

Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA

CORSO DI LAUREA IN INFORMATICA



Sviluppo di un modulo in Alfresco, sistema di gestione della conoscenza

Tesi di laurea triennale

Relatore

Prof. Gaggi Ombretta

Laureando

Trevisan Davide

ANNO ACCADEMICO 2016-2017

Abstract

Scopo di questa tesi di laurea è esporre il lavoro svolto dal laureando Trevisan Davide durante lo stage di trecentoventi ore presso l'azienda Ennova Research SRL (che nel presente documento spesso verrà abbreviata in "azienda") con sede in Venezia-Mestre. Il progetto di stage si è incentrato sullo sviluppo di alcune funzionalità connesse al sistema di gestione della conoscenza che l'azienda utilizza per gestire la documentazione relativa ai progetti assegnati dalle aziende clienti e monitorare l'operato dei dipendenti. Il sistema si basa sulla piattaforma Alfresco, un noto e molto utilizzato [KMS](#), la quale offre una SDK (che verrà descritta estensivamente ed esaurientemente in seguito) lanciabile in maniera autonoma tramite Maven che consente lo sviluppo di moduli per personalizzare e ottimizzare la piattaforma a proprio piacimento. Nella realizzazione del progetto è stato possibile, pertanto, approfondire le potenzialità, i difetti e le caratteristiche della piattaforma Alfresco nonché le fasi che hanno portato alla realizzazione di moduli per aggiungere ad Alfresco nuove funzionalità e una maggiore customizzazione del suo aspetto. Tutte le fasi, le problematiche e quanto prodotto durante il progetto sarà accuratamente esposto nei capitoli che compongono la presente tesi.

Indice

1	Introduzione	1
1.1	L'azienda	1
1.2	Lo stage	1
1.3	Organizzazione del testo	2
2	tecnologie e strumenti utilizzati	3
2.1	Tecnologie utilizzate	3
2.1.1	Linguaggi di Programmazione	3
2.1.2	Formato per l'interscambio di dati	4
2.1.3	Tecnica per l'interscambio di dati	4
2.1.4	Framework	4
2.1.5	Protocolli di servizi	5
2.1.6	Data Base Management System	5
2.2	Strumenti utilizzati	6
2.2.1	Strumenti di supporto all'attività di Codifica	6
2.2.2	Strumenti per il versionamento	6
2.2.3	Strumenti per <i>build automation</i>	6
3	Descrizione dello stage	9
3.1	Introduzione al progetto	9
3.1.1	Considerazioni preliminari	9
3.1.2	Considerazioni iniziali	9
3.1.3	Approccio alla piattaforma	10
3.1.4	Ambiente di test	10
4	Creazione del modulo	
	<i>E accenni al funzionamento di Alfresco</i>	11
	Glossary	13
	Acronyms	15
	Bibliografia	17

Elenco delle figure

Elenco delle tabelle

Capitolo 1

Introduzione

In questo capitolo verrà brevemente esposto il contesto in cui si è svolto lo stage, descrivendo le motivazioni che hanno spinto l'azienda a proporre questo stage

1.1 L'azienda

Ennova Research srl è un'azienda che opera nel settore [ICT](#) e realizza soluzioni informatiche altamente tecnologiche ed affidabili, che le permettono di agire con successo in settori di business come quello della Pubblica Amministrazione, delle grandi Corporazioni Bancarie, delle Multinazionali [ICT](#) e della Grande Distribuzione. È partner di grandi attori del mercato nazionale e internazionale quali Engineering, Toshiba, EMC, HP, Novell, Nvidia, ed altri . Ennova Research si distingue anche nel campo delle tecnologie open source utilizzate per la realizzazione di soluzioni multimediali avanzate destinate ai mercati [B2C](#) e [B2B](#) e di applicativi software destinati al settore del mobile, ad esempio Slash, che sfrutta le [Application Program Interface](#) di Twitter. L'azienda inoltre investe molto in ricerca e sviluppo ed è sempre pronta ad esplorare nuove tecnologie.

1.2 Lo stage

Il progetto di stage svoltosi all'interno dell'azienda Ennova Research è principalmente consistito nello sviluppo di moduli per Alfresco, che è il [KMS](#) che l'azienda ha incominciato da relativamente poco ad utilizzare e a cui intende sviluppare alcune funzionalità, quali ad esempio la gestione dei clienti e dei progetti, che al momento è svolta con l'ausilio di altri applicativi, e la rendicontazione, svolta anch'essa mediante l'utilizzo di altri applicativi. L'azienda tuttavia si pone l'obiettivo di creare non solo un prodotto necessario alle necessità interne, ma anche un prodotto vendibile e che generi profitto attraverso la vendita dello stesso ad aziende clienti. Il progetto è stato denominato Coral Tree e lo stage si è posto l'obiettivo di iniziare a porre le basi per questo progetto, attraverso la creazione di un modulo che introduca il nuovo tema per dare l'aspetto desiderato al prodotto e la creazione dei primi moduli che introducano le funzionalità precedentemente citate.

1.3 Organizzazione del testo

Il secondo capitolo descrive le tecnologie e gli strumenti utilizzati durante lo stage.

Il terzo capitolo descrive brevemente le condizioni e le metodologie utilizzate per lo stage.

Il quarto capitolo descrive brevemente come creare un modulo e l'architettura di Alfresco.

i tre capitoli successivi descriveranno in maniera separata i vari moduli sviluppati, dato che il modulo che aggiunge il nuovo tema differisce sostanzialmente nella sua realizzazione dai moduli che aggiungono funzionalità nel senso stretto. Si hanno quindi i capitoli

il modulo tema ;

il modulo clienti ;

il modulo progetti ;

Capitolo 2

tecnologie e strumenti utilizzati

In questo capitolo saranno esposte le principali tecnologie utilizzate e i principali strumenti che sono stati utilizzati per portare a compimento il progetto assegnato

2.1 Tecnologie utilizzate

In seguito verranno descritte più nel dettaglio le tecnologie apprese o particolarmente approfondite durante lo svolgimento dello stage, che vanno ad aggiungersi a quelle pregresse, come ad esempio l'HTML e il CSS, già noti e ampiamente utilizzati nel corso del progetto.

2.1.1 Linguaggi di Programmazione

Java

Java è un linguaggio di programmazione orientato agli oggetti a tipizzazione statica, specificatamente progettato per essere il più possibile indipendente dalla piattaforma di esecuzione. Uno dei principi fondamentali del linguaggio è espresso dal motto *write once, run anywhere*: il codice compilato che viene eseguito su una piattaforma non deve essere ricompilato per essere eseguito su una piattaforma diversa. Il prodotto della compilazione è infatti in un formato chiamato *bytecode* che può essere eseguito da una qualunque implementazione di un processore virtuale detto *Java Virtual Machine*. Alfresco usa Java per il lato backend, soprattutto per i *webscript*, per i quali è stato largamente utilizzato assieme alla sua API.

Javascript

JavaScript è un linguaggio di scripting orientato agli oggetti e agli eventi, comunemente utilizzato nella programmazione Web lato client per la creazione, in siti web e applicazioni web, di effetti dinamici interattivi, tramite funzioni di script invocate da eventi innescati, a loro volta, in vari modi dall'utente sulla pagina web in uso. Tali funzioni di script possono essere opportunamente inserite in file HTML, in pagine JSP o in appositi file separati con estensione *.js*, poi richiamati nella logica di business. Nel progetto esso è stato usato per popolare dinamicamente le pagine di cui si sono composti i vari moduli e per eseguire i controlli sui form una volta inviate le richieste.

2.1.2 Formato per l'interscambio di dati

JSON

JSON (JavaScript Object Notation) è un formato di scambio dati leggero e facile da leggere e scrivere sia per macchine che per umani, si basa su un sottoinsieme del linguaggio di programmazione JavaScript, anche se ne è indipendente, ed inoltre è un formato di testo completamente indipendente da qualsiasi linguaggio. Per queste caratteristiche è un linguaggio di scambio dati ideale. JSON è costruito su due strutture:

- * Una collezione di coppie nome/valore, spesso realizzato come un array associativo;
- * Un elenco ordinato di valori, spesso realizzato come una lista.

Nel progetto esso è stato usato per poter scambiare i dati tra il lato frontend e il lato backend attraverso il passaggio di oggetti JSON tramite AJAX

2.1.3 Tecnica per l'interscambio di dati

AJAX

AJAX, acronimo di Asynchronous Javascript And XML, è una tecnica di sviluppo software per la realizzazione di applicazioni web interattive. Lo sviluppo di applicazioni HTML con AJAX si basa su uno scambio di dati in background fra web browser e server, che consente l'aggiornamento dinamico di una pagina web senza esplicito ricaricamento da parte dell'utente. AJAX è asincrono, nel senso che i dati extra sono richiesti al server e caricati in background senza interferire con il comportamento della pagina esistente, anche se è possibile, ma sconsigliato, effettuare richieste sincrone. Normalmente le funzioni richiamate sono scritte con il linguaggio JavaScript. Tuttavia, e a dispetto del nome, l'uso di JavaScript e di XML non è obbligatorio, come non è detto che le richieste di caricamento debbano essere necessariamente asincrone. Nel progetto esso è stato usato per mettere in comunicazione il lato frontend e il lato backend, grazie anche all'utilizzo delle librerie AJAX messe a disposizione dell'API Javascript di Alfresco.

2.1.4 Framework

Spring

Spring è un framework per realizzare applicazioni web basate sul Modello MVC avendo come punti di forza l'inversion of control (tramite dependency injection) e la aspect oriented programming. Esso si occupa di mappare metodi e classi Java con determinati url, di gestire differenti tipologie di "viste" restituite al client, di realizzare applicazioni internazionalizzate e di gestire i cosiddetti temi per personalizzare al massimo l'esperienza utente. Questo framework è strutturato a livelli, e permette di scegliere quale dei suoi componenti usare, fornendo nello stesso momento un framework di ottima qualità per lo sviluppo di applicazioni distribuite. Questa architettura a livelli consiste in diversi moduli (o componenti) ben definiti, ognuno dei quali può rimanere da solo o essere implementato con altri. Su questo framework si basa la struttura di base di Alfresco.

Jquery

JQuery è un framework nato con il preciso intento di rendere il codice più sintetico e di limitare al minimo l'estensione degli oggetti globali per ottenere la massima compatibilità con altre librerie. Grazie a questo principio jquery è in grado di offrire un'ampia gamma di funzionalità, che vanno dalla manipolazione degli stili CSS e degli elementi HTML, agli effetti grafici, per passare a comodi metodi per chiamate AJAX cross-browser. Nel progetto esso è stato utilizzato per la creazione di popup ed effetti per l'utente.

Alfresco Surf

Alfresco Surf è un framework messo a disposizione della piattaforma Alfresco con cui è possibile creare interfacce grafiche, modelli e componenti sfruttando scripts server-side e templates. Anche se Alfresco stesso sta per abbandonarlo in favore di Angular2, esso è ancora il framework con cui viene distribuito il lato frontend nelle distribuzioni di Alfresco.

2.1.5 Protocolli di servizi

LDAP

LDAP è un acronimo che sta per LIGHTWEIGHT DIRECTORY ACCESS PROTOCOL. Come suggerisce il nome stesso, è un protocollo leggero per accedere ai servizi di directory, basati sul protocollo X.500. LDAP opera su TCP/IP o su altre connessioni orientate ai servizi di trasferimento. LDAP nasce per sostituire DAP in quanto molto oneroso dal punto di vista dell'impiego delle risorse ed è basato sul modello client-server: un client LDAP invia una richiesta ad un server LDAP, che processa la richiesta ricevuta, accede eventualmente ad un directory database e ritorna dei risultati al client. Il modello di informazioni di LDAP è basato sulle entry. Un'entry è una collezione di attributi aventi un unico nome globale: il Distinguished Name (DN). Il DN è usato per riferirsi ad una particolare entry, senza avere ambiguità. Ogni attributo dell'entry ha un tipo ed uno o più valori. In LDAP, le entry di una directory sono strutturate come in una struttura gerarchica di un albero. L'azienda fa uso di questo protocollo per la profilatura degli utenti in virtù della sua ottima integrazione con Alfresco, e quindi di conseguenza si è dovuta utilizzare questa tecnologia per avere accesso ai dati degli utenti registrati

2.1.6 Data Base Management System

PostgreSQL

PostgreSQL è un completo DBMS ad oggetti rilasciato con licenza libera; spesso viene abbreviato come "Postgres". PostgreSQL è una reale alternativa sia rispetto ad altri prodotti liberi come ad esempio MySQL, con i quali però comunque condivide molti aspetti del suo linguaggio, che a quelli a codice chiuso come ad esempio Oracle, poichè offre caratteristiche uniche nel suo genere che lo pongono per alcuni aspetti all'avanguardia nel settore dei database. È il DBMS usato in maniera predefinita da Alfresco e per questo motivo è stato scelto per questo progetto.

2.2 Strumenti utilizzati

2.2.1 Strumenti di supporto all'attività di Codifica

Eclipse

Eclipse è un ambiente di sviluppo integrato multi-linguaggio e multiplatforma. Ideato da un consorzio di grandi società quali Ericsson, HP, IBM, Intel, MontaVista Software, QNX, SAP e Serena Software, chiamato Eclipse Foundation. Eclipse è un software libero distribuito sotto i termini della Eclipse Public License ed è lo strumento che viene imposto dall'azienda ai suoi dipendenti, in quanto software libero, quindi la sua adozione per questo progetto è stata scontata. Questo IDE è stato utilizzato per l'importazione di un progetto allo scopo di estendere la piattaforma di Alfresco. Le funzionalità presentate da questo IDE in termini di aiuti al programmatore e la sua integrazione con Maven hanno permesso uno sviluppo rapido ed efficiente dei plugin.

2.2.2 Strumenti per il versionamento

Git

Il sistema di versioning adottato per questo progetto e in generale utilizzato dall'azienda è Git, che negli ultimi anni si è affermato come uno dei migliori sistemi di controllo di versione. Le caratteristiche per cui si è distinto dagli altri software sono:

- * L'architettura, progettata per essere totalmente distribuita, in modo da rendere possibile il lavoro e il versionamento offline ed anche un versionamento integrato con i principali IDE;
- * Le funzionalità di branching e merging potenti, rapide e comode, che permettono una efficiente gestione di progetti, anche di grandi dimensioni;
- * Le performance generalmente migliori e la possibilità di pubblicare i repository con i principali protocolli.

Ennova Research si appoggia nello specifico su GitLab, un servizio di hosting per progetti basato su Git.

2.2.3 Strumenti per *build automation*

Apache Maven

Maven è un software usato principalmente per la gestione di progetti Java e *build automation*. Per funzionalità è simile ad Apache Ant, ma basato su concetti differenti. Può essere usato anche in progetti scritti in C#, Ruby, Scala e altri linguaggi. Il progetto Maven è ospitato da Apache Software Foundation. Maven usa un costrutto conosciuto come Project Object Model (POM); un file XML che descrive le dipendenze fra il progetto e le varie versioni di librerie necessarie nonché le dipendenze fra di esse. In questo modo si separano le librerie dalla directory di progetto utilizzando questo file descrittivo per definirne le relazioni. Maven effettua automaticamente il download di librerie Java e plug-in Maven dai vari repository definiti scaricandoli in locale o in un repository centralizzato lato sviluppo. Questo permette di recuperare in modo uniforme i vari file JAR e di poter spostare il progetto indipendentemente da un ambiente all'altro avendo la sicurezza di utilizzare sempre le stesse versioni delle

librerie e la possibilità di impostare profili di esecuzione diversi. Per automatizzare la build e il lancio dell'SDK, Alfresco si basa appunto su Maven.

Capitolo 3

Descrizione dello stage

In questo capitolo si intende illustrare nel dettaglio le fasi e le considerazioni che hanno portato allo sviluppo del progetto e che sono alla base della proposta di stage

3.1 Introduzione al progetto

3.1.1 Considerazioni preliminari

Per prima cosa, è necessario premettere che l'azienda in cui si è svolto lo stage per i suoi prodotti software richiede l'impiego di procedure, regole e strumenti idonei a raggiungere il miglior risultato possibile. In particolare, durante lo stage sono stati richiesti i seguenti punti:

- * Tracciabilità: è stato spesso richiesto di rendere conto dello stato di avanzamento dei processi, delle componenti e delle attività in lavorazione, al fine di verificare il rispetto della tabella di marcia,
- * Documentazione: descrizione di quanto è stato fatto, delle strategie utilizzate e dei dettagli implementativi in modo puntuale e preciso, per far sì che quanto è stato fatto generi utilità all'azienda e sia riutilizzabile e replicabile in futuro, senza dover rifare tutte le attività, avendo già una strategia e una soluzione pronte e immediatamente utilizzabili.
- * Misurazione: per ogni attività eseguita, deve essere misurabile il suo costo (in termini di tempo, nel caso del progetto qui esposto), e la sua qualità complessiva, in base a vari indicatori.

Quanto esposto è stato tenuto in grande considerazione durante lo sviluppo, al fine di realizzare un prodotto quanto più possibile conforme agli standard di qualità aziendali

3.1.2 Considerazioni iniziali

Alfresco, soprattutto nella versione Community, utilizzata dall'azienda perché Open Source è dotata di pochissime utilità di base, dal momento che il profitto dell'azienda che produce il software proviene dalla consulenza tecnica fornita dai loro esperti e dalla vendita di moduli ulteriori che vanno ad estendere la dotazione di base, moduli disponibili, soprattutto nel caso dei più interessanti, solo tramite la sottoscrizione

di un abbonamento, che viene valutato di caso in caso ma è comunque costoso. L'azienda quindi punta a fornire moduli adatti alla gestione di clienti e progetti integrati direttamente in Alfresco, e punta quindi gradualmente ad accentrare in Alfresco più funzionalità possibili, sostituendo gradualmente gli applicativi attualmente utilizzati ad esempio per la gestione dei progetti e la rendicontazione, per poi in un futuro poter proporre a clienti terzi il prodotto finito. Per questo motivo il progetto di stage si è concretizzato in quanto esposto, cercando quindi di introdurre alcune funzionalità.

3.1.3 Approccio alla piattaforma

Le prime due settimane di stage sono state dedicate quasi interamente allo studio della complessa architettura e della SDK della piattaforma. Per affrontare quanto appena illustrato è stato effettuato un attento studio della SDK di Alfresco, scegliendo il tipo di archetipo da utilizzare. La scelta è ricaduta sull'archetipo all-in-one in quanto permette lo sviluppo sia del lato front-end che di quello back-end in simultanea, senza dover distribuire ed installare separatamente un modulo per ogni parte sviluppata. In seguito, si sono analizzate le richieste poste dall'azienda in merito alle funzionalità desiderate al fine di individuare le strategie e i metodi migliori per realizzare quanto richiesto.

3.1.4 Ambiente di test

Assieme allo sviluppo dei vari moduli, è proceduta a pari passo anche la creazione e il mantenimento di un ambiente di test dove poter testare i moduli in maniera sicura prima di portarli in produzione e applicare i moduli nel [KMS](#) aziendale. Ciò è stato espressamente richiesto dall'azienda e quindi i moduli prodotti, non appena la loro installazione era possibile, sono stati caricati e provati nel [KMS](#) di test che replicava quello dell'azienda, e successivamente, prima del rilascio finale, in una copia esatta del software del [KMS](#) di produzione.

Capitolo 4

Creazione del modulo

E accenni al funzionamento di Alfresco

In questo capitolo sarà esposto come creare un modulo e accenni all'architettura di Alfresco

Bisogna innanzitutto rispondere ai prerequisiti descritti in [questa pagina](#). Fatto questo il progetto può essere importato in un IDE, ad esempio per Eclipse in [questo sito](#). Le cartelle su cui andremo a sviluppare saranno la repo-amp e la share-amp. La share-amp verrà utilizzata per sviluppare l'interfaccia grafica del modulo mentre la parte di repo-amp per lo sviluppo del “back-end”, le altre folder create servono per replicare l'ambiente di Alfresco in fase di esecuzione del modulo, e quindi non ha senso lavorare su di esse. È necessario installare il driver jdbc se si intende lavorare nell'SDK interagendo con database terzi, e dato che non è incluso nell'SDK anche se presente nell'ambiente finale, va incluso seguendo quanto detto in [questo luogo](#). Una volta fatto questo, bisogna aprire nell'IDE o anche a mano all'indirizzo, partendo dalla cartella di installazione (sono riportati gli indirizzi utilizzando i nomi dell'esempio sopra citato: / alfresco-extension / acme-cms-poc / acme-cms-poc-share-amp / src / main / amp / config / alfresco / web-extension / site-webscripts / com / example / pages (questo porta a dove è situata la pagina dimostrativa). È sufficiente fermarsi due o tre livelli sopra se l'intenzione è quella di mettere la gerarchia per una pagina custom (come fatto in questo progetto). Una pagina fatta in alfresco share è composta da tre file, per semplicità prendiamo in considerazione la pagina di esempio già fornita: un file .xml, che descrive i dati essenziali della pagina:

```
<webscript>
  <shortname>Simple Page</shortname>
  <description>Simple page definition</description>
  <family>Share</family>
  <authentication>user</authentication>
  <url>/simple-page</url>
</webscript>
```

molto importanti sono il tag authentication, che specifica il livello di autenticazione necessario per accedere alla pagina, e il tag url, dal momento che definisce l'indirizzo della pagina, che sarà <ip>:port/share/page/hdp/ws<url> (hdp genera la pagina

con header e footer, dp la pagina grezza). Vi è poi il file .html che si comporta proprio come il body di una pagina html. Vi è infine un file .js, che serve per applicare delle modifiche via json al model. Quanto detto è sufficiente per creare una pagina che non necessiti di interazioni con java.

Per quanto riguarda una pagina che invece fa uso delle api di Alfresco il procedimento è più complicato: la pagina che verrà mostrata all'utente è collocata nella cartella `alfresco-extension / acme-cms-poc / acme-cms-poc-repo / amp / src / main / amp / config / alfresco / extension / templates / webscripts`. La pagina generata è situata all'indirizzo `<ip>:<porta>/alfresco/service<url>`. E' inoltre necessario creare un bean che fa da collegamento tra il codice java, situato in `alfresco-extension / acme-cms-poc / acme-cms-poc-repo-amp / src / main / java / <package>`, e la pagina che verrà generata. Per farlo seguire quanto descritto in [questo sito](#).

Glossario

API in informatica con il termine *Application Programming Interface API* (ing. interfaccia di programmazione di un'applicazione) si indica ogni insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per l'espletamento di un determinato compito all'interno di un certo programma. La finalità è ottenere un'astrazione, di solito tra l'hardware e il programmatore o tra software a basso e quello ad alto livello semplificando così il lavoro di programmazione. [1](#)

B2B Business-to-business, spesso indicato con l'acronimo *B2B*, in italiano commercio interaziendale, è una locuzione utilizzata per descrivere le transazioni commerciali elettroniche tra imprese, distinguendole da quelle che intercorrono tra le imprese e altri gruppi, come quelle oppure quelle tra una impresa e il governo. [1](#)

B2C Con Business to Consumer, spesso abbreviato in *B2C*, si indicano le relazioni che un'impresa commerciale detiene con i suoi clienti per le attività di vendita e/o di assistenza. [1](#)

ICT Information and communication technology:
Le tecnologie dell'informazione e della comunicazione (in inglese Information and Communications Technology, in acronimo ICT), sono l'insieme dei metodi e delle tecnologie che realizzano i sistemi di trasmissione, ricezione ed elaborazione di informazioni (tecnologie digitali comprese). [1](#)

KMS Knowledge management system:
I Knowledge management system sono sistemi software che supportano le fasi del ciclo dell'informazione e la comunicazione all'interno di una comunità di pratica (ad esempio un'azienda) o di apprendimento (ad esempio una classe "virtuale") anche disperse nello spazio. Dovrebbero assistere le persone ad esplicitare la conoscenza tacita, a reperirla, a condividerla, supportando in particolare le seguenti funzioni:

- * Cattura delle competenze collettive
- * Controllo per realizzare obiettivi comuni
- * Integrazione delle conoscenze frammentate

. [iii](#), [1](#), [10](#)

Acronimi

API [Application Program Interface](#). 13

B2B [Business-to-business](#). 13

B2C [Business-to-Business](#). 13

ICT [Information and Communication Tecnology](#). 13

KMS [Knowledge Managment System](#). 13

Bibliografia

Riferimenti bibliografici

James P. Womack, Daniel T. Jones. *Lean Thinking, Second Editon*. Simon & Schuster, Inc., 2010.

Siti Web consultati

Manifesto Agile. URL: <http://agilemanifesto.org/iso/it/>.