

Formulario de Aprobación Curso de Actualización

Asignatura: Calidad de Datos desde “small data” hacia “big data”: desafíos y técnicas
(Si el nombre contiene siglas deberán ser aclaradas)

Profesor de la asignatura¹: Dra. Cinzia Cappiello, Politecnico di Milano, Milán, Italia
(título, nombre, grado o cargo, Instituto o Institución)

Profesor Responsable Local¹: Dra. Adriana Marotta, Gr. 4, Instituto de Computación
(título, nombre, grado, Instituto)

Otros docentes de la Facultad:
(título, nombre, grado, Instituto)

Docentes fuera de Facultad:
(título, nombre, cargo, Institución, país)

Instituto ó Unidad: Instituto de Computación
Departamento ó Area: Programación

¹ Agregar CV si el curso se dicta por primera vez.
(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

Horas Presenciales: 17
(se deberán discriminar las mismas en el ítem Metodología de enseñanza)

Público objetivo y Cupos: El público objetivo son profesionales en informática que estén interesados en el área de Calidad de Datos. No tiene cupo.

Objetivos: El objetivo del curso es presentar a los estudiantes conceptos básicos sobre calidad de datos y luego los avances más recientes en investigación cuando se trasladan los problemas tradicionales de calidad de datos al contexto de “big data”, donde aparecen las características de volumen, variedad y velocidad.

Conocimientos previos exigidos: Bases de datos relacionales.

Conocimientos previos recomendados: Nociones sobre bases de datos no relacionales.

Metodología de enseñanza: El curso se desarrollará a través de 4 clases teórico-prácticas donde se introducirán los principales conceptos y luego los estudiantes tendrán la posibilidad de hacer un trabajo para obtener los créditos del curso. Durante el tiempo de desarrollo del trabajo la profesora estará disponible para contestar consultas de los estudiantes, la primer semana en forma presencial y luego via correo electrónico.

- Horas clase (teórico): 12
- Horas clase (práctico):
- Horas clase (laboratorio):
- Horas consulta: 5
- Horas evaluación:

-
- Subtotal horas presenciales: 17
 - Horas estudio: 10
 - Horas resolución ejercicios/prácticos:
 - Horas proyecto final/monografía: 20
 - Total de horas de dedicación del estudiante: 47

Forma de evaluación: A través de la entrega de un trabajo individual

Temario:

- Definición de Calidad de Datos
- Dimensiones de Calidad de Datos
- Métodos de evaluación de Calidad de Datos
- Métodos de mejora de Calidad de Datos
- Introducción a "Big Data"
- Calidad de Datos en "Big Data": cómo lidiar con volumen, variedad y velocidad

Bibliografía:

(título del libro-nombre del autor-editorial-ISBN-fecha de edición)

Carlo Batini, Anisa Rula, Monica Scannapieco, and Gianluigi Viscusi. 2015. From Data Quality to Big Data Quality. J. Database Manag. 26, 1 (2015), 60–82. DOI:<http://dx.doi.org/10.4018/JDM.2015010103>

Carlo Batini and Monica Scannapieco. 2016. Data and Information Quality - Dimensions, Principles and Techniques. Springer. DOI:<http://dx.doi.org/10.1007/978-3-319-24106-7>

Laure Berti-Equille and Javier Borge-Holthoefer. 2015. Veracity of Data: From Truth Discovery Computation Algorithms to Models of Misinformation Dynamics. Morgan & Claypool Publishers. DOI:<http://dx.doi.org/10.2200/S00676ED1V01Y201509DTM042>

Elisa Bertino. 2015. Data Trustworthiness—Approaches and Research Challenges. Springer Int. Publishing, Cham, 17–25.

M. Bovee, R.P. Srivastava, and B.R. Mak. September 2001. A Conceptual Framework and Belief-Function Approach to Assessing Overall Information Quality. In Proceedings of the 6th International Conference on Information Quality. Boston, MA.

Li Cai and Yangyong Zhu. 2016. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data Science Journal 14 (2016), 2.

Cinzia Cappiello and Fabio A. Schreiber. 2009. Quality and Energy-Aware Data Compression by Aggregation in WSN Data Streams. IEEE International Conference on Pervasive Computing and Communications - Workshops (PerCom Workshops 2009), 9-13 March 2009, Galveston, TX, USA. IEEE Computer Society, 1–6. DOI:<http://dx.doi.org/10.1109/PERCOM.2009.4912866>

Xin Luna Dong and Divesh Srivastava. 2015. Big Data Integration. Morgan & Claypool Publishers. DOI:<http://dx.doi.org/10.2200/S00578ED1V01Y201404DTM040>

Jorge Merino, Ismael Caballero, Bibiano Rivas, Manuel A. Serrano, and Mario Piattini. 2016. A Data Quality in Use model for Big Data. Future Generation Comp. Syst. 63 (2016), 123–130. DOI:<http://dx.doi.org/10.1016/j.future.2015.11.024>

Richard Y. Wang and Diane M. Strong. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. J. of Management Information Systems 12, 4 (1996), 5–33. <http://www.jmis-web.org/articles/> 1002

Wenlu Yang, A. Da Silva, and M. L. Picard. 2015. Computing data quality indicators on Big Data streams using a CEP. In Computational Intelligence for Multimedia Understanding (IWCIM), 2015 International Workshop on. 1–5. DOI:<http://dx.doi.org/10.1109/IWCIM.2015.7347061>



Facultad de Ingeniería Comisión Académica de Posgrado

Datos del curso

Fecha de inicio y finalización: 29/5 al 1/6

Horario y Salón: De lunes a jueves de 9:00 a 12:00 hs. Salón Marrón

Arancel: Sin costo
