

STANDARDEISTA NÄKYMÄTTÖMÄÄN VERKKOON

XML, SFX,... and other abbreviations: A Metalink to the future
EUSIDIC Spring meeting 22.-23.3.2001 Lille, Ranska

Pekka Olsbo
Jyväskylän yliopiston kirjasto
Julkaisuyksikkö



EUSIDIC (European Association of Information Services) on vuonna 1970 perustettu riippumaton asiantuntijaorganisaatio, jonka päätehtävänä on luoda foorumeita ja kommunikaatiokanavia informaatioalalla työskenteleville. Mukana toiminnassa on yli 100 erilaista organisaatioita 16 Euroopan maasta. EUSIDIC tarkemmin <http://www.eusidic.org/>

EUSIDIC:n järjestämään seminaaritapahtumaan Lillen oli kokoontunut yli 90 XML:stä ja muusta rakenteiseen julkaisemiseen liittyvästä kiinnostunutta kuulijaa ja yhteisön jäsentä 15 maasta. Seminaarissa käsitellyt asiat voidaan jakaa kolmeen toisiaan täydentävään osioon: 1. *XML-kielen kehitys ja XML-pohjaiset tietokantaratkaisut*, 2. *taksonomiat ja knowledge management*, sekä 3. *invisible web – hakukoneet ja niille näkymättömät verkkovarannot*. Seminaariaineistot ovat saatavissa EUSIDIC:n verkkosivuilta.

Seminaarissa esitetyt ratkaisut ja mallit perustuivat kaikki XML (Extensible Markup Language) –kieleen ja sen sovelluksiin. XML:n 1.0-version spesifikaatio julkistettiin helmikuussa 1998. Seuranneen kolmen vuoden aikana XML:stä on kehittynyt jo suhteellisen tunnettu ja monissa sovelluksissa käytetty formaatti. XML:ää pidetään formaattina, joka auttaa ratkaisemaan elektroniseen julkaisemiseen liittyvät ongelmat. Se on standardoitu ja sen avulla voidaan erottaa toisistaan dokumentin sisältö ja ulkoasu. Näin mahdollistetaan tehokkaat hakuominaisuudet sekä vastataan pitkäaikaisvarastoinnin vaatimuksiin. Näille ominaisuuksille on ollut jo pitkään selkeä tarve erityisesti teollisuuden ja muiden teknisiä raportteja ja manuaaleja tuottavien alojen piirissä. Myös yliopistomaailma ja erityisesti kirjastot ovat panostaneet XML:n tutkimukseen ja sen ominaisuuksien hyödyntämiseen. Viime vuosina XML:n ja sen laajennusten standardointi on edennyt merkittävästi ja sen käyttö on levinnyt käytännössä kaikille elektronisen julkaisemisen aloille. XML:ää pidetään formaattina, joka syrjäyttää ajan myötä HTML:n WWW:n rakennuskielenä.

Extensible Markup Language ja sen kehitys

EUSIDIC:n seminaarissa XML-kielen kehityksestä ja merkityksestä esitelmöi *FIZ Karlsruhe* IT-johtaja Josef Mattes. XML:n käyttö on levinnyt kaikkialle verkossa tapahtuvaan toimintaan. XML-sovelluksia on rakennettu niin dokumenttien, tietokantojen, verkkojulkaisemisen, e-kaupan kuin käyttöjärjestelmienkin tarpeisiin. Alkuperäinen syy XML:n kehittämisen oli verkkojulkaisemisen tarpeet. Tarvittiin standardoitu formaatti, joka olisi täysin sovellus- ja käyttöjärjestelmäriippumaton sekä mahdollistaisi käytännössä minkä tahansa tarvittavan muunnoksen.

Puhuttaessa XML:stä, puhutaan kokonaisuudesta XML-kieliperheestä. XML on kehitetty laajemmasta [SGML-standardista](#), josta myös esimerkiksi HTML on kehitetty. XML-standardien kehittämisestä vastaa [World Wide Web Consortium](#) (W3C), jonka piirissä kehitetään edelleen kymmeniä XML-pohjaisia standardeja.

XML:ään liittyvät standardit voidaan karkeasti jakaa kolmeen ryhmään. Näiden ryhmien nimeämisestä tuntuu olevan maailmalla erilaisia mielipiteitä. Josef Mattes jakaa XML-perheen kolmeen ryhmään seuraavasti (Linkit vievät W3C:n määrittelysivuille):

1. *XML Grammar* sisältää tiedon XML:n syntaksista ja siihen liittyvä joukko standardeja kuten [XML Name Spaces](#), nimiavaruudet (mekanismi, jonka avulla varmistetaan XML-dokumenttien loogisen rakenteen yksikäsitteisyys), [XSL](#)

[style sheets](#), tyylitiedostot, [XML Schema](#), [XPath](#), [Xlink](#) jne. Näillä määrittelyillä esitetään rajoitukset ja säännöt siitä, kuinka XML rakennetaan.

2. *XML Protocols* , jotka vastaavat tiedon liikkumisesta systeemien ja sovellusten välillä. Tällaisia ovat esimerkiksi RosettaNet, OASIS-ebXML sekä SOAP.
3. *XML Vocabularies* määrittelevät XML:n käytön eri sovelluksissa tai sisältävät skeemassa tai DTD (Document Type Definition):ssä (katso esimerkiksi W3C <http://www.w3.org/XML/1998/06/xmlspec-report.htm> ja DocBook <http://docbook.org/>) määriteltyjä erikoismäärittelyjä esimerkiksi tieteenalakohtaisesti. Tällaisia alakohtaisia sovelluksia ovat esimerkiksi CML (kemia), [MathML](#) (matematiikka), [SMIL](#) (multimedia) ja RoboML (Human-Robot interface Using Agents Communicating in an XML-based Markup Language).

Vastaavasti Airi Salminen nimeää edellä mainitut XML-perheet:

(Katso tarkemmin Salminen, Airi: Summary of the XML Family of W3C Languages <http://www.cs.jyu.fi/~airi/xmlfamily.html>)

1. XML Accessories,
2. XML Transducers
3. XML Applications.

W3C:n kehitysryhmä määritteli kehitystyölle alun perin kymmenen tavoitetta:

1. Kielen on sovelluttava hyvin internet-ympäristöön.
2. Se on tuettava laajaa joukkoa erilaisia sovelluksia kattavaan informaation kuvauksen rakenteisista tietokannoista aina multimedian julkaisemiseen
3. Kielen on oltava SGML-yhteensopiva.
4. XML-dokumentteja prosessoivien ohjelmien kirjoittaminen tulee olla helppoa.
5. Valinnaisia ominaisuuksia yhtä toimintoa kohden tulee olla mahdollisimman vähän.
6. XML-dokumenttien tulee olla selkeitä ja ymmärrettäviä.
7. Kieli on kehitettävä nopeasti.
8. Kielestä on kehitettävä formaali ja tiivis.
9. XML-dokumenttien luomisen tulee olla helppoa.
10. Merkkauksen niukkuus ei ole itse tarkoitus.

Josef Mattesin mielestä XML:n kehitys on ollut tavoitteiden mukaista. Suuria takaiskuja kehittämistyössä ei ole tapahtunut.

Christian Schöter *Software AG*:sta esitti mielenkiintoisen näkemyksen tulevaisuuden tietoyhteiskunnasta, jossa personoidut käyttöliittymät personoituihin tietovarantoihin tuotiin koko ihmiskunnan käyttöön. Tämä herätti voimakasta keskustelua eriarvoisuudesta ja käytettävien resurssien puutteesta. Mikä olisikaan mukavampaa, kuin antaa jokaiselle tulevaisuuden kämmenmikro, jolla hoidettaisiin kaikki arkipäiväiset asiat. Todellisuus on ikävä kyllä toinen. Näiden tulevaisuuskuvioiden ulottumattomissa elää 90 % maailmasta.

Teollistuneissa maissa käyttäjänsä tuntevat mobiili-sovellukset yleistyvät kuitenkin ajan myötä. Tämän mahdollistaa osaltaan juuri XML, koska se on tällä hetkellä ainoa formaatti, jonka katsotaan olevan niin yleismaailmallinen ja standardoitu, että kaikki sovellukset pystyvät sitä käsittelemään. Tämä ihannekuva edellyttää luonnollisesti suuren luokan sitoutumista kaikilta laite- ja sovellusvalmistajilta.

Schöter pitää tätä kuitenkin täysin mahdollisena ja ainoana mahdollisuutena, jolla pystytään varmistamaan erilaisten sovellusten kommunikointi keskenään.



(kuva Dagmar Marek)

XML ja mobiililaitteiden kehitys, katso lisää: <http://www.softwareag.com/xml/>

SFX ja tietosisältöjen linkitys

Tiedonhakijan ongelmana on usein, ettei hän saa haluamaansa tietoa käytettävissä olevista tietovarannoista. Tietoa jakavilla organisaatioilla on omat tietokantansa, joista saa vain niiden sisältämää tietoa. On olemassa myös organisaatioita, jotka on perustettu linkittämään ja kokoamaan tätä eri lähteistä tulevaa tietoa. Yleensä näitä tietoa kerääviä ja jakavia organisaatioita nimitetään kirjastoiksi.

Myös kirjastojen kautta saatavalla tiedolla on rajoituksensa. Tietokantoihin tehtävä haku kohdistuu yleensä vain joko yhteen tietokantaan tai yhden palvelun tarjoajan tarjoamiin tietokantoihin. Esimerkiksi elektroniset lehdet muodostuvat usean palvelun tarjoajan tietokannoista. Muun muassa [ABI Inform](#) ja [EBSCO HOST](#) ovat tällaisia palvelun tarjoajia.

Tehtaessä haku esimerkiksi EBSCO HOST:n tarjoamiin elektronisiin lehtiin, haku kohdistuu vain kyseisen palvelun lehtiin. Mikäli sama haku halutaan kohdistaa myös muihin tietokantoihin, on hakijan siirryttävä toisten tietokantojen pariin. Mutta eikö olisi hienoa, jos EBSCO:n haku tuottaisi samalla tiedot myös muista käytettävissä olevista tietokannoista? Sama hakutoiminto kohdistuisi EBSCO HOST:n tuottamien tietokantojen lisäksi myös esimerkiksi [Cambridge Scientific Abstracts](#):n ja [SilverPlatter](#):n tuottamiin tietokantoihin. Halutessaan hakija voisi lisäksi valita haettavaksi vaikka pelkästään metatiedot.

SFX, *Content Sensitive Reference Linking* on rakennettu ratkaisemaan nämä ongelmat. SFX:n avulla instituution elektroniset kokoelmat, sekä lisensioidut että vapaasti jaettavat, voidaan määritellä ja jakaa loppukäyttäjille halutulla tavalla. Tällöin jokainen käyttäjä identifioidaan sen organisaation tai ympäristön mukaan, mistä hän tulee. Käyttäjän tunnistus sekä käytettävien tietovarantojen metadata-objektien määrittely tapahtuu OpenURL:n avulla. OpenURL on standardi protokolla, joka varmistaa yhteentoimivuuden informaatiolähteen ja paikallisen palvelun tuottajan välillä. Katso lisää esimerkiksi <http://www.sfxit.com/openurl>

SFX-järjestelmä on käytössä Gentin yliopistossa, jossa on yhteiskäytössä yli 60 julkaisijan elektroniset lehdet sekä yli 30 muuta tietokantaa. SFX:ää voi käydä kokeilemassa Ex Libriksen SFX-sivuilla osoitteessa <http://www.sfxit.com>

XML-standardiperheen puolella [Xlink](#), XML Linking Language, on kehitetty tavoittelemaan samantapaisia tarpeita kuin SFX. XLink määrittelee rakenteet, jotka mahdollistavat linkkien luomiseen ja kuvaamiseen tarvittavien elementtien lisäämisen XML-dokumenttiin. Kunnianhimoisena (utopistisena?) tavoitteena on koko maailman tietoresurssien yhdistäminen. Kuten Bohdana Stoklasová Tšekin tasavallan kansalliskirjastosta totesi, on tällä matkalla monia mutkia. Kuinka saada kaikki käyttämään olemassa olevia standardeja? Puhumattakaan taloudellisesta ja myös intellektuaalisesta resurssipulasta, joka koskettaa suurinta osaa maailmaa. Vanhojen rakenteiden ja asenteiden kaataminen on usein hyvin suuren työn takana.

Taksonomia

Olipa käytössä kuinka tehokkaita hakujärjestelmiä tahansa, ei tieto löydy itsestään. Tieto tai tieto tiedosta täytyy aina järjestää siten, että sen löytäminen on ylipäänsä mahdollista. *Knowledge/Information Management* ja tiedon *luokitus* tai *taksonomia*

ovat saaneet uuden merkityksen yhä paisuvan elektronisen aineiston järjestämisessä. EUSIDIC:n seminaarissa järjestystä tähän kaaokseen loivat Susan Batley, *University of North London* ja Alan Gilchrist, *TFPL Ltd*.

Taksonomian ja luokituksen tarkoituksena on luoda järjestystä kaaokseen. On kyse tiedon yksinkertaistamisesta, jonka avulla yritetään löytää subjekteja yhdistäviä kuvailuja. Näiden kuvailutietojen avulla rakennetaan "subjektipuita", jotka luovat organisoitua järjestystä. Ilman organisoitua järjestystä oikean tiedon löytäminen on mahdotonta.

Ryhdyttyessä rakenteistamaan tietoa, sillä tiedon rakenteistamisestahan taksonomioissa on kyse, on ensin selvitettävä, mihin tarpeeseen (Needs) ja tehtävään (Task) taksonomiolla pyritään vastaamaan. Samoin on analysoitava resurssit, tiedot ja taidot (Resource) sekä aineiston sisältö, arvo (Value) ja käyttö (Subject analysis). Samasta rakenteistamisen määrittelemisestä on kyse määriteltäessä DTD esimerkiksi XML-pohjaiseen julkaisemiseen siirtymistä varten. DTD eli dokumentin tyyppimäärittely sisältää listan dokumentissa esiintyvistä elementeistä, attribuuteista, merkintätavoista ja entiteeteistä sekä niiden suhteista toisiinsa. DTD:t määrittelevät joukon dokumentin rakennetta koskevia sääntöjä (Harold 2000, 214). Samoin voidaan taksonomioiden ajatella kuvaavan tiedon rakennetta koskevia sääntöjä.

Susan Batleyn mukaan keskeisintä taksonomioiden määrittelyssä on yksiselitteisyys. Jokaiselle subjektille määritellään vain yksi kategoria ja nimeämisissä käytettävistä termeistä tulisi kaikilla olla yhteinen ymmärrys, mikäli se on ylipäänsä mahdollista.



(kuva Dagmar Marek)

Alan Gilchrist puolestaan toi esille joitakin faktoja, minkä vuoksi taksonomiat ja tiedon luokittaminen ovat saaneet uutta merkitystä ylenpalttisen informaation valtaamassa maailmassa. Kaikille avoin www sisältää tällä hetkellä noin 2,5 miljardia dokumenttia. Joka päivä määrä kasvaa noin 7,3 miljoonalla dokumentilla. Kun mukaan lasketaan niin sanottu "deep web" tai "invisble web", eli erilaiset tietokannat, intranetit ja muut hakukoneiden ulottumattomissa olevat dokumentit, saadaan verkossa olevien dokumenttien kokonaismääräksi arvioilta yli 500 miljardia. Pelkästään Microsoftin intranetissä on yli 3 miljoonaa dokumenttia. Puhutaan siis käsittämättömistä luvuista. Internetin kattavin hakukone Google on indeksoinut noin 1,35 miljardia web-sivua. Tämä antaa meille aavistuksen siitä, kuinka paljon verkossa olevasta aineistosta on hakukoneiden ulottumattomissa. Kuitenkin 95% tuosta yli 500 miljardista on kaikille täysin avointa aineistoa.

Kyse ei ole pelkästään siitä, että aineisto verkossa on huonosti kuvailtua tai sellaisessa formaatissa, etteivät hakukoneet sitä nykyisellään löydä. Kyse on myös siitä, ettei hakukoneita osata käyttää oikein. Gilchrist esitti myös hakukäyttäytymistä osoittavaa tilastotietoa. Miljardin hakukoneilla tehdyn haun otos osoitti, että yli 72 % hakulausekkeista sisälsi korkeintaan kaksi hakutermiä. Yli 20 % hakulausekkeista ei sisältänyt ainuttakaan hakutermiä. Lisäksi 80 % tiedonhakijoista ei yhdistänyt käyttämiään termejä millään operaattorilla. Tosin on otettava huomioon, että esim. Google lisää automaattisesti termien väliin AND-operaattorin, ellei toisin haluta. Taksonomioista lisää katso esimerkiksi TFPL Ltd: <http://www.tfpl.com>

Invisible Web

Mitä on sitten se suurin osa verkon materiaalista, jota nykyiset hakukoneet ja agentit eivät löydä, vaikka tekisi kuinka älykkäitä hakuja tahansa? Tähän EUSIDIC:n seminaarissa toi valotusta Chris Sherman [Search Wise.Net](#):stä.

Sherman arvioi, että "todellinen" näkymätön web on arviolta 2-50 kertaa suurempi kuin "näkyvä" web. Se miksi nämä arviot näkymättömän webin suuruudesta vaihtelevat näin paljon, johtuu Shermanin mukaan laskentatavasta. Laskentatavoissa, joissa näkymättömän webin dokumentti- tai tietoyksikkömääräksi on saatu yli 500 miljardia, on otettu huomioon myös esimerkiksi satelliittien maahan lähettämä data, jonka yksikkömääräinen laskeminen lienee melko mahdotonta.



(kuva Dagmar Marek)

Näkymätön web (Invisible Web) koostuu siis kaikesta siitä aineistosta, mitä hakukoneet eivät joko löydä, tai mitä niitä ei ole ohjelmoitu löytämään. Usein tämä löytymätön aineisto on tietosisällöltään huomattavasti korkeampaa kuin näkyvässä webissä, koska esimerkiksi rekisteröitymistä vaativat tietokannat ovat tällaista aineistoa. Se miksi hakukoneet eivät löydä näitä tietokantoja johtuu yksinkertaisesti siitä, että hakukoneet eivät osaa täyttää kirjautumislomaketta. Verkossa arvioidaan olevan yli 250 000 tällaista saavuttamatonta tietokantaa.

Muuta löytymätöntä aineistoa ovat mm. tietyt tiedostoformaattit (PDF, Flash-sovellukset, Office-tiedostot ym.), suurin osa real-time datasta (pörssi-, sää- ja lentoliikennetiedot) sekä dynaamisesti generoituvat sivut (cgi, javascript, asp ja sivut joiden URL-tunnuksessa on "?").

Koska näkymätön web edustaa niin merkittävää osaa koko verkossa olevasta materiaalista, on kehitetty palveluja, jotka ovat erikoistuneet jakamaan tietoa näistä näkymättömistä resursseista. Tällaisia palveluja tarjoavat esimerkiksi *Intelliseek* (<http://www.invisibleweb.com> sekä <http://beta.profusion.com>), *Complete Planet* (<http://www.completeplanet.com/>) ja *Librarian's Index to the Internet* (<http://www.iii.org>).

Complete Planetin sivuilta löytyy myös yleistä tietoa näkymättömästä web:stä.

Chris Sherman luetteloi myös tieteenalakohtaisesti merkittäviä "näkymättömiä" resursseja. Tässä niistä koottu lyhennetty listaus.

Computer Science

MacAfee World Virus Map: <http://www.mcafee.com>

- nimi kertoo mistä on kyse.

ResearchIndex: <http://www.researchindex.com>

- tieteellisen kirjallisuuden digitaalinen kirjasto, joka indeksoi mm. Postscript- ja PDF-muotoon talletettuja artikkeleita

Dictionaries & Languages

EuroDicAutom: <http://eurodic.ip.lu>

- käännöskone 12 Euroopan kielelle. Toimii, mutta melko hitaasti. (koukku on tanskaksi krog)

Verbix: <http://www.verbix.com/index.html>

- yli 50 kielen verbit. Käännökset muun muassa *jiwarlin* kielestä.

Art & Artists

ADAM (Art, Design, Architecture & Media Information Gateway): <http://www.adam.ac.uk/>

- toimii virtuaalikirjaston tapaan, eli etsii linkkejä sopiviin aineistoihin verkossa.

Artcyclopedia: <http://www.artcyclopedia.com/>

- taiteilijoiden, taideteosten ja taidemuseoiden hakukone.

Liitteessä täydellinen lista Shermanin linkeistä.

Tulevaisuus

Suurin osa tällä hetkellä hakukoneille näkymättömästä materiaalista on jossain muussa kuin webin perusformaattissa HTML:ssä. Esimerkiksi hyvin yleinen dokumenttien verkkojakelumuoto PDF on useimpien hakukoneiden ulottumattomissa. Hakukoneista Google on ensimmäisenä alkanut indeksoida myös verkossa olevia PDF-dokumentteja. Voisi kuvitella, että muut palvelut seuraavat pian perässä. On myös useita muita formaatteja, joiden indeksointiin hakukoneet eivät tällä hetkellä pysty.

Hakupalveluja tuottavilla organisaatioilla on siis edessä suurten panostusten aika. Sherman ennustaa, että tämän vuoden aikana ainakin kaksi nykyistä suurta hakupalvelujen tuottajaa joutuu lopettamaan toimintansa resurssien loppumisen vuoksi. Samanaikaisesti markkinoille tulee aivan uuden tyyppisiä palvelintarjoajia. *Alykkäät agentit*, *"killer applications"*, *Inktomi*, *Index Connect*, *Ultraseek*, *WhizBang* ja *wrappers* ovat termejä ja sovelluksia, joihin tulemme varmasti törmäämään.

Vaikka hakujärjestelmät koko ajan kehittyvät, ei näkymättömän webin ongelma koskaan häviä. Yhtä nopeasti kuin hakukoneet kehittyvät, kehittyvät myös uudet formaatit ja aina vain monimutkaisemmat sovellukset. Ainoa keino vähentää näkymättömän aineiston osuutta on lisätä standardeihin perustuvaa verkkojulkaisemista. Tämä vaatii myös ohjelmistotalojen ja sovelluskehittäjien sitoutumista. Ja aina kun vaaditaan sitoutumista, on odotettavissa ongelmia.

Pelastajaformaattina pidetään yleisesti XML:ää sen standartoinnin ja skaalautuvuuden vuoksi. On kuitenkin vielä vaikeaa sanoa jakaako XML niin usean standardin kohtalon: aluksi kaikki ovat yhtä mieltä standardin erinomaisuudesta mutta sen siirtyminen käytännön toimintaan on hidasta ja jää vajavaiseksi. Lopulta kaikki unohtavat koko standardin olemassaolon. Paitsi standardin kehittäjät. XML:n puolustukseksi on sanottava, että sen kehittyminen viime vuosina on ollut erittäin nopeaa ja käytännön sovelluksia on jo käytössä melko runsaasti. XML:llä on siis mahdollisuutensa ja monien mielestä XML on myös ainoa mahdollisuus.

Lähteet

Batley, Susan. Access to Electronic Documents: Taxonomies in Information and Knowledge Management. Esitelmä. Lille 23.3.2001. Saatavana: <http://www.eusidic.org>

Daubach, Marc. SFX: an open linking framework for libraries. Esitelmä. Lille 22.3.2001. Saatavana: <http://www.eusidic.org>

DocBook.org: <http://www.docbook.org>

EUSIDIC, European Association of Information Services. <http://www.eusidic.org>

ExLibris: <http://www.exlibris-usa.com/>

FIZ Karlsruhe: <http://www.fiz-karlsruhe.de/>

Gilchrist, Alan. Corporate Taxonomies: Different Approaches to their Construction and Use. Esitelmä. Lille 23.3.2001. Saatavana: <http://www.eusidic.org>

Harold, Eliotte Rusty, 2000. XML-tehokäyttäjän opas. Satku – Kauppakamari. Jyväskylä 2000

Inktomi Corporation: <http://www.inktomi.com/>

Mattes, Josef. The swift success of XML. Esitelmä. Lille 22.3.2001. Saatavana: . <http://www.eusidic.org>

Salminen, Airi, 2001. Summary of the XML Family of W3C Languages. Saatavana: <http://www.cs.jyu.fi/~airi/xmlfamily.html> (viitattu 5.4.2001)

Schöter, Christian. The Return of Magic or how mobile-Commerce applications will revolutionize

life on the planet ... again. Esitelmä. Lille 22.3.2001. Saatavana: <http://www.eusidic.org>

SearchWise.net: [http:// www.SearchWise.net](http://www.SearchWise.net)

Sherman, Chris. The Invisible Web. Esitelmä. Lille 23.3.2001. Saatavana: <http://www.eusidic.org>

SoftWare AG: <http://www.softwareag.com>

TFPL Ltd: <http://www.tfpl.com/index.html>

WhizBang! Labs, Inc. <http://www.WhizBang.com/>

World Wide Web Consortium: <http://www.w3.org>

Liite

Chris Shermanin Invisible Web linkit

(Reproduced with kind permission by Mr Chris Sherman)

Gateways

- *Intelliseek*

- <http://www.invisibleweb.com>

- <http://beta.profusion.com>

Computer Science

- *MacAfee World Virus Map*

- <http://www.mcafee.com>

- *ResearchIndex*

- <http://www.researchindex.com>

Company Research

- *European High_Tech Industry Database*

- <http://www.tornadoinsider.com/radar/>

- *Kompass*

- <http://www.kompass.com>

Intellectual Property

- *Delphion Intellectual Property Network*

- <http://www.delphion.com/>

- *ESP@CENET (European Patent Office) Patent Database*

- <http://ep.espacenet.com/>

Dictionaries & Languages

- *EuroDicAutom*

- <http://eurodic.ip.lu>

- *Verbix*

- <http://www.verbix.com/index.html>

Arts & Artists

- *ADAM (Art, Design, Architecture & Media Information Gateway)*

- <http://adam.ac.uk/>

- *Artcyclopedia*

- <http://www.artcyclopedia.com/>

Real-Time Information

- *Flight Tracker*

- <http://www.trip.com/ft/home/0,2096,1-1,00.shtml>

- *J-Track 3-D Satellite Locator*

- <http://liftoff.msfc.nasa.gov/realtime/Jtrack/Spacecraft.html>

Maps and Driving Directions

- *MapBlast*

- <http://www.mapblast.com>

- *Streetmap.co.uk*

- <http://www.streetmap.co.uk/>

Government Info

- *Parline Database*

- <http://www.ipu.org>

- *United Nations Daily Press*

- <http://www.un.org/news/>

Health & Medicine

- *Economics of Tobacco Control Database*

- <http://www1.worldbank.org/tobacco/database.asp>

- *International Digest of Health Legislation*

- <http://www.who.int>

News & Current Events

- *Cold North Wind Newspaper Archiva Project*

- <http://www.coldnorthwind.com>

- *Financial Times Global Archive*

- <http://www.globalarchive.ft.com>

Science

- *Great Barrier Reef Online Image Catalogue*

- http://www.gbrmpa.gov.au/corp_site/info_services/library/index.html

- *Nuclear Explosions Database*

- <http://ausseis.gov.au/databases>

Transportation

- *Equasis (Merchant Ships)*

- <http://www.equasis.org/>

- *World Aircraft Accident Summary (WAAS) Fatal Airline Accident Subset*

- <http://www.waasinfo.net/>

Pekka Olsbo
Jyväskylän yliopiston kirjasto
Julkaisuyksikkö