PRÁCTICA 2: SBC

INTELIGENCIA ARTIFICIAL

**Sistema de recomendación de libros**



Isaac Muñoz Busto, Ignasi Sant Albors, Pablo Fonoll Soto

ÍNDICE

[**1. Identificación**](#_q75qzevxhnph) **2**

[El problema](#_5q3aehh2bgh4) 2

[Viabilidad de la solución](#_gdw6n3nmcby5) 3

[Fuentes de conocimiento](#_uecynjb9ripc) 3

[Objetivos](#_23s07kcglhag) 3

[**2. Conceptualización**](#_ll2x6j4pa9b4) **3**

[Conceptos del dominio](#_10al47gom9bi) 3

[Problemas y Subproblemas que intervienen en la resolución](#_lqub0kr6v4e6) 4

[Proceso de resolución.](#_s4yoqd81ch) 6

[Organización problemas y subproblemas](#_w6tkelu8v5ej) 6

[**3. Formalización**](#_h1yq9j6a9lf3) **7**

[Desarrollo de la ontología](#_m59sc04a5zfo) 7

[Método de resolución](#_f7mems66czmc) 7

[**4. Implementación**](#_96z0owp2milt) **8**

[Construcción de la ontología](#_dz6pzacjxjp8) 8

[Módulos](#_bap848mm0x9o) 8

[Prototipos](#_dj8kccrbjp3i) 9

[**5. Prueba**](#_7xg7i49osul0) **9**

[**6. Conclusiones**](#_5nmrhetw0oku) **9**

# 

# **1.IDENTIFICACIÓN**

## **1.1.El problema**

No hay sensación más frustrante que la de tener que dejar un libro a medias porque no se ajusta a tus gustos. No solo por el dinero que has invertido en tu nuevo ejemplar sino también porque ese tiempo (que es más valioso que el dinero incluso) lo podrías haber dedicado a otras actividades o a leer otro libro que sí te guste.

Este problema para los lectores más novatos se asume como inevitable. Creen que debido a su poco conocimiento en este área sumado a la ingente cantidad de elecciones posibles van a tener que pasar por varios fracasos antes de que encuentren un libro que se ajuste a sus gustos.

Pero la verdad es que existe un método infalible para evitar todos estos contratiempos: los sistemas de recomendación basados en el conocimiento.

Esta herramienta de elección de libros es altamente recomendada tanto para los nuevos lectores como para los más expertos (ya que aunque seas un bibliófilo nunca podrás competir contra una base de datos).

Este tipo de sistema de recomendación está inspirado en el tradicional, es decir, cuando ibas a consultar al amigo que más sabía de libros qué libro te recomendaba para el verano y te fiabas de su sugerencia ya que conocía al dedillo tus preferencias.

Pero los SBC tienen una gran ventaja respecto al “sistema tradicional” y esta es la capacidad. Capacidad tanto para almacenar datos sobre millares de libros como para almacenar toda la información necesaria sobre el lector para una buena recomendación.

En el presente trabajo trataremos de generar dicho sistema de recomendación de libros, concretamente uno centrado en libros de ficción.

Para hacerlo, nos centraremos en analizar el perfil del lector al que queremos recomendar el libro (edad, temática preferida …) y el perfil de libro que quiere (número de páginas, dificultad , géneros preferidos ..).

La idea es que el sistema conozca al máximo el perfil del lector para que pueda hacer una recomendación parecida a la que nos haría un amigo aficionado a la lectura o nuestro librero de confianza. Se trata de dar la máxima calidez posible a la recomendación.

Una vez recibida toda la información necesaria, el sistema interpretará preferencias del usuario y dará una recomendación de hasta tres libros adecuados para él.

## **1.2.Viabilidad de construir una solución**

Antes de intentar resolver el problema, debemos saber identificar qué tipo de problema es y cómo puede resolverse, teniendo en cuenta cuestiones como: la información del dominio que se tiene, el tamaño, cómo es la solución que se ha de ofrecer, etc.

Podría enfocarse como un problema de búsqueda, ya que consiste en desplazarse por el espacio formado por los libros de los que se tiene información y construir un conjunto de ellos respetando las restricciones y preferencias del lector que hace la consulta. Sin embargo, no podría resolverse con los métodos de resolución conocidos dado que el espacio de exploración es bastante amplio, y los posibles conjuntos solución podrían ser muchos.

Tampoco lo podemos hacer de forma algorítmica, ya que una función no nos podrá representar todas las decisiones de exploración en este problema.

Sin embargo, puede verse que el volumen de información procedente de nuestras fuentes es muy rico, esto es una característica muy importante a la hora de resolver un problema mediante técnicas de SBC, que es lo que conocemos como conocimiento experto.

Además de la información acerca del dominio de este problema, también es necesario interpretar la información recogida sobre los gustos y características del lector, ya que deberemos evaluar de forma diferente dependiendo de, por ejemplo, su edad.

Por otra parte, necesitamos mostrar una solución justificada como resultado de la ejecución. Es decir, las razones por las cuales se recomienda un determinado libro para un determinado lector.

Por último, vemos que el problema tiene un tamaño adecuado que no constituye una tarea inabordable por su complejidad.

Por todas estas razones hemos llegado a la conclusión de que este problema es abordable a través de la construcción de un SBC.

## 

## **1.3.Fuentes de conocimiento**

Las fuentes del conocimiento para un SBC son un punto muy importante, estas son de donde el sistema “aprende” para poder tomar decisiones. En este apartado explicaremos las fuentes del conocimiento que se utilizan para resolver nuestro problema, estas son:

* Los **lectores**, estos indicarán sus características (como la edad, la cantidad de libros leídos,...) y sus preferencias sobre diferentes criterios de elección (como el número de páginas, el género del libro,...).
* Los **libros**, estos serán los agentes que iremos descartando o puntuando positiva/negativamente a medida que el lector va respondiendo las preguntas que le plantea el sistema.

## **1.4.Objetivos**

El sistema debería cumplir una serie de objetivos para ser capaz de resolver con éxito cualquier situación propuesta:

* **Interactuar** con el usuario para obtener todo el conocimiento relativo a sus gustos literarios.
* **Aprovechar** al máximo el aprendizaje obtenido mediante la base de datos que hemos creado y las respuestas de los usuarios.
* **Ordenar** los libros según cuáles pueden ser más adecuados para la recomendación.
* **Presentar** de forma adecuada y coherente la solución a los usuarios. De manera que se recomienden hasta 3 libros cuya selección es fruto de una interpretación acurada de la información proporcionada por el usuario. Además el sistema proporcionará una breve justificación a su elección

# **2.CONCEPTUALIZACIÓN**

## **2.1.Conceptos del dominio**

**2.1.1.Características de un Libro:**

* Título del libro
* Autor del libro
* Fecha de publicación
* Número páginas
* Temática
* Público
* Precio
* Perteneciente a una saga
* Género protagonista
* Mundo inventado

**2.1.2.Características de un Autor:**

* Nombre del autor
* Nacionalidad

**2.1.3.Características de un usuario:**

* Edad individuo
* Preferencia en cuanto a temáticas
* Preferencia en cuanto a épocas
* Preferencia en cuanto a sagas
* Preferencia en cuanto a dificultad
* Preferencia en cuanto a número de pàginas
* Grado de conocimiento literario
* Preferencia en cuanto a precio

Las preferencias del usuario junto a su perfil ayuda a escoger los libros más adecuados con el fin de ofrecerle una experiencia más personalizada.

## **2.2.Problemas y subproblemas que intervienen en la resolución**

Partimos de una librería diseñada por nosotros con una base de datos compuesta por libros, cada uno con sus respectivos atributos (título, nombre del autor, año de publicación, precio, número de páginas, temática , género del protagonista, perteneciente a una saga y si se desenvolupa en un mundo inventado)

Para cada recomendación, se quiere saber la edad del usuario, si se quiere leer una novela moderna o clásica, si se dispone de mucho tiempo para leer, si le importa el precio del libro y si es así qué precio estaría dispuesto a pagar por él. También se le pide cuánto tiempo dedica a la lectura, de esta manera podremos ajustar aún más el perfil de libro que le conviene.

A partir de toda esta información podemos suponer que conocemos un poco mejor el perfil del usuario y podemos proceder a una recomendación más detallada.

**2.2.1.División de los usuarios según la edad.**

Con el fin de proporcionar una solución más personalizada para la recomendación, una de las preguntas iniciales es “¿Qué edad tiene?”. El usuario nos indica si edad.

* Recomendación infantil (edad menor de 11 años).
* Recomendación adolescente (edad entre 11 y 16 años).
* Recomendación adulta (edad mayor de 16 años).

Hemos fijado un rango de respuesta que va del 5 a 100. El motivo de este rango es que una edad menor que 5 es ilógica ya que apenas se tiene conocimiento del lenguaje y a la vez una mayor de 100 no es usual.

**2.2.2.Creación del conjunto de libros disponibles en nuestra librería.**

Para formar la base de datos de los libros, nos hemos guiado por nuestros conocimientos literarios combinándolos con conocimientos que hemos obtenido de páginas como Amazon, La Casa del Libro, Wikipedia y Megustaleer. De esta manera nos aseguramos disponer de libros de diferentes épocas y estilos y no solo los más famosos.

**2.2.3.Cálculo de la dificultad de un libro.**

Con el fin de determinar la dificultad de un libro partimos de la idea de que la dificultad de un libro está directamente relacionada con el público hacia el que va dirigido, es decir, un libro infantil catalogado como difícil no significa que un adulto pueda encontrar dificultades en la lectura de este. Por otra parte también está relacionado con el número de páginas ya que entendemos que como más páginas tenga, más fácil es perderse en la história y por lo tanto la dificultad de este aumenta. Por último para evaluar la dificultad se han tenido en cuenta aspectos como la concentración que se requiere en la lectura y el tipo de vocabulario que se usa en la narración, siendo un libro difícil aquel en el que aparecen gran cantidad de tecnicismos y que por lo tanto requiere una mayor atención.

**2.2.4.Gestión de la información de la recomendación.**

Entre los problemas con los que hemos tenido que enfrentarnos está el de intentar crear un perfil del usuario para poder hacer una recomendación ajustada a este. Tenemos muchos criterios a tener en cuenta. Para empezar partimos de las preguntas iniciales que se les hace a cada recomendación.

1.¿Qué edad tiene?

Según la respuesta dividimos la recomendación en: recomendación infantil , adolescente o para adultos. Esta información será clave para poder determinar qué libros se adecuan a las capacidades del usuario.

2.¿Dispones de mucho tiempo para leer?

Dependiendo de la respuesta, la recomendación irá más enfocada a libros largos o cortos.

3.¿Te importa el precio de un libro?

Si la respuesta es afirmativa se pregunta lo siguiente:

3.1¿Cuántos euros te gustaria gastarte como mucho?

El rango válido de esa pregunta va des del precio del libro más barato hasta el precio del libro más caro. Los libros que no cumplen con la condición de precio se descartan de la recomendación.

4.¿Cómo te consideras a ti mismo?

Además del perfil del usuario, tenemos que basar nuestra recomendación según sus preferencias y gustos. Para hacerlo se le hacen una serie de preguntas después de las anteriores . En estas el usuario tiene la opción de responder “indiferente” si no quiere que su recomendación se guíe por el campo preguntado.

1. ¿Te gustaría leer un libro moderno o clásico?
2. ¿Te gustaría leer un libro protagonizado por una mujer o hombre?
3. ¿Te gustaría leer un libro perteneciente a una saga?
4. ¿Te gustaría que el libro transcurriese en un mundo imaginario o de fantasía?
5. Escoja sus categorías favoritas

Aquí el usuario puede escoger entre 13 subtemas de la novela de ficción.

1. ¿Le gustaría leer libros en español?

**2.2.5.Recogida de datos de la recomendación.**

Al inicio del sistema el usuario tiene que contestar una serie de preguntas para que se pueda ajustar la recomendación a su perfil ( edad, tiempo para leer, precio deseado y afición a la lectura).

Más tarde procederemos a preguntar sobre los gustos del usuario por lo que se plantean las preguntas vistas anteriormente para conocer sus preferencias respecto a la época de publicación, género del protagonista, espacio en el que transcurre la historia, temática y literatura española.

Todo este proceso facilita que el sistema experto forme una recomendación personalizada como la que nos podría hacer un amigo o nuestro librero.

**2.2.6.Análisis de las preferencias de la recomendación.**

El sistema se dedicara a analizar las respuestas del usuario. Se establecen dos consecuencias posibles a las diferentes preguntas.

La primera consecuencia es eliminar de la recomendación los libros que no siguen las condiciones establecidas. Atienden a esta consecuencia las respuestas de las preguntas de

“¿Qué edad tienes?” y “¿Cuántos euros te gustaria gastarte en el libro?”. Se ha escogido este criterio ya que no se quiere que a un niño se le pueda recomendar un libro de adulto ni tampoco se desea que a un adulto se le recomiende literatura infantil. Por otra parte el dinero que el usuario está dispuesto a gastarse en el libro también limita la recomendación ya que no tiene sentido recomendar un libro a un usuario que no está dispuesto a comprarlo.

La segunda consecuencia es otorgar una puntuación a los libros restantes de la recomendación. Se le otorga una mayor puntuación a esos libros que se acercan más a las preferencias del usuario. Responden a esta consecuencia todas las otras preguntas del sistema.

## **2.3.Ejempos del conocimiento experto**

En la ontología se han incluido 40 libros con sus características. En este apartado se explica cómo se han instanciado algunos de los libros.

**La Catedral del Mar**

Título : La Catedral del Mar

Autor : Ildefonso Falcones

Año de publicación : 2006

Categoría : Novela ficción histórica

Género protagonista : Masculino

Público : Adulto

Número de Páginas : 669

Mundo inventado : false

Saga : false

Dificultad : Difícil

Precio : 11€ (fuente google shopping)

Para completar los distintos campos necesarios para instanciar “La Catedral del Mar” se ha usado el conocimiento de los distintos miembros del grupo combinándolos con la información obtenida de internet. Se ha establecido como “Difícil” ya que la obra hace descripciones detalladas de varios edificios religiosos que requieren una completa atención por parte del lector. A parte, en la obra, se ve una gran cantidad de lenguaje técnico referente al mundo de la construcción y la arquitectura. Por último, al ser una obra de 669 páginas, con abundantes detalles entre sus textos que deben ser comprendidos para entender la obra, se ha decidido catalogarla como “Difícil”.

**Harry Potter**

Título : Harry Potter

Autor : J. K. Rowling

Año de publicación : 1997

Categoría : Novela ficción fantástica

Género protagonista : Masculino

Público : Adolescente

Número de Páginas : 255

Mundo inventado : ture

Saga : true

Dificultad : Fácil

Precio : 8€ (fuente google shopping)

Para completar los distintos campos necesarios para instanciar “Harry Potter” se ha usado el conocimiento de los distintos miembros del grupo combinándolos con la información obtenida de internet. El parámetro Mundo Inventado es “true” ya que las criaturas que aparecen en el libro serían impensables en la realidad. Se ha establecido como “Fácil” ya que la obra es fácil de entender y no contiene vocabulario técnico de ningún tipo. Además el número de páginas es compatible con que sea clasificado como fácil.

## **2.4.Proceso de resolución**

Después de recoger todos los datos necesarios, se recortan las posibles recomendaciones teniendo en cuenta el perfil del usuario. Por otra parte, se asigna una puntuación a los distintos libros que pertenecen a la recomendación, siendo esta más alta cuando se ajusta bien a las preferencias del usuario y siendo baja cuando son contrarias a sus preferencias. Al final le recomendamos los libros que han obtenido una mejor valoración

## **2.5.Organización problemas y subproblemas**

Empezamos con el conocimiento proporcionado por el usuario y sabemos que es un problema que puede ser resuelto por un Sistema Basado en el Conocimiento. Los pasos más destacados que se han de solucionar son: la abstracción de los datos , la obtención de los datos abstractos y la asociación heurística.

# **3. FORMALIZACIÓN**

## **3.1.Desarrollo de la ontología**

**3.1.1.Determinación del dominio y cobertura de la ontología**

En el apartado de Conceptualización hemos expuesto ya todos los conceptos del dominio, con lo que serán estos mismos conceptos los que nos guiaran para construir la ontología. La ontología tendrá que ser capaz de representar tanto la colección de libros y sus características como todos los datos del usuario, ya que su tarea será la de recomendar un libro a estos.

**3.1.2.Construcción de la Ontología**

La ontología se ha construido con la ayuda de Protégé. Se han definido 7 clases: Autor, Categoría, Llibres, País, Protagonista, PublicDestinatari y Dificultad. Después se incluyeron los diferentes Slots , mostrados en los apartados siguientes, y finalmente las Instancias.

Todo este proceso se hizo con sumo cuidado para poder representar todos los conceptos del dominio.

**3.1.3.Definición de clases**

* **Autor**

LlibresAutor: Instancia de Llibres

NomAutor : String

AutorTop: Boolean

PaisAutor : Instancia de Pais

Definimos toda la información referente al autor, lo asociamos a un libro y a un país. El Booleano “AutorTop” indica si el autor tiene cierta fama y es reconocido a nivel nacional si es español o a nivel internacional si es extranjero. Solo los Autores que cumplan esta restricción serán preguntados en el sistema.

* **Categoría**

LlibresCategoria : Instancia de Llibres

NomCategoris : String

Esta clase tiene la información referente a las distintas categorías o temáticas de los libros.

* **Llibres**

AnyPublicacio : Integer

AutorLlibre : Instancia de Autor

CategoriaLlibre : Instancia de Categoría

MundoInventado : boolean

NomLlibre : String

NumPagines : Integer

Preu : Integer

ProtagonizadoPor : Instancia de Protagonista

PublicLlibre : Instancia de PublicDestinatari

Saga : boolean

Esta clase tiene toda la información de un libro, lo asocia con un Autor, una categoría, un protagonista y un destinatario. MundoInventado hace referencia a si la historia sucede en un mundo mágico o irreal.

* **País**

AutorsPais : Instancia de Autor

NomPais : String

Contiene la información referente a los países de los autores.

* **Protagonista**

Género : String

LibrosProtagonizados : Instancia de Llibres

Esta clase informa del género del protagonista del libro.

* **PublicDestinatari**

LlibresPublic : Instancia de Llibres

Public : String

Contiene la información de a qué público va dirigida la obra.

## **3.2.Método de resolución**

**3.2.1.Abstracción de los datos**

En el problema que nos ocupa, consiste en pasar un conjunto de datos que el usuario proporciona a hechos generales como por ejemplo los autores y la temática literaria favoritas para encontrar sus datos generales y prioridades.

**3.2.2.Asociación heurística**

Se combinan con los datos abstractos obtenidos y se encuentra una solución también abstracta. Se hace una poda de las posibles recomendaciones descartando las que menos se ajustan al perfil del lector. Si un parámetro concreto es analizado, el índice de recomendación cambiará teniendo en cuenta las opciones que se hayan escogido, haciendo que el parámetro pueda ser más favorable para poder hacer una recomendación asociada a este.

**3.2.3.Refinamiento o Adaptación**

Transformamos la solución abstracta que hemos obtenido anteriormente en un resultado definitivo. En nuestro caso, todo parámetro que tenga un índice elevado es recomendable para el usuario y es lo que se tendrá en cuenta para hacer la recomendación, ya que estaremos interesados en encontrar los mayores índices.

## **3.3.Justificación método de resolución**

Se ha seleccionado el método de resolución de primero poda y después valoración ya que es el más adecuado y coherente para el problema que nos ocupa. Al descartar los libros que no se ajustan al perfil del usuario nos evitamos situaciones ilógica como podría ser la de recomendar a un adulto un libro infantil. Es cierto que podríamos haber implementado la misma idea con el método de puntuación, asignando puntuaciones exageradamente altas a los libros que se ajustan al perfil del usuario, aún así nos hemos decantado por hacer una poda de las posibles recomendaciones ya que hemos considerado que es una resolución más efectiva.

# **4. IMPLEMENTACIÓN**

## **4.1.Representación de la ontología**

A la hora de la realización de este apartado nos hemos encontrado con un problema que no hemos conseguido solucionar. Cuando intentábamos ejecutar el plugin para la visualización de la ontología nos daba un error de servidor. Hemos buscado como solucionarlo en foros pero no hemos dado con la solución.

## **4.2.Módulos**

**Main**

El módulo Main se compone de funciones útiles para otros módulos y la regla inicial que inicia la ejecución del sistema.

**recopilación-usuario**

Este módulo recopila la información referente a las preferencias del usuario: categoria, precio, público y pais del libro. Esta información se obtiene gracias a una serie de preguntas implementadas con las funciones del módulo MAIN.

**procesado**

Este módulo se encarga de todo el procesado de la información aportada por el usuario para elegir los libros que se le recomendaran. Después de la generación de hechos para facilitar el procesado de las siguientes operaciones, se asignan unas puntuaciones a los libros teniendo en cuenta las preferencias del usuario y su grado de conocimiento literario.

## **4.3.Prototipos**

Primero se hizo un prototipo inicial. La ontología de este consistia unos pocos libros con escasos atributos. Estos libros se diferenciaban mucho entre sí cosa que facilita el correcto desarrollo del prototipo y a la vez agiliza el testeo para ver el correcto funcionamiento de este.

Cuando se obtuvieron los resultados esperados del prototipo inicial se empezó a complicar la otologia, añadiendo más Clases, Slots e Instancias, y a su vez se fue adaptando incrementalmente el código CLIPS para atender a los nuevos requerimientos

Hay una representación adecuada de la ontología

La resolución del problema se divide en módulos usando los subproblemas identificados

La representación del proceso de resolución mediante reglas sigue el proceso de razonamiento establecido por la metodología de resolución de problemas escogida

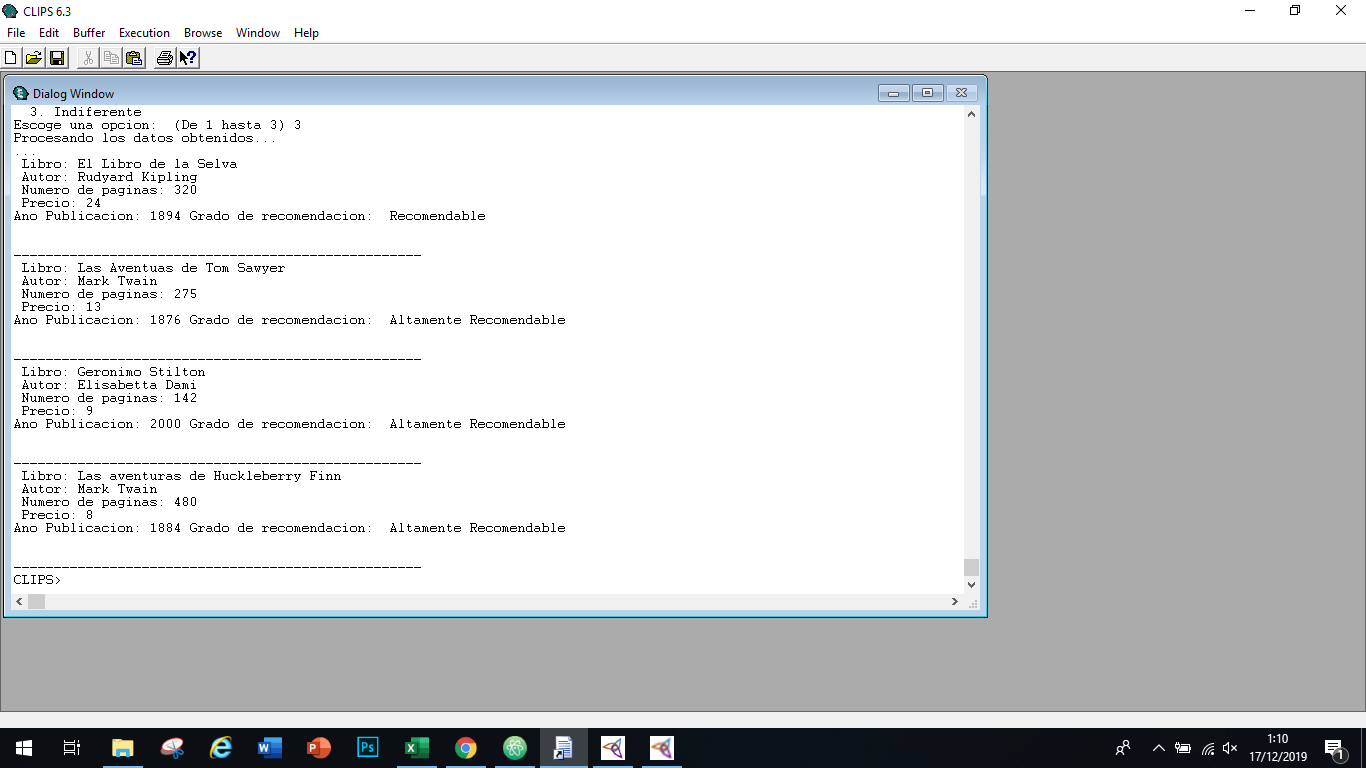
Se ha seguido una metodología incremental, desarrollándose prototipos y se describe como se ha hecho

# 

# 

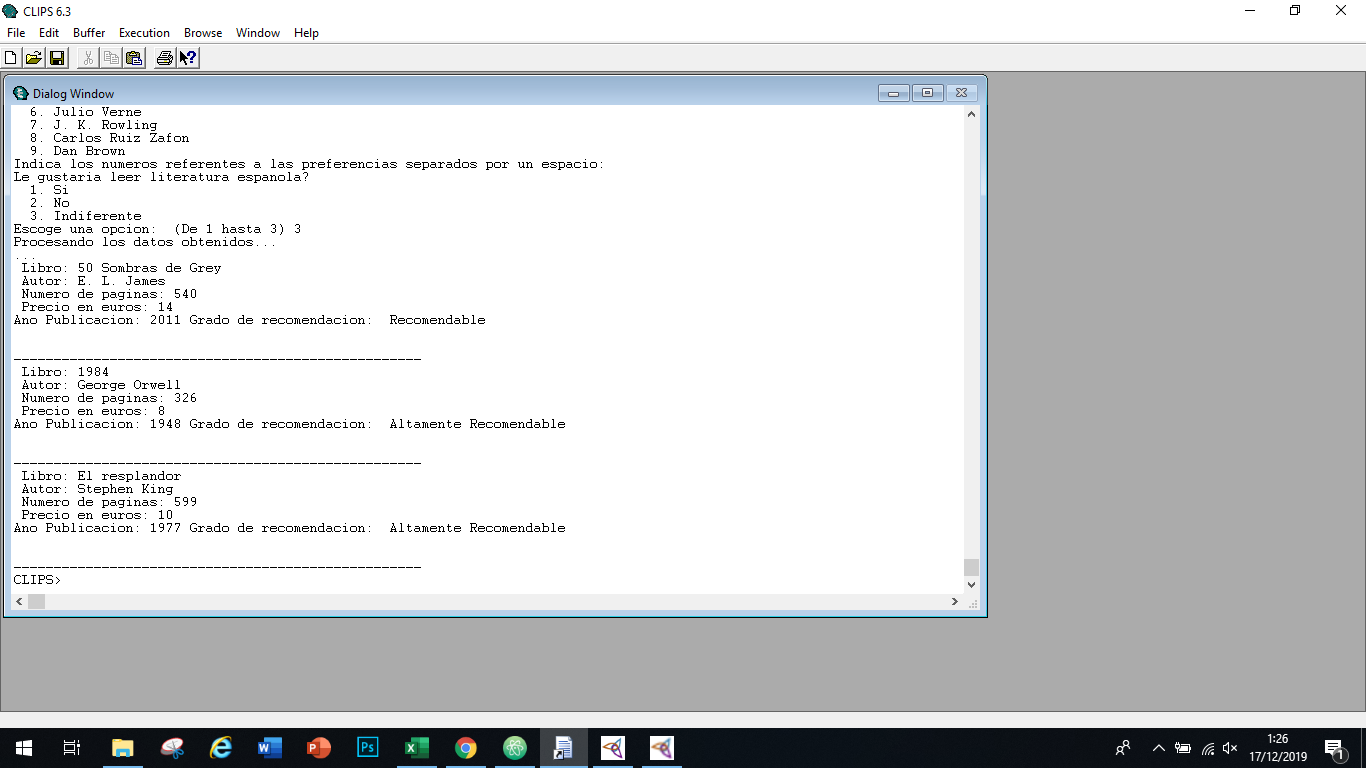
# **5.PRUEBA**

Para los juegos de prueba, probamos un seguido de respuestas para ver si el recomendador nos imprimía los resultados deseados. Por ejemplo, a la hora de comprobar que la restricción del público destino se cumplía, probamos a simplemente introducir la edad que teníamos y, fuesen cuales fuesen las respuestas en las demás preguntas, jamás nos podría aparecer un libro destinado a un público de mayor edad. Ejemplo, poniendo que teníamos 7 años obtenemos la siguiente salida:

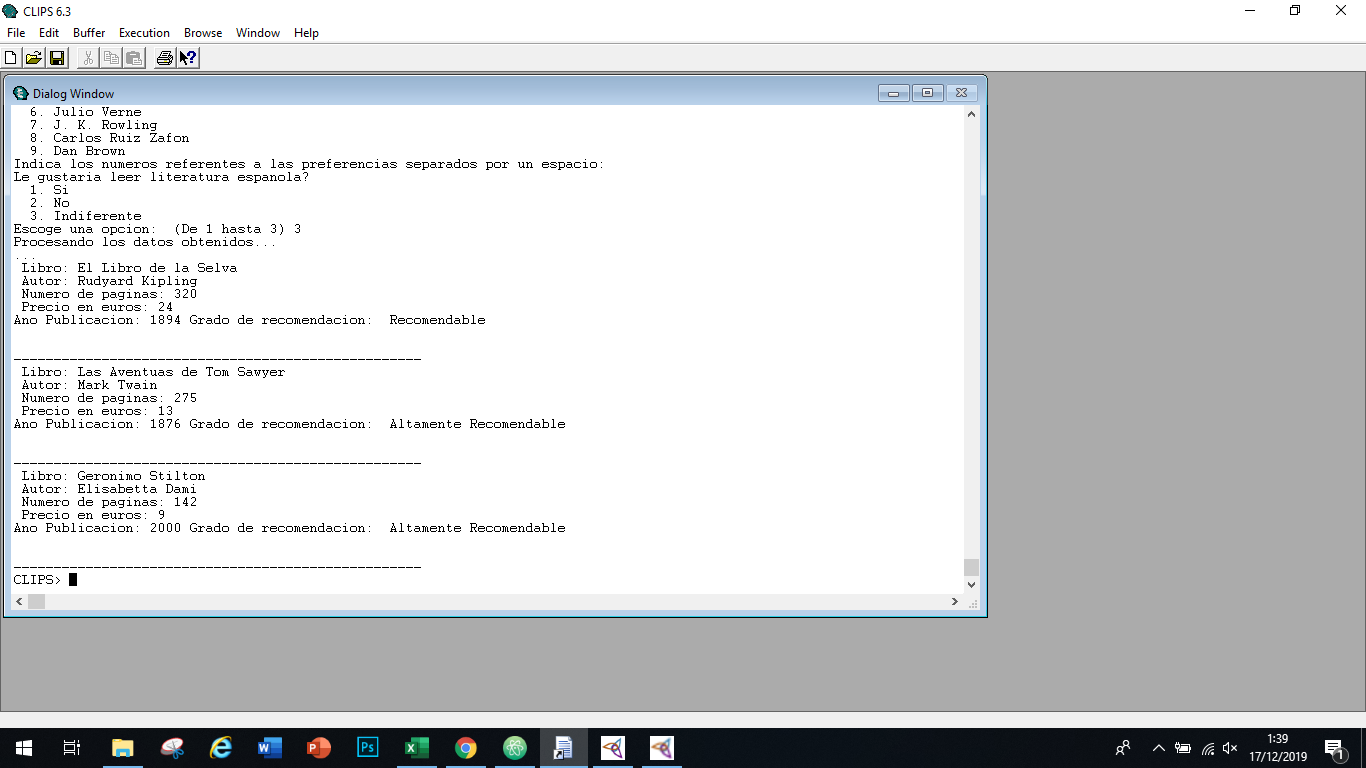


Como podemos observar, solo nos recomiendan libros destinados al público “Infantil”, a niños. También probamos lo mismo por todos los otros públicos, con éxito.

Para probar si el formato de pregunta con índice y con muchas preguntas válidas funcionaba, probamos a marcar varias categorías de libros, y en todas las demás opciones ser lo más neutrales posibles (siempre que se pudiese, marcar la opción “me es indiferente”), y comprobar que la impresión era la adecuada. Por ejemplo, al seleccionar las categorías Terror, Drama y Ciencia Ficción, la salida era la siguiente:



También probamos que si, por ejemplo, un niño quería leer un libro de la categoría Terror, no se mostrase ninguno de esta, ya que todos los libros que pertenecen a esta categoría son para adultos. Ejemplo al seleccionar la edad y categoría favorita terror:



Probamos muchos juegos de pruebas más, como comprobar que la preferencia de autor se puntuaba correctamente, probar que se relacionaba bien los niveles de abstracción( por ejemplo, si indicas que tienes poco tiempo para dedicarle a la lectura, se muestran libros de pocas páginas si es posible).

# **6.CONCLUSIÓN**

Para concluir este trabajo cabe destacar que se han resuelto los diferentes subproblemas para afrontar el problema principal . El sistema cumple con sus objetivos, ya que a partir de la información proporcionada por el usuario ( datos generales y preferencias), se encuentra una recomendación adecuada, es decir , un conjunto de libros que cumplen todas las preferencias o una gran mayoría de ellas.

Por otra parte cabe remarcar que se podrían ejercer varias mejoras al comportamiento del sistema. Por ejemplo podríamos mejorar los ratios del índice de recomendación para dar más importancia a algunos campos, se ha de destacar la dificultad de este punto ya que es algo muy subjetivo (¿A que le damos más importancia , a que la recomendación cumpla la preferencia en número de páginas o en precio?)

Finalmente queremos destacar la importancia de los SBC y de la gran cantidad de campos donde los podemos aplicar.

Su importancia reside en que podemos disponer del conocimiento experto necesario, preservarlo y usarlo en cualquier momento. Si que es cierto que el coste para crear estos sistemas es elevado y que hay que tener un problema con una complejidad suficiente que justifique el coste de su desarrollo, pero una vez creado podemos usar todo ese conocimiento por un coste muy reducido. Aún así también hemos de tener en cuenta la fragilidad de esos sistemas y la difícil integración de aprendizaje en estos.