

# Memoria Sistema Experto en Asesoría Legal

*Ingeniería del Conocimiento 2016/2017*

*Universidad de Granada*

Alumno: David Criado Ramón

D.N.I.:

# Contenido

1. RESUMEN.....	3
2. DESCRIPCIÓN DEL PROCESO SEGUIDO PARA EL DESARROLLO.....	4
2.a PROCEDIMIENTO SEGUIDO PARA OBTENER EL CONOCIMIENTO .....	4
<i>MÓDULO 1-SELECCIÓN DE LICENCIAS (REJILLA DE REPERTORIO)</i> .....	4
<i>MÓDULO 2-COMPATIBILIDAD DE LICENCIAS (ÁRBOL DE DECISIÓN)</i> .....	6
<i>MÓDULO 3-PROTECCIÓN DE DATOS</i> .....	6
2.b PROCEDIMIENTO SEGUIDO PARA VALIDAR Y VERIFICAR EL CONOCIMIENTO .....	7
3. DESCRIPCIÓN DEL SISTEMA DESARROLLADO .....	7
3.a Variables de entrada del problema .....	7
3.b Variables de salida del problema .....	7
3.c Conocimiento global del problema .....	7
3.d Módulos que se han desarrollado:.....	8
3.e Esquema de razonamiento del sistema:.....	8
3.f La lista de hechos que utiliza el sistema y la forma de representarlos:.....	8
3.g Las reglas y hechos de cada módulo:.....	9
4. BREVE MANUAL DE USO DEL SISTEMA.....	11

## 1. RESUMEN

El Sistema Experto desarrollado empieza con un menú que nos permite acceder a los 3 módulos desarrollados.

En el primer módulo, *Selección de Licencias*, el usuario tendrá que responder a varias preguntas con el fin de determinar qué licencia es la que más le interesa. Gracias a la información obtenida de la rejilla de repertorio (las licencias con sus constructores introducidas como conocimiento y las decisiones mediante las reglas del árbol de conceptos) sobre 7 licencias distintas el Sistema Experto va descartando posibilidades hasta dejar una única posibilidad donde para y volvemos al menú inicial.

En el segundo módulo, *Compatibilidad de Licencias*, el usuario indicará la licencia que desea usar para su software y la licencia de un software de terceros que desea usar de entre 20 licencias disponibles. Gracias a un árbol de decisión realizado con diferentes pares de licencias compatibles y no compatibles el árbol de decisión nos da lugar a las diferentes reglas que nos indicarán si dos licencias son compatibles o no. Una vez especificado volvemos al menú inicial.

En el tercer módulo, *Protección de Datos*, se realizará un procedimiento iterativo en el que:

- El usuario seleccionará de una lista datos que usa para su base de datos.
- El usuario indicará el tipo de organización, el área geográfica y el uso que va a dar los datos su organización.
- Se deducirá si los datos recogidos son suficientes para identificar a una persona.
- Basándose en toda la información recogida se determinarán cuáles son las medidas que se deben implementar para cumplir la LOPD.
- Al final, se rellenarán los datos para generar uno de los 4 formularios ARCO automáticamente que será generado en el directorio de trabajo y volvemos al menú inicial.

## 2. DESCRIPCIÓN DEL PROCESO SEGUIDO PARA EL DESARROLLO

### 2.a PROCEDIMIENTO SEGUIDO PARA OBTENER EL CONOCIMIENTO

#### *MÓDULO 1-SELECCIÓN DE LICENCIAS (REJILLA DE REPERTORIO)*

Para el proceso de obtención de información para determinar los constructores y calificarlos se han usado principalmente tanto el PDF disponible en la página web de la asignatura, como la tabla comparativa en Wikipedia<sup>1</sup> y el un artículo en Hipertextual<sup>2</sup> así como la descripción de las propias licencias. Este procedimiento se ha aplicado a las 20 licencias disponibles para el módulo 2 aunque la rejilla de repertorio sólo se ha aplicado a 7 de ellas.

En estos se consideraban los siguientes constructores:

- C1 - Uso enlazado (Ej.: Librería)  
1 – Dominio Público 1.5 – Limitado 2 – Permisivo 3 – CopyLeft 4 - Restringido 5- No
- C2 - Distribución  
0 – Dominio Público 0.5 – Limitado 1 – Permisivo 2 – CopyLeft 3 – No
- C3 – Modificación  
1 – Dominio Público 1.5 – Limitado 2 – Permisivo 3 – CopyLeft 4 - No
- C4 – Protección de Patentes  
1- No 2 - Manual 3 - Sí
- C5 – Uso Privado (Personal, no se puede publicar)  
1 – Sí 2 – No
- C6 – Tipo Sublicencia  
1 – Dominio Público 2 – Permisivo 3 – CopyLeft 4 – No

---

<sup>1</sup> [https://en.wikipedia.org/wiki/Comparison\\_of\\_free\\_and\\_open-source\\_software\\_licenses](https://en.wikipedia.org/wiki/Comparison_of_free_and_open-source_software_licenses)

<sup>2</sup> <https://hipertextual.com/archivo/2014/05/como-elegir-licencias-open-source/>

Licencia	C1	C2	C3	C4	C5	C6
Apache	2	1	2	3	1	2
BSD (2 cláusulas	2	1	2	2	1	2
Eclipse	1.5	0.5	1.5	3	1	1.5
GPL	4	2	3	3	1	3
MIT	2	1	2	2	1	2
Microsoft Office o Privativo	5	3	4	3	2	4
Mozilla Public License	2	2	3	3	1	3

Tabla 2.a.1: Rejilla de repertorio inicial

Tras el proceso de análisis de la rejilla y añadir una pregunta especial que no depende de las características para separar MIT de BSD que con las características puestas tienen 0 sale el siguiente árbol de conceptos:

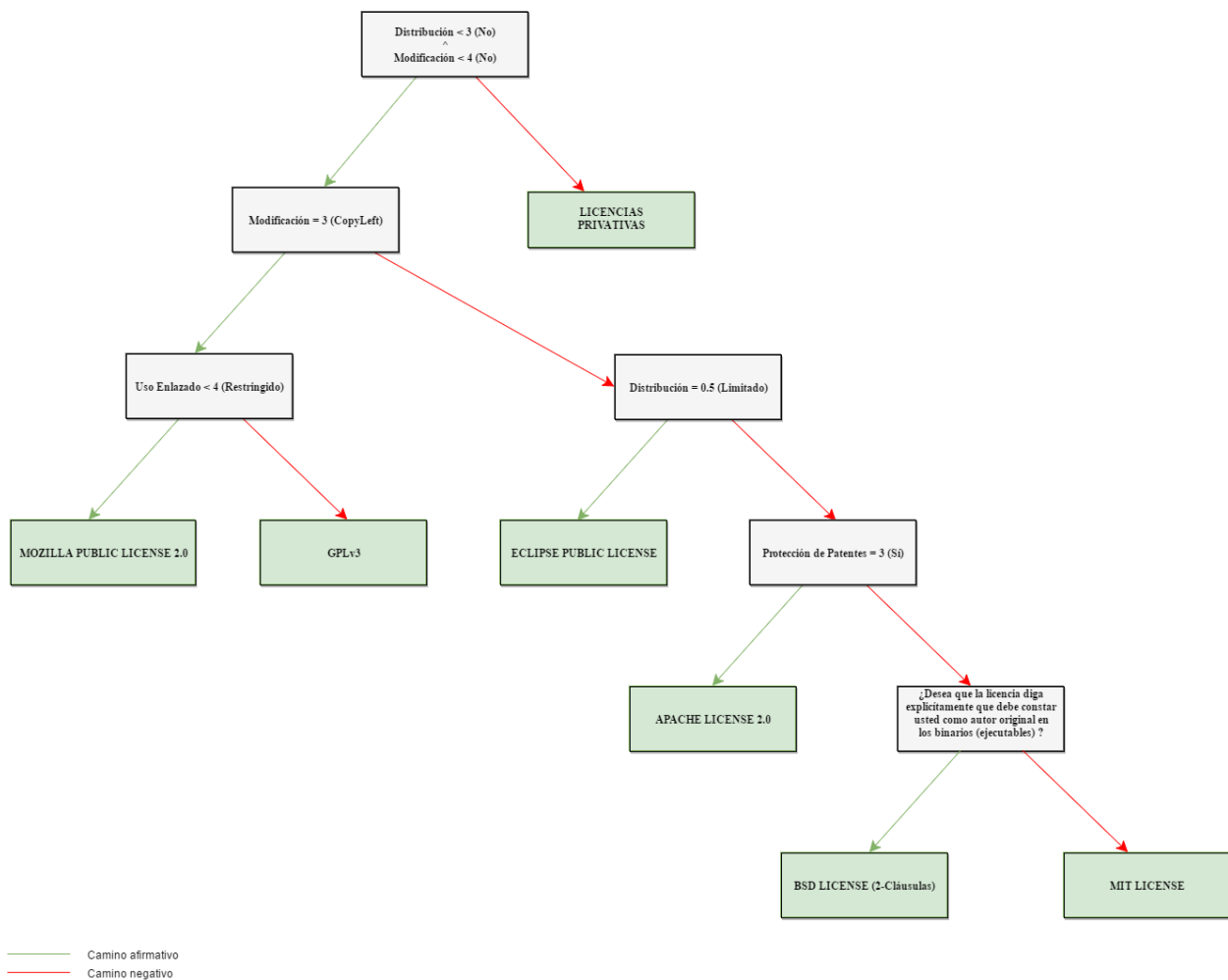


Figura 2.a.2: Árbol de conceptos

## MÓDULO 2-COMPATIBILIDAD DE LICENCIAS (ÁRBOL DE DECISIÓN)

Se han utilizado tanto las características de las anteriormente vistas como de otras para dar lugar a una tabla de pares de licencias y si estás son compatibles o no.

Para ello basándonos en un CSV (disponible en la carpeta otros dentro de la entrega y ejecutando el script en R en la misma carpeta) he obtenido el siguiente árbol de decisión (donde 0 implica no compatible y 1 compatible):

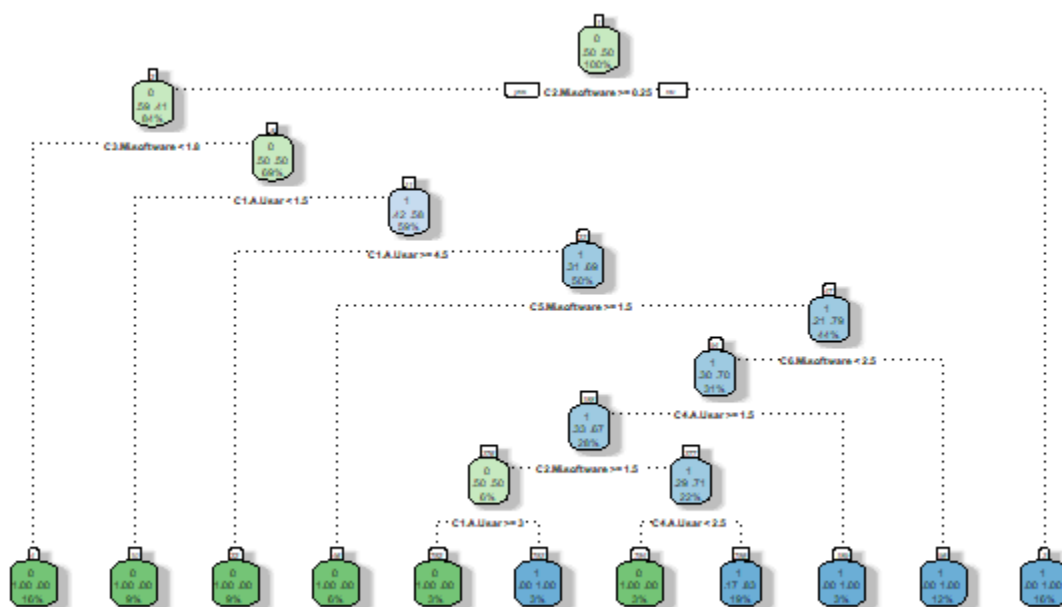


Figura 2.a.3: Árbol de decisión

Las reglas dadas por las hojas son las que se han implementado y se ha argumentado la decisión en base a cada uno de los pasos que da el árbol de decisión.

## MÓDULO 3-PROTECCIÓN DE DATOS

En este caso se ha realizado educación directa del experto en el tema proporcionado por la asignatura así como de las páginas web proporcionadas <sup>3 4 5</sup>

<sup>3</sup> <http://www.cuidatusdatos.com/infodatospersonales.html>

<sup>4</sup> <http://www.cuidatusdatos.com/infocomocumplirlop.html>

5

[http://www.agpd.es/portalwebAGPD/CanalDelCiudadano/derechos/principales\\_derchos/index-ides-idphp.php](http://www.agpd.es/portalwebAGPD/CanalDelCiudadano/derechos/principales_derchos/index-ides-idphp.php)

## 2.b PROCEDIMIENTO SEGUIDO PARA VALIDAR Y VERIFICAR EL CONOCIMIENTO

Una vez introducidas las reglas para cada caso que se han desarrollado se han probado todos los casos para comprobar que salen los resultados que cabía esperar (en aquellos que en los que entran muchos ejemplos probando aproximadamente un 40 % de ellos) y comprobando la traza de las reglas en la interfaz de CLIPS para Windows.

Además, se ha revisado que todos los mensajes sean comprensibles para los usuarios y que el sistema trate de alguna manera las entradas erróneas en cualquier punto de su ejecución.

## 3. DESCRIPCIÓN DEL SISTEMA DESARROLLADO

### 3.a Variables de entrada del problema

-Opción del menú principal

-Para cada pregunta del módulo 1 una respuesta binaria (sí/no)

-Para el módulo 2 selección de dos licencias de una lista (una para el software de terceros y otra la que se desea usar).

-Para el módulo 3:

- Selección de datos que se usan de una lista múltiple
- Selección del tipo de organización, uso y área geográfica de listas.
- Datos para rellenar uno de los formularios ARCO.

### 3.b Variables de salida del problema

-Para el módulo 1 la licencia que se recomienda usar.

-Para el módulo 2 si las licencias son compatibles o no.

-Para el módulo 3:

- Si los datos usados son suficientes para identificar a una persona.
- Las medidas que se deben implementar para el tratamiento de datos.
- El fichero ARCO solicitado.

### 3.c Conocimiento global del problema

El sistema dispone de:

-Todas las licencias (tanto del módulo 1 como del 2) en un *template* denominado *Licencia* con cada una de las características consideradas en la rejilla como *fields* y un *ID* para mostrarlo en las listas.

- Todos los datos ofertados en el módulo 3 en un *template* denominado *Dato* con su nombre, el tipo de dato y un ID para mostrarlo en las listas.
- Las opciones del menú global y un hecho de control (MostrarMenu) para este.
- Un hecho de control para el módulo 3 (MostrarDatos) y un hecho que indica que no se ha seleccionado un tipo de dato (OpcionDisponible <ID>).

### 3.d Módulos que se han desarrollado:

Módulo 1:

- Objetivo: Recomendar una licencia software.
- Conocimiento: Licencias activadas para este módulo.
- Deducciones: No.

Módulo 2:

- Objetivo: Comprobar si dos licencias son compatibles.
- Conocimiento: Todas las licencias disponibles.
- Deducciones: Las implicadas por las reglas dadas por el árbol de decisión.

Módulo 3:

- Objetivo: Medidas a aplicar para el tratamiento de datos en la LOPD y generación de ficheros ARCO
- Conocimiento: Tipos de datos y formato de los archivos ARCO.
- Deducciones: Qué medidas han de implementarse y cuáles no.

### 3.e Esquema de razonamiento del sistema:

En el menú inicial se escoge cuál es el módulo que se va a usar aunque el módulo 1 y el módulo 2 comparten conocimiento funcionan de manera totalmente independiente.

### 3.f La lista de hechos que utiliza el sistema y la forma de representarlos:

#### Una licencia software con sus características:

(deftemplate Licencia  
 (field Nombre)  
 (field Libreria)  
 (field Distribucion)  
 (field Modificacion)  
 (field Patente)  
 (field UsoPrivado)



```
(field SubLicencia)
(field ID)
)
```

**Una licencia software activa para el módulo 1 (debe tener el template anterior con el nombre asociado):**

(PosibleLicencia <Nombre>)

**Un tipo de dato del módulo 3:**

```
(deftemplate Dato
  (field nombre)
  (field tipo)
  (field ID)
)
```

(OpcionDisponible <ID>)

### 3.g Las reglas y hechos de cada módulo:

**En el módulo 1:**

- **ActivarModulo1** – Hace *assert* de los hechos que indican las licencias a usar para este módulo.
- **<Ordinal>Pregunta** – Hace referencia a cada una de las preguntas que se hacen para moverse por el árbol de conceptos. En esta regla se hace la pregunta y se lee la respuesta
- **Pregunta<X>Op<Y>** - Se encarga de resolver la pregunta X cuando se ha dado como respuesta la opción Y esto llevara a otra pregunta o a parar.
- **OpcionInvalidaMod1** – Se encarga de tratar las respuestas que no eran válidas para la pregunta en este caso se vuelve a preguntar lo mismo indicando que el valor introducido no era válido.
- **MostrarDisponiblesLabel, MostrarDisponibles y StopComprobar** – Tras cada respuesta muestra cuáles son las licencias que todavía está considerando el sistema.
- **StopMod1** – Cuando queda una única licencia se encarga de parar este módulo y volver al módulo inicial.

## En el módulo 2:

- **Modulo2MiSoft** – Muestra las licencias disponibles para seleccionar la que vas a usar en su software. Es necesario añadir aquí un printout con el id de la licencia si desea añadir una nueva, aparte de su template.
- **Modulo2Comp** - Muestra las licencias disponibles para seleccionar la del software de terceros que quiere usar. Es necesario añadir aquí un printout con el id de la licencia si desea añadir una nueva, aparte de su template.
- **Compatibilidad<X> o Incompatibilidad<X>** - Son las reglas del árbol de decisión en el orden obtenidos por el *asRules* de R adjuntado. Se llaman compatibilidad si ambas licencias son compatibles en el árbol o incompatibilidad si no lo son.
- **ErrorModulo2** – Trata el caso en el que las licencias no eran correctas. Vuelve al menú principal indicando que la entrada no era válida o se ha producido un error.

## En el módulo 3:

*Nota: entre cada uno de los grupos (se considera un único grupo mostrar y seleccionar una cosa concreta) mostrados existen hechos de control para garantizar que se ejecutan en el orden propuesto.*

- **IniciarModulo3** – Hace *assert* de todos los hechos que indican que un tipo de dato no está seleccionado
- **datosLabel y MostrarDatosPosibles** – Muestra los diferentes tipos de datos (no marcados ya) que el usuario puede seleccionar para indicar que los guarda en su base de datos.
- **SeleccionarDato, SeleccionCorrecta, SeleccionSeguir y SeleccionIncorrecta** – Se encargan de la selección de los datos del apartado anterior, comprobar si este es correcto y en el caso de que sea 0 seguir al siguiente apartado.
- **TiposOrganizaciones** – Muestra los diferentes tipos de organizaciones y lee la selección.
- **SeleccionarOrganizacion y SeleccionOrganizacionIncorrecta** – Tratan la selección de la organización del apartado anterior
- **Lo mismo que se aplica para las organizaciones se aplica para las áreas geográficas con hechos con TiposZonas, TratarZona y TratarZonaIncorrecta.**
- **Lo mismo que se aplica para las zonas ocurre con los usos con TiposUso, TratarUso y TratarUsoIncorrecto.**
- **ComprobarIdentificacion y NoIdentificacion** – Indican respectivamente si los datos usados son suficientes para identificar a una persona o no.
- **MedidasSi** - Indica que has de seguir las medidas de la web

- **MedidasNoPersonal** – Indica que no tiene datos que requieran de la LOPD
- **MedidasNo<Categoria>** - Indica que dicha categoría de medidas no tiene que implementarla y por qué.
- **SeguridadAlta y SeguridadOtra** – Indica el nivel de seguridad que se tiene que aplicar.
- **ARCOSolicitarDatosComunes** – Recoge los datos comunes a todos los datos comunes de los modelos ARCO.
- **LOPDMenu y ErrorMenuLOPD** – Se encarga de solicitar que modelo se desea crear y tratar errores en la selección
- **Crear<Modelo>** - Crear al archivo Modelo.txt y pone la parte inicial
- **EscribirParteComunInicial** – Rellena los datos comunes que todos los archivos tiene al principio tras el título
- **EscribirParteEspecifica<Modelo>** - Rellena los datos específicos del modelo
- **EscribirParteComunFinal** – Rellena la parte final común de todos los modelos, cierra el archivo abierto y vuelve al menú inicial.

#### 4. BREVE MANUAL DE USO DEL SISTEMA

Para iniciar el sistema se cargara en CLIPS el archivo proyecto.clp y se aplicará reset y run.

Para empezar, se seleccionar el número del módulo que se desea usar.

En el módulo 1, se recibirán preguntas de carácter binario para reducir a una única licencia las recomendaciones ofrecidas por el sistema. En cada paso se verán las licencias disponibles y por qué se han descartado las que se han descartado.

En el módulo 2, se seleccionará la licencia que quiere usar para su software y luego la licencia del software de terceros/librería que va a utilizar de una lista de licencias conocidas por el sistema. EL sistema deducirá si son compatibles o no e indicará la razón por la que lo son o no.

En el módulo 3, se iniciará seleccionando múltiples tipos de datos de una lista para indicar que se guardan en la base de datos. De diferentes listas se obtendrán, el tipo de organización, el área geográfica dónde está activa la organización y el uso que se dan a los datos. Una vez hecho el sistema indicará si hay datos suficientes para identificar a una persona según la LOPD y por qué, así como las medidas que se deben implementar para cumplir la LOPD. Luego se procederá a rellenar los datos comunes para los modelos ARCO y se seleccionará uno de ellos que será creado en formato en txt en el área de trabajo.

Tras la ejecución de cada módulo el sistema vuelve al menú inicial de selección de módulo.

Para finalizar el sistema deberá matar el proceso o usar halt en la GUI de Clips si se dispone de ella.