

Curso: Representación y organización de información científica en la web

Año lectivo: **2017**

Profesor a cargo: Dr. Pablo Fillottrani

Carga horaria: 30 hs.

Modalidad: A distancia

Fechas: 7 de marzo al 10 de abril de 2017

1. Fundamentación y objetivos

El curso pretende, en primer lugar, proporcionar conocimientos teóricos sobre la gestión semántica de información en Internet, y en segundo lugar, introducir a los alumnos en las más recientes tecnologías y herramientas desarrolladas para compartir e integrar coherentemente información en la web.

La interoperabilidad semántica de información puede definirse como la posibilidad para un sistema de información de funcionar coherentemente con otros sistemas, intercambiando información sin la necesidad de fijar restricciones a clientes y usuarios. Esta interoperabilidad está basada en el uso de ontologías, lo que permite el intercambio fluido de información sin pérdida de significado, la cooperación interna o externa, la composición de sistemas, etc., por lo que es un aspecto importante para el éxito de un sistema de información moderno. En su máximo nivel, la interoperabilidad se instrumenta a través del concepto de la Web Semántica. El empleo de herramientas de la Web Semántica permite desarrollar competencias en el acceso y el compartir información, el tratamiento, el dominio de diferentes estándares de representación y de consulta.

El objetivo del seminario es presentar el actual estado del arte en el campo de las Ontologías y Web Semántica, en particular aplicadas a repositorios de información científica y tecnológica, y explorar posibles perspectivas para el futuro.

2. Contenidos

Clase 1: Introducción a la Web Semántica

Objetivos. Antecedentes. Arquitectura de la Web. Lenguajes de marcado.

Normas y lenguajes básicos para representar información en la Web: UNICODE, URI, XML. Definición, ejemplos, usos. Espacios de nombres. XML Schema. Herramientas. Lenguaje de consulta: XQuery. Aplicación en bases de datos semiestructuradas.

Clase 2: Lenguaje simple para modelar datos compartidos: RDF

RDF: objetivos, sintaxis, introducción a la semántica. RDF Schema. Problemas. Herramientas existentes.

Lenguaje de consulta para RDF: SPARQL.

Normas "Linked Data": Objetivos. Alcance actual. Protocolos.

Herramientas. Ejemplo de aplicaciones.

Clase 3: Ontologías. Lógicas para la Descripción

Fundamentos. Propiedades. Ontologías como herramienta para Representación del Conocimiento. Representación formal: lógicas más usadas. Lógicas para la Descripción: definición, propiedades. Lógicas para la descripción y bases de datos.

Lenguaje para ontologías en la web: OWL. Formalización de OWL. Fragmentos (profiles). Propiedades. Problemas. Relación con diagramas ER, UML y lógica clásica. Estándares relacionados: MOF.

Clase 4: Aplicaciones en repositorios de información científica. Casos de estudio.

Herramientas para el diseño de ontologías.

Aplicaciones y herramientas para diseño conceptual, acceso a la información e integración de ontologías. Ejemplos de ontologías aplicadas a repositorios.

3. Metodología de trabajo y sistema de evaluación

El aula virtual será el espacio de enseñanza y aprendizaje en el cual se desarrollarán las clases. Los alumnos trabajarán en la resolución de ejercicios simples, y de ejercicios más complicados con la ayuda del profesor, utilizando la bibliografía recomendada. Se presentarán los temas del contenido con el nivel suficiente para que los alumnos comprendan la problemática y conozcan las posibles soluciones pero no se pretenderá el conocimiento detallado de cada una de ella ni el desarrollo de sistemas computacionales que las utilicen.

A continuación se detallan algunas actividades prácticas:

1- Representación de información semiestructurada en XML y vocabularios estándares para producción bibliográfica (Dublin Core Metadata). Uso de herramientas online para la edición y generación de metadatos en estos vocabularios. 2- Publicación de metadatos usando RDFa. Editores online. Vocabularios de citación. Descubrimiento de información publicada en la Web en este formato. 3. Publicación y consumo de metadatos en SPARQL endpoints. Uso y evaluación de las herramientas D2R Server. OpenLink Virtuoso y Triplify. 4- Integración de información bibliográfica en Mashups. Uso de Amazon API y Google Books API en un mashup de información bibliográfica. Se analizará la posibilidad de desarrollo de este tipo de herramientas en distintos contextos. 5- Elaboración de un informe sobre pautas de evaluación, y comparación de ontologías para la publicación de documentos. Se desarrollarán criterios y métricas para analizar las ontologías, se buscará sus versiones actuales y se evaluará su uso, aplicabilidad y extensibilidad.

Además de la realización y aprobación de las actividades por módulo, la aprobación del curso requiere la aprobación de un trabajo final. Los plazos de entrega y calificación seguirán las pautas generales de la carrera.

4. Bibliografía

Obligatoria

- Allemang, Dean; Hendler, James; Kaufmann, Morgan. (2011). *Semantic Web for the Working Ontologist, Second Edition: Effective Modeling in RDFS and OWL*, 2nd edition. ISBN-13: 978-0123859655
- Wood, David; Zaidman, Marsha; Ruth, Luke. (2014). *Linked Data*. Manning Publications. ISBN-13: 978-1617290398

- Antoniou, Grigoris Groth, Paul; van Harmelen, Frank. (2012). *A Semantic Web Primer*. 3rd edition. The MIT Press;. ISBN-13: 978-0262018289

Complementaria

- Szeredi, Péter; Lukácsy, Gergely; Benkő, Tamás. (2014). *The Semantic Web Explained: The Technology and Mathematics behind Web 3.0*. Cambridge University Press. ISBN-13: 978-0521700368
- Toby Segaran, Colin Evans, Jamie Taylor. (2009). *Programming the Semantic Web*. O'Reilly Media. ISBN-13: 978-0596153816
- Yu, Liyang . (2015). *A Developer's Guide to the Semantic Web*. 3rd edition. Springer. ISBN-13: 978-3662437957
- Hebel, John ; Fisher, Matthew; Blace, Ryan ; Perez-Lopez, Andrew. (2009). *Semantic Web Programming*. Wiley ISBN-13: 978-0470418017
- Sheth, Amit. (2013). *Semantic Web: Ontology and Knowledge Base Enabled Tools, Services, and Applications*. IGI Global. ISBN-13: 978-1466636101
- Davies, John ; Studer, Rudi; Warren, Paul. (2006). *Semantic Web Technologies: Trends and Research in Ontology-based Systems*. Wiley. ISBN-13: 978-0470025963
- Hitzler, Pascal ; Krötzsch, Markus ; Rudolph, Sebastian . (2009). *Foundations of Semantic Web Technologies*. Chapman and Hall/CRC . ISBN-13: 978-1420090505
- "Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata", Seth van Hooland, Ruben Verborgh. Amer Library Assn Editions (2014). ISBN-13: 978-0838912515
- Heath, Tom; Bizer, Christian ; Hendler, James. (2011). *Linked Data*. Morgan & Claypool Publishers. ISBN-13: 978-1608454303
- Hart, Glen; Dolbear; Catherine. (2013). *Linked Data: A Geographic Perspective*. CRC Press. ISBN-13: 978-1439869956
- Coyle, Karen. (2012). *Linked Data Tools: Connecting on the Web*. Amer Library Assn. ISBN-13: 978-0838958599
- Harth , Andreas; Hose, Katja; Schenkel, Ralf. (2014). *Linked Data Management*. Chapman and Hall/CRC. ISBN-13: 978-1466582408
- Konstantinou, Nikolaos; Emmanuel, Dimitrios. (2015). *Materializing the Web of Linked Data*. Springer. ISBN-13: 978-3319160733

- Hyvönen, Eero. (2012). *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. Morgan & Claypool Publishers. ISBN-13: 978-1608459971
- Erik T. Mitchell. (2013). *Library Linked Data: Research and Adoption*. Amer Library Assn. ISBN-13: 978-0838958964
- King, Brandy ; Reinold, Kathy. (2008). *Finding the Concept, Not Just the Word: A Librarian's Guide to Ontologies and Semantics*. Chandos Publishing . ISBN-13: 978-1843343189
- Akerkar, R. (2009). *Foundations of the Semantic Web: XML, RDF & Ontology*. Alpha Science Intl Ltd. ISBN-13: 978-1842655351
- Hepp, Martin ; Leenheer, Pieter de; Moor, Aldo de; Sure, York. (2007). *Ontology Management: Semantic Web, Semantic Web Services, and Business Applications*. Springer. ISBN-13: 978-0387698991
- Daconta, Michael; Obrst, Leo; Smith, Kevin. (2003). *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*. Wiley. ISBN 0-471-43257-1.
- Fensel, Dieter; Wahlster, Wolfgang; Lieberman, Henry; Hendler, James. (2002). *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press. ISBN 0-262-06232-1.
- Staab, Steffen; Studer, Rudi editors. (2004). *Handbook on Ontologies*. Heidelberg-Springer ISBN 3-540-40834-7.
- Berners-Lee, Tim ; Hendler, James; Lassila, Ora. (2001). *The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities"*. Scientific American. May.

Sitios web

- URI clarification. <http://www.w3.org/TR/uri-clarification/>
- URI schemes <http://www.iana.org/assignments/uri-schemes>
- OWL Web Ontology Language Overview. <http://www.w3.org/TR/owl-features/>
- World Wide Web Consortium RDF homepage. <http://www.w3.org/RDF/>
- RDF Primer. <http://www.w3.org/TR/rdf-primer/>
- World Wide Web Consortium XML homepage. <http://www.w3.org/XML/>