

Installation och konfiguration av ett webbaserat dokumentarkiv

Don Onnelainen

EXAMENSARBETE	
Arcada	
Utbildningsprogram:	Informations- och mediateknik
Identifikationsnummer:	4648
Författare:	Don Onnelainen
Arbetets namn:	Installation och konfiguration av ett webbaserat dokumentarkiv
Handledare (Arcada):	Thomas Forss
Uppdragsgivare:	CRIPS HEI ICI
<p>Sammandrag:</p> <p>Detta examensarbete behandlar ett installations- och konfigurationsprojekt av ett webbaserat dokumentarkiv för bl.a. slutarbeten, uppgifter och studiematerial. Projektet har beställts av University of Eastern Africa Baraton i samband med upprättandet av ett internationellt studieprogram inom global hälsovård.</p> <p>Arbetet presenterar beställarens kravställning på dokumentarkivet och jämför den med tre arkivplattformsföreslag som beställaren fört fram: SubjectsPlus, IR Plus och DSpace. Rapporten presenterar därefter närmare det arkiv som blivit valt som arkivplattform för projektet. Det valda arkivets installations- och konfigurationsprocess presenteras, och i samband med detta förs det även fram detaljer kring metoderna och hårdvaran som använts inom processen för att skraddarsy arkivet enligt beställarens krav och önskemål. Arbetet behandlar därefter arkivets övergång från utvecklingsmiljö till produktionsmiljö, och går även igenom det färdiga arkivets säkerhetsfaktorer såsom säker anslutning och säkerhetskopiering.</p> <p>Slutligen presenterar examensarbetet det färdigt skraddarsyddas arkivet och jämför det med en omodifierad installation. Det inleds även en diskussion om arkivet uppfyller de krav som beställaren ställt fram, och om de utmaningar som projektet presenterat behandlas.</p>	
Nyckelord:	dspace, installation, konfiguration, arkiv, UEAB
Sidantal:	39
Språk:	Svenska
Datum för godkännande:	27.04.2015

DEGREE THESIS	
Arcada	
Degree Programme:	Information and Media Technology
Identification number:	4648
Author:	Don Onnelainen
Title:	Installation and Configuration of a Web Based Document Repository
Supervisor (Arcada):	Thomas Forss
Commissioned by:	CRIPS HEI ICI
<p>Abstract:</p> <p>The topic of this thesis report is the installation and configuration of a web based document repository. The purpose of the repository is to store theses, assignments and educational material. The repository is a work order placed by the University of Eastern Africa Baraton (UEAB) in conjunction with the establishment of a master's program in global healthcare.</p> <p>This report reviews the requirements that UEAB has set in place for the repository, and compares them with three repository platforms suggested by UEAB: SubjectsPlus, IR Plus and DSpace. When the comparison is concluded this report presents a more detailed description of the repository platform that has been deemed to be the most fitting for UEAB's requirements. This is followed by a run-through of the installation, configuration and tailoring of the selected platform. Following the run-through this report also presents the details regarding the methods and hardware used in conjunction with the tailoring of the platform according to UEAB's requirements. Once the aspects regarding the tailoring of the platform have been explored the report also presents the process of migrating the completed repository platform from the development environment into production. Following the migration process this report briefly explores the security aspects like backups and secure connections relating to the repository.</p> <p>Concluding, this report showcases the completed tailored repository platform and compares it with an unmodified installation of the platform. The conclusion ends with a discussion regarding the challenges presented by the installation and configuration project and checks if UEAB's requirements for the repository have been met.</p>	
Keywords:	dspace, installation, configuration, repository, UEAB
Number of pages:	39
Language:	Swedish
Date of acceptance:	27.04.2015

INNEHÅLL

1	Inledning.....	9
1.1	Bakgrund	9
1.2	Syfte och målsättning	10
1.3	Metoder	10
1.4	Avgränsningar	11
2	Beställarens krav	11
2.1	Funktionella krav	11
2.1.1	<i>Användarhantering</i>	<i>12</i>
2.1.2	<i>Uppladdning av material.....</i>	<i>12</i>
2.1.3	<i>Granskning av uppladdat material.....</i>	<i>13</i>
2.1.4	<i>Sökning och visning av uppladdat material.....</i>	<i>13</i>
2.1.5	<i>Säkerhetskopiering.....</i>	<i>14</i>
2.2	Krav på utseende	14
3	Alternativ och val av arkivplattform	14
3.1	SubjectsPlus	15
3.2	IR Plus / IR+	15
3.3	DSpace	15
3.4	Slutligt plattformval	16
4	DSpace – Installation och konfiguration	16
4.1	Krav	17
4.1.1	<i>Mjukvarukrav</i>	<i>17</i>
4.1.2	<i>Hårdvarukrav</i>	<i>17</i>
4.1.3	<i>Kraven för DSpace-instansen vid UEAB.....</i>	<i>18</i>
4.2	Struktur	18
4.3	Komponenter	19
4.4	Utvecklingsmiljön.....	20
4.4.1	<i>Hård- och mjukvara</i>	<i>20</i>
4.4.2	<i>Installation och konfiguration av underliggande mjukvara</i>	<i>21</i>
4.4.3	<i>Grundinformation om DSpace-installationen.....</i>	<i>21</i>
4.4.4	<i>Installation av DSpace kortfattat.....</i>	<i>22</i>
4.4.5	<i>Installation av DSpace stegvis</i>	<i>23</i>
4.4.6	<i>Modifikation av DSpace.....</i>	<i>27</i>
4.5	Produktionsmiljön	29
4.5.1	<i>Hård- och mjukvara</i>	<i>29</i>
4.5.2	<i>Migration.....</i>	<i>29</i>

5 Arkivets datasäkerhet.....	32
5.1 Säkerhetskopiering.....	32
5.1.1 Säkerhetskopiering av filförråd.....	33
5.1.2 Säkerhetskopiering av relationsdatabas	33
5.2 Datasäker anslutning.....	33
6 Resultatredovisning	34
7 Slutsatser och diskussion.....	37
Källor	38
Bilagor	40

Figurer

Figur 1. Diagram över strukturkravet på arkivet.	13
Figur 2. DSpace-källpaketets struktur.	24
Figur 3. Utdrag ur förkonfigurationsfilen build.properties.....	25
Figur 4. Apache Maven-kompileringskommandot i rotkatalogen i Windows-miljö.	26
Figur 5. Innehållet i ROOT.xml.	27
Figur 6. Exempel på användning av PostgreSQL-funktionen pg_dumpall för att dumpa databasinnehåll.	31
Figur 7. Exempel på användning av PostgreSQL-funktionen psql för att återställa databasdump.	31
Figur 8. Indexeringskommandon i dspace.bat som måste köras efter migration.	32
Figur 9. Bild på det oförändrade Mirage-temats framsida.	34
Figur 10. Arkivets framsida med modifierat Mirage-tema.	35
Figur 11. Modifierat mobiltema i enlighet med Mirage-temats förändringar.	36

Förkortningar

API	Application Programming Interface
CSS	Cascading Style Sheets
GB	Gigabyte
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
IP	Internet Protocol
JDK	Java Development Kit
JSP	JavaServer Pages
OAI	Open Archive Initiative
RDP	Remote Desktop Protocol

SDK	Software Development Kit
UTF	Unicode Transformation Format
XHTML	Extensible Hypertext Markup Language
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language

FÖRORD

Jag vill tacka Yrkeshögskolan Arcada, Yrkeshögskolan Diakonia och University of Eastern Africa Baraton för att ha erbjudit mig möjligheten att jobba med installations- och konfigurationsarbetet som detta examensarbete behandlar. Jag vill även tacka Thomas Forss som fungerat som min handledare och Yrkeshögskolan Arcadas representant i samband med projektet.

1 INLEDNING

1.1 Bakgrund

Detta examensarbete utgör en del av ett större samarbete mellan Yrkeshögskolan Arcada, Yrkeshögskolan Diakonia samt University of Eastern Africa Baraton (UEAB) påbörjat i mars 2013. Samarbetsprojektet går under namnet *Capacity Building in Crisis Preparedness in Health Care Education* eller förkortat CRIPS HEI ICI.

Projektet har som sitt grundläggande mål att förbättra hälsomässiga omständigheter i mindre utvecklade områden inom Kenya genom att förbättra samhällets färdigheter för krishantering. För att uppnå dessa mål siktar CRIPS HEI ICI-projektet på att erbjuda en allmänt högre utbildningsnivå inom akut och allmän hälsovård.

Ett av projektets delmål är att upprätta ett internationellt utbildningsprogram vid namnet *Master's Program in Global Healthcare*. Utbildningsprogrammet har en planerad start i augusti 2014.

I samband med ovannämnda utbildningsprogram skall det upprättas ett digitalt arkiv på webben där bland annat slutarbeten, uppgifter och studiematerial kan laddas upp och som studenter inom programmet kommer ha tillgång till. Arkivet kommer även att innehålla de slutarbeten som studenterna inom det upprättade utbildningsprogrammet kommer att producera.

Yrkeshögskolan Arcada gavs då uppdraget att upprätta själva arkivet, och Arcada erbjöd då uppgiften vidare som en möjlighet till ett projektbaserat examensarbete för studenter inom utbildningsprogrammet i informations- och medieteknik. Genom denna process fick jag då som uppdrag att utföra det praktiska arbetet bakom upprättandet av arkivet, vilket då också utgör grunden för detta examensarbete.

1.2 Syfte och målsättning

Syftet med detta examensarbete är att fungera som en form av dokumentation av de olika aspekter som varit aktuella inom det tidigare nämnda praktiska installations- och konfigurationsarbetet. Examensarbetet presenterar de krav som beställaren av arkivet har ställt samt för fram utredningar om hur dessa krav uppfyllts inom mitt installations- och konfigurationsarbete genom mina egna beslut. Arbetet för även fram alla de utmaningar som stötts på i samband med projektet.

Målsättningen är att examensarbetet till största del behandlar och besvarar de utmaningar och frågor som uppstått under det praktiska installationsprojektet, och till slut för fram min egen (och förhoppningsvis kundens) bedömning gällande slutresultatet av installations- och konfigurationsarbetet.

1.3 Metoder

Beställaren har gett tre olika förslag på arkivplattformar som kunde användas för att upprätta arkivet och har visat ett stort intresse för en av dem. Trots det bad beställaren att jag tillsammans med min handledare skulle granska alla alternativ och bedöma hur bra de uppfyller, eller efter modifikation kommer att kunna uppfylla, de krav som beställaren har ställt på arkivet.

Efter att arkivplattformen slagits fast har jag arbetat med att göra förändringar i arkivets kod- samt konfigurationsfiler, för att då skräddarsy det enligt beställarens krav. Modifikationsarbetet har gjorts med hjälp av studier av arkivmjukvarans dokumentation. Här har också ingått direkt testning genom att observera ifall förändringarna man gjort lett till det resultat man önskat.

Utvecklingsarbetet på arkivet har till en början gjorts på en inom Yrkehögskolan Arcada på en dator som jag fysiskt arbetat vid. Vid ett senare skede har arkivet och därmed också utvecklingsarbetet flyttats till en dator som fysiskt varit placerad vid UEAB, vilket gjort arbetet till en form av distansarbete.

Som stöd för utvecklingsarbetet har kommunikation med personal från Yrkeshögskolan Diakonia samt representanter från UEAB använts. Detta har varit i form av rapporter om arbetets mellanskeden där feedback och förslag om önskvärda förändringar har ställts fram.

1.4 Avgränsningar

I målsättningen av examensarbetet ingår inte någon som helst fysisk installation av hårdvara. Den fysiska installationen av servern som är värddator för testinstanserna av Dspace har utförts av personal vid UEAB. Installationen av servern som är värddator för det färdiga arkivet har likaså utförts av personal vid UEAB.

Eftersom min del av projektet rätt så klart avgränsats till installationen och konfigurationen av arkivet, går detta examensarbete inte närmare in på några områden som handlar om arkivets konkreta användning i egenskap av en användare.

Även om arkivet kommer att finnas till för att ladda upp och därmed publicera skrivet material på, behandlar detta examensarbete inte lagfrågor kring upphovsrätt.

2 BESTÄLLARENS KRAV

I detta kapitel behandlas de funktionella och utseendemässiga krav som beställaren ställt fram gällande arkivet. De funktionella kraven handlar främst om arkivets hantering av dokument och användare. De utseendemässiga kraven berör främst möjligheten att forma om arkivets visuella aspekter så det passar ihop med beställarens visuella tema.

2.1 Funktionella krav

Beställaren ställde i sin arbetsbeställning (se Bilaga 1 – Thesis Work Order) fram målsättningen att arkivet skall ha funktionalitet för uppladdning av material samt stöd för sökning av uppladdat material för nedladdning och bläddring. De flesta av kraven i den ursprungliga projektbeställningen har varit mycket vaga vad beträffar detaljer kring egentlig funktionalitet och implementation. Dessa krav har under arbetets lopp vidare-

utvecklats genom förbättringsförslag i samband med demonstration av mellanskedet i det praktiska arbetet med arkivet.

Funktionella områden som kraven berör kan grovt uppdelas i följande:

- Användarhantering
- Uppladdning av material
- Granskning av uppladdat material
- Sökning och visning av uppladdat material
- Säkerhetskopiering

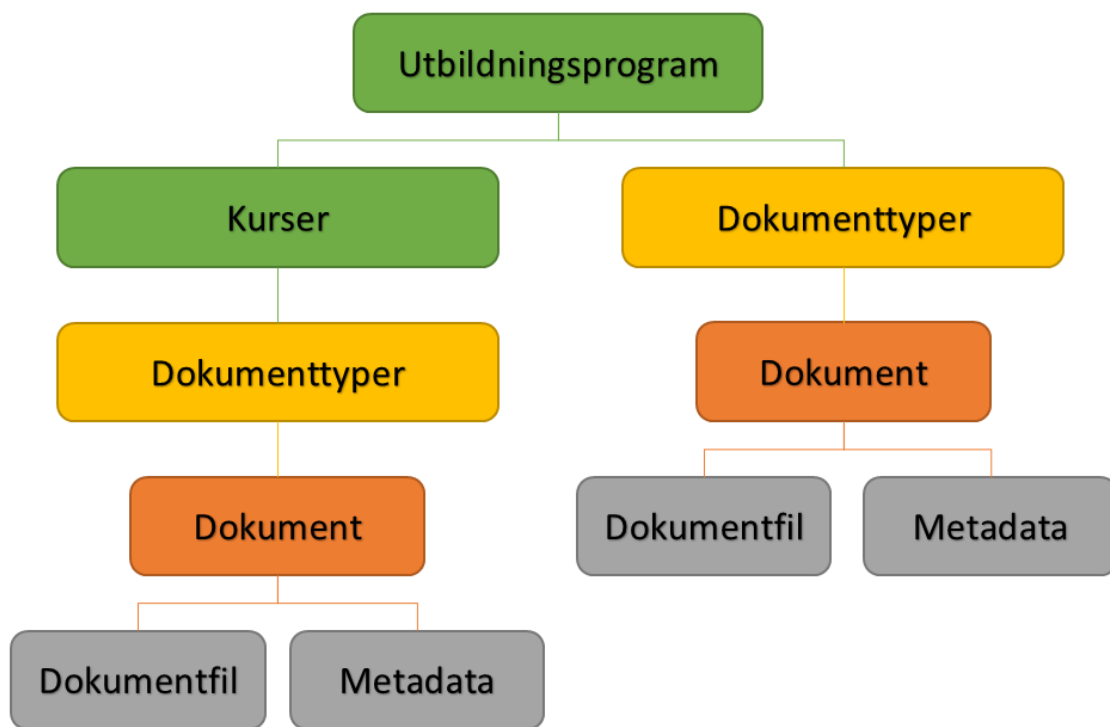
2.1.1 Användarhantering

Uppladdning av material och allmän administration av arkivet skall kontrolleras med någon form av användarhantering. Ett system som bygger på individuella användarkonton samt möjligheten att dela upp olika användare i grupper med specifika rättigheter torde enligt mig uppfylla detta krav.

2.1.2 Uppladdning av material

Personal och studenter vid UEAB skall kunna ladda upp dokument till arkivet. Vid uppladdningen skall användaren även kunna bifoga relevant metadata, dvs. information om exempelvis dokumentets författare, dokumenttyp, sammandrag, nyckelord etc. Arkivet måste godkänna dokument i PDF- och RTF-format, men ytterligare stöd för andra format såsom bilder och presentationer ses inte av beställaren som något förbjudet eller negativt.

Material som laddas upp till arkivet måste kunna indelas i olika sektioner eller kategorier, exempelvis kursvis. Kravet är i princip en trädstruktur där det finns en överkategori, såsom t.ex. ett utbildningsprogram. Under överkategorin finner man underkategorier såsom kurser. Inom underkategorierna så hittar man olika dokumenttyper såsom t.ex. studiematerial eller uppgifter. Det kan även finnas samlingar av dokumenttyper under en överkategori. Denna struktur visualiseras nedan i figur 1.



Figur 1. Diagram över strukturkravet på arkivet.

2.1.3 Granskning av uppladdat material

Arkivet skall erbjuda personal inom UEAB möjlighet att granska material samt dess metadata som laddas upp före det blir allmänt tillgängligt. Detta krav innebär även tekniskt att det måste finnas någon form av användarhantering inom arkivet, vilket då även ställts fram som ett krav ovan. Med personal avses främst föreläsare och examinatorer inom UEAB. Granskningsprocessen måste tillåta granskaren att avslå uppladdat material och tillåta användaren som laddat upp det att göra korrigerande förändringar både i/till materialet och dess metadata.

2.1.4 Sökning och visning av uppladdat material

Arkivet måste ha funktionalitet för sökning bland allt uppladdat material. Beställningen specificerade inte några exakta kriterier kring sökfunktionaliteten, men i samband med möten och diskussion fördes detaljer kring detta fram. Det bör finnas stöd för sökning av element i det uppladdade materialets metadata, såsom författare, nyckelord, publice-

ringsdatum mm. Man skall även kunna få förteckningar på allt material på basen av inom vilka kategorier samt avdelningar det laddats upp till arkivet.

Det är även viktigt att arkivet har ett mobiloptimerat gränssnitt för visning av innehållet i arkivet, eftersom en stor del av användarna kommer att använda sig av mobila apparater med små skärmar och låga överföringshastigheter på mobilnät. Det mobila användargränssnittet bör klara av sökning och visning av material som finns i arkivet, men uppladdning av material ställdes inte fram som ett krav.

2.1.5 Säkerhetskopiering

I den ursprungliga projektbeställningen behandlades inte frågor kring säkerhetskopiering, men det har lyfts fram av mig och ansetts vara en viktig del av helheten. Användarinformation, uppladdat material samt dess metadata bör säkerhetskopieras.

2.2 Krav på utseende

Arkivets utseende bör även vara modifierbart så det kan följa de krav som UEAB ställer kring arkivets titel samt utseendemässiga faktorer såsom logotyper, fonter, layout och färgval. Det har t.ex. specificerats att det finländska utbildningsministeriets stöd kring HEI CRIPS ICI projektet, och därmed också arkivet, konkret skall föras fram i form av en bild eller text.

3 ALTERNATIV OCH VAL AV ARKIVPLATTFORM

Den ursprungliga projektbeställningen förde fram arkivplattformen DSpace som grund för arkivet och gav inte från projektets början någon möjlighet till bruk av andra möjliga arkivplattformar. Under de första mötena med planeringsgruppen från UEAB fördes det dock fram förslag på några alternativ som man kunde ta en titt på: SubjectsPlus och IR Plus. Alla föreslagna plattformar har den gemensamma egenskapen att de distribueras gratis för alla bruksändamål.

3.1 SubjectsPlus

SubjectsPlus är ursprungligen utvecklat av Joyner-biblioteket vid East Carolina University och senare öppet utvecklat av individer vid bl.a. Ithaca College Library och University of Miami. [1] SubjectsPlus kan klassas som ett sökverktyg [2] och visade sig snabbt inte lämpa sig för projektets slutliga ändamål.

Plattformen fungerar snarare som ett index där man kan skapa listor som hänvisar till enskilda resurser eller samlingar av resurser inom exempelvis specifika kategorier. SubjectsPlus tillåter inte uppladdning av material i form av konkreta filer med metadata och är därmed i grund och botten inte heller ett arkiv.

3.2 IR Plus / IR+

Arkivplattformen IR Plus (eller IR+) bygger på öppen källkod, är utvecklat av individer vid University of Rochester, och beskriver sig själv som arkivmjukvara för institutioner. [3]

IR+ var ett förslag som kom mycket närmare projektbeställningens specifikationer. Mjukvarupaketet har stöd för användarkonton, uppladdning av filer tillsammans med metadata, granskning av uppladdat material och sökning i uppladdat material. Dock visade sig bl.a. granskningsprocessen inte vara uppdelad i faser i enlighet med beställarens funktionella krav.

3.3 DSpace

Arkivplattformen DSpace var det ursprungliga förslaget som projektbeställaren själv förde fram. Detta berodde till stor del på att UEAB:s bibliotekspersonal tidigare har jobbat med DSpace i andra sammanhang.

DSpace bygger också på öppen källkod och upprätthålls samt utvecklas av individer inom intressegruppen DURASPACE [4]. DSpace används inom ett stort antal organisat-

ioner och även av publikationsarkivet Theseus (<https://www.theseus.fi>) som har ett mycket liknande funktionalitetssyfte som detta arkivprojekt.

Plattformen innehåller alla de egenskaper som beställaren specificerat i sin kravspecifikation: användarhantering samt uppladdning, granskning, visning och sökning av dokument. Dessutom innebär den utsträckta användningen av DSpace att det finns mycket färdig dokumentation och allmän information om plattformen tillgänglig på nätet.

3.4 Slutligt plattformval

Arkivplattformerna SubjectsPlus och IR Plus saknade båda mer eller mindre den funktionalitet som beställaren önskade sig. SubjectsPlus var som tidigare nämnt inte en arkivplattform i sig, och det skulle vara resursmässigt meningslöst att försöka skriva om hela plattformen för att uppnå önskad funktionalitet. Även om IR Plus i varje fall var en arkivplattform, så saknade den som tidigare nämnt en del av önskvärd funktionalitet.

Som arkivplattform valdes till slut DSpace. Plattformen visade sig innehålla funktionalitet enligt beställarens krav och stöds av en välskött dokumentation som gör ett eventuellt modifikationsarbete enklare. DSpace var även som konstaterat beställarens främsta förslag från första början, eftersom personal vid UEAB bekantat sig med plattformen i ett tidigare skede.

4 DSPACE – INSTALLATION OCH KONFIGURATION

I detta kapitel behandlas de tekniska detaljerna såsom plattformens struktur och komponenter samt de krav som plattformen ställer på hårdvaran. Kapitlet behandlar installationsprocessen av själva arkivplattformen samt snuddar vid nödvändiga mindre modifikationer av den bakomliggande mjukvaran. Även om DSpace fyllt de flesta krav som beställaren ställt fram för arkivet, så är det som sådant trots allt inte perfekt och har därmed behövt modifiering. I detta kapitel ges även en grundläggande bild om var exempelvis utseendemässiga samt funktionella modifikationer i DSpace kan göras.

4.1 Krav

Arkivplattformen DSpace i sig är en form av webbapplikation som måste köras på underliggande mjukvara. På den officiella wikisidan för DSpace/Duraspace [5] med dokumentation uppräknas följande krav på mjukvaran och hårdvaran för att man skall kunna köra en DSpace-instans.

4.1.1 Mjukvarukrav

- Operativsystem: UNIX-baserat eller Microsoft Windows
- Java JDK/SDK: Oracle Java JDK 7 eller OpenJDK 7
- Java Build Tools: Apache Maven 3.x
- Java Build Tools: Apache Ant 1.8
- Relationsdatabas: PostgreSQL 8.4 till 9.1 eller Oracle 10g
- Servlet Engine: Apache Tomcat 7 eller senare

4.1.2 Hårdvarukrav

Dokumentationen nämner att DSpace i sin grundform kan köras på de flesta något moderna hårdvarukonfigurationer såvida mjukvarukraven uppfylls. [6] Ifall man ämnar köra DSpace för ett märkbart större användarantal så finns det tre olika konfigurationsförslag enligt mängden dokument som arkivet kommer att innehålla:

- Minimalnivå eller upp till 20 000 dokument:
 - 2 GB arbetsminne
 - 20 GB diskutrymme
- Medelnivå eller upp till 50 000 dokument:
 - 4 GB arbetsminne
 - 200 GB diskutrymme
- Högnivå eller mer än 50 000 dokument:
 - Fyrkärnig processor
 - 8 GB arbetsminne
 - NAS-diskutrymme med gigabitanslutning och RAID-konfiguration

Dokumentationen betonar upprepade gånger att ovannämnda rekommendationer är mycket grova uppskattningar och att de egentliga hårdvarukraven sist och slutligen beror på faktorer såsom antalet samtidigt aktiva användare, sökningar, nedladdningar och visningar av webbgränssnittet.

4.1.3 Kraven för DSpace-instansen vid UEAB

Det framkom inte några exakta eller uppskattade användarantal i projektbeställningen eller under möten med beställaren, men jag gjorde en personlig bedömning om att det totala antalet användare som kommer använda arkivet samtidigt i varje fall inte till en början kommer att överstiga 200. Därmed uppskattade jag också att mängden dokument i arkivet inte alltför snabbt kommer att nå det ovannämnda minimalkonfigurationsantalet på 20 000.

UEAB specificerade inom beställningen att de skulle erbjuda en server med hög prestanda som driftmiljö för det färdiga arkivet. De slutliga specifikationerna för hårdvaran framkom vid ett senare skede och blev följande: tvåkärnig processor på 2 GHz, 4 GB arbetsminne och 650 GB lagringsutrymme.

4.2 Struktur

DSpace är skrivet i Java och är i sin helhet en samling webbapplikationer som hanterar filförvaring och hantering av metadata, samt en kärna som binder ihop alla delar. Filer-na som laddas upp lagras på servern som kör DSpace och metadata lagras inom en relationsdatabas såsom Oracle 10g eller PostgreSQL.

De olika webbapplikationskomponenterna körs genom en servletmotor, såsom exempelvis Apache Tomcat, och hanterar de egentliga interna processerna inom arkivplattformen.

DSpace distribueras i två olika former: i ett ZIP-arkiv som innehåller de färdigt kompilerade webbapplikationskomponenterna, eller alternativt i ett arkiv som innehåller för komponenternas källkod som måste kompileras med kompileringsverktygen Apache

Maven och Apache Ant. Eftersom beställaren krävt vissa funktionella förändringar inom DSpace har jag valt att använda mig av distributionen som innehåller hela källkoden.

4.3 Komponenter

DSpace består alltså av ett antal komponenter i form av webbapplikationer som erbjuder helhetens funktionalitet. Ett antal av dessa hanterar inte direkt några funktionella uppgifter inom arkivet utan erbjuder istället extra funktionalitet för utvecklare i form av olika programmeringsgränssnitt. De olika DSpace-komponenterna presenteras kort nedan:

- **LNI** – *Lightweight Network Interface* är en form av nätverksgränssnitt som i sin grund är en del av DSpace API:n och ger fristående utvecklare möjligheten att skriva klienter som kan utföra operationer mot en DSpace-instans.
- **OAI** – Komponent som genom OAI-standarden tillåter delning av metadata mellan olika arkiv. [7]
- **REST** – Programmeringsgränssnitt för utvecklare som ger tillgång till material inom arkivet genom exempelvis fristående klienter istället för DSpace's egna webbgränssnitt.
- **SOLR** – SOLR är en av Apache Software Foundation utvecklad sökplattform som ansvarar för sökning och indexering av digitala resurser i form av en Java-webbapplikation. [8] Inom DSpace används SOLR främst som en bakomliggande infrastruktur för att skapa och hantera statistik kring användning av arkivet. Med detta menas alltså bl.a. annat antalet visningar samt nedladdningar för specifika dokument eller resurser inom arkivet. Denna komponent har visat sig vara viktig eftersom beställaren i samband med diskussioner visat intresse för möjligheten till noggrann användningsstatistik inom arkivet.
- **SWORD/SWORDv2** – Komponent som genom SWORD protokollet tillåter externa tjänster att mata in material i arkivet [9].
- **XMLUI och JSPUI** – Dessa är två olika variationer av användargränssnitt/webbgränssnitt som följer med DSpace-paketet. XMLUI bygger till största delen på Apache Cocoon-ramverket och består kodmässigt av XSL, XHTML,

CSS och även Java. JSPUI-gränssnittet bygger som namnet säger på JavaServer Pages, alltså JSP. [10]

4.4 Utvecklingsmiljön

Detta kapitel behandlar det tidiga installations- och konfigurationsarbetet samt miljöer som arbetets har utförts i. Utvecklingsarbetet har skett i två olika miljöer: på en bordsdator som fysiskt funnits vid Arcada och senare på en bordsdator vid UEAB i Kenya som jag har haft tillgång till genom fjärrskrivbordsanslutning (RDP).

I början av projektet bekantade jag mig främst med de av beställaren föreslagna olika arkivplattformerna som nämndes i kapitel 3 *Alternativ och val av arkivplattform*. Testningen av de olika plattformarna samt en stor del av den tidiga installationen och konfigurationen av DSpace gjordes på den bordsdatorn som fysiskt fanns vid Arcada.

Vid ett senare skede erbjöd UEAB en lokal dator inom deras campusområde där installations- och konfigurationsarbetet på en testinstans fortsatte. Testinstansen togs även i bruk för en mindre grupp användare för att ge personalen samt ett antal elever chansen att bekanta sig med arkivet.

4.4.1 Hård- och mjukvara

De två olika utvecklingsmiljöerna som har nämnts har kört på vanlig bordsdatorhårdvara som inte har överstigit normal konsumentnivå. DSpace-instansen har slutligen flyttats över till en hyllmonterad *bladeserver*.

Både den lokala datorn vid Arcada samt enheten vid UEAB har haft en Windows-baserad operativsystemsmiljö. Vilket operativsystem som används har inte någon större inverkan på den egentliga DSpace-instansen eftersom den underliggande mjukvaran finns tillgänglig för både Windows- och Unix-baserade operativsystem.

4.4.2 Installation och konfiguration av underliggande mjukvara

Mjukvarukraven för DSpace erbjöd ett antal alternativ vad gällde vissa underliggande delar såsom Java-plattform och relationsdatabas. Nedan presenteras mjukvarukonfigurationen för utvecklingsmiljön vid UEAB:

- Operativsystem: Microsoft Windows Server 2003 SP3
- Java JDK/SDK: Oracle Java JDK 7
- Java Build Tools: Apache Maven 3.x
- Java Build Tools: Apache Ant 1.8
- Relationsdatabas: PostgreSQL 8.4
- Servlet Engine: Apache Tomcat 7

Några större konfigurationsförändringar i dessa underliggande mjukvaror har inte gjorts. Däremot har dessa mindre modifieringar gjorts:

- Oracle Java JDK 7:
 - Tillsätta miljövariabler som hänvisar till Java JDK:ns hemkatalog
- Apache Maven och Apache Ant:
 - Tillsätta miljövariabler som hänvisar till Apache Mavens binärfiler så de kan köras inom Windows kommandotolk genom kommandot *maven* samt *ant*
- Apache Tomcat 7:
 - Miljövariabler för att allokeras mera minne till Tomcat-instansen i enlighet med DSpace-dokumentationens rekommendationer.
 - Konfigurera Tomcat så roten för webbapplikationer ligger i installationsmappen för DSpace
 - DSpace kräver även att UTF-8 används för filkodning, så detta måste även tas i bruk

4.4.3 Grundinformation om DSpace-installationen

Installationen av DSpace på Unix-baserad hårdvara är väldokumenterad i Duraspace's egna dokumentation. Installationsprocessen på Windows-baserad hårdvara skiljer sig

inte mycket från standardinstruktionerna och dokumentationen [11] betonar även de saker man borde tänka på ifall installationen görs på en Windows-plattform.

Detta delkapitel behandlar inte i någon större detalj installationen av den underliggande mjukvaran såsom exempelvis Oracle Java JDK eller den underliggande relationsdatabasen. Detta beror till stor del på att nämnda mjukvara till största del fungerar på önskat sätt med standardinställningar. Trots allt så kommer eventuella konfigurationsändringar inom den underliggande mjukvaran att kort behandlas ifall förändringen är nödvändig för funktionaliteten i en DSpace-instans.

Den tidigaste testinstallationen av DSpace gjordes med version 3.2 av DSpace. Det fortsatta utvecklingsarbetet på hårdvaran vid UEAB baserade sig på version 4.0 av DSpace. Den största skillnaden mellan dessa versioner var högre versionsnummer på den underliggande mjukvaran. Som tidigare nämnts erbjuder Duraspace en färdigt kompilerad version av DSpace som tillåter enkel modifikation av utseendemässiga faktorer såsom arkivets webbgränssnitt. För detta projekt har jag ändå använt mig av den s.k. ”source releasen” som före kompilering tillåter modifikation av källkod inom alla moduler som DSpace bygger på.

4.4.4 Installation av DSpace kortfattat

Installationen och ibruktagandet av en DSpace-instans går lätt och smidigt såvida all underliggande mjukvara fungerar som den skall och eventuella kommunikationshinder mellan servern som huserar instansen och Internet är eliminerade.

Installationen av både den till största delen färdigt kompilerade versionen och ”source releasen” följer samma steg för en basinstallation i enlighet med Duraspace:s dokumentation: [12]

1. Skapa en dedikerad användare inom operativsystemet för den bakomliggande mjukvarans samt DSpace’s processer och tjänster
2. Skapa en mapp där DSpace-instansens programfiler kommer finnas
3. Skapa användare samt databas för DSpace inom vald relationsdatabasmjukavara

4. Packa upp arkivet som innehåller källkodsfilererna för vald DSpace-instans
5. Modifiera förkonfigurationsfilen *build.properties* inom källkodskatalogen så att dess parametrar följer värden enligt tidigare steg, exempelvis information om relationsdatabasen
6. Kompilera källkoden i källkodskatalogen till ett installationspaket med Apache Maven
7. Installera det skapade installationspaketet med Apache Ant, som även skapar databasstrukturen till den valda relationsdatabasen
8. Flytta över Dspace-webbapplikationerna från installationsmappen till den valda webbapplikationsserverns rotkatalog för webbapplikationer ELLER
9. Hänvisa webbapplikationsservern till DSpace-installationens katalog som innehåller de kompilerade webbapplikationerna

Om installationen görs på hårdvara som använder sig av Windows bör man i konfigurationsfiler notera att man i sökstigar, trots operativsystemets praxis, skall använda sig av `"/` istället för `"\`.

4.4.5 Installation av DSpace stegvis

Detta delkapitel går igenom DSpace-installationen i mera detaljerade steg med hjälp av skärmdumpar och kommentarer. Eftersom målet är att modifiera utseende och funktionalitet till en viss grad så använder jag som tidigare konstaterat mig av hela källkodspaketet för Dspace. Nedan presenteras installationsprocessen stegvis.

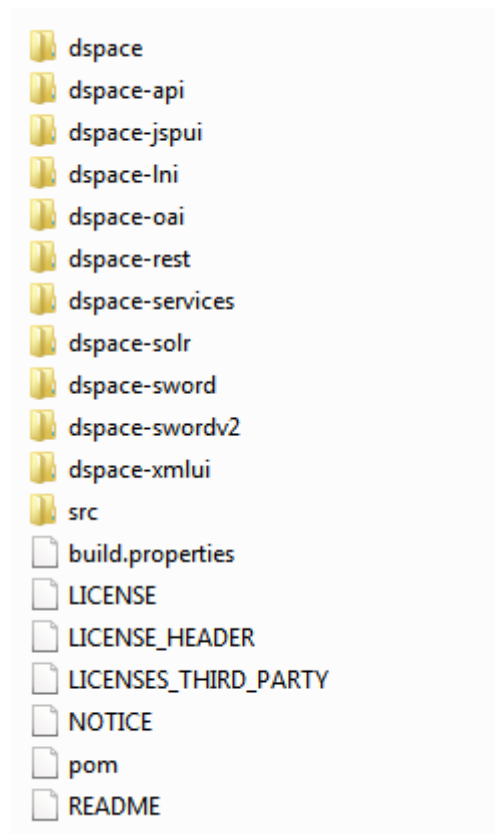
Skapa användare och tabeller i databasen

För denna installation bestämde jag mig för att använda PostgreSQL som bakomliggande relationsdatabas för DSpace. Detta gjordes till stor del på basen av en rekommendation från min handledares sida.

Inom databasen har jag skapat en databasanvändare med namnet *dspace* samt en databas med namnet *dspace* i enlighet med DSpace's egen dokumentation gällande installation. Användaren *dspace* skall även specificeras som ägare för den skapade *dspace*-databasen.

Uppackning av källkodspaket och dess struktur

Denna installation av DSpace använder sig av källkodspaketet *dspace-4.0-src-release* som distribueras i form av ett zip-arkiv som först måste packas upp. Paketets struktur presenteras nedan i figur 2.



Figur 2. DSpace-källpaketets struktur.

Modifiera förkonfigurationsfilen enligt installationsmiljön

Inom källpaketets struktur finns filen *build.properties* som används för att ange parametrar för DSpace-installationen. Där definieras bland information om relationsdatabasen samt vilken e-postserver som skall användas för systemmeddelanden. Övriga viktiga parametrar är t.ex. var själva DSpace-installationen skall ligga på maskinen samt vilken domän instansen kommer vara bunden till ute på nätet.

Även om man utför kompileringen och installationen med givna värden i *build.properties*-filen, kan man efter installation ändra på dem genom att modifiera DSpace-instansens huvudkonfigurationsfil *dspace.cfg* som behandlas i nästa delkapitel 4.4.6 *Modifikation av DSpace*.

I figur 3 nedan kan man se exempel på strukturen i förkonfigurationsfilen *build.properties* som följer samma modell som den nämnda huvudkonfigurationsfilen.

```
#####
# SERVER CONFIGURATION #
#####

# DSpace installation directory. This is the location where you want
# to install DSpace. NOTE: this value will be copied over to the
# "dspace.dir" setting in the final "dspace.cfg" file. It can be
# modified later on in your "dspace.cfg", if needed.
dspace.install.dir=C:/dspace

# DSpace host name - should match base URL. Do not include port number
dspace.hostname = localhost

# DSpace base host URL. Include port number etc.
dspace.baseUrl = http://localhost:8080

# Name of the site
dspace.name = DSpace Repository

# solr server
solr.server=http://localhost:8080/solr

# Default language for metadata values
default.language = en_US

#####
# DATABASE CONFIGURATION #
#####

# Database name ("oracle", or "postgres")
db.name=postgres

# Uncomment the appropriate block below for your database.
# postgres
db.driver=org.postgresql.Driver
db.url=jdbc:postgresql://localhost:5432/dspace
db.username=dspace
db.password=dspace
```

Figur 3. Utdrag ur förkonfigurationsfilen *build.properties*.

Kompilera installationspaketet från källkodspaketet med Apache Maven

Att kompilera DSpace-installationspaketet är en enkel process som i grund och botten innebär att köra Apache Maven-kompilatorn med *package*-parametern i DSpace-

källkodspaketets rot i enlighet med figur 4 nedan. Jag valde som tidigare nämnts att inkludera *mvn*-kompilatorn i operativsystemets miljövariabler.

```
C:\dSPACE-4.0-src-release>mvn package
```

Figur 4. Apache Maven-kompileringskommandot i rotkatalogen i Windows-miljö.

Installera med kompilerat installationspaket

Apache Maven kompilerar källan och skapar bland annat katalogen */dSPACE/target/dSPACE-4.0-build* i källkodspaketets rot inom vilken kompilerad kod ligger. I nämnda mapp kör man till nästa Apache ANT-installeraren med parametern *fresh_install*. Installationen skapar DSpace-instansen i filsystemet i enlighet med vilket installationsställe man angivit i förinstallationsfilen.

Man bör lägga märke till att parametern *fresh_install* enbart används då man gör den första installationen av DSpace-instansen, eftersom parametern leder till att Apache ANT-installeraren även gör förarbete i databasen man specificerat för DSpace. Vid vidareutvecklande förändringar i källpaketet skall man använda sig av parametern *update* för Apache ANT, varvid behandlas enbart de förändringar man gjort i källkoden och därmed sparas tid samt den redan skapade databasen förstörs inte.

Ta i bruk webbapplikationerna inom webbapplikationsservern

Installationen kompilerar alla de tidigare nämnda komponenterna som följer med DSpace till webbapplikationer för Apache Tomcat. Jag har valt att selektivt ta i bruk enbart de komponenter som är relevanta för instansen inom ramarna för detta projekt. Dessa är specifikt användargränssnittet XMLUI, indexerings- och statistikmodulen SOLR samt OAI.

Webbapplikationerna kan som tidigare nämnts tas i bruk på två olika sätt: kopiera över dem till Apache Tomcats webbapplikationsmapp eller konfigurera s.k. *contexts* inom Tomcat vilka hänvisar Tomcat till att ladda webbapplikationerna från ett spe-

cificerat ställe. För att göra eventuella mindre förändringar och därmed omkompileringar enklare har jag valt att hänvisa Apache Tomcat till att söka webbapplikationerna från DSpace-installationens *webapps*-mapp. Detta gör man som sagt genom att skapa samt modifiera XML-filer i Apache Tomcats installationsmapp, mera specifikt i *conf/Catalina/localhost*, där jag i mitt fall har skapat filerna *oai.xml* och *solr.xml* samt *ROOT.xml*. Modellinnehållet i *ROOT.xml* presenteras nedan i figur 5.

```
<?xml version='1.0'?>
<Context
  docBase="C:/dspace/webapps/xmlui"
  debug="0"
  reloadable="true"
  cachingAllowed="false"
  allowLinking="true"/>
```

Figur 5. Innehållet i *ROOT.xml*.

Som man kan observera ovan, anger filen i fråga alltså helt enkelt var webbapplikationsroten för Apache Tomcat-webbservern ligger. I detta fall leder alltså HTTP-förfrågan från en webbläsare mot arkivserverns IP-adress till DSpaces XMLUI-användargränssnitt. På liknande sätt binder *oai.xml* och *solr.xml* respektive webbapplikationer till */oai* samt */solr* i arkivets domän.

Efter dessa steg är arkivet inom en testmiljö tillgängligt genom att använda en webbläsare för att komma åt exempelvis *http://localhost* eller *http://localhost/oai* samt *http://localhost/solr* på själva utvecklingshårdvaran. Alternativt nås arkivet från Internet genom serverns IP-adress eller publika nätomän om man bundit webbapplikationsservern och DSpace till dem i samband med konfigurationen.

4.4.6 Modifikation av DSpace

En nyinstallerad omodifierad DSpace-instans uppfyller till stor del de funktionella kraven som förts fram av UEAB och som behandlades i kapitel 2 *Beställarens krav*. Majoriteten av ändringarna som har behövts göras på installationen har varit utseendemässiga förändringar och ibrukttagandet av ett mobilt användargränssnitt. Inom detta delkapitel behandlas kortfattat hur man kan modifiera DSpace-instansen enligt eget tyckte.

Metoden jag använt mig av för att utföra modifikationer är förändringar i filer inom källkodspaketet *dspace-4.0-src-release*, som använts för ursprunglig installation av DSpace-instansen, och genom omkompilering av paketet med *mvn package* samt *ant update* för att enbart uppdatera förändringarna i DSpace.

Modifikation av utseende

Som användargränssnitt för DSpace erbjuds det ett val mellan XMLUI och JSP, vilket presenterades i kapitel 4.3 *Komponenter*. XMLUI verkar numera vara det mer använda användargränssnittet och är även det gränssnitt som används som standard i en omodifierad installation av DSpace. XMLUI-användargränssnittet i sig är en stomme som erbjuder grundfunktionalitet för användning av arkivet. XMLUI har stöd för att man bygger så kallade teman, som har den egentliga uppgiften att ge arkivets användargränssnitt ett specifikt utseende och struktur. Alla teman inom XMLUI finner man i komponentens mapp i *dspace-4.0-src-release/dspace-xmlui/src/main/webapp/themes*, där de är uppdelade i egna undermappar.

Ett tema i sig bygger till stor del på XSL-/XML- och CSS-filer, som anger hur arkivets användargränssnitt ser ut både stilistiskt och strukturellt. Som standardtema använder XMLUI temat Mirage, som jag även valde att modifiera i enlighet med beställarens önskemål. Med XMLUI medföljer även ett färdigt tema som är gjort specifikt för användning på mobila enheter, vilket jag också tog i bruk samt modifierade per beställarens krav.

Konfigurationsförändringar

Den grundläggande konfigurationen av DSpace görs i samband med installationen som demonstrerades i föregående kapitel 4.4.5 *Installation av DSpace stegvis*. DSpace-konfigurationsfiler är dock även modifierbara efter installation och ligger i DSpaces installationsmapp under *dspace-4.0-src-release/dspace/config* och är i sig strukturerade textfiler.

Inom konfigurationsmappen hittar man bl.a. filer som definierar vilka metadatafält man kan fylla i vid uppladdning av nya dokument och vilket innehåll arkivets automatiserade e-postmeddelanden har. Dessa är exempel på konfigurationsändringar jag gjort i enlighet med beställarens önskemål.

Den största och främsta filen i konfigurationsmappen är dock *dspace.cfg*, som innehåller konfigurationsparametrar för DSpace i allmänhet samt parametrar för XMLUI- och JSP-användargränssnitten. Användargränssnittens parametrar omfattar inställningar som gäller för alla enskilda teman inom användargränssnitten.

4.5 Produktionsmiljön

Planen var som tidigare nämnts att ha arkivet på en server som erbjuder hög prestanda, vilket betydde hyllmonterad serverhårdvara. Som tidigare nämnts hade även utvecklingsinstansen av DSpace redan tagits i mindre bruk, vilket innebar att migration av data till en annan miljö var något man måste beakta. Detta delkapitel jämför de olika hårdvarumiljöerna samt behandlar migrationsprocessen från utvecklingsmiljön till produktionsmiljön.

4.5.1 Hård- och mjukvara

Under tidiga möten med arbetsgruppen från UEAB presenterades planen att den slutgiltiga hårdvaran för arkivet skulle ha ett Unix-baserat operativsystem. När jag slutligen fick ett meddelande om att den egentliga hårdvaran var klar för användning visade det sig dock att hårdvaran körde *Microsoft Windows Server 2012 R2*. Detta hade dock inte någon större betydelse, eftersom installationsprocessen och konfigurationen av DSpace i både Windows- och Unix-baserad miljö sist och slutligen är mycket lika varandra.

4.5.2 Migration

Eftersom utvecklingsversionen av DSpace tagits i testbruk under utvecklingen, var det önskvärt att även migrera all användarinformation, arkivstruktur samt uppladdade dokument till den slutgiltiga servern. Migrationsprocessen presenteras punktvis nedan.

Installation och konfiguration av underliggande mjukvara

Apache Tomcat, PostgreSQL samt Apache ANT och Apache Maven installeras och konfigureras på samma sätt som den ursprungliga installationen inom utvecklingsmiljön i enlighet med kapitel 4.4.2 *Installation och konfiguration av underliggande mjukvara* med vissa parameterundantag i enlighet med den nya miljön.

Kopiera över det modifierade DSpace-källkodspaketet

Eftersom alla modifieringar inom DSpace-instansen på utvecklingsmiljön gjorts genom att modifiera källkodspaketet *dspace-4.0-src-release*, behöver man enbart kopiera över och använda det modifierade paketet vid installationen av DSpace-instansen på den nya hårdvaran.

Kompilera och installera DSpace-instansen

Installationen av DSpace-instansen görs i enlighet med samma steg som presenterats inom kapitel 4.4.5 *Installation av DSpace stegvis* med nödvändiga parameterändringar i enlighet med den nya miljön. Efter installation lönar det sig att kontrollera ifall den rena installationen fungerar så som den skall. Apache Tomcat stängs ner för import av filförråd och databas i följande steg.

Kopiera över filförrådet från DSpaces installationsmapp

Filer som laddas upp till DSpace lagras i mappen *assetstore* som ligger i DSpace-installationens rotkatalog. Denna mapp kan rakt kopieras över och ersätta *assetstore*-mappen i den nya installationen.

Dumpa databasen från utvecklingsmiljön och återställ den i produktionsmiljön

Den relationella databasen som hanteras av PostgreSQL innehåller all användarinformation, strukturdata samt metadata för uppladdade dokument. Överflyttningen av denna databas bestämde jag mig för att göra genom att använda PostgreSQL-

funktionen *pg_dumpball*, som dumpar all databasinformation i en fil som sedan kan användas för att återställa databasen. Nedan i figur 6 visas kommandot samt dess parametrar. Det som anges i kommandot är databasanvändare och vilket filnamn databasen skall dumpas till; som exempel har namnet *DBdump* använts.

```
C:\dbDump>pg_dumpall -U postgres > DBdump
```

Figur 6. Exempel på användning av PostgreSQL-funktionen *pg_dumpall* för att dumpa databasinnehåll.

Återställningen av databasinnehållet på produktionsmiljön kan utföras med PostgreSQL-funktionen *psql*, som använder sig av den dumpade filen för att återskapa databasstrukturen. Före man återställer dumpen bör man se till att man har skapat databasanvändare enligt stegen i kapitel 4.4.5 *Installation av DSpace stegvis* samt tagit bort *dspace*-databasen som den rena installationen skapat för att inte stöta på felmeddelanden om att tabeller redan finns. Nedan i figur 7 presenteras hur *psql* kommandot används på produktionsmiljön för att återskapa PostgreSQL-databasstrukturen för DSpace. På samma sätt som *pg_dumpall*-kommandot så tar även *psql*-kommandot som parametrar databasanvändare och namnet på filen som dumpats.

```
C:\dbDump>psql -U postgres -f DBdump postgres
```

Figur 7. Exempel på användning av PostgreSQL-funktionen *psql* för att återställa databasdump.

Starta Apache Tomcat samt kör DSpace's indexeringskommandon

Efter att filförråd samt databas har flyttats över återstår att köra igång webbservern. För att DSpace skall visa det återställda innehållet måste man även köra indexeringskommandon som visas i figur 8 med hjälp av DSpace's eget kommando-promptsverktyg *dspace.bat*, som ligger i DSpace's installationskatalog i mappen *bin*.

```
C:\dspace\bin\dspace.bat index-discovery
C:\dspace\bin\dspace.bat index-lucene-init
C:\dspace\bin\dspace.bat index-lucene-update
```

Figur 8. Indexeringskommandon i *dspace.bat* som måste köras efter migration.

Indexeringen körs alltså genom att ange kommandona som parametrar till *dspace.bat* genom Windows' egen kommandotolk. Dessa kommandon leder till att DSpace-instansens interna sök- och bläddringsindex uppdateras och reflekterar det aktuella innehållet i arkivets filförråd och databas.

5 ARKIVETS DATASÄKERHET

I detta kapitel behandlas hur de naturliga datasäkerhetsfrågor som förekommit i samband med arkivet kan lösas. Dessa är bl.a. säkerhetskopieringen av arkivets filinnehåll och databas samt datasäkerheten i kommunikationen mellan klienter och DSpace-instansen.

5.1 Säkerhetskopiering

Även om säkerhetskopiering som tidigare nämnts inte ingick i beställarens ursprungliga krav så fanns nog tanken om säkerhetskopiering inom mig som en självklarhet. Ämnet togs förväntat upp av beställaren vid ett senare skede.

All data inom en DSpace-instans kan i sin grund hittas uppdelad på två olika ställen. Användardata, arkivets struktur och uppladdade dokumenters metadata finner man i den relationella databasen som används av DSpace. De konkreta uppladdade filerna förvaras i en separat mapp i DSpace-installationens rotkatalog. Detta innebär alltså att det som måste säkerhetskopieras är relationsdatabasen samt filförrådet. Processen är till stor del mycket lik de steg som måste utföras vid överflyttningen av utvecklingsmiljöns innehåll till produktionsmiljön, vilket behandlades i kapitel 4.5.2 *Migration*.

5.1.1 Säkerhetskopiering av filförråd

Inom DSpace-installationens mapp finner man mappen *assetstore* som mappseparerat innehåller alla de filer som laddats upp i samband med insättningen av nya dokument till arkivet. Säkerhetskopieringen av detta innehåll kan utföras genom att köra ett enkelt skript som gör en daterad kopia av filförrådsmappen. Skriptet körs regelbundet med hjälp av Windows-miljöns egen *Task Scheduler*.

5.1.2 Säkerhetskopiering av relationsdatabas

Säkerhetskopieringen av relationsdatabasen kan utföras på ett liknande sätt som säkerhetskopieringen av filförrådet i och med att ett regelbundet skript körs med hjälp av Windows' inbyggda *Task Scheduler*. PostgreSQL-dokumentationen för säkerhetskopiering och återställning av databaser presenterar ett antal inbyggda kommandon som kan användas för denna process [13]. Eftersom databasen enbart används för DSpace-instansen har jag bestämt mig för att använda mig av PostgreSQL-funktionen *pg_dumpall* som lagrar all databasinformation inom en valfritt namngiven fil.

Återställningen av innehållet görs med *psql*-funktionen som användes i samma syfte som vid migrationen av arkivet inom i kapitel 4.5.2 *Migration*. Efter återställningen måste man även köra de indexeringskommandon som behandlades i migrationsprocessen.

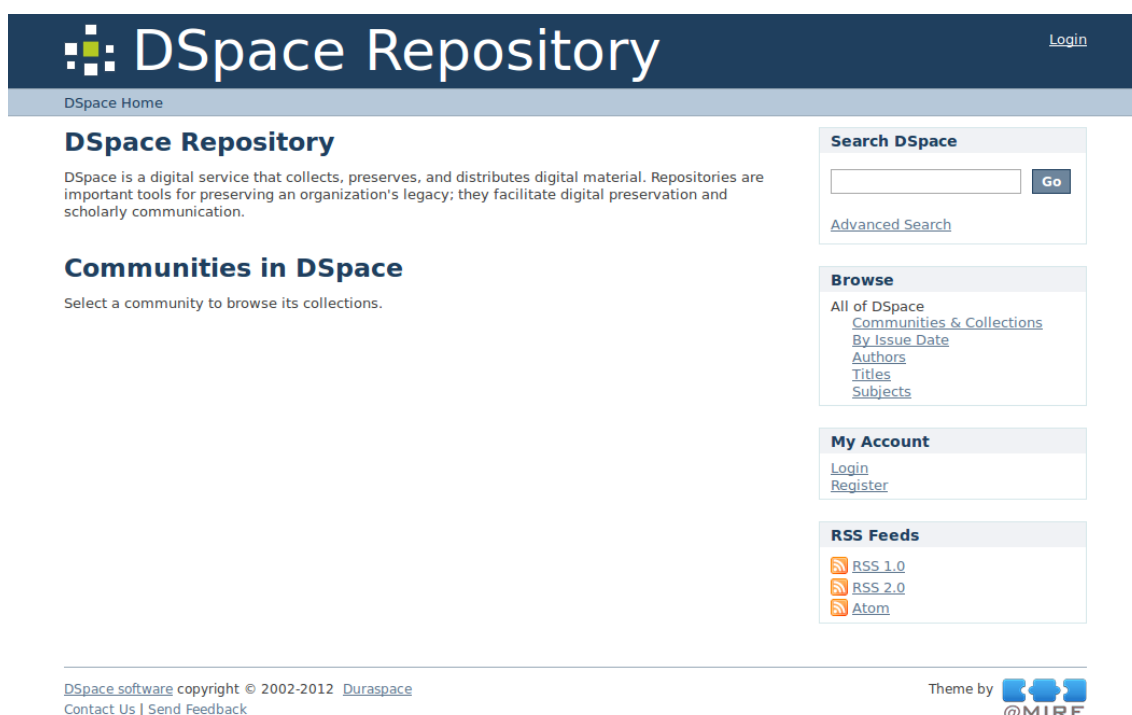
5.2 Datasäker anslutning

För att kunna göra annat i arkivet än att bara bläddra igenom material måste man som sagt skapa ett användarkonto som består av ett användarnamn och lösenord. Detta innebär att det är rekommenderbart att anslutningen mellan användaren och arkivet är krypterad i någon form, så att inloggningsinformationen inte kan uppsnappas och avläsas av externa parter.

En metod för att säkra anslutningen mellan användaren och arkivet är att ta i bruk HTTPS-anslutningar till webbapplikationerna inom Apache Tomcat som arkivet körs på. Processen för detta är välbeskriven i Apache Tomcat's egen dokumentation. [14]

6 RESULTATREDOVISNING

Som tidigare konstaterats i denna rapport, erbjöd DSpace genom sin basinstallation till stor del den önskade funktionaliteten samt en mycket bra stomme för användargränssnittet. Användargränssnittets tema i sig baserade sig på det nämnda *Mirage*-temat som används som standard för en omodifierad DSpace-installation genom XMLUI-användargränssnittkomponenten. I figur 9 nedan visas det oförändrade Mirage-temat för det fulla webbgränssnittet.



Figur 9. Bild på det oförändrade Mirage-temats framsida.

Mirage-temat modifierades i enlighet med beställarens önskemål. Dessa modifikationer handlade främst om färgtema, textinnehållsförändringar samt inplacering av extra element såsom bilder. Resultatet presenteras nedan i figur 10.

UNIVERSITY OF EASTERN AFRICA BARATON
Your Gateway to Wholistic Christian Education

Home Login

SUPPORTED BY OFFICIAL DEVELOPMENT AID FROM THE MINISTRY FOR FOREIGN AFFAIRS OF FINLAND

UEABSpace

The UEAB Institutional Repository called UEABSpace is an online collection of the UEAB research output. The Library collects, manages, and provides open access to the aggregated research projects of the UEAB community. We desire, as a learning community, to foster the open development and exchange of ideas by providing open access to our academic output. This free learning environment promotes access to our collections, and indicates in a tangible way, our social and academic impact in the world. We believe that sharing our research output with the world will increase the UEAB visibility, status and global recognition.

UEABSpace is supported by official development aid from the Ministry for Foreign Affairs of Finland

Communities in repository

Select a community to browse its collections.

- [Crisis preparedness & Global health care](#)
- [School of Education](#)
- [School of Humanities and Social Sciences](#)
- [UEAB REPOSITORY](#)

Recently Added

[Disaster Plan](#)
ohanga, christine; mwakesi, festus (2015-01-20)

[Disaster Plan for Moi Teaching and Referral Hospital](#)
ohanga, christine; mwakesi, festus (2015-01-20)

[Emergency Plan In a Local Health Facility in Rural Kenya](#)
Adega-Ayodo, Dorcas E. A. (2015-01-08)
The article describes the emergency plan for a sub health centre in Rural Kenya. With limited resources in terms of personnel and finance, they make good use of locally available resources and ensure a referral system to ...

[Ebola Epidemic Contingency Plan, Kenya](#)
Waihaka, Angela A. (2014-12-20)

[Emergency plan in Hospital of Nepal](#)
Khanal, Prakash (2014-12-14)

[View more](#)

Search repository

Browse

All of repository

- [Communities & Collections](#)
- [By Issue Date](#)
- [Authors](#)
- [Titles](#)
- [Subjects](#)

My Account

- [Login](#)
- [Register](#)

Discover

Author

- [Khanal, Prakash \(5\)](#)
- [Adega-Ayodo, Dorcas \(4\)](#)
- [Lax, Paula \(4\)](#)
- [Marshalls, Tuula J. \(4\)](#)
- [Kauppi, Susanna \(3\)](#)
- [muthumbi, kenneth \(3\)](#)
- [Nilsson, Ralli \(3\)](#)
- [Ohanga-Too, Appelles \(3\)](#)
- [Sandström, David \(3\)](#)
- [Waihaka, Angela A. \(3\)](#)
- ... [View More](#)

Subject

- [Kenya \(6\)](#)
- [Kendu Bay \(3\)](#)
- [accident \(2\)](#)
- [causes \(2\)](#)
- [culture \(2\)](#)
- [health system \(2\)](#)
- [interventions \(2\)](#)
- [motorcycle \(2\)](#)
- [Academic performance \(1\)](#)
- [aid worker \(1\)](#)
- ... [View More](#)

Date Issued

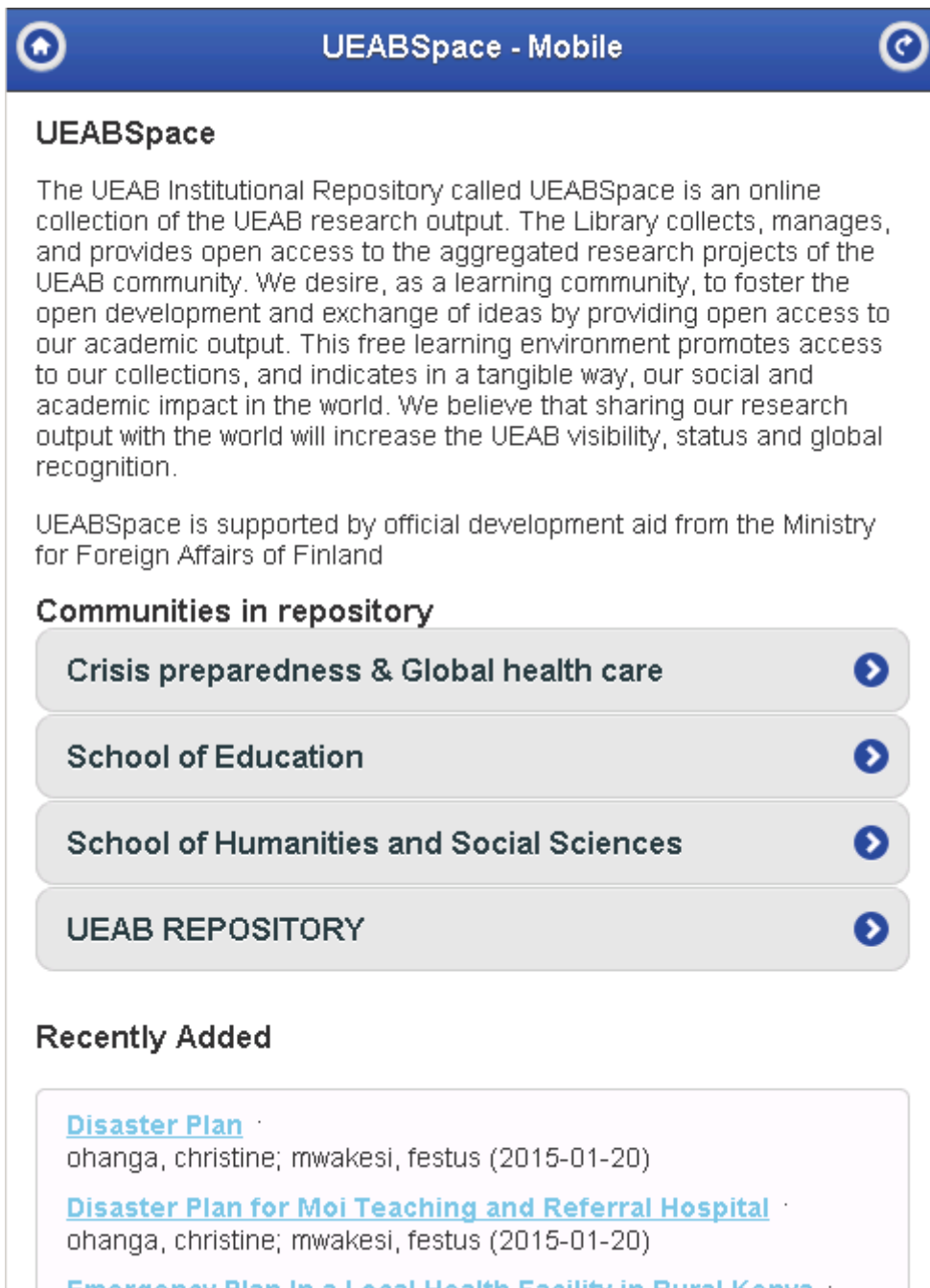
- [2010 - 2015 \(59\)](#)
- [2003 - 2009 \(1\)](#)

[DSpace software](#) copyright © 2002-2012 [Duraspace](#) Theme by

[Contact Us](#) | [Send Feedback](#) | [Mobile Site](#)

Figur 10. Arkivets framsida med modifierat Mirage-tema.

Som ytterligare krav hade beställaren även som tidigare nämnts ett mera minimalt användargränssnitt som skulle vara lämpligt att använda på mobila enheter såsom smarttelefoner. Detta var som tidigare konstaterat inte aktiverat i en omodifierad DSpace-installation och krävde konfigurationsförändringar för att tas i bruk. Efter förändringarna var det möjligt att komma åt ett mera mobilvänligt gränssnitt som presenteras i figur 11 nedan.



Figur 11. Modifierat mobiltema i enlighet med Mirage-temats förändringar.

Även arkivets struktur vad beträffar indelning av dokumenten och material som laddas upp har modifierats enligt beställarens önskemål och kan visuellt observeras i figur 10 ovan. Denna struktur kommer troligen leva och växa vidare enligt UEAB:s eget tycke.

Som slutresultat har jag alltså lyckats installera och konfigurera ett dokumentarkiv som till stor del uppfyller beställarens utseende- och funktionsmässiga krav. Det återstår ett

antal mindre önskemål kring förändringar från beställarens sida vilka framkommit vid ett mycket sent skede och därmed på grund av tidsbrist fallit utanför denna rapport.

7 SLUTSATSER OCH DISKUSSION

Detta examensarbete har varit ett intressant och annorlunda projekt att jobba med. Personligen tvekade jag till en början ifall man på basen av ett installations- och konfigurationsprojekt kunde få ihop ett examensarbete; tvekandet visade sig vara onödigt.

Under arbetets lopp har de främsta utmaningarna förekommit i samband med kommunikationen med beställaren och leveransen av hårdvaran för arkivet. Kommunikationshindren har till stor del handlat om den använda kommunikationskanalens, e-postens, tendens att nu och då leda till fördröjda svar på frågor från båda hållen. Den annorlunda naturen inom logistiken och handeln i beställarens miljö har bl.a. inneburit längre leveranstider av t.ex. arkivets serverhårdvara. Serverhårdvaran fördröjdes bland annat på grund av att det skickades ut tre olika offerter och att köpbeslut kunde göras först efter att alla offerter besvarats. Det tog i sin tur då längre än förväntat för svaren på offerterna att anlända, vilket fördröjde leveransen av hårdvaran. Dessa faktorer har därmed inneburit vissa fördröjningar i själva installations- och konfigurationsarbetet. Dessa utmaningar antar jag att kan räknas som typiska då man arbetar inom ett internationellt projekt.

Trots utmaningarna som nämnts har slutresultatet blivit ett fungerande dokumentarkiv som i detta skede till största del uppfyller beställarens krav och behov. Arbetsprocessen som detta examensarbete har velat realisera har alltså lyckats.

KÄLLOR

- [1] SubjectsPlus.com, 2014, *About*, Tillgänglig:
<http://www.subjectsplus.com/#section-5>, hämtad: 23.11.2014
- [2] EIFL.net, 2014, *SubjectsPlus Step-by-Step Guide*, Tillgänglig:
<http://www.eifl.net/resources/subjectsplus-step-step-guide>, hämtad: 24.3.2015
- [3] irplus, 2014, *Development And Testing*, Tillgänglig:
<https://code.google.com/p/irplus/wiki/DevelopmentAndTesting>, hämtad: 25.11.2014
- [4] Duraspace, 2014, *About DSpace*, Tillgänglig:
<http://www.dspace.org/introducing>, hämtad: 11.2.2015
- [5] Duraspace, 2014, *Dspace Wiki – Installing DSpace – Prerequisite Software*,
Tillgänglig:
<https://wiki.duraspace.org/display/DSDOC4x/Installing+DSpace#InstallingDSpace-PrerequisiteSoftware>, hämtad: 26.11.2014
- [6] Duraspace 2014, *Dspace Wiki – EndUserFaq*, Tillgänglig:
<https://wiki.duraspace.org/display/DSPACE/EndUserFaq>, hämtad: 11.2.2015
- [7] Openarchives.org, 2015, *Open Archives Initiative Protocol for Metadata Harvesting*, Tillgänglig: <http://www.openarchives.org/pmh/>, hämtad: 15.2.2015
- [8] Wikipedia, 2014, *Apache Solr*, Tillgänglig:
http://en.wikipedia.org/wiki/Apache_Solr, hämtad: 27.11.2014
- [9] SWORD, 2015, *About Sword*, Tillgänglig: <http://swordapp.org/about/>, hämtad: 15.2.2015

- [10] Duraspace, 2014, *Dspace Wiki – XMLUI Configuration and Customization*, Tillgänglig: <https://wiki.duraspace.org/display/DSDOC4x/XMLUI+Configuration+and+Customization>, hämtad: 27.11.2014
- [11] Duraspace, 2014, *Dspace Wiki – Installing Dspace – Windows Installation*, Tillgänglig: <https://wiki.duraspace.org/display/DSDOC4x/Installing+DSpace#InstallingDSpace-WindowsInstallation>, hämtad: 2.12.2014
- [12] Duraspace, 2014, *Dspace Wiki – Installing Dspace*, Tillgänglig: <https://wiki.duraspace.org/display/DSDOC4x/Installing+DSpace#InstallingDSpace-InstallationInstructions>, hämtad: 3.12.2014
- [13] The PostgreSQL Global Development Group, 2015, *PostgreSQL 9.4.1 Documentation – Chapter 24. Backup and Restore*, Tillgänglig: <http://www.postgresql.org/docs/9.4/static/backup.html>, hämtad: 14.2.2015
- [14] The Apache Software Foundation, 2015, *SSL/TLS Configuration HOW-TO*, Tillgänglig: <http://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>, hämtad: 2.3.2015

BILAGOR

Bilaga 1 – Thesis Work Order

Thesis work order

Information regarding client:

Organisation	CRIPS HEI ICI (Diak, Arcada, University of Baraton)
Contact	Thomas Forss
Telephone number	+358503435269
E-mail	thomas.forss@arcada.fi

Type of thesis:

Business – or marketing plan	<input type="checkbox"/>
Artistic work	<input type="checkbox"/>
Practical work	<input checked="" type="checkbox"/>
Product development	<input type="checkbox"/>
Service development	<input type="checkbox"/>
Investigation work	<input type="checkbox"/>
Scientific work	<input type="checkbox"/>

Time limits and payment:

Start date	16.09.2013
End date	01.03.2014
Practical work start date	16.09.2013
Practical work end date	03.12.2013
Payment:	
Paid work	<input checked="" type="checkbox"/>
Unpaid work	<input type="checkbox"/>

Description of the thesis work:

A high performance server will be acquired. The server will be hosting a repository based of the open source software DSpace (<http://www.dspace.org/>). The student will be required to install and configure the Dspace system. Additionally the student will have to configure a mobile version of the Dspace system that has special restrictions. The mobile version should be able to be accessed by slow internet connections (speeds can be as low as 0.1 Mb/s download and 0.05 Mb/s upload).

At the end of November a team from Kenya will be visiting Finland, and the student will be required to have a working prototype of both the normal repository page and the mobile repository page available. Discussions during the visit may result in more practical work.

Background to the thesis work:

Diak, Arcada and Baraton University are cooperating in a project known as HEI ICI. In the project we are setting up a new master's program in global health nursing. The thesis work has to do with increasing the capabilities of the e-learning platforms at Baraton University.

Our trip to Kenya during May 2013 showed that there is a need for a repository server that can handle all the e-learning and other related material both at the university but also in rural areas where internet connections are slower

Thesis goal:

A repository that can handle all e-learning needs for Baraton University. These needs include searching, downloading, and uploading material.

Extra information:

The student will get a great experience interacting with people of many nationalities and cultures. Something definitely worth putting onto a CV