



Escola Politècnica Superior
d'Enginyeria de Vilanova i la Geltrú

UNIVERSITAT POLITÈCNICA DE CATALUNYA

PROJECTE FINAL DE CARRERA

TÍTOL: PTKChA: Editor d'anotacions

AUTOR: Ezequiel Huguet Espargaró

TITULACIÓ: Enginyeria Tècnica d'Informàtica de Gestió

DIRECTOR: Jordi Turmo Borràs

PONENT: Lluís Padró Cirera

DEPARTAMENT: Llenguatges i sistemes informàtics (LSI)

DATA: Juny del 2010

TÍTOL: PTKChA: Editor d' anotacions

COGNOMS: Huguet Espargaró

NOM: Ezequiel

TITULACIÓ: Enginyeria Tècnica d' Informàtica

ESPECIALITAT: Informàtica de Gestió

PLA: 92

DIRECTOR: Jordi Turmo Borràs

PONENT: Lluís Padró Cirera

DEPARTAMENT: Llenguatges i sistemes informàtics (LSI)

QUALIFICACIÓ DEL PFC

TRIBUNAL

PRESIDENT

NEUS CATALÀ ROIG

SECRETARI

MERCÈ CLAVEROL AGUAS

VOCAL

JAVIER MASIP BRUIN

DATA DE LECTURA

06-07-2010

Aquest projecte té en compte aspectes mediambientals: Sí No

PROJECTE FINAL DE CARRERA

RESUM

El Grup de Processament del Llenguatge Natural, dins del Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (TALP), es dedica al tractament automàtic del llenguatge natural i desenvolupa diferents eines pel tractament d'aquest.

El projecte consisteix en modificar el PTKChA, que és una d'aquestes eines destinada a l'etiquetatge de textos.

El treball s'estructura en dues planificacions; en la primera es demana que, a partir d'uns requisits inicials, s'analitzi funcionalment l'aplicació i es determinin totes les modificacions que es poden fer per millorar-la. En la segona, s'especifiquen i es desenvolupen les modificacions.

Així doncs, l'objectiu d'aquest projecte és desenvolupar les modificacions.

PARAULES CLAU

TALP	ETIQUETATGE DE TEXTOS	PTKCHA	PERL	PROCESSAMENT DEL LLENGUATGE

SUMARI

En aquesta memòria s'han introduït dos annexes:

Annex 1: aquest primer annex fa referència al fitxer comprimit amb el codi font de l'aplicació.

Annex 2: es refereix a la llista final de requisits. Aquest es presenta en format paper i electrònic juntament amb la memòria.

ÍNDIX

1. INTRODUCCIÓ.....	7
2. OBJECTIUS.....	9
3. PLANIFICACIÓ I COST INICIAL.....	10
3.1 Planificació.....	10
3.2 Cost inicial.....	11
4. ANÀLISI DE FUNCIONALITATS.....	12
4.1 Menú aplicació.....	15
4.2 Gestió de projectes.....	16
4.3 Marcatges.....	18
4.4 Procés d'anotació.....	18
5. ELABORACIÓ DE LA NOVA LLISTA DE REQUISITS.....	20
5.1 Menú.....	20
5.2 Gestió de projectes.....	21
5.3 Gestió de marcatges.....	22
5.4 Gestió d'usuaris.....	24
5.5 Procés d'anotacions.....	24
6. REPLANIFICACIÓ.....	26
6.1 Temps estimat.....	26
6.2 Distribució de les tasques.....	26
6.3 Cost previst de la replanificació.....	28
7. MODEL CONCEPTUAL.....	29
7.1 Entitat administradors.....	30
7.2 Entitat projecte.....	30
7.3 Entitat etiquetatge.....	32
7.4 Entitat atribut.....	32
7.5 Entitat atribut.....	32
7.6 Entitat grup.....	32
7.7 Entitat plugin.....	32
7.8 Entitat anotació.....	35
8. CASOS D'ÚS.....	37
8.1 Menú. Punt de vista administrador.....	37
8.1 Menú. Punt de vista anotador.....	37
8.2 Gestió de projectes. Punt de vista administrador.....	38
8.2 Gestió de projectes. Punt de vista anotador.....	38

8.3 Gestió de marcatges. Punt de vista administrador.....	39
8.4 Gestió d'usuaris. Punt de vista administrador.....	40
8.5 Procés d'anotació. Punt de vista anotador.....	40
Casos d'ús del menú.....	41
Casos d'ús de gestió de projectes.....	43
Casos d'ús de gestió de marcatges.....	47
Casos d'ús de gestió d'usuaris.....	55
Casos d'ús de procés d'anotació.....	56
9. DIAGRAMA DE CLASSES.....	59
9.1 Menú.....	59
9.2 Gestió de projectes.....	59
9.3 Gestió de marcatges.....	60
9.4 Gestió d'usuaris.....	60
9.5 Procés d'anotació.....	60
10. DISSENY DE PANTALLES.....	61
11. MANUAL D'INSTAL·LACIÓ.....	62
11.1 Requeriments.....	62
11.2 Instal·lació.....	62
11.3 Executar el programa.....	62
12. MANUAL D'USUARI: ANOTADOR.....	63
12.1 Menú.....	63
12.2 Activar projecte.....	64
12.3 Pantalla principal.....	65
13. MANUAL D'USUARI: ADMINISTRADOR.....	67
13.1 Menú.....	67
13.2 Gestió de projectes.....	68
13.3 Gestió de marcatges.....	70
13.4 Gestió d'usuaris.....	78
13.5 Procés d'anotació.....	78
14. EXEMPLE D'ÚS.....	79
15. COST REAL.....	80
16. CONCLUSIÓ.....	81
17. REFERÈNCIES BIBLIOGRÀFIQUES.....	82
18. ANNEXES.....	83

1.- INTRODUCCIÓ

El Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (TALP) és un centre específic de recerca de la Universitat Politècnica de Catalunya (UPC).

L'àmbit tecnològic del Centre TALP és el tractament automàtic del llenguatge natural, tant oral com escrit, amb l'objectiu d'ajudar a superar les barreres lingüístiques i millorar l'accessibilitat dels sistemes d'informació.

El Centre TALP està format per dos grups d'investigació: el Grup de Processament del Llenguatge Natural (GPLN) i el Grup del Tractament de la Parla (VEU).

El Grup de Processament del Llenguatge Natural se centra en el tractament automàtic del llenguatge natural i concretament en els aspectes relacionats amb la construcció i explotació de recursos lèxics multilingües i en l'extracció d'informació de textos en llenguatge natural i el resum automàtic.

El Grup de Processament del Llenguatge Natural dedica especial atenció al desenvolupament d'eines de tractament del llenguatge : analitzadors morfològics i sintàctics, etiquetadors gramaticals, sistemes de tractament de corpus, mètodes genèrics d'aprenentatge automàtic per resoldre tasques de processament del llenguatge natural, etc..., i a la recopilació de recursos lingüístics: gramàtiques, diccionaris, base de dades lèxiques i bases de conceptuais.

En aquest ambient és quan sorgeix la proposta de Lluís Padró i Jordi Turmo, professors del Departament de Llenguatges i Sistemes Informàtics de la UPC, de modificar una eina per l'etiquetatge de textos: el PTKChA.

El PTKChA és una aplicació amb la qual, a partir d'uns fitxers d'entrada, es poden fer anotacions sobre aquests per després generar-ne de nous i ser utilitzats per altres sistemes per l'aprenentatge automàtic, per comprovar resultats, etc.

La idea del projecte que em varen proposar era modificar el PTKChA amb l'objectiu de millorar l'aplicació.

En Lluís i en Jordi em van oferir la tasca d'afegir la gestió dels marcatges, afegir tipus d'atributs, afegir tipus de relació, afegir gestió d'etiquetatges, controlar etiquetatges per projecte, poder fer anotacions aniuades i realitzar un procés d'unió d'etiquetatges.

De totes maneres, vam arribar a l'acord que, a banda d'especificar-me tots aquests requisits , havia de fer un anàlisi de les funcionalitats per veure què suposaven tots aquests canvis, a més de contemplar la possibilitat de trobar nous aspectes a millorar o modificar.

2.- OBJECTIUS

La idea principal del projecte, com ja he esmentat anteriorment, és modificar l'aplicació PTKChA en base a la petició dels professors Lluís Padró i Jordi Turmo, amb uns requisits inicials.

Aquests són:

- Manteniment de marcatges
- Afegir tipus d'atributs
- Afegir tipus de relacions
- Manteniment d'etiquetatges
- Control d'etiquetatges per projecte
- Anotacions aniuades
- Procés d'unió d'etiquetatges

A partir d'aquests requisits inicials, l'objectiu és fer un anàlisi de les funcionalitats per determinar què suposaran aquests canvis, així com detectar noves millores.

Acte seguit, s'elaborarà una nova llista la qual inclourà tots els canvis que s'hauran de dur a terme en l'aplicació PTKChA.

En definitiva, l'objectiu principal del projecte és complir amb el desenvolupament de les modificacions de la nova llista.

3.- PLANIFICACIÓ I COST INICIAL

3.1.- Planificació

L' inici de la realització del projecte es comença a dur a terme el dia 12 d'abril del 2010.

La planificació del projecte s'ha dividit en dues, dins de les quals hi ha determinades tasques a realitzar durant aquestes dates.

En la primera planificació, la feina a realitzar es divideix en aquestes tres grans tasques:

- Anàlisi dels requisits inicials
- Anàlisi de les funcionalitats
- Elaboració de la nova llista de requisits

Tal i com s'indica en el diagrama de Gantt l'anàlisi dels requisits inicials i el de les funcionalitats s'han portat a terme al mateix temps.

La temporalització d'aquesta primera planificació queda distribuïda de la següent manera:

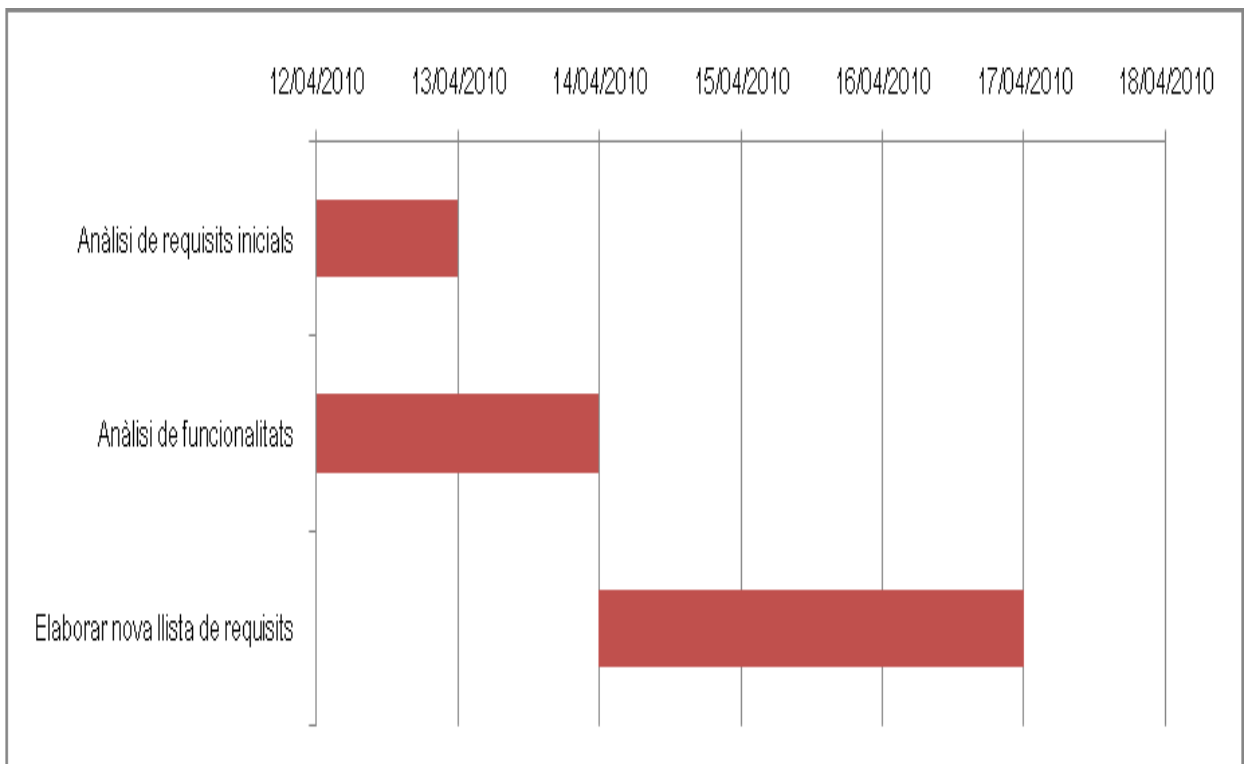


Figura 3.1: Diagrama de Gantt de la planificació

Amb la primera planificació, el que volem és planificar el temps per veure com funciona el PTKChA, és a dir, per poder entendre'l, per saber el que fa i determinar tot allò que finalment volem que faci.

Una vegada feta aquesta primera planificació, i sabent el que volem fer, passarem a replanificar (la segona planificació), és a dir, a fer una nova planificació sobre totes les modificacions que s'han de fer.

3.2.- Cost inicial de la primera planificació

Concepte	Hores	Total
Analista	40	2000€*
Total		2000€

Figura 3.2: taula de cost inicial en la primera planificació

*Partint de la base que l'hora de l'analista són 50€.

4-ANÀLISI DE FUNCIONALITATS

La figura següent descriu l'arquitectura del PTKChA :

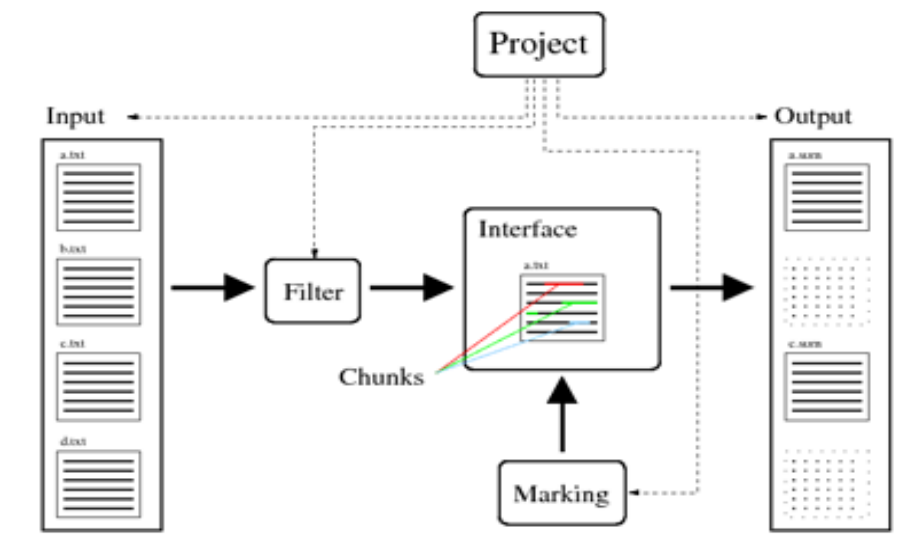


Figura 4.1: arquitectura del PTKChA

El treball d'anotació s'organitza en projectes, els quals venen definits per un directori d'entrada, una extensió, un filtre, un directori de sortida i un fitxer de marcatges.

L'interfície del PTKChA permet seleccionar els fitxers d'un projecte, pre-processats amb un filtre, per ser anotats. Les anotacions venen definides en el fitxer de marcatge. Finalment el fitxer anotat¹ és guardat en el directori de sortida per la seva posterior utilització en altres sistemes.

Les anotacions són fragments de text no sobreposats que es defineixen amb els seus atributs i les seves relacions, les quals estan definides en un sol fitxer de marcatge.

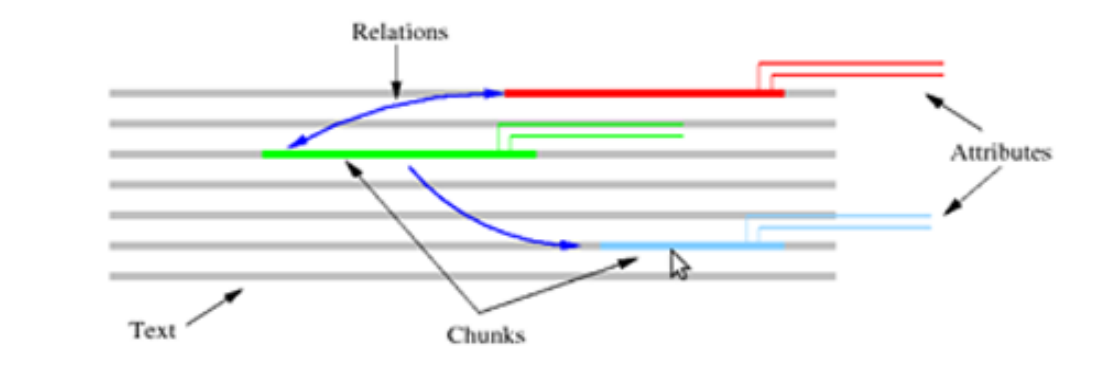


Figura 4.2: anotacions, atributs i relacions

¹ Veure pàgina 12 pel format de fitxer anotat

El marcatge és el conjunt d'elements que defineix les anotacions i diferents opcions pel procés d'anotació. Ve definit amb l'etiqueta *marking*; en l'exemple de la figura 4.3 es pot veure com obre i tanca tota la informació del marcatge.

Un marcatge ve definit per els següents elements:

- **Nom** : Indica el nom de l'etiqueta a utilitzar en el fitxer de sortida². El nom ve definit per l'etiqueta *marking* i l'atribut *element*. En l'exemple de la figura 4.3 es pot veure que el nom del marcatge es *ne* : `<marking element="ne">`.
- **Atributs** : Són les propietats que una anotació pot tenir. Cada atribut és assignat dins l'anotació amb el seu valor corresponent. Es defineixen amb l'etiqueta *attribute* i una llista d'etiquetes *value* amb atribut *v*. Com es pot veure en l'exemple de la figura 4.3, hi ha dos atributs : *class* i *multiword*, amb els seus valors.

```
<attribute name="class">  
  <value v="PERSON" />  
  <value v="ORGANIZATION" />  
  <value v="LOCATION" />  
  <value v="OTHERS" />  
</attribute>
```

```
<attribute name="multiword">  
  <value v="No" />  
  <value v="Yes" />  
</attribute>
```

- **Relacions**: Tal i com el seu nom indica, la funció que tenen és la de relacionar dues anotacions. D'aquestes dues, una serà definida com a origen i l'altra com a destí. La relació en sí té una propietat anomenada

² Veure exemple de fitxer de sortida a la pàgina 11

stereo, que indica si la relació és unidireccional, és a dir, que es produeix de l'origen al destí, o bé bidireccional, on la relació, aleshores, se suposa que és reflexiva. Les relacions es defineixen amb l'etiqueta *relation* i dos atributs, *name* i *stereo*. En l' exemple de la figura 4.3 està definida una relació amb nom *coref* i la propietat *stereo* amb valor *bi*.

```
< relation name="coref" stereo="bi" />
```

- **Els plugins:** Són els plugins que estan actius quan es treballa amb el marcatge. Les propietats dels plugins són el nom, el nom de la classe que implementa i el fitxer on es troba. Els fitxers es busquen en el directori per defecte del Perl³ i en el directori d'inclusió de les opcions generals. En l'exemple de la figura 4.3 està definit el plugin :

```
<plugin name="NE" class="PluginNE"  
file="Plugins::PluginNE::PluginNE" />
```

- **Grup:** És un conjunt de varies anotacions. En l'exemple de la figura 4.3 es pot veure definit l'opció de grup amb l'etiqueta *<clustered/>*.
- **Extres:** Indiquen altres etiquetes XML (*extensible markup language*) que poden estar presents en els fitxers d'entrada i que es volen mantenir en els fitxers de sortida. Els extres estan definits amb l'etiqueta *extra* i l'atribut *element*. En l'exemple de la figura 4.3 està definit un extra DOC : *<extra element="DOC" />*.

³ Llenguatge en que s'ha desenvolupat l'aplicació PTKChA

```

<marking element="ne">
  <attribute name="class">
    <value v="PERSON" />
    <value v="ORGANIZATION" />
    <value v="LOCATION" />
    <value v="OTHERS" />
  </attribute>
  <attribute name="multiword">
    <value v="No" />
    <value v="Yes" />
  </attribute>
  <relation name="coref" stereo="i" />
  <plugin name="NE" class="PluginNE"
    file="Plugins::PluginNE::PluginNE" />
  <clustered />
  <extra element="DOC"/>
</marking>

```

Figura 4.3: exemple de marcatge

Per a l'anàlisi de les funcionalitats de l'aplicació, s'han establert uns punts a seguir. Aquests són:

- **1. Menú de l'aplicació**
- **2. Gestió de projectes**
- **3. Marcatges**
- **4. Procés d'anotació**

A continuació es detallen les opcions que facilita cada punt:

1. Menú de l'aplicació: El menú de l'aplicació dona accés a:

1.1 Gestionar projectes : des d'aquesta opció accedim a la gestió de projectes.

1.2 Opcions generals : des d'aquesta opció es poden canviar tres valors de configuració generals: triar si vols expandir la selecció de text a una paraula en el moment que s'està anotant, si vols

forçar l'escriptura, és a dir, guardar qualsevol canvi en el fitxer de sortida, o bé escollir un directori on buscar els plugins.

1.3 Guardar fitxer actual, tornar a versió inicial i revertir canvis: des d'aquestes opcions es poden fer les esmentades operacions amb el fitxer actual d'anotació.

1.4 Vista per atributs: des d'aquesta, es poden veure les diferents anotacions del text actual segons l'atribut seleccionat.

1.5 Estadístiques d'anotacions: des d'aquí es pot veure la informació sobre les anotacions.

1.6 Plugins: des d'aquest punt del menú es poden utilitzar els plugins que estiguin definits en el marcatge.

- 2. Gestió de projectes:** Aquest apartat ens permet gestionar els projectes. Ens mostra els projectes actuals, el seu estat (si es pot treballar amb ell o no) i ens permet fer-ne de nous, modificar-los, esborrar-los, clonar-los i obrir-lo per treballar amb ell (anotar).

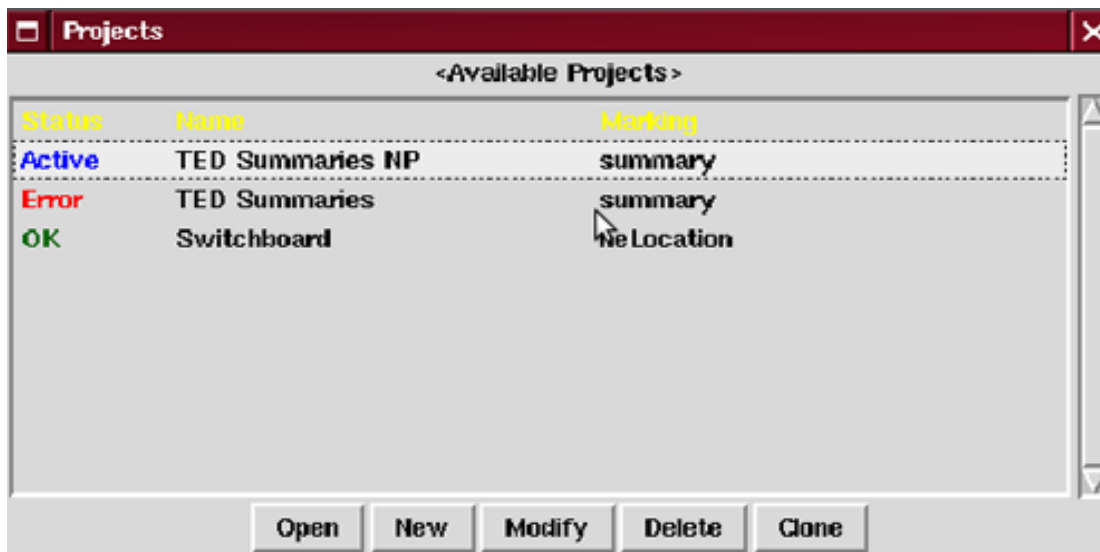


Figura 4.4: gestió de projectes

Les dades del projecte són guardades en un fitxer XML de configuració, juntament amb les opcions generals i els filtres.

```
<config version="1.1.xml">
  <option name="SelExp" value="1" />
  <option name="ForWri" value="1" />
  <option name="IncDir"
value="/home/administrador/PFC/ptkcha2/Plugins" />
  <filters>
    <filter name="bio" file="Filters::StandardFilters"
      class="Filters::FilterBIO" />
    <filter name="txt" file="Filters::StandardFilters"
class="Filters::FilterTxt" />
    <filter name="utf" file="Filters::ExtraFilters"
class="Filters::FilterUtf" />
    <filter name="utf_np" file="Filters::ExtraFilters"
class="Filters::FilterUtfNp" />
    <filter name="via" file="Filters::ExtraFilters"
class="Filters::FilterVia" />
    <filter name="xml" file="Filters::StandardFilters"
class="Filters::FilterXML" />
  </filters>
  <projects>
    <project name="Project One" dirIn="/home/input"
dirOut="/home/output" marking="markings/markings.xml" filter="txt"
extension="txt"/>
    <project name="Project Two" dirIn="/home/input2"
dirOut="/home/output2" marking="markings/markings.xml" filter="txt"
extension="txt"/>
  </projects>
</config>
```

Figura 4.5: fitxer de configuració de projecte

Les propietats del projecte són : nom, directori d'entrada, directori de sortida, fitxer de marcatge, format (filtre) i extensió.

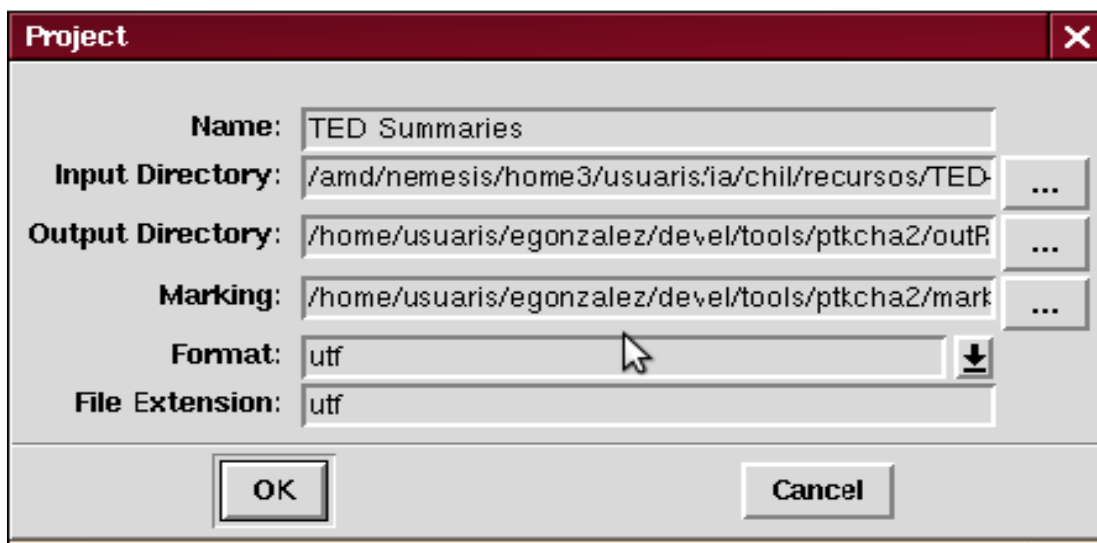


Figura 4.6: propietats del projecte

3. **Marcatges:** El marcatge ve definit per un fitxer de configuració on es defineixen tots els elements que es poden utilitzar en les anotacions. Els marcatges no es poden gestionar des d'una pantalla, s'han de fer des del mateix fitxer de configuració.

4. **Procés d'anotació:** La pantalla principal es divideix en quatre regions:

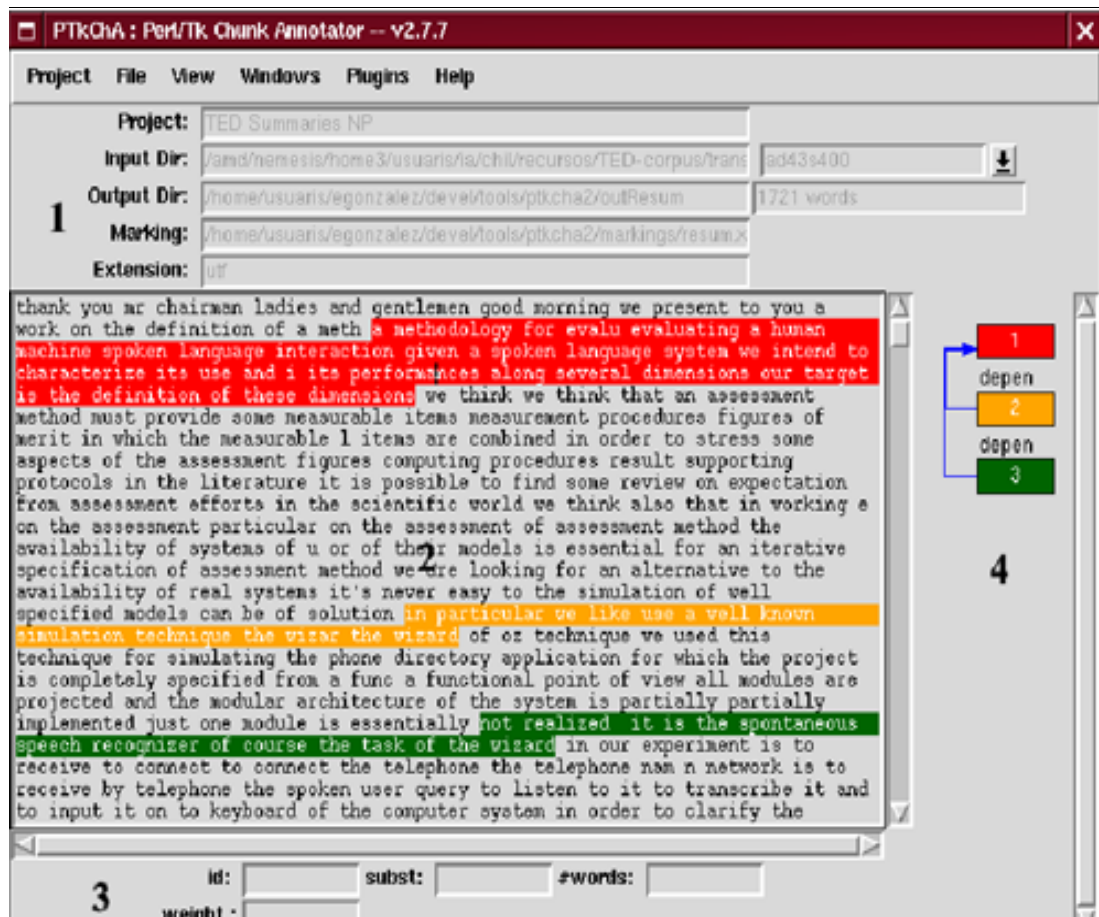


Figura 4.7: regions de la pantalla principal

4.1 Informació del projecte. Aquí es pot veure la informació del projecte actual. A la dreta es pot seleccionar el fitxer que es vol anotar i es poden veure les paraules del document (el nombre de paraules és informatiu, simplement conta els elements separats per espai en blanc).

4.2 Àrea d'anotació. Aquí es veu el text del fitxer seleccionat actualment i podem fer diferents operacions : crear una anotació nova, reduir una anotació existent, ampliar una anotació, editar-la,

canviar els valors dels atributs, afegir una relació entre les anotacions i esborrar una anotació.

4.3 Informació de l'anotació. Aquí es veu l'identificador d'anotació, si pertany a un grup, els valors dels atributs i el nombre de paraules de l'anotació.

4.4 Informació de les relacions. Aquí es veuen totes les relacions que té l'anotació actual.

FORMAT DE FITXER ANOTAT: El fitxer conté el text importat del directori d'entrada. Cada anotació es tanca dins d'un element XML. El tipus d'element ve determinat per l'atribut element del fitxer de marcatges. Els atributs de l'element són: un **ID** únic, un **SUBST** (si és un grup), els atributs i els valors de les anotacions.

Les relacions venen immediatament després de l'etiqueta d'obertura de l'anotació amb l'element **rels**, el qual consisteix en una llista d'elements **rel** que descriuen el tipus i el destí de la relació.

```
<ne id="1" subst="1" classe="PER"><rels><rel type="miembre_de" target="2" /></rels>Mas</ne>
```

concretó que las abstenciones que hicieron viable la aprobación por parte del Ejecutivo de sus dos grandes reformas responden a una necesidad de España, y no a una posición en la que <ne id="2" subst="2" classe="ORG"> CIU</ne> se sienta cómoda. "Si nos cierran el grifo desde fuera en España no se pagan ni las pensiones. Lo hicimos con desgana, pero el interés general está por delante de nuestras fobias y filias".

Figura 4.8: exemple de fitxer de sortida

En l'exemple, el text **Mas** està anotat amb el marcatge **ne**, amb **id** 1 i **subst** 1 i té l'atribut **classe** amb valor **PER**, i esta relacionat amb l'anotació 2.

El text **CIU** també està anotat amb el marcatge **ne**, amb **id** 2 i **subst** 2, té l'atribut **classe** amb valor **ORG**.

5.- ELABORACIÓ DE LA NOVA LLISTA DE REQUISITS⁴

Després d'avaluar les diferents funcionalitats i tenint en compte la llista inicial, s'elabora la nova llista de requisits.

Es decideix estructurar el procés d'elaboració de la llista en cinc apartats funcionals per poder agrupar les modificacions i així facilitar-ne l'anàlisi i el desenvolupament. Això és molt important, ja que durant tota la realització del projecte seguirem aquesta estructura.

Els cinc apartats són:

5.1. Menú

5.2 Gestió de projectes

5.3 Gestió de marcatges

5.4 Gestió d'usuaris

5.5 Procés d'anotacions.

5.1 Menú:

En aquest apartat es defineixen els accessos del menú.

En afegir la gestió d'usuaris (veure punt 4) s'ha afegit l'accés a aquest, segons el tipus d'usuari.

En afegir l'estat del projecte com a obert o tancat (veure punt 2) es dóna accés, des del menú, a poder activar i desactivar projectes com usuari anotador.⁵

En canviar la gestió de marcatges, la vista per atributs s'ha de modificar; s'han d'afegir els atributs, les relacions, o els grups de l'etiquetatge actiu en el menú per poder veure les anotacions segons els afegits. Igualment s'han de modificar les estadístiques dels atributs i els plugins.

Les modificacions, doncs, que s'han de realitzar dins l'apartat del menú són:

- Accedir a Gestió de projectes com a usuari administrador.
- Accedir a Gestió de marcatges com a usuari administrador.

⁴ Veure resum de la nova llista de requisits a *Annex 1*.

⁵ L'usuari administrador pot fer totes les funcions de l'anotador però no a l'inrevés.

- Accedir a Gestió d'usuaris com a usuari administrador.
- Activar el projecte com a usuari anotador.
- Desactivar el projecte com a usuari anotador.
- La vista d'atributs modificar segons etiquetatge actiu.
- Veure estadístiques per atributs.
- Accedir a plugins segons marcatges.

5.2 Gestió de projectes:

Les modificacions per a l'apartat de gestió de projectes són:

- Es decideix afegir un estat (obert, tancat) al projecte per poder controlar en quins projectes s'està treballant i en quins no. En conseqüència, l'acció d'obrir un projecte per anotar, li direm "activar".
- Poder veure la llista de projectes disponibles segons el seu estat.
- Poder afegir i treure etiquetatges per anotar (veure punt 3) al projecte.
- Poder modificar l'estat d'un projecte.
- Poder activar projectes per anotar.
- Poder desactivar el projecte actiu.
- Fer un control en esborrar un projecte. Un projecte que ja ha estat anotat no es pot esborrar.
- En activar un projecte, comprovar que els etiquetatges i els elements dels marcatges (veure punt 5.3 gestió de marcatges) de tots els seus fitxers estan definits.
- Es decideix, també, treure els filtres XML, bio i via del projecte i prendre com a requisit que els fitxers d'entrada són de text.

5.3 Gestió de marcatges:

En la gestió de marcatges el que es vol és poder mantenir-los des de la mateixa aplicació, sense necessitat de modificar-ne el fitxer manualment, a més a més de poder gestionar-ne més d'un alhora per poder ser utilitzats en el procés d' anotació.

Es decideix dividir-los en: etiquetatges, atributs, relacions, grups, plugins i extres.

Els etiquetatges defineixen el tipus d' anotacions, relacions o grups que es faran en el procés d' anotació. Són l'etiqueta element que definia el fitxer de marcatge i que ens permetran veure les anotacions que hi pertanyen en l'àrea d' anotació.

S'afegeix el concepte d'etiquetatge visible per poder indicar quines anotacions d'altres etiquetatges són visibles en el procés d' anotació.

Un etiquetatge estarà format o per atributs, o per relacions, o per grups. És a dir, només podrà tenir elements d'aquests tipus. Això es decideix perquè es vol que cada etiquetatge defineixi una tasca concreta a fer.

Els plugins i els extres, com que són elements que s'utilitzaran en el procés d' anotació, es permeten afegir a qualsevol etiquetatge però sense fer-ne una gestió, ja que aquests s'utilitzen en casos concrets. Són elements de l'etiquetatge però només d'un etiquetatge.

Les modificacions per cada un dels apartats de la gestió de marcatges són:

5.3.1 Gestió d'etiquetatges:

- Poder donar d'alta etiquetatges
- Poder modificar etiquetatges
- Poder esborrar etiquetatges
- Poder indicar als etiquetatges quins altres etiquetatges són visibles en el procés d' anotació.
- Poder afegir i treure elements a l'etiquetatge.

- Fer un control d'etiquetatges a l'hora de modificar o esborrar. No deixar esborrar un etiquetatge si està anotat en algun fitxer del projecte. No deixar treure elements de l'etiquetatge si aquest ja està anotat en algun fitxer del projecte.

5.3.2 Gestió d'atributs:

Els valors dels atributs que es permeten són de tipus enumerats, només es pot seleccionar un valor dels definits en el procés d'anotació. Volem afegir-ne de nous i que en el procés d'anotació, segons el tipus, es demani un valor a introduir.

- Afegir tipus d'atribut : Enter, cadena, enumerat (llista de valors) i rang (valor entra un mínim i un màxim).
- Poder donar d'alta atributs
- Poder modificar atributs
- Poder esborrar atributs
- Fer un control en modificar i esborrar atributs. Si estan en un etiquetatge, no deixar ni esborrar ni modificar.

5.3.3 Gestió de relacions:

Les relacions es poden fer entre dues anotacions qualssevol, es vol controlar que les anotacions origen i destí de la relació siguin del tipus d'etiquetatge que marca la relació.

- Afegir tipus de relació. Defineix els etiquetatges origen i destí de la relació.
- Poder donar d'alta relacions.
- Poder modificar relacions.
- Poder esborrar relacions.
- Fer un control en modificar i esborrar relacions. Si estan en un etiquetatge, no deixar ni esborrar ni modificar.

5.3.4 Gestió de grups:

Els grups són relacions entre varies anotacions, relació de grup.

Actualment, els grups no se'ls pot identificar amb un nom, són identificadors numèrics. És per això que es volen gestionar.

- Poder donar d'alta grups.
- Poder modificar grups.
- Poder esborrar grups.
- Fer un control en modificar i esborrar grups. No deixar ni modificar ni esborrar el grup si està en un etiquetatge.

5.4 Gestió d'usuaris

Actualment no hi ha gestió d'usuaris.

Es vol diferenciar dos tipus d'usuaris : Administradors i anotadors.

El primer podrà gestionar els projectes i els marcatges, igualment podrà fer tot el que pot fer l'anotador.

El segon podrà activar un projecte per anotar-lo.

Només es vol controlar els usuaris administradors, que es guardaran en el fitxer de configuració de l'aplicació. En entrar a l'aplicació, aquesta obtindrà l'usuari del sistema i si està a la llista d'administradors, podrà gestionar els projectes i els marcatges.

- Control d'usuari a l'entrar a l'aplicació.
- Definir les funcionalitats dels usuaris.
- Poder afegir usuari administrador.
- Poder treure usuari administrador.
- Al treure usuari administrador, si aquest és **root** o **administrador** o **administrator**, no deixar treure, és a dir, que siguin usuaris administradors per defecte.

5.5 Procés d'anotacions

En el procés d'anotació, les modificacions seran les següents:

- Aplicar tots els canvis de la gestió de marcatges al procés d'anotació.
 - Nova anotació
 - Només fer anotació si l'etiquetatge actiu és d'atributs.
 - Posar valors per defecte dels atributs de l'etiquetatge actiu.
 - Canviar valors dels atributs

- Mostrar atributs si l'etiquetatge de l'anotació és igual a l'etiquetatge actiu i és d'atributs.
- Demandar i validar valors segons tipus de l'atribut.
- Afegir relació entre anotacions
 - Mostrar les relacions possibles a fer si l'etiquetatge actiu és de relacions i mostrar-ne només aquelles on l'etiquetatge origen sigui igual a l'etiquetatge de l'anotació.
 - Validar origen i destí de la relació al fer-la.
- Afegir anotació a grup
 - Mostrar els grups possibles a fer si l'etiquetatge actiu és de grups.
 - Poder afegir anotacions al grup seleccionant-les.
- Veure anotacions en el text
 - Veure les anotacions de l'etiquetatge actiu i dels seus etiquetatges visibles.
 - Veure les anotacions segons els seus atributs. (Punt 5.1 menú vista atributs).
 - Veure les anotacions segons les seves relacions. (Punt 5.1 menú vista atributs).
 - Veure les anotacions segons els seus grups.
- Veure informació de les anotacions
 - Veure els atributs i els seus valors al seleccionar anotació.
 - Veure les relacions al seleccionar anotació.
 - Veure els grups al seleccionar anotació.
- Aplicar etiquetatges del projecte al procés d'anotació.
- Canviar el format del fitxer d'anotacions. Es vol separar en etiquetes XML el text anotat, les relacions i els grups.
- Poder seleccionar entre els fitxers anotats i no anotats.
- Anotacions aniuades.
 - Permetre anotacions aniuades
 - Diferenciar visualment l'anotació *pare* de les seves aniuades
- Procés d'unió d'etiquetatges.

6.- REPLANIFICACIÓ

6.1.-Temps estimat

Dates d'inici i finalització del projecte:

Per especificar i desenvolupar el projecte es disposa de cinquanta-un dies, que abasten des del dia 19 d'abril del 2010 fins el dia 28 de juny del 2010; data límit d'entrega del projecte.

A partir d'aquí i segons aquesta previsió, parlo amb en Lluís i en Jordi per treure funcionalitats al PTKChA , és a dir, treure les menys necessàries en aquest moment, i deixar-les per un posterior projecte.

Arribem a l'acord de:

- Treure l'opció de Veure estadístiques per atributs.
- Accedir a plugins segons marcatges. Deixem que els plugins es carreguin al menú, es deixa gestionar des de marcatges però no es modifiquen els plugins actuals a les modificacions de marcatges.
- Treure els extres, el control d'etiquetes XML no cal ja que els fitxers d'entrada son de format text.
- En les anotacions aniuades deixar la visualització d'anotacions per colors com està i no diferenciar-les per nivells.

6.2.- Distribució de les tasques

Tasques a realitzar:

- Model conceptual
- Casos d'ús
- Diagrama de classes
- Disseny de pantalles
- Implementació:
 - Menú
 - Gestió de projectes
 - Gestió de marcatges
 - Gestió d'usuaris
 - Procés d'anotació

- Manuals:
 - Manual d'instal·lació
 - Manual d'annotador
 - Manual d'administrador
- Exemples d'ús
- Memòria

Queda distribuïda, doncs de la següent manera:

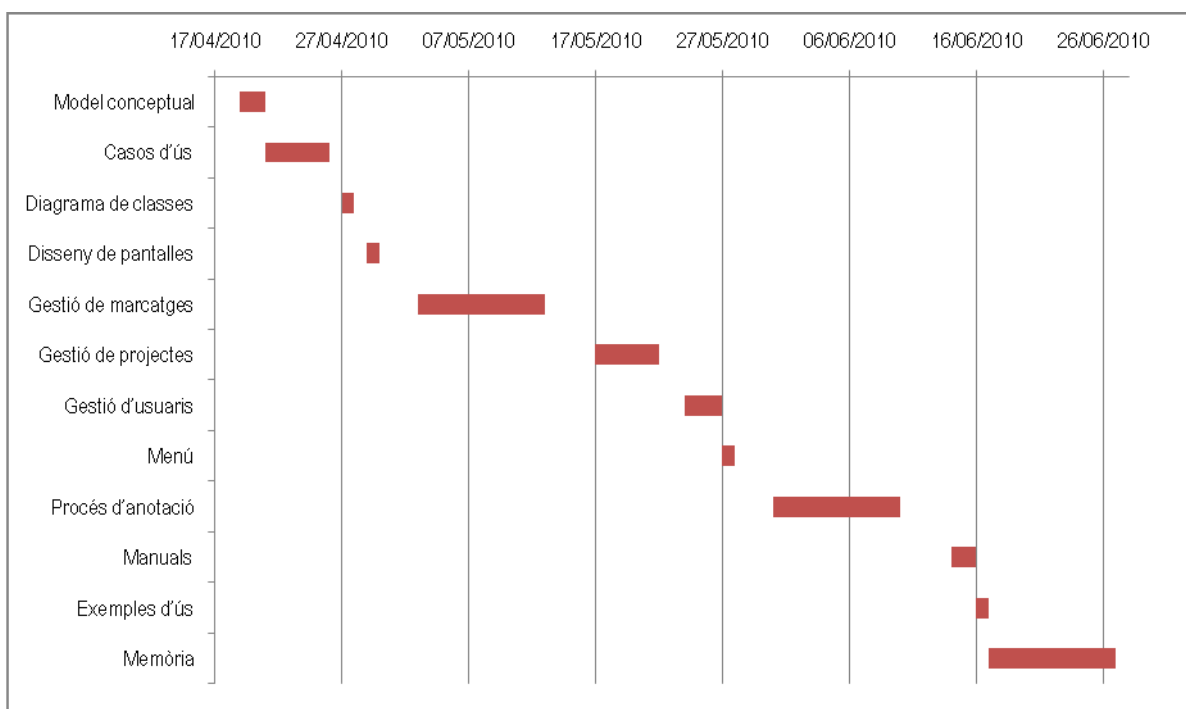


Figura 6.1: diagrama de Gantt de la replanificació

6.3.- Cost previst de la replanificació

Encara que tot el projecte ha estat realitzat per una sola persona, per calcular el preu aproximat del cost que suposaria realitzar aquest projecte amb un equip de desenvolupament format per un analista, un programador i un dissenyador, he tingut en compte aquestes tres últimes categories per repartir les tasques a fer.

D'aquesta manera:

- L'analista s'encarregarà de realitzar les tasques següents : model conceptual, casos d'ús, diagrama de classes.
- El programador s'encarregarà de realitzar les tasques següents: implementació, manuals, exemples d'ús.
- El dissenyador es dedicarà al disseny de les pantalles.

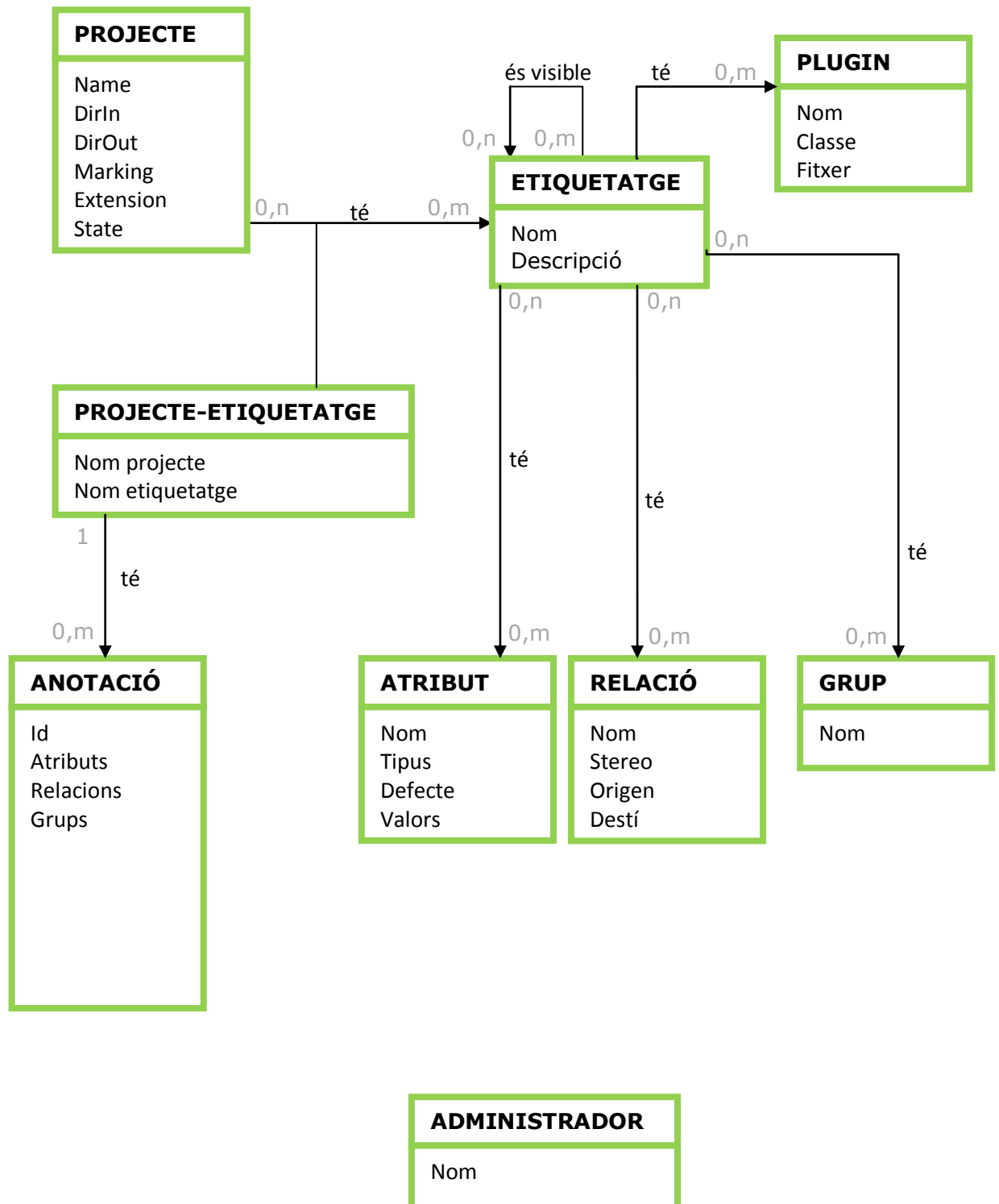
Els preus per hora de cadascuna de les categories professionals, és de:

- Analista: 50€/hora
- Programador: 35€/hora
- Dissenyador: 30€/hora

Concepte	Hores	Total
Analista	64	3200€
Programador	320	11200€
Dissenyador	8	240€
Total		15040€

Figura 6.2: taula de cost previst en la replanificació

7.- MODEL CONCEPTUAL



L'aplicació, actualment, té dos fitxers de configuració: el ptkcha.xml, on es guarda la informació dels projectes i opcions generals, i el markings.xml, on es guarda tota la informació dels marcatges.

El que he cregut convenient és fer el model conceptual per veure quins elements de la llista dels requisits són els que afecten als fitxers ptkcha.xml i markings.xml.

A continuació, es detalla de cada entitat quins són els canvis que faran als fitxers.

7.1. Entitat administradors: es decideix guardar els usuaris dins el fitxer de configuració ptkcha.xml, doncs només cal guardar el nom dels usuaris administradors, no cal guardar cap més informació. Aquests seran validats amb l'usuari del sistema. Si aquest usuari està a la llista d'usuaris administradors entrarà com a tal, si no, entrarà com a usuari anotador.

```
<users>
<user name="root" />
<user name="administrator" />
</users>
```

Figura 7.1:format xml dels usuaris administradors

7.2. Entitat projecte: s'afegeix la propietat "estat" al projecte. Així com també es defineix la relació entre projectes i etiquetatges i es modifica la propietat que guardava el fitxer de marcatge per tal que guardi l'etiquetatge actiu.

Es decideix guardar els etiquetatges d'un projecte en el fitxer actual de configuració ptkcha.xml, ja que també els projectes hi estan guardats.

```
<project name="NomProjecte" dirIn="RutaEntrada" dirOut="RutaSortida"
marking="EtiquetatgeActiu" extensio="ExtensioFitxersEntrada" state="EstatProjecte">
  <etiquetatge name="NomEtiquetatge" />
</project>
```

Figura 7.2: format xml d'un projecte

```
<config version="1.1.xml">
<markingfile file="markings/markings.xml" />
<option name="SelExp" value="1" />
<option name="ForWri" value="1" />
<option name="IncDir" value="/home/administrador/PFC/ptkcha2/Plugins" />
<project name="Projecte 1" dirIn=" home/administrador/PFC/ptkcha2/input" dirOut="
home/administrador/PFC/ptkcha2output" marking="ne" extensio="txt" state="EstatProjecte">
    <etiquetatge name="city" />
</project>
<users>
<user name="Eze" />
<user name="administrador" />
<user name="administrador" />
<user name="root" />
</users>
</config>
```

Figura 7.3: fitxer de configuració del PTKChA

7.3. Entitat etiquetatge: es defineix amb el nom i la descripció. Els etiquetatges seran guardats en el fitxer de configuració de marcatges⁶.

Els etiquetatges es relacionen amb ells mateixos per indicar quines anotacions de l'etiquetatge són visibles. Alhora, es relacionen amb els atributs, les relacions i els grups, de manera que aquests formen part dels etiquetatges.

Els plugins també formaran part de l'etiquetatge però sense gestionar-los, s'afegiran directament a l'etiquetatge.

7.4. Entitat atribut: en aquesta entitat, el que s'està fent és afegir la propietat "tipus", segons la qual el valor de l'atribut serà un enter, un valor cadena, un enumerat o un rang.

Els atributs es guarden en el fitxer de configuració de marcatges⁷.

7.5. Entitat relació: les propietats de les relacions són el nom, l'stereo i s'afegeix un origen i un destí. També es guarden en el fitxer de configuració de marcatges⁸.

7.6. Entitat grup: es defineix amb el nom i, com en els anteriors, es guarda en el fitxer de configuració de marcatges⁹.

7.7. Entitat plugin: els plugins formen part de l'etiquetatge però no es gestionen, sinó que s'afegeixen directament a l'etiquetatge.

⁶ Veure figura 7.4 i 7.5

⁷ Veure figura 7.4 i 7.5

⁸ Veure figura 7.4 i 7.5

⁹ Veure figura 7.4 i 7.5


```

<?xml encoding='UTF-8'?>

<!ELEMENT markings (etiquetatge|atribut|relacio|grup|plugin)*>

<!-- .... Atributs ..... -->

<!ATTLIST atribut nom ID #REQUIRED>
<!ATTLIST atribut tipus ('int'|'string'|'rang'|'enum') #REQUIRED>
<!ATTLIST atribut defecte #IMPLIED>
<!ELEMENT atribut (#PCDATA?|rang*|enum*)>

<!-- Rang-->

<!ELEMENT rang EMPTY>
<!ATTLIST rang min CDATA #REQUIRED>
<!ATTLIST rang max CDATA #REQUIRED>

<!-- Enumerats -->

<!ELEMENT enum (valor*)>
<!ELEMENT valor EMPTY>
<!ATTLIST valor v CDATA #REQUIRED>

<!-- .... Relacions .... -->

<!ELEMENT relacio EMPTY>
<!ATTLIST relacio nom ID #REQUIRED>
<!ATTLIST relacio stereo ('uni'|'bi') 'uni'>
<!ATTLIST relacio origen CDATA #REQUIRED>
<!ATTLIST relacio desti CDATA #REQUIRED>

<!-- .... Etiquetatges .... -->

<!ELEMENT etiquetatge (plugin|usaAtribut|usaRelacio|UsaEtiquetatge)*>

<!-- Etiquetatge-Atributs -->

<!ELEMENT usaAtribut EMPTY>
<!ATTLIST usaAtribut nom IDREF #REQUIRED>

<!-- Etiquetatge-Relacions -->

<!ELEMENT usaRelacio EMPTY>
<!ATTLIST usaRelacio nom IDREF #REQUIRED>

<!-- Etiquetatge-Etiquetatge -->

<!ELEMENT usaEtiquetatge EMPTY>
<!ATTLIST usaEtiquetatge nom IDREF #REQUIRED>

<!-- Plugins -->

<!ELEMENT plugin EMPTY>
<!ATTLIST plugin nom ID #REQUIRED>
<!ATTLIST plugin classe CDATA #REQUIRED>
<!ATTLIST plugin fitxer CDATA #REQUIRED>

```

Figura 7.4: fitxer DTD dels marcatges

```

<?xml version='1.0' encoding='UTF-8'?>
<!DOCTYPE markings SYSTEM 'markings.dtd'>
<markings>
<atribut nom='classe' tipus='enum' defecte='PER'>
  <enum>
    <valor v='PER' />
    <valor v='ORG' />
    <valor v='LOC' />
    <valor v='MISC' />
  </enum>
</atribut>
<atribut nom='place' tipus='enum' defecte='L_building'>
  <enum>
    <valor v='L_building' />
    <valor v='L_station' />
    <valor v='L_airport' />
    <valor v='L_port' />
  </enum>
</atribut>
<atribut nom='multiword' tipus='enum' defecte='no'>
  <enum>
    <valor v='yes' />
    <valor v='no' />
  </enum>
</atribut>
<atribut nom='translate' tipus='string' defecte='' />
<atribut nom='weight' tipus='enum' defecte='1'>
  <enum>
    <valor v='1' />
    <valor v='2' />
    <valor v='3' />
  </enum>
</atribut>
<atribut nom='_L' tipus='enum' defecte='L_city'>
</atribut>
<relacio nom='co-ne' stereo='bi' origen='ne' desti='ne' />
<relacio nom='co-ne-nel' stereo='uni' origen='ne' desti='neLocation' />
<grup nom='Anglesa' />
<grup nom='Catalana' />
<etiquetatge nom='city' des='Ciutats'>
  <usaGrup nom='Anglesa' />
  <usaGrup nom='Catalana' />
  <usaEtiquetatge nom='ne' />
  <usaEtiquetatge nom='neLocation' />
</etiquetatge>
<etiquetatge nom='coref' des='Correferencia'>
  <usaRelacio nom='co-ne' />
  <usaEtiquetatge nom='ne' />
</etiquetatge>
<etiquetatge nom='ne' des='Named Entity'>
  <usaAtribut nom='classe' />
  <usaAtribut nom='multiword' />
</etiquetatge>
<etiquetatge nom='summary' des='Summary'>
  <usaAtribut nom='weight' />
  <plugin nom='Summary' classe='PluginResum' fitxer='PluginResum' />
</etiquetatge>
</markings>

```

Figura 7.5: fitxer XML dels marcatges

7.8. Entitat anotació: de les anotacions es vol guardar a quin etiquetatge pertany, quins atributs amb els seus valors té, les relacions entre anotacions i els grups a que pertanyen.

Les anotacions actualment són guardades en el fitxer de sortida, i el que es vol fer és canviar el format d'aquest fitxer. Les anotacions, però, seran guardades amb el mateix format.

El text anotat s'englobarà amb una etiqueta XML. El nom d'aquesta etiqueta serà "text".

Les relacions no es guardaran juntament amb les anotacions, sinó que es guardaran amb una altra etiqueta XML, amb el nom de "relation". Cada relació anirà delimitada per l'etiqueta "rel".

Pels grups tindrem l'etiqueta "clusters" i cada grup anirà delimitat per l'etiqueta "cluster".

D'aquesta manera el format dels fitxers de sortida serà:

```
<text>
<ne id="1" classe="PER">Mas</ne> concretó que las abstenciones que hicieron viable la
aprobación por parte del Ejecutivo de sus dos grandes reformas responden a una necesidad
de España, y no a una posición en la que <ne id="2" classe"ORG"> CiU</ne> se sienta
cómoda. "Si nos cierran el grifo desde fuera en España no se pagan ni las pensiones. Lo
hicimos con desgana, pero el interés general está por delante de nuestras fobias y filias".
</text>
<relations>
  <rel type="membre_de" start="1" target="2" />
</relations>
<clusters>
  <cluster name="politic" id="1" />
</clusters>
```

Figura 7.6: fitxer XML de sortida o anotat

En l'exemple, el text **Mas** està anotat amb el marcatge **ne**, amb **id** 1 i té l'atribut **classe** amb valor **PER**.

El text **CIU** també està anotat amb el marcatge **ne**, amb **id** 2 i té l'atribut **classe** amb valor **ORG**.

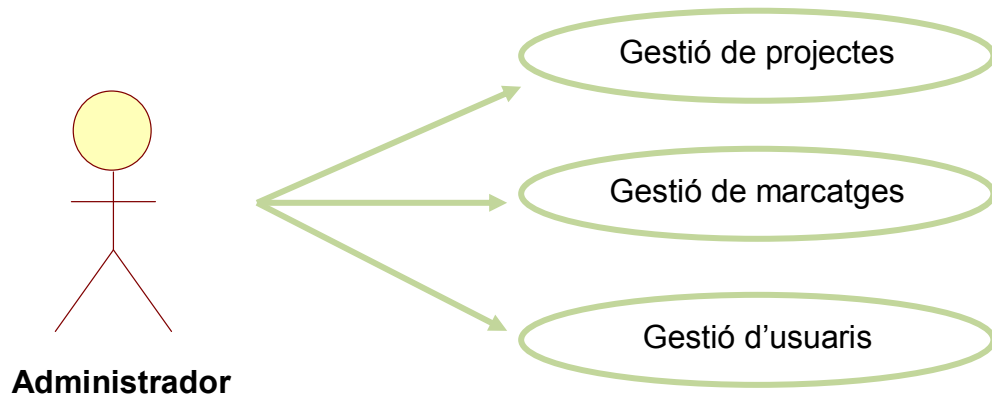
Hi ha una relació entre l'anotació 1 i l'anotació 2. Mas és membre de CIU.

L'anotació 1 pertany al grup polític. Mas està dins del grup polític.

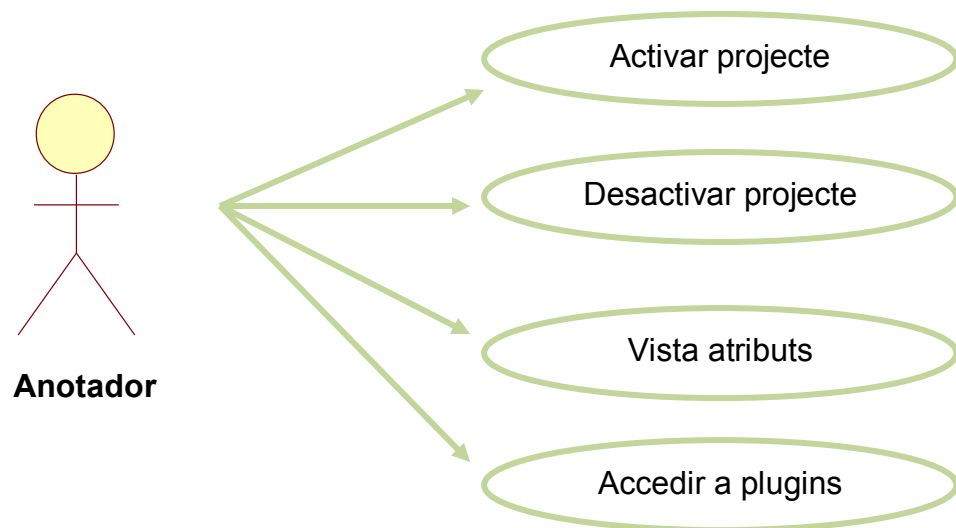
8.- CASOS D'ÚS

8.1 Menú

8.1.1 Punt de vista administrador

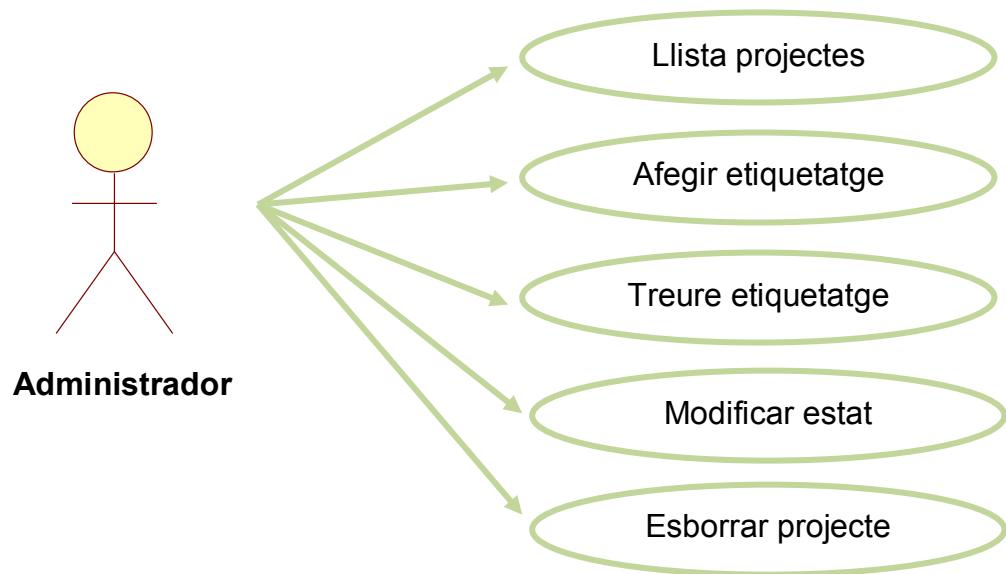


8.1.2 Punt de vista anotador

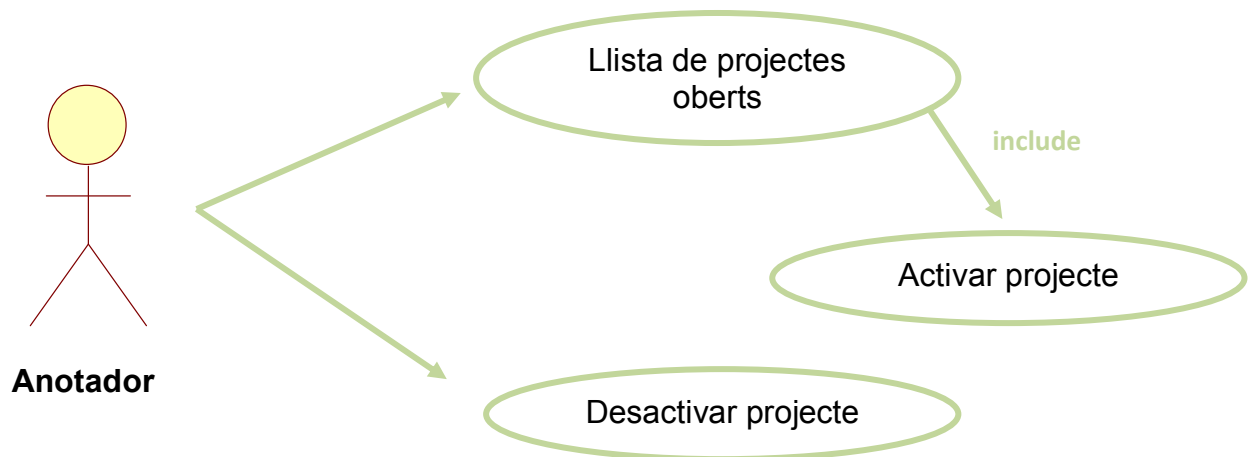


8.2 Gestió de projectes

8.2.1 Punt de vista administrador

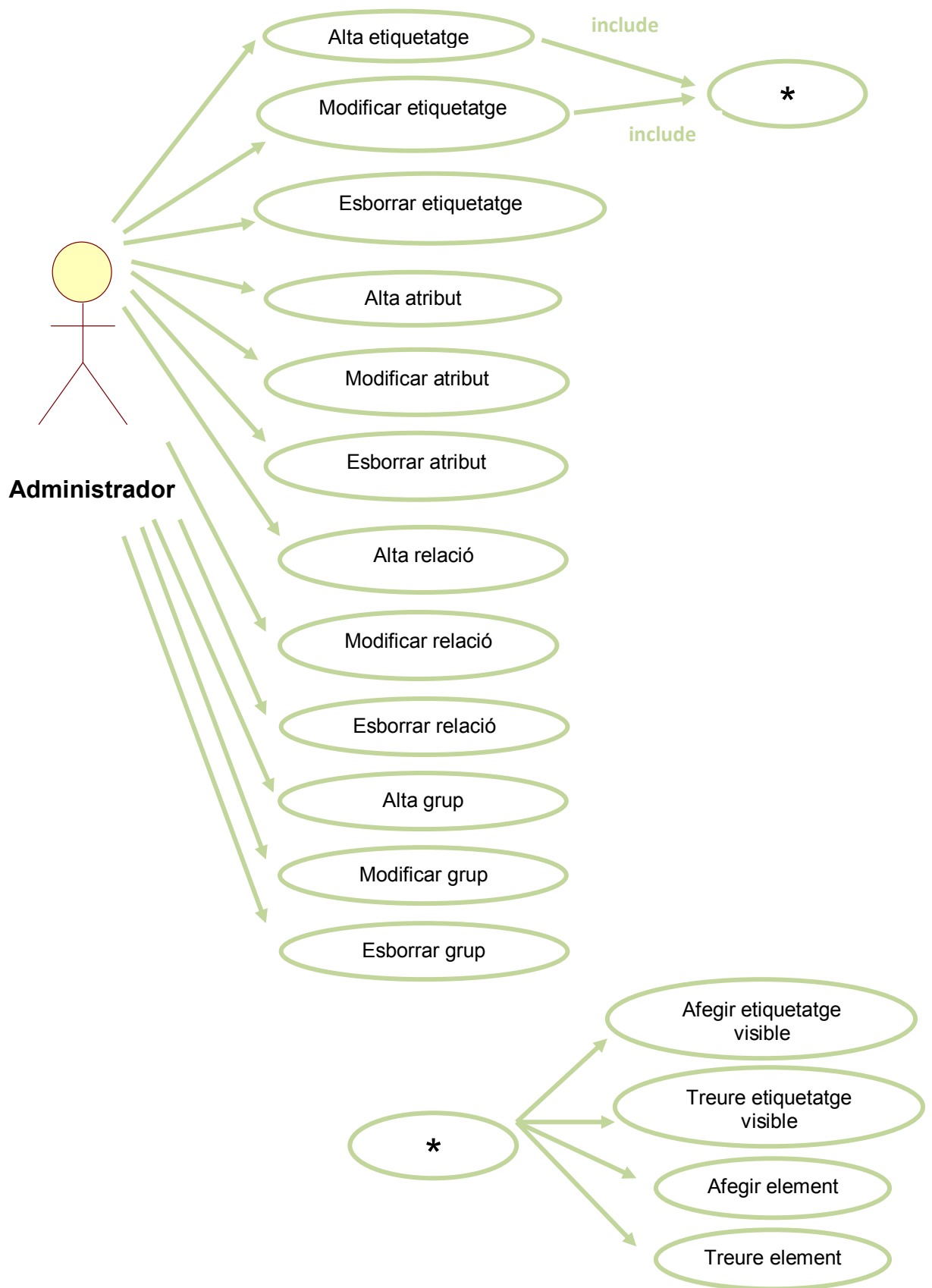


8.2.2 Punt de vista anotador



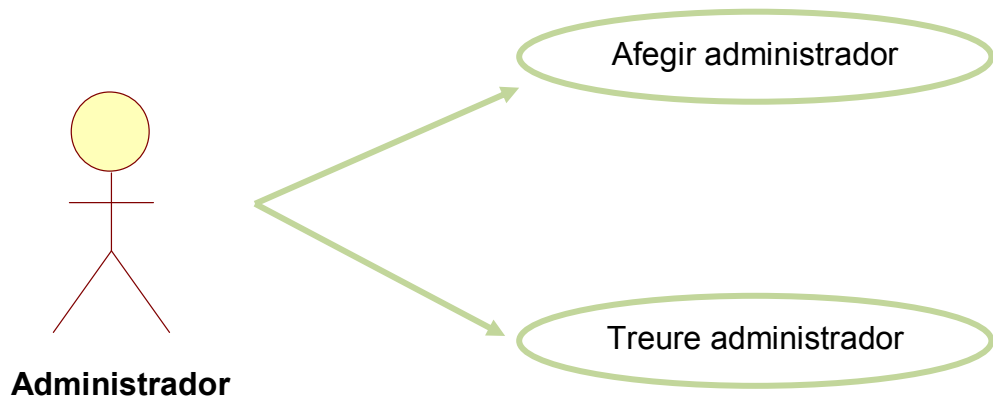
8.3 Gestió de marcatges

8.3.1 Punt de vista administrador



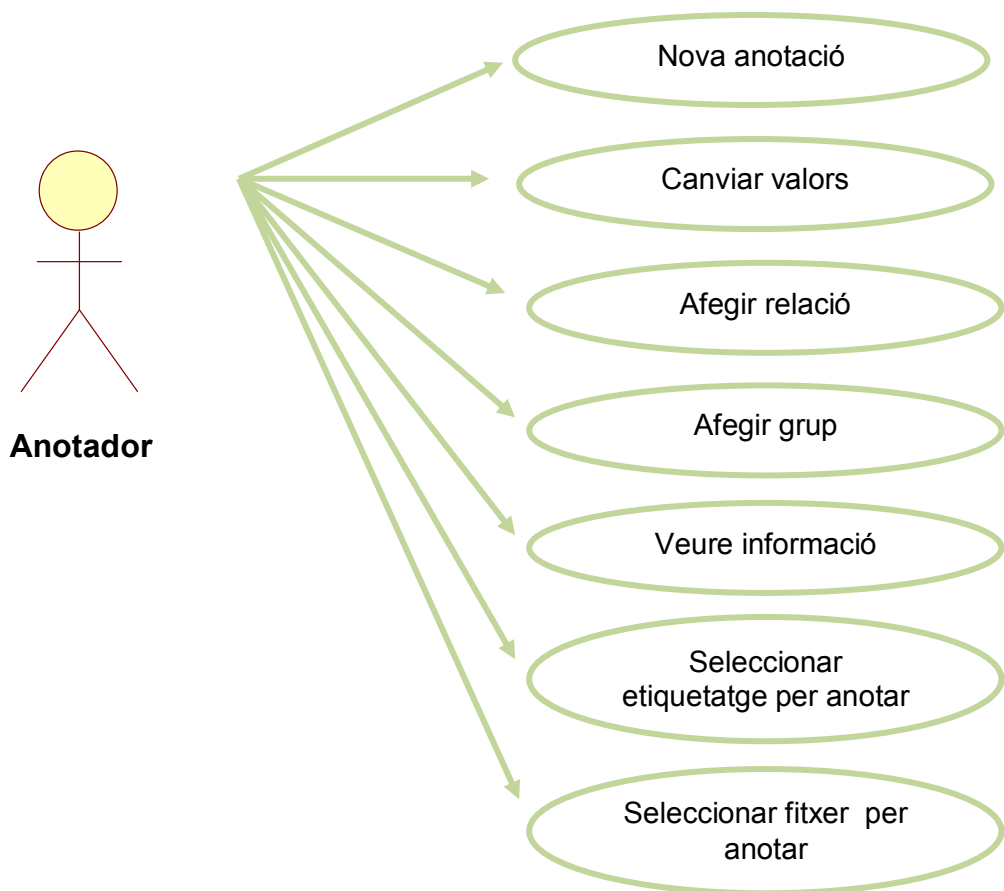
8.4 Gestió d'usuaris

8.4.1 Punt de vista administrador



8.5 Procés d'anotació

8.5.1 Punt de vista anotador



CASOS D'ÚS DEL MENÚ

Cas d'ús	Gestió de projectes
Actors	Administrador
Iniciador	Administrador
Descripció	Obrir la gestió de projectes
Precondició	
CURS NORMAL	
1. El cas d'ús comença quan l'administrador vol entrar a gestionar els projectes.	
	2. El sistema mostra la pantalla de gestió de projectes per usuari administrador i mostrar una llista dels projectes oberts.

Cas d'ús	Gestió de marcatges
Actors	Administrador
Iniciador	Administrador
Descripció	Obrir la gestió de marcatges
Precondició	
CURS NORMAL	
1. El cas d'ús comença quan l'administrador vol entrar a gestionar els marcatges.	
	2. El sistema mostra la pantalla de gestió de marcatges i llista els marcatges.

Cas d'ús	Gestió d'usuaris
Actors	Administrador
Iniciador	Administrador
Descripció	Obrir la gestió d'usuaris
Precondició	
CURS NORMAL	
1. El cas d'ús comença quan l'administrador vol entrar a gestionar els usuaris.	
	2. El sistema mostra la pantalla de gestió d'usuaris i llista els usuaris administradors.

Cas d'ús	Activar un projecte (include : Llista de projectes oberts)
Actors	Anotador
Iniciador	Anotador
Descripció	L'anotador activa un projecte
Precondició	
CURS NORMAL	
1. El cas d'ús comença quan l'anotador vol activar un projecte per treballar-hi.	
	2. include : Llista de projectes oberts.

Cas d'ús	Desactivar un projecte
Actors	Anotador
Iniciador	Anotador
Descripció	L'anotador desactiva el projecte actiu
Precondició	
CURS NORMAL	
1. El cas d'ús comença quan l'anotador vol desactivar un projecte.	
	2. El sistema desactiva el projecte.
CURS ALTERNATIU	
Cas 2: Si no hi ha cap projecte actiu no fer res i se surt.	

Cas d'ús	Vista atributs
Actors	Anotador
Iniciador	Anotador
Descripció	L'anotador vol veure les anotacions per colors segons els valors dels elements de l'etiquetatge actiu.
Precondició	Hi ha projecte activat.
CURS NORMAL	
1. El cas d'ús comença quan l'anotador selecciona un element del punt de menú.	
	2. El sistema mostra les anotacions de l'etiquetatge actiu i dels visibles pel color segons el valor de l'element seleccionat.

Cas d'ús	Accedir a plugins	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol cridar un plugin	
Precondició	Hi ha projecte activat i l'etiquetatge actiu té plugins	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol cridar un plugin		
2. L'anotador selecciona un plugin del punt de menú		
		3. El sistema crida el plugin

CASOS D'ÚS DE GESTIÓ DE PROJECTES

Cas d'ús	Llista projectes	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol veure els projectes per estat	
Precondició	Gestió de projectes obert	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol veure els projectes per estat.		
2. L'administrador selecciona un estat del projecte.		
		3. El sistema mostra els projectes de l'estat seleccionat.

Cas d'ús	Afegir etiquetatge	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol afegir un etiquetatge al projecte per ser utilitzat en el procés d'anotació.	
Precondició	S'està modificant el projecte	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol afegir un etiquetatge al projecte.		
		2. El sistema mostra una llista amb els etiquetatges disponibles.
3. L'administrador selecciona un etiquetatge		
		4. El sistema valida que l'etiquetatge es pugui afegir.
		3. El sistema afegeix l'etiquetatge al projecte.
CURS ALTERNATIU		
Cas 4 : Si l'etiquetatge ja està afegit, el sistema no fa res i torna al cas 1.		

Cas d'ús	Treure etiquetatge	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol treure un etiquetatge del projecte	
Precondició	S'està modificant el projecte	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol treure un etiquetatge del projecte.		
		2. El sistema mostra una llista dels etiquetatges en el projecte
3. L'administrador selecciona un etiquetatge		
		4. El sistema valida que l'etiquetatge es pugui treure.
		5. El sistema treu l'etiquetatge del projecte.
CURS ALTERNATIU		
Cas 4 : Si el projecte té algun fitxer anotat amb l'etiquetatge, no es treu i dona un missatge d'avís. Es torna al cas 2.		

Cas d'ús	Modificar l'estat del projecte	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol modificar l'estat del projecte	
Precondició	S'està modificant el projecte	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol modificar l'estat del projecte.		
		2. El sistema mostra les dades del projecte.
3. L'administrador modifica les dades		
		4. El sistema valida les dades
		5. El sistema modifica les dades del projecte.
CURS ALTERNATIU		
Cas 4 : Si hi ha alguna dada incorrecte del projecte, dóna un missatge d'error i es torna al cas 2.		

Cas d'ús	Esborrar projecte	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol esborrar un projecte	
Precondició	Gestió de projectes oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol esborrar un projecte.		
2. L'administrador selecciona un projecte		
		3. El sistema valida si es pot esborrar el projecte.
		4. El sistema esborra el projecte.
CURS ALTERNATIU		
Cas 3: Si el projecte té algun fitxer anotat no deixa esborrar el projecte, dóna un missatge d'error i es surt.		
Cas 3: Si hi ha més d'un projecte amb el mateix directori d'entrada o de sortida no deixa esborrar projecte.		

Cas d'ús	Llista de projectes oberts (include Activar projecte)	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'annotador vol veure els projectes oberts per activar-ne un	
Precondició		
CURS NORMAL		
1. El cas d'ús comença quan l'annotador vol veure els projectes oberts.		
		2. El sistema mostra la llista de projectes oberts.
3. L'annotador selecciona un projecte		
		4. Include (Cas d'ús Activar projecte)

Cas d'ús	Activar projecte	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'annotador vol activar un projecte	
Precondició	Gestió de projectes obert	
CURS NORMAL		
1. El cas d'ús comença quan l'annotador vol activar un projecte.		
2. L'annotador selecciona projecte.		
		2. El sistema comprova en tots els fitxers anotats del projecte que els etiquetages i els elements estan en els marcatges.
		3. El sistema activa el projecte. Mostra les dades del projecte actual, els fitxers per anotar i els anotats i el text del primer fitxer per anotar.
CURS ALTERNATIU		
Cas 2 : Si la comprovació dóna error es mostrar un missatge d'avís i se surt.		

Cas d'ús	Desactivar projecte	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'annotador vol desactivar el projecte actiu	
Precondició	Hi ha projecte actiu	
CURS NORMAL		
1.El cas d'ús comença quan l'annotador vol desactivar el projecte actiu.		
		2. El sistema desactiva el projecte actiu.

CASOS D'ÚS DE GESTIÓ DE MARCATGES

Cas d'ús	Alta etiquetatge	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol donar d'alta un etiquetatge	
Precondició	No hi ha projecte actiu i gestió de marcatge oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol donar d'alta un etiquetatge.		
		2. El sistema demana les dades de l'etiquetatge. Include (afegir etiquetatge visible, treure etiquetatge visible, afegir element i treure element).
3. L'administrador introdueix les dades de l'etiquetatge.		
		4. El sistema valida les dades de l'etiquetatge.
		5. El sistema guarda l'etiquetatge.
CURS ALTERNATIU		
Cas 4: Si les dades no són correctes es mostra un missatge d'error i es torna al cas 2.		

Cas d'ús	Modificar etiquetatge	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol modificar un etiquetatge	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol modificar un etiquetatge.		
2. L'administrador selecciona un etiquetatge.		
		3. El sistema mostra les dades a modificar. Include (afegir etiquetatge visible, treure etiquetatge visible, afegir element i treure element).
		4. El sistema modifica l'etiquetatge.
CURS ALTERNATIU		
Cas 3: Si les dades no són correctes es mostra un missatge d'error i es torna al cas 2.		

Cas d'ús	Esborrar etiquetatge	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol esborrar un etiquetatge	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol esborrar un etiquetatge.		
2. L'administrador selecciona un etiquetatge.		
		3. El sistema comprova si es pot esborrar l'etiquetatge.
		4. El sistema esborra l'etiquetatge.
CURS ALTERNATIU		
Cas 3: Si l'etiquetatge està anotat en algun fitxer del projecte, dona un missatge d'avís i se surt.		

Cas d'ús	Afegir etiquetatge visible	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol afegir un etiquetatge visible	
Precondició	S'està modificant l'etiquetatge	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol afegir un etiquetatge visible.		
		2. El sistema mostra tots els etiquetatges disponibles.
3. L'administrador selecciona un etiquetatge.		
		4. El sistema comprova que es pot afegir l'etiquetatge.
		5. El sistema afegeix l'etiquetatge a etiquetatge visible.
CURS ALTERNATIU		
Cas 4: Si l'etiquetatge ja hi és no es fa res i es torna al cas 2.		

Cas d'ús	Treure etiquetatge visible	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol treure un etiquetatge visible	
Precondició	S'està modificant etiquetatge	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol treure un etiquetatge visible.		
2. L'administrador selecciona un etiquetatge visible		
		2. El sistema treu l'etiquetatge visible.

Cas d'ús	Afegir element	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol afegir un element a l'etiquetatge	
Precondició	S'està modificant l'etiquetatge	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol afegir un element a l'etiquetatge.		
		2. El sistema valida quins tipus d'elements es poden afegir.
3. L'administrador selecciona quin tipus d'element vol afegir.		
		4. El sistema mostra els elements per afegir-los a l'etiquetatge.
5. L'administrador selecciona els elements que vol afegir.		
		6. El sistema valida que l'element es pugui afegir.
		7. El sistema afegeix l'element.
CURS ALTERNATIU		
Cas 6: Si l'element ja existeix no es fa res i es torna al cas 4.		

Cas d'ús	Treure element	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol treure un element de l'etiquetatge	
Precondició	S'està modificant l'etiquetatge	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol treure un element de l'etiquetatge.		
2. L'administrador selecciona un element.		
		3. El sistema comprova si es pot treure l'element.
		4. El sistema treu l'element.
CURS ALTERNATIU		
Cas 3: Si l'element ja està anotat en algun fitxer del projecte, es mostra un missatge d'error i se surt.		

Cas d'ús	Alta atribut	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol donar d'alta un atribut	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol donar d'alta un atribut.		
		2. El sistema demana les dades de l'atribut.
3. L'administrador introdueix les dades de l'atribut.		
		4. El sistema valida les dades de l'atribut.
		5. El sistema guarda l'atribut.
CURS ALTERNATIU		
Cas 4: Si les dades són incorrectes es mostra un missatge d'error i es torna al cas 2.		

Cas d'ús	Modificar atribut	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol modificar un atribut	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol modificar un atribut.		
2. L'administrador selecciona un atribut		
		3. El sistema comprova que es pot modificar l'atribut.
		4. El sistema mostra les dades de l'atribut.
5. L'administrador modifica les dades de l'atribut.		
		6. El sistema valida les dades de l'atribut.
		7. El sistema guarda l'atribut.
CURS ALTERNATIU		
Cas 3: Si l'atribut està en un etiquetatge es mostra missatge d'error i se surt.		
Cas 6: Si les dades són incorrectes es mostra missatge d'error i torna al cas 3.		

Cas d'ús	Esborrar atribut	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol esborrar un atribut	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol esborrar un atribut.		
2. L'administrador selecciona atribut		
		3. El sistema valida si es pot esborrar l'atribut.
		4. El sistema esborra l'atribut.
CURS ALTERNATIU		
Cas 3: Si l'atribut està en un etiquetatge es mostra un missatge d'error i se surt.		

Cas d'ús	Alta relació	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol donar d'alta una relació	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol donar d'alta una relació.		
		2. El sistema demana les dades de la relació.
3. L'administrador introdueix les dades de la relació.		
		4. El sistema valida les dades de la relació.
		5. El sistema guarda la relació.
CURS ALTERNATIU		
Cas 4: Si les dades són incorrectes es mostra un missatge d'error i es torna al cas 2.		

Cas d'ús	Modificar relació	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol modificar una relació	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol modificar una relació.		
2. L'administrador selecciona una relació.		
		3. El sistema valida si es pot modificar la relació.
		4. El sistema mostra les dades de la relació.
5. L'administrador modifica les dades de la relació.		
		6. El sistema valida les dades de la relació.
		7. El sistema guarda les dades de la relació.
CURS ALTERNATIU		
Cas 3: Si la relació està en un etiquetatge, dona un missatge d'error i se surt.		
Cas 4: Si alguna dada de la relació és incorrecta, dona un missatge d'error i es torna al cas 4.		

Cas d'ús	Esborrar relació	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol esborrar una relació	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol esborrar una relació.		
2. L'administrador selecciona una relació		
		3. El sistema valida si es pot esborrar la relació.
		4. El sistema esborra la relació.
CURS ALTERNATIU		
Cas 3: Si la relació està en un etiquetatge, dóna un missatge d'error i se surt.		

Cas d'ús	Alta grup	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol donar d'alta un grup	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol donar d'alta un grup.		
		2. El sistema demana les dades del grup.
3. L'administrador introdueix les dades del grup.		
		4. El sistema valida les dades del grup.
		5. El sistema guarda les dades del grup.
CURS ALTERNATIU		
Cas 4: Si les dades del grup són incorrectes, dóna un missatge d'error i es torna al cas 2.		

Cas d'ús	Modificar grup	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol modificar un grup	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol modificar un grup.		
2. L'administrador selecciona un grup.		
		3. El sistema valida si es pot modificar el grup.
		4. El sistema mostra les dades a modificar.
5. L'administrador modifica les dades.		
		6. El sistema valida les dades.
		7. El sistema guarda les dades.
CURS ALTERNATIU		
Cas 3: Si el grup està en un etiquetatge, dona un missatge d'error i se surt.		
Cas 6: Si les dades són incorrectes, dona un missatge d'error i es torna al cas 4.		

Cas d'ús	Esborrar grup	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol esborrar un grup	
Precondició	Gestió de marcatges oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol esborrar un grup.		
2. L'administrador selecciona un grup		
		3. El sistema valida si es pot esborrar el grup.
		4. El sistema esborra el grup.
CURS ALTERNATIU		
Cas 3: Si el grup està en un etiquetatge, dona un missatge d'error i se surt.		

CASOS D'ÚS DE GESTIÓ D'USUARIS

Cas d'ús	Afegir administrador	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol afegir un usuari com a administrador	
Precondició	Gestió d'usuaris	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol afegir un usuari com a administrador.		
		2. El sistema demana les dades de l'usuari administrador.
3. L'administrador introdueix les dades.		
		4. El sistema valida les dades.
		5. El sistema afegeix l'usuari com a administrador.
CURS ALTERNATIU		
Cas 4: Si les dades són incorrectes no es fa res i es torna al cas 2.		

Cas d'ús	Treure administrador	
Actors	Administrador	
Iniciador	Administrador	
Descripció	L'administrador vol treure un usuari administrador	
Precondició	Gestió d'usuaris oberta	
CURS NORMAL		
1. El cas d'ús comença quan l'administrador vol treure un usuari administrador.		
2.L'administrador selecciona un usuari		
		3. El sistema comprova si es pot treure l'usuari.
		4. El sistema treu l'usuari.
CURS ALTERNATIU		
Cas 3: Si l'usuari és root, administrador o administrador no es fa res i se surt.		

CASOS D'ÚS DE PROCÉS D'ANOTACIÓ

Cas d'ús	Nova anotació	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol fer una anotació nova	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol fer una anotació nova.		
2. L'anotador selecciona text i crea anotació		
		3. El sistema comprova que es pugui fer l'anotació.
		4. El sistema crea una anotació i assigna els valors dels atributs per defecte.
CURS ALTERNATIU		
Cas 3: Si l'etiquetatge actiu no és d'atributs, no es fa res i se surt.		

Cas d'ús	Canviar valors	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol canviar els valors dels atributs de l'anotació	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol canviar els valors dels atributs de l'anotació.		
2. L'anotador selecciona una anotació.		
		3. El sistema mostra els atributs.
4. L'anotador selecciona un atribut.		
		5. El sistema assigna el valor de l'atribut seleccionat a l'atribut de l'anotació.
CURS ALTERNATIU		
Cas 4: Si el tipus de l'atribut és "enter", "cadena" o "rang", es demana el valor i es valida. Si el valor és correcte s'assigna, si no, es mostra missatge d'error i es demana valor.		

Cas d'ús	Afegir relació	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol relacionar dues anotacions	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol relacionar dues anotacions.		
2. L'anotador selecciona una anotació.		
		3. El sistema mostra les relacions en les quals el seu origen és igual a l'etiquetatge de l'anotació.
4. L'anotador selecciona una relació.		
5. L'anotador selecciona l'anotació destí.		
		6. El sistema valida la relació.
		7. El sistema afegeix la relació entre les anotacions.
CURS ALTERNATIU		
Cas 6: Si la validació és incorrecta, dona un missatge d'error i se surt.		

Cas d'ús	Afegir grup	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol afegir anotacions al grup	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol afegir anotacions al grup.		
2. L'anotador seleccionar una anotació		
		3. El sistema mostra els grups de l'etiquetatge actiu.
4. L'anotador selecciona un grup.		
5. L'anotador selecciona una anotació per afegir-la al grup.		
		6. El sistema afegeix l'anotació al grup.
7. L'anotador decideix si vol continuar afegint anotacions al grup. Si decideix que sí, es torna al cas 5. Si no, se surt.		

Cas d'ús	Veure informació	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol veure la informació de l'anotació i de les seves relacions	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol veure la informació.		
2. L'anotador selecciona una anotació.		
		3. El sistema mostra els atributs i els seus valors i les relacions de l'anotació.

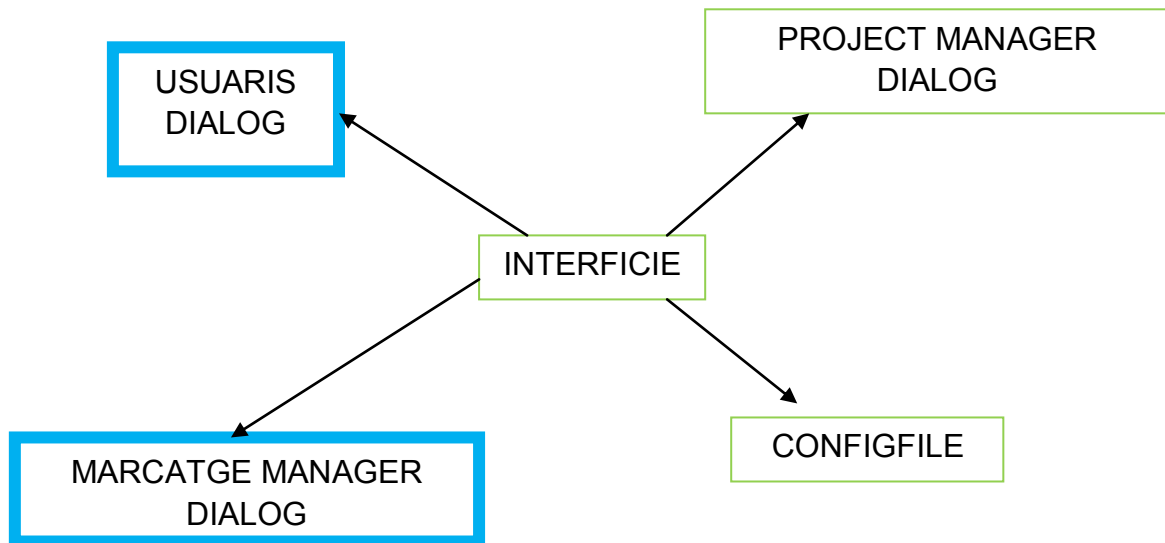
Cas d'ús	Seleccionar etiquetatge per anotar	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol seleccionar un etiquetatge per fer anotacions	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol anotar amb un etiquetatge.		
2. L'anotador selecciona un etiquetatge.		
		3. El sistema mostra els atributs i els seus valors i les relacions de l'anotació.

Cas d'ús	Seleccionar fitxer per anotar	
Actors	Anotador	
Iniciador	Anotador	
Descripció	L'anotador vol seleccionar un fitxer per fer anotacions	
Precondició	Hi ha projecte activat	
CURS NORMAL		
1. El cas d'ús comença quan l'anotador vol anotar en fitxer.		
2. L'anotador selecciona un fitxer.		
		3. El sistema mostra el text del fitxer seleccionat per anotar.

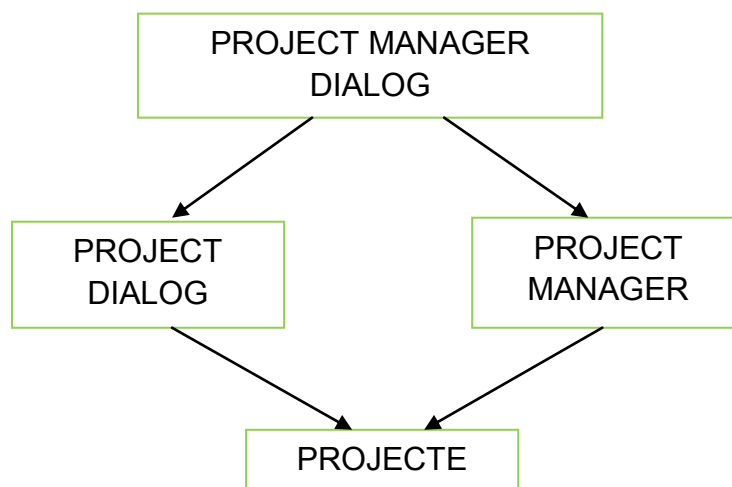
9.- DIAGRAMA DE CLASSES

En aquest punt s'indiquen les classes a modificar (de color verd) i les noves (de color blau) estructurat com sempre per menú, gestió de projectes, gestió de marcatges, gestió d'usuaris i procés d'anotació.

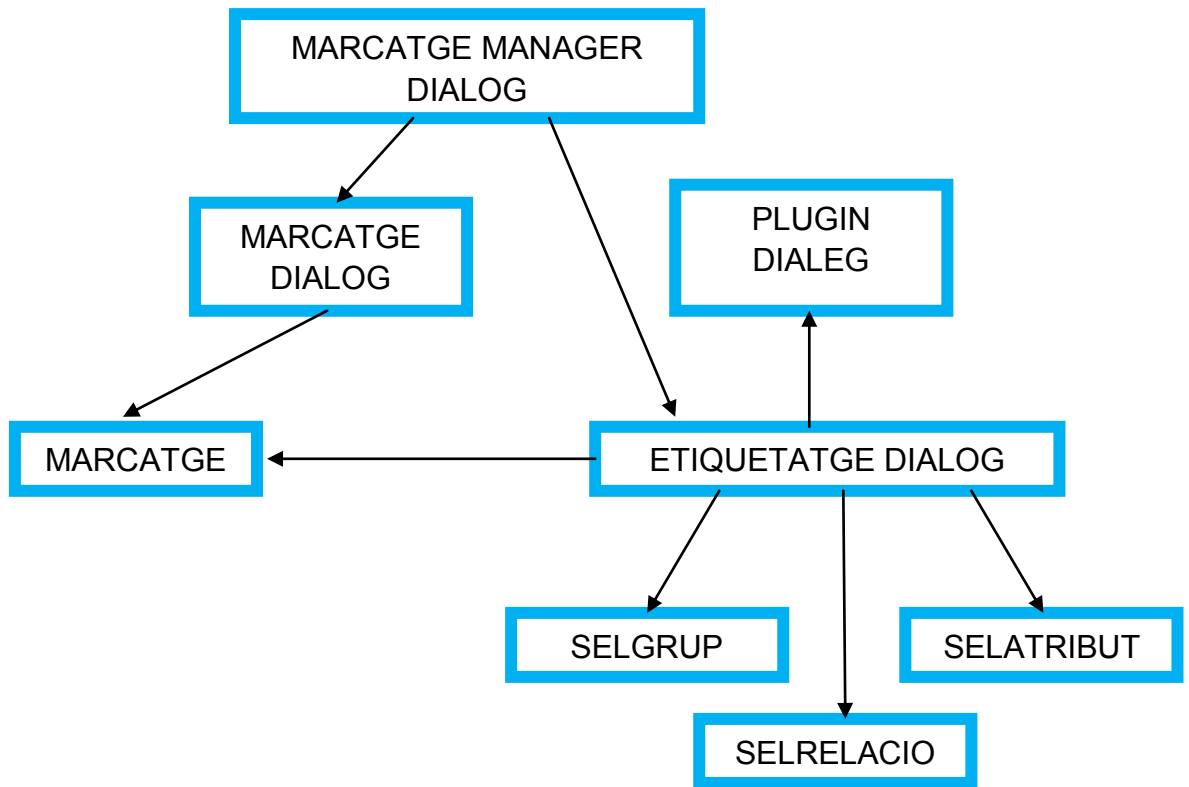
9.1 Menú



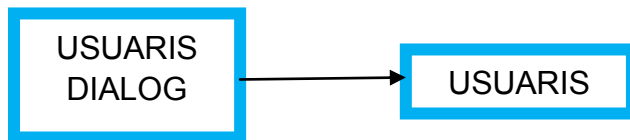
9.2 Gestió de projectes



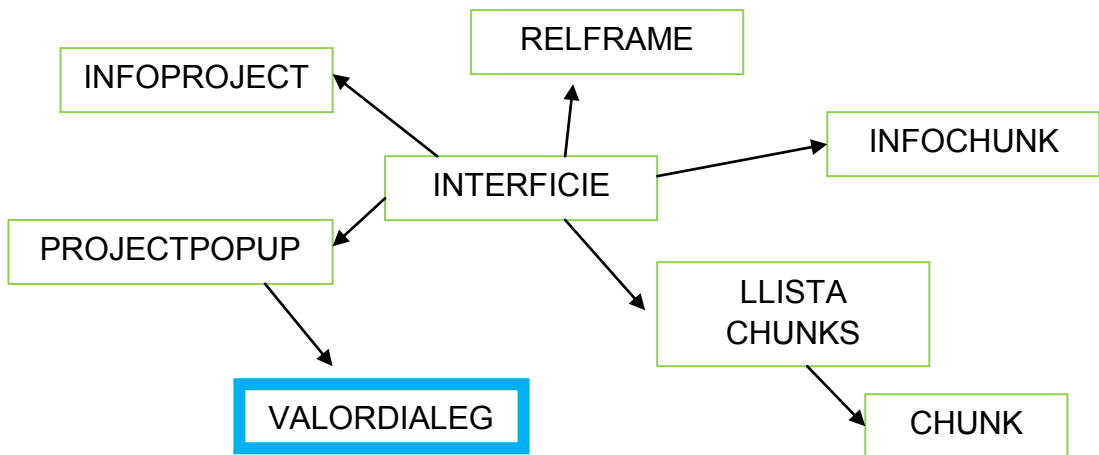
9.3 Gestió de marcatges



9.4 Gestió d'usuaris



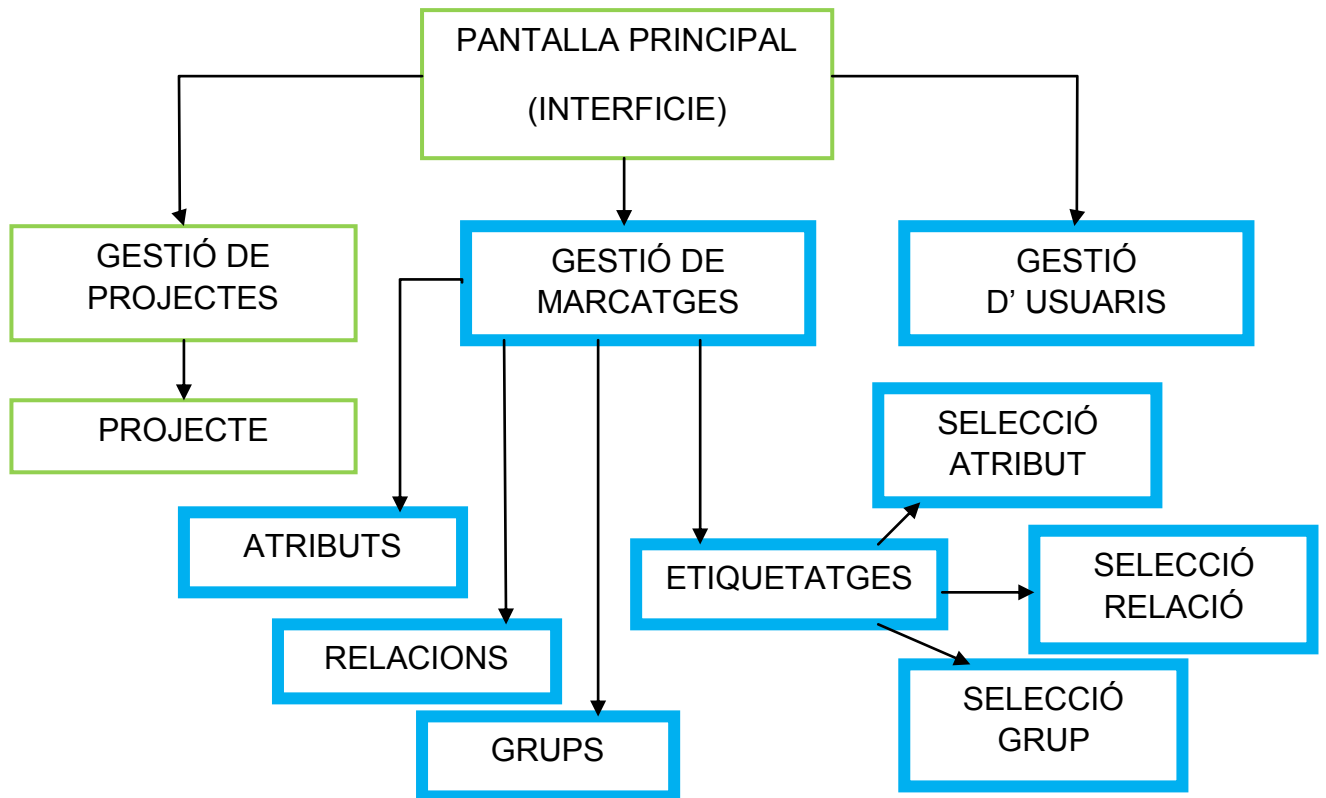
9.5 Procés d'anotació



10.- DISSENY DE PANTALLES

Pel disseny s'ha tingut en compte que estàvem modifiquen una aplicació ja existent, així que s'ha seguit mantenint la mateixa línia.

S'indica quines pantalles són les que es modificaran i quines són les que s'afegiran.



Amb el color **verd** s'indiquen les pantalles que es modifiquen i amb el **blau** les pantalles noves.

11.- MANUAL D'INSTAL·LACIÓ

11.1 Requeriments

Pel funcionament del PTKChA es necessita tenir instal·lat els següents components. Aquestes són les versions que hem fet servir, però és possible que altres versions puguin funcionar de la mateixa manera.

Les pàgines web des d' on es poden descarregar els components són:

- Perl-5.10: <http://www.perl.com>
- Expat-2.0.1-4: <http://expat.sourceforge.net>
- Tk-804.028 Mòdul de Perl: <http://search.cpan.org>
- XML-Parser-2.34 Perl Module: <http://search.cpan.org>
- Tk-DirSelect-1.11 Mòdul de Perl: <http://search.cpan.org>

De totes maneres, existeix la possibilitat de trobar els components en paquets.

Aquests són:

- perl-tk 1:804.028-5
- libtk-dirselect-perl 1.11
- libxml-parser-perl 2.36

11.2 Instal·lació

Només cal tenir els fitxers de codi font en alguna carpeta del sistema i que el fitxer de configuració .ptkcha.xml tingui permisos d'escriptura.

11.3 Executar el programa

Per iniciar el programa executar :

- perl ptkcha.pl

Si hi hagués algun problema amb la ruta d'instal·lació, es pot afegir al fitxer ptkcha.pl:

use lib "ruta d'instalació del ptkcha";

12-. MANUAL D'USUARI : ANOTADOR

12.1 MENÚ

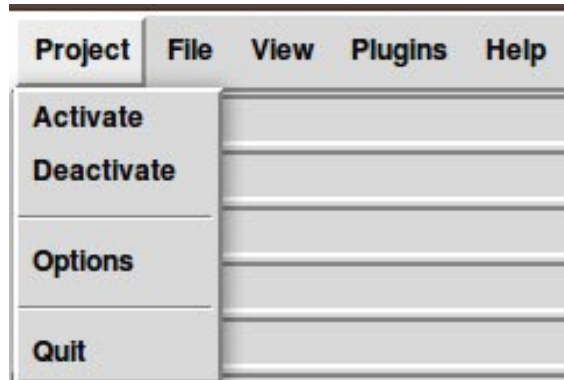


Figura 13.1: menú usuari anotador

12.1.1 Project

Activate. Obre la pantalla de gestió de projectes per activar un projecte. (veure activar projecte 13.2).

Deactivate. Desactiva el projecte actiu.

Options. Permet canviar els valors d'opcions globals :

-*Forced Writing.* Si es marca aquesta opció es força a guardar les anotacions en el fitxer de sortida quan hi ha alguna modificació en les anotacions.

-*Selection Expansion.* Expandeix la selecció de text quan es crea una anotació nova.

-*Include directories.* Directoris addicionals on es busquen els fitxers dels plugins.

Quit. Sortir de l'aplicació.

12.1.2 File

Save. Per guardar el text que actualment s'està anotant.

Revert. Desfà els últims canvis fets.

Initial versio. Torna a la versió inicial del fitxer. Copia el fitxer del directori d'entrada al directori de sortida.

12.1.3 View

En l'àrea d'anotació les anotacions es veuen per colors. Cada una d'elles en té un. Amb les opcions d'aquest menú es pot canviar la visualització de les anotacions. Selecciónant un atribut, relació o grup, depenent de l'etiquetatge actiu, les anotacions es mostren per colors segons els valors de l'atribut, relació o grup seleccionat.

12.1.4 Plugins

Inicialment aquest menú està buit. Quan es selecciona un etiquetatge després d'activar un projecte, es omple amb els plugins definits en l'etiquetatge.

12.2 ACTIVAR PROJECTE

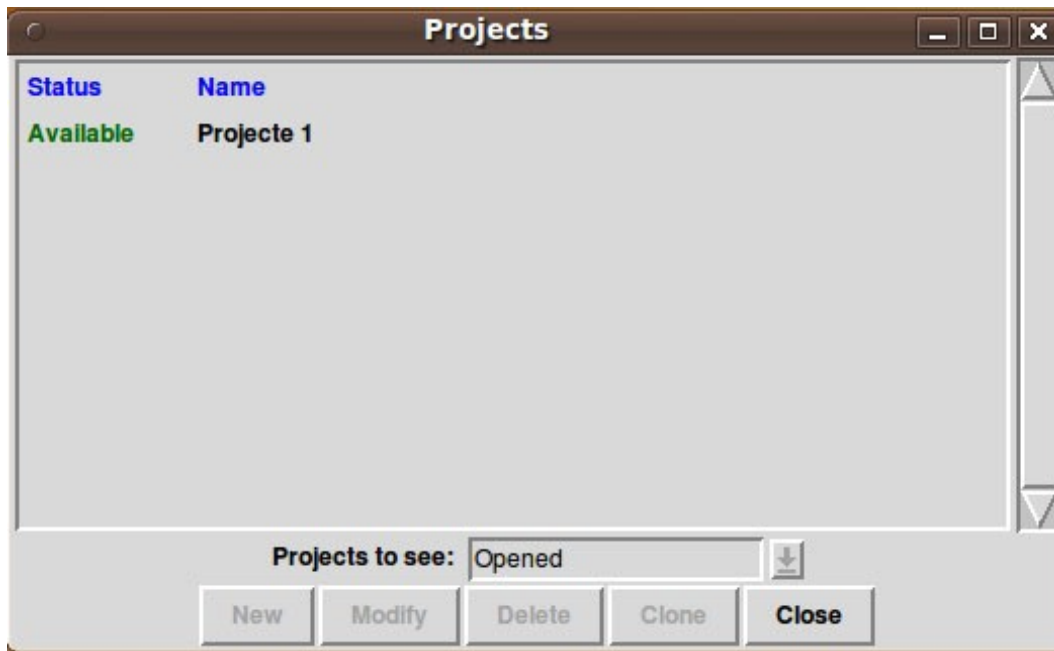


Figura 13.2: pantalla per activar projecte com a usuari anotador

La pantalla d'activar projectes mostra els projectes en estat obert i disponibles per ser activats per treballar. Fent doble clic sobre el projecte s'activa i s'actualitza la pantalla principal amb les dades del projecte per començar a anotar.

12.3 PANTALLA PRINCIPAL

La pantalla principal es divideix en quatre regions.

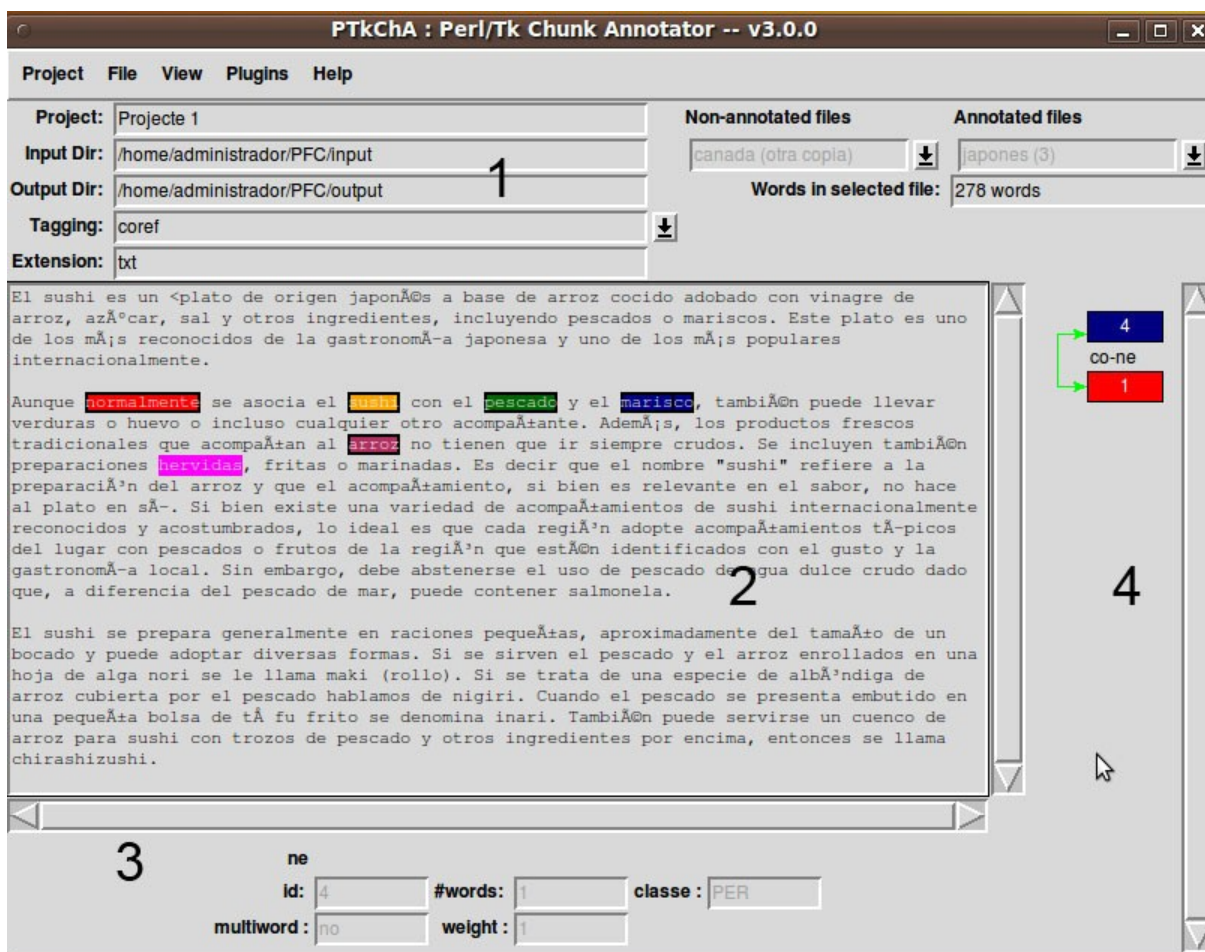


Figura 13.3: pantalla principal

1. Informació del projecte.

Aquí es veu la informació del projecte actual. Des d'aquesta regió es pot seleccionar un fitxer per anotar d'una de les dues llistes de fitxers del projecte, i seleccionar l'etiquetatge actiu per anotacions.

2. Area d'anotació.

En aquesta regió de la pantalla principal es mostra el text del fitxer actualment seleccionat, el qual es vol anotar, i es permet fer les diferents operacions d'anotació.

- a. Crear noves anotacions, seleccionant el text desitjat i o bé pressionant ENTER o fent clic amb el botó dret del ratolí. L'anotació es crea si l'etiquetatge actiu es d'atributs.

- b. Veure informació de l'anotació i de les seves relacions, clicant amb el botó dret sobre l'anotació. La informació surt a les regions 3 i 4.
 - c. Clicant el botó dret sobre l'anotació, apareix un menú que permet :
 - i. Canviar els valors de l'anotació sempre i quan l'etiquetatge de l'anotació sigui igual a la de l'etiquetatge actiu i d'atributs.
 - ii. Afegir relacions entre anotacions, sempre i quan l'etiquetatge actiu sigui de relacions. Es mostren les relacions que tenen el mateix origen que l'etiquetatge de l'anotació. Al afegir la relació es valida que l'origen i el destí d'aquesta es correspongui amb els etiquetatges de les anotacions origen i destí.
 - iii. Afegir grups, sempre i quan l'etiquetatge actiu sigui de grups. Seleccionar grup i anar clicant sobre les anotacions per afegir-les al grup.
 - iv. Esborrar l'anotació, sempre i quan l'etiquetatge de l'anotació sigui igual a etiquetatge l'actiu.
 - v. Treure grup de l'anotació, sempre i quan l'etiquetatge actiu sigui de grup.
- 3. Informació de l'anotació.** Aquesta regió mostra els atributs de l'anotació, si pertany a un grup i el nombre de paraules que té.
- 4. Informació de la relació.** Aquesta regió mostra la informació de les relacions de l'anotació actual. Clicant amb el botó dret del ratolí sobre el destí de la relació, aquesta es pot esborrar.

13-. MANUAL D'USUARI : ADMINISTRADOR

13.1 MENÚ

Per les opcions *activate*, *deactivate* i *options* del menú *Project*, per el menú *File*, *View* i *Plugins* veure manual d'usuari anotador.

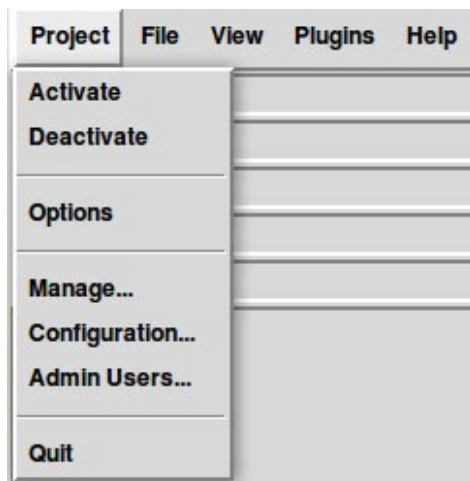


Figura 14.1: menú usuari administrador

13.1.1 Project

Manage... Obre la pantalla de gestió de projectes.

Configuration... Obre la pantalla de gestió de marcatges.

Admin Users ... Obre la pantalla de gestió d'usuaris administradors.

13.2 GESTIÓ DE PROJECTES

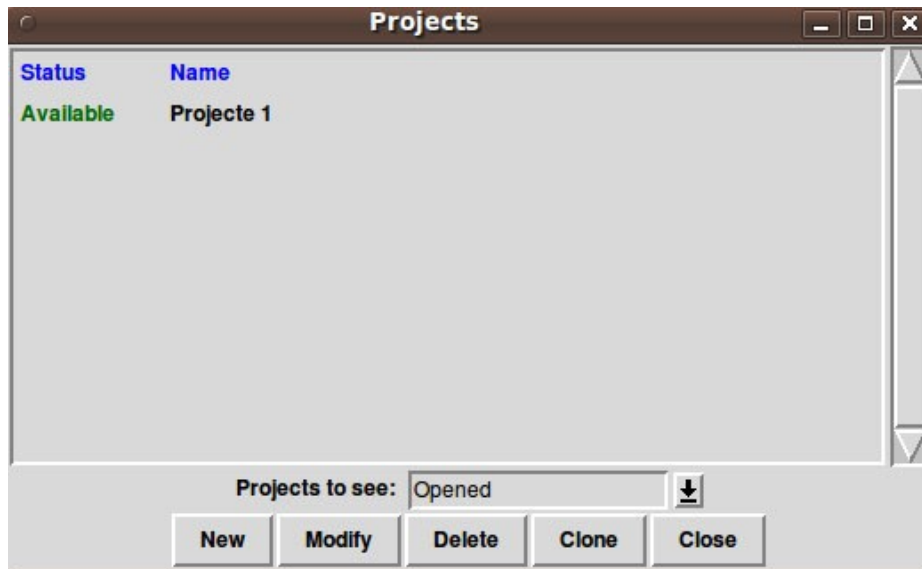


Figura 14.2: pantalla gestió de projectes

En aquesta pantalla es poden gestionar els projectes: crear projectes nous, modificar-los, esborrar-los i clonar-los.

Es poden veure els projectes per estat. Seleccionant l'estat es mostren els projectes oberts, tancats o tots.

Activar. Per activar un projecte s'ha de fer doble clic sobre ell. Al activar-lo es valida que els etiquetatges, els atributs, relacions i grups de tots els fitxers del projecte que estan anotats estiguin definits en el marcatge.

New. Crear un projecte nou. La pantalla de projecte es obre per demanar les dades del projecte. (veure 14.2.1 Projecte).

Modify. Modificar el projecte seleccionat. La pantalla de projecte es obre per modificar les dades del projecte. (veure 14.2.1 Projecte).

Delete. S'esborra el projecte seleccionat. El projecte actiu no es pot esborrar. Si el projecte té algun fitxer anotat no es pot esborrar.

Clone. Clonar el projecte seleccionat. Es crea un projecte nou amb les mateixes dades que el projecte actiu.

13.2.1 Projecte

Project info

Name:

Input Directory: ...

Output Directory: ...

File Extension:

State: ↓

Project taggings		All taggings	
Name	Description	Name	Description
ne	Named Entity	city	Ciutats
neLocation	Named Entity Location	coref	Correferencia
coref	Correferencia	ne	Named Entity
city	Ciutats	neLocation	Named Entity Location
		summary	Summary

Figura 14.3: pantalla per introduir o modificar les dades del projecte

En aquesta pantalla es poden introduir les dades del projecte o modificar-les, el nom, el directori d'entrada, el directori de sortida, l'extensió dels fitxers d'entrada i l'estat, a més de poder afegir quins etiquetatges podran ser utilitzats en el procés d'anotació.

Save. Guarda les dades del projecte.

Cancel. Surt de la pantalla sense guardar les dades del projecte.

13.3 GESTIÓ DE MARCATGES

13.3.1 Gestió d'etiquetatges

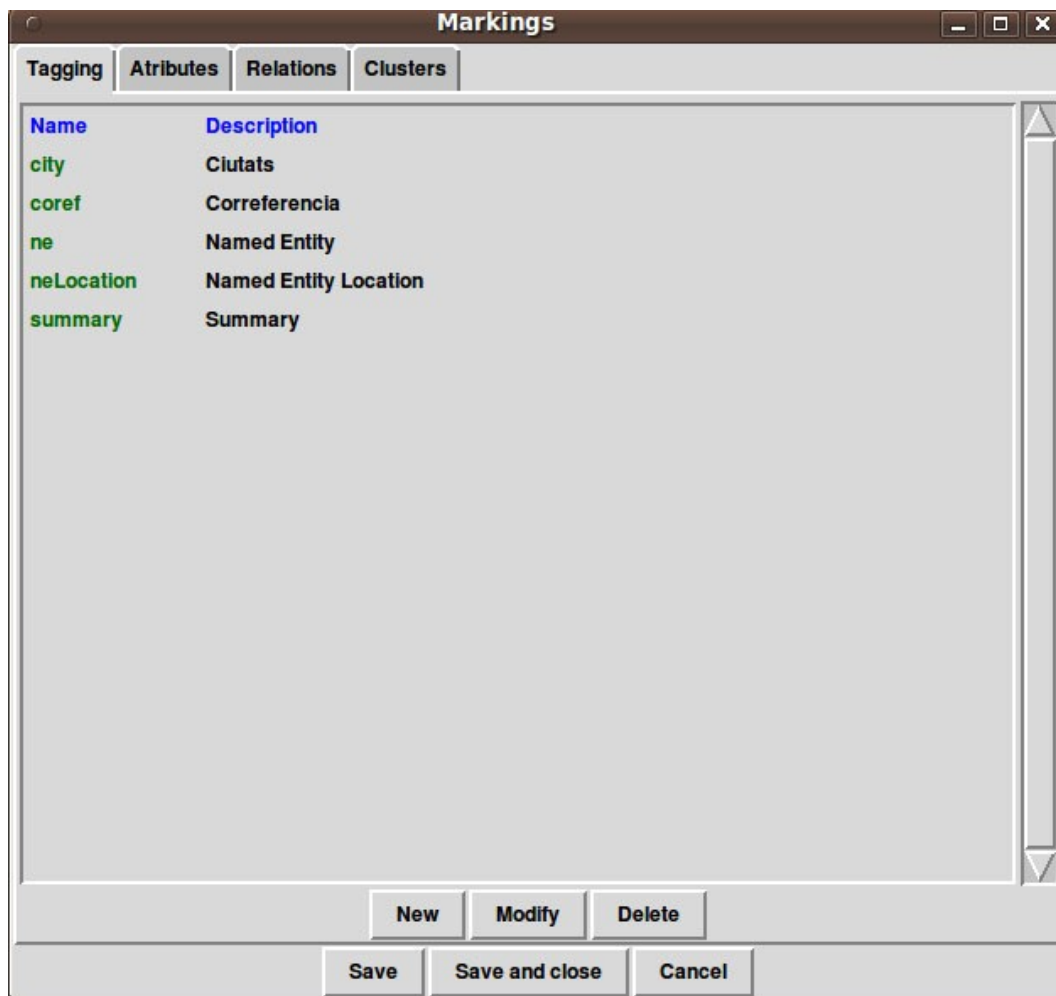


Figura 14.4: pantalla per gestionar els etiquetatges

En aquesta pantalla es poden veure els etiquetatges actuals, donar-ne d'alta, modificar-los o esborrar-los.

New. Crear un nou etiquetatge. La pantalla d'etiquetatge s'obre per demanar les dades de l'etiquetatge.

Modify. Modificar l'etiquetatge seleccionat. La pantalla d'etiquetatge s'obre per modificar les dades de l'etiquetatge. Si l'etiquetatge està en un projecte no es deixa modificar.

Delete. S'esborra l'etiquetatge seleccionat. Si l'etiquetatge està en un projecte no es deixa esborrar.

13.3.1.1 Etiquetatge

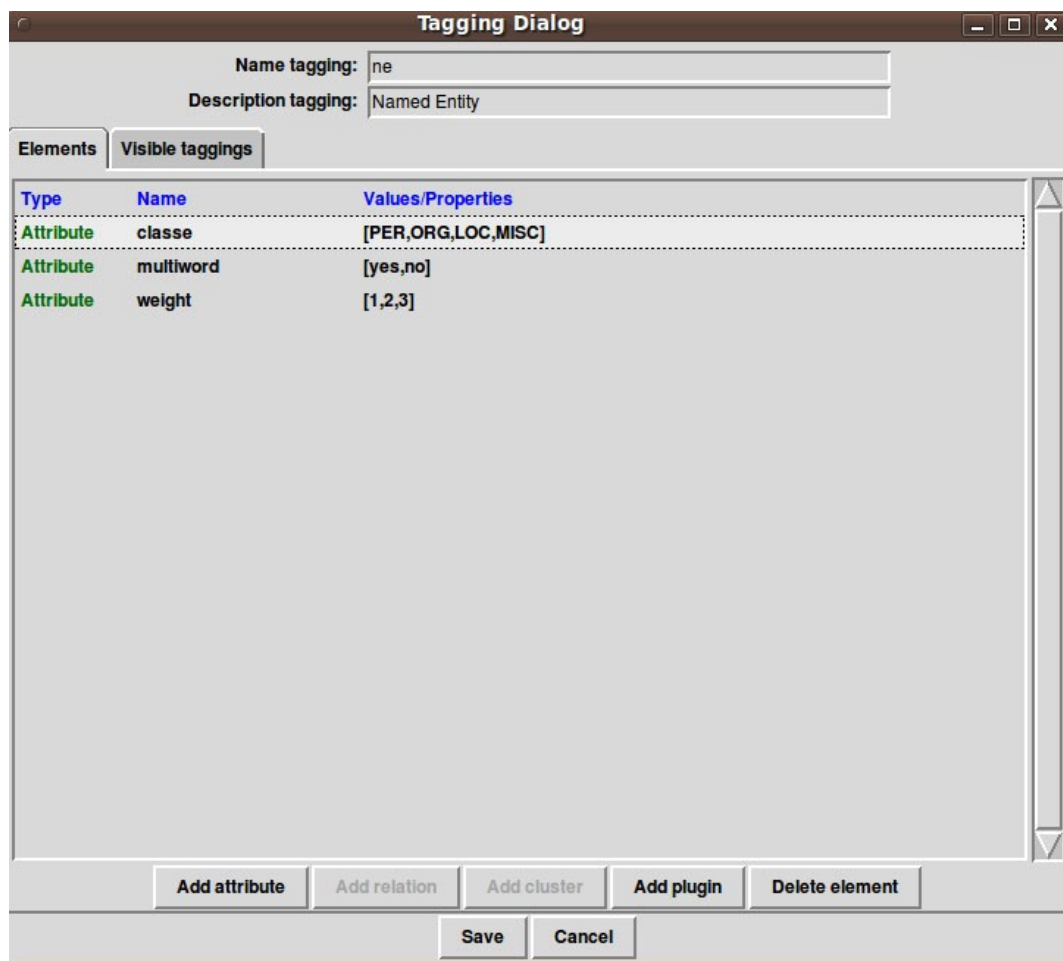


Figura 14.5: pantalla per introduir o modificar les dades de l'etiquetatge

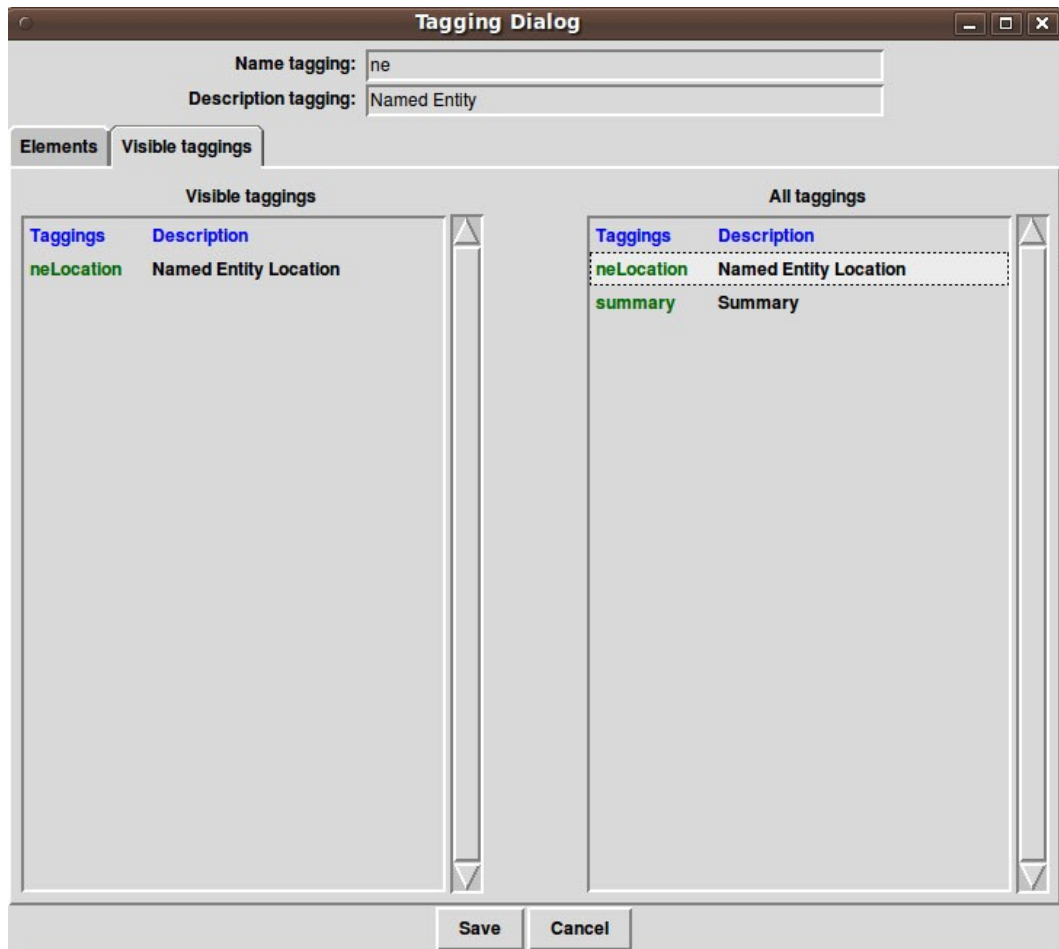


Figura 14.6: pantalla per afegir etiquetatges visibles

Des d'aquesta pantalla es poden introduir o modificar les dades de l'etiquetatge, afegir elements i afegir etiquetatges visibles.

Un etiquetatge només pot tenir elements d'un tipus. Un etiquetatge pot tenir o bé atributs o bé relacions o bé grups. Els plugins poden estar amb qualsevol altre element.

Segons els elements que l'etiquetatge ja tingui, els botons d'afegir atribut, relació i grup s'activen o desactiven.

Add attribute. Afegir atribut. S'obre la pantalla de selecció d'atributs per afegir-ne a l'etiquetatge.

Add relation. Afegir relació. S'obre la pantalla de selecció de relacions per afegir-ne a l'etiquetatge.

Add cluster. Afegir grup. S'obre la pantalla de selecció de grups per afegir-ne a l'etiquetatge.

Add plugin. Afegir plugin. S'obre la pantalla de plugin per afegir-lo a l'etiquetatge.

Delete element. Esborra l'element seleccionat. Si l'element està anotat en algun fitxer del projecte no es deixa esborrar.

Save. Guardar l'etiquetatge.

Cancel. Per sortir de la pantalla d'etiquetatge.

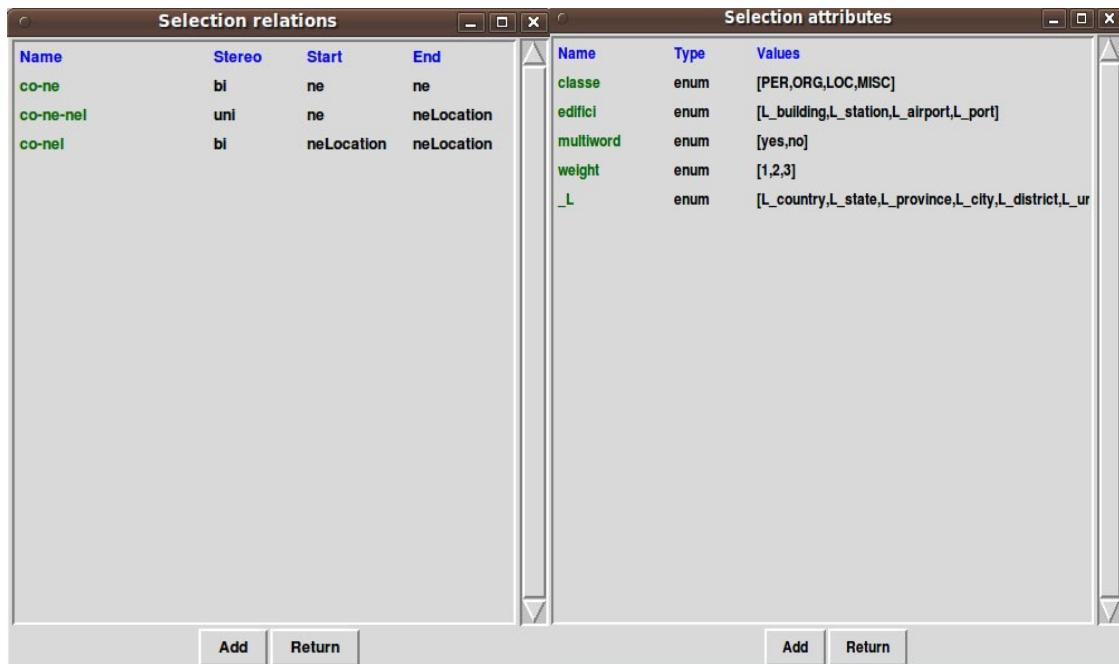


Figura 14.7: pantalles de selecció de relacions i atributs

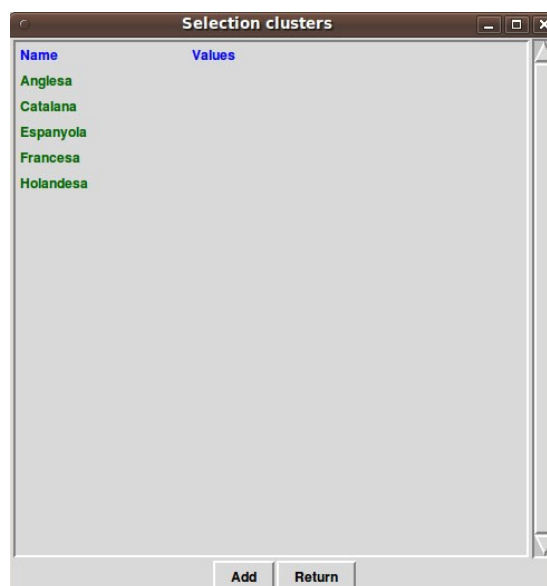


Figura 14.8: pantalla de selecció de grups

13.3.2 Gestió d'atributs

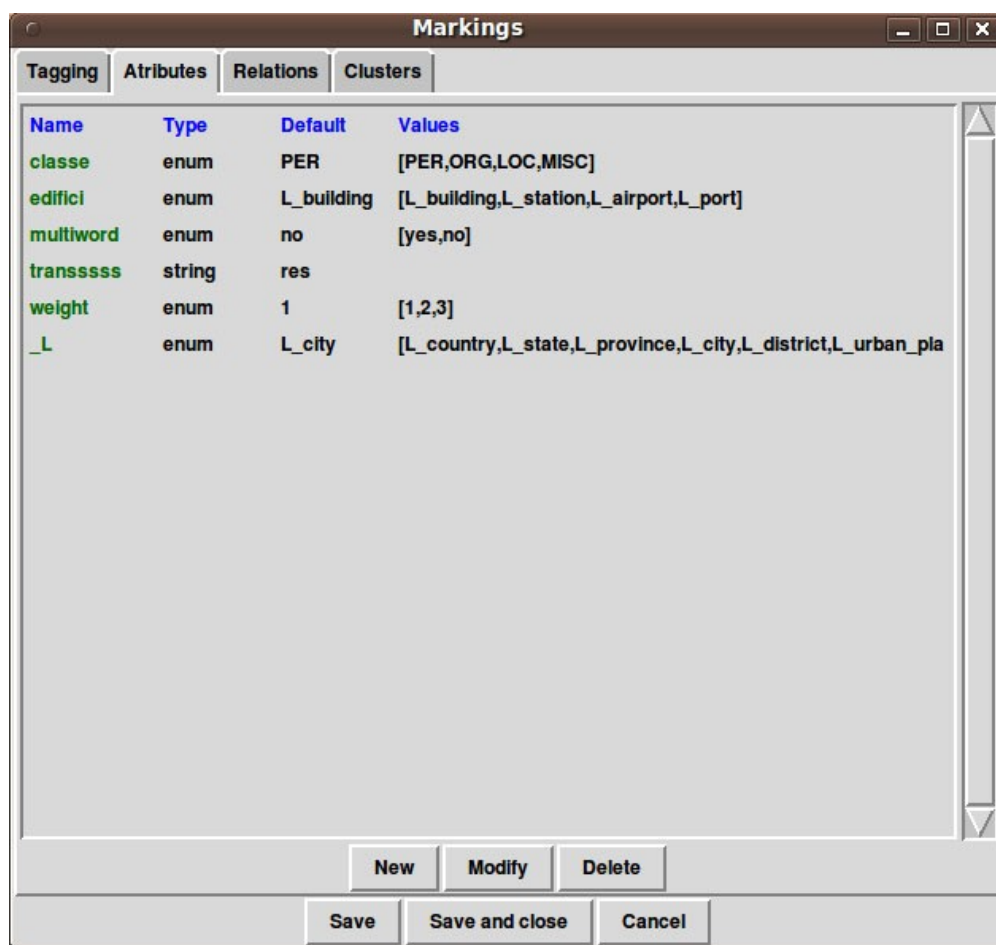


Figura 14.9: pantalla de gestió d'atributs

En aquesta pantalla es poden veure els atributs, donar-ne d'alta, modificar-los o esborrar-los.

New. Crear un nou atribut. La pantalla d'atribut s'obre per demanar les dades de l'atribut.

Modify. Modificar l'atribut seleccionat. La pantalla d'atribut s'obre per modificar les dades de l'atribut. Si l'atribut està en algun etiquetatge no es deixa modificar.

Delete. Esborrar l'atribut seleccionat. Si l'atribut està en algun etiquetatge no es deixa modificar.

13.3.2.1 Atribut

The dialog box is titled with fields for Name, Type, and Default. The Name field contains "_L", Type is set to "enum", and Default is "L_city". Below these are sections for "Rang type" with Min and Max input fields, and "Enum type" with a list of values: L_country, L_state, L_province, L_city, L_district, L_urban_place, L_road, and L_river. At the bottom, there are buttons for "Add enum", "Del enum", "Save", and "Cancel".

Figura 14.10: pantalla per introduir o modificar dades dels atributs

En aquesta pantalla es pot introduir o modificar les dades dels atributs.

13.3.3 Gestió de relacions

The "Markings" dialog box has tabs for Tagging, Atributes, Relations, and Clusters. The "Relations" tab is active, showing a table with columns Name, Stereo, Start, and End. The table contains three rows of relationship data.

Name	Stereo	Start	End
co-ne	bi	ne	ne
co-ne-ne1	uni	ne	neLocation
co-ne1	bi	neLocation	neLocation

At the bottom of the dialog are buttons for "New", "Modify", "Delete", "Save", "Save and close", and "Cancel".

Figura 14.11: pantalla de gestió de relacions

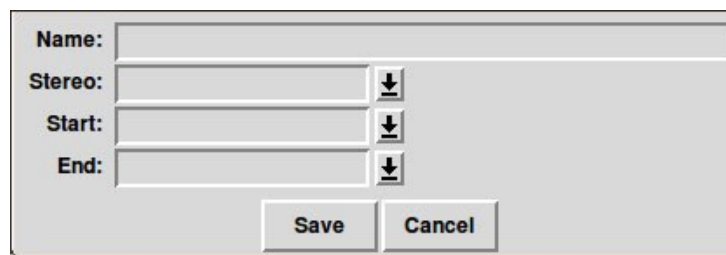
En aquesta pantalla es poden veure les relacions, donar-les d'alta, modificar-les o esborrar-les.

New. Crear una nova relació. La pantalla de relació s'obre per demanar les dades de la relació.

Modify. Modificar la relació seleccionada. La pantalla de relació s'obre per modificar les dades de la relació. Si la relació està en algun etiquetatge no es deixa modificar.

Delete. Esborrar la relació seleccionada. Si la relació està en algun etiquetatge no es deixa modificar.

13.3.3.1 Relació



The image shows a software interface for entering relationship data. It consists of a light gray rectangular window with a thin border. On the left side, there are four labels: 'Name:', 'Stereo:', 'Start:', and 'End:'. Each label is followed by a text input field. The 'Stereo:', 'Start:', and 'End:' fields have a small downward-pointing arrow icon to their right, indicating they are dropdown menus. At the bottom center of the window, there are two buttons: 'Save' and 'Cancel'.

Figura 14.12: pantalla de dades de les relacions

En aquesta pantalla es pot introduir o modificar les dades de les relacions.

13.3.4 Gestionar grups

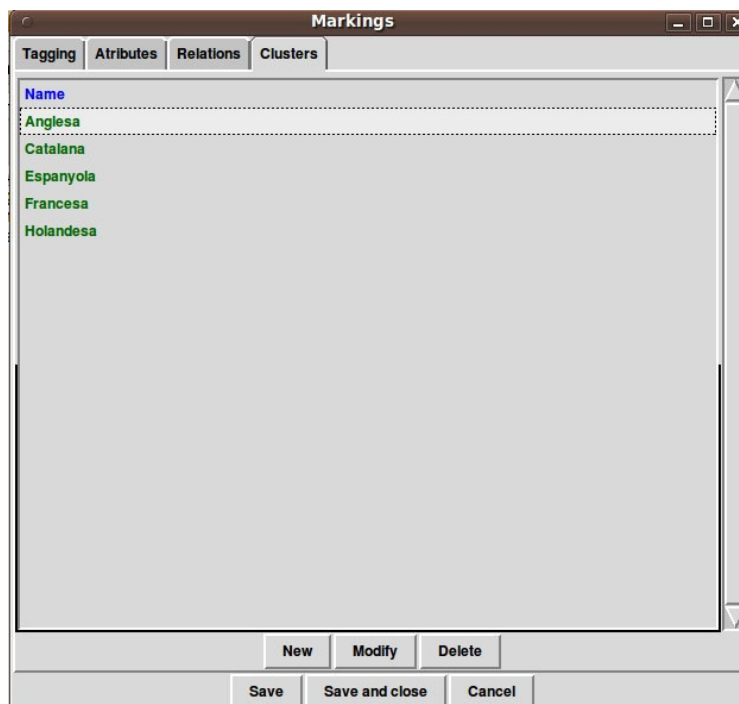


Figura 14.13: pantalla de gestió de grups

En aquesta pantalla es poden veure els grups, donar-los d'alta, modificar-los o esborrar-los.

New. Crear un nou grup. La pantalla de grup s'obre per demanar les dades del grup.

Modify. Modificar el grup seleccionat. La pantalla de grup s'obre per modificar les dades del grup. Si el grup està en algun etiquetatge no es deixa modificar.

Delete. Esborrar el grup seleccionat. Si el grup està en algun etiquetatge no es deixa modificar.

13.3.4.1 Grups

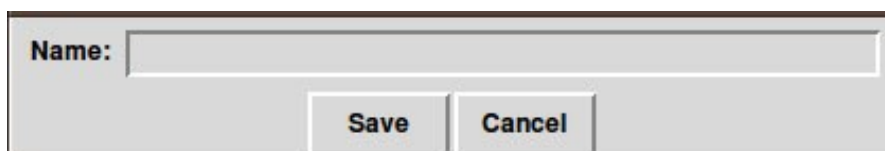


Figura 14.14: pantalla de dades dels grups

En aquesta pantalla es pot introduir o modificar el nom del grup.

13.4 GESTIÓ D'USUARIS

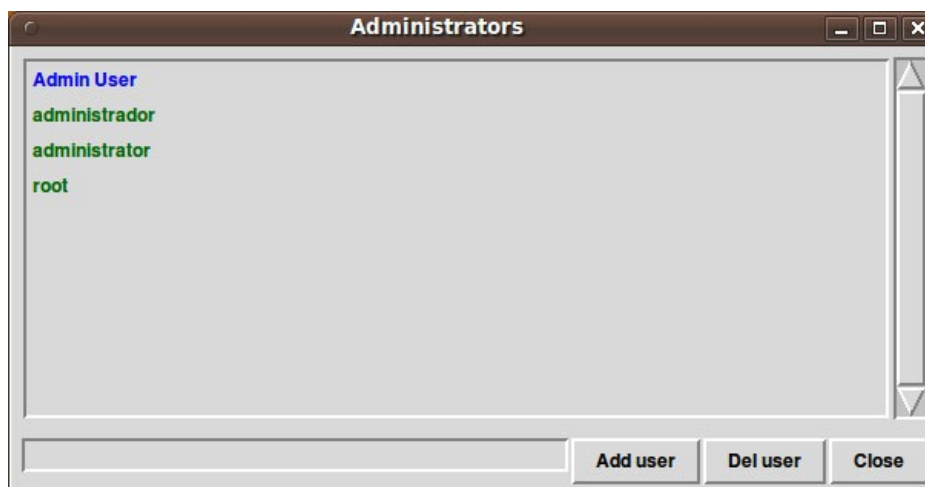


Figura 14.15: pantalla de gestió d'usuaris

En aquesta pantalla es poden afegir usuaris administradors i treure'ls. Els usuaris administradors definits per defecte no es poden esborrar.

Add user. Afegir usuari com administrador.

Del user. Esborrar l'usuari administrador seleccionat.

13.5 PROCÉS D'ANOTACIÓ¹⁰

¹⁰ Veure punt 13.3 del manual d'usuari anotador.

14.- EXEMPLE D'ÚS

Com a tall d'exemple, en un text el que es vol és etiquetar ciutats per visitar i saber a quin país i continent pertanyen. Una definició de marcatge en aquest exemple podria ser aquesta:

```
<atribut nom='lloc' tipus='enum' defecte='Ciutat' />
  <enum>
    <valor v='Ciutat capital' />
    <valor v='Ciutat' />
    <valor v='Poble' />
  </enum>
</atribut>
<atribut nom='pais' tipus='enum' defecte='Irlanda' />
  <enum>
    <valor v='França' />
    <valor v='Irlanda' />
    <valor v='Escòcia' />
    <valor v='Suïssa' />
    <valor v='Xina' />
  </enum>
</atribut>
<grup nom='Àsia' />
<grup nom='Europa' />
<etiquetatge nom="continent" des="Continent">
  <usaGrup nom="Àsia" />
  <usaGrup nom="Europa" />
</etiquetatge>
<etiquetatge nom="llocs" des="Llocs">
  <usaAtribut nom="lloc" />
  <usaAtribut nom="pais" />
</etiquetatges>
```

15.- COST REAL

En la fase d'implementació, després de desenvolupar la gestió de marcatges, es va detectar que la previsió havia estat incorrecta, ja que s'utilitzava un llenguatge nou per mi, el *Perl*, que m'estava costant més del compte.

Per tant, s'han hagut de fer hores extres, treballant caps de setmana inclosos, per poder fer les modificacions correctament i arribar al dia de l'entrega amb tot el que s'havia especificat.

Amb tot això, s'han hagut d'afegir 112 hores extres de programador quedant un cost real de:

Concepte	Hores	Total
Analista	104	5200€
Programador	320	11200€
Programador (hores extres)	112	7840€*
Dissenyador	8	240€
Total		24480€

Figura 16.1: cost real de tot el projecte

*Partint de la base que les hores extres, el programador les cobra a 70€.

El cost de la primera planificació més el de la segona, era de 16640€ (cost previst). Ara bé, com es pot apreciar en el quadre del cost real, hi ha hagut una variació de 7840€, que són les hores extres del programador.

16.- CONCLUSIÓ

Voldria remarcar, al finalitzar aquest treball, que em feia molta por endinsar-me en un projecte d'aquesta dimensió pel fet que feia molt temps que havia acabat la carrera i m'havia quedat pendent aquest projecte. De totes maneres, gràcies al Lluís i al Jordi, que em van animar i m'han animat durant tota la realització d'aquest, l'he pogut finalitzar.

Sí que m'agradaria mencionar alguns aspectes que m'han suposat difícils a l'hora de treballar. Possiblement l'etapa més difícil ha estat la de l'especificació, encara que no val a oblidar la de la implementació. En la primera, perquè saber exactament el que havia de fer m'ha sigut difícil, i en la segona, perquè el temps per a la realització, no s'ha ajustat amb el temps previst.

Tot i així i a la vegada estava i he estat molt il·lusionat, sobretot perquè la idea de fer alguna cosa nova i diferent, i programar amb *Perl* (que no ho havia fet mai) m'atreia.

17.- REFERÈNCIES BIBLIOGRÀFIQUES

17.1 Pàgines web consultades:

- <http://es.wikipedia.org>
- <http://search.cpan.org>
- <http://www.perl.com>
- <http://www.faqs.org/faqs/perl-faq/>

17.2 Llibres consultats:

- Barkakati, Naba: *Manual fundamental de Perl 5*
- Wyke, R. Allen: *Guía de referencia para programador*

ANNEX 2. LLISTA FINAL DE REQUISITS

MENÚ.

1. Accedir a Gestió de projectes com usuari administrador.
2. Accedir a Gestió de marcatges com usuari administrador.
3. Accedir a Gestió d'usuaris com usuari administrador.
4. Activar projecte com usuari anotador.
5. Desactivar projecte com usuari anotador.
6. Vista atributs modificar segons etiquetatge actiu.
7. Accedir a plugins segons marcatges.

GESTIÓ DE PROJECTES.

8. Afegir estat al projecte
9. Poder veure la llista de projectes disponibles segons el seu estat.
10. Poder afegir i treure etiquetatges per anotar al projecte.
11. Poder modificar l'estat d'un projecte.
12. Poder activar projectes per anotar.
13. Poder desactivar el projecte actiu.
14. No esborrar projecte si està anotat.
15. A l'activar un projecte comprovar que els etiquetatges i els elements dels marcatges de tots els seus fitxers estan definits.
16. Es decideix també treure els filtres del projecte i prendre com a requisit que els fitxers d'entrada son de text.

GESTIÓ DE MARCATGES.

GESTIÓ D'ETIQUETATGES

17. Poder donar d'alta etiquetatges.
18. Poder modificar etiquetatges.
19. Poder esborrar etiquetatges.
20. Poder indicar als etiquetatges quins altres etiquetatges son visibles en el procés d'anotació.
21. Poder afegir o treure elements a l'etiquetatge.

22. No deixar esborrar un etiquetatge si esta anotat en algun fitxer del projecte. No deixar treure elements de l'etiquetatge si aquest ja esta anotat en algun fitxer del projecte.

GESTIÓ D'ATRIBUTS

23. Afegir tipus d'atribut : Enter, cadena, enumerat i rang.
24. Poder donar d'alta atributs
25. Poder modificar atributs
26. Poder esborrar atributs
27. Si els atributs estan en un etiquetatge no deixar ni esborrar ni modificar.

GESTIÓ DE RELACIONS

28. Afegir tipus de relació. Quin origen i destí pot tenir aquesta relació.
29. Poder donar d'alta relacions.
30. Poder modificar relacions.
31. Poder esborrar relacions.
32. Si les relacions estan en un etiquetatge no deixar ni esborrar ni modificar.

GESTIÓ DE GRUPS

33. Poder donar d'alta grups.
34. Poder modificar grups.
35. Poder esborrar grups.
36. No deixar ni modificar ni esborrar el grup si està en un etiquetatge.
37. Control d'usuari a l'entrar a l'aplicació.
38. Definir les funcionalitats dels usuaris.
39. Poder afegir usuari administrador.
40. Poder treure usuari administrador.

GESTIÓ D'USUARIS

41. Aplicar tots els canvis de la gestió de marcatges al procés d'anotació.

-Nova anotació

-Només fer anotació si l'etiquetatge actiu és d'atributs.

- Posar valors per defecte dels atributs de l'etiquetatge actiu.
- Canviar valors dels atributs
 - Mostrar atributs segons el seu tipus si l'etiquetatge de l'anotació és igual a l'etiquetatge actiu i és d'atributs.
 - Demanar i validar valors segons tipus de l'atribut.
- Afegir relació entre anotacions
 - Mostrar les relacions possibles a fer si l'etiquetatge actiu és de relacions i mostrar-ne només les que l'etiquetatge origen sigui igual a l'etiquetatge de l'anotació.
 - Validar origen i destí de la relació al fer-la.
- Afegir anotació a grup
 - Mostrar els grups possibles a fer si l'etiquetatge actiu és de grups.
 - Poder afegir anotacions al grup seleccionant-les.
- Veure anotacions en el text.
 - Veure les anotacions de l'etiquetatge actiu i dels seus etiquetatges visibles.
 - Veure les anotacions segons els seus atributs. (Punt 5.1 menú vista atributs).
 - Veure les anotacions segons les seves relacions. (Punt 5.1 menú vista atributs).
 - Veure les anotacions segons els seus grups.
- Veure informació de les anotacions.
 - Veure els atributs i els seus valors al seleccionar anotació.
 - Veure les relacions al seleccionar anotació.
 - Veure els grups al seleccionar anotació.

PROCÈS D'ANOTACIÓ

42. Aplicar etiquetatges del projecte al procés d'anotació.
43. Canviar el format del fitxer d'anotacions. Es vol separar en etiquetes XML el text anotat, les relacions i els grups.
44. Poder veure els fitxers anotats i no anotats.
- 45.** Anotacions aniuades.