

---

**Curso Académico: ( 2022 / 2023 )****Fecha de revisión: 18-05-2022**

---

**Departamento asignado a la asignatura: Departamento de Informática****Coordinador/a: GENOVA FUSTER, GONZALO****Tipo: Obligatoria Créditos ECTS : 3.0****Curso : 1 Cuatrimestre : 2**

---

**REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)**

Programación de Datos (19138)

**OBJETIVOS**

Competencias Básicas:

- Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
- Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
- Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
- Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Competencias Generales:

- Capacidad para identificar, definir y formular problemas de las ciencias sociales y resolverlos mediante técnicas computacionales. Esta capacidad incluye la valoración simultánea de todos los factores en juego, no sólo técnicos, sino también legales.
- Capacidad de recopilar y analizar los conocimientos existentes en las diferentes áreas de las ciencias sociales computacionales y de hacer una propuesta de posibles soluciones a los problemas planteados.
- Capacidad para abordar de manera proactiva los problemas planteados bajo entornos nuevos o poco conocidos, dentro del contexto de las ciencias sociales computacionales.

Competencias Específicas:

- Capacidad de utilizar con nivel avanzado las herramientas computacionales específicas de las ciencias sociales computacionales.
- Capacidad de obtener, preparar, integrar y usar información procedente de fuentes secundarias y de la web.

Resultados del Aprendizaje:

- Conocimiento de las herramientas propias de las ciencias sociales computacionales.
- Conocimiento de los principios generales del diseño y funcionamiento de APIs, así como de los formatos más comunes de intercambio de información.
- Capacidad de identificar y acceder a APIs online para la descarga de datos observacionales de tipo social.
- Capacidad de compilar bases de datos estructuradas a partir de fuentes no estructuradas.

## DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Introducción al Web Scraping
  - ¿Qué es el Web Scraping?
  - Tipos de Web Scraping
  - Formatos de datos: XML y HTML
  - Acceso práctico a XML y HTML
  - Automatización de los programas de Web Scraping
  - Selenium y scraping basado en JavaScript
  - Cuestiones éticas del Web Scraping
  - Ejercicios prácticos
  
2. APIs de datos
  - Qué es una API
  - Fundamentos de la comunicación de la API
  - Introducción al formato JSON
  - Crea tu propia API (y compártela)
  - Arquitectura REST
  - Las API como forma de compartir y obtener datos (de cualquier tipo)
  - Automatización de las solicitudes de la API
  - Hablar con las bases de datos
  - Autenticación y acceso ético a las APIs
  - Ejercicios prácticos
  
3. Automatización de la adquisición de datos
  - ¿Por qué necesitamos la automatización?
  - Acceso a los servidores
  - Tecnologías para la automatización de programas
  - Automatización de trabajos cron
  - Tareas de registro
  - Ejercicios prácticos

## ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Actividades Formativas:

- Clases teórico-prácticas
- Tutorías
- Trabajo en grupo
- Trabajo individual del estudiante
- Exámenes parciales y finales

Metodologías Docentes:

- Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos.
- Lectura crítica de textos recomendados por el profesor de la asignatura: Artículos de prensa, informes, manuales y/o artículos académicos, bien para su posterior discusión en clase, bien para ampliar y consolidar los conocimientos de la asignatura.
- Resolución de casos prácticos, problemas, etc., planteados por el profesor de manera individual o en grupo.
- Exposición y discusión en clase, bajo la moderación del profesor de temas relacionados con el contenido de la materia, así como de casos prácticos.
- Elaboración de trabajos e informes de manera individual o en grupo.

## SISTEMA DE EVALUACIÓN

- Participación en clase (10%)
- Trabajos individuales o en grupo realizados durante el curso (70%)
- Examen final (20%)

**Peso porcentual del Examen Final:** 20

**Peso porcentual del resto de la evaluación:** 80

## BIBLIOGRAFÍA BÁSICA

- Barberá, P. & Steinert-Threlkeld, Z. How to use social media data for political science research. In The SAGE handbook of research methods in political science and international relations (Vol. 2, pp. 404-423). , SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781526486387>, 2020

- Freelon, D. Computational research in the post-API age. , Political Communication, 35(4), 665-668., 2018

- Nyhuis, D. Web data collection: potentials and challenges. In: The SAGE handbook of research methods in political science and international relations (Vol. 2, pp. 387-403). , SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781526486387>, 2020

- Perriam, J., Birkbak, A., & Freeman, A. Digital methods in a post-API environment. , International Journal of Social Research Methodology, 23(3), 277-290., 2020

#### BIBLIOGRAFÍA COMPLEMENTARIA

- Aydin, O. R Web Scraping Quick Start Guide: Techniques and tools to crawl and scrape data from websites., -, 2018

- Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. Automated data collection with R: A practical guide to web scraping and text mining. , John Wiley & Sons., 2014

#### RECURSOS ELECTRÓNICOS BÁSICOS

- . Application Programming Interfaces in R: [https://sicss.io/2020/materials/day2-digital-trace-data/apis/rmarkdown/Application\\_Programming\\_interfaces.html](https://sicss.io/2020/materials/day2-digital-trace-data/apis/rmarkdown/Application_Programming_interfaces.html)

- . Using APIs to get data: <https://cfss.uchicago.edu/notes/application-program-interface/>

- . Screen scraping with R: [https://cbail.github.io/ids704/screenscraping/rmarkdown/Screenscraping\\_in\\_R.html](https://cbail.github.io/ids704/screenscraping/rmarkdown/Screenscraping_in_R.html)