	<b>Curso:</b> CFGS Desarrollo de aplicaciones multiplataforma <b>Módulo:</b> M03 Programación <b>Nombre y apellidos:</b> _____ <small>Cualquier tipo de plagio o copia será penalizado siguiendo la normativa de la escuela.</small>	<b>Fecha:</b> 26/10/2022 <b>Tipo:</b> Actividad práctica  <b>P4: XML</b>
--	---	---

## P4: XML

**La actividad es individual, pero os podéis ayudar sin copiar.**

Esta entrega se realizará mediante la plataforma Moodle y con el control de versiones en [GitHub](#). En la entrega del Moodle, habrá que subir un **archivo .zip** con el siguiente nombre: **DAM\_POO\_Apellido1\_Apellido2\_Nombre.zip**. En el GitHub de la asignatura, añade los ejercicios de esta práctica.


Dentro del archivo comprimido, se deben tener las siguientes carpetas y archivos<sup>1</sup>:

- README.md
- EX\_01
  - App.java
  - EX\_01.gif
- EX\_02
  - App.java
  - EX\_02.gif
- EX\_03
  - App.java
  - EX\_03.gif
  - README.md

El archivo README.md tiene que seguir el formato Markdown para explicar cuál ha sido la solución del ejercicio, las dificultades encontradas y una sección con las funcionalidades no conseguidas.

---

<sup>1</sup> Se pueden añadir más archivos en caso de que se considere necesario.

	<b>Curso:</b> CFGS Desarrollo de aplicaciones multiplataforma <b>Módulo:</b> M03 Programación <b>Nombre y apellidos:</b> _____ Cualquier tipo de plagio o copia será penalizado siguiendo la normativa de la escuela.	<b>Fecha:</b> 26/10/2022 <b>Tipo:</b> Actividad práctica  <b>P4: XML</b>
--	--	---

1. Las páginas web suelen tener un documento XML que especifica la estructura de esta. En este documento se incluye información básica como la última modificación de las páginas y URLs a los principales enlaces. Podéis ver un ejemplo de la página de la escuela [https://www.eug.es/sitemap\\_index.xml](https://www.eug.es/sitemap_index.xml).

Para encontrar el sitemap de una web, podéis utilizar la ruta /sitemap.xml en la URL y, si está bien programado, mostrará la información.

Crea un programa en Java que obtenga el sitemap de xataka en XML y muestra por el terminal las diferentes URLs que encuentres.

2. Cada una de las URLs obtenidas, incluyen las URLs a las noticias.

Ejemplo de URLs obtenidas en el ejercicio anterior:

```
<?xml version="1.0" encoding="UTF-8" ?>
<sitemapindex xsi:schemaLocation="http://www.sitemaps.org/schemas
<sitemap>
  <loc>
    https://www.xataka.com/categoria/default/sitemap.xml
  </loc>
  <lastmod>2011-03-08T12:30:56Z</lastmod>
</sitemap>
<sitemap>
  <loc>https://www.xataka.com/categoria/otros/sitemap.xml</loc>
  <lastmod>2021-09-22T08:03:50Z</lastmod>
</sitemap>
```

Si accedemos a una de estas URL, veremos que contienen otro sitemap.xml con los enlaces a las noticias:

```
<?xml version="1.0" encoding="UTF-8" ?>
<loc>
  https://www.xataka.com/otros/soy-padre-unico-alumno-clase-14-anos-que-no-tiene-movil-1
</loc>
<lastmod>2021-09-22T08:03:50Z</lastmod>
<changefreq>daily</changefreq>
<priority>1</priority>
<image:image>
  https://i.blogs.es/267a24/gaelle-marcel-fldk3n-ygf4-unsplash/650_1200.jpg
</image:image>
</url>
```

Partiendo del programa anterior, procesa estas URLs de noticias y muestra por el terminal los nuevos enlaces.

3. Partiendo del programa anterior, crea una carpeta XMLs donde se guarden los documentos que utilizarán de path de la URL del sitemap correspondiente y dentro tendrán las URLs de las noticias. Por ejemplo:

Nombre	Fecha de modificación	Tipo	Tamaño
archivos_2018_04_sitemap.xml	25/09/2021 13:33	Documento XML	53 KB
archivos_2018_05_sitemap.xml	25/09/2021 13:33	Documento XML	52 KB
archivos_2018_06_sitemap.xml	25/09/2021 13:33	Documento XML	50 KB
archivos_2018_07_sitemap.xml	25/09/2021 13:33	Documento XML	48 KB
archivos_2018_08_sitemap.xml	25/09/2021 13:33	Documento XML	50 KB
archivos_2018_09_sitemap.xml	25/09/2021 13:33	Documento XML	48 KB

Podemos ver los archivos generados con sus respectivos nombres, y en la siguiente imagen, el resultado del contenido del documento.

```
archivos_2018_04_sitemap.xml
1 https://www.xataka.com/cine-y-tv/la-segunda-temporada-de-13-reasons-why-estrena-trailer-y-anuncia-su-fecha-de-estreno-18-de-mayo-en-netflix
2 https://www.xataka.com/privacidad/ian-koum-fundador-y-cep-de-whatsapp-renuncia-tras-susueas-presiones-de-facebook-para-debilitar-el-cifrado-de-datos
3 https://www.xataka.com/videojuegos/atari-vos-llegara-con-4k-hdr-a-60fps-algunas-especificaciones-finales-y-fechas-para-reservas-y-lanzamiento
4 https://www.xataka.com/servicios/llega-la-windows-10-april-2018-update-estas-son-las-novedades
5 https://www.xataka.com/servicios/twitter-le-vendio-datos-al-mismo-investigador-de-cambridge-analitica-que-se-los-robo-a-facebook
6 https://www.xataka.com/ciencia-ficcion/gene-wolfe-el-escritor-que-sono-con-un-futuro-donde-la-ciencia-y-la-magia-son-la-misma-cosa
```