

PROGRAMA DE CURSO

Código	Nombre			
CC6202	La Web de Datos			
Nombre en Inglés				
Linked Data				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3	-	7
Requisitos			Carácter del Curso	
CC3201 Bases de Datos			Electivo	
Resultados de Aprendizaje				
Al final del curso el estudiante conocerá los fundamentos de la revolución en la creación, manejo, procesamiento y consumo de datos, producto de la aparición de la Web y los avances en las capacidades del hardware. Asimismo se discutirá sobre el rol que los datos juegan en la sociedad y los nuevos desafíos técnicos. El alumno conocerá varias herramientas para manipular datos, modelarlos y visualizarlos, que le permitirán enfrentarse a grandes volúmenes de datos y crear aplicaciones novedosas sobre ellos.				

Metodología Docente	Evaluación General
El curso se desarrollara con clases presenciales, lecturas, presentaciones de alumnos y desarrollo de un proyecto de investigación (ya sea individual o grupal).	La evaluación constara de: 1. Lecturas 2. Informes escritos 3. Presentaciones 4. Proyecto final

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas	
1	El Diluvio de Datos	2	
Contenidos		Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
1. Tendencias en la generación, procesamiento, transformación, uso y archivo de datos. 2. Casos de estudio: gobierno, datos científicos.		El alumno deberá entender: * La revolución que se ha generado a comienzos de siglo XX en la producción de datos. * El rol que los datos juegan en el sistema de información actual (en la Web de datos). * Los desafíos que estos niveles de datos generan para la ciencia de la computación.	

Número	Nombre de la Unidad	Duración en Semanas	
2	Técnicas y Modelos	4	
Contenidos		Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
1. Modelos 2. Técnicas		El alumno deberá conocer: * Los modelos de datos básicos propuestos para conceptualizar la Web de datos y sus facetas. * El alumno deberá manejar las técnicas básicas usadas para desarrollar en la Web de datos	

Número	Nombre de la Unidad	Duración en Semanas
3	Áreas Particulares	9
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
Elegir un tema o área particular y desarrollar resultados teóricos o aplicativos.	El alumno deberá ser capaz de trabajar, en uno de los dos áreas: * Teórica: Ser capaz de entender modelos, de desarrollar funcionalidades, de detectar problemas de la Web de datos. * Aplicación: Ser capaz de diseñar (y si el tiempo permite, desarrollar) una aplicación que responda y resuelva problemas de la Web de datos detectados.	

Bibliografía
[1] R. Angles, C. Gutierrez, Survey of Graph Database Models, ACM Computing Surveys, Vol. 40, No. 1, February 2008.
[2] R. Agrawal et al., The Claremont Report on Database Research, 2008. http://db.cs.berkeley.edu/claremont/
[3] S. Abiteboul, V. Vianu, Queries and Computation on the Web, Theor. Comput. Sci. 239(2): 231-255 (2000).
[4] S. Abiteboul, P. Buneman, D. Suciu, Data on the Web. From Relations to Semistructured Data and XML, Morgan Kaufmann Publ. California, 2000.
[5] L. G. Alex Sung, N. Ahmed, R. Blanco, H. Li, M. Ali Soliman, D. Hadaller, A Survey of Data Management in Peer-to-Peer Systems}, Web Data Management, 2005.

Bibliografía

- [6] Tim Berners-Lee, Design Issues/Linked Data,
<http://www.w3.org/DesignIssues/LinkedData.html>
- [7] P. A. Bernstein, F. Giunchiglia, A. Kementsietsidis, J. Mylopoulos, L. Serafini, I. Zaihrayeu, Data Management for Peer-to-Peer Computing: A Vision, WebDB, Workshop on Databases and the Web, 2002.
- [8] G. Bell, J. Gray, A. Szalay, Petascale Computational Systems: Balanced CyberInfrastructure in a Data-Centric World, Computer, Volume 39, Issue 1 (January 2006), 110 - 112.
- [9] Ch. Bizer, T. Heath, T. Berners-Lee, Linked Data - The Story So Far, International Journal on Semantic Web and Information Systems, 3 (2009), pp. 1-22.
- [10] P. Buneman, Semistructured data, ACM PODS, 1997.
- [11] DATA.gov project, <http://www.data.gov/>
- [12] T. Green, V. Tannen, Models for Incomplete and Probabilistic Information, EDBT Workshops, Munich, Germany, March 2006.
- [13] S. Gribble, A. Halevy, Z. Ives, M. Rodrig, D. Suciu, What Can Databases Do for Peer-to-Peer?, WebDB, Workshop on Databases and the Web, 2001.
- [14] Alon Y. Halevy, M. J. Franklin, D. Maier, Principles of dataspace systems, PODS 2006: 1-9.
- [15] M. Hausenblas, M. Karnstedt, Understanding Linked Open Data as a Web-Scale Database, 1st Internat. Conf. on Advances in Databases, pp. 56-61, 2010.
- [16] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tompkins, E. Upfal, The Web as a Graph, PODS 2000, pp. 1-10.
- [17] Th. Lee, Attribution Principles for Data Integration: Technology and Policy Perspectives, Thesis (Ph. D.)--Massachusetts Institute of Technology, Engineering Systems Division, Technology, Management, and Policy Program, 2002.
- [18] LinkedData Project, <http://www.linkeddata.org>

Bibliografía

- [19] A. O. Mendelzon, The Web is not a Database, Workshop on Web Information and Data Management 1998.
- [20] A. O. Mendelzon, T. Milo, Formal Models of Web Queries, PODS 1997: 134-143.
- [21] A. O. Mendelzon, T. Milo, Formal Models of Web Queries, Inf. Syst. 23(8): 615-637 (1998).
- [22] A. Rubinstein, Modeling Bounded Rationality, MIT Press, 1998.
- [23] No SQL, <http://nosql-database.org/>
- [24] M. Spielmann, J. Tyszkiewicz, J. Van den Bussche, Distributed computation of web queries using automata, PODS 2002, pp. 97-108.
- [25] M. Stonebraker, S. Madden, D. J. Abadi, S. Harizopoulos, N. Hachem, and P. Helland, The end of an architectural era: (it's time for a complete rewrite), VLDB '07, 2007, pp. 1150-1160.
- [26] Tim O'Reilly, What Is Web 2.0}, <http://oreilly.com/web2/archive/what-is-web-20.html>

Vigencia desde:	Otoño 2014
Elaborado por	Claudio Gutiérrez