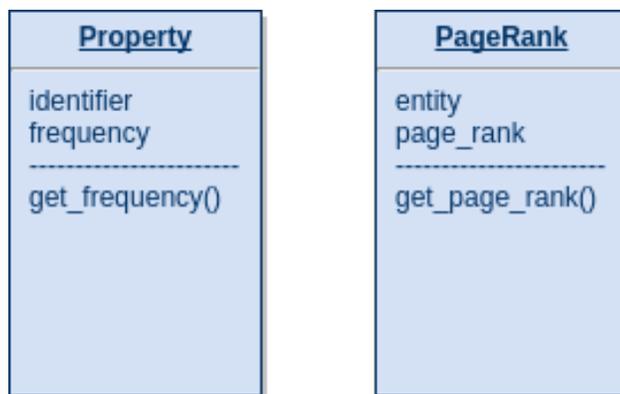


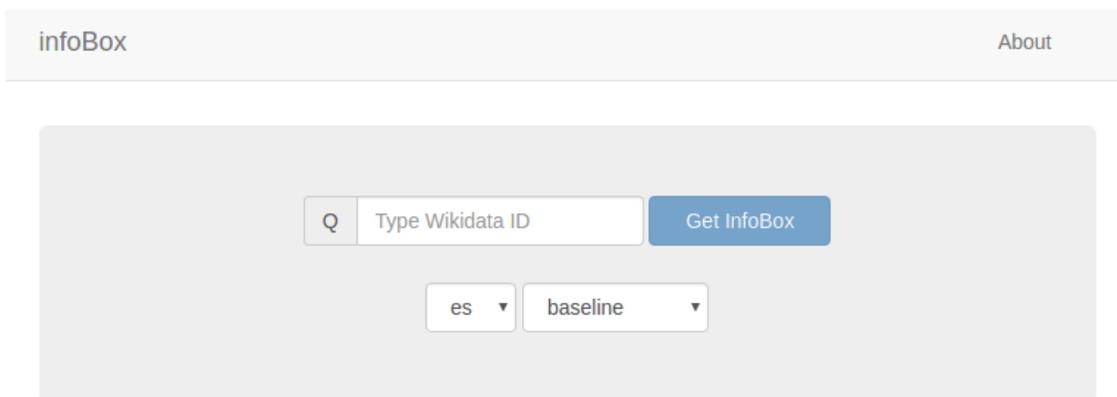
Figura 4.2: Modelo de datos

- **Property:** contiene el identificador de la propiedad (*identifier*) y la información del conteo de frecuencia (*frequency*).
- **PageRank:** contiene el identificador de la entidad (*entity*) y el puntaje PageRank (*page_rank*).

Esta información se ingresa a la base de datos mediante la utilización de *fixtures* en formato JSON, los cuales se encuentran adjuntos en el repositorio Github¹ de la API.

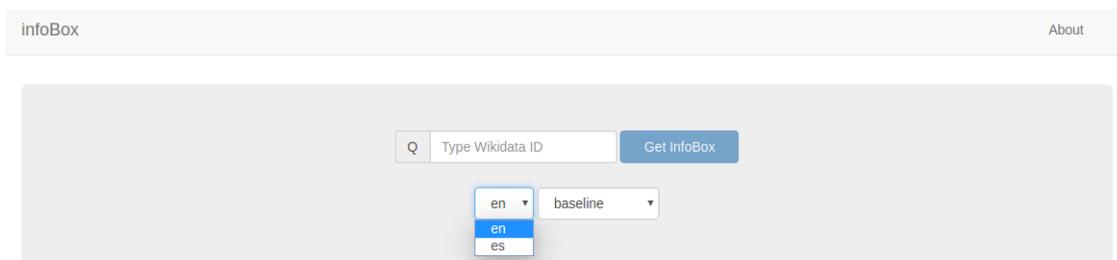
4.2. Graphical User Interface (GUI)

La interfaz gráfica en este proyecto busca generar una vista clara de las Infobox generadas, principalmente para llevar a cabo el proceso de evaluación de estrategias. Ésta consiste en una sola vista, la cual contiene el formulario que permite ingresar los valores de las variables a enviar junto con la petición HTTP a la API.



The screenshot shows the 'infoBox' application header with 'About' on the right. Below is a search form with a magnifying glass icon, a text input field containing 'Type Wikidata ID', and a blue 'Get InfoBox' button. Underneath are two dropdown menus: the first is set to 'es' and the second to 'baseline'.

Figura 4.3: Interfaz de usuario



This screenshot is similar to Figure 4.3 but shows the language dropdown menu open. The menu lists 'en', 'en', and 'es', with the top 'en' option highlighted in blue. The text input field still contains 'Type Wikidata ID' and the 'Get InfoBox' button is visible.

Figura 4.4: Selector de idioma

La interfaz, como puede observarse en la Figura 4.3, cuenta con los siguientes elementos:

- Campo de texto, que permite ingresar el identificador Wikidata de la entidad requerida. Se debe notar que solo se necesita ingresar el numero asociado, ignorando la letra *Q* inicial.
- Campo de selección para el idioma de la Infobox (Figura 4.4).
- Campo de selección para estrategia de orden y selección (Figura 4.5).
- Botón “*Get InfoBox*”, el cual al ser presionado confirma los datos del formulario y envía la petición HTTP. Este se encuentra desactivado por defecto, activándose solo si el campo de texto contiene exclusivamente números.

¹info-api Github repository, <https://github.com/tsaezb/info-api>

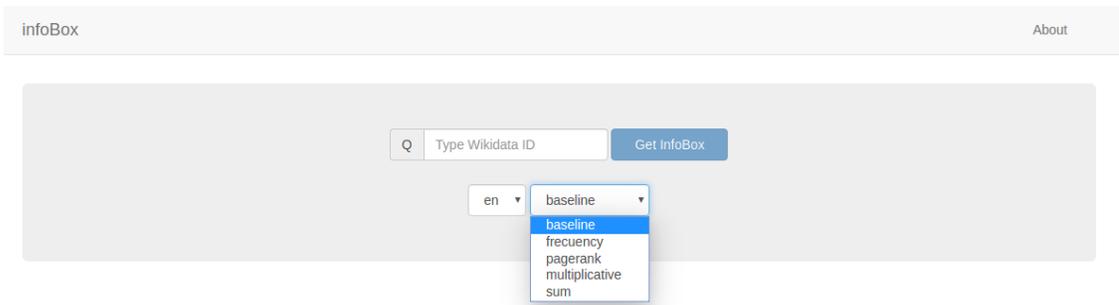


Figura 4.5: Selector de estrategia

Al presionar el botón se ejecuta una consulta HTTP que consulta por la Infobox de la entidad seleccionada, en el idioma de preferencia. La interfaz espera por la respuesta de la API:

- Si esta es negativa (ver Figura 4.6), se muestra un mensaje de error en la parte superior del formulario. Este mensaje puede ser cerrado mediante un click en su costado derecho.

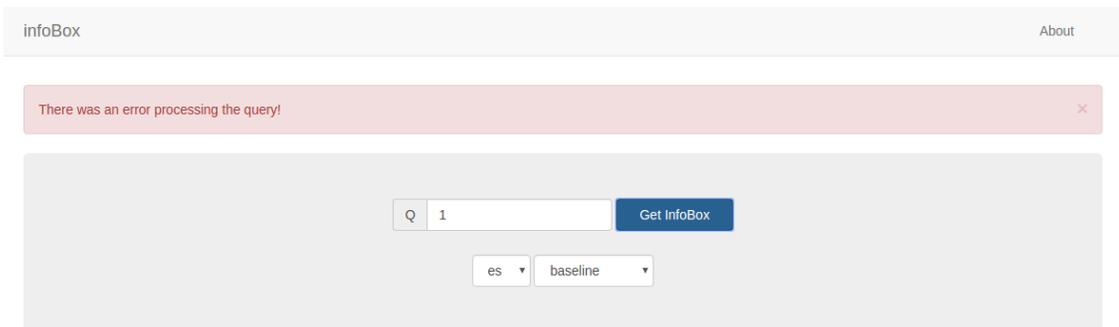


Figura 4.6: Mensaje de error

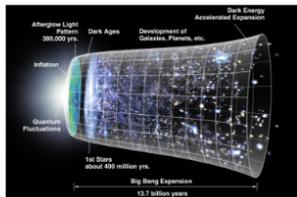
- Si la respuesta es positiva, se procede a mostrar la Infobox resultante en la parte inferior del formulario (Figura 4.7).

Cabe destacar las principales limitaciones en el sistema implementado:

1. Actualmente la GUI solo soporta consultas en español e inglés, debido a que estos fueron los lenguajes de uso para las pruebas preliminares y la evaluación. La extensión del sistema para múltiples lenguajes es simple, pues solo se necesita ingresar los códigos y el nombre de cada uno. La API no cuenta con esta limitación.
2. El sistema requiere que el usuario conozca de antemano el identificador Wikidata de la entidad. Esto podría ser solucionado utilizando la funcionalidad de autocompletado de Wikidata o construyendo un índice invertido para facilitar la obtención del identificador por parte del usuario. Por el momento, se asume que el usuario puede acceder a esta información mediante el uso de un motor de búsqueda externo.

Q 1 Get InfoBox

es baseline



	Universo
Description:	Totalidad del espacio-tiempo. La materia y la energía existentes
compuesto de:	sustancia física, espacio exterior, campo, espacio-tiempo
objeto astronómico hijo:	filamento galáctico
duplicado permanente:	Q22924128
fecha de inicio:	-13798000000-01-01T00:00:00Z
identificador de tema en Quora:	The-Universe
identificador BabelNet:	00022991n
identificador NDL:	00574074
evento clave:	Era hadrónica, Big Bang, Era quark
tiene partes de la clase:	Q28530079, objeto astronómico, materia oscura, energía oscura
identificador Cultureel	natuurkunde-scheikunde-en-

Figura 4.7: Vista de Infobox en la interfaz

Capítulo 5

Evaluación

5.1. Descripción

Uno de los objetivos de este trabajo es realizar una evaluación comparativa de las diferentes estrategias, para identificar cuáles son aquellas que realizan de mejor forma el proceso de ordenamiento y selección.

Para realizar la evaluación se seleccionaron 12 entidades de Wikipedia, procurando mantener una variedad de tipos. Para cada entidad se calcularon las 5 Infobox en español, una por cada estrategia implementada. La selección de los tipos de entidad se realizó utilizando las clases más comunes en Wikidata¹. Los tipos y entidades seleccionadas son:

- **3 Personas**
 - Douglas Adams (*id: Q42*)
 - Michele Bachelet (*id: Q320*)
 - David Lynch (*id: Q2071*)
- **2 Países**
 - Chile (*id: Q298*)
 - Zimbabwe (*id: Q954*)
- **2 Elementos químicos**
 - Oro (*id: Q897*)
 - Agua (*id: Q283*)
- **2 Animales**
 - Perro (*id: Q144*)
 - Ornitorrinco (*id: Q15343*)
- **2 Obras intelectuales**
 - La Biblia (*id: Q1845*)
 - 12 Angry Men (*id: Q2345*)

¹Wikidata:Statistics <https://www.wikidata.org/wiki/Wikidata:Statistics>

- **2 Objetos astronómicos**

- Marte (*id: Q111*)
- Betelgeuse (*id: Q12124*)

- **2 Construcciones**

- Torre Eiffel (*id: Q243*)
- Museo Guggenheim (*id: Q179199*)

Calculadas las Infobox, se procedió a colocar las 5 cajas de cada entidad en forma paralela (ver Figura 5.1) y distribuidas de forma aleatoria. Para mantener un tamaño razonable en el documento, se eliminan las imágenes de las Infobox. Se considera que esto no influye en el resultado de la evaluación pues la imagen es siempre la misma para las 5 estrategias. El resultado es un documento de 15 páginas que fue entregado a 12 entrevistados para evaluar la calidad de las Infobox, los cuales graduaron cada una en una escala de 1 a 7 (siendo 1 la peor calificación y 7 la mejor). Los criterios de evaluación fueron:

- Presencia de información de interés para el perfil de la entidad.
- Existencia de un orden coherente en la presentación de los datos.

Se obtuvieron 146 calificaciones por estrategia, con una desviación estándar promedio de 1,23, lo que habla de un nivel de acuerdo superior a una muestra aleatoria entre los evaluadores.

1	2	3	4	5																																																																																																																																																																																																								
<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr><td>Descripción:</td><td>escritor y humorista británico</td></tr> <tr><td>Identificador VIAF:</td><td>113230702</td></tr> <tr><td>Identificador Encyclopædia Britannica Online:</td><td>biography/Douglas-Adams</td></tr> <tr><td>lenguas habladas:</td><td>inglés</td></tr> <tr><td>Notable Names Database:</td><td>7314000023662</td></tr> <tr><td>Identificador CANTIC:</td><td>a10667040</td></tr> <tr><td>lugar de sepultura:</td><td>Cementerio de Highgate</td></tr> <tr><td>Identificador Kinopoisk de persona:</td><td>246164</td></tr> <tr><td>Identificador Bionet de autor:</td><td>45993</td></tr> <tr><td>Identificador SBN:</td><td>ITVCCURAAV034417</td></tr> <tr><td>Identificador ISFD de autor:</td><td>122</td></tr> <tr><td>empleador:</td><td>BBC</td></tr> <tr><td>madre:</td><td>Janet Adams</td></tr> <tr><td>ocupación:</td><td>guionista</td></tr> <tr><td>instancia de:</td><td>ser humano</td></tr> <tr><td>Identificador NUKAT:</td><td>n194004172</td></tr> <tr><td>altura:</td><td>1.96</td></tr> <tr><td>Identificador NKC:</td><td>j19990000029</td></tr> <tr><td>género artístico:</td><td>sátira</td></tr> <tr><td>alma mater:</td><td>Saint John's College</td></tr> <tr><td>Identificador Perlentaucher:</td><td>douglas-adams</td></tr> <tr><td>nombre en el idioma nativo:</td><td>Douglas Adams</td></tr> <tr><td>obra destacada:</td><td>Q002712. Informe sobre la Tierra: fundamentalmente rotundativa</td></tr> <tr><td>Identificador IMDb:</td><td>nm0010930</td></tr> <tr><td>país de nacionalidad:</td><td>Reino Unido</td></tr> </tbody> </table>	Douglas Adams		Descripción:	escritor y humorista británico	Identificador VIAF:	113230702	Identificador Encyclopædia Britannica Online:	biography/Douglas-Adams	lenguas habladas:	inglés	Notable Names Database:	7314000023662	Identificador CANTIC:	a10667040	lugar de sepultura:	Cementerio de Highgate	Identificador Kinopoisk de persona:	246164	Identificador Bionet de autor:	45993	Identificador SBN:	ITVCCURAAV034417	Identificador ISFD de autor:	122	empleador:	BBC	madre:	Janet Adams	ocupación:	guionista	instancia de:	ser humano	Identificador NUKAT:	n194004172	altura:	1.96	Identificador NKC:	j19990000029	género artístico:	sátira	alma mater:	Saint John's College	Identificador Perlentaucher:	douglas-adams	nombre en el idioma nativo:	Douglas Adams	obra destacada:	Q002712. Informe sobre la Tierra: fundamentalmente rotundativa	Identificador IMDb:	nm0010930	país de nacionalidad:	Reino Unido	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr><td>Descripción:</td><td>escritor y humorista británico</td></tr> <tr><td>instancia de:</td><td>ser humano</td></tr> <tr><td>sexo o género:</td><td>masculino</td></tr> <tr><td>ocupación:</td><td>guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción</td></tr> <tr><td>fecha de nacimiento:</td><td>1952-03-11T00:00:00Z</td></tr> <tr><td>nombre de pila:</td><td>Douglas</td></tr> <tr><td>país de nacionalidad:</td><td>Reino Unido</td></tr> <tr><td>categoría en Commons:</td><td>Douglas Adams</td></tr> <tr><td>lugar de nacimiento:</td><td>Cambridge</td></tr> <tr><td>fecha de fallecimiento:</td><td>2001-05-11T00:00:00Z</td></tr> <tr><td>Identificador Freebase:</td><td>m/0282x</td></tr> <tr><td>Identificador VIAF:</td><td>113230702</td></tr> <tr><td>sitio web oficial:</td><td>http://www.douglasadams.com/</td></tr> <tr><td>lugar de fallecimiento:</td><td>Santa Bárbara</td></tr> <tr><td>género artístico:</td><td>ciencia ficción, comedia, sátira</td></tr> <tr><td>lenguas habladas:</td><td>inglés, inglés británico</td></tr> <tr><td>alma mater:</td><td>Saint John's College, Brentwood School</td></tr> </tbody> </table>	Douglas Adams		Descripción:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	ocupación:	guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción	fecha de nacimiento:	1952-03-11T00:00:00Z	nombre de pila:	Douglas	país de nacionalidad:	Reino Unido	categoría en Commons:	Douglas Adams	lugar de nacimiento:	Cambridge	fecha de fallecimiento:	2001-05-11T00:00:00Z	Identificador Freebase:	m/0282x	Identificador VIAF:	113230702	sitio web oficial:	http://www.douglasadams.com/	lugar de fallecimiento:	Santa Bárbara	género artístico:	ciencia ficción, comedia, sátira	lenguas habladas:	inglés, inglés británico	alma mater:	Saint John's College, Brentwood School	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr><td>Descripción:</td><td>escritor y humorista británico</td></tr> <tr><td>instancia de:</td><td>ser humano</td></tr> <tr><td>sexo o género:</td><td>masculino</td></tr> <tr><td>país de nacionalidad:</td><td>Reino Unido</td></tr> <tr><td>lenguas maternas:</td><td>inglés, inglés británico</td></tr> <tr><td>lenguas habladas:</td><td>inglés, inglés británico</td></tr> <tr><td>residencia:</td><td>Londres, Santa Bárbara</td></tr> <tr><td>religión:</td><td>ateísmo</td></tr> <tr><td>ocupación:</td><td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td></tr> <tr><td>circunstancias de la muerte:</td><td>causas naturales</td></tr> <tr><td>instrumentos:</td><td>guitarra</td></tr> <tr><td>lugar de nacimiento:</td><td>Cambridge</td></tr> <tr><td>género artístico:</td><td>comedia, ciencia ficción</td></tr> <tr><td>causa de muerte:</td><td>infarto agudo de miocardio</td></tr> <tr><td>nombre de pila:</td><td>Douglas</td></tr> <tr><td>empleador:</td><td>BBC</td></tr> <tr><td>lugar de fallecimiento:</td><td>Santa Bárbara</td></tr> </tbody> </table>	Douglas Adams		Descripción:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas maternas:	inglés, inglés británico	lenguas habladas:	inglés, inglés británico	residencia:	Londres, Santa Bárbara	religión:	ateísmo	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	circunstancias de la muerte:	causas naturales	instrumentos:	guitarra	lugar de nacimiento:	Cambridge	género artístico:	comedia, ciencia ficción	causa de muerte:	infarto agudo de miocardio	nombre de pila:	Douglas	empleador:	BBC	lugar de fallecimiento:	Santa Bárbara	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr><td>Descripción:</td><td>escritor y humorista británico</td></tr> <tr><td>instancia de:</td><td>ser humano</td></tr> <tr><td>sexo o género:</td><td>masculino</td></tr> <tr><td>país de nacionalidad:</td><td>Reino Unido</td></tr> <tr><td>lenguas habladas:</td><td>inglés, inglés británico</td></tr> <tr><td>ocupación:</td><td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td></tr> <tr><td>lenguas maternas:</td><td>inglés, inglés británico</td></tr> <tr><td>lugar de nacimiento:</td><td>Cambridge</td></tr> <tr><td>nombre de pila:</td><td>Douglas</td></tr> <tr><td>residencia:</td><td>Londres</td></tr> <tr><td>género artístico:</td><td>comedia, ciencia ficción, sátira</td></tr> <tr><td>religión:</td><td>ateísmo</td></tr> <tr><td>circunstancias de la muerte:</td><td>causas naturales</td></tr> <tr><td>lugar de fallecimiento:</td><td>Santa Bárbara</td></tr> <tr><td>instrumento:</td><td>guitarra</td></tr> <tr><td>alma mater:</td><td>Saint John's College</td></tr> <tr><td>empleador:</td><td>BBC</td></tr> </tbody> </table>	Douglas Adams		Descripción:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas habladas:	inglés, inglés británico	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	lenguas maternas:	inglés, inglés británico	lugar de nacimiento:	Cambridge	nombre de pila:	Douglas	residencia:	Londres	género artístico:	comedia, ciencia ficción, sátira	religión:	ateísmo	circunstancias de la muerte:	causas naturales	lugar de fallecimiento:	Santa Bárbara	instrumento:	guitarra	alma mater:	Saint John's College	empleador:	BBC	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr><td>Descripción:</td><td>escritor y humorista británico</td></tr> <tr><td>instancia de:</td><td>ser humano</td></tr> <tr><td>sexo o género:</td><td>masculino</td></tr> <tr><td>país de nacionalidad:</td><td>Reino Unido</td></tr> <tr><td>lenguas habladas:</td><td>inglés, inglés británico</td></tr> <tr><td>lenguas maternas:</td><td>inglés</td></tr> <tr><td>ocupación:</td><td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td></tr> <tr><td>residencia:</td><td>Londres</td></tr> <tr><td>fecha de nacimiento:</td><td>1952-03-11T00:00:00Z</td></tr> <tr><td>nombre de pila:</td><td>Douglas</td></tr> <tr><td>lugar de nacimiento:</td><td>Cambridge</td></tr> <tr><td>categoría en Commons:</td><td>Douglas Adams</td></tr> <tr><td>fecha de fallecimiento:</td><td>2001-05-11T00:00:00Z</td></tr> <tr><td>Identificador Freebase:</td><td>m/0282x</td></tr> <tr><td>religión:</td><td>ateísmo</td></tr> <tr><td>Identificador VIAF:</td><td>113230702</td></tr> <tr><td>circunstancias de la muerte:</td><td>causas naturales</td></tr> <tr><td>género artístico:</td><td>comedia, ciencia ficción</td></tr> <tr><td>sitio web oficial:</td><td>http://www.douglasadams.com/</td></tr> </tbody> </table>	Douglas Adams		Descripción:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas habladas:	inglés, inglés británico	lenguas maternas:	inglés	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	residencia:	Londres	fecha de nacimiento:	1952-03-11T00:00:00Z	nombre de pila:	Douglas	lugar de nacimiento:	Cambridge	categoría en Commons:	Douglas Adams	fecha de fallecimiento:	2001-05-11T00:00:00Z	Identificador Freebase:	m/0282x	religión:	ateísmo	Identificador VIAF:	113230702	circunstancias de la muerte:	causas naturales	género artístico:	comedia, ciencia ficción	sitio web oficial:	http://www.douglasadams.com/
Douglas Adams																																																																																																																																																																																																												
Descripción:	escritor y humorista británico																																																																																																																																																																																																											
Identificador VIAF:	113230702																																																																																																																																																																																																											
Identificador Encyclopædia Britannica Online:	biography/Douglas-Adams																																																																																																																																																																																																											
lenguas habladas:	inglés																																																																																																																																																																																																											
Notable Names Database:	7314000023662																																																																																																																																																																																																											
Identificador CANTIC:	a10667040																																																																																																																																																																																																											
lugar de sepultura:	Cementerio de Highgate																																																																																																																																																																																																											
Identificador Kinopoisk de persona:	246164																																																																																																																																																																																																											
Identificador Bionet de autor:	45993																																																																																																																																																																																																											
Identificador SBN:	ITVCCURAAV034417																																																																																																																																																																																																											
Identificador ISFD de autor:	122																																																																																																																																																																																																											
empleador:	BBC																																																																																																																																																																																																											
madre:	Janet Adams																																																																																																																																																																																																											
ocupación:	guionista																																																																																																																																																																																																											
instancia de:	ser humano																																																																																																																																																																																																											
Identificador NUKAT:	n194004172																																																																																																																																																																																																											
altura:	1.96																																																																																																																																																																																																											
Identificador NKC:	j19990000029																																																																																																																																																																																																											
género artístico:	sátira																																																																																																																																																																																																											
alma mater:	Saint John's College																																																																																																																																																																																																											
Identificador Perlentaucher:	douglas-adams																																																																																																																																																																																																											
nombre en el idioma nativo:	Douglas Adams																																																																																																																																																																																																											
obra destacada:	Q002712. Informe sobre la Tierra: fundamentalmente rotundativa																																																																																																																																																																																																											
Identificador IMDb:	nm0010930																																																																																																																																																																																																											
país de nacionalidad:	Reino Unido																																																																																																																																																																																																											
Douglas Adams																																																																																																																																																																																																												
Descripción:	escritor y humorista británico																																																																																																																																																																																																											
instancia de:	ser humano																																																																																																																																																																																																											
sexo o género:	masculino																																																																																																																																																																																																											
ocupación:	guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción																																																																																																																																																																																																											
fecha de nacimiento:	1952-03-11T00:00:00Z																																																																																																																																																																																																											
nombre de pila:	Douglas																																																																																																																																																																																																											
país de nacionalidad:	Reino Unido																																																																																																																																																																																																											
categoría en Commons:	Douglas Adams																																																																																																																																																																																																											
lugar de nacimiento:	Cambridge																																																																																																																																																																																																											
fecha de fallecimiento:	2001-05-11T00:00:00Z																																																																																																																																																																																																											
Identificador Freebase:	m/0282x																																																																																																																																																																																																											
Identificador VIAF:	113230702																																																																																																																																																																																																											
sitio web oficial:	http://www.douglasadams.com/																																																																																																																																																																																																											
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																											
género artístico:	ciencia ficción, comedia, sátira																																																																																																																																																																																																											
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																											
alma mater:	Saint John's College, Brentwood School																																																																																																																																																																																																											
Douglas Adams																																																																																																																																																																																																												
Descripción:	escritor y humorista británico																																																																																																																																																																																																											
instancia de:	ser humano																																																																																																																																																																																																											
sexo o género:	masculino																																																																																																																																																																																																											
país de nacionalidad:	Reino Unido																																																																																																																																																																																																											
lenguas maternas:	inglés, inglés británico																																																																																																																																																																																																											
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																											
residencia:	Londres, Santa Bárbara																																																																																																																																																																																																											
religión:	ateísmo																																																																																																																																																																																																											
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																											
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																											
instrumentos:	guitarra																																																																																																																																																																																																											
lugar de nacimiento:	Cambridge																																																																																																																																																																																																											
género artístico:	comedia, ciencia ficción																																																																																																																																																																																																											
causa de muerte:	infarto agudo de miocardio																																																																																																																																																																																																											
nombre de pila:	Douglas																																																																																																																																																																																																											
empleador:	BBC																																																																																																																																																																																																											
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																											
Douglas Adams																																																																																																																																																																																																												
Descripción:	escritor y humorista británico																																																																																																																																																																																																											
instancia de:	ser humano																																																																																																																																																																																																											
sexo o género:	masculino																																																																																																																																																																																																											
país de nacionalidad:	Reino Unido																																																																																																																																																																																																											
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																											
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																											
lenguas maternas:	inglés, inglés británico																																																																																																																																																																																																											
lugar de nacimiento:	Cambridge																																																																																																																																																																																																											
nombre de pila:	Douglas																																																																																																																																																																																																											
residencia:	Londres																																																																																																																																																																																																											
género artístico:	comedia, ciencia ficción, sátira																																																																																																																																																																																																											
religión:	ateísmo																																																																																																																																																																																																											
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																											
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																											
instrumento:	guitarra																																																																																																																																																																																																											
alma mater:	Saint John's College																																																																																																																																																																																																											
empleador:	BBC																																																																																																																																																																																																											
Douglas Adams																																																																																																																																																																																																												
Descripción:	escritor y humorista británico																																																																																																																																																																																																											
instancia de:	ser humano																																																																																																																																																																																																											
sexo o género:	masculino																																																																																																																																																																																																											
país de nacionalidad:	Reino Unido																																																																																																																																																																																																											
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																											
lenguas maternas:	inglés																																																																																																																																																																																																											
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																											
residencia:	Londres																																																																																																																																																																																																											
fecha de nacimiento:	1952-03-11T00:00:00Z																																																																																																																																																																																																											
nombre de pila:	Douglas																																																																																																																																																																																																											
lugar de nacimiento:	Cambridge																																																																																																																																																																																																											
categoría en Commons:	Douglas Adams																																																																																																																																																																																																											
fecha de fallecimiento:	2001-05-11T00:00:00Z																																																																																																																																																																																																											
Identificador Freebase:	m/0282x																																																																																																																																																																																																											
religión:	ateísmo																																																																																																																																																																																																											
Identificador VIAF:	113230702																																																																																																																																																																																																											
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																											
género artístico:	comedia, ciencia ficción																																																																																																																																																																																																											
sitio web oficial:	http://www.douglasadams.com/																																																																																																																																																																																																											

Figura 5.1: Ejemplo de documento de evaluación
 (1) Aleatoria (2) Frecuencia (3) PageRank (4) Multiplicación (5) Suma

5.2. Evaluación comparativa

5.2.1. Grupo Control

La Figura ?? corresponde a un Boxplot de los promedios de puntaje de respuesta por entidad (Anexo A.1). En este se puede observar que los puntajes para la estrategia aleatoria son considerablemente menores que para el resto, constituyéndose como el caso de peor desempeño. Este resultado es totalmente esperable, pues el sistema aleatorio no es capaz de seleccionar consistentemente las propiedades más importantes ni proveer un orden coherente en la presentación de las mismas.

El hecho que las cuatro estrategias restantes evidenciaran un rendimiento considerablemente mayor establece un buen precedente inicial. Sin embargo, se debe considerar que esto se puede deber a las deficiencias del sistema aleatorio, más que a las virtudes de las restantes. Por esta razón la discusión se centra en los grupos experimentales con mayor detalle.

5.2.2. Frecuencia y PageRank

El conteo de frecuencia y el cálculo de PageRank ponen énfasis en distintas características del triple RDF. Mientras la frecuencia evalúa el peso de la propiedad a describir, PageRank pone atención en el valor asignado. El promedio observado en la evaluación le entrega una ventaja considerable al segundo (Frecuencia: $\bar{x} = 3,77$ y PageRank: $\bar{x} = 4,40$), pero a su vez le otorga una mayor dispersión de las calificaciones (Frecuencia: $\sigma = 1,30$ y PageRank: $\sigma = 1,47$).

Como se enunció anteriormente, el mayor problema que enfrenta el conteo de frecuencia es establecer un orden entre los valores de una misma propiedad. Pero al observar detenidamente las Infobox obtenidas mediante este método, podemos observar una falla fundamental en la lógica detrás de su funcionamiento.

fecha de fallecimiento:	2001-05-11T00:00:00Z
identificador Freebase:	/m/0282x
identificador VIAF:	113230702
sitio web oficial:	http://www.douglasadams.com/

Figura 5.2: Propiedades en Infobox de frecuencia

Observando la Figura 5.2, podemos ver insertos en la Infobox identificadores de la entidad para otras páginas web o bases de datos, las cuales no son de interés para el lector de un perfil de Wikipedia. Estos errores son inducidos por los procesos automáticos de edición en Wikidata. La hipótesis inicial establece que las propiedades más importantes serán ingresadas

con mayor regularidad por los editores de contenido, pero esto es cierto en la medida que el proceso sea llevado a cabo por seres humanos. Las maquinas editoras otorgan la posibilidad de ingresar un volumen mucho mayor de información, incluyendo propiedades poco relevantes como estos identificadores.

PageRank es inmune a este problema pues nos asegura que los valores seleccionados son aquellos de mayor importancia en el contexto de Wikidata, pero esta inmunidad conlleva un costo. PageRank no es capaz de medir la importancia de valores literales, pues estos no son considerados a la hora de ejecutar el algoritmo. Un buen ejemplo de este problema es la estatura y el peso de una persona; a ambas propiedades de les asigna como valor magnitudes numéricas, consideradas como literales en el modelo de datos de Wikidata, por lo que nunca serán incluidas en las Infobox calculadas mediante PageRank.

A pesar de esta desventaja, el hecho de que PageRank goce de una mejor calificación podría señalar que los usuarios prefieren un sistema que priorice valores por sobre propiedades.

5.2.3. Sistemas híbridos

En las evaluaciones se puede ver una clara tendencia favorable a la estrategia multiplicativa (Multiplicación: $\bar{x} = 4,77$ y Suma: $\bar{x} = 3,91$), la cual obtuvo el mayor promedio de toda la muestra; mientras la desviación estándar observado es el mismo ($\sigma = 1,44$).

La suma tiende a funcionar como un promedio directo entre ambos puntajes, por lo que inicialmente se esperaba que esta estrategia se posicionara como la mejor. Esta idea se ve reforzada al observar el comportamiento de la multiplicación de coeficientes en el intervalo $[0, 1]$, pues en este caso una baja magnitud en una de las dos variables tiene a tener mayor inferencia en el *score* final.

La tendencia a favor de la adición que se esperaba observar se justifica bajo la idea que ésta permitiría camuflar las debilidades de ambos sistemas de base al no castigar el puntaje final de una propiedad debido a un valor poco importante, o un valor gracias a una propiedad poco frecuente. El mejor ejemplo de este comportamiento son los valores literales, considerados como la debilidad más importante de PageRank, que generan un puntaje multiplicativo igual a *ceros* y un puntaje aditivo idéntico al *score* de frecuencia.

El hecho de que la multiplicación se adjudique un mayor puntaje puede indicar que los triples seleccionados para las Infoboxes deben contar al mismo tiempo con propiedades y valores sobresalientes en la escala.

5.3. Evaluación por entidad

Al observar la Figura 5.4, podemos identificar los promedios de puntaje de cada estrategia, separados por entidad.

Como se observa en la gráfica, la estrategia multiplicativa no solo goza del mejor promedio, sino que se posiciona como la nota más alta para la mayoría de las entidades evaluadas. Esto no es el caso para las siguientes entidades:

- **Agua y Perro**, donde el conteo de frecuencia mantiene un promedio similar.
- **Ornitorrinco**, donde el sistema aleatorio mantiene un promedio similar. Esto puede ser atribuido al hecho de que, dados muchos casos, la estrategia aleatoria logrará generar Infobox de mejor calidad en ciertas oportunidades.
- **Marte y Betelgeuse**, donde PageRank se establece como la mejor estrategia.
- **Torre Eiffel y Museo Guggenheim**, donde PageRank se establece como la mejor estrategia.

Estos dos últimos casos, nos lleva a pensar que cada estrategia puede tener distintos niveles de efectividad, dependiendo del tipo de entidad con el que desea trabajar. Si bien la muestra no es lo suficientemente grande para confirmar esta hipótesis, este podría ser el foco de un trabajo a futuro.

Podría ser esperable que aquellas entidades que tienden a tener descriptores literales de mayor importancia, como los elementos químicos o los objetos astronómicos, la estrategia de frecuencia obtuviera un mejor puntaje, pero la diferencia de promedios observada en estos casos es superior a la tendencia general. Eso hace pensar que, debido al diseño de la evaluación, los usuarios entrevistados no fueron capaces de reconocer propiedades faltantes; sino que solo se limitaron a evaluar las mostradas.

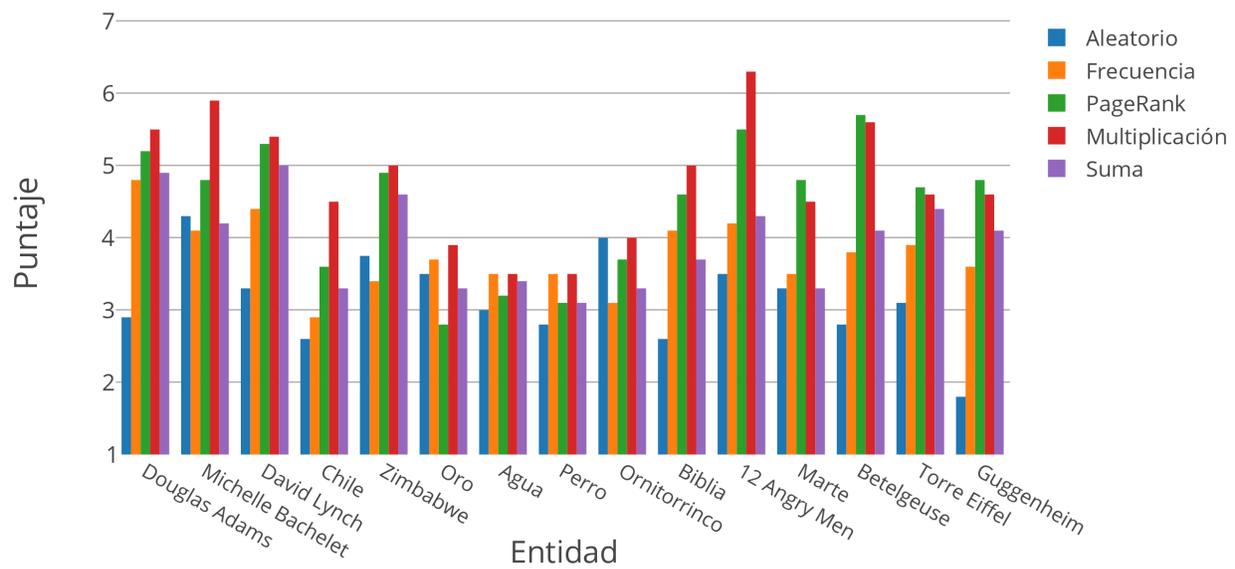


Figura 5.3: Promedios de puntaje por entidad

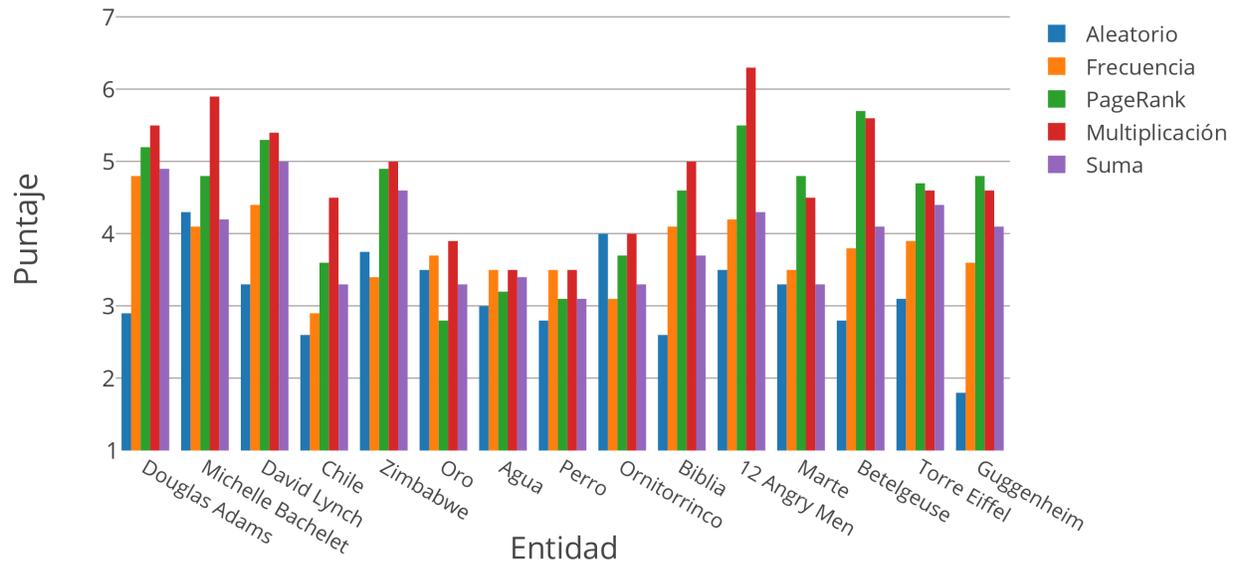


Figura 5.4: Promedios de puntaje por entidad

Capítulo 6

Conclusiones

La finalidad de este trabajo es crear una herramienta que permita disminuir la brecha existente entre la cantidad y calidad de los artículos de Wikipedia en diversos lenguajes. Se propone realizar un análisis de importancia sobre los pares propiedad-valor de un objeto almacenados en Wikidata, para crear una Infobox que resuma sus características y relaciones más importantes.

Al no existir un claro camino a seguir para realizar este proceso, se diseñan cuatro estrategias experimentales y un grupo control, para llevar a cabo un análisis comparativo de la efectividad de cada sistema. Esta evaluación comparativa muestra que existe una tendencia a preferir Infoboxes generadas con una estrategia híbrida, priorizando aquellos pares propiedad-valor que cuenten con un alto puntaje en ambos *score* de base (sistema multiplicativo). Esto era esperable, pues el conteo de frecuencia evalúa la relevancia de una propiedad, mientras PageRank mide la importancia de un valor; y, por tanto, pueden verse como estrategias complementarias.

Se considera que los objetivos planteados inicialmente se han cumplido en el transcurso de este trabajo, y que los resultados obtenidos arrojan una tendencia que sería interesante profundizar. Una forma de hacer esto sería establecer diferentes estrategias de orden y selección para distintos tipos de entidad, pues se observa un nivel de correlación entre los puntajes de entidades de igual tipo. Otras estrategias para extender este trabajo podrían considerar la aplicación de ingeniería inversa sobre las Infoboxes existentes en Wikipedia o utilizar estrategias de *machine learning*.

Adicionalmente se debe profundizar en la idea de un sistema multilingüaje, pues este trabajo no incluye ningún análisis respecto a la generación de Infobox en otros idiomas que no sean inglés o español. Si bien esto se hizo con la finalidad de disminuir las variables en la evaluación, la carencia de etiquetas observada para pares propiedad-valor en ciertos idiomas de Wikidata podría representar una seria amenaza al objetivo principal que se plantea al inicio de este trabajo.

Para finalizar, este trabajo abre más interrogantes sobre la creación automática de Infoboxes. Pero se considera que el trabajo presente establece una base sólida para desarrollar

trabajos futuros en el área.

Bibliografía

- [1] Baeza-Yates, Ricardo A. y Berthier A. Ribeiro-Neto: *Modern Information Retrieval – The concepts and technology behind search, Second edition*. Pearson Education Ltd., Harlow, England, 2011, ISBN 978-0-321-41691-9.
- [2] Harris, Steve, Andy Seaborne y Eric Prud’hommeaux: *SPARQL 1.1 Query Language*. W3C Recommendation, Marzo 2013. <http://www.w3.org/TR/sparql11-query/>.
- [3] Hogan, A.: *Linked Data Management*, capítulo Linked Data and the Semantic Web Standards. CRC Press, 2014.
- [4] Kleinberg, Jon M.: *Authoritative Sources in a Hyperlinked Environment*. J. ACM, 46(5):604–632, 1999. <http://doi.acm.org/10.1145/324133.324140>.
- [5] Page, Lawrence, Sergey Brin, Rajeev Motwani y Terry Winograd: *The PageRank Citation Ranking: Bringing Order to the Web*. Informe técnico 1999-66, Stanford InfoLab, November 1999.
- [6] Schreiber, Guus y Yves Raimond: *RDF 1.1 Primer*. W3C Working Group Note, Junio 2014. <http://www.w3.org/TR/rdf11-primer/>.
- [7] Vrandečić, Denny y Markus Krötzsch: *Wikidata: a free collaborative knowledgebase*. Communications of the ACM, 57(10):78–85, 2014.

Apéndice A

A.1. Tablas de promedios por entidad

	Aleatorio	Frecuencia	PageRank	Multiplicativo	Suma
Douglas Adams	2.90	4.80	5.20	5.50	4.90
Michelle Bachelet	4.30	4.10	4.80	5.90	4.20
David Lynch	3.30	4.40	5.30	5.40	5.00
Chile	2.60	2.90	3.60	4.50	3.30
Zimbabwe	3.75	3.38	4.88	5.00	4.63
Oro	3.45	3.73	2.82	3.91	3.27
Agua	3.00	3.55	3.18	3.55	3.36
Perro	2.82	3.45	3.09	3.45	3.09
Ornitorrinco	4.00	3.11	3.67	4.00	3.33
La biblia	2.56	4.11	4.56	5.00	3.67
12 Angry Men	3.50	4.20	5.50	6.30	4.30
Marte	3.30	3.50	4.80	4.50	3.30
Betelgeuse	2.80	3.80	5.70	5.60	4.10
Torre Eiffel	3.13	3.88	4.67	4.67	4.44
Museo Guggenheim	1.75	3.63	4.75	4.63	4.13

Tabla A.1: Tabla de promedios de estrategia por entidad

	Aleatorio	Frecuencia	PageRank	Multiplicativo	Suma
Douglas Adams	0.99	1.48	1.55	1.08	1.37
Michelle Bachelet	1.48	1.10	0.79	1.25	0.99
David Lynch	1.35	0.67	1.07	1.05	1.34
Chile	0.67	1.35	1.35	0.99	1.35
Zimbabwe	0.92	1.20	1.75	0.83	1.51
Oro	1.08	1.22	1.10	1.79	1.21
Agua	1.41	0.93	1.33	1.21	1.80
Perro	1.45	1.29	1.25	0.83	1.57
Ornitorrinco	1.83	2.12	1.58	1.58	1.41
La Biblia	1.41	1.24	1.73	1.76	1.33
12 Angry Men	1.16	1.35	1.62	1.18	0.82
Marte	1.16	1.27	0.92	0.71	0.95
Betelgeuse	0.82	1.55	1.52	0.92	1.17
Torre Eiffel	0.71	1.12	1.24	0.83	1.13
Museo Guggenheim	1.06	0.71	0.99	1.19	1.16

Tabla A.2: Desviación estándar de estrategia por entidad

A.2. Tabla total de promedios

	Aleatorio	Frecuencia	PageRank	Multiplicativo	Suma
Promedio	3.15	3.77	4.40	4.77	3.91
Desviación estándar	1.36	1.29	1.47	1.44	1.44

Tabla A.3: Tabla de promedios y desviación estándar total por estrategia

A.3. Infoboxes de evaluación

A continuación se juntan las Infoboxes generadas para la evaluación del trabajo

1	2	3	4	5																																																																																																																																																																																																										
<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>escritor y humorista británico</td> </tr> <tr> <td>identificador VIAF:</td> <td>113230702</td> </tr> <tr> <td>identificador Encyclopædia Britannica Online:</td> <td>biography/Douglas-Adams</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés</td> </tr> <tr> <td>Notable Names Database:</td> <td>73100023662</td> </tr> <tr> <td>identificador CANTIC:</td> <td>a10607040</td> </tr> <tr> <td>lugar de sepultura:</td> <td>Cementerio de Highgate</td> </tr> <tr> <td>identificador Kinopoisk de persona:</td> <td>246164</td> </tr> <tr> <td>identificador Bibliothèque de autor:</td> <td>45993</td> </tr> <tr> <td>identificador SBN:</td> <td>IT1CCLJRAVVQ34417</td> </tr> <tr> <td>identificador ISFD de autor:</td> <td>122</td> </tr> <tr> <td>empleador:</td> <td>BBC</td> </tr> <tr> <td>madre:</td> <td>Jane Adams</td> </tr> <tr> <td>ocupación:</td> <td>guionista</td> </tr> <tr> <td>instancia de:</td> <td>ser humano</td> </tr> <tr> <td>identificador NUKAT:</td> <td>r94004172</td> </tr> <tr> <td>altura:</td> <td>1.96</td> </tr> <tr> <td>identificador NK:</td> <td>j11999000029</td> </tr> <tr> <td>género artístico:</td> <td>sátira</td> </tr> <tr> <td>alma máter:</td> <td>Saint John's College</td> </tr> <tr> <td>identificador Perintaucher:</td> <td>Douglas-adams</td> </tr> <tr> <td>nombre en el idioma nativo:</td> <td>Douglas Adams</td> </tr> <tr> <td>obra destacada:</td> <td>¿Qué? 712. Informe sobre la Tierra: fundamentalmente inofensiva</td> </tr> <tr> <td>identificador IMDb:</td> <td>nm0010930</td> </tr> <tr> <td>país de nacionalidad:</td> <td>Reino Unido</td> </tr> </tbody> </table>	Douglas Adams		Description:	escritor y humorista británico	identificador VIAF:	113230702	identificador Encyclopædia Britannica Online:	biography/Douglas-Adams	lenguas habladas:	inglés	Notable Names Database:	73100023662	identificador CANTIC:	a10607040	lugar de sepultura:	Cementerio de Highgate	identificador Kinopoisk de persona:	246164	identificador Bibliothèque de autor:	45993	identificador SBN:	IT1CCLJRAVVQ34417	identificador ISFD de autor:	122	empleador:	BBC	madre:	Jane Adams	ocupación:	guionista	instancia de:	ser humano	identificador NUKAT:	r94004172	altura:	1.96	identificador NK:	j11999000029	género artístico:	sátira	alma máter:	Saint John's College	identificador Perintaucher:	Douglas-adams	nombre en el idioma nativo:	Douglas Adams	obra destacada:	¿Qué? 712. Informe sobre la Tierra: fundamentalmente inofensiva	identificador IMDb:	nm0010930	país de nacionalidad:	Reino Unido	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>escritor y humorista británico</td> </tr> <tr> <td>instancia de:</td> <td>ser humano</td> </tr> <tr> <td>sexo o género:</td> <td>masculino</td> </tr> <tr> <td>ocupación:</td> <td>guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción</td> </tr> <tr> <td>fecha de nacimiento:</td> <td>1952-03-11T00:00:00Z</td> </tr> <tr> <td>nombre de pila:</td> <td>Douglas</td> </tr> <tr> <td>país de nacionalidad:</td> <td>Reino Unido</td> </tr> <tr> <td>categoría en Commons:</td> <td>Douglas Adams</td> </tr> <tr> <td>lugar de nacimiento:</td> <td>Cambridge</td> </tr> <tr> <td>fecha de fallecimiento:</td> <td>2001-05-11T00:00:00Z</td> </tr> <tr> <td>identificador Freebase:</td> <td>/m/6292x</td> </tr> <tr> <td>identificador VIAF:</td> <td>113230702</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.douglasadams.com/</td> </tr> <tr> <td>lugar de fallecimiento:</td> <td>Santa Bárbara</td> </tr> <tr> <td>género artístico:</td> <td>ciencia ficción, comedia, sátira</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>alma máter:</td> <td>Saint John's College, Brentwood School</td> </tr> </tbody> </table>	Douglas Adams		Description:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	ocupación:	guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción	fecha de nacimiento:	1952-03-11T00:00:00Z	nombre de pila:	Douglas	país de nacionalidad:	Reino Unido	categoría en Commons:	Douglas Adams	lugar de nacimiento:	Cambridge	fecha de fallecimiento:	2001-05-11T00:00:00Z	identificador Freebase:	/m/6292x	identificador VIAF:	113230702	sitio web oficial:	http://www.douglasadams.com/	lugar de fallecimiento:	Santa Bárbara	género artístico:	ciencia ficción, comedia, sátira	lenguas habladas:	inglés, inglés británico	alma máter:	Saint John's College, Brentwood School	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>escritor y humorista británico</td> </tr> <tr> <td>instancia de:</td> <td>ser humano</td> </tr> <tr> <td>sexo o género:</td> <td>masculino</td> </tr> <tr> <td>país de nacionalidad:</td> <td>Reino Unido</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>residencia:</td> <td>Londres, Santa Bárbara</td> </tr> <tr> <td>religión:</td> <td>ateísmo</td> </tr> <tr> <td>ocupación:</td> <td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td> </tr> <tr> <td>circunstancias de la muerte:</td> <td>causas naturales</td> </tr> <tr> <td>instrumento:</td> <td>guitarra</td> </tr> <tr> <td>lugar de nacimiento:</td> <td>Cambridge</td> </tr> <tr> <td>género artístico:</td> <td>comedia, ciencia ficción</td> </tr> <tr> <td>causa de muerte:</td> <td>infarto agudo de miocardio</td> </tr> <tr> <td>nombre de pila:</td> <td>Douglas</td> </tr> <tr> <td>empleador:</td> <td>BBC</td> </tr> <tr> <td>lugar de fallecimiento:</td> <td>Santa Bárbara</td> </tr> </tbody> </table>	Douglas Adams		Description:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas habladas:	inglés, inglés británico	lenguas habladas:	inglés, inglés británico	residencia:	Londres, Santa Bárbara	religión:	ateísmo	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	circunstancias de la muerte:	causas naturales	instrumento:	guitarra	lugar de nacimiento:	Cambridge	género artístico:	comedia, ciencia ficción	causa de muerte:	infarto agudo de miocardio	nombre de pila:	Douglas	empleador:	BBC	lugar de fallecimiento:	Santa Bárbara	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>escritor y humorista británico</td> </tr> <tr> <td>instancia de:</td> <td>ser humano</td> </tr> <tr> <td>sexo o género:</td> <td>masculino</td> </tr> <tr> <td>país de nacionalidad:</td> <td>Reino Unido</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>ocupación:</td> <td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>lugar de nacimiento:</td> <td>Cambridge</td> </tr> <tr> <td>nombre de pila:</td> <td>Douglas</td> </tr> <tr> <td>residencia:</td> <td>Londres</td> </tr> <tr> <td>género artístico:</td> <td>comedia, ciencia ficción, sátira</td> </tr> <tr> <td>religión:</td> <td>ateísmo</td> </tr> <tr> <td>circunstancias de la muerte:</td> <td>causas naturales</td> </tr> <tr> <td>lugar de fallecimiento:</td> <td>Santa Bárbara</td> </tr> <tr> <td>instrumento:</td> <td>guitarra</td> </tr> <tr> <td>alma máter:</td> <td>Saint John's College</td> </tr> <tr> <td>empleador:</td> <td>BBC</td> </tr> </tbody> </table>	Douglas Adams		Description:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas habladas:	inglés, inglés británico	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	lenguas habladas:	inglés, inglés británico	lugar de nacimiento:	Cambridge	nombre de pila:	Douglas	residencia:	Londres	género artístico:	comedia, ciencia ficción, sátira	religión:	ateísmo	circunstancias de la muerte:	causas naturales	lugar de fallecimiento:	Santa Bárbara	instrumento:	guitarra	alma máter:	Saint John's College	empleador:	BBC	<table border="1"> <thead> <tr> <th colspan="2">Douglas Adams</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>escritor y humorista británico</td> </tr> <tr> <td>instancia de:</td> <td>ser humano</td> </tr> <tr> <td>sexo o género:</td> <td>masculino</td> </tr> <tr> <td>país de nacionalidad:</td> <td>Reino Unido</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés, inglés británico</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés</td> </tr> <tr> <td>lenguas habladas:</td> <td>inglés</td> </tr> <tr> <td>ocupación:</td> <td>guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante</td> </tr> <tr> <td>residencia:</td> <td>Londres</td> </tr> <tr> <td>fecha de nacimiento:</td> <td>1952-03-11T00:00:00Z</td> </tr> <tr> <td>nombre de pila:</td> <td>Douglas</td> </tr> <tr> <td>lugar de nacimiento:</td> <td>Cambridge</td> </tr> <tr> <td>categoría en Commons:</td> <td>Douglas Adams</td> </tr> <tr> <td>fecha de fallecimiento:</td> <td>2001-05-11T00:00:00Z</td> </tr> <tr> <td>identificador Freebase:</td> <td>/m/6292x</td> </tr> <tr> <td>religión:</td> <td>ateísmo</td> </tr> <tr> <td>identificador VIAF:</td> <td>113230702</td> </tr> <tr> <td>circunstancias de la muerte:</td> <td>causas naturales</td> </tr> <tr> <td>género artístico:</td> <td>comedia, ciencia ficción</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.douglasadams.com/</td> </tr> </tbody> </table>	Douglas Adams		Description:	escritor y humorista británico	instancia de:	ser humano	sexo o género:	masculino	país de nacionalidad:	Reino Unido	lenguas habladas:	inglés, inglés británico	lenguas habladas:	inglés	lenguas habladas:	inglés	ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante	residencia:	Londres	fecha de nacimiento:	1952-03-11T00:00:00Z	nombre de pila:	Douglas	lugar de nacimiento:	Cambridge	categoría en Commons:	Douglas Adams	fecha de fallecimiento:	2001-05-11T00:00:00Z	identificador Freebase:	/m/6292x	religión:	ateísmo	identificador VIAF:	113230702	circunstancias de la muerte:	causas naturales	género artístico:	comedia, ciencia ficción	sitio web oficial:	http://www.douglasadams.com/
Douglas Adams																																																																																																																																																																																																														
Description:	escritor y humorista británico																																																																																																																																																																																																													
identificador VIAF:	113230702																																																																																																																																																																																																													
identificador Encyclopædia Britannica Online:	biography/Douglas-Adams																																																																																																																																																																																																													
lenguas habladas:	inglés																																																																																																																																																																																																													
Notable Names Database:	73100023662																																																																																																																																																																																																													
identificador CANTIC:	a10607040																																																																																																																																																																																																													
lugar de sepultura:	Cementerio de Highgate																																																																																																																																																																																																													
identificador Kinopoisk de persona:	246164																																																																																																																																																																																																													
identificador Bibliothèque de autor:	45993																																																																																																																																																																																																													
identificador SBN:	IT1CCLJRAVVQ34417																																																																																																																																																																																																													
identificador ISFD de autor:	122																																																																																																																																																																																																													
empleador:	BBC																																																																																																																																																																																																													
madre:	Jane Adams																																																																																																																																																																																																													
ocupación:	guionista																																																																																																																																																																																																													
instancia de:	ser humano																																																																																																																																																																																																													
identificador NUKAT:	r94004172																																																																																																																																																																																																													
altura:	1.96																																																																																																																																																																																																													
identificador NK:	j11999000029																																																																																																																																																																																																													
género artístico:	sátira																																																																																																																																																																																																													
alma máter:	Saint John's College																																																																																																																																																																																																													
identificador Perintaucher:	Douglas-adams																																																																																																																																																																																																													
nombre en el idioma nativo:	Douglas Adams																																																																																																																																																																																																													
obra destacada:	¿Qué? 712. Informe sobre la Tierra: fundamentalmente inofensiva																																																																																																																																																																																																													
identificador IMDb:	nm0010930																																																																																																																																																																																																													
país de nacionalidad:	Reino Unido																																																																																																																																																																																																													
Douglas Adams																																																																																																																																																																																																														
Description:	escritor y humorista británico																																																																																																																																																																																																													
instancia de:	ser humano																																																																																																																																																																																																													
sexo o género:	masculino																																																																																																																																																																																																													
ocupación:	guionista, dramaturgo, comediante, escritor de literatura infantil, novelista, escritor de ciencia ficción																																																																																																																																																																																																													
fecha de nacimiento:	1952-03-11T00:00:00Z																																																																																																																																																																																																													
nombre de pila:	Douglas																																																																																																																																																																																																													
país de nacionalidad:	Reino Unido																																																																																																																																																																																																													
categoría en Commons:	Douglas Adams																																																																																																																																																																																																													
lugar de nacimiento:	Cambridge																																																																																																																																																																																																													
fecha de fallecimiento:	2001-05-11T00:00:00Z																																																																																																																																																																																																													
identificador Freebase:	/m/6292x																																																																																																																																																																																																													
identificador VIAF:	113230702																																																																																																																																																																																																													
sitio web oficial:	http://www.douglasadams.com/																																																																																																																																																																																																													
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																													
género artístico:	ciencia ficción, comedia, sátira																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
alma máter:	Saint John's College, Brentwood School																																																																																																																																																																																																													
Douglas Adams																																																																																																																																																																																																														
Description:	escritor y humorista británico																																																																																																																																																																																																													
instancia de:	ser humano																																																																																																																																																																																																													
sexo o género:	masculino																																																																																																																																																																																																													
país de nacionalidad:	Reino Unido																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
residencia:	Londres, Santa Bárbara																																																																																																																																																																																																													
religión:	ateísmo																																																																																																																																																																																																													
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																													
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																													
instrumento:	guitarra																																																																																																																																																																																																													
lugar de nacimiento:	Cambridge																																																																																																																																																																																																													
género artístico:	comedia, ciencia ficción																																																																																																																																																																																																													
causa de muerte:	infarto agudo de miocardio																																																																																																																																																																																																													
nombre de pila:	Douglas																																																																																																																																																																																																													
empleador:	BBC																																																																																																																																																																																																													
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																													
Douglas Adams																																																																																																																																																																																																														
Description:	escritor y humorista británico																																																																																																																																																																																																													
instancia de:	ser humano																																																																																																																																																																																																													
sexo o género:	masculino																																																																																																																																																																																																													
país de nacionalidad:	Reino Unido																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
lugar de nacimiento:	Cambridge																																																																																																																																																																																																													
nombre de pila:	Douglas																																																																																																																																																																																																													
residencia:	Londres																																																																																																																																																																																																													
género artístico:	comedia, ciencia ficción, sátira																																																																																																																																																																																																													
religión:	ateísmo																																																																																																																																																																																																													
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																													
lugar de fallecimiento:	Santa Bárbara																																																																																																																																																																																																													
instrumento:	guitarra																																																																																																																																																																																																													
alma máter:	Saint John's College																																																																																																																																																																																																													
empleador:	BBC																																																																																																																																																																																																													
Douglas Adams																																																																																																																																																																																																														
Description:	escritor y humorista británico																																																																																																																																																																																																													
instancia de:	ser humano																																																																																																																																																																																																													
sexo o género:	masculino																																																																																																																																																																																																													
país de nacionalidad:	Reino Unido																																																																																																																																																																																																													
lenguas habladas:	inglés, inglés británico																																																																																																																																																																																																													
lenguas habladas:	inglés																																																																																																																																																																																																													
lenguas habladas:	inglés																																																																																																																																																																																																													
ocupación:	guionista, novelista, dramaturgo, escritor de ciencia ficción, escritor de literatura infantil, comediante																																																																																																																																																																																																													
residencia:	Londres																																																																																																																																																																																																													
fecha de nacimiento:	1952-03-11T00:00:00Z																																																																																																																																																																																																													
nombre de pila:	Douglas																																																																																																																																																																																																													
lugar de nacimiento:	Cambridge																																																																																																																																																																																																													
categoría en Commons:	Douglas Adams																																																																																																																																																																																																													
fecha de fallecimiento:	2001-05-11T00:00:00Z																																																																																																																																																																																																													
identificador Freebase:	/m/6292x																																																																																																																																																																																																													
religión:	ateísmo																																																																																																																																																																																																													
identificador VIAF:	113230702																																																																																																																																																																																																													
circunstancias de la muerte:	causas naturales																																																																																																																																																																																																													
género artístico:	comedia, ciencia ficción																																																																																																																																																																																																													
sitio web oficial:	http://www.douglasadams.com/																																																																																																																																																																																																													

Figura A.1: Infoboxes automáticas para Douglas Adams

1		2		3		4		5	
Michelle Bachelet		Michelle Bachelet		Michelle Bachelet		Michelle Bachelet		Michelle Bachelet	
Description:	34.* y 36.* presidenta de Chile	Description:	34.* y 36.* presidenta de Chile	Description:	34.* y 36.* presidenta de Chile	Description:	34.* y 36.* presidenta de Chile	Description:	34.* y 36.* presidenta de Chile
instancia de:	ser humano	instancia de:	ser humano	instancia de:	ser humano	premio recibido:	Lista de Forbes de las 100 mujeres del mundo más poderosas, Real Orden de las Santitas, doctor honoris causa, Banda de la Orden Mexicana del Águila Azteca, 100 Mujeres de la BBC, Orden del Libertador	instancia de:	ser humano
lenguas habladas:	inglés, ruso, francés, alemán, español, portugués	sexo o género:	femenino	lenguas habladas:	inglés, ruso, francés, alemán, español, portugués	evento clave:	parto, exilio, interrogatorio	sexo o género:	femenino
sexo o género:	femenino	ocupación:	político, chileno, epidemiólogo	sexo o género:	femenino	lenguas habladas:	ruso, alemán, español, portugués, inglés	ocupación:	político, chileno, epidemiólogo
residencia:	Australá, Leipzig, Santiago de Chile, Potsdam	fecha de nacimiento:	1951-09-29T00:00:00Z	residencia:	Australá, Leipzig	número de hijos:	3	lenguas habladas:	inglés, ruso, francés, alemán, español, portugués
ocupación:	político, chileno	nombre de pila:	Michelle	lenguas habladas:	español	madre:	Angela Jara	país de nacionalidad:	Chile
lengua materna:	español	país de nacionalidad:	Chile	país de nacionalidad:	Chile	militante del partido político:	Partido Socialista de Chile	lengua materna:	español
país de nacionalidad:	Chile	categoría en Commons:	Michelle Bachelet	nombre de pila:	Michelle	lengua materna:	español	religión:	ateísmo
religión:	ateísmo	lugar de nacimiento:	La Cisterna	categoría en Commons:	Michelle Bachelet	cargo ocupado:	Presidente de Chile, Ministro de Salud de Chile	nombre de pila:	Michelle
premio recibido:	Gran Cruz de la Legión de Honor, doctor honoris causa, Collar de la Orden de Isabel la Católica	identificador VIAF:	10960703	lugar de nacimiento:	La Cisterna	residencia:	Santiago de Chile	premio recibido:	Gran Cruz de la Legión de Honor, doctor honoris causa, Collar de la Orden de Isabel la Católica, Collar de la Orden de Carlos III
alma máter:	Universidad Humboldt de Berlín, Universidad de Chile	identificador Freebase:	/m/03awr_	identificador Freebase:	/m/03awr_	identificador BNE:	XX1193225	alma máter:	Universidad Humboldt de Berlín, Universidad de Chile
cargo ocupado:	Presidente de Chile	sitio web oficial:	http://michellebachelet.cl/	religión:	ateísmo	hijo o hija:	Q3554330	cargo ocupado:	Presidente de Chile
evento clave:	parto	premio recibido:	Orden al Mérito de Chile, Q458895, Real Orden de las Santitas, Lista de Forbes de las 100 mujeres del mundo más poderosas, Q2297908, Q200740, Orden del Libertador, Orden Nacional de San Lorenzo, Q968041, Gran Cruz de la Legión de Honor, doctor honoris causa, caballero de Gran Cruz de la Orden al Mérito de la República Italiana	identificador VIAF:	10960703	alma máter:	Universidad de Chile		
nombre de pila:	Michelle	sitio web oficial:	http://michellebachelet.cl/	premio recibido:	Gran Cruz de la Legión de Honor, doctor honoris causa	sexo o género:	femenino		

Figura A.2: Infoboxes automáticas para Michelle Bachelet

1		2		3		4		5	
David Lynch		David Lynch		David Lynch		David Lynch		David Lynch	
Descripción:	director de cine, productor de música electrónica y guionista estadounidense	Descripción:	director de cine, productor de música electrónica y guionista estadounidense	Descripción:	director de cine, productor de música electrónica y guionista estadounidense	Descripción:	director de cine, productor de música electrónica y guionista estadounidense	Descripción:	director de cine, productor de música electrónica y guionista estadounidense
instancia de:	ser humano	instancia de:	ser humano						
sexo o género:	masculino	sexo o género:	masculino						
país de nacionalidad:	Estados Unidos	país de nacionalidad:	Estados Unidos						
lenguas habladas:	inglés	lenguas habladas:	inglés						
ocupación:	actor, pintor, músico, compositor, director de cine, guionista, actor de cine, actor de televisión, compositor de canciones, actor de voz, letrista, editor de cine, animador	ocupación:	actor, pintor, músico, compositor, director de cine, guionista, actor de cine, actor de televisión, compositor de canciones, actor de voz, letrista, editor de cine, animador	ocupación:	guionista, actor, compositor, animador, músico, compositor de canciones, letrista, pintor, actor de voz, director de cine, editor de cine, actor de televisión, actor de cine	ocupación:	actor, pintor, músico, compositor, director de cine, guionista, actor de cine, actor de televisión, compositor de canciones, actor de voz, letrista, editor de cine, animador	ocupación:	letrista, actor, director de cine, actor de televisión
nombre de pila:	David	nombre de pila:	David						
premio recibido:	Legión de Honor, Oficial de la Legión de Honor	premio recibido:	Legión de Honor, Oficial de la Legión de Honor	premio recibido:	Premios Saturn, Legión de Honor	premio recibido:	Legión de Honor	premio recibido:	Legión de Honor
miembro del partido político:	Partido Demócrata	miembro del partido político:	Partido Demócrata						
género artístico:	cine de terror	género artístico:	cine de terror						
alma máter:	Pennsylvania Academy of the Fine Arts, Universidad George Washington	alma máter:	Pennsylvania Academy of the Fine Arts	alma máter:	cine de terror	alma máter:	Pennsylvania Academy of the Fine Arts	alma máter:	Pennsylvania Academy of the Fine Arts, Universidad George Washington
residencia:	Los Ángeles	residencia:	Los Ángeles						
identificador genealogics.org de persona:	K00616201	identificador genealogics.org de persona:	K00616201						
identificador BNE:	XX832488	identificador BNE:	XX832488						
identificador GND:	119053316	identificador GND:	119053316						
número de hijos:	4	número de hijos:	4						
fecha de nacimiento:	1946-01-20T00:00:00Z	fecha de nacimiento:	1946-01-20T00:00:00Z						
lugar de nacimiento:	Los Ángeles	lugar de nacimiento:	Los Ángeles	lugar de nacimiento:	Missoula	lugar de nacimiento:	Missoula	lugar de nacimiento:	Missoula
identificador Freebase:	/m/026dx	identificador Freebase:	/m/026dx						
identificador VIAF:	12106923	identificador VIAF:	12106923						
sitio web oficial:	http://www.davidlynch.com/	sitio web oficial:	http://www.davidlynch.com/						
premio recibido:	Legión de Honor	premio recibido:	Legión de Honor						
identificador autoridades de la Biblioteca del Congreso de EE.UU.:	n92021345	identificador autoridades de la Biblioteca del Congreso de EE.UU.:	n92021345	identificador autoridades de la Biblioteca del Congreso de EE.UU.:	n92021345	identificador autoridades de la Biblioteca del Congreso de EE.UU.:	n92021345	identificador autoridades de la Biblioteca del Congreso de EE.UU.:	n92021345
identificador Discogs de artista:	47572	identificador Discogs de artista:	47572						
identificador NDL:	00484298	identificador NDL:	00484298						
identificador de tema en Quora:	David Lynch-Director-1	identificador de tema en Quora:	David Lynch-Director-1	identificador de tema en Quora:	David Lynch-Director-1	identificador de tema en Quora:	David Lynch-Director-1	identificador de tema en Quora:	David Lynch-Director-1

Figura A.3: Infoboxes automáticas para David Lynch

1

Chile	
Descripción:	país de América del Sur
relación diplomática:	Estados Unidos, Reino Unido, República Popular China, Francia, Alemania, Rusia, Indonesia, Australia, Italia, México, España, India, Brasil, Finlandia, Turquía, Dinamarca, Bielorrusia
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas
descrito en:	Diccionario Enciclopédico Brockhaus y Eron, Q11180675, Q4173137
país:	Chile
miembro de:	Organización de las Naciones Unidas

2

Chile	
Descripción:	país de América del Sur
teléfono de emergencia:	134 Q25648813
identificador New York Times de tema:	destruativochile
relación diplomática:	Haití, El Salvador, Brasil
duplicado permanente:	Q22929367
miembro de:	Organización para la Cooperación y el Desarrollo Económico
descrito en:	Q19180675
comparte fronteras con:	Perú
gentilicio:	chileno, chilena
mayoría de edad:	18
se conduce en la:	derecha
jefe de Estado:	Michelle Bachelet
punto más alto:	Nevado Ojos del Salado
tasa de desempleo:	6
subredes:	chile
categoría para las personas que fallecieron aquí:	Q6305403
Clasificación Decimal de Dewey:	2-83
punto más bajo:	océano Pacífico
huso horario:	UTC-5
historia:	Historia de Chile
población:	17619708
identificador de tema en Quora:	Chile

3

Chile	
Descripción:	país de América del Sur
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas
país:	Chile
descrito en:	Diccionario Enciclopédico Brockhaus y Eron, Q19180675, Q4173137
huso horario:	UTC-5, UTC-3
comparte fronteras con:	Argentina, Bolivia, Perú
miembro de:	Organización de las Naciones Unidas, Organismo Multilateral de Garantía de Inversiones, Banco Internacional de Reconstrucción y Fomento, Corporación Financiera Internacional, Asociación Internacional de Fomento, Organización Mundial del Comercio, Centro Internacional de Arreglo de Diferencias Relativas a Inversiones, Organización para la Cooperación y el Desarrollo Económico, Foro de Cooperación Económica Asia-Pacífico
relación diplomática:	Estados Unidos, Reino Unido, República Popular China
forma parte de:	América Latina

4

Chile	
Descripción:	país de América del Sur
instancia de:	país, Estado miembro de las Naciones Unidas, Estado soberano
país:	Chile
coordenadas:	Puñe(71.0 -31.0)
identificador Geonames:	3895114
categoría en Commons:	Chile
huso horario:	UTC-5, UTC-3
identificador Freebase:	/m/01j1v
identificador VIAF:	233665742
subdividido en (división administrativa):	Región de Ñuble
fecha de fundación o creación:	1818-01-01T00:00:00Z
identificador OND:	4009929-5
forma parte de:	América Latina
categoría principal del tema:	Categoría:Chile
comparte fronteras con:	Argentina, Perú, Bolivia
población:	17619708
miembro de:	Organización de las Naciones Unidas, Unión de Naciones Suramericanas, Organización Mundial del Comercio, Organización para la Cooperación y el Desarrollo Económico, Foro de Cooperación Económica Asia-Pacífico

5

Chile	
Descripción:	país de América del Sur
relación diplomática:	Estados Unidos, Reino Unido, República Popular China, Francia, Alemania, Rusia, Indonesia, Australia, Italia, México, España, India, Brasil, Finlandia, Turquía, Dinamarca, Bielorrusia
descrito en:	Diccionario Enciclopédico Brockhaus y Eron, Q19180675, Q4173137
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas
miembro de:	Organización de las Naciones Unidas
l idioma oficial:	español

Figura A.4: Infoboxes automáticas para Chile

1

Description:	país de África
instancia de:	país, república, Estado sin litoral, Estado miembro de las Naciones Unidas, Estado soberano
país:	Zimbabue
coordenadas:	Punt(30.016667 -19.016667)
identificador GeoNames:	878675
categoría en Commons:	Zimbabue
huso horario:	UTC+02:00
identificador Freebase:	/m/088q4
identificador VIAF:	128940544
subdividido en (división administrativa):	Mashonalandia Central, Masvingo, Harare, Bulawayo, Midlands, Matabelandia Septentrional, Mashonalandia Occidental, Manicalandia, Mashonalandia Oriental, Matabelandia Meridional
identificador GND:	4049850-5
categoría principal del tema:	Categoría:Zimbabue
comparte fronteras con:	Sudáfrica

2

Description:	país de África
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas, república, Estado sin litoral
huso horario:	UTC+02:00
país:	Zimbabue
miembro de:	Organización de las Naciones Unidas, Organismo Multilateral de Garantía de Inversiones, Banco Internacional de Reconstrucción y Fomento, Corporación Financiera Internacional, Asociación Internacional de Fomento, Organización Mundial del Comercio, Centro Internacional de Arreglo de Diferencias Relativas a Inversiones, Banco Africano de Desarrollo, Comunidad Híbrida de la Unión Africana y las Naciones Unidas en Darfur, Estados de África, del Caribe y del Pacífico
comparte fronteras con:	Sudáfrica, Zambia, Botswana, Mozambique
relación diplomática con:	República Popular China, Alemania, Rusia
idioma oficial:	inglés

3

Description:	país de África
identificador MusicBrainz de área:	691e8a96-2c09-3678-8495-1a2960bde055
se conduce en la:	izquierda
idioma oficial:	Idioma ndebele del norte, venda, tsonga
nombrado por:	Gran Zimbabue
punto más septentrional:	Punt(29.62 -15.6)
subdividido en (división administrativa):	Mashonalandia Oriental, Matabelandia Septentrional, Bulawayo
idioma usado:	inglés
electricidad doméstica:	220
FPIS 10-4 (países y regiones):	ZI
relación diplomática:	Sudáfrica
identificador Guardian de tema:	world/continent/
índice de desarrollo humano:	0.509
gentilicio:	ፖዋጋጋይ
punto más alto:	Monte Nyangani
identificador Catholic Encyclopedia:	13025a
geografía:	geografía de Zimbabue
identificador Tesoro de arte y arquitectura:	300293701
categoría principal del tema:	Categoría:Zimbabue
instancia de:	país
miembro de:	Centro Internacional de Arreglo de Diferencias Relativas a Inversiones
categoría para las personas que nacieron aquí:	Q8154341

4

Description:	país de África
idioma usado:	inglés
idioma oficial:	inglés
relación diplomática:	República Popular China, Alemania, Rusia, Australia, Israel, Grecia, Serbia, Angola, Pakistán
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas, república
huso horario:	UTC+02:00
miembro de:	Organización de las Naciones Unidas, Organismo Multilateral de Garantía de Inversiones, Banco Internacional de Reconstrucción y Fomento, Corporación Financiera Internacional, Asociación Internacional de Fomento, Organización Mundial del Comercio, Centro Internacional de Arreglo de Diferencias Relativas a Inversiones, Banco Africano de Desarrollo
se conduce en la:	izquierda

5

Description:	país de África
idioma oficial:	inglés
idioma usado:	inglés
relación diplomática:	República Popular China, Alemania, Rusia, Australia, Israel, Grecia, Serbia
instancia de:	Estado soberano, país, Estado miembro de las Naciones Unidas, república, Estado sin litoral
huso horario:	UTC+02:00
miembro de:	Organización de las Naciones Unidas, Organismo Multilateral de Garantía de Inversiones, Banco Internacional de Reconstrucción y Fomento, Corporación Financiera Internacional, Asociación Internacional de Fomento, Organización Mundial del Comercio, Centro Internacional de Arreglo de Diferencias Relativas a Inversiones, Banco Africano de Desarrollo
país:	Zimbabue
coordenadas:	Punt(30.016667 -19.016667)

Figura A.5: Infoboxes automáticas para Zimbabue

1		2		3		4		5	
oro		oro		oro		oro		oro	
Description:	elemento químico con número atómico 79	Description:	elemento químico con número atómico 79	Description:	elemento químico con número atómico 79	Description:	elemento químico con número atómico 79	Description:	elemento químico con número atómico 79
descrito en:	Enciclopedia soviética armenia	instancia de:	elemento químico, ingrediente culinario	instancia de:	elemento químico, ingrediente culinario	descrito en:	Enciclopedia soviética armenia	identificador BNE:	XX526003
catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	descrito en:	Enciclopedia soviética armenia	subclase de:	metal, elementos del grupo 11, elementos del periodo 6	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	subclase de:	metal
instancia de:	elemento químico, ingrediente culinario	subclase de:	metal, elementos del periodo 6, elementos del grupo 11	categoría en Commons:	Gold	instancia de:	elemento químico, ingrediente culinario	identificador Freebase:	/m/025n2z
subclase de:	metal, elementos del periodo 6, elementos del grupo 11	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	identificador Freebase:	/m/025n2z	subclase de:	metal, elementos del periodo 6, elementos del grupo 11	número ZVG:	8480
color:	oro	color:	oro	identificador GND:	4157819-3	categoría en Commons:	Gold	identificador NCF-RT:	N000022008
condición médica tratada:	artritis reumatoide	categoría principal del tema:	Categoría:Oro	categoría principal del tema:	Categoría:Oro	identificador Freebase principal del tema:	/m/025n2z	símbolo químico:	Au
categoría principal del tema:	Categoría:Oro	duplicado permanente:	Q27829821	identificador autoridades de la Biblioteca del Congreso de EELUJ:	in80505692	identificador GND:	4157819-3	definición de significado de Omegat:	5912
duplicado permanente:	Q27829821	condición médica tratada:	artritis reumatoide	identificador BnF:	133184606	categoría principal del tema:	Categoría:Oro	galera en Commons:	Gold
identificador BnF:	133184606	identificador BnF:	133184606	masa:	196.967	identificador autoridades de la Biblioteca del Congreso de EELUJ:	in80505692	código ISO 4217:	XAU
fórmula química:	Au	fórmula química:	Au	descrito en:	Enciclopedia soviética armenia	identificador BnF:	133184606	identificador PubChem:	23985
categoría en Commons:	Gold	categoría en Commons:	Gold	identificador de tema en Quora:	Gold	color:	oro	descrito en:	Enciclopedia soviética armenia
identificador NDL:	00565802	identificador NDL:	00565802	SMILES canónico:	[Au]	masa:	196.967	número CE:	231-165-9
identificador QID:	4157819-3	identificador QID:	4157819-3	hChKey:	PCH3UWPFVWCPO-UHFFFAOYSA-N	identificador de tema en Quora:	Gold	descubridor o inventor:	892043953
descubridor o inventor:	892043953	descubridor o inventor:	892043953	fórmula química:	Au	SMILES canónico:	[Au]	instancia de:	ingrediente culinario, elemento químico
número CAS:	7440-57-5	número CAS:	7440-57-5	hChE:	15/Au	hChKey:	PCH3UWPFVWCPO-UHFFFAOYSA-N	capacidad calorífica:	25.418
número CE:	231-165-9	número CE:	231-165-9	identificador PubChem:	23985	fórmula química:	Au	identificador autoridades de la Biblioteca del Congreso de EELUJ:	in80505692
hChKey:	PCH3UWPFVWCPO-UHFFFAOYSA-N	hChKey:	PCH3UWPFVWCPO-UHFFFAOYSA-N	color:	oro	hChE:	15/Au	categoría principal del tema:	Categoría:Oro
identificador autoridades de la Biblioteca del Congreso de EELUJ:	in80505692	identificador autoridades de la Biblioteca del Congreso de EELUJ:	in80505692	identificador ChemSpider:	22421	identificador PubChem:	23985	identificador Tesoro de la UNESCO:	concep111561
símbolo químico:	Au	símbolo químico:	Au	galera en Commons:	Gold	identificador ChemSpider:	22421	condición médica tratada:	artritis reumatoide
SMILES canónico:	[Au]	SMILES canónico:	[Au]	identificador NDL:	00565802	galera en Commons:	Gold	punto de ebullición:	5378
hChE:	15/Au	hChE:	15/Au	número CAS:	7440-57-5	identificador NDL:	00565802	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener
código ISO 4217:	XAU	código ISO 4217:	XAU	descubridor o inventor:	892043953	número CAS:	7440-57-5	número CAS:	7440-57-5
						identificador Nomisma:	iv		

Figura A.6: Infoboxes automáticas para Oro

1	2	3	4	5																																																																																																																																																		
<table border="1"> <thead> <tr> <th></th> <th>agua</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno</td> </tr> <tr> <td>color:</td> <td>marón oscuro, rojo</td> </tr> <tr> <td>Identificador Enciclopedia Británica Online:</td> <td>science/water</td> </tr> <tr> <td>densidad absoluta:</td> <td>0.999802, 0.998117, 0.99403, 0.96862</td> </tr> <tr> <td>Fórmula química:</td> <td>H₂O</td> </tr> <tr> <td>código NFPA de fuego:</td> <td>0</td> </tr> <tr> <td>Identificador KEGG:</td> <td>C00001</td> </tr> <tr> <td>Identificador BICF:</td> <td>0045</td> </tr> <tr> <td>código NFPA de salud:</td> <td>0</td> </tr> <tr> <td>base conjugada:</td> <td>hidróxido</td> </tr> <tr> <td>Identificador de tema en Quora:</td> <td>Water</td> </tr> <tr> <td>hChKey:</td> <td>XLVFNQVPLJNP-UHFFFAQYSA-N</td> </tr> <tr> <td>compuesto de:</td> <td>oxígeno</td> </tr> <tr> <td>caracterizado por:</td> <td>viscosidad, Anfitero, Q3007162, Q30100968</td> </tr> <tr> <td>código NFPA de reactividad:</td> <td>0</td> </tr> <tr> <td>Identificador de tema en TED:</td> <td>water</td> </tr> <tr> <td>Identificador PSH:</td> <td>12022</td> </tr> <tr> <td>Identificador CHEBI:</td> <td>15377</td> </tr> <tr> <td>punto de fase:</td> <td>punto crítico</td> </tr> <tr> <td>viscosidad dinámica:</td> <td>0.00101</td> </tr> </tbody> </table>		agua	Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno	color:	marón oscuro, rojo	Identificador Enciclopedia Británica Online:	science/water	densidad absoluta:	0.999802, 0.998117, 0.99403, 0.96862	Fórmula química:	H ₂ O	código NFPA de fuego:	0	Identificador KEGG:	C00001	Identificador BICF:	0045	código NFPA de salud:	0	base conjugada:	hidróxido	Identificador de tema en Quora:	Water	hChKey:	XLVFNQVPLJNP-UHFFFAQYSA-N	compuesto de:	oxígeno	caracterizado por:	viscosidad, Anfitero, Q3007162, Q30100968	código NFPA de reactividad:	0	Identificador de tema en TED:	water	Identificador PSH:	12022	Identificador CHEBI:	15377	punto de fase:	punto crítico	viscosidad dinámica:	0.00101	<table border="1"> <thead> <tr> <th></th> <th>agua</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno</td> </tr> <tr> <td>Instancia de:</td> <td>compuesto químico, medicamento, bebida, óxido, disolvente, ingrediente culinario</td> </tr> <tr> <td>categoría en Commons:</td> <td>Water</td> </tr> <tr> <td>compuesto de:</td> <td>hidrógeno, oxígeno</td> </tr> <tr> <td>hallado en el taxón:</td> <td>Homo sapiens</td> </tr> <tr> <td>Identificador Freebase:</td> <td>/m/0638F</td> </tr> <tr> <td>Identificador QND:</td> <td>4064689-0</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría:Agua</td> </tr> <tr> <td>masa:</td> <td>18.015288</td> </tr> <tr> <td>Identificador de tema en Quora:</td> <td>Water</td> </tr> <tr> <td>SMILES canónico:</td> <td>O</td> </tr> <tr> <td>hChKey:</td> <td>XLVFNQVPLJNP-UHFFFAQYSA-N</td> </tr> <tr> <td>Fórmula química:</td> <td>H₂O</td> </tr> <tr> <td>hChE:</td> <td>15HCOA1HQ</td> </tr> <tr> <td>Identificador PubChem:</td> <td>962</td> </tr> <tr> <td>color:</td> <td>verde, rojo, turquesa, marrón oscuro</td> </tr> <tr> <td>Identificador ChemSpider:</td> <td>937</td> </tr> </tbody> </table>		agua	Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno	Instancia de:	compuesto químico, medicamento, bebida, óxido, disolvente, ingrediente culinario	categoría en Commons:	Water	compuesto de:	hidrógeno, oxígeno	hallado en el taxón:	Homo sapiens	Identificador Freebase:	/m/0638F	Identificador QND:	4064689-0	categoría principal del tema:	Categoría:Agua	masa:	18.015288	Identificador de tema en Quora:	Water	SMILES canónico:	O	hChKey:	XLVFNQVPLJNP-UHFFFAQYSA-N	Fórmula química:	H ₂ O	hChE:	15HCOA1HQ	Identificador PubChem:	962	color:	verde, rojo, turquesa, marrón oscuro	Identificador ChemSpider:	937	<table border="1"> <thead> <tr> <th></th> <th>agua</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia:Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>Instancia de:</td> <td>compuesto químico, medicamento, ingrediente culinario, disolvente</td> </tr> <tr> <td>hallado en el taxón:</td> <td>Homo sapiens</td> </tr> <tr> <td>caracterizado por:</td> <td>color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968</td> </tr> <tr> <td>compuesto de:</td> <td>hidrógeno, oxígeno</td> </tr> <tr> <td>color:</td> <td>rojo, verde, turquesa, marrón oscuro</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría:Agua</td> </tr> <tr> <td>caracterizado por:</td> <td>color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q11286991</td> </tr> </tbody> </table>		agua	Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente	hallado en el taxón:	Homo sapiens	caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968	compuesto de:	hidrógeno, oxígeno	color:	rojo, verde, turquesa, marrón oscuro	categoría principal del tema:	Categoría:Agua	caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968	duplicado permanente:	Q11286991	<table border="1"> <thead> <tr> <th></th> <th>agua</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno</td> </tr> <tr> <td>Instancia de:</td> <td>compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido</td> </tr> <tr> <td>hallado en el taxón:</td> <td>Homo sapiens</td> </tr> <tr> <td>compuesto de:</td> <td>hidrógeno, oxígeno</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia:Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>color:</td> <td>rojo, verde, turquesa, marrón oscuro</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría:Agua</td> </tr> <tr> <td>caracterizado por:</td> <td>color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q11286991</td> </tr> </tbody> </table>		agua	Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno	Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido	hallado en el taxón:	Homo sapiens	compuesto de:	hidrógeno, oxígeno	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	color:	rojo, verde, turquesa, marrón oscuro	categoría principal del tema:	Categoría:Agua	caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968	duplicado permanente:	Q11286991	<table border="1"> <thead> <tr> <th></th> <th>agua</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia:Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>Instancia de:</td> <td>compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido</td> </tr> <tr> <td>hallado en el taxón:</td> <td>Homo sapiens</td> </tr> <tr> <td>caracterizado por:</td> <td>color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica</td> </tr> <tr> <td>compuesto de:</td> <td>hidrógeno, oxígeno</td> </tr> <tr> <td>categoría en Commons:</td> <td>Water</td> </tr> <tr> <td>color:</td> <td>rojo, verde, turquesa, marrón oscuro</td> </tr> <tr> <td>Identificador Freebase:</td> <td>/m/0638F</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría:Agua</td> </tr> <tr> <td>Identificador QND:</td> <td>4064689-0</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q11286991</td> </tr> </tbody> </table>		agua	Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno	catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido	hallado en el taxón:	Homo sapiens	caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica	compuesto de:	hidrógeno, oxígeno	categoría en Commons:	Water	color:	rojo, verde, turquesa, marrón oscuro	Identificador Freebase:	/m/0638F	categoría principal del tema:	Categoría:Agua	Identificador QND:	4064689-0	duplicado permanente:	Q11286991
	agua																																																																																																																																																					
Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno																																																																																																																																																					
color:	marón oscuro, rojo																																																																																																																																																					
Identificador Enciclopedia Británica Online:	science/water																																																																																																																																																					
densidad absoluta:	0.999802, 0.998117, 0.99403, 0.96862																																																																																																																																																					
Fórmula química:	H ₂ O																																																																																																																																																					
código NFPA de fuego:	0																																																																																																																																																					
Identificador KEGG:	C00001																																																																																																																																																					
Identificador BICF:	0045																																																																																																																																																					
código NFPA de salud:	0																																																																																																																																																					
base conjugada:	hidróxido																																																																																																																																																					
Identificador de tema en Quora:	Water																																																																																																																																																					
hChKey:	XLVFNQVPLJNP-UHFFFAQYSA-N																																																																																																																																																					
compuesto de:	oxígeno																																																																																																																																																					
caracterizado por:	viscosidad, Anfitero, Q3007162, Q30100968																																																																																																																																																					
código NFPA de reactividad:	0																																																																																																																																																					
Identificador de tema en TED:	water																																																																																																																																																					
Identificador PSH:	12022																																																																																																																																																					
Identificador CHEBI:	15377																																																																																																																																																					
punto de fase:	punto crítico																																																																																																																																																					
viscosidad dinámica:	0.00101																																																																																																																																																					
	agua																																																																																																																																																					
Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno																																																																																																																																																					
Instancia de:	compuesto químico, medicamento, bebida, óxido, disolvente, ingrediente culinario																																																																																																																																																					
categoría en Commons:	Water																																																																																																																																																					
compuesto de:	hidrógeno, oxígeno																																																																																																																																																					
hallado en el taxón:	Homo sapiens																																																																																																																																																					
Identificador Freebase:	/m/0638F																																																																																																																																																					
Identificador QND:	4064689-0																																																																																																																																																					
categoría principal del tema:	Categoría:Agua																																																																																																																																																					
masa:	18.015288																																																																																																																																																					
Identificador de tema en Quora:	Water																																																																																																																																																					
SMILES canónico:	O																																																																																																																																																					
hChKey:	XLVFNQVPLJNP-UHFFFAQYSA-N																																																																																																																																																					
Fórmula química:	H ₂ O																																																																																																																																																					
hChE:	15HCOA1HQ																																																																																																																																																					
Identificador PubChem:	962																																																																																																																																																					
color:	verde, rojo, turquesa, marrón oscuro																																																																																																																																																					
Identificador ChemSpider:	937																																																																																																																																																					
	agua																																																																																																																																																					
Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno																																																																																																																																																					
catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener																																																																																																																																																					
Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente																																																																																																																																																					
hallado en el taxón:	Homo sapiens																																																																																																																																																					
caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968																																																																																																																																																					
compuesto de:	hidrógeno, oxígeno																																																																																																																																																					
color:	rojo, verde, turquesa, marrón oscuro																																																																																																																																																					
categoría principal del tema:	Categoría:Agua																																																																																																																																																					
caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968																																																																																																																																																					
duplicado permanente:	Q11286991																																																																																																																																																					
	agua																																																																																																																																																					
Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno																																																																																																																																																					
Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido																																																																																																																																																					
hallado en el taxón:	Homo sapiens																																																																																																																																																					
compuesto de:	hidrógeno, oxígeno																																																																																																																																																					
catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener																																																																																																																																																					
color:	rojo, verde, turquesa, marrón oscuro																																																																																																																																																					
categoría principal del tema:	Categoría:Agua																																																																																																																																																					
caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica, Anfitero, Q3007162, Q30100968																																																																																																																																																					
duplicado permanente:	Q11286991																																																																																																																																																					
	agua																																																																																																																																																					
Descripción:	sustancia cuya molécula está compuesta por dos átomos de hidrógeno y uno de oxígeno																																																																																																																																																					
catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener																																																																																																																																																					
Instancia de:	compuesto químico, medicamento, ingrediente culinario, disolvente, bebida, óxido																																																																																																																																																					
hallado en el taxón:	Homo sapiens																																																																																																																																																					
caracterizado por:	color, Turbidez, dureza del agua, viscosidad, salinidad, capacidad calorífica																																																																																																																																																					
compuesto de:	hidrógeno, oxígeno																																																																																																																																																					
categoría en Commons:	Water																																																																																																																																																					
color:	rojo, verde, turquesa, marrón oscuro																																																																																																																																																					
Identificador Freebase:	/m/0638F																																																																																																																																																					
categoría principal del tema:	Categoría:Agua																																																																																																																																																					
Identificador QND:	4064689-0																																																																																																																																																					
duplicado permanente:	Q11286991																																																																																																																																																					

Figura A.7: Infoboxes automáticas para Agua

	1	2	3	4	5
Description:	mamífero carnívoro doméstico de la familia de los cánidos - lobos	mamífero carnívoro doméstico de la familia de los cánidos - lobos	mamífero carnívoro doméstico de la familia de los cánidos - lobos	mamífero carnívoro doméstico de la familia de los cánidos - lobos	mamífero carnívoro doméstico de la familia de los cánidos - lobos
catálogo:	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Wikipedia:Lista de artículos que toda Wikipedia debería tener	Wikipedia:Lista de artículos que toda Wikipedia debería tener
descrito en:	Enciclopedia de Olo				
instancia de:	nombre común				
subclase de:	animal doméstico, animal de compañía				
santo patrón:	Vito de Lucania, Roque de Montpellier, Huberto de Lejía	Vito de Lucania, Roque de Montpellier, Huberto de Lejía	Vito de Lucania, Roque de Montpellier, Huberto de Lejía	Vito de Lucania, Roque de Montpellier, Huberto de Lejía	Vito de Lucania, Roque de Montpellier, Huberto de Lejía
uso:	caza, animal de compañía, guarda, animal de servicio	caza, animal de compañía, guarda, animal de servicio	caza, animal de compañía, guarda, animal de servicio	caza, animal de compañía, guarda, animal de servicio	caza, animal de compañía, guarda, animal de servicio
categoría principal del tema:	Categoría:Perros	Categoría:Perros	Categoría:Perros	Categoría:Perros	Categoría:Perros
duplicado permanente:	Q21296839	Q21296839	Q21296839	Q21296839	Q21296839
nombre común:	Ἰκάλιας, Πέος, Πέος, Ἰκάλια, Dog, Köpek, 𐌱, Perri, Hond, Koer, 𐌱, 𐌱, Can, Can, Hando, Kyne, Hund, Hund, Hund	Ἰκάλιας, Πέος, Πέος, Ἰκάλια, Dog, Köpek, 𐌱, Perri, Hond, Koer, 𐌱, 𐌱, Can, Can, Hando, Kyne, Hund, Hund, Hund	Ἰκάλιας, Πέος, Πέος, Ἰκάλια, Dog, Köpek, 𐌱, Perri, Hond, Koer, 𐌱, 𐌱, Can, Can, Hando, Kyne, Hund, Hund, Hund	Ἰκάλιας, Πέος, Πέος, Ἰκάλια, Dog, Köpek, 𐌱, Perri, Hond, Koer, 𐌱, 𐌱, Can, Can, Hando, Kyne, Hund, Hund, Hund	Ἰκάλιας, Πέος, Πέος, Ἰκάλια, Dog, Köpek, 𐌱, Perri, Hond, Koer, 𐌱, 𐌱, Can, Can, Hando, Kyne, Hund, Hund, Hund
Descripción:					
catálogo:					
descrito en:					
instancia de:					
subclase de:					
categoría en Commons:					
Identificador Freebase:					
Identificador GND:					
nombre común:					
uso:					
categoría principal del tema:					
duplicado permanente:					
nombre común:					
Descripción:					
catálogo:					
descrito en:					
instancia de:					
subclase de:					
categoría en Commons:					
Identificador Freebase:					
Identificador GND:					
categoría principal del tema:					
nombre común:					
Descripción:					
catálogo:					
descrito en:					
instancia de:					
subclase de:					
categoría en Commons:					
Identificador Freebase:					
Identificador GND:					
categoría principal del tema:					
nombre común:					

Figura A.8: Infoboxes automáticas para Perro

1		2		3		4		5	
Description:	especie de mamífero semiacuático	Description:	especie de mamífero semiacuático	Description:	especie de mamífero semiacuático	Description:	especie de mamífero semiacuático	Description:	especie de mamífero semiacuático
instancia de:	taxón	instancia de:	taxón	instancia de:	taxón	instancia de:	taxón	instancia de:	taxón
taxón superior inmediato:	Ornithorhynchus	taxón superior inmediato:	especie	taxón superior inmediato:	especie	taxón superior inmediato:	especie	taxón superior inmediato:	especie
nombre del taxón:	Ornithorhynchus anatinus	taxón superior inmediato:	Australia	estado de conservación IUCN:	especie casi amenazada	estado de conservación IUCN:	especie casi amenazada	estado de conservación IUCN:	especie casi amenazada
categoría taxonómica:	especie	taxón superior inmediato:	Ornithorhynchus	taxón superior inmediato:	Ornithorhynchus	taxón superior inmediato:	Ornithorhynchus	taxón superior inmediato:	Ornithorhynchus anatinus
categoría en Commons:	Ornithorhynchus anatinus	nombre del taxón:	Ornithorhynchus anatinus	mapa de distribución de un taxón:	http://commons.wikimedia.org/wiki/Special:FilePath:Distribution_of_Ornithorhynchus_anatinus.jpg	mapa de distribución de un taxón:	http://commons.wikimedia.org/wiki/Special:FilePath:Distribution_of_Ornithorhynchus_anatinus.jpg	mapa de distribución de un taxón:	http://commons.wikimedia.org/wiki/Special:FilePath:Distribution_of_Ornithorhynchus_anatinus.jpg
identificador GBIF:	2433376	categoría en Commons:	Ornithorhynchus anatinus	nombre del taxón:	Ornithorhynchus anatinus	nombre del taxón:	Ornithorhynchus anatinus	nombre del taxón:	Ornithorhynchus anatinus
identificador EOL:	323858	identificador GBIF:	2433376	categoría en Commons:	Ornithorhynchus anatinus	categoría en Commons:	Ornithorhynchus anatinus	categoría en Commons:	Ornithorhynchus anatinus
identificador Freebase:	/m/05syb	identificador EOL:	323858	identificador UICN:	40488	identificador UICN:	40488	identificador UICN:	40488
identificador ITIS:	552524	identificador Freebase:	/m/05syb	identificador Fossilworks:	m/05syb	identificador Fossilworks:	m/05syb	identificador Fossilworks:	m/05syb
nombre común:	カモノハシ, Platypus, 鴨嘴獸, Утконос, ⵏⴰⵏⴰⵏ ⵏ ⵏⴰⵏⴰⵏ, Platynonca, Nattodyr, Schrabbeiler, Duck-billed platypus, Duck-billed Platypus, Ornitorinco, Nökköom, Ornitorhinus, Vesinokkakäinen, Ornithorynque, ⵏⵓⵏⵓⵏⵓⵏ	identificador ITIS:	552524	identificador ITIS:	552524	identificador ITIS:	552524	identificador ITIS:	552524
		nombre común:	カモノハシ, Platypus, 鴨嘴獸, Утконос, ⵏⴰⵏⴰⵏ ⵏ ⵏⴰⵏⴰⵏ, Platynonca, Nattodyr, Schrabbeiler, Duck-billed platypus, Duck-billed Platypus, Ornitorinco, Nökköom, Ornitorhinus, Vesinokkakäinen, Ornithorynque	identificador EOL:	323858	identificador EOL:	323858	identificador EOL:	323858
				identificador Fossilworks:	168505	identificador Fossilworks:	168505	identificador Fossilworks:	168505
				identificador GBIF:	2433376	identificador GBIF:	2433376	identificador GBIF:	2433376
				galería en Commons:	Ornithorhynchus anatinus	galería en Commons:	Ornithorhynchus anatinus	galería en Commons:	Ornithorhynchus anatinus
				identificador de especies MSW:	1030000	identificador de especies MSW:	1030000	identificador de especies MSW:	1030000
				identificador Encyclopædia Britannica Online:	animalplatypus	identificador Encyclopædia Britannica Online:	animalplatypus	identificador Encyclopædia Britannica Online:	animalplatypus
				nombre común:	カモノハシ, Platypus, 鴨嘴獸, Утконос, ⵏⴰⵏⴰⵏ ⵏ ⵏⴰⵏⴰⵏ, Platynonca, Nattodyr	nombre común:	カモノハシ, Platypus, 鴨嘴獸, Утконос, ⵏⴰⵏⴰⵏ ⵏ ⵏⴰⵏⴰⵏ, Platynonca, Nattodyr	nombre común:	カモノハシ, Platypus, 鴨嘴獸, Утконос, ⵏⴰⵏⴰⵏ ⵏ ⵏⴰⵏⴰⵏ, Platynonca, Nattodyr

Figura A.9: Infoboxes automáticas para Ornitorrinco

1	2	3	4	5																																																																																											
<table border="1"> <thead> <tr> <th colspan="2">Biblia</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento</td> </tr> <tr> <td>instancia de:</td> <td>libro, obra literaria, libro sagrado</td> </tr> <tr> <td>descrito en:</td> <td>Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963</td> </tr> <tr> <td>compuesto de:</td> <td>Nuevo Testamento, Antiguo Testamento</td> </tr> <tr> <td>idioma de la obra o del nombre:</td> <td>koíné, hebreo bíblico, arameo</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría Biblia</td> </tr> <tr> <td>tiene partes de la clase:</td> <td>libro de la Biblia</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q2425089</td> </tr> <tr> <td>edición:</td> <td>Septuaginta, Q11145376, Biblia de Lutero, Q5083751</td> </tr> </tbody> </table>	Biblia		Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento	instancia de:	libro, obra literaria, libro sagrado	descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963	compuesto de:	Nuevo Testamento, Antiguo Testamento	idioma de la obra o del nombre:	koíné, hebreo bíblico, arameo	categoría principal del tema:	Categoría Biblia	tiene partes de la clase:	libro de la Biblia	duplicado permanente:	Q2425089	edición:	Septuaginta, Q11145376, Biblia de Lutero, Q5083751	<table border="1"> <thead> <tr> <th colspan="2">Biblia</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento</td> </tr> <tr> <td>descrito en:</td> <td>Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963</td> </tr> <tr> <td>instancia de:</td> <td>libro, obra literaria, libro sagrado</td> </tr> <tr> <td>idioma de la obra o del nombre:</td> <td>koíné, hebreo bíblico, arameo</td> </tr> <tr> <td>compuesto de:</td> <td>Nuevo Testamento, Antiguo Testamento</td> </tr> <tr> <td>estudiado por:</td> <td>estudios bíblicos</td> </tr> <tr> <td>tiene partes de la clase:</td> <td>libro de la Biblia</td> </tr> <tr> <td>edición:</td> <td>Septuaginta, Q11145376, Biblia de Lutero, Q5083751, Biblia del rey Jacobo</td> </tr> </tbody> </table>	Biblia		Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento	descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963	instancia de:	libro, obra literaria, libro sagrado	idioma de la obra o del nombre:	koíné, hebreo bíblico, arameo	compuesto de:	Nuevo Testamento, Antiguo Testamento	estudiado por:	estudios bíblicos	tiene partes de la clase:	libro de la Biblia	edición:	Septuaginta, Q11145376, Biblia de Lutero, Q5083751, Biblia del rey Jacobo	<table border="1"> <thead> <tr> <th colspan="2">Biblia</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento</td> </tr> <tr> <td>instancia de:</td> <td>libro, obra literaria, libro sagrado</td> </tr> <tr> <td>descrito en:</td> <td>Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963</td> </tr> <tr> <td>categoría en Commons:</td> <td>Bible</td> </tr> <tr> <td>compuesto de:</td> <td>Nuevo Testamento, Antiguo Testamento</td> </tr> <tr> <td>identificador Freebase:</td> <td>/m/013j7</td> </tr> <tr> <td>identificador VIAF:</td> <td>174425434</td> </tr> <tr> <td>identificador GND:</td> <td>4006406-2</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Categoría Biblia</td> </tr> <tr> <td>identificador BnF:</td> <td>12008248x</td> </tr> <tr> <td>idioma de la obra o del nombre:</td> <td>araméico, koíné, hebreo bíblico</td> </tr> <tr> <td>SUDOC:</td> <td>028186054</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia Ortodoxa Teológica, Vlastenský slovník historický, Q20089963, Enciclopedia Católica, Diccionario Enciclopédico Brockhaus y Efron, Encyklopédia Britannica (1911), Q1910746, Q4086271, Q4173137, Q19180675</td> </tr> </tbody> </table>	Biblia		Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento	instancia de:	libro, obra literaria, libro sagrado	descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963	categoría en Commons:	Bible	compuesto de:	Nuevo Testamento, Antiguo Testamento	identificador Freebase:	/m/013j7	identificador VIAF:	174425434	identificador GND:	4006406-2	categoría principal del tema:	Categoría Biblia	identificador BnF:	12008248x	idioma de la obra o del nombre:	araméico, koíné, hebreo bíblico	SUDOC:	028186054	descrito en:	Enciclopedia Ortodoxa Teológica, Vlastenský slovník historický, Q20089963, Enciclopedia Católica, Diccionario Enciclopédico Brockhaus y Efron, Encyklopédia Britannica (1911), Q1910746, Q4086271, Q4173137, Q19180675	<table border="1"> <thead> <tr> <th colspan="2">Biblia</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento</td> </tr> <tr> <td>edición:</td> <td>Q9170121, Biblia Abama, Q1240868, Q20160290, Q20157294, Q9170725, Q9170718, Q10931146, Q4294434, Q1786906, Edición de la Biblia, Biblia de Bresl, Biblia de Lutero</td> </tr> <tr> <td>Clasificación Decimal de Dewey:</td> <td>220</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia Ortodoxa Teológica, Q4086271, Enciclopedia Católica</td> </tr> <tr> <td>portal principal del tema:</td> <td>Q5088115</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q2425089</td> </tr> <tr> <td>instancia de:</td> <td>libro</td> </tr> <tr> <td>identificador PSH:</td> <td>7794</td> </tr> <tr> <td>galera en Commons:</td> <td>Bible</td> </tr> <tr> <td>SUDOC:</td> <td>028186054</td> </tr> <tr> <td>identificador de la Gran Enciclopedia Catalana:</td> <td>0009910</td> </tr> <tr> <td>identificador NLA:</td> <td>10638269</td> </tr> </tbody> </table>	Biblia		Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento	edición:	Q9170121, Biblia Abama, Q1240868, Q20160290, Q20157294, Q9170725, Q9170718, Q10931146, Q4294434, Q1786906, Edición de la Biblia, Biblia de Bresl, Biblia de Lutero	Clasificación Decimal de Dewey:	220	descrito en:	Enciclopedia Ortodoxa Teológica, Q4086271, Enciclopedia Católica	portal principal del tema:	Q5088115	duplicado permanente:	Q2425089	instancia de:	libro	identificador PSH:	7794	galera en Commons:	Bible	SUDOC:	028186054	identificador de la Gran Enciclopedia Catalana:	0009910	identificador NLA:	10638269
Biblia																																																																																															
Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento																																																																																														
instancia de:	libro, obra literaria, libro sagrado																																																																																														
descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963																																																																																														
compuesto de:	Nuevo Testamento, Antiguo Testamento																																																																																														
idioma de la obra o del nombre:	koíné, hebreo bíblico, arameo																																																																																														
categoría principal del tema:	Categoría Biblia																																																																																														
tiene partes de la clase:	libro de la Biblia																																																																																														
duplicado permanente:	Q2425089																																																																																														
edición:	Septuaginta, Q11145376, Biblia de Lutero, Q5083751																																																																																														
Biblia																																																																																															
Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento																																																																																														
descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963																																																																																														
instancia de:	libro, obra literaria, libro sagrado																																																																																														
idioma de la obra o del nombre:	koíné, hebreo bíblico, arameo																																																																																														
compuesto de:	Nuevo Testamento, Antiguo Testamento																																																																																														
estudiado por:	estudios bíblicos																																																																																														
tiene partes de la clase:	libro de la Biblia																																																																																														
edición:	Septuaginta, Q11145376, Biblia de Lutero, Q5083751, Biblia del rey Jacobo																																																																																														
Biblia																																																																																															
Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento																																																																																														
instancia de:	libro, obra literaria, libro sagrado																																																																																														
descrito en:	Diccionario Enciclopédico Brockhaus y Efron, Q19180675, Q4173137, Enciclopedia Católica, Q4086271, Vlastenský slovník historický, Encyklopédia Britannica (1911), Q1910746, Enciclopedia Ortodoxa Teológica, Q20089963																																																																																														
categoría en Commons:	Bible																																																																																														
compuesto de:	Nuevo Testamento, Antiguo Testamento																																																																																														
identificador Freebase:	/m/013j7																																																																																														
identificador VIAF:	174425434																																																																																														
identificador GND:	4006406-2																																																																																														
categoría principal del tema:	Categoría Biblia																																																																																														
identificador BnF:	12008248x																																																																																														
idioma de la obra o del nombre:	araméico, koíné, hebreo bíblico																																																																																														
SUDOC:	028186054																																																																																														
descrito en:	Enciclopedia Ortodoxa Teológica, Vlastenský slovník historický, Q20089963, Enciclopedia Católica, Diccionario Enciclopédico Brockhaus y Efron, Encyklopédia Britannica (1911), Q1910746, Q4086271, Q4173137, Q19180675																																																																																														
Biblia																																																																																															
Descripción:	conjunto de libros canónicos del cristianismo que abarcan el Antiguo y el Nuevo Testamento																																																																																														
edición:	Q9170121, Biblia Abama, Q1240868, Q20160290, Q20157294, Q9170725, Q9170718, Q10931146, Q4294434, Q1786906, Edición de la Biblia, Biblia de Bresl, Biblia de Lutero																																																																																														
Clasificación Decimal de Dewey:	220																																																																																														
descrito en:	Enciclopedia Ortodoxa Teológica, Q4086271, Enciclopedia Católica																																																																																														
portal principal del tema:	Q5088115																																																																																														
duplicado permanente:	Q2425089																																																																																														
instancia de:	libro																																																																																														
identificador PSH:	7794																																																																																														
galera en Commons:	Bible																																																																																														
SUDOC:	028186054																																																																																														
identificador de la Gran Enciclopedia Catalana:	0009910																																																																																														
identificador NLA:	10638269																																																																																														

Figura A.10: Infoboxes automáticas para la Biblia

1	2	3	4	5																																																																																																																																																						
<table border="1"> <thead> <tr> <th colspan="2">Doce hombres en pugna</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>película estadounidense de 1957</td> </tr> <tr> <td>país de origen:</td> <td>Estados Unidos</td> </tr> <tr> <td>idioma original:</td> <td>inglés</td> </tr> <tr> <td>instancia de:</td> <td>película</td> </tr> <tr> <td>fecha de publicación:</td> <td>1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z</td> </tr> <tr> <td>título:</td> <td>12 Angry Men</td> </tr> <tr> <td>categoría en Commons:</td> <td>12 Angry Men (1957 film)</td> </tr> <tr> <td>color:</td> <td>blanco y negro</td> </tr> <tr> <td>identificador IMDb_1:</td> <td>inv0m_1</td> </tr> <tr> <td>identificador VIAF:</td> <td>198561370</td> </tr> <tr> <td>género artístico:</td> <td>cine dramático</td> </tr> <tr> <td>miembro del reparto:</td> <td>Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec, Edward Binns, Rudy Bond</td> </tr> </tbody> </table>	Doce hombres en pugna		Descripción:	película estadounidense de 1957	país de origen:	Estados Unidos	idioma original:	inglés	instancia de:	película	fecha de publicación:	1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z	título:	12 Angry Men	categoría en Commons:	12 Angry Men (1957 film)	color:	blanco y negro	identificador IMDb_1:	inv0m_1	identificador VIAF:	198561370	género artístico:	cine dramático	miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec, Edward Binns, Rudy Bond	<table border="1"> <thead> <tr> <th colspan="2">Doce hombres en pugna</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>película estadounidense de 1957</td> </tr> <tr> <td>director de fotografía:</td> <td>Boris Kaufman</td> </tr> <tr> <td>clasificación FSK:</td> <td>FSK 12</td> </tr> <tr> <td>fecha de publicación:</td> <td>1957-09-04T00:00:00Z, 1957-04-10T00:00:00Z</td> </tr> <tr> <td>nominado a:</td> <td>Oscar a la mejor película, Globo de Oro al mejor actor - Drama, Oscar al mejor director</td> </tr> <tr> <td>identificador Cinema:</td> <td>1300240</td> </tr> <tr> <td>ISAN:</td> <td>0000-0000-0000-0000-0000-0000-0</td> </tr> <tr> <td>identificador Allmovie de película:</td> <td>v51289</td> </tr> <tr> <td>puntuación de la crítica:</td> <td>4/4</td> </tr> <tr> <td>miembro del reparto:</td> <td>Henry Fonda, Robert Webber, Jack Klugman, Jack Warden</td> </tr> <tr> <td>identificador Freebase:</td> <td>inv0m_1</td> </tr> <tr> <td>tema principal de la obra:</td> <td>juicio con jurado</td> </tr> <tr> <td>título:</td> <td>12 Angry Men</td> </tr> <tr> <td>idioma original:</td> <td>inglés</td> </tr> <tr> <td>identificador Kinopoisk de película:</td> <td>346</td> </tr> <tr> <td>premio recibido:</td> <td>Grand Prix</td> </tr> <tr> <td>identificador GND:</td> <td>4749801-8</td> </tr> <tr> <td>identificador OFDb:</td> <td>4902</td> </tr> <tr> <td>identificador PORT de película:</td> <td>2457</td> </tr> <tr> <td>certificado de distribución cinematográfica:</td> <td>19677</td> </tr> </tbody> </table>	Doce hombres en pugna		Descripción:	película estadounidense de 1957	director de fotografía:	Boris Kaufman	clasificación FSK:	FSK 12	fecha de publicación:	1957-09-04T00:00:00Z, 1957-04-10T00:00:00Z	nominado a:	Oscar a la mejor película, Globo de Oro al mejor actor - Drama, Oscar al mejor director	identificador Cinema:	1300240	ISAN:	0000-0000-0000-0000-0000-0000-0	identificador Allmovie de película:	v51289	puntuación de la crítica:	4/4	miembro del reparto:	Henry Fonda, Robert Webber, Jack Klugman, Jack Warden	identificador Freebase:	inv0m_1	tema principal de la obra:	juicio con jurado	título:	12 Angry Men	idioma original:	inglés	identificador Kinopoisk de película:	346	premio recibido:	Grand Prix	identificador GND:	4749801-8	identificador OFDb:	4902	identificador PORT de película:	2457	certificado de distribución cinematográfica:	19677	<table border="1"> <thead> <tr> <th colspan="2">Doce hombres en pugna</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>película estadounidense de 1957</td> </tr> <tr> <td>instancia de:</td> <td>película</td> </tr> <tr> <td>fecha de publicación:</td> <td>1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z</td> </tr> <tr> <td>título:</td> <td>12 Angry Men</td> </tr> <tr> <td>categoría en Commons:</td> <td>12 Angry Men (1957 film)</td> </tr> <tr> <td>identificador IMDb_1:</td> <td>inv0m_1</td> </tr> <tr> <td>identificador Freebase:</td> <td>198561370</td> </tr> <tr> <td>idioma original:</td> <td>inglés</td> </tr> <tr> <td>miembro del reparto:</td> <td>Henry Fonda, Jack Klugman, Martin Balsam, Ed Begley, Lee J. Cobb, Jack Warden, Rudy Bond, E. G. Marshall, Robert Webber, Joseph Sweeney, John Fiedler, Jif Voskovec, Edward Binns, Q4913993</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.mgm.com/tour-sites/2061012-Angry-Men-(1957)</td> </tr> </tbody> </table>	Doce hombres en pugna		Descripción:	película estadounidense de 1957	instancia de:	película	fecha de publicación:	1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z	título:	12 Angry Men	categoría en Commons:	12 Angry Men (1957 film)	identificador IMDb_1:	inv0m_1	identificador Freebase:	198561370	idioma original:	inglés	miembro del reparto:	Henry Fonda, Jack Klugman, Martin Balsam, Ed Begley, Lee J. Cobb, Jack Warden, Rudy Bond, E. G. Marshall, Robert Webber, Joseph Sweeney, John Fiedler, Jif Voskovec, Edward Binns, Q4913993	sitio web oficial:	http://www.mgm.com/tour-sites/2061012-Angry-Men-(1957)	<table border="1"> <thead> <tr> <th colspan="2">Doce hombres en pugna</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>película estadounidense de 1957</td> </tr> <tr> <td>país de origen:</td> <td>Estados Unidos</td> </tr> <tr> <td>idioma original:</td> <td>inglés</td> </tr> <tr> <td>instancia de:</td> <td>película</td> </tr> <tr> <td>color:</td> <td>blanco y negro</td> </tr> <tr> <td>lugar de la narración:</td> <td>Manhattan</td> </tr> <tr> <td>género artístico:</td> <td>cine dramático, cine judicial</td> </tr> <tr> <td>tema principal de la obra:</td> <td>pena de muerte</td> </tr> <tr> <td>nominado a:</td> <td>Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director, Globo de Oro al mejor actor - Drama, BAFTA a la mejor película, Globo de Oro al mejor actor de reparto, Anexo:Globo de Oro al mejor director</td> </tr> <tr> <td>clasificación FSK:</td> <td>FSK 12</td> </tr> <tr> <td>clasificación CNC:</td> <td>todo público</td> </tr> <tr> <td>premio recibido:</td> <td>National Film Registry, Oso de oro</td> </tr> <tr> <td>distribuidor:</td> <td>Q17126013</td> </tr> <tr> <td>productor:</td> <td>Henry Fonda</td> </tr> <tr> <td>miembro del reparto:</td> <td>Henry Fonda, Martin Balsam, Ed Begley</td> </tr> <tr> <td>director:</td> <td>Sidney Lumet</td> </tr> </tbody> </table>	Doce hombres en pugna		Descripción:	película estadounidense de 1957	país de origen:	Estados Unidos	idioma original:	inglés	instancia de:	película	color:	blanco y negro	lugar de la narración:	Manhattan	género artístico:	cine dramático, cine judicial	tema principal de la obra:	pena de muerte	nominado a:	Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director, Globo de Oro al mejor actor - Drama, BAFTA a la mejor película, Globo de Oro al mejor actor de reparto, Anexo:Globo de Oro al mejor director	clasificación FSK:	FSK 12	clasificación CNC:	todo público	premio recibido:	National Film Registry, Oso de oro	distribuidor:	Q17126013	productor:	Henry Fonda	miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley	director:	Sidney Lumet	<table border="1"> <thead> <tr> <th colspan="2">Doce hombres en pugna</th> </tr> </thead> <tbody> <tr> <td>Descripción:</td> <td>película estadounidense de 1957</td> </tr> <tr> <td>instancia de:</td> <td>película</td> </tr> <tr> <td>país de origen:</td> <td>Estados Unidos</td> </tr> <tr> <td>idioma original:</td> <td>inglés</td> </tr> <tr> <td>género artístico:</td> <td>cine dramático, cine judicial, Q11209409</td> </tr> <tr> <td>color:</td> <td>blanco y negro</td> </tr> <tr> <td>tema principal de la obra:</td> <td>pena de muerte</td> </tr> <tr> <td>lugar de la narración:</td> <td>Manhattan</td> </tr> <tr> <td>premio recibido:</td> <td>National Film Registry, Oso de oro</td> </tr> <tr> <td>nominado a:</td> <td>Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director</td> </tr> <tr> <td>miembro del reparto:</td> <td>Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec</td> </tr> <tr> <td>clasificación FSK:</td> <td>FSK 12</td> </tr> </tbody> </table>	Doce hombres en pugna		Descripción:	película estadounidense de 1957	instancia de:	película	país de origen:	Estados Unidos	idioma original:	inglés	género artístico:	cine dramático, cine judicial, Q11209409	color:	blanco y negro	tema principal de la obra:	pena de muerte	lugar de la narración:	Manhattan	premio recibido:	National Film Registry, Oso de oro	nominado a:	Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director	miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec	clasificación FSK:	FSK 12
Doce hombres en pugna																																																																																																																																																										
Descripción:	película estadounidense de 1957																																																																																																																																																									
país de origen:	Estados Unidos																																																																																																																																																									
idioma original:	inglés																																																																																																																																																									
instancia de:	película																																																																																																																																																									
fecha de publicación:	1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z																																																																																																																																																									
título:	12 Angry Men																																																																																																																																																									
categoría en Commons:	12 Angry Men (1957 film)																																																																																																																																																									
color:	blanco y negro																																																																																																																																																									
identificador IMDb_1:	inv0m_1																																																																																																																																																									
identificador VIAF:	198561370																																																																																																																																																									
género artístico:	cine dramático																																																																																																																																																									
miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec, Edward Binns, Rudy Bond																																																																																																																																																									
Doce hombres en pugna																																																																																																																																																										
Descripción:	película estadounidense de 1957																																																																																																																																																									
director de fotografía:	Boris Kaufman																																																																																																																																																									
clasificación FSK:	FSK 12																																																																																																																																																									
fecha de publicación:	1957-09-04T00:00:00Z, 1957-04-10T00:00:00Z																																																																																																																																																									
nominado a:	Oscar a la mejor película, Globo de Oro al mejor actor - Drama, Oscar al mejor director																																																																																																																																																									
identificador Cinema:	1300240																																																																																																																																																									
ISAN:	0000-0000-0000-0000-0000-0000-0																																																																																																																																																									
identificador Allmovie de película:	v51289																																																																																																																																																									
puntuación de la crítica:	4/4																																																																																																																																																									
miembro del reparto:	Henry Fonda, Robert Webber, Jack Klugman, Jack Warden																																																																																																																																																									
identificador Freebase:	inv0m_1																																																																																																																																																									
tema principal de la obra:	juicio con jurado																																																																																																																																																									
título:	12 Angry Men																																																																																																																																																									
idioma original:	inglés																																																																																																																																																									
identificador Kinopoisk de película:	346																																																																																																																																																									
premio recibido:	Grand Prix																																																																																																																																																									
identificador GND:	4749801-8																																																																																																																																																									
identificador OFDb:	4902																																																																																																																																																									
identificador PORT de película:	2457																																																																																																																																																									
certificado de distribución cinematográfica:	19677																																																																																																																																																									
Doce hombres en pugna																																																																																																																																																										
Descripción:	película estadounidense de 1957																																																																																																																																																									
instancia de:	película																																																																																																																																																									
fecha de publicación:	1957-04-10T00:00:00Z, 1957-07-29T00:00:00Z, 1957-09-04T00:00:00Z, 1958-10-10T00:00:00Z																																																																																																																																																									
título:	12 Angry Men																																																																																																																																																									
categoría en Commons:	12 Angry Men (1957 film)																																																																																																																																																									
identificador IMDb_1:	inv0m_1																																																																																																																																																									
identificador Freebase:	198561370																																																																																																																																																									
idioma original:	inglés																																																																																																																																																									
miembro del reparto:	Henry Fonda, Jack Klugman, Martin Balsam, Ed Begley, Lee J. Cobb, Jack Warden, Rudy Bond, E. G. Marshall, Robert Webber, Joseph Sweeney, John Fiedler, Jif Voskovec, Edward Binns, Q4913993																																																																																																																																																									
sitio web oficial:	http://www.mgm.com/tour-sites/2061012-Angry-Men-(1957)																																																																																																																																																									
Doce hombres en pugna																																																																																																																																																										
Descripción:	película estadounidense de 1957																																																																																																																																																									
país de origen:	Estados Unidos																																																																																																																																																									
idioma original:	inglés																																																																																																																																																									
instancia de:	película																																																																																																																																																									
color:	blanco y negro																																																																																																																																																									
lugar de la narración:	Manhattan																																																																																																																																																									
género artístico:	cine dramático, cine judicial																																																																																																																																																									
tema principal de la obra:	pena de muerte																																																																																																																																																									
nominado a:	Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director, Globo de Oro al mejor actor - Drama, BAFTA a la mejor película, Globo de Oro al mejor actor de reparto, Anexo:Globo de Oro al mejor director																																																																																																																																																									
clasificación FSK:	FSK 12																																																																																																																																																									
clasificación CNC:	todo público																																																																																																																																																									
premio recibido:	National Film Registry, Oso de oro																																																																																																																																																									
distribuidor:	Q17126013																																																																																																																																																									
productor:	Henry Fonda																																																																																																																																																									
miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley																																																																																																																																																									
director:	Sidney Lumet																																																																																																																																																									
Doce hombres en pugna																																																																																																																																																										
Descripción:	película estadounidense de 1957																																																																																																																																																									
instancia de:	película																																																																																																																																																									
país de origen:	Estados Unidos																																																																																																																																																									
idioma original:	inglés																																																																																																																																																									
género artístico:	cine dramático, cine judicial, Q11209409																																																																																																																																																									
color:	blanco y negro																																																																																																																																																									
tema principal de la obra:	pena de muerte																																																																																																																																																									
lugar de la narración:	Manhattan																																																																																																																																																									
premio recibido:	National Film Registry, Oso de oro																																																																																																																																																									
nominado a:	Oscar a la mejor película, Oscar al mejor guion adaptado, Oscar al mejor director																																																																																																																																																									
miembro del reparto:	Henry Fonda, Martin Balsam, Ed Begley, Jack Warden, Lee J. Cobb, John Fiedler, Jack Klugman, Robert Webber, E. G. Marshall, Jif Voskovec																																																																																																																																																									
clasificación FSK:	FSK 12																																																																																																																																																									

Figura A.11: Infoboxes automáticas para 12 Angry Men

1

Marte
Descripción: Cuarto planeta del Sistema Solar
portal principal del tema: Q14615881
identificador Wikcionario de tema: scien:elmas
punto más alto: Monte Olimpo
radio: 3389.5
precedido por: Tierra
descrito en: Diccionario Enciclopédico Brockhaus y Efron, Q19180675
semble mayor: L32979, 227936637
forma parte de: sistema solar interior
identificador RST: 492
masa: 641.71
Identificador autoridades de la Biblioteca del Congreso de EEUU: in05081548
identificador NDL: 000711940
instancia de: planeta interior
longitud del nodo ascendente: 49.5936391
vídeo: http://commons.wikimedia.org/wiki/Special:FilePath:Mars.jpg
fecha de fundación o creación: -454000000-01-01T00:00:00Z
categoría principal del tema: Categoría:Marte
sucedido por: Júpiter
identificador NAF: 499
pronunciación (archivo de sonido): http://commons.wikimedia.org/wiki/Special:FilePath:En-us-Mars.ogg
especificidad subclase: 0.093941
categoría en Commons: Mars (planet)
identificador BNC Thesgl: 0a204f73-8b34-465a-8b76-2b7a4125a51

2

Marte
Descripción: Cuarto planeta del Sistema Solar
instancia de: planeta interior
categoría en Commons: Mars (planet)
compuesto de: Atmósfera de Marte
Identificador Freebase: /m/09kvs
Identificador VIAF: 316741886
fecha de fundación o creación: -454000000-01-01T00:00:00Z
Identificador GND: 4037687-4
forma parte de: sistema solar, sistema solar interior
categoría principal del tema: Categoría:Marte
Identificador autoridades de la Biblioteca del Congreso de EEUU: in05081548
Identificador BnF: 120752119
SUDOC: 029643387
precedido por: Tierra
sucedido por: Júpiter
masa: 641.71
descrito en: Diccionario Enciclopédico Brockhaus y Efron, Q19180675
identificador de tema en Quora: Mars-planet
nombrado por: Marte
galera en Commons: Mars
Identificador NDL: 00565094
Identificador NLE: 000711940
Identificador JSTOR de tema: mars
Identificador TGN: 7030644

3

Marte
Descripción: Cuarto planeta del Sistema Solar
catálogo: Wikipedia:Lista de artículos que toda Wikipedia debería tener
descrito en: Diccionario Enciclopédico Brockhaus y Efron, Q19180675
precedido por: Tierra
forma parte de: sistema solar
objeto astronómico padre: Sol
sucedido por: Júpiter
instancia de: planeta interior
nombrado por: Marte
portal principal del tema: Q14615881
objeto astronómico hijo: Fobos, Deimos
compuesto de: Atmósfera de Marte
categoría principal del tema: Categoría:Marte
punto más alto: Monte Olimpo
notación: Simbolo de Marte
duplicado permanente: Q11201870
subredir: Mars
vídeo: http://commons.wikimedia.org/wiki/Special:FilePath:Mars.jpg
Identificador de tema en Quora: Mars-planet
Identificador JSTOR de tema: mars
Identificador VIAF: 316741886
Identificador GND: 4037687-4
Identificador autoridades de la Biblioteca del Congreso de EEUU: in05081548
Identificador BnF: 120752119

4

Marte
Descripción: Cuarto planeta del Sistema Solar
descrito en: Diccionario Enciclopédico Brockhaus y Efron, Q19180675
instancia de: planeta interior
precedido por: Tierra
forma parte de: sistema solar
catálogo: Wikipedia:Lista de artículos que toda Wikipedia debería tener
sucedido por: Júpiter
compuesto de: Atmósfera de Marte
categoría principal del tema: Categoría:Marte
nombrado por: Marte
objeto astronómico padre: Sol
duplicado permanente: Q11201870
punto más alto: Monte Olimpo
objeto astronómico hijo: Fobos, Deimos
portal principal del tema: Q14615881
notación: Simbolo de Marte
subredir: Mars
vídeo: http://commons.wikimedia.org/wiki/Special:FilePath:Mars.jpg
Identificador de tema en Quora: Mars-planet
Identificador JSTOR de tema: mars
Identificador VIAF: 316741886
Identificador GND: 4037687-4
Identificador autoridades de la Biblioteca del Congreso de EEUU: in05081548
Identificador BnF: 120752119

5

Marte
Descripción: Cuarto planeta del Sistema Solar
catálogo: Wikipedia:Lista de artículos que toda Wikipedia debería tener
instancia de: planeta interior
descrito en: Diccionario Enciclopédico Brockhaus y Efron, Q19180675
precedido por: Tierra
categoría en Commons: Mars (planet)
forma parte de: sistema solar, sistema solar interior
compuesto de: Atmósfera de Marte
Identificador Freebase: /m/09kvs
objeto astronómico padre: Sol
Identificador VIAF: 316741886
fecha de fundación o creación: -454000000-01-01T00:00:00Z
Identificador GND: 4037687-4
categoría principal del tema: Categoría:Marte
Identificador autoridades de la Biblioteca del Congreso de EEUU: in05081548
sucedido por: Júpiter
Identificador BnF: 120752119
SUDOC: 029643387
masa: 641.71
Identificador de tema en Quora: Mars-planet
nombrado por: Marte
galera en Commons: Mars
Identificador NDL: 00565094
Identificador NLE: 000711940

Figura A.12: Infoboxes automáticas para Marte



Figura A.13: Infoboxes automáticas para Betelgeuse

1	2	3	4	5																																																																																																																																																																																																																																												
<table border="1"> <thead> <tr> <th colspan="2">Torre Eiffel</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>monumento en París, Francia</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia de Olo</td> </tr> <tr> <td>país:</td> <td>Francia</td> </tr> <tr> <td>propietario:</td> <td>París</td> </tr> <tr> <td>estatus patrimonial:</td> <td>monumento histórico inscrito</td> </tr> <tr> <td>huso horario:</td> <td>Hora central europea</td> </tr> <tr> <td>situado en la entidad territorial administrativa:</td> <td>VII Distrito de París</td> </tr> <tr> <td>instancia de:</td> <td>destino turístico, torre de observación, punto de referencia, Q1440476</td> </tr> <tr> <td>forma parte de:</td> <td>Riberas del Sena en París</td> </tr> <tr> <td>nombrado por:</td> <td>Gustave Eiffel</td> </tr> <tr> <td>constructor:</td> <td>Gustave Eiffel</td> </tr> <tr> <td>materias empleadas:</td> <td>hierro forjado</td> </tr> <tr> <td>ingeniero estructural:</td> <td>Émile Nouguier, Maurice Koechlin</td> </tr> <tr> <td>calle:</td> <td>Campo de Marte</td> </tr> <tr> <td>arquitecto:</td> <td>Stephen Sauvestre</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Q6741108</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q25689717</td> </tr> <tr> <td>Identificador VIAF:</td> <td>236835198</td> </tr> <tr> <td>Identificador GND:</td> <td>4102872-7</td> </tr> <tr> <td>Identificador autoridades de la Biblioteca del Congreso de EEUU:</td> <td>sh85136247</td> </tr> <tr> <td>Identificador BnF:</td> <td>11941003z</td> </tr> </tbody> </table>	Torre Eiffel		Description:	monumento en París, Francia	catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener	descrito en:	Enciclopedia de Olo	país:	Francia	propietario:	París	estatus patrimonial:	monumento histórico inscrito	huso horario:	Hora central europea	situado en la entidad territorial administrativa:	VII Distrito de París	instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476	forma parte de:	Riberas del Sena en París	nombrado por:	Gustave Eiffel	constructor:	Gustave Eiffel	materias empleadas:	hierro forjado	ingeniero estructural:	Émile Nouguier, Maurice Koechlin	calle:	Campo de Marte	arquitecto:	Stephen Sauvestre	categoría principal del tema:	Q6741108	duplicado permanente:	Q25689717	Identificador VIAF:	236835198	Identificador GND:	4102872-7	Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247	Identificador BnF:	11941003z	<table border="1"> <thead> <tr> <th colspan="2">Torre Eiffel</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>monumento en París, Francia</td> </tr> <tr> <td>país:</td> <td>Francia</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia de Olo</td> </tr> <tr> <td>estatus patrimonial:</td> <td>monumento histórico inscrito</td> </tr> <tr> <td>huso horario:</td> <td>Hora central europea</td> </tr> <tr> <td>instancia de:</td> <td>destino turístico, torre de observación, punto de referencia, Q1440476</td> </tr> <tr> <td>situado en la entidad territorial administrativa:</td> <td>VII Distrito de París</td> </tr> <tr> <td>propietario:</td> <td>París</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>forma parte de:</td> <td>Riberas del Sena en París</td> </tr> <tr> <td>materias empleadas:</td> <td>hierro forjado</td> </tr> <tr> <td>nombrado por:</td> <td>Gustave Eiffel</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Q6741108</td> </tr> <tr> <td>calle:</td> <td>Campo de Marte</td> </tr> <tr> <td>arquitecto:</td> <td>Stephen Sauvestre</td> </tr> <tr> <td>duplicado permanente:</td> <td>Q25689717</td> </tr> <tr> <td>constructor:</td> <td>Gustave Eiffel</td> </tr> <tr> <td>ingeniero estructural:</td> <td>Émile Nouguier, Maurice Koechlin</td> </tr> <tr> <td>Identificador VIAF:</td> <td>236835198</td> </tr> <tr> <td>Identificador GND:</td> <td>4102872-7</td> </tr> <tr> <td>Identificador autoridades de la Biblioteca del Congreso de EEUU:</td> <td>sh85136247</td> </tr> <tr> <td>Identificador BnF:</td> <td>11941003z</td> </tr> </tbody> </table>	Torre Eiffel		Description:	monumento en París, Francia	país:	Francia	descrito en:	Enciclopedia de Olo	estatus patrimonial:	monumento histórico inscrito	huso horario:	Hora central europea	instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476	situado en la entidad territorial administrativa:	VII Distrito de París	propietario:	París	catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener	forma parte de:	Riberas del Sena en París	materias empleadas:	hierro forjado	nombrado por:	Gustave Eiffel	categoría principal del tema:	Q6741108	calle:	Campo de Marte	arquitecto:	Stephen Sauvestre	duplicado permanente:	Q25689717	constructor:	Gustave Eiffel	ingeniero estructural:	Émile Nouguier, Maurice Koechlin	Identificador VIAF:	236835198	Identificador GND:	4102872-7	Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247	Identificador BnF:	11941003z	<table border="1"> <thead> <tr> <th colspan="2">Torre Eiffel</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>monumento en París, Francia</td> </tr> <tr> <td>catálogo:</td> <td>Wikipedia Lista de artículos que toda Wikipedia debería tener</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia de Olo</td> </tr> <tr> <td>país:</td> <td>Francia</td> </tr> <tr> <td>instancia de:</td> <td>destino turístico, torre de observación, punto de referencia, Q1440476</td> </tr> <tr> <td>coordenadas:</td> <td>Point(2.2953 48.858)</td> </tr> <tr> <td>situado en la entidad territorial administrativa:</td> <td>VII Distrito de París</td> </tr> <tr> <td>estatus patrimonial:</td> <td>monumento histórico inscrito</td> </tr> <tr> <td>huso horario:</td> <td>Hora central europea</td> </tr> <tr> <td>propietario:</td> <td>París</td> </tr> <tr> <td>Identificador GeoNames:</td> <td>6254976</td> </tr> <tr> <td>categoría en Commons:</td> <td>Eiffel Tower</td> </tr> <tr> <td>Identificador Freebase:</td> <td>/m/02j81</td> </tr> <tr> <td>Identificador VIAF:</td> <td>236835198</td> </tr> <tr> <td>código postal:</td> <td>75007</td> </tr> <tr> <td>fecha de fundación o creación:</td> <td>1889-01-01T00:00:00Z</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.toureiffel.paris</td> </tr> <tr> <td>forma parte de:</td> <td>Riberas del Sena en París</td> </tr> <tr> <td>Identificador GND:</td> <td>4102872-7</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Q6741108</td> </tr> <tr> <td>Identificador autoridades de la Biblioteca del Congreso de EEUU:</td> <td>sh85136247</td> </tr> <tr> <td>materias empleadas:</td> <td>Hierro forjado</td> </tr> <tr> <td>Identificador BnF:</td> <td>11941003z</td> </tr> </tbody> </table>	Torre Eiffel		Description:	monumento en París, Francia	catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener	descrito en:	Enciclopedia de Olo	país:	Francia	instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476	coordenadas:	Point(2.2953 48.858)	situado en la entidad territorial administrativa:	VII Distrito de París	estatus patrimonial:	monumento histórico inscrito	huso horario:	Hora central europea	propietario:	París	Identificador GeoNames:	6254976	categoría en Commons:	Eiffel Tower	Identificador Freebase:	/m/02j81	Identificador VIAF:	236835198	código postal:	75007	fecha de fundación o creación:	1889-01-01T00:00:00Z	sitio web oficial:	http://www.toureiffel.paris	forma parte de:	Riberas del Sena en París	Identificador GND:	4102872-7	categoría principal del tema:	Q6741108	Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247	materias empleadas:	Hierro forjado	Identificador BnF:	11941003z	<table border="1"> <thead> <tr> <th colspan="2">Torre Eiffel</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>monumento en París, Francia</td> </tr> <tr> <td>instancia de:</td> <td>Q1440476, punto de referencia</td> </tr> <tr> <td>SUDOC:</td> <td>027348857</td> </tr> <tr> <td>coordenadas:</td> <td>Point(2.2953 48.858)</td> </tr> <tr> <td>Identificador autoridades de la Biblioteca del Congreso de EEUU:</td> <td>sh85136247</td> </tr> <tr> <td>descrito en:</td> <td>Enciclopedia de Olo</td> </tr> <tr> <td>propietario:</td> <td>París</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.toureiffel.paris</td> </tr> <tr> <td>Identificador NDL:</td> <td>01193444</td> </tr> <tr> <td>Identificador Mérimée:</td> <td>PA00088001</td> </tr> <tr> <td>Google Doodle:</td> <td>120th anniversary of the public opening of the Eiffel tower</td> </tr> <tr> <td>fecha de inauguración oficial:</td> <td>1889-03-31T00:00:00Z</td> </tr> <tr> <td>galería en Commons:</td> <td>Tour Eiffel</td> </tr> <tr> <td>ingeniero estructural:</td> <td>Émile Nouguier</td> </tr> <tr> <td>Identificador Magillary:</td> <td>SQZP5-MQZLJ3M9N-pT44bghu02906</td> </tr> <tr> <td>estatus patrimonial:</td> <td>monumento histórico inscrito</td> </tr> <tr> <td>Identificador PSS:</td> <td>FR-75056-19</td> </tr> <tr> <td>fecha de fundación o creación:</td> <td>1889-01-01T00:00:00Z</td> </tr> <tr> <td>Identificador FAST:</td> <td>1153112</td> </tr> <tr> <td>huso horario:</td> <td>Hora central europea</td> </tr> <tr> <td>nombre de usuario en Twitter:</td> <td>LaTourEiffel</td> </tr> <tr> <td>ocupante:</td> <td>Q16008668</td> </tr> <tr> <td>nombrado por:</td> <td>Gustave Eiffel</td> </tr> <tr> <td>Identificador Chemins de mémoire:</td> <td>la-tour-eiffel</td> </tr> <tr> <td>Identificador Freebase:</td> <td>/m/02j81</td> </tr> </tbody> </table>	Torre Eiffel		Description:	monumento en París, Francia	instancia de:	Q1440476, punto de referencia	SUDOC:	027348857	coordenadas:	Point(2.2953 48.858)	Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247	descrito en:	Enciclopedia de Olo	propietario:	París	sitio web oficial:	http://www.toureiffel.paris	Identificador NDL:	01193444	Identificador Mérimée:	PA00088001	Google Doodle:	120th anniversary of the public opening of the Eiffel tower	fecha de inauguración oficial:	1889-03-31T00:00:00Z	galería en Commons:	Tour Eiffel	ingeniero estructural:	Émile Nouguier	Identificador Magillary:	SQZP5-MQZLJ3M9N-pT44bghu02906	estatus patrimonial:	monumento histórico inscrito	Identificador PSS:	FR-75056-19	fecha de fundación o creación:	1889-01-01T00:00:00Z	Identificador FAST:	1153112	huso horario:	Hora central europea	nombre de usuario en Twitter:	LaTourEiffel	ocupante:	Q16008668	nombrado por:	Gustave Eiffel	Identificador Chemins de mémoire:	la-tour-eiffel	Identificador Freebase:	/m/02j81	<table border="1"> <thead> <tr> <th colspan="2">Torre Eiffel</th> </tr> </thead> <tbody> <tr> <td>Description:</td> <td>monumento en París, Francia</td> </tr> <tr> <td>instancia de:</td> <td>destino turístico, torre de observación, Q1440476, punto de referencia</td> </tr> <tr> <td>país:</td> <td>Francia</td> </tr> <tr> <td>coordenadas:</td> <td>Point(2.2953 48.858)</td> </tr> <tr> <td>situado en la entidad territorial administrativa:</td> <td>VII Distrito de París</td> </tr> <tr> <td>Identificador Geonames:</td> <td>6254976</td> </tr> <tr> <td>categoría en Commons:</td> <td>Eiffel Tower</td> </tr> <tr> <td>Identificador Freebase:</td> <td>/m/02j81</td> </tr> <tr> <td>Identificador VIAF:</td> <td>236835198</td> </tr> <tr> <td>código postal:</td> <td>75007</td> </tr> <tr> <td>fecha de fundación o creación:</td> <td>1889-01-01T00:00:00Z</td> </tr> <tr> <td>sitio web oficial:</td> <td>http://www.toureiffel.paris</td> </tr> <tr> <td>forma parte de:</td> <td>Riberas del Sena en París</td> </tr> <tr> <td>Identificador GND:</td> <td>4102872-7</td> </tr> <tr> <td>categoría principal del tema:</td> <td>Q6741108</td> </tr> <tr> <td>Identificador autoridades de la Biblioteca del Congreso de EEUU:</td> <td>sh85136247</td> </tr> <tr> <td>materias empleadas:</td> <td>Hierro forjado</td> </tr> <tr> <td>Identificador BnF:</td> <td>11941003z</td> </tr> <tr> <td>SUDOC:</td> <td>027348857</td> </tr> <tr> <td>nombre nativo:</td> <td>tour Eiffel</td> </tr> <tr> <td>altura:</td> <td>300</td> </tr> </tbody> </table>	Torre Eiffel		Description:	monumento en París, Francia	instancia de:	destino turístico, torre de observación, Q1440476, punto de referencia	país:	Francia	coordenadas:	Point(2.2953 48.858)	situado en la entidad territorial administrativa:	VII Distrito de París	Identificador Geonames:	6254976	categoría en Commons:	Eiffel Tower	Identificador Freebase:	/m/02j81	Identificador VIAF:	236835198	código postal:	75007	fecha de fundación o creación:	1889-01-01T00:00:00Z	sitio web oficial:	http://www.toureiffel.paris	forma parte de:	Riberas del Sena en París	Identificador GND:	4102872-7	categoría principal del tema:	Q6741108	Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247	materias empleadas:	Hierro forjado	Identificador BnF:	11941003z	SUDOC:	027348857	nombre nativo:	tour Eiffel	altura:	300
Torre Eiffel																																																																																																																																																																																																																																																
Description:	monumento en París, Francia																																																																																																																																																																																																																																															
catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener																																																																																																																																																																																																																																															
descrito en:	Enciclopedia de Olo																																																																																																																																																																																																																																															
país:	Francia																																																																																																																																																																																																																																															
propietario:	París																																																																																																																																																																																																																																															
estatus patrimonial:	monumento histórico inscrito																																																																																																																																																																																																																																															
huso horario:	Hora central europea																																																																																																																																																																																																																																															
situado en la entidad territorial administrativa:	VII Distrito de París																																																																																																																																																																																																																																															
instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476																																																																																																																																																																																																																																															
forma parte de:	Riberas del Sena en París																																																																																																																																																																																																																																															
nombrado por:	Gustave Eiffel																																																																																																																																																																																																																																															
constructor:	Gustave Eiffel																																																																																																																																																																																																																																															
materias empleadas:	hierro forjado																																																																																																																																																																																																																																															
ingeniero estructural:	Émile Nouguier, Maurice Koechlin																																																																																																																																																																																																																																															
calle:	Campo de Marte																																																																																																																																																																																																																																															
arquitecto:	Stephen Sauvestre																																																																																																																																																																																																																																															
categoría principal del tema:	Q6741108																																																																																																																																																																																																																																															
duplicado permanente:	Q25689717																																																																																																																																																																																																																																															
Identificador VIAF:	236835198																																																																																																																																																																																																																																															
Identificador GND:	4102872-7																																																																																																																																																																																																																																															
Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247																																																																																																																																																																																																																																															
Identificador BnF:	11941003z																																																																																																																																																																																																																																															
Torre Eiffel																																																																																																																																																																																																																																																
Description:	monumento en París, Francia																																																																																																																																																																																																																																															
país:	Francia																																																																																																																																																																																																																																															
descrito en:	Enciclopedia de Olo																																																																																																																																																																																																																																															
estatus patrimonial:	monumento histórico inscrito																																																																																																																																																																																																																																															
huso horario:	Hora central europea																																																																																																																																																																																																																																															
instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476																																																																																																																																																																																																																																															
situado en la entidad territorial administrativa:	VII Distrito de París																																																																																																																																																																																																																																															
propietario:	París																																																																																																																																																																																																																																															
catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener																																																																																																																																																																																																																																															
forma parte de:	Riberas del Sena en París																																																																																																																																																																																																																																															
materias empleadas:	hierro forjado																																																																																																																																																																																																																																															
nombrado por:	Gustave Eiffel																																																																																																																																																																																																																																															
categoría principal del tema:	Q6741108																																																																																																																																																																																																																																															
calle:	Campo de Marte																																																																																																																																																																																																																																															
arquitecto:	Stephen Sauvestre																																																																																																																																																																																																																																															
duplicado permanente:	Q25689717																																																																																																																																																																																																																																															
constructor:	Gustave Eiffel																																																																																																																																																																																																																																															
ingeniero estructural:	Émile Nouguier, Maurice Koechlin																																																																																																																																																																																																																																															
Identificador VIAF:	236835198																																																																																																																																																																																																																																															
Identificador GND:	4102872-7																																																																																																																																																																																																																																															
Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247																																																																																																																																																																																																																																															
Identificador BnF:	11941003z																																																																																																																																																																																																																																															
Torre Eiffel																																																																																																																																																																																																																																																
Description:	monumento en París, Francia																																																																																																																																																																																																																																															
catálogo:	Wikipedia Lista de artículos que toda Wikipedia debería tener																																																																																																																																																																																																																																															
descrito en:	Enciclopedia de Olo																																																																																																																																																																																																																																															
país:	Francia																																																																																																																																																																																																																																															
instancia de:	destino turístico, torre de observación, punto de referencia, Q1440476																																																																																																																																																																																																																																															
coordenadas:	Point(2.2953 48.858)																																																																																																																																																																																																																																															
situado en la entidad territorial administrativa:	VII Distrito de París																																																																																																																																																																																																																																															
estatus patrimonial:	monumento histórico inscrito																																																																																																																																																																																																																																															
huso horario:	Hora central europea																																																																																																																																																																																																																																															
propietario:	París																																																																																																																																																																																																																																															
Identificador GeoNames:	6254976																																																																																																																																																																																																																																															
categoría en Commons:	Eiffel Tower																																																																																																																																																																																																																																															
Identificador Freebase:	/m/02j81																																																																																																																																																																																																																																															
Identificador VIAF:	236835198																																																																																																																																																																																																																																															
código postal:	75007																																																																																																																																																																																																																																															
fecha de fundación o creación:	1889-01-01T00:00:00Z																																																																																																																																																																																																																																															
sitio web oficial:	http://www.toureiffel.paris																																																																																																																																																																																																																																															
forma parte de:	Riberas del Sena en París																																																																																																																																																																																																																																															
Identificador GND:	4102872-7																																																																																																																																																																																																																																															
categoría principal del tema:	Q6741108																																																																																																																																																																																																																																															
Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247																																																																																																																																																																																																																																															
materias empleadas:	Hierro forjado																																																																																																																																																																																																																																															
Identificador BnF:	11941003z																																																																																																																																																																																																																																															
Torre Eiffel																																																																																																																																																																																																																																																
Description:	monumento en París, Francia																																																																																																																																																																																																																																															
instancia de:	Q1440476, punto de referencia																																																																																																																																																																																																																																															
SUDOC:	027348857																																																																																																																																																																																																																																															
coordenadas:	Point(2.2953 48.858)																																																																																																																																																																																																																																															
Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247																																																																																																																																																																																																																																															
descrito en:	Enciclopedia de Olo																																																																																																																																																																																																																																															
propietario:	París																																																																																																																																																																																																																																															
sitio web oficial:	http://www.toureiffel.paris																																																																																																																																																																																																																																															
Identificador NDL:	01193444																																																																																																																																																																																																																																															
Identificador Mérimée:	PA00088001																																																																																																																																																																																																																																															
Google Doodle:	120th anniversary of the public opening of the Eiffel tower																																																																																																																																																																																																																																															
fecha de inauguración oficial:	1889-03-31T00:00:00Z																																																																																																																																																																																																																																															
galería en Commons:	Tour Eiffel																																																																																																																																																																																																																																															
ingeniero estructural:	Émile Nouguier																																																																																																																																																																																																																																															
Identificador Magillary:	SQZP5-MQZLJ3M9N-pT44bghu02906																																																																																																																																																																																																																																															
estatus patrimonial:	monumento histórico inscrito																																																																																																																																																																																																																																															
Identificador PSS:	FR-75056-19																																																																																																																																																																																																																																															
fecha de fundación o creación:	1889-01-01T00:00:00Z																																																																																																																																																																																																																																															
Identificador FAST:	1153112																																																																																																																																																																																																																																															
huso horario:	Hora central europea																																																																																																																																																																																																																																															
nombre de usuario en Twitter:	LaTourEiffel																																																																																																																																																																																																																																															
ocupante:	Q16008668																																																																																																																																																																																																																																															
nombrado por:	Gustave Eiffel																																																																																																																																																																																																																																															
Identificador Chemins de mémoire:	la-tour-eiffel																																																																																																																																																																																																																																															
Identificador Freebase:	/m/02j81																																																																																																																																																																																																																																															
Torre Eiffel																																																																																																																																																																																																																																																
Description:	monumento en París, Francia																																																																																																																																																																																																																																															
instancia de:	destino turístico, torre de observación, Q1440476, punto de referencia																																																																																																																																																																																																																																															
país:	Francia																																																																																																																																																																																																																																															
coordenadas:	Point(2.2953 48.858)																																																																																																																																																																																																																																															
situado en la entidad territorial administrativa:	VII Distrito de París																																																																																																																																																																																																																																															
Identificador Geonames:	6254976																																																																																																																																																																																																																																															
categoría en Commons:	Eiffel Tower																																																																																																																																																																																																																																															
Identificador Freebase:	/m/02j81																																																																																																																																																																																																																																															
Identificador VIAF:	236835198																																																																																																																																																																																																																																															
código postal:	75007																																																																																																																																																																																																																																															
fecha de fundación o creación:	1889-01-01T00:00:00Z																																																																																																																																																																																																																																															
sitio web oficial:	http://www.toureiffel.paris																																																																																																																																																																																																																																															
forma parte de:	Riberas del Sena en París																																																																																																																																																																																																																																															
Identificador GND:	4102872-7																																																																																																																																																																																																																																															
categoría principal del tema:	Q6741108																																																																																																																																																																																																																																															
Identificador autoridades de la Biblioteca del Congreso de EEUU:	sh85136247																																																																																																																																																																																																																																															
materias empleadas:	Hierro forjado																																																																																																																																																																																																																																															
Identificador BnF:	11941003z																																																																																																																																																																																																																																															
SUDOC:	027348857																																																																																																																																																																																																																																															
nombre nativo:	tour Eiffel																																																																																																																																																																																																																																															
altura:	300																																																																																																																																																																																																																																															

Figura A.14: Infoboxes automáticas para la Torre Eiffel

1		2		3		4		5	
Museo Guggenheim Bilbao									
Description:	museo de arte moderno y contemporáneo en Bilbao, España	Description:	museo de arte moderno y contemporáneo en Bilbao, España	Description:	museo de arte moderno y contemporáneo en Bilbao, España	Description:	museo de arte moderno y contemporáneo en Bilbao, España	Description:	museo de arte moderno y contemporáneo en Bilbao, España
instancia de:	edificio, museo de arte								
país:	España								
coordenadas:	Punt(2.934286111, 43.268605559)								
situado en la entidad territorial administrativa:	Bilbao	situado en la entidad territorial administrativa:	Bilbao	situado en la entidad territorial administrativa:	Bilbao	situado en la entidad territorial administrativa:	Bilbao	situado en la entidad territorial administrativa:	Bilbao
identificador GeoNames:	6452771								
categoría en Commons:	Guggenheim Museum Bilbao								
identificador Freebase:	/m/0n88f								
identificador VIAF:	128744734								
fecha de fundación o creación:	1997-10-18T00:00:00Z								
sitio web oficial:	http://www.guggenheim-bilbao.es								
premio recibido:	Premio Vasco Universal								
identificador GND:	5210044-3								
forma parte de:	Fundación Solomon R. Guggenheim	forma parte de:	Fundación Solomon R. Guggenheim	forma parte de:	Fundación Solomon R. Guggenheim	forma parte de:	Fundación Solomon R. Guggenheim	forma parte de:	Fundación Solomon R. Guggenheim
identificador autoridades de la Biblioteca del Congreso de EEUU:	n95006602	identificador autoridades de la Biblioteca del Congreso de EEUU:	n95006602	identificador autoridades de la Biblioteca del Congreso de EEUU:	n95006602	identificador autoridades de la Biblioteca del Congreso de EEUU:	n95006602	identificador autoridades de la Biblioteca del Congreso de EEUU:	n95006602
ISNI:	0000 0004 1775 8168								
identificador Bnf:	135589733								
SUDOC:	052502554								
identificador de tema en Quora:	Guggenheim-Museum								
nombre de usuario en Twitter:	MuseoGuggenheim								
identificador ULAN:	500304957								
identificador NLI:	000096579								
identificador BNE:	XX114466								
identificador Encyclopædia Britannica Online:	/m/0n88f								
movimiento:	arquitectura postmoderna								

Figura A.15: Infoboxes automáticas para el Museo Guggenheim

A.4. Fichas de evaluación

A continuación se adjuntan dos fichas de respuesta, a modo de ejemplo.

5

Nombre de primera entidad B

		I	II	III	IV	V
1	oro	3	6	3	7	4
2		2	3	1	1	1
3	uma	4	4	1	4	4
4	5"	2	4	6	2	4
5	53	4	3	3	5	3
6	3	2	3	2	7	6
7		4	2	5	4	3
8	1.7	5	3	1	4	7
9	150	5	3	6	3	4
10		4	1	4	6	5
11		1	5	2	4	3
12		1	4	4	2	6
13		3	5	3	5	4
14		3	2	3	4	3
15		5	4	2	5	3

Observaciones generales o sugerencias:
inmunda

Hay campos que deberían ser simulados, entre estos estos:

- Instancia de ser humano
- catalogo: Wikipedia...
- Los identificadores en general, los cuantos no son lo internos para el usuario promedio

El campo "descrito en:" es lo internos pero debería estar más abajo en la tabla



Nombre de primera entidad:

Doce hombres en pugna (C-312)

	I	II	III	IV	V
1	5	4	2	3	6
2	4	3	6	5	2
3	6	4	5	5	6
4	5	6	4	2	3
5	3	2	4	6	5
6	3	2	6	5	4
7	4	2	3	5	6
8	6	5	3	4	2
9	4	2	5	3	4
10	2	6	3	5	4
11	2	2	2	2	2
12	2	4	3	5	1
13	3	3	3	3	3
14	3	2	4	4	5
15	3	4	1	2	1

Observaciones generales o sugerencias:

Creo que deberían eliminar siempre los identificadores de freebase y los datos en otros idiomas (árabs...)
Igual, podrían eliminar los datos en forma de URI