

Tulevaisuuden verkkotekniikoiden ja -palveluiden tietoturva

Juha-Matti Berg

Tietotekniikan koulutusohjelman opinnäytetyö
Ohjelmistotekniikka
Insinööri (AMK)

KEMI 2012

ALKUSANAT

Kiitokset työn ohjaamisesta Tapani Ruokaselle. Lisäksi kiitän henkisestä tuesta, kannustuksesta ja neuvoista avovaimoa ja äitiä.

TIIVISTELMÄ

KEMI-TORNION AMMATTIKORKEAKOULU

Tekniikan yksikkö

Tekijä:	Juha-Matti Berg
Opinnäytetyön nimi:	Tulevaisuuden verkkotekniikoiden ja -palveluiden tietoturva
Sivumäärä:	47+1 liite
Päiväys:	7.5.2012
<p>Opinnäytetyön tavoitteena oli tutkia HTML5:n uusia tulevaisuudessa yleistäviä verkkotekniikoita ja nousevia verkkopalveluiden suuntauksia sekä selvittää miltä tulevaisuuden tietoturva näyttää verkkotekniikoiden ja palvelusuuntausten perusteella. Tehtävänä oli siis selvittää, millä tekniikoilla tuleva HTML5-standardi pyrkii uudistamaan verkkoa, mitkä ovat nousevat trendit verkkopalveluissa ja miten muutokset vaikuttavat Internetin yleiseen tietoturvaan. Lisäksi työssä käsitellään Internetin käyttäjän ja sivuntekijän roolia muutoksissa.</p> <p>HTML5 on kaivattu päivitys HTML-standardiin, joka on pysynyt neljännessä versiossaan jo lähes 20 vuotta. Tällä hetkellä HTML5 on luonnosvaiheessa, jossa määritellään HTML5:n sisältämiä teknologioita. Merkittävimmät selainten valmistajat ovat mukana luonnostyössä toteuttaen määriteltyjä tekniikoita. HTML5 pyrkii vastaamaan verkkopalveluiden todellisiin teknologiatarpeisiin tuomalla useita ominaisuuksia HTML:n osaksi. Suosituimmat verkkopalvelut alkavat nyt käyttää uusia tekniikoita, jolloin tietoturva nousee tärkeäksi osaksi verkon tulevaisuutta.</p> <p>Tutkimus on toteutettu perustutkimuksena, jossa tutkimusaineistona on käytetty tuoreimpia julkaisuja ja uutisia liittyen HTML5-määritelmiin. Suurin osa aineistosta on englanninkielistä.</p> <p>Tutkimuksen keskeisimmät tulokset ovat nähtävissä tässä opinnäytetyössä. Yleisenä päätelmänä voidaan todeta, että HTML5 parantaa yksittäisiä tekniikoiden tietoturvaominaisuuksia, mutta tuo mukanaan myös uusia haavoittuvuuksia ja potentiaalisesti kasvattaa WWW-sivujen hyökkäyspinta-alaa.</p>	
Asiasanat: HTML, tietosuoja, tietoturva, verkkopalvelut, WWW-sivustot.	

ABSTRACT

KEMI-TORNIO UNIVERSITY OF APPLIED SCIENCES

Technology

Name:	Juha-Matti Berg
Title:	Information Security of the Future Web Technologies and Services
Pages:	47+1 appendix
Date:	7 May 2012
<p>The aim of this study was to examine new web technologies provided by HTML5 and to find out what the future trends in web services are. In addition, the object was to analyze, on the basis of new trends in web technologies and services, what the information security concerns of the future web are. In short, the task was to find out what technologies HTML5 uses to modernize the web, what the emerging trends in web services are and how these changes affect the information security of the web. This study also provides an insight into the roles of designers and users of the changing web.</p> <p>HTML5 is a welcome update to the HTML standard which has remained in its fourth version for almost 20 years. At the moment, HTML5 is in draft stage, which means that the technologies are still being defined. All major browser vendors are participating in the drafting by implementing the defined technologies. The purpose of HTML5 is to provide technologies for the actual technical requirements of the web services directly as part of the HTML standard. The most popular web services are beginning to use the new web technologies and thus, information security rises to play an important role in the future of the web.</p> <p>This research was conducted as a basic research, where the research material includes the most recent publications and news regarding HTML5 technologies and web services. The bulk of the research material is in English.</p> <p>The key results of the research are available in this report. In general, it can be said that HTML5 improves some information security aspects of the web technologies but, at the same time, brings along new vulnerabilities and potentially increases the attack surface of websites.</p>	
Keywords: data protection, HTML, information security, web services, websites.	

SISÄLLYSLUETTELO

ALKUSANAT	2
TIIVISTELMÄ	3
ABSTRACT	4
SISÄLLYSLUETTELO.....	5
KÄYTETYT MERKIT JA LYHENTEET:	6
1. JOHDANTO	7
2. UUDET VERKKOTEKNIIKAT	9
2.1. HTML:n kehitys.....	9
2.2. HTML5	10
2.3. HTML5-määritelmien uudet tekniikat	12
2.3.1. HTML-elementit, rakenne ja semantiikka	13
2.3.2. Skriptauskielten käyttö	17
2.3.3. CSS3	18
2.3.4. Multimediatekniikat.....	19
2.3.5. Muistiominaisuudet	21
2.3.6. Paikannus (Geolocation).....	23
2.3.7. Muut teknologiat.....	23
3. VERKKOPALVELUSUUNTAUKSET.....	25
3.1. Semanttisuus ja rikkaat verkkosovellukset	26
3.2. Yhdistelmäsovellukset ja -palvelut	26
3.3. Pilvipalvelut	27
3.4. Sosiaalinen media ja sovellukset	28
4. TIETOTURVA	29
4.1. Tietoturva yleisesti ja muutamia käsitteitä	29
4.2. Tietoturvakulttuuri	30
4.3. Tutkimus ja uutisointi	31
5. LÄHITULEVAISUUDEN VERKON TIETOTURVA.....	32
5.1. Selaimet ja tulevaisuuden verkko	32
5.1.1. Selain sovellusalustana	33
5.1.2. Selainversioiden erot.....	34
5.2. Uhkat ja haavoittuvuudet	35
5.2.1. HTML5:n ominaisuuksien periaatteelliset riskitekijät	35
5.2.2. Yhtenäistämisen ongelmat.....	37
5.3. Kehittäjän ja käyttäjän vastuu verkossa.....	38
5.4. Haavoittuvuuksien tahallinen käyttö ja murtautuminen	39
6. YHTEENVETO	41
7. LÄHDELUETTELO.....	43
8. LIITELUETTELO	47

KÄYTETYT MERKIT JA LYHENTEET:

API	Application Programming Interface, ohjelmointirajapinta
CORS	Cross-Origin Resource Sharing, selaintekniikka, jolla sallitaan resurssipyyntöjä eri domainien välillä
CSS	Cascading Style Sheets, kaskadinen tyyliohjejärjestelmä
DOM	Document Object Model, dokumenttioliomalli
ENISA	European Network and Information Security Agency
GPS	Global Positioning System, paikannusjärjestelmä
(X)HTML	(eXtensible) Hypertext Markup Language, WWW:n sivunkuvauskieli
RDFa	Resource Description Framework in Attributes, tietomalli tietosisältöjen ja semanttisten suhteiden esittämiseen
RIA	Rich Internet Applications, rikkaat Internet-sovellukset
SQL	Structured Query Language, tietokantojen kyselykieli
URL	Uniform Resource Locator, merkkijono, jota käytetään osoittamaan WWW-sivuja
W3C	World Wide Web Consortium
WebGL	Web Graphics Library, JavaScript-pohjainen ohjelmointirajapinta 3D-grafiikkaa varten
WHATWG	Web Hypertext Application Technology Working Group
WWW	World Wide Web
XML	Extensible Markup Language, tiedon kuvaamisen formaatti
XSS	Cross-Site Scripting, Internetissä esiintyvä tietoturva-aukko

1. JOHDANTO

Internet on yleistymisensä jälkeen ollut tärkeä ja jokapäiväinen osa länsimaalaisen ihmisen elämää. HTML, verkkosivujen perustana pidetty rakennuspalikka, on pysynyt samanlaisena neljännessä versiossaan jo lähes kahden vuosikymmenen ajan. Nyt HTML:stä ollaan kehittämässä seuraavaa standardiversiota, jonka tarkoituksena on samaan aikaan yksinkertaistaa ja monipuolistaa verkkoa ja sen sisällön rakennetta.

HTML5-standardi on tarkoitus saada lopulliseen muotoonsa ja käyttöön muutamien vuosien aikana. On tärkeää muistaa, että HTML5 on tällä hetkellä luonnosasteella, eli määritelmät voivat muuttua radikaalistikin ennen standardin julkaisemista. Samaan aikaan verkon tarjoamat palvelut muuttuvat ajan hengen mukaisesti, ja näillä näkymin sosiaalisen median sovellukset sekä mm. pilvipalvelut ovat yleistymässä. Internet on myös muuttumassa vahvasti sisältöpainotteiseksi.

HTML5 on nykyään tuttu aihe tietotekniikkapiireissä. HTML5-standardin ja Internetin seuraavan sukupolven mukana tulevia tietoturva-asioita ja -ominaisuuksia käsitellään mediassa ja verkon foorumeilla yhä tiuhempaan tahtiin. Opinnäytetyön tavoitteena on siis selvittää mitkä ovat HTML5-luonnosten merkittävimmät uudet ominaisuudet ja lähitulevaisuuden verkkopalvelusuuntaukset sekä selvittää lähitulevaisuuden verkkopalveluiden käyttäjiä ja suunnittelijoita koskevia tietoturvanäkökohtia. Tietoturvaa käsitellään luonnollisesti selvittämällä merkittävimmät tietoturvauhat sekä verkkotekniikoista että -palveluista, mutta tutkimusta on tehty myös periaatetasolla palveluihin liittyen siten, että verkon käyttäjät, suunnittelijat ja kehittäjät voivat saada tutkimuksesta hyödyllistä ja yleistä tietoa lähitulevaisuuden verkosta.

Tietoturvaa ei tutkita palvelin- tai päätelaitetasoisesti, eli tutkimuksessa ei ole tarkasteltu sitä, miten laitteiden tietoturvaa ylläpidetään. Luvussa 4 käsitellään kuitenkin lyhyesti tietoturvaa terminä ja esitellään muutamia käsitteitä tietoturvasta.

HTML5:n osalta selvityksessä keskitytään merkittävimpiin uudistuksiin verkkotekniikoissa. Internet-selaimet ovat tutkimuksessa suurimmassa roolissa, koska ne soveltavat käytännössä tulevan standardin tekniikoita, sekä tarjoavat palveluita

verkon käyttäjille. Tutkimuksessa ei ole keskitytty niinkään teknisiin yksityiskohtiin, vaan opinnäytetyössä käsitellään HTML5:n periaatteita ja verkkopalveluita enemmän yleisellä tasolla.

Alan kirjallisuudessa termillä ”käyttäjä-agentti” (user agent) viitataan useasti verkkoselaimiin. Tässä opinnäytetyössä käyttäjää ei sovi sekoittaa selaimiin tai muihin verkon päätesovelluksiin, vaan käyttäjällä tarkoitetaan yksinkertaisesti ihmistä, joka käyttää verkon palveluita ja sivuja. Selaimista puhuttaessa tarkoitetaan Internetin tarkasteluun käytettäviä selaimia, joita löytyy nykyään työpöytä- ja mobiilikäytöstä.

Käyttäjän tietoturvaa ei tutkita tässä opinnäytetyössä ottaen kantaa siihen, mihin palveluihin käyttäjä henkilökohtaisiakin tietojansa antaa, eli tekstissä luotetaan käyttäjien omaan harkintaan palveluita kohtaan. Tietoturvalla, tietosuojalla ja yksityisyydellä tarkoitetaan yksinkertaisesti sitä, että tietoja ei käytetä luvattomasti tai ilman käyttäjän tietoisuutta.

2. UUDET VERKKOTEKNIIKAT

Tässä luvussa tarkastellaan taustatietoja HTML:stä, esitellään viimeisimpien HTML5-määritelmien uusimmat merkittävät tekniikat aihepiireittäin järjesteltynä ja tutustutaan siihen, miten HTML5 päivittää HTML-standardia. Lisäksi tarkastellaan lyhyesti HTML:n kehitystä.

2.1. HTML:n kehitys

HTML:n kehittämisestä vastaa nykyisin kaksi organisaatiota; W3C ja WHATWG. WHATWG perustettiin Applen, Mozilla Foundationin ja Opera Softwaren toimesta sen jälkeen, kun yritykset totesivat W3C:n kehittävän HTML-standardia huonosti todellisia verkon tarpeita silmälläpitäen. W3C ja WHATWG ylläpitävät kumpikin omaa versiotansa muuttuvista HTML5-luonnoksista, mutta luonnosten sisältö on suhteellisen sama. (WHATWG 2012, hakupäivä 7.3.2012)

Kun tässä opinnäytetyössä puhutaan HTML5-määritelmistä, -spesifikaatioista tai -luonnoksista, tarkoitetaan W3C:n ja WHATWG:n luonnosvaiheessa olevia HTML5-määritelmiä tai HTML5-määritelmiksi yleisesti luettavia tekniikkamääritelmiä. Työssä ei tehdä erottelua siitä, kuuluvatko yksittäiset määritelmät kummankin organisaation HTML5-määritelmiin, koska organisaatioiden luonnokset ja määritelmät ovat suhteellisen samanlaisia ja vielä voimakkaassa muutoksessa.

HTML:n kehittyminen on kulkenut ennen HTML5-määrittelyn alkamista W3C:n johtamana pitkälti epäsuorasti selainkehittäjien vetämänä. Kuvauskieltä on standardoitu ja dokumentoitu W3C:n toimesta vuodesta 1994 alkaen usein vasta sen jälkeen, kun selaimet ovat jo laajasti toteuttaneet uusia ominaisuuksia. WHATWG perustettiin usean selainkehittäjän toimesta juuri siitä syystä, että selainkehittäjät haluavat olla mukana kehittämässä HTML5-standardia samanaikaisesti, kun toimintoja voidaan testata käytännössä ja toteuttaa selaimissa. W3C ja WHATWG ovat kehittäneet HTML5-spesifikaatiota vuodesta 2007 alkaen yhteistyössä. (Korpela 2011, 24)

HTML:n rinnalla on verkkosivujen kehityksessä kulkenut XHTML, jonka juuret ovat tietojen rakenteen kuvaamiseen tarkoitettussa XML-kielessä. XHTML on käytettävissä myös HTML5:n rinnalla verkkosivujen kehityksessä. XHTML:n ja HTML:n käytön seurauksena syntyy verkkosivuja ja verkkosovelluksia kummallakin samaan tapaan, mutta XHTML-parserit ovat selaimissa hyvin tarkkoja sivujen oikein muodostetusta rakenteesta ja ohjelmointikielen kieliopista. XHTML valitaan kehityskieleksi siis lähinnä silloin, kun halutaan olla jostain syystä erityisen tarkkoja verkkosivujen rakenteen oikeasta muodosta. Verkkosivujen kehityksessä käytetään siis yleisimmin pelkkää HTML:ää, koska se antaa paremmin pienet sivujen virheet anteeksi. (Förster & Öggl 2011, 15-16)

2.2. HTML5

HTML5:n on tarkoitus puuttua siihen tosiasiaan, että Internet ja sen palvelut ovat muuttuneet huomattavasti vuodesta 1997, jolloin HTML:n nykyinen standardiversio on otettu käyttöön. Tulevalla standardilla vaikutetaan erityisesti multimedian ja sisällön esittämiseen. Myös verkkosivujen yleinen rakenne on muuttunut huomattavasti sisältökeskeisemmäksi. Verkon käyttäjät julkaisevat yhä enemmän omaa materiaalia blogeissa, sosiaalisessa mediassa ja yhteisöpalveluissa, joita ei kymmenen vuotta sitten ollut vielä olemassa.

Adoben kehittämä Flash-ympäristö on teknologia, jota on yleisesti käytetty median esittämiseen WWW-sivuilla. Flash-esitysten näyttämistä varten selaimen täytyy olla yhteensopiva Flash-pluginin kanssa, ja käyttäjän on pitänyt asentaa Flash-ympäristö laitteellensa. (Gil, hakupäivä 12.3.2012)

Flash-tekniikasta ollaan aina oltu kahta mieltä – se joko rikastuttaa ja parantaa verkkosivun antia tai tekee sivusta epäammattimaisen, käyttäjälle vaikean ja parhaimmillaan toimimattoman, jos Flash-ohjelmistoa ei voi käyttää käyttäjän laitteella (Nielsen 2000, hakupäivä 12.3.2012). Useat kannettavat laitteet eivät nykyisin toista Flash-sisältöä kannettavien laitteiden määrän samaan aikaan kasvaessa kiihtyvään tahtiin.

Flash-ympäristön käyttäminen on mobiililaitteilla rajoittunutta, koska laitteiden käyttämät selaimet eivät aina tue Flash-tekniikoiden käyttämistä. Adobe on itse myöntänyt jättävänsä mobiililaitteiden RIA-sovellusten kehittämisen HTML5-tekniikoiden harteille. (Vaughan-Nichols 2011, hakupäivä 12.3.2012)

HTML5 pyrkii vähentämään WWW-sivuilla tarvittavien lisäosien määrää, ja näyttäisi siltä, että esimerkiksi Flash-ympäristön käyttäminen on jäämässä HTML5:n multimediaominaisuuksien jalkoihin suosiossa. Flash on formaattina toki auttamattomasti vanha verrattuna alalla hehkutettuun HTML5:een, mutta HTML5 on myös mobiilialustojen kannalta kätevämpi ratkaisu sovellusten ja WWW-sivujen kehittäjille, koska palveluita ei välttämättä tarvitse tehdä kahtena versiona mobiili- ja työpöytäkäyttöön.

Kaikki merkit viittaavat vahvasti siihen, että HTML5-tekniikoiden käyttäminen tulee yleistymään niin työpöytäkäytössä kuin myös mobiililaitteilla ja -selaimilla. Multimedian esittäminen yhden formaatin alla on helpompaa kehittäjille ja suunnittelijoille, ja tarjoaa selkeämpää sisältöä verkon käyttäjille. Suuret organisaatiot ovat alkaneet jo siirtyä verkon seuraavalle aikakaudelle – esimerkiksi Google ja Microsoft ovat erittäin kiinnostuneita HTML5:n tulevaisuudesta, ja ovat mukana kehittämässä standardia kopiosuojauksen merkeissä (Laakso 2012, hakupäivä 22.3.2012).

HTML5:tä käytetään puhekielessä ja ammattislangissa samanlaisena hyvin laajana ilmaisuna kuin aikaisemmin IT-alalla yleisesti käytössä ollutta termiä ”Web 2.0” tai ”Web 3.0”. Termejä Web 2.0 ja Web 3.0 käytetään yleisesti kuvaamaan verkon aikakausia ominaisuuksien tai palvelusuuntausten valossa, esimerkiksi Web 2.0:n pääominaisuuksia ovat mm. tiedon jakaminen ja yhteistyö verkossa, käyttäjäkeskeinen sivujen suunnittelu sekä sosiaalinen media (O’Reilly 2005, hakupäivä 22.3.2012). Puhekielessä HTML5:llä viitataan usein virheellisesti kaikkiin mahdollisiin uusiin verkkotekniikoihin.

HTML5 on kuitenkin virallisesti itsessään HTML:n seuraavan version määritelmä, ja termiin HTML5 liitetään usein muitakin teknisiä määrittelyitä (Korpela 2011, s. 13). Tämän takia HTML5:n jakaminen teknologiana Web 2.0:een tai 3.0:een on vaikeata ja

tietyssä mielessä turhaa ainakin tämän opinnäytetyön tutkimuksen ja selvityksen kannalta.

Tuleva HTML5-standardi pyrkii yksinkertaistamaan HTML-kielen rakennetta siten, että se olisi enemmän ihmisen luettavissa ja ymmärrettävissä, mutta samaan aikaan paremmin koneiden tulkittavissa ja luokiteltavissa semanttisesti. Kieltä yritetään yksinkertaistaa ja samaan aikaan monipuolistaa siten, että erilaisten medioiden upottaminen olisi entistä helpompaa verkkosivuille mahdollisimman pitkälle ilman erillisiä ohjelmakirjastoja tai ajuripaketteja.

HTML5:n suunnittelua haittaa kaikkien kehittävien osapuolien harmiksi se, että Internet on ollut olemassa jo pitkään, ja kehittäjiä ja sivujen tekijöitä on ollut ajan saatossa mukana monia. Ongelma tässä on se, että vanhojenkin verkkosivujen pitäisi näkyä kaikissa selaimissa sellaisina kuin tekijä halusi niiden näkyvän, eikä kaikkia sivuja suinkaan ole tehty täysin yhteisten pelisääntöjen ja standardien mukaan. Tällöin HTML5-määritelmien ja selaimien tulee sisältää ominaisuuksia, joiden avulla vanhatkin (ja mahdollisesti virheelliset) sivut näkyvät käyttäjälle oikein. HTML5-sivujen pitää siis olla takaisinpäin yhteensopivia, ja jotain pitäisi sivuilla näyttää myös käyttäjille, jotka käyttävät vanhaa selainversiota, joka ei tue HTML5-ominaisuuksia. Uusilla standardeilla on siis lyhyesti sanottuna kovia vaatimuksia.

HTML5-määritelmien tilanne on vuonna 2012 edelleen häilyvä, eikä HTML5-standardin julkaisemiselle ole asetettu tarkkaa aikataulua. Seuraaviin alalukuihin on kasattu opinnäytetyön kirjoitushetkellä määritelmässä selkeästi mukana olevat ominaisuudet ja teknologiat, jotka muuttavat HTML-standardia oleellisesti, eikä aikaisemmin poistettuihin tai vanhentuneisiin HTML5-ominaisuuksiin oteta kantaa. Poistuneet ominaisuudet voivat tehdä myöhemmin paluun HTML5-määritelmiin, mutta palaavien ominaisuuksien ennustaminen olisi mahdotonta.

2.3. HTML5-määritelmien uudet tekniikat

Opinnäytetyössä ei tarkastella selaimien toteuttamia perustoimintoja, kuten URL:ien käyttämistä sivujen lataamisessa, vaikka ne kuuluvatkin HTML5:n spesifikaatioihin ja

näin ollen HTML:n ja HTML5:n ominaisuuksiin. Käsittelyssä ovat vain HTML:n kannalta uudet teknologiat ja ominaisuudet, sekä HTML5-termiin vahvasti liitetyt tekniikat. Lueteltuja ominaisuuksia on toteutettu selaimissa vaihtelevasti, ja jotkin ominaisuuksista voivat jäädä luonnosprosessin tuloksena kokonaan pois lopullisesta HTML5-standardista.

2.3.1. HTML-elementit, rakenne ja semantiikka

HTML-elementeistä ja kielen syntaksista on tehty joustavampaa siten, että selvistä dokumentin virheistä huolimatta selaimen täytyy pyrkiä näyttämään sivu mahdollisimman lähelle sitä, mitä sivun suunnittelija toivoi. Uudet elementit ovat myös taaksepäin yhteensopivia siten, että vanhemmat selaimet korvaavat uudet selaimelle tuntemattomat elementit mahdollisimman lähelle yhteensopivalla elementillä tai jättävät elementin kohdan dokumentista kokonaan pois.

Verkkosivujen kirjoittaminen ja suunnittelu HTML-dokumenttina on edelleen hyvin samanlaista kuin aikaisemmissa HTML-versioissa. HTML-dokumentit koostuvat elementeistä, jotka on määritelty dokumenttiin ”aloitus- ja lopetustageilla”, joiden väliin syötetään elementtien tarvitsemat tiedot. HTML-dokumentin alussa mainitaan dokumenttityyppi, joka HTML5:ssä merkitään yksinkertaisesti ”`<!DOCTYPE html>`”, minkä jälkeen HTML5:ssä sivun rakenne on melko vapaa. Selain rakentaa HTML-dokumentin pohjalta DOM:n, jota voidaan tarvittaessa muokata ohjelmallisesti ja dynaamisesti komentosarjakielten avulla. Elementit rakentuvat selaimella nähtäväksi verkkosivuksi DOM:n tietojen perusteella. (Tukkinen 2011, hakupäivä 15.3.2012)

Tyypillinen HTML5-pohjainen WWW-sivu koostuu vähintään liitteessä 1 nähtävistä elementeistä lukuun ottamatta liitteessä 1 olevaa sähköpostielementtiä, joka ei ole välttämätön. Liitteessä on myös kyseisen sivun muodostama DOM.

Ennen hyvin tarkkaan määritettyjen merkintöjen tekeminen on tehty lyhyemmäksi ja helpommin muistettavaksi. Esimerkiksi monissa HTML-elementeissä käytettävän type-attribuutin käyttämistä on vähennetty, jolloin selaimet tulkitsevat elementin tyyppimerkinnän puuttuessa elementin tyyppiä oletusarvoisen tyyppin. Esimerkiksi

tyylitiedostoa määriteltäessä ei tarvitse enää kirjoittaa ”<style type=’text/css’>”, vaan pelkkä style-tagin käyttäminen ilman type-määrettä riittää, ja selaimet tulkitsevat tyylin olevan oletuksena CSS-tyyliä. Oletusarvojen käyttäminen vaatii toki sitä, että sivun suunnittelija on maininnut HTML-sivun alussa dokumenttityyppiä määrittäessään käyttävänsä HTML5-tilaa. (Korpela 2011, 32)

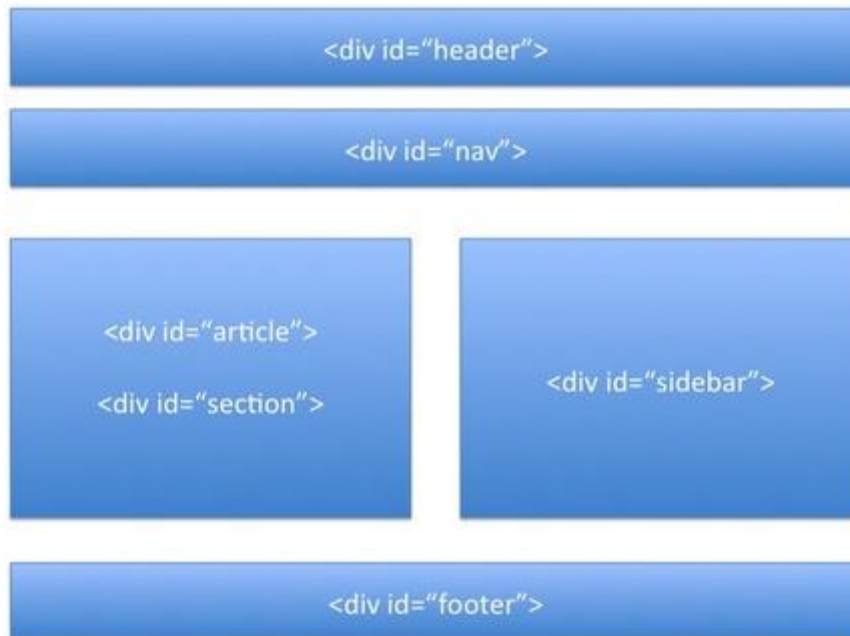
HTML5 tulee tukemaan verkon semanttisuutta tarjoamalla useita tapoja merkata sivuston tietoja semanttista käsittelyä tai luokittelua varten. Semantiikalla tarkoitetaan verkossa sitä, että koneet, hakukoneet ym. pystyvät ymmärtämään ja luokittelemaan verkkosivujen sisältöä. Sisällön semanttinen merkkäminen saa sivun näkymään hakukonehauissa paremmin. Semanttisella merkkäamisellä voi myös olla muita hyötyjä tulevaisuudessa, mutta tekniikoiden toteutus on vielä keskeneräistä. (Tukkinen 2011, hakupäivä 15.3.2012)

HTML5-määritelmistä löytyy tällä hetkellä tukea semantiikalle nimikkeillä mikrodata, RDFa ja mikroformaatit. Kaikilla semantiikkaa lisäävillä tekniikoilla lisätään HTML-dokumenttiin metatietoa, jota ei tavallisesti välttämättä olisi HTML-dokumentin elementeissä. Metatietoja käytetään aineiston luokitteluun siten, että koneet pystyvät luokittelemaan sivustojen tietoja eri tavoin. Lisätyt tiedot eivät välttämättä näy selaimen näyttämällä sivulla, vaan metatietoja käytetään esimerkiksi hakukoneissa. (Lawson 2011, hakupäivä 27.3.2012)

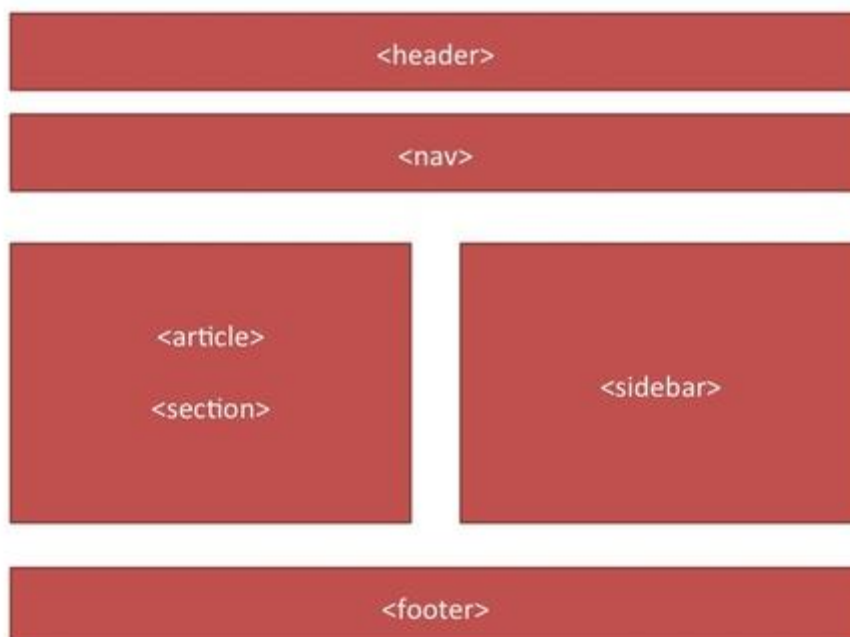
HTML5 tarjoaa uusia HTML-elementtejä HTML-dokumenttien rakenteen muodostamista varten. Uusilla elementeillä toteutetaan sellaisia rakenteita, joita on ennen yleisesti toteutettu kömpelösti ja vaikeammin div- tai span-elementeillä. Uusille rakenne-elementeille on varattu omat roolinsa dokumentin rakenteessa. Esimerkiksi footer-elementtiä käytetään alatunnisteiden esittämiseen ja nav-elementtiä dokumentin navigointiosaan. (Korpela 2011, 86)

Vanha tapa muodostaa HTML-dokumenttien rakennetta on nähtävillä kuvassa 1. On huomaamisen arvoista, että kaikki elementit ovat div-elementtejä, joille on vain annettu eri id-attribuutti, jonka avulla voidaan määritellä esimerkiksi elementtien tyyliä. Kuvassa 2 puolestaan nähdään HTML5:n tarjoama dokumentin rakenteen merkkäustapa, jossa rakenne on toteutettu uusilla HTML5-elementeillä. Tarvittavat

viittaukset tehdään elementteihin niille varatuilla uusilla nimikkeillä, eikä enää ole tarvetta käyttää kömpelöitä div-elementeistä koostuvia rakenteita.



Kuva 1. Vanha WWW-sivujen rakenteen merkintätapa (Sulkko 2011, hakupäivä 17.3.2012)



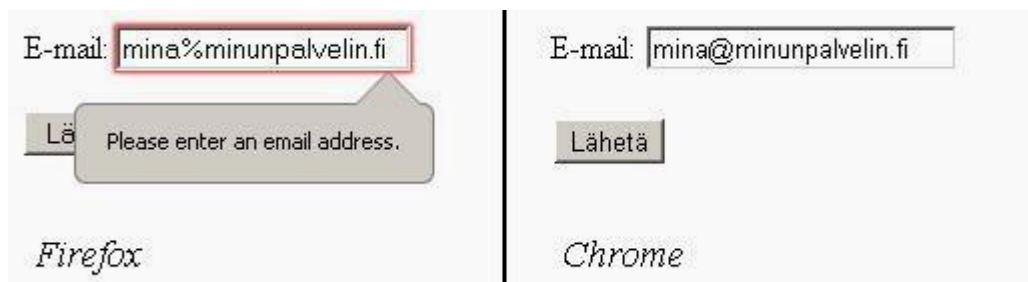
Kuva 2. HTML5-määritelmien mukainen sivun rakenteen merkintätapa (Sulkko 2011, hakupäivä 17.3.2012)

Uudistetut lomakekentät

Viimeisimpien HTML5-määritelmien osaksi on liitetty myös uudet lomakekentät, joita käytetään verkkosivun käyttäjän syötteen saamiseksi. Lomakekentät voi myös ymmärtää HTML-elementteinä. Syöte voi olla esimerkiksi jokin valinta (esimerkiksi väri- tai vaihtoehtovalinta), tekstiä, numeroita tai tiedostoja. HTML5 tarjoaa monipuolisempia ja käytännöllisempiä lomakekenttiä kuin HTML:n osalta on verkossa aiemmin nähty, eikä lomakekenttien tarkistuksen tekemiseen tarvita välttämättä enää toiminnallisuutta toteuttavaa ohjelmakoodia selaimessa. (Korpela 2011, 144-146)

Kehittäjän kannalta on hienoa ja käytännöllistä, että mobiililaitteen selaimen ollessa HTML5-määritelmien mukainen, toimivat uudet lomakekentät myös kannettavan laitteen selaimessa.

Kuvassa 3 on esimerkkinä Mozilla Firefox -selaimen (kuvassa vasemmalla) ja Google Chrome -selaimen (oikealla) versiot sähköpostilomakkeesta. Firefox-selaimesta käytössä oli versionumero 11 ja Chromesta versio 17. Kuvassa vasemmalla sähköpostiosoite on syötetty väärässä muodossa, jolloin sähköpostiosoitetta ei oteta jatkokäsittelyyn painettaessa kuvassa näkyvää lähetä -painiketta. Oikealla sähköpostiosoite on oikeassa muodossa, jolloin syöte voidaan lähettää jatkokäsittelyyn palvelimelle lähetä -nappia painamalla. Uudet HTML5-lomakkeet vähentävät näin huomattavasti ohjelmallisten skripteillä tehtävien tarkistusten tarvetta selaimissa. Käytetyn lomakkeen ohjelmallisuus on nähtävillä liitteessä 1 yksinkertaisena HTML5-dokumenttina.



Kuva 3. Sähköpostilomake kahdessa selaimessa

HTML5-määritelmä antaa selainkehittäjille tarkat vaatimukset lomakkeiden toteutuksesta, jolloin kenttien pitäisi olla toiminnallisuudeltaan samanlaisia kaikissa HTML5:ttä tukevissa selaimissa. Kuvasta 3 on ainakin sähköpostikenttien osalta nähtävissä se, että lomakekentät myös näyttävät graafisesti selaimissa suhteellisen samanlaisilta selainversiosta huolimatta. Vanhoissa HTML5-tuettomissa selaimissa sähköpostikenttä näytettäisiin tavallisena tekstikenttänä (Korpela 2011, 144).

HTML5 tarjoaa kaiken kaikkiaan vanhojen lomakekenttien lisäksi nykyisin värivalinnan, ajan valitsimia (kellonaika-, päivä-, viikko-, kuukausi-, ja vuosivalitsimet), lukualueen valitsimen, hakusanakentän, puhelinnumerokentän sekä URL-kentän vanhojen tuttuun HTML 4.01 -aikaisten lomakekenttien lisäksi. (Powers 2011, hakupäivä 20.3.2012)

2.3.2. Skriptauskielten käyttö

Verkkosivujen toiminnallisuuden tekemisen helpottamiseksi on HTML5:ssä edelleen mahdollista käyttää erilaisia komentosarjakieliä, joita kutsutaan myös yleisesti skriptikieliksi. HTML5-spesifikaatioissa suunnittelijoita kehoitetaan käyttämään jotain HTML5-spesifikaation mukaista elementtiä ja sen ominaisuuksia komentosarjan sijaan, jos se on sivuston rakenteen ja toiminnallisuuden kannalta mahdollista (W3C 2012b, hakupäivä 29.3.2012). Tällä tavalla erillisten skriptikielten ja laajennusten käyttäminen pyritään pitämään mahdollisimman vähäisenä.

Skriptauskieliä käytetään yleisesti silloin, kun jonkin toiminnon tekeminen pelkällä HTML-kielellä tai CSS:n avulla olisi liian vaikeaa tai työlästä. Yleinen käyttökäytännö on verkkosivun DOM-rakenteen muuttaminen dynaamisesti komentosarjan avulla. Komentosarjat asetetaan HTML-dokumenteissa script-tagien sisään, jolloin HTML5-ympäristössä selaimet ymmärtävät oletuksena komentosarjan olevan JavaScript-kieltä. Muitakin skriptikieliä voidaan käyttää, mutta tällöin HTML-dokumentissa on ilmoitettava erikseen käytettävä kieli.

Komentosarjakielten käyttäminen WWW-sivuilla ei sinänsä ole mitenkään uutta, mutta HTML5-määritelmässä sivujen toiminnallisuutta kannustetaan tekemään JavaScriptillä.

Joitain HTML5:n sisältämiä edistyneempiä ominaisuuksia ohjataan yksinomaan JavaScriptillä. JavaScript käsitetäänkin yleisesti yhdeksi perustavanlaatuiseksi osaksi HTML5-tekniikoita yhdessä CSS3:n ja muiden HTML5-määritelmien kanssa (Google Chrome Team & Wichary 2012, hakupäivä 25.3.2012).

Verkkosivujen kehittäjiä ja suunnittelijoita kannustetaan käyttämään JavaScript-kieltä HTML5:n yhteydessä, jos skriptin käyttäminen on verkkosivun toiminnallisuuden kannalta väistämätöntä. JavaScript on oliopohjainen ja funktionaalinen ohjelmointikieli, jota tulkitaan suoraan selaimessa (Jaakkola 2011, 12). Oliopohjaisuus ja funktionalisuus mahdollistavat monimutkaisiakin komentosarjoja, jotka suoritetaan suoraan verkkosivun käyttäjän selaimessa, jolloin ei rasiteta verkkosivua tarjoavaa palvelinta erillisellä käänös- tai tulkkauksella.

JavaScriptin lisäksi verkkosivujen toiminnallisuuden tekemiseen voidaan käyttää muitakin komentosarjakieliä, esimerkiksi palvelinpuolella tulkittavaa PHP:tä tai Pythonia.

2.3.3. CSS3

HTML:n kanssa on jo pitkään käytetty elementtien tyylimäärittelyitä CSS:n muodossa. HTML5-käsitteen alle luetaan usein myös CSS3, joka on CSS:n uusin versio. CSS:n kolmas versio ei muuta CSS-kielen rakennetta, vaan versiossa lisätään vanhojen ominaisuuksien päälle uusia ominaisuuksia. CSS3 on laaja kokonaisuus, joka koostuu pienemmistä osista, joita kehitetään eri tahtiin. (Korpela 2011, 284-285)

CSS:n avulla toteutetaan edelleen vanhaan tuttuun tyyliin verkkosivujen muotoilu ja presentaatiotyylit ja määritetään esimerkiksi HTML-elementtien leveys- ja pituusattribuutteja. CSS3 tarjoaa uutuutena edistyneitä graafisia esitysmuotoja kuten erilaisia transformaatioita (elementtien kiertoja, kallistuksia jne.), elementtien manipulointia kolmiulotteisesti (ponnahdus- ja upotusefektit), animaatioita ja tekstisisällön helppoa graafista manipulointia (HTML5 Rocks 2012b, hakupäivä 25.3.2012).

Uusien CSS-ominaisuuksien käyttäminen on sivuston suunnittelijan tai kehittäjän vastuulla – tuki uusille ominaisuuksille ja toiminnoille vaihtelee vielä suuresti verkkoselaimesta riippuen.

2.3.4. Multimediatekniikat

Yksi HTML5:n eniten näkyvyyttä saanut teknologia on median upottamisen helpottaminen, jonka pitäisi HTML5:n avulla onnistua myöhemmin täysin ilman lisäosia tai kirjastoja. HTML5 pyrkii multimediaominaisuuksien lisäämisellä vapauttamaan Internetiä Flash-ympäristön massakäytöstä tarjoamalla median upottamista verkkosivuille natiivisti HTML:n avulla. Ainakin luonnosvaiheessa median upottamisen ongelmana on ollut tuettujen mediaformaattien alhainen lukumäärä.

Mediaelementit

HTML5-määritelmässä on määritelty video- ja audio-elementit median toistamista varten. Elementit ovat helposti käytettävissä HTML-dokumentissa samannimisten <video> ja <audio> -tagien avulla. Mediaelementtien toistokontrollit voidaan määrittää elementeille erikseen siten, että kontrollit voidaan vaikka sijoittaa sivulla eri paikkaan kuin missä median toisto tapahtuu. Kontrollit voidaan kuitenkin asettaa helposti mediaelementtien avulla suoraan oletusmuodossa, jolloin kontrollien ulkonäkö riippuu käytettävästä verkkoselaimesta.

Ongelmia median upottamisessa aiheuttaa vielä se, että audio- ja videoelementit eivät vielä tue suurta määrää mediamuotoja, jolloin kehittäjä voi joutua käyttämään sivullaan lisäosaa tai kolmannen osapuolen kirjastoa median toistamiseen. Videoelementti tukee tällä hetkellä (käytettävästä selaimesta riippuen) MP4-, WebM-, ja Ogg-tiedostomuotoja. Audioelementti toistaa MP3-, Wav-, ja Ogg-formaatissa olevia äänitiedostoja. (Mozilla Developer Network 2012, hakupäivä 29.3.2012)

Mediaelementtien käyttäminen suoraan HTML-dokumentissa helpottaa palveluiden kehittäjien töitä, mutta näyttää peruskäyttäjälle tavalliselta mediasoittimelta.

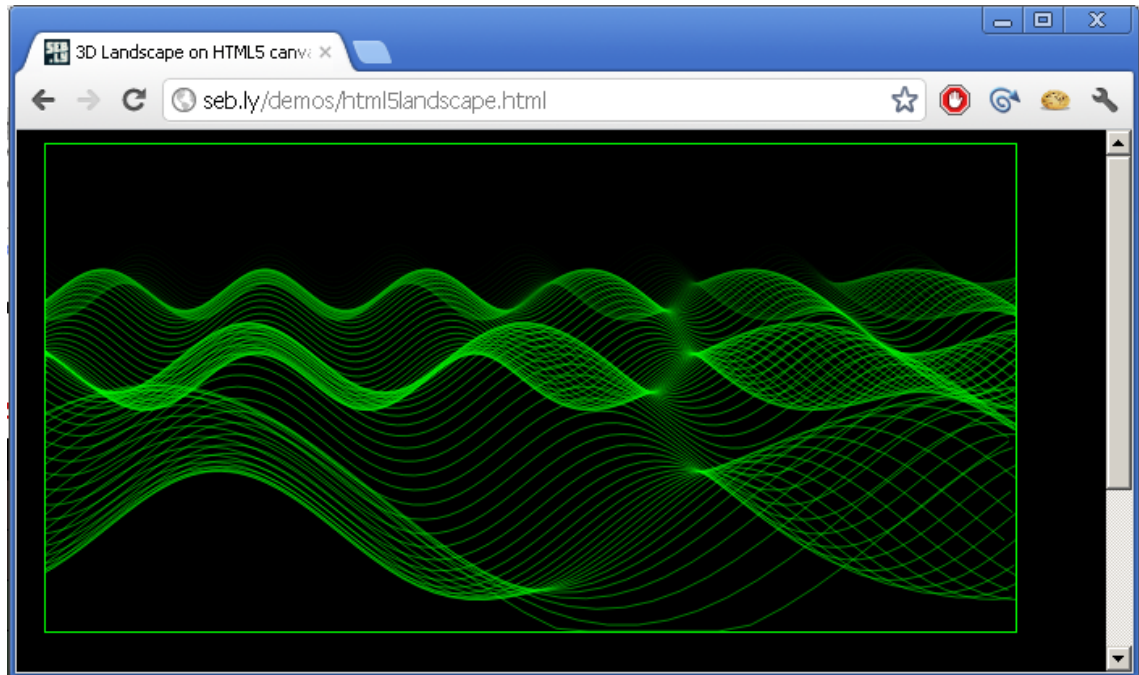
Mediaelementtien käyttö voi olla tulevaisuudessa kuitenkin käytännöllistä, jos tukea erilaisille tiedostoformaateille saadaan enemmän.

Canvas-elementti

Canvas-elementtiä käytetään HTML5:ssä rajaamaan verkkosivulta piirtoalue, jolla on tarkoitus näyttää 2D- tai 3D-grafiikkaa. Elementin avulla ei pysty yksinään tekemään juuri mitään, vaan grafiikan piirtämiseen täytyy käyttää jotain komentosarjakieltä. HTML5-määritelmässä suositellaan käytettäväksi JavaScriptiä. Canvas-elementille voi piirtää koordinaatistopohjaisesti yksinkertaisia pisteitä, viivoja, neliöitä, ympyröitä ja muita kuvioita sekä asettaa ja käyttää erilaisia taustakuvia, efektejä ja värivalintoja. (Förster & Öggl 2011, 108-174)

3D-grafiikan piirtoa varten on järkevää käyttää jotain tarkoitukseen sopivaa JavaScript-kirjastoa sen sijaan, että alkaisi itse ohjelmoimaan kolmiulotteisten mallien hallintaa alusta lähtien. WebGL-rajapinta on suunniteltu canvas-elementissä esitettävää 3D-grafiikkaa varten. WebGL:n käyttö on ilmaista ja sitä voi käyttää millä tahansa toiminnallisella skriptauskielellä, jolla pystyy muokkaamaan verkkodokumentin DOM-rakennetta. WebGL toimii laiteläheisesti siten, että rajapinnan kautta pääsee käsiksi selaimen, mutta myös tarvittaessa esimerkiksi laitteen näytönohjaimen ominaisuuksiin. Käyttö onnistuu siis esimerkiksi JavaScriptillä tai Javalla. Selainkehittäjiä Google, Opera, Mozilla ja Apple ovat yhteistyössä kehittämässä WebGL-rajapintaa. (Khronos 2011, hakupäivä 28.3.2012)

Kuvassa 4 on nähtävillä esimerkkinä canvas-elementin monipuolisesta toiminnasta canvas-elementti, joka piirtää JavaScript-kielen avulla yksinkertaista animoitua 3D-kuviota.



Kuva 4. Canvas-elementti toiminnassa (Lee-Delisle 2009, hakupäivä 4.4.2012)

2.3.5. Muistiominaisuudet

Verkkokehittäjät ovat pitkään kaivanneet käytännöllisempiä tallennus- ja muistiominaisuuksia selaimilta, mutta standardien puuttuessa ei muistiominaisuuksia ole tähän mennessä ollut pääosin muuten kuin evästeiden ominaisuudessa. HTML5-määritelmät tarttuvat tähän epäkohtaan useilla tavoilla, joita käsitellään seuraavaksi. Informaation tallennuksen menetelmät ovat määritelmien tässä vaiheessa vielä hyvin kirjavia ja luultavasti hyvin muuttuvia. Muistiominaisuuksien määritelmiä voidaan yhdistellä tai poistaa HTML5-standardin julkaisuun mennessä.

Paikallinen tietosäilö (Web Storage)

Paikallinen selaimen muistialueen sisäinen tietosäilö (HTML5-määritelmässä nimillä Web Storage ja Application Cache) tarjoaa tapoja suurienkin tietomäärien tallentamiseen verkkosivuilla vierailleen käyttäjän laitteen muistiin. Tiedon tallentamisen avulla verkkosivut voidaan saada latautumaan käyttäjän selaimessa jouhevammin ja kätevämmiin kuin aiemmin yleisten evästeiden avulla. Tietosäilöön on

mahdollista tallentaa kokonaisia verkkosivuja tai -sovelluksia siten, että sivuja tai sovelluksia on mahdollista käyttää selaimessa myös silloin, kun laitteella ei ole yhteyttä verkkoon. (W3Schools 2012a, hakupäivä 24.3.2012)

Verkkosivuilta vierailijan laitteelle tallennettava informaatio on täysin sivuston ylläpitäjän tai kehittäjän päätettävissä ja voidaan tallentaa helposti ilman erillisiä kirjastoja tai laajennuksia. Sivujen latautumisen nopeutuminen ja offline-ominaisuudet ovat tervetulleita ominaisuuksia mobiililaitteiden käyttäjille.

Tietokantaominaisuudet

HTML5-määritelmät tarjoavat nykyisessä versiossa myös mahdollisuutta tallentaa sisältöä verkkosivun käyttäjän selaimen tietokantamuodossa. Tietokantaominaisuudet löytyvät määritelmistä sekä täydellisenä relaatiotietokantana nimellä Web SQL Database että indeksoituna tietokantana nimellä IndexedDB, johon tiedot tallennetaan objektimuodossa. Tietokantaominaisuudet on lisätty erilaisten selainsovellusten kehittämisen helpottamiseksi. Tietokantoja käytetään ohjelmallisesti JavaScriptin tai muun toiminnallisen komentosarjakielen avulla.

Tietokantoja tarjotaan tällä hetkellä periaatteessa kolmessa erilaisessa muodossa ja nähtäväksi jää, mitkä tietokantatyypeistä jäävät lopulliseen standardiin. Tietokantoja on toteutettu tällä hetkellä vaihtelevasti eri selaimissa.

Selaushistoriarajapinta (History API)

HTML5 tarjoaa yksinkertaisen ja helppokäyttöisen rajapinnan verkkoselaimen sivuhistorian ohjelmallista tarkistusta ja manipulaatiota varten. JavaScriptin avulla on mahdollista liikkua sivuhistoriassa ohjelmallisesti samaan tapaan, kuin käyttäjä painaisi selaimen ”takaisin” -painiketta, ja poimia tarvittavia tietoja käsittelyä varten. Verkkosovelluksen on siis mahdollista etsiä vain aiemmin samassa ikkunassa (välilehdessä tai sivulla) vierailuilta sivuilta informaatiota. (Mozilla Developer Network 2011, hakupäivä 29.3.2012)

2.3.6. Paikannus (Geolocation)

Geolocation-tekniikan avulla verkkosivu voi käyttää vierailevan käyttäjän sijaintitietoa hyödykseen näyttämällä sijainnin käyttäjälle tai prosessoimalla sijaintitietoa jossain muussa tarkoituksessa. Paikannus ei ole verkossa uusi idea, mutta nyt ominaisuutta ollaan ensimmäistä kertaa liittämässä vahvasti HTML-määritelmiin.

Paikannusominaisuus on erittäin käytännöllinen esimerkiksi mobiililaitteessa, jossa on käytössä GPS tai jokin muu paikannustapa. Sijaintitiedon saamiseksi verkkosivu joutuu kuitenkin kysymään lupaa käyttäjältä joka kerta ennen paikannuksen tekemistä. Geolocation-rajapinta tarjoaa myös muita ominaisuuksia paikannusta varten, kuten paikannustietojen automaattisen käsittelyn määrärajoittain, minkä avulla voidaan esimerkiksi piirtää mobiililaitteen näytölle käyttäjän liikkeitä janaana. (W3Schools 2012b, hakupäivä 24.3.2012)

2.3.7. Muut teknologiat

HTML5 määrittelee myös muita uusia verkkotekniikoita HTML:n yhteyteen, esimerkiksi kommunikaatiota ja työpöytäsovellusten kaltaisia tiedostorajapintoja selaimiin.

Filesystem- ja File-rajapinnat tarjoavat selaimelle yksinkertaisten pääsyn laitteen tiedostojärjestelmään. File API mahdollistaa tiedostojen vetämisen (drag and drop) laitteen työpöydältä suoraan selainikkunaan käytettäväksi. Ominaisuus on hyödyllinen erityisesti kosketusnäytöllisissä laitteissa. Esimerkiksi Googlen Gmail käyttää File-rajapintaa liitteiden lisäämiseen (Boor 2010, hakupäivä 22.3.2012).

Filesystem API puolestaan tekee mahdolliseksi tiedostojen kirjoittamisen selainsovelluksen omaan varattuun muistialueeseen ja lukemisen muistialueelta mahdolliseksi verkkosovellukselle. Selainten tuki ominaisuudelle on kuitenkin vielä suhteellisen puutteellista, koska ainoastaan Google Chrome tukee ominaisuutta tänä päivänä. (HTML5 Rocks 2012a, hakupäivä 22.3.2012)

HTML5-luonnoksissa määritellään Messaging API -niminen rajapinta dokumenttien välistä viestintää varten. Dokumentit voivat sijaita verkossa eri sivustoilla, eli rajapinnan sisältämällä tekniikoilla voidaan sallia rajoitetusti sellaista toimintaa, minkä selaimet yleensä kieltävät tietoturvasyistä. W3C on julkaissut myös CORS-määritelmän, jossa määritellään selaimelle tapoja hyväksyä yhteyksiä useasta lähteestä. Selaimet ovat tyypillisesti aikaisemmin hyväksyneet yhteyksiä vain yhdestä lähteestä dokumenttia kohden. (Korpela 2011, 289; W3C 2012a, hakupäivä 3.4.2012)

Lisäksi HTML5:n yhteydessä on kehitetty WebSocket API -sovellusliittymää, jonka tarkoituksena on muodostaa selaimen ja palvelimen välille yksinkertaista protokollaa käyttävä nopea ja jatkuva yhteys. Yhteyden datana lähetetään vain sovelluksen kannalta tarpeelliset tiedot, jolloin yhteys pysyy kevyenä ja nopeana. (Korpela 2011, 292)

3. VERKKOPALVELUSUUNTAUKSET

Verkon palvelut ovat monipuolistuneet voimakkaasti 2000-luvulla, eikä palveluiden monimuotoisuuden kasvun pysähtymistä ole vielä havaittavissa. Seuraavaksi kuvaillaan yleistyvimmit verkkopalvelumuodot ja kehityssuuntauksat, joiden tarpeisiin HTML5-määritelmät yrittävät uusilla teknologioilla vastata. Verkkopalveluita tullaan toteuttamaan tulevaisuudessa yhä enenevässä määrin 2. luvussa käsitellyillä tekniikoilla. 3. luvussa käsitellään siis tulevaisuuden verkkopalveluiden suuntauksia.

Palveluita pyritään suuntaamaan nykyisin ja lähitulevaisuudessa erityisesti kannettaville päätelaitteille, koska mobiililaitteiden sekä laitteiden tiedonsiirron määrä on kirrkaassa nousussa, eli kannettavien laitteiden määrä päätelaitemarkkinoilla on nousussa (Schwartz 2010, hakupäivä 25.3.2012). Kevyille päätelaitteille sopii erityisesti pilvipalveluiden konsepti, jota käsitellään tässä luvussa.

Staatistien sivujen aika on ollut Internetissä ohi jo kauan sitten ja verkossa on siirrytty sosiaalisen webin aikakauteen. Verkkopalvelusuuntauksia kuvaavien Web 2.0- ja Web 3.0 -termien asteikolla ollaan menossa Web 2.0:ssa ja ollaan vähitellen siirtymässä kohti Web 3.0:aa, joka tähtää semanttisen verkon laajempaan toimintaan. Lähitulevaisuuden selkeitä näkymiä verkkopalvelusuuntauksissa ovat mm. sosiaalinen web, RIA-sovellukset, mashup-sovellukset, pilvipalvelut ja semanttinen web. (Jaakkola 2011, 14-19)

Internet on kokonaisuudessaan siirtymässä seuraavalle aikakaudelle tai sukupolvelle, joita voidaan kutsua monella nimellä. Semanttisen webin aikakaudeksi sitä ei voi vielä täysin kutsua, mutta Web 3.0 käsitetään yleisesti juuri tämän aikakauden termiksi. Web 3.0:ssa ollaan siirtymässä verkkosivujen ideasta verkkosovelluksiin ja -palveluihin. Web-versionumerointi on kuitenkin edelleen sumeaa siten, että tuskin kukaan tarkkaan tietää mitä kuuluu esimerkiksi Web 3.0:aan, eikä näitä termejä ole koskaan määritelty tarkasti. (Iskold 2007, hakupäivä 26.3.2012)

3.1. Semanttisuus ja rikkaat verkkosovellukset

Verkko on kehittynyt alun staattisista tekstiä ja kuvia sisältävistä verkkosivuista dynaamisiin lisäosien avulla toteutettuihin graafisiin ja toiminnallisiin sivuihin, ja nyt on siirrytty RIA-aikakaudelle. RIA:t ovat uudenlaisia rikkaita verkkosovelluksia, joiden kehittämiseen ei tulevaisuudessa tarvita verkkosivulla enää välttämättä lisäosia tai liittäviä kuten Flash tai Silverlight. RIA:t ovat mahdollistaneet monenlaisten palveluiden tarjoamisen verkossa, kuten tekstinkäsittelyn, laskentaohjelmistojen ja muiden työpöytäsovellusmaisten sovellusten käyttämisen verkossa suoraan selainympäristössä. (Taivalsaari 2009, hakupäivä 26.3.2012)

Semanttisella webillä tarkoitetaan sellaista verkkoa, jossa laitteet ja ohjelmat pystyvät tarjoamaan ihmisille automaattisesti parempia ehdotuksia ja tekemään automaattisesti uusia palveluita tiedon paremman luokittelun takia. Verkon palveluista merkitään tulevaisuuden semanttisessa verkossa niin tarkat tiedot, että laitteet pystyvät ymmärtämään ja luokittelemaan palveluita itsenäisesti. (Davies, Fensel & Richardson 2004, hakupäivä 26.3.2012)

Semanttinen verkko -käsite ei ole vielä RIA-aikakaudella päässyt täysiin oikeuksiinsa tarvittavien tekniikoiden puuttuessa. Tulevaisuuden näkymät ovat siis osittain havaittavissa: kun tarvittavat teknologiat saadaan käyttövalmiiksi, voivat laitteet toimia verkossa tehokkaammin siten, että uutta sisältöä kehittyy laitteiden ja sovellusten avulla lähes automaattisesti käyttäjää varten. Tähän pisteeseen on kuitenkin vielä selvästi pitkä matka, koska semantiikkaa tukevat tekniikat ovat yleistyneet hitaasti.

3.2. Yhdistelmäsovellukset ja -palvelut

RIA-tyylisten sovellusten jälkeen on näköpiirissä jo seuraava palvelusuuntaus, jota voidaan lyhyesti kutsua mashup-sovelluksiksi, eli sovelluksiksi, jotka yhdistelevät aikaisempien palveluiden ja sovellusten tuloksia muodostaen uutta sisältöä.

Mashup-sovellukset ovat jo käytössä ja näkyvissä esimerkiksi karttapalveluissa, jotka esittävät kulkureittejä tai sijainteja Google Maps -karttojen avulla. Google Maps ei sisällä käytettyjä sijainti- tai kulkureittitietoja, vaan tiedot haetaan jostain muualta. Tällöin käytetään useamman palvelun tuloksia hyväksi luoden uutta sisältöä verkkopalveluihin. (Taivalsaari 2009, hakupäivä 26.3.2012)

Selain on muuttumassa palvelurakenteen tarpeiden muutoksen myötä yhä enemmän sovellusalustaksi tavallisten verkkosivujen esittämisen lisäksi. Tämä tekee verkon sisällöstä monipuolisempaa, mutta tuo varmasti samalla esille uusia vielä tuntemattomia ja arvaamattomia riskejä. Tavallisen WWW-sivun ja WWW-sovelluksen raja alkavat vähitellen hämärtyä tavalliselle verkon käyttäjälle, eikä jaottelu verkkosivuihin ja -sovelluksiin ole välttämättä tulevaisuudessa enää tarpeellista sisällön ollessa tärkeämmässä roolissa.

3.3. Pilvipalvelut

Pilven määritelmä ei ole selkeä, vaan siitä on monenlaisia tulkintoja. Yleisesti voidaan kuitenkin sanoa, että pilvellä tarkoitetaan Internet-verkkoa itsessään. Perusajatusmallina on siis se, että käyttäjän kuluttamat palvelut sijaitsevat jossain muualla kuin käyttäjän laitteessa – sisältöineen. Pilvipalveluiden idea ei ole mitenkään uusi: monet verkkopalvelut ovat tarjonneet pilvipalveluiden kaltaisia sovelluksia ja palveluita jo pitkään.

Pilvipalveluita voidaan tarjota verkon käyttäjille useissa muodoissa. Yksinkertaisin tapa on palvelu, jossa käyttäjälle tarjotaan tallennustilaa omalle informaatiolle. On myös mahdollista, että pilvipalvelun käyttäjälle tarjotaan kokonaista infrastruktuuria (esimerkiksi käyttöjärjestelmää tai jotain sovellusta), jossa käyttäjä ei kuitenkaan ole tarkkaan tietoinen siitä, miten henkilökohtaisten tietojen tallennus toteutetaan tai missä päin maailmaa palvelu fyysisesti sijaitsee (Grance & Mell 2011, hakupäivä 27.3.2012).

3.4. Sosiaalinen media ja sovellukset

Ihmiset käyttävät yhä enemmän aikaa Internetissä ja tapaavat toisiaan virtuaalisesti verkkopalveluiden avulla. Sosiaalisesti mediaksi voidaan käsittää yleisesti kaikkisellainen verkon sisältö tai verkkopalvelut, missä keskeisimpänä osana ovat käyttäjien oma sisältö ja käyttäjien vuorovaikutus.

Sosiaalisessa mediassa informaation suunta on muuttunut yksisuuntaisesta kaksisuuntaiseksi siten, että käyttäjät eivät enää pelkästään etsi tietoa verkosta, vaan sisältöä tuotetaan myös itse. Tämä suuntaus on nähtävillä erityisen hyvin erilaisissa Internetin blogeissa, joissa käyttäjien on tarkoitus tuottaa kaikki sisältö itse. (Jaakkola 2011, 15-16)

Sosiaalisen median sovelluksille on tyypillistä, että käyttäjillä on mahdollisuus tuottaa ja muokata sisältöä, kommentoida ja keskustella sekä jakaa aineistoa verkottuen samalla keskenään (Rongas 2012, hakupäivä 3.4.2012). Sosiaalisten sovellusten käyttö ei näytä olevan ainakaan vielä hiipumassa ja nähtäväksi jää se, miten uudet tekniikat tulevat vaikuttamaan sosiaalisen webin sovellusten kehitykseen. Sosiaalinen media on ilmiönä vielä niin uusi, että sosiaalisen median palvelut saattavat muuttua arvaamattomallakin tavalla tarjoten täysin uusia sovelluksia.

4. TIETOTURVA

Luvussa tarkastellaan muutamia avainkäsitteitä tietoturvan maailmasta ja perehdytään tietoturvan taustoihin. Luvussa käydään myös läpi ajankohtaisia tietoturva-asioita tutkivat organisaatiot ja informaation lähteet.

4.1. Tietoturva yleisesti ja muutamia käsitteitä

Tietoturvalla tarkoitetaan yleisesti sitä perustaa, jolle tärkeän informaation käsittely rakentuu. Tietosuoja puolestaan tarkoittaa enemmän yksilön yksityisyyden suojaa, eli sitä, miten ihmisen henkilötietoja sekä tietoja henkilön henkilökohtaisesta toiminnasta kerätään ja käytetään. Tietoturvan ja tietosuojan käsitteitä ei sovi sekoittaa keskenään, vaikka ne liittyvätkin läheisesti toisiinsa. Tekniikoiden kehittyessä tietoturva ja -suoja ovat nousseet esille mediassa yhä tiheämpään tahtiin. (Järvinen 2002, 21)

Verkkotekniikoiden ja erityisesti verkkopalveluiden osalta tietoturva ja -suoja ovat tärkeitä asioita. Verkkopalvelun käyttäjätietojen vuotaminen tietokannan murtamisen tai teknisen vian takia, luvaton käyttäminen markkinoinnissa ja muut tietoturvan negatiiviset puolet ovat verkossa yhä suuremmassa ja näkyvämmässä roolissa.

Tietoturvaa käsiteltäessä puhutaan usein termeistä resurssi, uhka, hyökkäys ja tunkeutuminen. Resursseilla tarkoitetaan yleisesti turvattavia kohteita, kuten laitteita, ohjelmistoja, informaatiota sekä ihmisiä. Uhka -termillä tarkoitetaan mitä tahansa sellaista, jolla on potentiaalia murtaa, tuhota tai muuttaa resursseja jollakin tavalla, eli uhalla on ei-toivottuja vaikutuksia resursseihin. Hyökkäykseksi ja tunkeutumiseksi katsotaan sellainen toiminta, jossa hyökkääjä pyrkii voimakeinoin tai tunkeutuen pääsemään käsiksi resursseihin. (Allen 2002, 13)

Tietojärjestelmien ja palveluiden olisi hoidettava tietoturvasa siten, että resurssit on luottamuksellisesti ja eheästi suojattu ei-halutuilta tahoilta. Ulkopuolisten ei pitäisi pystyä muokkaamaan resursseja ja niiden tulisi aina olla käytettävissä esteettömästi palveluiden käyttäjille. Nykyisin käytetään todennusta niin, että käyttäjät pystytään

toteamaan oikeaksi verkkopalveluissa ja toisaalta myös niin päin, että käyttäjä pystyy toteamaan verkkopalvelun olevan aito. Käyttäjän tunnistaminen tehdään yleensä sellaisen tiedon avulla, minkä ainoastaan oikea käyttäjä tietää (yleensä salasana tai -lause). Käyttäjä puolestaan pystyy toteamaan palvelun olevan aito varmenteiden ja suojattujen yhteyksien tietojen avulla. (Järvinen 2002, 22-27)

Lyhyesti sanottuna asetelma on verkkopalveluiden ja -tekniikoiden kehittäjien asemassa seuraavanlainen: Resursseja pyritään suojaamaan uhilta, hyökkäyksiltä ja tunkeutumisilta.

4.2. Tietoturvakulttuuri

Verizonin tutkimuksen mukaan hyökkäykset ja tunkeutumiset ovat entisestään lisääntyneet vuosien aikana. Yrityksiä ja palveluntarjoajia vastaan tehdyt tietomurrot tulevat yhä useammin ulkopuolelta, eli kyseessä ei ole yleensä organisaation sisäinen oma työntekijä tai henkilö. Hyökkääjä tekee tietomurron yleensä oman hyödyn vuoksi, joko rahallisen tai muun henkilökohtaisen hyödyn vuoksi. Tunkeutumiset ja hyökkäykset kohdistuvat yhä useammin suuriin organisaatioihin, jolloin jo yhdellä onnistuneella tietomurrolla hyökkääjä voi saada merkittäviä henkilökohtaisia hyötyjä itselleen tai muuttaa verkkopalvelun resursseja jollakin tavalla. (Verizon 2012, hakupäivä 20.3.2012)

Tietojen kalastelu ja varastaminen ovat kokonaan oma taiteen lajinsa, ja hyökkäykset ja tunkeutumiset ovat suurimmaksi osaksi ammattilaisten tekemiä. Tekijät ovat siis ammattitietoudeltaan samassa pisteessä kuin organisaatioiden omat työntekijät tai tietävät jopa enemmän teknisistä ja tietoturvan asioista, jolloin hyökkääjät ovat etulyöntiasemassa. Hyökkääjien toiminta on siis yleensä hyvin pitkälle organisoitunutta.

Vuonna 2011 yleisimmät keinot tehdä tietomurtoja olivat erilaiset hakkerointitekniikat tai haittaohjelmien käyttäminen. Hyökkääjien tietomurtojen yritykset ovat muuttuneet vähitellen yhä passiivisemmiksi, eli verkkopalveluiden käyttäjätietoja varastetaan esimerkiksi verkkopalveluiden kirjautumissivuilta hyökkääjien sivuille asettamien haittaohjelmien avulla. On myös yleistä, että haittaohjelma on saatu jotain kautta

käyttäjän laitteelle tai selaimen, esimerkiksi haittaohjelman sisältäneen sähköpostin tai WWW-sivun kautta. Tyypillinen vaihtoehto käyttäjätietojen saamiseksi on myös käyttää tekniikoita, joissa käyttäjän syöttämät kirjautumistiedot lähetetään salaisesti hyökkääjän palvelimelle käyttäjän tietämättä. Täytyy myös muistaa, että hyökkääjä tarvitsee vain yhden käyttäjätiedot päästäkseen kirjautumaan luvattomasti verkkopalveluun. (Verizon 2012, hakupäivä 20.3.2012)

Palveluntarjoajat pystyvät ottamaan käyttöön uusia tekniikoita hitaasti siksi, että uusien tekniikoiden tietoturvaominaisuuksista ja luotettavuudesta täytyy olla varmuus ennen kuin tekniikoita voidaan ottaa käyttöön. HTML5-määritelmät kannustavat kehittäjiä ottamaan tekniikoita käyttöön samaan tapaan hitaasti pala kerrallaan, eli ei sillä tavalla, että organisaatio siirtyisi yhdellä kertaa ”HTML5-aikakaudelle”.

4.3. Tutkimus ja uutisointi

Kun suuret organisaatiot ottavat uusia verkkoteknologioita käyttöön käyttäjille suunnatuilla sivuillaan, voi tietoturva-aukkoja syntyä tärkeisiin palveluihin nopeastikin. Informaation nopea leviäminen tietoturvatapauksista on tärkeää.

Ajantasaista ja virallista informaatiota Suomen tietoturvatilanteesta on saatavilla Viestintäviraston tietoturvatyöryhmältä, CERT-FI:ltä osoitteesta www.cert.fi/. Kansainvälistä informaatiota aiheesta tarjoavat esimerkiksi CERT osoitteessa www.cert.org/ ja ENISA osoitteessa www.enisa.europa.eu/. Lisäksi uutisia on kattavasti saatavilla useilta tekniikka-aiheisilta uutissivustoilta.

5. LÄHITULEVAISUUDEN VERKON TIETOTURVA

Luvussa käsitellään tulevaisuuden verkon tietoturva HTML5-määritelmien ja yleistyvien verkkopalveluiden valossa. Erityisesti tarkastellaan sitä, miten HTML5-määritelmien kuvaamat uudet tekniikat vaikuttavat verkon yleiseen tietoturvaan. Luku sisältää pohdintaa ja yleiskuvauksen verkon tämänhetkisistä ja tulevaisuuden tietoturvanäkökohdista verkon käyttäjien, suunnittelijoiden ja kehittäjien kannalta. Teknisiä termejä, jotka mainitaan tässä luvussa, on selitetty verkkotekniikoiden osalta luvussa 2. Tietoturvatерmejä, hyökkääjien motiiveja ja hyökkäystekniikoita on käsitelty luvussa 4.

5.1. Selaimet ja tulevaisuuden verkko

Web-selaimet tulevat olemaan tulevaisuudessa merkittävässä roolissa Internetin kehityksessä, koska suurin osa teknologiamuutoksista painottuu selainsovellusten rakenteen muutoksiin tai erilaisiin laajennuksiin. Selaimista ollaan tekemässä merkittävää sovellusalustaa työpöytäsovellusten tapaisille sovelluksille ja selaimilla tullaan tulevaisuudessa varmasti näkemään sellaisia sovelluksia, joita ei ole ennen nähty Internet-käytössä. Yhä suurempi osa yhteiskunnan erilaisista hallintatoimista tapahtuu selainikkunassa, jolloin selaimet ja selainten tietoturva ovat kriittisessä asemassa (De Ryck, Desmet, Philippaerts & Piessens 2011, hakupäivä 29.3.2012).

Erikoisen tilanteesta tekee myös se, että ensimmäistä kertaa historian aikana on olemassa sellaisia kannettavia laitteita, joiden verkkoselaimet ovat toiminnallisuudeltaan ainakin lähes samanlaisia kuin järeämissä tietokoneissa toimivat selaimet. HTML5-määritelmät yrittävät tuoda verkkosivujen suunnittelutapoja yhteen siten, että suunnittelijoiden ei tarvitsisi tehdä eri versioita suunniteltavista sivuista mobiililaitteille ja työpöytäselaimilla tarkasteltaviksi. Sivujen oikein näkyminen ja tietoturvaominaisuudet ovat kuitenkin nykyään vahvasti liitoksissa siihen, kuinka käyttäjän selain on suunniteltu ja kehitetty.

5.1.1. Selain sovellusalustana

Selaimesta ollaan tekemässä uusien määritelmien mukaisesti verkkopalvelujen tarpeisiin sovellusalustaa, jonka päällä voivat toimia myös samankaltaiset sovellukset kuin työpöytäkäytössä. Aikaisemmin tarkasteltujen uusien tekniikoiden avulla sovellusten kehittäminen selaimen helpottuu huomattavasti, koska selain itsessään tulee tarjoamaan sovelluksen käyttöön muistialuetta ja ainakin osittaisen pääsyn laiteläheisiin resursseihin, kuten näytönohjaukseen.

Selaimet ovat ennen HTML5-määritelmiä olleet niin sanottuja kevyempiä päätesovelluksia, joiden taltioimat tiedot ovat olleen suhteellisen vähäisiä. Sivustot ovat taltioineet käyttäjien selaimiin tarvittavia tietoja evästeiden avulla, eikä evästeissä yleensä ole ollut tarkkoja tietoja selaimen käyttäjästä henkilökohtaisella tasolla. Kun selaimelle asetetaan vähitellen yhä suurempaa roolia verkossa, on hyökkääjille luvassa uudenlaisia mahdollisuuksia käyttäjien henkilökohtaisten tietojen saamiseksi. HTML5-määritelmät tarjoavat verkkosivuille ja -sovelluksille useanlaisia tapoja tallentaa tietoa käyttäjien selaimiin. Sovelluksia ja sivuja voidaan esimerkiksi tallentaa kokonaan offline-käyttöä varten, jolloin hyökkääjät voivat yrittää murtautua erilaisten haavoittuvuuksien kautta suoraan käyttäjien selaimiin henkilökohtaisten tietojen toivossa. (Lyne 2011, hakupäivä 23.3.2012)

Tämä tarkoittaa tiivistettynä sitä, että hyökkääjillä ei ole välttämättä entiseen tapaan tarvetta hakkeroida verkkopalvelun palvelimia tai yrittää päästä käyttäjien käyttöjärjestelmään varastamaan tietoja, vaan hyökkääjät voivat suunnitella laajoja tietomurtoja suoraan käyttäjien selaimen esimerkiksi oikealta näyttävien valesivujen, erilaisten selaimiin suunniteltujen haittaohjelmien tai muiden selainten haavoittuvuuksien avulla. Riskiryhmässä ovat tällöin erityisesti suosituimpien selainten käyttäjät.

Hyvänä esimerkkinä uhkan todellisuudesta käy jo se, että useat Internetin sähköpostisovellukset tarjoavat sähköpostien tallennusta selaimen offline-käyttöön (Lyne 2011, hakupäivä 23.3.2012). Uusien verkkopalveluiden kannalta tällaisten ominaisuuksien olemassaolo on toki hieno asia, koska se mahdollistaa uudenlaisen sisällön ja sovellusten tuottamisen verkkoon. Tavallisen verkon käyttäjän kannalta se,

että käyttäjän tietoja tallennetaan käyttäjän selaimen, voi kuitenkin olla merkittävä uhka. Lisäksi uudet hyökkäysmallit vaikuttavat selainten tulevaisuuden kehitykseen siten, että selainkehittäjien on suunniteltava uudenlaisia suojauksia ja suojausrakenteita käyttäjien tietojen suojaamiseksi (Lyne 2011, hakupäivä 23.3.2012).

5.1.2. Selainversioiden erot

Erilaisten verkkoselaimien määrä on ollut 2000-luvun jälkeen kasvussa. Erityisesti erilaisia kevyempiä mobiiliverkkoselaimia alkaa olla runsaasti. Selaimet voivat käyttää samoja taitto- ja renderöintimoottoreita kääntääkseen HTML- ja CSS-koodin käyttäjälle näkyväksi verkkosivuksi, mutta toiminnallisuuksia tehdään web-selaimissa silti hyvin eri tavoin. HTML5-määritelmät tuovat useita uusia ominaisuuksia ja toiminnallisuksia, joiden implementaatiosta selainkehittäjät vastaavat itse. Selaimet voivat tukea uusia ominaisuuksia vaihtelevasti, jolloin WWW-sivujen kehittäjien on varauduttava siihen, että joillakin selaimilla uudentyyppinen sisältö ei välttämättä toimi.

Tätä ongelmaa sotkee entisestään se, että verkossa on edelleen vanhojen selaimien käyttäjiä. Esimerkiksi vuonna 2001 julkaistulla Internet Explorer 6:lla on edelleen käyttäjiä, vaikkakin käyttömäärät ovat laskemassa vähäisiksi (StatCounter 2012, hakupäivä 30.3.2012).

On todennäköistä, että uusien tekniikoiden käyttöönoton alkaessa massiivisessa mittakaavassa löytyy erilaisista selainversioista monenlaisia haavoittuvuuksia. Todennäköisyyttä nostaa se, että uudet ominaisuudet tehdään itsenäisesti jokaisen selainkehittäjän ehdoilla, jolloin uusien ominaisuuksien toteutustyyliä on erilaisissa selaimissa monenlaisia.

Haavoittuvuudet eivät tällöin kuitenkaan ole samaan aikaan jokaisessa selaimessa. Suurten tietomurtojen tai hyökkäysten kohdentaminen on tällöin vaikeampaa, ja hyökkääjät pystyvät toteutustavoista riippuvien haavoittuvuuksien avulla uhkaamaan vain pientä osaa selaimista ja verkon käyttäjistä. Ominaisuuksien erilaiset toteutustavat ja potentiaaliset tietoturva-aukot vaikuttavat siis tietoturvaan siten, että suuria massatietovarkauksia tapahtuu vähemmän, mutta joidenkin selainten käyttäjät voivat

olla tietämättään hetkellisesti alttiina tietomurroille jonkin yksittäisen verkkosivuilla olevan elementin kautta. Selainkehittäjät korjaavat yleensä tietoturvaan liittyviä puutteita nopeasti. Selainversioiden monimuotoisuuden ei näin ollen pitäisi nousta erityiseksi ongelmaksi tulevaisuuden verkkotietoturvan kannalta.

5.2. Uhkat ja haavoittuvuudet

Tulevaisuuden verkon tietoturvaa ja yksilöiden tietosuojaa koettelevat useat uhkakuvat ja haavoittuvuudet. Osa haavoittuvuuksista liittyy HTML5-määritelmien uusiin periaatteisiin, kun taas osa uhkista liittyy verkon käyttäjiin ja verkkosivujen kehittäjien osaamiseen. Kenties suurin Internetin tietoturvaongelma ovat aina olleet käyttäjät itse tai puolivillaisesti suunnitellut käyttäjille suunnatut verkkopalvelut, joiden tietoturva on ollut helppo murtaa, tai palvelu on ollut käyttäjille liian vaikeaselkoinen.

5.2.1. HTML5:n ominaisuuksien periaatteelliset riskitekijät

HTML5-määritelmät kasvattavat edellä käsitellyn selaimen roolimutoksen lisäksi myös eräiden muiden tietoturvariskien potentiaalia tulla käytetyksi väärin tarkoituksiin. Kaikkia HTML5-tekniikoihin liittyviä yksittäisiä teknisiä tietoturva-aukkoja ei ole järkevää tarkastella tässä opinnäytetyössä, koska määritelmien teknisiin yksityiskohtiin tulee muutoksia jatkuvasti standardien ollessa luonnosvaiheessa. Seuraavaksi tarkastellaan siis muutamia sellaisia pysyvän oloisia tietoturvariskejä HTML5-määritelmistä, missä ongelmat liittyvät HTML5-määritelmien syvempiin periaatteisiin.

Samaan aikaan kun kannettavien laitteiden määrä ja käyttö lisääntyvät kiihtyvään tahtiin, HTML5-määritelmät tuovat mobiililaitteiden laitteiston paremmin selaimen käytettäväksi siten, että selaimen ja laitteeseen voidaan tallentaa käyttäjän viimeisimpiä paikkatietoja. Selainsovellukset pystyvät myös tarvittaessa käyttäjän luvalla käyttämään laitteiden kameroita ja mikrofoneja hyödykseen. Tähän mennessä on kuitenkin paljastunut jo useita sellaisia tapauksia, joissa laitteen laitteistoa on päästy käyttämään vahingollisesti ilman käyttäjän suostumusta. Tulevaisuudessa löytyy siis

melko varmasti lisää tietoturva-aukkoja, joiden avulla selaimen kautta voi päästä käsiksi käytettyyn laitteistoon. (Lyne 2011, hakupäivä 23.3.2012)

Selaimen muistiominaisuuksista on tehty suhteellisen joustavia ja riippuvaisia implementointimenetelmistä. Toisin sanottuna verkkosivun- tai sovelluksen selaimeen tallentama informaatio ja tietorakenne riippuvat suuresti siitä, miten verkkopalvelun kehittäjä on suunnitellut tiedon tallentamisen. Tallennettujen tietojen rakenne ei ole HTML5:n muistiominaisuuksissa selaimen käyttäjälle enää niin selkeää kuin evästeisiin tallennetuilla tiedoilla. (Lyne 2011, hakupäivä 23.3.2012)

HTML5 tuo mukanaan muutoksia dokumenttien välisen tiedonsiirron rajoituksiin. Aikaisemmista HTML:n tarkoista rajoituksista ollaan luopumassa window.postMessage-rajapinnan myötä, minkä avulla tietoa voidaan vaihtaa vapaammin eri sivujen tai ikkunoiden välillä. Tietojen alkuperän tarkistus jää tällöin verkkosivun kehittäjän vastuulle. (Weiss 2010, hakupäivä 28.3.2012)

HTML5-määritelmiin sisällytetyt multimediaominaisuudet yrittävät vapauttaa selaimet erilaisilta multimedian toistamiseen tarkoitetuilta laajennuksilta, joiden on aiemmin todettu olevan verkkosivujen tietoturvan kannalta ongelmallisia. Verkkosivujen kehittäjät vastaavat kuitenkin multimedian käytöstä sivuillaan, jolloin audio- ja videoelementtien käyttö voi olla kirjavaa erilaisilla verkkosivuilla ja avata uusia tietoturva-aukkoja. (Weiss 2010, hakupäivä 28.3.2012)

Lisää tietoturva-aukkoja tarjoavat HTML5:n multimedian rikastamista varten luodut erilaiset laajennukset. Esimerkiksi haitallisessa tarkoituksessa verkkosivuilla oleva canvas-elementti, joka käyttää monipuolisen multimedian esitykseen WebGL:ää, voi aiheuttaa verkkosivuilla käyneen käyttäjän laitteistossa näytönohjaimen täydellisen jumiutumisen, jolloin koko laitteisto muuttuu käyttökelvottomaksi ainakin hetkellisesti (Forshaw 2011, hakupäivä 31.3.2012). Tässäkin yhteydessä voi siis todeta, että palvelimen ylläpitäjän on harkittava melko tarkasti sitä, mitä uusia HTML5-ominaisuuksia sallitaan käytettäväksi palvelimella.

HTML5-määritelmiin lisättyjen semanttisuutta parantavien ominaisuuksien tietoturvaa on vielä hyvin vaikeaa analysoida, koska semanttiset ominaisuudet eivät ole vielä olleet

suuressa suosiossa. Internetin käyttäjille semanttisuus olisi varmasti piristävä ominaisuus, koska esimerkiksi hakukoneet osaisivat tehdä tarkempia ehdotuksia käyttäjän hakusanoille ja tarjota tarvittaessa juuri niitä palveluita, joita käyttäjä on hakemassa. Kukaan ei kuitenkaan vielä täysin varmasti osaa sanoa, miten semanttisuus vaikuttaa verkon tietoturvaan.

HTML5-määritelmät itsessään sisältävät monia tietoturvaongelmia periaatetasolla, mutta määritelmät ovat vielä luonnosasteella. On ainoastaan positiivista, että HTML5-määritelmien tietoturvaongelmat nousevat verkossa esille tässä vaiheessa HTML5-ominaisuuksien ja -tekniikoiden tullessa suosituimmiksi. HTML5 alkaa olla uutena ilmiönä siinä mittakaavassa, että luonnoksia on tutkimassa jatkuvasti kasvava aktiivinen harrastelijajoukko ja luonnoksia tutkitaan kriittisesti myös tietoturvaorganisaatioissa. Tämä tarkoittaa lähitulevaisuuden kannalta sitä, että luonnoksista löytyviä ongelmia voidaan korjata yhdessä luonnoksien kehittäjien kanssa.

Lisäksi on yleisesti hyvä asia, että HTML5 pyrkii suuntaamaan verkkopalveluiden kehitystä niin, että kolmansien osapuolien lisäosien ja laajennusten käyttäminen saadaan laskuun. Multimedian esitykseen ja RIA-käyttöön tarkoitettut pluginit ovat olleet pitkään yksi verkkopalvelujen kompastuskivistä tietoturvamielessä, koska lisäosien kautta on löytynyt hyökkääjille helppoja keinoja murtaa verkkosivujen suojauksia. Rikkaan verkkosisällön esittämismuotojen tuominen osaksi HTML-standardia tuo kuitenkin uudenlaisia potentiaalisia tietoturvariskejä, joita käsitellään seuraavassa alaluvussa.

5.2.2. Yhtenäistämisen ongelmat

HTML5:ttä rakennetaan vanhojen HTML-standardien päälle sen sijaan, että oltaisiin kehittämässä täysin uutta standardia. HTML-standardiin ollaan tuomassa kertaheitolla useita uusia ominaisuuksia monesta aihepiiristä ja myös entisiin kehitystyyleihin yritetään vaikuttaa joillakin muutoksilla. Edellisten HTML-versioiden tietoturvaongelmia ei ole suinkaan ratkaistu, vaan ongelmien hyväksikäyttöä vastaan on tehty useanlaisia ohjeita ja suosituksia sen sijaan, että standardeja olisi muutettu. HTML:n periaatteisiin kun on aina kuulunut joustavuus sivujen suunnittelussa. Suuri osa menneisyyden tietoturvaongelmista tulee jäämään standardeihin HTML:n vanhojen

ominaisuuksien muodossa, joita uuden HTML5:n on tuettava, jotta Internetin vanhatkin sivut toimisivat oikein nykyisissä selainversioissa.

Kun osittain tietoturvan kannalta ongelmalliseksi todettuja elementtejä tuodaan suuri määrä mukaan HTML-standardiin HTML5:n ominaisuuksien parantamiseksi, on selvää, että tulevan HTML5-standardin ja siihen liittyvien määritelmien hyökkäyspinta-ala kasvaa. (Eng 2010, hakupäivä 20.3.2012)

Hyökkäyspinta-alan kasvu tarkoittaa sitä, että verkkorikollisilla on enemmän potentiaalia löytää HTML5-määritelmien kuvaamista ominaisuuksista tietoturvaongelmia. Uusien verkkotekniikoiden yleistyminen on jo hyvässä vauhdissa tunnetuimmilla verkkopalveluntarjoajilla ja tulevaisuudessa jääkin näin nähtäväksi, kuinka paljon tietomurtoja tulee tapahtumaan.

5.3. Kehittäjän ja käyttäjän vastuu verkossa

Uudet verkkotekniikat kuten HTML5 eivät tule huolehtimaan tietoturvasta ja tietosuojasta itsestään, vaan verkkopalveluiden kehittäjien on itse luotava sellaiset puitteet palveluille, että ne olisivat mahdollisimman turvallisia. HTML5 sisältää monia sellaisia periaatteellisia ongelmia, joiden käyttöä tietomurtoihin voi estää suunnittelemalla verkkopalvelun WWW-sivut järkevästi huolehtien tarvittavista tietojen tarkistuksista ja siitä, ovatko uudet tekniset ominaisuudet vielä oikeasti täysin käyttövalmiita suurelle yleisölle kohdennetulle verkkosivulle. HTML5:n tarjoamat uudet verkkotekniikat ovat vielä suureksi osaksi luonnosasteella ja muuttuvat parempaan suuntaan kehityksen myötä myös tietoturvan osalta.

Uusilla verkkotekniikoilla toteutettujen mashup-palveluiden ja sosiaalisen median sovellusten tilanne voi olla tulevaisuudessa erityisen mielenkiintoinen tietoturvamielessä, koska palveluissa yhdistellään monenlaisia tekniikoita. Useiden erilaisten tekniikoiden yhdisteleminen samalla verkkosivulla voi johtaa uudenlaisten tietoturva-aukkojen löytämiseen. Tekniikoiden yhdistely kokeilumielessä on tärkeää HTML5:n ollessa vasta luonnosvaiheessa, koska tietoturvaongelmien löytyminen

myöhemmässä vaiheessa voisi johtaa suureenkin informaation vuotamiseen tapahtuessaan väärään aikaan.

Erityisen kriittisessä asemassa tulee olemaan edelleen jo verkkopalveluiden kehittäjille aikaisemmista kokemuksista tuttu käyttäjien syötteen validointi. HTML5 päivittää WWW-sivuilla käytettävät lomakekentät seuraavaan versioon, jossa käyttäjän syötteen rakennetta tarkistetaan jo käyttäjän selaimessa. Selaintarkistukset vähentävät toki virheellisten lomaketietojen käsittelyä palvelimen puolella ja ovat käyttäjänkin kannalta hyödyllisiä, mutta sivuntekijöiden ei pidä luottaa pelkästään uudistettujen lomakekenttien tarkistuksiin. Kaikki käyttäjien lomakkeisiin syöttämä tieto ja sivujen oikeellisuus on edelleen tarkistettava verkkopalvelun palvelimella, koska lomakekenttien tarkistuksia on mahdollista kiertää monin erilaisin keinoin ja lomakkeita on mahdollista manipuloida, jos verkkopalvelussa ei tehdä tarvittavia tietoturvatarkastuksia (McArdle 2011, hakupäivä 4.4.2012).

Käyttäjän on luonnollisesti edelleen oltava tarkkana verkossa sen suhteen, mitä tietoja kannattaa erilaisiin verkkopalveluihin luovuttaa. Olisi viisasta siis pitää mielessä ajatus siitä, että suuri osa käyttäjän verkossa olevasta sisällöstä voi joutua väärin käsiin tai kadota.

5.4. Haavoittuvuuksien tahallinen käyttö ja murtautuminen

HTML5 pyrkii ottamaan uusissa ominaisuuksissa myös tietoturvan ja käyttäjien tietosuojan huomioon, mutta on syytä muistaa, että hyvätkin suojaukset voidaan yleensä kiertää tavalla tai toisella. Sivuntekijän on hyvin vaikeaa tehdä sivuja niin täydellisesti, että yhtään tietoturva-aukkoa ei jää uhkaamaan suojattavia resursseja. Verkkosivujen kehittäjien on siis parempi varautua tietomurtoihin, koska on hyvin todennäköistä että suojauksia yritetään murtaa jossakin pisteessä – ainakin sitten jos kehitettävä verkkopalvelu on suosittu ja sisältää hyökkääjien kannalta arvokkaita resursseja. Kehittäjien kannattaa siis tehdä toimintasuunnitelmat tietomurtojen varalta jo hyvissä ajoin.

HTML5:n monia hyväksikin todettuja tietoturvaominaisuuksia on mahdollista kiertää, joskin syy on yleensä sivujen puutteellisessa implementaatiossa tai kehittäjän tietämättömyydessä. On osoitettu, että useat HTML5:n sisältämät ominaisuudet sisältävät sellaisia puutteita, joita voidaan hyödyntää erilaisten hyökkäystekniikoiden avulla. Yleisiä uusia ominaisuuksia vastaan käytettyjä hyökkäystekniikoita ovat edelleen XSS-tyyppiset hyökkäykset, erilaiset injektio menetelmät (esimerkiksi haitallisen JavaScript-koodin tai SQL-injektion käyttö) ja käyttäjätietoa keräävien haittaohjelmien käyttö. Lyhyesti sanottuna selain on noussut välineeksi, jonka kautta yritetään päästä käsiksi resursseihin, ja selain tulee olemaan tulevaisuudessa tietoturvalan ammattilaisten ja kyberkriminaalien välinen taistelukenttä. (McArdle 2011, hakupäivä 4.4.2012)

6. YHTEENVETO

Yhteenvetona voidaan todeta, että HTML5 pyrkii korjaamaan määritelmien avulla tällä hetkellä jo joitakin tietoturvan ja tietosuojan puutteita, mutta avaa samalla uusia mahdollisuuksia käyttää tekniikoita hyväksi väriin tarkoituksiin. Verkkosivujen ja -palveluiden kehittäjien on mietittävä tarkkaan, mitä uusia tekniikoita voi ottaa käyttöön turvallisesti. HTML5 työntää siis suureksi osaksi tietoturva-asioista huolehtimisen verkkopalveluiden suunnittelijoille ja kehittäjille.

HTML5 siirtää verkon todellisten nykytarpeiden mukaisesti toiminnallisuutta sopivasti asiakas-palvelin -suhteessa asiakaspäähän eli verkkoselaimille. Selaimen merkityksen kasvattaminen tuo uusia haasteita niin verkkopalveluiden kehittäjille kuin myös selainkehittäjille tietoturvamielessä. Osa tulevista tietoturvaongelmista on suurten muutosten myötä varmasti myös niin arvaamattomia, että niitä on vielä mahdotonta ennustaa. Positiivista on se, että HTML5 mahdollistaa monien asioiden helpomman käsittelyn verkkopalveluissa. Nähtäväksi jää kuitenkin se, miten uudet tekniikat toimivat yhdisteltäessä uusiksi verkkopalveluiksi.

HTML5-suositukset ja -määritelmät eivät ehkä kerro tällä hetkellä sivuntekijöille tarpeeksi selkeästi, miten sivujen ominaisuuksia tulisi tehdä uusilla tekniikoilla. HTML5 antaa suuren vastuun verkkosivujen suunnittelijoille, koska heidän pitäisi huomata itse koska on tarpeellista tehdä tarkistuksia verkkosivuilla käsitellylle informaatiolle. Internetin käyttäjät puolestaan joutuvat itse määrittelemään, ovatko verkkosivut luotettavia vai eivät. Kaikki käyttäjät eivät välttämättä omaa riittävän kriittistä näkökulmaa Internetistä, jolloin tietosuojan murtuminen on todennäköistä.

Jos unohdetaan tietoturvataarkastelu, HTML5:n hyödynnettävyys verkkopalveluissa näyttää hyvältä. HTML5 tarjoaa useita käytännöllisiä ominaisuuksia, kuten paremman sivujen rakenteen kuvauksen ja uudistetut lomakekentät. Verkkoselainten vuorovaikutteisuus näyttäisi paranevan HTML5:n määrittelemien ominaisuuksien myötä – esimerkiksi tiedoston vetäminen ja pudottaminen suoraan selainikkunaan käsiteltäväksi on käyttäjälle vuorovaikutteista ja loogista. Erityisesti mobiililaitteiden

käyttäjille näyttäisi olevan luvassa hyviä kokemuksia, koska HTML5 tekee verkkopalveluiden tuottamisen mobiilialustoille helpommaksi.

Jos esille nousseet tietoturvaongelmat otetaan huomioon riittävän vakavasti HTML5:n jatkoluonnostelussa ja verkkokehittäjille saadaan yhtenäiset ja tarkat ohjeet tekniikoiden käytöstä ja riskien minimoimisesta, vain verkkokehittäjien luovuus on rajana erilaisille uusille verkkopalveluille. Tällä hetkellä HTML5:n suosiota verkkopalveluissa rajoittavat vielä ongelmat tietoturvaominaisuuksissa ja osittain vanhojen selainversioiden käyttäjien määrä.

Tämän opinnäytetyön aiheena olleen tutkimuksen voisi sanoa olleen hyvin laaja. Työssä mainitun aineiston parsiminen ja valinta aiheutti työn tekijälle haasteita opiskelutekniikoiden ja ajankäytön osalta, mutta tutkiminen opetti samaan aikaan uusista verkkotekniikoista paljon uusia asioita, mikä toivottavasti välittyy tästä dokumentista myös lukijalle.

7. LÄHDELUETTELO

- Allen, Julia H. 2002. CERT – Verkkotietoturvan hallinta. Helsinki: Edita.
- Boor, Adam 2010. Drag and drop attachments onto messages. Hakupäivä 22.3.2012.
<<http://gmailblog.blogspot.com/2010/04/drag-and-drop-attachments-onto-messages.html>>
- Davies, John & Fensel, Dieter & Richardson, Marc 2004. The Future of Web Services. Hakupäivä 26.3.2012.
<http://www1.cs.unicam.it/insegnamenti/2007_reti2/Readings/05092713223030758.pdf>
- De Ryck, Philippe & Desmet, Lieven & Philippaerts, Pieter & Piessens, Frank 2011. A Security Analysis of Next Generation Web Standards. Hakupäivä 29.3.2012.
<www.enisa.europa.eu/activities/Resilience-and-CIIP/critical-applications/web-security/a-security-analysis-of-next-generation-web-standards/at_download/fullReport>
- Eng, Chris 2010. HTML5 Security in a Nutshell. Hakupäivä 20.3.2012.
<www.veracode.com/blog/2010/05/html5-security-in-a-nutshell/>
- Forshaw, James 2011. WebGL - A New Dimension for Browser Exploitation. Hakupäivä 31.3.2012. <www.contextis.com/resources/blog/webgl/>
- Förster, Klaus & Öggl, Bernd 2011. HTML5 Guidelines for Web Developers. Boston: Addison-Wesley.
- Gil, Paul. What is 'Flash'? Is that the same as 'Adobe Flash'?. Hakupäivä 12.3.2012
<<http://netforbeginners.about.com/od/f/f/flash.htm>>
- Google Chrome Team & Wichary, Marcin 2012. HTML5 - Web Development to the next level. Hakupäivä 25.3.2012. <<http://slides.html5rocks.com/#landing-slide>>
- Grance, Timothy & Mell, Peter 2011. The NIST Definition of Cloud Computing. Hakupäivä 27.3.2012. <<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>>
- HTML5 Rocks 2012a. HTML5 Features: File Access. Hakupäivä 22.3.2012.
<http://www.html5rocks.com/en/features/file_access>
- HTML5 Rocks 2012b. HTML5 Features: Presentation. Hakupäivä 25.3.2012.
<<http://www.html5rocks.com/en/features/presentation>>

Iskold, Alex 2007. Web 3.0: When Web Sites Become Web Services. Hakupäivä 26.3.2012.

<http://www.readwriteweb.com/archives/web_30_when_web_sites_become_web_services.php>

Jaakkola, Tapio 2011. Arvio HTML5:n uusien ominaisuuksien sovellettavuudesta. Turun yliopisto. Ohjelmistotekniikan koulutusohjelma. Diplomityö.

Järvinen, Petteri 2002. Tietoturva & Yksityisyys. Jyväskylä: Docendo.

Khronos 2011. Getting Started. Hakupäivä 28.3.2012.

<www.khronos.org/webgl/wiki/Getting_Started>

Korpela, Jukka K. 2011. HTML5 – Uudet ominaisuudet. Jyväskylä: Docendo.

Laakso, Henri 2012. Microsoft ja Google havittelevat kopiosuojausta HTML5:een.

Hakupäivä 22.3.2012.

<www.mikropc.net/kaikki_uutiset/microsoft+ja+google+havittelevat+kopiosuojausta+html5een/a784451>

Lawson, Bruce 2011. HTML5 Semantics. Hakupäivä 27.3.2012.

<<http://coding.smashingmagazine.com/2011/11/18/html5-semantics/>>

Lee-Delisle, Seb 2009. 3D landscape on HTML5 Canvas. Hakupäivä 4.4.2012.

<<http://seb.ly/demos/html5landscape.html>>

Lyne, James 2011. HTML5 and security on the new web. Hakupäivä 23.3.2012.

<www.sophos.com/medialibrary/PDFs/other/sophosHTML5andsecurity.pdf>

McArdle, Robert 2011. HTML5 Overview: A Look at HTML5 Attack Scenarios.

Hakupäivä 4.4.2012. <www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/reports/rpt_html5-attack-scenarios.pdf>

Mozilla Developer Network 2011. Manipulating the browser history. Hakupäivä 29.3.2012.

<https://developer.mozilla.org/en/DOM/Manipulating_the_browser_history>

Mozilla Developer Network 2012. Using HTML5 audio and video. Hakupäivä

29.3.2012. <https://developer.mozilla.org/en/Using_HTML5_audio_and_video>

Nielsen, Jakob 2000. Flash: 99% Bad. Hakupäivä 12.3.2012.

<www.useit.com/alertbox/20001029.html>

O'Reilly, Tim 2005. What Is Web 2.0?. Hakupäivä 22.3.2012.

<<http://oreilly.com/web2/archive/what-is-web-20.html>>

- Powers, David 2011. Understanding HTML5 intelligent forms – Part 1: New input elements. Hakupäivä 20.3.2012. <www.adobe.com/devnet/dreamweaver/articles/html5-forms-pt1.html>
- Rongas, Anne 2012. Sosiaalisen median palvelut ja välineet hyötykäyttöön. Hakupäivä 3.4.2012. <http://www.edu.fi/materiaaleja_ja_tyotapoja/tvt_opetuksessa/mika_ihmeen_sosiaalinen_media>
- Schwartz, Ewan I. 2010. The Mobile Device Is Becoming Humankind's Primary Tool (Infographics Feature). Hakupäivä 25.3.2012. <www.technologyreview.com/business/26653/?ref=rss>
- StatCounter 2012. Top 12 Browser Versions from July 2008 to March 2012. Hakupäivä 30.3.2012. <http://gs.statcounter.com/#browser_version-ww-monthly-200807-201203>
- Sulkko, Samuli 2011. Html5 – mikä se on ja mitä se tarkoittaa?. Hakupäivä 17.3.2012. <www.tietoviikko.fi/msareena/msblogit/bittiviidakko/html5+ndash+mika+se+on+ja+mita+se+tarkoittaa/a688822>
- Taivalsaari, Antero 2009. Mashware: The Future of Web Applications. Hakupäivä 26.3.2012. <http://labs.oracle.com/techrep/2009/sml_i_tr-2009-181.pdf>
- Tukkinen, Pia 2011. HTML5 ja WWW-sivut. Hakupäivä 15.3.2012. <https://noppa.aalto.fi/noppa/kurssi/t-111.1100/luennot/T-111_1100_kotisivut_-_html_-_perusteet.pdf?state:CourseLectures=BrO0ABXcOAAAAQAAB3Nob3dBbGxzcgARamF2YS5sYW5nLkVjb2xlYW7NIHKA1Zz67gIAAVoABXZhbHVleHAA>
- Vaughan-Nichols, Steven J. 2011. Flash is dead, Long live HTML5. Hakupäivä 12.3.2012. <www.zdnet.com/blog/networking/flash-is-dead-long-live-html5/1633>
- Verizon 2012. 2012 Data Breach Investigations Report. Hakupäivä 20.3.2012. <http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2012_en_xg.pdf>
- W3C 2012a. Cross-Origin Resource Sharing: W3C Working Draft 3 April 2012. Hakupäivä 3.4.2012. <<http://www.w3.org/TR/2012/WD-cors-20120403/>>
- W3C 2012b. HTML5: Edition for Web Authors: W3C Working Draft 29 March 2012. Hakupäivä 29.3.2012. <dev.w3.org/html5/spec-author-view/spec.html>

W3Schools 2012a. HTML5 Application Cache. Hakupäivä 24.3.2012.
<www.w3schools.com/html5/html5_app_cache.asp>

W3Schools 2012b. HTML5 Geolocation. Hakupäivä 24.3.2012.
<www.w3schools.com/html5/html5_geolocation.asp>

Weiss, Aaron 2010. Top 5 Security Threats in HTML5. Hakupäivä 28.3.2012.
<www.esecurityplanet.com/trends/article.php/3916381/Top-5-Security-Threats-in-HTML5.htm>

WHATWG 2012. FAQ. Hakupäivä 7.3.2012
<<http://wiki.whatwg.org/wiki/FAQ>>

8. LIITELUETTELO

Liite 1 (2 sivua): Yksinkertainen HTML5-sivu ja DOM

Yksinkertainen HTML5-sivu ja DOM

Tässä HTML-dokumentissa kuvataan yksinkertainen HTML5-sivu, jossa on lomakekenttä sähköpostiosoitetta varten. Lomakkeeseen kuuluu myös lähetä -painike.

Koska kyseinen esimerkkisivu pyrkii olemaan mahdollisimman yksinkertainen, ei määritellä esimerkiksi merkistökoodausta. Tässä on näkyvillä vain kaikki peruselementit, joita HTML5 tarvitsee toimiakseen. ASP-muotoinen lomakkeen käsittelijä on tässä vain esimerkin vuoksi osoittamassa sen, että lomakkeiden syötteet täytyy vielä käsitellä erikseen.

```
<!DOCTYPE html>
<html>
<head>
<title>HTML5-sivu</title>
</head>
<body>

<form action="email_kasittelija.asp">
  E-mail: <input type="email" name="kayttajan_email" /><br /><br />
  <input type="submit" value="Lähetä" />
</form>

</body>
</html>
```


Kyseinen yksinkertainen HTML5-sivu tuottaa seuraavanlaisen dokumenttioliomallin, jota kuvataan DOM:n solmurakenteena. Rakenteessa on nähtävissä olioita, joilla on erilaisia attribuutteja kuten tyyppejä ja tekstejä.

```
└─ HTML
  └─ HEAD
    └─ #text:
      └─ TITLE
        └─ #text: HTML5-sivu
      └─ #text:
    └─ #text:
  └─ BODY
    └─ #text:
    └─ FORM action="email_kasitteliija.asp"
      └─ #text: E-mail:
      └─ INPUT type="email" name="kayttajan_email"
      └─ BR
      └─ BR
      └─ #text:
      └─ INPUT type="submit" value="Lähetä"
      └─ #text:
    └─ #text:
```