

Toimittanut Inkeri Salonharju

ELEKTRA – Elektroninen julkaiseminen ja julkaisujen verkkovälitys

Projektin loppuraportti

Helsingin yliopiston kirjasto
Kirjastotoimen verkkopalvelut
Helsinki 1998

Elektra - Elektroninen julkaiseminen
ja julkaisujen verkkovälitys: projektin
loppuraportti / toim. Inkeri Salonharju.
- Helsinki : Helsingin yliopiston kirjasto,
Kirjastotoimen verkkopalvelut, 1998
ISBN 951-45-8572-0 (elektroninen, PDF)

ESIPUHE

ELEKTRA – elektroninen julkaiseminen ja verkkojakelu on opetusministeriön Tietoyhteiskunta –ohjelmaan liittyvä kehittämishanke. Sen päätavoitteena on kotimaisen elektronisen julkaisu-toiminnan teknisten ja sopimuksellisten edellytysten kehittäminen.

ELEKTRA-hanke päätti ensimmäisen vaiheensa asiantuntijaseminaariin, joka järjestettiin 20.3.1998 Tieteiden talolla. ELEKTRAn loppuraportti ilmestyy sekä elektronisena verkkojulkaisuna osoitteessa <http://linna.helsinki.fi/elektra>, että painettuna. Painetun julkaisun on kustantanut Yliopistopaino, joka on näin käynnistänyt ELEKTRA-aineistolle soveltuvan tarvepainatuspalvelun kehittämisen.

ELEKTRA-loppuraportin julkaisutoimikuntana on toiminut ELEKTRAn johtoryhmä. Se on hyväksynyt kaikki ELEKTRAn osaraportit, joista on koostettu loppuraportti. ELEKTRAn johtoryhmään kuuluvat ylikirjastonhoitaja Esko Häkli Helsingin yliopiston kirjastosta, joka on toiminut myös johtoryhmän puheenjohtajana, julkaisupäällikkö Eeva-Liisa Aalto Tieteellisten seurain valtuuskunnasta, ylikirjastonhoitaja Päivi Kytömäki Oulun yliopiston kirjastosta, johtaja Sauli Laitinen VTT Tietopalvelusta, myyntipäällikkö Karri Laurila Yliopistopainosta, johtaja Juhani Lokki Suomen Tietokirjailijat Ry:stä, asiakaspalvelupäällikkö Jukka-Pekka Rissanen Helsingin yliopiston atk-keskuksesta, osastonjohtaja Irja-Leena Suhonen Helsingin yliopiston kirjastosta, osastopäällikkö Jukka-Pekka Timonen Kopiosto ry:stä ja osastonjohtaja Inkeri Salonharju Helsingin yliopiston kirjastosta, joka on toiminut myös ELEKTRAn projektipäällikkönä.

Loppuraportin osaraporttien pohjalta on toimittanut projektipäällikkö Inkeri Salonharju ja hän on laatinut myös luvut 1 ja 2, joissa esitellään ELEKTRA-hankkeen lähtökohtia ja tuloksia sekä jatkokehittämistarpeita.

Luku 3 esittelee kansainvälisiä elektronisen julkaisemisen hankkeita ja sen ovat laatineet Riitta Alkula ja Teemu Rautanen VTT Tietopalvelusta. Elektronisen aineiston tiedostomuotoihin ja ohjelmistoihin liittyvän luvun 4 on laatinut Teemu Rautanen.

Luvun 5, jossa selvitetään kotimaista verkkojulkaisukehitystä ovat kirjoittaneet Inkeri Salonharju ja Heli Salonen Helsingin yliopiston kirjastosta. Lukuun sisältyvän osan Oulun yliopiston väitöskirjakokeilusta on laatinut Ville Varjonen Oulun yliopiston kirjastosta.

Luvut 6 ja 7, jotka kuvaavat ELEKTRAn artikkelipilotin teknistä ympäristöä ja aineiston digitointia on kirjoittanut Petri Heliniemi Helsingin yliopiston kirjastosta.

Elektronisia maksujärjestelmiä luvussa 8 ovat selvittäneet Teemu Rautanen VTT Tietopalvelusta ja Juha Hakala Helsingin yliopiston kirjastosta.

Tekijänoikeusperiaatteisiin ja korvauksiin liittyvät luvut 9 ja 10 on laatinut Jukka-Pekka Timonen Kopiosto ry:stä.

Luku 11 käsittelee Internetin tietoturvaa ja sen on laatinut Petri Heliniemi. Raportin päätteeksi luvussa 12 raportoidaan hankkeeseen liittyneestä käyttökokeilusta ja sen ovat laatineet Inkeri Salonharju ja Heli Salonen.

Raportin muuntamisesta verkkojulkaisuksi ja sen ulkoasusta ovat vastanneet Heli Salonen, Timo Hellgren ja Jani Stenvall Helsingin yliopiston kirjastosta.

Kiitän kaikkia ELEKTRA-projektin osallistujia hyvästä matkaseurasta uuteen tuntemattomaan elektronisen julkaisemisen maailmaan. Matka on ollut laajalle osallistujajoukolle huikea oppimiskokemus, jonka tuloksena syntyy toivottavasti pysyvä elektroninen julkaisu- ja välityspalvelu. Toivon rakentavan yhteistyön jatkuvan laajapohjaisena ELEKTRAn jatkohankkeessa.

Helsingissä 10.9.1998

Inkeri Salonharju

Projektipäällikkö

SISÄLTÖ

TIIVISTELMÄ	1
1. JOHDANTO	7
1.1 ELEKTRA-PROJEKTIN TAVOITTEET JA TEHTÄVÄT.....	8
1.2 OSALLISTUJAT JA OSAPUOLTEN VASTUUT	9
1.3 ELEKTRA-PROJEKTIN TUOTTAMAT HYÖDYT	10
1.4 PROJEKTIN RESURSSIT, ORGANISOINTI JA TOTEUTUS	10
1.5 PROJEKTISTA JA TULOKSISTA TIEDOTTAMINEN.....	11
2. YHTEENVETO PROJEKTIN TULOKSISTA, JOHTOPÄÄTÖKSET JA JATKOKEHITTÄMINEN	12
2.1 YHTEYDET JULKAISIJOIHIN JA AINEISTON HANKINTA	12
2.2 DIGITOINTIYMPÄRISTÖ JA JULKAISUPALVELUT	13
2.3 JULKAISUJEN VERKKOJAKELU JA KÄYTTÖKOKEILU	13
2.4 TEKIJÄNOIKEUDET JA SOPIMUSKÄYTÄNNÖT	16
2.5 MAKSULLISUUS JA KORVAUSTEN HALLINNOINTI.....	17
2.6 ELEKTRAN YHTEYDET MUIHIN HANKKEISIIN	18
2.7 YHTEENVETO TULOKSISTA	18
2.8 SUOSITUKSET ELEKTRAN JATKOHANKETTA VARTEN.....	20
2.9 ELEKTRAN JATKOHANKE	21
2.10 ELEKTRA II:N TUOTTAMAT TULOKSET JA NIISTÄ SAATAVAT HYÖDYT	22
3. KANSAINVÄLISIÄ ELEKTRAN KANNALTA KIINNOSTAVIA HANKKEITA	24
3.1 JOHDANTO	24
3.2 WEBDOC	24
3.3 EU:N RAHOITTAMAT HANKKEET	25
3.4 ENGLANNIN ELIB-HANKKEET	26
3.5 BRITISH LIBRARY DIGITAL LIBRARY PROJECT	29
3.6 NCSTRL - NETWORKED COMPUTER SCIENCE TECHNICAL REPORTS LIBRARY	30
3.7 YHDYSVALTAIN SUURET DIGITAL LIBRARIES -HANKKEET	30
3.8 MUITA KIINNOSTAVIA PROJEKTEJA.....	31
4. ELEKTRONISEN JULKAISEMISEN TIEDOSTOMUODOT JA OHJELMISTOT	33
4.1 JOHDANTO	33
4.2 ELEKTRONISEN JULKAISEMISEN TIEDOSTOMUODOISTA	33
4.3 KUVAFORMAATEISTA	41
4.4 LIIKKUVAN KUVAN JA ÄÄNEN FORMAATEISTA	43
4.5 ADOBEN ACROBAT-OHJELMISTA	45
4.6 TEKSTINTUNNISTUSOHJELMISTA	47
4.7 TEKSTINKÄSITTELY- JA JULKAISUOHJELMISTA.....	48
4.8 KUVANKÄSITTELY- JA PIIRTO-OHJELMISTA	52
5. ELEKTRAN AINEISTOT JA INTERNETIN KAUTTA SAATAVILLA OLEVAT KOTIMAISET VERKKOJULKAISUT	55
5.1 ELEKTRAN ARTIKKELIAINEISTON VALINTAPERUSTEET.....	55
5.2 SELVITYS KOTIMAISISTA VERKKOLEHDISTÄ	55
5.3 KOTIMAISSIA LEHTILISTOJA INTERNETISSÄ	57
5.4 KYSELY TIETEELLISTEN JULKAISIJOIDEN VERKKOKEHITYSTILANTEESTA JA KIINNOSTUKSESTA OSALLISTUA ELEKTRA-KOKEILUUN.....	59
5.5 ELEKTRAN ARTIKKELIPILOTISSA MUKANA OLEVAT JULKAISUT	60
5.6 VÄITÖSKIRJOJEN ELEKTRONINEN JULKAISEMINEN OULUN YLIOPISTOSSA	62
6. ELEKTRAN PILOTTIEN TEKNINEN YMPÄRISTÖ	65

6.1	ELEKTRONINEN JULKAISU- JA DIGITOINTIYMPÄRISTÖ	65
6.2	DOKUMENTTIEN HAKU JA SELAUS	65
6.3	TULOSTUS.....	66
6.4	ARKISTOINTI.....	66
7.	AINEISTON DIGITOINTI.....	68
7.1	PILOTTIAINEISTON TOIMITUS JA KÄYTETYT TIEDOSTOMUODOT.....	68
7.2	PAINETUSSA MUODOSSA TOIMITETTU AINEISTO	69
7.3	ELEKTRONISESSA MUODOSSA TOIMITETTU MATERIAALI	72
7.4	DIGITOINTITYÖN EDELLYTTÄMÄT RESURSSIT	77
7.5	TIEDOSTOJEN KOOT	78
7.6	JOHTOPÄÄTÖKSET	80
8.	ELEKTRONISET MAKSUJÄRJESTELMÄT	82
8.1	JOHDANTO	82
8.2	AVANT-KORTTIRAHA.....	84
8.3	EUNET ECASH.....	87
8.4	YHTENVELO MAKSUJÄRJESTELMIEN KUSTANNUKSISTA KAUPPIAALLE.....	89
8.5	TULOSTEN VERTAILU	94
9.	TEKIJÄNOIKEUSPERIAATTEET.....	95
9.1	TEOSTEN UDELLENJULKAISU DIGITAALISESSA MUODOSSA.....	95
9.2	DIGITAALISTEN KÄYTTÖOIKEUKSIEN LUOVUTUKSISTA	95
9.3	MITÄ OIKEUKSIA ELEKTRA TARVIKSEE ?.....	96
9.4	KÄYTÖN EHDOT JA KORVAUKSET	96
9.5	OIKEUKSIEN HANKINTA JA HALLINNOINTI	96
9.6	OIKEUDENHALTIJOIDEN KANSSA SOLMITTAVAT SOPIMUKSET	96
9.7	KORVAUKSEN MAKSAMINEN OIKEUDENHALTIJOILLE	97
9.8	ELEKTRA -PROJEKTIN PERIMÄT KORVAUKSET	97
9.9	LOPPUKÄYTTÄJÄT JA TEOSTEN KÄYTTÖKORVAUKSET.....	97
9.10	KÄYTTÖTAPAHTUMIEN REKISTERÖINTI	98
9.11	KÄYTTÖTIETOJEN JA KÄYTTÖKORVAUSTEN VÄLITTÄMINEN OIKEUDENHALTIJOILLE	98
9.12	KOPIOSTON ROOLI ELEKTRA -PROJEKTISSA.....	98
10.	TEKIJÄNOIKEUSKORVAUKSET JA NIIDEN HALLINNOINTI.....	99
10.1	PALVELUJEN HINNOITTELU.....	99
10.2	HINTASUOSITUKSET TALLENNUSKORVAUKSEKSI JA TULOSTEKOHTAISEKSI KORVAUKSEKSI.....	99
10.3	HINTASUOSITUS TEKNISEKSI TULOSTUSKORVAUKSEKSI.....	99
10.4	EHDOTUKSET KORVAUKSEN JAOSTA.....	100
10.5	OIKEUDENHALTIJAN VALTUUTUS KOPIOSTOLLE VÄLITTÄÄ KORVAUSTEN TILITYS	102
10.6	SOPIMUSTEN MUKAINEN RAPORTOINTI JA TILITYKSET.....	102
10.7	LASKUTTAMINEN JA SUMMIEN TARKASTAMINEN	103
10.8	KORVAUSTEN JAKO JA TILITYSRAPORTIT.....	104
10.9	ALUSTAVA TIETOKANTAKUVAUS	105
11.	TIETOTURVA.....	108
11.1	TIETOTURVAN OSA-ALUEET JA TIETOTURVAPOLITIikka	108
11.2	SUOJAUTUMINEN	109
11.3	IP- JA TCP-PROTOKOLLATASOJEN TURVAMEKANISMIT.....	110
11.4	TUNNETTUJA TURVALLISUUSAUUKKOJA PROTOKOLLISSA	111
11.5	UNIX/LINUX KÄYTTÖJÄRJESTELMIEN RAKENNE JA SUOJAUKSET.....	112
11.6	UNIXIN TUNNETTUJA TURVALLISUUSAUUKKOJA	113
11.7	INTERNET JA WWW	114
11.8	INTERNETIN TURVARATKAISUJA	115
11.9	JOITAKIN WWW:N TURVA-AUKKOJA	119
11.10	TIETOTURVA ELEKTRA-PROJEKTISSA	120

12. KÄYTTÖKOKEMUKSET	124
12.1 KÄYTTÖYMPÄRISTÖ	124
12.2 KOEKÄYTTÖ	124
12.3 KÄYTTÄJÄPALAUTE JA PALVELUN KEHITTÄMINEN	125
LIITE 1: PDF-DOKUMENTIN PAKKAUSMENETELMÄT	127
LIITE 2: MINIATYYRIKUVAT JA KIRJANMERKIT ACROBAT READERISSA	128
LIITE 3: KYSELY TIETEELLISTEN SEURAIN JÄSENORGANISAATIOILLE/JULKAISUJEN TOIMITUSKUNNILLE	129
LIITE 4: JULKAISIJAKYSELYN TULOKSET	134
LIITE 5: ELEKTRAN KÄYTTÄJÄPALAUTTEEN VASTAUKSET	141

TIIVISTELMÄ

Lähtökohdat ja tavoitteet

ELEKTRA-projekti on opetusministeriön Suomi tietoyhteiskunnaksi –ohjelmaan kuuluva yhteistyöhanke, jonka osapuolina on yliopistokirjastoja, kustantajia, tieteellisen julkaisutoiminnan ja elektronisen julkaisemisen tekniikan asiantuntijoita sekä tekijänoikeusjärjestöjen edustaja. Laajan konsortion avulla koko kotimainen tieteellinen julkaisukenttä on saatu mukaan hankkeeseen ja näin voidaan selvittää rinnan elektronisen julkaisemisen teknisiä, taloudellisia ja sopimuksellisia seikkoja, solmia tarvittavia julkaisijakontakteja, kehittää sopimusmalleja ja saada kokemusta elektronisen julkaisemisen kokonaisprosessista. Kattava osallistujapohja antaa ainutlaatuisen mahdollisuuden kehittää elektronista julkaisu- ja jakelupalvelua, joka vastaa niin tekijöiden, julkaisijoiden kuin tiedon käyttäjienkin tarpeita.

Projektia koordinoi Helsingin yliopiston kirjasto ja osallistajat ovat Oulun yliopiston kirjasto, Tieteellisten seurain valtuuskunta, Kopioisto ry, VTT Tietopalvelu, Gaudeamus Kirja, Yliopistopaino ja Helsingin yliopiston atk-keskus. ELEKTRA on osallistujien yhteisrahoitteinen hanke ja se on saanut avustusta opetusministeriöltä. Hankkeen kokonaisbudjetti vuosille 1996-97 on ollut noin 1,7 miljoonaa markkaa, josta osallistujien omarahoitus on kattanut lähes puolet. Projekti on käynnistynyt 1.6.1997 ja ensimmäisen vaiheen kesto on ollut 18 kuukautta. ELEKTRA-projektin toista vaihetta varten on tehty uusi kahden vuoden mittainen hanke-ehdotus. Tässä raportissa esitellään projektin ensimmäisen vaiheen tulokset sekä lyhyesti kuvataan toisen vaiheen hanke-ehdotusta.

ELEKTRA-projektin tavoitteina on

- toteuttaa julkaisujen haku- ja välityspalvelu verkkoympäristössä ja evaluoida tuloksia todellisilla käyttäjillä
- luoda malli korkeakoulujen elektronista kurssimateriaali- ja väitöskirjapalvelua varten
- selvittää tekijänoikeuskysymyksiä ja kehittää sopimusmallit erityisesti tieteellisten julkaisujen välittämiseen
- selvittää maksullisuusperiaatteita sekä laskutus- ja suojauskäytäntöjä
- luoda yhteyksiä kansainvälisiin elektronisen julkaisemisen hankkeisiin.

Kotimainen tieteellinen julkaisutoiminta ja ELEKTRA

Elektroninen julkaiseminen on prosessi, joka kattaa dokumentin luomisen, tallentamisen, esittämisen ja jakelun tietokoneen avulla. Vaikka elektroninen julkaiseminen lisääntyy nopeasti, on vasta murto-osa julkaisutuotannosta käytettävissä joko verkkojulkaisuina, cd-rom- tai muilla fyysisillä välineillä levitettävänä elektronisina julkaisuina. Verkkojulkaisujen määrä kasvaa kuitenkin nopeasti ja valtaa alaa perinteiseltä tieteelliseltä julkaisutoiminnalta potentiaalisen suuren levikkinsä ja kehittyneiden teknisten ominaisuuksiensa, kuten helpon päivitettävyyden ja hypertekstiominaisuuksien ansiosta.

Kotimainen tieteellinen julkaisijakenttä on varsin hajanainen ja kaupallisten kustantajien lisäksi suuresta osasta julkaisutoimintaa vastaavat yliopistot ja tieteelliset seurat. Suomessa julkaistaan vuosittain noin 800 väitöskirjaa. Väitöskirjojen lisäksi yliopistot julkaisevat erilaisia tieteellisiä

raporttisarjoja. Tieteellisiä seuroja Suomessa on noin 200, ja ne julkaisevat yhteensä lähes sataa tieteellistä aikakauslehteä, vuosikirjaa tai muuta sarjaa. Merkittävä osa väitöskirjoista julkaistaan tieteellisten seurojen sarjoissa.

Mikään kotimainen yliopisto ei ole vielä merkittävässä määrin siirtänyt julkaisuautoimintaansa elektroniseen muotoon. Pienimuotoisia kokeiluja on ollut ELEKTRA-projektissa mukana olevien Helsingin ja Oulun yliopistojen lisäksi Teknillisen korkeakoulun sekä Vaasan ja Jyväskylän yliopistojen piirissä. Kaikissa näissä kokeiluhankkeissa yliopistokirjastoilla on ollut tärkeä rooli.

ELEKTRA -projektin tekemän selvityksen mukaan vain harvoilla tieteellisillä seuroilla on kokemusta elektronisesta julkaisemisesta tai omia konkreettisia suunnitelmia käynnistää elektroninen julkaisuautoiminta lähivuosina. Valtaosa julkaisijoista oli kuitenkin kiinnostuneita oman julkaisunsa muuntamisesta verkossa jaeltavaan muotoon ja julkaisukokeilusta. ELEKTRA-projektin artikkelikokeiluun saatiin aineistoa 36 lehdestä. Projektin ansiosta jo lähes puolet julkaisevista tieteellisistä seuroista on saanut kokemusta elektronisesta julkaisuautoiminnasta. Tällä voi olla pysyvää merkitystä kotimaisen elektronisen sisältötuotannon käynnistämisessä ja kokeilu voi jouduttaa innovaatiota vuosilla yksittäisiä tieteellisiä julkaisijoita ajateltaessa.

ELEKTRAn ensimmäinen vaihe on keskittynyt varsin voimakkaasti elektronisten julkaisijapalvelujen keittämiseen ja julkaisijakontaktien luomiseen. ELEKTRAn toisen vaiheen pääpaino on taloudellisten mallien kehittämisessä ja jakelupalvelujen laajentamisessa ja parantamisessa. Nykyisestä rinnakkaisesta julkaisemisesta painetun version kanssa pyritään myös ensisijaisesti elektroniseen julkaisemiseen. Tämä edellyttää kuitenkin sopimusten edelleenkehittelyä ja aineiston arkistointikriteereiden tarkistamista. Elektronisen aineiston pitkäaikainen säilytys ja siihen liittyvät ongelmat ratkaistaan uudistettavana olevan vapaakappalelain tultua voimaan Kansalliskirjaston tarjoamana vapaakappalepalveluna.

Julkaisutekniikka

Elektronisen julkaisemisen tekninen infrastruktuuri on jo varsin hyvin kehittynyt. Verkojulkaisemiseen ja jakeluun on mahdollisuus lähes kaikilla hyvin varustetuilla Internetin käyttäjillä. Koska vapaa verkko tekee kaikista tiedeyhteisön jäsenistä ei pelkästään potentiaalisia julkaisijoita vaan myös kustantajia, on vaara että tieto pirstoutuu verkkoihin vaikeasta hallitavaksi ja laadultaan kirjavaksi kaaokseksi. Teknisten standardien suuri määrä ja epäyhtenäinen soveltaminen heikentävät tiedon käytettävyyttä. Tieteellisten julkaisujen laadunvalvontaan kehitetty referee-järjestelmä ei myöskään ulotu vapaaseen verkkoaineistoon. Samoin verkkodokumenttien autentisoinnin ja käyttäjien elektronisen identiteetin sekä automaattisten maksujärjestelmien kehittäminen on kesken.

ELEKTRAn julkaisupalvelu on kehitetty julkaisijoiden tarpeita vastaavaksi. Artikkelipilottia varten aineistoa on saatu painettuna sekä erilaisilla tekstinkäsittely- ja julkaisuohjelmistoilla tuotettuina tiedostoina. Pitkäaikaista säilytystä varten aineisto on tallennettu alkuperäisessä muodossa, ja jakelua varten siitä on tuotettu html- tai pdf-tiedostot, jotka on asetettu verkkokäyttöön Kansalliskirjaston arkistopalvelimelle. Kokeiluyistä osa aineistoa on tallennettu julkaisijan omalle palvelimelle. Oulun väitöskirjapilotissa on kokeiltu html- ja pdf-tiedostomuotojen lisäksi dokumentin muuntaamista sgml-standardin mukaiseen muotoon.

ELEKTRAn jatkohankkeessa aikaisempaa suurempi osuus aineistosta pyritään saamaan valmiiksi jossain projektin suosittelemassa tiedostomuodossa. Julkaisijoita varten laaditaan ohjeet, jotka sisältävät myös tyylimääritykset ja tallennusohjeet dokumenttiin itseensä sisällytettävissä metatdatakuvailuja varten. Yhteistyötä julkaisijoiden kanssa on tarkoitus kehittää niin, että do-

kumentit edellyttävät vain teknisiä konversioita. Tällöin julkaisukustannukset alentuvat tuntuvasti ja samalla dokumenttien laatu kohoaa.

ELEKTRAN tekijänoikeusperiaatteet

ELEKTRA-projektissa uudelleenjulkaistaan kirjallisia teoksia ja niihin liittyvää kuvitusta, valokuvia ja mahdollisesti luetteloina suojattuja teoksia. Tämä aineisto on tekijänoikeuden alaista ja edellyttää sopimuksia oikeudenhaltijoiden ja ELEKTRA-projektin välillä.

Tekijänoikeuksien hankintaa varten on laadittu kirjallinen mallisopimus sekä tekijöitä että kustantajia varten. Sopimuksen mukaan Helsingin yliopiston kirjasto saa oikeuden dokumenttien saattamiseen digitaaliseen muotoon ja rajoitettuun verkkolevityskokeiluun. Oikeus on rajattu kestämään ELEKTRA-projektin ajan. Lähtökohtana on, että teosten oikeudenhaltijat voivat kukin päättää sekä teoksen käytöstä perittävän korvauksen että aineiston käyttöehdot. Vaikka elektroniset maksujärjestelmät vielä puuttuvat, on oikeudenhaltijalla mahdollisuus antaa oikeus dokumentin lukemiseen ja tulostuksen tai rajata oikeus pelkkään lukemiseen. Dokumentin tulostus tiedostoksi ei ole sallittua ja se on pyritty teknisesti estämään.

Tekijänoikeussopimusten laadinnasta ja oikeuksien hankinnasta ELEKTRA-projektissa on vastannut Kopiosto. Oikeuksien hankintaan on osallistunut oman jäsenkuntansa osalta aktiivisesti myös Tieteellisten seurain valtuuskunta.

Kustannusrakenne ja maksullisuus

ELEKTRAN julkaisu- ja jakelupalvelusta aiheutuvat kustannukset on katettu osin OPM:n avustuksella ja osittain osallistujien omarahoituksella, joka on tavallisimmin sisältänyt henkilötyötä. Julkaisijoille ja tekijöille kokeiluun osallistuminen on ollut maksutonta. Suurimman menoeran projektissa ovat aiheuttaneet julkaisijoille tarjottu digitointipalvelu, tekijänoikeussopimusten laadinta ja oikeuksien hankinta sekä laitteisto- ja ohjelmistohankinnat.

ELEKTRAN haku- ja dokumentinvälityspalvelu on ollut käyttäjille osin maksutonta ja osin maksullista. Dokumenttien hakeminen ja lukeminen näytöltä on ollut asiakkaille maksutonta, mutta tulostuksesta on veloitettu sivujen määrään perustuva hinta. Helsingin yliopiston kirjasto on maksanut oikeudenhaltijoille korvauksen teosten tallentamisesta ELEKTRAN dokumenttipalvelimelle. Tallennuskorvauksen suuruus on ollut lehden niiden vuosikertojen hinta, joita on sisällynyt pilottiaineostoihin.

Perityt korvaukset on suositeltu jaettavaksi tasan kustantajan ja tekijän kesken. Tämä suositus on hyväksytty laajalla rintamalla, ja sen taakse on saatu Suomen tiedekustantajien liitto ry, Suomen tietokirjailijat ry ja Suomen tiedetoimittajain liitto ry. Oikeudenhaltijoilla on ollut myös mahdollisuus itse hinnoitella oma tuotteensa. Dokumenttien käytön raportointia ja korvausten laskutusta ja tilitystä varten on kehitetty erillinen tietojärjestelmä. Sen avulla voidaan tulostaa sopimusten edellyttämät käytön seuranta- ja tilitysraportit oikeudenhaltijoille. Raportointi tehdään teoskohtaisesti ja se sisältää tiedot julkaisujen katselukerroista ja tulosteista.

ELEKTRAN julkaisupalvelun kustannukset samoin kuin arkistoinnista aiheutuneet kulut ovat hankevaiheessa Helsingin yliopiston kirjaston vastuulla ja ne on pääosa voitu kustantaa OPM:n avustuksella. Jatkuvaa julkaisupalvelua suunniteltaessa nämä kustannukset on kuitenkin voitava siirtää palvelun käyttäjille, eli kustantajille ja loppukäyttäjille. Aineiston digitointia voidaan tarjota maksullisena palveluna ja pääosa aineiston arkistoinnista voitaneen tehdä elektronisten

vapaakappalepalvelujen yhteydessä. Digitointikustannuksia voidaan myös tuntuvasti alentaa rationalisoimalla työtä yhteistyössä julkaisijoiden kanssa.

Merkittävän kustannustekijän aiheuttaa myös tekijänoikeuksien hallinnointi, mikäli oikeudet joudutaan hankkimaan jälkikäteen, kuten projektin ensimmäisen vaiheen aikana on tehty. Mikäli oikeudet voitaisiin hankkia keskitetysti julkaisuvaiheessa yhteistyössä kustantajien kanssa, voitaisiin kustannuksia alentaa tuntuvasti. Tämä edellyttää sopimusmekanismien edelleenkehittämistä.

Julkaisujen jakelukanavat ja ELEKTRAn käyttökokeilu

ELEKTRAn haku- ja dokumentin välityspalvelu perustuu nykyisen artikkelitietokannan, ARTOn näyttöluettelotoiminnon laajennukseen. Tarkoituksena on, että asiakkaat voivat hyödyntää näyttöluettelon hyviä hakuominaisuuksia ja korkeatasoisia bibliografisia- ja sisällönkuvailevia, jotka jo nyt tuotetaan tieteellisten kirjastojen ja muiden asiantuntijaorganisaatioiden laajana yhteistyönä. Palvelun taso parantuu oleellisesti, kun asiakkaat saavat tiedonhaun yhteydessä käyttöönsä itse varsinaiset dokumentit.

Hakupalvelu on teknisesti toteutettu VTLS-hakutietokannan ja www-hakupalvelimen avulla. Testivaihetta varten on perustettu erillinen ELEKTRA-tietokanta. Www-käyttöliittymänä on myös Linnean yhteisluettelotietokannoissa käytössä oleva Hyperlib-ohjelmisto, johon on kehitetty ELEKTRA-palvelua varten uusia ominaisuuksia, kuten lokiseuranta. Viitetiedoista päästään itse dokumenttiin URN-osoitteen avulla. ELEKTRA-projekti on luonut myös elektronisen aineiston Finmarc-kuvailun edellyttämät luetteloinnin sovellusohjeet ja luettelointikäytännöt, joita on ELEKTRAn lisäksi hyödynnetty jo nyt laajasti Linnea-yhteisluetteloinnissa. ELEKTRA-aineistoon on osittain tallennettu myös Dublin Core formaatin mukaisia metadata-tietoja, jotka sisällytetään itse dokumenttiin.

ELEKTRA-tietokanta koekäytössä seitsemässä tieteellisessä kirjastossa, Helsingin-, Jyväskylän- ja Oulun yliopistokirjastoissa, Teknillisen korkeakoulun kirjastossa, VTT Tietopalvelussa ja Helsingin yliopiston Maatalous- ja Luonnontieteiden kirjastossa. ELEKTRA-tietokannan lukukäytöstä on kerätty seurantatietoja käyttöliittymään yhdistetyn lokiseurantaohjelman avulla, joka raportoi teoskohtaisesti artikkeleiden ja muiden dokumenttien lukemiskerrat. Lisäksi kokeilukirjastot raportoivat manuaalisesti kerätyt tiedot teoskohtaisista sivutulostemääristä. ELEKTRAn kokeilukäyttö aloitettiin kirjastoissa 15.6.1997 ja varsinaisesti se pääsi käyntiin vasta syyslukukauden alussa 1.9.1997. Tietojen keruu päättyi vuoden 1997 lopussa. Kokeilukausi oli varsin lyhyt ja sen antamat tiedot viitteellisiä. ELEKTRA-dokumentteja on luettu kokeilukirjastoissa yhteensä 200 kertaa. Lehdittäin tarkasteltuna eniten käyttökertoja kohdistui Yliopistolehteen, mutta muuten käyttö hajaantuu eri lehtien kesken.

Palautetietoja kerättiin käyttöliittymän yhteydessä olevalla suoraikäyttölomakkeella ja kirjastoissa jaettavalla paperilomakkeella. Kyselyn tavoitteena oli hankkia asiakkaiden arvioita käyttöliittymän toimivuudesta, palvelun sisältöön liittyvistä toiveista ja reaktioista maksullisuuteen. Vastajat pitivät ELEKTRA-palvelua tervetulleena, mutta käyttökokeilun liittyvää tiedotusta oli lähes kaikkien vastaajien mielestä ollut liian vähän. Kokeilukirjastoissa vain puolet vastaajista oli löytänyt ELEKTRA-työaseman helposti. Palvelun käyttöliittymää ja opasteita sensijaan pidettiin helppoina ja selkeinä. Käyttäjistä puolet ilmoitti, että pelkkä lukeminen ilman tulostusta on riittävä vaihtoehto. Hintoja vajaa puolet vastaajista piti kohtuullisina, neljännes kalliina ja kolmannes ei osannut muodostaa mielipidettään. Jotkut vastaajista toivoivat myös palvelun laajentamista yliopiston paikallisverkkoon.

Palvelun liian suppea testaus ja riittämätön markkinointi ovat epäilemättä olleet ELEKTRA-projektin ensimmäisen vaiheen suurimmat heikkoudet. Resurssit keskitettiin kuitenkin tietoisesti julkaisupalvelun kehittämiseen ja aineiston digitointiin. ELEKTRAN rajattu koekäyttömahdollisuus samoin kuin käyttökertoihin perustuva hinnoittelu ovat vastanneet kustantajien ja tekijöiden tämänhetkisiä tarpeita. Elektronisen aineiston käyttöä voidaan kuitenkin edistää ja palveluja kehittää käyttäjien tarpeita paremmin vastaaviksi monin tavoin. Palvelusta tiedottamalla ja käyttäjiä kouluttamalla saataneen nopeimmin toivottuja tuloksia. ELEKTRAN toisessa vaiheessa on jo suunniteltu keskittymistä käyttäjäpalvelujen kehittämiseen. Sopimusmekanismeja on tarkoitus kehittää niin, että käyttöä voidaan kokeilla organisaatiolisenssein yliopistojen ja muiden oppilaitosten paikallisverkoissa.

Yhteydet muihin hankkeisiin

Yhtenä ELEKTRAN tavoitteena oli luoda yhteyksiä vastaaviin koti- ja ulkomaisiin hankkeisiin. Verkostoituminen on onnistunut yli odotusten, ja ELEKTRA on avannut partnereilleen mahdollisuuksia osallistua useisiin muihinkin hankkeisiin. ELEKTRA on päässyt testihankkeeksi EU:n kirjasto-ohjelman tekijänoikeuksia selvittävään Tecup-hankkeeseen, jossa ELEKTRAN partnereista ovat mukana HYK ja Kopiosto. ELEKTRAssa luotu osaaminen on avannut HYK:ille ovet myös saman ohjelman Dieper-hankkeeseen. Sen tavoitteina on digitoida eurooppalaisia tieteellisiä aikakauslehtiä ja tarjota yhteinen pääsy tähän tietovarantoon. Euroopan neuvoston Adapt-ohjelmaan liittyvässä Uusi kirjatalous -hankkeessa on mukana useita ELEKTRAN partnereita asiantuntijatahoina.

ELEKTRAssa on kehittynyt avainhanke, joka on synnyttänyt ympärilleen myös uusia hankkeita, joiden kanssa on luotu luonteva työnjako ja hankkeet tukevat toisiaan. Elektronisen aineiston arkistointiproblematiikka edellytti erillistä hanketta, jota varten perustettiin EVA, elektronisen verkkoaineiston hankinta ja arkistointiprojekti. Siinä ratkaistaan myös ELEKTRAN osalta arkistointikysymykset samalla kun ELEKTRA tuottaa EVAa varten tarvittavaa testiaineistoa. Helsingin yliopisto on käynnistänyt vuoden 1998 alussa oman Elektroninen julkaisupankki-hankeensa, jonka tavoitteena on siirtää merkittävä osa Helsingin yliopiston julkaisutuotannosta elektroniseen muotoon aina laitossarjoista tutkielmiin ja väitöskirjoihin. Hankkeen koordinointi on HYK:in tehtävänä, ja ELEKTRAN kokemuksia hyödynnetään tässä vaativassa hankkeessa. Julkaisupankin aineistosta se osa, mikä edellyttää kontrolloitua jakelua, ohjataan julkaistavaksi ELEKTRAssa. Opetusministeriön käynnistämä Kansallisen elektronisen kirjaston ohjelma edistää elektronisten verkkoaineistojen käyttöä yliopistoissa ja tutkimuslaitoksissa. ELEKTRA on ainoa kotimainen hanke, joka kehittää jakelukanavaa tekijänoikeudellista tieteellistä kotimaista verkkoaineistoa varten.

Suosituksat ELEKTRAN jatkohanketta varten

1. ELEKTRA on julkaisukanavana kiinnostava tekijöiden ja julkaisijoiden kannalta ja merkittävä osa tiedekustantajista on saatu mukaan kokeiluun. Julkaisijakontakteja tulee vahvistaa sopimuksin ja yhteistyötä kehittämällä.
2. Mahdollisimman suuri osa aineistosta tulee pyrkiä saamaan palveluun valmiiksi elektronisessa muodossa, jotta tuotantokustannukset saadaan hallittua.
3. Julkaisijoita ja tekijöitä varten tulee laatia julkaisuohjeet ja työkalut dokumenttiin itseensä sisällytettäviä kuvailuja ja tarvittavia tyyliäärityksiä varten. Näin parannetaan dokumenttien laatua.

4. Palveluun tulee integroida tarvepainatus.
5. ELEKTRAN sopimusmekanismi on ollut vakuuttava kustantajien ja tekijöiden kannalta, mutta sitä tulee kehittää edelleen julkaisujen avoimempaa välitystä varten.
6. ELEKTRAN jakelupalvelu nykyisessä rajoitetussa muodossaan ei ole riittävästi kiinnostanut käyttäjiä, ja jakelua tulee laajentaa sopimuskirjastoista rajoitettuun verkkokäyttöön organisaatiolisenssein. Myös muita jakeluvaihtoehtoja tulee selvittää.
7. ELEKTRA on profiloitunut tekijänoikeudellisen, etupäässä maksullisen aineiston dokumentinvälityspalveluksi, ja sen suhteet vapaan verkkoaineiston välitykseen väitöskirjapalvelua varten tulee tarkentaa.
8. Aineiston maksullisuus ei käyttäjäpalautteen mukaan ole ongelma, mutta tulostuksen vähäisyys osoittanee sen, että hintatasoa täytyy edelleen harkita.
9. Automaattista käytön laskutus-, seuranta- ja raportointiohjelmistoa tulee kehittää edelleen jo rakennetun ohjelmiston pohjalta.
10. Käyttäjäpalvelusta tulee tiedottaa enemmän ja kouluttaa tarpeen mukaan käyttäjiä.
11. Vastaavaa kansainvälistä kehitystä tulee edelleen seurata sekä osallistumalla muihin yhteistyöhankkeisiin että solmimalla aktiivisesti kontakteja uusien hankkeiden kanssa.
12. Elektronisen identiteetin ja automaattisten maksujärjestelmien kehitystä tulee seurata ja osallistua kotimaisiin kokeiluhankkeisiin heti kun se on mahdollista.
13. Elektronisen aineiston pitkäaikainen tallennus ja siihen liittyvät tekniset ratkaisut tulee selvittää myös väitöskirjojen osalta elektronisen aineiston hankinta- ja arkistointihankkeessa Evassa ja aikanaan osana uudistuvan vapaakappalelain määrittämää peruspalvelua.

1. JOHDANTO

Elektroninen julkaiseminen ja julkaisujen jakelu verkkoyhteyksin on yleistynyt nopeasti. Kotimainen tieteellisen aineiston lukijakunta on suppea ja levikki usein vaatimaton, vaikka painokustannukset ovat korkeat. Tieteellisestä julkaisutoiminnasta suuri osa on yliopistojen ja tieteellisten seurojen vastuulla. Vaikka elektronisen julkaisemisen kehittelykustannukset ovat suhteellisen korkeat, uskotaan sen kuitenkin pitkällä aikavälillä muodostuvan painettua julkaisemista taloudellisemmaksi. Vielä merkittävämpiä etuja ovat kuitenkin verkkojakelun mahdollistama nopea tiedonvälitys ja potentiaalisen lukijakunnan moninkertaistuminen.

Internet-verkosta löytyy jo suuri määrä vapaassa jakelussa olevaa tieteellistä ja muuta aineistoa, jonka laatutaso on kuitenkin kirjavaa. Vapaan verkkoaineiston löytyvyyttä edistävät verkkoindeksit, mutta suurten aineistomäärien vuoksi tiedonhaku niiden kautta on usein ongelmallista. Suuret tiedekustantajat ovat myös alkaneet siirtää julkaisutoimintaansa, erityisesti aikakauslehtiä laajassa mitassa elektroniseen muotoon. Kaupallisen, tekijänoikeuksin suojatun aineiston jakelu tapahtuu erillisten dokumenttipalvelimien kautta, joiden yhteyteen on liitetty tehokkaat hakumahdollisuudet. Palvelu on tavallisimmin maksullista ja tarjolla ovat edelleen myös julkaisujen painetut rinnakkaisversiot, jolloin elektronisten palvelujen käyttösovimukset perustuvat organisaatioiden sisäisiin lisensseihin.

Kotimaiset väitöskirjat julkaistaan edelleen painettuina, mutta tavoitteeksi on tullut väitösopinäyte, jota voidaan säilyttää ja tarjota käyttöön elektronisessa muodossa. Myös nopeasti uusiutuvan oppimateriaalin jakeluun elektroninen välittäminen todennäköisesti sopii nykyisiä kanavia paremmin. Aikakauslehtiartikkelien saatavuutta voidaan oleellisesti parantaa verkkojakelulla. Kirjastot ovat tuottaneet laajoja viitetietokantoja, joiden hakupalveluihin voidaan helposti liittää myös varsinaiset elektroniset dokumentit. Näin tiedonhakija saa käyttöönsä myös itse dokumentin tiedonhaun yhteydessä.

Tekniikan luomien mahdollisuuksien hyväksikäyttöä elektronisten julkaisupalvelujen yleistymisessä ovat kuitenkin jarruttaneet standardien moninaisuus ja tekijänoikeudelliset kysymykset. Digitaalitekniikka mahdollistaa periaatteessa dokumentin rajattoman kopioinnin, mikä on asettanut elektronisen aineiston tekijänoikeuksille tiukemmat vaatimukset kun painettuun aineistoon sovelletut käytännöt. Myös elektronisen identiteetin ja verkkorahan kehitys ovat olleet odotettua hitaampia ja lukuisat kilpailevat rinnakkaiset maksujärjestelmät vaikeuttavat kaupallisten palvelujen kehittämistä. Myös epätietoisuus elektronisten dokumenttien säilyvyydestä ja arkistoinnin vastuutahoista tulevia sukupolvia varten on jarruttanut perinteisestä julkaisemisesta luopumista.

ELEKTRA on opetusministeriön Suomi tietoyhteiskunnaksi -ohjelmaan liittyvä yhteistyöhanke, joka pyrkii edistämään kotimaisen tieteellisen aineiston elektronista julkaisutoimintaa. Tätä varten selvitetään koko elektronisen julkaisemisen, julkaisujen jakelun ja arkistoinnin edellyttämät teknisiä, sopimuksellisia ja taloudellisia prosesseja.

ELEKTRA-projektiin on koottu osallistujatahot, jotka edustavat monipuolisesti alan asiantuntemusta. Mukana on kirjastojen lisäksi kustantajien, tieteellisen julkaisutoiminnan, elektronisen julkaisemisen ja atk-tekniikan sekä tekijänoikeuskysymysten keskeiset asiantuntijatahot. Tämä mahdollistaa eri ongelmien kattavan rinnakkaisen tarkastelun, ja näin teknisten ongelmien lisäksi myös vaikeasti ratkaistavat tekijänoikeus- ja maksullisuusperiaatteet voidaan sisällyttää kokeiluun ja etsiä niille valtakunnallisia ratkaisuja. Projekti pyrkii myös aktiivisesti luomaan yhteisiä muihin vastaaviin eurooppalaisiin ja kansainvälisiin hankkeisiin.

1.1 ELEKTRA-projektin tavoitteet ja tehtävät

Projektin tavoitteina on:

- toteuttaa julkaisujen haku- ja välityspalvelu verkkoympäristössä ja evaluoida käyttöä todellisilla käyttäjillä
- luoda malli korkeakoulujen elektronista kurssimateriaali- ja väitöskirjapalvelua varten
- selvittää tekijänoikeuskysymyksiä ja kehittää sopimusmallit erityisesti tieteellisten julkaisujen välittämiseen
- selvittää maksullisuusperiaatteita sekä laskutus- ja suojauskäytäntöjä ja
- luoda yhteyksiä kansainvälisiin elektronisen julkaisemisen hankkeisiin.

Projektin toteutukseen tarvittava infrastruktuuri on pääosin jo valmiina. Jakelukanavana käytetään Internetin WWW-palvelua ja valtakunnallinen LINNEA-järjestelmä puolestaan tarjoaa hakupalvelut. VTLS-ohjelmisto sisältää jo nyt tekstin ja kuvan linkittämiseen tarvittavat uudet piirteet. Monista elektronisen julkaisemisen kokeiluista poiketen projektissa julkaisut linkitetään Marc-tietueisiin, mistä on etuna monipuoliset hakumahdollisuudet. Kokeiluun sisältyy samanaikaisesti myös ratkaisut sekä valtakunnallista yhteisluettelo- että paikallista kirjastoluetteloympäristöä varten, jolloin projektin tulokset ovat suoraan sovellettavissa kaikissa yliopistokirjastoissa.

Projektin tarkoituksena on kytkeä Helsingin yliopiston kirjaston yhteisluettelotietokantojen ja Oulun yliopiston kirjastotietokannan bibliografisiin viitteisiin itse julkaisuja. Yhteisluettelotietokannoista ensisijaisena pilottiympäristönä käytetään ARTOa, joka on julkisessa käytössä oleva artikkeliviitetietokanta. ARTO sisältää tietoja yli 1000 kotimaisesta aikakauslehdestä vuodesta 1994 lähtien, ja viitemäärä vuoden 1997 lopussa on noin 200 000. Aineistoa ARTOon tallentaa 37 asiantuntijaorganisaatiota.

Hakupalvelut sijaitsevat HP3000-laitteella ja ohjelmistona on VTLS. WWW-käyttöliittymä on kehitetty tarkoitusta varten ja se sisältää runsaasti ELEKTRAN käyttöön tarkoitettuja lisäpiirteitä. Julkaisut kuvaillaan tietokantaan Marc-formaatin avulla, mikä helpottaa tiedonhakua verrattuna nykyisiin WWW-indeksointipalveluihin, joissa yleensä indeksoidaan julkaisun koko teksti tai pelkkä otsikko. Atk-tallenteiden FINMARC-formaatti on valmistunut projektin aikana ja sitä sovelletaan sekä projektissa että myös muussa Linnea-yhteisluetteloinnissa.

Eri aineistotyypeistä kokeiluun otetaan artikkeleja, yliopiston omien julkaisusarjojen raportteja, väitöskirjoja sekä oppimateriaalia. Aineistot valitaan siten, että pääosa saadaan käyttöön valmiiksi elektronisessa muodossa. Tekijänoikeuskysymykset ratkaistaan sopimuksin ja nimellisiin tekijänoikeuskorvauksiin varaudutaan. Väitöskirjakokeilu toteutetaan ensisijaisesti Oulun yliopiston kirjastossa.

Testiaineisto annetaan todellisten tietokannan käyttäjien koekäyttöön, kerätään käyttäjäpalaute ja analysoidaan sen tulokset. Palautteen avulla selvitetään käytön määrää ja käyttäjien reaktioita palvelun sisältöön, maksullisuusperiaatteisiin ja käyttöliittymiin.

Projektin aikana selvitetään lisäksi tietoturvakysymyksiä sekä maksullisuuskäytäntöjä ja laskutusohjelmia. EU:n tutkimusohjelmissa on käynnissä useita projekteja, esim. Desire, Copinet ja Copycat, joissa kehitetään paitsi laskutus- myös suojausohjelmia. Pilottiprojektissa suojaukset voidaan hoitaa käyttäjätunnusten ja salasanojen avulla.

Marc-formaattiin on hiljattain lisätty kenttä 856 (Electronic Location and Access), johon voidaan tallentaa elektronisen dokumentin URL eli Uniform Resource Locator -tunnus sekä muuta hyödyllistä tietoa, kuten tiedoston koko. Käyttöliittymän ohjelma lukee *URL:n* ja noutaa sen perusteella halutun dokumentin ja esittää sen käyttäjälle. VTLs-ohjelmassa kentän 856 käyttö on jo nyt mahdollista. URL-tunnukseen perustuva viitteen ja siihen liittyvän dokumentin linkkaus on hyvin joustavaa. Itse dokumentti voi sijaita joko HYK:in ylläpitämällä arkistopalvelimella tai esim. kustantajan palvelimella. HYK:in palvelinten tallennuskapasiteettia mitoitettaessa on varauduttu projektin tarpeisiin.

Eri aineistotyypeille soveltuvat tiedosto- ja kuvaformaattit selvitetään projektin aikana. Formaattien valinnassa kiinnitetään huomiota mm. niiden yleisyyteen ja tilavaatimuksiin. WWW:ssä yleisesti käytetty HTML (Hypertext Markup Language) soveltuu sellaisten dokumenttien esittämiseen, jotka sisältävät pääasiassa tekstiä ja kuvia. Näyttävän painoasun omaaville tai esim. matemaattista esitystä sisältäville dokumenteille saattaa Adoben Acrobat-formaatti (PDF) osoittautua käyttökelpoiseksi. Projektissa otetaan mahdollisuuksien mukaan huomioon myös yleisen SGML-standardin soveltuvuus tallennettaville dokumenteille ja sen käyttöä testataan väitöskirja-aineistolla. Kuvien tallennukseen valitaan yleinen, tilaa säästävä, mutta samalla riittävän tarkkuuden säilyttävä formaatti (esim. JPEG). Kuvien tallennuksessa otetaan lisäksi huomioon verkkojulkaisemisen ja painetun julkaisun asettamat tarkkuuserot.

1.2 Osallistajat ja osapuolten vastuut

Projektin osallistajat ovat

- Helsingin yliopiston kirjasto ja atk-keskus
- Oulun yliopiston kirjasto
- Kopiosto ry
- Tieteellisten seurain valtuuskunta
- VTT Tietopalvelu
- Gaudeamus kirja
- Yliopistopaino

Helsingin yliopiston kirjasto koordinoi ja johtaa projektia, tarjoaa tuotantoympäristön elektronisen julkaisemisen kokeilua varten ja tarjoaa pääosan työhön tarvittavista henkilöresursseista sekä vastaa maksullisuus- ja laskutuskäytäntöjen, tietoturvaan ja julkaisujen arkistointiin liittyvistä teknisistä selvityksestä sekä käyttäjäpalautteen analysoinnista. Oulun yliopiston kirjasto osallistuu kokeiluun väitöskirja-aineiston osalta ja tarjoaa ympäristön kokeilua varten.

Julkaisijoita projektissa edustavat Gaudeamus Kirja ja Yliopistopaino, jotka tarjoavat elektronisessa muodossa olevaa oppikirja- ja muuta aineistoa sekä toimitusteknistä ja julkaisemisen asiantuntemusta projektin käyttöön.

Kopiosto vastaa projektissa tekijänoikeuksiin liittyvästä asiantuntemuksesta, sopimuskäytäntöihin sekä maksullisuusperiaatteisiin liittyvistä periaatteellisista selvityksistä sekä hoitaa yhteyksiä oikeudenhaltijoihin. Tieteellisten seurain valtuuskunta vastaa myös osaltaan tieteellisen julkaisutoiminnan asiantuntemuksesta, osallistuu aineiston valintaan ja toimittamiseen sekä käyttöoikeuksien hankintaan.

VTT Tietopalvelu on pohjoismainen elektronisen julkaisemisen osaamiskeskus, ja VTT tarjoaa teknistä asiantuntemusta ja koulutusta elektroniseen julkaisemiseen liittyvissä asioissa. Helsingin yliopiston atk-keskus tarjoaa verkkokäytön, asiakasohjelmistojen ja laitteistojen teknistä asiantuntemusta.

1.3 ELEKTRA-projektin tuottamat hyödyt

ELEKTRA-projektin avulla saavutettuina hyötynä voidaan pitää seuraavia seikkoja :

- mahdollistaa tieteellisen aineiston julkaisemisen ja välittämisen uudella teknologialla suhteellisen laajassa mitassa
- tarjoaa valtakunnallisen kanavan korkeakouluissa julkaistavan aineiston verkkokäyttöön ja mallin elektronisten väitöskirjojen välittämiseen
- antaa viitetietokantojen käyttäjille mahdollisuuden saada primaarijulkaisut tiedonhaun yhteydessä
- nopeuttaa merkittävästi tiedon julkaisua ja välitystä
- antaa kokemusta tekijänoikeuksien hankinnasta ja hallinnoinnista
- antaa kokemusta maksullisen välityspalvelun kehittämisistä ja taloudellisista malleista
- antaa kokemusta menettelytavoista aineiston autenttisuuden ja saatavuuden varmistamiseksi
- antaa kokemusta elektronisen julkaisemisen sekä julkaisujen käyttöön asettamisen ja välittämisen kokonaisprosessista
- tarjoaa kilpailukykyisen palvelun, joka vastaa tekijöiden, julkaisijoiden ja tiedon käyttäjien tarpeita
- luo kansainvälisiä yhteyksiä muihin vastaaviin kehittämishankkeisiin

1.4 Projektin resurssit, organisointi ja toteutus

Projekti kuuluu opetusministeriön Tietoyhteiskunta -ohjelmaan ja saa tukea tämän ohjelman kautta. Projektin budjetin muodostavat osallistujien oma rahoitus budjettivaroin sekä opetusministeriön avustus. Projektin kokonaismenoarvio vuosille 1996-97 on ollut 1 730 000 markkaa.

Projektin kuluessa selvityksistä ja osavaiheista on laadittu raportit ja osaraporttien tulokset on koostettu kokonaisraportiksi. Pääosa raporteista on julkaistu ELEKTRAN WWW-sivuilla osoitteessa <http://linnea.helsinki.fi/elekra/raport.html>. Opetusministeriö on seurannut projektin etenemistä ja arvioinut jatkorahoitustarpeen.

Projektilla on ollut ohjausryhmä, jossa kaikki osallistujat ovat edustettuina. Johtoryhmä on toiminut asiantuntijaelimenä, joka on ohjannut projektin työtä, hyväksynyt suunnitelmat, päättänyt resurssien käytöstä ja hyväksynyt tulokset.

Projektin taloudesta vastaavana koordinaattorina on toiminut Inkeri Salonharju, joka on raportoinut sekä johtoryhmälle että opetusministeriölle. Projektia on johtanut projektipäällikkö, jonka vastuulla on ollut projektin toteutus, projektin vaiheiden dokumentointi ja tiedotus. Oulun pilotilla on ollut oma vastuullinen tutkija ja projektiryhmä.

Projektin asiantuntijat ovat osallistuneet tehtäväkohtaiseen työryhmätyöskentelyyn ja antaneet muita projektin edellyttämiä asiantuntijapalveluja. Projektiryhmä on toteuttanut projektin tehtävät ja toiminut tiiviissä yhteistyössä taitotiedon kehittämiseksi ja jakamiseksi.

Aineiston tuottajien vastuuna on ollut tarjota pilottiaineistoa projektin käyttöön, mutta kustantajaosapuolet eivät ole tässä tehtävässään onnistuneet, eikä suunnitelmien mukaista oppimateriaalia ole saatu mukaan kokeiluun.

1.5 Projektista ja tuloksista tiedottaminen

ELEKTRA-projektin tiedotusta varten on laadittu erillinen suunnitelma, ja päätiedotuskanavana on käytetty omia WWW-sivuja. Kotisivut sisältävät sekä julkisen osan, jolla tiedotetaan projektin etenemisestä sekä suomen- että englanninkielisille verkkokäyttäjille ja salasanalla suojatun, projektin osallistujien omaan käyttöön tarkoitetun arkiston.

Hanketta on esitelty useilla messuilla ja konferensseissa, kuten Tieto97 ja Tieto 98, Turun kirjamessut ja IFLA-konferenssin näyttelyssä Kööpenhaminassa sekä yhteispohjoismaisessa teki-jänoikeusjärjestöjen tapaamisessa Moldessa 1998. Projektista on laadittu lehdistötiedote sekä suomen ja englanninkieliset esitteet, joita on jaeltu laajasti.

Projektista on laadittu lukuisia osaraportteja, joiden tulokset kootaan loppuraportiksi. Loppuraportti julkaistaan sekä verkkojulkaisuna että painettuna. Projektiin liittyvä asiantuntijaseminaari järjestettiin Helsingissä 20.3.1998.

2. YHTEENVETO PROJEKTIN TULOKSISTA, JOHTOPÄÄTÖKSET JA JATKOKEHITTÄMINEN

2.1 Yhteydet julkaisijoihin ja aineiston hankinta

Internetin kautta saatavia verkkojulkaisuja selvitettiin VTT Tietopalvelun ja Helsingin yliopiston kirjaston ylläpitämien lehtilistojen pohjalta. Projektin valintakriteereiden mukaisia verkkojakelussa olevia dokumentteja löytyi kuitenkin niin vähän, että päädyttiin tekemään erillinen kysely Tieteellisten seurain valtuuskunnan jäsenorganisaatioiden verkkojulkaisutilanteesta ja halukkuudesta antaa aineistoaan ELEKTRA-projektin käyttöön.

Selvitysten tuloksista ilmeni, että vuoden 1997 lopussa kotimaisia verkkolehtiä ilmestyi yhteensä 347. Niistä 83 % oli aikakauslehtiä, 14 % sanomalehtiä ja 3 % sarjajulkaisuja. Lehdistä 55 % oli rinnakkaisjulkaisuja painetun version kanssa, 25 % paperilehden liitteitä ja 19 % vain elektronisessa muodossa ilmestyviä lehtiä (66 lehteä). Tieteellisten seurain jäsenorganisaatiot julkaisivat noin 100 lehteä, joista neljä ilmestyy elektronisena, kolme painetun lehden rinnakkaisjulkaisuna ja yksi pelkästään verkossa. Näistä lehdistä 64 indeksoidaan HYK:in ARTO-tietokantaan. ELEKTRA-projektissa on tähän mennessä julkaistu artikkeleja 36 julkaisusta, joista 35 liittyy tieteellisten seurojen piiriin.

Julkaisijoille tehtyyn kyselyyn vastasi 40 julkaisijaa. Heistä 30 (75 %) oli kiinnostuneista kokeilusta ja vain neljä suhtautui hankkeeseen kielteisesti. Aineistoaan verkkojakeluun halusi 29 (73 %) vastaajista, arkistoitavaksi HYK:in palvelimelle 25 (63 %) ja digitoitavaksi 19 (48 %). Valtaosa julkaisijoista on toimittanut lehtensä tekstinkäsittelytiedostojen (32) tai sivuntaitto-ohjelmien (21) pohjalta. Taitto on tehty kokonaan kirjapainossa 11 julkaisijalla ja elektroniseen muotoon aineistoa on tuottanut vain 3 julkaisijaa. Internetistä löytyi kotisivut kolmannekselta vastaajista.

ELEKTRA-projektin ansiosta noin joka kolmas tieteellisten seurojen lehtijulkaisija on saanut kokemusta elektronisesta julkaisutoiminnasta ja aineistonsa verkkojakelusta. Tämä on enemmän kuin kyselyn perusteella voitiin odottaa. Valtaosa aineistosta on jouduttu digitoimaan, mutta jatkohankkeessa on tarkoitus ohjeistaa julkaisijoita tuottamaan aikaisempaa enemmän aineistoaan suoraan elektroniseen muotoon. Tämä alentaa palvelun tuotantokustannuksia ja parantaa aineiston laatua.

Projektin tavoitteiden mukaisesti testiaineistoksi on valittu artikkeli-, väitöskirja- ja muuta monografia-aineistoa. Suunnitellusta oppimateriaalipilotista jouduttiin luopumaan, koska testiaineistoa ei saatu kustantajilta. Aika ei vielä ollut kypsä kaupallisen oppimateriaalin aineiston verkkojulkaisukokeiluun.

Helsingin yliopiston kirjaston pilottiin on sisällytetty sekä artikkeleja että monografioita, mutta Oulun yliopiston kirjaston pilotissa keskityttiin väitöskirjoihin ja muihin monografioihin. Pilotin sisältöalaa ei ole rajattu millekään erikoisalalle, vaan mukaan on otettu aineistoa eri aloilta. Näin on saatu simuloitua mahdollisimman hyvin todellinen palveluympäristö ja kartoitettua tuottajien verkkokehitystilanne ja kiinnostus elektronisen palvelun kehittämiseen, markkinapotentiaali ja käyttäjien reaktiot.

Artikkeliaineiston valinnan yhtenä kriteerinä on käytetty lehden indeksointia ARTO-tietokantaan ja ensisijaisesti on otettu mukaan niitä lehtiä, joiden julkaisijat ovat Tieteellisten seurain val-

tuuskunnan jäseniä. ARTOon indeksoidaan tällä hetkellä yli 1000 kotimaista aikakauslehteä. Lehtiä on kaikilta tieteen aloilta ihmislääketiedettä lukuun ottamatta ja mukana on myös harastus- ja paikallislehtiä.

Kokeiluun on saatu mukaan yhteensä 36 julkaisun aineistoa. Alat kattavat bio-, elintarvike- ja ympäristötieteitä, geologiaa, kasvatustieteitä, kieli- ja kirjallisuustiedettä, oikeustieteitä, metsätiedettä, yhteiskuntatieteitä, maantiedettä, hammas- ja eläinlääketiedettä, kirkkohistoriaa, tulevaisuudentutkimusta ja yleistieteellisiä lehtiä. Mukana on myös joitakin tiedotuslehtiä. Alakirjo on laaja ja julkaisijat ovat olleet erityisen aktiivisia oikeus- ja metsätieteiden alalla. Molemmilta aloilta aineistoa on saatu mukaan kaikista keskeisistä julkaisuista.

Väitöskirjapilotissa OYK on julkaissut eri alojen uusia väitöskirjoja ja HYK on tehnyt yhteistyötä Suomen historiallisen seuran kanssa, jonka julkaisemat uudet väitöskirjat on saatu kokeiluun. HYK on toteuttanut myös väitöskirjojen julkisesti nähtäväksi asettamisen verkossa. Väitöskirjat ovat olleet vapaasti luettavissa verkossa 10 päivää ennen väitöstilaisuutta, mutta tulos on estetty. Tämän jälkeen väitöskirjojen käyttö on rajoitettu vain ELEKTRAN testikirjastoihin. Laskurin mukaan ensimmäisten väitöskirjojen sivuja on luettu noin 400 kertaa.

2.2 Digitointiympäristö ja julkaisupalvelut

Pilotteja varten saatavan valmiiksi elektronisessa muodossa olevan aineiston määrä on osoittautunut ennakoitua vähäisemmäksi, jolloin digitoinnin tarve ja sen edellyttämät resurssit ovat muodostuneet keskeisiksi projektissa. Pilottien julkaisuympäristö koostuu Pentium-tasoisista PC-työasemista, jotka on varustettu 64 Mt:n keskusmuistilla ja 1,7 Gt:n kiintolevyillä. Lisäksi laitteistoon kuuluu HP ScanJet 4C väriskanneri.

Artikkelipilottiin sisällytettyjen julkaisujen tekninen taso ja tiedostomuodot ovat vaihdelleet suuresti. Osa artikkeleista on saatu erilaisilla tekstinkäsittely- tai suvuntaitto-ohjelmilla tuotetuissa tiedostomuodoissa (Microsoft Word, Corel WordPerfect, Corel Ventura, Adobe PageMaker tai QuarkXpress) ja osa paperikopioina. Artikkelit on digitointiprosessissa muutettu joko HTML- tai PDF-muotoon ja väitöskirjoja on muunnettu myös SGML-muotoon.

Painetun aineiston digitoinnissa on käytetty lähdekielestä ja aineiston graafisesta ulkoasusta riippuen joko Visioneer PaperPort, OmniPage tai Adobe Acrobat Capture-ohjelmia. Kuvat ovat tuottaneet ongelmia ja niiden tallennusmuotona on JPG, TIFF tai GIF. Valmiiksi tiedostomuodossa olevan aineiston käsittely on periaatteessa helpompaa ja lopputulos laadukkaampaa kuin skannaamalla, mutta käsittelyajat ja samalla kustannukset vaihtelevat tälläkin aineistolla suuresti. Parhaaseen lopputulokseen päästäisiin jos julkaisijat toimittaisivat aineiston valmiiksi elektroniseen muotoon konvertoituna, verkkojakelua varten PDF- tai HTML-tiedostoina ja pitkäaikaista arkistointia ja paperitulostusta varten PostScript-muodossa.

Valmiina digitoidussa muodossa on 1.3.1998 yli 1300 artikkelia, 10 väitöskirjaa ja 10 muuta monografiaa.

2.3 Julkaisujen verkkojakelu ja käyttökokeilu

ELEKTRAN haku- ja dokumentin välityspalvelu perustuu nykyisen artikkelitietokannan, ARTOn näyttöluettelotoiminnon laajennukseen. Tarkoituksena on, että asiakkaat voivat hyödyntää näyttöluettelon hyviä hakuominaisuuksia ja korkeatasoisia bibliografisia- ja sisällönkuvailuja, jotka jo nyt tuotetaan tieteellisten kirjastojen ja muiden asiantuntijaorganisaatioiden laajana

yhteistyönä. Palvelun taso parantuu oleellisesti, kun asiakkaat saavat tiedonhaun yhteydessä käyttöönsä itse varsinaiset dokumentit.

Pilottiaineiston koekäyttöä varten on perustettu testitietokanta ELEKTRA HP-3000 tietokantapalvelimelle VTLS-ympäristöön. Uuden tietokannan perustamiseen tarvittava lisenssi on saatu VTLS Inc:iltä, ja se oikeuttaa testitietokannan käytön projektin ajan. Elektronisen aineiston luettelointiohjeet on laadittu projektissa, ja ELEKTRAN lisäksi niitä sovelletaan jo laajasti LINNEA-kirjastoissa. Tietokantaan on tallennettu testiaineiston bibliografiset tietueet, jotka ensin tallennetaan ARTO-tietokantaan ja kopioidaan sieltä testitietokantaan. Tietueisiin lisätään URL-osoitteet kenttään 856 hypertekstilinkkien luomiseksi. Luetteloidun aineiston määrä on yli 1300 tietuetta.

ELEKTRA-projekti on luonut myös elektronisen aineiston (atk-tallenteiden) Finmarc-kuvailun edellyttämät luetteloinnin sovellusohjeet ja luettelointikäytännöt, joita on ELEKTRAN lisäksi hyödynnetty jo nyt laajasti Linnea-yhteisluetteloinnissa. ELEKTRA-aineistoon on osittain tallennettu myös Dublin Core formaatin mukaisia metadatatietoja, jotka sisällytetään itse dokumenttiin. Työkalut HTML-julkaisujen DC-kuvailuja varten ovat koekäytössä ja kokeilua jatketaan myös ELEKTRAN aineistolla.

Julkaisujen jakelukanavana käytetään Internetin WWW-palvelua. Elektronisten dokumenttien arkistointia ja WWW-palvelinohjelmiston käyttöä varten on hankittu Pentium Pro mikrotietokone, joka on varustettu 64 Kt:n keskusmuistilla ja 2 Gt:n SCSI-kiintolevyllä. Käyttöjärjestelmänä on Linux, ja koneessa toimii WWW-palvelin ja WWW-pohjainen hakupalvelun käyttöliittymäohjelma. Käyttöliittymäohjelma on Jyväskylän yliopistossa kehitetty Hyperlib, johon on kehitetty runsaasti lisäpiirteitä, mm. käytön seuranta ELEKTRAN tarpeita varten.

Testitietokannan tietoturva on ratkaistu mahdollistamalla käyttö vain tarkoitusta varten varuilla laitteilla IP-osoitteen avulla. Työasemilta on mahdollisuus aineiston lukemiseen ja tulostukseen oheiskirjoittimelle, mutta tiedostojen kopiointi on estetty. Työasemilla tarvitaan WWW-selainohjelma Netscape sekä Adoben Acrobat Reader -ohjelma PDF-muotoisten dokumenttien lukemista varten.

ELEKTRA-tietokannan koekäyttö on meneillään seitsemässä tieteellisessä kirjastossa, Helsingin, Jyväskylän ja Oulun yliopiston kirjastossa, Teknillisen korkeakoulun kirjastossa, VTT Tietopalvelussa ja Helsingin yliopiston Maatalous- ja Luonnontieteiden kirjastossa. Helsingin yliopiston luonnontieteiden kirjastossa palvelua ei ole sopimuksista huolimatta kuitenkaan käynnistetty lainkaan. Koekäyttöä varten ko. kirjastojen kanssa on solmittu kirjallinen sopimus ja kirjastoille on tuotettu opaskansio, joka sisältää kaikki ELEKTRA-tietokannan käyttöön liittyvät ohjeet sekä testikäytön yleiset periaatteet ja raportointiohjeet. ELEKTRAN WWW-käyttöliittymään on myös sisällytetty kaikki käyttäjän tarvitsemat opastetiedot, jotka käsittelevät tekijänoikeuksia, tiedonhakua ja tulosteiden hinnoittelua. Kirjastoille ei ole kuitenkaan järjestetty varsinaista aiheeseen liittyvää koulutusta, mikä olisi todennäköisesti ollut tarpeen palvelujen paikallista tiedotusta varten.

ELEKTRA-tietokannan lukukäytöstä on kerätty seurantatietoja käyttöliittymään yhdistetyn lokiseurantaohjelman avulla, joka raportoi teoskohtaisesti artikkeleiden ja muiden dokumenttien lukemiskerrat. Lisäksi kokeilukirjastot raportoivat manuaalisesti kerätyt tiedot teoskohtaisista sivutulostemääristä. ELEKTRAN kokeilukäyttö aloitettiin kirjastoissa 15.6.1997 ja varsinaisesti se pääsi käyntiin vasta syyslukukauden alussa 1.9.1997. Kokeilukausi on näin muodostunut varsin lyhyeksi. Marraskuun alusta käynnistettiin palautetietojen keräys ELEKTRAN käyttäjiltä käyttöliittymän yhteydessä olevalla suorakäyttölomakkeella ja kirjastoissa jaettavalla paperilo-

makkeella. Palautelemakkeen avulla haluttiin asiakkaiden arvioita käyttöliittymän toimivuudesta, palvelun sisältöön liittyvistä asiakkaiden toiveista ja reaktioista maksullisuuteen.

Koekäytön tulokset olivat varsin vaatimattomat. ELEKTRA-dokumentteja on luettu kokeilukirjastoissa yhteensä 200 kertaa. Eniten palvelua on käytetty Oulun yliopiston kirjastossa, yhteensä 80 kertaa ja seuraavana ovat HYK 47 kertaa ja Maatalouskirjasto 43 kertaa. Teknillisen korkeakoulun kirjastossa eikä VTT:llä palvelua ole käytetty juuri lainkaan ja Jyväskylän yliopiston kirjastossa vain 19 kertaa. Myöskään maksullista tulostuspalvelua ei ole käytetty kuin yhdessä kirjastossa.

Lehdittäin tarkasteltuna eniten käyttökertoja kohdistui jo muutenkin avoimessa verkkokäytössä olevaan Yliopistolehteen. Myös HYK:in tiedotuslehtiä on käytetty jonkin verran. Muuten käyttö hajaantuu eri lehtien kesken. Tulosten perusteella voi päätellä, että palvelun käyttäjät ovat olleet lähinnä kirjaston henkilökuntaa. Palvelusta tiedottamista kirjastot eivät ole kokeneet omaksi tehtäväkseen eivätkä ole sitä tehneet. Palvelun löytäminen on jäänyt satunnaisten aktiivisten asiakkaiden omatoimisuuden varaan.

Käyttäjäpalautekyselyyn on saatu vastaukset 31.1.1998 mennessä 21 vastaajalta. Tämä on noin 10% ELEKTRA-palvelun käyttäjistä. Kirjastoja aktivoimalla olisi todennäköisesti päästy pienin ponnistuksin kattavampaan tulokseen. Käytettävissä oleva otos on kuitenkin niin pieni, että tulokset ovat vain viitteellisiä, mutta vastaukset antavat arvokasta lisävalaistusta lokiseurannan kylmille numeroille.

Vastaajista puolet oli opiskelijoita ja kolmannes kirjastojen työntekijöitä. Kustantajien edustajiin otettiin henkilökohtaisesti yhteyttä, ja heille annettiin testikäyttömahdollisuus omalta työpaikalta pariksi päiväksi ja tätä mahdollisuutta käytti hyväksi kolme vastaajaa. Vastaajista 7 tuli Maatalouskirjastosta, 5 TKK:n kirjastosta, 4 Oulun yliopiston kirjastosta ja 2 HYK:istä. ELEKTRAan liittyvää tiedotusta oli lähes kaikkien vastaajien mielestä ollut liian vähän ja vain puolet oli löytänyt ELEKTRAN käyttöön tarkoitetun työaseman helposti. ELEKTRAN käyttöliittymää ja opasteita sen sijaan pidettiin helppoina ja selkeinä.

Vähäistä tulostusta selitti kysymys, jossa 19 vastaajaa ilmoitti että ei ollut tulostanut, 7 koska siihen ei ollut tarvetta ja 2 koska se oli liian kallista. Käyttäjistä puolelle siis tuntui riittävän pelkkä lukeminen. Hintoja vajaa puolet vastaajista piti kohtuullisina, neljännos kalliina ja kolmannos ei osannut muodostaa mielipidettään. ELEKTRA-palvelun ihanteellisena ympäristönä puolet piti yliopistokirjastoja, kolmannos maksullista Internet-levitystä ja vain 2 vastaajaa yleisiä kirjastoja. Kehittämishdotuksia tuli neljältä vastaajalta, ja niissä palvelua sinänsä pidettiin tervetulleena ja sen laajentamista vapaaseen verkkojakeluun toivottavana suuntana. Myös linkkejä suoraan lehdistä artikkeleihin toivottiin nykyistä selkeämmiksi.

Vaatimaton menestys kokeilukäytössä voidaan katsoa ELEKTRA-projektin ensimmäisen vaiheen suurimmaksi heikkoudeksi, vaikka resursseja on tietoisesti keskitetty julkaisupalvelun kehittämiseen ja aineiston digitointiin. ELEKTRAN kontrolloitu ja rajattu käyttömahdollisuus samoin kuin käyttökertoihin perustuva hinnoittelu ovat vastanneet kustantajien ja tekijöiden tämänhetkisiä tarpeita. Malli ei kuitenkaan mahdollista palvelun käyttöä riittävän laajasti. Jos sekä lukemis- että tulostusmahdollisuudet keskitetään vain kirjastoihin, joissa sama materiaali on myös painettuna ja asiakkaiden valokopioitavissa elektronista vaihtoehtoa halvemmalla hinnalla, syntyy palvelun kannalta epäedullinen kilpailutilanne. Sitä mukaa kun elektroninen julkaiseminen alkaa korvata painetun version, tilanne tietysti pikkuhiljaa muuttuu.

Elektronisen aineiston käyttöä voidaan kuitenkin edistää ja palveluja kehittää käyttäjien tarpeita paremmin vastaaviksi monin tavoin. Palvelusta tiedottamalla ja käyttäjiä kouluttamalla saata-

neen nopeimmin toivottuja tuloksia. ELEKTRAn toisessa vaiheessa on jo suunniteltu keskitytävän käyttäjäpalvelujen kehittämiseen. Sopimuskirjastoissa rajoitetun käytön mukaiseen hinnoitteluun perustuvan kokeilun rinnalle tarvitaan laajempia jakelukanavia. Jatkohankkeessa on tarkoitus kehittää sopimusmekanismeja organisaatiolisensseihin perustuvaan yliopistoissa ja vaikkapa myös muissa oppilaitoksissa kuten ammattikorkeakouluissa ja lukioissa tapahtuvaa avointa paikallisverkkokäyttöä varten. Kokeilu voisi tapahtua osana Kansallisen elektronisen kirjaston kokeiluhanketta ja tarvittaessa pilotti voitaisiin rajata alueelliseksi kokeiluksi.

Myös kansalaisen sähköisen identiteetin kehitystä seurataan ja ollaan valmiita toimimaan kotimaisten elektronisten maksujärjestelmien testihankkeena.

2.4 Tekijänoikeudet ja sopimuskäytännöt

ELEKTRAn tekijänoikeusperiaatteiden luomiseksi Kopiosto on laatinut selvityksen, joka käsittelee digitaalisten käyttöoikeuksien luovutusta, käytön ehtoja ja korvauksia, oikeuksien hankintaa ja hallinnointia, sopimuksia, korvauksia sekä käyttötapahentumien rekisteröintiä ELEKTRAssa. Selvityksen perusteella ELEKTRA-projektissa uudelleenjulkaistaan kirjallisia teoksia ja niihin liittyvää kuvitusta, valokuvia ja mahdollisesti luetteloina suojattuja teoksia. Tämä aineisto on tekijänoikeuden alaista ja edellyttää sopimuksia oikeudenhaltijoiden ja ELEKTRA-projektin välillä.

Painettuja julkaisuja varten kustantaja tekee kustannussopimuksen tekijän kanssa. Sopimus voi olla kirjallinen tai suullinen tai ns. hiljainen sopimus. Sopimus kattaa vain ne käyttötavat ja –muodot joista on nimenomaisesti sovittu. Epäselvissä tapauksissa sopimuksia tulkitaan yleensä tekijän eduksi. ELEKTRA-projektissa Kopiosto selvittää kenellä teoksen oikeudet ovat. Ellei kustantaja nimenomaan näytä, että myös digitaaliset oikeudet ovat siirtyneet hänelle, sopimus julkaisuun sisältyvien artikkeleiden käytöstä tehdään niin tekijöiden kuin kustantajan kanssa.

Tekijänoikeuksien hankintaa varten on laadittu kirjallinen mallisopimus sekä tekijöitä että kustantajia varten. Sopimuksen mukaan Helsingin yliopiston kirjasto saa oikeuden dokumenttien saattamiseen digitaaliseen muotoon ja rajoitettuun verkkolevityskokeiluun. Oikeus on rajattu kestäväseen ELEKTRA-projektin ajan. Projektin päämääränä on, että teosten oikeudenhaltijat voivat kukin määrätä sekä teoksen käytöstä perittävästä korvauksista että aineiston käyttöehtojista. Vaikka elektroniset maksujärjestelmät vielä puuttuvat, on oikeudenhaltijalla mahdollisuus antaa oikeus sekä dokumentin lukemista että tulostusta varten tai rajata oikeus pelkkään lukemiseen. Dokumentin tulostus tiedostoksi on kielletty ja teknisesti pyritty estämään. Oikeudenhaltijalla on myös mahdollisuus määrittää teoksensa tulostamisesta perittävä korvaus.

ELEKTRA-pilotin aineiston välityskokeilu on käynnissä seitsemässä tieteellisessä kirjastossa, jotka ovat solmineet koekäytöstä kirjallisen sopimuksen Helsingin yliopiston kirjaston kanssa. Sopimuksen mukaan kirjasto tarjoaa asiakkaiden käyttöön yhden tai useamman työaseman, joka on määritelty IP-osoitteen avulla. Työasemalla ei ole mahdollisuutta tallentaa dokumentteja tiedostoiksi eikä tulostaa aineistoa muualle kuin asiakaspalvelun verkkotulostimelle. Kirjasto sitoutuu perimään hinnaston mukaiset korvaukset asiakkailta ja tulouttamaan ne HYK:ille. Kirjasto sitoutuu myös raportoimaan tulosteiden määrät teoskohtaisesti.

Tekijänoikeussopimusten laadinnasta ja oikeuksien hankinnasta ELEKTRA-projektissa on vastannut Kopiosto. Oikeuksien hankintaan on osallistunut oman jäsenkuntansa osalta aktiivisesti myös Tieteellisten seurain valtuuskunta. Kirjallinen sopimus on allekirjoitettu jo yli 1300 oikeudenhaltijan kanssa, ja lähes kaikki tekijät ja kustantajat ovat hyväksyneet joko suositellut sopimusehdot tai rajoitetut oikeudet esim. pelkkään lehden lukemiseen ilman tulostusoikeutta.

Oikeuksien hankinta ja hallinnointi, jossa joudutaan käyttämään runsaasti ATK-teknisiä sovelluksia, on kuitenkin vaatinut resursseja enemmän kuin alkuperäisessä suunnitelmassa oletettiin. Näihin tehtäviin on jouduttu palkkaamaan projektia varten kokopäivätoiminen henkilö. Tämä huomioidaan jatkohankkeen resurssisuunnitelmissa.

2.5 Maksullisuus ja korvausten hallinnointi

ELEKTRAn julkaisu- ja jakelupalvelusta aiheutuvat kustannukset on katettu osin OPM:n avustuksella ja osittain osallistujien omarahoituksella, joka on tavallisimmin sisältänyt henkilötyötä. Julkaisijoille ja tekijöille kokeiluun osallistuminen on ollut maksutonta. Suurimman menoerän projektissa ovat aiheuttaneet julkaisijoille tarjottu digitointipalvelu, tekijänoikeuksien hankinta ja hallinnointi sekä laitteisto- ja ohjelmistohankinnat.

ELEKTRAn haku- ja dokumentinvälityspalvelu on ollut käyttäjille osin maksutonta ja osin maksullista. Dokumenttien hakeminen ja lukeminen näytöltä on ollut asiakkaille maksutonta, mutta tulostuksesta on veloitettu sivujen määrään perustuva hinta. Helsingin yliopiston kirjasto on maksanut oikeudenhaltijoille korvauksen teosten tallentamisesta ELEKTRAn dokumenttipalvelimelle. Tallennuskorvauksen suuruus on ollut lehden niiden vuosikertojen hinta, joita on sisällytynyt pilottiaineestoihin.

Artikkelitulosteiden hinnoittelu perustuu kunkin julkaisun painetun vuosikerran yhteisötilaus hintaan ja kokonaissivumäärään. ELEKTRAn suositus on, että tulosteesta peritään artikkeli-kohtainen hinta, joka saadaan jakamalla vuosikertatilaushinta paperiversion kokonaissivumäärällä. Kukin oikeudenhaltija voi kuitenkin määrittää julkaisulleen itse sopivaksi katsomansa hinnan. Suosituksen mukaiset sivun tulostehinnat vaihtelevat 0,50–2,00 markan välillä ja suositus minimikorvaukseksi on 0,50 markkaa. Oikeudenhaltijoilla on ollut myös mahdollisuus itse hinnoitella oma tuotensa. Käytännössä kuitenkin lähes aina on hyväksytty hintasuositus tai vaihtoehtoisesti ilmoitettu, että teos annetaan käyttöön korvauksetta. Jotkut lehdet ovat antaneet oikeuden vain aineiston lukemiseen näytöltä, jolloin tulostekorvauksia ei luonnollisestikaan muodostu.

Korvausten jakoperusteeksi ELEKTRAssa ovat tieteellisten artikkeleiden tekijöitä ja kustantajia edustavat järjestöt (Suomen tiedekustantajien liitto ry, Suomen tietokirjailijat ry ja Suomen tiedetoimittajien liitto) suositelleet tekijänoikeusmaksujen jakamista puoliksi kustantajille ja tekijöille. Kokeilukirjastolle korvauksesta jää 0,30 markkaa sivulta. Asiakkaalta veloitettava suositusten mukainen sivukorvaus on korkeimmillaan 2,30 markkaa ja alimmillaan 0,30 markkaa, mikäli tekijä ei halua korvausta lainkaan.

ELEKTRAn käytön raportointia ja korvausten laskutusta ja tilitystä varten on kehitetty erillinen tietojärjestelmä. Sen avulla voidaan tulostaa sopimusten edellyttämät käytön seuranta- ja tilitysraportit oikeudenhaltijoille. Raportointi tehdään teoskohtaisesti ja se sisältää tiedot julkaisujen katselukerroista ja tulosteista.

ELEKTRAn julkaisupalvelun kustannukset samoin kuin arkistoinnista aiheutuneet kulut ovat hankevaiheessa Helsingin yliopiston kirjaston vastuulla ja ne on pääosa voitu kustantaa OPM:n avustuksella. Jatkovaa julkaisupalvelua suunniteltaessa kustannukset on kuitenkin voitava siirtää palvelun käyttäjille, eli kustantajille. Aineiston digitointia voidaan tarjota maksullisena palveluna ja pääosa aineiston arkistoinnista voitaneen tehdä elektronisten vapaakappalepalvelujen yhteydessä. Digitointikustannuksia voidaan myös tuntuvasti alentaa rationalisoimalla työtä yhdessä julkaisijoiden kanssa.

Tekijänoikeuksien hallinnointi, oikeuksien hankinnasta hinnoitteluun ja korvausten jakoon, on ollut tärkeä osa projektia ja sitonut merkittävästi resursseja. On ilmeistä, että se tulee myös jatkossa olemaan eräs painopistealue. Sopimusmekanismien kehittyminen ja oikeuksien hallinnointitapojen vakiintuminen tulevat pidemmällä tähtäimellä pienentämään sen osuutta. Digitaalioikeuksien hankinta on todettu varsin työlääksi myös vastaavissa ulkomaisissa hankkeissa, kuten englantilaisessa eLibissä, missä on päädytty suosittelemaan valtakunnallista digitointikeskusta, joka huolehtisi keskitetysti myös tekijänoikeuksien hankinnasta (<http://ukoln.bath.uk/elib/>).

2.6 ELEKTRAN yhteydet muihin hankkeisiin

Yhtenä ELEKTRAN tavoitteena oli luoda yhteyksiä vastaaviin koti- ja ulkomaisiin hankkeisiin. Verkostoituminen on onnistunut yli odotusten, ja ELEKTRA on avannut partnereilleen mahdollisuuksia osallistua useisiin muihinkin hankkeisiin. ELEKTRA on päässyt testihankkeeksi EU:n kirjasto-ohjelman tekijänoikeuksia selvittävään Tecup-hankkeeseen, jossa ELEKTRAN partnereista ovat mukana HYK ja Kopiosto. ELEKTRAssa luotu osaaminen on avannut HYK:ille ovet myös saman ohjelman Dieper-hankkeeseen. Sen tavoitteina on digitoida eurooppalaisia tieteellisiä artikkeleita ja tarjota yhteinen pääsy tähän tietovarantoon. Euroopan neuvoston Adapt-ohjelmaan liittyvässä Uusi kirjatalous -hankkeessa on mukana useita ELEKTRAN partnereita asiantuntijatahoina.

ELEKTRAssa on kehittynyt avainhanke, joka on synnyttänyt ympärilleen myös uusia hankkeita, joiden kanssa on luotu luonteva työnjako ja hankkeet tukevat toisiaan. Elektronisen aineiston arkistointiproblematiikka edellytti erillistä hanketta, jota varten perustettiin EVA, elektronisen verkkoaineiston hankinta ja arkistointiprojekti. Siinä ratkaistaan myös ELEKTRAN osalta arkistointikysymykset samalla kun ELEKTRA tuottaa EVAa varten tarvittavaa testiaineistoa. Helsingin yliopisto on käynnistänyt vuoden 1998 alussa oman Elektroninen julkaisupankki -hankkeensa, jonka tavoitteena on siirtää merkittävä osa Helsingin yliopiston julkaisutuotannosta elektroniseen muotoon aina laitossarjoista tutkielmiin ja väitöskirjoihin. Hankkeen koordinointi on HYK:in tehtävänä, ja ELEKTRAN kokemuksia hyödynnetään tässä vaativassa hankkeessa. Julkaisupankin aineistosta se osa, mikä edellyttää kontrolloitua jakelua, ohjataan julkaistavaksi ELEKTRAssa.

Opetusministeriön käynnistämän Kansallisen elektronisen kirjaston tehtävänä on tarjota yhteinen pääsy elektronisiin aineistoihin ja edesauttaa tutkimuksen ja tuotantoelämän käyttöön tarvittavan aineiston hankintaa kansallisin yhteislisenssein. ELEKTRA voi osaltaan edistää kotimaisen tieteellisen verkkoaineiston tuotantoa ja kehittää aineiston jakelua varten vaihtoehtoisen kontrolloidun kanavan, jonka periaate sisältyy myös Kansallisen elektronisen kirjaston ehdotukseen.

2.7 Yhteenveto tuloksista

No	Tavoite	Tulokset
1	1.1 Toteuttaa julkaisujen haku- ja välityspalvelu	<ul style="list-style-type: none"> Selvitys tiedostomuodoista-, ohjelmistoista ja laitteista Selvitys valmiista verkkoaineistoista ja julkaisijoiden valmiuksista

		<ul style="list-style-type: none"> • Digitointi- ja elektronisen julkaisemisen palveluympäristö • Julkaisujen haku- ja välityspalvelu • Julkaisijakontaktit ja aineiston hankinta • Artikkel- ja monografiapilotti, aineiston digitointi • <i>Elektronisen</i> aineiston Marc-kuvailun sovellusohjeet • Elektronisen aineiston arkisto • Käyttökokeilu 7 sopimuskirjastossa
	1.2 Evaluoida tuloksia	<ul style="list-style-type: none"> • Käyttäjäkysely ja raportointi • Lokiseurantatietojen keräys ja raportointi
2	2.1 Luoda malli yliopistojen elektronista oppimateriaalipalvelua varten	
	2.2 Luoda malli yliopistojen elektronista väitöskirjapalvelua varten	<ul style="list-style-type: none"> • Selvitys tiedostomuodoista • Väitöskirjapilotti
3	3.1 Selvittää tekijänoikeusky-symyksiä	<ul style="list-style-type: none"> • Selvitys ELEKTRAN tekijänoikeusperiaatteista
	3.2 Kehittää sopimusmallit tieteellisten julkaisujen välittämiseen	<ul style="list-style-type: none"> • Kustantajien ja HYK:in välinen sopimus • Tekijöiden ja HYK:in välinen sopimus • Välittäjäkirjastojen ja HYK:in välinen sopimus • Oikeuksien hankinta oikeudenhaltijoilta
4	4.1 Selvittää maksullisuusperiaatteita	<ul style="list-style-type: none"> • Selvitys maksullisuusperiaatteista ja hinnoittelu
	4.2 Selvittää laskutuskäytäntöjä	<ul style="list-style-type: none"> • Selvitys elektronisista maksujär-

		<p>jestelmistä</p> <ul style="list-style-type: none"> • Tietojärjestelmä korvausten perintää, laskutusta ja tuloutusta varten • Raportointi ja korvausten tuloutus oikeudenhaltijoille
	4.3 Selvittää suojauskäytäntöjä	<ul style="list-style-type: none"> • Selvitys tietoturvasta • ELEKTRAN tietoturvaperiaatteet
5	5.1 Luoda yhteyksiä kansainvälisiin vastaaviin hankkeisiin	<ul style="list-style-type: none"> • Selvitys vastaavista ulkomaisista elektronisen julkaisemisen hankkeista • Partneri EU:n TECUP-hankkeessa • Partneri EU:n DIEPER-hankkeessa • Yhteyksiä mm. eLib, Decomate, Imprimatur
	5.2. Verkottuminen kansallisiin hankkeisiin	<ul style="list-style-type: none"> • EVA, testiaineistoa • Uusi kirjatalous, tekijänoikeudet, julkaisuohjeet • HY Elektroninen julkaisupankki, välityskanava
	5.3 Tiedotus	<ul style="list-style-type: none"> • Osaraportit, loppuraportti, WWW-sivut, esitteet, messut, seminaarit, lehdistötiedote, asiantuntijaseminaari

2.8 Suositukset ELEKTRAN jatkohanketta varten

1. ELEKTRA on julkaisukanavana kiinnostava tekijöiden ja julkaisijoiden kannalta ja merkittävä osa tiedekustantajista on saatu mukaan kokeiluun. Julkaisijakontakteja tulee vahvistaa sopimuksin ja yhteistyötä kehittämällä.
2. Mahdollisimman suuri osa aineistosta tulee pyrkiä saamaan palveluun valmiiksi elektronisessa muodossa, jotta tuotantokustannukset saadaan hallittua.
3. Julkaisijoita ja tekijöitä varten tulee laatia julkaisuohjeet ja työkalut dokumenttiin itseensä sisällytettäviä kuvailuja ja tarvittavia tyylimäärityksiä varten. Näin parannetaan dokumenttien laatua.
4. Palveluun tulee integroida tarvepainatus.

5. ELEKTRAn sopimusmekanismi on ollut vakuuttava kustantajien ja tekijöiden kannalta, mutta sitä tulee kehittää edelleen julkaisujen avoimempaa välitystä varten.
6. ELEKTRAn jakelupalvelu nykyisessä rajoitetussa muodossaan ei ole riittävästi kiinnostanut käyttäjiä, ja jakelua tulee laajentaa sopimuskirjastoista rajoitettuun verkkokäyttöön organisaatiolisenssein. Myös muita jakeluvaihtoehtoja tulee selvittää.
7. ELEKTRA on profiloitunut tekijänoikeudellisen, etupäässä maksullisen aineiston dokumentinvälityspalveluksi, ja sen suhteet vapaan verkkoaineiston välitykseen väitöskirjapalvelua varten tulee tarkentaa.
8. Aineiston maksullisuus ei käyttäjäpalautteen mukaan ole ongelma, mutta tulostuksen vähäisyys osoittanee sen, että hintatasoa täytyy edelleen harkita.
9. Automaattista käytön laskutus-, seuranta- ja raportointiohjelmistoa tulee kehittää edelleen jo rakennetun ohjelmiston pohjalta.
10. Käyttäjäpalvelusta tulee tiedottaa enemmän ja kouluttaa tarpeen mukaan käyttäjiä.
11. Vastaavaa kansainvälistä kehitystä tulee edelleen seurata sekä osallistumalla muihin yhteistyöhankkeisiin että solmimalla aktiivisesti kontakteja uusien hankkeiden kanssa.
12. Elektronisen identiteetin ja automaattisten maksujärjestelmien kehitystä tulee seurata ja osallistua kotimaisiin kokeiluhankkeisiin heti kun se on mahdollista.
13. Elektronisen aineiston pitkäaikaista tallennusta ja siihen liittyvät tekniset ratkaisut tulee selvittää myös väitöskirjojen osalta elektronisen aineiston hankinta- ja arkistointihankkeessa Evassa ja aikanaan osana uudistuvan vapaakappalelain määrittämää peruspalvelua.

2.9 ELEKTRAn jatkoahanke

ELEKTRA-projekti on kunnianhimoinen hanke, joka on ottanut selvittääkseen koko elektronisen julkaisu- ja jakeluprosessin ja piloteissaan testannut palvelun teknistä ja taloudellista perustaa. Uuden palvelun luominen edellyttää kuitenkin useiden vuosien kehittämistyötä ja myös omia investointeja. Opetusministeriön myöntämä tuki vajaan kahden vuoden ajaksi ei ole riittävä varsinaisen tuotantosovelluksen aikaansaamiseksi, vaan konsortiopohjaista kehittämistyötä on edelleen tarpeen jatkaa. Tästä syystä on suunniteltu kahden vuoden mittainen ELEKTRA II-projekti, jota varten on anottu jatkorahoitusta Opetusministeriöltä.

ELEKTRA II tavoitteina on

- toteuttaa aineiston elektroninen julkaisupalvelu pyrkien kustannusvastaavuuteen
- laajentaa elektronisen aineiston sopimuskäytäntöjä ja kehittää malli organisaatiolisenssiksi
- laajentaa haku- ja välityspalvelua käytettäväksi sopimuskirjastojen lisäksi organisaatiolisenssein
- laajentaa väitöskirjapalvelua kattamaan sekä maksullista että vapaasti verkossa levitettävää aineistoa ja selvittää mallin soveltuvuutta myös oppimateriaalin välitykseen
- testata automaattista maksujärjestelmää
- selvittää artikkeliaineiston kokotekstitietokantaratkaisuja
- integroida palveluun tarvepainatus

ELEKTRA II:n aineiston jakelu toteutetaan Internetin välityksellä yleiskäyttöisillä WWW-selaimilla. Nykyinen Hyperlib-käyttöliittymäohjelmisto pyritään korvaamaan kehittyneemmällä tuotteella ja käytön seuranta- laskutusohjelma päivitetään. Dokumenttien kuvailu tehdään Marc-muodossa erilliseen hakutietokantaan ja jakeluformaatti valitaan niin, että tekijäsopimusten

ehdot voidaan täyttää. Vapaaseen verkkolevitykseen tarkoitettun HTML-muotoinen aineisto kuvaillaan dokumentteihin sisällytettävän metadatan avulla ja EVA:n kehittämien työkalujen avulla. ELEKTRAssa on testattu erilaisten tiedostomuotojen tuottamista ja käyttöä. Oulun väitöskirjapilotissa on keskitytty SGML-, HTML- ja PDF-formaatteihin ja niiden edellyttämien työkalujen käyttökokeiluun. HYK:in pilotissa aineisto on saatu julkaisijoilta lukuisissa erilaisissa tiedostomuodoissa, ja jakelumuotoina on käytetty joko PDF- tai HTML-formaatteja aineiston lähdetiedostosta ja sopimusehdoista riippuen. Tiedostomuotokokeiluja jatketaan edelleen ELEKTRAN II-vaiheessa. Tarkoituksena on saada aikaan malli ja julkaisuohjeet yliopistoissa tapahtuvaa monimuotoista elektronista julkaisemista varten.

ELEKTRAN I-vaiheessa luovuttiin suunnitellusta oppikirjakokeilusta, koska riittävän edustavaa oppimateriaaliaineistoa ei ollut tarjolla. Oppimateriaalikokeilun suunnittelua jatketaan. Yksittäisiä kustantajia ei enää oteta mukaan projektin konsortioon, vaan aineistoa otetaan mukaan eri kustantajilta. Väitöskirjakokeilun yhteydessä tarvepainatus on osoittautunut välttämättömäksi ja sen integrointi verkossa toimivaan jakelupalveluun toteutetaan. Tarvepainatuskokeilusta hankkeessa vastaa Yliopistopaino.

ELEKTRAN käyttökokeilussa laskutus on hoidettu manuaalisesti, koska tarvittavia automaattisia laskutusohjelmia ei ole ollut käytettävissä. Sellaiset kaupalliset palvelut kuin Avant tai EUNET Ecash eivät sovellu pienten summien laskuttamiseen. Myöskään EU-hankkeet kuten Copinet eivät ole saaneet markkinoille tuotteita, joita olisi voitu soveltaa ELEKTRAssa. Kotimaisen verkkorahan kehittäminen on jo kuitenkin pitkällä. Helsingin yliopistossa on aloitettu sellaisen toimikortin käyttökokeilu, joka sopii myös elektroniseen rahastukseen. Uutena EU-hankkeena vuoden alussa käynnistynyt Candle tähtää myös yleiskäyttöisen ja edullisen maksusovelluksen luomiseen ja sen tuloksia seurataan. ELEKTRAN jatkohankkeessa on tarkoitus testata tarkoitukseen soveltuvaa elektronista maksujärjestelmää.

Markkinoilla on myös paljon erilaisia elektronisia palveluympäristöjä, kuten esimerkiksi Xeroxin Xdod/DocuWEB, ICL:n Commands ja IBM:n Digital Library, jotka tarjoavat välineet sekä julkaisemiseen että jakeluun. Osa järjestelmistä perustuu kokotekstitietokantaan, josta haku voidaan tehdä dokumentin sisällön perusteella ilman suuritöistä kuvailutyötä. Monet järjestelmät tarjoavat myös hyvät linkitysmahdollisuudet dokumenttien sisällä ja välillä. Tämä mahdollisuus on luonnollisesti käytettävissä myös vapaaseen verkkojakeluun asetetuissa HTML-dokumenteissa. Hankkeessa kartoitetaan näitä tuotteita ja analysoidaan niiden palveluja ja kustannuksia.

Julkaisujen välityskokeilun laajentaminen toteutettavaksi yliopistojen paikallisverkossa edellyttää uuden sopimuskäytännön luomista. Tarvitaan organisaatiolisenssejä, joilla yliopistokirjastot voivat tarjota aineiston lukuoikeuden paikallisverkoissaan. Tulostusoikeus rajataan edelleen vain kirjastoihin. Myös oppimateriaalin sisällyttäminen kokeiluun edellyttää nykyisiä laajempia jakelumahdollisuuksia, jotka voitaisiin toteuttaa uudella sopimusmallilla.

2.10 ELEKTRA II:n tuottamat tulokset ja niistä saatavat hyödyt

Tulokset

- kontrolloitu dokumentinvälityspalvelu tekijänoikeuksien alaista tieteellistä aineistoa varten
- laaja elektronisessa muodossa oleva sisältöaineisto, joka kattaa artikkeleja, väitöskirjoja ja oppimateriaalia
- tarvepainatuspalvelu

- joustavat tekijänoikeussopimukset, jotka mahdollistavat dokumenttien vapaan tai kontrolloidun välityksen
- helppo käyttöliittymä
- vakiintuneet suhteet julkaisijoihin aineiston saamiseksi pääosin suoraan elektronisessa muodossa
- elektronisen maksujärjestelmän kokeilu
- selvitys tekijänoikeuksien hallintajärjestelmistä (CMS)
- selvitys kokotekstitietokantaratkaisuista
- käyttäjäpalaute
- loppuraportti

Projektin avulla saavutettavat hyödyt

- nopeuttaa kotimaisen tieteellisen sisältötuotannon luomista verkkoon
- mahdollistaa tieteellisen aineiston julkaisemisen joko rinnakkaisversioina tai ilman painettua vastinetta
- tarjoaa valtakunnallinen kanava tekijänoikeuden alaisen aineiston kontrolloituun dokumentinvälitykseen
- tarjoaa helpon pääsyn ja monipuoliset tiedonhakupmahdollisuudet palveluun integroituun aineistoon
- parantaa osallistujien kansainvälistä näkyvyyttä ja luo eurooppalaisia yhteistyösuhteita
- tukee tuloksillaan lukuisia muita projekteja, kuten EVA, Helsingin yliopiston elektroninen julkaisupankki, Tecup ja Dieper

3. KANSAINVÄLISIÄ ELEKTRAN KANNALTA KIINNOSTAVIA HANKKEITA

3.1 Johdanto

Tässä luvussa esitellään projekteja, joissa tutkitaan samantyyppisiä kysymyksiä kuin ELEKTRA-projektissa. Tällaisia teemoja ovat dokumenttien välittäminen elektronisessa muodossa tietoverkkojen (Internet) välityksellä, elektronisten (full-text) dokumenttien liittäminen näyttöluetteloiden viitetietueisiin ja elektronisten dokumenttien välitykseen liittyvät tekijänoikeuskysymykset ja -korvaukset.

Alla olevaan listaan on pyritty valikoimaan projekteja, jotka ovat jo jonkin aikaa olleet käynnissä ja joissa siten on käytännön kokemuksia jo saatu. Vasta alkamassa olevat projektit on yleensä kuvattu vain lyhyesti, jos niistä ei vielä tällä hetkellä ole muuta tietoa kuin projektisuunnitelma. Samaten on mainittu lyhyesti sellaisiakin projekteja, joissa on kiinnitetty huomiota vain yhteen yllämainituista teemoista, jos projekti muuten on kiinnostava.

Toisaalta listasta on jätetty pois hankkeita, jotka sinänsä ovat kiinnostavia, mutta eivät enää ELEKTRA-projektin kannalta ajankohtaisia. Esimerkiksi TULIP-projektissa välitettiin dokumentteja elektronisessa muodossa, mutta skannattuina kuvatiedostoina, joita katseltiin projektia varten suunnitelluilla selaimilla. (<http://www.elsevier.nl:80/homepage/about/resproj/tulip.shtml>).

Nykyisissä projekteissa ei enää suunnitella omia erillisiä selainohjelmia, vaan pyritään yleiskäyttöisyyteen: dokumenttien halutaan olevan selailtavissa ja poimittavissa verkosta samoilla työvälineillä, joita käyttäjät muutenkin tarvitsevat (tyypillisimmillään WWW-selainohjelma). Kun suuret kansainväliset kustantajat ovat uudistaneet tuotantoprosessiaan ja niiden julkaisut tuotetaan valmiiksi elektronisessa muodossa kuvia myöten, skannattujen kuvatiedostojen käyttö projektien testiaineistona on vähentymään päin ja valmiista tiedostoista tuotettavien tallennus- ja jakelumuotojen (esimerkiksi Adobe Acrobat eli PDF-tiedostomuoto) suosio lisääntymässä.

3.2 WebDOC

WebDOC-projektissa tutkitaan, kuinka elektronisessa muodossa olevat dokumentit voidaan asettaa tiedontarvitsijoiden ulottuville Ww:n välityksellä. Projektin ytimenä on keskitetty kirjastoluettelo WebCAT, johon elektronisessa muodossa saatavilla olevat dokumentit luetteloidaan. Tätä kirjastoluetteloa ylläpitää hollantilainen kirjastojen atk-palvelukeskus Pica. Projekti alkoi helmikuussa 1995 ja sen pilottivaihe kestää noin kaksi vuotta.

Projektiin osallistuu Pican lisäksi joukko korkeakoulukirjastoja ja pari tieteellistä kustantajaa. Korkeakoulukirjastot tarjoavat projektin käyttöön korkeakoulujen omia julkaisuja kuten väitöskirjoja ja tutkimusraportteja, kaupalliset kustantajat taas tieteellisten lehtien artikkeleita elektronisessa muodossa. Projektiin osallistuvat kirjastot luettelivat aineistonsa itse keskitettyyn kirjastoluetteloon Pican käyttämässä luettelointimuodossa.

Toisin kuin monessa muussa projektissa (esimerkiksi alempana esitettävissä COPINET- ja ELSA-projekteissa), dokumentteja ei tallenneta yhteen keskitettyyn ja paikalliseen järjestelmään, vaan kukin projektin osapuolista rakentaa ja ylläpitää itse omaa dokumenttitietokantaan-

sa, jota WebDOC-järjestelmä hyödyntää. Kokeilun alkuvaiheessa käytetyt tiedostomuodot olivat HTML, PDF, TIFF ja PostScript.

Tiedontarvitsija voi käyttää keskitettyä kirjastoluetteloa WWW-selaimellaan. Kun käyttäjä löytää luettelosta kiinnostavan dokumentin viitetietueen, hän painaa tilauspainiketta. Käyttäjä ei kuitenkaan saa dokumenttia heti nähtävilleen, vaan ensin järjestelmä tarkistaa dokumentin käyttöoikeuden ja tilanteesta riippuen toimii seuraavasti:

Dokumentti on vapaasti käytettävissä, joten tilauspyyntö lähetetään kyseiseen dokumenttietokantaan ja dokumentti toimitetaan asiakkaalle

Dokumentin käytöstä on tehty lisenssisopimus kyseisen tilaajan oman korkeakoulukirjaston ja kustantajan välillä ja se on vapaasti kyseisen korkeakoulun opiskelijoiden ja henkilökunnan käytettävissä; tilauspyyntö lähetetään kyseiseen dokumenttietokantaan ja dokumentti toimitetaan asiakkaalle

Dokumentti on suojattu, joten tilaajan on syötettävä järjestelmään kaukolainatilinsä koodi (PIN-koodi-menetelmä, jolla hollantilaiset yksityishenkilöt voivat hankkia kirjastosta erityisen tilin, jolta kaukolainatilaukset veloitetaan). Kun tililtä on veloitettu tarvittava summa, tilaajalle lähetetään avain, jolla hän saa haettua dokumentin tietokannasta.

Lisätietoja: <http://www.pica.nl/docs/en/webdoc/webproj.html>

3.3 EU:n rahoittamat hankkeet

DECOMATE - Delivery of Copyrighted Materials to End-users

DECOMATE on EU:n rahoittama projekti (Telematics for Libraries 3078), jossa välitetään tekijänoikeussuojan alaisia elektronisia julkaisuja tiedontarvitsijoille, tässä tapauksessa yliopistojen henkilökunnalle ja opiskelijoille. Projekti alkoi maaliskuussa 1995 ja kestää kaksi vuotta.

Projektissa mukana olevat kolme yliopistoa saavat yli sata tieteellistä lehteä elektronisessa muodossa käyttöönsä suoraan niiden kustantajalta Elsevieriltä. Julkaisu- ja artikkelitietokantaa käytetään WWW-selaimen avulla. Dokumentit on tallennettu Adobe Acrobat- eli PDF-muodossa.

DECOMATE-järjestelmä sisältää ohjelmistomodulin, jolla määritellään käyttäjät ja näiden käyttöoikeudet sekä seurataan järjestelmän toimintoja siten, että käytönseurantatietojen perusteella voidaan laskea tekijänoikeuskorvaukset.

Lisätietoja: <http://www.lse.ac.uk/decomate/>

COPINET - Billing System for Open-Access Networked Information Resources

COPINET on EU:n kirjasto-ohjelman rahoittama projekti (Telematics for Libraries 3-3033), jossa tutkitaan elektronisten julkaisujen hakua, välittämistä ja laskutusta Internetin välityksellä. Erityisesti tutkitaan sitä, miten palveluja voidaan tarjota myös satunnaisille organisaation ulkopuolisille käyttäjille, jotka maksavat vain haluamistaan julkaisuista. Projekti alkoi helmikuussa 1995 ja kestää kaksi vuotta.

Projektin testitietokantana on 100 000 viitteen otos IEE:n INSPEC-tietokannasta, jonka 10 000 viitetietueeseen on linkitetty myös itse julkaisut. Tietokantaan on rakennettu WWW-käyttöliittymä, joten tiedontarvitsija voi hakea sitä normaalilla WWW-selaimellaan. Jos viitteen on linkitetty artikkeli, käyttäjä voi saada myös itse artikkelin nähtävilleen. Artikkelien tilaus on maksullista (joskaan todellisia rahalaskuja ei vielä testivaiheessa lähetetä). Laskutukseen on useita vaihtoehtoja: rekisteröityjen tiliasiakkaiden tiliä veloitetaan tai heille lähetetään lasku, rekisteröimättömät käyttäjät voivat maksaa luottokortilla tai DigiCash-rahakkeilla. Artikkelit on tallennettu tietokantaan kuvamuodossa, jolloin käyttäjä voi halutessaan tilata ja maksaa vain haluamansa sivut - koko artikkelin tilaus kerralla tulee kuitenkin halvemmaksi kuin sivujen ostaminen yksittäin.

Lisätietoja: <http://www.mari.co.uk/copinet/>

ELSA - Electronic Library SGML Applications

EU:n rahoittamassa (Telematics for Libraries 2006) ELSA-projektissa tutkitaan tieteellisten lehtien artikkeleiden tallennusta ja hakua elektronisessa muodossa. Artikkelien teksti on koodattu SGML-muotoon ja ne sekä niihin liittyvät kuvat tallennetaan julkaisutietokantaan.

Testiaineistona on Elsevierin noin sadasta lehdestä kokoama 5000 artikkelin kokoelma. Jouve S. A. on projektin koordinaattori ja kehittää projektissa käytettävät asiakas/palvelin-ohjelmat ja De Montfortin yliopisto käyttäjien tarvitsemat käyttöliittymät; lisäksi De Montfort huolehtii käyttäjätestauksista.

Lisätietoja: <http://elsa.dmu.ac.uk/>

IMPRIMATUR - Intellectual Multimedia Property Rights Model and Terminology for Universal Reference

IMPRIMATUR (alkuaan latinaa, 'painolupa') on EU:n ESPRIT-ohjelman projekti 20676. (ESPRIT-ohjelmaa koordinoi Euroopan komission direktoraatti DG III). Projekti alkoi 1996 ja kestää vuoteen 1998.

Projektin tarkoituksena on suunnitella ja suositella menetelmiä, joiden avulla on mahdollista myydä erilaisia tekijänoikeuksin suojattuja elektronisia julkaisuja, multimediatuotteita ym. siten, että sekä tekijöiden että käyttäjien edut ja oikeudet otetaan huomioon. Siinä on tarkoitus rakentaa Internet-palvelimen prototyyppi, jossa kaupankäyntiin liittyvien tekijänoikeuskorvausten hallinnointi otetaan huomioon.

Lisätietoja: <http://www.imprimatur.alcs.co.uk/>

3.4 Englannin eLib-hankkeet

Monien eLib-projektien laajuus ja niiden kohderyhmä ovat hyvin samanlaiset kuin ELEKTRA-projektissa. eLib-hankkeissa on myös tyypillistä, että samaan aihealueeseen voi liittyä useita, osin päällekkäisiä projekteja (<http://www.ukoln.ac.uk/elib.html>).

eLib-hankkeiden perustana on Follettin työryhmän raportti jossa esitettiin suosituksia siitä, miten kirjastoista tulisi tietotekniikan avulla kehittää elektronisia tietovarantoja (http://ukoln.bath.ac.uk/follett/follett_report.html). Tällaiset tietovarannot voisivat tietoverkkojen kautta olla yleisesti mahdollisimman monen korkeakoulun käytettävissä ja siten säästäisivät resursseja. eLib-hankkeita rahoittaa Joint Information Systems Committee (JISC; <http://www.niss.ac.uk/education/jasper/intro.html>).

InfoBike

InfoBike-projektissa tiedontarvitsija voi hakea kiinnostavia artikkeleita viitetietokannoista ja saada löytämänsä artikkelit elektronisessa muodossa omalle työasemalleen. Projektissa myös testataan, miten tällainen elektronisen dokumentin hakeminen, tilaaminen, toimittaminen ja selailu käytännössä toimii. Järjestelmän testauksen on tarkoitus alkaa tammikuussa 1997.

Järjestelmän osat ovat seuraavat:

- dokumenttien tilaus InfoBike Gatewayn kautta (tai BIDSin kautta)
- WWW-käyttöliittymä kustantajien julkaisuluetteloihin ja dokumenttien jakelu WWW-selaimen avulla (kustantajien luettelot sisältävät julkaisujen metadatakuvaukset eli julkaisujen bibliografiset tiedot)
- elektronisten lehtien tietokanta; aluksi artikkelien tallennusmuotona on Adobe Acrobat PDF-tiedostomuoto, mutta projektin myöhemmässä vaiheessa mukaan on tarkoitus ottaa myös muita tiedostomuotoja (multimediajulkaisuja); elektroniset lehdet on tallennettu joko paikalliselle palvelimelle tai muualla verkossa sijaitseville palvelimille
- tapahtumaseuranta, jonka avulla voidaan hallita lehtien käyttölisenssejä ja maksujärjestelmiä; käyttäjä voi maksaa kustannukset esimerkiksi luottokortilla, mutta on myös mahdollista, että kustantajat ja laitokset tekevät lisenssisopimuksen, jonka mukaan järjestelmän käyttö on ko. laitosten henkilökunnalle ja opiskelijoille ilmaista (jolloin taas järjestelmän on osattava tunnistaa tällaiset rekisteröidyt käyttäjät).

InfoBike-järjestelmän on tarkoitus olla avoin siinä mielessä, että siinä voidaan sallia monia erilaisia tiedostomuotoja, mukaanlukien erilaiset multimediatiedostot (kuva-, ääni- ja videotiedostot, SGML, perus-ASCII jne.). Samaten tiedostot voivat olla jo alunperin elektroniseen muotoon tuotettujen julkaisujen lisäksi myös kuvatiedostoja, jotka on tuotettu skannaamalla alkuperäinen paperimuotoinen julkaisu kuvanlukijalla. Testiaineiston projektin käyttöön antavat Academic Press ja Blackwell Science.

Lisätietoja: <http://www.bids.ac.uk/elib/infobike/homepage.html>

SCOPE - Scottish Collaborative On-demand Publishing Enterprise

SCOPE-projektissa välitetään kurssimateriaalia Internetin välityksellä skotlantilaisten yliopistojen opiskelijoille. Alunperin kurssimateriaali välitettiin valokopioina, mutta projektin painopiste on siirtymässä elektronisiin julkaisuihin. Elektronisten julkaisujen tallennusmuotona on pääosin Adoben Acrobat- eli PDF-muoto. Projekti alkoi toukokuussa 1995 ja kestää kolme vuotta.

Olenainen osa projektia on tekijänoikeusmaksujärjestelmän rakentaminen (Cactus, SCOPE Electronic Copyright Management System). Cactus-järjestelmän avulla on mahdollista kontrol-

luida käyttöoikeuksia, järjestää tapahtumaseuranta, laskea tekijänoikeusmaksut sekä välittää sallitut dokumentit rekisteröidyille käyttäjille. SCOPE-järjestelmän on tarkoitus olla mahdollisimman yleinen niin, että siihen voidaan ottaa yhteys erilaisilla laitteisto- ja ohjelmistokokoonpanoilla.

Lisätietoja: <http://www.stir.ac.uk/infoserv/scope/>

ACORN - Access to Course Readings via Networks

ACORN-projektissa kehitetään järjestelmä, jossa opiskelijat saavat helposti kurssikirjallisuutensa sisältyvistä artikkeleista kopion elektronisessa muodossa. Projekti alkoi elokuussa 1996 ja päättyi tammikuussa 1998. Olennainen osa ACORN-järjestelmää on tapahtumaseuranta, johon tekijänoikeusmaksujen maksatus perustuu. Projektissa oleva kaupallinen dokumenttien välitykseen erikoistunut yritys, Swets, huolehtii tekijänoikeussopimuksien laatimisesta ja royaltien maksamisesta.

Järjestelmä muodostuu erillisestä dokumenttitietokannasta, joka on linkitetty kirjaston näyttöluetteloon. Näyttöluetteloa on mahdollista käyttää myös WWW:n kautta (TALIS Web OPAC). Varsinainen dokumenttitietokanta on tallennettu Sun-palvelimelle, jossa dokumentteja hallinnoidaan Sybase-relaatiotietokannan avulla. Dokumentit (artikkelit) on tarkoitus tallentaa tietokantaan PDF-muodossa. Järjestelmän käyttöoikeus on vain rekisteröidyillä käyttäjillä. Projektiin liittyy myös käyttäjätutkimus, jossa selvitetään opiskelijoiden näkemyksiä elektronisen aineiston ja järjestelmän käytettävyydestä.

Lisätietoja: <http://liba.lboro.ac.uk/acorn/acorn.htm>; Kingston, P. 1996. Project Acorn: Using Internet tools to provide local solutions. *Managing Information*, vol. 3, no. 10, pp. 42 - 44.

EDDIS - Electronic Document Delivery: the Integrated Solution

EDDIS-hankeessa rakennetaan elektronisten dokumenttien paikantamis- ja välitysjärjestelmän pilottiversio, jota tiedontarvitsijat (loppukäyttäjät) voivat itse käyttää. Järjestelmän avulla kuka tahansa Englannin korkeakoulun opiskelija tai tutkija voi etsiä häntä kiinnostavia julkaisuja, saada tietoja niiden saatavuudesta, ja tilata haluamiaan julkaisuja järjestelmän välityksellä. Tyyppillisesti kyseessä ovat julkaisut, joista asiakkaalle toimitetaan kopio (non-returnable items). EDDIS-järjestelmän pilottiversiota on myös tarkoitus kokeilla käytännössä.

EDDIS-järjestelmän periaatteena on, että tiedontarvitsija hakee julkaisutietoja bibliografisista tietojärjestelmistä. Jos kiinnostava julkaisu löytyy, eikä sitä ole paikallisessa kirjastossa, käyttäjä voisi lähettää tilauksen johonkin välityspalveluun (kuten British Library Document Supply Centre) ja määritellä, missä muodossa julkaisu hänelle toimitetaan. Jos kyseessä on artikkeli, se voidaan lähettää suoraan tiedontarvitsijan omalle työasemalle elektronisessa muodossa.

Järjestelmään integroitu tilausten hallinta ja laskutus tapahtuu normaalisti loppukäyttäjältä piilossa, kirjaston hallinnoimana. Laskutusjärjestelmä pystyy käsittelemään kaikki tavanomaiset dokumenttien maksutavat. Myös tekijänoikeusmaksut (copyright clearance fees) on otettu huomioon.

Lisätietoja: <http://www.ukoln.ac.uk/elib/flyers/eddis.html>
<http://www.lancs.ac.uk/users/library/news.d/eddis.htm>

ERIMS - Electronic Readings In Management Studies

ERIMS-projektin tavoitteet ovat monessa suhteessa samanlaiset kuin ELEKTRA-projektin, paitsi että ERIMS-hankkeessa dokumentit eivät ole varsinaisesti elektronisia julkaisuja, vaan perinteisiä paperijulkaisuja digitoituina. Paperijulkaisu on skannattu ja sivut on tallennettu tietokantaan kuvamuodossa.

Projektissa luodaan asiakas-palvelin-järjestelmä, jonka avulla käyttäjät (opiskelijat, korkeakoulujen ja kirjastojen henkilökunta) löytävät helposti haluamaansa aineistoa, voivat selailta sitä ruudulla ja tarvittaessa tulostaa. ERIMS-järjestelmä on erillinen, vain rekisteröityjen käyttäjien käytettävissä oleva järjestelmä (käyttöön tarvitaan käyttäjätunnus ja salasana). Tutkimustietokantaan on tarkoitus tallentaa noin 20 000 sivua (20 gigatavua). Projektissa selvitetään myös vaihtoehtoisia jakelutapoja, kuten CD-ROM- tai verkkojakelua.

Keskeinen osa projektia on ollut **käyttöoikeuksista sopiminen** (käyttölisenssi) kustantajien kanssa. Tekijänoikeuskysymykset otetaan huomioon siten, että järjestelmä tilastoi, mitä julkaisuja on katseltu ja tulostettu. Näiden tilastojen perusteella käyttäjiä voidaan laskuttaa ja tekijänoikeuksien omistajille (kirjoittajat ja kustantajat) maksaa käyttöön perustuva korvaus. Myöhemmin järjestelmää on tarkoitus laajentaa niin, että laskutus on mahdollista tehdä automaattisesti.

ERIMS-järjestelmä koostuu seuraavista osista:

- kuvatietueista koostuva varsinainen julkaisutietokanta; kuvia voidaan selailta ruudulla ja haluttaessa tulostaa paperille
- bibliografiset tiedot, jotka on poimittu korkeakoulujen kurssikirjallisuusluetteloista (nimeke, tekijä, julkaisija, aihealue, tiivistelmä)
- oikeudenomistajat eli tietokantaan tallennettujen julkaisujen kustantajien tai muiden kustannusoikeuden omistavien tahojen yhteystiedot
- dokumenttien copyright-tiedot: kuka kunkin dokumentin oikeudet omistaa ja millaiset ehdot sen käytöstä on annettu (royalit, käyttömaksut)
- järjestelmän hallintatiedot: käyttäjät, salasanat, käyttöoikeudet; käyttötilastot

Lisätietoja: <http://www.templeton.ox.ac.uk/www/college/library/erims/intro.htm>

3.5 British Library Digital Library Project

British Library on käynnistämässä hanketta elektronisten kirjastopalvelujen kehittämiseksi ja tarjoamiseksi. Palvelut suunnitellaan sekä englantilaisia käyttäjiä varten että kansainväliselle käyttäjäkunnalle. Ensimmäinen tarjouskierros on suoritettu. Hankkeen odotetaan käynnistyvän syksyllä 1998. Hankkeen tarkoituksena on mm.

- laajentaa ja parantaa kirjaston artikkelikopioiden välitystä ottamalla käyttöön elektroninen arkisto
- laajentaa patenttipalveluita
- parantaa kirjaston historiallisten kokoelmien käyttömahdollisuuksia

Lisätietoja: <http://www.bl.uk/services/ric/diglib/digilib.html>

3.6 NCSTRL - Networked Computer Science Technical Reports Library

NCSTRL on avoin, tietoverkkojen käyttöön perustuva dokumenttien tallennus- ja jakelujärjestelmä, jossa hajautetusti eri palvelimille tallennettuja dokumentteja (tässä tapauksessa teknisiä raportteja) voidaan hakea ja välittää keskitetysti. Sen pohjana on ns. Dienst-protokolla, joka määrittelee, miten tällaisia dokumentti-tietokantoja voidaan käyttää WWW:n kautta. (<http://www.ncstrl.org/Dienst/htdocs/Info/protocol4.html>)

Dienst sallii dokumenttien olevan periaatteessa missä tahansa tiedostomuodossa. Teknisten raporttien bibliografiset tiedot on indeksoitu omaan tietokantaansa, josta käytetään nimitystä metatietokanta. Tiedontarvitsija voi hakea haluamaansa tietoa metatietokannasta ja saada haun tuloksena saamansa dokumentin nähtävilleen ilman, että hänen tarvitsee tietää, minkä organisaation mille tietokoneelle kyseinen dokumentti on tallennettu. Tekijänoikeusmaksuja ei NCSTRL-järjestelmässä peritä, vaan dokumentit ovat julkisia ja periaatteessa kenen tahansa käytettävissä. Organisaatiot voivat toteuttaa NCSTRL-järjestelmän joko laajana normaaliversiona tai suppeana versiona:

- Normaalitasolla järjestelmä muodostuu WWW-palvelimesta ja siihen liittyvästä toisesta palvelimesta, joka sisältää varsinaisen dokumenttitietokannan; hakuohjelmiston, jolla viite- eli metatietoja haetaan; ja käyttöliittymän, jolla ollaan yhteydessä muuhun (eli muualla sijaitsevaan hajautettuun) NCSTRL-kokoelmaan. Lisäksi dokumentit on mahdollista tallentaa useissa tiedostomuodoissa ja tulostaa paikallisesti paperille.
- Suppealla tasolla järjestelmään tallennetaan dokumentit vain yhdessä tiedostomuodossa (kuten Post-Script) ja järjestelmä on käytännössä vain tiedoston tallennus- ja hakupalvelin (FTP-palvelin). Organisaatio ei indeksoi dokumentteja itse, vaan toimittaa bibliografiset tiedot palvelimelle, joka huolehtii indeksoinnista keskitetysti.

Lisätietoja: <http://www.ncstrl.org/>; <http://cs-tr.cs.cornell.edu/>

Euroopassa NCSTRL-hankkeessa kehitetyn tekniikan on ottanut käyttöön tietotekniikan ja matematiikan tutkimuslaitosten yhteenliittymä ERCIM, European Research Consortium for Informatics and Mathematics. ERCIM aikoo SAMOS-hankkeessa rakentaa digitaalisen kirjaston, joka muodostuu sen jäsenorganisaatioiden dokumenttitietokannoista ja näiden metatietokannasta (The Ercim Digital Library Initiative).

Lisätietoja: <http://www.area.pi.cnr.it/ErcimDL/>

3.7 Yhdysvaltain suuret Digital Libraries -hankkeet

NSF/DARPA/NASA Digital Libraries Initiative Projects

Yhdysvalloissa on käynnissä kuusi suurta projektia, joissa tutkitaan elektronisen aineiston hallintaa ja verkkovälitystä yms. Rahoittajina ovat National Science Foundation (NSF), the Department of Defense Advanced Research Projects Agency (DARPA) ja National Aeronautics and Space Administration (NASA).

Tyypillistä hankkeille on käsiteltävän aineiston suuri määrä. Näiden hankkeiden tavoitteena on löytää ratkaisuja siihen, miten valtavia tietovolyymejä voidaan tallentaa, käsitellä ja asettaa tarjolle digitaalisessa muodossa. Keskeistä on myös, miten haluttu yksittäinen tieto voidaan

löytää ja saada tiedontarvitsijan käyttöön valtavan aineistomäärän joukosta. Hankkeet ja niiden tavoitteet eivät tästä syystä ole suoraan hyödynnettävissä ELEKTRA-projektissa, mutta ne mainitaan tässä, koska ne ovat olleet monien myöhempien projektien esikuvina.

Hankkeeseen sisältyvät projektit ovat:

University of Michigan: Intelligent Agents for Information Location,
<http://www.si.umich.edu/UMDL/HomePage.html>

University of Illinois: "Building the Interspace: Digital Library Infrastructure for a University Engineering Community" — Federating Repositories of Scientific Literature;
<http://dli.grainger.uiuc.edu/default.htm>

University of California, Berkeley: "The Environmental Electronic Library: A Prototype of a Scalable, Intelligent, Distributed Electronic Library"; *<http://elib.cs.berkeley.edu/>*

Carnegie Mellon University: "Informedia: Integrated Speech, Image and Language Understanding for Creation and Exploration of Digital Video Libraries";
<http://www.informedia.cs.cmu.edu/>

Stanford University: Interoperation Mechanisms Among Heterogeneous Services;
<http://walrus.stanford.edu/diglib/>

University of California, Santa Barbara: "The Alexandria Project: Towards a Distributed Digital Library with Comprehensive Services for Images and Spatially Referenced Information" — Spatially-referenced Map Information; *<http://alexandria.sdc.ucsb.edu/>*

Lisätietoja: *<http://dli.grainger.uiuc.edu/national.htm>*;
<http://www.cise.nsf.gov/iris/DLHome.html>; Special Issue on Digital Libraries in the Institute of Electrical and Electronics Engineers, IEEE Computer Magazine:
<http://www.computer.org/pubs/computer/dli/>

3.8 Muita kiinnostavia projekteja

eLib: CATRIONA II, Approaches to University Management of Electronic Resources

Projektissa tutkitaan, miten korkeakoulut voivat laatia ja hallita opetusmateriaaleja yms. tietovarantoja elektronisessa muodossa niin, että ne saadaan kampuksella (ja mahdollisesti muuallakin) helposti käyttäjien ulottuville. Projektissa myös linkitetään opetusmateriaaleja, luentoja ym. aineistoja näyttöluetteloon (Z39.50) niin, että näyttöluettelosta löydetty dokumentti on mahdollista saada saman tien käyttäjän työasemalle.

Lisätietoja: *<http://www.ukoln.ac.uk/elib/flyers/catriona.html>*

eLib: eOn, Interinstitutional networking of learning materials

Projektissa tuotetaan oppimateriaalipaketteja, joista opettaja voi koota haluamansa koosteen (tarvejulkaiseminen). Aineisto on saatavilla verkossa ja lisäksi bibliografiset tiedot luetteloidaan

mukana olevien korkeakoulujen kirjastoluetteloihin. Julkaisujärjestelmässä on tarkoitus ottaa huomioon myös tekijänoikeus- ja käyttömaksut.

Lisätietoja: <http://eon.unn.ac.uk/info.htm>

eLib: Phoenix, On-demand publishing project under the eLib programme

Phoenix-projektissa tutkitaan kurssikirja-aineiston tarvejulkaisemista. Paperimuotoinen kurssikirjallisuus skannataan, muunnetaan tekstiksi ja näin saatu raakateksti koodataan SGML/HTML-tiedostoiksi. Phoenix-järjestelmässä kirjastoluettelo on linkitetty paikallisverkoissa olevaan julkaisutietokantaan. Jos kirjastoluettelosta löytyy julkaisu, joka on tässä julkaisutietokannassa, käyttäjä voi tilata sen tulostettavaksi paikallisverkon välityksellä lähimmälle tulostimelle. Järjestelmä on maksullinen, käyttäjä maksaa kopion asettamalla kopiointikortin työasemaan kytkettyyn maksulaitteeseen. Maksuun sisältyvät myös tekijänoikeuskorvaukset. Projektissa on mukana yliopistoja (erityisesti De Montfort), jotka ovat olleet aktiivisia erilaisissa digitaalisten kirjastojen projekteissa: South Bank, De Montfort, Huddersfield ja Open Universities. Yritysosapuolet ovat Digital Equipment, Rank Xerox and BLCMP.

Lisätietoja: <http://www.sbu.ac.uk:80/~litc/phoenix/>

eLib: ERCOMS, Electronic Reserve Copyright Management Systems

Hankkeessa rakennetaan järjestelmä, jolla voidaan kontrolloida ja tilastoida elektronisten tietovarantojen käyttöä ja copyright-korvauksia. Sen on tarkoitus olla yleinen siten, että sitä voidaan hyödyntää missä tahansa tietojärjestelmässä, jossa tarvitaan copyright-oikeuksien hallinnointia.

Lisätietoja: <http://ford.mk.dmu.ac.uk/Projects/ERCOMS/>

eLib: Sources of Digital Information

Marc Freskon British Librarylle tekemä selvitys erilaisista digitaalisessa muodossa olevista julkaisuista ja muista tiedonlähteistä. Selvityksessä on listattu yli kolme sataa erilaista tiedonlähdettä, myös muita kuin Internet-tiedostoja (paikallisesti käytettäviä tietokantoja, CD-tietolevyjä yms.). Suurin osa kuvauksista on kuitenkin tehty jo vuonna 1994, joten tiedot saattavat olla vanhentuneita.

Lisätietoja: <http://ukoln.bath.ac.uk/papers/bl/sdi/intro.html/> – tiedosto on poimittavissa tästä osoitteesta Word for Windows 2.0 -tiedostona; virallinen julkaisu on British Library R&D Report 6102.

EU/ESPRIT: COPEARMS

Projektin tarkoituksena on auttaa muita EU-projekteja erityisesti multimedia-aineiston välittämiseen liittyvissä tekijänoikeuskysymyksissä ja tekijänoikeuksien hallintaan tarkoitettujen järjestelmien kehittämisessä (electronic copyright management systems, ECMS). Projekti alkoi tammikuussa 1996 ja kestää kolme vuotta.

4. ELEKTRONISEN JULKAISEMISEN TIEDOSTOMUODOT JA OHJELMISTOT

4.1 Johdanto

Elektronisella julkaisemisella tarkoitetaan kaikkea sitä julkaisemista, jossa lukemiseen tarvitaan elektronisia apuvälineitä. Elektroninen julkaisu muodostuu tiedostosta tai joukosta tiedostoja ja sitä jaetaan tietovälineillä tai verkon välityksellä. Tietovälineinä käytetään nykyään tavallisimmin levykettä tai CD-ROM-levyjä, verkko on useimmiten Internet. Julkaisujen laatimisessa käytetään monenlaisia ohjelmia: tekstinkäsittely- ja julkaisuohjelmia, kuvankäsittely- ja piirto-ohjelmia sekä erityisesti verkkojulkaisemiseen tarkoitettuja ohjelmia. Elektronisten julkaisujen tiedostomuotoja (formaatteja) on myös suuri joukko, ja tietty julkaisu voi olla saatavana useammassa tiedostomuodossa.

4.2 Elektronisen julkaisemisen tiedostomuodoista

Tässä esiteltävistä tiedostomuodoista vanhin, mutta edelleen laajalti käytössä oleva muoto on Adoben PostScript. Se on ennen kaikkea tulostusmuoto, esimerkiksi valtaosa tulostimista tukee PostScript-muotoa. Koska PostScript-tiedostomuoto ei sovellu tiedostojen esittämiseen näyttöruudulla, Adobe kehitti sen pohjalta uuden tiedostomuodon Portable Document Format'in (PDF). PDF-muodossa olevia julkaisuja voi WWW-selaimen ja ilmaisen apuohjelman avulla lukea ruudulla ja tavallisesti myös tulostaa halutut sivut.

WWW-sivujen tavallisin esitysmuoto on HTML eli Hyper Text Markup Language. HTML perustuu rakenteisia dokumentteja koskevaan yleiseen merkkäuskieleen Standard Generalized Markup Language eli SGML.

Post Script

PostScript on Adoben kehittämä sivunkuvauskieli. Sivut voivat tekstin ohella sisältää grafiikkaa ja skannattuja kuvia. PostScript-muodossa olevat tiedostot on tarkoitettu tulostettaviksi paperille, eikä niinkään luettavaksi ruudulla.

PostScript-formaatin ominaisuuksia

- Adoben kehittämä sivunkuvauskieli.
- Laitteistoriippumaton tulostuslaitteiden ohjauskieli.
- ks. myös <http://www.adobe.com/prodindex/postscript/overview.html#whatpostscript>.

+	Säilyttää alkuperäisen dokumentin ulkoasun	-	Tarkoitettu lähinnä tulostettavaksi, ei luettavaksi ruudulta
+	Laitteistoriippumaton	-	Tiedostojen suuri koko
		-	Tiedostoa ei voi muuttaa

Adobe Acrobat PDF (Portable Document Format)

Adoben PDF-formaatti on tarkoitettu dokumenttien siirtoon eri järjestelmien ja käyttäjien välillä. Erityisesti se soveltuu dokumenttien jakamiseen ja julkaisemiseen tilanteissa, joissa on olennaista säilyttää alkuperäisen dokumentin ulkoasu riippumatta siitä, missä käyttöympäristössä dokumenttia käytetään (luetaan tai tulostetaan).

WWW-julkaisemisessa PDF-dokumentit ovat kuin mitä tahansa tiedostoja, eli niihin voidaan viitata normaalilla URL-viittauksella¹. Jos lukijan tietokoneeseen on asennettu PDF-lukuohjelma, dokumentti näkyy automaattisesti joko lukuohjelman omassa ikkunassa tai selainohjelman ikkunan sisällä.

PDF-dokumenttien tuottamiseen tarvitaan maksullinen ohjelma *PDF Writer*, *Distiller* tai *Capture* (kaikki Adoben tuotteita). Kun PDF-dokumenttien tuottamiseen tarkoitettujen ohjelmien asennettuina tietokoneeseen, mikä tahansa ohjelma, joka pystyy tulostamaan paperille, pystyy tuottamaan myös PDF-tiedostoja, joko suoraan (PDF Writer) tai PostScript tulostuksen kautta (Distiller). PDF-dokumentin muokkaus, esimerkiksi linkkien lisääminen, suoritetaan *Exchange*-ohjelmassa.

PDF-formaatin ominaisuuksia

- Perustuu PostScript-kieleen, säilyttää dokumentin ulkoasun.
- PDF-dokumentti sisältää tiedon dokumentissa käytetyistä kirjasimista. Tämä mahdollistaa kirjasimen korvaamisen saman tyyppisellä kirjasimella tapauksissa, joissa lukijan tietokoneessa ei ole PDF-dokumentissa käytettyä kirjasinta. Kirjasimet voidaan myös sijoittaa PDF-dokumentin sisään (*embed*), jolloin kirjasinten korvausta ei suoriteta vaikka lukijan tietokoneessa ei olisikaan ko. kirjasinta, vaan käytetään PDF-dokumenttiin sijoitettua kirjasinta. Kirjasinten sijoittaminen PDF-dokumenttiin kuitenkin kasvattaa PDF-tiedoston kokoa.
- PDF-dokumentti ei ole *kuva*² sivuista, vaan se sisältää myös dokumentin varsinaisen tekstisisällön, joka on valittavissa ja kopioitavissa, ellei sitä ole erikseen estetty. Dokumentin tekstiin voidaan suorittaa myös hakuja. Lukuohjelman (tai selaimen) ruudulla näkyvää tekstiä ja kuvia voidaan suurentaa ja pienentää tekstin laadun kärsimättä.
- Kuvat ovat dokumentin sisällä eikä viittauksina, kuten esimerkiksi HTML:ssä tai SGML:ssä. PDF-dokumentti on siis yksi kokonaisuus (yksi tiedosto).
- PDF-dokumenttia (tekstiä ja kuvia) voidaan pakata dokumentin luontivaiheessa. Tekstille / grafiikalle, väri- ja harmaasävykuville sekä mustavalkokuville voidaan valita erilaiset pakkausmetodit ja pakkauksen taso (pakkausmetodeista katso liite 1).
- PDF-dokumentti voi sisältää dokumentin sisäisiä hypertekstilinkkejä tai linkkejä WWW-sivuihin (normaali URL viittaus), sivujen miniatyyrikuvat (*thumbnails*, ks. liite 2, kuva 1) ja kirjanmerkkejä (*bookmarks*, ks. liite 2, kuva 2). PDF-dokumenttiin voidaan sijoittaa myös elektronisia muistilappuja ja digitaalista videota (QuickTime tai AVI).

¹Esimerkiksi HTML-dokumentista voidaan viitata PDF-dokumenttiin linkillä ``, jolloin linkin osoittaminen lataa PDF-dokumentin, ja jos lukuohjelma on asennettu oikein, käynnistää lukuohjelman. Vastaavasti PDF-dokumentista voi olla linkki toiseen kohtaan dokumentin sisällä, toiseen PDF-dokumenttiin tai HTML-sivuun.

²PDF-dokumentti voidaan tarvittaessa tallettaa myös PDF-ImageOnly formaatissa Capture-ohjelmalla, jos skannatulle paperidokumentille ei haluta suorittaa tekstintunnistusta (OCR). Tällöin dokumentti koostuu skannattujen sivujen kuvista ja dokumentin teksti ei ole haettavissa tai valittavissa. Skannattu sivu kuvana PDF-dokumentissa on kooltaan pienempi kuin vastaava sivu TIFF-kuvatiedostona.

- PDF-dokumentille voidaan määritellä erilaisia salausasteita, esimerkiksi estää dokumentin avaaminen ilman salasanaa, estää dokumentin muokkaaminen (esimerkiksi muistilappujen tai kirjanmerkkien ja linkkien lisääminen) tai estää tulostaminen paperille.
- PDF-dokumenttien lukemiseen tarvitaan erillinen lukuohjelma (Adoben *Acrobat Reader* tai jokin kolmannen osapuolen lukuohjelma. Acrobat Reader, <http://www.adobe.com/acrobat/readstep.html>). Acrobat Reader 3.0 integroituu (*plug-in*) WWW-selainohjelmiin kuten Netscape Navigator (<http://www.netscape.com/>) tai Microsoft Internet Explorer (<http://www.microsoft.com/ie/>), jolloin PDF-dokumentti on luettavissa selainohjelman ikkunassa.
- Acrobat versio 3.0:ssa on mahdollista optimoida dokumentit verkkojakelua varten, jolloin dokumentti jaetaan lukijalle sivu kerrallaan, ts. koko dokumenttia ei ladata kerralla.
- ks. myös Portable Document Format Reference Manual, <http://www.adobe.com/supportservice/devrelations/PDFS/TN/PDFSPEC.PDF>

+	Säilyttää alkuperäisen dokumentin ulkoasun	–	Dokumentin looginen rakenne ei ole eksplisiittisesti ilmaistu
+	Luettavissa eri käyttöympäristöissä (UNIX, PC/WIN, MAC), laitteisto- ja käyttöjärjestelmäriippumaton. Ilmainen lukuohjelma	–	Liiallinen sitoutuminen yhteen ohjelmistotuottajaan (Adobe)? PDF-formaatti on avoin, eli myös kolmannet osapuolet voivat valmistaa PDF-ohjelmistoja, mutta ainakin tällä hetkellä lähes kaikki ohjelmat ovat Adoben tuotteita
+	Hypertekstiominaisuudet, sivujen zoomaus	–	Tiedon löytyminen verkossa olevista PDF-dokumentista: Hakupalvelut, kuten AltaVista tai Infoseek Ultra eivät indeksoi PDF-dokumenttien sisältöä. Meta-datan sijoittaminen PDF-dokumenttiin?
+	Helppo tuottaa tekstinkäsittely- ja julkaisuohjelmista käsin	–	PDF-dokumentin sisältöä ei voi suoraan muokata, vaan muutokset on tehtävä alkuperäiseen dokumenttiin ja luoda siitä uusi PDF-tiedosto
		–	Arkistointi: Voidaanko nykyisiä PDF-dokumentteja lukea kymmenen vuoden päästä silloisilla ohjelmaversioilla?

SGML (Standard Generalized Markup Language)

SGML on merkkäuskieli, jolla kuvataan dokumentin *looginen rakenne*. Jokaisen dokumentissa esiintyvän elementin alkaminen ja loppuminen ilmaistaan deskriptiivisillä tunnisteilla, jotka ilmaisevat *mikä elementti on*, ei *miltä sen tulisi näyttää*.

Dokumentissa sallitut elementit (dokumentin osat, esimerkiksi otsikot, listat jne.) ja näiden elementtien väliset suhteet, järjestys ja toistettavuus määritellään dokumenttityypin määrittelyssä, DTD:ssä (Document Type Definition). Elementteihin ja niitä kuvaaviin tunnisteisiin voidaan myös liittää elementin ominaisuuksia kuvaavia attribuutteja. Dokumenttityypin määrittelyn kirjoittaja siis määrittelee millä nimillä dokumentissa esiintyviä elementtejä kutsutaan ja samalla millä tunnisteella elementin alkaminen ja loppuminen dokumentissa ilmaistaan.

Tunniste on merkkijono, joka ilmaisee elementin alkamisen (alkutunniste) ja elementin loppumisen (lopputunniste), esimerkiksi **<OTSIKKO>** ja **</OTSIKKO>**. **<** ja **>** merkit erottavat tunnisten varsinaisesta dokumentin tekstisisällöstä.

SGML-merkkaukielen ja SGML-dokumenttien ominaisuuksia

- SGML on ISO 8879:1986 standardi.
- Dokumenttityypin määrittely (DTD) määrittelee dokumentin loogisen rakenteen, joka ilmaistaan dokumentissa DTD:n mukaisilla tunnisteilla. Jokaista dokumenttityyppiä varten voidaan luoda oma DTD (esimerkiksi muistio, kirje, kirja jne.). Verkosta löytyy myös valmiita DTD:tä erilaisille dokumenttityypeille, joita voi hyödyntää joko sellaisenaan tai lähtökohtana oman DTD:n kirjoittamiselle.
- SGML-dokumentin mukana toimitetaan ko. dokumentin DTD tai viittaus käytetyn DTD:n nimeen ja sijaintipaikkaan.
- SGML on meta-kieli, ts. sillä voidaan määritellä toisia merkkaukieliä kuten esimerkiksi HTML.
- SGML-dokumentin käsittelyyn liittyy yleensä dokumentin loogisen rakenteen oikeellisuuden tarkistaminen DTD:n perusteella (*parsing, validating*). SGML-editoreissa tämä yleensä tarkoittaa, että kirjoittajaa ohjataan kirjoittamaan dokumentti loogisen rakenteen kannalta oikein. Kirjoittaja ei voi esimerkiksi sijoittaa kirjan luvun sisällä otsikkotasoa 3 välittömästi otsikkotason 1 jälkeen, koska välistä puuttuu otsikkotaso 2³.
- SGML ei määrittele dokumentin ulkoasua. Ulkoasu voidaan määritellä erillisillä, loogisen rakenteen osiin liittyvillä tyylimäärityksillä. Samaan dokumenttiin liittyviä tyylimäärityksiä voi olla useita erilaisia, esimerkiksi paperitulostetta ja elektronisessa muodossa julkaistavaa versiota varten.
- DTD:t ja SGML-dokumentit ovat puhtaassa ASCII-muodossa, eli *periaatteessa* muokattavissa millä tahansa tekstieditorilla. Käytännössä kuitenkin monimutkaisten SGML-dokumenttien muokkaamiseen tarvitaan jokin SGML:ää ymmärtävä editori.
- ks. myös <http://www.sil.org/SGML/SGML.html>.

³ Olettaen, että DTD:ssä on määritelty, että otsikkotaso 3 ei voi esiintyä välittömästi otsikkotason 1 jälkeen. Tätä oikeellisuuden tarkistamista ei ole useimmissa HTML -editoreissa, vaikka HTML onkin SGML:n osajoukko.

+	Dokumentin looginen rakenne on ilmaistu eksplisiittisesti	–	Monimutkainen: Dokumenttityypin määrittelyn kirjoittaminen vaikeaa monimutkaisille dokumenteille, esimerkiksi kirjoille
+	Dokumentissa olevan tekstin loogisen rakenteen oikeellisuus voidaan tarkistaa	–	Matemaattisten kaavojen ja taulukoiden ilmaiseminen SGML:ssä. Esimerkiksi ilmaistaanko taulukko riveinä ja soluina (rivi 5, solu 3), vai taulukossa olevan solun sisällön perusteella ("Liikevaihto")?
+	Loogista rakennetta voidaan hyödyntää monella tavalla; rakenteeseen kohdistuva tiedonhaku, tiedon suodattaminen, osadokumenttien luominen samasta lähdemateriaalista, erilaiset näkymät dokumenttiin jne.	–	Kaupalliset SGML-sovellukset suhteellisen kalliita
+	Laitteisto- ja ohjelmisto- ja käyttöjärjestelmäriippumaton: SGML-muodossa olevan tiedon elinkaari on pitkä, arkistointi	–	SGML-dokumenttien selainohjelmat harvinaisia (vrt. WWW-selaimet); SGML-dokumenttien jakelun hyöty WWW:ssä kyseenalaista
+	Kansainvälinen standardi	–	Tavallisen tekstidokumentin, esimerkiksi Word-dokumentin, konvertoiminen SGML-muotoon vaikeaa

XML (Extensible Markup Language)

XML on SGML:n yksinkertaistettu versio ja tarkoitettu nimenomaan WWW käyttöön. Samoin kuin SGML:ssä, myös XML:ssä voidaan määritellä omia dokumenttityypin määrittelyjä. XML dokumentin merkkaukseen käytetyt tunnisteet eivät siis ole ennalta tarkoin määriteltyjä kuten HTML:ssä.

Tällä hetkellä XML:n kehitystyö on kesken, eikä XML:ää ymmärtäviä selainohjelmia ole vielä juurikaan saatavilla.

XML-kielen ominaisuuksia

- ks. myös <http://www.sil.org/SGML/XML.html>.

HTML (Hyper Text Markup Language)

HTML on merkkaukieli, jolla ilmaistaan HTML-dokumenttityyppejä edustavien hypertextidokumenttien looginen rakenne. HTML:n tärkein ominaisuus on mahdollisuus luoda laitteisto-, ohjelmisto ja käyttöjärjestelmäriippumattomia hypertextidokumentteja, joissa ovat elementit

(kuvat, äänet, videot yms.) voivat olla erillisinä tiedostoina ja sijaita fyysisesti eri puolilla maailmaa olevissa WWW-palvelinkoneissa.

HTML-dokumentin ulkoasu luodaan selainohjelmassa jolla dokumenttia luetaan, ko. selainohjelman määritysten ja ikkunan koon mukaan. HTML-dokumentin kirjoittajalla ei siis ole kovin paljon mahdollisuuksia vaikuttaa siihen miltä dokumentti näyttää lukijan ruudulla.

HTML-dokumentin kirjoittaja voi kuitenkin liittää dokumenttiin tyylimäärittelyksiä joko dokumentin HEAD-elementin sisälle⁴ tai viittauksena tyylitiedostoon. Tyylitiedostolla voidaan määrittellä HTML-elementtien muotoilu (marginaalit, kirjasimet, värit, taustat ym.) elementtikohtaisesti. Useat HTML-tiedostot voivat käyttää samaa tyylimäärittelyä, joten kokonaisen WWW-palvelimen ulkoasun muuttaminen onnistuu yhtä tyylitiedostoa muuttamalla. Microsoft Internet Explorer ja Netscape Navigator 4.0 ymmärtävät Cascading Style Sheet (CSS, <http://www.cern.ch/WebOffice/Courses/CSS/CSSFAQ.html>) tyylimäärittelyksiä. Dokumentin lukija voi kuitenkin omassa selainohjelmassaan määrittellä, että selaimen asetukset ylittävät dokumentissa mahdollisesti olevat tyylimäärittelyt. Myös Dynamic HTML (DHTML, esim. <http://www.microsoft.com/workshop/author/dHTML/>) mahdollistaa perus-HTML:ää monipuolisemmat mahdollisuudet dokumentin ulkoasun määrittelemiseen.

Koska HTML-dokumentit ovat puhtaassa ASCII-tekstimuodossa, dokumenttien siirtäminen verkossa on suhteellisen nopeaa, joskin suurten kuva- yms. tiedostojen liittäminen dokumenttiin hidastaa sivujen latautumista.

HTML-merkkaukielen ja HTML-dokumenttien ominaisuuksia

- HTML-merkkaukielessä sallitut tunnisteet on määritelty SGML-syntaksin mukaisessa HTML-DTD:ssä⁵, eli HTML-dokumenttien dokumenttityypin määrittelyssä. HTML-dokumentti on siis samalla myös SGML-dokumentti.
- Selainohjelmissa on ”sisäänrakennettuna” tieto HTML-DTD:stä, joten DTD:tä ei tarvitse toimittaa HTML-dokumentin mukana (useimmiten DTD:hen ei edes viitata HTML-dokumenteissa).
- HTML-dokumentin sisältö on puhdasta ASCII-tekstiä, ja on siis muokattavissa millä tahansa tekstieditorilla. HTML-editoreita ja konverttereita on runsaasti saatavilla, sekä kaupallisina, että shareware- tai freeware-ohjelmina.
- Dokumentin ulkoasu muotoutuu selainohjelman ja sen asetusten sekä ikkunan koon mukaan. Kirjoittajalla ei ole kovin suurta kontrollia sen suhteen miltä dokumentti näyttää lukijan ruudulla.
- Käyttämällä tyylitiedostoja voidaan kuitenkin jossain määrin vaikuttaa dokumentin ulkoasuun (kirjasinkoot, marginaalit, värit ym.).
- Koska HTML:ssä olennaista on dokumentin looginen rakenne, A4-orientoituneiden (alunperin paperille suunniteltujen) dokumenttien esittäminen HTML-muodossa voi olla mahdotonta niin että alkuperäinen ulkoasu säilyisi.
- Hypertekstiviittaukset mihin tahansa verkossa olevaan dokumenttiin; muut HTML-dokumentit, PDF-dokumentit, äänitiedostot, digitaalinen video jne.

⁴ Esimerkiksi HTML-dokumentin HEAD-elementin sisällä: `<STYLE>H1 {font-family: helvetica; font-size: 12pt; color:blue}</STYLE>`

⁵ Draft-versio uusimmasta HTML-DTD:stä (4.0) löytyy osoitteesta <http://www.w3.org/TR/PR-HTML40/>.

- Kaikkia dokumentteja, joihin voidaan viitata ei voida näyttää suoraan selainohjelman ruudulla, vaan tarvitaan apuohjelmia⁶.
- Kuvat sijoitetaan HTML-dokumenttiin viittauksella kuvatiedostoon, ts. ruudulla näkyvä lopputulos voi koostua useista eri tiedostoista, jotka voivat fyysisesti sijaita missä tahansa verkkoon kytketyssä koneessa.
- HTML-standardi on jatkuvassa käymistilassa ja lisäksi selainvalmistajien omat epästandardit laajennukset sekoittavat tilannetta. Viimeisin virallinen HTML-standardi on versio 3.2.
- ks. myös <http://www.w3.org/MarkUp/>.

+	Dokumentin looginen rakenne on ilmaistu eksplisiittisesti	–	Kirjoittajalla ei ole kontrollia dokumentin ulkoasun suhteen. Monia ulkoasumäärittäyksiä ei edes ole mahdollista koodata HTML:llä
+	Dokumentissa oleva teksti haettavissa erilaisilla ilmaisilla hakupalveluilla, haut voidaan kohdistaa osaan HTML-dokumentin loogisen rakenteen osista	–	Matemaattisten kaavojen esittämiseen ei ole standardia, yleensä kaavat sijoitettava HTML-dokumenttiin kuvana
+	Ilmaiset selainohjelmat (myös maksullisia saatavilla). WWW:n ja HTML:n suuri suosio julkaisukanavana	–	Kaikki teksti ja kuvat HTML-dokumentissa ovat kopioitavissa ja tulostettavissa
+	Dokumenttien kirjoittaminen ei (välttämättä) vaadi erikoisohjelmia. Saatavilla runsaasti ilmaisia tai shareware HTML-editoreita ja konverttereita	–	HTML-tunnisteet laajalti väärinkäytettyjä ulkoasun, ei rakenteen tuottamiseen
+	Standardi (joskin myös epästandardeja tunnisteita käytössä, standardilla jatkuvia muutospaineita).	–	Selainvalmistajien (Microsoft, Netscape) omat, epästandardit laajennukset HTML-standardiin. HTML-dokumentit, joissa on käytetty selainspesifejä tunnisteita eivät välttämättä ole luettavissa toisen valmistajan selaimella ainakaan siinä muodossa kuin kirjoittaja on tarkoittanut

T_EX ja L^AT_EX

T_EX on tekstinkäsittely- ja ladontaohjelma, joka sopii erityisesti matemaattisia lausekkeita, eri kirjasinlajeja ja erikoissymboleja sisältävien julkaisujen tuottamiseen. T_EX on pohjimmiltaan sivunkuvauskieli, joka mahdollistaa käyttäjän omien makrojen teon, jolloin nämä muunnokset

⁶ Riippuu luonnollisestikin käytettävästä selainohjelmasta mitä dokumentteja ohjelma osaa näyttää ilman apuohjelmia. Suurin osa tarvittavista apuohjelmista on ilmaisia. Plug-in -tyyppiset apuohjelmat integroituvat selainohjelmaan ja näyttävät dokumentit selainohjelman ikkunassa.

voivat tulla varsin mutkikkaiksi. T_EX:in yhteyteen on tehty useita makropaketteja kuten L_AT_EX, joka helpottaa laajojen dokumenttien tuottamista sisältäen välineitä mm. tekstin lukujen automaattiseen numerointiin, sisällysluettelon automaattiseen tuottamiseen, hakemiston tekemiseen, tekstin sisäisten viittausten sivunumeroiden automaattiseen tuottamiseen sekä kuvioiden tuottamiseen. L_AT_EX -muotoinen teksti on geneerisesti merkattua ja ASCII-muotoista, jolloin muutokset esimerkiksi HTML:ään ja SGML:ään ovat mahdollisia. Perus-T_EX:in sijasta kannattaakin yleensä käyttää L_AT_EX:ia, jolloin monet asiat voidaan tehdä puuttumatta teknisiin yksityiskohtiin. Tällöin toteutus tapahtuu kuitenkin oletusarvojen mukaan, jotka eivät aina tuotakkaan haluttua tulosta. T_EX on eräajotyypinen ohjelma. Teksti kirjoitetaan jollakin editorilla tiedostoon siten, että varsinaisen tekstin joukossa on ohjausmerkkejä, jotka säätelevät mm. tekstin ulkoasua. Aikaansaatu tiedosto annetaan syötteenä T_EX-ohjelmalle, joka tuottaa siitä DVI-tiedoston, joka on periaatteessa laitteistoriippumaton ja näin ollen siirrettävissä oleva tiedostomuoto, eikä se ole ihmisen suoraan luettavissa. Ongelman muodostavat kuvat, jotka edellyttävät käytetyltä DVI-ohjelmalta tukea käytetylle kuvatekniikalle. DVI-tiedosto ei myöskään sisällä käytettyjä fontteja, vaan ainoastaan viitteen niihin. Siksi kyseiset fontit on joko siirrettävä dokumentin mukana tai loppukäyttäjällä on oltava ne valmiiksi asennettuina. Aikaansaatu DVI-tiedosto voidaan eri ohjelmilla tulostaa erilaisille laitteille ja siitä voidaan tuottaa myös PostScript-tiedosto.

Lisätietoja: <http://www.tex.ac.uk/CTAN/latex/>
<http://www.cogs.susx.ac.uk/cgi-bin/texfaq2HTML>
<ftp://ftp.funet.fi/pub/TeX/CTAN/>

Yhteenveto: Eri tiedostomuotojen vertailua

	PostScript	PDF	HTML	SGML
Säilyttää dokumentin ulkoasun	✓	✓	(✓)*	(✓)*
Dokumentin looginen rakenne eksplisiitisti ilmaistu			✓	✓
Valmis dokumentti kirjoittajan muokattavissa			✓	✓
Ilmaiset ohjelmat dokumenttien tuottamiseen	✓		✓	(✓)†
Ilmaisia lukuohjelmia saatavilla	✓	✓	✓	(✓)†
WWW-julkaiseminen		✓	✓	(✓)*
CD-ROM julkaiseminen	✓	✓	✓	✓
Dokumentin suojaaminen		✓	(✓)*	(✓)*
Dokumentin pakkaaminen		✓		

- * Ulkoasua on mahdollista kontrolloida tyylimääriyksillä (style-sheets). Silti alkuperäisen dokumentin ulkoasu voi olla mahdotonta toistaa.
- † Suurin osa ohjelmista maksullisia, mutta myös joitain ilmaisia on saatavilla.
- ‡ Dokumentit voidaan suojata WWW-palvelimen salasanasuojauksella, mutta varsinaista dokumentikohtaista salasanasuojausmahdollisuutta ei ole kuten PDF:ssä.

4.3 Kuvaformaateista

Elektroniset julkaisut sisältävät usein myös kuvia, sekä viivapiirroksia että sävykuvia. Nämä voivat olla joko sulautettuina itse julkaisuun (julkaisun tiedostomuotona esim. PostScript tai PDF) tai linkitettyinä siihen (julkaisun tiedostomuotona esim. HTML tai SGML). Kuville tarkoitettuja tiedostomuotoja on kymmeniä. Tässä esitellään verkkojulkaisemisen kannalta keskeiset tiedostomuodot.

TIFF (Tagged Image File Format)

- Bittikartta
- 24-bittinen (16 milj. väriä)
- Pakkausmenetelmät: Uncompressed, PackBits RLE, LZW, JPEG, CCITT Group 3 & 4
- <http://www.cis.ohio-state.edu/hypertext/faq/usenet/graphics/fileformats-faq/part3/faq-doc-148.html>.

TIFF ei sovellu verkkojulkaisemiseen kovin hyvin, koska selainohjelmat eivät osaa näyttää TIFF-muodossa olevia kuvia suoraan selaimen ruudulla ilman soveltuvaa plug-iniä. TIFF-tiedostot ovat myös kooltaan huomattavasti suurempia kuin esimerkiksi JPEG-muodossa olevat kuvat.

TIFF-muotoa on käytetty melko laajasti esimerkiksi skannattujen paperisivujen tallettamiseen kuvamuodossa, myös verkkojulkaisemisessa. TIFF-formaatin ongelmana on kuitenkin kuvatie-dostojen suuri koko ja TIFF-formaatin erilaiset ”murteet”; kaikki ohjelmat eivät osaa avata kaikkia TIFF-muodon variantteja.

Suurin osa kuvankäsittelyohjelmista osaa avata (ja tallettaa) TIFF-muotoa, joten TIFF-muodossa oleva kuva voidaan muuntaa esimerkiksi GIF- tai JPG- muotoon.

JPEG (Joint Photographic Experts Group)

- Bittikartta
- 24-bittinen (16 milj. väriä)
- Pakkausmenetelmä: JPEG, eri tasoilla. Pakkausalgoritmi on hävittävä (lossy), eli kuvan laatu huononee pakkaustason kasvaessa. Kuvan huononeminen pienillä pakkaustasoilla ei ole silmin havaittavaa.
- <ftp://rtfm.mit.edu/pub/usenet/news.answers/jpeg-faq/part1>.

JPEG on tarkoitettu täysväri- tai harmaasävykuvien pakkaamiseen, eikä sovellu kovin hyvin esimerkiksi bittikarttamuodossa olevien viivapiirrosten tai kuvien, joissa on laajoja yhtenäisiä värialueita, tallettamiseen. JPEG-formaatin käyttämisestä 8-bittisille kuville (256 väriä) ei ole suurtakaan hyötyä, vaikka se toki onkin mahdollista.

JPEG on yleisesti käytetty kuvaformaatti WWW-julkaisemisessa, koska kuvien koko saadaan pysymään melko pienenä⁷ valokuvatyypisissä kuvissa. Lähes kaikki kuvankäsittelyohjelmat pystyvät avaamaan ja tallettamaan JPEG-formaattia.

GIF (Graphics Interchange Format)

- Bittikartta
- 8-bittinen (256 väriä tai harmaan sävyä)
- Pakkausmetodi: LZW (automaattisesti), hukkaamaton (lossless)
- <http://www.faqs.org/faqs/graphics/fileformats-faq/part3/section-57.html>.

Yleisin kuvaformaatti WWW-julkaisuissa. GIF-formaatti sopii hyvin bittikarttamuodossa olevan grafiikan, esimerkiksi viivapiirrosten tallettamiseen. GIF-formaattia voidaan käyttää myös valokuvien tallettamiseen, esimerkiksi jos vaaditaan jonkin tietyn kuvassa olevan värin läpinäkyvyyttä. JPEG-muodossa oleviin kuviin läpinäkyvyyttä ei voida määritellä.

GIF-formaatista on olemassa kaksi versiota; GIF87a ja GIF89a. GIF87a on perus-GIF formaatti, mutta GIF89a:ssa on joitain lisäominaisuuksia, kuten esimerkiksi interlace-metodin⁸ käyttömahdollisuus talletuksessa, tietyn värin läpinäkyvyys ja GIF-animaatiot.

PNG (Portable Network Graphics)

- Bittikartta
- maksimissaan 48-bittinen (harmaasävykuvat 16-bittinen)
- Pakkausmetodi: RLE (Run-length), hukkaamaton (lossless)
- <http://www.boutell.com/boutell/png/>

PNG on kehitetty korvaamaan GIF-formaatti. Ominaisuudet ovat muuten samat kuin GIF-formaatissa, mutta lisäksi täysvärikuvat voivat olla jopa 48-bittisiä ja harmaasävykuvat 16-bittisiä.

Useimmat selaimet eivät vielä osaa näyttää PNG-kuvia selaimen ruudulla, vaan niiden katselamiseen tarvitaan jokin PNG:tä osaava apuohjelma. Osa kuvankäsittelyohjelmista, esimerkiksi Paint Shop Pro 4.0 osaa kuitenkin lukea ja tallettaa PNG-formaattia.

⁷ Esimerkiksi TIFF-kuva, jonka koko on 4,1 Mt voidaan tallettaa JPEG-kuvana, jolloin tiedoston kooksi tulee 1,1 Mt (pakkaustaso "Good") tai 0,05 Mt (pakkaustaso "Fair") (*Jarkoskuvan tulostusohjelma*, 1996). Tiedostojen kokoerot ovat viitteellisiä, koska kuva- ja värialueiden ym. laajuus ja määrä vaikuttavat JPEG-formaatin soveltuvuuteen ko. kuvan pakkaamiseen.

⁸ Interlace-metodi tarkoittaa, että kuva on talletettu linja kerrallaan. Kuvaa ladattaessa esimerkiksi selaimen ruudulle, kuva latautuu joka neljäs linja kerrallaan. Lukija saa siis käsityksen siitä mitä kuva esittää ennenkuin se on kokonaan latautunut, ja voi päättää haluaako odottaa kuvan täydellistä latautumista.

BMP (Windows BitMap Format)

- Bittikartta
- 24-bittinen (16 milj. väriä)
- pakkausmetodi: RLE
- <http://www.faqs.org/faqs/graphics/fileformats-faq/part3/section-18.html>.

Windows-kuvaformaatti. Selaimista ainoastaan Microsoft Internet Explorer osaa näyttää BMP-kuvat selaimen ruudulla ilman apuohjelmia.

Yhteenveto: Kuvaformaattien vertailua

	TIFF	JPEG	GIF	PNG	BMP
Bittikartta	✓	✓	✓	✓	✓
bittii/pikseli (maksimi)	24-bit	24-bit	8-bit	48-bit	24-bit
Kuvassa olevan värin läpinäkyvyys			✓	✓	
Interlace-metodi			✓	✓	
Kuvan pakkaus	✓	✓	✓	✓	✓
Soveltuvuus					
Valokuvat	✓	✓	(✓)	✓	(✓)
Viiivapiirrokset, laajoja yhteisiä värialueita			✓		✓
Verkojulkaiseminen (WWW)	(✓)*	✓	✓	(✓)	(✓)

* Vaatii soveltuvan plug-inin

4.4 Liikkuvan kuvan ja äänen formaateista

AVI (Video for Windows)

AVI (Audio Video Interleave) on Microsoftin liikkuvan kuvan tiedostomuoto. Tiedostot ilmaistaan *.AVI tarkentimella. AVI-tiedosto voi sisältää myös kuvien kanssa synkronoitua ääntä.

Windows 95 käyttöjärjestelmässä on mukana *Video for Windows* -ajurit, joiden avulla voidaan katsella AVI-tiedostoja esimerkiksi Media Player-ohjelmalla. Windows 3.1-ympäristöön Video for Windows on asennettava erikseen.

AVI-tiedosto voidaan liittää HTML-dokumenttiin joko normaalina linkkinä (A HREF) tai upottaa sivulle (inline, EMBED).

Lisätietoja: <http://www.rahul.net/jfm/avi.html>.

MPEG (Moving Picture Experts Group)

MPEG on liikkuvan kuvan ja äänen kompressoitimenetelmä. MPEG-tiedostojen tarkennin on *.MPG liikkuvalla kuvalla (video) ja *.MP2 tai *.MP3 äänitiedostoille. MPEG-videotiedosto (*.MPG) voi sisältää myös kuvan kanssa synkronoitua ääntä.

MPEG-tiedostojen katselamiseen tarvitaan erillinen ohjelma, esimerkiksi *InterVU MPEG Player*-plug-in Netscape Navigatoriin (<http://www.intervu.net/>).

Lisätietoja: <http://www.mpeg1.de/mpegfaq/>.

QuickTime

QuickTime on Applen kehittämä liikkuvan kuvan tiedostomuoto. Tiedosto koostuu useista kuvista (frame), jotka esitetään järjestyksessä peräkkäin. Tiedosto voi sisältää myös ääntä, joka on synkronoitu kuvien esitykseen. QuickTime-tiedostot ilmaistaan tiedoston nimessä tarkentimella *.MOV tai *.QT.

QuickTime-tiedostojen katseluun tarvitaan erillinen Movie Player-ohjelma, joka voidaan asettaa esimerkiksi Netscapen apuohjelmaksi, jolloin osoitettaessa linkkiä joka viittaa QuickTime-tiedostoon käynnistetään ko. ohjelma automaattisesti. Movie Player-ohjelma voi toimia myös plug-inina ja QuickTime-tiedosto voidaan upottaa HTML-dokumenttiin (EMBED)⁹. Movie Player-ohjelma on ilmainen ja löytyy osoitteesta <http://www.quicktime.apple.com/>.

AVI- ja QuickTime-muodossa olevia videoita voi sijoittaa PDF-dokumentteihin Exchange-ohjelman Movie-työkalulla.

WAV (Windows Wave Format)

WAV on Windows-ääniformaatti. Netscape Navigator 3.0:ssa on sisäinen tuki AIFF, AU, MIDI ja WAV-äänitiedostojen soittamiseen.

Äänikorttien mukana tulee usein myös ohjelmat, joilla ääniä voidaan digitoida WAV-formaattiin (tai myös muihin ääniformaatteihin) esimerkiksi audio-CD-levyiltä¹⁰. Äänen kuunteleminen kuitenkin edellyttää, että koneessa on äänikortti ja kaiuttimet tai kuulokkeet.

MIDI (Musical Instrument Digital Interface)

MIDI on musiikin SGML. Toisin kuin digitoiduissa äänitiedostoissa, kuten esimerkiksi WAV tai AU-formaateissa, MIDI-tiedostossa on ainoastaan ohjeet kuinka ääntä tuotetaan äänikortil-

⁹ Esimerkiksi <EMBED src="esimerkki.mov" WIDTH=100 Height=100 AUTOPLAY="TRUE" CONTROLLER="FALSE">.

¹⁰ Äänen käsittelyohjelmia löytyy esimerkiksi osoitteesta: <http://www.eunet.fi/tucows/sound95.html>.

la¹¹ tai syntetisaattorilla; mitä soitinääntä, millä sävelkorkeudella ja kuinka pitkään ääni soi. MIDI siis soveltuu ainoastaan musiikin esittämiseen.

MIDI-tiedostojen tuottamiseen tarvitaan yleensä sekvensseri-ohjelma ja koskettimisto, jolla soitetaan jokainen haluttu instrumenttiosalle kanavalleen.

MIDI-tiedoston etu verrattuna digitoituun ääneen on tiedostojen pienempi koko. MIDI-tiedosto voidaan upottaa HTML-tiedostoon EMBED-tunnisteella. Microsoft Internet Explorer tukee myös taustääntä, esimerkiksi <BGSOUND SRC="esimerkki.mid" LOOP=INFINITE>.

Lisätietoja: http://www.midisoft.com/HTML/tech_support/faq/faq-midi.htm

4.5 Adoben Acrobat-ohjelmista

Laitteistovaatimukset (3.0)

Windows	Macintosh
<ul style="list-style-type: none">• Vähintään 386-prosessori• Windows 3.1, Windows 95 tai Windows NT 3.51 (tai uudempi)• Keskusmuistia 8 MB (Reader, Exchange ja Catalog, NT 16 MB), 16 MB (Distiller tai Capture, NT 24 MB)• CD-ROM-asema (ohjelmat saatavissa 1,44 MB levykkeillä erikoispyynnöstä)• Suositellaan Pentium-pohjaista mikroa, jossa Windows 95 tai Windows NT, vähintään 16 MB keskusmuistia ja 40 MB levytilaa	<ul style="list-style-type: none">• 68020 prosessori tai parempi• System 7.0 tai uudempi• Reader 3.5 MB keskusmuistia (5 MB PowerMac), Exchange 4 MB (6 MB), Distiller 6 MB (8 MB), Capture plug-in 16 MB (ainoastaan PowerMac)• CD-ROM asema• Suositellaan Power Macintosh, jossa System 7.1.2 tai uudempi, 16 MB keskusmuistia ja 40 MB kiintolevytilaa.

Adobe Acrobat Reader

Tarkoitettu PDF-tiedostojen lukemiseen ja tulostamiseen. Saatavilla ilmaiseksi Adoben WWW-sivuilta

Versiot eri käyttöjärjestelmille: Windows 95 ja NT, Windows 3.1 ja Windows for Workgroups, Macintosh, UNIX, LINUX ja OS/2.

Adobe Acrobat PDF Writer

Kirjoitinajuri, jolla voidaan tulostaa mistä tahansa ohjelmasta avoinna oleva dokumentti PDF-tiedostoksi.

¹¹ Toistettaessa MIDI-tiedostoja äänikortilla, soitinäänien laatu on huomattavasti huonompi kuin syntetisaattorilla (äänikortista riippuen).

Tulostusvaiheessa voidaan valita pakkausmenetelmä ja dokumenttiin mahdollisesti sijoitettavat kirjasimet. PDF Writerilla ei voi kuitenkaan sijoittaa PDF-dokumenttiin linkkejä, kirjanmerkkejä tai muita lisäarvoa tuovia ominaisuuksia eikä asettaa dokumentin salasanasuojauksia. PDF Writerilla tuotettu PDF-dokumentti on tulostetun paperiversion elektroninen kopio.

Adobe Acrobat Distiller

Ohjelma, jolla tuotetaan PostScript-tiedostosta PDF-dokumentteja.

Joissain tapauksissa¹² on PDF-dokumentin laadun kannalta parempi tulostaa dokumentti ensin PostScript-tiedostoksi (valitaan tulostimeksi Acrobat Distiller), jonka jälkeen luodaan PDF-tiedosto Distiller-ohjelmalla. Distiller voidaan asettaa tarkkailemaan käyttäjän määrittelemiä hakemistoja ja luomaan automaattisesti PDF-tiedosto ko. hakemistoihin sijoitettavista PostScript-tiedostoista. Distiller-ohjelmassa voidaan PDF-dokumentin asetuksiin vaikuttaa hieman laajemmin kuin PDF Writerilla. Lisäksi ohjelma voi luoda automaattisesti miniatyyrikuvat dokumentin sivuista.

Adobe Acrobat Exchange

Exchange on tarkoitettu jo olemassa olevien PDF-dokumenttien muokkaamiseen, esimerkiksi kirjanmerkkien, linkkien, muistilappujen ja videoiden lisäämiseen PDF-dokumenttiin.

Myös salasanasuojaukset asetetaan Exchange-ohjelmassa.

Adobe Acrobat Capture

Capture on tarkoitettu paperimuodossa olevien sivujen skannaamiseen optisella kuvanlukijalla, tekstin tunnistamiseen (OCR) ja PDF-muotoon tallettamiseen.

Capture tunnistaa TIFF-muodossa olevasta kuvasta (sivusta) siinä käytetyt kirjasimet sekä tekstin sanaston perusteella. Tällä hetkellä Capturen suurin ongelma on siinä, että saatavilla ei ole suomenkielistä sanastoa, joten suomenkielisen tekstin tunnistus ei onnistu. Capture osaa hyödyntää ainoastaan mustavalkoisia (200–600 dpi) ja harmaasävykuvia (200–400 dpi) PDF-dokumentin tuottamiseen, joten värillisinä skannatut kuvat on muutettava harmaasävykuviksi, jolloin menetetään alkuperäisen dokumentin väri-informaatio. Capturen tekstintunnistuksessa on myös ongelmia kuvissa ja diagrammeissa olevien tekstien kanssa, koska ohjelma yrittää tunnistaa ne tekstinä, eikä kuvana.

Capturella skannatut sivut voidaan tallettaa myös PDF ImageOnly muodossa, jolloin sivu talletetaan PDF-tiedostoon kuvana, ilman kirjasinten ja tekstin tunnistusta. Ruudulla näkyvä lopputulos ei ole yhtä hyvä kuin jos kirjasimet ja teksti olisi tunnistettu, mutta riittää mainiosti tulostarkoituksiin¹³.

¹² Esimerkiksi jos dokumentissa on EPS-muotoisia kuvia tai matemaattisia erikoismerkkejä.

¹³ Esimerkki skannatuista ja Capturella PDF ImageOnly muodossa talletetusta julkaisuista löytyy osoitteesta <http://www.inf.vtt.fi/PDF/tiedotteet/1982/T115.PDF>

Adobe Acrobat Catalog

Catalog-ohjelmalla voidaan indeksoida PDF-dokumenttien tekstisisältö, ja käyttää sitä hakujen suorittamiseen useista PDF-dokumenteista yhtäaikaaisesti.

Lisätietoja: <http://www.adobe.com/prodindex/acrobat/main.html>

Compose 2.0

Ambia Corporationin Compose on tarkoitettu PDF-dokumenttien kirjanmerkkien ja linkkien luomisen automatisoimiseen. Esimerkiksi kirjanmerkkien automaattinen luominen perustuu PDF-dokumentissa oleviin kirjasimiin, niiden kokoon ja tyyliin (lihavoitu, kursivoitu).

Compose on Exchange-ohjelmaan liitettävä plug-in ja näkyy ohjelmaikkunassa uutena valikkona. Ilmainen 30 päivän demoversio löytyy osoitteesta <http://www.ambia.com/>

Myös muita vastaavia muiden kuin Adoben valmistamia ohjelmia on saatavilla, esimerkiksi osoitteessa <http://www.emrg.com/> on koottuna linkkejä erilaisiin kaupallisiin ja shareware tai freeware apuohjelmiin.

4.6 Tekstintunnistusohjelmista

Caere OmniPage Limited Edition ja Pro 8.0

Omnipage on tarkoitettu paperimuodossa olevien sivujen skannaamiseen optisella kuvanlukijalla, tekstin tunnistamiseen (OCR) ja tallettamiseen useisiin eri tekstinkäsittelyohjelmien, PageMakerin ja Excelin tiedostomuotoihin sekä HTML-muotoon.

Omnipage tunnistaa eri kuvankäsittelymuodoissa talletetusta sivusta (.bmp, .dcm, .jpg, .pcx, .tif, .met) siinä käytetyt kirjaimet ja merkit. Ohjelma tunnistaa hyvin myös skandinaaviset merkit ja yrittää Adobe Acrobat Capturen tapaan tunnistaa tekstiä myös kuvista, mutta ei läheskään niin hanakasti kuin Adoben Capture tekee. LE-versiossa kuvat eivät tule mukaan lopulliseen dokumenttiin, mutta uudessa ohjelmaversiossa myös kuvat saadaan automaattisesti mukaan.

Limited Edition -versiolla tuotetut dokumentit vaativat melko paljon jatkokäsittelyä, oikoluvun ja myös muotoilua, sillä joidenkin merkkien tai merkkiyhdistelmien tunnistaminen sekä palstoitettu teksti tuottaa ohjelmalle suuria hankaluuksia. Uudessa versiossa on parannettu kirjainten tunnistusominaisuuksia ja palstoitetun tekstin tunnistus onnistuu myös paremmin. Ohjelmalle voidaan "opettaa" tunnistamaan haluttuja merkkejä. Uudella versiolla saadaan melko hyvin säilytettyä myös paperiversion ulkonäkö Adoben Capturen tapaan. Kaksipuolisten sivujen käsittely onnistuu myös paremmin kuin LE-versiossa, eikä vaadi enää PaperPort-ohjelman käyttöä.

Projektissa on käytetty toistaiseksi ainoastaan Limited Edition -versiota, mutta uusi PRO 8.0 -versio on hankittu vuodenvaihteessa 97/98, joten siitä ei ole vielä käyttökokemuksia.

Laitteistovaatimukset (Pro 8.0)

- 486-prosessori

- Windows 95 tai NT
- 8 MB keskusmuistia (suositellaan 16 MB)
- 33 MB kiintolevytilaa
- CD-ROM-asema
- **Lisätietoja:** <http://www.caere.com>

Visioneer PaperPort

PaperPort on tarkoitettu elektronisten dokumenttien arkistoinnin apuvälineeksi. Ohjelman kautta voidaan paperimuodossa olevaa materiaalia skannata optisella kuvanlukijalla, tunnistaa teksti (OCR) ja tallentaa tulos halutussa muodossa. Käsiteltävän materiaalin tallennusmuoto voidaan valita sen perusteella, mitä ohjelmia koneelle on asennettu, mahdollisuuksia ovat esimerkiksi WordPerfect, Word, Excel, kuvankäsittelyohjelmat, Lotus Notes, Adobe Acrobat. Tiedostot voidaan myös lähettää sähköpostina tai faxilla.

Sivut voidaan skannata yhtenä nippuna, kaksipuoleiset sivut on skannattava kahdesti. Ohjelma tekee skannatuista sivuista käyttöliittymässä näkyvän pinon. Kaksipuolisten sivujen pinot voidaan yhdistää yhdeksi pinoksi. Tämän jälkeen pino raahataan sen ohjelman ikonin päälle, millä dokumenttia halutaan jatkokäsitellä. OCR tapahtuu tässä vaiheessa jonkin tähän tarkoitukseen soveltuvan ohjelman (esimerkiksi OmniPage) avulla ja lopuksi dokumentti aukeaa valitussa ohjelmassa jatkokäsittelyä varten.

Ohjelmassa on mahdollista yhdistellä skannattuja elementtejä sekä tuoda dokumenttiin ulkopuolisia osia eli ohjelmaa voidaan pitää myös jonkinasteisena julkaisuohjelmana.

Laitteistovaatimukset:

- 386-prosessori
- Windows 3.1
- 4 MB keskusmuistia
- 1 MB kiintolevytilaa
- Lisätietoja: <http://www.visioneer.com>

4.7 Tekstinkäsittely- ja julkaisuohjelmista

Tavallisin työkalu verkkojulkaisujen lähdemateriaalina käytettävien dokumenttien tuottamisessa lienee edelleenkin tekstinkäsittely- tai julkaisuohjelma. Monien tekstinkäsittely- ja julkaisuohjelmien uusimmissa versioissa on huomioitu verkkojulkaisemisen tarpeet siten, että ohjelmasta voidaan tuottaa suoraan jotain verkkojulkaisemiseen sopivaa tiedostomuotoa, esimerkiksi HTML:ää tai PDF:ää, *Save As...*, *Export* tms. komentoa käyttämällä. Jos kirjoittaja käyttää johdonmukaisesti tekstinkäsittely- tai julkaisuohjelman tyylimäärittelyitä ja sijoittaa kuvat ja kaa-viot osaksi julkaisua, voidaan julkaisun verkkoon sopiva elektroninen muoto tuottaa myös erilaisilla konverttereilla tai apuohjelmilla.

Microsoft Word

Laitteistovaatimukset (Windows)

Word 7.0	Word 97
<ul style="list-style-type: none">• Vähintään 386DX-prosessori (suositellaan 486-prosessoria)• Windows 95 tai Windows NT 3.51 (tai uudempi)• 6 MB keskusmuistia (Windows 95), 16 MB (NT)• 8 MB kiintolevytilaa (suppea asennus, 16 MB tyypillinen asennus ja 35 MB täysasennus)• Levykeasema	<ul style="list-style-type: none">• Vähintään 486-prosessori• Windows 95 tai NT 3.51 + Service Pack 5 (tai uudempi)• 8 MB keskusmuistia (Windows 95), 16 MB (NT)• 20-60 MB kiintolevytilaa (tyypillinen asennus 46 MB)• CD-ROM asema

PDF-dokumentin tuottaminen

Jos tietokoneeseen on asennettu Acrobat-ohjelmistot, PDF-dokumentin tuottaminen onnistuu Wordistä hyvin helposti. Tämä kuitenkin edellyttää, että kaikki dokumentissa olevat osat (kuvat, taulukot jne.) ovat sijoitettuna Word-dokumenttiin, ts. sanoen kaikki mikä näkyy ruudulla tuostuu myös PDF-dokumenttiin.

Kirjanmerkit voidaan luoda automaattisesti otsikkotyylilien perusteella käyttäen Wordiin saatavilla olevia mallipohjia (Document Template) ja niissä olevia makroja. Mallipohjia löytyy verkosta, esimerkiksi *WordMark* (<http://www.brookhouse.co.uk/company/markone.htm>).

HTML-dokumentin tuottaminen

HTML-dokumenttien tuottamiseen Word-tiedostoista on olemassa Microsoftin oma apuohjelma *Internet Assistant*, joka on kopioitavissa ilmaiseksi osoitteesta <http://www.microsoft.com/msword/internet/ia>¹⁴.

HTML-käännös perustuu alkuperäisessä Word-dokumentissa käytettyihin tyylimääriytyksiin, esimerkiksi *Heading 1* tyyllillä merkitty teksti ympäröidään HTML-tunnisteilla **<H1>** ja **</H1>**. Word-dokumentissa olevat kuvat käännetään GIF-muotoon. HTML-käännöksen onnistuminen edellyttää, että alkuperäisessä dokumentissa on käytetty johdonmukaisesti Wordin tyylimääriytyksiä (heading 1 jne.). Käännös ei kuitenkaan useimmissa tapauksissa ole täydellinen, eikä sulje kokonaan pois manuaalista HTML-koodausta.

Word dokumenttien kääntämiseen HTML-muotoon voidaan käyttää myös esimerkiksi *HotMetal Pro 4 HTML-editoria* ja monia shareware-konverttereita (ks. esim. http://www.w3.org/Tools/Word_proc_filters.html#MSWord).

¹⁴Vastaavat Microsoftin apuohjelmat HTML-dokumenttien tuottamiseen löytyvät myös esimerkiksi Exceliin (<http://www.microsoft.com/msoffice/msexcel/internet/ia/>) ja PowerPointiin (<http://www.microsoft.com/msoffice/mspowerpoint/internet/ia/>).

SGML-dokumenttien tuottaminen

SGML-dokumenttien tuottamiseen Wordillä voidaan käyttää esimerkiksi Microstarin *Near&Far*-ohjelmaa (http://www.microstar.com/ta_index.html). Ohjelmalla voidaan suunnitella DTD graafisena, puumaisena esityksenä ja luoda DTD:hen perustuvia SGML-dokumentteja.

Dokumenttia muokataan normaalissa Wordin WYSIWYG¹⁵ -näyttötilassa. Sijainti dokumentin rakenteessa näkyy graafisessa rakennepuussa, mikä helpottaa dokumentin rakenteen ymmärtämistä ja elementtien oikeellista sijoittamista. Ohjelmalla voidaan myös tarkistaa dokumentin oikeellisuus suhteessa valittuun DTD:hen.

Myös Microsoftilla on oma Wordin SGML-editori, *Microsoft SGML Author for Word*, jossa Wordin tyylit muutetaan SGML-tunnisteiksi. Lisätietoja:

<http://www.microsoft.com/syspro/technet/desk/office/word/SGMLtool/backgrnd.htm>.

Lisätietoja: http://www.microsoft.com/products/prodref/188_ov.htm.

Adobe PageMaker

Laitteistovaatimukset (PageMaker 6.0)

Windows	Macintosh
<ul style="list-style-type: none">• 486-prosessori• Windows 95 tai Windows 3.1 enhanced-tilassa• 8 MB keskusmuistia (Windows 95) tai 10 MB (Windows 3.1)• 24 MB levytilaa	<ul style="list-style-type: none">• Vähintään 68030 prosessori• System 7.1 tai uudempi• 16 MB keskusmuistia• 20 MB kiintolevytilaa

Laitteistovaatimukset (PageMaker 6.5)

Windows	Macintosh
<ul style="list-style-type: none">• 486-prosessori (suositellaan Pentium)• Windows95 tai NT 4.0• 8 MB keskusmuistia (16 MB NT)• 26 MB kiintolevytilaa• CD-ROM asema	<ul style="list-style-type: none">• Vähintään 68030 prosessori (suositellaan PowerMac)• System 7.1 tai uudempi• 6 MB keskusmuistia (9 MB PowerMac)• 26 MB kiintolevytilaa• CD-ROM asema

¹⁵ What You See Is What You Get, dokumentti näkyy muotoiltuna ruudulla.

PDF-dokumentin tuottaminen

PageMakerissa on valikko PDF-dokumenttien tuottamiseen, ”Luo Adobe PDF-tiedosto...”. PDF-dokumentti luodaan kuitenkin Distillerin kautta, joten sen on oltava asennettuna. PageMaker 6.5 mukana tulee myös Distiller-ohjelma, joten sitä ei tarvitse erikseen hankkia.

PageMakerin etu verrattuna esimerkiksi Wordiin PDF-dokumenttien tuottamisessa on, että PageMaker-dokumentin tyyleihin perustuvan sisällysluettelon perusteella voidaan luoda automaattisesti kirjanmerkit PDF-dokumenttiin. Myös varsinainen sisällysluetteloteksti linkitetään vastaaviin kohtiin PDF-dokumentissa.

HTML-dokumentin tuottaminen

PageMakerissa on HTML-editori-apuohjelma, jonka avulla voidaan määrittellä PageMaker-tyyliin vastaavuudet HTML-tunnisteisiin, sekä lisätä linkkejä.

Lisätietoja: <http://www.adobe.com/prodindex/pagemaker/main.html>

FrameMaker +SGML 5.5

FrameMaker +SGML on Adoben SGML-ominaisuuksilla varustettu julkaisuohjelma, jolla voidaan tuottaa SGML-dokumentteja ulkoasumääritysten mukaisessa WYSIWYG-tilassa. Dokumenttia voidaan tarkastella myös puumaisena rakennekuvauksena.

Käytettyyn DTD:hen voidaan yhdistää monta erilaista tyylimäärittelyä dokumentin ulkoasun määrittelemiseksi.

FrameMaker dokumenttien kääntämiseen HTML-muotoon on olemassa lukuisia konverttereita, http://www.w3.org/Tools/Word_proc_filters.html#Framemaker.

Lisätietoja: <http://www.adobe.com/prodindex/framemaker/prodinfoSGML.html>

Laitteistovaatimukset

Windows	Macintosh
<ul style="list-style-type: none">• Vähintään 486-prosessori• Windows95 tai NT 4.0• 24 MB keskusmuistia• 22-90 MB kiintolevytilaa• PostScript- tai PCL-tulostin• Adobe Type Manager	<ul style="list-style-type: none">• 68040-prosessori tai parempi (suositellaan PowerMac)• System 7.5 tai uudempi• 24 MB keskusmuistia• 55-90 MB kiintolevytilaa• Adobe Type Manager

4.8 Kuvankäsittely- ja piirto-ohjelmista

Adobe PhotoShop 4.0

Adobe PhotoShop on ammattimaisen kuvankäsittelyn perusohjelma. Pelkkiin kuvaformaateista toiseen muutokseen PhotoShop ei ole paras mahdollinen valinta, sillä tuettuja kuvaformaatteja on huomattavasti vähemmän kuin esimerkiksi Paint Shop Pro:ssa. PhotoShopin oma tiedostoformaatti PSD ei sovellu verkkojulkaisemiseen, koska selaimet eivät tue sitä.

Lisätietoja: <http://www.adobe.com/prodindex/photoshop/main.html>

Laitteistovaatimukset

Windows	Macintosh
<ul style="list-style-type: none">• Vähintään 386-prosessori• Windows 3.1, Windows 95 tai Windows NT 3.51 (tai uudempi)• 16 MB keskusmuistia (suositellaan 32 MB)• 25 MB levytilaa• 8-bittinen näytönohjainkortti (256 väriä) tai parempi• CD-ROM-asema	<ul style="list-style-type: none">• 68030 prosessori tai parempi (suositellaan PowerMac)• System 7.1 tai uudempi• 16 MB keskusmuistia (suositellaan 32 MB)• 25 MB kiintolevytilaa• 8-bittinen näytönohjainkortti (256 väriä) tai parempi• CD-ROM asema

Corel PhotoPaint 7

Corel PhotoPaint on samantyyppinen kuvankäsittelyohjelma kuin PhotoShop, ja soveltuu vaativiin kuvankäsittelytehtäviin.

Lisätietoja: <http://www.corel.com/products/graphicsandpublishing/photopaint7/index.htm>

Laitteistovaatimukset

- 486-prosessori (suositellaan Pentium)
- Windows 95 tai NT 4.0
- 16 MB keskusmuistia (suositellaan 32 MB)
- 42 MB kiintolevytilaa
- CD-ROM-asema

Paint Shop Pro ja LView Pro

Paint Shop Pro (JASC Inc., <http://www.jasc.com/psp.html>) on shareware-ohjelma, joka lukee ja kirjoittaa 35 kuvaformaattia. Paint Shop Pro:ssa on myös erittäin laajat kuvankäsittelyominaisuudet. Ohjelmassa on myös mahdollisuus eräajotyypiseen kuvaformaattista toiseen muuntamiseen. Ohjelma on saatavissa ainoastaan Windows-ympäristöön (3.1, 95 ja NT).

Samoin kuin Paint Shop Pro, **LView Pro** (<http://www.lview.com/>) on myös shareware-ohjelma joka soveltuu yleisimpiin kuvaformaattista toiseen muunnoksiin. Kuvankäsittelyominaisuudet ovat suppeammat kuin Paint Shop Pro:ssa.

Corel Draw 7

PhotoShop ja Corel PhotoPaint on tarkoitettu bittikarttamuodossa olevien kuvien käsittelyyn. Piirto-ohjelmilla, kuten esimerkiksi Corel Drawilla tuotetuissa kuvissa kuva muodostuu graafisista objekteista, esimerkiksi suorakaiteesta jonka vasen yläkulma sijaitsee koordinaateissa X1,Y1 ja oikea alakulma X2, Y2. Kuva muodostetaan piirtämällä tarvittavat viivat näiden pisteiden kautta. Piirto-ohjelmilla tuotettujen kuvien etu on kuvien hyvä skaalattavuus, koska kuvakoon muuttuessa objekti piirretään uudestaan koordinaattien perusteella.

Corel Draw:ssa objektit, esimerkiksi tekstit tai suorakaiteet yms., voidaan täyttää myös bitmap -kuvilla.

Corel Draw:n oma kuvaformaatti CDR ei sovellu sellaisenaan verkkojulkaisemiseen, vaan vaatii joko kuvan tallettamista bitmap-muodossa (GIF, JPG) tai kuvan tallettamisen CMX-muodossa (Corel Presentation Exchange) ja soveltuvan plug-inin¹⁶ CMX-muodossa olevien kuvien katseluun selaimen ruudulla.

Lisätietoja: <http://www.corel.com/products/graphicsandpublishing/draw7/index.htm>

Laitteistovaatimukset

- 486DX-prosessori (suositellaan Pentium 120MHz)
- Windows 95 tai NT 4.0
- 16 MB keskusmuistia (suositellaan 32 MB)
- 40 MB kiintolevytilaa
- CD-ROM-asema

¹⁶ <ftp://ftp.corel.com/pub/misc/Corelcmx.exe>

Yhteenveto: Kuvankäsittely- ja piirto-ohjelmien soveltuvuus

	PhotoShop	Corel PhotoPaint	Paint Shop Pro	LView Pro	Corel Draw
Kuvaformaattien kon- vertointi	(✓)	(✓)	✓	✓	
Vaativa kuvankäsittely	✓	✓	(✓)		
Kuvan läpinäkyvyyssi- formaation tallettaminen (GIF89a)	✓		✓	✓	
Viivapiirrokset					✓

5. ELEKTRAN AINEISTOT JA INTERNETIN KAUTTA SAATAVILLA OLEVAT KOTIMAISET VERKKOJULKAISUT

5.1 ELEKTRAN artikkeliaineiston valintaperusteet

ELEKTRAN artikkelipilottia suunniteltaessa aineiston valintakriteereinä käytettiin lehden tieteellistä tai kulttuurillista alaa ja laadun varmistavaa julkaisijatahoa. Aineistossa ei haluttu rajoittaa mihinkään tiettyyn tieteenalaan, vaan mukaan pyrittiin saamaan aineistoa kattavasti eri aloilta, niin että voitaisiin selvittää julkaisijoiden verkkokehitystilanne ja simuloida todellista tieteellistä julkaisu-ympäristöä. Kokeiluun haluttiin sisällyttää myös jo valmiiksi Internet-jakelussa olevia lehtiä, niin että voitaisiin kokeilla linkittämistä.

Koska kokeilun lähtökohtana oli, että jakelu tapahtuu olemassa olevien yhteisluettelotietokantojen kautta, valittiin etupäässä lehtiä, jotka jo valmiiksi sisältyvät kansalliseen artikkeliviitteiden tietokantaan ARTOon. Siihen indeksoitavien kotimaisten lehtien määrä oli vuoden 1997 lopussa yli 1000 ja indeksointityöhön osallistuu 38 tiedontuottajaa, jotka ovat erikoisalojen kirjastoja tai tutkimuslaitoksia. Tiedontuottajina toimivat asiantuntijaorganisaatiot takaavat tietokannan korkean laadun niin aineiston valinnan kuin syvällisen sisällönkuvailun avulla. Hajautettu tuotanto-organisaatio vaatii todennäköisesti keskitettyä enemmän resursseja, mutta syntyvä lopputuote on tasokkaampi. Lehtiä Artossa on kaikilta tieteenaloilta ihmislääketiedettä lukuunottamatta ja mukana on myös harraste- ja aluelehtiä. Artikkeliviitteiden määrä vuoden 1997 lopussa oli noin 200 000.

Toisena artikkelipilotin lähestymissuuntana olivat Tieteellisten seurain valtuuskunnan (TSV) jäsenjärjestöjen julkaisemat lehdet. Niiden osalta kontaktit julkaisijoihin voitiin solmia Tieteellisten seurain valtuuskunnan välityksellä. Tieteellisillä seuroilla on omat julkaisutoimikuntansa jotka vastaavat julkaisujen laadusta ja toisaalta vakiintunut toimitus, joka viimeistelee julkaisun painovalmiiksi. TSV:n jäsenjärjestöt julkaisevat yhteensä noin 100 tieteellistä lehteä, joista 64 sisältyy ARTOon indeksoitaviin julkaisuihin. Lisäksi seurat julkaisevat vuosikirjoja ja monografiatyypisiä julkaisuja.

Lehtivalikoimaa laajennettiin ottamalla mukaan myös Helsingin yliopiston kirjaston omia julkaisuja (Helsingin yliopiston kirjaston tiedotuslehti ja Tietolinja), Suomen hammaslääkäriliiton Hammaslääkäri-lehti sekä Suomen Maantieteellisen Seuran julkaisema Terra. Ne indeksoidaan jo valmiiksi ARTOon.

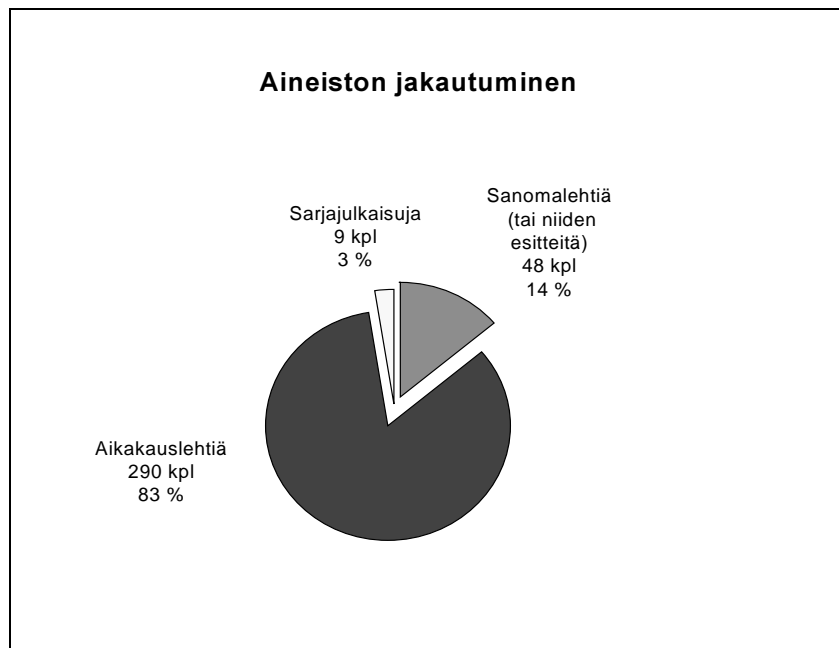
Artikkelivolyymin kasvattamiseksi ja linkittämisen kokeilemiseksi mukaan otettiin myös Yliopisto-lehti, joka ilmestyi elektronisena ja sen vapaa verkkojakelu oli toteutettu. Yliopisto-lehteä ei kuitenkaan ole valmiiksi indeksoitu ARTOon ja ELEKTRAN pilotissa vuoden 1997 aikana siitä on luetteloitu yhteensä 191 artikkelia. Syksyllä 1997 lehti muutti elektronista julkaisuun entistä enemmän ajankohtaisaineistoa ja vähemmän kokotekstiartikkeleita sisältäväksi. Nykyisin elektronisen julkaisun yhdestä numerosta löytyy kokotekstiartikkeleita enää vain noin 30% painetun artikkeleiden määrästä.

5.2 Selvitys kotimaisista verkkolehdistä

Internetin kautta saatavia verkkojulkaisuja selvitettiin VTT Tietopalvelun ja Helsingin yliopiston kirjaston ylläpitämien lehtilistojen pohjalta. VTT on tehnyt valikoidun luettelon pohjoismai-

sista verkkolehdistä, ja luettelossa on noin 60 suomalaista lehteä. Kaikkiin ei kuitenkaan vielä sisälly kattavasti artikkeleja, vaan lehdestä voi olla pelkät esittely- ja tilaustiedot. Lista löytyy osoitteesta <http://www.vtt.fi/inf/nordep/projects/webpilot/journals/>.

Helsingin yliopiston kirjaston kausijulkaisulistan tavoitteena on osoitetietojen välittämisen lisäksi muodostaa kokonaisnäkemyks Suomessa julkaistavista kotimaisista verkkolehdistä sekä niiden lukumäärästä ja nimekkeiden kasvuvauhdista. Lehtilistan kokoamisessa on verkkolehtiin sovellettu kolmea valintakriteeriä: 1) niissä tulee olla eri numeroita kokoava yhteisnimeke, 2) jatkuva numerointijärjestelmä ja 3) lehden pääartikkeleiden on oltava verkkojakelussa kokotekstinä. Listalla oli vuoden 1997 lopussa 347 lehteä. Lehtilista löytyy osoitteesta http://www.helsinki.fi/~hyk_ml/lehdet.htm.



Kuva 1. Lehtiaineiston tyypit

Lehtilistaa analysoimalla ilmeni, että vuoden 1997 lopussa kotimaisia verkkolehtiä ilmestyi yhteensä 347. Niistä 83 % oli aikakausslehtiä, 14 % sanomalehtiä ja 3 % sarjajulkaisuja.



Kuva 2. Verkkolehtien rinnakkaisversiot

Lehdistä 55 % (189) oli rinnakkaisjulkaisuja painetun version kanssa, 25 % (92) paperilehden liitteitä, mainoksia tai esitteitä ja 19% (66) vain elektronisessa muodossa ilmestyviä lehtiä. Tieteellisten seurain jäsenorganisaatiot julkaisemista noin 80 lehdestä kaksi ilmestyi painetun lisäksi myös elektronisena rinnakkaisversiona (*Skeptikko* ja *Tähdet ja avaruus*) ja samoin ilmestyi yksi monografiasarja (*Annales Academiae scientiarum Fennicae. Series A 1, Mathematica*). Yksi lehti, *Elektroloristi*, ilmestyi pelkästään elektronisena.

Projektin valintakriteereiden mukaisia verkkojakelussa olevia dokumentteja löytyi kuitenkin niin vähän, että päädyttiin tekemään erillinen kysely Tieteellisten seurain valtuuskunnan jäsenorganisaatiolle heidän verkkojulkaisutilanteestaan ja halukkuudestaan antaa aineistoaan ELEKTRA-projektin käyttöön.

ELEKTRA-projektin ansiosta puolet tieteellisten seurojen lehtijulkaisijoista on saanut kokemusta elektronisesta julkaisutoiminnasta ja aineistonsa verkkojakelusta kokeilun aikana. Tämä on enemmän kuin tehdyn kyselyn perusteella voitiin odottaa. Valtaosa aineistosta on jouduttu digitoimaan, mutta jatkohankkeessa on tarkoitus ohjeistaa julkaisijoita tuottamaan aikaisempaa enemmän aineistoaan suoraan elektroniseen muotoon. Tämä alentaa palvelun tuotantokustannuksia ja parantaa aineiston laatua.

5.3 Kotimaisia lehtilistoja Internetissä

LISKOn aiheenmukainen lehtilista (<http://www.lisko oulu.fi/lisko/viestinta.htm#lehdet>)

Telen Keskuskadun uutislehtilista (<http://www.inet.fi/keskuskatu/uutiset.htm>)

Kaapelisolmun lehtihakemisto (<http://www.kaapeli.fi/suositt/lehtia.html>)

RadioCityn Linkkioppaan lehtilista (<http://www.linkkiopas.fi/z-lehdet.html>)

Helsingin yliopiston lehtilista (<http://www.helsinki.fi/lehdet/>)

Tampereen yliopiston lehti- (<http://www.uta.fi/verkkopalvelut/#uutiset>) **ja sarjakuvalista**
(<http://www.uta.fi/yhteydet/sarjikkeet.html>)

VTT:n Tietopalvelun pohjoismaisten WWW-lehtien valikoima
(<http://www.vtt.fi/inf/nordep/projects/webpilot/journals/>)

Vanhan ylioppilastalon WWW-sivuilla oleva lista Sähköisiä lehtiä Suomessa ja maailmalla
(<http://www.vanha.fi/tekstit/e-news.htm>)

Sanomalehtien Liiton Linkit verkkosanomalehtiin (<http://www.sanomalehdet.fi/linkit.html>)

Hämeenlinnan kaupunginkirjaston Makupalojen lehtihakemisto
(<http://www.htk.fi/kirjasto/viestin2.htm>)

Alku - Suomen linkkiaseman lehtilista (http://alku.hrsk.edu.fi/viihde_ja_vapaa-aika/ajankohtaislehdet.html)

Serverin hakemisto Viihde ja kulttuuri/Lehdet (<http://www.serveri.net/hakemisto/1/6/>)

Journalismi verkossa-sivujen lehtilista (<http://www.wysiwyg.fi/jv/>)

Oulun kaupunginkirjaston Elektronisen lehtisalin Aakkosellinen lehtiluettelo
(<http://www.ouka.fi/kirjasto/aakkole.htm>)

Listalla on runsaasti on kulttuuriin, talouteen, tietotekniikkaan, uskontoon ja erilaisiin harrastuksiin liittyvää aineistoa, joka kiinnostanee laajempaakin lukijakuntaa. Valtaosa listalla olevista aikakauslehdistä on opiskelijalehtiä sekä erilaisten harrastus-, uskonnollisten ja aatteellisten yhdistysten eli asiastaan innostuneiden suhteellisten pienten ryhmien julkaisuja. Varsinaisia tieteellisiä lehtiä listalta löytyy vain kymmenkuntaa nimekettä (EJBO, Elektroloristi, Poliitiikka, Suomen lääkärilehti, Suomen Proviisorilehti, Terra, Tieteessä tapahtuu, Virittäjä ja Yliopisto).

ELEKTRAan on saatu verkkolehdistältä mukaan suurin osa tieteellisistä lehdistä. Tietopalvelualan lehdistä mukana ovat Tiedokas ja Tietolinja. Poliitiikalla, Terralla ja Virittäjällä ei ole ollut aikaisemmin Internetissä kokotekstiartikkeleita, vaan lehden yleisesittelyn lisäksi tarjolla on ollut sisällysluetteloita ja abstrakteja.

ELEKTRA-kokeilun aineiston valintakriteerit, tieteellisyys ja ARTO-luettelointi ovat täyttyneet hyvin. Elektronisten verkkojulkaisujen valmiiksi digitoidusta ja pääosin ARTOon luetteloidusta artikkeliaineistosta on pystytty luomaan suhteellisen nopeasti pohja ELEKTRA-tietokannalle. Elektroloristin, Tieteessä tapahtuu-, Tietolinja- ja Yliopisto-lehtien artikkelit muodostavat tällä hetkellä noin neljäsosan luetteloidusta aineistosta. Projektin alkuvaiheessa kustantajilta ja julkaisijoilta saatua aineistoa on saatu oletettua enemmän perinteisessä paperimuodossa. Paperilehtien osuus tulee pienenevän, kun materiaalia saadaan vastaisuudessa yhteistyön vakiinnuttua yhä enemmän digitaalisessa muodossa.

5.4 Kysely tieteellisten julkaisijoiden verkkokehitystilanteesta ja kiinnostuksesta osallistua ELEKTRA-kokeiluun

Tieteellisten julkaisijoiden verkkokehitystilanteesta kartoitettiin vuoden 1997 alussa lomakekyselynä. Kohderyhmäksi valittiin Tieteellisten seurain valtuuskuntaan (TSV) kuuluvat järjestöt, jotka olivat myös tieteellisten aikakauslehtien kustantajia tai julkaisijoita. Kyselyn tarkoituksena oli selvittää julkaisijoiden valmiuksia elektroniseen julkaisemiseen kyselyajankohtana ja lähitulevaisuudessa sekä kiinnostusta osallistua ELEKTRA-hankkeeseen testiaineiston tuottajana. Kysely lähetettiin noin sadalle TSV:n jäsenjärjestölle ja vastauksia saatiin 40.

Kyselyyn vastasi 40 julkaisijaa. Heistä 30 (75%) oli kiinnostuneista kokeilusta ja vain neljä suhtautui hankkeeseen epäroivästi. Aineistoaan verkkojakeluun halusi 29 (73%) vastaajista, arkistoitavaksi HYK:in palvelimelle 25 (63%) ja digitoitavan 19 (48%). Valtaosa julkaisijoista on toimittanut lehtensä tekstinkäsittelytiedostojen (32) tai sivuntaitto-ohjelmien (21) pohjalta. Taitto on tehty kokonaan kirjepainossa 11 julkaisijalla ja elektroniseen muotoon aineistoa on tuottanut vain 3 julkaisijaa. Internetistä löytyi kotisivut kolmannekselta vastaajista. Joulukuussa 1997 aineistoa toimitti jo 36 kustantajaa (90 % vastaajista). Kustantajakyselyn luvuissa eivät ole mukana Elektroloristia lukuunottamatta lehdet, joiden artikkelit sijaitsevat julkaisijan omalla palvelimella kuten Annales Botanici Fennici, Annales Zoologici Fennici, Tietolinja, Tieteessä tapahtuu ja Yliopisto.

Aineiston toimittamiseen liittyvistä yksityiskohdista halusi neuvotella 28 (70%) vastaajaa, viisi (13%) ei pitänyt neuvottelua tarpeellisena ja kysymykseen jätti vastaamatta seitsemän (17%). Projektin edetessä on tehty konkreettista yhteistyötä kokeiluun osallistuvien julkaisijoiden kanssa. Yhteyksiä on hoidettu kirjein, sähköpostitse ja puhelimitse sekä järjestämällä tiedotustilaisuuksia projektin etenemisestä. Kustantajat ovat halutessaan voineet kokeilla ELEKTRA-palvelun käyttöä omilta työkoneiltaan käsin. He ovat myös voineet vaikuttaa mielipiteillään palvelun edelleen kehittämiseen.

ELEKTRAn palveluista kiinnostavimpana pidettiin aineiston asettamista asiakkaiden ulottuville verkkokäyttöön, artikkelien kuvailua ja luettelointia tietokantaan sekä Internetissä toimivan hakuohjelman käyttömahdollisuutta. Tällaisesta palvelusta oli kiinnostunut 29 vastaajaa (73%).

Aineiston arkistointi ja säilyttäminen Helsingin yliopiston kirjaston palvelimella kiinnosti 25 julkaisijaa (63%). Aineiston digitointi skannaamalla sekä tekstinkäsittely- ja muiden tiedostojen konvertointi verkkoon sopivaan tiedostomuotoon kiinnosti 19 vastaajaa (48%). Muita palveluita, kuten tiedotusta projektin etenemisestä, tiivistelmien digitoimista, tekijänoikeussopimusten hoitamista sekä koulutuksen järjestämistä toivoi viisi kustantajaa (13%).

Kokeiluun osallistuminen jollain muulla tavalla kuin aineistoa toimittamalla kiinnosti kuutta vastaajaa (15%). Vaihtoehtoinen osallistumistapa on artikkelien tallentaminen omalle palvelimelle kuten ennenkin, jolloin ELEKTRA-projektin tehtävänä on artikkelitietueiden poimiminen ARTOsta, linkkien luominen tietueiden ja artikkeleiden välille sekä tekijänoikeusasioiden hoitaminen.

Valtaosa kyselyyn vastanneista tieteellisistä lehdistä toimittaa julkaisunsa tekstitiedostojen pohjalta. Peräti 32 (80%) lehteä toimitetaan tekstinkäsittelyohjelmilla (esim. Word, WP) tuotettujen tiedostojen pohjalta. Sivuntaitto-ohjelmia (esim. Ventura, Page Maker) käytetään lisäksi 21 lehdessä (53%). Elektronisen julkaisemisen tiedostomuotoja (HTML, PDF) on käyttänyt vain 3 kustantajaa (8%). Taitto tehdään kokonaan kirjepainossa 11 vastaajan (28%) lehdessä.

Muita kuin edellä mainittuja ohjelmia kuten kuvankäsittelyohjelmia (esim. Photoshop, Paint Shop Pro), on käyttänyt kuusi lehdentekijää (15%). Kuvitusta on mukana valtaosassa lehdistä eli 32:ssa (80%). Originaalikuivat käsitellään ja liitetään julkaisuun kirjapainossa 20 tapauksessa (50%) ja kuvat liitetään elektronisessa muodossa tekstinkäsittely- tai taitto-ohjelmilla tuotettuihin tiedostoihin 17 tapauksessa (43%).

Mainoksia, esittelyjä ja julkaisun yhteystiedot sisältävä kotisivu löytyy Internetistä 14:ltä (35%) kyselyyn vastanneelta julkaisulta, jotka ovat AFinLAN vuosikirja, Agricultural and Food Science in Finland, Elektroloristi, Futura, Kasvatus, Musiikin suunta, Poliitikka, Savon Luonto, Sosiologia, Suo, Suomen eläinlääkärilehti, Terra, Tähdet ja avaruus ja Virittäjä. Sisällysluettelo löytyy 13 lehden (33%) sivuilta, artikkelitivistelmät kuudelta (15%), valikoituja artikkeleita kokotekstinä kolmelta lehdeltä (8%) ja kattavasti artikkeleita (kolmelta (8%) lehdeltä.

5.5 ELEKTRAN artikkelipilotissa mukana olevat julkaisut

ELEKTRAN pilotissa on artikkeleja yhteensä 37 lehdestä ja 1.3.1998 luetteloitujen artikkelien määrä oli noin 1300 ja digitoitujen noin 1500. Valtaosa aineistosta on vuosilta 1996 ja 1997. Jotkut lehdet, kuten Kasvatus, Metsätieteen aikakauskirja Folia Forestalia ja Silva Fennica ovat halunneet irtonumeromyyntinsä turvaamiseksi antaa testiin vain vanhempaa, vuoden 1995 aineistoa. Pääosin kuitenkin aineisto on ajantasaista ja saadaan jatkuvasti uuden numeron ilmestyessä.

Tieteenaloittain lehtiä on lähes kaikilta keskeisiltä aloilta. Erityisen kattavasti mukana ovat oikeustiede ja kasvatustieteet, joskaan kaikkiin oikeustieteellisiin julkaisuihin ei ole saatu tulos- tusoikeutta lainkaan. Myös metsätiede ja eläinlääketiede ovat olleet aktiivisesti mukana.

Julkaisun nimi	Artikkelit (kpl)	Vuosikerta	Tiedostomuoto
Lakimies	78	96-97	PDF
Yliopisto	191	96-97	HTML
Suomen eläinlääkärilehti	44	96-97	PDF
Arkhimedes	57	96-97	PDF
Tieteessä tapahtuu	82	96-97	HTML
HYK:in tiedotuslehti	73	96-97	HTML
Kasvatus	60	95	PDF
Aikuiskasvatus	58	97	PDF
Suomen kirkkohistoriallisen seuran vuosikirja	57	95-96	PDF
Gerontologia	53	96-97	PDF
Kirjallisuudentutkijain seuran vuosikirja	45	96	HTML
Genos	48	96-97	PDF

Luonnon tutkija	38	96	PDF
Terra	38	96-97	PDF
Suo	32	95-97	PDF
Ympäristöjuriikka	30	97	PDF, ei tulostusta
Politiikka	30	97	PDF
Science Studies	28	95-98	PDF
Geologi	28	96	PDF
Tietolinja	22	95-96	HTML ja PDF
Elektroloristi	22	96	HTML
Sosiologia	19	96-97	PDF
Agricultural and Food Science in Finland	18	96-97	PDF
Siivekäs	15	95-96	PDF
Savon Luonto	14	95-96	PDF
AfinLan vuosikirja	13	96	PDF
Metsätieteen aikakauskirja Folia Forestalia	9	95	HTML ja PDF
Suomen Antropologi	8	96	HTML
Lounais-Hämeen Luonto	8	95-96	PDF
Källan	8	96	HTML
Sukutieto	8	97	PDF
Annales Botanici Fennici	7	96	PDF, ei tulostusta
Geophysica	6	97	PDF
Musiikin suunta	6	97	PDF
Tiedokas	5	96	PDF, ei ARTOssa
Silva Fennica	4	95	PDF
Acta Forestalia Fennica	1	96	PDF
YHTEENSÄ (1.3.1998)	1263		

5.6 Väitöskirjojen elektroninen julkaiseminen Oulun yliopistossa

Väitöskirjojen pääasiallinen julkaisukanava on Oulun yliopistossa *Acta Universitatis Ouluensis* -sarja. Oulun yliopiston kirjaston tehtävänä ELEKTRAssa oli luoda julkaisumalli väitöskirjojen elektroniselle julkaisemiselle ja luontevaa oli keskittyä ACTA -sarjoissa julkaistaviin väitöksiin. Malli on yleistettävissä muihinkin sarjoihin ja julkaisuihin.

Elektroninen julkaiseminen voidaan määritellä useilla tavoilla. Tässä elektronisella julkaisemisella tarkoitetaan julkaisuprosessia, jossa dokumentti on koko elinkaarensa ajan käytettävissä elektronisessa muodossa. Usein elektronisella julkaisemisella tarkoitetaan prosessia, jossa dokumentti on ainoastaan jossain vaiheessa elektronisessa muodossa, lopullinen dokumentti on usein olemassa ainoastaan painetussa muodossa.

Näiden lähestymistapojen ero on tärkeä. Jos tavoitteena on ainoastaan painetun dokumentin tuottaminen, ei käytetyillä menetelmillä ja tiedostomuodoilla ole lopputuloksen kannalta ratkaisevaa merkitystä. Tilanne on kuitenkin toinen, kun halutaan dokumentin olevan käytettävissä elektronisessa muodossa koko elinkaarensa ajan. Väitöskirjojen kohdalla voidaan puhua jopa ”ikuisesta” elinkaaresta.

Elinkaaren varmistaminen edellyttää käytettäviltä teknisiltä ratkaisuilta paljon. Tyypillisesti erilaisten ohjelmistojen ja laitteistojärjestelmien elinkaari on huomattavasti lyhyempi kuin niillä tehtyjen dokumenttien elinkaari. Tästä johtuen keskeinen vaatimus käytettäville tiedostomuodoille on laitteisto-, käyttöjärjestelmä- ja ohjelmistoriippumattomuus, joka varmistaa dokumentin käyttökelpoisuuden vielä silloinkin, kun dokumentin tuottamiseen käytetyt järjestelmät on vanhentuneina hylätty. Seuraavassa kuvataan dokumentin vaiheet käsikirjoituksesta valmiiksi julkaisuksi.

1. Vaihe: Ohjeistus julkaisusarjoille.

Väitöskirjat julkaistaan *Acta Universitatis Ouluensis*-sarjassa. Yliopiston julkaisutoimikunta on yhdessä kirjaston kanssa laatinut uudistetut ohjeet sarjajulkaisuille, joissa on huomioitu elektronisen julkaisemisen vaatimukset. Keskeisimmät seikat ovat seuraavat: elektroninen sarjajulkaisu saa oman ISSN-tunnuksen ja sillä on oltava oma kotisivu.

2. Vaihe: Ohjeistus sarjassa julkaistaville käsikirjoitukselle

ACTA-sarjan ohjeet kirjoittajille uudistetaan. Tarkoituksena on tehdä sarjan yleisohjeet, joita voidaan tarvittaessa muokata eri tieteenalojen käytäntöihin sopiviksi yhdessä eri tiedekuntien ja laitosten kanssa. Tämä on pitkä ja jatkuva prosessi ja yhteistyö laitosten kanssa onkin välttämätöntä tarkkojen ohjelmistokohtaisten ohjeiden ja mallipohjien tuottamiseksi.

Alkuvaiheessa tullaan suosittamaan käsikirjoitusten viimeistelyn suorittamista yliopistopainon toimesta, jolloin voidaan varmistaa dokumenttien sopiminen sekä painamiseen että elektroniseen julkaisemiseen.

3. Vaihe: Julkaisusopimus

Kirjoittajan saatua julkaisuluvan tehdään julkaisusopimus, jossa on otettava huomioon myös elektroniseen julkaisemiseen liittyvät kysymykset. Sopimuslomakkeet ovat saatavissa elektroni-

sessä muodossa WWW:n kautta ja osa julkaisun tiedoista saadaan jo tässä vaiheessa elektronisessa muodossa sarjatoimittajan käyttöön.

4. Vaihe: Julkaisuun liittyvät sarja- ja muut tiedot

Sarjan toimitussihteeri tekee julkaisun kansi- ja nimiöisivut. Koska julkaisujen tietoja tullaan käyttämään useissa erilaisissa yhteyksissä, esimerkkinä uutuuksiluettelot ja julkaisuluettelot, kehitetään tietokantapohjainen järjestelmä julkaisutietojen hallintaan. Tietokantaratkaisu mahdollistaa julkaisun kansi- ja nimiöisivujen automaattisen taittamisen, samoin julkaisun kotisivun automaattisen generoinnin ja ylläpidon.

5. Vaihe: Julkaisun arkistointi ja jakelu

Valmis julkaisu luetteloidaan normaalisti ja painetun julkaisun osalta ei mitään muutosta vanhoihin menetelmiin tule. Elektronisen version jakelu tapahtuu kirjaston ylläpitämällä palvelimella ja julkaisun arkistointiin tarkoitettu versio siirretään omalle arkistopalvelimelle. Varsinaisena arkistoversiona pidetään kuitenkin toistaiseksi painettua julkaisua ja ainoastaan elektronisesta julkaisusta tulostetaan mahdollisuuksien mukaan paperinen arkistokappale.

Elektronisella julkaisemisella ja WWW-jakelulla voidaan huomattavasti parantaa julkaisun saatavuutta. Tietyn julkaisun löytäminen WWW:n valtavasta tarjonnasta voi kuitenkin olla ongelmallista, seuraavilla keinoilla pyritään helpottamaan tätä ongelmaa:

Julkaisulla on oma pysyvä kotisivu ja URL. Julkaisun URL on muotoa: *http://herkules.oulu.fi/"julkaisun ISBN-tunnus"/*. Julkaisun kotisivulla käytetään myös Dublin Core -metadata formaattia, joka helpottaa dokumentin löytymistä hakukoneiden avulla.

Pilotissa käytettyjä jakeluformaatteja on kaksi: HTML erityisesti nopeaan selailuun ja ruudulta lukemiseen, PDF dokumentin laadukkaaseen tulostamiseen. Tekniikan nopean kehityksen myötä tarvetta jakeluformaattien muuttamiseen ja lisäämiseen varmastikin ilmenee ja toivottavaa olisikin yhteen jakeluformaattiin pääseminen. Toistaiseksi tämä on kuitenkin teknisesti mahdotonta, mikäli halutaan mahdollistaa hyvä luettavuus ruudulta ja korkeatasoinen painojälki.

Standard Generalized Markup Language (SGML) ja sen uusi yksinkertaistettu versio eXtensible Markup Language (XML), ovat mahdollistavia tekniikoita, joiden avulla dokumentin linkaari on mahdollista turvata. SGML ja XML ovat metakieliä, joilla määritellään varsinaisia dokumentin merkkauškieliä. Tällaisia merkkauškieliä kutsutaan SGML:n sovelluksiksi, joista tunnetuin ja laajimmassa käytössä oleva on WWW:ssä käytettävä HyperText Markup Language (HTML).

SGML mahdollistaa käytännössä mielivaltaisen dokumentin mielivaltaisen merkkauksen, joten ymmärrettävästi standardi on laaja ja jopa vaikeaselkoinen. Oleellista on huomata, että dokumentit ja niiden käyttö määrittelee käytettävän SGML-sovelluksen monimutkaisuuden. Rakenteeltaan ja käytöltään yksinkertaiset dokumentit eivät edellytä monimutkaista sovellusta. Joissakin tapauksissa jopa HTML-pohjaiset ratkaisut voivat olla riittäviä. Dokumentin rakenteen manipuloitavuus ja dokumenteille asetettävien vaatimusten kasvaessa käytettävän sovelluksenkin monimutkaisuus luonnollisesti kasvaa.

Oulun yliopiston kirjaston tavoitteena oli perehtyä elektroniseen julkaisemiseen ja kehittää väitöskirjoille optimoitu julkaisujärjestelmä. Erityisesti perehdyttiin SGML:n käyttöön elektronisessa julkaisemisessa.

SGML-pohjaiseen järjestelmään siirtyminen yliopiston tapaisessa avoimessa ja hajautetussa ympäristössä on ongelmallista, koska jokaisella kirjoittajalla on mahdollisuus tuottaa tekstinsä parhaaksi katsomallaan tavalla ja ainoastaan erilaisia suosituksia tekstin tuottamiseen voidaan antaa. SGML-pohjaiseen järjestelmään tullaan todennäköisesti siirtymään vaiheittain. Aluksi siirrytään yliopistopainossa tapahtuvassa käsikirjoituksen viimeistelyssä, taitossa, SGML-pohjaiseen järjestelmään. Tällöin järjestelmä saadaan testattua ja viimeistelyä todellisilla käyttäjillä.

SGML:n ja erityisesti XML tulevat todennäköisesti muuttamaan huomattavasti dokumenttien verkkovälitystä. XML:n tavoitteena on geneerisen SGML:n käyttö WWW:ssä samalla tavalla, kun nykyisin käytetään HTML:ää. Eräs tärkeimmistä ja mielenkiintoisimmista XML:n sovelluksista on Mathematical Markup Language (MathML), jonka tavoitteena on mahdollistaa matemaattisen tekstin laite- ja ohjelmistoriippumaton käyttö. XML on kuitenkin vielä hyvin uusi standardi ja toistaiseksi sen sovellukset ja ohjelmistotarjonta ovat rajallisia. XML:n tukijoksi ovat ilmoittautuneet lähes kaikki merkittävät alan yritykset ja todennäköistä onkin sovellusten ja ohjelmistojen määrän nopea kasvu.

6. ELEKTRAN PILOTTIEN TEKNINEN YMPÄRISTÖ

6.1 Elektroninen julkaisu- ja digitointiympäristö

ELEKTRAN julkaisuympäristönä toimii PC-palvelin, jossa on 200 MHz:n Pentium-prosessori ja ATX Pentium Pro -emolevy. Koneessa on 64 Kt RAM-muistia ja 2 Gt:n SCSI-2-kiintolevy, 15” SVGA-monitori, CD-ROM- ja 1,44 Mt 3,5” levyke- sekä nauha-asema varmistuksia varten. Verkkoyhteyksiä varten on Etherlink-verkkokortti. Käyttöjärjestelmänä on Linux (RedHat 4.0) ja WWW-palvelimena on NCSA 1.4.2-Linux. Lisäksi koneessa on dokumenttien hakupalvelua varten kehitetty Hyperlib-käyttöliittymäohjelma.

Digitointiympäristönä toimii PC-työasema, jossa on 133 MHz:n Pentium-prosessori ja ATX -emolevy. Koneessa on 64 Kt RAM-muistia ja 1,7 Gt:n EIDE-kiintolevy, 17” SVGA-monitori, CD-ROM ja 1,44 3,5” levyke-asema. Verkkoyhteyksiä varten on Intelin EtherExpress-verkkokortti. Helsingin pilotin käytössä on ollut neljä tällaista työasemaa. Käyttöjärjestelmänä on Microsoftin Windows95 ja ohjelmistona on Microsoftin Office-paketti, versio 7.0 (Word, Excel, Powerpoint), kuvankäsittelyohjelmista Corelin PhotoPaint 3, Adoben Photoshop 4.0 ja LE, PaintShopPro 4.0 ja LViewPro 1.D2, sivuntaitto-ohjelmista Adoben PageMaker 6.0 Alduksen PageMaker 5.0 ja Corelin Ventura 7.0, pääteyhteysohjelmista Reflection 1, julkaisuohjelmista Adoben Acrobat 3.0 (Distiller, Exchange, Catalog, Reader), HTML-editoreista mm. HotMetal Pro 3.0, HotDog 2.53, ja HomeSite 1.2, tekstintunnistusohjelmista Adoben Acrobat Capture 1.0 ja Caeren OmniPage 8.0 ja LE, HTML-selaimista mm. Netscapen 3.0, 4.0 ja Communicator, Microsoftin Explorer ja arkistointiohjelmista Visioneerin PaperPort (käytetty OmniPage LE:n kanssa).

Projektin edetessä on laitteistokokoonpanossa ja laitteiden määrässä sekä ohjelmistossa tapahtunut joitakin muutoksia. Ohjelmaversiot on päivitetty useimmissa tapauksissa uusimmiksi, mutta myös vanhemmat versiot on säilytetty. Yksi kone on korvattu 200 MHz:n Pentium II -koneella ja yhteen koneeseen on hankittu lisää kiintolevykapasiteettia. Lisäksi on hankittu PC:hen liitettävä ZIP-levykeasema sekä Macintosh 6500/250, johon hankitaan tarpeelliseksi katsottavat ohjelmat

Projektin laitteistoon kuuluu lisäksi HP ScanJet 4c -väriskanneri ja HP 5M LaserJet -verkkolasertulostin.

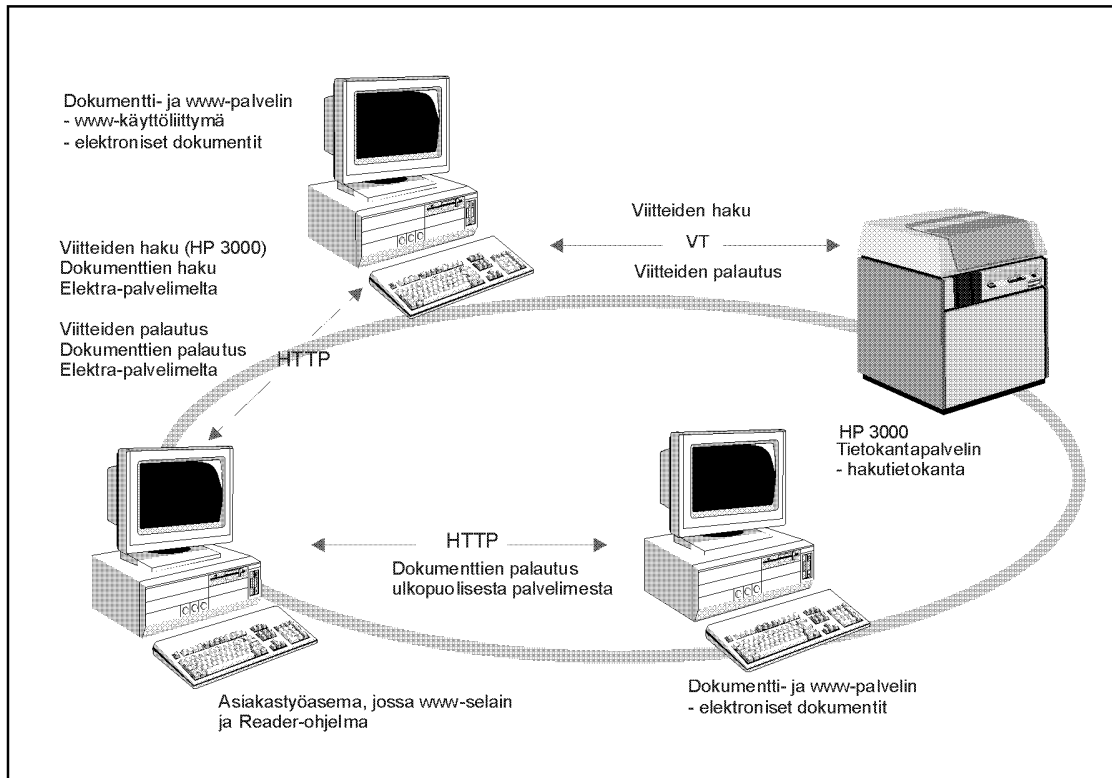
6.2 Dokumenttien haku ja selaus

ELEKTRAN hakupalvelua varten on perustettu erillinen bibliografisten kuvailujen tietokanta Helsingin yliopiston kirjaston HP 3000 -tietokoneelle, jossa hakuohjelmistona on VTLS. Dokumenttien arkisto ja WWW-palvelin on sijoitettu erilliselle palvelinkoneelle, jonka käyttöjärjestelmänä on Linux. WWW-palvelun käyttöliittymäohjelmana on tarkoitusta varten kehitetty Hyperlib, joka toimii Netscape ja Internet Explorer -selainohjelmien kanssa. Kommunikointi HP 3000 - koneen kanssa tapahtuu VT-protokollaa käyttäen.

Järjestelmän käyttäjällä on työasema, jossa toimii Netscape- tai Internet Explorer WWW-selainohjelma. Lisäksi tarvitaan myös Adoben Acrobat Reader -ohjelma PDF-muotoisten dokumenttien lukemista varten. Käyttäjä ottaa yhteyden ELEKTRAN WWW-palvelimeen ja tekee Hyperlib-käyttöliittymän avulla tiedonhaun ELEKTRA-tietokannasta. Viitteessä olevaa URL-osoitetta klikkaamalla näytölle saadaan itse elektroninen dokumentti. Yhteysprotokollana on

HTTP. ELEKTRAN arkistopalvelimena toimii sama palvelin, jolla myös käyttöliittymäohjelmisto sijaitsee. Dokumentit on joko tallennettu omalle palvelimelle tai sijaitsevat omilla WWW-palvelimilla.

Seuraavassa kuvassa on esitetty dokumenttien haku ja selaus.



Kuva 3. Dokumenttien haku ja selaus.

6.3 Tulostus

ELEKTRAN käyttökokeilussa dokumenttien lukemista ja tulostusta ohjaavat tekijänoikeussopimukset. Dokumenttien lukeminen näytöltä on maksutonta ja vapaasti käytettävissä kokeilukirjastojen ELEKTRA-työasemilta. Dokumenttien tallennus tiedostoksi on sen sijaan estetty. Tulostus on sopimuksissa määritelty maksulliseksi, mutta koska valmiita rahastusohjelmistoja on käytettävissä, peritään korvaukset manuaalisesti. Tulostimia ei ole tästä syystä sijoitettu työasemien yhteyteen, vaan käytössä ovat verkkotulostimet. Tämä mahdollistaa tulostuksen kontrolloinnin ja tulostusmaksujen perimisen kirjaston asiakaspalvelussa. Teknisesti dokumenttien tulostus tapahtuu WWW-selainten omia tulostusohjelmia käyttäen.

6.4 Arkistointi

ELEKTRAN tiedostojen arkistopalvelimelle tehtävää tallennusta varten projektissa luotiin hakemistorakenne ja tiedostojen nimeämiskäytäntö. Nimipalvelussa arkistopalvelin nimettiin ELEKTRA.helsinki.fi. Kaikki omalle dokumenttipalvelimelle tallennettujen dokumenttien MARC-kuvailutietueisiin tallennetut URL-osoitteet käyttävät tätä nimeä. Kustantajien palvelimella sijaitseviin dokumentteihin viitataan näiden arkistopalvelimine osoitteiden avulla.

ELEKTRAn dokumenttipalvelimella oleva aineisto on tallennettu palvelinkoneen WWW-serverin juurihakemiston alle. Tallennettava materiaali on jaettu hakemistoihin seuraavasti:

Ylimmällä tasolla aineisto on jaettu FINMARC-formaatin aineistotyyppien mukaan (esim. monografiat, kausijulkaisut ja niiden lyhenteitä käytetään kuten formaatissa, eli bk = kirjat, se = kausijulkaisut). Kussakin aineistossa voidaan soveltaa sille parhaiten sopivia jaotteluperusteita; tässä tekstissä käsitellään jatkossa vain artikkeleita ja niiden sisältämiä kuvia (jos ne on tallennettava erillään tekstistä) sekä monografioita. Seuraavalla tasolla artikkeliaineisto jaetaan lehden nimen mukaan kirjaimittain (hakemisto a, hakemisto b ..., hakemisto z). Skandinaavisten merkien käyttö nimeämisessä tuottaa ongelmia, ja siitä syystä esim. a ja ä on yhdistetty hakemistoon a, jne.

Artikkeleiden hakemistopuu on jaettu ISSN-tunnuksella, joka on tallennettu tavuviivan kanssa (esim. 0036-8733). Vastaavasti monografioille on käytetty ISBN-tunnusta. Mikäli julkaisulla ei ole ISSN-tunnusta, on käytetty julkaisun nimeä tai sen osaa. Oleellista on että nimi on riittävän lyhyt ja helposti muistettava. Lehden nimen sanat yhdistetään tarvittaessa alaviivoilla, koska tyhjämerkkien käyttö ei ole mahdollista.

Seuraava jakoperusteena on käytetty vuosikertaa ja lehden numeroa. Mikäli lehden vuosikertaa ei ole ilmoitettu, käytetään julkaisuvuotta. Senkin puuttuessa käytetään pelkkää numerointia.

Alimmalla hakemistotasolla on annettu tiedostonimi, joka sisältää tietoa artikkelin nimekkeestä ja tiedostomuodosta (html, pdf). Erotteluperusteeksi riittää usein artikkelin nimen 8 ensimmäistä merkkiä. Tarvittaessa käytetään artikkelin nimekkeen ja tekijän sukunimen yhdistelmää (neljä merkkiä kummankin alusta). Koska hakemistorakenteessa on päädytty siihen, että saman artikkelin eri tiedostomuodossa tallennetut versiot kannattaa tallentaa samaan hakemistoon, ei hakemistopuun alemmilla tasoilla ole tehty jaottelua formaatin mukaan. Hakemistojen nimeämissä isojen kirjainten käyttö ei ole sallittua.

Esimerkki artikkelin URL:stä: <http://elektra.helsinki.fi/se/g/0016-6898/67/1/theslofd.pdf>

Pdf-dokumenteissa kuvat tallennetaan osaksi perustiedostoa, mutta html-tekstien kuvat ovat erillisiä tiedostoja. Html-dokumenttien sisältämät kuvat tallennetaan samaan hakemistoon itse dokumenttitekstien kanssa. Kuvien nimissä ei ole käytetty absoluuttisia (koko URL mukana) osoitteita.

Monografioissa jaotteluperusteena on käytetty aineistotyyppiä osoittavan bk-tunnisteen jälkeen sitä osaa ISBN-tunnuksesta, joka yksilöi kustantajan (esimerkiksi 951710). Tavuviivaa ei tallenneta. Jos kustantajalla ei ole ISBN-tunnusta, on käytetty kustantajan nimeä tai sen lyhennettä. Kustantajan jälkeen hakemisto on jaettu vielä vuoden mukaan. Lopuksi on annettu tiedostolle nimi, joka yksilöi teoksen. Tämä on tehty samalla periaatteella kuin artikkelien nimeäminen.

Esimerkki monografian URL:stä: <http://elektra.helsinki.fi/bk/951710/1997/vanhansu.pdf>

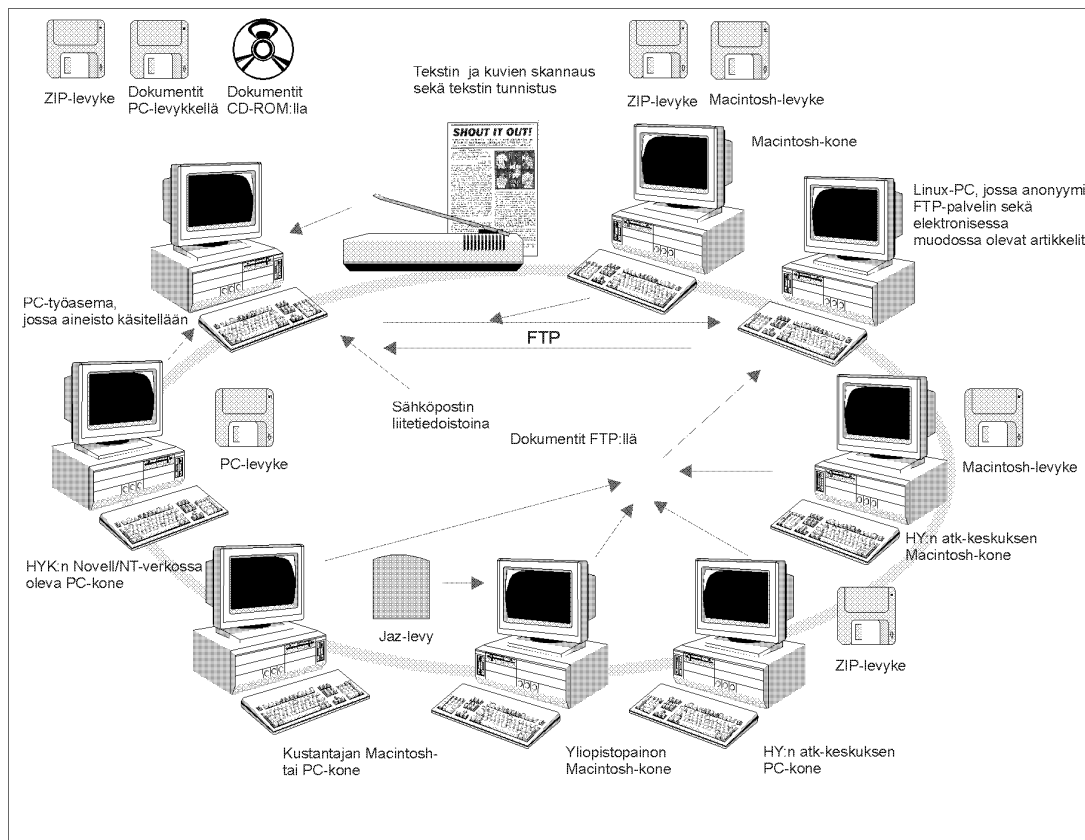
7. AINEISTON DIGITOINTI

7.1 Pilottiaineiston toimitus ja käytetyt tiedostomuodot

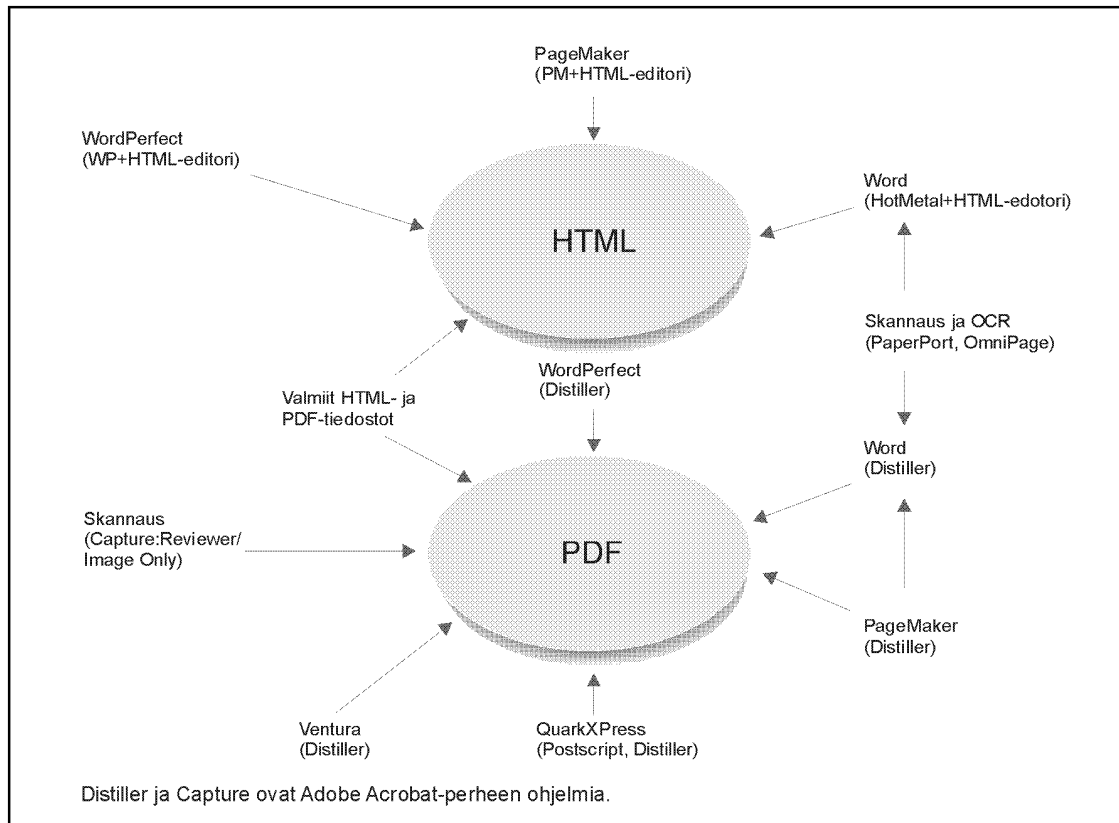
ELEKTRAn pilotit sisältävät artikkeli- ja monografia-aineistoa. Aineistoa on toimitettu projektille erilaisissa tiedostomuodoissa levykkeillä, sähköpostin liitetiedostoina ja FTP:llä sekä painetussa muodossa. Artikkelit on digitointiprosessissa muutettu joko HTML- tai PDF- muotoon. Pääosa materiaalista on vain jommassa kummassa muodossa, mutta jonkin verran aineistoa on vertailumielessä tallennettu molemmissakin tiedostomuodoissa.

Tavallisimmin tallennusmuotona on PDF, koska se säilyttää julkaisun alkuperäisen ulkoasun, sen tuottaminen on suhteellisen helppoa ja se mahdollistaa dokumentin tulostuksen rajoittamisen. Myös silloin, kun dokumentti sisältää kaavoja, jotka voitaisiin esittää html-muodossa vain kuvina, on käytetty PDF-formaattia. Se osa aineistoa, joka on käytössä suoraan julkaisijoiden omilta palvelimilta, on tavallisimmin html-muodossa. Oulun väitöskirja-pilotissa on kokeiltu myös SGML- ja XML-formaattien käyttömahdollisuutta. Silloin kun aineisto pilotteja varten on saatu tuotettuna jollain tekstinkäsittely- tai julkaisuohjelmistolla, on säilytetty myös alkuperäinen tiedosto. Tämä on tarpeen pitkäaikaista säilytystä varten.

Seuraavissa kuvissa on esitetty ne kanavat, joita on käytetty pilottiaineiston toimituksessa ja konvertointi- ja digitointiprosessin eri vaihtoehdot.



Kuva 4. Aineiston toimituskanavat ja digitointiprosessi



Kuva 5. Konvertointi- ja digitointiprosessin eteneminen.

7.2 Painetussa muodossa toimitettu aineisto

Silloin kun dokumentit on saatu ainoastaan painetussa muodossa, tapahtuu niiden digitointi skannaamalla ja tekstintunnistusohjelmien (OCR, Optical Character Recognition) avulla. Varsinaisessa skannauksessa on käytetty HP:n DeskScan II- ohjelmaa ja tekstin tunnistusohjelman valinnassa on otettu huomioon mm. dokumenttien kieli ja ulkoasulle asetettavat vaatimukset.

OmniPage

Mikäli artikkeli on suomenkielinen ja sisältää pääosin pelkkää tekstiä, on dokumentin tekstin tunnistaminen tehty Visioneer PaperPort -ohjelman kautta OmniPage (Limited Edition) -ohjelmalla.

OmniPage tunnistaa tekstin ja myös skandinaaviset merkit melko hyvin. Paperport-ohjelma helpottaa skannatun materiaalin hallintaa ja jatkokäsittelyä. Joidenkin kirjainyhdistelmien tunnistaminen sitävastoin tuottaa OmniPagelle suuria vaikeuksia. Läheisesti toisiaan muistuttavat merkit tai niiden yhdistelmät (aakkoset ja numerot) sekoittuvat helposti. Esimerkiksi $rn \Rightarrow m$, $m \Rightarrow rn$, $ll \Rightarrow H$, $ll \Rightarrow ff$, $b \Rightarrow 6$, $i \Rightarrow l$. Tekstissä olevan tavutuksen ohjelma osaa hoitaa kohtalaisesti poistamalla ylimääräiset tavuviivat. Fontin koko saattaa vaihdella kesken tekstin, vaikka se painetussa versiossa olisikin koko ajan saman kokoista. Tekstin väliin saattaa myös jäädä ylimääräisiä tyhjiä rivejä. Tekstissä olevat ylä- ja alaviitteet muuttuvat muun tekstin kaltaiseksi, fonttikoon vaihdellessa. Taulukoiden käsittely aiheuttaa runsaasti lisätyötä, sillä taulukon sarakkeet eivät tahdo pysyä paikoillaan. Sivujen pinojäljen taso ja tekstityyppi vaikuttavat merkittä-

västi tunnistuksen onnistumiseen. Normaali teksti tunnistetaan kohtalaisen hyvin, mutta esimerkiksi kursivoidun tekstin tunnistaminen on huomattavasti hankalampaa. Palstoitus tuottaa myös ongelmia, eli ohjelma saattaa pitää palstoilla olevia tekstin osia yhtenä normaalina rivinä, jolloin teksti menee sekaisin.

PaperPort-ohjelma kokoaa artikkelin sivut pinoiksi. Kaksipuoliset sivut on skannattava erikseen ja muodostuneet kaksi pinoa yhdistettävä. Sivupinoa voidaan jatkokäsittellä halutulla ohjelmalla, esimerkiksi Word-dokumenttina, jossa artikkelille voidaan suorittaa oikoluku sekä tehdä muut tarvittavat korjaukset, muotoilut, muutokset ja/tai lisäykset. Tämän jälkeen dokumentti voidaan konvertoida haluttuun muotoon (PDF tai HTML).

Mikäli artikkelissa on kuvia, on ne skannattava erikseen jonkin kuvankäsittelyohjelman (Adobe Photoshop, Corel Photo-Paint) avulla ja liitettävä joko tekstinkäsittelytiedostoon PDF konversiota varten tai linkkeinä HTML-tiedostoon. Vaikka kuvat eivät tulekaan mukaan lopulliseen dokumenttiin, pyrkii OmniPage silti tunnistamaan tekstiä myös kuvista, tosin ei läheskään yhtä innokkaasti kuin esimerkiksi Adoben Acrobat Capture.

Projektin käyttöön on vuoden 1997 lopulla hankittu uusin versio 8 OmniPage-ohjelmasta. Kokemukset ohjelman käytöstä ovat vielä melko vähäisiä, mutta tunnistusominaisuuksien luvataan olevan parempia ja ohjelma tarjoaa runsaasti lisäominaisuuksia esimerkiksi palstoitetun tekstin käsittelyyn. Kuvat saadaan myös mukaan ja alkuperäisen sivun muotoilut voidaan säilyttää, joten ohjelma lähenee ominaisuuksiltaan Adoben Acrobat Capturea.

Adobe Acrobat Capture

Toinen skannatun aineiston tunnistusohjelma on Adoben Acrobat Capture. Käytössä oleva Capturen ohjelmaversio ei kuitenkaan tunnista skandinaavisia merkkejä eikä myöskään suomen kieltä (tarkistus perustuu sanaston ja kirjasinten tunnistukseen). Mikäli dokumentti halutaan tallentaa tekstinä eikä pelkästään kuvana, on sen jatkokäsittely työlästä. Englanninkielisen aineiston käsittely on Capturella on sen sijaan nopeaa ja vaivatonta. Ohjelman uudempi versio, joka projektille ollaan hankkimassa, tunnistaa myös skandinaaviset kirjaimet ja sisältää ruotsinkielisen sanaston.

Dokumentti voidaan tallettaa myös kuvana. Tällöin kirjasinta ja tekstiä ei tunnisteta, vaan jokainen sivu tallennetaan yhtenä kokonaisuena TIFF-muotoisena tiedostona. Tätä menettelyä käyttäen tiedostojen koot kasvavat kuitenkin melko suuriksi, eikä teksti laadultaan ole yhtä hyvää kuin tunnistetun tekstin. Kuvaruudulta luettuna ero on selvä, mutta paperille tulostettuna ero pienee. Capture tunnistaa myös kuvat, mutta niiden laatu ei ole kovin hyvä ja kuvat tulevat mukaan ainoastaan harmaasävyisinä, jolloin mahdollinen kuvien väri-informaatio menetetään.

Tunnistettu teksti ja teksti kuvana

Tunnistetun tekstin ja kuvina käsiteltävän tekstin suurimmat erot ovat hakumahdollisuuksissa, dokumentin tulostustarkkuudessa ja tiedostojen koossa ja käsittelyajassa. Ohessa on esimerkkejä tunnistetun tekstin ja kuvana käsiteltävän tekstin eroista tiedostojen koon ja käsittelyajan suhteen.

Suomen kirkkohistoriallisen seuran vuosikerran kaksi artikkelia muunnettiin skannaamalla PDF-muotoon. Tekstin tunnistuksen ja oikoluvun jälkeen (PaperPort, OmniPage ja Word) artikkelien koot ovat 736 ja 569 kilotavua ja Capturella kuvina tallennettuina vastaavasti 3122 ja 2720 kilotavua. Käsittelyajat näille kahdelle tiedostolle olivat yhteensä vastaavasti 8.23 ja 1.14

tuntia. Eli kun tiedoston kokoa saatiin pienennettyä 4-5 kertaisesti kasvoi aineiston käsittelyaika noin seitsenkertaiseksi. Toisena esimerkkinä on Genos-lehti, jota alkuun saatiin ainoastaan painetussa muodossa, mutta myöhemmin myös valmiiksi elektronisessa muodossa (Word), ja joka myös talletettiin PDF:ssä. Capturen TIFF-kuvina yhden lehden artikkelien koot olivat: 439, 959 489, 43 ja 339 kilotavua ja lehden käsittelyaika oli 51 minuuttia. Kun artikkelien teksti tunnustettiin ja artikkelit tallennettiin tekstinä, vei tämä aikaa 10.24 tuntia ja tiedostojen koot olivat jälleen huomattavasti pienemmät (27, 50, 29, 30 ja 22). Vastaavasti elektronisessa muodossa toimitetun saman lehden artikkelien koot olivat: 24, 48, 26, 27 ja 19 kilotavua ja käsittelyaika oli myös 51 minuuttia.

Mikäli käytössä on riittävästi resursseja, eikä työstä aiheutuva kustannus ole este, on luonnollisesti paras vaihtoehto tunnistaa teksti, jolloin tiedostojen koot pienenevät, tekstin laatu paranee ja tekstiin voidaan kohdistaa hakuja. Artikkelien tarkistus ja oikoluku on kuitenkin paljon resursseja vaativaa ja vieraan ammattisanaston ansiosta vaikeaa työtä. Oikoluvusta huolimatta artikkeleihin jäi varmasti jonkin verran virheitä, sillä oikolukuohjelmat eivät suinkaan löydä kaikkia virheitä, vaan käytännössä artikkeli on luettava kokonaan. Pieni tiedoston koko on selvä etu sekä arkistoinnin että artikkelien käytön kannalta. Artikkelien säilytys vie vähemmän tilaa ja ne latautuvat lukijan näytölle nopeammin ja jälki on myös parempaa ainakin ruudulta luettaessa. Paperille tulostettujen artikkelien ulkonäön ja luettavuuden laatuun eri menetelmät eivät vaikuta yhtä selvästi.

Painettuun materiaaliin kanssa ongelmia aiheuttavat sellaiset sivut, joissa uusi artikkeli alkaa samalta sivulta kuin edellinen päättyy. Sivulla voi myös olla mainoksia tai muuta itse artikkeliin kuulumatonta tekstiä tai kuvamateriaalia. Jos skannatut sivut tallennetaan TIFF-kuvina, voidaan ne avata jossakin kuvankäsittelyohjelmassa ja itse artikkeliin kuulumattomat osat voidaan poistaa. Lopputulos saattaa kuitenkin olla epätydyttävä, jos itse artikkeli alkaa vasta sivun puolivälistä tai jos artikkelin ensimmäisellä sivulla on vain yksi sivun oikeassa laidassa oleva palsta esimerkiksi kolmen palstan asemesta. Tämä saattaa turhaan kasvattaa artikkelin sivumäärää sitä tulostettaessa. Lyhyiden artikkelien osalta tekstiosioita on mahdollista siirrellä kuvankäsittelyohjelmassa, jolloin sivuille ei jää tyhjää tilaa, mutta pitkien artikkelien kohdalla tehtävä on mahdoton.

Koska aineiston skannaus paperimuodosta aiheuttaa monia ongelmia ja on melko työlästä, olisi-kin paras vaihtoehto saada materiaali valmiiksi jossakin elektronisessa muodossa.

Kuvat

Painetussa muodossa toimitetun aineiston ongelmia aiheuttavat myös kuvat, jotka sisältävät paljon pieniä yksityiskohtia usein pienifontisella tekstillä yhdistettynä. Näiden yksityiskohtien merkitys kuvassa saattaa olla hyvinkin suuri (esimerkiksi luonnontieteellinen aineisto), mutta tällaisten kuvien esittämiseen ei normaali tekstin tunnistava skannaustarkkuus riitä. Mikäli tarkkuutta kasvatetaan, kasvaa myös kuvatiedoston koko ja samalla kuvan koko näyttöruudulla. Valokuvien osalta skannaus ei ole ongelma, mutta piirrosten ja erilaisten diagrammien skannaus ei onnistu hyvin edes siten, että niitä ensin suurennettaisiin esimerkiksi valokopioimalla.

Skannaustulokseen voidaan vaikuttaa mm. valitsemalla skannaustarkkuus, esim. 300 dpi (dots Per inch). Tulokseen vaikuttaa myös kirkkauden (brightness) ja kontrastin (contrast) säädöllä, joka voidaan tehdä itse tai automaattisesti ohjelman toimesta. Adoben Capturessa käytetään yleisesti tarkkuutta 300 dpi ja aineistoa käsitellään mustavalkoisina tai harmaasävykuvina. Tarkkuuden nostaminen (ohjelman rajoitukset: mustavalkokuvat 200–600 dpi ja harmaasävyt 200–400 dpi) ei sanottavasti vaikuta varsinkaan piirroksen omaisten kuvien kuten taulukoiden ja

diagrammien laatuun, vaan kuvainformaatiota menetetään. Tarkkuuden nostaminen hidastaa huomattavasti itse skannausprosessia.

Capturella skannatuissa dokumenteissa, joissa on pienehköjä kuvia, on ratkaisuna ongelmaan skannata kuvat erikseen ja suurentaa kuvaa jonkin verran skannausvaiheessa ja tallentaa kuvat erillisiksi HTML-sivuiksi, joihin Adoben Exchange-ohjelmalla voidaan varsinaisesta artikkelista tehdä linkit perustuen normaaliin URL-käytäntöön. Suurempien kuvien osalta ratkaisu on ei ole näin yksinkertainen, koska tarkkuutta ja suurennosta kasvatettaessa tiedoston koko kasvaa helposti niin suureksi, että kuvan latautuminen nopeillakin yhteyksillä kestää huomattavan kauan. Koko kuva ei myöskään mahdu yhdellä kertaa näytölle, vaan vain osa siitä. Yksityiskohdan täyttäessä koko ruudun, on vaikea hahmottaa kokonaisuutta.

Elektronisessa muodossa toimitettu aineisto on sisältänyt melko harvoin kuvia valmiina, eli kuvat on ensin skannattu ja liitetty artikkeleihin ja sen jälkeen dokumentti on muunnettu joko PDF- tai HTML-muotoon. HTML:n osalta kuvat tallennetaan erillään itse tekstistä.

Kuvan käyttötarkoituksesta (PDF, HTML) ja laadusta (valokuva, harmaasävykuva, piirros jne.) on riippunut, millä tarkkuudella skannaus on tehty. HTML-tiedostojen kuvissa on käytetty etupäässä tarkkuutta 75/75 (Piirroset ja harmaasävykuvat/Valokuvat) ellei kuvan laatu (esimerkiksi pieniä yksityiskohtia sisältävä piirros) ole edellyttänyt suurempaa tarkkuutta. Tekstinkäsittely- ja sivuntaitto-ohjelmien kautta PDF:ksi muunnettavissa artikkeleissa on käytetty tarkkuutta 300/100. HTML-muotoisten artikkelien kuvien tallennusmuotona on joko JPG tai GIF kuvan laadusta riippuen (valokuvat JPG, muut GIF). PDF:ksi muunnettavissa artikkeleissa on JPG:n ja GIF:n ohella käytetty myös EPS- ja TIFF-kuvia. Valtaosa kuvista on skannattu painetusta julkaisusta eli alkuperäisiä kuvia ei juurikaan ole ollut käytettävissä. Alkuperäisten kuvien käyttö tuottaisi kuitenkin paremman lopputuloksen kuin painetussa muodossa olevasta kuvasta saatava versio.

7.3 Elektronisessa muodossa toimitettu materiaali

Projektissa on käytetty mahdollisuuksien mukaan aina ohjelmien uusimpia versioita, koska yhteensopivuus vanhemmilla versioilla tuotettujen dokumenttien kanssa on ainakin periaatteessa taattu.

Yleisimmillä tekstinkäsittelyohjelmilla (MS Word, WordPerfect) tuotettujen dokumenttien muuttaminen HTML- tai PDF-muotoon onnistuu melko vaivattomasti. Word- ja WP-dokumentit voidaan Adoben Acrobat PDF Writer -ajurin tai Distiller-ohjelman avulla muuttaa suoraan PDF-muotoon valitsemalla tulosta-komennon asetuksista normaalin kirjoittimen sijasta tulostus PDF:ksi. Käytännössä Distillerillä ja PDF Writerilla ei ole suuria eroja, mutta Distiller antaa enemmän mahdollisuuksia vaikuttaa muunnosprosessiin ohjelman asetusten avulla ja Distillerillä tuotettu materiaali on joissain tapauksissa laadukkaampaa ja tiedostojen koot ovat hieman pienempiä.

Microsoft Word

MS Word-dokumentin konvertointi HTML-muotoon on tehty HTML-editorilla (esim. HotMetal). Konvertoinnin jälkeen tarvitaan lähes aina korjauksia HTML-tiedostoon, sillä HTML-editorit eivät osaa konvertoida dokumenttia täysin oikein. Yleisimmät virheet ilmenevät otsikoissa, listoissa ja taulukoissa. Jotta artikkelin konvertointi onnistuisi mahdollisimman vähällä lisätyöllä, tulee sen olla kirjoitettu Wordin tyyliääritysten mukaisesti, mitä projektille toimi-

tetussa aineistossa ei juurikaan ole tehty. Kuvat voivat olla joko valmiina Word-dokumentissa, jolloin konvertoitaessa muodostetaan linkit ko. kuviin tai kuvat voidaan lisätä myös jälkikäteen käsityönä, jolloin tulos usein on parempi.

Wordin vanhemmilla ohjelmaversioilla kirjoitetut artikkelit eivät ole aiheuttaneet ongelmia. Macintoshin Wordillä kirjoitetut artikkelit käytiin avaamassa Helsingin yliopiston atk-keskuksen Macintosh-laitella ja ne talletettiin PC:n Word-muotoon PC-alustetuille levykkeille tai siirrettiin FTP:llä Linux-koneelle ja siitä edelleen PC:lle. Tämä menettely ei ole aiheuttanut suuria ongelmia. Tällaisen materiaalin määrä on ollut pieni ja tiedostot olivat sisällöltään normaalia tekstiä, jossa ei esiintynyt mitään erikoismerkkejä.

Muutamassa Word-muodossa toimitetussa artikkelissa fontti on vaikuttanut kovin pieneltä, eikä se normaaleilla fontin kokoon muunnoksilla ole juurikaan kasvanut. Lopulta kävi ilmi, että koko teksti olikin jossain vaiheessa muuttunut yliviitteeksi. Yhden julkaisun materiaali toimitettiin Wordin txt-muodossa ja kun tiedostot avattiin Word-ohjelmalla, olivat kaikki skandinaaviset merkit muuttuneet muiksi merkeiksi. Useimmiten skandinaaviset sekä näppäimistöön kuulumattomat merkit korvautuvat sellaisilla muilla merkeillä, joita ei saada tekstinkäsittelyohjelman korvaustoiminnon avulla, vaan korvaus on tehtävä käsityönä.

Corel WordPerfect

Corelin WordPerfectin versiosta 7.0 voidaan tallentaa dokumentti suoraan HTML-muodossa. Konvertointi onnistuu usein myös paremmin kuin Word-dokumentista jollakin HTML-editoreilla, esimerkiksi listat ja taulukot konvertoituvat huomattavasti paremmin, mutta konvertointiohjelma lisää dokumenttiin myös turhia tyhjiä rivejä, jotka on poistettava.

Vanhemmilla ohjelmaversioilla tuotettu materiaali on aiheuttanut jonkin verran ongelmia. Palstoituksen osalta teksti on siirtynyt siten, että sivuilla saattaa olla tyhjiä osuuksia. Tekstin tasaaminen normaaliksi ei ole aina onnistunut. Eräässä aineistossa oli lisäksi runsaasti erilaisia, melko suurikokoisia taulukoita, mikä hankaloitti tilannetta lisää. Aineisto käsiteltiin avaamalla tiedostot MS Wordissä. Samalla poistettiin palstoitus, jolloin lopullisessa muodossa (PDF) olevan tekstin luettavuus näytöltä myös parani, koska yhdellä sivulla ei tällöin tarvitse liikkua pystysuunnassa kuin yhden kerran (useampipalstaisessa tekstissä = palstojen lukumäärä). Tekstiä oli siirreltävä taulukoiden ympärillä, jotta taulukot saatiin kokonaisuudessaan yhdelle sivulle. Konversiot välillä WP-Word aiheutti sen, että skandinaaviset merkit tulivat muusta fontista erottuvina, lihavoituina. Leipätekstissä tämä ei PDF:ssä erottunut, mutta kaikissa otsikoissa skandinaaviset merkit muistuttivat PDF-tiedostossa kuvia eli ne olivat huomattavan epätarkkoja. Ongelma ratkesi muuttamalla otsikkojen fonttia (Palatino ⇒ Times New Roman). Sama ongelma esiintyi, vaikkei konversiota edes tehty, vaan vanhemmalla WP-versiolla tuotettu dokumentti avattiin projektin WP-ohjelmalla (versio 7.0). Itse leipätekstistä katosivat kaikki skandinaaviset merkit, kun aineisto konvertoitiin PDF-formaattiin. Dokumentti on kirjoitettu käyttämällä Palatino-fonttia, joka selvästi tuottaa ongelmia PDF-konversiossa. Ongelmasta päästiin muuttamalla dokumentin kirjasin toiseksi. Myös rivien täyttötasaus muuttui vasemman reunan tasaukseksi, eivätkä sivut vastanneet alkuperäisiä, vaan teksti siirtyi joko eteen tai taaksepäin. Skandinaaviset merkit tulivat mukaan, mutta merkkejä on mennyt päällekkäin. Kirjainten tai merkkien päällekkäinmeno näyttää olevan satunnaista, eikä se näytä liittyvän ainoastaan skandinaaviisiin kirjaimiin tai erikoismerkkeihin. Ongelmaa ei ole vielä saatu poistettua.

Suosittelavaa olisikin, että artikkeleissa käytettäisiin tavallisia ja yleisimpiä kirjaimia ”eksoottisempien” kirjainten sijaan. WordPerfectin vanhempia versioita avattaessa on tekstin

joukossa esiintyy ohjausmerkkejä, jotka on poistettava, sillä muuten ne tulostuisivat myös PDF-muotoiseen dokumenttiin.

Muilla tekstinkäsittelyohjelmilla kirjoitettua aineistoa ei projektille ole toimitettu.

QuarkXpress

Projekti on saanut yhden lehden aineiston Macintoshin QuarkXpress-ohjelmistolla tehtynä. Ohjelmaa ei hankittu projektille, koska oli epävarmaa toimisiko Macintoshilla tehty materiaali PC-ympäristössä. Ohjelmiston hinta on myös melko korkea. Tämän materiaalin osalta on toimitettu siten, että lehden kuvat on skannattu projektin omilla resursseilla, jonka jälkeen aineisto on toimitettu kokonaisuudessaan Yliopistopainolle, joka on suorittanut kuvien liittämisen artikkeleihin ja aineiston muunnon PostScript-muotoon, josta on itse edelleen jalostettu PDF-tiedostoja.

PageMaker

PageMakerillä tehdyt dokumentit voidaan myös konvertoida suoraan PDF-muotoon, mikäli koneessa on asennettuna Adoben Distiller-ohjelma tai HTML:ksi PageMakerin omalla apuohjelmalla. Tämä on kuitenkin monimutkainen prosessi ja aiheuttaa runsaasti lisätöitä, eikä siksi ole kovin suositeltava tapa. PageMakerillä tuotettua aineistoa ei ole muutettu HTML:ksi muuten kuin kokeilumielessä. Parempi tapa on siirtää teksti ensin johonkin tekstinkäsittelyohjelmaan ja tehdä konversio HTML-muotoon sieltä käsin.

Merkittävin ongelma PageMakerillä toimitetun aineiston osalta on ollut se, että aineiston toimittajilla on melko laajasti käytössä Macintosh-tietokoneita. Macintoshilla tehdyt PageMaker-dokumentit aukeavat periaatteessa myös Windows-ympäristössä. Mikäli tiedostot toimitetaan levykkeellä, on edellytyksenä se, että levyke on alustettu PC:tä varten. Käytännössä PC:llä päästään käsittelemään Macintoshilla tehtyjä tiedostoja vain sillä edellytyksellä, että ohjelmaversiot ovat samat, esimerkiksi konvertointi Macintosh versiosta 5.0 suoraan Windows-versioon 6.0 ei onnistu, vaan tarvitaan ensin muunnos Windows-versioon 5.0 ja vasta sitten muunnos versioon 6,0, jos vanhempaa versiota ei, kuten ELEKTRAssa aluksi oli tilanne, ole käytettävissä.

Vanhemmalla kuin 6.0-versiolla tuotettu aineisto, joka on toimitettu PC-alustetuilla levykkeillä tai FTP:llä on käyty avaamassa sellaisella koneella, jossa on PageMakerin vanhempi versio (yleensä 5.0) ja siirretty aineisto sitten projektin tuotantokoneille jatkokäsittelyä varten. Vuoden 1997 lopulla projektille saatiin oma 5.0-ohjelmaversio, joten aineiston käsittely on nopeutunut ja yksinkertaistunut, koska nyt PDF-tiedostot voidaan muuntaa suoraan versiosta 5.0. Mikäli aineisto on toimitettu Macintosh-alustetuilla levykkeillä, on tiedostot käyty avaamassa Yliopiston atk-keskuksen Macintosh-koneilla ja siirretty sitten FTP:llä Linux-palvelimelle ja siitä edelleen jollekin PC:lle käsiteltäväksi.

Mikäli PageMakerin tekstissä on käytetty normaalista merkistöstä poikkeavia merkkejä, eivät ne välttämättä säily oikeina muunnettaessa Macintoshissa tuotettua aineistoa PC:lle. Eräissä aineistossa esimerkiksi merkit š, č ja ž muuttuivat merkeiksi £, \$ ja ¥. Lisäksi Macintoshissa käytetään sivujen muotoilussa yleisesti tuumamitoitusta kun taas PC:ssä mitoitus on oletuksena millimetreissä. Usein tekstiä on muotoiltu tiivistämällä ja erilaiset mittajärjestelmät sotkevat sivujen muotoilua. Myös erilainen fontti muutti alkuperäistä taittoa liu'uttamalla tekstiä eteenpäin, mikä taas yhdessä väitöskirjassa vaikutti esimerkiksi kirjallisuusviitteisiin siten, että itse viite sijaitsikin eri sivulla kuin itse viittaus. Hämmästyttävää kyllä, että jotkut virheet, esimer-

kiksi tekstissä olevien erilaisten elementtien siirtyminen tapahtuikin vasta konvertoitaessa PC:n PageMaker versiosta 5.0 versioon 6.0, eikä konversiossa Macintosh versiosta PC-versioon. Saman versionkaan käyttö ei poistanut kaikkia ongelmia, vaan esimerkiksi otsikon muotoilut ja kirjasimet on muuttuivat PDF-muodossa melkoisesti ja kirjaimia on myös kasautunut WP:llä tehdyn materiaalin tapaan päällekkäin.

Osassa toimitettua materiaali on ollut kyrillistä merkistöä (myös tekstinkäsittelyohjelmilla tuotetussa aineistossa), joka tuottaa myös ongelmia. HTML-muodossa kyrillisen tekstin esittäminen on melko hankalaa. Toteutus vaatii KOI-8-standardin mukaisten fonttien asentamista kaikille niille koneille, joista kyrillistä tekstiä sisältävää tekstiä halutaan lukea. Lisäksi käyttäjä joutuu itse muuttamaan sivujen koodauksen selaimen Options-valikosta. Uusimmissa selaimissa tämä tapahtuu jo automaattisesti. Siksi PDF on järkevämpi valinta, mutta senkään käyttö ei ole ongelmatonta. Käytössä on useita eri standardeja ja kaupallisten fonttien lisäksi myös omia vityksiä, jotka toimivat ainoastaan siinä ympäristössä (Macintosh, MS-DOS), mihin ne on alunperin tehty. Pelkkä uusien kirjasimien asentaminen tuotantokoneelle ei välttämättä riitä, vaan tarvitaan myös näppäinohjelmisto, jotta oikeanlaiset merkit saadaan aikaiseksi. PC:n ja Macintoshin välillä saman näppäimen painallus tuottaa myös eri merkin ruudulle.

Koska tilanne on näin kirjava, katsottiin, että projektille ei hankita uusia ohjelmia ja fontteja tällaisen aineiston käsittelemiseen, vaan ongelmat pyritään ratkaisemaan muulla tavalla. Mikäli kyrillisiä merkkejä sisältävät artikkelit ovat lyhyehköjä, hoidetaan ne skannaamalla ja tallentamalla sivut kuvina (Adobe Acrobat Capture). Mikäli artikkelit ovat pitkiä, on pyydetty julkaisijalta/kirjapainolta aineisto valmiiksi joko PDF tai PostScript muodossa. Ensimmäiset kokemukset tällaisesta toimituksesta ovat lupaavia eli olipa aineisto tuotettu missä laitteistoympäristössä tahansa tai millä ohjelmistolla tahansa, säilyivät kirjaimet ja merkit oikeina. Voidaan siis todeta, että PDF- ja PostScript-tiedostojen siirrettävyys on hyvä. Aineiston saamista, varsinkin juuri vaikeampien aineistojen kohdalla, edellä mainituissa muodoissa, voidaan pitää projektin jatkossa yhtenä tärkeimpänä tavoitteena.

Koska PageMaker-tiedostot vievät melko paljon tilaa, on yhden lehden taitto usein usealla erillisellä levykkeellä. Muutamassa tapauksessa artikkeli on ollut kahdella eri levykkeellä. Tällaisten artikkelien yhdistäminen on osoittautunut hankalaksi. Jos valitaan PageMakerin book-ominaisuus, jolla voidaan erillisinä tiedostoina olevat osat koota yhteen, ei tietyn osan tulostusta kokonaisuudesta olekaan mahdollista. Mikäli lisätään artikkelin loppuun uusi sivu ja kopioidaan artikkelin loppu toisesta tiedostosta, ei teksti siirrykään uudelle sivulle alkuperäisessä asussaan, vaan teksti vie huomattavasti enemmän tilaa. Yksi mahdollisuus on muuntaa koko artikkeli book-yhdistelyn kautta PDF:ksi ja Adoben Acrobat Exchange -ohjelmalla erotella artikkelit omiksi PDF-tiedostoiksi.

Joissain tapauksissa PDF:ksi muutettu sivu ei ole vastannut PageMakerin alkuperäistä sivua, vaan tekstin muotoilu on muuttunut sivulla ja/tai osa tekstistä on siirtynyt seuraavalle sivulle. Samaa tekstin siirtymistä on esiintynyt myös Word-tekstinkäsittelyohjelmassa. Wordissä on auttanut tulostuksen ominaisuuksissa sivukoon muuttaminen suuremmaksi. Värikuvat eivät aina ole tulleet PDF-tiedostoon värillisinä, vaikka kaikkien asetusten mukaan niiden olisi pitänyt.

Page-Makerilla toimitettu aineisto on saatu tavallisimmin ilman kuvia. Ne on ensin skannattu ja lisätty dokumentteihin ja sen jälkeen konvertoitu PDF-muotoon. Erään julkaisun kohdalla kuvien lisääminen ei onnistunut vaikka kokeiltiin kaikkia mahdollisia kuvien tiedostomuotoja sekä kuvien sijoitustapoja. Ruudulla kuva näytti olevan artikkelissa oikealla paikallaan, mutta kun artikkeli joko tulostettiin paperille tai muutettiin PDF-muotoon, ei kuvia enää ollutkaan mukana sivuilla, vaan kuvan paikalla oli ainoastaan tyhjä laatikko, jonka yli on vedetty henkselit. On-

gelmaa selvitetään edelleen. EPS-muotoisten kuvien kanssa on myös esiintynyt ongelmia eli jostain syystä kuvat eivät näy myöskään PDF-dokumentissa.

Muutamissa julkaisuissa on otsikot ja kuvat taitettu koko aukeaman levyisiksi. Tämä ei tietenkään toimi artikkelin elektronisessa versiossa, vaan teksti ja kuvat on siirretty tekstinkäsittely-ohjelmaan (Word) ja muutettu sitä kautta PDF:ksi. PageMaker taitossa ongelmia ovat aiheuttaneet myös sellaiset julkaisut, joissa uusi artikkeli alkaa jostain kohtaa samalta sivulta kuin edellinen loppuu, joko palstoitettuna tai horisontaalisessa suunnassa. Näissä tapauksissa on toimitettu samalla tavalla kuin koko aukeaman ylittävissä artikkeleissa tai edeltävän/seuraavan artikkelin osat on poistettu PM:ssä ja muunnettu artikkeli PDF-muotoon, jolloin artikkeli saattaa alkaa sivun alalaidasta tai yksipalstaisena sivun oikeasta reunasta (vrt. painetussa muodossa toimitettu aineisto).

Corel Ventura

Kokemukset Corelin Ventura-taitto-ohjelmasta ovat vaihtelevia. Kaikkia tällä ohjelmalla tuotettuja tiedostoja ei ole saatu muutettua PDF:ksi. Ongelma johtuu todennäköisesti ohjelmien eri versiosta. Venturan hyvänä puolena on kaavaeditori, jolla monimutkaisetkin kaavat saadaan kirjoitettua.

Projektilla on käytössään uusin versio 7.0 ja ainakin versiolla 3.0 tuotetun materiaalin kanssa on ollut ongelmia. Uudemmissa versioissa, ainakin versiosta 4.2 lähtien, tuotettu aineisto koostuu PageMakerin tapaan yhdestä tiedostosta, eikä sen käsittely ole aiheuttanut ongelmia, mutta esimerkiksi Venturan 3.0-versiolla tehty aineisto rakentuu useista eri osista; teksti voi esimerkiksi olla erillään omana tiedostonaan esimerkiksi Word-muodossa. Kun tällainen taitto avataan versiolla 7.0, häviää tekstiä satunnaisesti sivu pari sieltä täältä. Muutokset eivät synny PDF:ksi muunnettaessa, vaan jo itse Venturassa. Kuvien liittäminen vanhalla versiolla tehtyyn materiaaliin myös hankalaa, kuva- ym. teksti siirtyi kuvan ympärille, ellei kuva vienytkin tarkasti tilaa sivun oikeasta marginaalista vasempaan. Kaavoja sisältävää materiaalia käsiteltäessä löytyi ohjelman kaavaeditorista virhe, jota ei aiemmissa versioissa ole ollut. Virhe saatiin kuitenkin kierrettyä valmistajan ohjeiden avulla. Version 3.0 ongelmia todennäköisesti lisäsi myös se, että materiaali on toimitettu Macintosh-koneessa käytettävällä Venturalla. Ohjelma kaatuu melko usein, varsinkin versiolla 3.0 taitetun aineiston kanssa, joten käsiteltävään tiedostoon tehdyt muutokset piti tallentaa välittömästi.

Aineiston toimitus

ELEKTRA-projektissa on pilottiaineistoa otettu vastaan julkaisijoiden toimittamassa muodossa. Mitään rajoituksia ei siis ole asetettu, vaan aineisto on käsitelty aina kun se on ollut teknisesti mahdollista. Jatkoprojektissa keskityttäne lähinnä valmiiksi elektronisessa muodossa olevaan aineistoon. Kuten aiemmin todettiin, pyrkimys on saada aineisto valmiina PDF-tiedostoina, mutta tähän tuskin koskaan täysin päästään, josta johtuen tulee projektin ohjelmisto- ja ohjelmaversiokantaa laajentaa tarpeen mukaan. Laitteistoa täydennetään Macintosh-työasemalla, jotta aineiston vastaanotossa ja käsittelyssä ohjelmien ja laitteiden yhteensopimattomuudesta aiheutuvat ongelmat voidaan minimoida.

Aineistoa on saatu painettujen dokumenttien lisäksi Post Script-muodossa ja PDF-tiedostoina. Näiden aineistojen vastaanotto ja jatkokäsittely on onnistunut ongelmitta.

Projektille on toimitettu aineistoa myös sellaisilla medioilla, joiden lukemiseen ei ole projektin laitteilla ilman lisähankintoja onnistunut. Tällaisia medioita ovat olleet esimerkiksi PC:n ja

Macintoshin 100 Mt:n ZIP-levykkeet sekä Macintoshin JAZ-levy. PC:n ZIP levyke on saatu avattua Helsingin yliopiston atk-keskuksen laitteilla. Macintoshin JAZ-levy on purettu Yliopistopainossa ja aineisto siirrettiin FTP:llä Linux-koneen kautta PC:lle. Macintoshin ZIP-levyketä ei ole kyetty lukemaan. ELEKTRAN käyttöön on nyt hankittu oma ZIP-levykeasema pc:lle sekä Macintosh-laitteisto, jossa on myös ZIP-asema.

7.4 Digitointityön edellyttämät resurssit

Aineiston käsittelystä on kerätty runsaasti erilaista seurantatietoa. Esimerkiksi artikkelien käsittelyn edellyttämä aika, käsittelymenetelmät, saadun aineiston ja tuotetun aineiston muoto, aineiston koko painettuina sivuina, valmiin aineiston koko sekä kaikki esiintyneet ongelmat on kirjattu ylös. Tarkkoja artikkelikohtaisia tietoja ei kuitenkaan ole saatu aina rekisteröityä, koska useampi henkilö, joilla on ollut myös erilainen osaamistaso, on osallistunut aineiston käsittelyyn, eikä kaikkien keräämä tieto ole täysin yhteismitallista. Painettu materiaali on lähes aina skannattu yhtenä kokonaisuutena, samoin julkaisuissa esiintyvät kuvat.

Dokumenttien käsittelyaikaan vaikuttaa merkittävästi se muoto, jossa aineisto on toimitettu projektille. Koko aineistossa yhden sivun käsittelyyn käytetyn ajan keskiarvo on 3.15 minuuttia. Käsittelyajat vaihtelevat kuitenkin suuresti. Mikäli aineisto on valmiiksi elektronisessa muodossa, eikä artikkeleihin ole tarpeen tehdä minkäänlaisia muutoksia tai korjauksia, on konvertointiprosessi PDF-tiedostoksi erittäin nopeaa. Noin 20-sivuinen artikkeli konvertoituu parissa minuutissa. Ääriarvot sivukohtaisissa testiaineiston käsittelyajoissa vaihtelevat 11 sekunnin ja 15 minuutin välillä.

Paras tulos on saavutettu väitöskirjalla, joka toimitettiin neljällä PC-levykkeellä ja joka oli taitettu PageMaker-ohjelmalla. Väitöskirjassa oli 196 sivua, eikä ainuttakaan kuvaa. Teos oli kuitenkin jaettu 14:ta erilliseen tiedostoon, jotka ensin piti PageMakerissä yhdistää yhdeksi kokonaisuudeksi. Adoben Exchange -ohjelmassa tehtiin lopuksi sisällysluettelo manuaalisesti sekä sivujen pienoisikonit ja muutettiin artikkelin käyttöön vaikuttavia oikeuksia (tulostuskielto). Käsittelyaika kokonaisuudessaan oli 38 minuuttia eli yhden sivun osalta noin 11 sekuntia. Toisessa ääripäässä on painettuna toimitettu suomenkielinen julkaisu, jossa oli 82 sivua ja runsaasti kuvia. Teksti ja kuvat skannattiin, tekstille suoritettiin tunnistus, teksti oikoluettiin Wordin oikolukuohjelmalla, kuvat liitettiin mukaan ja lopuksi muunnettiin osa artikkeleista PDF:ksi ja osa HTML:ksi. Yhden sivun käsittely vei keskimäärin 15.03 minuuttia. PDF:n osalta tilanne oli hieman parempi (13.50 minuuttia/sivu) kuin HTML:n (16.26 minuuttia/sivu). Muu projektissa käsitelty aineisto sijoittuu näiden ääripäiden väliin.

Verrattaessa PDF:n ja HTML:n käsittelyaikoja, on PDF:n tuottaminen helpompaa ja nopeampaa. Konvertoinnin jälkeen PDF on aina valmis, mutta HTML-tiedostoille joudutaan lähes aina tekemään korjailutoimenpiteitä, vaikka alkuperäinen dokumentti olisikin mahdollisimman hyvin esimerkiksi Wordin tyylimääritysten mukaan kirjoitettu.

Seuraavassa taulukossa on esitetty eri käsittelymenetelmien keskimääräisiä aikoja artikkeliaineiston osalta. Menetelmien vertailu on vaikeaa. Vaikka käsittelymenetelmä olisikin sama, ei käsittelyn viemä aika ole suoraviivaisesti sama kaikilla julkaisuilla. Samankin käsittelytavan ajoissa voi olla suuria eroja esimerkiksi riippuen siitä, onko julkaisussa kuvia ja mikäli on, niin onko ne erikseen skannattava ja liitettävä dokumenttiin, vai ovatko ne jo valmiina dokumentissa. Alkuperäisen lähdetiedoston tuotantoympäristöllä on myös suuri vaikutus käsittelyaikaan. Näiden seikkojen vaikutuksia tarkemmin erittelemättä ovat taulukossa esitetyt ajat keskimääräisiä ja ainoastaan suuntaa antavia. Kaikkia menetelmiä, paitsi Word muunnoksista HTML-muotoon, on testattu. Tällaisen aineiston käsittely vienee hieman WordPerfectillä tuotettua ai-

neistoa pitempään, koska WP:n oma konversio tuottaa valmiimpaa HTML-koodia kuin mitä tyylimäärityksiä käyttämättömistä Word-tiedostoista saadaan. Suora konversio PageMakeristä HTML:ksi oli sen verran hankalaa, johtuen mm testimateriaalin palstoituksesta, että käytännöllisempi tapa oli muuntaa teksti ensin Word-tiedostoksi ja vasta siitä HTMLksi.

Alkuperäismateriaalin toimitusmuoto	Käsittelyohjelma	Tuotettu tiedostomuoto	Käsittelyaika/sivu
Ventura	Distiller	PDF	43 sek
PM	Distiller	PDF	47 sek
Word	Distiller	PDF	2.42 min
WP	Distiller	PDF	3.37 min
painettu	Capture, ImageOnly	PDF	1.54 min
painettu	Capture, Reviewer	PDF	3.30 min
painettu	OCR, Word	PDF	7.06 min
PM	Word, HTML-editori	HTML	7.27
Word	HTML-editori	HTML	-
WP	WP, HTML-editori	HTML	3.07
painettu	OCR, Word, HTML-editori	HTML	11.20

Taulukko 1. Artikkeliaineiston keskimääräisiä käsittelyaikoja.

Saman käsittelymenetelmän vaatiman ajan ero tulee selvimmän esille, kun tarkastellaan kahta väitöskirjaa. Molemmat väitöskirjat oli taitettu PageMaker-ohjelmalla, toinen Macintosh- ja toinen PC-ympäristössä. Macintoshilla taitettu työ toimitettiin nauhalla, jollaisen lukemiseen sopiva nauha-asema löytyi lopulta Yliopistopainosta. Konversio sotki kirjasimet aiemmin kerrotulla tavalla. Väitöskirjassa oli 322 sivua ja se sisälsi myös joitakin kuvia, osa oli mukana valmiina, osa oli skannattava. Käsittely vei aikaa kaikkiaan noin 28 tuntia eli 5.13 minuuttia/sivu. Pc:llä taitetussa väitöskirjassa oli 196 sivua, eikä se sisältänyt yhtään kuvaa. Käsittely vei kaikkiaan 38 minuuttia eli 11 sekuntia/sivu. PageMaker-tiedostona toimitettu monografia, jossa lähes kaikki kuvat olivat valmiina mukana käsittelyaika oli 2.15 minuuttia/sivu.

Aineistojen käsittely on mahdollista saada suhteellisen vaivattomaksi ja nopeaksi, jos aineisto tuotetaan valmiiksi elektroniseen muotoon tarvittavia tyylimäärityksiä noudattaen ja käytetään ohjelmien uusimpia versioita. Näin säästetään myös palvelun tuotantokustannuksia.

7.5 Tiedostojen koot

Dokumenttien tiedostokoot vaihtelevat huomattavasti. Jos tiedosto sisältää pelkästään tekstiä, on HTML-versio kooltaan hieman pienempi kuin PDF, mutta mikäli mukana on myös kuvia kas-

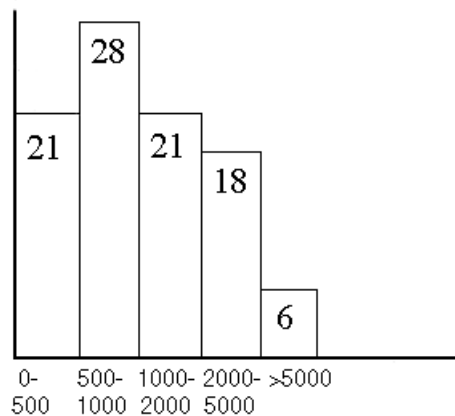
vaa HTML-tiedoston koko suuremmaksi kuin vastaava PDF-tiedosto. Tämä johtuu PDF:n pakkausominaisuuksista. Muutama esimerkki tästä.

Erään julkaisun pelkkää tekstiä sisältävien tiedostojen koot HTML-muodossa olivat 16, 17, 27, 35, 73 ja 75 Kt ja vastaavasti PDF-muodossa 23, 24, 39, 43, 76 ja 81Kt. Kuvia sisältävässä julkaisussa HTML-muodossa olevien tiedostojen koot (teksti + kuva/kuvat) 55, 75, 85 ja 98 Kt ja samat PDF-muodossa 34, 39, 48 ja 51 Kt. Esimerkissä olevat tiedostot ovat melko pieniä ja sisältävät vain yhden tai kaksi kuvaa, joten erot eivät vielä ole kovin suuria, mutta suuremmissa ja ennen kaikkea paljon kuvia sisältävässä materiaalissa erot kasvavat jo merkittäviksi. Taulukossa 7.2 on eri toimitusmuodossa ja eri käsittelytavoilla tuotetun aineiston keskimääräiset koot laskettuna sivua kohden.

Toimitusmuoto	Käsittelytapa/tiedostomuoto	Tiedoston koko/sivua
Ventura	Distiller/PDF	11,3 Kt
PM	Distiller/PDF	25,7 Kt
Word	Distiller/PDF	7,5 Kt
WP	Distiller/PDF	6,3 Kt
Painettu	Capture, ImageOnly/PDF	65,7 Kt
Painettu	Capture, Reviewer/PDF	10,2 Kt
Painettu	OCR, Word/PDF	43,1 Kt (runsaasti kuvia) 6,2 Kt (pelkkää tekstiä)
PM	Word, HTML-editori/HTML	25,7 Kt
Word	HTML-editori/HTML	—
WP	WP, HTML-editori/HTML	2,7 Kt
Painettu	OCR, Word, HTML-editori/HTML	4,8 Kt

Taulukko 2. Tiedostojen keskimääräiset koot/sivu eri toimitusmuodoista eri käsittelymenetelmillä tuotettuna.

ELEKTRAN testidokumenttien koko on keskimäärin 1,29 Mt. Julkaisu voi sisältää yhden tai useampia artikkeleita. Pienin lehden yhden numeron sisältävä tiedosto on 144 Kt (PDF) ja suurin 10,5 Mt (PDF). Kuvassa 1 on esitetty julkaisujen kokojen jakauma. Väitöskirjojen osalta pienin tiedostonkoko on 817 Kt (196 sivua, ei kuvia) ja suurin 15,342 Mt (537 sivua ja 78 kuvaa). Suurimmillaan suurikokoisin väitöskirja oli PostScript-muodossa lähes 70 Mt. Tämän kokoisten (15 Mt) dokumenttien latautuminen on jo kohtalaisen hidasta, mutta Adoben Exchange -ohjelman uusimman version määrittelyillä ja uusimmilla selaimilla saadaan dokumentit latautumaan myös yksi sivu kerrallaan. Seuraavassa kuvassa on tiedostokokojen jakauma koko ELEKTRAN testiaineistossa:



Kuva 6. Julkaisujen kokojakauma.

7.6 Johtopäätökset

Jotta aineiston muuntaminen HTML-muotoon olisi nopeaa ja tapahtuisi mahdollisimman automaattisesti, tulee toimitettavan materiaalin olla valmiiksi jossakin elektronisessa muodossa. Paras vaihtoehto on jokin tekstinkäsittelyohjelma, jolla tuotettu materiaali on kirjoitettu tyyli-määritysten mukaisesti. Tällöin konversiossa syntyy vähiten korjaustyötä. Pitkän, painetussa muodossa vastaanotetun dokumentin muuntaminen HTML-muotoon on kallista, koska oikolu-kuohjelmista huolimatta tekstin tarkistaminen vaatii paljon käsityötä.

PDF-muotoon konvertoitaessa painetun materiaalin käsittely ei ole aivan yhtä hidasta, mikäli sivut voidaan tallettaa kuvina. Ongelmaksi muodostuvat kuitenkin tiedostojen suuri koko sekä tekstin huonompi laatu. Tiedostojen koko kasvattaa arkiston kokoa ja tuntuu käyttäjille doku-menttien hitaana latautumisenä. Valmiiksi jossain elektronisessa muodossa toimitettavan ai-neiston tiedostomuodolle ei ole rajoituksia, kun sitä halutaan muuntaa PDF-muotoon.

Digitaalisessa muodossa vastaanotettava aineisto asettaa kuitenkin aineiston käsittelylle joitakin vaatimuksia. Projektin aikana kertyneet kokemukset osoittavat, että uusien versio kaikista tarvit-tavista ohjelmista ei välttämättä ole riittävä, vaan projektilla tulee olla käytössä myös kaikki ne vanhemmat ohjelmaversiot, joilla materiaalia toimitetaan. Yhteensopivuus eri ohjelmaversioi-den välillä ei käytännössä aina tunnu toteutuvan. Samoin projektilla tulee olla PC- sekä Macin-tosh- työasemat.

Aineistoa on viimeaikoina toimitettu enenevässä määrin myös erilaisilla pakkaavilla levykkeil-lä, joista yleisin näyttää olevan 100 Mt:n ZIP-levyke. Tällaisia levykkeitä lukevan levykease-man on hankittu projektin käyttöön ja samaa, ulkoista levykeasemaa voidaan käyttää sekä PC-että Macintosh-koneissa. Projektille toimitettua alkuperäismuodossa olevaa aineistoa voidaan myös tallentaa näille levykkeille.

Kirjasinkokoelman tulee olla melko kattava. Esimerkiksi PageMaker-tiedostoina toimitetun aineiston osalta on useissa artikkeleissa käytetty jotakin korvaavaa fonttia, kun tiedostot on avattu ELEKTRAN PC:llä. Kyrilliikkaa sisältävää aineistoa on jäänyt kokonaan muuntamatta verkkokäyttöiseen muotoon, koska oikeaa fonttia ei ole ollut käytettävissä. Artikkeleissa, joissa on ollut muutamia yksittäisiä kyrillisiä merkkejä sisältäviä sanoja, on konvertoitu käyttämällä PC:n mukana valmiina olevia lisä-/erikoismerkkejä.

Aineistoa käsiteltäessä niiden tekstinkäsittely- ja julkaisuohjelmien käytön osaaminen, joilla aineisto on alunperin kirjoitettu ja/tai taitettu, tulee olla hyvää ja syvällistä ja ohjelmien käyttökoulutukseen kannattaa kiinnittää huomiota.

8. Elektroniset maksujärjestelmät

Tässä luvussa käsitellään ELEKTRA-projektin kannalta käyttökelpoisia elektronisia maksujärjestelmiä elektronisessa muodossa olevien dokumenttien, esimerkiksi PDF-muodossa olevan artikkelin, käytön rahastamiseen. Maksujärjestelmillä tarkoitetaan ohjelmistoja ja laitteita, joilla voidaan hoitaa dokumenttien käytön, esimerkiksi tulostamisen hallinta ja maksaminen joko asiakaspääteeltä tai asiakkaan omalta koneelta WWW-palvelun kautta.

Koska maksettavat summat useimmissa tapauksissa ovat erittäin pieniä, pankkikortti-, luottokortti-, pankkisiirto- tai laskutus pohjaiset ratkaisut eivät ole soveltuvia. Lisäksi useimmissa tapauksissa asiakas haluaa saada tuotteen käyttöönsä välittömästi.

8.1 Johdanto

Internetissä käytettiin jo vuoden 1996 lopulla 27 erilaista maksujärjestelmää. Järjestelmäsekaannus ei ole sittemmin ainakaan lieventynyt. Edelleenkin on erittäin vaikea suosittaa mitään standardia yleiseen käyttöön, koska ”standardimarkkinoilla” ei ole ainakaan toistaiseksi selvää voittajaa. Eri standardeihin perustuvat järjestelmät eivät keskustelee toistensa kanssa; sekä asiakas- että palvelinohjelman pitää olla samaa perhettä. Jos siis halutaan tukea useampaa standardia, pitää hankkia useampia järjestelmiä. Käytännössä tämä tulisi palvelujen ylläpitäjille liian kalliiksi. Palvelujen käyttäjille kyse ei ole niinkään kustannuksista – muutamat palvelut ovat käyttäjälle hyvin halpoja – vaan käyttömukavuudesta. Opiskelijoiden ja tutkijoiden kannalta olisi varsin ongelmallista, jos he joutuisivat maksamaan eri yliopistoihin eri järjestelmää käyttäen.

Luotettavien maksullisten palvelujen luonti Internetiin on vasta käynnissä. Verkossa on toki käyty kauppaa jo hyvän aikaa, mutta menetelmin jotka vaativat asiakkaalta rohkeutta. eLib-projektien suosituksessa (katso <http://www.ukoln.ac.uk/services/elib/papers/other/standards/stand2.html>) todetaan, että on ihme, että Internetin kaltaisessa turvattomassa verkossa käydään niinkin paljon kauppaa kuin on tehty. Vuoden 1997 aikana tilanne on parantunut, koska muutamia suuria järjestelmiä – kuten Visan SET – on tullut testikäyttöön, ja muutamat sovellukset ovat rutiinikäytössä. Yleisesti ottaen tilanne on kuitenkin edelleen sekava, ja johtopäätös sama kuin se mihin eLib-raportissa päädyttiin: ei ole mahdollista antaa kansallista suositusta valittavasta standardista. Tässä dokumentissa onkin vertailtu kahta Suomessa jo tuotantokäytössä olevaa järjestelmää, pankkien Avant-korttirahaa ja EUNETin Ecashia.

Käyttäjien autentisointi

Jos maksujärjestelmiä on tarjolla liiankin kanssa, muutamia dokumenttien käytön valvontajärjestelmän perusosia puuttuu edelleen. Kansallisen järjestelmän laadinnassa yksi perusedellytys on käyttäjien autentisointi.

Tätä kirjoitettaessa käyttäjien autentisointi perustuu yksinkertaisiin ja osin tehottomiin menetelmiin. Perinteinen käyttäjätunnukseen ja salasanaan perustuva autentisointi on erittäin raskas menettely, jos käyttäjiä on paljon. Kustantajat eivät ole edes yrittäneet vaatia käyttäjätunnuksiin perustuvaa suojausta kansallisten lisenssien nojalla käytettäviin tietokantoihin, koska tuhansien k-tunnusten ja salasanojen luonti ja ylläpito vaatii paljon työtä. Lisäksi, jos käytetään protokollaa, joka ei hyödynnä salausta, kuten Telnetiä tai HTTP:tä, tunnukset ja salasanat ovat helposti luettavissa verkosta ja lähestulkoon hyödyttömiä.

Käytön valvonta IP-osoitteen perusteella on suosittu menetelmä helppoutensa vuoksi – määrittäminen tarvitsee tehdä periaatteessa vain kerran, jos lisenssi on hankittu koko organisaatiolle. Tietoresurssin käyttäjän ei tarvitse tehdä mitään, edellyttäen että hänen laitteensa on määritelty suojaohjelmaan oikein. Tämän lähestymistavan ongelma on, että IP-osoite on helppo väärentää. Käyttäjän näkökulmasta tilanne on ongelmallinen vain, jos ”varas” on tietokannassa samaan aikaan oikean käyttäjän IP-osoitteella, koska tämäntyyppinen päällekkäisyys sotkee verkkoliikenteen. Käytännössä IP-osoitteiden päällekkäinmeno on erittäin harvinaista, joten asiakkaan kannalta IP-osoitteen käyttö on hyvä ratkaisu. Palvelun ylläpitäjä tarvitsee kuitenkin pitkällä aikavälillä tehokkaamman suojan.

Tuotannossa olevissa maksujärjestelmissä identifiointi perustuu ID-tunnukseen, joka toimii vain asianomaisen järjestelmän sisällä. Tämä on yksi keskeisistä syistä maksujärjestelmien yhteismitattomuuteen. Suomessa pyritään ottamaan käyttöön kansalaisten sähköinen identiteetti asteittain vuodesta 1998 alkaen. Tuotantokäyttöön päästyään tämä palvelu tarjoaa erinomaisen perustan luotettavien dokumenttien välitysjärjestelmien rakentamiselle. Aineiston välitys voi olla kahdensuuntaista: kirjastosta käyttäjälle tai päinvastoin; jälkimmäistä proseduuria tarvitaan esimerkiksi vapaakappaleaineiston tai ELEKTRAn aineiston lähettämisessä. Dokumentit lähetetään salattuna, ja vastaanottaja voi helposti varmistua siitä että lähettäjä on juuri se henkilö tai organisaatio joka hän (se) väittää olevansa. Myös nykyiset ja tulevat maksujärjestelmät voidaan ainakin periaatteessa muokata käyttämään kansallista sähköistä identiteettiä tai tarjoamaan sen yhtenä vaihtoehtona.

Käyttäjien autentisointi on yksi luotettujen järjestelmien rakentamisen ehto. Niiden avulla elektronista aineistoa voidaan lainata käyttäjille painetun aineiston tapaan. Itse käyttö voisi olla maksutonta sen jälkeen kun materiaali on hankittu kirjaston kokoelmiin. Luotettu järjestelmä kontrolloi kattavasti sitä mitä aineistolle tehdään, ja laina-ajan kuluttua umpeen dokumentin käyttöoikeus lakkaa, ja vastaavasti kirjaston virtuaaliseen hyllyyn tulee yksi lainattava virtuaalinen nide lisää.

ELEKTRAn tarpeet

ELEKTRA-tietokannan käytön valvonta perustuu toistaiseksi IP-osoitteiden käyttöön. Käyttäjien kannalta tämä ratkaisu toimii kohtuullisen hyvin.

Käyttäjät pääsevät viitetiedoista itse dokumenttiin klikkaamalla viitteessä olevaa URL-tunnusta. Linkki toimii niin kauan kun dokumenttien sijaintipaikka ei muutu. Dokumentteja ei ole identifioitu; käytännössä artikkeliaineiston identifiointiin voitaisiin käyttää SICI-koodia ja monografia-aineiston koodaukseen ISBN:ää tai NBN-tunnusta. Pitkällä aikavälillä monografia-aineiston tunnistus voidaan viedä kirjan lukujen tasolle BICI-koodilla, jos tähän on tarvetta.

Aineisto on tallennettu joko PDF- tai HTML-muodossa. PDF-tiedostojen katselu on ilmaista, mutta tulostamisesta pitää maksaa. Käytön kontrollointi tehdään PDF-tiedostomäärittysten avulla. Jos siis ELEKTRAA halutaan laajentaa organisaatioihin, jotka haluavat vain katsella artikkeleita, tarvitaan nykyisellä tekniikalla jokaisesta dokumentista kaksi kopiota, joista toisessa on tulostuksen esto. Toinen, tulostettava kopio, pitää suojata IP-osoitteen avulla siten, että aineiston siirto onnistuu vain niille kirjastoille, jotka haluavat tulostusoikeuden. Eri kopiot voidaan tallentaa esimerkiksi eri alihakemistoihin kontrolloinnin mahdollistamiseksi.

ELEKTRAn maksujärjestelmän valintaa rajoittava tekijä on se, että artikkelikohtaiset maksut ovat varsin pieniä. Tämä sulkee pois luottokortteihin perustuvat maksujärjestelmät kuten esi-

merkiksi SET:in. Niissä oheiskulut ovat sen verran korkeat, että minimimaksu on käytännössä noin 20 mk.

Nykyisillä maksujärjestelmillä on ELEKTRAN kannalta se puute, että ne toimivat aineiston lähetysvaiheessa, ei siis tulostettaessa. Tämä ongelma voidaan kiertää edellä kuvatuilla kahden dokumentin järjestelmällä, jolloin siis katseltava dokumentti ei ole maksullinen, mutta tulostettava on. Kirjaston ei tarvitsisi hoitaa laskutusta, vaan se menisi automaattisesti ”kortilta”. Jos tulostusta halutaan edelleen kontrolloida siten, että sitä tehdään keskitetysti lainaustoimistossa, voidaan maksamiseen tietenkin käyttää kenen tahansa validia korttia, mutta maksajan on tultava lainaustoimiston koneelle, jonka IP-osoitteen ELEKTRA hyväksyy. Vaihtoehto on, että tulostuksen voi tehdä kuka tahansa jolla on hyväksyttävä kortti, ja hyväksytään se riski että artikkelista jää kopio asiakkaan koneelle.

Pitkällä aikavälillä ELEKTRAssa luotetut järjestelmät tarjoavat mielenkiintoisen vaihtoehdon nykykäytännölle. Toistaiseksi on kannattavaa jatkaa tuotantoa nykyiseen tapaan, ja suorittaa pilottikokeilu tulostettavien artikkelikopioiden jakelussa hyödyntäen jompaa kumpaa tai molempia jäljempänä kuvatuista Avant- ja Ecash-järjestelmistä. Järjestelmätoimittajien halukkuus pilottikokeiluihin alennetuin kustannuksin kannattaa selvittää.

ELEKTRAssa on syytä seurata saman alan EU-hankkeita päällekkäistyön välttämiseksi. Relevantteja ovat varsinkin SEMPER (<http://www.semper.org/>), Secure Electronic Marketplace for Europe ja Copinet (<http://www.dtv.dk/copinet/>). Muita saman alueen hankkeita ovat Copicat (<http://www.mari.co.uk/copicat/>), Gaia (<http://www.syspace.co.uk/GAIA/>), Imprimatur (<http://www.imprimatur.alcs.co.uk/>) ja Decomate (<http://www.lse.ac.uk/DECOMATE/>). Esimerkiksi Copinet käyttää Ecashia silloin kun dokumentin tilaaja on ns. ulkopuolinen.

8.2 Avant-korttiraha

Avant-korttiraha on tavalliseen pankkikorttiin liitettävä mikrosiru, johon voi siirtää pankkiautomaatilla haluamansa määrän rahaa (10–2000 mk). Avant-korttiraha -ominaisuus on saatavissa Meritan (<http://www.merita.fi/kortit/>), Osuuspankin (<http://www.osuuspankki.fi/akoht/kraha.htm>) ja Postipankin (<http://www.psp.fi/su/yksit/avant.htm>) pankkikortteihin. Esiladattuja, kertakäyttöisiä Avant-kortteja voi ostaa R-kioskeista.

Korttirahalla maksettaessa kortti luetaan kortin lukulaitteessa ja kortin mikrosirulta vähennetään ostoksen summa. Periaatteessa maksettavalla ostoksella ei ole minimihintaa, mutta käytännössä maksuprovisioista johtuen pienin hyväksyttävä yksittäisen ostoksen summa on 20 penniä¹⁷.

Kortilla ostaminen vastaa käteisellä rahalla ostamista, koska myyjälle ei välity tietoa maksajan henkilöllisyydestä. Myyjälle tosin välittyy kortin tunnistenumero, joka voidaan yhdistää käyttäjän henkilötietoihin, jos käyttäjä ne ilmoittaa esimerkiksi tavaroiden toimittamista varten.

Avant-korttirahaa voidaan käyttää myös maksettaessa ostoksia WWW:n kautta. Tietokoneeseen liitettävä kortinlukulaite maksaa tällä hetkellä 390 mk (hintaa sisältää postikulut ja alv:n, tilauslomake on osoitteessa <http://www.avant.fi/verkkomaksu/tilaus.html>). Kortinlukulaitteen mukana tulee myös käyttäjän koneella tarvittava AvantClient-ohjelma. Kortinlukulaite on saatavissa

¹⁷ Pienin siirrettävissä oleva rahayksikkö on 1 penni, joka on 5% provisio 20 pennin ostoksesta.

Windows 95 ja NT-käyttöjärjestelmiin. Kortinlukulaitteiden hintojen arvellaan laskevan, kun markkinoille tulee uusien valmistajien laitteita.

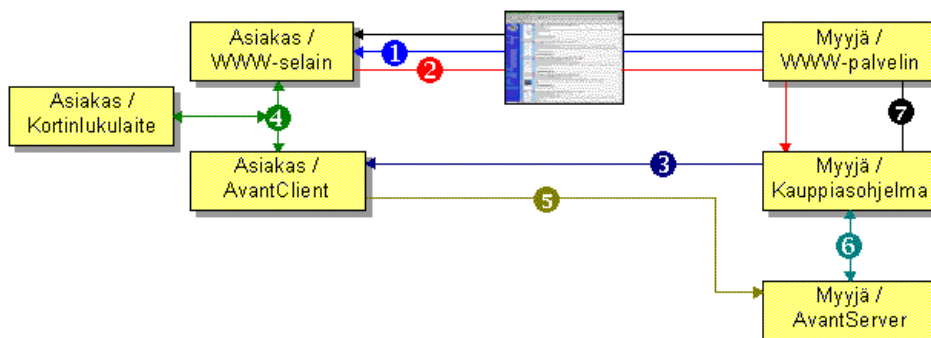
Vaadittavat ohjelmistot ja laitteet WWW-käytössä

Asiakkaalta Avant-korttirahan käyttäminen vaatii siis ainoastaan Avant-kortin ja internet-yhteyksillä varustetun PC-tietokoneen, johon on liitetty kortinlukulaite sekä AvantClient-ohjelmisto.

Myyjältä Avant-korttirahamaksujen vastaanottaminen vaatii sopimuksen Automatia Rahakortit Oy:n kanssa, sekä WWW-palvelinohjelmiston (Microsoft, Netscape), WWW-kauppiasohjelmiston ja AvantServer-ohjelmiston. Lisäksi myyjän kortille kertyneiden maksujen siirtämiseen clearing-järjestelmään tarvitaan X.25-yhteensopiva tiedonsiirto-ohjelma.

Avant-maksutapahtuma WWW:ssä

Kun asiakas haluaa maksaa ostoksensa Avant-kortilla hän laittaa Avant-korttinsa kortinlukulaiteeseen ja osoittaa hiirellä maksutapaa ilmaisevaa painiketta (Avant logo). Tämän jälkeen WWW-kauppiasohjelmisto käynnistää asiakkaan koneella olevan AvantClient-ohjelman ostokseen liittyvillä parametreilla (esimerkiksi tuote- ja hintatiedot). Käyttäjän hyväksyttyä maksun AvantClient lähettää maksun myyjän AvantServer-ohjelmalle (myyjän Avant-kortille) ja vähentää ostoksen hinnan asiakkaan kortilta, jonka jälkeen kauppiasohjelmisto toimittaa tuotteen asiakkaan ruudulle.



1. Asiakas saa WWW-selaimensa ruudulle luettelon (esimerkiksi hakua vastaavista) dokumenteista.
2. Asiakas laittaa Avant-korttinsa kortinlukijaan, valitsee maksutavaksi Avant-korttirahan ja klikkaa haluamansa dokumenttia, jolloin tieto halutusta dokumentista ja sen hinnasta menee myyjän WWW-palvelimella olevaan kauppiasohjelmaan.
3. Kauppiasohjelma kutsuu AvantClient -ohjelmaa antaen ohjelmalle parametreina tietoa ostettavasta tuotteesta ja ostoksen kokonaishinnan.
4. AvantClient -ohjelma pyytää asiakkaalta vahvistuksen ostoksen maksamiselle ja vähentää summan asiakkaan kortilta.
5. AvantClient -ohjelma siirtää rahat myyjän Avant-kortille.
6. AvantServer -ohjelma ilmoittaa kauppiasohjelmalle, että maksu on suoritettu ja tilattu dokumentti voidaan toimittaa.
7. Kauppiasohjelma toimittaa dokumentin asiakkaan ruudulle luettavaksi, tulostettavaksi tai levykkeelle talletettavaksi. Tapahtumatiedot tallentuvat asiakkaan kortille.

Kuva 7. Kaaviokuva Avant maksutapahtumasta.

Kustannukset

Käyttäjän kustannukset muodostuvat kortinlukulaitteesta (tällä hetkellä 390 mk) ja rahan lataamisesta kortille (2 mk/kerta). Pankkien perimät maksut Avant-ominaisuuden liittämiseksi pankkikorttiin vaihtelevat: Meritan pankkikorttiin Avant-omaisuus maksaa 50 mk/kortti kertamaksuna, Osuuspankissa maksu on 30 mk/vuosi (maksua ei peritä alle 18-vuotiailta) ja Postipankissa ensimmäiseen korttiin ilmainen ja seuraaviin kortteihin 30mk/3 vuotta. Esiladatut kertakäyttöiset kortit maksavat käyttäjälle hieman enemmän kuin kortilla oleva rahamäärä: 50 mk:n kortti maksaa 55 mk ja 100 mk:n kortti 105 mk.

Asiakkaan käyttäessä omaa tietokonettaan kustannuksiin on luonnollisesti laskettava myös normaalit tietoliikennekulut. Jos asiakas käyttää tiedostojen ostamiseen kirjaston asiakaspäätettä, asiakkaan kustannuksista jää luonnollisesti pois kortinlukulaitteen ja yhteysajan hinta.

AvantServer-ohjelmisto maksaa 5000 mk + alv sisältäen lisäksi Avant-verkkomaksamisen rajapintamäärittäykset, AvantClient-ohjelman, 2 kortinlukijaa, 2 testiturvamodulia ja 6 testikorttia. AvantServer-ohjelmiston käyttö ei ole sidottu mihinkään tiettyyn WWW-kauppiasohjelmistoon.

Myyjän Avant-kortille kertyneiden rahojen siirtämiseen Automatia Rahakortit Oy:n clearing-järjestelmään tarvitaan X.25 pakettiverkkoyhteys sekä soveltuva tilitysohjelma¹⁸.

Maksutapahtumista peritään provisio, joka vaihtelee maksun suuruuden mukaan. Pienimmillään provisio on 1% (yli sadan markan ostotapahtuma), 20p + 1% ostoksen summasta (alle sadan markan ostotapahtuma), mutta kuitenkin maksimissaan 5% ostotapahtuman summasta. AvantServer-ohjelmiston vuosilisenssi on 200 mk.

Turvallisuus

Avant-korttiraha on käyttäjälle yhtä turvallista kuin käteinen raha. Jos käyttäjä hukkaa korttinsa, kortilla oleva raha on kenen hyvänsä käytettävissä, koska kortilla maksettaessa maksajan henkilöllisyyttä ei mitenkään tarkisteta.

Maksutapahtuman yhteydessä rahakortti ja myyjän turvamoduli tarkistavat toistensa aitouden. Ladattavaan rahakorttiin on talletettu kortin tunniste-, kontrolli-, saldo- ja tapahtumatietoja mm. rahan liikkeellelaskijan tunnus, kortin voimassaoloaika, kortin saldo, valuuttakoodi- ja yksikkö sekä 10 viimeisimmän maksutapahtuman tiedot.

Verkossa siirtyvä maksumääräys (rahakortilta myyjälle) siirtyy salattuna ja sinetöitynä tiedostona.

Lähteet

- <http://www.avant.fi>
- Avant-korttiraha koulutustilaisuuden (26.11.1997) materiaali

¹⁸ Automatian ilmoituksen mukaan korttirahan lähettämiseen soveltuvia ohjelmia myyvät ainakin seuraavat yritykset: ADS-Anker Data System Oy, Jukka Nieminen (09-5497 1435) ja Analyste Oy, Vesa Sarvala (09-268 4111).

8.3 EUNET Ecash

EUNET Ecash on digitaalinen raha, joka voi sijaita joko palvelimella olevassa Ecash-puskurissa tai käyttäjän koneella olevassa Ecash-lompakko-ohjelmassa. Käyttäjä siirtää rahaa Ecash-puskuriin normaalilla pankkisiirroilla, josta rahat edelleen siirretään käyttäjän tietokoneen kiintolevylle lompakko-ohjelmaan. Ecash-maksutapahtumat suoritetaan ja hyväksytään käyttäen lompakko-ohjelmaa.

Ecash-käyttäjäksi ryhtyminen, lompakko-ohjelma ja rahan siirtäminen puskuuriin ei maksa käyttäjälle mitään (paitsi mahdolliset pankkien perimät maksut tilisiirroista). Puskurissa voi olla maksimissaan 100 markkaa rahaa, mutta lompakossa olevan rahan määrää ei ole rajoitettu.

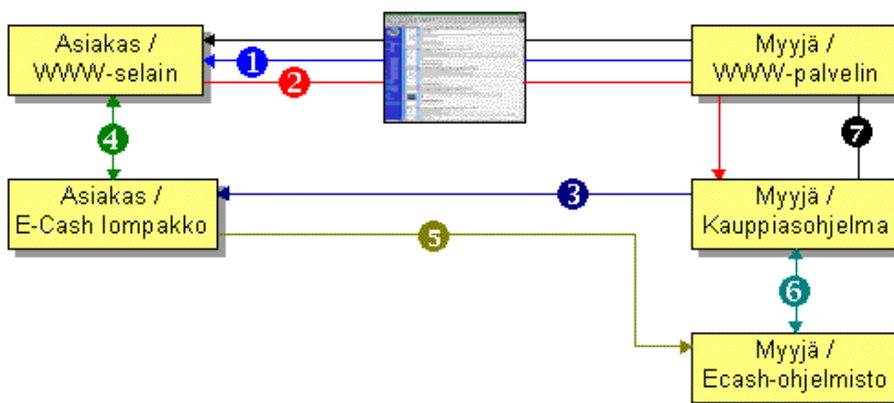
Vaadittavat ohjelmistot ja laitteet

Käyttäjältä vaaditaan tietokone, johon lompakko-ohjelma asennetaan, internet-yhteys maksujen suorittamiseen lompakko-ohjelmalla ja oma sähköposti-osoite. Lompakko-ohjelma on saatavissa Macintosh-, Unix- ja Windows-käyttöympäristöihin. Ecash-käyttäjäksi voi liittyä osoitteessa <http://www.eunet.fi/ecash/>

Myyjä tekee sopimuksen EUNETin kanssa Ecash kauppiaksi ryhtymisestä. Myyjän Ecash ohjelmisto voi sijaita joko myyjän omalla palvelinkoneella tai EUNETin palvelimella. Ecash-ohjelmisto on saatavissa useille eri WWW-palvelimille eri käyttöympäristöihin (ks. http://www.digicash.com/ecash/shop_win.html#support ja <http://www.digicash.com/ecash/eu/buildshop.html>).

Tuotteiden toimittamiseen ei välttämättä tarvita WWW-kauppiasohjelmistoa, koska Ecash-kauppiasohjelmisto voi huolehtia myös maksetun elektronisessa muodossa olevan tuotteen toimittamisesta. Jos myyntitapahtumaan halutaan liittää ostoskori-ominaisuuksia tms. jokin WWW-kauppiasohjelma on suotava.

Ecash-maksutapahtuma WWW:ssä



1. Asiakas saa WWW-selaimensa ruudulle luettelon (esimerkiksi hakua vastaavista) dokumenteista.
2. Asiakas valitsee maksutavaksi Ecashin ja klikkaa haluamansa dokumenttia, jolloin tieto halutusta dokumentista ja sen hinnasta menee myyjän WWW-palvelimella olevaan kauppiasohjelmaan.
3. Kauppiasohjelma kutsuu asiakkaan Ecash lompakko-ohjelmaa antaen ohjelmalle parametreina tietoa ostettavasta tuotteesta ja ostoksen kokonaishinnan.
4. Asiakkaan Ecash-ohjelma pyytää asiakkaalta vahvistuksen ostoksen maksamiselle.

5. Rahat siirretään asiakkaan lompakosta myyjälle. Rahojen oikeellisuus tarkistetaan pankista.
6. Myyjän Ecash-ohjelma ilmoittaa kauppiasohjelmalle, että maksu on suoritettu ja tilattu dokumentti voidaan toimittaa.
7. Kauppiasohjelma toimittaa dokumentin asiakkaan ruudulle luettavaksi, tulostettavaksi tai levykkeelle tallettavaksi. Tapahtumatiedot tallettavat asiakkaan lompakko-ohjelman lokiin.

Kuva 8. Kaaviokuva Ecash maksutapahtumasta.

Kustannukset

Asiakkaan kustannukset muodostuvat ainoastaan yhteysajan tietoliikennekustannuksista ja mahdollisista pankkien maksuista siirrettäessä rahaa tilisiirtona Ecash-puskuriin.

Myyjän kustannukset vaihtelevat sen mukaan sijaitseeko myyjän Ecash-ohjelmisto myyjän omalla vai EUNETin palvelinkoneella.

Myyjän omalla koneella olevan Ecash kaupan kustannukset ovat seuraavat: Alkukustannus 1000 mk, kuukausiveloitus 350 mk ja provisio maksusuorituksista 2%. EUNETin palvelimella vastaavat hinnat ovat: Alkukustannus 500 mk, kuukausiveloitus 350 mk ja provisio maksusuorituksista 2% (hinnat ilman arvonlisäveroa). Uudet kauppiat voivat saada pilottisopimuksen, jossa ensimmäisen kuuden kuukauden ajalta ei peritä kuukausimaksuja.

Turvallisuus

Ecash perustuu sähköisiin kolikoihin, joille luodaan satunnaisia sarjanumeroita. Sarjanumerot suojataan "kirjekuoriin" ja lähetetään allekirjoitettavaksi pankkiin. Pankin allekirjoituksen jälkeen ne lähetetään takaisin ja "kirjekuori" poistetaan. Pankin allekirjoitus jää kolikoihin. Jokaisen kolikon sarjanumero on erilainen, joten samaa kolikkoa ei voida käyttää kahdesti, koska kaikki Ecashillä maksetut suoritukset tulevat pankin tietoon. Maksajan anonymiteetti kuitenkin säilyy, koska kolikot ovat pankin allekirjoitushetkellä "kirjekuoressa", joten pankki ei tiedä kenen kolikoista on kysymys.

Maksutapahtuman yhteydessä, ennen tuotteen toimittamista, myyjän ohjelma lähettää asiakkaan toimittamat kolikot pankkiin ja odottaa pankin hyväksyntää kolikoille. Kun pankki on hyväksynyt kolikot, myyjän ohjelmisto voi toimittaa tilatun tuotteen. Kaikista Ecash-tapahtumista jää merkintä asiakkaan lompakko-ohjelman lokiin.

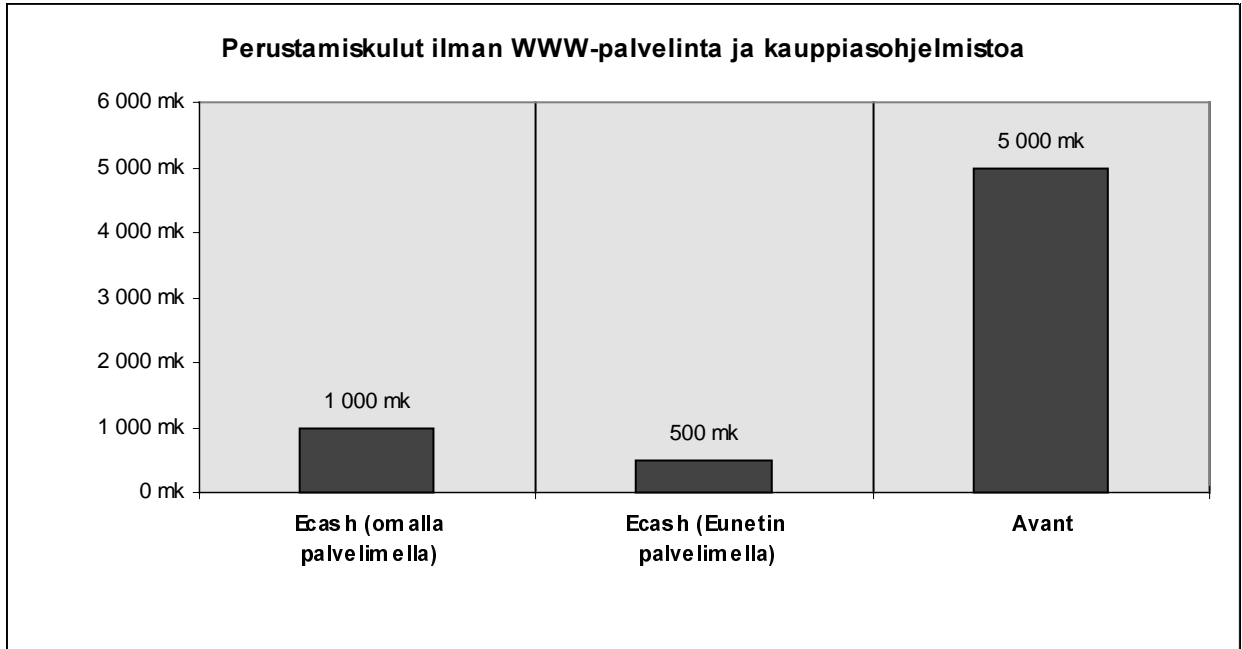
Lompakko-ohjelman käyttö ja rahojen siirtäminen puskurista lompakko-ohjelmaan on suojattu salasanalla. Kolikoiden siirto tapahtuu salatussa muodossa (RSA).

Lähteet

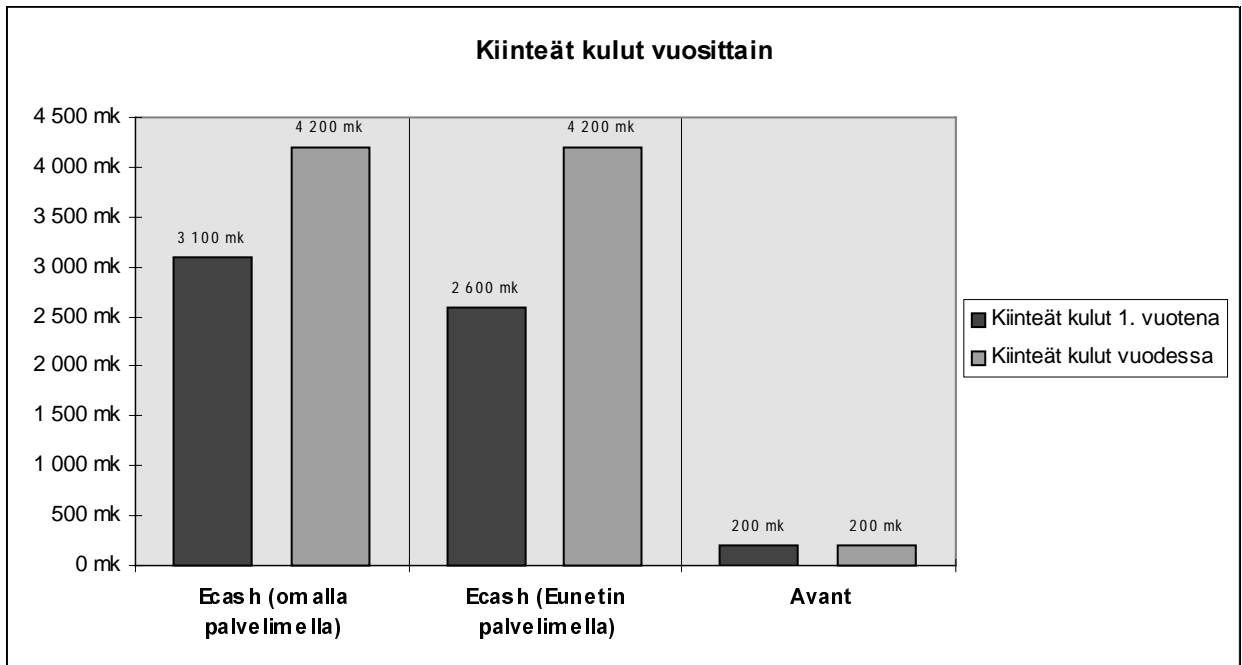
- <http://www.eunet.fi/ecash/>
- <http://www.digicash.com/ecash/>

8.4 Yhteenveto maksujärjestelmien kustannuksista kauppiaille

Perustamiskustannukset



Kuva 9. Perustamiskulut ilman WWW-palvelintä ja kauppiasohjelmistoa. (Hinnat ilman arvonlisäveroa).



Kuva 10. Kiinteät kulut (kuukausi- tai vuosimaksut) ensimmäisenä vuotena ja vuodessa ilman perustamiskuluja. Ecashissä huomioitu pilottikauppiashinta, eli 6 ensimmäistä kuukautta ilman kuukausimaksuja. (Hinnat ilman arvonlisäveroa).

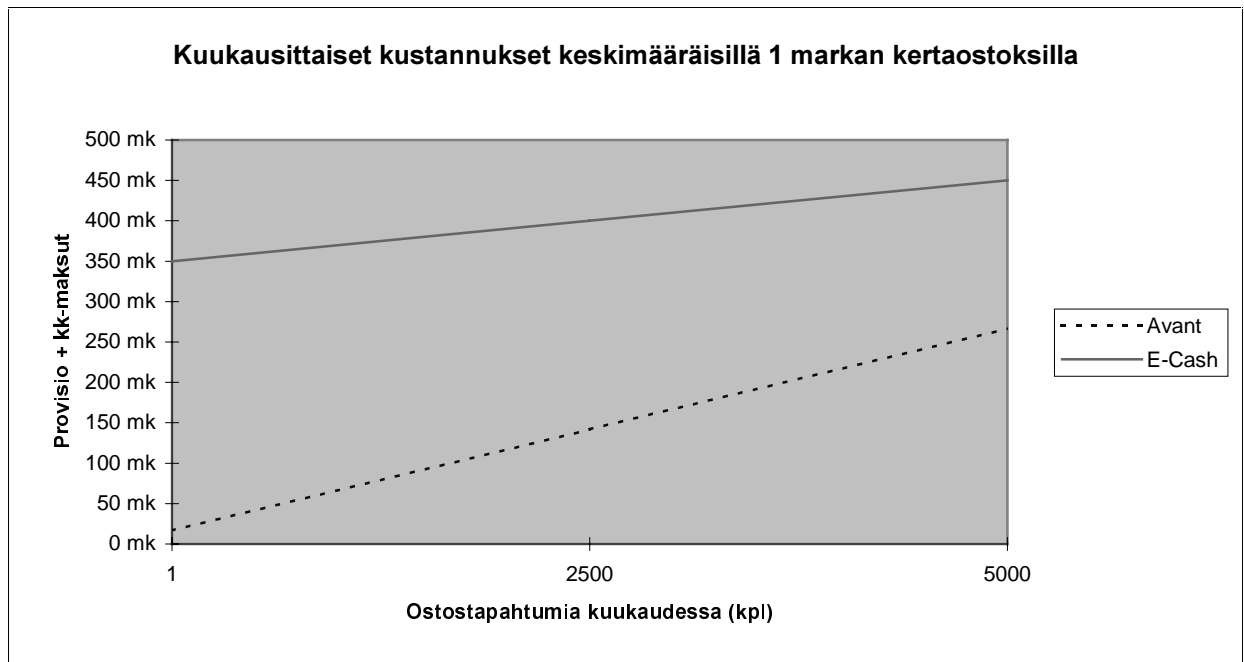
Kuukausittaiset kustannukset eri suuruisilla keskimääräisillä kertaostoilla

Seuraavissa diagrammeissa on laskettu kauppiaan kustannuksia eri hintaisilla keskimääräisillä ostostapahtumilla ja ostostapahtumien määrällä kuukaudessa.

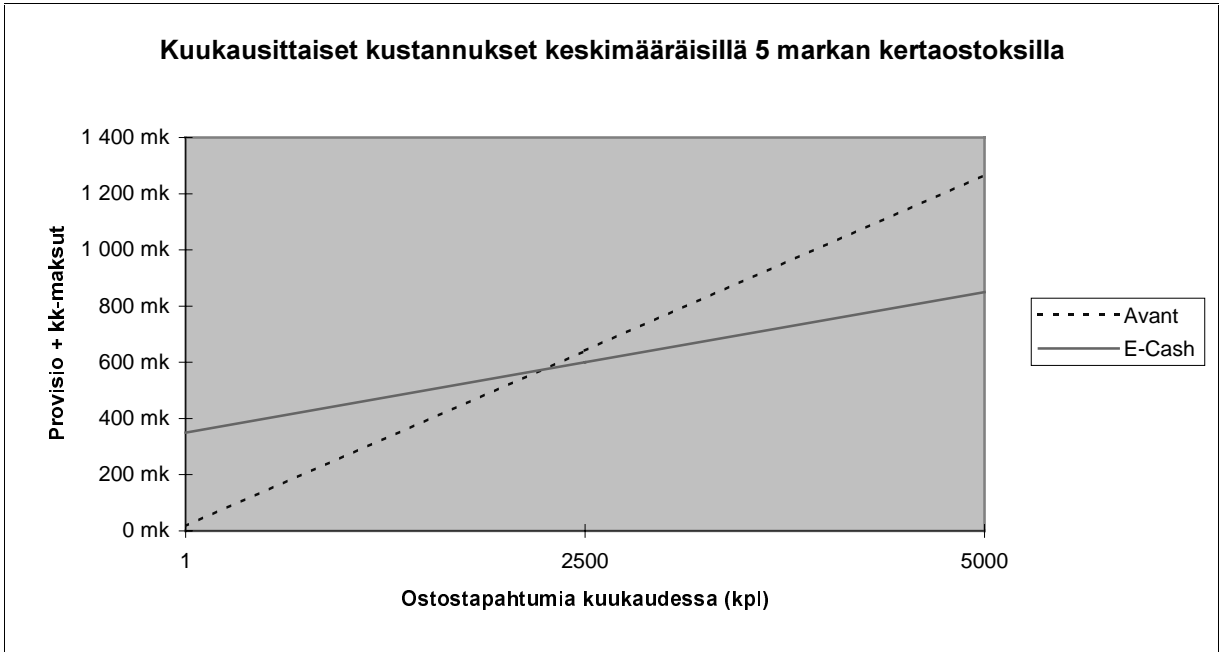
Eunetin Ecashissa proviisio on kiinteä 2% ostoksen summasta. Avant-korttirahalla proviision suuruus riippuu maksutapahtuman summasta. Proviisio on alle sadan markan ostoksilla 20 penniä + 1%, kuitenkin enintään 5% ostoksen summasta. Alle viiden markan ostoksista proviisio on siis aina 5% (20p + 1% = 5% viiden markan ostoksesta).

Kustannuksiin on laskettu mukaan kuukausimaksut, jotka Ecashissa ovat 350 mk/kk ja Avantil- la 16,67 mk/kk. Avantin kuukausimaksu on laskettu jakamalla AvantServer-ohjelmiston vuosil- lisenssi 200 mk kahdellatoista.

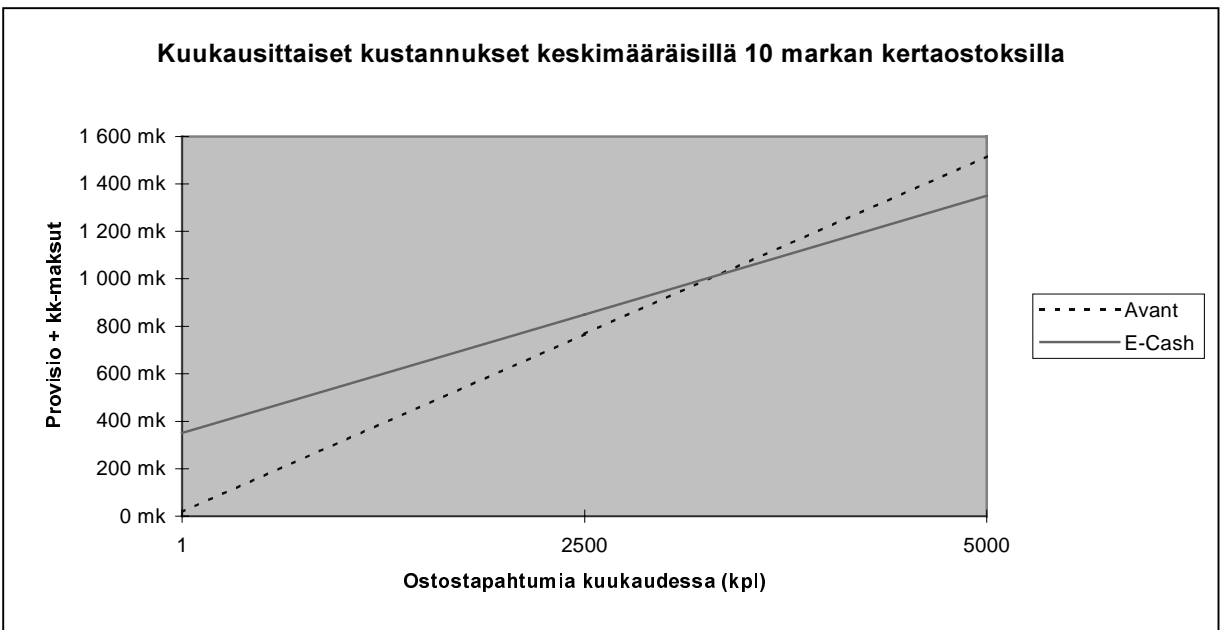
Diagrammeista on selvästi nähtävissä, että Ecashin korkeampi kuukausimaksu aiheuttaa Avan- tia korkeammat kustannukset jos yksittäiset ostostapahtumat ovat hinnaltaan pieniä ja ostosta- pahtumia on vähän. Ostostapahtuman hinnan ja ostostapahtumien määrän kasvaessa Avant on kuitenkin kalliimpi alle 50 markan kertaostoksilla korkeammasta provisiosta johtuen (esimer- kiksi 5 markan kertaostoksilla jos ostostapahtumia on yli 2200).



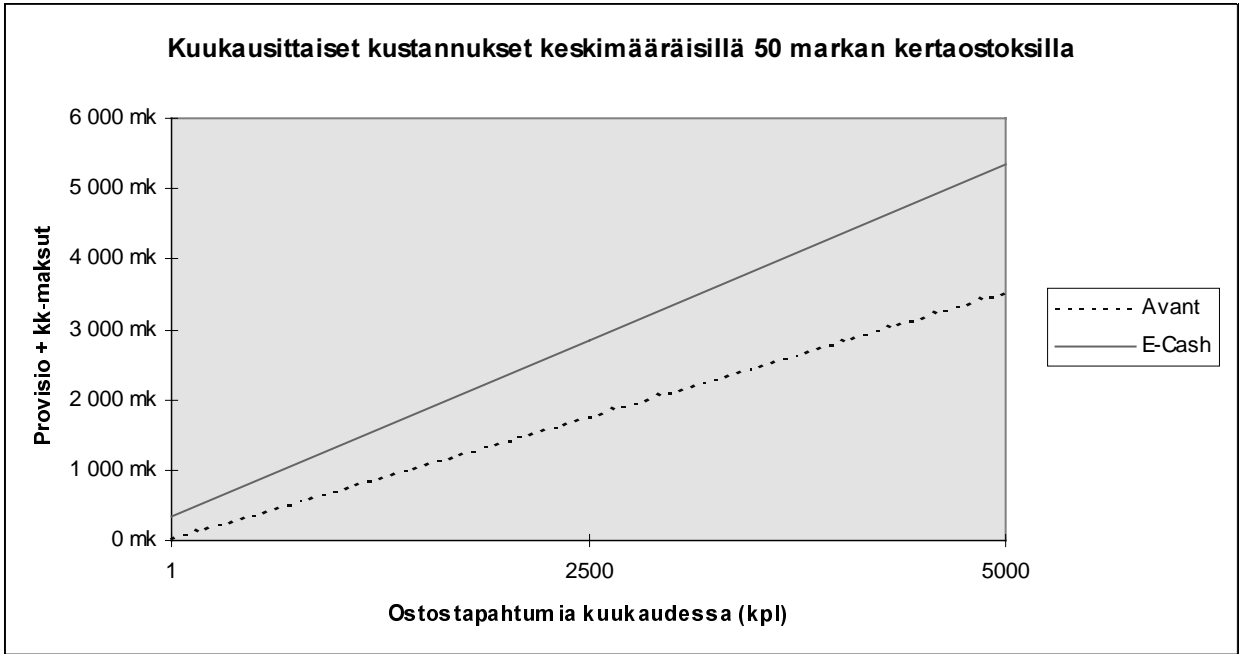
Kuva 11. Kuukausittaiset kustannukset keskimääräisillä 1 markan kertaostoksilla (kuukausimaksut + provisiot).



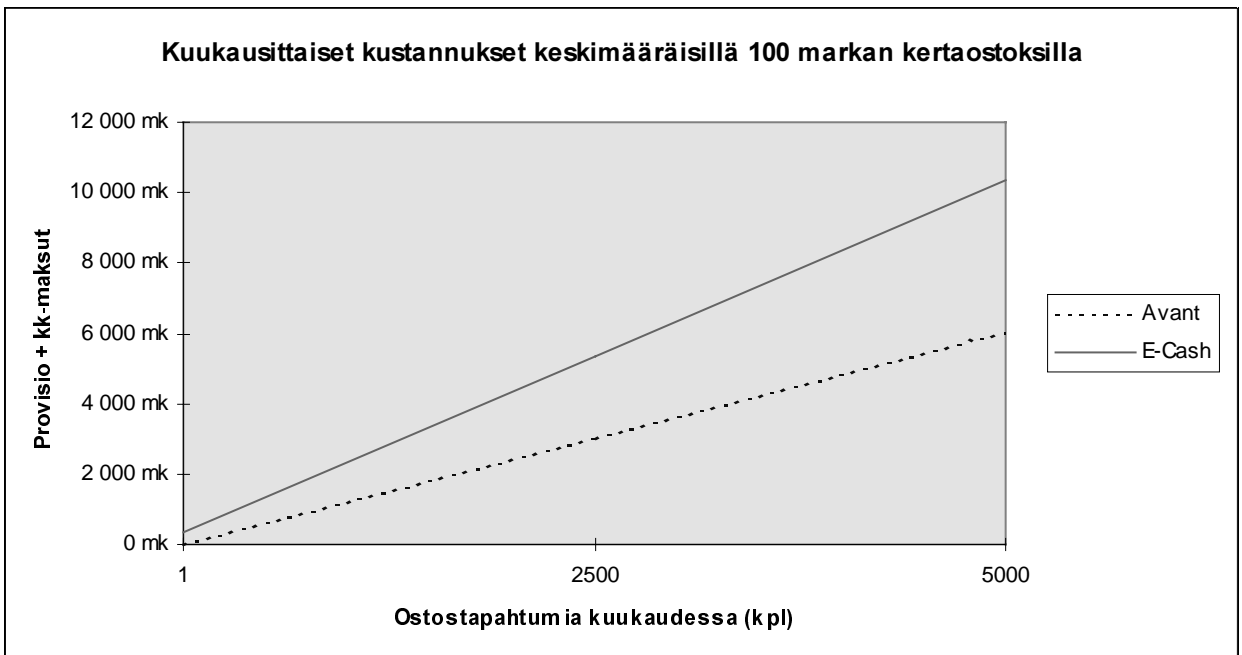
Kuva 12. Kuukausittaiset kustannukset keskimääräisillä 5 markan kertaostoksilla (kuukausimaksut + provisiot).



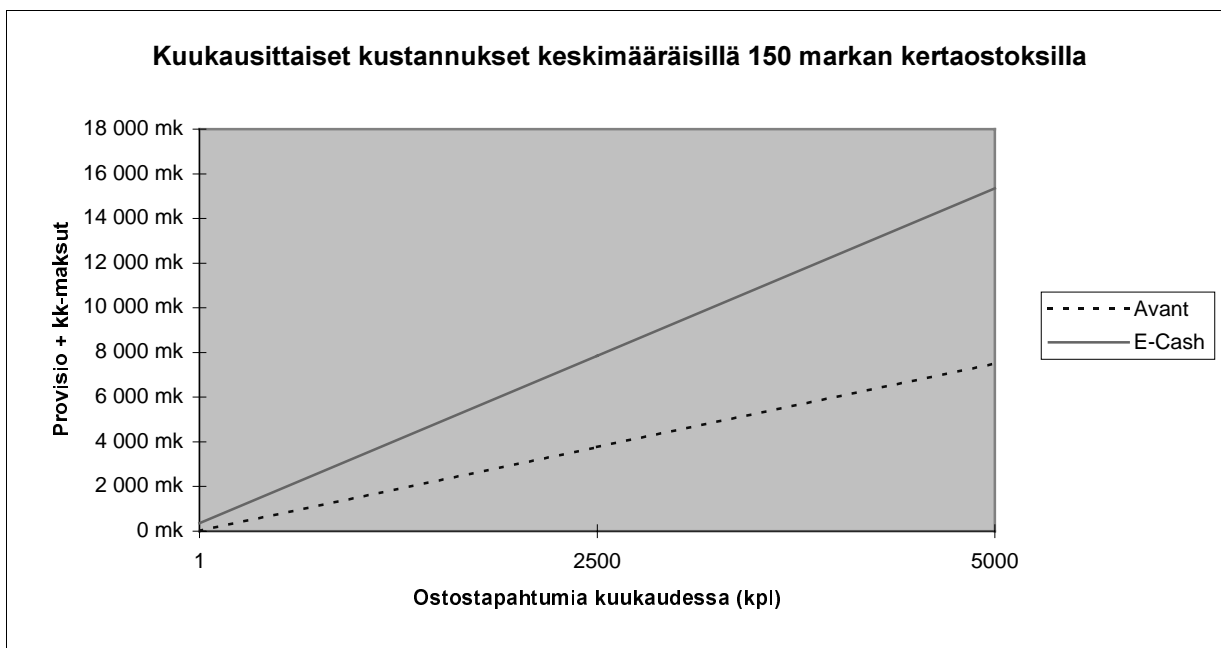
Kuva 13. Kuukausittaiset kustannukset keskimääräisillä 10 markan kertaostoksilla (kuukausimaksut + provisiot).



Kuva 14. Kuukausittaiset kustannukset keskimääräisillä 50 markan kertaostoksilla (kuukausimaksut + provisiot).



Kuva 15. Kuukausittaiset kustannukset keskimääräisillä 100 markan kertaostoksilla (kuukausimaksut + provisiot).



Kuva 16. Kuukausittaiset kustannukset keskimääräisillä 150 markan kertaostoksilla (kuukausimaksut + provisiot).

Pienillä kertaostosten hinnoilla ja vähäisillä ostostapahtumilla on mahdollista, että myyntitulot eivät kata maksujärjestelmän aiheuttamia kuluja. Kuvasta 11 voidaan todeta, että keskimääräisellä 1 markan kertaostoksella Avant-korttirahalla ostostapahtumia täytyy olla kuukaudessa 18 kappaletta ja Ecashillä 357 kappaletta, jotta myyntitulot kattavat maksujärjestelmän provisiot ja kuukausimaksut. Tässä ei ole otettu huomioon perustamiskustannuksia vaan ainoastaan kuukausikulut ja provisiot.

Hinta	Tarvittavien ostostapahtumien määrä (kpl)	
	Avant	Ecash
1 mk	18	357
5 mk	4	71
10 mk	2	36
50 mk	0	7
100 mk	0	4
150 mk	0	2

Taulukko 3. Kuukaudessa tarvittavien ostostapahtumien määrä erisuuruisilla keskimääräisillä ostoksilla.

8.5 Tulosten vertailu

Molemmat maksujärjestelmät sopivat hyvin pienten kertamaksujen välittämiseen WWW-palvelussa. Ecash ei kuitenkaan sovellu asiakaspäättekäyttöön, joten jos maksujärjestelmänä on Ecash, dokumenttien käytön ja maksujen suorituksen on tapahduttava asiakkaan omalta tietokoneelta. Asiakkaiden omilta koneilta tapahtuvaan käyttöön ja maksujen välittämiseen Ecash on käyttäjälle halvempi kuin Avant-korttiraha. Asiakkaan kannalta luonnollisesti paras ratkaisu olisi, että valittavissa olisi molemmat maksujärjestelmät, mutta myyjälle kustannukset saattavat nousta liian suuriksi.

Myyjän kannalta järjestelmät eroavat toisistaan mm. perustamis- ja kuukausikuluiltaan, provi-sioiltaan sekä ohjelmistojen saatavuudelta eri käyttöjärjestelmäympäristöihin.

Avant-korttirahan perustamiskustannukset ovat suuremmat kuin Ecashin, mutta Ecashin kuukausimaksut ovat puolestaan huomattavasti kalliimmat kuin Avantin. Pienillä yksittäisillä mak-suilla Avantin provisiot ovat kalliimmat kuin Ecashissa, maksujen kasvaessa Avantin provisio pienenee, ollen yli 100 markan ostoksilla pienempi kuin Ecashin provisio.

Avant-korttiraha -ohjelmistot AvantClient ja AvantServer ovat tällä hetkellä saatavissa valmiina ainoastaan Windows 95 ja NT -ympäristöön (AvantServer ainoastaan NT:lle). AvantServer-oh-jelmisto on käännettävissä Unix-ympäristöön, mutta se on suoritettava tapauskohtaisesti. Ecash-ohjelmistot ovat saatavissa Macintosh, Unix ja Windows-ympäristöihin lukuisille eri WWW-palvelinohjelmistoille.

9. TEKIJÄNOIKEUSPERIAATTEET

9.1 Teosten uudelleenjulkaisu digitaalisessa muodossa

Tekijänoikeus syntyy samalla kuin teoskin. Tekijänoikeuden saa se henkilö, joka teoksen on luonut. Teoksia ovat esimerkiksi kirjat, lehdet, tietokoneohjelmat, valokuvat, näytelmät, elokuvat ja maalaukset. Suojan kohteena on teoksen ulkomuoto ja ilmaisutapa. Tietoa, ideaa tai rakennetta ei suojata, joten ne ovat vapaasti jokaisen käytettävissä. Tekijänoikeus on voimassa tekijän eliniän ja 70 vuotta hänen kuolinvuotensa päättymisestä.

ELEKTRA -projektissa uudelleenjulkaistaan tekijänoikeuden suojaamia kirjallisia teoksia sekä niihin liittyvää kuvitusta. Oikeudenhaltijat voivat tarjota ELEKTRAn puitteissa graafiseen ja/tai elektroniseen primaarijulkaisuun sisältyviä artikkeleita tietoverkon välityksellä yleisön käyttöön. Teokset tallennetaan tietokantaan, josta ne ovat käytettävissä Helsingin yliopiston kirjaston bibliografisten tietokantojen, lähinnä ARTO-tietokannan, välityksellä. Tämä bibliografinen tietokanta on julkisten kirjastojen sekä eräiden yritysten kirjastojen käytössä. Käyttö edellyttää nimenomaista sopimusta HYK:n kanssa. Käytöstä maksetaan korvaus. ELEKTRA -projektissa bibliografisiin tietoihin liitetään itse teokset. Tähän tarvitaan oikeudenhaltijoiden luvat.

Kopioston tehtäviä projektissa ovat tekijänoikeuslupiin liittyvä neuvonta sekä tarvittaessa avustaminen käyttöoikeuksien hankinnassa.

9.2 Digitaalisten käyttöoikeuksien luovutuksista

Tekijänoikeus syntyy aina teoksen tekijälle. Tekijä voi sopimuksella luovuttaa taloudelliset oikeutensa osittain tai kokonaan. Digitaalisten käyttöoikeuksien luovutuksista ei ole olemassa vakiintunutta tapaa, mallisopimuksia tai vastaavia. Alalla on käytössä erilaisia oikeuksien luovutustapoja, joita voidaan kuvata mm. seuraavasti:

1. Tekijän ja kustantajan välinen (kustannus)sopimus kattaa vain printtijulkaisuoikeudet. Digitaali-oikeudet ovat tekijällä.
2. Digitaali-oikeudet ovat siirtyneet sopimuksella kokonaan kustantajalle; sopimus voi olla joko alkuperäinen (kustannus) sopimus tai erillinen sopimus
3. Digitaali-oikeudet ovat siirtyneet osittain kustantajalle, mutta eivät kata (kaikkia) projektissa tarvittavia käyttötapoja
4. Tekijä ja kustantaja hallinnoivat yhteisesti digitaali-oikeuksia.
5. Kustantajalle on varattu digitaalisia käyttömuotoja koskeva etuosto- (optio) oikeus.

Erikseen on selvitettävä teokseen liittyvien kuvien, piirrosten ja taulukoiden käyttöoikeudet.

Kopiosto edustaa sekä tekijöitä että kustantajia. Kopiosto ei ota kantaa osapuolten välisiin oikeuksien luovutuksista koskeviin sopimuksiin sinänsä. ELEKTRA -projektille on laadittava sellaiset sopimusmallit, että ne ovat sovellettavissa kaikkiin edellämäinnittäisiin tilanteisiin.

9.3 Mitä oikeuksia ELEKTRA tarvitsee ?

Hankittavien oikeuksien tulee vastata palvelun tarjoajan (ELEKTRA -projekti), välittäjien (kirjastot) sekä loppukäyttäjien (kirjaston asiakkaat) tarvitsemia käyttöoikeuksia.

Projektissa tarvitaan seuraavat oikeudet:

1. Reproduktio-oikeus, joka käsittää oikeuden valmistaa teoksesta kappale tietokoneella luettavaan muotoon rinnakkaisjulkaisun valmisteluprosessin aikana, asiakkaiden käyttäessä teosta tietokoneen välityksellä sekä teoksen tallentamiseksi pysyvästi ELEKTRAn hallinnoimaan tietokantaan (ns. elektroninen arkistointi)
2. Oikeus palvelun toteuttamiselle välttämättömien teknisten muutosten tekoon.
3. Oikeus saattaa teos yleisön (asiakkaiden) saataville eli teoksen käyttö tietokannan välityksellä
4. Tulostamiseen liittyvä reproduktio-oikeus eli tulostus paperille.

9.4 Käytön ehdot ja korvaukset

Projektin yleisenä päämääränä on, että teosten oikeudenhaltijat voivat kukin itse määrätä aineistonsa käyttöehdoista ja käytöstä perittävistä korvauksista. Projektin pilottivaiheessa tämä on kuitenkin vain rajoitetusti mahdollista.

ELEKTRA -projekti hankkii oikeudenhaltijoilta käyttöoikeuden. Oikeudenhaltijoiden asettamat ehdot voivat koskea mm. seuraavia seikkoja:

1. Käyttötiedot; mitä tietoja rekisteröidään ja kuinka niistä raportoidaan
2. Korvauksen peruste. Peruste voi vaihdella käyttötavoittain esimerkiksi tietokantaan tallentamisen ja paperille tulostamisen kohdalla
3. Korvauksen määrä. Korvaus voi projektin puitteissa olla myös nimellinen
4. Korvauksen maksutapa

9.5 Oikeuksien hankinta ja hallinnointi

Projektin tarvitsemat käyttöoikeudet on lueteltu kohdassa 3. Projekti voi hankkia nämä oikeudet joko suoraan tekijöiltä, kustantajalta tai osapuolten edustajalta. Edustus voi perustua tekijän, kustantajan tai molempien antamaan toimeksiantoon. Kopiosto voi tarvittaessa toimia oikeuksien välittäjänä. ELEKTRA-projekti on syytä rakentaa siten, että kaikki vaihtoehtoiset tavat hankkia oikeuksia ovat mahdollisia. Näin turvataan tarvittavat käyttöoikeudet kaikissa tilanteissa. Käyttöoikeuksien hallinnointi edellyttää teos- ja oikeudenhaltijakohtaista rekisteröintiä. Saman hallinnointijärjestelmän avulla voidaan tilittää myös käytöstä perittävät korvaukset.

9.6 Oikeudenhaltijoiden kanssa solmittavat sopimukset

On luontevaa, että ELEKTRA -projekti solmii sopimuksen tarvittavista käyttöoikeuksista suoraan tekijöiden tai kustantajien kanssa silloin, kun oikeudet ovat tällä tavoin helposti hankittavissa ja hallinnoitavissa. Mikäli oikeuksista ei vallitse varmuutta tai mikäli muut käytännölliset seikat, esimerkiksi monitekijyys, antavat aiheutta, Kopiosto voi toimia oikeuksien välittäjänä ja ottaa vastuulleen oikeuksien selvittämisen.

Teoksen käyttöoikeus luovutetaan ELEKTRA -projektille erillisellä sopimuksella. Sopimus voi olla pysyvä tai määräaikainen, kuitenkin kokeiluprojektin luonne huomioonottaen luontevasti määräaikainen.

9.7 Korvauksen maksaminen oikeudenhaltijoille

Projektin eräänä lähtökohtana on, että oikeudenhaltijat voivat kukin itse määrätä teoksensa käytöstä perittävistä korvauksista. Mikäli käyttöoikeuden luovutuksen ehtona on (nimellinen) korvaus, ELEKTRA projekti ylläpitää teos- ja oikeudenhaltijakohtaista rekisteriä siitä, kenelle korvauksia tilitetään. Mikäli Kopiosto toimii oikeuksien välittäjänä, se ottaa vastuun näiden oikeudenhaltijoiden osalta rekisteröinnistä ja korvausten tilittämisestä.

9.8 ELEKTRA -projektin perimät korvaukset

ELEKTRA -projektin pyrkimyksenä on tarjota kustannuksiltaan kohtuullinen, erityisesti kirjastolaitokselle ja tieteelliselle yhteisölle suunnattu palvelu, jota osin on tuettu julkisin varoin. Tällöin on luontevaa, ettei projekti sinänsä hanki oikeuksia tarvittavaa kokeiluprojektia laajemmin tai peri käytöstä liikevoittoon tähtääviä korvauksia. ELEKTRA-projektin asiakkailta perimät korvaukset koostuvat tietokannan käytöstä perittävistä korvauksista sekä korvauksista teosten oikeudenhaltijoille.

Bibliografisen ARTO-tietokannan käyttö edellyttää sopimusta Helsingin yliopiston kirjaston kanssa sekä korvausta käytöstä. Vain ARTO:n käyttöön oikeutetut voivat käyttää siihen liitettäviä teoksia. ELEKTRAn käytöstä asiakkaalta perittävät korvaukset jakautuvat näin periaatteessa kahtia: bibliografisen ARTO-tietokannan käyttö ja sen korvausosuus, josta on olemassa hinta- ja sopimuskäytännöt, ja suunniteltu teosten käyttö ARTO - tietokannan kautta ELEKTRA-projektin hallinnoimissa puitteissa. Tässä esityksessä käsitellään vain jälkimmäistä eli teosten käytöstä perittävää korvausta.

9.9 Loppukäyttäjät ja teosten käyttökorvaukset

Seuraavassa on tarkasteltu loppukäyttäjien korvausperusteisiin liittyviä kysymyksiä silloin, kun kysymyksessä ovat julkiset kirjastot ja niiden asiakkaat, jotka käyttävät kirjaston tiloissa sijaitsevia tietokoneita.

Kirjastojen asiakkailleen tarjoamat palvelut voivat olla seuraavia:

1. Oikeus tutustua teokseen tietokoneen välityksellä. Tämän käyttömuodon kohdalla voidaan kysyä, onko projektin puitteissa aiheellista periä siitä korvausta ja jos peritään, millä perusteella? Kopiosto suosittaa ettei kokeiluvaiheessa peritä erillistä korvausta teokseen tutustumisesta kirjaston tiloissa.
2. Oikeus tulostaa teos paperille. Tässä käyttömuodossa kopion hinta ja tekijänoikeuskorvaus ovat periaatteessa helposti määriteltävissä ja perittävissä

9.10 Käyttötapahtumien rekisteröinti

Kopioston ehtona projektiin osallistumiselle on ollut, että käytettävä tekniikka tuottaa teoskohtaisia tietoja käytön määrästä. Eri käyttömuotoja seurataan todennäköisesti eri tavoin. Seuraavassa on oletettu hahmotelma:

1. Tutustumista teokseen tietokoneen välityksellä tulee projektin aikana pyrkiä seuraamaan ainakin julkaisukohtaisesti. Kopioston lähtökohtana on ollut, ettei projektin aikana tästä käyttömuodosta peritä korvauksia. Käyttötietojen raportointi on tältä osin ensisijaisesti palvelu projektiin osallistuville ja väline projektin kehittämiseen.
2. Paperitulosteiden seuranta rakennetaan teoskohtaisesti. Käyttö raportoidaan oikeudenhaltijoille yhdessä korvauksen tilityksen kanssa.

9.11 Käyttötietojen ja käyttökorvausten välittäminen oikeudenhaltijoille

ELEKTRA -projekti välittää oikeudenhaltijoille heidän kanssaan sovitut tiedot käytöstä ja tilittää samalla oikeudenhaltijan korvaukset. Korvausten tilitys oikeudenhaltijoille edellyttää tietoja sekä itse teoksen että siinä käytettyjen elementtien (esimerkiksi kuvitus) oikeudenhaltijoista ja heidän osuuksistaan korvauksista. Tilitysjärjestelmään kuuluvat oikeudenhaltioiden yhteystiedot mukaanlukien tiedot mahdollisista perikunnista. Oikeuksien hallinnoinnissa tarvittavien tietokantojen rakentaminen ja ylläpito ovat tavanomaisia tekijänoikeusjärjestön tehtäviä. Kopioston rekisterit ovat tarvittaessa ELEKTRA -projektin käytössä.

9.12 Kopioston rooli ELEKTRA -projektissa

Kopiosto laatii ELEKTRA-projektille sopimusmallin tai -malleja koskien oikeudenhaltijan (tekijä, kustantaja, perikunta tms.) omistaman tai hallinnoiman tekijänoikeuden siirtoa projektin käyttöön. Malli koskee ELEKTRA-projektissa tarvittavia oikeuksia. Sopimusmallissa käsitellään oikeuksien luovutukseen liittyvät yleiset ehdot ja se on siten käyttökelpoinen riippumatta siitä, keneltä oikeudet hankitaan.

Kopiosto voi tarvittaessa toimia välittäjänä hankittaessa ELEKTRA -projektille käyttöoikeuksia tekijöiltä, kustantajalta tai molemmilta, aina kulloinkin kysymyksessä olevan tapauksen mukaan. Kopiosto pitää tarkoituksenmukaisena, että kaikki käyttöoikeuksien siirrot ja niitä koskevat valtuutukset rajataan tässä vaiheessa vain ELEKTRA -projektin tarvitsemiin oikeuksiin.

10. TEKIJÄNOIKEUSKORVAUKSET JA NIIDEN HALLINNOINTI

10.1 Palvelujen hinnoittelu

ELEKTRA-projektissa tutkitaan maksullisten, kirjastojen asiakkaille suunnattujen palveluiden teknistä, sopimusteknistä sekä liiketaloudellista toimivuutta.

Palveluiden asiakaskunta on kolmitasoinen:

1. tasolla HYK:n asiakkaana ovat yksittäiset tekijät ja kustantajat, jotka antavat käyttöluvan HYK:lle
2. tasolla asiakkaina ovat hankkeen piirissä olevat sopimuskirjastot
3. tasolla asiakkaina ovat kirjastojen asiakkaat eli loppukäyttäjät.

Asiakkailta perittävät palvelumaksut muodostuvat tallennuskorvauksesta, tulostekohtaisista korvauksista sekä teknisistä tulostusmaksuista. Suurin osa oikeudenhaltijoista näyttäisi hyväksyvän ELEKTRA-projektin suositteleman hinnoittelumallin. Oikeudenhaltija voi halutessaan esittää myös oman ehdotuksensa korvaussummista sekä siitä, miten korvaukset tulisi jakaa.

10.2 Hintasuositukset tallennuskorvaukseksi ja tulostekohtaiseksi korvaukseksi

ELEKTRA-projektissa suositellun hinnoittelumallin mukaan kunkin julkaisun vuosikerran tallentamisesta peritään HYK:lta tallennusmaksu, joka on yhtä suuri kuin julkaisun vuosikerran yhteisötilaushinta artikkelien julkaisuvuodelta. Tämä korvaus kattaa aineiston digitoinnin sekä aineiston saatavilla olon koko projektin ajalta (projektin arvioitu kesto kattaa aikavälin heinäkuusta 1996 heinäkuuhun 1999).

Useimmat kustantajat ovat hyväksyneet tallennuskorvauksen hinnoittelumallin.

- ELEKTRA-hankkeessa mukana olevat sopimuskirjastot eivät maksa tallennusmaksua, vaan HYK maksaa tallennusmaksun kirjastojen puolesta keskitetysti.
- Jotkut oikeudenhaltijat eivät peri korvausta vaan myöntävät tallennusluvan ilmaiseksi.

ELEKTRA-projektin suositteleman ehdotuksen mukaan kustakin kirjaston asiakkaiden tulostamasta sivusta peritään hinta, joka muodostuu teknisestä tulostusmaksusta sekä tulostekohtaisesta tekijänoikeuskorvauksesta.

10.3 Hintasuositus tekniseksi tulostuskorvaukseksi

Tekninen tulostusmaksu TTM peritään asiakkaalta jokaisesta tulostetusta sivusta ja se on suuruudeltaan 0,30 mk. Tulostekohtainen tekijänoikeuskorvaus TK TOK lasketaan jakamalla julkaisun vuosikerran yhteisötilaushinta YH (julkaisuvuodelta) julkaisun kokonaissivumäärällä SM sekä kertomalla saatu tulo tulostettujen sivujen määrällä S. Tulostekohtaiselle tekijänoikeuskorvaukselle on lisäksi asetettu alaraja AR (= 0,50 mk).

$$\text{TK TOK} = \text{MAX} (\text{YH}/\text{SM} * \text{S}, \text{AR})$$

10.4 Ehdotukset korvauksen jaosta

Taulukko 4 ELEKTRA-projektin suositus tulostekorvaukseksi.

JULKAISU	VUOSI	VUOSI-KERRAN TILAUSHINTA, mk	JULKAISUN KOKONAISIVUMÄÄRÄ	ASIAKKAALTA PERITTÄVÄ TULOSTEKORVAUS
Acta Forestalia Fennica	1996	280	322	1,20
AfinLAN vuosikirja	1996	70	166	0,80
Agricultural and Food Science in Finland	1996	300	616	0,80
Annales Botanici Fennici	1996	400	319	1,55 tulostus kielletty
Annales Zoologici Fennici	1996	400	723	0,85 tulostus kielletty
Arkhimedes	1995	60	345	0,80
— ” —	1996	60	160	0,80
Finnish Economic Papers	1996	150	156	1,25
Finnisch-Ugrische Forschungen	1996	150	329	0,80
Folkmålsstudier	1996	80	177	0,80
Futura	1996	200	295	1,00
Genos	1996	150	188	1,10
Geologi	1996	85	176	0,80
Geophysica	1996	200	100	2,30
Gerontologia	1996	200	285	1,00
Helsingin yliopiston Kirjaston tiedotuslehti	1996	maksuton	178	0,80
Historiallinen arkisto	1996	240	555	0,80
Kasvatus	1995	150	532	0,80
— ” —	1996	170	525	0,80
Kehittyvä Elintarvike	1996	150	192	1,10
Kirjallisuudentutkijain Seuran vuosikirja	1996	150	234	0,95
Kuu taivaalta - opettajan opas		63	80	1,10
Källan	1996	82	191	0,80
Lakimies	1996	500	1294	0,80

Lounais-Hämeen Luonto	1996	75	108	1,00
Luonnon tutkija	1996	120	160	1,05
Memoranda soc.pro Fauna et Flora Fennica	1996	75	265	0,80
Metsätieteen aikakauskirja - Folia Forestalia	1995	200	347	0,90
— ” —	1996	200	449	0,80
Musiikin suunta	1996	150	236	0,95
Politiikka	1996	250	260	1,25
Rakenteiden Mekaniikka - Journal of Structural Mechanics	1996	50	247	0,80
Savon Luonto	1996	30	40	1,05
Science studies	1996	160	160	1,30
Siivekäs	1995	70	92	1,10
— ” —	1996	85	72	1,50
Silva Fennica	1995	200	315	0,95
— ” —	1996	200	489	0,80
Sosiologia	1996	200	344	0,90
Suo	1996	160	152	1,35
Suomalais-Ugrilaisen Seuran Aikakauskirja	1996	140	255	0,80
Suomen Antropologi	1996	130	192	1,00
Suomen eläinlääkärilehti	1996	450	762	0,90
Suomen Hammaslääkärilehti	1996	715	1208	0,90
Suomen kirkkohistoriallisen seuran vuosikirja	1996	120	427	0,80
Terra	1996	140	272	0,80
Tidskrift utgiven av Juridiska Föreningen i Finland	1996	300	530	0,90
Tiedokas	1996	maksuton	40	0,80
Tieteessä tapahtuu	1996	maksuton	430	0,80
Tietolinja	1995	maksuton	104	0,80
— ” —	1996	maksuton	88	0,80
Tähdet ja avaruus	1996	140	328	0,80
Virittäjä	1996	250	635	0,80

Yliopisto	1996	252	1380	0,80
Ympäristöjuridiikka	1996	120	171	1,00 tulostus kielletty

Poikkeukset ELEKTRA-suositukseen

1. Jotkut oikeudenhaltijat ovat kokonaan kieltäneet artikkeliansa tulostuksen.
2. Jotkut oikeudenhaltijat eivät vaadi lainkaan tekijänoikeuskorvauksia artikkelien tulostamisesta (tämä tilanne voitaneen mallintaa asettamalla sivuhinta nollassi).
3. Jotkut oikeudenhaltijat esittävät, että tulostekohtainen tekijänoikeuskorvaus lasketaan jollakin muulla (sivu)hinnalla kuin projekti ehdottaa.
4. Jotkut oikeudenhaltijat eivät vaadi tallennuskorvausta.
5. Jotkut oikeudenhaltijat ehdottavat muuta kuin projektin suosittamaa tallennuskorvausta.

10.5 Oikeudenhaltijan valtuutus Kopiostolle välittää korvausten tilitys

Helsingin yliopiston kirjastolla ei toistaiseksi ole käytössään tilityksessä tarvittavia oikeudenhaltijarekistereitä ja tietojärjestelmiä. Ensimmäistä laskutusjärjestelmän osaa HYK:lle rakennetaan parhaillaan jäljempänä esiteltävän Software Vision Oy:n tekemän mallin pohjalta.

Helsingin yliopiston kirjasto ei toistaiseksi ole myöskään itse hoitanut yhteyksiä suoraan oikeudenhaltijoihin. Kopiosto on vastannut sopimusliikenteestä kustantajille ja tekijöille. Korvausten teknisestä tilityksestä HYK:n tallentamien käyttöraporttien mukaan on samoin vastannut Kopiosto siltä osin kuin oikeudenhaltijat ovat sitä toivoneet ja antaneet tehtävään valtuuden. Tilityksestä ei peritä palkkiota.

Korvausten tilitystä koskeva valtuutus koskee vain ELEKTRA-projektia ja raukeaa ELEKTRA-projektin päättyessä tai jos oikeudenhaltija irtisanoo julkaisunsa/teoksensa käyttöä koskevan sopimuksen.

10.6 Sopimusten mukainen raportointi ja tilitykset

ELEKTRA-hankkeen piiriin otetut artikkelit, julkaisut ja tekijät on tallennettu ELEKTRA-tietokantaan. Tiedot oikeudenhaltijoista on periaatteessa mahdollista siirtää käytettävään tilitysjärjestelmään joko yksitellen taikka eräajona.

ELEKTRA-kannan vakiomuotoisen käyttöliittymän ongelmana on, ettei vakiomuotoinen artikkeliluettelo sisällä tietoa artikkeleiden tekijöistä ja kustantajista. Nämä tiedot ovat saatavilla ainoastaan tarkasteltaessa artikkeleita erikseen ns. korttimuodossa, josta niiden siirtäminen toiseen kantaan on hankalaa. Myös julkaisuluettelossa on sama ongelma: Vakiomuotoinen luettelo ei sisällä tietoa julkaisujen kustantajista. Vastaavasti ELEKTRA-tietokantaan kytketyn soveluksen vakiomuotoinen tekijäluettelo ei sisällä tekijöiden yhteystietoja.

Näistä puutteista huolimatta ELEKTRA-kantaa lienee mahdollista hyödyntää tilitystä toteutettaessa. Tämä edellyttää kuitenkin a) HYK:n atk-osaston myötävaikutusta tai b) tietojen siirtämistä yksitellen tallentamalla.

Artikkelin käyttötapaukset merkitään kirjastoissa siten, että kirjaston henkilökunta ottaa kunkin tulostetun artikkelin kannesta ylimääräisen kopion. Tähän kopioon kirjoitetaan käsin artikkelitulosteen sivuhinta sekä tulostettu sivumäärä. Käyttötapauksien atk-tallennus tapahtuu HYK:ssä

ELEKTRA-sopimuksen mukaisesti kirjastojen on ”noudatettava tulostuksesta annettuja ohjeita”. Tulostusohjeisiin voidaan sisällyttää näin myös tulostustapahtumien raportointiohjeet. Seuraavassa on eräs malli, miten kirjastot voisivat tallentaa tulostustapahtumia kuvaavat tiedot:

Artikkelien kansilehtisten sisältämät tiedot tulisi tallentaa kirjastoissa määrämukotoisiin Excel-tilukoihin, jotka lähetetään HYK:lle tammikuussa. HYK voi näin meneteltäessä yhdistää eri kirjastoilta saadut käyttöraportit yhdeksi kokonaisraportiksi, jonka pohjalta sopimuskirjastoille lähetetään laskut ja laaditaan tilitystä varten käyttöraporttien yhteenvetoraportti (esim. Excel-tiludostona).

Kirjastojen HYK:lle lähettämässä käyttöraportissa tulisi mainita

- artikkelin nimi
- tulostettu sivumäärä
- asiakkaalta peritty tulostekohtainen (tekijänoikeus-) korvaus (0,50 - 2,30 mk/sivu)
- tulostettujen sivujen määrä

Näiden tietojen avulla voidaan laskea tulostumistapahtumaan liittyvä tekijänoikeuskorvaus sekä jakaa tämä summa oikeudenhaltijoille.

Mikäli ELEKTRA-sopimuskirjastot eivät ole valmiit edellä kuvattuun tai vastaavaan tyyppiseen hajautettuun tallennusratkaisuun, vastuu tallennustyön organisoinnista jäänee HYK:lle. Tällöin HYK:n on kerättävä ja toimitettava tilitystä varten edellä kuvatut tiedot.

Tietojen siirto HYK:lta edelleen tilitystä varten tilitystietokantaan tapahtuisi helpoimmin siten, että tulostustiedot tallennetaan standardimukotoiseen atk-rekisteriin (yksi tietue kutakin tulostustapahtumaa kohden). Tietojen siirto tilitysjärjestelmään voidaan tällöin rakentaa automaattiseksi, mikäli niin halutaan.

10.7 Laskuttaminen ja summien tarkastaminen

Tallennuskorvauksen laskuttaminen

Kopiosto voi laskea ja laskuttaa HYK:lta tallennuskorvauksen niiden tekijöiden ja kustantajien puolesta jotka ovat tällaista palvelua Kopiostolta pyytäneet. On myös mahdollista, että HYK tilittää oma-aloitteisesti tekijänoikeuskorvaukset. Kun tallennuskorvaus on julkaisukohtainen kertamaksu, ei sen laskuttaminen tuota ongelmia.

Tulostekohtaisten korvausten laskeminen ja laskuttaminen

Mikäli HYK kerää kirjastojen tulosteraportit yhteen ja tallentaa ne esimerkiksi Excel-malliin, on HYK:n helppo

- laskuttaa tulostekohtaiset korvaukset sopimuskirjastoilta sekä
- raportoida konekielisessä muodossa tulostuskorvausten kertyminen (tapahtumittain sekä yhteenlaskettuna).

Korvausten tilityksen ja teosten käyttöä koskevien raporttien toimittamiseksi oikeudenhaltijoille on joka tapauksessa keskeistä, että tulostustapahtumaraportti on konekielinen.

Summien tarkastaminen

Perittyjen teoskohtaisten tekijänoikeuskorvausten vertaaminen käyttöraportteihin tapahtuneen projektin aikana lähinnä pistokokeilla.

Jos tulostekohtaisen tekijänoikeuskorvauksen laskemisessa käytetyt sivuhintatiedot halutaan kuitenkin tarkastaa, on tämä mahdollista joko a) paperitulostetta ja näytöllä olevaa tietokantaa vertailemalla tai b) kahta tietokantaa näytöllä selaten. Mitä enemmän tarkastusmenettelyä pyritään automatisoimaan, sen kalliimmaksi järjestelmän perustaminen tulee.

10.8 Korvausten jako ja tilitysraportit

Korvausten jako kirjoittajille ja kustantajille

Eräiden tekijä- ja kustantajajärjestöjen suositteleman jakomallin mukaisesti kirjastoilta perityt korvaukset on ELEKTRA-projektin suositussopimuksessa jaettu tekijöiden ja kustantajien kesken tasan puoliksi (50% / 50%). Tämä jako voi kuitenkin periaatteessa vaihdella tilanteen mukaisesti: esimerkiksi eräät kustantajat ovat hankkineet teoksiin niin laajat oikeudet, että koko korvausosuus menee kustantajalle.

Tekijöiksi määritellään tässä kirjoittajat, kääntäjät ja kuvittajat. Tekijöille koitua tekijänoikeuskorvauksen kokonaissumma on suositussopimuksessa ehdotettu jaettavaksi siten, että tekstintuottajat (eli kirjoittajat ja kääntäjät) saavat 85% ja kuvittajat 15%.

Mikäli artikkelilla on monta tekstintuottajaa, tekijänoikeuskorvaukset esitetään jaettavaksi pääluvun mukaisessa suhteessa. Sama pätee myös kuvitukseen: jos artikkelilla on monta kuvittajaa, kuvittajille kuuluva tekijänoikeuskorvaus jaetaan pääluvun mukaisessa suhteessa. Kun kyse on lähes yksinomaan tieteellisistä julkaisuista, lähes kaikki kuvituksen käyttö perustuu tekijänoikeuslain kuvasitaattia tieteellisessä esityksessä koskevaan lainkohtaan. Kuvan käyttö on tällöin vapaata.

Mikäli artikkelilla on erikseen kustantaja ja julkaisija, korvaukset jaetaan näille osapuolille tasan. Tällaisia tapauksia on kuitenkin vain harvoja.

Korvausten maksamiseen liittyvät tulosteet

Oikeudenhaltijoille eli tekijöille ja kustantajille maksetaan tekijänoikeuskorvauksia, joista on samalla suoritettava ennakonpidätykset.

Maksetuista suorituksista on lisäksi voitava tuottaa raportti, joka samalla kuvaa

1. mitä teoksia on tulostettu ja kuinka paljon
2. mitkä kirjastot ovat tulostaneet ja kuinka paljon

Lisäksi on mahdollista, että oikeudenhaltijat haluavat tietää, minkälaisista eristä heidän tekijänoikeuskorvauksensa ovat muodostuneet. Tämän palvelun laajuus ja taso on ratkaistava erikseen.

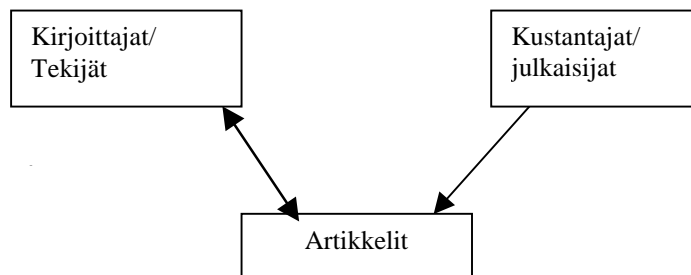
Teosten käyttämistä koskevat raportit oikeudenhaltijoille

HYK on sitoutunut toimittamaan oikeudenhaltijoille raportit myös teosten käytöstä. Sopimuksen mukaan raportointi tapahtuu joko teos- tai julkaisukohtaisesti, riippuen käytetystä tekniikasta. Koska projektin puitteissa perityt korvaukset ovat lähinnä nimellisiä, on syytä pyrkiä toimittamaan käyttöraportit kullekin sopimuksen tehneelle oikeudenhaltijalle teoskohtaisesti. Käytettävän tallennus- ja tilitysjärjestelmän tulisi näin kyetä käsittelemään laskutukseen ja käytöstä perittyjen korvausten jakoon liittyvän datan lisäksi puhdasta, teoskohtaista käyttöraporttietoa esim. julkaisun lukukerroista.

10.9 Alustava tietokantakuvaus

Kirjoittajat, kustantajat ja artikkelit

Kunkin artikkelin osalta on tiedettävä tekijät ja kustantaja, joille tekijänoikeusmaksuja ollaan tilittämässä. (Mikäli kustantajalla on kaikki oikeudet, ei artikkelin kirjoittajia välttämättä tarvitse tietää).



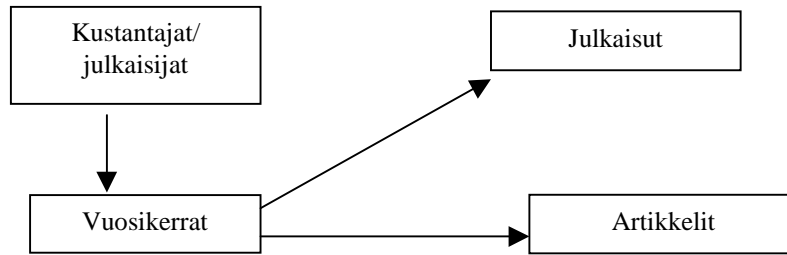
Kullakin artikkelilla voi olla useita kirjoittajia/tekijöitä ja kukin kirjoittaja voi olla kirjoittajana useassa eri artikkelissa.

Yhdellä artikkelilla on yleensä kuitenkin vain yksi kustantaja. Teoreettinen mahdollisuus, jonka mukaan kustantajan lisäksi voi olla olemassa erillinen julkaisija, voidaan alustavan selvityksen mukaisesti käsitellä siten, että julkaisija tallennetaan kustantajakohtaisiin lisätietoihin. Julkaisijoiden esittäminen erillisesti on kuitenkin myös mahdollista.

Julkaisut, vuosikerrat ja artikkelit

Kukin artikkeli voi kuulua vain yhteen vuosikertaan ja kukin vuosikerta voi edustaa vain yhtä julkaisua (lehteä). Julkaisuilla voi olla erilainen sivuhinnoittelu (ks. kohta 2.1 ELEKTRA-palvelujen hinnoittelu).

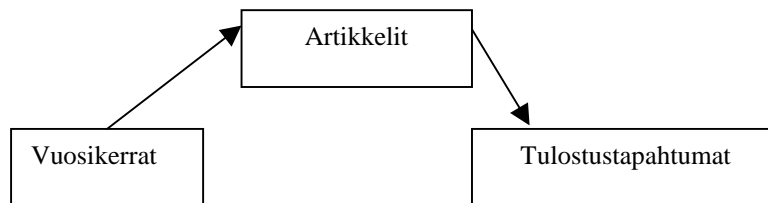
Kullakin vuosikerralla voi olla yleensä vain yksi kustantaja. Tämä merkitsee sitä, että artikkelin kustantajaa ei tarvitse välttämättä kunkin artikkelin osalta erikseen tallentaa, koska kustantaja voidaan päätellä "moni yhteen" -relaatioiden perusteella.



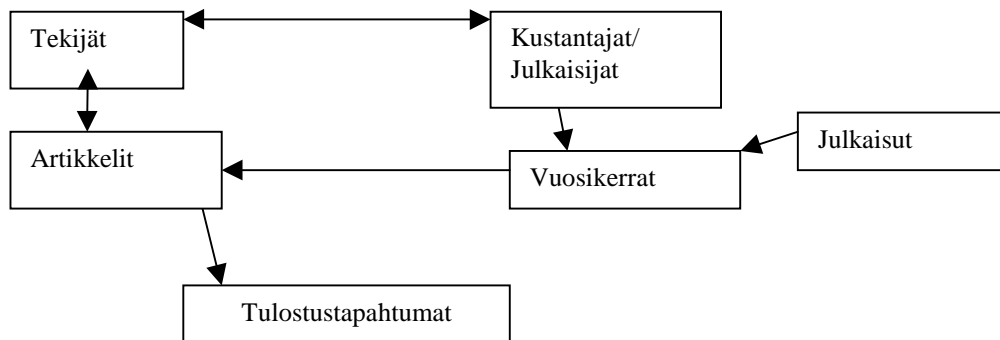
Järjestelmän määrittelytyön yhteydessä on kuitenkin vielä selvítettävä, voiko julkaisun kustantaja vaihtua siten, että eri vuosikerroilla on eri kustantaja. (Tämä tieto vaikuttaa vuosikertojen, kustantajien ja julkaisujen välisten relaatioiden rakenteeseen).

Tulostustapahtumat

Kuhunkin artikkeliin voi liittyä useita tulostustapahtumia. Tulostustapahtumiin liittyvä keskeinen tallennettava tieto on tulostettu sivumäärä. Tulostusvolyymin mukaisten käyttömaksujen laskuttamiseksi ja tilittämiseksi on tiedettävä myös, missä vuosikerrassa artikkeli on ilmestynyt, jotta tulosteiden osalta sovellettava sivuhinta saadaan selville. Tämä tieto ilmenee tulostustapahtumaan liittyvän vuosikerran tiedoista.



Yhteenveto tietokannasta -kaavio



Esitutkimuksessa sovelletut käsitelmääritelmät

Artikkeli on lehdessä julkaistu kirjoitus, joka on esitetty lehdessä itsenäisenä kokonaisuutena.

Teos on luovan työn omaleimainen tulos. Teoslajeja ovat mm. tekstit, kuvat, elävä kuva ja musiikki.

Tekijä on teoksen tekijä tai teoksen tekemiseen osallistunut henkilö, joka on työpanoksensa kautta saanut teokseen tekijänoikeuden. Tekijät luokitellaan tekstintekijöihin ja kuvantekijöihin.

Tekstintekijöiksi luokitellaan kirjoittajat sekä kääntäjät.

Kuvantekijöiksi luokitellaan ne kuvittajat, joilla on oikeuksia tarkasteltavaan artikkeliin. (Mikäli artikkelissa on sitaattioikeuden sallima kuva- tai kaaviolainaus, kyseisen kuvion tekijää ei luokitella kuvantekijöiden joukkoon).

11. TIETOTURVA

Luvun tavoitteena on perehdyttää lukijat aluksi tietoturvaan yleisemmällä tasolla ja tarkastella sen jälkeen tarkemmin tietoturvaa Internetissä, ottaen huomioon ELEKTRA-projektin ratkaisut ja tarpeet tällä hetkellä sekä tulevaisuudessa. Teksti toimii ohjeena kaikille projektin parissa työskenteleville ja siinä esitetyjä periaatteita ja sääntöjä tulee kaikkien noudattaa, jotta palvelun turvallisuus varmistetaan. Tietoturvaperiaatteiden noudattaminen vähentää häiriöiden ja virheiden todennäköisyyttä ja vaikutusta sekä vaikeuttaa tai estää kokonaan tietokantojen ja muun datan luvattoman käytön.

Seuraavassa esitetään tunnettuja tietoturvaa heikentäviä tietoturva-aukkoja, jotka uhkaavat myös ELEKTRA-palvelua sekä ratkaisuja näiden uhkien poistamiseksi tai vähentämiseksi. Tietoverkon turvallisuusratkaisut ovat seuranneet muun kehityksen jäljessä ja uusissa järjestelmissä ja ohjelmistoissa ilmenee jatkuvasti tietoturva-aukkoja. Verkon välityksellä käytävän kaupan liisääntyminen on kasvattanut vaatimuksia turvalliselle liikennöinnille ja uusia tietoturvaa parantavia ratkaisuja onkin alkanut tulla markkinoille samalla kun vanhempien järjestelmien turvallisuutta pyritään parantamaan.

11.1 Tietoturvan osa-alueet ja tietoturvapoliittikka

Tietoturva koostuu useista eri osa-alueista, joita ovat hallinto, henkilöstö, rakenteellinen ja tekninen turvaaminen, käyttötoiminta, toiminnan varmistaminen, jatkuvuussuunnittelu, käyttövaltuuksien hallinta ja asiakirjaturvallisuus, tietoliikenteen turvaaminen ja oheispalvelut, laadunvarmistus, dokumentointi ja järjestelmän muuttaminen sekä vakuutukset, sopimukset ja vastuut.

Hallinnon tietoturvavastuusiin kuuluu tietoturvastrategian luonti sekä strategian mukaiset toimet turvallisuuspolitiikan luomiseksi ja koulutuksen järjestämiseksi. Hallinnon vastuulla ovat myös valitun strategian organisointi-, vastuu- ja valvontakysymykset.

Oma henkilöstö on aina riski. Siksi onkin tärkeää ylläpitää henkilöstön voimavaroja, motivaatiota ja luotettavuutta, määrittellä vastuut ja huolehtia jatkuvasta koulutuksesta, jotta taataan henkilöstön tietopääoman säilyvyys ja jatkuvuus. Koko henkilöstön on myös sitouduttava organisaation tietoturvastrategiaan.

Rakenteelliseen ja tekniseen turvaamiseen kuuluu suojautuminen tulipaloilta, vesivahingoilta, sähkökatkoksilta, ilkvallalta ja varkauksilta. Suojautumistapoja ovat rakenteellisesti turvalliset tilat, paloturvallisuus, kulunvalvonta ja tietovälineiden säilytys.

Tietojen ja järjestelmien käytön sekä operoinnin tulee tapahtua sovittujen menettelytapojen ja ohjeiden mukaisesti. Vikatilanteisiin tulee varautua, järjestelmän käytettävyyttä ja siinä esiintyviä häiriöitä tulee jatkuvasti seuranta. Erilaisten teknisten varmistuskeinojen avulla taataan laitteiden toiminta, jatkuva sähkön saanti, riittävä ilmastointi sekä itse tiedot ja ohjelmistot sekä niitä sisältävät mediat. Jatkuvuussuunnittelun tarkoituksena on valmistautua pitkäaikaisien keskeytysten varalle ja varmistaa keskeytyksistä toipuminen. Tietojen suojaus ja valtuuttamattoman käytön esto tapahtuvat etupäässä teknologian avulla. Menetelmiä ovat salaus, valtuuksien kontrollointi sekä kulunvalvonta.

Tietoliikenteen turvaaminen tapahtuu myös etupäässä teknisten ratkaisujen avulla, mutta unohdetaan ei sovi myöskään inhimillistä tekijää. Laadunvarmistuksen, dokumentoinnin ja järjestelmän muuttamisen tavoitteena on turvallisempien tietojärjestelmien suunnittelu, ylläpito ja käyttö.

Sopimus- ja vastuukysymykset vaikuttavat tietoturvaan. Organisaation kriittisten järjestelmien keskeytyksistä aiheutuviin menetyksiin voidaan varautua myös vakuutuksin.

Tietoturvaliikalla tarkoitetaan turvatason asettamista tietoturvalle sekä keinoja tason saavuttamiseksi. Turvallisuus maksaa aina, joten sopiva taso on syytä valita suhteessa turvattavan materiaalin arvoon. Liian korkea taso voi myös heikentää ja vaikeuttaa toimintaa. Tietoturvassa ei kannata pyrkiä täydellisyyteen, vaan optimaaliseen tasoon, joka voidaan määritellä esimerkiksi uhka-analyysin perusteella. Optimitasoon vaikuttavat eniten turvattavan materiaalin arvo niin sen omistajalle kuin mahdolliselle hyökkääjällekin.

Vahingot voivat olla potentiaalisia, tietoihin kohdistuvia tai laitteistoon/ohjelmistoon kohdistuvia. Potentiaalisiin vahinkoihin kuuluvat järjestelmälle harmittomat, mutta toimintaa vaarantavat tapahtumat, kuten varmistuskopioiden laiminlyönti, virheitä tai viruksia sisältävien ohjelmien käyttö, ulkopuoliset tunkeutajat, ulkopuolisten suorittama resurssien tuhlaus ja järjestelmän käyttö rikolliseen toimintaan. Tietoihin kohdistuvia vahinkoja ovat tiedon valtuuttamaton käyttö ja varastaminen, tiedon tuhoutuminen ja/tai vahingoittuminen (*korruptoituminen*) sekä väärän tiedon luonti ja levitys. Laitteistoon tai ohjelmistoon kohdistuvia vahinkoja ovat ohjelman muuttaminen tai tuhoaminen, laitteiston muuttaminen, tuhoaminen tai varastaminen ja satunnaiset virhetoiminnot laitteistossa.

11.2 Suojautuminen

Loogiset menetelmät

Ohjelmiston avulla toteutettavat menetelmät, joilla teknisesti estetään järjestelmään tunkeutuminen ja tiedon vuotaminen. Tärkeimpiä tehtäviä ovat:

- Tiedon eheyden varmistaminen eli, että tietoa ei ole tahallisesti tai vahingossa muutettu ja että tiedolla on sellainen tekninen ja sisällöllinen laatu, että se on käyttökelpoista tarkoitukseensa ja täyttää siihen asetetut odotukset (*integrity*).
- Tiedon luottamuksellisuuden varmistaminen eli, että tieto säilyy vain niiden henkilöiden käytettävissä, joille se on tarkoitettu (*confidentiality*).
- Tapahtuman/tiedon alkuperän kiistämättömyys. Kiistämättömyydellä tarkoitetaan sitä, että tiedon lähettäjä ei jälkikäteen voi kiistää lähetettyä tietoa (*non-repudiation*).
- Käyttäjän/tiedon alkuperän tunnistus ja todennus (*authentication*).
- Pääsyn valvonta eli käyttäjien pääsyä koneessa olevaan tietoon rajoitetaan ja valvotaan (*access control*).
- Järjestelmän/tiedon käytettävyyden varmistaminen ja valvonta sekä ajantasaisuus (*availability*).

Fyysiset menetelmät

Mekaaniset ja rakenteelliset menetelmät rajoittaa pääsyä. Keinoja ovat mm. kulunvalvonta, rakenteellisesti turvalliset tilat ja kaapeleiden ja laitteiden sijoitus.

Ylläpidon toimet

Teknisten ja mekaanisten keinojen toteutus sekä järjestelmän jatkuva seuranta, virheiden korjaaminen, varmuuskopiointi ja raportointi.

Hallinnolliset menetelmät

Oman henkilöstön muodostaman uhkan minimoiminen käyttöä tarkkailemalla ja poikkeustilanteita selvittämällä sekä henkilöstöä kouluttamalla, motivoimalla ja lisäämällä sitoutumista sekä selkeyttämällä vastuita. Tietoturvaan kiinnitetään entistä enemmän huomiota myös lainsäädäntöä ja standardeja kehitettäessä.

11.3 IP- ja TCP-protokollatasojen turvamekanismit

IP-osoitteet

Perusprotokollatasolla IP-osoitteet muodostavat käytännössä ainoan turvamekanismin, johon edes jossain määrin voi luottaa. Tämä perustuu ajatteluun, että lähettäjä ei voi saada vastausta, ellei IP-osoite ole oikea. TCP-yhteyden lähdeosoite on kuitenkin helppo väärentää ja reititysprotokollien avulla on mahdollista ohjata IP-paketit aivan eri paikkaan, eikä aina ole edes tarpeen saada vastauspaketteja perille.

Luotetut portit

Luotetut portit ovat Berkeleyn yliopiston kehittämä menetelmä, jolla huolehditaan siitä, että alle 1024:n olevan porttinumeron voi varata vain prosessi, jolla on pääkäyttäjän oikeudet. Eli, jos toinen osapuoli on Unix-järjestelmä ja yhteys tulee portista, jonka numero < 1024, voidaan periaatteessa luottaa siihen, että toisessa päässä oleva prosessi on pääkäyttäjän hyväksymä, eikä valehtelee. Menetelmä ei ole kovinkaan laajassa käytössä.

TCP-sekvenssinumerot

TCP-yhteyden alussa molemmat liikennöivät osapuolet ilmoittavat toisilleen sekvenssinumerot, joita käytetään vuonohjaukseen ja yhteyden aiempien pakettien tunnistukseen. Mahdollisen hyökkääjän on siis pyrittävä saamaan selville tämä numero eli hyökkääjän on saatava paluupaketit näkyviin (salakuuntelu ja/tai reitityksen sotkeminen) tai yritettävä arvata paluupaketin sekvenssinumero. Käytännössä molemmat vaihtoehdot ovat melko hankalia.

IP Security Option

Tämä menetelmä lisää jokaiseen pakettiin turvaleiman, joka kertoo turvatason ja organisaation, mutta nämäkin välitetään selväkielisinä.

Palomuurit

Palomuurien avulla pyritään lisäämään sisäpuoleisen verkon turvallisuutta ulkopuoleiseen nähdän. Palomuurien perusidea on liikenteen suodattaminen, joka perustuu pääosin IP-osoitteisiin.

11.4 Tunnettuja turvallisuusaukkoja protokollissa

ARP-väärennys

ARP (*Address Resolution Protocol*) on verkossa olevien laitteiden IP-osoitteiden selvittämiseen tarkoitettu protokolla. ARP-kysely on broadcast-tyyppinen Ethernet-paketti, jolla kysytään, mikä Ethernet osoite vastaa tätä IP-osoitetta? Hyökkäys järjestelmään voi tapahtua Ethernet-segmenttiin liitetyn koneen kautta, yrittämällä vastata ARP-kyselyyn ennen oikeaa konetta. Onnistuessaan hyökkäyksen kohde kuvittelee keskustelevansa eri koneen kanssa kuin mitä itse asiassa tapahtuu.

IP-lähdeosoitteiden väärentäminen

IP-lähdeosoitteita väärentämällä voidaan kiertää IP-osoitteisiin perustuvat autentikointimenetelmät.

IP-lähdereitityksen hyväksikäyttö

IP-protokolla sisältää mahdollisuuden kiertää reitittimien keskenään sopiman reitityksen esim. tilanteissa, jolloin reititystaulut ovat menneet sekaisin. Jos halutaan väärentää oma IP-osoite siten, että saadaan myös vastauspaketit takaisin, tämä onnistuu helpoiten juuri lähdereitityksen avulla.

Reititysprotokollien hyväksikäyttö

Reitittimen reititystauluja on mahdollista muuttaa. Tällöin hyökkäyksen kohteen olevan koneen paluupaketit voidaan kääntää tulemaan hyökkääjän koneelle.

Sähköpostin väärentäminen

Koska sähköpostiviestit välitetään selväkielisinä, eikä yhteyden ottavaa konetta tarkisteta, on sähköpostiviestin väärentäminen erittäin helppoa esim. ottamalla TELNET-yhteys sopivan koneen SMTP-porttiin.

Verkon salakuuntelu

Mikäli hyökkääjä pääsee fyysisesti käsiksi johonkin verkon osaan, on liikenteen salakuuntelu mahdollista lähes millä tahansa tietokoneella kaupallisten diagnostiikkaohjelmien avulla.

Nimipalvelun käänteisrelaatio

Internetin nimipalvelun DNS (*Domain Name Service*) avulla muutetaan domain-nimet vastaaviksi IP-osoitteiksi ja käänteisrelaatioissa toiseen suuntaan. Tavallinen nimipalvelu on melko luotettava, mutta käänteisrelaatio voi palauttaa periaatteessa minkä IP-osoitteen tahansa. Useimmat TCP/IP-ohjelmat suojautuvat käänteiseltä nimipalvelun kautta tapahtuvilta hyökkäyksiltä tekemällä aina käänteisen nimipalvelun osana tavallisen nimipalvelun kyselyn.

Verkkomyrskyt

Järjestelmään voidaan tarkoituksellisesti lähettää niin paljon liikennettä, että koneen kaikki prosessointikapasiteetti kuluu tämän liikenteen käsittelyyn. Hyökkäys ei ole varsinainen tunkeutuminen, mutta se estää koneen muun käytön.

11.5 Unix/Linux käyttöjärjestelmien rakenne ja suojaukset

Käyttöjärjestelmien turvallisuus perustuu laitetasolla toteutettuihin suojausratkaisuihin, joita käyttöjärjestelmä hyödyntää. Unix-järjestelmissä suojaukset kohdistuvat prosesseihin ja tiedostoihin eli käyttäjiin.

Jokaisella Unix-järjestelmän oliolla on omistaja ja omistajaryhmä. Tiedostotasolla käyttäjäoikeuksia voidaan määritellä kolmelle eri käyttäjäluokalle; omistaja, omistajan ryhmä ja kaikki muut. Prosesseille ei ole määritelty omia käyttöoikeuksia, koska prosessin käyttöoikeus on vain prosessin käynnistäjällä ja ylikäyttäjällä. Jokaisella luokalla voi olla kolme erillistä oikeutta:

r: oikeus lukea

w: oikeus kirjoittaa

x: oikeus suorittaa

Kaikki järjestelmän toiminnot tapahtuvat prosesseina, verkkopalveluita tarjoavat daemonit ovat siis myös prosesseja. Jokaisella prosessilla on oma osoiteavaruutensa ja prosessit saavat käyttöjärjestelmältä aikaa tehtävänsä suorittamiseen. Prosessien osoiteavaruus koostuu kolmesta osasta: ohjelmakoodista, datasta ja pino-osuudesta.

Tiedostojen suojaukset tarkistetaan aina niiden käsittelyä yrittävän prosessin suhteen. Ylikäyttäjä (*root*) voi tehdä mitä tahansa muutoksia käyttöoikeuksista ja omistajista riippumatta.

Käyttäjä- ja systeemitilat

Käyttäjä- ja systeemitilojen tulee olla erilliset. Käyttäjätila on käyttäjän prosessien suorittamista varten ja systeemitila järjestelmän omaa toimintaa varten. Prosessori on jokaisella hetkellä jommassa kummassa tilassa ja systeemitilassa sillä on täysi käsittely- ja hallintaoikeus koko laitteistoon.

Käyttäjätilassa voidaan määritellä looginen osoiteavaruus, jonka ulkopuolelle prosessori ei voi viitata. Prosessi voi pyytää käyttöjärjestelmältä palveluita systeemikutsujen avulla. Tällöin tarkistetaan kunkin kutsun laillisuus.

Systemikutsut

Prosessi pyytää palveluja systeemikutsujen avulla. Tämän seurauksena prosessin tila vaihtaa systeemitilaan. Systeemikutsut ovat ainoa tapa, joilla prosessit voivat vaikuttaa oman osoiteavaruutensa ulkopuolelle. Jokainen systeemikutsu tarkistaa pyydetyn toimenpiteen laillisuuden. Tämä tapahtuu parametrien arvon suhteen sekä katsomalla, onko pyydetty toimenpide sallittu suojausmekanismin suhteen.

Suojauskeskeytykset

Tietokoneen muut osat voivat keskeyttää keskusyksikön toiminnan tarvitessaan palvelua. Mikäli käyttäjätilassa oleva ohjelma yrittää lukea tai kirjoittaa tietoa oman osoiteavaruutensa ulkopuolelta/ulkopuolelle tai oheislaitteilta/oheislaitteille tai suorittaa vain systeemitilassa sallitun käskyn, aikaansaa laitteisto keskeytyksen. Tällaiset tilanteet johtuvat yleensä ohjelmointivirheistä.

Salasanat

Käyttäjälle näkyvin suojausmenettely on salasanan käyttö, joka Unixissa on toteutettu käyttöjärjestelmän ytimen ulkopuolisella sovellusohjelmalla. Unix-järjestelmässä käyttäjätunnukset on talletettu */etc/passwd*-tiedostoon ja kryptatut salasanat sijaitsevat */etc/shadow*-tiedostossa, joka ei passwd-tiedoston tapaan ole kaikkien luettavissa.

Lokit

Unixin lokijärjestelmä kerää tietoja järjestelmän tapahtumista ja siihen kirjautuvat myös mahdolliset murtautumisyrietykset järjestelmään. Unix-koneissa käytetään yleensä syslog-järjestelmää. Unix-koneissa voidaan turvallisuusriskejä analysoida erilaisilla ohjelmilla, jotka paljastavat aukkoja niin käyttöjärjestelmässä kuin ohjelmissakin. Esimerkkinä tällaisista ohjelmista on SATAN (*Security Administrator Tool for Analyzing Systems*).

Varmuuskopiointi

Varmuuskopiojärjestelmä ei sinänsä liity tietoturvaan, mutta se suojaa vahingoilta ja helpottaa tietoturvamurron vahinkojen selvittelyä sekä mahdollisesti helpottaa murron selvitystä ja tekijän kiinnisaantia.

11.6 Unixin tunnettuja turvallisuusaukkoja

Unix-järjestelmän turvallisuusaukot liittyvät useimmin suid-ohjelmien ja verkkopalveluohjelmien ohjelmointivirheisiin.

Suid/sgid-ohjelmien virheet

useiden Unix-palvelinohjelmien on pakko jossain vaiheessa toimia pääkäyttäjän (*root*) oikeuksin. Lisäksi ne pääsevät koko järjestelmän tiedostoavaruuteen. Tällaisessa tilassa ohjelmassa

olevien virheiden seuraukset voivat olla kohtalokkaita ja turva-aukkoja syntyy helposti, mikäli ohjelmat eivät toimi oikein.

Verkko-ohjelmistojen virheet

Nämä virheet ovat saman tyyppisiä kuin suid/sgid-ohjelmien virheet.

Kilpailutilanne ja siitä aiheutuvat suunnittelu- ja määrittelyvirheet sekä lapsukset

Uusien ohjelmien ja tuotteiden mahdollisimman nopea markkinoille saattaminen aiheuttaa puutteita niin ohjelmien suunnittelussa kuin toteutuksessakin. Eivätkä virheet tule välttämättä esille ohjelmien testauksessakaan, mikäli se tehdään nopeasti ja liian pienellä testiaineistolla, joka ei kata kaikkia mahdollisia tilanteita. Melkeinpä jokaisesta ohjelmasta löytyy tällaisia virheitä. Ne eivät välttämättä ole turvallisuusaukkoja ja niiden aiheuttamat ongelmatilanteet saattavat esiintyä vain tietyissä (poikkeavissa) tilanteissa.

Hyvillekin ohjelmoijille sattuu tahattomia virheitä. Eräs yleinen paikka ovat puskurit, joihin tuotavan datan kokoa ei tarkisteta.

11.7 Internet ja WWW

Alunperin Internet on suunniteltu hyvinkin avoimeksi verkoksi, joten siinä ei ole luotettavia mekanismeja, joilla verkkoon liitetyt koneet tai verkossa liikkuva tieto voitaisiin suojata. Menetelmiä on myöhemmin kehitetty, mutta mitään täysin luotettavaa menetelmää ei ole olemassa.

Internetin voimakas laajeneminen aiheuttaa omat riskinsä. Sen tulo yhä useampien ulottuville kasvattaa potentiaalisten väärinkäyttäjien määrää. Verkossa liikkuvan tiedon oikeellisuudesta tai alkuperästä ei ole varmuutta ja suoranainen verkkorikollisuus on myös kasvussa. Murtautujien ensisijaisia kohteita ovat moniajojärjestelmät ja murtautuessa käytetään hyväksi ohjelmien virheitä ja vääriä konfigurointeja. Koska verkko on kansainvälinen, ei sillä ole selvää yksittäistä tai nimettävissä olevaa ylläpitäjää.

On syytä tehdä ero itse verkon turvaominaisuuksien ja siihen liitettävän yksittäisen tietokoneen tai tietokoneiden turvaominaisuuksien välille. Useimmiten riskit liittyvätkin itse tietokoneeseen, eikä niinkään verkkoon. Riskit sijoittuvat seuraaviin kategorioihin:

käytettävä tietoliikenneverkko ja sen solmut, tietoliikenneprotokollat, verkon hallinnointimekanismit, verkkoon liitetyt tietokonejärjestelmät (palvelinohjelmien ja käyttöjärjestelmien turvaaukot), huonosti valitut salasanat tai esille jätetyt salasanat sekä ylläpidon puutteellisuus.

Kukin kone voidaan erikseen suojata siten, ettei asiaton pääse verkon kautta koneeseen. Suojaus voidaan myös keskittää yhteen paikkaan, joko rajoittamalla sisään tulevia yhteyksiä reitittimen pääsilystoilla tai erityisellä suojamuurikoneella. Turvallinen Internet-liittymä rakennetaan toisiinsa täydentävillä menetelmillä ja työkaluilla. Keskitetyllä suojauksella voidaan siis täydentää yksittäisten laitteiden suojausta. Erittäin tärkeä tekijä Internet-liittymässä on myös järjestelmän jatkuva ja asiantunteva ylläpito.

11.8 Internetin turvaratkaisuja

Palomuurit (*Firewall*)

Palomuuuri on laitteisto tai ohjelma tai niiden yhdistelmä, joka rajoittaa Internetin ja lähiverkon liikennettä. Kaikki liikenne rajoitetaan kulkemaan tämän yhden pisteen kautta. Palomuurilla voidaan korvata verkon protokollien puutteita. Lisäksi voidaan kerätä ja taltioida tietoa tapahtuneesta liikenteestä. Palomuuuri-ratkaisussa lokitiedot kerätään sisäisessä verkossa olevalle palvelimelle. Tällöin lokitietoja ei voida muuttaa tai lukea palomuuuri-järjestelmän kautta.

Emme tarkastele palomuuureja tässä yhteydessä tämän tarkemmin, koska palomuuureja ei ole käytössä yliopiston liittymissä Internetiin.

Pääsyylistat

Yleisin tapa hoitaa Internet-liittymän suojaus ovat pääsyylistat liittymää hoitavassa reitittimessä tai palvelimen sisältävässä koneessa. Yhteyksiä voidaan rajoittaa joko osoitteiden tai palveluprotokollien perusteella. Reitittimet osaavat yleensä havaita, kummalta puolelta yhteydenavaus on tapahtunut. Pääsyylistojen suunnittelu ja ylläpito on kuitenkin melko raskasta.

WWW-palvelinohjelmisto sisältää itsessään ominaisuuksia, joilla voidaan rajata tietyt html-sivut näkymään vain halutuille katselijoille. Toteutukset perustuvat tavallisimmin kyselijän osoitteen tarkistamiseen, eli tarkistetaan, mistä koneesta pyyntö tulee. Eri valmistajien http-palvelimien suojausominaisuudet vaihtelevat (toteutukseltaan), mutta toimivat edellä mainittujen periaatteiden mukaisesti.

Koneen osoitteeseen perustuvat pääsyylistat tehdään yksittäisen koneen tai osoiteryhmän tarkkuudella. Suojaus määritellään joko koneen nimen tai IP-osoitteen perusteella. Sisääntulevasta IP-paketista selviää kutsuvan koneen osoite. DNS-palvelin selvittää tätä osoitetta vastaavan nimen, jossa on silloin mukana Internet-domain osa (esim. pc2.helsinki.fi). Koneen nimien perusteella tehdyt listat ovat raskaampia käsitellä kuin IP-osoitteisiin (esim. 128.214.71.58) perustuvat listat.

Käyttäjätunnukset ja salasanat

Käyttäjätunnusten avulla tehdään vain tietyille käyttäjille tarkoitettut WWW-sivut. Käyttäjälle määritellään käyttäjätunnus ja salasana, joilla suojattua sivu päästään katselemaan. Käyttäjien tunnistus on rakennettu HTML-protokollan sisälle. Suojatun palvelun hakupyntö ei onnistu ensimmäisellä kerralla, mutta vastauksessa käyttöliittymä kysyy tunnistetiedot (käyttäjätunnus ja salasana) käyttäjältä ja toteuttaa hakupyynnön uudelleen, mikäli tunnistetiedot ovat oikeat.

Tiedostossa `/usr/local/etc/httpd/conf/access.conf` määritellään suojaukset kullekin alihakemistolle. Toinen mahdollisuus on tehdä kuhunkin hakemistoon `.htaccess`-tiedosto, jossa hakemisto-kohtaiset suojaukset määritellään.

Käyttäjätunnukset ja salasanat kulkevat salaamattomina tai korkeintaan kevyesti sekoitettuina, joten niiden selvittäminen on helppoa.

WWW-sivut voivat palvelimella sijaita useissa eri hakemistoissa ja kullekin hakemistolle on mahdollista määritellä erilliset suojausparametrit.

WWW-yhteydet

Myös WWW-yhteydet käyttävät siirtoteitä, jotka eivät ole täysin turvallisia. Siksi myös WWW-liikennettä on mahdollista salata. Ratkaisuja tähän ovat Netscapen SSL (*Secure Sockets Layer*) sekä Secure HTTP. Salauksen lisäksi voidaan toteuttaa myös käyttäjien tunnistaminen sekä tiedon kiistämättömyys elektronisten allekirjoitusten avulla.

SSL on Netscape Communicationsin kehittämä ja toteuttama protokolla, jolla voidaan toteuttaa salatut yhteydet kahden koneen välillä. Toteutus on tehty kuljetustasolla, joten samaa suojausta voidaan käyttää myös muiden protokollien yhteydessä. HTTPS on yksi SSL-salausta käyttävä protokolla, jossa HTTP-komentoja kuljetetaan SSL-salatun yhteyden yli. HTTPS ja SSL on toteutettu Netscape Navigator WWW-client:ssa ja Netscape Commerce Server -palvelimessa. Nykyinen toteutus salaa vain työaseman ja palvelimen välisen liikenteen. Käyttäjien tunnistus saataneen myöhemmin. SSL on jo osoittautunut haavoittuvaksi.

S-HTTP on EIT:n (*Enterprise Integration Technologies*), RSA:n ja NCSA:n kehittämä ratkaisu. Siinä toteutuvat salattu siirto ja käyttäjien tunnistus elektronisilla allekirjoituksilla. Toteutus perustuu PGP- ja PEM-menetelmien soveltamiseen HTTP-protokollassa. Toteutusta ei ole saatavana johtuen Yhdysvaltojen vientisäännöksistä. Tekstiä tunnistettaessa automaattisesti tunnistettavissa, onko WWW-sivu väärennetty. Sivut on myös mahdollista salata niin, että vain tietyt lukijat voivat lukea. Ratkaisu on vasta suunnitteilla ja kehitteillä.

Kryptografiset menetelmät

Salausteknisillä ratkaisuilla tavoitellaan korkeampia tietoturvaratkaisuja kuin mitä protokolilla tai sovelluksiin sisältyvillä ratkaisuilla saavutetaan. Salauksella voidaan toteuttaa seuraavia palveluja:

- Luottamuksellisuus eli tieto ei ole luvattomien henkilöiden, olioiden tai prosessien saatavissa, eikä sitä paljasteta niille. Tieto muutetaan sellaiseen muotoon, että vain tietyn avaimen haltija pystyy sen tulkitsemaan.
- Todentaminen eli käyttäjän, koneen, ohjelmistokomponentin tai minkä tahansa muun identiteetin sekä tietolähteen varmistaminen. Tiedon muuttaminen tiettyyn muotoon on mahdollista vain tietyn avaimen haltijalle.
- Todennus- ja eheystarkisteet varmistavat lähteen alkuperän ja muuttumattomuuden siirron aikana.
- Salaustekniset tarkistusfunktiot.
- Sähköiset allekirjoitukset. Allekirjoitettava tieto muutetaan allekirjoitetuksi sähköiseksi dokumentiksi allekirjoitusalgoritmia ja omaa salaista avainta käyttäen. Vastaavalla julkisella avaimella voidaan todentaa tiedon aitous. Sähköistä allekirjoitusta käytetään tietolähteen, eheyden ja kiistämättömyyden varmistukseen.
- Kiistämättömyys.
- Automaattinen avainten jakelu ja vaihto.
- Pääsyn valvonta

- Tietoturvatiivisteet. Tiivistä algoritmeilla (*hash function*) lasketaan mielivaltaisen pituisesta tiedosta lyhyt, määrätyn kokoinen tiiviste.

USA:n vientirajoitukset estävät useiden turvallisuutta lisäävien tuotteiden maastaviennin. Monesta tuotteesta on kuitenkin olemassa myös vientiversiot, jotka on toteutettu lyhyemmillä avaimilla. USA on hiljattain lieventänyt vientikieltoa siten, että tällä hetkellä sallitaan jo 128-bittisten avainten maastavienti.

Nykyiset kryptografiset menetelmät voidaan jakaa kahteen päätyyppiin:

Symmetriset menetelmät

Symmetrisessä menetelmässä lähettäjällä, joka suorittaa tietojen salausmuunnoksen lähetystä tai arkistointia varten, ja vastaanottajalla on sama salainen avain. Symmetriset menetelmät ovat joko jono- tai lohkosalaajia. Jonosalaajat voidaan edelleen jakaa synkronisiin ja synkronoituviin. Jonosalaaja salaa tietoa lyhyt lohko (bitti tai tavu) kerrallaan ja kukin lohko eri muunnoksella. Lohkosalaaja salaa dataa pitkä lohko (64–128 bittiä) kerrallaan samalla monimutkaisella, avaimesta riippuvalla muunnoksella.

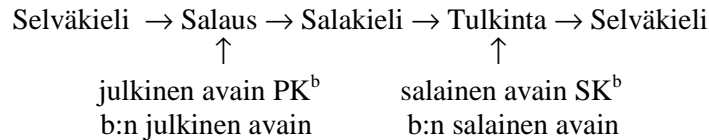
Laitevalmistajien salausalgoritmit ovat yleensä jonosalaajia ja suurin osa julkistetuista algoritmeista ovat lohkosalaajia. Synkroninen jonosalaaja sietää hyvin siirtotiellä syntyneitä virheitä vastaan, joten se ei siksi sovellu käytettäväksi todentamismekanismeissa. Lohkosalaajia voidaan käyttää myös eheyden tarkistamiseen sekä tiivistysfunktiona. Symmetriset lohkosalaajat käyvät kolmannen osapuolen välityksellä myös kiistämättömyyspalveluun. Jonosalaajat ovat nopeita ja voivat käsitellä kaikenlaista tietoa, kun taas lohkosalaajat edellyttävät tiettyä lohkomuotoa. Esimerkkejä

symmetrisistä menetelmistä ovat DES (*Data Encryption Standard*) ja IDEA (*International Data Encryption Algorithm*). DES on kohtalaisen turvallinen. Salauksessa käytetään 56-bittistä avainta, joka mahdollistaa 256 eri avainyhdistelmää. IDAEA:ssa puolestaan käytetään 128-bittistä avainta, joka mahdollistaa 2128 avainyhdistelmää. USA:n NSA (*National Security Agency*) kykenee todennäköisesti purkamaan DES-kryptatut viestit. Esimerkki symmetrisestä salauksesta:



Epäsymmetriset menetelmät (julkinen avain)

Avaimet ovat käyttäjäkohtaisia. Julkisten avainten menetelmässä avaimisto muodostuu julkisista ja salaisista avaimista. Kukin käyttäjä pitää oman salaisen avaimen omana tietonaan. Kuka tahansa, jolla on tiedossaan tietyn käyttäjän julkinen avain, voi salata tätä avainta käyttäen tiedot niin, että ainoastaan kyseinen käyttäjä, jolla on vastaava salainen avain, pystyy tulkitsemaan sen. Esimerkki epäsymmetrisestä lohkosalausmenetelmästä on RSA (*Rivest-Shamir-Adelman*), joka on lähes mahdoton murtaa. Menetelmä perustuu suurien lukujen jakamiseen tekijöihin. Esimerkki epäsymmetrisestä salaamisesta:



RSA menetelmää käyttävistä sovelluksista mainittakoon PGP (*Pretty Good Privacy*) ja PEM (*Privacy Enhanced Mail*). PGP on sähköpostin salausjärjestelmä, joka toimii MS-DOS-, UNIX- ja VAX/VMS-alustoilla ja perustuu RSA/IDEA salausalgoritmeihin ja julkisiin avaimiin. Ohjelma on yksityiskäyttäjille ilmainen. PGP:llä aikaansaadaan kohtuullinen luotettavuus eri avaimien oikeellisuudesta ja se mahdollistaa dokumentin allekirjoittamisen kryptografisesti. Vastaanottaja voi siis varmistua saamansa aineiston alkuperästä. Julkisen avaimen talletus ja jakelu tapahtuu ns. kolmannen osapuolen toimesta. PGP on toistaiseksi murtamaton. Sen käytön leviämistä hidastaa kuitenkin hankala käytettävyys. PEM on RSA-kryptaukseen perustuva sähköpostin salausmenetelmä, jolla voidaan toteuttaa myös lähetyksen kiistämättömyys.

Toinen RSA sovellus on pääteyhteyksien SSH (*Secure Shell*), jossa kaikki liikenne salataan, kommunikoivat osapuolet tunnistetaan ja tiedon muuttumattomuus taataan. Edellytyksenä on, että kohdekoneessa on SSH-palvelin ja yhteys muodostetaan SSH-ohjelmalla.

Epäsymmetrisillä RSA-avaimilla voidaan toteuttaa datan selauksen lisäksi myös liikennöivien osapuolten tunnistaminen. Tämä toteutetaan liittämällä julkisiin RSA-avaimiin tunnistetiedot, jonka perusteella nähdään, kenen avain kyseessä on. Näin saadaan nk. Sertifikaatti, joka sisältää seuraavat tiedot:

Palvelimen julkinen avain, palvelimen ja organisaation nimi, avaimen sarjanumero, voimassaoloaika ja sertifikaatin varmentajan tunnistetiedot ja allekirjoitus.

Sertifikaatin varmentaa kolmas, ulkoinen, molempien osapuolten luottama elin, joka omalla allekirjoituksellaan varmistaa avaimen kuuluvan sertifikaatissa mainitulle omistajalle.

Avaimet

Salaustekniikan keskeinen ongelma ja tehtävä on julkisten ja salaisten avainten ylläpito ja jakelu. Jakeluorganisaationa toimii ns. luotettu kolmas osapuoli. Avainten elinikä on rajoitettu. Epäsymmetrisen menetelmän avainten hallinta laajassa hajautetussa ympäristössä on tarkoituksenmukaista järjestää yhden tai useamman varmenneviranomaisen toimesta. Julkisen avaimen menetelmän avaimen salaisen osan voi muodostaa joko käyttäjä itse tai varmenneviranomainen. EU:n Infosec-ohjelman toinen vaihe keskittyy tähän problematiikkaan ja Suomessa on meneillään TIVEKE-hanke saman asian tiimoilta.

Sähköinen identiteetti tarkoittaa henkilön luotettavaa sähköistä tunnistamista. Se on perusedellytys useimmille uusille ja jo käytössä oleville elektronisille palveluille. Tunnistamisen tarve ja sen vahvuus vaihtelevat sovelluksesta riippuen. Sähköinen identiteetti määritellään yhdistelmänä seuraavista:

Käyttäjän yksikäsitteinen nimi sähköisen tietojärjestelmän käyttäjärekisterissä ja käyttäjän avainpari julkisen avaimen menetelmässä.

Sekä Suomessa että muualla maailmassa on meneillään kehittämishankkeita sähköisen tunnistamisen ratkaisemiseksi.

11.9 Joitakin WWW:n turva-aukkoja

Palvelinohjelmissa esiintyvien ohjelmointivirheiden lisäksi tietoturvariskejä voivat aiheuttaa seuraavat tekijät:

CGI-scriptit

CGI-scriptit ovat HTML-palvelimen käynnistämiä ohjelmia, jotka suorittavat oman tehtävänsä itsenäisesti. Kaikki CGI-scriptit kannattaa keskittää yhteen hakemistoon, jolloin palvelimen ylläpitäjän on helpompi seurata ohjelmien toimintaa.

JAVA

JAVA-ohjelmat toimivat siten, että etäkone siirtää tekstimuotoisen ohjelman, joka suoritetaan asiakkaan koneessa. JAVAsta on löydetty useita turvallisuusaukkoja ja sen kehittäminen on vielä osin keskeneräinen ja testaamatta. Uusia JAVA-versioita on jouduttu julkaisemaan tiiviiseen tahtiin johtuen paljastuneista virheistä. Jotkut asiantuntijat ovat arvelleet, ettei kaikkia JAVAssa esiintyviä turva-aukkoja voida koskaan nykyisellä rakenteella poistaa.

Automaattiset hakemistolistaukset

Kun hakemistosta puuttuu index.html-tiedosto, voi http-palvelin automaattisesti generoida listan hakemiston sisällöstä, mikä voi paljastaa dokumentteja, jotka eivät ole tarkoitettu ulkopuolisten nähtäviksi.

Symbolisten linkkien seuranta

Linkkien käyttö mahdollistaa kuvaukset hakemistoihin, jotka eivät ole suoraan HTDOCS-hakemistohaaran alla. Tällöin on vaarana, että joku tekee viittauksen hakemistoon, jonka ei ole tarkoitus näkyä WWW-palvelimen kautta.

Selainohjelmien virheet

Uudessa Microsoftin Internet Explorer 4.0:ssa on havaittu olevan ammottava turvareikä, joka tarjoaa hakkereille mahdollisuuden muuttaa verkossa surffailijoiden koneissa lähdekoodia lähes mielin määrin. Explorerin käyttöä ei siksi suositella ennen kuin Microsoft korjaa selaimen puutteen.

Virukset

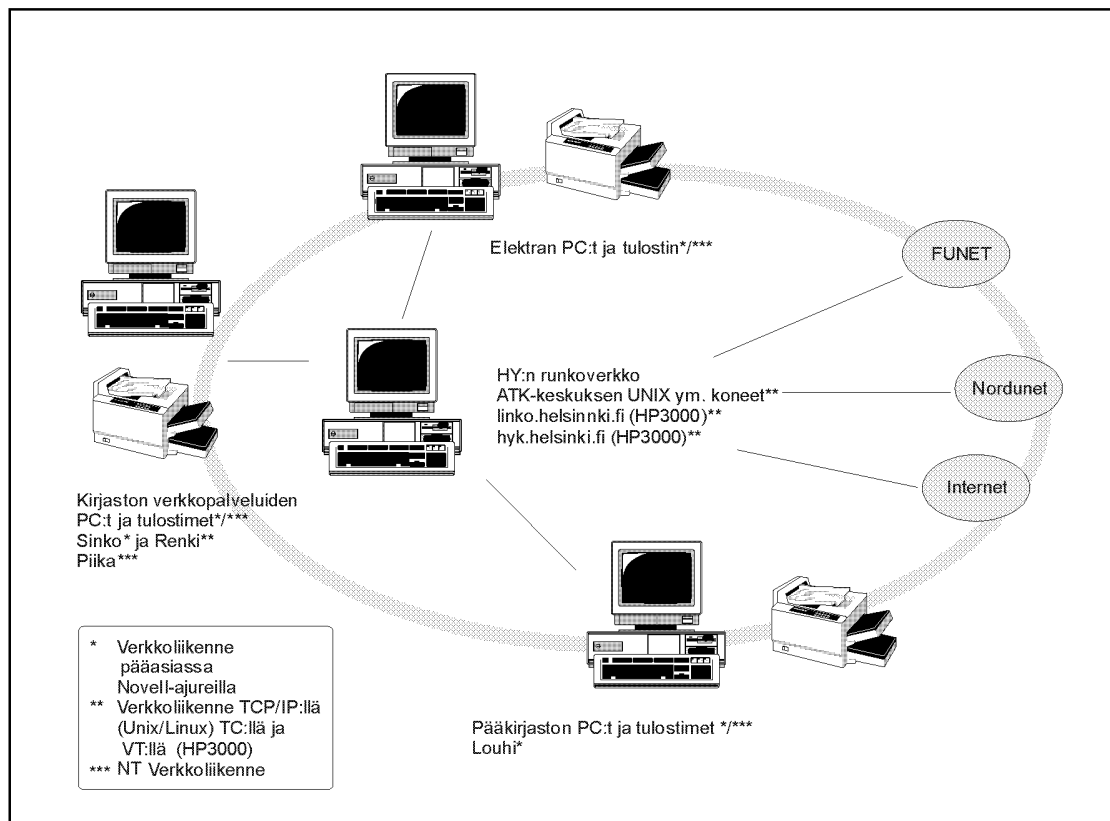
Internetin kautta leviävien virusten lukumäärä on toistaiseksi melko pieni, mutta uudentyyppisiä, myös verkon kautta leviäviä viruksia on varmasti tulossa. Esimerkkinä tämän tyyppisistä viruksista ovat sähköpostin liitetiedostoina lähetettävät datatiedostojen makroiin piilotetut virusohjelmat. Paras virusten torjuntakeino on luotettavien lähteiden käyttö ja tiedon tietojen tarkistus ennen käyttöä esimerkiksi automaattisten virusten torjunta- ja etsintäohjelmien avulla.

Tulevaisuudessa Internet-apuohjelmissa on sisäänrakennettuna perustarkistukset tai virusten etsintäohjelmat. Nykyisistä Internetin virustorjuntaohjelmista mainittakoon WebScan.

11.10 Tietoturva ELEKTRA-projektissa

ELEKTRA-projektin arkisto- ja WWW-palvelimessa, joka on Pentium Pro -tasoinen PC-laitteisto ja jonka käyttöjärjestelmänä on Linux, sijaitsevat NSCA:n WWW-palvelinohjelmisto, Hyperlib-käyttöliittymäohjelmisto sekä valtaosa palvelun kautta saatavilla olevista dokumenteista PDF- ja HTML-muodossa. Palvelimelta on poistettu joukko tarpeettomia ohjelmia ja siinä on edellä mainittujen lisäksi Internetiin liittyvistä ohjelmistoista vain FTP-palvelin, sähköpostiohjelma sekä WWW-selain. Palvelimelle ei saa asentaa käyttäjien henkilökohtaisia ohjelmia, eikä tiedostoja.

Palvelin on osa Helsingin yliopiston verkon laitekantaa. Yliopiston verkko kuuluu FUNETin kautta koko Internetiin (kuva 11.1). ATK-keskus on toistaiseksi noudattanut nk. avoimien ovien politiikkaa ja se vastaa myös verkon liikenteen tarkkailusta. Jos ulkopuolisten pääsy haluttaisiin estää tai rajoittaa, voitaisiin se toteuttaa palomuurina toimivan reitittimen avulla, mutta tähän tuskin kannattaa ryhtyä, sillä verkon tarkkailu vaatii runsaasti voimavaroja.



Kuva 47. ELEKTRAN tietokoneiden liittyminen Internetiin.

Koska kyseessä on kaupallinen palvelu, on palvelimen turvallisuuteen kiinnitettävä erityistä huomiota. Mahdolliset hyökkäykset palvelimeen tapahtuvat todennäköisesti hakkerien tai crackerien toimesta joko harmittomasti tai vahingontuottamistarkoituksessa, mutta ne eivät toden-

näköisesti kohdistu vartavasten palvelimen sisältämän materiaalin luvattomana käyttöönä. Kaikenlaisia hyökkäyksiä vastaan on kuitenkin varauduttava.

Mitään istuntoja, varsinkaan *root*-oikeuksin, niin muilta työasemilta kuin itse Linux-koneeltakaan ei saa jättää avoimiksi sen jälkeen kun koneen operointi on saatu tehtyä. Palvelimen kaikki tiedot ovat palautettavissa joko varmistusnauhoilta tai projektin muilta PC-laitteilta, joissa tallella ovat niin alkuperäiset PDF- ja HTML-tiedostot kuin osa niistä lähdetiedostoista, joista PDF- ja HTML-tiedostot on konvertoitu, joko PC:n omalla kiintolevyllä tai levykkeillä pakatussa muodossa. Elektronisessa muodossa toimitettu aineisto on tallella siltä varalta, että konvertoinnissa tai muuten syntyneet virheet on helppo korjata projektille asetetun velvoitteen mukaisesti. Tietyn virhemarginaalin salliminen elektronisissa julkaisuissa tuntuisi järkevältä, sillä varsinkin painetusta materiaalista skannaamalla digitoidussa materiaalissa esiintyy runsaasti virheitä, joiden löytäminen ja korjaaminen on työlästä. Sähköisen aineiston arkistointi ja pitkäaikaissäilytys ovat sinällään osa tietoturvaa ja vaativat omat ratkaisunsa. Saman tiedon säilyttäminen useassa eri muodossa vaatii erillisen arkistointipalvelimen, koska talletettavan materiaalin tarvitseman muistin koko kasvaa nopeasti. Pitkäaikaissäilytyksen ratkaisujen tulee taata myös aineiston käytettävyys pitempienkin aikojen kuluttua. Arkistointiin perehdytään tarkemmin Helsingin yliopiston kirjaston johtamassa Eva-projektissa sekä Helsingin yliopiston kirjaston elektronisen julkaisupankin yhteydessä.

ELEKTRA-projektissa pääsy WWW-palvelimelle on rajoitettu määrittelemällä laitteistotasolla Linux-järjestelmän *access*-tiedostoon ne IP-osoitteet, joista voidaan olla yhteydessä palvelimelle. IP-osoitteisiin perustuva pääsynä rajoitetaan itseasiassa kaksinkertaisesti, sillä Hyperlib-ohjelmassa on vastaavanlainen tiedosto. Kaksinkertaisen järjestelmän ylläpito on luonnollisesti myös kaksin verroin työläämpää, mutta toistaiseksi ylläpidettävien IP-osoitteiden määrä on ollut niin pieni, ettei tehtävä ole ongelma. Kaksinkertaista pääsyn rajoitusta voidaan perustella sillä, että kaikki halutut käyttäjät voidaan päästää palvelun aloitussivuille katsomaan sieltä löytyviä tietoja; lehtilistaa, hinnastoa ja tietoja tekijänoikeuksista, ja itse palvelun käyttö voidaan rajata tätä pienemmän ryhmän käytettäväksi. Tällä hetkellä kaikki jotka pääsevät läpi laitteiston IP-tarkistuksesta, pääsevät myös käyttämään itse palvelua.

Linux-kone ja WWW-palvelin on otettu käyttöön ja niitä ylläpidetään tässä raportissa esitettyjen turvallisuusnäkökohtien mukaisesti. Koneesta otetaan täysvarmuuskopiot kerran kuukaudessa ja artikkelien ja Hyperlibin osalta varmuuskopiot kerran viikossa.

Linux tuottaa normaaleja lokitietoja. Hyperlib-ohjelmisto tuottaa lisäksi lokitietoa niistä IP-osoitteista, joista palvelimeen on oltu yhteydessä. Järjestelmän käytettävyyttä seurataan jatkuvasti ja loppukäyttäjät voivat raportoida järjestelmässä mahdollisesti esiintyvistä ongelmista sähköpostitse.

Raportissa esitettyjen turvaratkaisujen käytön tarve tulee ELEKTRAssa lisääntymään jos ja kun palveluun voidaan liittää jokin nykyistä manuaalista järjestelmää kehittyneempi maksujärjestelmä. Tällöin tulee suurempi tarve tunnistaa liikennöivät osapuolet sekä salata tapahtuva liikenne. Myös tietojen eheyden varmistamisen tarve kasvaa. Nykyisessä järjestelmässä asiakas voi varmistua tiedon laadusta tulosteita maksaessaan, mutta mikäli tulostusmahdollisuudet laajenevat ja maksu kerätään etukäteen, on tiedon oikeellisuus ja eheys varmistettava.

Projektille voidaan toimittaa materiaalia myös FTP:n avulla. Linux-kone toimii FTP-palvelimena, jonne materiaalin toimittajat voivat siirtää omia tiedostojaan. Tätä kautta tulevan materiaalin eheys todetaan käymällä se läpi. Siirrettävät tiedostot voivat olla jo valmiiksi PDF-tai HTML-muodossa, jolloin niiden aukeaminen ja sisältö tarkistetaan oikeaksi. Muun materi-

aalin osalta mahdolliset ongelmat havaitaan konvertoitaessa materiaalia PDF- tai HTML-muotoon. Saapuneen materiaalin alkuperä tulee samalla varmistettua.

Anonyymi FTP-palvelin on uhka turvallisuudella, sillä sen kautta järjestelmään voi päästä tiedostoja, joihin on piilotettu virus tai muita viruksen tyyppisiä ohjelmakoodeja. Mitään tiedostoja ei kuitenkaan avata Linux-koneessa, vaan ne siirretään ensin PC-laitteelle, jossa tiedostot tarkistetaan virusten torjuntaohjelmalla ja vasta sitten konvertoidaan haluttuun muotoon tai siirretään valmiit PDF- ja HTML-tiedostot takaisin niille määrättyyn hakemistoon Linux-palvelimella. Tiedostojen siirron turvallisuutta voidaan parantaa muuttamalla FTP-palvelu toimimaan käyttäjätunnuksen ja salasanan yhdistelmällä, mikä onkin toteutettu turvatarkastuksen yhteydessä, jossa yhteydessä koneesta myös poistettiin muutamia turhia ohjelmia ja koneen turvallisuutta muutenkin paranneltiin.

ELEKTRA-palvelun tietokanta on Helsingin yliopiston kirjaston Kirjastotoimen verkkopalvelujen HP-3000 -tietokoneessa. Tämän koneen turvallisuusratkaisuja ei tässä yhteydessä kuvata tarkemmin. Organisaation omaan käyttöön tarkoitettussa tietoturvasuunnitelmassa kerrotaan enemmän tämän koneen tietoturvasta.

Tämän raportin periaatteita ja ohjeita tulee soveltaa myös projektin tuotantokäytössä oleviin PC-laitteisiin, jotka on liitetty sekä atk-keskuksen että kirjaston omaan mikroverkkoon. PC-laitteiden tiedostot varmuuskopioidaan mikroverkon Novell-palvelimelle ja siitä edelleen nauhalle tai levykkeille. Organisaation tietoturvasuunnitelmassa on ohjeita mikroverkon käytöstä ja kaikkea kirjastossa tapahtuvaa tietojenkäsittelytoimintaa koskevat myös Helsingin yliopiston laatimat säännöt ja ohjeet.

Lähteet

Nikander, Pekka; Peltonen, Tapio; Viljanen, Lea. Internet tietoturva. Espoo: Suomen ATK-kustannus, 1996. ISBN 951-762-377-1.

Porvari, Paavo. Tietoturvallisuuden tason arviointi: kysymyssarja tietoturvatoimien tarkastamiseen ja suunnitteluun. Helsinki: Pohjola-yhtiöt. (Pohjola-yhtiöiden julkaisuja 18). ISBN 952-6949-04-5

Elektroninen kaupankäynti ja Tietoturva -seminaarin materiaali. VTT-Tietotekniikka 25.10.1996 Espoo.

WWW-palvelun ohjelmointi -kurssin materiaali. Relevantum 20.3.1997 Tampere.

Hakala, Juha. Tietoturvasuunnitelma. Helsingin yliopiston kirjasto: Kirjastotoimen verkkopalvelut.

Elektroninen julkaisupankki -työryhmän muistio 1997.

Lukuisat Internetistä löytyvät aihetta käsittelevät WWW-sivut ja artikkelit:

www.cert.funet.fi/cert.html

www.cert.funet.fi/rfc/rfc1704.txt

www.telmo.fi/tiveke/raportit/pelkrap.htm

www.telmo.fi/tiveke/raportit/tietolyh.htm

www.telmo.fi/tiveke/turva.htm

www.telmo.fi/tiveke/tiveke/trap_tt.htm

www.telmo.fi/tiveke/sty/tturav/inetsec.htm

www.telmo.fi/tiveke/sty/tturv/serysa.htm

www.telmo.fi/tiveke/sty/tturv/stand.htm

www.vn.fi/vm/suomi/muuta/vahti/suositus.htm

www.sigma.net/tdunn/aaa2/main.html

www.niksula.cs.hut.fi/~nhaatain/prtcls.htm

www.it.spt.fi/~mpenkkim/tturva/tturva.htm

www.cs.auc.dk/~johnson/secure.html

www.epm.ornl.gov/~batsell/NSC.html

csc.nsl.nict.gov/nistpubs/800-7/node113.html

cip.physik.uni-wuerzburg.de/www-security/

Edellä mainituilta sivuilta löytyy edelleen runsaasti linkkejä tietoturvaan käsitteleville WWW-sivuille.

12. KÄYTTÖKOKEMUKSET

12.1 Käyttöympäristö

ELEKTRAN haku- ja dokumentin välityspalvelu perustuu nykyisen artikkelitietokannan, ARTO:n näyttöluettelotoiminnon laajennukseen. Tarkoituksena on, että asiakkaat voivat hyödyntää näyttöluettelon hyviä hakuominaisuuksia ja korkeatasoisia bibliografisia- ja sisällönkuvailuja, jotka jo nyt tuotetaan tieteellisten kirjastojen ja muiden asiantuntijaorganisaatioiden laajana yhteistyönä. Palvelun taso parantuu oleellisesti, kun asiakkaat saavat tiedonhaun yhteydessä käyttöönsä itse varsinaiset dokumentit.

Hakupalvelu on teknisesti toteutettu VTLS-hakutietokannan ja WWW-hakupalvelimen avulla. Pilottiaineiston koekäyttöä varten on perustettu erillinen ELEKTRA-tietokanta HP-3000 tietokantapalvelimelle. Uuden tietokannan perustamiseen tarvittava lisenssi on saatu VTLS Inc:iltä, ja se oikeuttaa testitietokannan käytön projektin ajan. Tietokantaan on tallennettu testiaineiston bibliografiset tietueet, jotka ensin tallennetaan ARTO-tietokantaan ja kopioidaan sieltä testitietokantaan. Elektronisen aineiston luettelointiohjeet on laadittu projektissa, ja ELEKTRAN lisäksi niitä sovelletaan jo laajasti LINNEA-kirjastoissa. Tietueisiin lisätään mm. URL-osoitteet kenttään 856 hypertekstilinkkien luomiseksi jolloin viitetiedoista päästään lukemaan itse dokumenttia. WWW-käyttöliittymänä on myös Linnean yhteisluettelotietokannoissa käytössä oleva Hyperlib-ohjelmisto, joka on kehitetty Jyväskylän yliopistossa. ELEKTRAN tarpeita varten Hyperlibiin on kehitetty runsaasti lisäpiirteitä, mm. käytön seuranta.

Testitietokannan tietoturva on ratkaistu mahdollistamalla käyttö vain tarkoitusta varten varatuilla laitteilla IP-osoitteen avulla. Työasemilta on mahdollisuus aineiston lukemiseen ja tulostukseen oheiskirjoittimelle, mutta tiedostojen kopiointi on estetty. Työasemilla tarvitaan WWW-selainohjelma Netscape sekä Adoben Acrobat Reader -ohjelma PDF-muotoisten dokumenttien lukemista varten.

12.2 Koekäyttö

ELEKTRA-tietokannan koekäyttö on meneillään seitsemässä tieteellisessä kirjastossa, Helsingin, Jyväskylän ja Oulun yliopiston kirjastossa, Teknillisen korkeakoulun kirjastossa, VTT Tietopalvelussa ja Helsingin yliopiston Maatalous- ja Luonnontieteiden kirjastossa. Helsingin yliopiston luonnontieteiden kirjastossa palvelua ei ole sopimuksista huolimatta kuitenkaan käynnistetty lainkaan. Koekäyttöä varten ko. kirjastojen kanssa on solmittu kirjallinen sopimus ja kirjastoille on tuotettu opaskansio, joka sisältää kaikki ELEKTRA-tietokannan käyttöön liittyvät ohjeet sekä testikäytön yleiset periaatteet ja raportointiohjeet. ELEKTRAN WWW-käyttöliittymään on myös sisällytetty kaikki käyttäjän tarvitsemat opastetiedot, jotka käsittelevät tekijänoikeuksia, tiedonhakuja ja tulosteiden hinnoittelua. Kirjastoille ei ole kuitenkaan järjestetty varsinaista aiheeseen liittyvää koulutusta, mikä olisi todennäköisesti ollut tarpeen palvelujen paikallista tiedotusta varten.

ELEKTRA-tietokannan lukukäytöstä on kerätty seurantatietoja käyttöliittymään yhdistetyn lokiseurantaohjelman avulla, joka raportoi teoskohtaisesti artikkeleiden ja muiden dokumenttien lukemiskerrat. Lisäksi kokeilukirjastot raportoivat manuaalisesti kerätyt tiedot teoskohtaisista sivutulostemääristä. ELEKTRAN kokeilukäyttö aloitettiin kirjastoissa 15.6.1997 ja varsinaisesti se pääsi käyntiin vasta syyslukukauden alussa 1.9.1997. Kokeilukausi on näin muodostunut varsin lyhyeksi.

Koekäytön tulokset olivat varsin vaatimattomat. ELEKTRA-dokumentteja on luettu kokeilu-kirjastoissa yhteensä 200 kertaa. Eniten palvelua on käytetty Oulun yliopiston kirjastossa, yhteensä 80 kertaa ja seuraavana ovat HYK 47 kertaa ja Maatalouskirjasto 43 kertaa. Teknillisen korkeakoulun kirjastossa eikä VTT:llä palvelua ole käytetty juuri lainkaan ja Jyväskylän yliopiston kirjastossa vain 19 kertaa. Myöskään maksullista tulostuspalvelua ei ole käytetty kuin yhdessä kirjastossa.

Lehdittäin tarkasteltuna eniten käyttökertoja kohdistui jo muutenkin avoimessa verkkokäytössä olevaan Yliopistolehteen. Myös HYK:in tiedotuslehtiä on käytetty jonkin verran. Muuten käyttö hajaantuu eri lehtien kesken. Tulosten perusteella voi päätellä, että palvelun käyttäjät ovat olleet lähinnä kirjaston henkilökuntaa. Palvelusta tiedottamista kirjastot eivät ole kokeneet omaksi tehtäväkseen eivätkä ole sitä tehneet. Palvelun löytäminen on jäänyt satunnaisten aktiivisten asiakkaiden omatoimisuuden varaan.

12.3 Käyttäjäpalaute ja palvelun kehittäminen

Marraskuun alusta käynnistettiin palautetietojen keräys ELEKTRAn käyttäjiltä käyttöliittymän yhteydessä olevalla suorakäyttölmakkeella ja kirjastoissa jaettavalla paperilomakkeella. Palautelomakkeen avulla haluttiin asiakkaiden arvioita käyttöliittymän toimivuudesta, palvelun sisältöön liittyvistä asiakkaiden toiveista ja reaktioista maksullisuuteen. Käyttäjäpalautekyselyyn on saatu vastaukset 31.1.1998 mennessä 21 vastaajalta. Tämä on noin 10% ELEKTRA-palvelun käyttäjistä. Kirjastoja aktivoimalla olisi todennäköisesti päästy pienin ponnistuksin kattavampaan tulokseen. Käytettävissä oleva otos on kuitenkin niin pieni, että tulokset ovat vain viitteellisiä, mutta vastaukset antavat arvokasta lisävalaistusta lokiseurannan kylmille numeroille.

Vastaajista puolet oli opiskelijoita ja kolmannes kirjastojen työntekijöitä. Kustantajien edustajiin otettiin henkilökohtaisesti yhteyttä, ja heille annettiin testikäyttömahdollisuus omalta työpaikalta pariksi päiväksi ja tätä mahdollisuutta käytti hyväksi kolme vastaajaa. Vastaajista 7 tuli Maatalouskirjastosta, 5 TKK:n kirjastosta, 4 Oulun yliopiston kirjastosta ja 2 HYK:istä. ELEKTRAn liittyvää tiedotusta oli lähes kaikkien vastaajien mielestä ollut liian vähän ja vain puolet oli löytänyt ELEKTRAn käyttöön tarkoitetun työaseman helposti. ELEKTRAn käyttöliittymää ja opasteita sen sijaan pidettiin helppoina ja selkeinä.

Vähäistä tulostusta selitti kysymys, jossa 19 vastaajaa ilmoitti että ei ollut tulostanut, 7 koska siihen ei ollut tarvetta ja 2 koska se oli liian kallista. Käyttäjistä puolelle siis tuntui riittävän pelkkä lukeminen. Hintoja vajaa puolet vastaajista piti kohtuullisina, neljännes kalliina ja kolmannes ei osannut muodostaa mielipidettään. ELEKTRA-palvelun ihanteellisena ympäristönä puolet piti yliopistokirjastoja, kolmannes maksullista Internet-levitystä ja vain 2 vastaajaa yleisiä kirjastoja. Kehittämisehdotuksia tuli neljältä vastaajalta, ja niissä palvelua sinänsä pidettiin tervetulleena ja sen laajentamista vapaaseen verkkojakeluun toivottavana suuntana. Myös linkkejä suoraan lehdistä artikkeleihin toivottiin nykyistä selkeämmiksi.

Vaatimaton menestys kokeilukäytössä voidaan katsoa ELEKTRA-projektin ensimmäisen vaiheen suurimmaksi heikkoudeksi, vaikka resursseja on tietoisesti keskitetty julkaisupalvelun kehittämiseen ja aineiston digitointiin. ELEKTRAn kontrolloitu ja rajattu käyttömahdollisuus samoin kuin käyttökertoihin perustuva hinnoittelu ovat vastanneet kustantajien ja tekijöiden tämänhetkisiä tarpeita. Malli ei kuitenkaan mahdollista palvelun käyttöä riittävän laajasti. Jos sekä lukemis- että tulostusmahdollisuudet keskitetään vain kirjastoihin, joissa sama materiaali on myös painettuna ja asiakkaiden valokopioitavissa elektronista vaihtoehtoa halvemmalla hin-

nalla, syntyy palvelun kannalta epäedullinen kilpailutilanne. Sitä mukaa kun elektroninen julkaiseminen alkaa korvata painetun version, tilanne tietysti pikkuhiljaa muuttua.

Elektronisen aineiston käyttöä voidaan kuitenkin edistää ja palveluja kehittää käyttäjien tarpeita paremmin vastaaviksi monin tavoin. Palvelusta tiedottamalla ja käyttäjiä kouluttamalla saata-
neen nopeimmin toivottuja tuloksia. ELEKTRAN toisessa vaiheessa on jo suunniteltu keskityt-
tävän käyttäjäpalvelujen kehittämiseen. Sopimuskirjastoissa rajoitetun käytön mukaiseen hin-
noitteluun perustuvan kokeilun rinnalle tarvitaan laajempia jakelukanavia. Jatkohankkeessa on
tarkoitus kehittää sopimusmekanismeja organisaatiolisensseihin perustuvaan yliopistoissa ja
vaikkapa myös muissa oppilaitoksissa kuten ammattikorkeakouluissa ja lukioissa tapahtuvaa
avointa paikallisverkkokäyttöä varten. Kokeilu voisi tapahtua osana Kansallisen elektronisen
kirjaston kokeiluhanketta ja tarvittaessa pilotti voitaisiin rajata alueelliseksi kokeiluksi

Myös kansalaisen sähköisen identiteetin kehitystä seurataan ja ollaan valmiita toimimaan koti-
maisten elektronisten maksujärjestelmien testihankkeena. Palvelu on tarkoitus ottaa käyttöön
Suomessa asteittain kuluva vuodesta alkaen. Tuotantokäyttöön päästyään tämä palvelu tarjo-
aa erinomaisen perustan maksullisten verkkopalvelujen rakentamiselle. Merkittävänä kansalli-
sena hankkeena ELEKTRA voi laajentaa sähköisen identiteetin kokeilualuetta huomattavasti
tieteellisen tiedonvälityksen suuntaan.

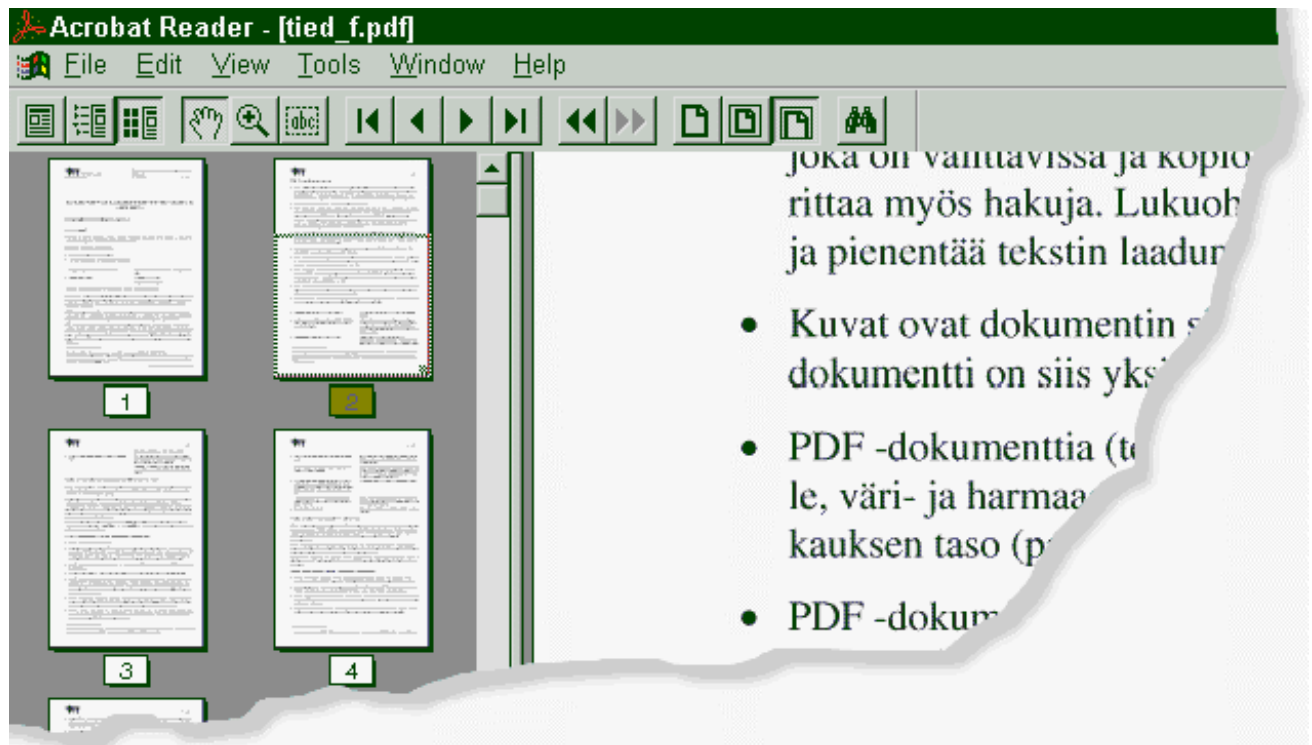
Liite 1: PDF-dokumentin pakkausmenetelmät

Tekstille/grafiikalle, väri-/harmaasävykuville ja mustavalkokuville on omat pakkausmenetelmänsä.

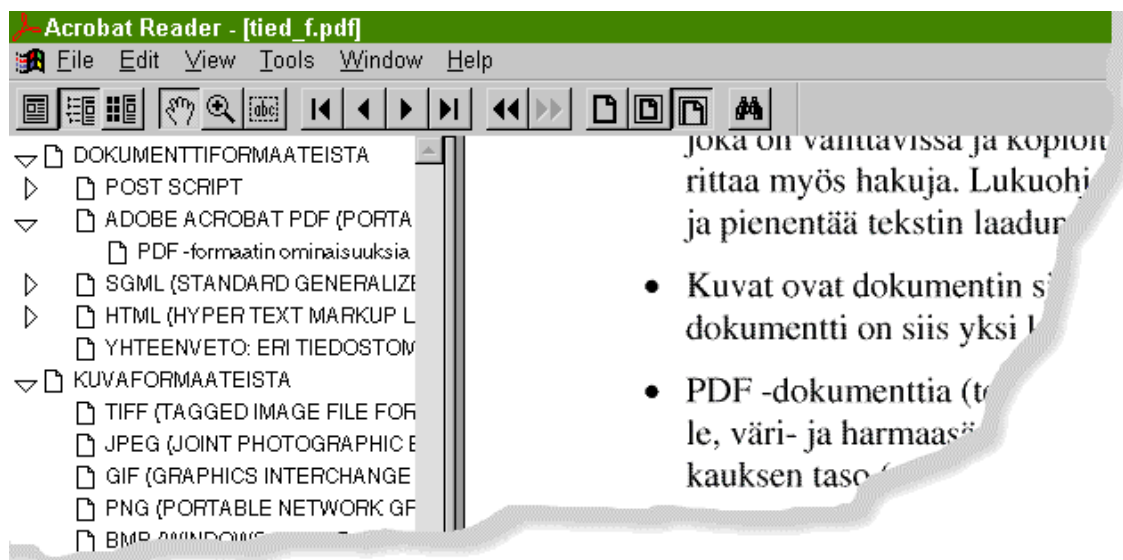
- Teksti / grafiikka: LZW-metodi (Lempel-Ziv-Welch). Tekstiä / grafiikkaa on kaikki muu paitsi bittikarttakuvat
- Väri- / harmaasävykuvat (bittikarttakuvat):
 - Kuvien tarkkuutta on mahdollista laskea PDF-tiedoston tuottamisvaiheessa (*downsampling*), jolloin kuvan koko pienenee. Samalla myös kuvan laatu huononee jonkin verran.
 - 16 tai 24-bittisille kuville käytetään JPEG-pakkausmenetelmää. JPEG-pakkausmenetelmä sopii parhaiten sävykuville, esimerkiksi skannatuille valokuville. JPEG ei sovellu hyvin sellaisten kuvien pakkaamiseen, joissa on laajoja yksivärisiä alueita tai viivapiirrosten pakkaamiseen. JPEG pakkaustasoja on valittavissa viisi; High - Medium High - Medium - Medium Low - Low.
 - LZW-pakkausmenetelmä sopii paremmin viivapiirrosten ja kuvien, joissa on laajoja yhtenäisiä värialueita, pakkaamiseen. Valokuvan pakkaaminen LZW-menetelmällä saattaa jopa kasvattaa kuvan kokoa.
- Mustavalkokuvat:
 - Tarkkuuden vähentäminen (*downsampling*)
 - CCIT Group 3, pakataan rivi kerrallaan (sama kuin useimmissa fakseissa)
 - CCIT Group 4, yleiskäyttöinen pakkausmenetelmä
 - LZW, paras kuville, joissa on toistuvia täyttökuvioita (*pattern*)
 - Run Length, paras kuville joissa on laajoja yhtenäisiä valkoisia tai mustia alueita

Pakkausasetukset asetetaan ennen PDF-tiedoston luomista, joko PDW Writerin tai Distillerin asetuksista

Liite 2: Miniatyyrikuvat ja kirjanmerkit Acrobat Readerissa



Kuva 1. Miniatyyrikuvat (thumbnails) Acrobat Readerin ruudulla



Kuva 2. Kirjanmerkit (bookmarks) Acrobat Readerin ruudulla

Liite 3: Kysely Tieteellisten seurain jäsenorganisaatioille/julkaisujen toimituskunnille

Kyselyn vastaukset palautetaan 20.2.1997 mennessä joko lomakkeella osoitteeseen Helsingin yliopiston kirjasto, ELEKTRA-projekti, PL 26, 00014 Helsingin yliopisto tai fax 09 - 708 44341.

1. Yhteystiedot

1.1 Julkaisija	
1.2 Yhteyshenkilö	
1.3 Osoite	
1.4 Puhelin	
1.5 Telefax	
1.6 Sähköposti	

2a. Julkaisu 1

2.1 Lehden/julkaisun nimi	
2.2 Ilmestymistiheys	
2.3 Painosmäärä/levikki kpl	
2.4 Levitys	
2.4.1 Jäsenjakeluna kpl	
2.4.2 Maksulliset tilaukset kpl	
2.4.3 Vaihdot kpl	

2b. Julkaisu 2

2.1 Lehden/julkaisun nimi	
2.2 Ilmestymistiheys	
2.3 Painosmäärä/levikki kpl	
2.4 Levitys	
2.4.1 jäsenjakeluna kpl	
2.4.2 maksulliset tilaukset kpl	
2.4.3 vaihdot kpl	

2c. Julkaisu 3

2.1 Lehden/julkaisun nimi	
2.2 Ilmestymistiheys	
2.3 Painosmäärä/levikki kpl	
2.4 Levitys	
2.4.1 Jäsenjakeluna kpl	
2.4.2 Maksulliset tilaukset kpl	
2.4.3 Vaihdot kpl	

3. Kokeiluun osallistuminen ja testiaineisto

- 3.1 Oletteko kiinnostunut toimittamaan testiaineistoa sovitussa muodossa ELECTRA-projektia varten kyllä ei
- 3.2 Haluatte neuvotella yksityiskohdista ja toivotte, että teihin otetaan yhteyttä kyllä ei
- 3.3 Haluatte osallistua kokeiluun muulla tavalla, miten

4. Kiinnostus yhteistyöhön ELEKTRA-projektin kanssa

Mistä ELEKTRAN tarjoamista palveluista olette kiinnostunut

4.1 Aineiston digitointi (skannaaminen painetuista julkaisuista, tekstinkäsittely- tai muiden tiedostojen käsittely)	kyllä	ei
4.2 Aineiston arkistointi (elektronisten dokumenttien arkistointi ja säilyttäminen HYK:in arkistopalvelimilla)	kyllä	ei
4.3 Aineiston asettaminen asiakkaiden verkkokäyttöön (mm. kuvailutietokanta-, hakuohjelma-, elektronisten dokumenttien lukeminen, käytön tuki)	kyllä	ei
4.4 Toivotte muita palveluja, mitä _____		

5. Julkaisun/lehden/lehtien toimittamisessa käytettävä tekniikka

Käytättekö toimitustyössä

5.1 Tekstinkäsittelyä (esim. Word, WP)	kyllä	ei
5.2 Sivuntaitto-ohjelmia (esim. Ventura, PageMaker)	kyllä	ei
5.3 Elektronisen julkaisemisen ohjelmia (esim. html Adobe acrobat)	kyllä	ei
5.4 Toimitatte aineiston erilaisissa muodoissa kirjapainolle, joka tekee myös taiton	kyllä	ei
5.5 Muita ohjelmia, mitä _____		
5.6 Sisältääkö julkaisunne kuvitusta	kyllä	ei
Jos kyllä, niin missä muodossa kuvat liitetään julkaisuunne		
5.6.1 Originaalikuvat käsitellään kirjapainossa	kyllä	ei
5.6.2 Kuvat liitetään tekstinkäsittely- tai sivuntaitto-ohjelmilla tuotettuihin tiedostoihin	kyllä	ei
5.6.3 Kuvat ovat erillisinä tiedostoina linkitettynä julkaisuun	kyllä	ei
5.6.4 Muuten, miten _____		

5.6.5 Mitä tiedostomuotoa käytätte kuvien käsittelyyn (esim. Tiff, Gif, JPEG)		

5.6.6 Kuvatkaa lyhyesti julkaisunne teknistä toimitusprosessia		

6. Julkaisunne/julkaisujenne elektronisen jakelun nykytila

Onko tällä hetkellä verkkojakelussa (Internet)

6.1 Lehden/julkaisun kotisivu kyllä ei

6.2 Sisällysluettelo kyllä ei

6.3 Artikkelitiivistelmät kyllä ei

6.4 Koko artikkelit valikoiden kyllä ei

6.5 Koko artikkelit kattavasti kyllä ei

Jos aineistoanne on verkkojakelussa, niin

6.6 Mistä lähtien (kuukausi, vuosi) _____

6.7 Mihin aineistonne on arkistoitu _____

Muut elektroniset julkaisunne

6.8 CD-ROM:it ja muut mahdolliset julkaisut _____

7. Muuta elektroniseen julkaisuun liittyvää kehitystoimintaanne, kommentteja tai muuta mielenkiintoista

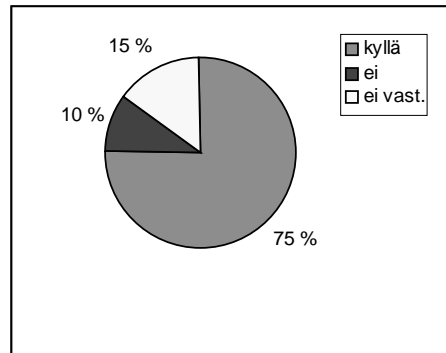
Kiitoksia osallistumisesta kyselyymme. Toivomme yhteistyömme jatkuvan. Kyselyn tuloksista kerromme ELEKTRA-projektin WWW-palvelimella, osoitteessa <http://linnea.helsinki.fi/elektra>.

Liite 4: Julkaisijakyselyn tulokset

3. Kokeiluun osallistuminen ja testiaineisto

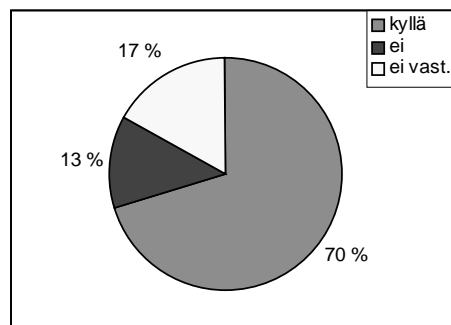
3.1 Oletteko kiinnostunut toimittamaan testiaineistoa sovitussa muodossa ELEKTRA-projektia varten?

	Yht.	%
kyllä	30	75
ei	4	10
ei vast.	6	15
	40	



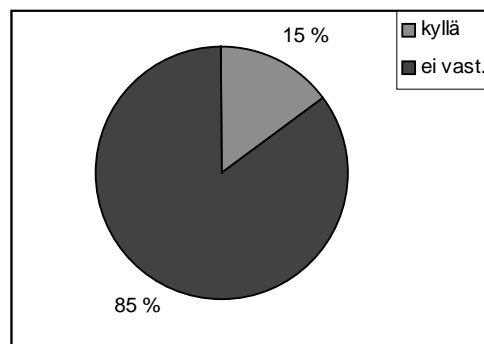
3.2 Haluatte neuvotella yksityiskohdista ja toivotte, että teihin otetaan yhteyttä?

	Yht.	%
kyllä	28	70
Ei	5	13
ei vast.	7	17
	40	



3.3 Haluatte osallistua kokeiluun muulla tavalla?

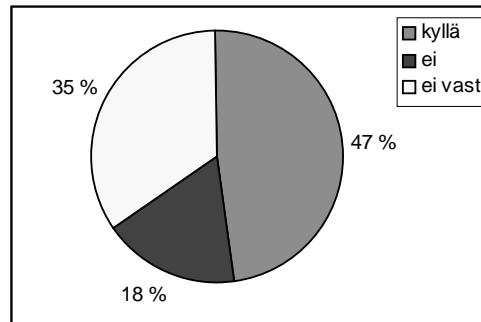
	Yht.	%
kyllä	6	15,0
ei vast.	34	85,0
	40	



**4. Kiinnostus yhteistyöhön ELEKTRA-projektin kanssa
Mistä ELEKTRAN tarjoamista palveluista olette kiinnostunut?**

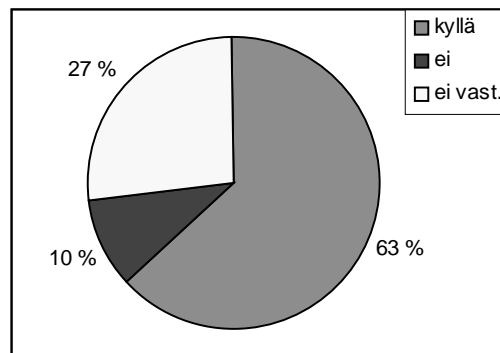
4.1 Aineiston digitointi (skannaaminen painetuista julkaisuista, tekstinkäsittely- tai muiden tiedostojen käsittely)

	Yht.	%
Kyllä	19	48
Ei	7	18
ei vast.	14	35
	40	



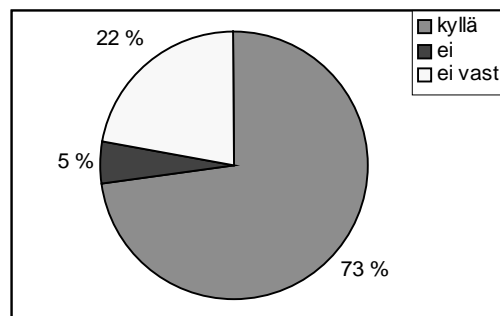
4.2 Aineiston arkistointi (elektronisten dokumenttien arkistointi ja säilyttäminen HYK:n arkistopalvelimilla)

	Yht.	%
Kyllä	25	63
Ei	4	10
ei vast.	11	27
	40	



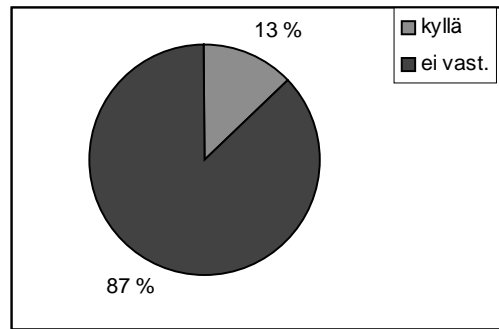
4.3 Aineiston asettaminen asiakkaiden verkkokäyttöön (mm. kuvailutietokanta-, hakuohjelma-, Elektronisten dokumenttien lukeminen, käytön tuki)

	Yht.	%
Kyllä	29	73
Ei	2	5
ei vast.	9	22
	40	



4.4 Toivotte muita palveluja

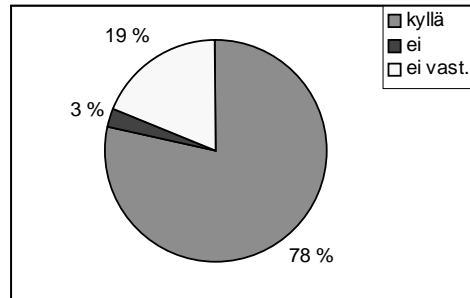
	Yht.	%
Kyllä	5	13
ei vast.	35	87
	40	



5. Julkaisun/lehden toimittamisessa käytettävä tekniikka? Käytättekö toimistustyössä

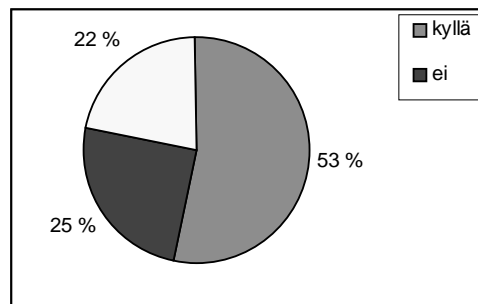
5.1 Tekstinkäsittelyä (esim. Word, WP)

	Yht.	%
Kyllä	32	70
Ei	1	3
ei vast.	7	17
	40	



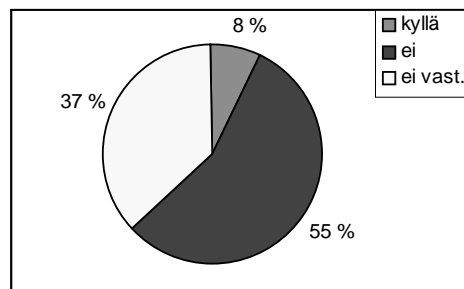
5.2 Sivuntaitto-ohjelmia (esim. Ventura, PageMaker)

	Yht.	%
Kyllä	21	53
Ei	10	25
ei vast.	9	22
	40	



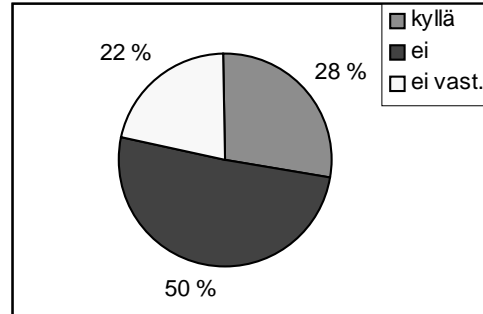
5.3 Elektronisen julkaisemisen ohjelmia (esim. html Adobe acrobat)

	Yht.	%
Kyllä	3	8
Ei	22	55
ei vast.	15	37
	40	



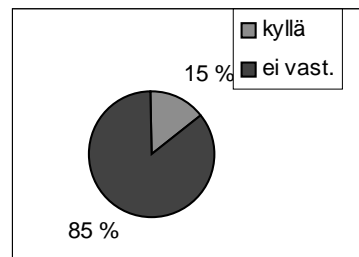
5.4 Toimitatte aineiston erilaisissa muodoissa kirjapainolle, joka tekee myös taiton

	Yht.	%
Kyllä	11	28
Ei	20	50
ei vast.	9	22
	40	



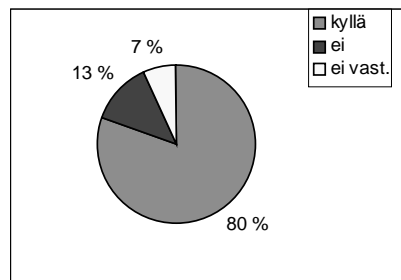
5.5 Käytättekö toimitustyössä muita ohjelmia?

	Yht.	%
Kyllä	6	15
ei vast.	34	85
	40	



5.6 Sisältääkö julkaisunne kuvitusta?

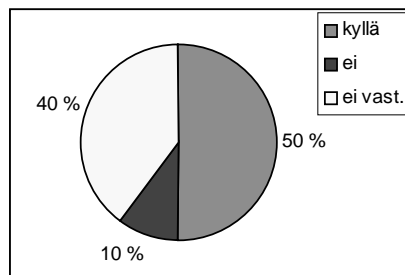
	Yht.	%
Kyllä	32	80
Ei	5	13
ei vast.	3	7
	40	



Jos kyllä, niin missä muodossa kuvat liitetään julkaisuunne?

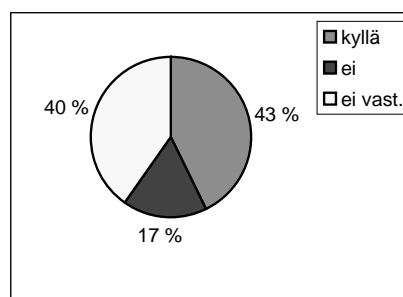
5.6.1 Originaalikuvat käsitellään kirjapainossa

	Yht.	%
Kyllä	20	50
Ei	4	10
ei vast.	16	40
	40	



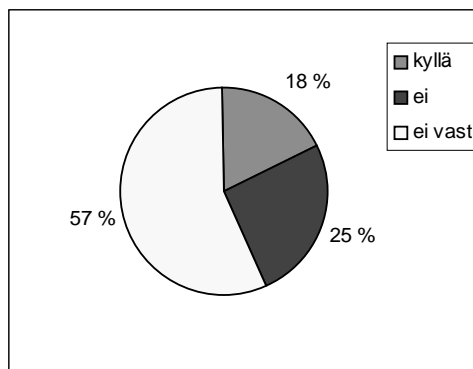
5.6.2 Kuvat liitetään tekstinkäsittely- tai sivuntaitto-ohjelmilla tuotettuihin tiedostoihin

	Yht.	%
Kyllä	17	43
Ei	7	17
ei vast.	16	40
	40	



5.6.3 Kuvat ovat erillisinä tiedostoina linkitettyinä julkaisuun

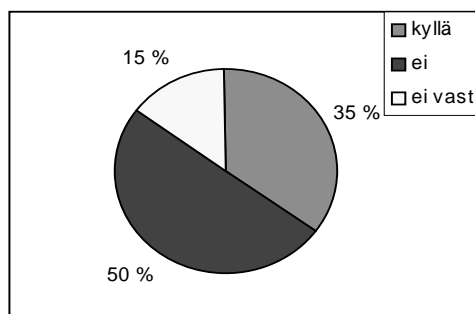
	Yht.	%
Kyllä	7	18
Ei	10	25
ei vast.	23	57
	40	



6. Julkaisunne/julkaisujenne elektronisen jakelun nykytila Onko tällä hetkellä verkkojakelussa (Internet)

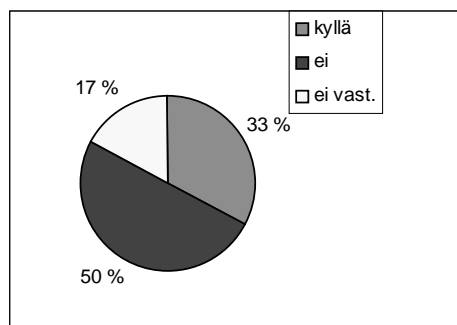
6.1 Lehden/julkaisun kotisivu

	Yht.	%
Kyllä	14	35
Ei	20	50
ei vast.	6	15
	40	



6.2 Sisällysluettelo

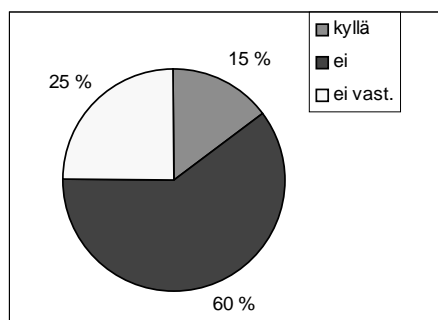
	Yht.	%
Kyllä	13	33
Ei	20	50
ei vast.	7	17
	40	



Onko tällä hetkellä verkkojakelussa (Internet)?

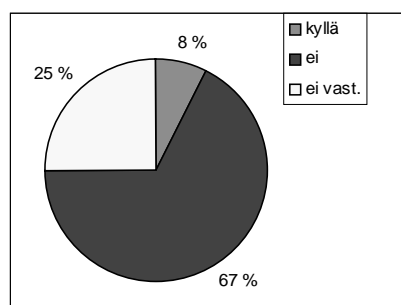
6.3 Artikkelitiivistelmät

	Yht.	%
Kyllä	6	15
Ei	24	60
ei vast.	10	25
	40	



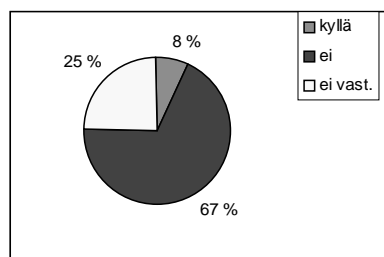
6.4 Koko artikkelit valikoiden

	Yht.	%
Kyllä	3	8
Ei	27	67
ei vast.	10	25
	40	



6.5 Koko artikkelit kattavasti

	Yht.	%
Kyllä	3	8
Ei	27	67
ei vast.	10	25
	40	

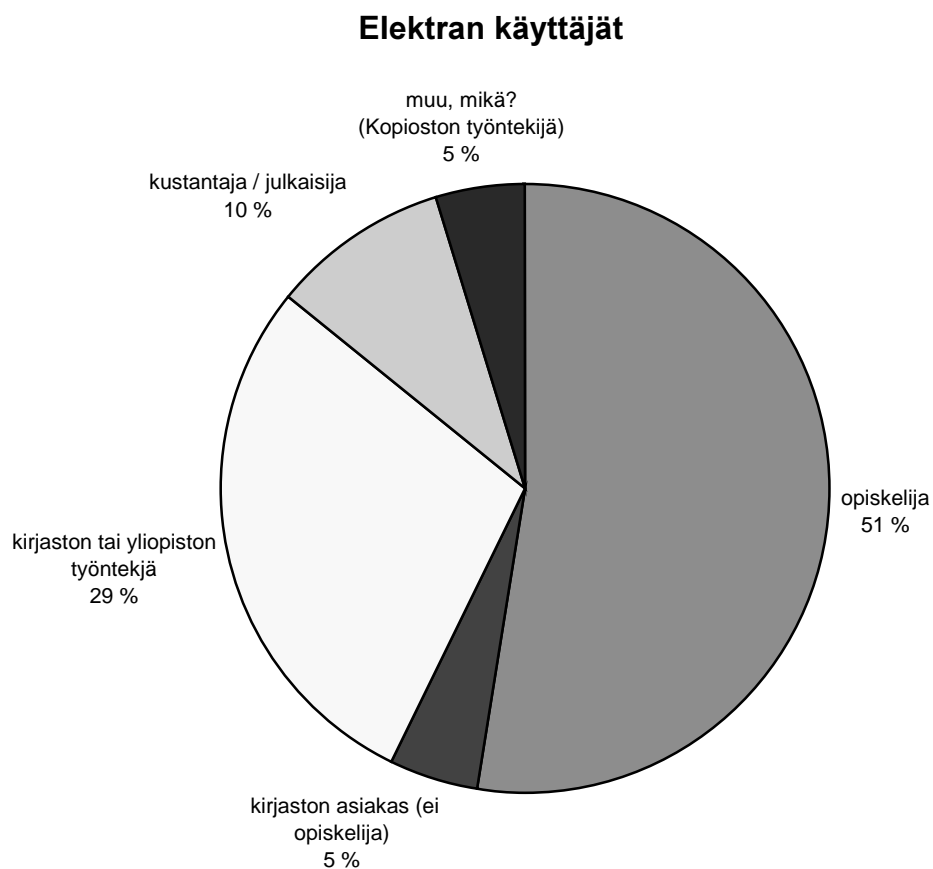


Liite 5: Elektran käyttäjäpalautteen vastaukset

(31.1.1998 mennessä vastanneet)

1. Taustatietoja

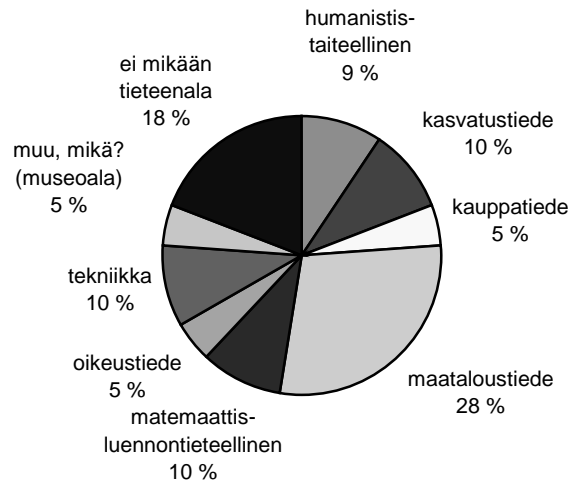
A. Käyttäjryhmät



Opiskelija yliopistossa tai korkeakoulussa	11
Kirjaston tai yliopiston työntekijä	6
Kokeiluun osallistuvan lehden kustantaja/julkaisija	2
Kirjaston asiakas (ei opiskelija)	1
Muu, mikä (Kopioston työntekijä)	1
Kokeiluun osallistuvan lehden artikkelin kirjoittaja	0
Yhteensä	21

B. Elektra-palvelun käyttäjän edustama tieteenala

Käyttäjän edustama tieteenala



Maataloustiede	6
Ei edusta mitään tieteenalaa	4
Humanistis-taiteellinen	2
Kasvatustiede	2
Matemaattis-luonnontieteellinen	2
Teknillinen tieteenala	2
Kauppatiede	1
Oikeustiede	1
Muu, mikä (Museoala)	1
Metsätiede	0
Psykologia	0
Yhteiskuntatieteet	0
Yhteensä	21

2. Palvelun käytön syy kyselyhetkellä

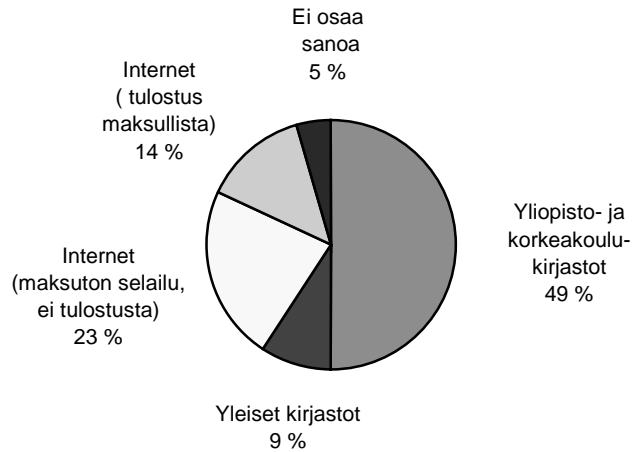
Kokeilu	13
Oman alan tiedon etsiminen	6
Muu syy (pyynnöstä 1, arvonta 1)	2
Yhteensä	21

3. Tiedotusta (esim. lehdissä ja kirjastoissa)

Liian vähän	10
Jonkin verran	8
Riittävästi	2
En osaa sanoa	1
Yhteensä	21

4. ELEKTRA-palvelun toimivin ja käyttökelpoisin ympäristö

Elektran ihanteellisin käyttöympäristö



Yliopisto- ja korkeakoulukirjastot	11
Internet (maksuton, ei tulostusmahdollisuutta)	5
Internetissä (tulostus maksullista, verkkorahastus)	3
Yleiset kirjastot	2
Ei osaa sanoa	1
Yhteensä (Yksi mainitsi kaksi ylintä vaihtoehtoa)	22

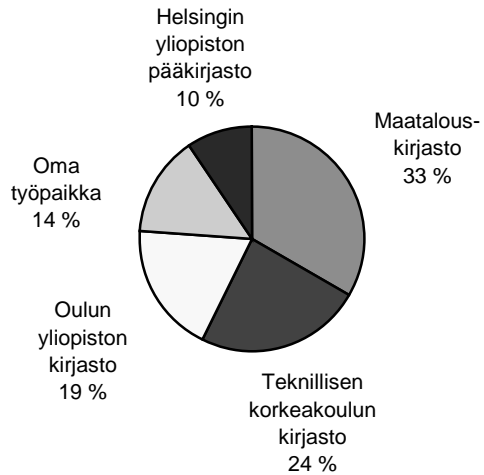
Lisäksi kommentteja 4 kpl:

1. Laajemman yleisön saatavilla olettaen, että mukaan tulee myös muita kuin tieteellisiä julkaisuja.
2. Minusta olisi hyvä jos materiaali olisi kaikkien selattavissa Internetissä. Tuorein materiaali jää ehkä pois, jos materiaalista ei peritä maksua. Maksun tulostus kustantajalle olisi tällöin sovitettava.
3. Jos Elektra-palvelu on Internetissä vapaasti selailtavana, tulostusmahdollisuus on.
4. Internetissä muodossa tai toisessa lehtien olisi oltava, koska muuten käyttö jää vähäiseksi. Esim. yliopistoissa lehtien käyttöön voisi olla lisenssi samaan tapaan kuin Academic Pressin lehdissä.

Arvio Elektra-palvelun laadusta

Elektra-palvelun kokeilupaikka

Kokeilukirjasto tai -paikka



Maatalouskirjasto	7
Teknillisen korkeakoulun kirjasto	5
Oulun yliopiston kirjasto	4
Oma työpaikka	3
Helsingin yliopiston pääkirjasto	2
Luonnontieteiden kirjasto	0
Jyväskylän yliopiston kirjasto	0
VTT Tietopalvelu	0
Yhteensä	21

5. Elektra-kokeiluun varatun päätteen löytyminen

Helppoa	10
Löytyi sattumalta	5
Ei vastausta	5
Vaikeaa	1
Yhteensä	21

6. Onko Elektran aineistosta kerrottu selkeästi?

Kyllä	17
Ei	3
Ei vastausta	1
Yhteensä	21

7. ELEKTRA-haun ohjeet ovat

Selkeät ja ymmärrettävät	20
En osaa sanoa	1
Huonot ja epäselvät	0
Yhteensä	21

8. Arvio ELEKTRAn käyttöliittymästä ja tietokannan käytettävyydestä

A. Käyttöliittymä

Selkeä ja helppokäyttöinen	17
Ei osaa sanoa	4
Sekava ja epäkäytännöllinen	0
Yhteensä	21

Lisäksi kommentteja 4 kpl:

- Ohjelma palaa Hyperlib-käynnistyksen kautta perushakulomakkeelle, kun vahingossa painaa selaimen Back-näppäintä. Tämä on hidasta. Kun käyttää selainta paljon, ei tahdo muistaa, että pitäisi käyttää ohjelman omia linkkejä eikä Back-näppäintä.*
- Käyttöliittymä on O.K. Omalta kohdaltani Acrobat Readerin lataus ja käyttö selailuun tuntuu hankalalta, kun en ole tottunut surffailija ja Internet-materiaalin käyttäjä.*
- Selainohjelman Back-nappia pitäisi pystyä käyttämään. Kun nyt palataan Hyperlibin Paluuvalinnalla takaisin hakuun, aiemmin käytetty hakuehto tuhoutuu ja sen joutuu kirjoittamaan (yleensä melkein samanlaisena) uudestaan! Sanahaussa ei voi hakea esim. ohjelmoi* sarjaport* eli tarkoituksena oli hakea tietoa sarjaportin ohjelmoinnista (sanahaku ei hyväksy kuin yhden sanan!)*
- Viimeistelyssä parantamisen varaa.*

B. ELEKTRA-haun käytettävyys

Helppo käyttää	17
Ei osaa sanoa	2
Vaikea käyttää	1
Ei vastausta	1
Yhteensä	21

Arvio tiedonhaun tulosten käytettävyydestä

9. Tiedonhaku

A. Käytettyjä hakuvaihtoehtoja

Tekijähaku	14
Nimekehaku	13
Asiasanahaku	12
Ei ole hakenut vielä mitään	1
Muu, mikä	0
Vastaajia yhteensä 21 , monet mainitsivat 2-3 vaihtoehtoa	

B. Muita tiedonhaun perusteita

Tietty aihe (artikkelit)	12
Tietty lehti (artikkelit)	7
Muuta, mitä (tekijät)	1
Ei ole hakenut vielä mitään	1
Väitöskirjoja ja opinnäytteitä	0
Yhteensä	21

C. Hakutuloksen saaminen ruudulle

Riittävän nopeaa	17
Ei kiinnittänyt huomiota asiaan	2
Liian hidasta	1
Ei vastausta	1
Yhteensä	21

D. Aineiston lataaminen ruudulle

Riittävän nopeaa	13
Liian hidasta	4
Ei kiinnittänyt huomiota asiaan	3
Ei vastausta	1
Yhteensä	21

10. Tulostus

A. Aineiston tulostus paperille

Ei tulostanut	19
Miksi, vastauksia	12
Ei ollut tarvetta	7
Liian kallista	2
Kopioi muualla paperiversiosta tai teki kynällä muistiinpanoja	2
Ei saanut PDF-artikkelia esiin	1
Kyllä (tulosti)	2
Miksi, vastauksia	2
Uteliaisuudesta	1
Aineisto vaikutti hyvältä	1
Yhteensä	21

B. ELEKTRA-palvelun sivukohtaiset tulostushinnat

Kohtuulliset	9
Ei osaa sanoa	7
Kalliit	5
Halvat	0
Yhteensä	21

11. Elektronisen aineiston luettavuus

A. Verkkokäyttöön sopiva tiedostomuoto

Ei osaa sanoa	12
HTML	4
Sekä HTML että PDF	4
PDF	1
Yhteensä	21

B. Hankaluudet aineiston käytössä

Palstoitus häyttasi lukemista	3
Tekstistä ei saanut selvää	2
Kuvista ei saanut selvää	1
Tekstissä oli kirjoitusvirheitä	0
Vastauksia 5, joista yksi mainitsi kaksi asiaa	

12. Aikooko käyttää ELEKTRA-tietokantaa uudelleen

Kyllä	14
--------------	-----------

Miksi, vastauksia 8 kpl

- Aion käyttää, mutta jos hakuni tuottavat liian vähän "hittejä", menee tyhjän hakemiseen liian paljon aikaa. Elektraan tallennettavia lehtiä pitäisi olla PALJON enemmän. Nykyisellään artikkeleita on liian vähän, jotta aineistoa löytyisi. Jos Elektrasta jatkossa peritään maksua, pitäisi sen kohdistua aineiston lataamiseen, ei sen tulostamiseen. Tämä siksi, että aineiston lataamisen valvonta on paljon helpompaa kuin sen tulostamisen. (Aineiston lataamisen jälkeen sen voi tallettaa ja tulostaa missä vain.) Kirjasto ei näytä paljoa panostavan koko kokeiluun, kun PDF-muotoisia artikkeleita ei edes pysty avaamaan. Tosin en tiedä, onko tämä ollut koko ajan, vai onko joku käyttäjä käynyt sotkemassa Netscapen asetukset. Hienoa, että olette saaneet tällaisen systeemin pystyyn. Toivottavasti jaksatte kehittää tätä edelleen.*
- Aion käyttää, jos asiakkaat tarvitsevat Elektran lehteä, jota ei ole Maatalouskirjastossa. Palvelusta pitäisi tiedottaa hirveästi ennen kuin ihmiset muistaisivat sen. Myöskään minä en muista, että asiakkaan etsimä lehti voisi olla Elektrassa. Koska löytyy vain yhdeltä päätteeltä, luulen, että asiakkaat eivät viitsi tulla kirjastoon asti käyttämään sitä. Elektrassa on myös varsin vähän Maatalouskirjaston asiakkaita kiinnostavia lehtiä, jotka eivät jo tule Maatalouskirjastoon. Jos lehti tulee kirjastoon, se on nopeampaa ja halvempaa hakea hyllystä. Sinänsä palvelu on ihan helppo käyttää, sen vain pitäisi olla henkilökunnan ja opiskelijoiden omilta työasemilta käytettävissä ja silloin pitäisi olla joku muu rahastuskäytäntö.*
- Nopea tapa saada artikkeleja omalle työpöydälle.*
- Helppokäyttöinen, mielenkiintoisia artikkeleita.*
- Gradu tekeillä.*
- Käytän jos tarvetta on. Hakukatetta voisi yrittää laajentaa vielä ns. epävirallisempiin julkaisuihin, ammattilehtiin jne. erikoislehtiin.*

Ei	3
----	---

Miksi ei, vastauksia 2 kpl

1. Yleinen historia puuttuu.
2. Saan asiakkaalta viitteet suoraan

Ei osaa sanoa	4
Yhteensä	21

Lisäksi vapaamuotoista palautetta lehtien kustantajilta ja kirjastohenkilökunnalta (yht. 4 kpl):

1. *Projektia tarkemmin tuntematta jäin kaipaamaan yksittäisiin lehtiin kohdistettavia hakuja, minkä tarve korostuu mitä enemmän aineistoa tietokannassa on. Sinänsä Acrobatilla toteutettu taltiointi toimii mielestäni varsin erinomaisesti, tosin en tiedä miten sivukohtaisen las-
kutuksen voi toteuttaa (ellei joudu maksamaan dokumentin avaamisesta, ei niinkään sen tu-
lostamisesta). Lisäksi dokumentit voi helposti samalla kopioida itselleen, mikä voi olla tämän
'järjestelmän hengen vastaista'. Epäselväksi jäi miten eri vuosia on mahdollista plärätä läpi
(lähinnä näytti siltä kuin kaikki lehden artikkelit olisi olleet listassa yhtä aikaa).*
 - *ohjeita luetaan aina tilanteen yhteydessä (ei siis ennen jotain tilannetta/tarvetta), joten oikea
paikka olisi kontekstin yhteydessä eli hakusivulla (tai sitten sen voisi lisätä myös itse hakuva-
likkoon omaksi vaihtoehdokseen, mikä olisi suurempi eli käyttäjän näkökulmasta parempi to-
teutustapa).*
2. *ELEKTRAn aineiston käyttämahdollisuus sähköisessä muodossa on eriomainen asia ja toi-
vottavasti laajenee edelleen.*
 - *mitä vapaammin aineisto on käytettävissä, sen parempi (esim. kirjastojen lisäksi avoimena
Internetissä)*
 - *satunnaisen käyttäjän on vaikea hahmottaa mitä on tarjolla, esim. en löytänyt yhteenvetoa
siitä, mitä artikkeleita ja väitöskirjoja on nähtävissä*
 - *lehtiluetteloon voisi olla selkeä oma linkki (hinnaston lisäksi)*
 - *hakuohjelmassa käytetyt termit eivät ole itsestäänselviä (esim. termi ”Osakohteisiin”)*
 - *PDF-tiedostot ovat varsin hitaita lukea, mutta ehkä hitaus koskee vain etäkäyttöä modeemin
kautta*
 - *lehden nimi, vuosikerta, vuosi ja sivut olisi mukava olla artikkelin alussa*
3. *lehtiluettelo selkeämmin eteen, ei hinnaston taakse*
 - *kun hakee lehden nimellä, on hankalaa, kun artikkelilistauksessa ei näe, missä numerossa ar-
tikkelit on julkaistu*
 - *haun lopetus on hiukan hankala löytää opastuksen tai yhteyden katkaisun kautta (yhteyden
katkaisuhan voi merkitä poistumista koko Elektrasta)*
 - *tulostuksen hinnoittelu on moninainen*
4. *palvelun pitäisi olla julkisesti käytettävissä Internetissä, ei vain kokeilukirjastoissa*