

Formulario de aprobación de curso de posgrado/educación permanente

Asignatura: Calidad de Datos e Información

(Si el nombre contiene siglas deberán ser aclaradas)

Modalidad:

(posgrado, educación permanente o ambas)

Posgrado

Educación permanente

Profesor de la asignatura ¹: Dra. Adriana Marotta, grado 4, Instituto de Computación

(título, nombre, grado o cargo, instituto o institución)

Profesor Responsable Local ¹:

(título, nombre, grado, instituto)

Otros docentes de la Facultad: Dra. Lorena Etcheverry, grado 3, Instituto de Computación, Ing. Camila

Sanz, grado 2, Instituto de Computación, Ing. Sebastián García, grado 2, Instituto de Computación

(título, nombre, grado, instituto)

Docentes fuera de Facultad:

(título, nombre, cargo, institución, país)

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)

[Si es curso de posgrado]

Programa(s) de posgrado: Maestría en Informática (Pediciba), Doctorado en Informática (Pediciba),
Maestría en Ciencia de Datos y Aprendizaje Automático

Instituto o unidad: Instituto de Computación

Departamento o área: Grupo Gestión, Modelado y Análisis de Datos (Gema)

Horas Presenciales: 45

(se deberán discriminar las horas en el ítem Metodología de enseñanza)

Nº de Créditos: 8

[Exclusivamente para curso de posgrado]

(de acuerdo a la definición de la UdelaR, un crédito equivale a 15 horas de dedicación del estudiante según se detalla en el ítem Metodología de enseñanza)

Público objetivo: Estudiantes de posgrado interesados en el área de calidad de datos.

Cupos: máximo 10, selección por sorteo

(si corresponde, se indicará el número de plazas, mínimo y máximo y los criterios de selección. Asimismo, se adjuntará en nota aparte los fundamentos de los cupos propuestos. Si no existe indicación particular para el cupo máximo, el criterio general será el orden de inscripción, hasta completar el cupo asignado)

Objetivos: Calidad de datos es una temática de creciente relevancia, dada la gran importancia que han cobrado los datos para cualquier actividad y el incremento del volumen y disponibilidad de estos. En

particular, el análisis de datos orientado a la toma de decisiones es una tarea cada vez más requerida y para que ésta pueda ser exitosa la consideración de la calidad de los datos constituye una condición fundamental.

El objetivo principal de esta unidad curricular es brindar al estudiante conocimientos básicos sobre calidad de datos. Se espera que el estudiante conozca los principales fundamentos y técnicas existentes para poder abordar la problemática de la calidad de los datos y de la información, tanto en las organizaciones como en tareas de la vida cotidiana de las personas. Además, se espera que conozca en profundidad algunos de los avances científicos más importantes en el área.

Conocimientos previos exigidos: Conocimientos de Bases de Datos y Programación.

Conocimientos previos recomendados:

Metodología de enseñanza:

(comprende una descripción de la metodología de enseñanza y de las horas dedicadas por el estudiante a la asignatura, distribuidas en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

Descripción de la metodología: Se dictarán clases teóricas de exposición de los conceptos básicos del área. Los estudiantes deberán realizar dos tareas, en las cuales serán guiados por los docentes a través de clases prácticas y de monitoreo. La primer tarea será una tarea práctica de laboratorio donde se aplicarán los conceptos aprendidos a un caso de estudio. La segunda tarea consistirá en el estudio y análisis de un tema específico del área de calidad de datos.

Además, los estudiantes realizarán presentaciones orales de las tareas realizadas.

Detalle de horas:

- Horas de clase (teórico): 20
- Horas de clase (práctico):
- Horas de clase (laboratorio):
- Horas de consulta: 20
- Horas de evaluación: 5
 - Subtotal de horas presenciales: 45
- Horas de estudio: 20
- Horas de resolución de ejercicios/prácticos:
- Horas proyecto final/monografía: 60
 - Total de horas de dedicación del estudiante: 125

Forma de evaluación: La evaluación se realizará a partir de las tareas realizadas, en base a los informes y presentaciones/defensas. La instancia de defensa final será individual.

Temario:

- Introducción
 - Concepto de Calidad de Datos
 - Problemas y consecuencias de la mala calidad
 - Necesidades en distintas áreas de aplicación
 - Iniciativas académicas e industriales

- Dimensiones de calidad
 - Multi-dimensionalidad de la calidad
 - Estudio de dimensiones y métricas de calidad
 - Relaciones entre dimensiones

- Técnicas de calidad de datos
 - Modelos de calidad de datos
 - Evaluación de calidad de datos
 - Limpieza de datos, reparación e imputación de datos

- Gestión de calidad de datos
 - Metodologías de gestión de calidad de datos
 - Contexto en calidad de datos

Bibliografía:

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

- Data and Information Quality. Carlo Batini, Monica Scannapieco. Springer. ISBN: 978-3-319-24104-3. 2016.

 - Big Data Integration. Synthesis Lectures on Data Management. Xin Luna Dong, Divesh Srivastava. Morgan & Claypool Publishers, 2015. ISBN: 1627052240, 9781627052245
-

Datos del curso

Fecha de inicio y finalización: Agosto

Horario y Salón: A definir

Arancel:

[Si la modalidad no corresponde indique "no corresponde". Si el curso contempla otorgar becas, indíquelo]

Arancel para estudiantes inscriptos en la modalidad posgrado:

Arancel para estudiantes inscriptos en la modalidad educación permanente:
