

# **Yhteenvedodokumentti**

Halaan-ryhmä

Helsinki 18.12.2006

Ohjelmistotuotantoprojekti

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

**Kurssi**

581260 Ohjelmistotuotantoprojekti (6 ov)

**Projektiryhmä**

Paula Kemppi  
Liisa Oikarinen  
Ville Palkosaari  
Maria Rinta-Opas  
Jaakko Sorri  
Marjaana Välisalo

**Asiakas**

Terttu Nevalainen

**Johtoryhmä**

Ilari Moilanen  
Kimmo Simola

**Kotisivu**

<http://www.cs.helsinki.fi/group/ohtu/halaan>

**Versiohistoria**

Versio	Päiväys	Tehdyt muutokset
0.1	13.12.2006	Hahmottelua
0.9	15.12.2006	Melkein valmis
1.0	17.12.2006	Projektin päättöanalyysi lisätty

# Sisältö

<b>1 Johdanto</b>	<b>1</b>
<b>2 Sanasto</b>	<b>1</b>
<b>3 Lopputuotteen esittely</b>	<b>2</b>
3.1 Korpustietojen haku . . . . .	2
3.2 Hakutulosten näyttö . . . . .	2
3.3 Korpustietojen tallennus omalle koneelle . . . . .	3
3.4 Korpustietojen tuominen tiedostosta . . . . .	3
3.5 Korpustietojen päivitys web-lomakkeella . . . . .	3
3.6 Käyttäjätietojen hallinta . . . . .	3
<b>4 Dokumenttien tiivistelmät</b>	<b>3</b>
4.1 Projektisuunnitelma . . . . .	3
4.2 Vaatimusdokumentti . . . . .	4
4.3 Suunnitteludokumentti . . . . .	4
4.4 Testaussuunnitelma . . . . .	4
4.5 Käyttöohje . . . . .	4
4.6 Ylläpitodokumentti . . . . .	5
<b>5 Projektin päättöanalyysi</b>	<b>5</b>
5.1 Projektin kulku . . . . .	5
5.2 Työtunnit . . . . .	5

# 1 Johdanto

Halaan-projektin tarkoituksena oli kehittää hakukone laajalle annotoidulle tekstiaineistolle VARIENG-tutkimusyksikön käyttöön. Tutkijat ovat koonneet elektronisen tekstiaineiston eli korpuksen vanhoista englanninkielisistä kirjeistä. Kirjeiden lisäksi tutkimusyksiköllä on kerättyä muutakin kirjeisiin liittyvää dataa, kuten tietoja kirjeiden kirjoittajista ja vastaanottajista. Tutkijat tarvitsevat hakukoneen, joka mahdollistaisi sosiolingvistiset haut, eli kirjeiden ja niihin liittyvien tietojen hakemisen tietokannasta esimerkiksi kirjoittajan sukupuolen, asuinpaikan ja kirjoitusajankohdan perusteella.

Projektin tehtävänä oli koota metadatan sisältävät tiedostot yhdeksi tietokannaksi ja tuottaa järjestelmä, joka toimii hakukoneena edellä mainitulle aineistolle. Järjestelmä tarjoaa myös muun muassa mahdollisuuden muokata hakutuloksen esitystä ja tallentaa haulla rajattu osa-aineisto tai vain siihen liittyvä metadata. Korpukseen voidaan lisätä uutta aineistoa ja olemassa olevaa voidaan päivittää. Koska järjestelmän käyttö edellyttää kirjautumista, käyttäjätietojen hallinnointi on myös osa järjestelmää.

Tämä dokumentti esittelee projektin tuloksena syntyneen ohjelmiston sekä muut projektin tuottamat raportit. Dokumentti sisältää myös projektin päätösanalyysin. Dokumentin tavoitteena on antaa yleiskuva projektin tuotoksista, ennen kuin lukija tutustuu muihin dokumentteihin, joista löytyvät tarkemmat yksityiskohdat.

## 2 Sanasto

**Apache Tomcat** Java Servlet-suoritusympäristö

**Container-olio** Olio, johon voidaan tallettaa tietoa.

**CSS** *Cascading Style Sheets* Tyyllisivut, jotka määrittelevät, miten HTML-elementit esitetään

**Data** Yhteydestä riippuen viittaa joko tekstidataan (kirjeiden sisältö) tai metadataan (data, johon haut kohdistuvat).

**Hakulomake** Hakusivu, josta käyttäjä voi valita hakuun vaikuttavat parametrit. Haku kohdistuu joko kirjeisiin tai kirjeiden kirjoittajiin tai vastaanottajiin.

**Hakutulos** Hakuehtoihin täsmävä *lista* joko kirjeistä tai niiden kirjoittajista tai vastaanottajista. Viitattaessa hakutuloksen tallennettavaan versioon pyritään puhumaan *taulukosta* tai CSV-tiedostosta. Hakutuloksesta voidaan tehdä myös *graafinen esitys*.

**Java** Oliopohjainen ohjelmointikieli

**JavaScript** Skriptikieli, jolla voidaan lisätä web-sivuille dynaamista toiminnallisuutta.

**JSP** *JavaServer Pages* Javapohjainen palvelintekniikka dynaamisten web-sivujen tuottamiseen.

**JUnit** Javapohjainen testikehys, jota voidaan käyttää testien ajamiseen ja joka mahdollistaa testien automatisoinnin.

**Korpus** Elektroninen tekstiaineisto.

**Luokitteleva parametri** Graafisen esityksen X-akselille sijoittuva parametri. Tämä parametri on oletuksena Aika, mutta se voi olla jokin muukin parametri, kunhan valitun parametrin arvot ovat koodi- tai boolean-tyyppisiä.

**Metadata** Kirjeteksteihin liittyvä tieto, joka kuvaa itse kirjettä ja sen kirjoittajaa tai vastaanottajaa. Data, jota parametrit kuvaavat ja johon haut kohdistuvat.

**Korpustiedot** Korpukseen kuuluvaa tekstiaineistoa tai siihen liittyvää metadataa.

**Web-sovellus** Selaimen välityksellä käytettävä palvelinsovellus.

**Projektiryhmä** Ohjelmistotuotantoprojektin ryhmä, joka toteutti tämän dokumentin kuvaaman ohjelmiston ja tuotti siihen liittyvän dokumentaation.

**Käyttäjätunnus** Käyttäjän nimimerkki, jonka avulla käyttäjä tunnistetaan.

**Käyttäjän oikeudet** Eri käyttäjätunnuksilla on erilaiset oikeudet ohjelmiston tarjoamien palvelujen käyttöön.

**Pääkäyttäjä** Käyttäjä, jolla on järjestelmän kaikki oikeudet. Pääkäyttäjää voi olla useita.

## 3 Lopputuotteen esittely

### 3.1 Korpustietojen haku

Ohjelmiston pääasiallinen tarkoitus on toimia hakukoneena asiakkaan keräämälle datalle. Data koostuu korpuksesta ja sen sisältöä luokittelevasta metadatasta. Järjestelmän pääsivu tarjoaa lomakkeen, jolla käyttäjä voi määrittellä hakuparametrit ja haun kohteen (henkilö, kokoelma, kirje) sekä suorittaa haun. Parametrien merkitykset on kuvattu käyttöohjeessa.

### 3.2 Hakutulosten näyttö

Haun tulokset näytetään listana haettuja objekteja. Listan järjestyserustetta voi vaihtaa ja siinä näkyviä kenttiä piilottaa. Listan jokaisella rivillä on linkki, jonka avulla pääsee tarkastelemaan yksittäisen objektin tarkempia tietoja. Lisäksi käyttäjä voi tallentaa omalle koneelleen hakutuloksen (metadatan) tai siihen liittyvien kirjeiden sisällöt (katso kohta 3.3). Tulosta voi myös havainnollistaa kuvan avulla.

Mikäli käyttäjällä on riittävät oikeudet, voi hän siirtyä tulossivulta tietojen muokkaustilaan, jossa voi tehdä pysyviä muutoksia tietokantaan (katso kohta 3.5). Myös käytettyjen hakuparametrien tallentaminen järjestelmään on mahdollista.

### **3.3 Korpustietojen tallennus omalle koneelle**

Käyttäjä voi halutessaan tallentaa haun tulokset omalle koneelleen. Metadata tallennetaan sellaisessa muodossa, että tietoja voi tarkastella taulukkolaskentaohjelmalla. Mikäli haun kohteena on ollut kirjeet, voi käyttäjä tallentaa myös kirjeiden sisällön. Tarkemmat määritelmät käytetyistä tiedostomuodoista löytyvät käyttöohjeesta.

### **3.4 Korpustietojen tuominen tiedostosta**

Korpustietoja voi tuoda järjestelmään lataamalla tiedostoja omalta koneelta. Sekä uusien tietojen tuominen että vanhojen päivittäminen on mahdollista. Ladatun tiedoston sisältämää dataa voi muokata web-lomakkeen avulla ennen varsinaista tallentamista.

Tiedostojen muotoon liittyvät määritelmät löytyvät käyttöohjeesta.

### **3.5 Korpustietojen päivitys web-lomakkeella**

Tietokantaan tallennettua tietoa voi muokata web-lomakkeen avulla. Muokkaus tapahtuu suorittamalla normaali haku ja siirtymällä sen jälkeen tietojen muokkaustilaan. Muokkaustilassa käyttäjä voi valita listasta haluamansa objektin ja siirtyä muokkaamaan sitä web-lomakkeen avulla. Myös tietojen poistaminen on mahdollista.

### **3.6 Käyttäjätietojen hallinta**

Järjestelmä mahdollistaa käyttäjien hallinnan web-sovelluksessa. Oikeus käyttäjien hallintaan on vain pääkäyttäjillä eli Admin-käyttäjryhmään kuuluvilla käyttäjillä. Käyttäjätietojen muokkaustilassa näytetään listaus käyttäjistä ja näiden tiedoista. Käyttäjän oikeuksien ja muiden tietojen muuttaminen on mahdollista web-lomakkeen avulla. Myös käyttäjien lisääminen ja poistaminen on mahdollista.

## **4 Dokumenttien tiivistelmät**

### **4.1 Projektisuunnitelma**

Projektisuunnitelma on ensimmäinen projektin tuottama dokumentti. Se on tarkoitettu lähinnä projektiryhmän sisäiseen käyttöön projektin aikana. Projektin eteneminen ja työtavat määritellään projektisuunnitelmassa, joten se ohjaa projektin kulkua alusta loppuun.

Projektisuunnitelma sisältää projektille etukäteen laaditun aikataulun ja suunnitelman työn osituksesta. Lisäksi projektisuunnitelmassa on projektin jäsenten ja sidosryhmien esittely, riskianalyysi ja arvioita projektin vaativuudesta.

## 4.2 Vaatimusdokumentti

Vaatimusdokumentti laadittiin yhteistyössä asiakkaan kanssa projektin alkuvaiheessa. Sen tehtävänä on ohjelmistolle asetettujen toiminnallisten ja laadullisten vaatimusten määrittely. Vaatimusdokumentti siis määrittelee toteutetun ohjelmiston ja toimii näin sopimuksena asiakkaan ja projektiryhmän välillä.

Vaatimusdokumentissa esitetään kaikki ohjelmistolta vaaditut toiminnot. Toiminnot on johdettu asiakkaan toivomien käyttötapauksien pohjalta. Toteuttamatta jääneet vaatimukset on dokumentoitu ylläpitodokumenttiin.

## 4.3 Suunnitteludokumentti

Suunnitteludokumenttiin koottiin suunnitteluvaiheessa tehdyt ohjelman toteutusta koskevat päätökset. Dokumentti toimii projektiryhmän ohjeena toteutusvaiheen aikana. Sen tavoitteena oli kuvata järjestelmä sellaisella tarkkuudella, että toteutus oli suoraviivaista.

Suunnittelun tuloksena saatiin kuvaukset toteutettavasta ohjelmistosta, sen arkkitehtuurista ja tietokannasta, järjestelmän tarvitsemista ja tuottamista tiedoista, ohjelman eri osien välisistä rajapinnoista sekä järjestelmän komponenteista. Näin suunnitteludokumentti toimii myös toteutetun järjestelmän dokumentaationa mahdollisesti tulevalle jatkokehitykselle. Toteutuksen aikana tapahtuneet poikkeamat suunnitelmasta on dokumentoitu ylläpitodokumenttiin.

## 4.4 Testaussuunnitelma

Testaussuunnitelma määrittää käytetyt testausstrategiat. Siinä selvitetään jokaisesta testausvaiheesta lähestymistapa eli miten testataan, testattavat kohteet eli mitä testataan ja hyväksymiskriteerit eli milloin on testattu tarpeeksi.

Jos jossain testausvaiheessa järjestelmästä löytyi virheitä, on ne pyritty korjaamaan, jos se on aikataulun puitteissa ollut mahdollista. Jos virheitä ei korjattu, on ne dokumentoitu ylläpitodokumenttiin.

## 4.5 Käyttöohje

Käyttöohjeessa annetaan ohjeet ohjelmiston asennukseen, käyttöönottoon ja käyttämiseen. Siinä kuvataan ohjelmiston kaikki toiminnot niin, että ohjelmistoa tuntematon loppukäyttäjä voi opetalla käyttämään ohjelmistoa täydellisesti sen avulla. Ohjelman toimintojen lisäksi käyttöohjeessa on määritelty hauissa käytettyjen parametrien merkitykset sekä asiakkaan kanssa sovitut vaatimukset ohjelmistoon tuotavien tiedostojen muodosta ja sisällöstä.

## 4.6 Ylläpitodokumentti

Ylläpitodokumentti käsittelee sellaisia asioita, joita asiakas tai mahdolliset järjestelmän jatkokehittäjät tarvitsevat muokataksaan ohjelmaa. Dokumenttia ei ole tarkoitettu luettavaksi yksinään, vaan yhdessä suunnitteludokumentin ja ohjelmakoodin kanssa. Dokumentin tehtävänä on esittää kaikki ylläpitoon ja jatkokehitykseen liittyvät seikat ja raportoida toteutuksen eroavaisuuksista suunnittelu- ja vaatimuskäytäntöihin nähden.

# 5 Projektin päättöanalyysi

## 5.1 Projektin kulku

Suunnittelu- ja toteutusvaiheille oli varattu aikaa suunnilleen yhtä paljon ja vaatimusmäärittelylle hieman niitä vähemmän. Vaatimusmäärittely saatiin hoidettua hyvin aikataulussa, mutta suunnittelussa kesti hieman aiottua pidempään, koska ohjelmisto sisälsi osia, joista kenelläkään projektiryhmäläisellä ei ollut kokemusta. Tällaisia olivat muun muassa tiedostojen käsittelyyn ja graafien luomiseen liittyvät tekniikat. Ryhmäläisillä ei myöskään ollut kokemusta toisesta toteutuskielistä eli jsp:stä.

Toteutus ja testaus jaettiin niin, että neljä ryhmäläistä ohjelmoi ja kaksi vastasi testauksesta. Toteutus ei aivan pysynyt aikataulussaan ja myös yksikkötestien teko jUniteilla kesti huomattavasti suunniteltua kauemmin, joten järjestelmätestejä päästiin tekemään vasta projektin viimeisellä viikolla. Näin ollen järjestelmätestausta ei ehditty tehdä niin kattavasti kuin oli tarkoitus. Toteutusvaiheessa aika meinasi loppua kesken, mutta saimme kuitenkin toteutettua lähes kaikki vaatimuskäytännössä määritellyt vaatimukset. Käyttöliittymän käytettävyyteen ja ulkoasuun ei valitettavasti jäänyt aikaa paneutua kunnolla.

## 5.2 Työtunnit

Koska projektin laajuus on 6ov/9op, vastaa se noin 240 tunnin työmäärää eli noin 20h/viikko. Näistä noin 4h/viikko tulee kokouksista. Kukin projektiryhmäläinen piti kirjaa tehdyistä työtunneistaan ja kirjasi ne ohjelmistotuotantoprojektin tietojärjestelmään.

Nimi	Kirjatut tunnit
Paula Kemppi	306,5
Liisa Oikarinen	152
Ville Palkosaari	241,5
Maria Rinta-Opas	178,5
Jaakko Sorri	253
Marjaana Välisalo	308,5
Yhteensä	1445

Ryhmän yhteenlasketut tunnit vastaavat hyvin arvioitua työmäärää  $240h \cdot 6 = 1440h$ , mutta töiden jakaantumisessa ryhmäläisten kesken oli huomattavasti hajontaa. Kun osa teki selvästi vähemmän kuin osuutensa, joutuivat muut tekemään enemmän.