

Pràctica de Sistemes Basats en el Coneixement  
Intel·ligència Artificial

Bernat Farrero Badal  
Jordi Romero de Villalonga

12 d'Abril de 2010

# Índex

<b>1</b>	<b>Introducció</b>	<b>3</b>
1.1	Metodologia per l'anàlisi del problema . . . . .	3
<b>2</b>	<b>Identificació del problema</b>	<b>5</b>
2.1	Identificació del problema . . . . .	5
2.2	El SBC és viable . . . . .	5
2.3	Fonts de coneixement . . . . .	5
2.4	Objectius del sistema . . . . .	6
<b>3</b>	<b>Conceptualització</b>	<b>7</b>
3.1	Elements del domini . . . . .	7
3.2	Descomposició en subproblemes i flux de raonament . . . . .	8
3.3	Abstraccions, Evidències i hipòtesis . . . . .	9
3.3.1	Edat de l'usuari . . . . .	9
3.3.2	Ocupació del lector . . . . .	9
3.3.3	Sexe . . . . .	9
3.3.4	Gèneres . . . . .	9
3.3.5	Llocs de lectura . . . . .	10
3.3.6	Temps de lectura . . . . .	10
3.3.7	Influència del preu . . . . .	10
3.3.8	Format del llibre . . . . .	10
3.3.9	Popularitat dels llibres . . . . .	11
3.3.10	Dates de publicació dels llibres . . . . .	11
3.3.11	Estil del llibre . . . . .	11
3.3.12	Autor preferit . . . . .	11
3.3.13	Llengües . . . . .	11
3.3.14	Vendes . . . . .	12

<b>4</b>	<b>Formalització del problema</b>	<b>13</b>
4.1	Formalisme de representació del coneixement . . . . .	13
4.2	Mida de l'espai de cerca . . . . .	13
4.3	Metodologia de resolució del problema . . . . .	14
4.4	Tractament de la incertesa . . . . .	14
<b>5</b>	<b>Implementació</b>	<b>16</b>
5.1	Procés d'implementació . . . . .	16
<b>6</b>	<b>Ontologia</b>	<b>17</b>
6.1	Fases de desenvolupament de l'ontologia . . . . .	17
6.2	Canvis conceptuals i decisions preses . . . . .	17
6.3	Descripció formal . . . . .	18
6.3.1	Frame Llibre . . . . .	18
6.3.2	Frame Autor . . . . .	19
6.3.3	Frame Gènere . . . . .	19
6.3.4	Relació pertany_a . . . . .	19
6.3.5	Relació escrit_per . . . . .	19
<b>7</b>	<b>Descripció dels mòduls</b>	<b>20</b>
7.1	Principal . . . . .	20
7.2	Perfil Lector . . . . .	20
7.3	Preguntes Comunes . . . . .	20
7.4	Preguntes Específiques . . . . .	20
7.5	Associacions Incondicionals . . . . .	21
7.6	Esborrar instàncies . . . . .	21
7.7	Associacions heurístiques . . . . .	21
7.8	Refinament . . . . .	21
<b>8</b>	<b>Casos de prova</b>	<b>22</b>
8.1	Experiment Bernat . . . . .	22
8.2	Experiment Lucia . . . . .	26
8.3	Experiment Gerard . . . . .	29
8.4	Experiment Jordi . . . . .	33
8.5	Experiment Xavier . . . . .	37
8.6	Experiment Serafin . . . . .	40
8.7	Experiment Jordi 2 . . . . .	44
<b>9</b>	<b>Conclusions</b>	<b>48</b>
<b>10</b>	<b>Annex: Entrevista a un llibreter</b>	<b>49</b>

# Capítol 1

## Introducció

En aquesta pràctica es planteja el problema de crear un Sistema Basat en el Coneixement per recomanar llibres adients als usuaris segons les seves preferències i característiques. En l'actualitat hi ha una enorme oferta de llibres de tots els tipus i gèneres, precisament per això resulta difícil per un usuari desinformat destriar entre tots ells quin és el llibre perfecte. Evidentment, el procés de recomanació d'un llibre és quelcom que va més enllà d'elegir gènere, tema, popularitat del llibre, etc. Per recomanar adequadament un llibre cal conèixer bé l'usuari i emular el procés de recomanació d'un llibreter que ens coneix de tota la vida.

En el nostre cas el domini de cerca es restringeix a les novel·les de ficció, tot i que dins d'aquest gènere hi ha molts subgèneres i llibres de molt diferent naturalesa. El Sistema Basat en el Coneixement que construïm recollirà eficientment informació de l'usuari i restringirà progressivament el domini. El sistema ha de comptar amb una base de dades d'instàncies de llibres significativa per poder donar respostes adequades.

En primer lloc, el sistema començarà identificant l'usuari, li realitzarà preguntes comunes i després preguntes específiques pel seu perfil, traurà conclusions esborrant opcions no adequades, realitzarà l'associació heurística segons els criteris extrets, es refinarà el resultat (categoritzant entre opcions més recomanables que altres) i, finalment, se li retornarà a l'usuari les 3 millors opcions de lectura recomanades ordenades per adequació. Evidentment, si la base de dades no ha estat capaç de trobar cap llibre per aquell usuari, se li retornarà menys llibres o bé cap.

### 1.1 Metodologia per l'anàlisi del problema

La metodologia seguida per afrontar el problema es basa en un cicle de vida en cascada que també encaixa alhora amb la metodologia basada en el prototipat ràpid i el disseny incremental. Consta de les següents fases:

- Identificació del problema: On es determina la viabilitat de la construcció del SBC i la disponibilitat de les fonts de coneixement.

- Conceptualització: Descripció semiformal del coneixement del domini del problema i descomposició en subproblemes, segons la visió d'un expert.
- Formalització: Cal definir el mecanisme adequat de representació del coneixement, en aquest cas segons la visió de l'enginyer de coneixement.

Seguidament a aquestes fases, ve la implementació i les proves del sistema. Com que es tracta d'un procés incremental, s'ha començat amb un prototipat senzill, que s'ha anat extenent i refinant a mesura que s'ha anat provant, amb els canvis que això ha suposat al model final d'ontologia, la descomposició de subproblemes i el desplegament de casos.

## Capítol 2

# Identificació del problema

### 2.1 Identificació del problema

La identificació del problema és la primera fase de l'anàlisi del problema. En primer lloc, consisteix en analitzar la viabilitat del sistema de coneixement (SBC) proposat. Cal justificar el sistema, demostrar que és adequat i que és capaç d'arribar a solucions satisfactòries. És també necessari especificar quines seran les fonts de coneixement que tindrem i quin d'aquest coneixement serà necessari pel sistema. Determinarem quins són els objectius del sistema, i quins tipus de sortides serà capaç d'oferir. També detallarem les hipòtesis prèvies que assumim.

### 2.2 El SBC és viable

Si es vol que el sistema sigui pràctic de cara a l'usuari, les bases de dades de llibres amb les que treballa han de ser molt grans, parlariem de l'ordre de milions de llibres. Evidentment, si el sistema estigués basat en qualsevol algorisme de cerca exhaustiva sobre tots aquests llibres, els resultats serien dolents i es trigaria molt en arribar a una solució. Com que el problema es defineix com un conjunt de restriccions i preferències, deduïm fàcilment que es pot resoldre amb un sistema basat en regles, com ara el que ofereix CLIPS.

En quant a la viabilitat econòmica del projecte, considerem que el mercat de les ventes de llibres per Internet es troba en constant creixement, on cada dia es presenta una nova plataforma de venda de llibres en línia. La majoria d'aquestes plataformes ofereixen avançats sistemes de cerca i molta informació a l'usuari, però no ofereixen sistemes intel·ligents de recomanació. Per tant, seria una aportació útil i innovadora pel mercat.

### 2.3 Fonts de coneixement

La informació disponible dels llibres l'hem extret de diverses llibreries en línia. En particular, hem usat Amazon.com, La Central, Barnes&Noble i La Casa

del Llibre. Les instàncies particulars les hem extret (mitjançant la programació d'un programa *scraper*) dels llocs web esmentats.

Evidentment, podem comptar a priori només amb aquella informació que és pública: la descripció dels llibres i tots els seus camps característics i de l'autor, tot i que no podem conèixer a priori detalls personals de l'autor ni quines influències ha tingut per comparar-les amb l'usuari. Tampoc detalls sobre la trama del llibre, que caldrà possiblement estudiar. Per altra banda, tenim la possibilitat en aquesta pràctica de permetre que el sistema aprengui noves regles a força d'ús. Per exemple, seria interessant que el sistema pogués rebre una valoració positiva quan una recomanació és adequada o negativa quan no ho és, i corregir o reescriure automàticament algunes de les seves regles (sistema conegut com aprenentatge basat en explicacions).

Les evidències, les hipòtesis, el mètode d'abstracció de dades (definicional, qualitatiu i de generalització) i els heurístics usats sobre el problema els hem obtingut contactant amb un expert, que en aquest cas és un lector habitual. En particular, hem entrevistat el pare d'un de nosaltres, Jordi Romero (entrevista a l'annex, capítol 10) que publica un blog de lectura. També hem obtingut informació de la valoració que ens han fet companys i amics i de consells que hem trobat en llibreries en línia.

El coneixement necessari pel sistema es divideix en aquell coneixement referent als llibres, els seus gèneres i els seus autors, però també el coneixement de l'usuari, que tant pot ser la descripció del seu perfil com les seves preferències personals en quant al domini que ens afecta.

## 2.4 Objectius del sistema

L'objectiu del sistema és generar una llista de **tres** llibres recomanats ordenat per grau d'adequació. Ara bé, si l'usuari ha respost de manera clarament negativa totes les preguntes llavors no se li recomanarà cap llibre. De la mateixa manera, si l'usuari respon amb indiferència totes les preguntes, la recomanació vindrà donada només per les característiques del seu perfil. S'entén que en condicions normals, i amb un domini suficientment gran, haurien d'existir almenys tres llibres recomanables per qualsevol combinació de preferències i restriccions de l'usuari.

## Capítol 3

# Conceptualització

En aquesta fase s'intenta adquirir la perspectiva de l'expert. De la interacció amb l'expert o experts en la recomanació de llibres n'hem d'extreure la dimensió final del domini i l'ontologia. A la vegada, hem de descomposar el problema en subproblemes i realitzar un anàlisi per refinaments successius amb l'objectiu d'obtenir la relació jeràrquica entre les diferents fases de resolució fins als operadors de raonament més elementals. Cal, doncs, identificar quin és el flux de raonament necessari en la resolució del problema des del punt de vista de l'expert.

### 3.1 Elements del domini

Els primers elements a considerar són aquells que formen part de l'ontologia permanent del sistema, és a dir, els llibres, els autors i els gèneres.

- Els llibres: El seu títol, el preu, l'any de publicació, el format, l'idioma, el número de pàgines, si és un best-seller o no ho és, l'enquadrernació, l'edat a la qual està orientat, el número de vendes, la trama, el vocabulari usat i si hi apareixen un gran nombre de personatges en són les seves característiques directes principals.
- Els autors: El seu nom, la nacionalitat i el seu any de naixement.
- Els gèneres: Dins del gènere de ficció, hi ha molts gèneres diferents. En guardem el seu nom.

També forma part del domini l'usuari: les seves característiques i preferències comunes i específiques (en funció de les respostes a les preguntes comunes).

- Característiques: El seu nom, l'edat, el sexe, la llengua materna i el seu perfil d'ocupació.
- Preferències comunes: La freqüència i el temps de lectura, el gènere preferit dins de tots els disponibles al sistema, autor preferit dels que hi ha al sistema, si li agrada experimentar amb nous gèneres, lloc habitual de



lectura, tipus d'enquadració preferida, importància del preu, preferència per llibres actuals, preferència per best-sellers, preferència per llibres gruixuts, preferència per llibres amb vocabulari senzill, preferència per llibres de trama enrevesada, preferència per llibres amb molts personatges, origen preferit de l'autor i llengua de lectura preferida.

- Preferències específiques: Preferència pels tipus de lletra grans, per aquells que no són ni nens petits ni persones grans; preferència per llibres lleugers per aquells que llegeixen fora de casa; preferència per llibres amb contingut explícit, per usuaris que no són nens; preferència per llibres més venuts per aquells que no són fans dels best sellers; preferència per una franja de preu determinada, per aquells a qui influeix el preu.

Per últim forma part també del domini les recomanacions emeses pel sistema, que estan ordenades per ordre de més apropiades a menys. Les tres millors són les que es retornen a l'usuari.

### 3.2 Descomposició en subproblemes i flux de raonament

Partim de la presumpció que tot llibre és recomanable per l'usuari en un inici.

En la primera interacció amb l'usuari se li demanarà les característiques que el defineixen. Un cop assimilades, caldrà descartar, si s'escau, aquells llibres incompatibles amb el seu perfil. Per exemple, si es tracta d'una persona major de 60 anys es descarten aquells llibres de butxaca o amb lletra molt petita o bé si es tracta d'un infant se li recomanaran llibres de gènere infantil i per tant es descarten els altres. La resta de llibres no descartats podran veure modificada la seva puntuació de cara a la recomanació. Per exemple, si el perfil d'ocupació és acadèmic, la puntuació dels llibres clàssics o històrics es veurà incrementada.

Després es procedirà a fer les preguntes comunes. En funció de la rotunditat o incertesa de les respostes es procedirà sumar punts, restar-ne o bé esborrar aquells llibres en funció de les evidències i hipòtesis prèviament definides. Un cop respostes les preguntes comunes, es tindrà un *ranking* limitat de llibres (sempre en funció del nivell de certesa de les respostes) i estarà ordenat de més recomanable a menys.

Finalment es passarà a realitzar les preguntes específiques, en funció del perfil de preferències que el sistema ha obtingut de l'usuari. Per exemple, l'usuari que ha respost que li influeix el preu se li pregunta entre quins marges es pot moure, o l'usuari que llegeix fora de casa se li pregunta fins a quin punt el pes del llibre pot resultar un problema. Igual que en l'anterior fase, les respostes que dona l'usuari serviran per eliminar llibres o per variar-ne la seva puntuació. Quan s'han acabat de fer totes les preguntes s'analitzarà quins llibres són els més adequats i s'establirà l'ordre final. S'oferirà els 3 millors llibres a l'usuari com a recomanació.

### 3.3 Abstraccions, Evidències i hipòtesis

Després d'entrevistar el nostre expert, n'hem extret unes quantes hipòtesis i evidències que s'acostumen a tenir en compte en un procés de recomanació de llibres i que serviran en la següent fase per elaborar el procés d'associació heurística i refinament. Aquí les detallarem per temes, en aquesta fase es mostren tal com l'expert ens les ha especificat.

#### 3.3.1 Edat de l'usuari

Evidències:

- Als nens se'ls recomana orientats a públic infantil.
- Als joves se'ls recomana orientats a públic juvenil.
- Als adults se'ls recomana llibres orientats a públic adult.
- Als nens no se'ls recomanen llibres amb contingut explícit.
- Els llibres orientats a públic juvenil se'ls recomana als joves principalment i als adults en menys grau.

#### 3.3.2 Ocupació del lector

Hipòtesis:

- Als desocupats i als estudiants sovint els preocupa més el preu dels llibre

#### 3.3.3 Sexe

Hipòtesis:

- Als usuaris homes se'ls recomana en més grau la novel·la històrica i en menys grau la novel·la romàntica.
- Als usuaris dona se'ls recomana en més grau la novel·la thriller i la novel·la negra.

#### 3.3.4 Gèneres

Evidències:

- Se li recomanen preferentment els llibres del gènere preferit de l'usuari.

Hipòtesis:

- Si a l'usuari li agrada experimentar, se li poden recomanar altres gèneres a part del seu preferit.

- Als usuaris amb gènere preferit la ciència ficció també els pot agradar la fantasia (i viceversa).
- Als usuaris amb gènere preferit la novel·la històrica també els pot agradar els clàssics (i viceversa).

### 3.3.5 Llocs de lectura

Hipòtesis:

- Els usuaris que llegeixen en transports públics prefereixen llibres de butxaca o tapa tova.

### 3.3.6 Temps de lectura

Hipòtesis:

- Als usuaris que llegeixen molt se'ls recomanen llibres amb trama i vocabulari més complexos. Altrament als usuaris que llegeixen poc.
- Als usuaris que dediquen més temps a llegir se'ls recomana llibres gruixuts.

### 3.3.7 Influència del preu

Evidències:

- Els usuaris que els influeix molt el preu convé recomanar-los llibres de butxaca i preguntar-los en quina franja es poden moure.

### 3.3.8 Format del llibre

Hipòtesis:

- Els usuaris que són persones grans prefereixen un tipus de lletra gran.
- Els usuaris que són nens prefereixen un tipus de lletra gran.
- Els usuaris que prefereixen un tipus de lletra gran no se'ls recomana llibres de butxaca.
- Els usuaris que prefereixen els llibres lleugers se'ls recomanen les edicions de butxaca.
- Els usuaris que prefereixen els llibres ben enquadernats se'ls recomana en més grau els llibres de tapa dura i els de tapa tova, i en menys grau els de butxaca.
- Hi ha usuaris que prefereixen llibres gruixuts, i cal preguntar-los-ho.

### 3.3.9 Popularitat dels llibres

Hipòtesis:

- A tots els usuaris que no especifiquin el contrari, se'ls recomana un llibre més venut abans que un llibre menys venut.
- Als usuaris que no els agraden els best-sellers, no se'ls recomana llibres només perquè siguin els més venuts.

### 3.3.10 Dates de publicació dels llibres

Hipòtesis:

- Usuari que li agraden les novetats se li recomanen els llibres de l'últim any.
- Usuari que li agraden els llibres actuals se'ls recomanen els llibres dels últims 10 anys.

### 3.3.11 Estil del llibre

Hipòtesis:

- Als usuaris amb perfil acadèmic se'ls recomana novel·la històrica, clàssics o bé llibres amb trama complexa o vocabulari complex.
- Cal preguntar als usuaris si prefereixen llibres amb vocabulari senzill o més complicat i recomanar-los llibres d'acord amb la seva resposta.
- Als usuaris que els agrada prendre notes al costat dels llibres se'ls recomana preferentment llibres de trama i vocabulari més complex.

### 3.3.12 Autor preferit

Hipòtesis:

- L'usuari que té per autor preferit X, li agradaran els llibres del mateix estil que l'autor X.
- Hi ha usuaris que prefereixen llibres escrits per autors d'una nacionalitat concreta, cal preguntar-ho i recomanar-los d'acord amb les respostes.

### 3.3.13 Llengües

Hipòtesis:

- Es recomanen amb més grau els llibres en la llengua materna de l'usuari.
- Cal preguntar als usuaris en quin idioma prefereixen llegir i recomanar-los llibres d'aquell idioma.

### **3.3.14 Vendes**

Abstraccions:

- Els llibres amb menys de 500.000 exemplars venuts són relativament poc venuts.
- Els llibres entre 500.000 i 5.000.000 d'exemplars venuts són bastant venuts.
- Els llibres amb més de 5 milions d'exemplars venuts són molt venuts.

## Capítol 4

# Formalització del problema

### 4.1 Formalisme de representació del coneixement

Per representar la informació del banc de llibres usarem una ontologia. En la fase anterior hem estudiat quins són els elements del nostre domini i ara determinem quins d'aquests elements seran representats per l'ontologia. En el cas del nostre problema en concret, l'ontologia la formaran els llibres, els gèneres i els autors. Tota aquella informació permanent que l'adquirent del nostre SBC mantindrà, usará i actualitzarà estarà estructurada dins l'ontologia. Tanmateix, aquella informació provisional com és el perfil de l'usuari i les recomanacions emeses pel sistema es trobaran en el propi llenguatge de regles. El detall concret de l'ontologia es troba en el capítol 6 d'ontologia.

Per programar l'aplicació s'usará el programa basat en regles CLIPS. CLIPS és una eina per la producció i execució de sistemes experts escrita en C. Les regles descrites s'expliquen en el capítol 7 d'aquest document.

El problema proposat és un problema d'anàlisi. Parteix d'una situació inicial on tot és solució i mitjançant l'abstracció de dades, l'aplicació de regles (associació heurística) i posterior refinament s'arriba a una solució bona per l'usuari. Aquesta metodologia de resolució de problemes s'anomena classificació heurística i en els propers apartats s'explicarà amb més detalls.

### 4.2 Mida de l'espai de cerca

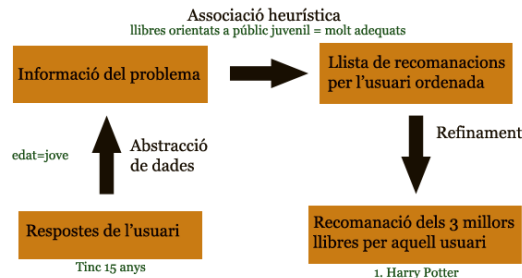
L'espai de cerca consta bàsicament dels llibres del sistema. Com que els llibres s'hauran de recórrer per cada criteri que es pregunta, l'espai de cerca serà equivalent a la multiplicació dels criteris pel nombre de llibres. Evidentment, el fet d'eliminar llibres de la llista reduirà l'espai de cerca. Per tant, un usuari que sigui molt taxatiu descartant llibres farà minvar ràpidament el domini de llibres a recomanar.

### 4.3 Metodologia de resolució del problema

La metodologia de resolució del problema que s'usa és la **classificació heurística**. La seva característica principal consisteix en el fet que l'expert escull una categoria concreta d'un conjunt de solucions prèviament enumerat. Aquest és precisament el cas que ens afecta. Es parteix d'un conjunt de solucions que són tots els llibres i mitjançant un procés d'anàlisi i de classificació (basat en les respostes de l'usuari) anem descartant i avaluant el domini.

La classificació heurística es divideix en tres etapes. La primera etapa és l'**abstracció de les dades**. Aquesta pot ser definicional, per exemple un usuari que li agrada prendre notes al costat d'un llibre es defineix com a usuari complex, o bé pot ser qualitativa, per exemple l'edat de l'usuari al sistema es classifica com a nen, jove, adult o persona gran. Per últim pot ser de generalització, com és el cas dels usuaris a qui agraden els best-sellers, s'enregistren com a amants dels top-ventes. La segona etapa a realitzar és l'**associació heurística** i es realitzarà mitjançant les regles definides de les hipòtesis i evidències extretes de la interacció amb l'expert en l'apartat de conceptualització. Per últim, resta l'etapa de **refinament**, on s'acabaran ordenant i escollint els millors tres llibres per retornar a l'usuari, sempre i quan no hagin estat tots els llibres descartats. Un petit exemple es pot veure en la figura 4.1.

Figura 4.1: Esquema del procés resolució del problema, amb un petit exemple



### 4.4 Tractament de la incertesa

Per fer un sistema que sigui natural per l'usuari s'ha d'acceptar la incertesa habitual del món real. Això vol dir que les respostes a les preguntes no han de ser si o no, sinó que s'ha d'oferir un més ampli espectre possible de certesa. En el nostre cas, aquesta incertesa l'introduïm en el segon i tercer blocs de preguntes. En el primer no s'accepta perquè preguntem valors deterministes, com per exemple el nom i l'edat. El segon i tercer bloc de preguntes, en canvi, s'admetran les respostes Molt, Bastant, Indiferent, Poc o Gens. En cas que sigui gens generalment es procedirà a eliminar els llibres de la llista, si es tracta de poc es restarà punts, si el criteri és indiferent es deixarà igual, i si és bastant o molt es sumaran punts als llibres coincidents.

Per tal de no cometre errors en el refinament, el que hem fet en una segona iteració del procés incremental de desenvolupament del nostre sistema és mantenir dos atributs de valoració dels llibres. Un per quantificar les valoracions negatives i l'altre per quantificar les valoracions positives. Els casos en que la resposta és Indiferent no aporta cap informació nova al sistema.



## Capítol 5

# Implementació

### 5.1 Procés d'implementació

La implementació s'ha dut a terme en el següent ordre. S'ha començat creant una versió preliminar de l'ontologia mitjançant el programa Protégé. Un cop feta, s'han instanciat uns pocs llibres i autors d'exemple (també gràcies al programa Protégé). Llavors s'han exportat a CLIPS i s'ha passat a implementar el mòdul del perfil d'usuari. Posteriorment s'ha implementat el mòdul de preguntes comunes i finalment el de preguntes específiques. Mentre s'ha anat fent s'ha revisat constantment l'ontologia i les regles coincidint amb cada fase de la metodologia d'anàlisi del problema. Les regles s'han anat implementant progressivament i s'ha anat provant amb petits casos fixats cada vegada. Quan l'ontologia ha estat considerada definitiva s'ha creat un nombre més gran d'instàncies que podien representar tots els possibles casos del domini. A mesura que s'han seguit implementant noves regles, el procés d'associació heurística i refinament s'ha anat provant i assegurant-se que els resultats obtinguts eren coherents. Això s'emmarca dins de la metodologia de disseny incremental i prototipatge ràpid que hem decidit seguir des d'un inici.

## Capítol 6

# Ontologia

### 6.1 Fases de desenvolupament de l'ontologia

En un principi no vam dedicar temps significatiu a l'anàlisi de l'ontologia. Vam fer un model simple que es pogués usar i que ens permetés avançar en les diferents fases de resolució del problema. Tanmateix, superada la fase de conceptualització, ja teníem suficient coneixement del domini com per definir el que seria l'ontologia definitiva.

Per desenvolupar l'ontologia vam començar determinant el domini i la cobertura que havia de tenir, realitzant-nos a nosaltres i a l'expert les anomenades preguntes de competència, és a dir, quin tipus de preguntes i respostes es requeria fer sobre el domini. Posteriorment, vam buscar ontologies de llibres i autors existents, per si era possible reutilitzar algunes parts (sobretot basant-nos en les llibreries en línia). Llavors vam procedir a enumerar tots els termes importants que havia de tenir l'ontologia (pensant les propietats rellevants que podia tenir un llibre o un autor), ajudats de la interacció amb l'expert. Vam traslladar això a classes i establir-ne una jerarquia, tot i que en el nostre cas no vam trobar cap generalització i la comunicació de classes es resol mitjançant relacions. Per últim, quan ja estava ben definida, ens vam dedicar a crear instàncies.

### 6.2 Canvis conceptuals i decisions preses

En un inici les característiques dels llibres de la nostra ontologia eren purament físiques. Parlàvem del pes, de la mesura, l'enquadrernació, el número de pàgines, etc. però no parlàvem del contingut, ja que això era quelcom difícil d'extreure de les llibreries en línia. Com que teníem la intenció que el nostre sistema de recomanació fos acurat, vam decidir introduir-hi informació sobre la trama, l'estil i els personatges dels llibres, malgrat que això compliqués el procés d'instanciació (que esdevindria manual i molt més lent). Creiem que fent preguntes més personals a l'usuari es pot determinar amb més exactitud quin tipus de llibre li escau més. També hem introduït l'any de publicació per resoldre el problema dels llibres clàssics, que generalment se'n fan moltes edicions i la data d'edició

no es correspon amb l'any de publicació. Per això hem creat dos atributs per ambdues coses.

Una altra decisió presa és l'entitat de gènere. En un primer moment ens vam plantejar que gènere fos un atribut multivaluat dins de Llibre però després vam veure que, si l'ontologia s'havia de poder mantenir, s'hauria de poder instanciar nous gèneres i que el sistema continués responent. Per això, i per fer més eficient les cerques per gènere, vam decidir donar l'entitat de classe a gènere, que tindria un atribut nom. La relació entre gènere i llibre seria de pertinença de gènere a llibre.

Una altra important decisió presa és la de no incorporar l'usuari a l'ontologia. Com hem explicat anteriorment, la nostra ontologia reflexa tota aquella informació que és permanent (o anterior al procés de recomanació) i que ha de mantenir el propi llibreter o l'adquirent del SBC. La instància de l'usuari viu durant el procés de recomanació i mor amb aquest.

## 6.3 Descripció formal

### 6.3.1 Frame Llibre

- **Títol**
- **Preu:** L'import que costa el llibre.
- **Any edició:** L'any en que l'editorial va editar el llibre.
- **Any publicació:** Any en que es va crear la primera edició del llibre.
- **Format** [gran, mitja, petit]: És la mesura del llibre, informació disponible en qualsevol llibreria.
- **Pes** [gran, mitja, petit]: El pes del llibre, abstret qualitativament a aquestes 3 opcions.
- **Idioma** [català, castellà, altres]: L'idioma en que està escrit el llibre.
- **Número de pàgines**
- **Best-seller?:** Si és considerat com a *Best Seller* pels grans venedors.
- **Enquadrernació** [tapadura, tapatova, butxaca]
- **ISBN:** Codi únic d'identificació del llibre.
- **Orientació** [infantil, juvenil, adult]: El públic al qual està orientat el llibre.
- **Vendes anuals:** Número aproximat de vendes anuals del llibre.
- **Trama:** [simple, complexa]: Com és la trama del llibre.
- **Vocabulari:** [simple, complex]: Com és el vocabulari i l'estil narratiu del llibre.

- **Personatges:** [molts, pocs]: Si es tracta d'una novel·la amb molts personatges o bé pocs.
- **Contingut explícit?:** Si conté violència o sexe.

### 6.3.2 Frame Autor

- **Nom:** El nom de l'autor.
- **Any naixement:** L'any de naixement de l'autor.
- **Nacionalitat:** La nacionalitat d'on l'autor és originari.

### 6.3.3 Frame Gènere

- **Nom:** En un principi instanciaré Ficció, Fantasia, Ciència Ficció, Terror, Thriller, Clàssics, Històrica i Social.

### 6.3.4 Relació pertany\_a

- **Domini:** Llibre
- **Rang:** Gènere
- **Cardinalitat:** 1
- **Inversa:** té\_llibres

### 6.3.5 Relació escrit\_per

- **Domini:** Llibre
- **Rang:** Autor
- **Cardinalitat:** 1
- **Inversa:** ha\_escrit

## Capítol 7

# Descripció dels mòduls

La implementació del sistema basat en el coneixement s'ha dividit en mòduls per tal que cada tasca estigui en una secció de codi i el programa salti d'un mòdul a l'altre depenent de la tasca a realitzar.

A continuació es llisten els mòduls implementats junt amb una breu descripció de la seva tasca.

### 7.1 Principal

És el mòdul que s'executa primer. Es defineixen les funcions auxiliars, templates, assumpcions inicials i es mostra un missatge de benvinguda.

### 7.2 Perfil Lector

En el primer mòdul on es fan preguntes, se li demana a l'usuari que respongui un seguit de preguntes per formar un perfil de lector. Se li demana el nom per cortesia, edat, sexe, llengua materna i ocupació. Amb aquestes dades després podem decidir fer determinades preguntes o prendre determinades conclusions, com expliquem a l'apartat d'Evidències i Hipòtesis.

### 7.3 Preguntes Comunes

En aquest mòdul es demana a l'usuari que respongui un seguit de preguntes per permetre al sistema obtenir informació sobre preferències en els hàbits de lectura, gèneres preferits i conceptes relacionats amb el domini del problema.

### 7.4 Preguntes Específiques

Depenent del perfil del lector, o de respostes en el mòdul anterior, es procedeix a concretar preferències o fer preguntes que s'obvien si el perfil és un altre.

## 7.5 Associacions Incondicionals

Lligat amb l'apartat anterior, aquí s'extreuen conclusions a partir d'informació del perfil o de preguntes comunes, com per exemple quan no cal fer una pregunta específica perquè una característica del perfil del lector ja implica una conclusió.

## 7.6 Esborrar instàncies

Quan no es compleixen els requisits mínims establerts per l'usuari s'esborra la instància del llibre. D'aquesta manera es redueix l'espai de cerca per al pas posterior.

## 7.7 Associacions heurístiques

En el mòdul on s'apliquen les associacions heurístiques és on es creen totes les possibles recomanacions de llibres i les respostes de l'usuari classifiquen aquestes recomanacions. Es qualifiquen diferents indicadors per finalment en el proper mòdul obtenir-ne l'ordre final.

## 7.8 Refinament

Després d'haver aplicat les associacions heurístiques sobre el conjunt inicial de recomanacions, en aquest mòdul s'ordenen valorant els diferents indicadors de cada recomanació i se n'escullen les tres millors, o les millors en cas de que s'hagin descartat gairebé tots els llibres. En aquest mòdul s'ensenyà per pantalla a l'usuari la solució del problema.

## Capítol 8

# Casos de prova

### 8.1 Experiment Bernat

El Bernat és un estudiant català d'informàtica amant de les novel·les de ciència ficció (tot i que li agrada experimentar amb nous gèneres). Té poc temps per llegir, així que llegeix unes 4 hores al mes generalment, excepte en períodes de vacances. Llegeix a tot arreu, al bus o a casa tot i que no li preocupa especialment el pes del llibre. No el motiven ni desmotiven especialment els llibres més venuts, i intenta evitar els *totxos* degut a la manca de temps. Li agrada prendre notes d'alguna que altra frase interessant que li ha quedat i no li importa la llengua en que estigui escrit el llibre ni l'origen de l'autor. El seu autor preferit ha sigut sempre Tolkien i intenta evitar els llibres que costin més de 15 euros.

```
=====
Sistema basat en el coneixement recomanador de llibres

Nom? Bernat

Edat?  (número) 24

Sexe? (home dona) home

Llengua materna? (catala castella altres) catala

Ocupació? (estudiant academic professional desocupat altres)
estudiant

En quina freqüència acostumes a llegir? (diaria setmanal mensual
indiferent) mensual

Quantes hores acostumes a llegir en aquesta freqüència? (número) 4

Gènere:  (classics fantasia ficcio historica negra romantica
```

scifi social terror thriller indiferent) scifi

T'agrada experimentar amb nous gèneres? (molt bastant indiferent poc gens) bastant

On llegeixes habitualment? (casa exterior transport indiferent) indiferent

Prefereixes els llibres ben enquadernats? (molt bastant indiferent poc gens) poc

Prefereixes els llibres actuals? (molt bastant indiferent poc gens) bastant

T'agraden els best-sellers? (molt bastant indiferent poc gens) indiferent

Estàs disposat a llegir llibres gruixuts? (molt bastant indiferent poc gens) poc

Prefereixes que el vocabulari usat sigui senzill? (molt bastant indiferent poc gens) poc

T'agrada prendre notes al costat dels llibres? (molt bastant indiferent poc gens) bastant

T'agraden els llibres amb molts personatges? (molt bastant indiferent poc gens) indiferent

Quina és la teva preferència per la nacionalitat de l'autor? (catala espanyol catala-espanyol estranger indiferent) indiferent

En quina llengua prefereixes llegir? (catala castella altres indiferent) indiferent

Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla



15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jimenez
22. Xavier Montaña
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown
52. Carlos Ruiz Zafón
53. Alice Sebold

Indica el numero corresponent: 40

Prefereixes els llibres lleugers? (molt bastant indiferent poc gens)  
indiferent

Prefereixes que la lletra sigui gran? (molt bastant indiferent poc gens)  
indiferent

Fins quan estàs disposat a gastar-te en un llibre? (número) 15

Prefereixes els llibres amb contingut explícit (violència o sexe)?

(molt bastant indiferent poc gens) bastant

=====  
Bernat,  
aquests llibres són la recomanació del SBC:  
=====

Titol: Tots tenim secrets  
Autor: Carme Lafay  
Preu: 11.8 euros  
ISBN: 8497916824  
Gènere: scifi  
Nacionalitat autor: Espanya  
Any de publicació: 2007  
Número de pàgines: 230  
Enquadernació: tapatova  
Estil trama: simple  
Estil vocabulari: simple  
Quantiat de personatges: molts  
=====

Titol: El restaurant de la fi del món  
Autor: Douglas Adams  
Preu: 6.0 euros  
ISBN: 9788433912411  
Gènere: scifi  
Nacionalitat autor: Regne Unit  
Any de publicació: 2006  
Número de pàgines: 208  
Enquadernació: butxaca  
Estil trama: simple  
Estil vocabulari: simple  
Quantiat de personatges: molts  
=====

Titol: El Hobbit  
Autor: J. R. R. Tolkien  
Preu: 8.0 euros  
ISBN: 9712346932283  
Gènere: fantasia  
Nacionalitat autor: Regne Unit  
Any de publicació: 1937  
Número de pàgines: 220  
Enquadernació: butxaca  
Estil trama: simple  
Estil vocabulari: complexa  
Quantiat de personatges: molts  
=====

Els llibres s'ajusten molt al seu perfil, i de fet 2/3 són llibres que ja ha llegit. Són llibres de preu ajustat, del gènere de ciència ficció o fantasia, enrevessats i amb

molts personatges, tal i com li agrada. En el seu cas creiem que la recomanació és existosa.

## 8.2 Experiment Lucia

La Lucia és una noia italiana de 22 anys desocupada que ha estudiat Treball Social. És bastant llegidora, de fet llegeix dos hores al dia. Li interessa sobretot les trames socials, enrevessades als quals sovint emplena de notes o s'apunta informació en llibretes. Evita llibres cars i no li importa el format o l'enquadració. Llegeix a tot arreu. No li importa que els llibres siguin best-sellers, ni el vocabulari que s'hi usi ni que abusin de contingut explícit. Tampoc li importa si són gruixuts o lleugers. Sempre li ha agradat Isabel Allende.

```
=====
Sistema basat en el coneixement recomanador de llibres

Nom? Lucia

Edat?  (número) 22

Sexe? (home dona) dona

Llengua materna? (catala castella altres) altres

Ocupació? (estudiant academic professional desocupat altres)
desocupat

En quina freqüència acostumes a llegir? (diaria setmanal mensual
indiferent) diaria

Quantes hores acostumes a llegir en aquesta freqüència? (número) 2

Gènere:  (classics fantasia ficcio historica negra romantica scifi
social terror thriller indiferent) social

T'agrada experimentar amb nous gèneres? (molt bastant indiferent
poc gens) bastant

On llegeixes habitualment? (casa exterior transport indiferent)
indiferent

Prefereixes els llibres ben enquadrats? (molt bastant indiferent
poc gens) poc

Prefereixes els llibres actuals? (molt bastant indiferent poc
gens) indiferent

T'agraden els best-sellers? (molt bastant indiferent poc gens)
```

indiferent

Estàs disposat a llegir llibres gruixuts? (molt bastant  
indiferent poc gens) indiferent

Prefereixes que el vocabulari usat sigui senzill? (molt bastant  
indiferent poc gens) indiferent

T'agrada prendre notes al costat dels llibres? (molt bastant  
indiferent poc gens) molt

T'agraden els llibres amb molts personatges? (molt bastant  
indiferent poc gens) indiferent

Quina és la teva preferència per la nacionalitat de l'autor?  
(catala espanyol catala-espanyol estranger indiferent) indiferent

En quina llengua prefereixes llegir? (catala castella altres  
indiferent) indiferent

Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jiménez
22. Xavier Montañó
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon

30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown
52. Carlos Ruiz Zafón
53. Alice Sebold

Indica el numero corresponent: 50

Prefereixes els llibres lleugers? (molt bastant indiferent poc gens) bastant

Prefereixes que la lletra sigui gran? (molt bastant indiferent poc gens) bastant

Fins quan estàs disposat a gastar-te en un llibre? (número) 15

Prefereixes els llibres amb contingut explícit (violència o sexe)? (molt bastant indiferent poc gens) indiferent

=====

Lucia,  
aquests llibres són la recomanació del SBC:

=====

Titol:	La casa de los espíritus
Autor:	Isabel Allende
Preu:	14.0 euros
ISBN:	0060951303
Gènere:	ficció
Nacionalitat autor:	Chile
Any de publicació:	1995
Número de pàgines:	464
Enquadernació:	tapatova

Estil trama: complexa  
Estil vocabulari: complexa  
Quantiat de personatges: pocs

=====

Titol:	I, Robot
Autor:	Isaac Asimov
Preu:	6.0 euros
ISBN:	9755586932283
Gènere:	scifi
Nacionalitat autor:	Rússia
Any de publicació:	1950
Número de pàgines:	211
Enquadernació:	butxaca
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	pocs

=====

Titol:	Dune
Autor:	Frank Herbert
Preu:	8.95 euros
ISBN:	9711116932283
Gènere:	scifi
Nacionalitat autor:	Estats Units
Any de publicació:	1965
Número de pàgines:	636
Enquadernació:	butxaca
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	molts

=====

Els llibres que li han tocat són diferents dels que està acostumada a llegir, però es mostra oberta receptiva i amb ganes de provar-los. Li agrada especialment que siguin ben coneguts i que siguin barats i enrevessats. El primer llibre ja l'ha llegit i li va agradar molt.

### 8.3 Experiment Gerard

El Gerard és un noi de 15 anys que just ha descobert el món de la lectura. Li agrada llegir totxos com més grans millors, i sobretot li encanten les sèries de novel·les famoses, com *El Codi Davinci* o similars. Llegeix bastant i diu que de tant en tant es pot permetre comprar-se llibres de 20 euros, tot i que sovint li regalen o li deixa sa mare. Li agraden els llibres "macos", ben enquadernats i amb lletra gran. Confia bastant amb el número de vendes si té algun dubte sobre un llibre i, si pot, escull autors estrangers ja que creu que "acostumen a ser millors que el d'aquí".

Sistema basat en el coneixement recomanador de llibres

Nom? Gerard

Edat? (número) 15

Sexe? (home dona) home

Llengua materna? (catala castella altres) catala

Ocupació? (estudiant academic professional desocupat altres) estudiant

En quina freqüència acostumes a llegir? (diaria setmanal mensual indiferent) setmanal

Quantes hores acostumes a llegir en aquesta freqüència? (número) 2

Gènere: (classics fantasia ficcio historica negra romantica scifi social terror thriller indiferent) historica

T'agrada experimentar amb nous gèneres? (molt bastant indiferent poc gens) indiferent

On llegeixes habitualment? (casa exterior transport indiferent) casa

Prefereixes els llibres ben enquadernats? (molt bastant indiferent poc gens) bastant

Prefereixes els llibres actuals? (molt bastant indiferent poc gens) bastant

T'agraden els best-sellers? (molt bastant indiferent poc gens) molt

Estàs disposat a llegir llibres gruixuts? (molt bastant indiferent poc gens) molt

Prefereixes que el vocabulari usat sigui senzill? (molt bastant indiferent poc gens) indiferent

T'agrada prendre notes al costat dels llibres? (molt bastant indiferent poc gens) gens

T'agraden els llibres amb molts personatges? (molt bastant indiferent poc gens) poc

Quina és la teva preferència per la nacionalitat de l'autor? (catala espanyol catala-espanyol estranger indiferent) estranger

En quina llengua prefereixes llegir? (catala castella altres indiferent) catala

## Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jimenez
22. Xavier Montañó
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay



- 49. J. K. Rowling
- 50. Isabel Allende
- 51. Dan Brown
- 52. Carlos Ruiz Zafón
- 53. Alice Sebold

Indica el numero corresponent: 51

Si estàs entre dos llibres, prefereixes aquell que es ven més?  
(molt bastant indiferent poc gens) bastant

Prefereixes que la lletra sigui gran? (molt bastant indiferent  
poc gens) bastant

Fins quan estàs disposat a gastar-te en un llibre? (número) 20

Prefereixes els llibres amb contingut explícit (violència o sexe)?  
(molt bastant indiferent poc gens) indiferent

Gerard,  
aquests llibres són la recomanació del SBC:

=====

Titol:	Eclipse
Autor:	Stephenie Meyer
Preu:	16.0 euros
ISBN:	1603960228
Gènere:	terror
Nacionalitat autor:	Estats Units
Any de publicació:	2007
Número de pàgines:	637
Enquadernació:	tapatova
Bestseller	
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	pocs

=====

=====

Titol:	Luna Nueva
Autor:	Stephenie Meyer
Preu:	15.0 euros
ISBN:	9705800235
Gènere:	terror
Nacionalitat autor:	Estats Units
Any de publicació:	2007
Número de pàgines:	674
Enquadernació:	tapatova
Bestseller	
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	pocs

=====

Titol: Amanecer  
Autor: Stephenie Meyer  
Preu: 20.0 euros  
ISBN: 607110033X  
Gènere: terror  
Nacionalitat autor: Estats Units  
Any de publicació: 2008  
Número de pàgines: 829  
Enquadernació: tapatova  
Bestseller  
Estil trama: simple  
Estil vocabulari: simple  
Quantitat de personatges: pocs

=====  
Els llibres recomanats són, casualment, els últims llibres que ha llegit i que li han agradat molt. Diu que sempre s'enganxa a aquest tipus de triologies...

## 8.4 Experiment Jordi

A continuació tenim la prova d'un lector jove i estudiant, acostumat a llegir regularment i amb la majoria d'idees clares sobre els seus gustos.

=====  
Sistema basat en el coneixement recomanador de llibres  
  
Nom? Jordi  
  
Edat? (número) 23  
  
Sexe? (home dona) home  
  
Llengua materna? (catala castella altres) catala  
  
Ocupació? (estudiant academic professional desocupat altres)  
estudiant  
  
En quina freqüència acostumes a llegir?  
(diaria setmanal mensual indiferent) diaria  
  
Quantes hores acostumes a llegir en aquesta freqüència?  
(número) 1  
  
Gènere: (classics fantasia ficcio historica negra romantica  
scifi social terror thriller indiferent) ficcio

T'agrada experimentar amb nous gèneres? (molt bastant indiferent  
poc gens) bastant

On llegeixes habitualment? (casa exterior transport indiferent) casa

Prefereixes els llibres ben enquadernats? (molt bastant  
indiferent poc gens) poc

Prefereixes els llibres actuals? (molt bastant indiferent poc  
gens) bastant

T'agraden els best-sellers? (molt bastant indiferent poc gens) poc

Estàs disposat a llegir llibres gruixuts? (molt bastant indiferent  
poc gens) bastant

Prefereixes que el vocabulari usat sigui senzill? (molt bastant  
indiferent poc gens) indiferent

T'agrada prendre notes al costat dels llibres? (molt bastant indiferent  
poc gens) poc

T'agraden els llibres amb molts personatges? (molt bastant indiferent  
poc gens) bastant

Quina és la teva preferència per la nacionalitat de l'autor? (catala  
espanyol catala-espanyol estranger indiferent) estranger

En quina llengua prefereixes llegir? (catala castella altres  
indiferent) indiferent

Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain

19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jimenez
22. Xavier Montaña
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown
52. Carlos Ruiz Zafón
53. Alice Sebold

Indica el numero corresponent: 1

Prefereixes que la lletra sigui gran? (molt bastant indiferent  
poc gens) indiferent

Fins quan estàs disposat a gastar-te en un llibre? (número) 25

Prefereixes els llibres amb contingut explícit (violència o sexe)?  
(molt bastant indiferent poc gens) indiferent

=====

Jordi,  
aquests llibres són la recomanació del SBC:

=====

Titol: Petits plaers

Autor: Jane Bowles  
Preu: 13.0 euros  
ISBN: 9788492562862  
Gènere: ficció  
Nacionalitat autor: Regne Unit  
Any de publicació: 2009  
Número de pàgines: 173  
Enquadernació: tapatova  
Contingut explícit  
Estil trama: simple  
Estil vocabulari: simple  
Quantitat de personatges: molts

=====

Títol: 2666  
Autor: Roberto Bolaño  
Preu: 20.0 euros  
ISBN: 0307475956  
Gènere: ficció  
Nacionalitat autor: Xile  
Any de publicació: 2009  
Número de pàgines: 1136  
Enquadernació: tapatova  
Estil trama: complexa  
Estil vocabulari: simple  
Quantitat de personatges: molts

=====

Títol: El amor en los tiempos del cólera  
Autor: Gabriel García Márquez  
Preu: 15.0 euros  
ISBN: 0307387267  
Gènere: ficció  
Nacionalitat autor: Mèxic  
Any de publicació: 2007  
Número de pàgines: 464  
Enquadernació: tapatova  
Estil trama: complexa  
Estil vocabulari: complexa  
Quantitat de personatges: pocs

=====

En el cas de les recomanacions, les tres han estat satisfactòries, i la segona opció era un llibre que el lector ja havia llegit recentment. Tots els paràmetres que s'han contestat amb rotunditat es veuen reflectits en les respostes, per tant el lector està satisfet amb els tres llibres recomanats.

## 8.5 Experiment Xavier

El Xavier és professor d'un institut de 57 anys. Sempre ha sigut molt llegidor, tot i que actualment té poc temps i llegeix un parell d'hores setmanals. Li agrada molt la novel·la històrica, tot i que està obert a nous gèneres. No el criden massa els llibres més populars ni els més venuts, i intenta evitar els llibres gruixuts. Els llibres que llegeix han d'estar en català tot i que la procedència de l'autor i li és indiferent. No li importa que el llibre sigui enrevessat, tot i que li molesta que hi hagi masses personatges perquè diu que "es perd". Omple els llibres de notes i usa llibretes també. Demana sempre que la lletra sigui gran perquè diu que "no hi veu". Per ell el preu és un factor important i diu que no es vol gastar més de 20 euros en un llibre.

=====

Sistema basat en el coneixement recomanador de llibres

Nom? Xavier

Edat? (número) 57

Sexe? (home dona) home

Llengua materna? (catala castella altres) catala

Ocupació? (estudiant academic professional desocupat altres)  
academic

En quina freqüència acostumes a llegir? (diaria setmanal mensual  
indiferent) setmanal

Quantes hores acostumes a llegir en aquesta freqüència? (número) 2

Gènere: (classics fantasia ficcio historica negra romantica scifi  
social terror thriller indiferent) historica

T'agrada experimentar amb nous gèneres? (molt bastant indiferent  
poc gens) bastant

On llegeixes habitualment? (casa exterior transport indiferent)  
casa

Prefereixes els llibres ben enquadernats? (molt bastant indiferent  
poc gens) indiferent

Prefereixes els llibres actuals? (molt bastant indiferent poc  
gens) bastant

T'agraden els best-sellers? (molt bastant indiferent poc gens) poc

Estàs disposat a llegir llibres gruixuts? (molt bastant

indiferent poc gens) poc

Prefereixes que el vocabulari usat sigui senzill? (molt bastant  
indiferent poc gens) poc

T'agrada prendre notes al costat dels llibres? (molt bastant  
indiferent poc gens) bastant

T'agraden els llibres amb molts personatges? (molt bastant  
indiferent poc gens) poc

Quina és la teva preferència per la nacionalitat de l'autor?  
(catala espanyol catala-espanyol estranger indiferent) indiferent

En quina llengua prefereixes llegir? (catala castella altres  
indiferent) catala

Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jiménez
22. Xavier Montaña
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry

33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown
52. Carlos Ruiz Zafón
53. Alice Sebold

Indica el numero corresponent: 39

Prefereixes que la lletra sigui gran? (molt bastant indiferent  
poc gens) bastant

El preu és un factor prioritari? (si no indiferent) si

Fins quan estàs disposat a gastar-te en un llibre? (número) 20

Prefereixes els llibres amb contingut explícit (violència o sexe)?  
(molt bastant indiferent poc gens) indiferent

=====

Xavier,  
aquests llibres són la recomanació del SBC:

=====

Titol:	1984
Autor:	George Orwell
Preu:	9.0 euros
ISBN:	9344444932283
Gènere:	scifi
Nacionalitat autor:	Regne Unit
Any de publicació:	1949
Número de pàgines:	180
Enquadernació:	tapatova
Contingut explícit	
Estil trama:	complexa
Estil vocabulari:	simple
Quantitat de personatges:	pocs



```

=====
Titol:                El caçador d'estels
Autor:                Khaled Hosseini
Preu:                 8.95 euros
ISBN:                 9788499081106
Gènere:               historica
Nacionalitat autor:  Persa
Any de publicació:   2006
Número de pàgines:   255
Enquadernació:       butxaca
Bestseller
Estil trama:          complexa
Estil vocabulari:     simple
Quantiat de personatges: pocs
=====
Titol:                L'elegància de quan es dorm
Autor:                Vizconde de Lascano Tegui
Preu:                 19.0 euros
ISBN:                 978-84-935927-5-2
Gènere:               social
Nacionalitat autor:  Argentina
Any de publicació:   2008
Número de pàgines:   185
Enquadernació:       tapadura
Estil trama:          simple
Estil vocabulari:     complexa
Quantiat de personatges: pocs
=====

```

Els primer llibre és un dels seus preferits, així que es considera un encert. El segon li ha recomanat força gent, i creu que pot ser molt interessant. I el tercer no l'ha sentit anomenar, però les característiques les valora bastant interessants.

## 8.6 Experiment Serafín

Aquesta prova s'ha fet amb una persona adulta, psicòleg, que no llegeix tant com voldria i li bastant tots els llibres populars i els best-sellers.

```

=====

Sistema basat en el coneixement recomanador de llibres

Nom? Serafin

Edat?  (número) 57

Sexe?  (home dona) home

```

Llengua materna? (catala castella altres) castella

Ocupació? (estudiant academic professional desocupat altres)  
professional

En quina freqüència acostumes a llegir? (diaria setmanal mensual  
indiferent) setmanal

Quantes hores acostumes a llegir en aquesta freqüència?  
(número) 4

Gènere: (classics fantasia ficcio historica negra romantica  
scifi social terror thriller indiferent) historica

T'agrada experimentar amb nous gèneres? (molt bastant  
indiferent poc gens) bastant

On llegeixes habitualment? (casa exterior transport indiferent)  
casa

Prefereixes els llibres ben enquadernats? (molt bastant  
indiferent poc gens) indiferent

Prefereixes els llibres actuals? (molt bastant indiferent poc  
gens) indiferent

T'agraden els best-sellers? (molt bastant indiferent poc gens)  
bastant

Estàs disposat a llegir llibres gruixuts? (molt bastant  
indiferent poc gens) bastant

Prefereixes que el vocabulari usat sigui senzill? (molt bastant  
indiferent poc gens) indiferent

T'agrada prendre notes al costat dels llibres? (molt bastant  
indiferent poc gens) poc

T'agraden els llibres amb molts personatges? (molt bastant  
indiferent poc gens) indiferent

Quina és la teva preferència per la nacionalitat de l'autor?  
(catala espanyol catala-espanyol estranger indiferent) indiferent

En quina llengua prefereixes llegir? (catala castella altres  
indiferent) castella

Autor preferit:

1. Cap dels anteriors

2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson
21. Francisco Jimenez
22. Xavier Montañó
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown

52. Carlos Ruiz Zafón

53. Alice Sebold

Indica el numero corresponent: 1

Si estàs entre dos llibres, prefereixes aquell que es ven més?  
(molt bastant indiferent poc gens) indiferent

Prefereixes que la lletra sigui gran? (molt bastant indiferent  
poc gens) bastant

El preu és un factor prioritari? (si no indiferent) no

Prefereixes els llibres amb contingut explícit (violència o sexe)?  
(molt bastant indiferent poc gens) indiferent

Serafin,  
aquests llibres són la recomanació del SBC:

=====

Titol:	Los Borgia
Autor:	Jose Antonio Cebrián
Preu:	18.5 euros
ISBN:	978847518888
Gènere:	historica
Nacionalitat autor:	Espanya
Any de publicació:	2010
Número de pàgines:	700
Enquadernació:	tapatova
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	molts

=====

=====

Titol:	El símbolo perdido
Autor:	Dan Brown
Preu:	25.0 euros
ISBN:	8408089250
Gènere:	thriller
Nacionalitat autor:	Estats Units
Any de publicació:	2009
Número de pàgines:	619
Enquadernació:	tapatova
Bestseller	
Estil trama:	complexa
Estil vocabulari:	complexa
Quantiat de personatges:	pocs

=====

=====

Titol:	Dune
Autor:	Frank Herbert
Preu:	8.95 euros
ISBN:	9711116932283

=====

Gènere: scifi  
Nacionalitat autor: Estats Units  
Any de publicació: 1965  
Número de pàgines: 636  
Enquadernació: butxaca  
Estil trama: complexa  
Estil vocabulari: complexa  
Quantitat de personatges: molts

=====

El resultat ha estat molt satisfactori. Al lector li han encantat les dues primeres propostes i li ha sorprès la tercera. Molt satisfet per les recomanacions ha comentat que les prendria en consideració.

## 8.7 Experiment Jordi 2

En aquesta prova hem preguntat a un arquitecte de 61 anys, àvid lector amb moltes opinions formades al respecte. Llegeix un promig de 2 a tres llibres la setmana i manté un blog amb opinions del que va llegint.

=====

Sistema basat en el coneixement recomanador de llibres

Nom? Jordi

Edat? (número) 61

Sexe? (home dona) home

Llengua materna? (catala castella altres) catala

Ocupació? (estudiant academic professional desocupat altres)  
professional

En quina freqüència acostumes a llegir? (diaria setmanal mensual indiferent) diaria

Quantes hores acostumes a llegir en aquesta freqüència? (número) 4

Gènere: (classics fantasia ficcio historica negra romantica  
scifi social terror thriller indiferent) ficcio

T'agrada experimentar amb nous gèneres? (molt bastant indiferent poc gens) molt

On llegeixes habitualment? (casa exterior transport indiferent)

casa

Prefereixes els llibres ben enquadernats? (molt bastant  
indiferent poc gens) indiferent

Prefereixes els llibres actuals? (molt bastant indiferent poc  
gens) molt

T'agraden els best-sellers? (molt bastant indiferent poc gens)  
gens

Estàs disposat a llegir llibres gruixuts? (molt bastant indiferent  
poc gens) molt

Prefereixes que el vocabulari usat sigui senzill? (molt bastant  
indiferent poc gens) poc

T'agrada prendre notes al costat dels llibres? (molt bastant  
indiferent poc gens) poc

T'agraden els llibres amb molts personatges? (molt bastant  
indiferent poc gens) indiferent

Quina és la teva preferència per la nacionalitat de l'autor?  
(catala espanyol catala-espanyol estranger indiferent) indiferent

En quina llengua prefereixes llegir? (catala castella altres  
indiferent) indiferent

Autor preferit:

1. Cap dels anteriors
2. Stephenie Meyer
3. Raymond Carver
4. Paulo Coelho
5. Gabriel García Márquez
6. Thomas Pynchon
7. Boris Pahor
8. Julian Barnes
9. Joseph Roth
10. Jose Antonio Cebrián
11. Parramon
12. Juan Bonilla
13. Quim Monzó
14. Josep Pla
15. Bernhard Schlink
16. Khaled Hosseini
17. Glen Cooper
18. Mark Twain
19. Roberto Bolaño
20. Stieg Larsson

21. Francisco Jimenez
22. Xavier Montaña
23. Robert Luis Stevenson
24. Richard Yates
25. Maria Àngels Anglada
26. Xavier Bosch
27. Camilla Lackberg
28. Jane Bowles
29. Thomas Pynchon
30. Stefano Bordiglioni
31. Donata Pizzato
32. Antoine de Saint-Exupéry
33. Laura Esquivel
34. Sandra Cisneros
35. Hector Abad Faciolince
36. Italo Calvino
37. Isaac Asimov
38. Frank Herbert
39. George Orwell
40. J. R. R. Tolkien
41. Clive Staples Lewis
42. R. A. Salvatore
43. Jules Verne
44. Douglas Adams
45. Orson Scott Card
46. Arthur C. Clarke
47. Vizconde de Lascano Tegui
48. Carme Lafay
49. J. K. Rowling
50. Isabel Allende
51. Dan Brown
52. Carlos Ruiz Zafón
53. Alice Sebold

Indica el numero corresponent: 1

El preu és un factor prioritari? (si no indiferent) no

Prefereixes els llibres amb contingut explícit (violència o sexe)?  
(molt bastant indiferent poc gens) indiferent

=====

Jordi,  
aquests llibres són la recomanació del SBC:

=====

Titol:	El amor en los tiempos del cólera
Autor:	Gabriel García Márquez
Preu:	15.0 euros
ISBN:	0307387267
Gènere:	ficció
Nacionalitat autor:	Mèxic

Any de publicació: 2007  
Número de pàgines: 464  
Enquadernació: tapatova  
Estil trama: complexa  
Estil vocabulari: complexa  
Quantiat de personatges: pocs

=====

Títol: Cien años de soledad  
Autor: Gabriel García Márquez  
Preu: 16.0 euros  
ISBN: 8420471836  
Gènere: ficció  
Nacionalitat autor: Mèxic  
Any de publicació: 2007  
Número de pàgines: 756  
Enquadernació: tapadura  
Estil trama: complexa  
Estil vocabulari: complexa  
Quantiat de personatges: pocs

=====

Títol: 2666  
Autor: Roberto Bolaño  
Preu: 20.0 euros  
ISBN: 0307475956  
Gènere: ficció  
Nacionalitat autor: Xile  
Any de publicació: 2009  
Número de pàgines: 1136  
Enquadernació: tapatova  
Estil trama: complexa  
Estil vocabulari: simple  
Quantiat de personatges: molts

=====

Les recomanacions no han estat molt útils, ja que el lector ja havia llegit els tres llibres, però això reflecteix la qualitat de les recomanacions. En un programa sense la restricció de tres llibres com aquest, es podria haver demanat més i més recomanacions fins trobar llibres que no hagués llegit ja el lector.



## Capítol 9

# Conclusions

La recomanació de llibres és un procés més complex i personal del que pot semblar aparentment. Per satisfer les necessitats de l'usuari no n'hi ha prou en fixar-se en les seves característiques disponibles dels llibres, sinó que caldria entrar en el seu contingut i analitzar-lo i crear una base de dades molt particular. En general, molts heurístics usats estan basats en estadístiques i poden no ser aplicables en tots els casos. Com més acurat es vol ser en el procés d'associació heurística més complex esdevé el problema.

Aquesta pràctica ens ha obert les portes d'un nou coneixement que, fins ara, gairebé no havíem tingut experiència a la carrera. La programació de llenguatges de regles, que s'havia vist lleugerament amb Prolog, obliga a canviar radicalment la perspectiva del programador. Cal codificar el programa pensant en les implicacions de les regles i no en procediments o funcions com estem acostumats. Gràcies a això, es pot expressar cada hipòtesi o limitació del Sistema Basat en el Coneixement de forma natural i senzilla. Tanmateix, cal dir que el llenguatge CLIPS resulta bastant ferragós i costa entrar-hi en contacte. A més, creiem que és especialment difícil de debugar i costa sortir-se'n quan un es queda encallat en algun punt.

## Capítol 10

# Annex: Entrevista a un llibreter

Notes: Al final li he fet una entrevista de veritat a un amic (no és el meu pare). Ho completaré amb altres preguntes patillades i d'altres al meu pare.

**Què és el primer que preguntes quan algú et demana que li recomanis un llibre?**

Quin tipus de llibre llegeixes normalment. Una pregunta molt bona, considero jo, és quins llibres t'has llegit més d'un cop.

**I si et diuen que no llegeixen gaire? Per on els faries començar?**

Home, pots preguntar interessos en altres camps (com la televisió, cine, ...) per fer-te una idea.

**Recomanes coses diferents als homes i a les dones?**

Depèn del cas, pero en general sí.

**Quins d'aquests punts creus que són més importants: l'edat del lector, el sexe, els gèneres que diu que són els seus preferits, l'estona que dedica a llegir, on acostuma a llegir (al llit, al metro, als bars), ...**

Els gèneres preferits i el temps disponible (bàsicament perquè si hi dedica molta es fàcil que sigui millor lector i li pots recomanar coses mes difícils)

**Quines altres preguntes creus que ajuden a encertar-la recomanant llibres?**

Preguntaria si com a lector t'agrada que t'enganyin i et portin per terrenys desconeguts, o t'agraden les coses properes i conegudes.

**Quan algú et diu que és molt d'un gènere en concret, acostumes a intentar allunyar-lo d'aquest i suggerir-li coses diferents? O li fas cas i et cenyeixes al que ell et diu? (relacionat amb la pregunta anterior)**

Jo li obriria el ventall. Sobretot si és un lector habitual i confio que acceptarà les alternatives.

Una altra pregunta que crec que diferencia bastant als lectors és si prenen notes en el llibre o no.

**Creus que el tipus d'enquadrernació afecta a l'elecció del llibre?**

Jo crec que els lectors més reincidentes poden ser una mica fetitxistes i preferir llibres de tapa dura, sobretot amb els seus autors preferits o llibres de capçalera. Una mica com passa amb els vinils.

**I com seria el teu llibre perfecte?**

Llibre de poesia, tapa dura, un llibre gran però de poemes curts, complex, en català o castellà. la nacionalitat de l'autor no és massa important, però en cas de desconeguts tiro més per un país exòtic que proper. I si pot ser escrit al segle vint.