

UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

SISTEMAS BASADOS EN CONOCIMIENTOS

Tema:

Conflictos entre países.

Integrantes:

Jonathan Yaguana Jordy Sarango

Ing. Janneth Chicaiza

Loja-Ecuador

Objetivos

- Representar los datos referentes al conflicto bélico de Siria de manera estructurada mediante web semántica.
- Recolectar datos y definir variables relevantes al problema del terrorismo en Siria.
- Proponer una solución semántica que unifique la información disponible en sitios web de ONG's.
- Construir una aplicación web que permita visualizar los datos del conflicto de bélico de Siria de manera precisa y ordenada.

Justificación

La información referente al conflicto bélico de Siria, al no encontrarse estructurada ni representada, conlleva al problema de la distorsión de la misma. Esta misma información se encuentra publicada en páginas que utilizan tecnologías como HTML. Haciendo de esto una mala presentación de los datos a través de HTML.

El presente trabajo consiste en recolectar, analizar y representar la información referente al terrorismo con herramientas de web semántica, de manera particular la situación actual de Siria. Se ha considerado como principal fuente de datos organizaciones sin nes de lucro y proyectos colaborativos anes al tema. Los beneficiarios son todos los internautas que deseen hacer consultas acerca del conflicto en Siria

Descripción del alcance

El presente trabajo consiste en recolectar, analizar y representar la información referente al terrorismo con herramientas de web semántica, de manera particular la situación actual de Siria. Se ha considerado como principal fuente de datos de organizaciones sin fines de lucro y proyectos colaborativos afines al tema(VDC, The Carter Center[1],[2]).

Trabajos Relacionados

- AmCAT, es un trabajo de tesis que consiste en herramienta de almacenaje, análisis y consulta de artículos periodísticos y otros mensajes de los medios de comunicación vía web semántica.[3]
- SWEET. Ontología y herramienta de búsqueda denominada GCMD para la consulta de términos referentes a ciencias de la Tierra. [4]
- Tienda de Compras, Ventas y Alquiler, descrita con FOAF.[5]

Descripción de las fuentes de datos:

Se realizó un análisis de las fuentes de datos que servirán para la generación del nuevo conocimiento. Las mismas que se a tomado para el presente trabajo han

sido referentes a Siria y los conflictos bélicos que presentan. Para tomar en consideración se ha decidido indagar sobre fuentes confiables:

Carter Center: https://www.cartercenter.org/
 Organización no gubernamental, que ayuda a mejorar la vida de las personas en más de 80 países mediante la resolución de conflictos y prevenir enfermedades; promueve la democracia y derechos humanos. Organización que maneja un gran volumen de información acerca de conflictos,



enfermedades y corrupción que se suscita alrededor del mundo.

Violations Documentation Center in Syria http://vdc-sy.net/en/

El Centro de Documentación de Violaciones en Siria (VDC) que garantiza la documentación cuidadosa e independiente de todo tipo de violaciones de los derechos humanos en Siria. Presenta su información con el objetivo era convertirse en una referencia no solo para los medios, sino también para todos los procedimientos



futuros de rendición de cuentas y justicia para Siria.

 REGIANAL REFUGEE & RESILIENCE PLAN 2018- 2019 (3RP) http://www.3rpsyriacrisis.org/crisis/

El Plan Regional de Refugiados y Resiliencia (3RP) página que presenta información acerca de los planes elaborados para garantizar la protección, asistencia humanitaria y fortalecer la resiliencia en naciones como la República Árabe de Egipto, la República de Iraq, el Reino



Hachemita de Jordania, la República Libanesa y la República de Turquía

Informe sobre la extracción/obtención de datos

Para la obtención de los datos de lo realizó una búsqueda de información que sea formal, es decir correcta. Nos dirigiremos a cada una de las páginas de los proveedores (3RP, Carter Center y Violations Documentation Center in Syria) para la extracción, y transformación de datos e información fiable. Para que la

recolección de esta información sea precisa y clara. Y el usuario tenga la facilidad de entender y este a gusto con los resultados.

Para lograr esto se utilizó la técnica de de extracción web scraping, la misma que fue de gran utilidad para obtener la información presentada mediante tablas en cada una de las páginas de los proveedores. Información acerca de:

- Los ataques o conflictos
- Grupos Armados
- Gobiernos en los cuales se dieron dichos ataques o conflictos
- Personas que han sido víctimas (Muertos, Desaparecidos, Detenidos)
- Lugares donde se suscitó el conflicto.
- Iniciador y objetivo del conflicto.
- Medios de comunicación que presentaron la información anterior

Limpieza de los datos

Para el proceso de limpieza de los datos se utilizó la herramienta de OpenRefine para la eliminación de caracteres especiales como: tildes, signos de interrogación, entre otros de los campos (provincia, lugar de nacimiento, fecha de desaparición, nombre, fuente de información, iniciador, objetivo entre otros) esto con el fin de generar datos confiables, eliminar espacios en blanco dentro de la data ya que estos valores no son valores nulos sino que están constando como un carácter de espacio en blanco, normalización del texto ya que en varias de las columnas se presentan diferentes formatos entre Mayúsculas/Minúsculas para esto se cambio los textos estableciendo que todos empiecen con una letra Mayúscula, eliminar espacios al inicio y al final dentro de las celdas esto con la finalidad de eliminar los espacios que se pueden mostrar en el momento de generar el RDF.



Documento de especificación de requerimientos

1 PROPÓSITO

La ontología sobre conflictos entre países que será empleada como vocabulario para describir los términos asociados al dominio indicado.

2 ALCANCE

La ontología está enfocada en el dominio de los problemas que existe en el estado de Siria y la involucración de varios países para un uso general. El nivel de granularidad está directamente relacionado con el conflicto de Siria y los países involucrados que se especifican después y demás términos indicados.

3 LENGUAJE DE IMPLEMENTACIÓN

La ontología será implementada en el lenguaje OWL (ontology Web language), usado Protégé como herramienta de edición de ontologías.

4 POSIBLES USUARIOS FINALES

1. Los internautas que deseen hacer consultas acerca de los conflictos de Siria.

5 **POSIBLES USOS**

- 1. Búsqueda de información sobre las víctimas que han dejado los conflictos de Siria.
- 2. Búsqueda sobre los ataques sucedidos en cada gobernación.
- 3. Presentar cifras generales sobre los conflictos sucedidos.

6 **REQUERIMIENTOS**

a. Requerimientos no funcionales

- La ontología se desarrollará inicialmente en el idioma español, con proyección para implementación en inglés

b. Requerimientos funcionales: Cuestiones de competencia

- 1. ¿En qué país se está dando un conflicto bélico?
- 2. ¿Cuál es la causa de muerte más común?
- 3. ¿Dónde vivía X victima antes de morir?
- 4. ¿Qué fuente reportó X ataque?
- 5. ¿Qué grupo armado fue causante de la muerte de X víctima?
- 6. ¿Qué grupos armados han atacado X gobernación?
- 7. ¿Cuántas personas han sido reportadas fallecidas?
- 8. ¿Cuántas personas han sido reportadas detenidas?
- 9. ¿Cuántas personas han sido reportadas desaparecidas?
- 10.¿Qué gobernación reporta más ataques?

7	Pre-glosario de términos					
	Términos de las cuestiones de competencia					
	Siria	involucrados	causa de muerte	inicio		
	conflicto	víctimas	reporte	civiles		
	paises	ataques	grupo armado			
	estado	gobernación	atacantes			
	bélico	cifras generales	personas encontradas			

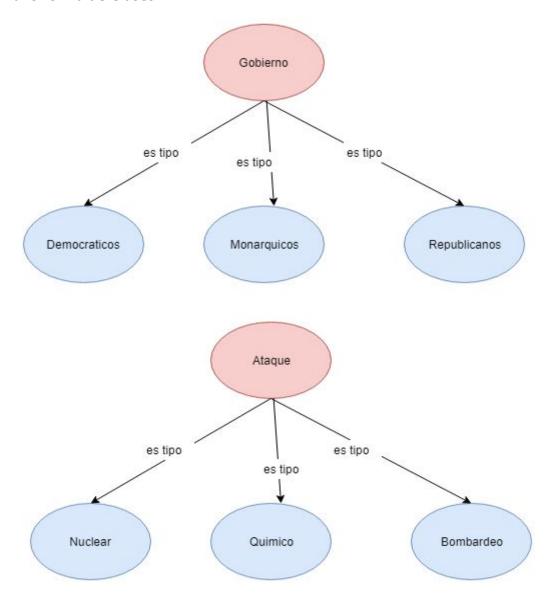
Extracto del glosario de términos

Nombre	Sinónimos	Acrónim os	Descripción	Tipo
Siria			República Árabe Siria, país soberano del Oriente Próximo, sumida en una Guerra Civil Siria.	Concepto
conflicto			Oposición o desacuerdo entre personas o cosas.	Concepto
paises			Territorio con características geográficas y culturales propias.	Relación
estado			Organización política forzosa e involuntaria constituida por	Concepto

		instituciones burocráticas y administrativa s estables.	
bélico	Guerrero, agresivo, militarista	 Hace referencia a aquello perteneciente a la guerra.	Concepto
involucrados		 Que está provisto de involucrado, o que forma parte de algo.	Concepto
víctimas		 Un ser viviente sacrificado o destinado a sacrificio, o quien recibe daño por caso fortuito o culpa ajena.	Concepto
ataques		 Temas para los cuales existen becas	Concepto
gobernación		 País al que pertenece una persona	Concepto
cifras generales			Concepto
causa de muerte			
reporte			Concepto
grupo armado			Concepto
atacantes			Relación

personas encontradas		
inicio		
civiles		

Taxonomía de Clases



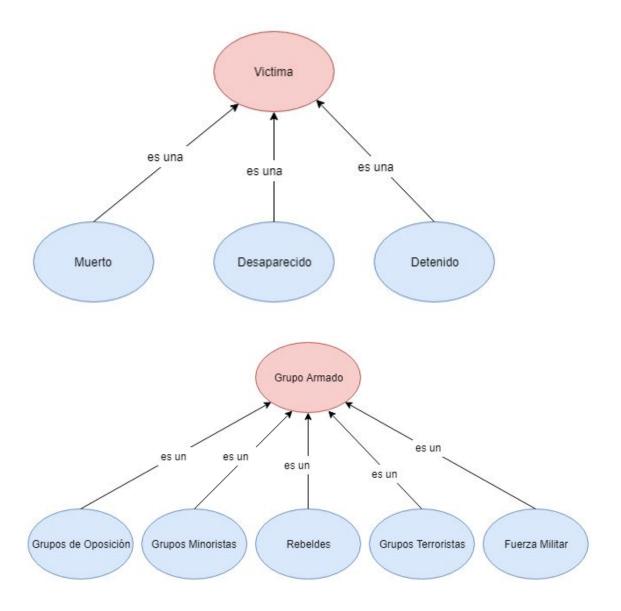
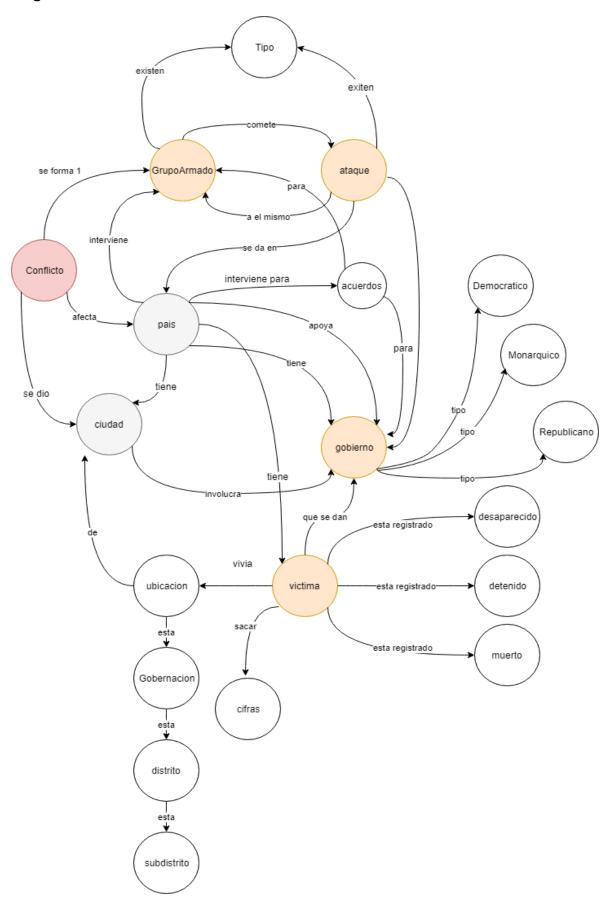


Diagrama de Relaciones Binarias



Diccionario de Clases

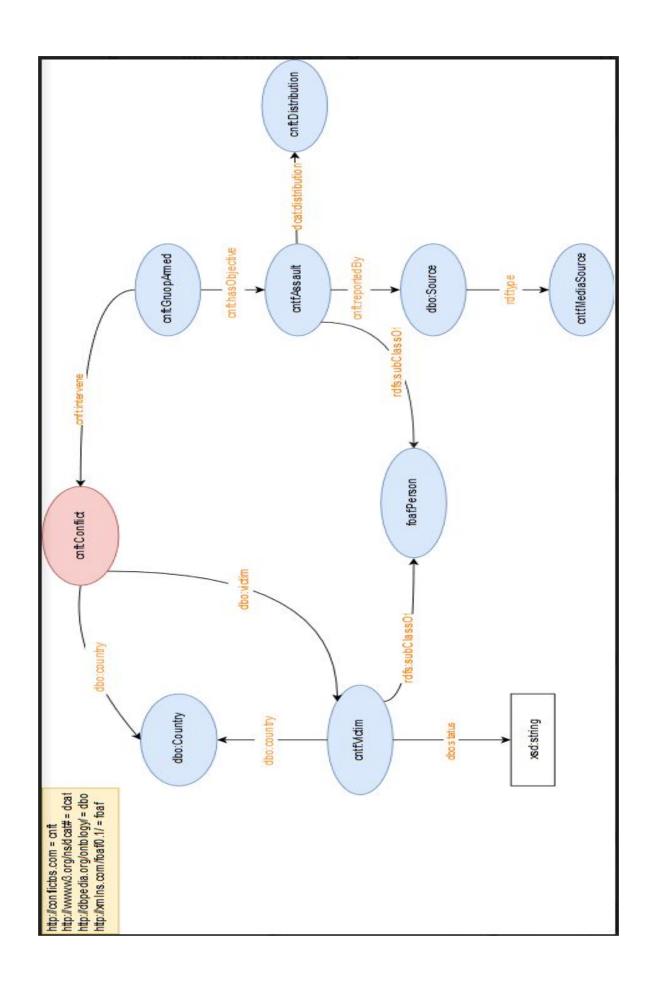
NOMBRE DEL CONCEPTO	INSTANCIAS		ATRIBUTOS DE CLASE	ATRIBUTOS DE INSTANCIA	RELACIONE S
Gobierno				Nombre Sede	
Republicanos	Vietnam Popular democrática d corea	de de			es un
Monárquicos	Monarquía constitucional Noruega Monarquía constitucional Suecia Monarquía constitucional España				es un
Democráticos	Alianza Pais				es un
Conflicto				Nombre	
Conflictos Internos	Guerras civiles Guerras Territoriales Lucha por poder				
Conflictos Internacionale s	Guerras Territoriales				
Víctima				Nombre	
Muertos					es una
Desaparecido s					es una

Detenidos			es una
Grupos Armados	Grupos de Oposición Rebeldes Grupos minoristas Grupos terroristas Fuerza militar	Nombre Sede	Comente
Ataque	Nuclear Químico Bombardeos		a el mismo se da en ocurre contra

Matriz de Atributos

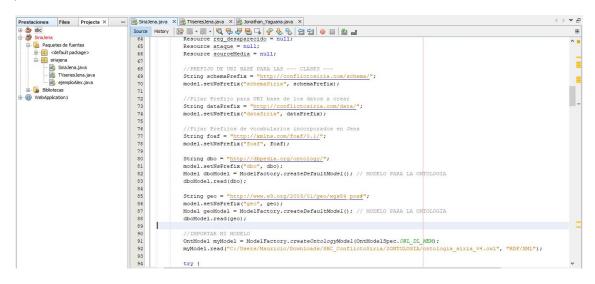
Nombre de Atributo	Concepto	Tipo de Valor	Rango de Valores	Cardinalida d
Nombre	Gobierno Grupo Armado Víctima	cadena de caracteres cadena de caracteres cadena de caracteres		(1,1)
Dirección	Víctima	cadena de caracteres	Desaparecid o Muerto Detenido	(1,N)
Sede	Ciudad País	cadena de caracteres cadena de caracteres		(1,1)
Sexo	Víctima	cadena de caracteres	Desaparecid o Muerto Detenido	(1,1)
Lugar de Nacimiento	Víctima	Fecha	Desaparecid o	(1,1)

			Muerto Detenido	
Fecha desaparición	Víctima	Fecha	Desaparecid o	(1,1)
Fecha Encontrado	Víctima	Fecha	Desaparecid o	(1,1)
Fecha detención	Víctima	Fecha	Detenido	(1,1)
Causa	Ataque Víctima	cadena de caracteres cadena de caracteres		(1,N)
Descripción	Grupo Armado	cadena de caracteres		(1,1)
Iniciador	Gobierno Grupo Armado País	cadena de caracteres cadena de caracteres cadena de caracteres		(1,N)
fecha de muerte	Víctimas	fecha	Muertos	(1,1)
ciudad	Grupo Armado Víctima Ataque	cadena de caracteres cadena de caracteres cadena de caracteres		(1,N)
tipo de ciudadano	víctima	cadena de caracteres	Muertos Detenidos Desaparecid os	(1,1)



Generación de datos RDF

Para la generación de los datos de trabajo con Jena que es un marco de Java que proporciona un entorno programático para RDF, RDFS y OWL, SPARQL e incluye un motor de inferencia basado en reglas. Dentro de la cual los datos se generarán, correctamente armados en tripletas, para realizar la aplicación funcional. El cual tiene como objetivo, informar al usuario sobre los aspectos más generales acerca de los conflictos bélicos que se dan en los países. Para publicar los datos como Linkend Data.



Informe sobre evaluación de repositorios RDF y almacenamiento de datos

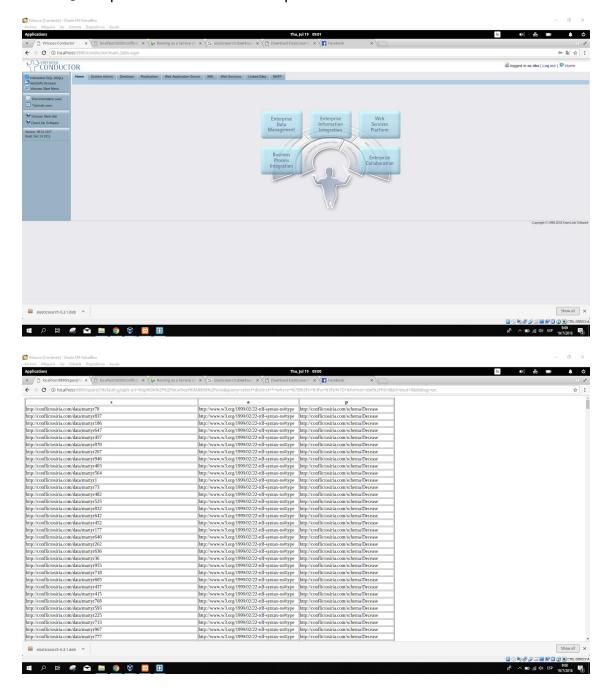
Para el almacenamiento de los datos RDF se a utilizado el servidor Virtuoso la cual nos permite desarrollar soluciones eBusiness que utilizan XML como una capa básica de acceso a datos común que proporciona acceso transparente a datos estructurados y no estructurados. Se utilizará Virtuoso para garantizar que las URI's generadas sean simples y manejables y se pueda acceder a los datos de mejor manera.

Además del almacenamiento de los datos, Virtuoso nos permite generar las consultas Sparql para verificar que los datos este nativamente en RDF y obtener todas las tripletas posibles en base a las consultas realizadas y con las mismas dar solución a las preguntas planteadas para este caso de estudio.

Preguntas como:

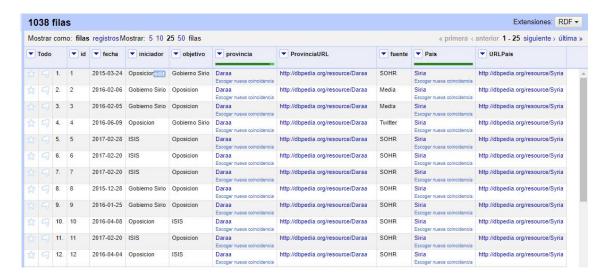
- ¿Cuál es la causa de muerte más común?
- ¿Cuántas personas han sido víctimas de
- fallecimiento/detención/desaparición ?
- ¿Cuáles son los grupos armados implicados en el conflicto?
- ¿Cuántos ataques ha recibido una gobernación?

- ¿Cuántas personas civiles han muerto hasta ahora?
- ¿En qué gobernación se registra la mayor cantidad de fallecidos por
- género?
- ¿Cuántos ataques se registran por cada gobernación o grupo armado?
- ¿Por quién fue iniciado un ataque?



Reconciliación de datos

Para poder realizar el trabajo de la reconciliación de los datos se ha utilizado OpenRefine el cual nos provee de algunas herramientas para realizar este proceso. Para esto se a tomado en cuenta lo que es la Provincia y el País para poderlos enlazar con la DBpedia y estos puedan ser accedidos mediante la búsqueda de los mismos.



Enriquecimiento de los datos

Aplicación

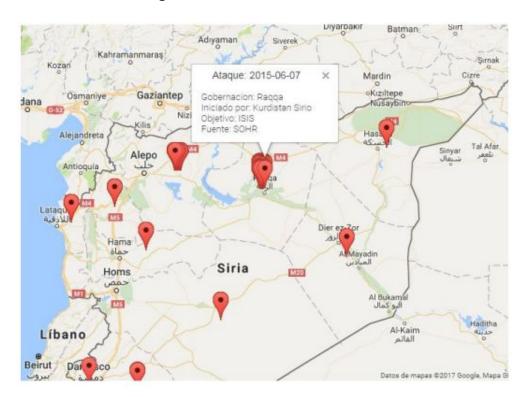
Para habilitar la visualización de los datos enlazados referentes al conflicto en Siria se ha tomado en cuenta, en primer lugar que la ontología pueda responder a las siguientes preguntas:

- ¿Cuál es la causa de muerte más común?
- ¿Cuántas personas han sido víctimas de
- fallecimiento/detención/desaparición ?
- ¿Cuáles son los grupos armados implicados en el conflicto?
- ¿Cuántos ataques ha recibido una gobernación?
- ¿Cuántas personas civiles han muerto hasta ahora?
- ¿En qué gobernación se registra la mayor cantidad de fallecidos por
- género?
- ¿Cuántos ataques se registran por cada gobernación o grupo armado?
- ¿Por quién fue iniciado un ataque?

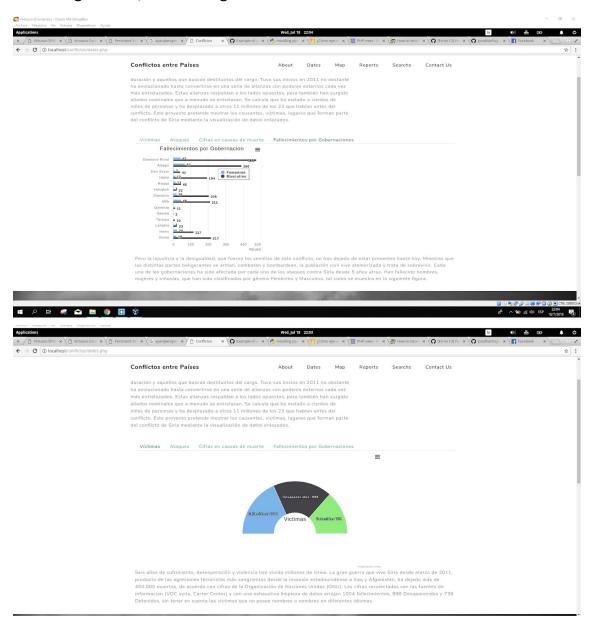
Para esto se pretende construir una APP para describir relaciones entre recursos web referentes al terrorismo en un país. La obtención de datos desde sitios web afines a países que sean víctimas del terrorismo, mediante técnicas de extracción de datos de la web como scrapping. Para lo cual la aplicación web que permita visualizar los datos del conflicto de bélico de manera precisa y ordenada a través de:

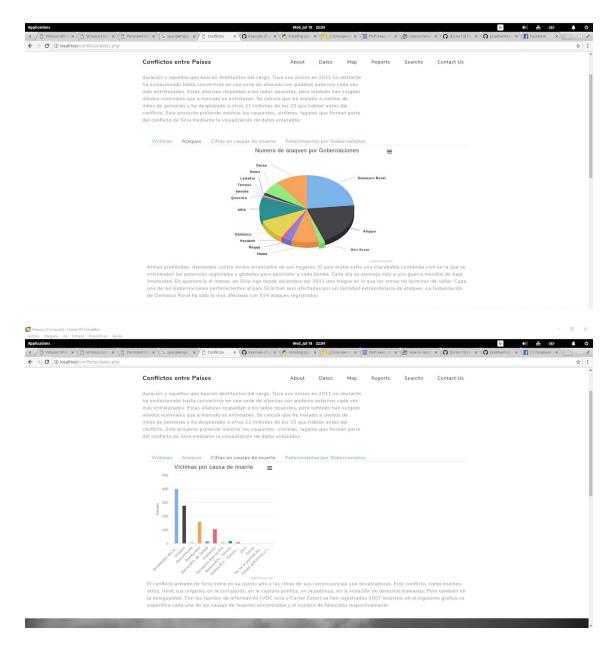


• Mapas representativos: Para representar los ataques sucedidos en cada gobernación, rescatando datos relevantes como fecha de ataque, grupo armado iniciador y grupo armado objetivo, fuente de la noticia, url, tipo de medio de comunicación. Esta información se mostrará a través de un mapa interactivo que indique qué grupos armados controlan determinada gobernación.

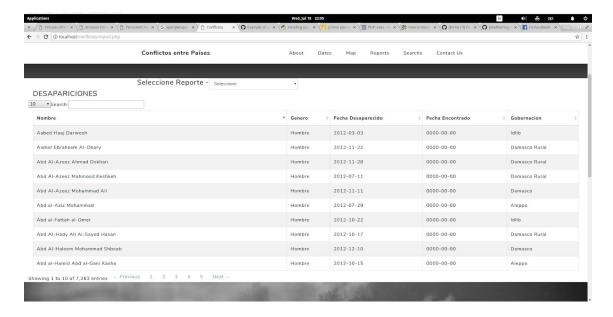


 Gráficas estadísticas: Para representar cifras generales del conflicto mediante gráficos estadísticos, tales como número de defunciones, detenciones y desapariciones, gobernaciones afectadas, ataques registrados, noticias registradas.





 Reportes: Para representar la información referente a las víctimas como nombre, genero, causa de muerte, fecha de muerte, fecha de arresto, número de detenciones, antecedentes penales, fecha de desaparición, fecha de encontrado. Esta información se representará mediante datatables dinámicos.



Resultados

- Modelado de una ontología que representa el conflicto bélico de los Países, rescatando las entidades más principales y tomando como referencia a Siria para este caso de estudio.
- Se extrajo y procesó datos referentes al conflicto que se suscitan en Siria.
- Construcción de una aplicación web que mediante visualización de datos permitiera responder a las preguntas planteadas para la ontología propuesta.

Referencias

- 1. VDC, "VDC Violations Documentation Center in Syria," 2011.16.
- 2. The Carter Center, "The Carter Center Waging Peace. Fighting Disease. BuildingHope.," 2011.
- 3. W. van Atteveldt, Semantic Network Analysis. PhD thesis, Charleston SC, 2008.4.
- 4. R. G. Raskin and M. J. Pan, "Knowledge representation in the semantic web for Earth and environmental terminology (SWEET)," Computers and Geosciences, vol. 31, no. 9, pp. 1119–1125, 2005
- 5. R. C. Fernandez, "Representación del Conocimiento . Web Semántica," Universidad Carlos III de Madrid, 2005

1. ¿En qué país se está dando un conflicto bélico?

Esta consulta no va al caso, ya que solo se toma a Siria como ejemplo.

2. ¿Cuál es la causa de muerte más común?

Consulta

```
select COUNT(*) as ?n ?causa
where
{
     ?n <http://conflictosiria.com/schema/deathCause> ?causa .
}
```

Resultado

n	causa
399	"Bombardeo del avión de guerra"@es
277	"Disparo"@es
8	"Desconocido"@es
158	"Bombardeo"@es
16	"Ejecución de campo"@es
105	"Explosion"@es
10	"Secuestro-Ejecución"@es
17	"Detención - Tortura"@es
8	"Detención - Ejecución"@es
2	"Otro"@es
2	"Cerco"@es
1	"No se le permite buscar ayuda médica"@es
1	"Gases químicos y tóxicos"@es

3. ¿Dónde vivía X victima antes de morir?

Consulta

```
PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/>
select COUNT(*) as ?m ?genero ?nombregov where{
?m <a href="http://xmlns.com/foaf/0.1/gender">http://xmlns.com/foaf/0.1/gender</a> ?genero.
?m <a href="http://conflictosiria.com/schema/livedIn">http://conflictosiria.com/schema/livedIn</a> ?prov.
?prov dbo:province ?nombregov.
}
```

Resultado

m	nombregov		
514	"Damasco Rural"@es		
453	"Aleppo"@es		
44	"Deir Ezzor"@es		
204	"Hama"@es		
57	"Raqqa"@es		
24	"Hasakeh"@es		
223	"Damasco"@es		
253	"Idlib"@es		
11	"Quneitra"@es		
3	"Sweida"@es		
11	"Tartous"@es		
26	"Lattakia"@es		
139	"Homs"@es		
235	"Daraa"@es		

4. ¿Qué fuente reportó X ataque?

```
PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/>
PREFIX geo: <a href="http://www.w3.org/2003/01/geo/wgs84_pos#">http://www.w3.org/2003/01/geo/wgs84_pos#</a>
select ?ataque ?nombrereporte {
    ?ataque <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>
<a href="http://conflictosiria.com/schema/reportedBy">http://conflictosiria.com/schema/reportedBy</a> ?reporte .
    ?reporte <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a> ?tiporeporte <a href="http://xmlns.com/foaf/0.1/name">http://xmlns.com/foaf/0.1/name</a> ?nombrereporte .
}
```

ataque	nombrereporte
http://conflictosiria.com/data/assault859	"SOHR"@es
http://conflictosiria.com/data/assault803	"SOHR"@es
http://conflictosiria.com/data/assault669	"Media"@es
http://conflictosiria.com/data/assault613	"SOHR"@es
http://conflictosiria.com/data/assault479	"Media"@es
http://conflictosiria.com/data/assault423	"Media"@es
http://conflictosiria.com/data/assault289	"Media"@es
http://conflictosiria.com/data/assault233	"SOHR"@es
http://conflictosiria.com/data/assault1033	"Media"@es
http://conflictosiria.com/data/assault975	"Media"@es
http://conflictosiria.com/data/assault785	"SOHR"@es
http://conflictosiria.com/data/assault891	"SOHR"@es
http://conflictosiria.com/data/assault595	"SOHR"@es
http://conflictosiria.com/data/assault302	"SOHR"@es
http://conflictosiria.com/data/assault112	"Twitter"@es
http://conflictosiria.com/data/assault938	"Twitter"@es
http://conflictosiria.com/data/assault748	"Media"@es
http://conflictosiria.com/data/assault854	"SOHR"@es
http://conflictosiria.com/data/assault558	"Twitter"@es
http://conflictosiria.com/data/assault664	"Media"@es
http://conflictosiria.com/data/assault368	"Media"@es
http://conflictosiria.com/data/assault474	"Twitter"@es
http://conflictosiria.com/data/assault178	"SOHR"@es
http://conflictosiria.com/data/assault284	"Media"@es
http://conflictosiria.com/data/assault1025	"SOHR"@es

5. ¿Víctimas por tipo de muerte de X víctima?

n	causa
399	"Bombardeo del avión de guerra"@es
277	"Disparo"@es
8	"Desconocido"@es
158	"Bombardeo"@es
16	"Ejecución de campo"@es
105	"Explosion"@es
10	"Secuestro-Ejecución"@es
17	"Detención - Tortura"@es
8	"Detención - Ejecución"@es
2	"Otro"@es
2 2 1	"Cerco"@es
1	"No se le permite buscar ayuda médica"@es
1	"Gases químicos y tóxicos"@es

6. ¿Qué grupos armados han atacado X gobernación?

```
PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/>
select COUNT(*) as ?n ?nombregov where{
?n <a href="http://conflictosiria.com/schema/livedIn">http://conflictosiria.com/schema/livedIn</a> ?prov.
?prov dbo:province ?nombregov.
}
```

n	nombregov
514	"Damasco Rural"@es
453	"Aleppo"@es
44	"Deir Ezzor"@es
204	"Hama"@es
57	"Raqqa"@es

7. ¿Cuántas personas están desaparecidas hasta el momento?

```
select COUNT(*) as ?count where{
    ?count <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://conflictosiria.com/schema/Disapperance</a>.
}
```



8. ¿Cuántas personas están reportadas fallecidas hasta el momento?

```
select COUNT(*) as ?count where{
     ?count <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://conflictosiria.com/schema/Decease</a>.
}
```

count 1004

9. ¿Cuántas personas están reportadas detenidas hasta el momento?

```
select COUNT(*) as ?count where{
    ?count <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://conflictosiria.com/schema/Detention</a>.
}
```



10. ¿Qué gobernación reporta más ataques?

```
PREFIX dbo: <a href="http://dbpedia.org/ontology/">http://dbpedia.org/ontology/>
select COUNT(*) as ?n ?nombregov where{
?n <a href="http://conflictosiria.com/schema/livedIn">http://conflictosiria.com/schema/livedIn</a> ?prov.
?prov dbo:province ?nombregov.
}
```

