

# Prueba de parseo

El ejercicio consiste en hacer un script en PHP que lea la información de unos cuantos cursos de una web, hacer un matching con los que están en un archivo CSV y guardar la información extraída en archivos XML.

Se pueden usar las librerías PHP que se crea conveniente.

Usar programación orientada a objetos.

El objetivo es parsear los cursos que aparecen en este listado:

<https://www.masterd.es/cursos-de-formacion-mantenimiento-industrial-g11>

## 1. Parsear listado de cursos

El primer paso es obtener los títulos y los enlaces de los cursos que aparecen en el listado. En el momento de preparar el ejercicio hay 13 cursos, pero puede variar si la web añade o quita cursos.



The screenshot shows the MasterD website's 'Cursos de Mantenimiento Industrial' page. The header includes the MasterD logo, navigation links (Oposiciones, Cursos, Centros, Blog, Opiniones MasterD, TV), and a contact number (900 30 40 30). The main heading is 'Cursos de Mantenimiento Industrial', followed by a subtext: '¿Quieres trabajar en el Sector Industrial? Obtén toda la información sobre nuestros cursos de mantenimiento y automatización industrial, autómatas programables y electrónica.' Below this, there is a list of three courses, each with a circular icon, a title, a red arrow pointing left, and an orange 'MÁS INFORMACIÓN' button.

Curso	Acción
MÁSTER EN MANTENIMIENTO INDUSTRIAL	MÁS INFORMACIÓN
MÁSTER EN AUTOMATIZACIÓN INDUSTRIAL Y ROBÓTICA	MÁS INFORMACIÓN
CURSO AUTOMATIZACIÓN INDUSTRIAL	MÁS INFORMACIÓN

**master.D** Oposiciones Cursos Centros Blog Opiniones MasterD TV Teléfono de información 900 30 40 30

 CURSO DOMÓTICA  [MÁS INFORMACIÓN](#)

 CURSO ORGANIZACIÓN Y GESTIÓN DE ALMACENES  [MÁS INFORMACIÓN](#)

**NOTICIAS DE CURSOS DE MANTENIMIENTO INDUSTRIAL**

<b>ESTUDIAR ELECTRÓNICA Y COMUNICACIONES</b> ¿Alguna vez te has preguntado, o has buscado razones, por las que estudiar electrónica o	<b>NUEVO CURSO DE LOGÍSTICA PARA TRABAJAR EN ORGANIZACIÓN Y GESTIÓN DE ALMACENES</b> La gestión de almacenes es una profesión que evoluciona a gran velocidad, que obliga a saber	<b>ESTUDIAR AUTOMATIZACIÓN INDUSTRIAL EN CÓRDOBA</b> ¿Quieres saber cómo estudiar Automatización Industrial en Córdoba? Gracias al Curso de
--	--	--

## 2. Obtener temario

Una vez tenemos los cursos, hay que entrar en la página de cada curso y leer la descripción. En este caso como descripción cogeremos todo el texto que hay desde debajo de la imagen del header hasta el final. Sólo nos interesa el texto, no las imágenes ni los vídeos, y los tags HTML se pueden quitar.

**master.D** Oposiciones Cursos Centros Blog Opiniones MasterD TV

**MASTER EN MANTENIMIENTO INDUSTRIAL**

- ✓ **CERTIFICADO SIEMENS**  
En automatización y automatismos
- ✓ **CAMPUS ONLINE**  
Lo mejor en tecnología para formación
- ✓ **LICENCIA TIA PORTAL**  
1 año Gratis de Licencia
- ✓ **PRÁCTICAS EN EMPRESAS**  
Accede al mercado laboral con MasterD

### Máster en Mantenimiento Industrial de MasterD

**MasterD** facilita al alumno una preparación especializada y profesional en el sector de la industria.

Este **Máster en mantenimiento industrial** está impartido por el instituto tecnológico de MasterD en centros colaboradores de SIEMENS y en el que participan importantes empresas del sector. Este **Máster profesional de Mantenimiento** está dirigido a especialistas que quieran sacar el máximo partido a sus conocimientos, asegurándose una mejora en su vida laboral y facilitando el acceso al sector industrial gracias a la preparación en tecnologías actuales como son la **industria 4.0**.

La duración total del **Máster en mantenimiento industrial** es de 1.500 horas, que corresponden a 60 créditos ECTS.

 **MÁSTER**

Nombre

0034

Email

Esna

- Conocer estándares de las redes de comunicación.
- Conocer de forma general qué es un dispositivo HMI y un sistema SCADA. Tener una visión general de la arquitectura de computadoras.

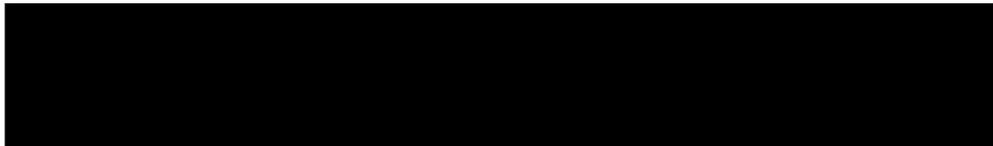
## Prácticas en empresas y bolsa de empleo

Gracias a los convenios obtenidos por MasterD, podrás realizar **prácticas profesionales en las más importantes empresas de la industria 4.0** como por ejemplo **FMI System** o **Aralux**. También tendrás disponible una **bolsa de empleo exclusiva** con la que mantenerte informado de las **ofertas laborales de mantenimiento industrial** que mejor se ajusten con tu formación.

## Opiniones Máster en mantenimiento industrial

Serafin estudió el **Máster en mantenimiento industrial de MasterD**. Nos cuenta su experiencia, como ha mejorado en su puesto de trabajo con esta titulación y **destaca la flexibilidad en horarios** para compatibilizar su empleo con los estudios del Máster.

¿Por qué estudiar un Máster en mantenimiento industrial?



### 3. Hacer matching con cursos existentes

El siguiente paso es mirar si alguno de los cursos existe en el archivo `cursos_masterd.csv`. Este archivo es un CSV con tres columnas: el id, el título y el tipo de impartición. Hay que comparar el título de los cursos que acabamos de parsear y mirar si coincide con alguno de los del CSV. Si coincide cogemos el id y se lo asignamos al curso.

### 4. Guardar resultados

Finalmente guardamos la información en dos archivos XML. Uno con los cursos que hemos encontrado matching y tenemos su id, y el segundo archivo con los cursos que no hemos hecho matching. El formato de los archivos XML tiene que ser el siguiente:

```
<courses>
  <course>
    <title></title>
    <url></url>
    <id></id>
    <description></description>
  </course>
</courses>
```