The Semantic Web: Background and Introduction

Gregory L. Heileman

Informatics Research Group
Department of Electrical & Computer Engineering
University of New Mexico

April 19, 2013
The Internet – Web 1.0

How the Internet works:

Client is a browser, server is a web server.
How the Internet works:

- Client is a browser, server is a web server.
- The request/response interchange is mediated via the HTTP protocol.

Most commonly, an HTML document. It’s the hyperlinks in the collection of HTML documents available through the Internet that creates the World Wide Web.
How the Internet works:

- Client is a browser, server is a web server.
- The request/response interchange is mediated via the HTTP protocol.
- The response (in a GET request) involves the delivery of a resource.
How the Internet works:

- Client is a browser, server is a web server.
- The request/response interchange is mediated via the HTTP protocol.
- The response (in a GET request) involves the delivery of a resource. Most commonly, an HTML document.
- It’s the hyperlinks in the collection of HTML documents available through the Internet that creates the World Wide Web.
The Internet – Web 1.0

How the Internet works:

- Client is a browser, server is a web server.
- The request/response interchange is mediated via the HTTP protocol.
- The response (in a GET request) involves the delivery of a resource. Most commonly, an HTML document.
- It’s the hyperlinks in the collection of HTML documents available through the Internet that creates the World Wide Web.
- With Web 1.0 (when you guys were in diapers), static HTML content was delivered.
The Internet – Web 2.0

It’s still client-server:

- The server side is a lot more complicated, and the client side a lot more capable.
The Internet – Web 2.0

It’s still client-server:

- The server side is a lot more complicated, and the client side a lot more capable.
- What makes it Web 2.0?

There’s a lot more interactivity.

Web applications look a lot more like the desktop applications were used to using on our PCs – we can order stuff online, execute transactions with our bank, take online classes, etc.

"Web 2.0 generally refers to a set of social, architectural, and design patterns resulting in the mass migration of business to the Internet as a platform."

—Governor, Hinchcliffe and Nickull, 2009
It’s still client-server:

- The server side is a lot more complicated, and the client side a lot more capable.
- What makes it Web 2.0? There’s a lot more interactivity.

"Web 2.0 generally refers to a set of social, architectural, and design patterns resulting in the mass migration of business to the Internet as a platform."
—Governor, Hinchcliffe and Nickull, 2009
The server side is a lot more complicated, and the client side a lot more capable.

What makes it Web 2.0? There’s a lot more interactivity. Web applications look a lot more like the desktop applications were used to using on our PCs – we can order stuff online, execute transactions with our bank, take online classes, etc.
The Internet – Web 2.0

It’s still client-server:

- The server side is a lot more complicated, and the client side a lot more capable.
- What makes it Web 2.0? There’s a lot more interactivity. Web applications look a lot more like the desktop applications were used to using on our PCs – we can order stuff online, execute transactions with our bank, take online classes, etc.
- “Web 2.0 generally refers to a set of social, architectural, and design patterns resulting in the mass migration of business to the Internet as a platform.”
  —Governor, Hinchcliffe and Nickull, 2009
The Internet – Semantic Web

“A web of data that can be processed directly and indirectly by machines.”
—Tim Berners-Lee

The current web consists of unstructured or semi-structured documents. Conversion to the semantic web involves adding semantic content to web pages. This semantic content allows actors to find, share and combine information more easily.

Humans can easily use the Web to search for and order the lowest priced book on a given topic — machines cannot.

“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web — the content, links, and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The ‘intelligent agents’ people have touted for ages will finally materialize.”
—Tim Berners-Lee

UNM Informatics
The Semantic Web: Background and Introduction
The Internet – Semantic Web

- “A web of data that can be processed directly and indirectly by machines.”
  —Tim Berners-Lee
- The current web consists of unstructured or semi-structured documents. Conversion to the semantic web involves adding semantic content to web pages.
“The Internet – Semantic Web

- “A web of data that can be processed directly and indirectly by machines.”
  —Tim Berners-Lee
- The current web consists of unstructured or semi-structured documents. Conversion to the semantic web involves adding semantic content to web pages.
- This semantic content allows actors to find, share and combine information more easily.
“A web of data that can be processed directly and indirectly by machines.”
—Tim Berners-Lee

The current web consists of unstructured or semi-structured documents. Conversion to the semantic web involves adding semantic content to web pages.

This semantic content allows actors to find, share and combine information more easily. Humans can easily use the Web to search for and order the lowest priced book on a given topic — machines cannot.
The Internet – Semantic Web

“
A web of data that can be processed directly and indirectly by machines.”
— Tim Berners-Lee

The current web consists of unstructured or semi-structured documents. Conversion to the semantic web involves adding semantic content to web pages.

This semantic content allows actors to find, share and combine information more easily. Humans can easily use the Web to search for and order the lowest priced book on a given topic — machines cannot.

“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web — the content, links, and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The ‘intelligent agents’ people have touted for ages will finally materialize.”
— Tim Berners-Lee
The concept of a *semantic network model* as a means of representing semantically structured knowledge was developed in the early 1960’s by cognitive scientists and psychologists.
The concept of a *semantic network model* as a means of representing semantically structured knowledge was developed in the early 1960's by cognitive scientists and psychologists.

In a semantic network model, knowledge representation is stored in the form of semantic relations — a graph, where the nodes are concepts, and edges are relations between them.
Tim Berners-Lee coined the usage of this term in the WWW.
The Semantic Web – History

- Tim Berners-Lee coined the usage of this term in the WWW.
- His vision was a set of technologies that extends the network of hyperlinked human-readable web pages (the WWW) by inserting machine-readable metadata about pages and how they are related to each other, enabling automated agents to access the Web more intelligently and perform tasks on behalf of users.
The Semantic Web – Terms

- **Resource Description Framework (RDF)** – a family of W3C specifications for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

- **RDF Data Model** – a specification for making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as triples. RDF expressions are often collected in triple stores.

- **SPARQL (SPARQL Protocol and RDF Query Language)** – an RDF Query language.

- **RDF Schema (RDFS)** – An ontology defined using the RDF data model – a set of classes with defined properties. I.e., it's a language for RDF vocabulary sharing.

- **OWL (Web Ontology Language)** – A family of knowledge representation languages, endorsed by the W3C, for authoring ontologies.
Resource Description Framework (RDF) – a family of W3C specifications for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

RDF Data Model – a specification for making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as *triples*. RDF expressions are often collected in *triple stores*.
The Semantic Web – Terms

- **Resource Description Framework (RDF)** – a family of W3C specifications for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

- **RDF Data Model** – a specification for making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as *triples*. RDF expressions are often collected in *triple stores*.

- **SPARQL (SPARQL Protocol and RDF Query Language)** – an RDF Query language.
The Semantic Web – Terms

- **Resource Description Framework (RDF)** – a family of W3C specifications for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

- **RDF Data Model** – a specification for making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as *triples*. RDF expressions are often collected in *triple stores*.

- **SPARQL (SPARQL Protocol and RDF Query Language)** – an RDF Query language.

- **RDF Schema (RDFS)** – An ontology defined using the RDF data model – a set of classes with defined properties. I.e., it’s a language for RDF vocabulary sharing.
The Semantic Web – Terms

- **Resource Description Framework (RDF)** – a family of W3C specifications for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

- **RDF Data Model** – a specification for making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as *triples*. RDF expressions are often collected in *triple stores*.

- **SPARQL (SPARQL Protocol and RDF Query Language)** – an RDF Query language.

- **RDF Schema (RDFS)** – An ontology defined using the RDF data model – a set of classes with defined properties. I.e., it’s a language for RDF vocabulary sharing.

- **OWL (Web Ontology Language)** – A family of knowledge representation languages, endorsed by the W3C, for authoring ontologies.
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.

- SQL was developed before the advent of the Internet, SPARQL after.

- SQL queries operate over a given database, there is no standard mechanism for query federation.

- SPARQL supports the ability to federate queries across different repositories.

- SPARQL can be used to access relational data, as well as RDF data.

- A SPARQL endpoint is a web service that accepts SPARQL queries, e.g., DBPedia.

- DBPedia has a web form where you can submit queries, and results are returned in HTML:
  
  http://dbpedia.org/snorql/
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, SPARQL after.
- SPARQL supports the ability to federate queries across different repositories.
- SPARQL can be used to access relational data, as well as RDF data.
- A SPARQL endpoint is a web service that accepts SPARQL queries, e.g., DBPedia.
- DBPedia has a web form where you can submit queries, and results are returned in HTML: http://dbpedia.org/snorql/
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, SPARQL after.
- SQL queries operate over a given database, there is no standard mechanism for query federation.
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, SPARQL after.
- SQL queries operate over a given database, there is no standard mechanism for query federation.
- SPARQL supports the ability to federate queries across different repositories.

SPARQL can be used to access relational data, as well as RDF data. A SPARQL endpoint is a web service that accepts SPARQL queries, e.g., DBPedia. DBPedia has a web form where you can submit queries, and results are returned in HTML: http://dbpedia.org/snorql/
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, SPARQL after.
- SQL queries operate over a given database, there is no standard mechanism for query federation.
- SPARQL supports the ability to federate queries across different repositories.
- SPARQL can be used to access relational data, as well as RDF data.

A SPARQL endpoint is a web service that accepts SPARQL queries, e.g., DBPedia. DBPedia has a web form where you can submit queries, and results are returned in HTML: [http://dbpedia.org/snorql/](http://dbpedia.org/snorql/)
SQL vs. SPARQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), while SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, whereas SPARQL was developed after.
- SQL queries operate over a given database, with no standard mechanism for query federation.
- SPARQL supports the ability to federate queries across different repositories.
- SPARQL can be used to access relational data, as well as RDF data.
- A **SPARQL endpoint** is a web service that accepts SPARQL queries, e.g., DBpedia.

**Example:**

DBPedia has a web form where you can submit queries, and results are returned in HTML:

http://dbpedia.org/snorql/
SPARQL vs. SQL

- SQL is used for querying relational data (rows of data collected into tables, along with relationships between them), SPARQL is for querying RDF data.
- SQL was developed before the advent of the Internet, SPARQL after.
- SQL queries operate over a given database, there is no standard mechanism for query federation.
- SPARQL supports the ability to federate queries across different repositories.
- SPARQL can be used to access relational data, as well as RDF data.
- A **SPARQL endpoint** is a web service that accepts SPARQL queries, e.g., DBPedia.
- DBPedia has a web form where you can submit queries, and results are returned in HTML: [http://dbpedia.org/snorql/](http://dbpedia.org/snorql/)
Sample SPARQL Queries

All of the states in the US:

```
SELECT ?state WHERE {
}
```
NBA players who played for Wake Forest:

PREFIX p: <http://dbpedia.org/property/>
SELECT ?player ?team

WHERE {
    p:team ?team .
}
User Applications

A number of applications using DBpedia:

- **SemLens** – uses scatter plots for the analysis of DBpedia data and semantic lenses for further exploration.
User Applications

A number of applications using DBpedia:

- **SemLens** – uses scatter plots for the analysis of DBpedia data and semantic lenses for further exploration.
- **RelFinder** – enter objects and find out what connects them.
User Applications

A number of applications using DBpedia:

- **SemLens** – uses scatter plots for the analysis of DBpedia data and semantic lenses for further exploration.
- **RelFinder** – enter objects and find out what connects them.
- **DayLikeToday** presents facts from DBpedia scoped to a given day and presented via a timeline control.
Ex: Retrieve all the universities in Lebanon with their respective locations.
Ex: Retrieve all the universities in Lebanon with their respective locations.

Analysis:
Most of the universities in Lebanon are located in Beirut city.
Ex: Retrieve relations between Albert Einstein and Kurt Gödel
Ex: Retrieve events happened on May 16

Kengir uprising (War)

The Kengir uprising was a prisoner uprising that took place in the Soviet prison labor camp Kengir in May and June 1954. Its duration and intensity distinguished it from other Gulag uprisings in the same period. After the murder of some of their fellow prisoners by guards, Kengir inmates launched a rebellion and proceeded to seize the entire camp compound, holding it for weeks and creating a period of freedom for themselves unique in the history of the Gulag.

Read more...
Take a look at DBpedia...
References
