

## A1. Almacenamiento de la información

### Índice.

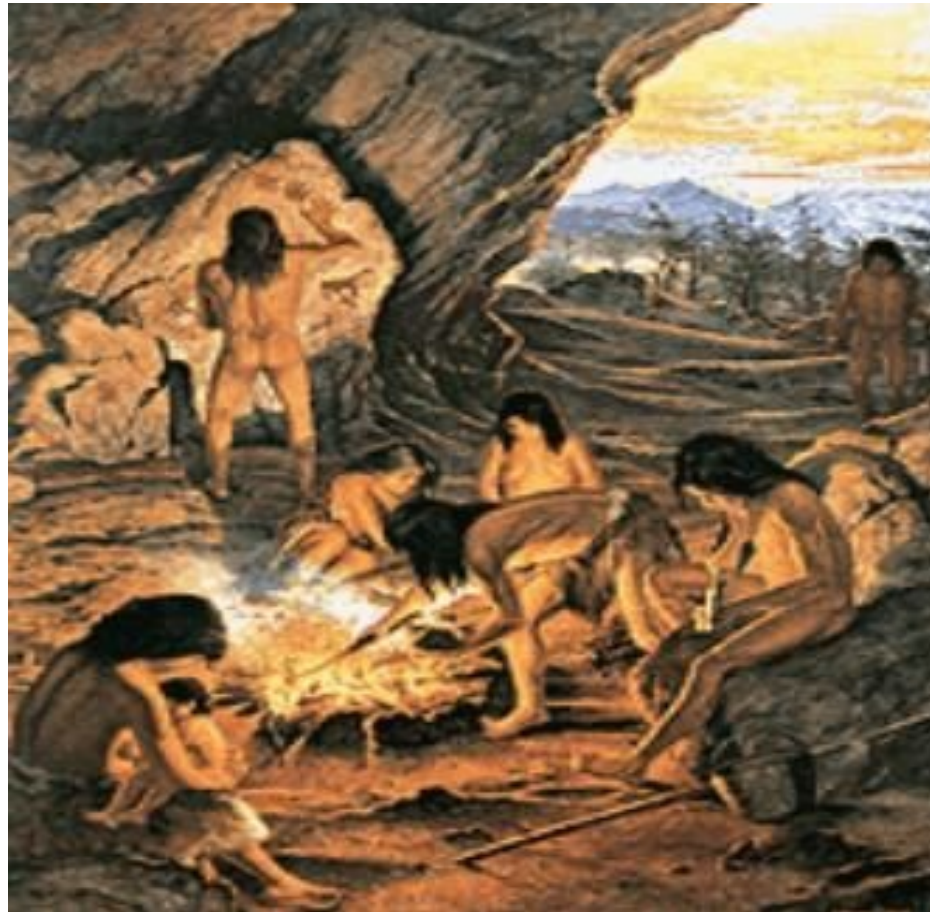
1.	Sistemas de almacenamiento.....	2
2.	¿Qué es un fichero?.....	4
3.	Operaciones con ficheros.....	7
4.	Tipos de ficheros.....	8
5.	Soportes de información.....	10
6.	Métodos de acceso.....	11
7.	Tipos de organización de ficheros.....	12
7.1.	Organización secuencial.....	12
7.2.	Organización secuencial encadenada.....	13
7.3.	Organización secuencial indexada.....	14
7.4.	Organización directa o aleatoria.....	15
8.	Parámetros de utilización.....	16
9.	Otros sistemas de almacenamiento.....	17

## A1. Almacenamiento de la información

### 1. Sistemas de almacenamiento.

---

Un sistema de almacenamiento de información es un conjunto de elementos y dispositivos capaces de contener información y, debido a que almacenan datos con gran valor, una buena parte de ellos es esencial.



## A1. Almacenamiento de la información

### 1. Sistemas de almacenamiento.

La evolución histórica de los **sistemas de almacenamiento** ha sido la siguiente:

- **Fichero manual** → hasta hace sólo unos 30 años, la información se guardaba y se hacía su seguimiento mediante soportes de papel. Se utilizaba un sistema de ficheros manual que se componía de un conjunto de carpetas etiquetadas que se guardaban en un armario o archivo.
- **Sistema de ficheros** → las primeras aplicaciones que manejaban datos con el ordenador, los guardaban en ficheros de forma secuencial y se accedía a dichos datos para realizar tareas que ayudaban en la toma de decisiones.



- **Sistema de bases de datos** → las innovaciones tecnológicas de los programas informáticos fueron evolucionando y, a fin de los 60, surgieron las bases de datos como conjuntos de archivos relacionados entre sí, compartidos entre procesos y usuarios, y dotados de estructuras complejas.



## A1. Almacenamiento de la información

### 2. ¿Qué es un fichero?

En la década de los 70, las empresas realizaban una serie de procesos básicos relacionados con la contabilidad y la facturación, cuyas necesidades de almacenamiento y de tratamiento de información se cumplían con un reducido número de archivos en papel (normalmente agrupados y ordenados) → los típicos ficheros.



La informática transformó esos datos en formato de papel para que fuesen accesibles de forma más rápida a través de ordenadores. En ese momento, la informática adaptó sus herramientas para que los elementos con los que el usuario trabajase se pareciesen a los del ordenador → ficheros, formularios, carpetas, directorios, ...

El elemento clave que permitió el almacenamiento de forma permanente en dispositivos de memoria masiva fue el **fichero** o **archivo**.

## A1. Almacenamiento de la información

### 2. ¿Qué es un fichero?

Un **fichero** (o archivo) es un conjunto organizado de información almacenada en un soporte común.

Los **ficheros** están formados por **registros lógicos** que contienen datos relativos a un mismo elemento.

Los **registros** están divididos en **campos** que contienen cada una de las informaciones elementales del registro.

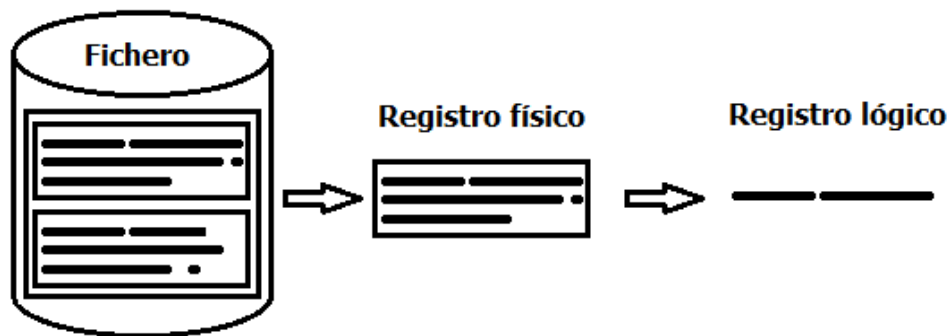
Nombre Proyecto	Contacto	Correo	Teléfonos
Eagle	anell y adrian	eagleog@cantv.net	0212-2365502 - 4162800
TLC	Ambar Texeira	boutiquepoventas2003@gmail.com	043-07-85
Tedexis	Luis Poggi	lpoggi@gmail.com	4141638624
FYC	Carlos Marrero	carlos.marrero@fyccorp.com	2327611
NGS camcaroni	Alfonso Mora	amora@ngs.com.ve	41-86905269
Tecnoquim	Lenin Marques	grupotecnquim@gmail.com	7616267, 7628563, 0416 808-0706
Tuscomprasporinternet	Rafael Monserrat	bufetemontserrat@hotmail.com	4147968204

## A1. Almacenamiento de la información

### 2. ¿Qué es un fichero?

Los ficheros suelen ser muy voluminosos, por lo que no se suelen cargar en memoria principal en su totalidad, sino en una pequeña parte llamada **registro físico** (o bloque).

Un **registro físico** (o bloque) es la cantidad de información que se se transfiere entre el fichero y la memoria principal en cada operación de lectura/grabación.



Un **registro lógico** es la unidad mínima de proceso de una aplicación, mientras que un **registro físico** es la unidad mínima de acceso al soporte.

Cada operación de lectura/escritura transfiere varios registros del fichero (un bloque suele contener varios registros lógicos).

El **factor de bloqueo** es el número de registros que entran en cada bloque.

El **bloqueo de registros** es la operación de agrupar varios registros en un mismo bloque.

## A1. Almacenamiento de la información

### 3. Operaciones con ficheros.

---

Las operaciones básicas o normales con los ficheros son las siguientes:

- **Crear** → se crea el descriptor físico del fichero en función del descriptor lógico asociado.
- **Borrar** → se elimina el descriptor físico del fichero.
- **Renombrar** → se cambia el nombre del fichero.
- **Abrir** → se asocia el descriptor lógico del fichero a su descripción física para acceder a la información contenida.
- **Cerrar** → se disocia el descriptor lógico del fichero de su descripción física, grabando previamente toda la información que estuviera en caché.
- **Grabar** → se guarda los registros lógicos que estén en memoria dentro de los registros físicos del fichero, para su uso posterior.
- **Leer** → se lee los registros físicos del fichero y se pasan a registros lógicos que estarán en caché.
- **Posicionarse** → se ubica el descriptor lógico del fichero en una determinada posición para leer el contenido del registro asociado.

# A1. Almacenamiento de la información

## 4. Tipos de ficheros.

Los ficheros se pueden clasificar de muy diversas formas:

- **Según la forma de guardar la información:**

- **Ficheros de texto** → la información se guarda secuencialmente y en forma de caracteres.
- **Ficheros binarios** → La información se guarda tal y como está en memoria principal, en ceros y unos.

- **Según la longitud de los registros:**

- **Longitud fija** → todos los registros del fichero tienen la misma longitud.
- **Longitud variable** → la longitud de cada registro varía entre un mínimo y un máximo.
- **Longitud delimitada** → la longitud de los registros es en función de un carácter especial que indica fin del registro.
- **Longitud indefinida** → la longitud de los registros es variable y cada uno tiene delimitadores de inicio y fin del registro.



- **Según el uso de los registros:**

- **Ficheros permanentes** → información necesaria para el funcionamiento de una aplicación, con larga vida.
  - × Ficheros maestros o de situación → con información que refleja el estado actual de los datos, actualizándose periódicamente.
  - × Ficheros constantes → con datos fijos que no variarán o si lo hacen es de forma ocasional.
  - × Ficheros históricos → con datos que fueron actuales en situaciones anteriores.
- **Ficheros temporales** → información útil durante el funcionamiento de una aplicación, con vida breve.
  - × Ficheros intermedios → con los resultados de una aplicación que se utilizarán en otra aplicación.
  - × Ficheros de maniobras → con datos que no se pueden mantener en memoria por razón de espacio.
  - × Ficheros de resultados → con datos que serán transferidos a un dispositivo de salida.



# A1. Almacenamiento de la información

## 4. Tipos de ficheros.

---

Actualmente, todo sistema operativo trata los ficheros desde dos puntos de vista:

- Contenido → texto o datos binarios.
- Tipo → imágenes, ejecutables, vídeos, ...

Bajo esta perspectiva aparece una nueva clasificación:

- **Ficheros de texto (planos o ascii)** → ASCII es un estándar que asigna un valor numérico a cada carácter, los documentos se representan en texto plano y son directamente legibles. No necesitan de un formato especial para su interpretación pero su extensión identifica el tipo de texto que contiene.
  - Ficheros de configuración → ini, inf, conf, ....
  - Ficheros de código fuente → sql, c, java, ...
  - Ficheros de páginas web → html, php, csc, xml, ...
  - Ficheros de formato enriquecido → rtf, ps, tex, ....
- **Ficheros binarios** → los que no son de texto, es decir, si requieren de un formato para su interpretación:
  - De imagen → jpg, gif, tiff, bmp, wmf, png, pcx, ...
  - De vídeo → mpg, mov, avi, gt, ...
  - Comprimidos o empaquetados → zip, Z, gz, tar, lhz, ...
  - Ejecutables o compilados → exe, com, xgi, o, a, ...
  - Procesadores de textos → doc, odt, ...

## A1. Almacenamiento de la información

### 5. Soportes de información.

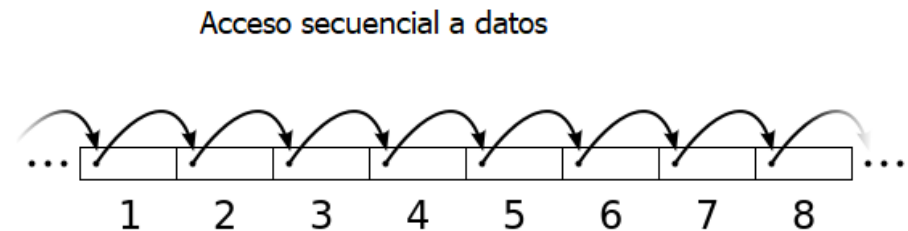
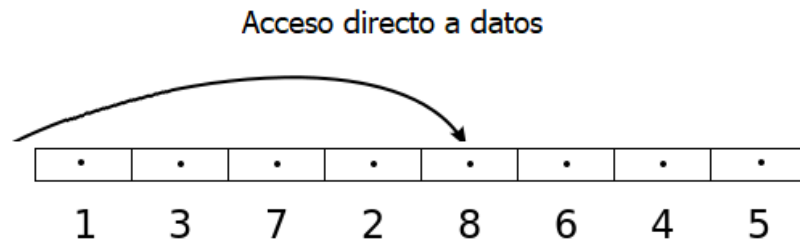
Un **soporte de información** es un conjunto de dispositivos que permiten almacenar la información en formato electrónico y, normalmente, es de fácil transporte.

Los soportes más utilizados son los siguientes:

- Discos duros internos y externos.
- Cintas y discos de copias de seguridad.
- Unidades USB o pendrives.
- Tarjetas de memoria (SD, microSD, ...).
- Discos ópticos (CD, DVD).
- Discos de estado sólido (SSD).

Básicamente, todos estos soportes se pueden agrupar en dos tipos:

- Soportes de acceso directo a los datos → los más empleados y el acceso se realiza de forma directa.
- Cintas y discos de copias de seguridad → se suelen utilizar para copias de seguridad.



## A1. Almacenamiento de la información

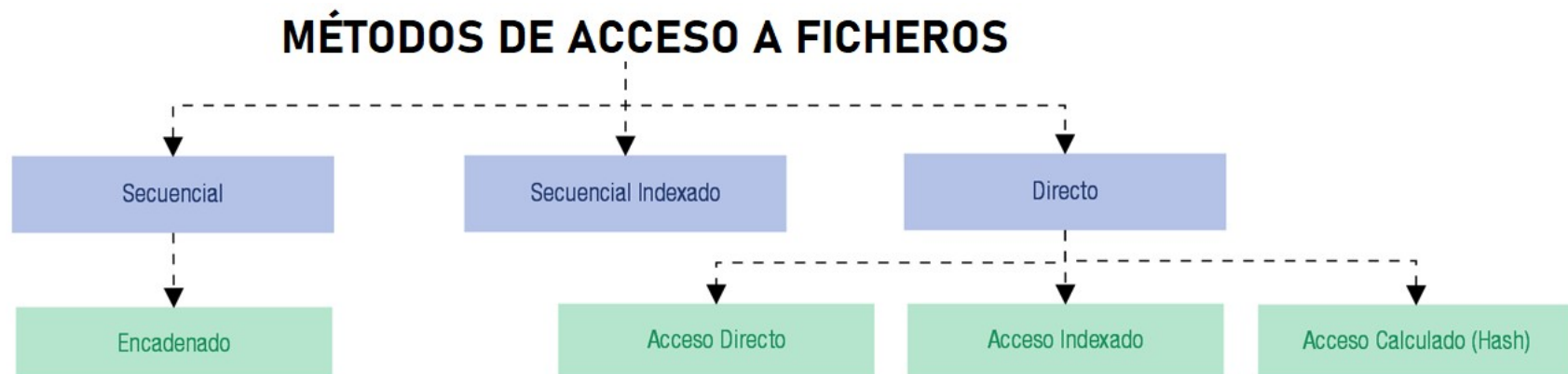
### 6. Métodos de acceso.

Los **métodos de acceso** a los distintos tipos de ficheros han variado enormemente gracias a la evolución de la tecnología, que también lo ha hecho principalmente debido a los avances en el hardware.

Los **objetivos** fundamentales de estas modificaciones se pueden resumir en:

- Proporcionar acceso rápido a los registros.
- Reducir costes de almacenamiento.
- Facilitar la actualización de los registros.
- Permitir que la estructura refleje la organización real de la información.

Hay distintas formas de organizar los ficheros en los soportes de memoria (secuencia, directo y secuencial indexado), cada una de las cuales tiene asociado un tipo distinto de método de acceso (encadenado, indexado, calculado).



## A1. Almacenamiento de la información

### 7. Tipos de organización de ficheros.

#### 7.1. Organización secuencial.

Un **fichero con organización secuencial** tiene todos sus registros almacenados de forma contigua y el acceso se realiza yendo un registro tras otro.

Suele tener una marca indicadora de fin de fichero, denominada EOF (End Of File).

Este tipo de ficheros se utiliza en dispositivos o soportes no direccionales o de acceso secuencial, como son las cintas magnéticas de almacenamiento de datos.



Los registros se identifican por medio de una información ubicada en uno de los campos y que se suele denominar **clave** o **llave**.

Si se ordena el archivo secuencial por su clave se realizará antes cualquier operación.

Las **operaciones** que se pueden realizar son las siguientes:

- Agregar → sólo al final.
- Consultar → se leen los registros secuencialmente desde el principio hasta encontrar el buscado .
- Borrado lógico → se marca el registro de forma que al leerlo nuevamente se identifique como no válido.
- Actualizar → las operaciones de agregar en una posición, modificar y eliminar físicamente obligan a reescribir nuevamente el fichero.

## A1. Almacenamiento de la información

### 7.2. Organización secuencial encadenada.

Un **fichero con organización secuencial encadenada** tiene todos sus registros almacenados con un puntero al siguiente registro, según el orden lógico del fichero.



Las **operaciones** que se pueden realizar son las siguientes:

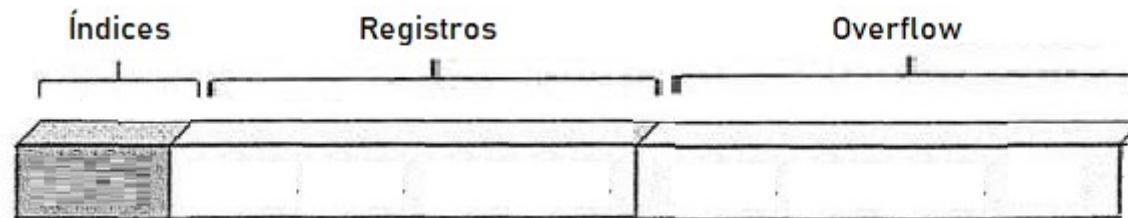
- Agregar → en cualquier posición.
- Consultar → se leen los registros secuencialmente desde el principio hasta encontrar el buscado .
- Borrado lógico → el registro anterior apuntará al registro que apunte el que queremos borrar.
- Actualizar → se agrega el registro al final y se modifican los apuntadores asociados.

## A1. Almacenamiento de la información

### 7.3. Organización secuencial indexada.

Un **fichero con organización secuencial indexada** tiene dos estructuras o zonas, una para índices y otra para registros que, a su vez, se va a descomponer en dos (registros y desbordamiento):

- Zona de índices → estructura secuencial con campo clave y campo dirección de un registro del fichero.
- Zona de registros → todos los registros del fichero, ordenados según el valor de una llave.
- Zona de desbordamiento u overflow → estructura secuencial.



Las **operaciones** que se pueden realizar son las siguientes:

- Agregar → en su posición o en el overflow si ya está ocupada.
- Consultar → se leen el registro en función de su índice.
- Borrado lógico → elimina el registro deseado.
- Actualizar → se agrega el registro modificado a la zona de desbordamiento.

# A1. Almacenamiento de la información

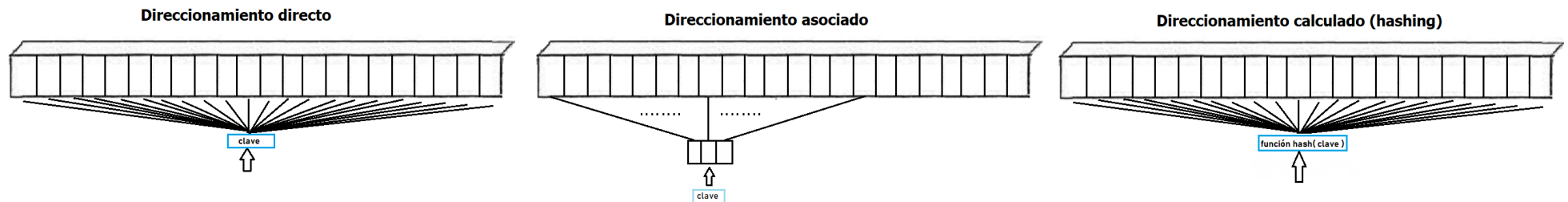
## 7.4. Organización directa o aleatoria.

Un **fichero con organización directa o aleatoria** ubica cada registro en una dirección de memoria calculada aplicando una fórmula (o algoritmo matemático).

Se busca que haya el mínimo número de sinónimos → registros con llaves diferentes acceden a una misma dirección física.

Principal problema → elección de una buena función de transformación o método de direccionamiento:

- **Direccionamiento directo** → asocia cada llave a una dirección lógica.
- **Direccionamiento asociado** → asocia cada llave a una dirección lógica a través de una tabla.
- **Direccionamiento calculado (hashing)** → asocia cada llave a una dirección lógica calculada matemáticamente en función de su valor, produciendo sinónimos.



Las **operaciones** básicas son las siguientes:

- Agregar.
- Consultar .
- Borrar.
- Modificar.

## A1. Almacenamiento de la información

### 8. Parámetros de utilización.

---

La creación de un fichero implica especificar su organización y el tipo de operaciones que se van a realizar. La forma para determinar la organización más adecuada y eficiente se realiza estudiando las características de su utilización, llamadas como **parámetros de referencia o de utilización**:

- **Capacidad o volumen** → espacio que ocupa (en caracteres).
- **Actividad** → número de consultas y modificaciones a lo largo de su vida.
  - **Tasa de consulta o modificación (%)** → registros consultados (o modificados) en cada operación con respecto al total de registros.
  - **Frecuencia de consulta o modificación** → veces que se accede al fichero para consultar (o modificar) por unidad de tiempo.
- **Volatilidad** → número de inserciones y borrados a lo largo de su vida.
  - **Tasa de renovación (%)** → registros renovados en cada operación con respecto al total de registros.
  - **Frecuencia de renovación** → veces que se accede al registro para modificar por unidad de tiempo.
- **Crecimiento** → variación de la capacidad del fichero.
  - **Tasa de crecimiento (%)** → altas y bajas con respecto al total de los registros.



# A1. Almacenamiento de la información

## 9. Otros sistemas de almacenamiento.

Además de los sistemas de almacenamiento vistos, hay otros sistemas para guardar también las bases de datos, como son los siguientes:

- **XML** → almacenan gran cantidad de información en formato XML sobre documentos de fácil edición.

```
<?xml version="1.0" standalone="no"?>
<!DOCTYPE peliculas SYSTEM "Peliculas.dtd">

<peliculas>
  <pelicula tipo="comedia" sistema="PG-13" ejemplares="5" año="1987">
    <titulo>Raising Arizona</titulo>
    <guionista>Ethan Coen</guionista>
    <guionista>Joel Coen</guionista>
    <productor>Ethan Coen</productor>
    <director>Joel Coen</director>
    <actor>Nicolas Cage</actor>
    <actor>Holly Hunter</actor>
    <actor>John Goodman</actor>
    <comentarios>Una película clásica de Comedia.</comentarios>
  </pelicula>

  <pelicula tipo="ciencia-ficción" sistema="PG-13" ejemplares="4" año="1989">
    <titulo>The Abyss</titulo>
    <guionista>James Cameron</guionista>
    <productor>Gale Anne Hurd</productor>
    <director>James Cameron</director>
    <actor>Ed Harris</actor>
    <actor>Mary Elizabeth Mastrantonio</actor>
    <comentarios>Una buena película.</comentarios>
  </pelicula>
</peliculas>
```

- **Servicio de directorios** → un directorio es una base de datos optimizada para la lectura, navegación y búsqueda basada en atributos y con capacidades de filtrado. Hay muchas formas de proveer este servicio de directorio: NIS, LDAP, Active Directory, ...

