hyväksymispäivä arvosana

arvostelija

# Hyperlinkkityypit

Ilpo Lyytinen

Helsinki 21. maaliskuuta 2002

Hypermediaseminaarin kirjallinen esitys

HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

# Sisältö

# 1 Johdanto

Hypermediaa voi hyvin kuvata verkkomallilla, jossa verkon solmut esittävät sisältöä - mahdollisesti erilaisissa formaateissa olevaa tekstiä, kuvaa, ääntä - ja verkon kaaret sisältösolmut toisiinsa liittäviä linkkejä. Tällaisen verkon solmujen välisistä suhteista on vaikea saada kovin selvää kuvaa, mikäli sen linkit edustavat ainoastaan yhteyttä yhdestä solmusta toiseen. Rakenteellisella tasolla ei voida tehdä eroa tärkeiden ja vähemmän tärkeiden sidosten välillä. Sisältösolmujen tutkiminen toki auttaa asiaa, mutta vaatii aikaa ja vaivaa. Toisaalta analysoimalla eri solmuihin tulevien ja näistä lähtevien linkkien määriä ja toisaalta solmujen välisiä linkityssuhteita, voidaan tehdä päätelmiä solmujen välisten kytkösten luonteesta. Tällaiset menetelmät tuottavat kuitenkin parhaimmillaankin vain valistuneen arvauksen.

Pelkkien sisältösolmujen välisien kytkösten sijaan hypermedialinkkejä voisi ajatella eri solmujen välisien suhteiden kuvauksina. Tässä tapauksessa linkkiin liitettävää ominaisuutta, linkin *tyyppiä*, voisi luontevasti käyttää kertomaan, millainen linkin yhdistämien solmujen välinen suhde on. Esimerkiksi hypertekstijärjestelmässä tällaisia tyypejä voisivat olla muun muassa "Sisällysluetttelo", "Seuraava luku", "Edellinen luku"ja "Lähdeviite". Tyypitettyjen linkkien avulla hyperavaruuden tarkastelusta saisi ainakin rakenteellisella tasolla paljon enemmän irti. Toisaalta hyvin määriteltyjen linkkityyppien huomioiminen intuitiivisella tavalla hypermediajärjestelmän käyttöliittymässä voisi auttaa järjestelmän käyttäjiä hahmottamaan paremmin ympäristöään. Tämä auttaisi osaltaan vähentämään turhien linkkien seuraamisen aiheuttamaa käyttäjien turhautumista sekä helpottamaan pahamaineista hyperavaruuteen eksymisen ongelmaa.

Hyvän linkkien tyypitysjärjestelmän määrittely ei kuitenkaan ole helppoa. Liian laaja tyyppijoukko saattaa olla tavalliselle käyttäjälle tai hypermedian tuottajalle liian monimutkainen hahmotettavaksi ja menettää näin merkityksensä. Liian suppea joukko puolestaan voi turhauttaa käyttäjän, joka ei löydä tarkoitukseensa sopivaa tyyppiä kuvaamaan tuotostensa suhteita. Entä pitäisikö käyttäjien sallia omien tyyppien määrittely? Mitään yleistä kaikkiin tilanteisiin sopivaa sääntöä ei varmaan ole mahdollista keksiä, vaan tyyppijaottelua pitää tarkastella aina tilanteen mukaan.

Tässä dokumentissa tarkastellaan joitakin ehdotettuja menetelmiä linkkien jakamiseksi eri tyyppeihin. Monet näistä menetelmistä on toteutettukin ainakin jollain asteella. Dokumentin sisältö on jaettu sen mukaan, missä ympäristössä linkkityyppejä on tarkasteltu. Luvussa 2 esitellyt menetelmät on suunniteltu tiettyjä ympäristöjä varten. Näistä kaksi ensimmäistä edeltävät WWW:n tuloa tavallisten käyttäjien tietoisuuteen. Luvussa 3 puolestaan tarkastellaan tyypitettyjen linkkien käyttöä WWW:ssä. Loppusanat ja yhteenveto käsitellyistä asioista löytyvät luvusta 4.

# 2 Joitakin linkkityyppien määrittelyjä

Tässä luvussa esitellään erilaisiin ympäristöihin kehiteltyjä tapoja jakaa hyperlinkkejä eri tyyppeihin linkkien ominaisuuksien perusteella. Eri järjestelmien käsitys siitä, mitä linkit ovat ja mihin niitä käytetään, poikkevat selvästi toisistaan. Eri järjestelmien linkkikäsityksiä voisi jo sinänsä pitää tyyppijakona, mutta tässä keskitytään lähinnä järjestelmien sisäisten jakojen tarkasteluun. Esimerkkijärjestelmiä tai niiden linkkijakoja ei ole tarkoitus käydä kokonaisuudessaan läpi, vaan niistä pyritään esittelemään kiinnostavimpia piirteitä.

dokumenttia.

## 2.1 Triggin linkkityypit tieteellisessä hypertekstissä

Yhden tarkimmista kehitetyistä linkkien tyypitysmenetelmistä esitteli Randall Trigg väitöskirjassaan vuonna 1983 [Tri83]. Triggin jaottelu on erittäin yksityiskohtainen, mutta rajoittuu tarkastelemaan tieteellisiä kirjoituksia sisältävää hypertekstijärjestelmää. Koko Triggin käyttämä tyyppijoukko on taulukossa 1.

Ylimmällä tasolla Trigg jakaa linkit kahteen ryhmään: *tavallisiin* ja *kommentti*linkkeihin[1]. Tavalliset linkit edustavat erilaisia kytkentöjä eri artikkelien välillä. Kommenttilinkit puolestaan yhdistävät artikkelin siihen kohdistuviin kommentteihin. Ideana on, että artikkelin lukijat (tai miksei tekijäkin) voivat kirjata mielipiteensä erilliseen kommenttiartikkeliin ja liittää sen sitten lähdeartikkeliin kirjoituksensa sävyä kuvaavalla kommenttilinkillä. Sekä tavalliset että etenkin kommenttilinkit jakautuvat edelleen aliryhmiin. Esimerkiksi artikkelin kysymyksenasettelua arvioiville kommenteille on oma linkkiryhmänsä.

Triggin linkkijoukko tarjoaa hyvin ilmaisuvoimaisen välineen hypertekstin suhteiden kuvaamiseen. Pelkästään artikkeliin liitettyjä kommenttilinkkejä tarkastelemalla voi saada hyvän kuvan siitä, miten muut ovat artikkeliin suhtautuneet. Varsinaisia kommentteja ei tarvitse edes lukea. Lähdeviitteiden jakaminen useammantyyppisiin voi taas helpottaa johonkin aiheeseen liittyvän materiaalin etsimistä.

Toisaalta Triggin yksityiskohtainen lähestymistapa aiheuttaa myös ongelmia. Laaja linkkityyppien valikoima pakottaa kirjoittajan miettimään tarkemmin luomiensa linkkien luonnetta. Sopivan tyypin löytyminen voi olla vaikeaa, koska yksittäisillä tyypeillä on melko rajallinen käyttöalue. Varsinkin kommenttien kodalla taas samaan kommenttiin voi sopia useampikin kommenttilinkkityyppi. Pitäisikö tällöin kommentti pilkkoa pienempiin osiin vai linkittää se kohdeartikkeliin usealla eri kommenttilinkillä? Kumpikaan vaihtoehto ei tunnu kovin houkuttelevalta kommentin kirjoittajan kannalta. Tällaiseen matalan tason jaotteluun on myös helppoa keksiä lisää tyyppejä, joita järjestelmään voisi

[1] Kommenttien sijasta voisi hyvin myös puhua *kritiikistä*, sillä Triggin kommenttityypit ovat enimmäkseen sävyltään negatiivisia.

Normal link types

| | | |
|---|---|---|
| Citation | Generalization/Specification | Summarization/Detail |
| C-source | Abstraction/Example | Alternate-view |
| C-pioneer | Formalization/Application | Rewrite |
| C-credit | | |
| C-leads | Argument | Simplification/Complication |
| C-epon | A-deduction | Explanation |
| | A-induction | |
| Background | A-analogy | Correction |
| Future | A-intuition | Update |
| | | |
| Refutation | Solution | Continuation |
| Support | | |
| | | |
| Methodology | | |
| Data | | |

Commentary link types

| | | |
|---|---|---|
| Comment | *Points* | *Data* |
| Critics | Pt-comment | D-comment |
| Supportive | Pt-trivial | D-inadequate |
| | Pt-unimportant | D-dubious |
| *Environment* | Pt-irrelevant | D-ignored |
| E-comment | Pt-redherring | D-irrelevant |
| E-misrepresent | Pt-contradict | D-inapplicable |
| E-vacuum | Pt-dubious | D-misinterpreted |
| E-ignored | Pt-counter | |
| E-Isupersede | Pt-inelegant | *Style* |
| E-Irefute | Pt-simplistic | S-comment |
| E-Isupport | Pt-arbitrary | S-boring |
| E-Irepeat | Pt-unmotivated | S-unimaginative |
| | | S-incoherent |
| *Problem Posing* | *Arguments* | S-arrogant |
| P-comment | A-comment | S-rambling |
| P-trivial | A-invalid | S-awkward |
| P-unimportant | A-insuff | |
| P-impossible | A-immaterial | |
| P-ill-posed | A-mislead | |
| P-solved | A-alternate | |
| P-ambitious | A-strawman | |

Taulukko 1: Triggin linkkityypit

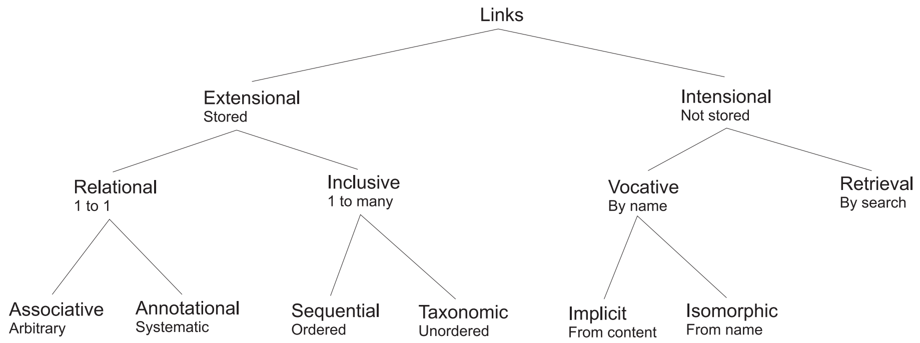
lisätä. Esimerkiksi kirjoitustyyliä voidaan varmasti kritisoida useammalla kuin kuudella määrittelyllä. Kommenttilinkkien eri aliryhmiin kuuluukin aina "yleinen ryhmään kuuluva kommentti-vaihtoehto, mutta tällainen erikoistuneiden ja yleisten linkkityyppien sekoittaminen voi turhauttaa kirjoittajia ja hämmentää lukijoita.

## 2.2 DeRosen hierarkkinen malli

Steven DeRosen 1989 esittämä jaottelu [DeR89] liikkuu Triggiä huomattavasti yleisemmällä tasolla. Siinä missä Trigg esittelee yli 70 hyvin rajattua suhdetta kuvaavaa linkkityyppiä, joilla on ehkä osin päällekkäisiäkin merkityksiä, on DeRosella vain 7 selvästi toisistaan poikeavaa ryhmää, joihin linkit jakautuvat. Hänen jaottelunsa ei perustu linkkien ilmaisemaan suhteeseen lähteen ja kohteen välillä, vaan lähinnä siihen, millä tavalla linkit käyttäytyvät. Esitetyn tyypitysjärjestelmän hyödyntäminen vaatisi erityisesti sitä tukevia käyttöliittymäratkaisuja.

DeRosen motivaationa on ollut erityisesti vanhojen tekstien tutkimisen tukeminen hypertekstijärjestelmässä. Esimerkkeinä hän mainitsee Raamatun ja muinaisten kreikkalaisten tekstien tutkimuksen. Tämän lähtökohdan takia esitetyssä järjestelmässä on erityisesti pyritty huomioimaan, että tutkittavista teksteistä on tyypillisesti olemassa useita eri versioita ja käännöksiä useille eri kielille. DeRosen käsitys siitä, mitä linkit ylipäätään ovat, poikkeaa osin huomattavasti siitä, mitä esimerkiksi WWW:n yhteydessä mielletään linkeiksi. Hänen hierarkkisen tyyppijaottelunsa (kuva 1) ylimmän tason kaksi ryhmää ovat laajennuslinkit (extensional links), jotka lähinnä vastaavat perinteistä linkkikäsitystä ja ovat kiinteä osa hypertekstiä, sekä tarkoituslinkit (intensional links), joita ei talleteta erikseen, vaan joiden kohde on jonkin funktion arvo. Ajatuksena on, että laajennuslinkit kuvaavat sellaisia hypertekstin solmujen välillä olevia yhteyksiä, joita on vaikea tai mahdoton nähdä ilman linkkejä.

DeRosella on jaottelussaan useita linkkityyppejä, joilla on useampi kuin yksi kohde. Laajennuslinkkien puolelta nämä on edelleen jaettu kahteen ryhmään: *järjestys-* (sequential) ja *ryhmittely*linkkeihin (taxonomic). Järjestyslinkkien tärkein käyttötarkoitus on do-




Kuva 1: DeRosen jaottelu linkkityyppeihin.

kumentin rakenteen kuvaus. Esimerkiksi jokin dokumentin luku voitaisiin kytkeä kaikkiin erillisissä dokumenteissa sijaitseviin alilukuihinsa yhdellä järjestyslinkillä. Järjestyslinkkien sovellutukseksi DeRose esittää, että niiden osoittamat dokumentit voisi saada käsiteltäväksi lineaarisessa muodossa. Erityisesti tämä voisi olla käyttökelpoista, mikäli dokumentti on jaettu muuten hyvin pieniin osiin. Tällaisessa tapauksessa dokumentin lukija joutuu muuten siirtymään usein dokumentin osien välillä, mikä haittaa varsinaiseen asiaan keskittymistä. Ryhmittelylinkit puolestaan yhdistävät dokumentin järjestämättömään joukkoon muita dokumentteja. Käyttäjän kannalta tämä tarkoittaa hakukoneen käyttöä vastaavaa operaatiota, jossa tulosjoukko on linkin tekijän määrittelemä.

Tarkoituslinkkeihin kuuluvat *hakulinkit* (retrieval link) vastaavat oikeastaan täysin nykyisin WWW:stä tuntemiemme hakukoneiden toimintaa. Linkin kohteet haetaan tutkimalla esimerkiksi niiden nimeä, sisältöä tai rakennetta. Sen sijaan toiset tarkoituslinkkien ryhmät, *implisiittiset* ja *isomorfiset* linkit, ovat ajatuksena selvästi oudompia. Implisiittisten linkkien ajatuksena on yhdistää lähdedokumentin sisältöä kohteisiin, joiden nimessä valittu kohta lähteestä esiintyy. Esimerkkinä tästä on mainittu sanakirja- ja karttahaut. Jokainen dokumentin sana voidaan liittää implisiittisellä linkillä kyseisen sanan selitykseen sanakirjassa, tai vaikka vastaavaan vieraan kielen termiin. Paikannimi liittyisi tämän lisäksi paikkaa esittävään karttaan. Mahdollisuuksia olisi valtavasti, mutta käytännössä tämänkaltainen implisiittisten linkkien käyttö voinee toimia vain rajatussa ympäristössä, jossa joitakin tiettyjä implisiittisiä yhteyksiä tuetaan. Toisaalta DeRose luokittelee myös esimerkiksi kirjallisuus- ja dokumentin sisäiset viittaukset implisiittisiksi linkeiksi. Toki nämä voivat esiintyä muissakin kirjoittajan määrittelemissä muodoissa, mutta jos tekstissä lukee vaikkapa "katso lukua 10", niin tästä pitäisi päästä automaattisesti siirtymään kyseiseen paikkaan ilman, että tekstin kirjoittaja on erikseen tehnyt siihen linkin. Isomorfiset linkit poikkeavat implisiittisistä siinä, että ne yhdistävät eri dokumenttien saman*nimisiä* elementtejä toisiinsa. Tätä voitaisiin käyttää esimerkiksi tutkittaessa rinnan jonkin tekstin eri versioita nämä tekstit sisältävien ikkunoiden synkronointiin. Isomorfisten linkkien toteuttaminen vaatisi tiukkaa nimeämiskäytäntöjen määrittelyä ja noudattamista, mutta teskstianalyysin lisäksi esimerkiksi kielten opiskelussa ne voisivat olla varsin hyödyllisiä.

## 2.3 Allan ja linkkien automaattinen generointi

James Allan on puolestaan pohtinut hypertekstilinkkien automaattista generoimista ja tyypittämistä [All96]. Aluksi hän jakaa linkit kolmeen pääryhmään sen mukaan, miten vaikeaa niiden havaitseminen automaattisesti on. Helpoimmin havittaviin *malliin sopiviin* (pattern-matching) linkkeihin hän määrittelee yksinkertaiset yhteydet, kuten sanan määrittelyn hakemisen sanakirjasta, sekä dokumentin loogista rakennetta kuvaavat linkit - esimerkiksi linkit alilukuihin tai kirjallisuusviitteissä mainittuihin dokumentteihin. Vaikeimmin automaattisesti havaittavia linkkejä Allan nimittää käsin tehtäviksi (manual) linkeiksi. Tähän joukkoon kuuluvat dokumentit, joiden välinen yhteys selviää vain luonnollisen kielen ymmärryksellä. Varsinaisen kiinnostuksen kohteena Allanilla ovat sellaiset yhteydet dokumenttien välillä, jotka voidaan tunnistaa automaattisesti dokumentteja analysoimalla ja joihin voidaan liittää yhteyttä kuvaava tyyppimääritys. Nämä automaattisesti tunnistettavat linkkityypit on lueteltu taulukossa 2.

Allanin menetelmässä linkkien löytämiseksi määritellään aluksi aihe tai lähdedokumentti ja haetaan sitten joukko tähän tavalla tai toisella liittyviä dokumentteja esimerkiksi jollakin hakukoneella. Tämän jälkeen selvitetään saadun dokumenttijoukon keskinäiset yh-

| Tyyppi | Vastakkainen tyyppi |
|---|---|
| Versio (Revision) | |
| Kooste (Summary) | Laajennus (Expansion) |
| Yhtäläisyys (Equivalence) | |
| Vastaavuus (Comparison) | Poikkeavuus(Contrast) |
| Tiiviimpi käsittely (Condensed treatment) | Laajempi käsittely (Expanded treatment) |
| Tangentti (Tangent) | |
| Ryhmittely (Aggregate) | |

Taulukko 2: Allanin automaattisesti löydettävät linkkityypit

teydet. Ne dokumentit, joilla ei ole kytkentöjä muualle kuin lähteeseen, luokitellaan lähteeseen *tangentiaalisesti*liittyviksi ja poistetaan jatkotarkasteluista. Esimerkiksi haettaessa pilviin liittyviä dokumentteja "pilvet-nimistä maalausta käsittelevä artikkeli voisi tulla joukkoon, mutta sillä ei luultavasti olisi paljoakaan yhteistä muiden dokumenttien kanssa.

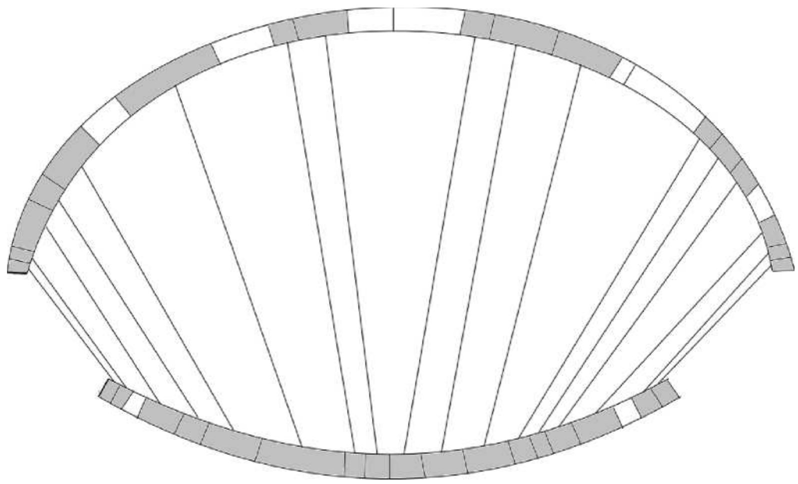
Jäljelle jääneet dokumentit paloitellaan osiin, ja niiden samankaltaisten osien välille muodostetaan metalinkkejä. Näille annetaan painoarvot sen mukaan, miten merkittäväksi niiden välinen yhteys katsotaan. Kahden dokumentin välisen linkin tyyppi selvitetään seuraavalla prosessilla:

1. Tutkitaan vain dokumenttien välillä olevia vahvoja metalinkkejä.

    - Mikäli metalinkit eivät juurikaan mene ristiin, eli dokumenttien toisiaan vastaavat osat ovat kummassakin dokumentissa samassa järjestyksessä ja kummassakaan dokumentissa ei ole runsaasti sellaisia osioita, joista ei lähde metalinkkiä, on dokumenttien välinen linkki tyypiltään *versiolinkki*
    - Jos toisessa dokumentissa on paljon enemmän sellaisia osioita, joista ei lähde metalinkkiä, on tämä toisen *laajennus*. Vastaavasti toiseen suuntaan dokumenttien välinen linkki on tyypiltään *kooste*
    - Jos kumpikaan näistä ei päde, katsotaan, ettei dokumenttien välillä ole selvää suhdetta, vaikka niissä onkin vahvasti samantapaista aiheen käsittelyä. Tällöin dokumentit yhdistetään *yhtäläisyys*linkeillä.

2. Jos dokumenttien välillä on vain vähän vahvoja metalinkkejä, toimitaan seuraavien askelten mukaan:

    - Mikäli metalinkit eivät juurikaan mene ristiin on yhteys dokumenttien välillä jälleen *yhtäläisyys*.
    - Mikäli kummankin dokumentin päädyissä on paljon tekstiä, josta ei lähde metalinkkejä, linkitetään dokumentit *poikkeavuus*linkillä.
    - Kuten kohdassa 1., tarkastellaan metalinkittämättömien osuuksien määrää dokumenteissa. Mikäli tässä on dokumenttien välillä huomattava ero, nimetään harvemmin metalinkitettyyn dokumenttiin viittaava linkki *Laajemmaksi käsittelyksi* ja tiheämmin metalinkitettyyn *Tiiviimmäksi käsittelyksi*[2].
    - Jos mikään edellä esitetyistä ehdoista ei toteudu, laitetaan dokumenttien välillä olevan linkin tyypiksi *Vastaavuus* Tämä siis lähinnä tarkoittaa, että dokumenttien välillä on jonkinlainen relevantti yhteys.



Kuva 2: Kaksi dokumenttia, jotka on jaettu osiin ja metalinkitetty keskenään.

[2]Tämä on ainakin minun tulkintani. Lähteessä tyyppejä "Tiiviimpi käsittely"ja "Laajempi käsittely"ei ole mainittu tyyppien merkityksiä selitettäessä. Tässä on oletettu, että kyseessä on artikkelin kirjoittajan huoimattomuus, mutta voi olla myös, että kyseiset tyypit on tarkoitettu "Koosteen"ja "Laajennuksen"synonyymeiksi.

Esimerkkinä kuvassa 2 on esitetty kaksi dokumenttia, joiden eri osien väliset metalinkit on piirretty näkyviin. Aluksi voidaan todeta, että metalinkit eivät mene ollenkaan ristiin. Dokumentissa A näyttäisi kuitenkin olevan selvästi enemmän metalinkittämättömiä osioita kuin dokumentissa B, joten A on B:n laajennus ja B puolestaan A:n kooste. (Olettaen, että linkit ovat enimmäkseen vahvoja likkejä).

Tätä menettelyä tarkasteltaessa voisi odottaa kohdassa 1. löydettyjen vahvoihin metalinkkeihin perustuvien tyypitysten tuottavan kohtalaisen hyviäkin tuloksia. Sen sijaan, kun joudutaan turvautumaan heikkojen metalinkkien arviointiin, tunnutaan kyllä olevan heikommalla pohjalla - varsinkin poikkeavuuslinkkien kanssa. Samalla voidaan todeta tulosten riippuvan siitä, miten metalinkkien vahvuus määritetään.

Allanin tyyppijaottelu on melko yksinkertainen. Esimerkiksi Triggin tyyppijakoon verrattuna sen ilmaisuvoima ei ole kovin suuri. Toisaalta hienojakoisemmassa mallissa automaattinen linkkien tunnistus olisi vaikeampaa ja virhetulkintojen todennäköisyys kasvaisi. Allanin linkkityypit ovat myös helposti käyttäjän yksiselitteisesti ymmärrettävissä, kun taas monimutkaisemmissa järjestelmissä käyttäjällä saattaa olla vaikeuksia erottaa eri tyyppien merkityksiä toisistaan.

# 3 Linkkityypit WWW:ssä

Internetissä hyperteksteihin yleisemminkin liitetty eksymisen ongelma on erityisen vaikea valtavan dokumenttien ja niiden erilaisten muotojen määrän takia. Ongelma on kyllä tiedossa, mutta siihen ehdotetut ratkaisut ovat olleet ohjeita selkeiden käyttöliittymien, sivujen ja sivustojen suunnittelemiseksi. Tämä on toki hyödyllistä, mutta mahdollisuuteen, että linkkeihin itseensä liitettäisiin navigoimista helpottavaa tietoa, on kiinnitetty melko vähän huomiota.

Erilaiset linkit näyttävätkin internetissä käyttäjälle usein täysin identtisiltä. Ellei dokumenttiin ole lisätty linkin oheen tietoa sen kohteesta, on käyttäjän mahdotonta päältä päin sanoa, millaiseen tietoon linkki viittaa ja missä suhteessa linin kohde on lähdedokumenttiin. Weinrichin ja Lamersdorfin paperissa [WL00] viitataankin Hampurin yliopistossa tehtyyn tutkimukseen, jossa todettiin käyttäjien usein turvautuvan linkin kohteen URL:n tulkitsemiseen ennen kuin päättävät, seuratako linkkiä vai ei.

Käyttäjän selaaminen helpottuisi selvästi, jos selainohjelmat pystyisivät tekemään tätä tulkintatyötä käyttäjän puolesta ja esittämään selvitetyn tiedon käyttäjälle intuitiivisesti käyttöliittymään liitettynä. Kohdedokumentin muoto selviää usein helposti vain tarkastelemalla kohdetiedoston nimen pääteosaa. Lähdedokumentin ja kohteen välisestä suhteestakin voidaan päätellä ainakin jotakin analysoimalla URL:n polkua kokonaisuudessaan. Tällainen automaattinen arviointi voisi kuitenkin tuottaa vääriä johtopäätöksiä, jotka voisivat viedä käyttäjien luottamuksen linkeistä saamaansa informaatioon. Parempi olisi, jos linkkien tekijät liittäisivät linkkeihin niiden tyyppiä kuvaavan informaation.

## 3.1 Linkkityypit HTML:ssä

HTML-määrittelyn versiossa 4.01 [(W302] linkeille on kuin onkin määritelty tyyppiominaisuuksia, vaikka tätä harvemmin verkossa liikkuessa huomaa. Jo määrittelyltään poikkeavia tyyppejä on kaksi: jokaiselle HTML:n kirjoittajalle tuttu *A*-elementti sekä hieman oudompi *LINK*. A-linkit on tarkoitettu tukemaan käyttäjän liikkumista verkossa. Niihin liittyy jokseenkin aina ankkuri, esimerkiksi pätkä tekstiä tai kuva, jonka valitsemalla käyttäjä pyytää linkin kohdetta käyttöönsä. LINK-elementti puolestaan on tarkoitettu kuvaamaan dokumenttien välisiä suhteita. LINK-elementtejä voi esiintyä vain HTML-dokumentin HEAD-osiossa, eikä siihen voi liittää ankkuria. LINK-elementteihin voisi laittaa viiteitä, jotka liittyvät koko dokumenttiin sen jonkin tietyn osan sijasta. HTML-standardi mainitsee LINK-elementtien käyttötapaukseksi informaation tarjoamisen hakukoneille. Esimerkiksi printteriystävälliseen ps-versioon tai toisenkieliseen versioon voisi tarjota tällaisen dokumenttitason viitteen. Selaimetkin voisivat kyllä hyödyntää LINK-määrittelyjä tarjoamalla käyttäjälle niihin sisältyvää informaatiota muodossa tai toisessa, mutta tietääkseni yksikään selain ei näin tee.

Sekä A- että LINK-elementteihin voi liittää samoja niiden ominaisuuksia tarkemmin kuvaavia määreitä. Näistä *type*-määrittelyä käytetään ilmoittamaan kohdedokumentin tyyppi. Suositeltavaa olisi käyttää MIME määrittelyn mukaisia tyyppejä, joista esimerkkejä ovat "text/html", "text/css", "image/png", "video/mpeg", ja "audio/basic"[3]. Nämä tyypit voisi usein helposti päätellä suoraan linkin kohteen nimen päätteestä, mutta erillinen tyypin ilmoittaminen säästää tietysti tulkinnan vaivalta ja ongelmilta oltaessa tekemisissä poikkeavien nimeämiskäytäntöjen kanssa.

Likitettyjen dokumenttien välisestä suhteesta puolestaan voidaan kertoa linkkielementin *rel-* ja *rev*-määreillä. Näiden erona on se, että rel-määre ilmaisee suhdetta lähdedokumentista kohteeseen ja rev puolestaan kohteesta takaisin lähteeseen. Esimerkiksi dokumentista A voisi olla dokumenttiin B linkki, jonka rel-määreen arvo olisi "next"ja rev-määreen arvo "prev"(mikä tosin tuntuisi hieman ylityypittämiseltä). HTML-määrittelyssä esitetyt linkkityypit on esitetty taulukossa 3. Kuten voidaan huomata, rajoittuvat nämä lähes täysin yhden useampaan osaan jaetun dokumentin osien välisten suhteiden kuvaamiseen. Tällainen jako on usein helppo tehdä, mutta valittu tyyppijoukko tuntuu melko rajoittuneelta. Varsinkin verkon selaajalle esitetyistä tyypeistä on varsin vähän hyötyä, sillä vastaava informaatio on sangen helppo toteuttaa osaksi hypertekstidokumenttia ja tällaisia käytetään jo nyt yleisesti ympäri Internetiä. HTML-määrittelyssä sanotaan myös, että yhteen rev- tai rel-määrittelyyn voi liittyä useampia tyyppejä, mutta tämänhetkinen tyyppijoukko sisältää tyypillisesti toisensa pois sulkevia tyyppejä.

Linkkien toivoisi tulevaisuudessa tarjoavan verkossa liikkujalle enemmän informaatiota kohteestaan. Oikeastaan on aika merkillistä, etteivät yleisimmät selaimet tarjoa visuaalista informaatiota edes linkin kohteen tyypistä, tämä kun selviäisi vähällä vaivalla kohteen nimen päätettä tarkastelemalla. Lähteen ja kohteen välisten suhteiden tulkitseminen pelkän URL:n pohjalta on puolestaan selvästi hankalampaa ja näiden hyödyntäminen vaatisi sitä, että hypertekstin tekijät todella liittäisivät linkkeihinsä myös tyyppimäärittelyt.

Jotta linkkityppejä hyödyntävästä järjestelmästä saataisiin mahdollisimman paljon irti,

[3] Kattava lista MIME-tyypeistä löytyy osoiteesta *ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/media-types*

| Tyyppi |
| --- |
| Alternate |
| Stylesheet |
| Start |
| Next |
| Prev |
| Contents |
| Index |
| Glossary |
| Copyright |
| Chapter |
| Section |
| Subsection |
| Appendix |
| Help |
| Bookmark |

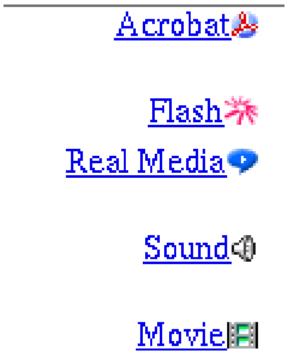
Taulukko 3: HTML4.01:n määrittelemät linkkityypit.

pitäisi löytää sopiva tyyppeihin jako. Tämän pitäisi samalla kertaa kattaa kaikki yleiset tyypit ja olla riittävän selkeä, ettei se vieroittaisi hypertekstin luojia tekemällä sopivan tyypin valinnasta rasittavaa valintatilannetta. Hyvä standardijaottelu lienee ehtona sille, että sovellukset, lähinnä selaimet, voisivat alkaa tukea linkkityyppejä käyttöliittymissään.

## 3.2 QBullets

Erilaisia ideoita linkkien ilmaisuvoiman parantamiseksi on kyllä kehitelty varsinkin käyttöliittymätasolla. Microsoftin Internet Explorerin uusien versioiden tapa näyttää linkkiin liitettävä informatiivisen title-attribuutin sisältö linkin yhteydessä, kun linkin ankkuria osoitetaan hiirellä. Tästä saatava hyöty on kuitenkin täysin title-attribuutin tekijän varassa. Samaa ideaa ovat kehitelleet laajemmin Weinreich ja Lamersdorf [WL00]. Heidän HyperScout-järjestelmässään linkkiä osoitettaessa ilmestyvässä tietoikkunassa voidaan kertoa title-attribuutin lisäksi muun muassa kohteen koko ja tyyppi, koska kohteessa on viimeksi vierailtu tai onko palvelin, jolla kohde sijaitsee toiminnassa.

Edellä mainittujen menetelmien ongelmana on, että käyttäjä joutuu erikseen osoittelemaan mahdollisesti kiinnostavia linkkejä saadakseen niistä lisätietoa. QBullets [Med02] tarjoaa ratkaisua, jossa käyttäjälle tarjotaan informaatiota linkeistä suoraan niiden yhteydessä. Idea on hyvin yksinkertainen. Erilaisia linkin ominaisuuksia kuvaamaan on kehitetty joukko pieniä kuvia, jotka voi liittää dokumenttiin heti linkin perään. Näistä muutama on esitetty kuvassa 3 Määritellyt tyypit selityksineen on esitelty taulukossa 4.




Kuva 3: Esimerkki QBullet-kuvista.

QBulletsin tarjoamat tyypit tuntuvat enimmäkseen varsin järkeenkäyviltä. Tosin ne kuvaavat lähinnä linkin kohteen ominaisuuksia, eivät niinkään lähteen ja kohteen välistä suhdetta. Esimerkiksi HTML-spesifikaatiossa annetuille linkkityypeille ei löydy kuvaajia QBulletien joukosta.

# Viitteet

All96 James Allan. Automatic hypertext link typing. In *Proceedings of the Seventh ACM Conference on Hypertext*, Autonomous Hypertext Systems and Link Discovery, pages 42–52, 1996.

DeR89 Steven J. DeRose. Expanding the notion of links. In *UK Conference on Hypertext*, pages 249–257, 1989.

Med02 Matterform Media. Qbullets, 2002.

| Navigointivihjeet | |
|---|---|
| OutLink | Uudelle sivustolle |
| Scroll Up | Ylemmäs samalla sivulla |
| Scroll Down | Alemmas samalla sivulla |
| New Window | |
| **Multimediavihjeet** | |
| Acrobat | PDf-tiedosto |
| Flash | Flashia käyttävä sivu |
| Real Media | Real media video- tai musiikkitiedosto |
| Sound | Äänitiedosto |
| Movie | Videotiedosto |
| 3D Media | 3D laseille suunniteltu sivu |
| Word | Word-tiedosto |
| Excel | Excel-tiedosto |
| PowerPoint | PowerPoint-tiedosto |
| **Sisältövihjeet** | |
| Definition | Määrittely- tai sanastosivu |
| Note | Lyhyt huomautus, kuten alaviite |
| Help | Ohjeita |
| Info | Käytetään lähinnä erottamaan tietosivulle johtavia linkkejä tiedostolinkeistä. Esim "Tietoa tuotteesta"vs. "Imuroi tuote" |
| Form | Kaavakkeen sisältävä sivu |
| Search | Hakusivu |
| Edit/Change | Mahdollisuus muuttaa linkitetyn tekstin arvoa |
| **Verkkopalveluvihjeet** | |
| Secure Page | Turvattu sivu, usein tilauslomake |
| Log-in page | Sivu vaatii salasanatunnistuksen |
| Email Link | Sähköpostiosoite |
| Download File | Yleinen linkki imuroitavaan tiedostoon |
| FTP Directory | FTP-hakemistoon |
| News | Uutissivulle |

Taulukko 4: QBullets-järjestelmän linkkityypit.

Tri83 R. Trigg. A network-based approach to text handling for the online scientific community (4. luku online www:ssä), 1983.

(W302 World Wide Web Consortium (W3C). Html 4.01 määrittely, 2002.

WL00 Harald Weinreich and Winfried Lamersdorf. Concepts for improved visualization of web link attributes. *WWW9 / Computer Networks*, 33(1-6):403–416, 2000.