
Web scraping

PID_00256969

Laia Subirats Maté
Mireia Calvo González

**Laia Subirats Maté**

Ingeniera de Telecomunicaciones por la Universidad Pompeu Fabra (2008), máster en Telemática por la Universidad Politécnica de Cataluña (2009) y doctora en Informática por la Universidad Autónoma de Barcelona (2015). Desde 2009, trabaja como investigadora en Eureka (Centro Tecnológico de Cataluña) aplicando la ciencia de datos a distintas áreas como son la salud, el medio ambiente o la educación. Desde 2016, colabora con la UOC como docente en el máster de Data Science y en el grado de Informática. Es especialista en inteligencia artificial, ciencia de datos, *eHealth* y representación del conocimiento.

**Mireia Calvo González**

Ingeniera de Telecomunicaciones por la Universidad Politécnica de Cataluña (2011), máster en Ingeniería Biomédica por la Universidad de Barcelona y la Universidad Politécnica de Cataluña (2014) y doctora en Procesamiento de Señales y Telecomunicaciones por la Universidad de Rennes 1 y en Ingeniería Biomédica por la Universidad Politécnica de Cataluña (2017). Desde 2012, ha trabajado como investigadora en diferentes entornos académicos, clínicos e industriales, aplicando el procesamiento de datos al estudio de diferentes enfermedades cardíacas y respiratorias. Desde 2017, colabora con la UOC como docente en el máster de Data Science.

Primera edición: febrero 2019

© Mireia Calvo González y Laia Subirats Maté

Todos los derechos reservados

© de esta edición, FUOC, 2019

Av. Tibidabo, 39-43, 08035 Barcelona

Diseño: Manel Andreu

Realización editorial: Oberta UOC Publishing, SL

Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares del copyright.

Contenidos

Módulo didáctico 1

Web scraping

Laia Subirats Maté, Mireia Calvo González e Isabel Guitart Hormigo

1. ¿Por qué y cómo realizar *web scraping*?
2. Primeros pasos para realizar *web scraping*
3. *Web scraping* de contenido gráfico y audiovisual
4. Almacenamiento y compartición de datos
5. Prevención del *web scraping*
6. Resolución de obstáculos en *web scraping*
7. Aspectos legales
8. Mejores prácticas y consejos
9. Ejemplos de *web scraping* y casos de éxito

