

**APLICACIÓ WEB PER ESTRUCTURAR EN XML LA
INFORMACIÓ RECOLLIDA DURANT LES
INTERACCIONS SÍNCRONES DE GRUPS VIRTUALS**

(PROJECTE Ix2xml)

----- Memòria del projecte -----

Universitat Oberta de Catalunya

**Treball Fi de Carrera
Enginyeria Tècnica en Informàtica de Sistemes**

**Ramon Govern Carrillo
Consultor: Santi Caballé Llobet
01 de Juny de 2006**

Índex

1	Pla de treball	5
1.1	Introducció	5
1.1.1	Resum executiu	5
1.1.2	Antecedents	5
1.1.3	Motiu de la tria	6
1.1.4	Anàlisi de riscos	6
1.2	Descripció del projecte	6
1.2.1	Abast	7
1.2.2	Objectius	7
1.3	Organització	7
1.3.1	Relació de tasques	9
1.3.2	Fites	9
1.4	Planificació	11
2	Anàlisi	13
2.1	Introducció	13
2.1.1	Nomenclatura	13
2.1.2	Mètode i model	13
2.2	Especificació i anàlisi	14
2.2.1	Informació inicial	14
2.2.2	Model de domini	14
2.2.3	Glosari del model de negoci	15
2.3	Paquets	15
2.4	Casos d'ús	16
2.4.1	Diagrama de casos d'ús	16
2.4.2	Documentació textual dels casos d'ús	16
2.4.3	Detall del cas d'ús principal (diagrama de col.laboració)	18
2.4.4	Diagrama estàtic de l'anàlisi	19
2.4.5	Interfície d'usuari	20
2.4.6	Eines, tecnologies i entorns de desenvolupament	22
3	Disseny	24
3.1	Introducció	24
3.1.1	Nomenclatura	24
3.1.2	Mètode i model	24
3.2	Disseny dels casos d'ús	25
3.2.1	Diagrama de casos d'ús (etapa anàlisi)	25
3.2.2	Cas d'ús Triar log	25
3.2.3	Cas d'ús Transformar xml	26
3.2.4	Cas d'ús presentar html	26
3.2.5	Descripció de la classe de control Transformar_xml	27
3.3	Revisió del diagrama estàtic de classes	28
3.4	Fitxes CRC	29
3.5	Diagrama de classes de disseny definitiu	31
3.5.1	Consideracions al diagrama de classes	33
3.6	Disseny de la persistència	33
3.7	Disseny de la interfície gràfica d'usuari	34

3.7.1	Guia d'estil	34
3.7.2	Disseny	35
3.7.3	Ajut i documentació	38
3.8	Eines, tecnologies i entorns de desenvolupament	39
3.8.1	Tecnologia client-servidor.....	39
3.8.2	Llenguatge i entorn de desenvolupament	39
3.8.3	Eines.....	39
3.8.4	Consideracions tècniques	40
4	Implementació	41
4.1	Introducció	41
4.2	Consideracions al disseny	41
4.2.1	Diagrama de disseny de classes.....	41
4.2.2	Ús de jsp i javaBeans	41
4.2.3	Funcionalitats al codi versus funcionalitats a les tecnologies xslt, xsd, xml.....	41
4.3	Detalls de la implementació	42
4.3.1	Mètode doTradueix (classe Transforma_xml.java)	42
4.3.2	Mètode doValida (classe Transformar_xml.java)	43
4.3.3	Mètode doMostrar (classe Presentar_html.java).....	44
4.3.4	Jsp triar_log	45
4.3.5	Jsp transformar_xml	46
4.3.6	Jsp presentar_html.jsp	48
4.4	Test de funcionament	49
4.4.1	Test 1: Proves amb type1_logxat_normal.log	49
4.4.2	Test 2: fitxer type1_logxat_errorOrdreCamps.log	51
4.4.3	Test 3 : Fitxer type1_logxat_BigSize.log	51
4.4.4	Test 4: fitxer type5_logxat.log.....	52
4.5	Possibles millores	53
5	Conclusions.....	54
6	Agraïments.....	54
7	Bibliografia i recursos.....	55
8	Anexe 1 – Manual d'usuari	57
8.1	Introducció	57
8.2	Inici de l'aplicació.....	57
8.3	Capçalera de l'aplicació.....	58
8.4	Pantalla Menú Principal (index.jsp)	58
8.5	Pantalla triar_log	59
8.6	Pantalla Transformar_xml	60
8.7	Pantalla presentar_html.....	61
8.8	Veure Fitxers	62
8.9	Carregar fitxer al servidor (Copyright © JavaZOOM 1999-2005)	62
8.10	Pantalla d'error	63
8.11	Especificacions Xat TFC-UOC	63
8.12	Fitxer xml	64
8.13	Expressions Regulars.....	65

8.14	Fitxer de Configuració	65
8.15	Com definir un nou tipus de log ?	67
8.16	Fitxers schema	67
9	<i>Annex 2 - Especificacions log xat TFC-UOC.....</i>	70

1 Pla de treball

1.1 Introducció

1.1.1 Resum executiu

L'aprenentatge virtual, pilar fonamental sobre el que s'erigeix el model educatiu de la UOC té com un dels seus principals objectius l'ús i potenciació d'eines que permetin el treball col·laboratiu no presencial.

Dins de l'àmbit dels entorns col·laboratius d'aprenentatge virtual, darrerament s'ha produït un augment en quan a ús i importància dels identificats com a síncrons (xat, missatgeria instantània, etc...). Aquest augment fa que apareguin unes necessitats en quan a conceptes d'anàlisi i mesura de l'ús que la comunitat fa d'aquests sistemes.

Algunes, les més evidents són de fàcil obtenció:

- Participants al xat
- Durada.
- Freqüència.
- Nombre d'intervencions.

D'altres, no ho són tan, però la seva obtenció aportaria un munt d'informació valuosa alhora de l'anàlisi:

- Anàlisi de continguts exposats pels participants.
- Estadístiques de participació segons tipificació de continguts.
- Valoració qualitativa de les aportacions dels membres (nombre de preguntes o nous fils de conversa introduïts a la sessió, nombre de respostes aportades...).

Aquesta necessitat d'anàlisi d'informació (fora de l'àmbit d'aquest projecte) ens portarà al requeriment de disposar d'un sistema organitzat i estructurat per manejar de forma eficient la informació i facilitar-ne el seu anàlisi.

És aquí on apareix el llenguatge XML com a estàndard àmpliament reconegut i utilitzat per l'organització d'informació en documents estructurats.

Així doncs tenim dos funcionalitats a grans trets identificades, la implementació de les quals ens facilitarà les tasques d'anàlisi anteriorment esmentades.

1. Enregistrament d'informació específica en els logs de les sessions.
2. Transformació i emmagatzemant dels registres originals en un format de dades estructurat, flexible que permeti el seu tractament dins d'entorns web.

El primer punt, descriu les especificacions del TFC corresponent a la proposta anterior a aquest tema ("Desenvolupament d'una aplicació web per a la recollida i classificació d'informació en entorns virtuals d'interacció síncrona").

El punt 1.2 identifica les especificacions del TFC triat i que en els propers apartats, desenvoluparem més abastament.

1.1.2 Antecedents

Aquest TFC es va iniciar el Setembre del 2005 amb l'objectiu de ser finalitzat i presentat al Gener del 2006.

Un dels punts que van motivar la tria (la presa de contacte amb un entorn i eines fins la data desconeguts) va esdevenir des del inici el principal obstacle, ja que el temps invertit en la recollida d'informació, test i familiarització amb aquests productes va triplicar les previsions inicials.

Per aquests motius i d'altres de caire personal, al desembre del 2005 l'estat del projecte no es corresponia amb la planificació prevista i s'entreveia molt difícil que a per les dates fixades de presentació fos possible entregar un treball amb el grau de qualitat que es correspon a un TFC.

Així doncs, es va prendre la decisió de postergar la presentació del projecte fins el proper semestre (Juny 2006).

1.1.3 Motiu de la tria

Hi ha diverses raons que em van portar a prendre la decisió d'enfocar el meu TFC dins d'aquesta àrea, entre tots ells, en destaco els de més rellevància:

- Tot i que la branca d'estudis d'ETI cursada (Sistemes) fa que els continguts d'aquesta àrea no siguin els dominants del meu currículum, és precisament, la motivació per la immersió en entorns i eines noves i la necessitat d'ampliació de coneixements que implica el projecte, el que esdevé un dels factors claus al moment de la tria.
- Com a professional des de fa 15 anys de les TI, m'atrau de forma especial, qualsevol projecte enfocat a l'obtenció d'informació complexa que millori la presa de decisions moltes vegades basades en dades simples, insuficients per modelar de forma eficaç la realitat.
- Aquesta àrea de TFC està directament relacionada amb l'entorn d'aprenentatge virtual de la UOC i a la recta final dels meus estudis, m'atrau profundament poder col·laborar amb l'entitat que m'ha permès obtenir aquesta fita a la meva carrera acadèmica.

1.1.4 Anàlisi de riscos

Tot i que tots aquests riscos han disminuït respecte el semestre passat, es evident que continuen sent punts crítics que caldrà tenir en compte.

- Reduït coneixement teòric de les eines, sistemes i tècniques relacionades amb l'àmbit del projecte. Cal aprofundir en alguns aspectes que no van ser poder tractats el semestre passat.
- Reduïda experiència pràctica o professional en els entorns i eines relacionades amb el projecte, que implicarà un desavantatge en la fase de tria d'eines a emprar i una despesa en temps d'aprenentatge d'ús de les mateixes.
- El temps disponible per l'acabament del projecte és molt curt (aproximadament 3,5 mesos). Tenint en compte la manca de coneixements previs en algunes de les àrees que es tractaran i la complexitat de la especificació, caldrà molta cura en el compliment de la planificació.
- Caldrà un esforç addicional per refer les fases d'anàlisi i disseny incorporant-hi les millores que puguin sorgir de l'experiència del passat semestre i tractant al mateix temps d'aprofitar al màxim la feina feta fins ara.

1.2 Descripció del projecte

Dissenyar una aplicació web que permeti al client l'obtenció d'informació estructurada i de fàcil manipulació a partir de l'enregistrament de sessions xat que contindran les converses categoritzades exposades a la sessió.

- El client iniciarà l'aplicació amb la selecció d'un fitxer log que contindrà informació referent al desenvolupament d'una sessió xat.
- S'executarà un algorisme genèric mitjançant llibreries específiques a tal efecte per transformar els logs dels xats en documents XML que quedaran guardats de forma persistent.
- S'executarà un procés de validació sobre l'XML resultant mitjançant patrons preestablerts.
- La informació obtinguda, emmagatzemada en format de fitxer XML i podrà ser tractada en un futur per la obtenció d'estadístiques.
- La informació obtinguda, serà retornada al client en codi html segons diferents criteris seleccionats.

Les eines que s'utilitzaran per la implementació del projecte, estan basades en el software de programari lliure i prèvia definició a la fase d'anàlisi es preveu que siguin les següents:

- S'implementarà un servidor d'aplicacions (Tomcat) que contindrà l'aplicació web construïda amb tecnologia servlets java i jsp.
- S'usarà el XML com a format final dels fitxer log tractats.
- S'usarà schemas XML (xsd) per la validació de la informació continguda en la estructura XML.
- S'utilitzaran fulles d'estil (xslt) per formatar la sortida segons els criteris seleccionats els documents html finals que es presentaran al client.
- Es testarà la part client per navegador Iexplorer.

1.2.1 Abast

- L'entrada de l'aplicació (logs originats en sessions xat) estarà circumscrita a les especificacions que es defineixin a la fase de anàlisi i compliran les normes, definicions i categoritzacions allà establertes.
- Les funcionalitats que l'aplicatiu proporcionarà en la fase de retorn de la informació seran limitades i quedaran definides a la fase de documentació d'especificacions.
- L'aplicatiu estarà dissenyat i testat per funcionar en un entorn real en producció.
- Juntament amb el codi del producte, es farà entrega de tota la documentació requerida en les especificacions del TFC i un manual bàsic d'usuari.
- S'utilitzaran eines de desenvolupament basades en el software de programari lliure.

1.2.2 Objectius

- Adquirir coneixements sobre les aplicacions web i les tecnologies i eines que s'usen en el seu desenvolupament (servidor d'aplicacions Tomcat, java servlets i jsp).
- Adquirir coneixements sobre l'estàndard XML.
- Aplicar els conceptes i tècniques estudiats a l'assignatura d'Enginyeria de Programari.
- Dissenyar i implementar una eina web que permeti, prèvia transformació a format XML i emmagatzemat, l'obtenció d'informació continguda en fitxers logs resultants de l'enregistrament de sessions xat en grups virtuals d'aprenentatge.

1.3 Organització

El desenvolupament del projecte es dividirà en 4 fases la fi de les quals coincidirà amb les dates de lliurament de les respectives PAC's i el lliurament final del projecte. Aquestes fases són:

- Fase preliminar
- Especificació i Anàlisi
- Disseny
- Implementació
- Documentació

Les tasques de “recerca tecnològica” i “Estudi + proves d’eines i productes” no es subscriuen a cap fase en concret i es duran a terme des de els inicis del projecte fins haver assolit els coneixements i la familiarització necessaris amb les eines i entorns a utilitzar.

La fase d’implementació, s’iniciarà abans de finalitzar el disseny seguint el consell dels consultors per evitar imprevistos en la consecució dels objectius dins les dates fixades.

També seguint les recomanacions dels consultors, el procés de documentació serà indicat des del inici del projecte amb un feedback constant a mida que avanci la implantació.

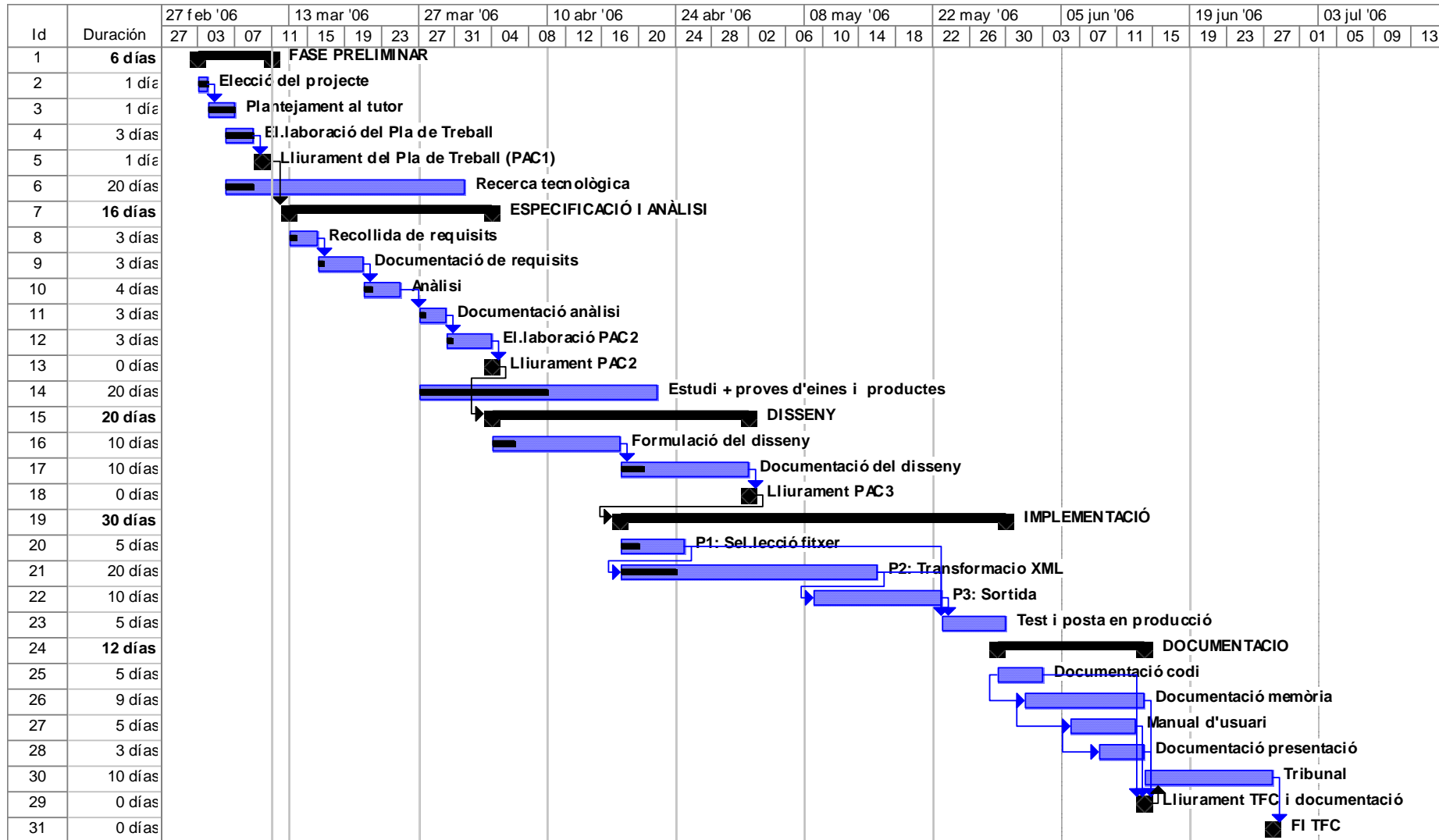
1.3.1 Relació de tasques

Id		Nombre de tarea	Duración	Comienzo	Fin	% completado
1		FASE PRELIMINAR	6 días	vie 03/03/06	vie 10/03/06	100%
2		Elecció del projecte	1 día	vie 03/03/06	vie 03/03/06	100%
3		Plantejament al tutor	1 día	sáb 04/03/06	lun 06/03/06	100%
4		El.laboració del Pla de Treball	3 días	lun 06/03/06	mié 08/03/06	100%
5		Lliurament del Pla de Treball (PAC1)	1 día	vie 10/03/06	vie 10/03/06	100%
6		Recerca tecnològica	20 días	lun 06/03/06	vie 31/03/06	15%
7		ESPECIFICACIÓ I ANÀLISI	16 días	lun 13/03/06	lun 03/04/06	25%
8		Recollida de requisits	3 días	lun 13/03/06	mié 15/03/06	25%
9		Documentació de requisits	3 días	jue 16/03/06	lun 20/03/06	25%
10		Anàlisi	4 días	mar 21/03/06	vie 24/03/06	25%
11		Documentació anàlisi	3 días	lun 27/03/06	mié 29/03/06	25%
12		El.laboració PAC 2	3 días	jue 30/03/06	lun 03/04/06	25%
13		Lliurament PAC2	0 días	lun 03/04/06	lun 03/04/06	0%
14		Estudi + proves d'eines i productes	20 días	lun 27/03/06	vie 21/04/06	50%
15		DISSENY	20 días	mar 04/04/06	lun 01/05/06	25%
16		Formulació del disseny	10 días	mar 04/04/06	lun 17/04/06	25%
17		Documentació del disseny	10 días	mar 18/04/06	lun 01/05/06	25%
18		Lliurament PAC3	0 días	lun 01/05/06	lun 01/05/06	0%
19		IMPLEMENTACIÓ	30 días	mar 18/04/06	lun 29/05/06	15%
20		P1: Sel.lecció fitxer	5 días	mar 18/04/06	lun 24/04/06	40%
21		P2: Transformació XML	20 días	mar 18/04/06	lun 15/05/06	20%
22		P3: Sortida	10 días	mar 09/05/06	lun 22/05/06	0%
23		Test i posta en producció	5 días	mar 23/05/06	lun 29/05/06	0%
24		DOCUMENTACIÓ	12 días	lun 29/05/06	mar 13/06/06	0%
25		Documentació codi	5 días	lun 29/05/06	vie 02/06/06	0%
26		Documentació memòria	9 días	jue 01/06/06	mar 13/06/06	0%
27		Manual d'usuari	5 días	mar 06/06/06	lun 12/06/06	0%
28		Documentació presentació	3 días	vie 09/06/06	mar 13/06/06	0%
29		Lliurament TFC i documentació	0 días	mar 13/06/06	mar 13/06/06	0%

1.3.2 Fites

Id		Nombre de tarea	Duración	Comienzo	Fin	% completado
1		FASE PRELIMINAR	6 días	vie 03/03/06	vie 10/03/06	100%
5		Lliurament del Pla de Treball (PAC1)	1 día	vie 10/03/06	vie 10/03/06	100%
7		ESPECIFICACIÓ I ANÀLISI	16 días	lun 13/03/06	lun 03/04/06	25%
13		Lliurament PAC2	0 días	lun 03/04/06	lun 03/04/06	0%
15		DISSENY	20 días	mar 04/04/06	lun 01/05/06	25%
18		Lliurament PAC3	0 días	lun 01/05/06	lun 01/05/06	0%
29		Lliurament TFC i documentació	0 días	mar 13/06/06	mar 13/06/06	0%
31		FI TFC	0 días	mar 27/06/06	mar 27/06/06	0%

1.4 Planificació



2 Anàlisi

2.1 Introducció

El present capítol, conté el resultat de les fases d'especificació i anàlisi del TFC "Aplicació web per estructurar en xml la informació recollida durant les interaccions síncrones de grups virtuals".

2.1.1 Nomenclatura

S'ha triat el mot lx2xml (log xat to xml) com a àlies del projecte " Aplicació web per estructurar en xml la informació recollida durant les interaccions síncrones de grups virtuals". D'ara en endavant, per referenciar aquest projecte es podran utilitzar ambdues terminologies indistintament.

2.1.2 Mètode i model

S'ha triat el model de desenvolupament de programari basat en la programació orientada a objectes degut a que és el més idoni per representar el context en que la aplicació haurà de funcionar (entorn web model client servidor).

Per la representació formal dels diagrames s'utilitzarà el model UML (Unified Modeling Language).

Per l'execució dels processos d'especificació i anàlisi, s'han seguit les pautes i models proposats a l'assignatura "Enginyeria del Programari" dels estudis d'Enginyeria Tècnica en Informàtica de Sistemes de la Universitat Oberta de Catalunya .

2.2 Especificació i anàlisi

2.2.1 Informació inicial

A continuació reproduïm el contingut de l'especificació a alt nivell del projecte que ja introduïem en el pla de treball:

Dissenyar una aplicació web que permeti al client l'obtenció d'informació estructurada i de fàcil manipulació provinent de l'enregistrament de sessions xat que contindran les converses categoritzades exposades a la sessió.

- El client iniciarà l'aplicació amb la selecció d'un fitxer log que contindrà informació referent al desenvolupament d'una sessió xat.
- El fitxer haurà estat generat per una aplicació web creada dins l'entorn TFC UOC destinada a proveir els entorns col·laboratius amb una eina que permeti l'execució de sessions de comunicació síncrona i el seu posterior enregistrament en format fitxer de text.
- El fitxer tindrà un format pre-establert que quedarà concretat en data posterior a l'entrega del present anàlisi. Alguns dels paràmetres que establiran aquest format poden ser:
 - Delimitadors d'intervenció.
 - Nombre de camps per intervenció.
 - Delimitadors de camp.
 - Tipus de camp (obligatori / opcional).
- S'executarà un algorisme mitjançant llibreries específiques a tal efecte per transformar els logs dels xats en documents XML que quedaran emmagatzemats de forma persistent.
- S'executarà un procés de validació sobre l'XML resultant mitjançant patrons preestablerts.
- La informació obtinguda, s'emmagatzemarà en format XML i podrà ser tractada en un futur per la obtenció d'estadístiques.
- La informació obtinguda, serà retornada al client en codi html segons diferents criteris seleccionats per l'usuari.

Puntualitzacions:

- Tot i que aquest és l'objectiu genèric del projecte serà tractar un model concret de logs (provinents de un aplicatiu xat dissenyat en un altre TFC del àrea), pot ser útil que el programa proporcioni una certa flexibilitat per tractar lleugeres variacions del log pre-establert (provinents potser de futures modificacions a l'aplicatiu de xat). Si més no, que es puguin ajustar certs paràmetres en mode configuració per permetre l'anàlisi de diferents tipus de logs.
- El programa, ha de permetre la possibilitat a l'usuari de triar el formateig de la informació resultant (sortida html) entre diferents tipus. Bàsicament, aquest tipus de sortida, formatarà la informació entregada a l'usuari classificada de 3 maneres diferents:
 - per intervencions.
 - per tipologia d'intervencions
 - per usuaris.

2.2.2 Model de domini

A primer cop d'ull s'identifiquen les següents classes, les quals formen part del domini de l'aplicació:

- FitxerLog

- FitxerXml
- SortidaHtml
- Peticio_de_traduccio
- FitxerXslt
- FitxerXsd

No s'identifiquen rel.lacions d'associativitat entre elles. Més endavant, com a conseqüència de la interpretació de la descripció dels casos d'ús, anirem desenvolupant aquestes classes fins arribar al diagrama estàtic de classes, on hi reproduïrem una informació més detallada d'aquest model.

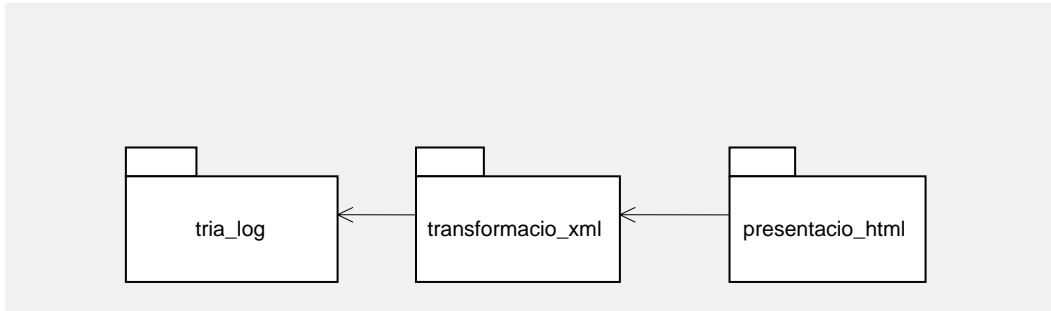
2.2.3 Glosari del model de negoci

DTD	(Documen Type Definition) Descripció d'un conjunt de regles XML per la representació d'un document XML d'un determinat tipus.
HTML	(Hipertext Markup Language) Es un llenguatge de marcatge dissenyat per l'estructuració de textos i per la seva presentació en format hipertext (estàndart pàgines web).
Intervenció	Cada una de les aportacions que fan els usuaris al xat.
Log	Fitxer (generalment de text) que conté informació corresponents al desenvolupament i contingut d'una sessió xat.
OMG	(Object Management Group) Organització no oficial d'estandarització de tot allò relacionat amb la programació orientada a objectes.
Parser	Analitzador sintàctic.
Participant	idem "usuari".
Processador xslt	Component lògic que realitza la funció de generar una transformació sobre un fitxer xml basant-se en les regles especificades a un fitxer xslt.
Sessió	Espai de temps contingut entre l'inici i fi d'un xat.
SGML	(Standart Generalized Markup Language) Normalitzat per ISO, consisteix en un estàndart d'organització i etiquetatge de documents.
UML	(Unified Modeling Language) Model de representació per la construcció de programari orientat a objectes proposats per l'OMG.
Usuari	(en el context de l'aplicació) L'executor dels casos d'ús del programa.
W3C	(World Wide Web Consortium) Organització que produeix estàndarts per la World Wide Web.
Xat	Comunicació escrita a través d'Internet que es realitza entre dues o més persones de forma instantànea.
XML	(Extensible Markup Language) Llenguatge de marcatge generalitzat. Deriva del SGML i ha estat desenvolupat pel W3C. Es basa en l'estructuració semàntica d'un document per facilitar el seu tractament i estandarització.
XPath	Especificació responsable de la sintaxi d'accés i consulta als elements d'un document XML
XSD	(Esquema) Document de definició estructural (similar al DTD) que compleix amb l'estàndart XML. Són útils per validar que un document XML s'adapta a una determinada estructura i regles.
XSL	(Extensible Stylesheet Language). Conjunt de les especificacions XSL-FO, XSLT i XPath.
XSL-FO	Estandart basat en XML amb la funció de donar format a la informació.
XSLT	(Extensible Stylesheet Language Transformation) Llenguatge basat en XML que permet definir un conjunt de regles per transformar els elements d'un document XML.

2.3 Paquets

Distingim 3 paquets d'anàlisi principals els quals engloben aspectes comuns de l'aplicació:

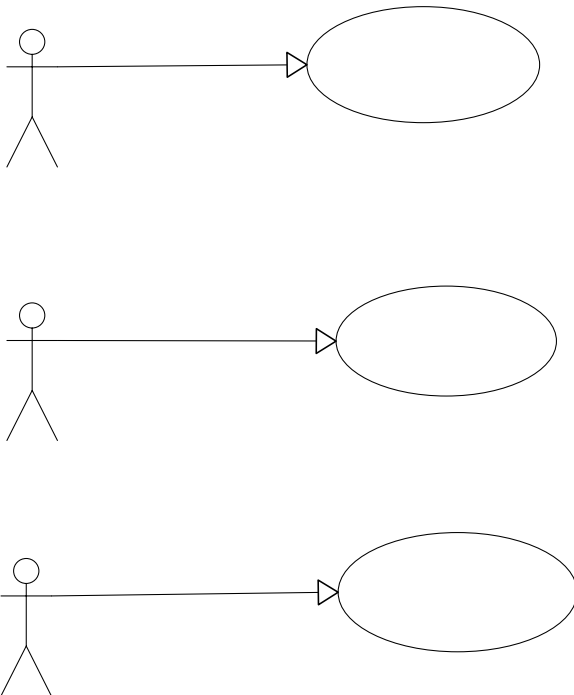
- Tria_log: Contindrà el conjunt d'elements destinats a la selecció del fitxer a tractar i al seu processament previ a la transformació.
- Transformacio_xml: Contindrà el conjunt de subsistemes destinats a la transformació del fitxer triat en un fitxer xml, la seva validació com a fitxer xml correcte i la obtenció de la seva persistència.
- Presentacio_html: Contindrà els elements responsables del retorn d'informació provinent del fitxer xml generat en format html i segons criteris i format seleccionats per l'usuari.



2.4 Casos d'ús

En un anàlisi inicial, s'identifica un cas d'ús a cadascun dels paquets. A continuació mostrem el diagrama que els representa.

2.4.1 Diagrama de casos d'ús



2.4.2 Documentació textual dels casos d'ús

A continuació procedim a la descripció textual dels casos d'ús definits al diagrama anterior:

Cas d'ús número 1: "Triar log"

Resum de la funcionalitat: Permet a l'usuari la selecció d'un fitxer de text per la seva posterior transformació. I la tria d'un nom pel fitxer xml de sortida que generarà el següent cas d'ús.

Paper dins el treball de l'usuari: És el cas d'ús principal del paquet *Tria_log*.

Actors: usuari

Casos d'ús relacionats: *Transformar_xml*

Precondició: Existeix un fitxer de text amb el contingut enregistrat d'una sessió xat.

Postcondició: Els fitxers *log* i sortida xml han estat seleccionats.

Flux d'events principal:

1. L'actor introdueix el nom de fitxer *log* corresponent a l'enregistrament d'una sessió xat per a ser tractat.
2. L'actor introdueix el nom del fitxer xml que generarà el programa. L'actor disposa d'una opció
3. L'actor valida l'operació i els fitxers seleccionats es passen com a paràmetre al següent cas d'ús.

Flux alternatiu:

- 3b. Si el fitxer no existeix, el sistema mostra un missatge d'error a l'usuari.
- 4b. El sistema torna a sol·licitar el nom del fitxer.

Cas d'ús número 2: "Transformar xml"

Resum de la funcionalitat: transforma un fitxer de text que conté l'enregistrament d'una sessió xat a un fitxer xml segons les regles d'una configuració pre-establerta.

Paper dins el treball de l'usuari: És el cas d'ús principal del paquet *transforma_xml*.

Actors: usuari.

Casos d'ús relacionats: *Presentar html* i *Triar log*.

Precondició: Existeix un fitxer seleccionat que es rep com a paràmetre d'entrada. Existeix un conjunt de regles pre-establertes que permetran la transformació del contingut del fitxer en una estructura xml.

Postcondició: S'ha executat la transformació del fitxer. S'ha executat la validació del fitxer. Si el resultat de la validació és correcte, s'ha amagat un fitxer xml amb el contingut de la transformació de forma persistent.

Flux d'events principal:

1. El sistema transforma el fitxer rebut com a paràmetre a una estructura xml seguint les regles pre-establertes en la configuració del programa.
2. Es grava l'estructura xml resultant a disc en un fitxer xml.
3. Es valida el fitxer xml.
4. Si el resultat de la validació és correcte, s'informa a l'actor i el sistema arrenca l'execució del proper cas d'ús *Presentar html*.

Flux d'events alternatiu:

4b. Si el resultat de la validació no és correcte, el sistema mostra un missatge d'error a l'usuari amb la informació referent als errors que provoquen la no validació. El sistema retorna al inici del programa.

Cas d'ús número 3: "Presentar html"

Resum de la funcionalitat: presenta a l'actor el contingut del fitxer xml resultant del cas d'ús *Transformar xml* en format html segons uns criteris d'agrupació i ordenació triats per l'actor.

Paper dins el treball de l'usuari: És el cas d'ús principal del paquet *presentar html*.

Actors: usuari.

Casos d'ús relacionats: *Transformar xml*.

Precondició: Existeix un fitxer xml vàlid per ser processat. Existeix un conjunt de criteris preestablerts que permetran a l'actor seleccionar diferents formats de sortida.

Postcondició: S'ha mostrat a l'actor el contingut del fitxer xml en format html segons els criteris que l'actor a triat.

Flux d'events principal:

1. El sistema presenta a l'actor una pantalla on seleccionar el criteri de presentació del fitxer.
2. L'actor selecciona un criteri.
3. El sistema mostra en format html el contingut del fitxer xml segons els criteris d'ordenació i agrupació triats per l'actor.
4. El sistema demana a l'actor si vol transformar un altre fitxer.
5. Si la resposta de l'actor es afirmativa, el sistema retorna al cas d'ús *Triar log*.

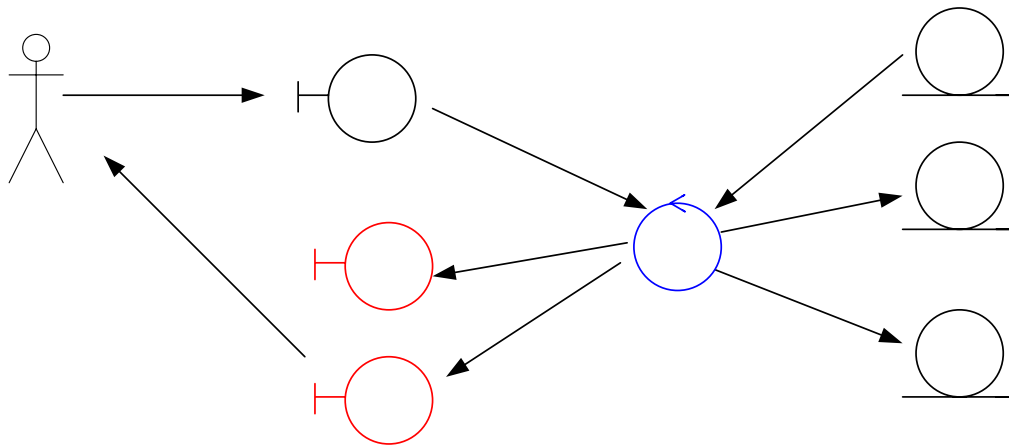
Flux d'events alternatiu:

5b. Si la resposta és negativa, el sistema tanca l'aplicació.

2.4.3 Detall del cas d'ús principal (diagrama de col·laboració)

Degut a que considerem que es tracta del cas d'ús principal de l'aplicació, procedim a definir amb més detall les seves característiques. Per aquesta tasca, farem ús del diagrama de col·laboració, on hi identificarem les classes control, frontera i entitat.

Cas d'ús número 2 "Transformar xml"

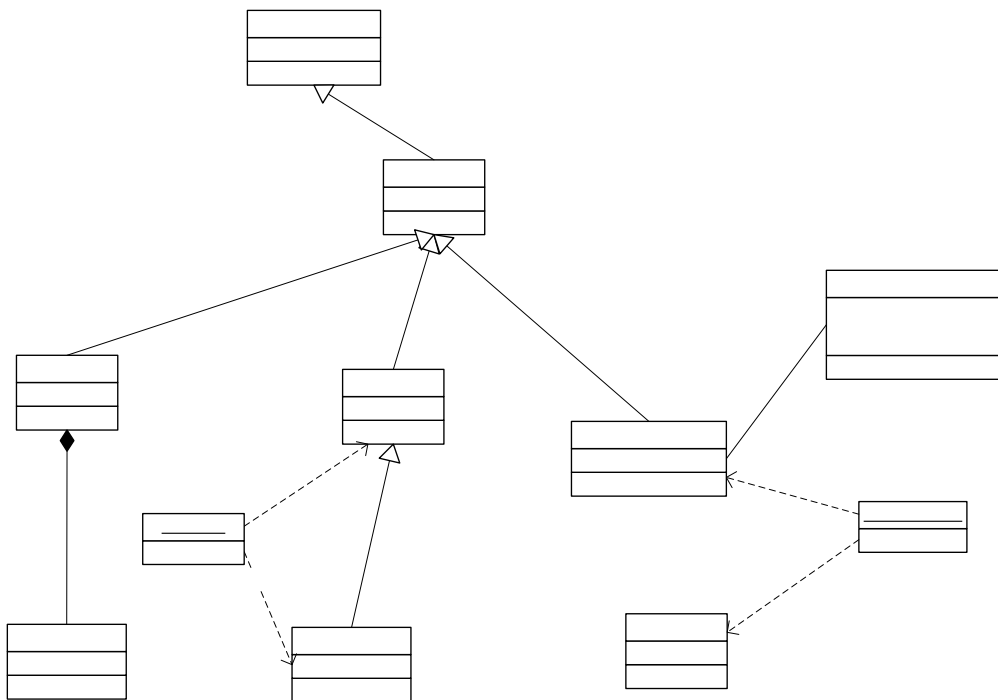


Flux d'events principal:

1. L'usuari arrenca l'operació transformació_log, el sistema transformarà el fitxer log en un fitxer xml segons els paràmetres preestablerts.
2. El sistema mostra a l'actor el resultat de la transformació.
3. El sistema validarà el fitxer xml resultant utilitzant un fitxer xml esquema.
4. El sistema dóna el resultat de la validació a l'actor.

2.4.4 Diagrama estàtic de l'anàlisi

Tot seguit, procedirem a identificar les classes d'entitat i enumerar-ne les seves principals característiques. En el següent diagrama de classes les podem veure representades amb els seus atributs i interrelacions corresponents.



2.4.4.1 Comentaris

Existeix la possibilitat de que a l'etapa de disseny, les classes *FitxerLog*, *FitxerXml* i *FitxerXslt* siguin eliminades ja que no representen cap diferència entre elles (atributs) ni cap especialització respecte a la classe *FitxerText*. Tampoc es preveu que puguin contenir mètodes diferents de les operacions comunes (lectura, escriptura, creació etc...) que incorpora un objecte d'aquest tipus. Aquestes classes, han sorgit de la interpretació textual de les descripcions dels casos d'ús i considerem aclaridor representar-les en el diagrama en aquesta fase d'anàlisi.

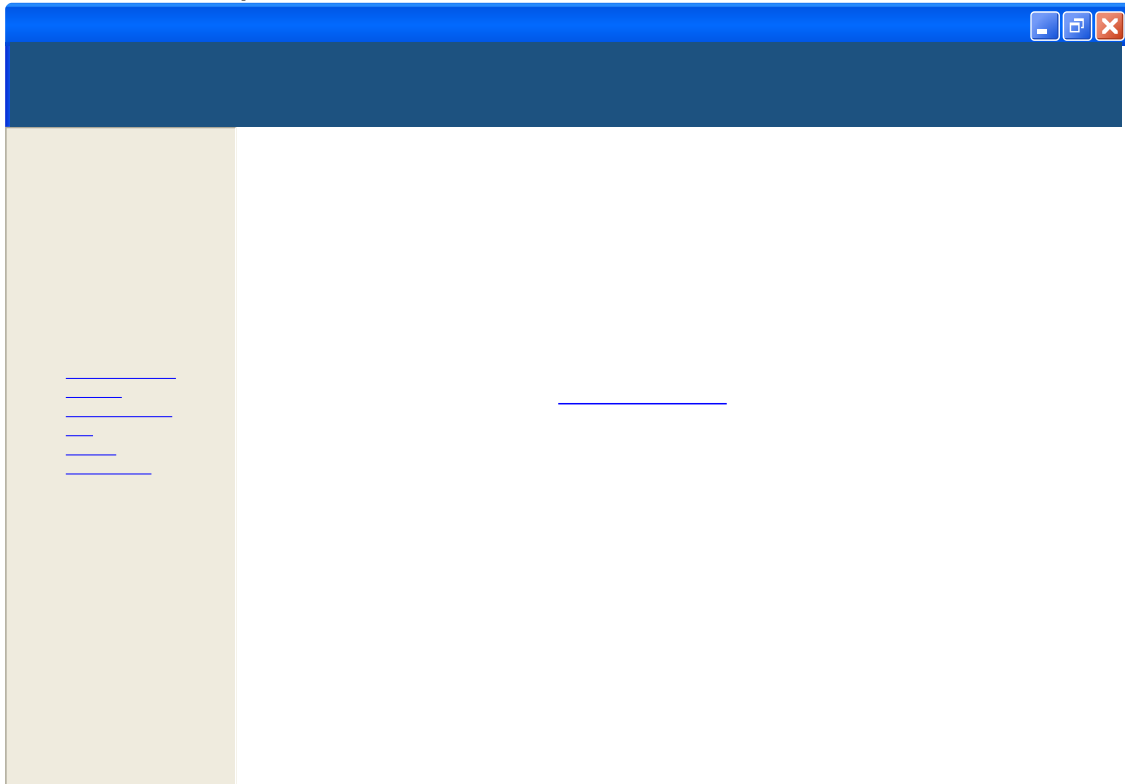
Validador i *Transformador* són objectes únics que en la etapa de disseny seran substituïts per les respectives crides als motors de validació i transformació de les llibreries triades al respecte. Aquest objectes usaran respectivament les classes *FitxerXml*, *FitxerXsd* i *FitxerXslt*, *Sortida*.

En l'etapa de disseny, probablement apareguin més classes com a conseqüència del desenvolupament de l'algorisme de traducció a xml.

2.4.5 Interfície d'usuari

Passem ara a donar una aproximació de com serà la interfície d'usuari pel que respecte al cas d'ús principal "**Transformar log**". Aquesta aproximació, més que preveure el format gràfic (pendent de les llibreries gràfiques del llenguatge de desenvolupament triat), fa referència al contingut i funcionalitats de la interfície:

Classe MenuPrincipal



Classe PantallaSeleccioFitxer:



Classe PantallaErrorFitxerNoTrobat:



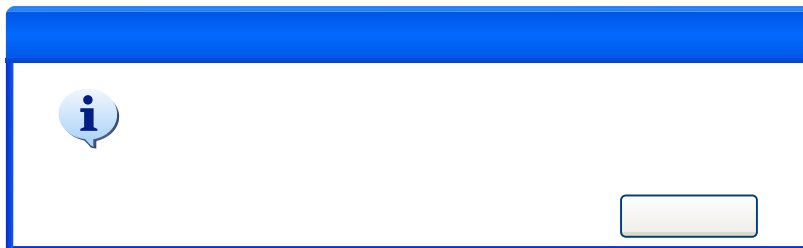
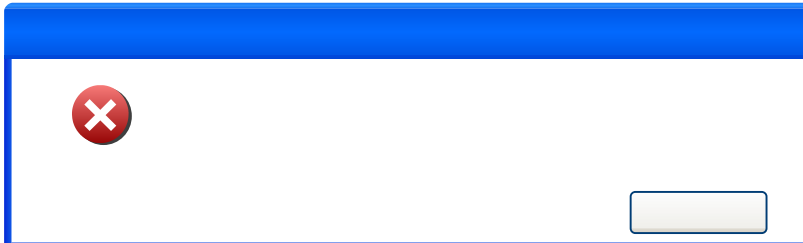
Classe PantallaResultatTransformacio:



Seleccio fitxer

Introdueix el nom del fitxer

Classe PantallaErrorValidacio:



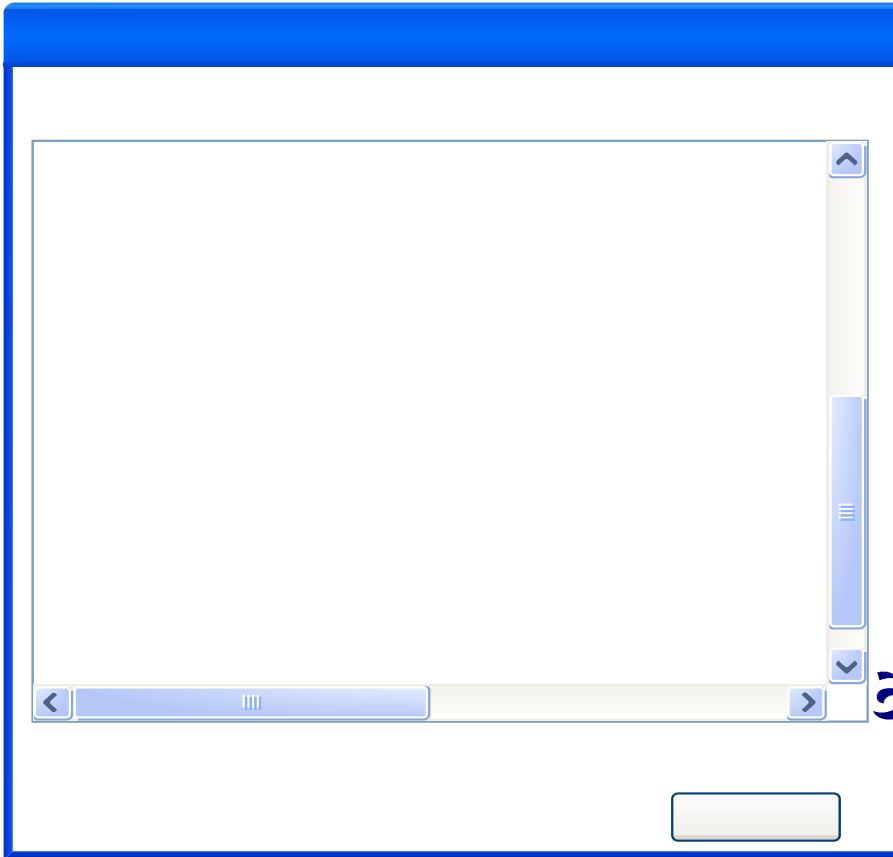
Classe PantallaSeleccioSortida:



no trobat

ATENCIÓ !
El nom del fitxer no pot ser buit.
Prem "Acceptar" per continuar.

Classe PantallaSortidaHtml:



a.html

Sortida fitxer.log orden

2.4.6 Eines, tecnologies i entorns de desenvolupament

Encara que pendents de concretar algunes de les opcions en la propera etapa del desenvolupament, ja podem avançar algunes de les tries realitzades, si més no una aproximació:

2.4.6.1 Tecnologia client-servidor

- Servidor d'aplicacions Apache Tomcat (Projecte Jakarta) v 5.0.28 amb tecnologia web d'implementació de servlets v 2.4
- El suport pel tractament de xml en el servidor l'incorpora Xerces Java Parser 1.4.4 release.
- El motor de transformació el proporcionarà JDOM treballant amb objectes SAX (TransformerFactory i Transformer).

2.4.6.2 Llenguatge i entorn de desenvolupament

- J2SE (Java 2 Platform Standart Edition) v 1.4.2
- JAXP (Java Api for XML)
- SAX v 2.0 (Simple API for XML parsing)
- XSLT 1.0
- JDOM

2.4.6.3 Eines

Desenvolupament:

USUARI HORA

=====

laura 20:26

laura 20:30

mon1 20:25

mon1 20:30

mon1 20:31

dubtes ?

mon1 20:32

manel 20:29

- Desenvolupament en les pròpies eines del J2SE amb l'entorn integrat de desenvolupament NetBeans IDE v 4.1

Documentació:

Microsoft Word XP
Microsoft Visio 2003 Professional Edition
Microsoft Power Point XP
Adobe Acrobat

Entorns de proves (servidor):

Microsoft Windows XP Professional

Entorns de proves (clients):

Microsoft Windows XP Professional (browser: Internet Explorer v 6.0)

3 Disseny

3.1 Introducció

El present capítol, conté el resultat de la fase de disseny del TFC “Aplicació web per estructurar en xml la informació recollida durant les interaccions síncrones de grups virtuals” .

3.1.1 Nomenclatura

S’ha triat el mot lx2xml (log xat to xml) com a àlies del projecte " Aplicació web per estructurar en xml la informació recollida durant les interaccions síncrones de grups virtuals". D’ara en endavant, per referenciar aquest projecte es podran utilitzar ambdues terminologies indistintament.

3.1.2 Mètode i model

S’ha triat el model de desenvolupament de programari basat en la programació orientada a objectes degut a que és el més idoni per representar el context en que l’aplicació haurà de funcionar (entorn web model client servidor).

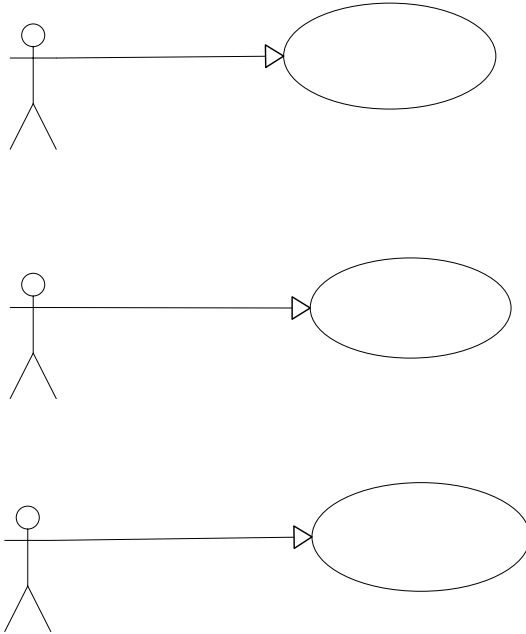
Per la representació formal dels diagrames s’utilitzarà el model UML (Unified Modeling Language).

Per l’execució dels processos de disseny, s’han seguit les pautes i models proposats a l’assignatura "Enginyeria del Programari" dels estudis d’enginyeria Tècnica en Informàtica de Sistemes de la Universitat Oberta de Catalunya i les orientacions del consultor d’aquest TFC.

3.2 Disseny dels casos d'ús

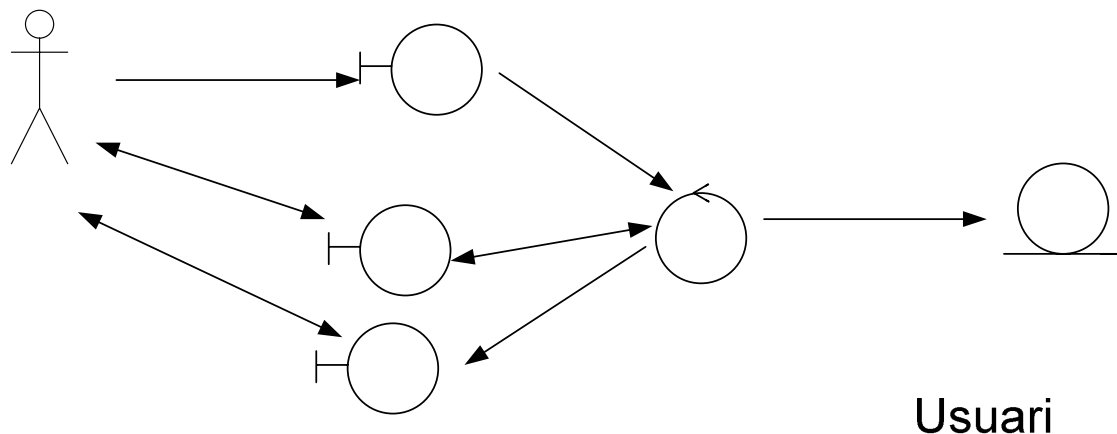
3.2.1 Diagrama de casos d'ús (etapa anàlisi)

Partim del següent diagrama elaborat a l'etapa d'anàlisi:



Anem a descriure a continuació cadascun d'ells de forma detallada tal i com correspon al nivell de l'etapa de disseny.

3.2.2 Cas d'ús Triar log



Quan l'usuari escull l'opció corresponent a aquest cas d'us del menú principal, es presenta per pantalla un formulari (quadre de text) per que l'usuari introdueixi el nom del fitxer a transformar.

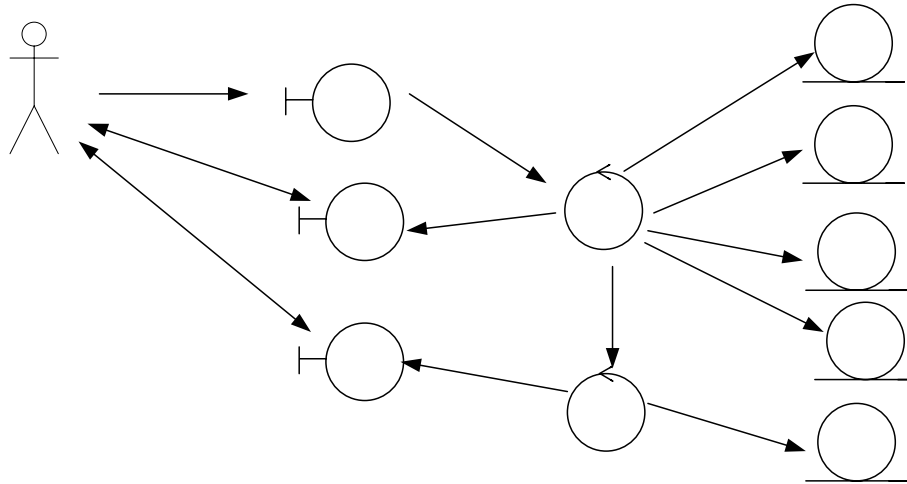
Adicionalment, el sistema presenta a l'usuari la possibilitat d'accedir a una pantalla on es llisten tots els fitxers continguts al directori de treball amb la finalitat de que pugui veure el seu nom i contingut i fer una selecció del fitxer log triat tot clicant sobre el seu nom.

Un cop introduït el nom del fitxer, el sistema proposa el nom del fitxer xml de sortida, l'usuari té la possibilitat d'acceptar-lo o modificar-lo.

Quan l'usuari, accepta els paràmetres del formulari, s'arrenca l'execució del següent cas d'ús *Transformar xml* passant-li els noms dels fitxers com a paràmetre.

3.2.3 Cas d'ús Transformar xml

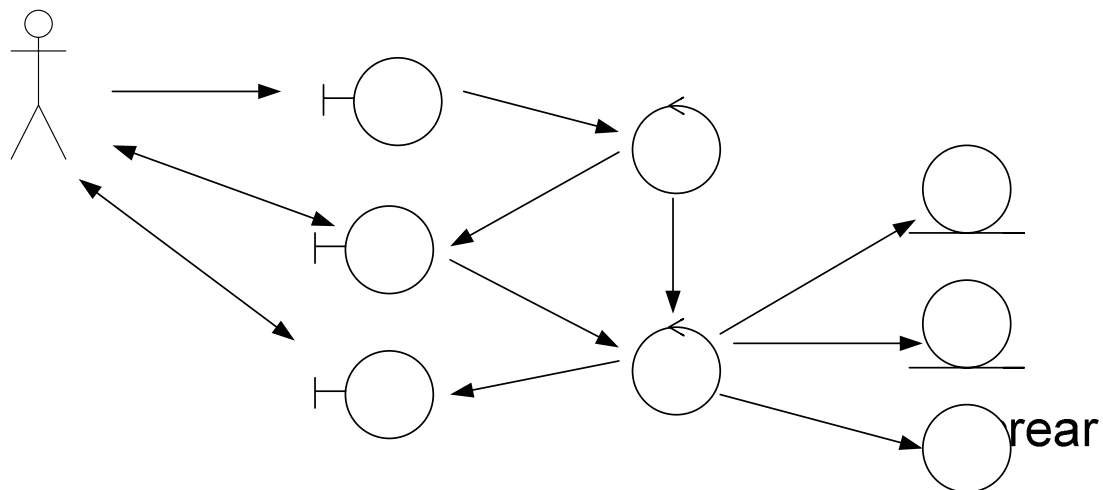
Passem a dissenyar el cas d'ús principal *Transformar xml*. Basant-nos en el diagrama de col·laboració obtingut en l'etapa d'anàlisi procedim a ampliar-lo per obtenir més detall.



Transformar xml recull els paràmetres del cas d'ús anterior (el noms dels fitxers log i xml) es procedeix a la traducció del fitxer. Es genera un fitxer xml a disc en base al fitxer log i als paràmetres que el defineixen. Posteriorment, s'executa *Validar_xml* que procedeix a la validació del xml generat al mètode anterior .

S'envia a l'usuari la pantalla *PResultatValidacio* que mostra el resultat de la validació (correcte o no correcte) i informació sobre els errors en cas de no correcte. L'usuari pot visualitzar el contingut del fitxer xml generat o bé el contingut del fitxer que conté les línies descartades del procés de transformació. Finalment, l'usuari pot retornar a la pantalla *triar_log* per realitzar un altre tria de fitxer o bé procedir a la presentació hmtl, on es crida el cas d'ús *presentar html* .

3.2.4 Cas d'ús presentar html



S'envia a l'usuari la pantalla *PSeleccioSortida* que permet triar entre tres formats de sortida html pre-establerts. Quan l'usuari tria una de les opcions, s'executa la classe de control *Transformar_html* la qual recull un fitxer de normes de format en funció de l'opció triada i utilitzant un motor de transformació, inicia la transformació del fitxer xml. Es crea una pantalla que presenta a l'usuari la sortida html formatada segons l'opció triada.

L'usuari pot sol·licitar altres tipus de presentació o repetir-ne algun fins que desitgi finalitzar aquest cas d'ús. Podrà dirigir-se al menú principal per sortir de l'aplicació o anar directament a la pantalla de triar_log per iniciar un altre selecció.

3.2.5 Descripció de la classe de control *Transformar_xml*

Passem a descriure amb més detall aquesta classe, ja que conforma el nucli de la funcionalitat de l'aplicació.

Se li passen 3 paràmetres: un objecte de la classe *TipusLog* que conté tots els paràmetres identificadors del log (nº de camps, delimitadors, etc..) , un objecte de la classe *fitxer* que ens apuntarà el fitxer de text emmagatzemat a disc que conté el log a tractar i un objecte de la classe *fitxerxml* que conté el nom del fitxer xml que es crearà per emmagatzemar-hi el resultat de la transformació.

El mètode comença instanciant un objecte arbre xml el qual contindrà l'estructura xml que anirem creant a partir de la lectura seqüencial del fitxer log.

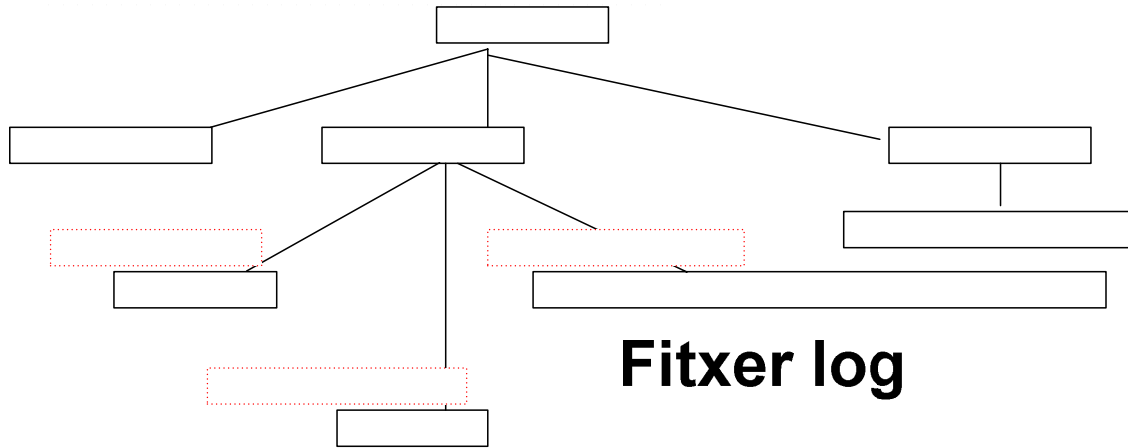
El mètode crea dos objectes patró que contenen les característiques formalitzades en llenguatge d'expressions regulars del tipus de log a tractar segons dos casos: línies especials i intervencions.

El mètode obra el fitxer log i mentre no arriba al seu final anirà llegint línies del fitxer que es correspondran amb possibles entrades vàlides del log.

Aquesta intervenció serà comparada amb el patró de les línies especials, si coincideix es guardarà el seu contingut a l'arbre xml en un element *linia_especial*.

Si no coincideix, es procedirà a comprovar si quadra amb el patró de intervencions, en cas afirmatiu es crearà un element de l'arbre xml *intervenció* es procedirà a obtenir els camps segons les agrupacions que marca el patró. Per cada intervenció vàlida obtenida i tans cops com nombre de camps s'hagin definit, crearem un element xml *camp* anirem recullim les agrupacions i generant un element de l'arbre xml que contindrà el nom del camp i el seu contingut, l'element serà inserit com a fill de l'element *intervenció*.

Si no coincideix amb cap dels dos patrons, la intervenció serà gravada al fitxer de línies rebutjades.



Fitxer log

10:22 maria: Hola a tothom

Al arribar al final del fitxer log, ja no tindrem més intervencions i ja podrem bolcar l'estructura xml a un fitxer de text (format xml sense validar).

10:23 user1: Hola Maria, e

3.3 Revisió del diagrama estàtic de classes

Procedim a la revisió del diagrama estàtic de classes que varem obtenir a l'etapa d'anàlisi. Per aquesta revisió, tenint en compte el llenguatge (Java) i la tecnologia (Servlets i JSP) escollits hem considerat els següents punts:

10:24 user1 → desconect

Normalització de noms: Tot i que ja varem tenir en compte aquest punt a l'etapa d'anàlisi hem revisat la nomenclatura de les classes segons la normativa de noms Java i la coherència del nostre disseny.

Reutilització de classes: Ja varem esmentar a l'etapa de disseny que algunes classes podrien desaparèixer si consideràvem que no aportaven cap avantatge respecte a les seves classes pare. Així doncs, les classes *FitxerLog*, *FitxerXml*, *FitxerXsd* i *FitxerXslt* desapareixen deixant la seva classe pare *Fitxer*.

Adaptació de la herència: Java suporta el nivell d'herència que havíem definit al diagrama de classes. No s'observen casos d'herència múltiple.

Substitució d'interfícies: No aplica.

Canvis per millorar el rendiment: No s'han incorporats canvis en el diagrama en base a aquest criteri.

Especificació d'operacions implícites: Totes les classes de l'aplicació tindran les operacions constructores, i els mètodes get i set corresponents als cadascun dels seus atributs. Aquestes operacions no seran relaxades al diagrama de classes per motius de visibilitat.

Cohesió i acoblament: La classe *Intervenció* desapareix ja que no aporta cap avantatge i la seva funcionalitat serà proporcionada per un atribut tipus String dins la classe control *transformar_xml*.

Intervenció

Atribut camp=data

3.4 Fitxes CRC

A continuació mostrem en format fitxes CRC (Classe, responsabilitats, col·laboracions) la representació de les classes més característiques de l'aplicació:

Lx2xml
inicia l'aplicacio inicialitza espai recull parametre opcions menu principal
Triar_log

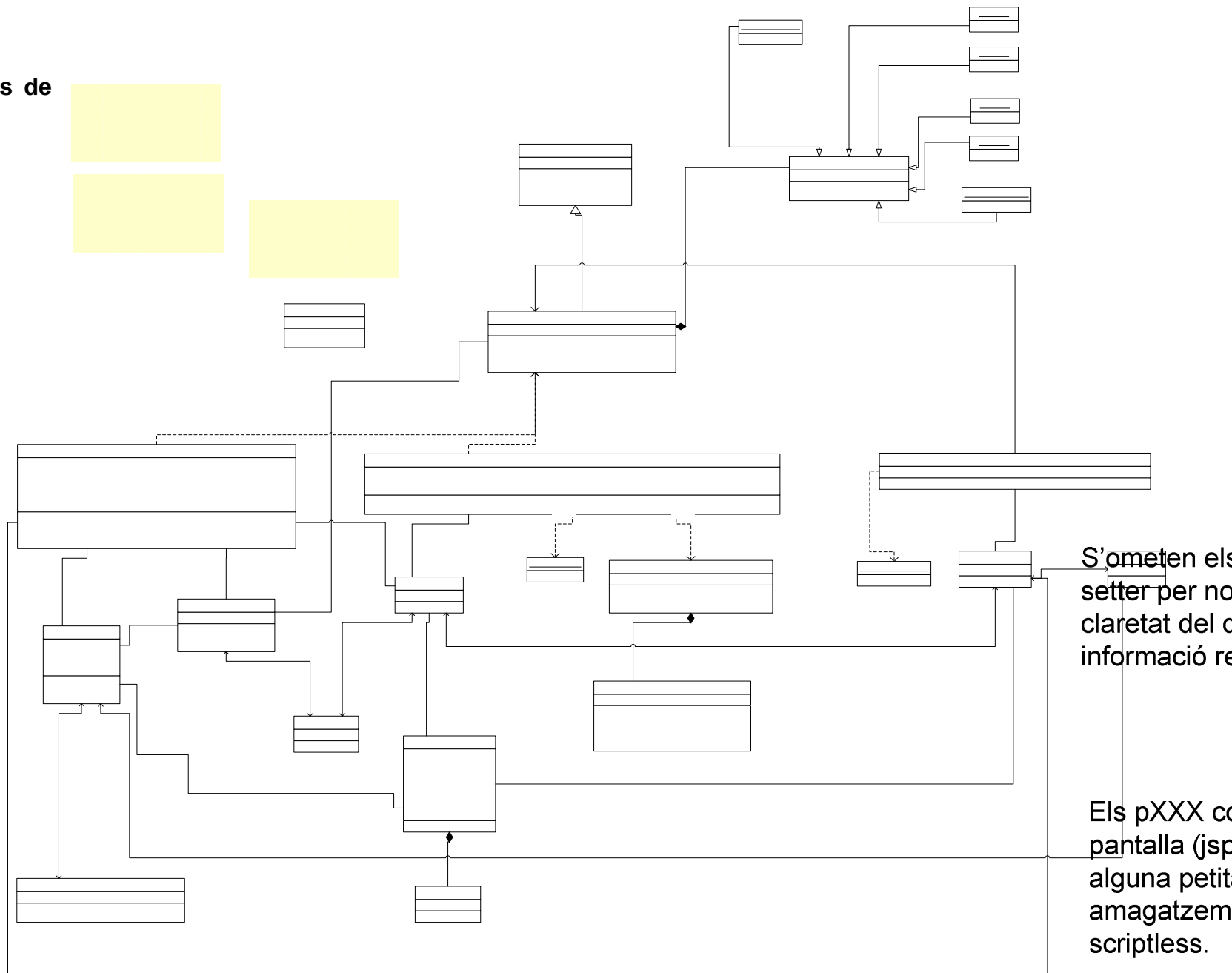
Triar_log
recull noms fitxer log i xml
Transformar_xml, Lx2xml

Transformar_xml
recull parametres del fitxer log obre el fitxer log crea el fitxer xml executa mètode transformador a fitxer xml
Presentar_html, Triar_log

Validar_xml
Presentar_html, Triar_log

Presentar_html
Sel.lecciona opcions de sortida valida fitxer xml amb fitxer esquema mostra resultat de la validació
triar_log, Lx2xml

3.5 Diagrama de classes de disseny definitiu



3.5.1 Consideracions al diagrama de classes

- Hem especificat el nivell d'herència a les classes en que podia aportar informació significativa en el disseny.
- Hem especificat els mètodes i atributs que en el moment del disseny considerem necessaris a les classes pròpies implementades a l'aplicació.
- Hem especificat els mètodes i atributs que en el moment del disseny considerem útils per l'aplicació a les classes utilitzades d'altres API's.
- La classe *Document* es correspon a una estructura xml en forma d'arbre que conté elements tipus *Element*. Aquesta estructura serà construïda durant el procés de transformació i finalment persistida utilitzant les funcions de les classes *Fitxer* i *GestorFitxers*.
- Les classes *Motor_Validacio* i *Motor_Transformació* es correspondran amb els respectius motors de la tecnologia triada que permetran realitzar la comprovació del fitxer xml i la seva posterior transformació a format html

3.6 Disseny de la persistència

Tractarem la persistència de la classe *FitxerTexte* que es correspondrà als següents objectes:

Fitxers log: Es considera que aquests fitxers log estaran ubicats a una carpeta del servidor en format fitxer de text.

Fitxers xml: Contindran el resultat de la traducció del fitxer log a xml. Es guardaran en format fitxer de text a una carpeta del servidor.

Fitxer esquema: Contindrà el conjunt de normes que validaran la correcta construcció del fitxer xml.

Fitxer format: Contindrà el conjunt de normes que transformaran el fitxer xml resultant de la traducció en codi html.

Fitxer html: Contindrà la sortida del contingut del fitxer xml segons el format establert al Fitxer format.

Fitxer de configuració: S'utilitzarà per emmagatzemar paràmetres de l'aplicació i permetre variacions al format de logs tractats per l'aplicació.

Tot i que el format dels logs d'entrada de l'aplicació serà preestablert i obeirà a un conjunt de normes fixades, es preveu l'opció de realitzar lleugeres variacions a nivell d'administració per permetre l'adaptació de l'aplicació a canvis en els fitxers logs d'entrada.

El format del fitxer podria ser quelcom com ara:

- Paràmetres generals:
- Directoris per defecte dels fitxers que usa l'aplicació.
- Limitadors (Ex: màxim n^o de línies del fitxer log d'entrada)
- Identificació del Fitxer Esquema.
- Identificació dels Fitxer Format
- Paràmetres del fitxer log (delimitadors, n^o de camps, noms etc...)

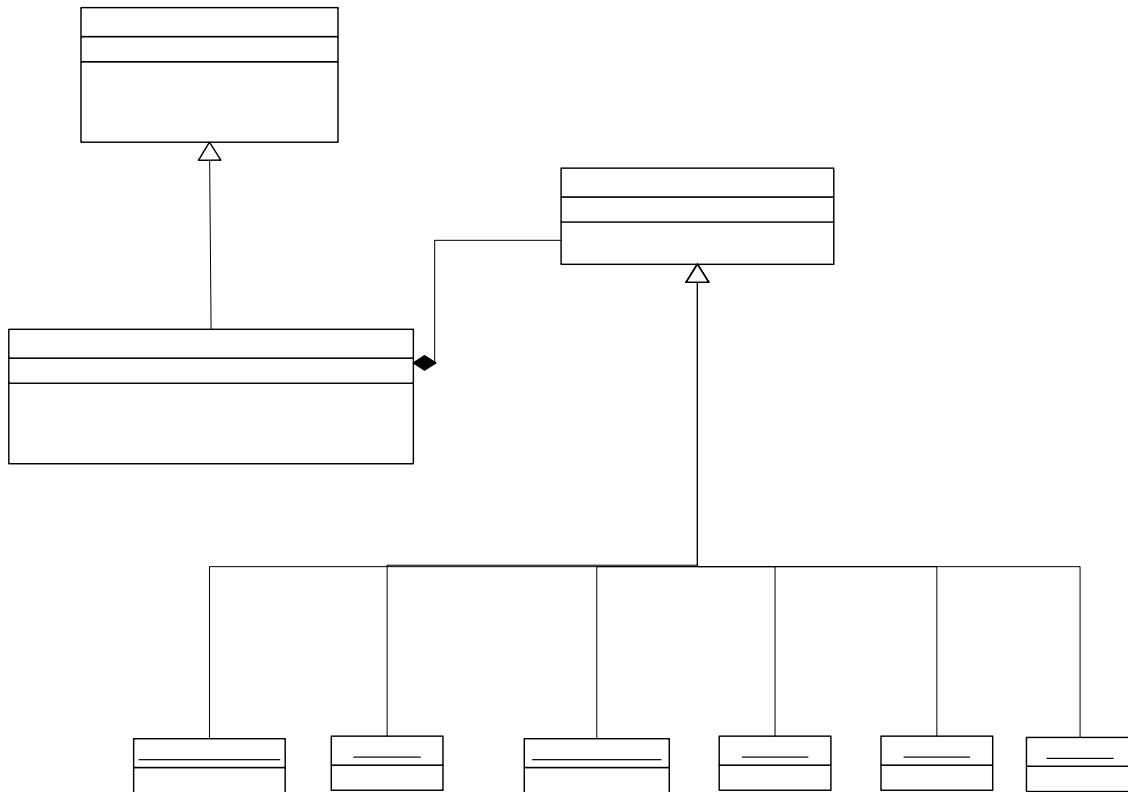
La classe *Fitxer* contindrà les operacions bàsiques de un fitxer de text, que ens permetran gestionar el seu contingut:

- llegirLinia():String
- escriureLinia():String

La gestió de la persistència d'aquests objectes s'implementarà en base a operacions d'una classe *GestorDisc* filla de la classe abstracta estàndard *GestorGenèric*. Aquesta classe ens

permetrà implementar les operacions bàsiques per manejar la persistència de fitxers de text al disc de la màquina on s'executa l'aplicació:

- crear un fitxer nou a disc per escriptura
- esborrar un fitxer de disc
- obrir un fitxer per lectura
- visualitzar la llista de fitxers existents a un directori del disc



3.7 Disseny de la interfície gràfica d'usuari

GestorGenèric

3.7.1 Guia d'estil

Es seguirà una guia d'estil basada en el model Campus UOC.

No s'usarà cap entorn gràfic pel desenvolupament de la interfície.

Degut a la quantitat de línies de text html que cal incorporar al codi del programa per la generació d'aquests elements, es definirà un estil senzill i sobri.

La pàgina principal de l'aplicació tindrà funcions de portal de documentació al mateix temps que contindrà el menú principal d'inici de la mateixa.

+crear()

+gravar()

+llegir(entrada id) : Object

+materialitzar(entrada id) : Object

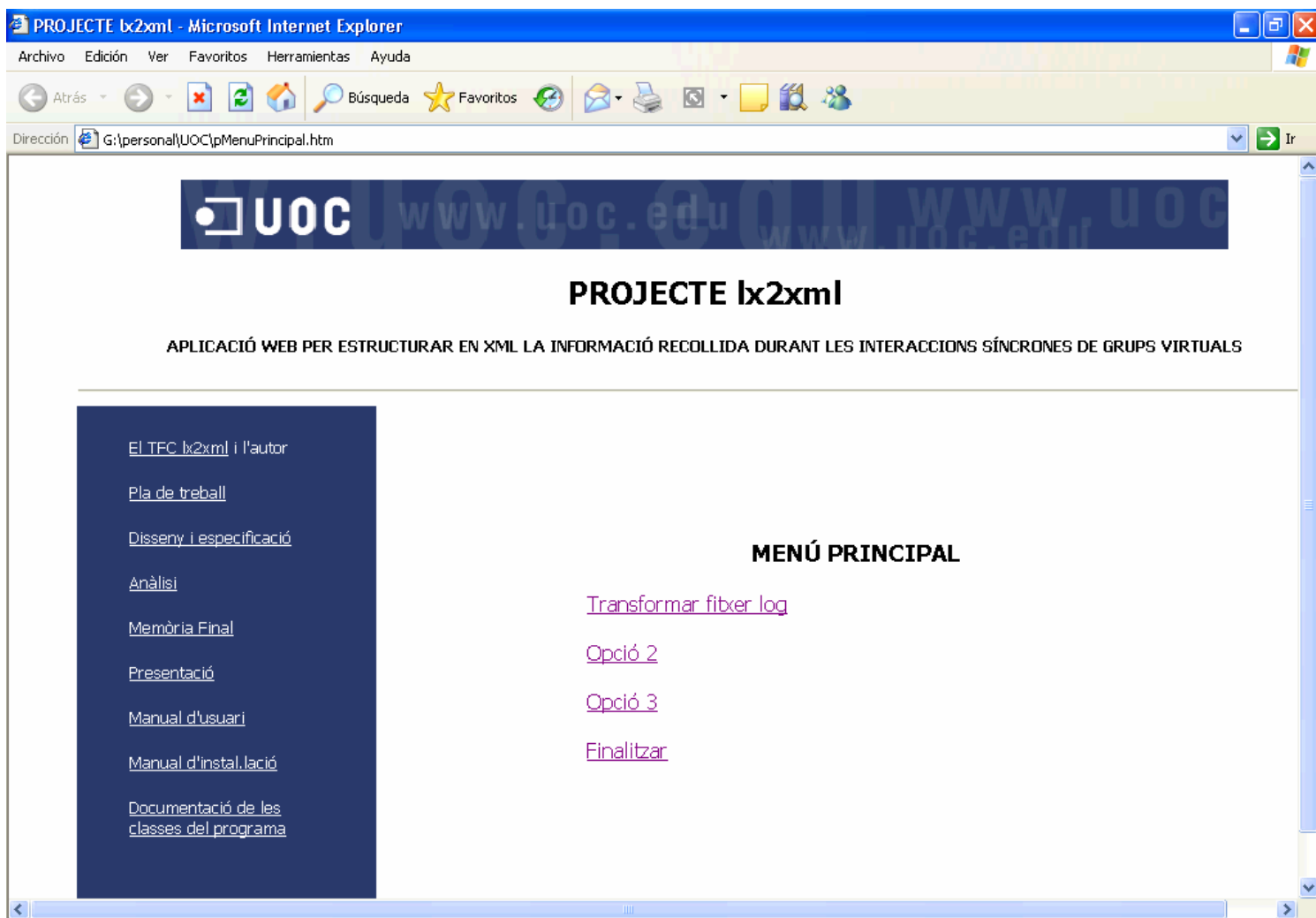
- Les dades de tipus textual es presentaran en forma d'àrees de text amb barra de desplaçament vertical.
- Els camps de longitud fixa es presentaran com a camp de text quan siguin paràmetres de sortida. En el cas dels paràmetres d'entrada, si el nombre de valors possibles és reduït, s'utilitzaran controls "llista" (combo o llista limitada).
- Per les dades booleanes es faran servir checkboxes.
- Per la tria de una opció entre varies s'utilitzaran conjunt de botons de ràdio.

- Els missatges informatius (informatius, error o advertiment) tindran el format típic de finestra de text amb el missatge corresponent i disposaran dels botons (acceptar, cancel·lar, o ambdós) segons correspongui.

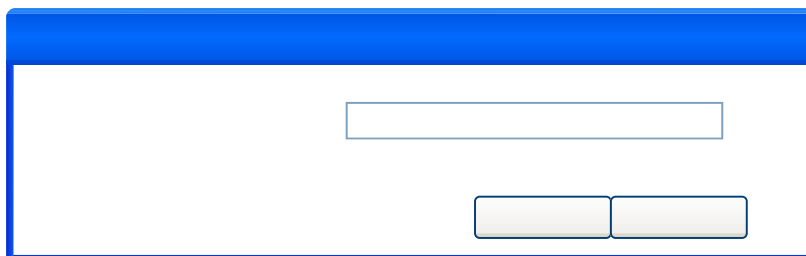
3.7.2 Disseny

A continuació presentem els prototips corresponents al disseny de la interfície del programa. Cal advertir que les possibilitats de l'eina de prototipatge emprada no permetia simular l'estil final que tindrà aquesta interfície.

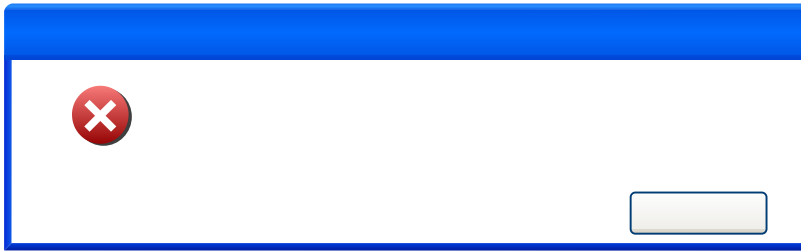
Pantalla PmenuPrincipal



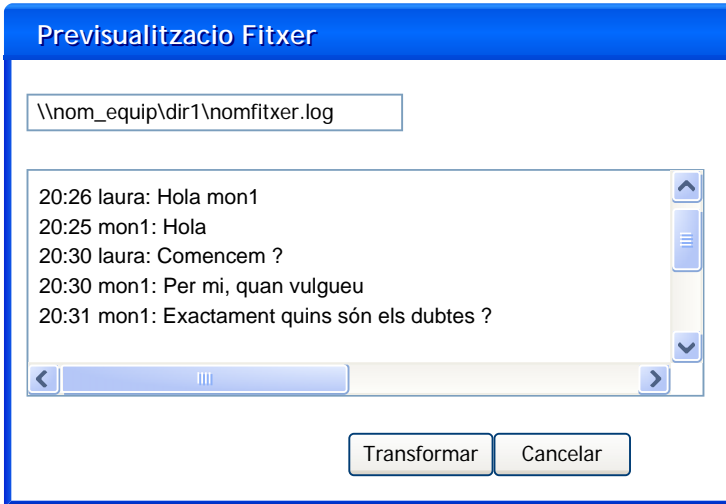
PantallaSeleccioFitxer:



PantallaErrorFitxerNoTrobat:

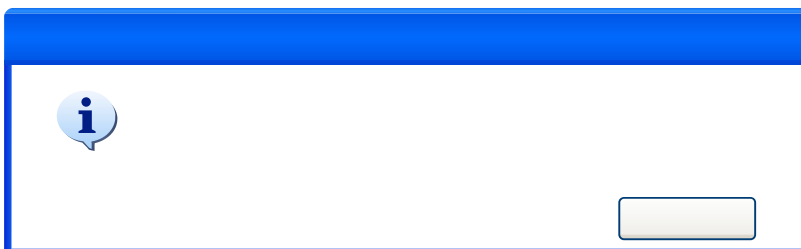
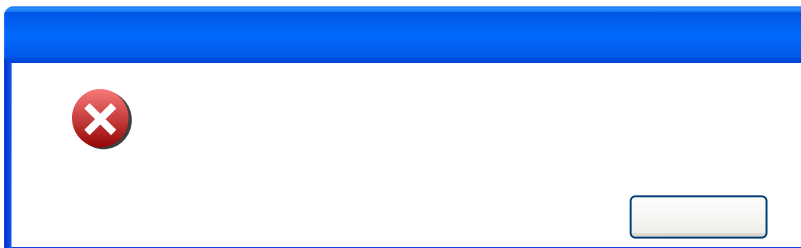


PantallaPrevisualFitxer:



er no trobat

PantallaResultatValidacio:



PantallaSeleccioSortida:



PantallaSortidaHtml:

ATENCIÓ !
El nom del fitxer
Prem "Aceptar"



ida html

Sortida fitxer.log ordenat per

3.7.3 Ajut i documentació

No es contempla la creació d'un mòdul d'ajut en línia ja que es considera que l'aplicació no es prou complexa a nivell d'interfície i funcionalitat com per fer necessària aquesta eina.

Tal i com s'ha indicat al punt 3.7.2 la pàgina html inicial farà de "portal" de l'aplicació, proporcionant links a tot l'entorn informatiu del programa:

- **Link el TFC i l'autor:** Accés a un breu currículum de l'autor i informació general sobre el TFC.
- **Links documentació:** Contindrà un accés a tota la documentació del TFC (pla de treball, especificacions, anàlisi, disseny, presentació etc..).
- **Links documentació classes del programa:** Documentació "JavaDoc" sobre les classes del programa.
- **Link manual d'usuari:** Contindrà un breu manual d'usuari sobre la utilització de l'aplicació.
- **Link instal·lació:** Contindrà un manual d'instal·lació de l'aplicació (també proporcionat amb format text en el paquet del programa).

USUARI	HORA	TIPU
laura	20:26	S
laura	20:30	P
mon1	20:25	S
mon1	20:30	R
mon1	20:31	P
dubtes ?		
mon1	20:32	R
manel	20:29	S
retard.		
manel	20:30	R
manel	20:31	P
tota la cosa ?		

3.8 Eines, tecnologies i entorns de desenvolupament

A l'etapa d'anàlisi apuntàvem entorns, eines i versions que s'estaven avaluant, ara, a l'etapa de disseny ja les tenim concretades. Eines que, si no es detecta cap mancança o impediment a l'etapa de desenvolupament, seran les definitives:

3.8.1 Tecnologia client-servidor

- Servidor d'aplicacions Apache Tomcat (Projecte Jakarta) v 5.0.28 amb tecnologia web d'implementació de servlets v 2.4
- El suport per la validació del fitxer xml l'incorpora Xerces Java Parser 1.4.4 release que serà cridat des de JDOM 1.0
- Per la transformació de xml a html s'utilitzarà JDOM 1.0 i JAXP

3.8.2 Llenguatge i entorn de desenvolupament

- J2SE (Java 2 Platform Standart Edition) v 1.4.2
- JAXP (Java Api for XML)
- XSLT 1.0
- JDOM 1.0

Pel que respecta a la tria de la tecnologia que ens permetrà tractar i generar el document xml, teníem dues opcions:

DOM (Document Object Model): El model DOM , tracta l'estructura xml com un arbre. Aquest tractament permet disposar d'unes funcionalitats que un tractament seqüencial no permet (insertar, esborrar i cercar elements dins l'estructura sense tenir que refer-la de nou). El model DOM carrega a memòria l'estructura xml, això podria ser un inconvenient en cas d'estructures molt grans (en el nostre cas, fitxers de log molt grans).

SAX (Simple API for XML): Aquest model, tracta el xml com una estructura d'accés seqüencial. És una API que té com a avantatge la seva senzillesa i la reducció del consum de memòria (ja que només carrega parts del document a memòria). Com a desavantatges, destaquem la limitació en les operacions que requeririen un tractament no seqüencial de l'estructura.

JDOM és una API de codi obert basada en el model DOM creada en i per Java pel tractament, manipulació, creació i serialització de documents xml. JDOM va ser creat per Brett McLaughlin i Jason Hunter l'any 2000.

Tot i que les operacions que requerirem fer amb la nostra aplicació no són d'especial complexitat i podrien ser satisfetes per l'API SAX, hem escollit JDOM pels següents motius:

- La facilitat i intuïtivitat de les operacions que aquesta API incorpora.
- La utilitat de funcions auxiliars (ex: mètodes de chequeix que col·laboren en la creació d'un document xml ben format).
- La possibilitat d'incorporar noves funcionalitats bassades en aquesta API en un futur.

3.8.3 Eines

Desenvolupament:

- Desenvolupament en les pròpies eines del J2SE amb l'entorn integrat de desenvolupament NetBeans IDE v 4.1.

Documentació:

Microsoft Word 2002
Microsoft Visio 2003 Professional Edition
Microsoft Power Point 2002
Adobe Acrobat

Entorns de proves (servidor):

Microsoft Windows XP Professional
Microsoft Windows 2000 Server
Linux Knoppix 3.4 Debian

Entorns de proves (clients):

Microsoft Windows XP Professional (browser: Internet Explorer v 6.0)
Linux Knoppix 3.4 Debian (browser: Mozilla Firefox)

3.8.4 Consideracions tècniques

En base a la tecnologia triada, esmentem algunes qüestions tècniques que seran dutes a terme durant la fase d'implementació.

La captura de paràmetres als servlets s'efectuarà sobreescrivint els mètodes *doPost()* o *doGet()* heretats de la classe *HttpServlet()* i utilitzant el mètode *getParam()*.

L'enviament d'informació al client s'efectua a través del canal de sortida del servlet *HttpServletResponse*.

La construcció de l'arbre jdom document es realitzarà insertant elements (parells camp/valor) a l'estructura amb les funcions *element(String)*, *element.setText(String)* i *element.addContent(element)* corresponents a la classe *Element* que ens permetran anar creant l'estructura d'arbre xml a memòria.

La validació del document xml amb esquema xsd es realitzarà amb el parser que incorpora Xerces i les utilitats per cridar-lo des de JDOM en el moment de construir el document:

- s'instancia un objecte *SAXBuilder("path del xerces", true)* Amb el valor "true" indiquem que desitgem activar la validació en el moment de construir el document.
- *builder.setFeature("s'indica que es valida amb esquema")*
- *builder.setProperty("es defineixen propietats del esquema, ubicació, nom etc..")*
- *Document doc=builder.build(fitxerXml)*
- Interceptant les excepcions de *JDOMException* obtenim el resultat de la validació i informació sobre els possibles errors apareguts (línia, tipus etc...)

La transformació xslt es farà aprofitant el suport de les API's JDOM i JAXP (paquets de transformació):

- s'instancia un objecte *Transformer transformer = TransformerFactory.newInstance().newTransformer(new StreamSource(fitxerXslt))*
- s'instancien un parell d'objectes *JDOMResult, JDOMSource(fitxerXml)*
- *transformer.transform(source, result)*
- *Document doc2 = result.getDocument()* Ja tenim el resultat a memòria, ara podem enviar-lo cap el canal de sortida que desitgem (en el nostre cas, serà el canal de resposta del servlet).

4 Implementació

4.1 Introducció

En el present capítol, documentarem el procés d'implementació del projecte. En base a la darrera revisió del disseny i les tecnologies triades hem implementat l'aplicació obtenint el producte, documentació i resultats que a continuació exposem.

- **Producte:** aplicació lx2xml publicada a <http://lx2xml.no-ip.org:8084/lx2xml/lx2xml/index.jsp> i entregat i paquetitzat (fonts i executables) segons requeriments a l'aula de TFC en la data de presentació del projecte.
- **Documentació:** Documentació de les classes generada amb JavaDoc. Disponible a <http://lx2xml.no-ip.org:8084/lx2xml/lx2xml/index.jsp> (opció capçalera JavaDoc).
- **Documentació addicional:** Manual d'usuari, manual d'instal·lació, memòria (aquest document). Disponible a <http://lx2xml.no-ip.org:8084/lx2xml/lx2xml/index.jsp> (opcions capçalera).

4.2 Consideracions al disseny

4.2.1 Diagrama de disseny de classes

Les etapes d'anàlisi i disseny s'han revisat en base als canvis apareguts durant l'etapa d'implementació.

Al diagrama de disseny de classes, [veure diagrama](#), s'ha redefinit en base a la tecnologia d'implementació triada (jsp, javaBeans) com a objectes amb les següents característiques:

- Funció principal mostrar informació a l'usuari. Per això, tots implementen el mètode `mostrarPantalla()`
- Capacitat per emmagatzemar variables (capacitats paràmetres html).
- Capacitat per executar codi (scriptless de java i javascript).
- Capacitat per realitzar crides a altres objectes pantalla (crides html href)
- Capacitat per relacionar-se amb classes java amb `useBean`

4.2.2 Ús de jsp i javaBeans

Amb la finalitat de facilitar la implementació del codi s'ha optat per l'ús de la tecnologia Java Server Pages que permet alternar codi html i scriptless java dins d'una pàgina. D'aquesta manera s'evita treballar amb servlets que obliguen a dirigir tota la entrada / sortida cap els canals request / response en forma de strings.

JSP però, pot provocar l'aparició de pàgines poc estructurades de difícil manteniment amb la barreja de codi html i java simultàniament. Per mantenir el nivell de presentació el màxim d'independent del nivell d'execució hem utilitzat la tecnologia JavaBeans que permet definir objectes independents als jsp's amb unes propietats que poden ser modificades i consultades en qualsevol moment del codi html.

4.2.3 Funcionalitats al codi versus funcionalitats a les tecnologies xslt, xsd, xml

Ja que aquestes tecnologies són de per si uns entorns suficients per desenvolupar-hi funcionalitats i representen un standard reutilitzable i conegut, durant l'etapa d'implementació s'ha presentat varies vegades el dubte de implementar determinades funcionalitats a través del

codi java de les classes o jsp's o bé, aprofitar la potència de les eines addicionals amb les que l'aplicació ja tenia contacte establert.

Finalment, tenint en compte criteris de simplicitat i per mantenir el codi el màxim de generalista, s'ha optat per implementar el màxim de funcionalitats a les eines auxiliars.

Així doncs, s'ha deixat la major part del treball de traducció a l'entorn de les expressions regulars, que esdevenen el nucli de la funcionalitat de l'aplicació. La seva potència ens ha permès configurar el patró per detectar i separar els camps del log del xat TFC-UOC especificat, però al mateix temps, ens obren la porta a poder capturar infinitats de tipus de logs diferents tan sols variant una cadena String al fitxer de configuració de l'aplicació.

Una decisió similar s'ha pres en l'etapa de validació, que es podia haver deixat en un 90% sota la responsabilitat de codi que comprovés les intervencions abans de ser introduïdes al arbre xml. En el seu lloc, s'ha deixat tota la feina per l'esquema xsd, el qual, si es dota de tota la informació necessària, pot arribar a filtrar i validar amb una efectivitat immillorable, mantenint-nos una vegada més independents del codi.

El cas de la sortida formatada html, tot i que per motius de temps no va poder ser implementada tal i com es pretenia inicialment, també representa un avantatge al deixar tota la tasca en mans del full xslt que permet les ordenacions necessàries i en futures millores podia ser el responsable de tasques com la visualització de la data (originalment en format hora Unix) en un format de calendari Gregorià.

4.3 Detalls de la implementació

Passem a llistar les classes més importants de l'aplicació. Al tractar-se de codi comentat, considerem que no s'escau afegir més comentaris. Per ampliar detalls de les funcionalitats, veure JavaDoc.

4.3.1 Mètode doTradueix (classe Transforma_xml.java)

```
public void doTradueix(String nomFitxerLog, String nomFitxerXml,
    TipusLog tpLog, String dirFitxers, String nomFitLinReb) throws
IOException,FileNotFoundException,JDOMEException {

    // obrim fitxer log d'entrada
    BufferedReader fln = new BufferedReader(new
InputStreamReader(gDisc.obrirFitxer(dirFitxers+"/"+nomFitxerLog));
    // creem un fitxer de sortida xml. Si ja existeix l'esborra
    PrintWriter fLinRebutjades = new PrintWriter(new FileWriter(dirFitxers+"/"+nomFitLinReb));
    String linia=null;
    int contLinies=0;

    // creem un element de l'arbre jdom (el primer)
    Element root=new Element("log");

    // li afegim caracteristiques de capçalera de fixter xml (tipus, namespace i el nom del fixter
    // xsd schema que el validara
    Namespace xsi=Namespace.getNamespace("xsi", "http://www.w3.org/2001/XMLSchema-instance");
    root.addNamespaceDeclaration(xsi);
    root.setAttribute("noNamespaceSchemaLocation", "file:/"+dirFitxers+"/"+tpLog.getNomFitxerXsd(),xsi);

    //creem un document jdom
    Document doc=new Document(root);

    // creem un patró d'expressions regulars amb l'atribut que defineix les expressions regulars del
    // tipus de log actiu
    Pattern patro = Pattern.compile(tpLog.getRegExp());
    Pattern patroLinEspecial = Pattern.compile(tpLog.getRegExpLinEspecial());

    //Comença el proces de transformació
    // Mentre llegim linies al fitxer log
    while ((linia=fln.readLine())!=null){
```

```

// Si el contingut es correspon al patro de línies especial es van capturar
//i es van introduint a partir de l'element arrel
contLinies++;
Matcher matchLinEspecial=patroLinEspecial.matcher(linia);
if (matchLinEspecial.find()){
    setNumLinEspecial(getNumLinEspecial()+1);
    Element elem2=new Element("linia_especial");
    elem2.setText(linia);
    root.addContent(elem2);
}

// si el contingut es correspon al patro de intervencions es crea un element intervencio i s'inserta a
// sota l'element root.
else {
    Matcher match=patro.matcher(linia);
    if (match.find()){
        setNumIntervencionsOK(getNumIntervencionsOK()+1);
        Element elem=new Element("intervencio");
        root.addContent(elem);

        // per cada camp de la intervencio, es crea un element camp
        // anem agafant els grups del patro, s'assigna el seu contingut a l'element camp
        // i s'assigna el seu nom a l'atribut nomCamp de l'element camp.
        for (int i=1;i<=tpLog.getNombreCamps();i++) {

            Element elem1=new Element("camp");
            elem1.setAttribute("nom", tpLog.getCamp(i-1).getNomCamp());
            elem1.setText(match.group(i));
            // s'inserta l'element camp sota l'element intervencio
            elem.addContent(elem1);
        }
    } else {
        // Si la linia no ha coincidit amb el patro definit, es rebutja i s'introdueix al fitxer
        // de línies rebutjades.
        setNumLinError(getNumLinError()+1);
        fLinRebutjades.println(("Linia "+String.valueOf(contLinies)+" :"+linia));
    }
}
}

// tanquem fitxers
fLinRebutjades.close();
fln.close();

// assignem un atribut per que el document xml tingui un aspecte "elegant"
Format format=Format.getPrettyFormat();

// volquem l'arbre al fitxer xml creat mitjançant un objecte XMLOutputter (JDOM)
XMLOutputter outputter = new XMLOutputter(format);
FileOutputStream fitXml=gDisc.crearFitxer(dirFitxers+"//"+nomFitxerXml);
outputter.output(doc,fitXml);

// tanquem fitxer xml
fitXml.close();
}

```

4.3.2 Mètode doValida (classe Transformar_xml.java)

```

/**
 * Metode que valida la correctesa xml del fitxer generat segons full esquema "xsd schema" contingut
 * a la capçalera del fitxer xml que es tracta.
 * Treballa amb les API's de Xerces ver 1.4.4 i objectes de SAX (SAXBuilder)
 * @param dirFitxers Valor del nom del directori de treball
 * @param nomFitxerXml Valor del nom del fitxer xml a tractar
 * @throws java.io.IOException Les excepcions tipus input/output son llençades cap a
 * metodes o jsp superiors per que siguin tractades.
 * @throws java.io.FileNotFoundException Les excepcions tipus fitxer no trobat son llençades cap
 * metodes o jsp superiors per que siguin tractades.
 * @return El valor del string que conté el resultat de la validacio
 * i qualsevol missatge d'erro resultant del proces.
 */
public static String doValidaXml(String nomFitxerXml,String dirFitxers)throws IOException,FileNotFoundException {
    String result=null;

```

```

    FileInputStream fitXml=gDisc.obrirFitxer(dirFitxers+"//"+nomFitxerXml);

    // creem l'objecte SAXBuilder amb les propietats necessaries per que validi el document amb schema
    SAXBuilder builder = new SAXBuilder("org.apache.xerces.parsers.SAXParser");
    builder.setValidation(true);
    builder.setFeature("http://apache.org/xml/features/validation/schema", true);

    // Capturem les excepcions perque necessitem saber si hi han hagut errors en el proces
    // de validacio.
    try {
        // executem la validacio. Si es correcte no es produeix cap excepcio
        builder.build(fitXml);
        result="el fitxer es valid (XML)";
        fitXml.close();

        // S'ha produït un error en la validacio. Capturem l'excepcio per mostrarla
        // A la finestra de l'aplicacio
    } catch (JDOMException e){
        result="el fitxer no es valid (XML) "+e.getMessage()+" "+e.getCause();
    }

    return(result);
}

```

4.3.3 Mètode doMostrar (classe Presentar_html.java)

```

/**
 * Aquest metode arrenca l'execucio del motor de transformació xslt
 * Utilitza les APIs de JDOM i SAX per la implementació dels objectes transformadors.
 * Crea un objecte TransformerFactory que engega la transformació basant-se en
 * el fitxer xml i el fitxer xslt passats com parametres al metode.
 * El resultat de la transformació el retorna en forma de ByteArrayOutputStream cap el
 * canal de resposta.
 * @param fitxerXml Nom del fitxer xml a tractar
 * @param fitxerXslt Nom del fitxer xslt que actuara com a plantilla de la transformació.
 * @param dirFitxers Atribut que representa el nom de la carpeta fisica de disc on s'ubiquen
 * els fitxers de treball. Juntament amb l'atribut WorkPath de la sessio, conformen
 * el path real del directory de treball.
 * @return ByteArrayOutputStream resultStream. Es retorna el ByteArrayOutputStream
 * que conté la sequencia de caracters generats pel motor de transformació.
 * Aquesta sortida va directament sobre el canal de resposta de la pagina jsp
 * que ha realitzat la crida.
 * @throws javax.xml.transform.TransformerConfigurationException Llença aquests tipus d'excepcions ocorreguts
 durant el proces de configuracio
 * del motor de transformació en cas de que alguna de les propietats configurades
 * no fos correcta.
 * L'excepcio sera recollida per la classe o jsp superior que cridi el metode.
 * @throws javax.xml.transform.TransformerException Llença una excepcio del tiups TransformerException en cas
 de que ocorreixi
 * algun error en el proces de transformació.
 * L'excepcio sera recollida per la classe o jsp superior que cridi aquest metode.
 */
public static String doMostrarHtml(String fitxerXml,String fitxerXslt,String dirFitxers)
throws TransformerConfigurationException,TransformerException{

    ByteArrayOutputStream resultStream=new ByteArrayOutputStream();

    // utilitzem dos objectes Source que contenen els fitxers xml i xslt
    Source xmlSource = new StreamSource(dirFitxers+"//"+fitxerXml);
    Source xsltSource = new StreamSource(dirFitxers+"//"+fitxerXslt);

    // Creem un objecte TransformerFactory
    TransformerFactory transFact = TransformerFactory.newInstance();
    // Li assignem el fitxer xslt
    Transformer trans = transFact.newTransformer(xsltSource);
    // Executem la transformació. Obtenim la sortida al ByteArrayOutputStream
    trans.transform(xmlSource, new StreamResult(resultStream));

    // retornem la sortida de la transformació al metode superior
    return(resultStream.toString());
}
}

```

4.3.4 Jsp triar_log

```

<!-- Jsp triar_log.jsp
@autor: Ramon Govern
Aquest jsp s'ocupa de la presentacio del formulari per recollir els parametres
d'inici de la transformacio: nomFitxerLog d'entrad i nomFitxerXml de sortida.
Crida als metodes del bean utils de la classe Lx2xmlUtils.java que contenen totes les utilitats
de l'aplicacio destinades al control, inicialitzacio i gestio de variables.
Utilitza el bean tpLog de la classe TipusLog, per amagatzemar-hi els parametres corresponents al tipus
del log que es tractara recollits de la carrega del fitxer de configuracio.
Presenta la possibilitat de cridar dos jsp addicionals que afegeixen funcionalitats addicionals:
veureFitxer.jsp : en la busqueda visualitzacio i sel.leccio del fitxer log.
SimpleUpload.jsp: en la carrega d'un fitxer del client local al servidor.
-->
<!-- Grup de directives destinades a determinar les caracteristiques de la pagina.
S'indica la pagina d'error que apareixera en cas de rebre excepcions dels metodes cridats
de les classes (tots els que no capturen les excepcions.
S'importen les classes necessaries per l'execucio del jsp
Es defineixen els beans que s'utilitzaran i es determinen els seus valors en cas que
sigui necessari.
-->
<%@page contentType="text/html"%>
<%@page pageEncoding="UTF-8"%>
<%@page import="java.io.*"%>
<%@page import="java.util.*"%>
<%@page errorPage="error.jsp"%>

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">

<html>
<jsp:useBean id="tpLog" scope="session" class="lx2xml.TipusLog" />
<jsp:useBean id="utils" scope="session" class="lx2xml.Lx2xmlUtils" />

<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<title>UOC - Projecte lx2xml</title>
</head>
<body>
<%@include file="headerPage.html"%>

<!-- fragment de codi que permet, en cas de retorn carregar el formulari amb els valors que s'havien posat
anteriorment.
-->
<%utils.initEntorn(tpLog,application.getRealPath(""));
String valorNom="";
if (request.getParameter("nomFitxerLog")!=null) {
    valorNom=request.getParameter("nomFitxerLog");
} else {
    if (utils.getFitxerLog()!=null) {
        valorNom=utils.getFitxerLog();
    }
}
%>
<!-- Inicialitzem els valors del bean per evitar que es processin en cas d'enviar el formulari
sense valors
-->
<jsp:setProperty name="utils" property="fitxerLog" value="<%=null%>" />
<jsp:setProperty name="utils" property="fitxerXml" value="<%=null%>"/>

<!-- Visualitzem el formulari amb cuadro informatiu dels valors del tipus de log actiu
i tots els controls necessaris per la captacio de dades
-->
<br>
<table bgcolor="white" >
<form method="post" title="Arrenca el procés de transformació" action="transformar_xmlFrames.jsp" >
<tr bgcolor="#EDED" style="background-color:#EDED;color:blue;font-size:10pt;font:arial" >
<td colspan="4">
PARAMETRES DEL TIPUS DE LOG: <jsp:getProperty name="tpLog" property="nomTipusLog"/><br>
<jsp:getProperty name="tpLog" property="descripcio" /><br>
Nombre de Camps: <jsp:getProperty name="tpLog" property="nombreCamps"/> Nom dels camps:
<%=tpLog.getNomCampstoString()%><br>
Expressió regular: <code><big><jsp:getProperty name="tpLog" property="regExp" /></big></code><br>
Exemple: <jsp:getProperty name="tpLog" property="exemple" /><br>
Expressió regular línies especials: <jsp:getProperty name="tpLog" property="regExpLinEspecial" /><br>
    
```

```

        <br>
      </td>
    </tr>
    <tr>
      <td width="5%">
        Nom del Fitxer Log
      </td>
      <td width="50%">
        <input type="text" name="nomFitxerLog" size="60" title="Introduir el nom del Fitxer log a processar"
value=<%=valorNom%>>
      </td>
      <td width="15%" s>
        <a href="escullFitxerFrames.jsp" title="Veure els fitxers del directori de treball"
style="font:arial;color:white;background-color:#29396b" > Veure Fitxers </a>
      </td>
      <td width="30%">
      </td>
    </tr>
    <tr>
      <td>
        Nom del Fitxer Xml (sortida):
      </td>
      <td>
        <input type="text" title="Introduir el nom del Fitxer XML resultant" name="nomFitxerXml" size="60"
value=<%=utils.canviaExtensio(valorNom)%> >
      </td>
      <td colspan="2">
        <a href="SimpleUpload.jsp" title="Permet carregar fitxers locals a servidor" style="font:arial;
color:white;background-color:#29396b"> Carregar fitxer al servidor </a>
      </td>
    </tr>
    <tr>
      <td>
      </td>
      <td>
      </td>
      <td>
        <input align="right" type="submit" style="color:white;background-color:#29396b" value="Transformar a
XML">
      </td>
    </tr>
  </form>
</tr>
<tr>
  <td>
    <form method="get" action="index.jsp">
      <input align="left" type="submit" title="Torna al menú principal" style="color:white;background-
color:#29396b" value="Tornar al Menú Principal">
    </form>
  </td>
</tr>
</table>
</body>
</html>

```

4.3.5 Jsp transformar_xml

```

<!-- Jsp transformar_xml.jsp
@author: Ramon Govern
Aquest jsp s'ocupa del control del procés de traducció i validació mitjançant les crides al bean
transforma de la classe Transforma.java.
Paralelament utilitza els beans utils de la classe Lx2xmlUtils i tpLog de la classe TipusLog per la gestió
dels valors amagatzemats en aquesta classe.
Presenta per pantalla els missatges corresponents al desenvolupament del procés.
Al finalitzar, permet a l'usuari la visualització de dos fitxers resultants del procés:
El fitxer xml generat i el fitxer que conte les línies que han estat rebutjades pel procés de
transformació.
-->
<!-- Grup de directives destinades a determinar les característiques de la pàgina.
S'indica la pàgina d'error que apareixerà en cas de rebre excepcions dels mètodes cridats
de les classes (tots els que no capturen les excepcions.
S'importen les classes necessàries per l'execució del jsp
Es defineixen els beans que s'utilitzaran i es determinen els seus valors en cas que
sigui necessari.
-->
<%@page contentType="text/html"%>

```

```

<%@page pageEncoding="UTF-8"%>
<%@page errorPage="error.jsp"%>
<%@page import="java.util.*, lx2xml.*, java.io.*"%>

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">

<jsp:useBean id="utils" scope="session" class="lx2xml.Lx2xmlUtils" />
<jsp:useBean id="tpLog" scope="session" class="lx2xml.TipusLog" />
<jsp:useBean id="transforma" scope="session" class="lx2xml.Transforma" />

<!-- Inicialitzem aquests beans per extreure dades sobre el procés de traducció
al finalitzar
-->
<jsp:setProperty name="transforma" property="numIntervencionsOK" value="0" />
<jsp:setProperty name="transforma" property="numLinError" value="0" />
<jsp:setProperty name="transforma" property="numLinEspecial" value="0" />

<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<title>UOC - Projecte lx2xml</title>
</head>
<body>
<%@include file="headerPage.html" %>

<!-- Iniciem el proces de transformació. Anem treien dades informatives
pel seguiment de l'estat
-->
<br>
<TEXTAREA NAME="log" COLS="140" ROWS="10" style="align:left;color:blue;font:courier new;font-size:8pt" >
Processant el fitxer: <%=utils.getFitxerLog()%>
<%

transforma.doTradueix(utils.getFitxerLog(),utils.getFitxerXml(),tpLog,utils.getDirFitxers(),utils.getNomFitxerLinRebutjades
s());
%>
Segons el tipus de log: <%=tpLog.getNomTipusLog()%>
Gravant sortida al fitxer: <%=utils.getFitxerXml()%>
Resultat de la transformació
*****
Nombre de línies processades:
<%=transforma.getNumIntervencionsOK()+transforma.getNumLinEspecial()+transforma.getNumLinError()%>
Nombre de línies acceptades: <%=transforma.getNumIntervencionsOK()%>
Nombre de línies especials: <%=transforma.getNumLinEspecial()%>
Nombre de línies no acceptades: <%=transforma.getNumLinError()%>
Validant fitxer: <%=utils.getFitxerXml()%>
Resultat de la validació
*****
El resultat de la validació es:
<%String s=transforma.doValidaXml(utils.getFitxerXml(),utils.getDirFitxers());
out.println(s);%>
</textarea>
<!-- Hem cridat al metode que valida el fitxer XML. El seu retorn es entrega qualsevol
missatge d'error (Exception) de la validació.
-->
<table>
<tr>
<td align="left">
<form method="get" action="veureFitxer.jsp" target="f5">
<input type="hidden" name="nomFitxer" value="<%=utils.getNomFitxerLinRebutjades()%>">
<input align="left" title="Fitxer amb les línies rebutjades durant el procés" type="submit"
style="color:white;background-color:#29396b" value="Veure fitxer línies no acceptades">
</form>
</td>
<td align="left">
<form method="get" action="veureFitxer.jsp" target="f5">
<input type="hidden" name="nomFitxer" value="<%=utils.getFitxerXml()%>">
<input align="left" type="submit" title="Veure el contingut del fitxer XML generat"
style="color:white;background-color:#29396b" value="Veure fitxer xml">
</form>
</td>
<td align="right">
<form method="get" action="index.jsp" target="_top">
<input align="left" title="Tornar al menu principal" type="submit" style="color:white;background-
color:#29396b" value="Tornar al Menu Principal">

```

```

        </form>
    </td>
    <td align="left">
        <form align="left" method="get" action="triar_log.jsp" target="_top">
            <input type="submit" title="Tornar a la pantalla Triar Log" style="color:white;background-color:#29396b"
value="Triar Log">
        </form>
    </td>
    <td align="left">
        <form method="get" action="presentar_html.jsp" target="_top">
            <input type="hidden" name="validat" value="<%=s%>" >
            <input title="Activa la presentacio html" type="submit" style="color:white;background-color:#29396b"
class="button" value="Presentar Sortida Html">
        </form>
    </td>
</tr>
</table>
</body>
</html>

```

4.3.6 Jsp presentar_html.jsp

```

<!-- Jsp presentar_html.jsp
@autor: Ramon Govern
Aquest jsp s'ocupa del control del procés de presentacio del contingut del fitxer xml en format html
Utilitza el bean presenta de la classe Presenta.java que conte els metodes destinats a la gestio de la
transformació xslt.
Utilitza els beans utils i tpLog per la gestio i manteniment dels atributs allotjats a la classe TipusLog
La pagina presenta un control tipus select destinat a la tria de les opcions de presentacio html, corresponent
a la sel.leccio d'un dels fitxers xslt que es llença com a font de la transformació.
-->
<!-- Grup de directives destinades a determinar les caracteristiques de la pagina.
S'indica la pagina d'error que apareixerà en cas de rebre excepcions dels metodes cridats
de les classes (tots els que no capturen les excepcions.
S'importen les classes necessaries per l'execucio del jsp
Es defineixen els beans que s'utilitzaran i es determinen els seus valors en cas que
sigui necessari.
-->
<%@page contentType="text/html"%>
<%@page pageEncoding="UTF-8"%>
<%@page import="java.util.*, lx2xml.*, java.io.*"%>
<%@page errorPage="error.jsp"%>

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">

<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<title>UOC - Projecte lx2xml</title>
</head>
<body>
<%@include file="headerPage.html" %>
<jsp:useBean id="tpLog" scope="session" class="lx2xml.TipusLog" />
<jsp:useBean id="utils" scope="session" class="lx2xml.Lx2xmlUtils" />
<jsp:useBean id="presenta" scope="session" class="lx2xml.Presenta" />

<!-- Es presenta el formulari amb la informacio sobre el resultat de la validacio
per que l'usuari conegui si està treballant amb un fitxer no validat.
Es tria l'opcio de tipus de sortida i s'assigna un fitxer xslt.
-->
<%String validat=request.getParameter("validat");%>
<p style="color:blue"> Fitxer:&nbsp;&nbsp;&nbsp;<%=utils.getFitxerXml()%> &nbsp;&nbsp;&nbsp; Validacio:&nbsp;&nbsp;&nbsp;<%=validat%> </p>
<table bgcolor="white">
<tr>
<td align="center">
<form method="post">
<td width="25%" valign="top">
<select title="Sel.lecciona la opció de sortida html" name="TipusSortida" >
<option value="1" >Ordenat per usuari i data
<option value="2" >Ordenat per data
<option value="3" >Ordenat per tipus d'intervencio i usuari
<option value="4" >Llistat de linies especials
</select>
</td>
<td width="25%" valign="top">

```



```

*****
Nombre de línies processades: 40
Nombre de línies acceptades: 24
Nombre de línies especials: 16
Nombre de línies no acceptades: 0
Validant fitxer: type1_logxat_normal.xml
Resultat de la validació
*****
El resultat de la validació es:
el fitxer es valid (XML)

```

Fitxer: type1_logxat_normal.xml

```

<linia_especial>desconnectat||1105376660||2||</linia_especial>
<intervencio>
  <camp nom="idUsuari">1</camp>
  <camp nom="data">1105376705</camp>
  <camp nom="idCategoria">3</camp>
  <camp nom="contingut">Ja s'ha c aquí de nou. Per fer la PAC3 podríem fer servir l'Eclipse, no?</camp>
</intervencio>
<linia_especial>desconnectat||1105378483||4||</linia_especial>
<intervencio>
  <camp nom="idUsuari">1</camp>
  <camp nom="data">1105376705</camp>
  <camp nom="idCategoria">3</camp>

```

Resultat: S'accepten totes les línies, el fitxer es vàlid (XML). La sortida html es correcta.

Fitxer: type1_logxat_normal.xml Validació: el fitxer es valid (XML)

idUsuari	data	idCategoria	contingut
2	1105376519	1	Bona tarda Josep, com va tot?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
4	1105376613	1	Hola a tothom!
4	1105376613	1	Hola a tothom!
4	1105376613	1	Hola a tothom!
4	1105376613	1	Hola a tothom!
2	1105376640	1	Ara torno!
2	1105376640	1	Ara torno!
2	1105376640	1	Ara torno!

Es llisten les línies de control:

Fitxer: type1_logxat_normal.xml Validació: el fitxer es valid (XML)

idUsuari	data	idCategoria	contingut
connectat 1105376468 1			
connectat 1105376496 2			
desconnectat 1105376660 2			
connectat 1105376660 2			
desconnectat 1105378483 4			
desconnectat 1105378491 1			
desconnectat 1105378500 2			
desconnectat 1105376660 2			
connectat 1105376660 2			
desconnectat 1105378483 4			
desconnectat 1105376660 2			
connectat 1105376660 2			

4.4.2 Test 2: fitxer type1_logxat_errorOrdreCamps.log

Conté línies errònies i una línia amb els camps canviats d'ordre.

The screenshot shows a validation tool interface. The top panel displays the following text:

```
Nombre de línies acceptades: 24
Nombre de línies especials: 16
Nombre de línies no acceptades: 4
Validant fitxer: type1_logxat_errorOrdreCamps.xml
Resultat de la validació
*****
El resultat de la validació es:
el fitxer no es valid (XML) Error on line 39: cvc-pattern-valid: Value '1105376705' is not facet-valid with respect to pattern '[0-9]{1,2}' for type 'campUser_type'. org.xml.sax.SAXParseException: cvc-pattern-valid: Value '1105376705' is not facet-valid with respect to pattern '[0-9]{1,2}' for type 'campUser_type'.
```

Below the text are four buttons: "Veure fitxer línies no acceptades", "Veure fitxer xml", "Triar Log", and "Presentar Sortida Html".

The bottom panel shows the file name "Fitxer: lx2xml_rebutjades.log" and the following error messages:

```
Linia 8 :això provocara un error del 13
Linia 22 :això provocara un error del 13
Linia 31 :això provocara un error del 13
Linia 40 :això provocara un error del 13
```

El fitxer és processat, es descarten les línies errònies però no supera el resultat de la validació.

Nota: Si aquest fitxer es processa (modificant el fitxer de configuració) amb el schema no restrictiu, el resultat de la validació és correcte.

The screenshot shows the same validation tool interface. The top panel displays the following text:

```
Nombre de línies processades: 44
Nombre de línies acceptades: 24
Nombre de línies especials: 16
Nombre de línies no acceptades: 4
Validant fitxer: type1_logxat_errorOrdreCamps.xml
Resultat de la validació
*****
Al resultat de la validació es:
el fitxer es valid (XML)
```

The text "el fitxer es valid (XML)" is circled in red. Below the text are four buttons: "Veure fitxer línies no acceptades", "Veure fitxer xml", "Triar Log", and "Presentar Sortida Html".

The bottom panel shows the file name "Fitxer: type1_logxat_errorOrdreCamps.xml" and the following XML code:

```
</intervencio>
<linia_especial>desconnectat||1105378483||4||</linia_especial>
<intervencio>
  <camp nom="idUserari">1105376705</camp>
  <camp nom="data">1</camp>
  <camp nom="idCategoria">3</camp>
  <camp nom="contingut">AQUESTA LINIA TE L'ORDRE DELS DOS PRIMERS CAMPS CANVIAT I PROVOCA UNA NO VALIDACIO :-)</camp>
</intervencio>
<intervencio>
  <camp nom="idUserari">1</camp>
  <camp nom="data">1105376708</camp>
  <camp nom="idCategoria">4</camp>
```

El fitxer és validat tot i tenir una intervenció on els valors s'han creuat.

Nota: S'obté un resultat similar amb el fitxer type1_logxat_buit.log (que no conté cap línia).

4.4.3 Test 3 : Fitxer type1_logxat_BigSize.log

El fitxer té 2000Kb. de grandària 39899 línies. El resultat del procés és tot correcte:

```

*****
Nombre de línies processades: 39989
Nombre de línies acceptades: 23993
Nombre de línies especials: 15994
Nombre de línies no acceptades: 2
Validant fitxer: type1_lograt_bigSize.xml
Resultat de la validació
*****
El resultat de la validació es:
el fitxer es valid (XML)
    
```

Fitxer: type1_lograt_bigSize.xml

```

<camp nom="data">1105376613</camp>
<camp nom="idCategoria">1</camp>
<camp nom="contingut">Hola a tothom!</camp>
</intervencio>
<intervencio>
<camp nom="idUsuari">2</camp>
<camp nom="data">1105376640</camp>
<camp nom="idCategoria">1</camp>
<camp nom="contingut">Ara torno!</camp>
</intervencio>
<linia_especial>desconnectat||1105376660||2||</linia_especial>
<linia_especial>connectat||1105376660||2||</linia_especial>
    
```

Fitxer: type1_lograt_bigSize.xml Validació: el fitxer es valid (XML)

idUsuari	data	idCategoria	contingut
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?

L'aplicació comença a presentar problemes de memòria a partir del tractament de fitxers de 4 Mb. tot i que aquesta dada depèn d'altres paràmetres del sistema que executi l'aplicació. Els primers problemes (excepció out.of.memory) apareixen en les funcions de validació i transformació.

4.4.4 Test 4: fitxer type5_lograt.log

Fitxer de configuració:

```

nomTipus=Type5_Xat_Internet
descripcio=logs tipus estandar (extret de Internet)
exemple=[19:05] &lt;Ricardo Vaamonde&gt; Hola, soy Ricardo Vaamonde. Saludos a todos
nombreCamps=3
regExp=(.*) (?:(.*)>)? (.*)
regExpLinEspecial=^[.]
camps=data,idUsuari,contingut
fitxerXsd=lx2xml_logSchema.xsd
fitxersXslt=lx2xml_perUsuari.xsl,lx2xml_perHora.xsl
    
```

Aquest fitxer no es correspon al estàndard definit a les especificacions, però podem tractar-lo activant la configuració Type5_Xat_Internet. El fitxer conté un camp opcional. En un moment determinat del xat, apareixen línies que no contenen aquest camp. Aquesta opció està contemplada en la configuració del tipus i el resultat del procés és OK.

The screenshot shows a web application interface with three main sections:

- Top Section:** A text area displaying the results of processing a log file. It shows statistics such as 'Nombre de línies processades: 21', 'Nombre de línies acceptades: 19', and 'Nombre de línies no acceptades: 2'. Below the text are four buttons: 'Veure fitxer línies no acceptades', 'Veure fitxer xml', 'Triar Log', and 'Presentar Sortida Html'.
- Middle Section:** A text area showing the XML content of the log file. The XML structure includes elements like <intervencio>, <camp nom="data">, <camp nom="idUsuari" />, and <camp nom="contingut">. A red circle highlights a specific entry with the date [19:05] and the user Ricardo Vaamonde.
- Bottom Section:** A table displaying the processed log data. The table has three columns: 'data', 'idUsuari', and 'contingut'. A red circle highlights the entry for [19:07] with user Ricardo Vaamonde and the message 'Vale'.

En el paquet de proves hi han altres fitxers que es corresponent a diferents tipus de xats amb diferents casuístiques. Tots ells, disposen de la seva configuració desactivada (comentada amb símbols # al fitxer de configuració).

4.5 Possibles millores

A continuació, fem esment d'una sèrie de millores o funcions addicionals que en la planificació de la implementació varem considerar d'importància lleu en relació a les especificacions inicials, però que caldria tenir en compte de cara a futures versions de l'aplicació:

- Tot i que les especificacions es basen en el tractament de tipus de log generat per el xat TFC-UOC veure [Anexe-2 especificacions log xat TFC-UOC](#) i la modificació del fitxer de configuració per tractar altres tipus de log hauria de ser una acció no habitual, no es descarta la possibilitat d'afegir la possibilitat de modificar i carregar online aquestes configuracions.

- El fitxer de configuració està definit en format fitxer de propietats de java (properties). Tenint en compte l'obsolescència i limitacions d'aquest format i les possibilitats del xml, seria interessant implementar el fitxer de configuració en format xml.
- Afegir un control d'accés a l'aplicació. Validar usuari i contrasenya al inici i posteriorment validar les dades de sessió i idUsuari abans de mostrar cada plana.
- Millorar la informació que rep l'usuari al produir-se excepcions, afegint-hi informació contextual i amigable sobre possibles causes de l'error i les accions correctores a realitzar.
- No aturar el procés de validació al produir-se el primer error sinó continuar el procés recollint els errors en un objecte específic per aquesta tasca i mostrar al finalitzar l'operació tota la informació.
- Parametritzar el full xslt per que rebí els noms dels camps que l'usuari pugui introduir al afegir nous tipus de logs al fitxer de configuració.
- Recollir (actualment és present al fitxer de configuració però no es tracta) la propietat fitxers xslt (associats a un tipus de log) per cada tipus de log i presentar a l'usuari la selecció de possibles opcions de sortida en funció d'aquesta dada.
- Crear un full xslt específic pel tipus de log TFC-UOC que transformi el contingut del camp data (en format hora Unix) a un format més amigable.

5 Conclusions

Pesem que la finalització d'aquest TFC ha complert amb escriure els objectius que inicialment ens varem fixar al [Pla de Treball](#):

- Hem dissenyat una eina que permet complir amb les especificacions del pla de treball.
- Hem assolit uns coneixements i experiències amb entorns i tecnologies que fins ara i degut a la branca triada dels estudis d'Enginyeria Tècnica d'informàtica de Sistemes no havia tingut cabuda en el pla docent.
- Hem aplicat els coneixements obtinguts en algunes assignatures cursades durant la carrera tan en les etapes d'anàlisi i disseny com en l'etapa d'implementació.

Hem descobert (amb sorpresa) que l'xml i tecnologies associades (xslt, xsd) són de per sí unes eines de gran potència i projecció futura cap estàndards de comunicació entre entorns i sistemes d'emmagatzemen i estructuració de la informació. De fet, les hores invertides en navegar pels recursos utilitzats (manuais, fòrums, bibliografia, entitats...) no han fet més que fer-nos entreveure que estàvem només fent una introducció en el univers d'aquests entorns.

El contacte amb les tecnologies d'aplicacions web a la banda de servidor (tomcat, servlets, jsp...) ens han ajudat també a comprendre que hi ha al darrera del cada dia més utilitzat món Internet des d'un punt de vista que fins ara no havíem pogut conèixer durant la carrera.

Finalitzem el TFC engrescats, motivats, satisfets i també cansats. Ja que com sempre, les planificacions i temporitzacions s'estrenyen a la part final dels projectes i a darrera hora ha estat necessari un esforç personal i familiar important.

6 Agraïments

Voldríem agrair l'ajut i suport del Santi Caballé (consultor del TFC) que en tot moment ens ha proporcionat.

A totes les comunitats i persones que de forma altruista publiquen els seus coneixements i experiències i proporcionen una font inesgotable de coneixement i recursos que faciliten la realització de tasques com aquest TFC.

..i de tot cor, a la família, pel suport, la paciència i sobretot per les hores que els he deixat sols, que d'ara en endavant caldrà recuperar.

7 Bibliografia i recursos

Tot i que tots han estat de molta utilitat, destaquem amb el símbol 👍 els més visitats i interessants segons la nostra opinió.

General:

👍 Enciclopedia wikipedia
<http://es.wikipedia.org/wiki/Portada>

👍 World Wide Web Consortium
<http://www.w3c.org>

👍 Benet Campderich Falgueras
Modul Enginyeria del Programari
UOC

👍 O'reilly Media
<http://www.oreilly.com/>

Programació:

Desarrollo web: Introducción a los lenguajes del web
<http://www.desarrolloweb.com/manuales/27/>

👍 Programación en castellano
<http://www.programacion.com>

El rincón del programador
<http://www.elrincondelprogramador.com/>

Javaworld
<http://www.javaworld.com/>

Javazoom
<http://www.javazoom.net/>

Sun, tomcat...

👍 Pàgina oficial de SUN Microsystems
<http://java.sun.com/>

Pàgina oficial d'Apache
<http://tomcat.apache.org>

Pàgina oficial Xerces
<http://xerces.apache.org/>

Netbeans
<http://www.netbeans.org/kb/41/>

XML,xslt xsd schema:

Santanach Delisau F.: Lleguatsges de marques XML
Modul UodE del curs Llenguatges de Marques XML.

Vazquez, A.: Navegar en Internet: XML

W3 Schools. XML Tutorial
<http://www.w3schools.com/xml/default.asp>

Desarrollo web: Introducción a XML
<http://www.desarrolloweb.com/manuales/18/>

Maria Isabel Garcia Arenas. Curso XML Introducción.
<http://geneura.ugr.es/~maribel/xml/introduccion/index.shtml>

Eva Vilagínés Aguila
XML / XQueryLanguages: Estat de l'Art
TFC UOC 2003-2004
Biblioteca UOC

Processing XML with Java
<http://www.cafeconleche.org/books/xmljava/chapters/index.html>

Xfront
<http://www.xfront.com/>

<http://www.xml.com/>

JDOM

<http://www.jdom.org>

RegExp:

ZonaSiete
<http://www.zonasiete.org/>

Libreria de expresiones regulares
<http://regexlib.com/>

<http://osteele.com/tools/rework/>

8 Anexe 1 – Manual d'usuari

8.1 Introducció

El present manual pretén ser una guia d'ajut ràpid i eficaç a l'hora d'introduir l'usuari en l'ús de l'aplicació lx2xml.

L'aplicació lx2xml permet traduir fitxers logs provinents de l'enregistrament de sessions xat a fitxers XML .

Els fitxers traduïts, són testejats mitjançant un full xsl (xlm schema) per validar la seva correctesa.

L'aplicació permet a l'usuari obtenir en format html el contingut del fitxer xml generat segons diferents opcions de presentació:

- ordenació per usuari i data
- ordenació per data
- ordenació per tipus d'intervenció
- llistat de línies especial

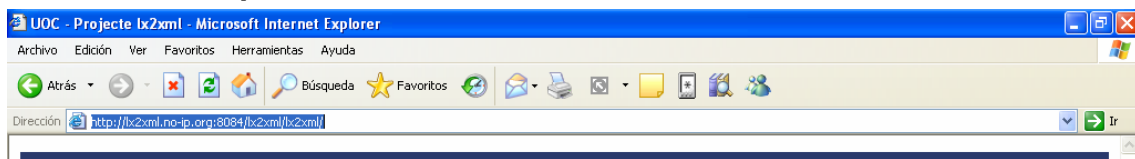
L'aplicació ve configurada inicialment per tractar logs tipificats segons les especificacions del xat TFC-UOC, una eina desenvolupada per estudiants de la UOC que permet l'execució i enregistrament de sessions xat. Veure [Especificacions Xat TFC-UOC](#).

L'edició del fitxer de configuració lx2xml_config.properties permet d'introducció de nous tipus de xat a tractar. [Veure Fitxer de configuració](#).

L'aplicació presenta dues funcionalitats auxiliars encaminades a donar suport a l'usuari:

- **Upload de fitxers des de màquina client a servidor:** Basada en codi lliure obtingut de Internet <http://www.javazoom.net/jzservlets/uploadbean/uploadbean.html> Permet pujar al directori de càrrega de fitxers qualsevol fitxer que l'usuari vulgui transformar.
- **Visualització de fitxers:** Utilitat addicional a les funcionalitats de l'aplicació que dona suport a la UI per permetre de forma fàcil la visualització, consulta i selecció dels fitxer continguts a la carpeta de treball.

8.2 Inici de l'aplicació



Accedirem a l'aplicació depenen de la instal·lació de l'entorn d'execució que estigui muntat.

- En cas d'un entorn local, hi accedirem posant al navegador l'adreça: <http://localhost:PPPP/lx2xml/lx2xml> on PPPP correspon al nº de port habilitat al servidor d'aplicacions.
- En cas d'accedir a l'entorn de proves a Internet, introduïrem la següent adreça al navegador: <http://lx2xml.no-ip.org:8084/lx2xml/lx2xml>

Nota: L'entorn de proves a Internet es basa en una infraestructura local (equip domèstic, comunicacions ADSL i serveis de DNS gratuïts per ip no fixa www.no-ip.org). Tot i que es farà el possible per mantenir aquesta infraestructura online durant el període de presentació del TFC, no es garanteix la seva disponibilitat a qualsevol hora.

8.3 Capçalera de l'aplicació

La capçalera resta present durant tota l'execució del programa i permet, en cas de necessitat, accedir a les següents funcions documentals de l'aplicació:

- Presentació: accés a la presentació virtual del TFC en format PowerPoint.
- Memòria: accés al document memòria del TFC en format Word.
- Manual d'usuari: accés a aquest document.
- JavaDoc: accés a la API de les classes del programa en format html JavaDoc.

8.4 Pantalla Menú Principal (index.jsp)



Des d'aquesta pantalla podem cridar la principal funcionalitat de l'aplicació "Transformar Log". Representa el punt d'inici del programa i és el punt de retorn en cas de detecció d'errors recuperables d'execució.

Les altres dues opcions del menú representen funcions auxiliars a la funcionalitat principal:

- **Editar fitxer de configuració:** Utilitat freeware *Copyright (C) 2003,2004, Boris von Loesch*. Permet editar el fitxer de configuració de l'aplicació `lx2xml_config.properties` per variar els paràmetres en que l'aplicació es basa per la transformació de logs. Veure *Fitxer de configuració*.
Cal dirigir-se al directori fitxers localitzar el fitxer `lx2xml_config.properties` i prémer edit. Un cop modificat el fitxer es pot seleccionar l'opció de mantenir una còpia de backup per més seguretat.

- **Finalitzar aplicació:** Permet finalitzar l'aplicació tancant la finestra i la corresponent sessió amb el servidor.

Nota: per motius de seguretat, l'opció *Editar fitxer de configuració* pot restar desactivada en l'entorn de proves d'Internet. En el seu lloc es proposarà un mètode alternatiu per poder fer proves amb diferents configuracions.

8.5 Pantalla triar_log

Aquesta pantalla ens mostra a la part superior un conjunt de dades informatives referents a la configuració dels paràmetres de traducció que en aquests moments està activa. Veure *Fitxer Log*.

Aquesta informació és:

- Nom del tipus de log definit
- Descripció
- Nombre (màxim) de camps que conté cada intervenció del log.
- Expressió regular que defineix el tipus de log. Veure *Expressions regulars*
- Exemple de possible línia d'intervenció d'un log d'aquest tipus.
- Expressió regular que defineix les línies especials (línies de control, accés al xat, desconnexió etc..)

La pantalla presenta dos camps de text:

- **Nom Fitxer Log:** On l'usuari pot introduir directament el nom del fitxer log a traduir.
- **Nom Fitxer Xml:** On l'usuari pot introduir el nom que tindrà el fitxer xml generat.

En cas de no introduir el nom del fitxer log a tractar i prémer el botó de Transformar a XML, l'aplicació mantindrà aquesta pantalla activa.

En cas d'introduir només el nom del fitxer log, al prémer el botó de Transformar a XML, l'aplicació proposa un nom estàndard (nom del fitxer log amb extensió .xml) per el fitxer de sortida. L'usuari té l'oportunitat de validar modificar aquests valors o validar-los per iniciar la traducció.

Adicionalment, es disposa de dues funcionalitats auxiliars per la interfície d'usuari:

- **Veure fitxers:** Crida a un menú que facilita la visualització del contingut del directori de treball. Per més detalls, consultar l'apartat [Veure fitxers](#).
- **Carregar fitxer al servidor:** Crida a una pantalla que permet la càrrega (upload) d'arxius ubicats a la màquina local cap al directori de treball del servidor. Per més detalls veure [Carregar fitxer al servidor](#).

8.6 Pantalla Transformar_xml

Direcció http://localhost:8084/lx2xml/lx2xml/transformar_xmlFrames.jsp Ir

UOC UOC -- Universitat Oberta de Catalunya
Projecte lx2xml (transformació de logs de xat a xml)

[Presentació](#) | [Memoria](#) | [Manual d'usuari](#) | [Manual d'instal·lació](#) | [JavaDoc](#)

Nombre de línies processades: 40
Nombre de línies acceptades: 2
Nombre de línies especials: 0
Nombre de línies no acceptades: 38
Validant fitxer: type1_logxat_normal.xml
Resultat de la validació

El resultat de la validació es:
el fitxer es valid (XML)

Veure fitxer línies no acceptades Veure fitxer xml Triar Log Presentar Sortida Html

Fitxer: type1_logxat_normal.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<log xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="file:/C:/dades\TFC\lx2xml\build\web\lx2xml\fitxers\lx2xml_logSchema.xsd">
<intervencio>
<camp nom="data">1|1105376708|4||Ara fare una "putada" a l'aplicacio que transforma els xats...</camp>
<camp nom="idUsuari">...</camp>
<camp nom="contingut">Ja ! son símbols prohibits</camp>
</intervencio>
<intervencio>
<camp nom="data">1|1105376708|4||Aquesta es forta !</camp>
```

La pantalla es carrega mostrant-nos el resultat del procés de transformació que s'està realitzant. Obtenim algunes dades sobre el procés:

- **Nombre de línies processades:** El nombre total de línies del fitxer.
- **Nombre de línies acceptades:** El nombre de línies que s'han acceptat com a intervencions correctes segons el format d'expressió regular que defineix el tipus de log. Veure [Expressions Regulars](#) i [Fitxer de Configuració](#).
- **Nombre de línies especials:** El nombre de línies que s'han identificat com a línies especials (control del xat), segons el format d'expressió regular que defineix el tipus de log. Veure [Expressions Regulars](#) i [Fitxer de Configuració](#).
- **Nombre de línies no acceptades:** El nombre de línies que no han estat acceptades com a intervencions vàlides. Veure [Expressions Regulars](#).

La pantalla ens mostra finalment el resultat de la validació del fitxer xml generat contra l'esquema corresponent al tipus de log actiu (veure [Fitxer de Configuració](#)). En cas de que el resultat de la validació sigui "no correcte", afegeix el missatge d'error (línia i motius) generat pel motor de validació.

Des d'aquesta pantalla, l'usuari pot realitzar les següents accions:

- **Veure el contingut del fitxer xml:** Es mostra el contingut del fitxer generat.
- **Veure el contingut del fitxer de línies rebutjades:** Es un fitxer que mostra el nº de línia i el seu contingut per les línies que no han estat acceptades en el procés de traducció.
- **Triar Log:** Es torna a la pantalla per procedir a la tria d'un altre fitxer.
- **Presentar sortida html:** Es passa a la pantalla següent que permetrà extreure el contingut del fitxer xml en diferents formats html.

8.7 Pantalla presentar_html

idUsuari	data	idCategoria	contingut
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376568	4	Bé. Em podràs passar els apunts de mates quan puguis?
1	1105376705	3	Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?
1	1105376705	3	Plou i fa sol :-)
1	1105376705	3	Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?
1	1105376705	3	Fa sol ? O bé creieu que hauriem d'agafar un paraigua? :-)
1	1105376705	3	Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?
1	1105376705	3	Fa sol ? O bé creieu que hauriem d'agafar un paraigua? :-)
1	1105376705	3	Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?
1	1105376705	3	Fa sol ? O bé creieu que hauriem d'agafar un paraigua? :-)

La pantalla mostra a l'usuari el resultat de la validació ja que es poden tractar fitxers que no han superat el procés de validació xml. [Veure Fitxers schema](#).

En aquesta pantalla l'usuari disposa d'un control tipus "select" per procedir a triar l'opció de sortida html del fitxer processat.

Cada opció es correspon a un fitxer xstl. Les opcions disponibles són:

- **Per usuari i data:** La sortida s'ordena alfabèticament per usuari i data.
- **Per data:** La sortida s'ordena només per data.
- **Per tipus d'intervenció i usuari:** La sortida s'ordena per tipus d'intervenció primer i per usuari en segon terme.

NOTA: Aquesta funcionalitat no ha estat parametrizada a nivell de fitxer de configuració (tot i que el paràmetre hi figura, el programa no el tracta). D'introducció de nous tipus de logs en el fitxer de configuració hauria de coincidir amb els valors dels noms dels camps dels fitxers xstl actuals (definitos pel tipus de xat TFC-UOC) que són (idUsuari, data, idCategoria, contingut).

Des d'aquesta pantalla podem tornar a la pantalla de tria d'un altre log o bé podem anar al menú principal per sortir de l'aplicació o iniciar una nova traducció de fitxer.

8.8 Veure Fitxers

lx2xml_rebutjades.log	Selecciona	Veure
type1_logSchema_restrictiu.xsd	Selecciona	Veure
type1_logxat_ambErrors.log	Selecciona	Veure
type1_logxat_ambErrors.xml	Selecciona	Veure
type1_logxat_bigSize.log	Selecciona	Veure
type1_logxat_buit.log	Selecciona	Veure
type1_logxat_buit.xml	Selecciona	Veure
type1_logxat_errorOrdreCamps.log	Selecciona	Veure

Fitxer: type1_logxat_bigSize.log

```
1|1105376705|3|Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?  
desconnectat|1105378483|4|  
1|1105376705|3|Fa sol ? O bé creieu que hauriem d'agafar un paraigues? :-)  
1|1105376568|4|Bé. Em podràs passar els apunts de mates quan puguis?  
4|1105376613|1|Hola a tothom!  
2|1105376640|1|Ara torno!  
desconnectat|1105376660|2|  
connectat|1105376660|2|  
1|1105376705|3|Ja sóc aquí de nou. Per fer la PAC3 podrém fer servir l'Eclipse, no?  
desconnectat|1105378483|4|  
1|1105376705|3|Fa sol ? O bé creieu que hauriem d'agafar un paraigues? :-)  
1|1105376568|4|Bé. Em podràs passar els apunts de mates quan puguis?
```

Aquesta pantalla permet a l'usuari visualitzar el contingut de la carpeta seleccionada com a carpeta de treball. Els fitxers es mostren ordenats alfabèticament i clicant sobre els links Seleccionar i Veure respectivament podem :

- Seleccionar el nom del fitxer com a entrada pel control quadre de text de la pantalla triar_log.
- Veure el contingut d'un fitxer qualsevol de la carpeta.

Adicionalment, l'usuari pot tornar a la pantalla triar_log sense haver realitzat cap tria.

8.9 Carregar fitxer al servidor (Copyright © [JavaZOOM](#) 1999-2005)

Selecciona el Fitxer a pujar :

Aquesta pantalla permet pujar un fitxer de la màquina client que està executant la sessió cap a la carpeta de treball del servidor.

- **Botó examinar:** es visualitza el contingut del disc local de la màquina client.
- **Botó Pujar Fitxer:** s'inicia el upload del fitxer. Al finalitzar, es mostren els bytes copiat.
- **Botó Tornar:** es retorna a la pantalla triar_log.

8.10 Pantalla d'error



Aquesta pantalla apareix quan l'aplicació detecta un error "conegut" en qualsevol punt del programa i no pot corregir-lo.

L'objectiu d'aquesta pantalla és donar el màxim d'informació sobre el tipus d'error i les possibles causes. Es mostra informació capturada pels mètodes del sistema que manegen les excepcions.

Des d'aquesta es poden realitzar les següents accions:

- **Tornar al menú principal:** per iniciar una nova transformació de fitxer o tancar l'aplicació.
- **Enviar un correu al responsable de l'aplicació:** per informa-li del problema.

Nota: Per motius de seguretat, l'opció d'enviament de correu apunta a una adreça fictícia en l'entorn de proves d'Internet.

8.11 Especificacions Xat TFC-UOC

Els fitxers logs són fitxers de tipus text.

El volum dels fitxers no pot superar els 4 Mb. (

Hi han dos tipus d'entrades al log:

- **Entrades intervenció:** corresponents a les aportacions dels usuaris al xat.
- **Entrades línies especials:** corresponents als missatges de control que l'aplicació xat enregistra al xat.

Cada entrada (intervenció o línia especial) correspon a una línia del fitxer log. El delimitador doncs, de les intervencions és el control de retorn de línia.

El nombre de camps contingut a cada intervenció és 4. Tots són obligatoris.

El ordre dels camps no pot variar.

El nom dels camps són (per ordre d'aparició a la intervenció): idUsuari, data, idCategoria, contingut.

El separador de camp es el conjunt de caràcters ||

El format de la intervenció és: idUsuari || data || idCategoria || contingut

Ex: 2||1105376519||1||Bona tarda Josep, com va tot?

La seqüència que identifica i separa una intervenció en el llenguatge d'expressions regulars és: `(.*)\\|\\|(.*)\\|\\|(.*)\\|\\|(.*)`

El contingut i tipus dels camps és el següent:

- **idUsuari:** 1 o 2 dígits del 0 al 9
- **data:** 10 dígits del 0 al 9
- **idCategoria:** 1 dígit del 0 al 9

- **contingut:** text, qualsevol caràcter a excepció de la seqüència del separador de camps
||

El format de la línia especial és : línia_especial (tot el text junt)

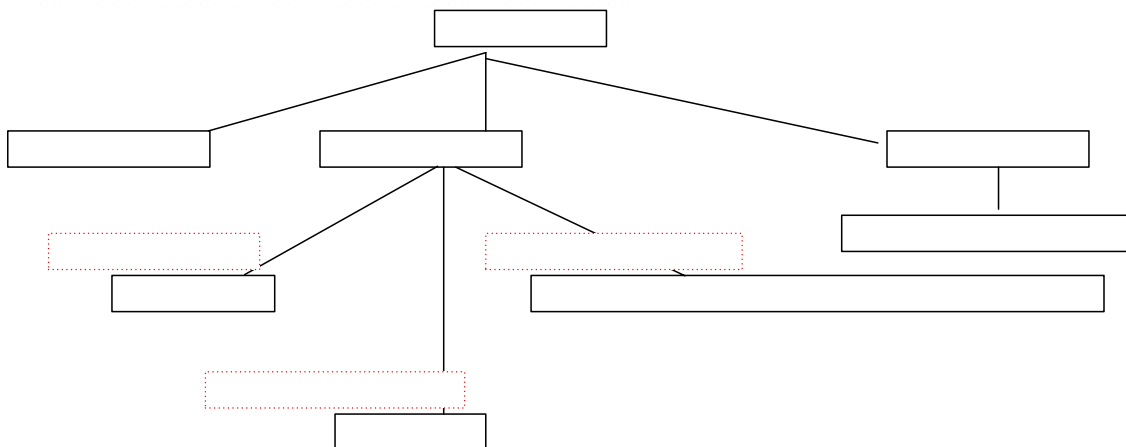
La seqüència (expressió regular) que identifica una línia especial és: connectat | desconnectat

Ex: connectat||1105376496||2||

La validació del fitxer xml es realitza amb el fitxer type1_logSchema_restrictiu.xsd i conté els paràmetres necessaris per garantir les especificacions anteriors. L'associació entre el fitxer xml i el fitxer esquema la realitza l'aplicació, però queda de forma permanent ja que el fitxer xml porta introduït el nom i localització del esquema a la seva capçalera per posteriors validacions. Per més detalls sobre el xml esquema veure el capítol bibliografia de la memòria. Veure [Fitxers Schema](#).

8.12 Fitxer xml

L'aplicació genera un fitxer amb estructura xml seguint els criteris següents:



A continuació, es mostra un fragment d'un fitxer xml generat per l'aplicació:

```
<?xml version="1.0" encoding="UTF-8"?>
<log xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="file:/C:/dades/TFC/lx2xml/build/web/lx2xml/fitxers/type1_logSchema_restrictiu.xsd"
>
<linia_especial>connectat||1105376468||1||</linia_especial>
<linia_especial>connectat||1105376496||2||</linia_especial>
<intervencio>
<camp nom="idUserari">2</camp>
<camp nom="data">1105376519</camp>
<camp nom="idCategoria">1</camp>
<camp nom="contingut">Bona tarda Josep, com va tot?</camp>
</intervencio>
<intervencio>
<camp nom="idUserari">1</camp>
<camp nom="data">1105376568</camp>
<camp nom="idCategoria">4</camp>
<camp nom="contingut">Bé. Em podràs passar els apunts de mates quan puguis?</camp>
</intervencio>
<intervencio>
<camp nom="idUserari">4</camp>
<camp nom="data">1105376613</camp>
```



```
<camp nom="idCategoria">1</camp>
<camp nom="contingut">Hola a tothom!</camp>
</intervencio>
<intervencio>
  <camp nom="idUsuari">2</camp>
  <camp nom="data">1105376640</camp>
  <camp nom="idCategoria">1</camp>
  <camp nom="contingut">Ara torno!</camp>
</intervencio>
<linia_especial>desconnectat||1105376660||2||</linia_especial>
<linia_especial>connectat||1105376660||2||</linia_especial>
</intervencio>
```

8.13 Expressions Regulars

El llenguatge de les expressions regulars defineix la seqüència de caràcters que identifiquen les intervencions i separen el seu contingut correctament.

Tot i que l'abast sobre les funcionalitats i aplicacions d'aquest llenguatge queda molt lluny de l'objectiu d'aquest manual (veure *bibliografia*), afegirem alguns exemples encaminats a guiar a l'usuari en d'introducció d'un nou tipus de log:

Indicacions generals:

.	qualsevol caràcter
*	qualsevol nombre d'aparicions
()	agrupació ==> camp
	or
\	símbol per indicar que al darrera hi va un caràcter especial (expressions regulars) que no ha de ser tractat com a tal.
<(.*)?>	Opcional: pot aparèixer o no. Però apareixeran els separadors buits < >
(?:<(.*)>)?	Opcional: < > pot aparèixer o no. No apareixeran els separadors quan el camp no hi sigui present.
^	inici de línia
(^a)	negació ==> tot menys a

Exemples:

1.- intervenció de 3 camps amb separador < al camp central i espai entre els camps.
(.*) <(.*)> (.*)

2.- intervenció de 4 camps amb 2 separadors diferents. Un dels camps és opcional:

<(.*)> \[(.*)\] (?:<(.*)>)? <(.*)>

3.- intervenció amb un sol camp i dues possibles cadenes d'inici:

^connectat | ^desconnectat

4.- Tot:
(^.*)

8.14 Fitxer de Configuració

El nom del fitxer de configuració és Ix2xml_config.properties. Es basa en el forma properties de la classe properties de java. <http://java.sun.com/j2se/1.4.2/docs/api/java/util/Properties.html>

Conté dos apartats:

- **Paràmetres generals:** destinats a parametritzar el funcionament general de l'aplicació.

- **Tipus de log:** Grups de paràmetres etiquetats com a comentaris que al ser descomentats passant a determinar el tipus de log que l'aplicació tracta.

Nota: La manipulació del fitxer de configuració ha de ser duta a terme amb precaució. I només per usuaris amb perfil d'administració. Qualsevol errada en la introducció dels paràmetres pot provocar errors en l'execució del programa.

Fragment del fitxer de configuració:

```
#####
# Sample ResourceBundle properties file
# Ramon Govern -- TFC UOC -- lx2xml
# Fitxer de configuracio de l'aplicacio
#
# Cal comentar/descomentar una secció per fer-la no_activa / activa
# Per afegir un nou tipus de log cal inserir (copy/paste) una secció
# complerta. Mes detalls al manual d'usuari
#####
#####
# Aquesta seccio es de parametres generals de la aplicacio
# sempre ha d'estar descomentada
#####

dirFitxers=lx2xml/fitxers
nomFitxerLinRebutjades=lx2xml_rebutjades.log

#####
# Tipus definit pel tractament dels logs que compleixen
# les especificacions del TFC
# 4 camps: idUsuari, data (format Unix), idTipusIntervencio, contingut
# Camps Opcionals: Cap. Tots son obligatoris.
# Ex: 2||1105376519||1||Bona tarda Josep, com va tot?
# El fitxer log es molt restrictiu. Chequeja la estructura del document,
# el contingut tipus dels camps i la seva longitud. Si un log conté errors
# no superarà el proces de validació.
#####

nomTipus=Type1_Standard_TFC_UOC
descripcio=logs que compleixen les especificacions del TFC
exemple=2||1105376519||1||Bona tarda Josep, com va tot?
nombreCamps=4
regExp=c
regExpLinEspecial=connectat|desconnectat
camps=idUsuari,data,idCategoria,contingut
fitxerXsd=type1_logSchema_restrictiu.xsd
fitxersXslt=lx2xml_perCategoria.xsl,lx2xml_perHora.xsl,lx2xml_perUsuari.xsl

#####
# Tipus definit pel tractament dels logs que compleixen
# les especificacions del TFC
# 4 camps: idUsuari, data (format Unix), idTipusIntervencio, contingut
# Camps Opcionals:
# Ex: 2||1105376519||1||Bona tarda Josep, com va tot?
# El fitxer esquema es molt poc restrictiu. Nomes chequeja la correctesa
# de la estructura xml. Els fitxers log mal formats superen la validació.
#####

#nomTipus=Type1_Standard_TFC_UOC
#descripcio=logs que compleixen les especificacions del TFC
#exemple=2||1105376519||1||Bona tarda Josep, com va tot?
#nombreCamps=4
#regExp=(.*)\\|\\|(\\.*)\\|\\|(\\.*)\\|\\|(\\.*)
#regExpLinEspecial=connectat|desconnectat
#camps=idUsuari,data,idCategoria,contingut
#fitxerXsd=lx2xml_logSchema.xsd
#fitxersXslt=lx2xml_perCategoria.xsl,lx2xml_perHora.xsl,lx2xml_perUsuari.xsl

#####
# Tipus definit per tractament d un altre tipus de log
# 3 camps: Data, user, intervencio Obligatoriis
```

```
# Ex: 10:33:43 <Emmy> Has avanzado? necesitas ayuda?
# No conté línies especials
# S'utilitza el esquema de validació poc restrictiu
#####

#nomTipus=Type2_Xat_Internet
#descripcio=logs tipus estandar (extret de Internet)
#exemple=10:33:43 &lt;Emmy&gt; Has avanzado? necesitas ayuda?
#nombreCamps=3
#regExp=(.*) <(.*)> (.*)
#regExpLinEspecial=^[.]
#camps=data,idUsuari,contingut
#fitxerXsd=lx2xml_logSchema.xsd
#fitxersXslt=lx2xml_perUsuari.xsl,lx2xml_perData.xls

#####
# Tipus definit per tractament d un altre tipus de log
# 2 camps: user, intervencio Obligatoris
# Ex: <RMartinez> Rogelio, este es un proyecto que implica
# No conté línies especials
# S'utilitza el esquema de validació poc restrictiu
#####

#nomTipus=Type3_Xat_Internet
#descripcio=logs tipus estandar (extret de Internet)
#exemple=&lt;RMartinez&gt; Rogelio, este es un proyecto que implica
#nombreCamps=2
#regExp=<(.*)> (.*)
#regExpLinEspecial=^[.]
#camps=idUsuari,contingut
#fitxerXsd=lx2xml_logSchema.xsd
#fitxersXslt=lx2xml_perUsuari.xsl
```

8.15 Com definir un nou tipus de log ?

Cal afegir un apartat complet de tipus de log amb els corresponents valors a cada propietat:

- **nomTipus:** nom del tipus de log
- **descripció:** text informatiu que apareixerà en la pantalla de l'aplicació
- **nombreCamps:** nombre de camps que conté la intervenció (màxim en cas d'haver-hi camps opcionals).
- **regExp:** expressió regular que caracteritza la intervenció.
- **regExpLinEspecial:** expressió regular que caracteritza les línies especials.
- **camps:** llista dels noms dels camps, ordenada segons aparició a la intervenció.
- **fitxerXsd:** fitxer esquema que validarà el fitxer xml
- **fitxerXslt:** llista de fitxers xslt que poden ser aplicats un determinat tipus de log.

Nota: la propietat fitxerXslt no està implementada, aquesta funcionalitat està fixada pel codi de l'aplicació tot i que s'ha mantingut el paràmetre en base a futures millores.

Cal que els noms dels camps coincideixin amb els camps definits als fitxers xlst utilitzats. Actualment, aquest noms són: idUsuari, data, idCategoria i contingut. S'aconsella començar utilitzant l'esquema lx2xml_Schema.xsd. És molt poc restrictiu i validarà la majoria de logs. Posteriorment es pot personalitzar aquest esquema pel tipus de log concret per fer-lo més restrictiu.

8.16 Fitxers schema

Per la resta de tipus de log definits actualment al fitxer de configuració, s'ha creat un fitxer esquema genèric molt poc restrictiu que només valida la correctesa de l'estructura xml. No realitza cap validació sobre el contingut dels camps.

Si es volgués aprofundir en la validació d'alguna altre tipologia de log, caldria definir el corresponent esquema <http://www.w3.org/XML/Schema> amb el màxim de restriccions possibles (contingut dels camps, obligatorietat, ordre etc...) i introduir-lo en l'apartat corresponent del fitxer de configuració.

Contingut del fitxer esquema type1_logSchema_restringit.xsd:

```
<?xml version="1.0" encoding="UTF-8"?>
<!--
  Document : type1_logSchema_restringit.xsd
  Created on : 10 / juny / 2006, 00:24
  Author : ramon govern
  Description: Esquema restringit basat amb les especificacions del xat TFC-UOC
  Purpose of the document follows.
-->
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="log">
    <xsd:complexType>
      <xsd:sequence maxOccurs="unbounded">
        <xsd:choice>
          <xsd:element ref="linia_especial" minOccurs="1" maxOccurs="unbounded"/>
          <xsd:element ref="intervencio" minOccurs="1" maxOccurs="unbounded" />
        </xsd:choice>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="linia_especial" type="linia_especial_type"/>
  <xsd:element name="intervencio" >
    <xsd:complexType >
      <xsd:sequence>
        <xsd:element name="camp" >
          <xsd:complexType>
            <xsd:simpleContent>
              <xsd:extension base="campUser_type">
                <xsd:attribute name="nom" type="xsd:string" use="required" />
              </xsd:extension>
            </xsd:simpleContent>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="camp" >
          <xsd:complexType>
            <xsd:simpleContent>
              <xsd:extension base="campData_type">
                <xsd:attribute name="nom" type="xsd:string" use="required" />
              </xsd:extension>
            </xsd:simpleContent>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="camp" >
          <xsd:complexType>
            <xsd:simpleContent>
              <xsd:extension base="campTipus_type">
                <xsd:attribute name="nom" type="xsd:string" use="required" />
              </xsd:extension>
            </xsd:simpleContent>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="camp" >
          <xsd:complexType>
            <xsd:simpleContent>
              <xsd:extension base="campContingut_type">
                <xsd:attribute name="nom" type="xsd:string" use="required" />
              </xsd:extension>
            </xsd:simpleContent>
          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:simpleType name="linia_especial_type">
    <xsd:restriction base="xsd:string">
      <xsd:pattern value="((connectat)|(desconnectat))\|([0-9]{10})\|([0-9]{1,2})\|\" />
    </xsd:restriction>
  </xsd:simpleType>

```

```
<xsd:simpleType name="campData_type">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="[0-9]{10}"/>
  </xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name="campUser_type">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="[0-9]{1,2}"/>
  </xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name="campTipus_type">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value="[0-9]{1}"/>
  </xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name="campContingut_type">
  <xsd:restriction base="xsd:string">
    <xsd:pattern value=".*"/>
  </xsd:restriction>
</xsd:simpleType>
</xsd:schema>
```

9 Annex 2 - Especificacions log xat TFC-UOC

Durant l'execució de les etapes de disseny i anàlisi, es va rebre unes breus especificacions corresponents a una tipologia de log generat per un aplicatiu provinent d'un TFC de la UOC. Aquestes especificacions determinen les característiques del tipus de log que tracta l'aplicació (tot i que es deixa una via oberta a la modificació via configuració d'aquestes especificacions). Aquestes especificacions han estat ampliadetes i concretades a criteri propi segons les indicacions del consultor, degut a que per motius personals, el responsable de l'aplicació xat TFC-UOC no va poder realitzar aquesta tasca.

A continuació, les exposem:

Els fitxers logs són fitxers de tipus text.

La grandària dels fitxer no pot superar els 4 Mb. (

Hi han dos tipus d'entrades al log:

- **Entrades intervenció:** corresponents a les aportacions dels usuaris al xat.
- **Entrades línies especials:** corresponents als missatges de control que l'aplicació xat enregistra al xat.

Cada entrada (intervenció o línia especial) correspon a una línia del fitxer log. El delimitador doncs, de les intervencions és el control de retorn de línia.

El nombre de camps contingut a cada intervenció és 4. Tots són obligatoris.

El ordre dels camps no pot variar.

El nom dels camps són (per ordre d'aparició a la intervenció): idUsuari, data, idCategoria, contingut.

El separador de camp es el conjunt de caràcters ||

El format de la intervenció és: idUsuari || data || idCategoria || contingut

Ex: 2||1105376519||1||Bona tarda Josep, com va tot?

La seqüència que identifica i separa una intervenció en el llenguatge d'expressions regulars és:

`(.*)\\|\\|(.*)\\|\\|(.*)\\|\\|(.*)`

El contingut i tipus dels camps és el següent:

- **idUsuari:** 1 o 2 dígits del 0 al 9
- **data:** 10 dígits del 0 al 9
- **idCategoria:** 1 dígit del 0 al 9
- **contingut:** text, qualsevol caràcter a excepció de la seqüència del separador de camps
||

El format de la línia especial és : línia_especial (tot el text junt)

La seqüència (expressió regular) que identifica una línia especial és: connectat | desconnectat

Ex: connectat||1105376496||2||

La validació del fitxer xml es realitza amb el fitxer type1_logSchema_restrictiu.xsd i conté els paràmetres necessaris per garantir les especificacions anteriors. L'associació entre el fitxer xml i el fitxer esquema la realitza l'aplicació, però queda de forma permanent ja que el fitxer xml porta introduït el nom i localització del esquema a la seva capçalera per posteriors validacions. Per més detalls sobre el xml esquema veure el capítol bibliografia de la memòria. Veure [Fitxers Schema](#).