

```

C:\Python27\Scripts\postUGR>scrapy crawl postUGR_spider -o postUGR_items.json -t json
2015-06-22 00:53:40+0200 [scrapy] INFO: Scrapy 0.24.6 started (bot: postUGR)
2015-06-22 00:53:40+0200 [scrapy] INFO: Overridden settings available: ssl, http11
2015-06-22 00:53:40+0200 [scrapy] INFO: Overridden settings: {'NEWSPIDER_MODULE': 'postUGR.spiders', 'FEED_FORMAT': 'json', 'SPIDER_MODULES': ['postUGR.spiders'], 'FEED_URI': 'postUGR_
T_NAME': 'postUGR'}
2015-06-22 00:53:40+0200 [scrapy] INFO: Enabled extensions: FeedExporter, LogStats, TelnetConsole, CloseSpider, WebService, CoreStats, SpiderState
2015-06-22 00:53:41+0200 [scrapy] INFO: Enabled downloader middlewares: Http11Middleware, DownloadTimeoutMiddleware, UserAgentMiddleware, RetryMiddleware, DefaultHeadersMiddleware, M
ware, HttpCompressionMiddleware, RedirectMiddleware, CookieMiddleware, ChunkedTransferMiddleware, DownloaderStats
2015-06-22 00:53:41+0200 [scrapy] INFO: Enabled spider middlewares: HttpErrorMiddleware, OffsiteMiddleware, RefererMiddleware, UrlLengthMiddleware, DepthMiddleware
C:\Python27\lib\site-packages\scrapy\contrib\pipeline\_init_.py:21: ScrapyDeprecationWarning: ITEM_PIPELINES defined as a list or a set is deprecated, switch to a dict
category=ScrapyDeprecationWarning, stacklevel=1)
2015-06-22 00:53:41+0200 [scrapy] INFO: Enabled item pipelines: PostugrPipeline, CUSExport, XmlExportWithLabels, XmlExportWithOutLabels
2015-06-22 00:53:41+0200 [postUGR_spider] INFO: Spider opened
2015-06-22 00:53:41+0200 [postUGR_spider] INFO: Crawled 0 pages (at 0 pages/min), scraped 0 items (at 0 items/min)
2015-06-22 00:53:41+0200 [scrapy] DEBUG: Telnet console listening on 127.0.0.1:6023
2015-06-22 00:53:41+0200 [scrapy] DEBUG: Web service listening on 127.0.0.1:6080
2015-06-22 00:53:43+0200 [postUGR_spider] DEBUG: Crawled (200) <GET http://osl.ugr.es> (refer: None)

```

postUGR_items.html

contenido	autor	categorías	etiquetas	fondo
<p>[Si no eres miembro de la UGR] Ya... organizado por la Oficina de Software Libre y el resto de la Delegación del Rector de la UGR para las TIC para el Gabinete de Acción Social. Foto Finish V CISEL. Como en ocasiones anteriores, el objetivo es que las mujeres y niñas aprendan a crear con los ordenadores en un ambiente lúdico y pedagógico. El campus va dirigido a niñas y niños de 7 a 14 años, y se desarrollará del 29 de junio al 10 de julio. Como en anteriores ediciones, el horario será, de 9 de la mañana a las 2 de la tarde. Los niños podrán estar en el campus de 1 a 3 de la tarde con presencia de los monitores. Los niños se organizarán en tres grupos de 25 personas. Principantes, 7 a 9 años. Para quien no haya asistido antes al campus. Intermedios, de 9 a 11 años, también para quienes no hayan asistido en otra ocasión al campus. Veteranos, de 11 a 14 años y menores que hayan asistido en varias ocasiones al campus. Este año el máster de trabajo estará, basado en la ejecución de un proyecto en grupo, desde los niños grandes a los más pequeños. Los contenidos explicados en la pizarra serán máximos y se utilizarán, sobre todo, la resolución personalizada con las dudas que vayan surgiendo y animar a los niños y niñas a que vayan descubriendo por sí solos, preguntando a los profesores y a los compañeros, las posibilidades que tienen las herramientas con las que van a trabajar. Ya hemos contado detalles confunde se acorpan las fechas.</p>	Pablo Hinojosa	Campus infantil	Campus infantil	Campus Infantil de Software Libre 2015
	Renato Luis Ramirez Rivero	Eventos	Eventos , open data , OSL	Archivo diario:
<p>El próximo Lunes día 15 se celebra la final del... Serán en el Aula 1-2 (el aula Polivalente) de la... donde, a lo largo de la mañana, se presentarán los proyectos finalistas y habrá una serie de charlas sobre temas relacionados con el software libre. Posteriormente... anunciarán los ganadores y se hará la entrega de premios. El Certamen de proyectos libres de la UGR cuenta con el apoyo de... y la... aparte de la organización y apoyo de la... y la...</p>	Pablo Hinojosa	Certamen UGR de proyectos Libres	Certamen UGR de proyectos Libres, OSL	Final del Certamen de Proyectos Libres de la UGR 2014-2015
	Renato Luis Ramirez Rivero	Campus infantil	Campus infantil, Formaci3n, OSL , proyectos , open data, Software Libre	Archivo diario:
<p>Este año la Oficina de Software Libre de la Universidad de Granada no hemos podido participar en las (las #JPD15) debido a que coincidía con otros eventos ya programados, pero felicitamos a los Organizadores y animamos a que aquellas personas que se puedan acercar a dichas Jornadas en Madrid, Barcelona, Almería y Pamplona. La #JPD15 del 2 al 7 de junio de 2015, lo hacen, siendo esta una experiencia enriquecedora.</p>	Renato Luis Ramirez Rivero	Eventos	Eventos , open data , OSL	III Jornadas de Periodismo de Datos y Open Data

[illegible]

postUGR_items.json

```
[{"autor": ["Pablo Hinojosa"], "titulo": ["Campus Infantil de Software Libre 2015 "], "categorias": [" Campus infantil "], "contenido": ["(Si no eres mie", "autor": ["Renato Luis Ramirez Rivero"], "titulo": ["Archivo diario: "], "categorias": [" Eventos "], "contenido": [], "etiquetas": [" Eventos ", " ope", "autor": ["Pablo Hinojosa"], "titulo": ["Final del Certamen de Proyectos Libres de la UGR 2014-2015"], "categorias": [" Certamen UGR de proyectos Libre", "autor": ["Renato Luis Ramirez Rivero"], "titulo": ["Archivo diario: "], "categorias": [" Campus infantil "], "contenido": [], "etiquetas": [" Campus i", "autor": ["Renato Luis Ramirez Rivero"], "titulo": ["III Jornadas de Periodismo de Datos y Open Data"], "categorias": [" Eventos "], "contenido": ["Est", "autor": ["Pablo Hinojosa"], "titulo": ["Archivo diario: "], "categorias": [" Certamen UGR de proyectos Libres "], "contenido": [], "etiquetas": [" Cer", "autor": ["Manu"], "titulo": ["Archivo diario: "], "categorias": [" Software Libre "], "contenido": [], "etiquetas": [" Software Libre ", " asociacione", "autor": ["Renato Luis Ramirez Rivero", "Pablo Hinojosa"], "titulo": ["Archivo diario: "], "categorias": [" Cursos ", " Certamen UGR de proyectos Libre", "autor": ["Renato Luis Ramirez Rivero"], "titulo": ["Oferta formativa primer cuatrimestre curso 2015-2016"], "categorias": [" Cursos "], "contenido": [", "autor": ["Renato Luis Ramirez Rivero"], "titulo": ["Novidades del Campus Infantil de Software Libre 2015"], "categorias": [" Campus infantil "], "cont", "autor": ["Pablo Hinojosa"], "titulo": ["Ganadores del Certamen de Proyectos Libres de la UGR 2014-2015"], "categorias": [" Certamen UGR de proyectos L", "autor": ["Pablo Hinojosa"], "titulo": ["Ganadores del Certamen de Proyectos Libres de la UGR 2014-2015"], "categorias": [" Certamen UGR de proyectos L
```

postUGR_withLabel.xml

```
<?xml version="1.0" encoding="utf-8"?>
<items><item><autor><value>Pablo Hinojosa</value></autor><titulo><value>Campus Infantil de Software Libre 2015 </value></titulo><categorias><value>Camp
</value><value>Arduino básico.</value><value>
19 de octubre de 2015</value><value>
</value><value>Gestión de contenidos con Joomla.</value><value>
26 de octubre de 2015</value><value>
</value><value>Introducción al procesador de textos científicos LaTeX.</value><value>
3 de noviembre de 2015</value><value>
</value><value>Programación cliente-servidor en Javascript.</value><value>
26 de octubre de 2015</value><value>
</value><value>Raspberry Pi básico.</value><value>
26 de octubre de 2015</value><value>
</value><value>Scratch: Iniciación a la programación para enseñanza primaria y secundaria.</value><value>
3 de noviembre de 2015</value><value>
</value><value>Es solo un adelanto y una oportunidad para </value></contenido><etiquetas><value> Cursos </value><value> Formación </value><value> OSL </
```