

Dokumentaatio: WWW-sovellukset

Webgraph – Datavisualisointia webissä

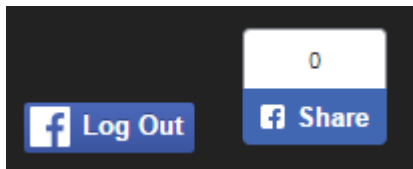
Joel Suomalainen 0438633



Työn tarkoituksena oli luoda moderneja web-tekniikoita hyödyntäen web-sovellus, joka keskittyy datan visualisoimiseen graafeiksi. Aiheen valitsin omien intressieni data-analyysiä kohtaan pohjalta ja taka-ajatuksena oli luoda tässä harkkatyön puitteissa hyvä pohja sivustolle, jota olisi helppo sitten päivitellä ja ehostaa myöhemmin. Ajattelin, että näpsäkkä sivu, jossa on hyvä pohja esitellä omia tuottamia graafeja sekä demonstroida web-kehitystaitojaan kiinnostaisi myös työnantajia vaikkapa kesätöitä hakiessa.

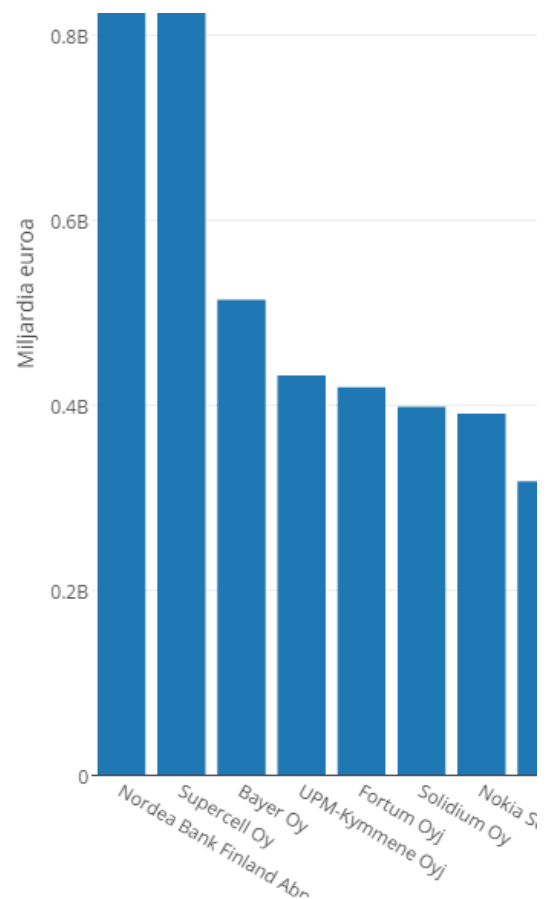
Tekniseltä toteutukselta työn pohjan muodostaa PHP:llä toteutettu backend, joka pyörii Digital Oceanin pilvipalvelussa hostatulla palvelimella LAMP-stackilla. LAMP-stack tarkoittaa siis Linuxilla pyörivää Apache-palvelinta, MySQL-tietokannan käyttöä sekä PHP:tä ohjelmointikielenä. PHP:n lisäksi työssä on hyödynnetty Javascriptiä esimerkiksi graafeja piirtävän plotly.js:n upottamisessa sivustolle, sekä Facebook-ominaisuuksien implementoinnissa.

Sivuston ulkoasun pohjana toimii Bootstrap ja sen päälle rakennetun templatien laajentaminen. Tämä mahdollistaa tyylikkään yksisivuisen (login- ja rekisteröitymissivuja lukuunottamatta) ulkoasun sekä reaktiivisuuden melko yksinkertaisesti. Bootstrapia varten palvelimenkin sielunelämään tuli tutustuttua paremmin sitä ja sen tarvitsemia packet managereita asennellessa.



Ensinnäkin sivulla on navigaatiopalkki, joka ohjailee eri kohtaan sivua, josta löytyy erilaisia graafeja. Navigaatiopalkin yhteydessä ovat mahdollisuudet kirjautua sisälle sivulle, joko luomalla uudet sivukohtaiset tunnukset tai sitten vaihtoehtoisesti Facebook-kirjautumista hyödyntämällä. Kirjautumisen lisäksi Facebook-ominaisuuksista sivulle on integroitu mahdollisuus jakaa sivusto Facebook-seinälleen, onhan sen verran hieno tekele kyseessä! Sivun oma tunnusten luominen toimii yksinkertaisen rekisteröitymissivun kautta, jossa luodaan käyttäjätunnus sekä valitaan sille salasana. Salasanan oikeinkirjoitus vielä varmistetaan käyttäjältä pyytämää toistamaan se. Selkotekstistä salasanaa ei tietenkään tallenneta tietokantaan sellaisenaan, vaan salasana hashataan `password_hash()`-funktion läpi, joka myös lisää turvallisuuden lisäämiseksi siihen suolauksen. Tietokantakutsuissa on myös huomioitu tietoturvaa käyttämällä `bindParamia`, sen sijaan että lähetettäisiin suoraan käyttäjän syötteet. Syynä tähän on SQL-injektion estäminen, eli että käyttäjä saisi ovelilla syötteillä vastauksena tietokannasta tietoa, johon hänellä ei ole oikeutta eikä pääsyä. Salasanan tarkistuksessa käytetään puolestaan `password_verify()`-funktioita, eli tietoturvan puolesta ainakin selkeimmät best practiset on huomioitu.

Graafien piirtämiseen applikaatiossa on käytetty `plotly.js`-kirjastoa, joka tarjoaa melko simppelellä tavalla upottaa aika tyylikkäitä datavisualisaatioita sivustolle. Tässä työssä on käytetty pelkästään Javascript-ominaisuuksia, mutta tulevaisuudessa tutustuminen tähän tarjoaisi hyvän mahdollisuuden hyödyntää samaa työkalua myös yhteistyössä Pythonin tai R-kielen kanssa tehtyjen datamanipulaatioiden visualisaatioon. Esimerkin omaisena graafina on käytetty pylväsdiaagrammia, joka on muodostettu [https://www.vero.fi/fi-FI/Avoim_data\(29357\)](https://www.vero.fi/fi-FI/Avoim_data(29357)) osoitteesta löytyvästä avoimesta datasta vuoden 2014 Yhteisöjen tuloverotuksen julkisista tiedoista listauksen top 20 eniten veroja maksaneista yrityksistä.



Ensimmäiseen viivadiagrammiin puolestaan käyttäjä voi syöttää arvoja, jotka sitten piirtyvät Ajaxin avulla graafiin ilman koko sivun päivitystä. Grafiikat ovat SVG-muodossa ja tarjoavat plotlyn myötä valmiiksi helppoja työkaluja datan tarkasteluun.

Työn päällepäin näkymättömissä osissa on käytetty esimerkiksi SEO-menetelmiä eli hakukoneoptimointia. Tätä on tehty muodostamalla urllist.txt (lista sivuston sivuista), sitemap.xml (xml-tiedosto sivuista sisältäen infoa esimerkiksi muokkauspäivistä), robots.txt (web-crawlereita ja botteja varten, jotka tekevät mm. indeksointia) sekä hyödyntämällä metadataan syötettyjä sivuston kuvauksia ja avainsanaluetteloa.

Lähdekoodipaketista on sensuroitu app-id:t pois sekä tietokantayhteydet. Tarvittavia kirjastoja koodin ajamiseen ovat Bootstrap, plotly.js, jQuery.

Ominaisuus reaaliaikaiseen datan plottaamiseen jäi ajanpuutteen takia toteuttamatta, mutta se on ehdottomasti seuraavana listalla, sillä se on varmasti tällaisella sivulla se käyttäjää eniten kiinnostava ominaisuus. Myöskin käyttäjäystävällisyyden lisääminen on ehdottomasti listalla, sillä tällä hetkellä esimerkiksi error-viestejä tulee aika harvakseltaan. Esimerkiksi epäonnistuneesta login-yrityksestä viestii oikeastaan vain se ettei käyttäjä pääse kirjautumaan sisälle vaan ohjautuu takaisin kirjautumissivulle.

Ominaisuuslistaus:

Responsiivinen ulkoasu	5p
Tietokannan käyttö tietojen tallennukseen	5p
Käyttäjän autentikointi	5p
SEO	3p
Kolmannen osapuolen palvelu (FB)	5p
Toinen kolmannen osapuolen palvelu (plotly.js)	3p
jQueryn käyttö	3p
Dynaamisen SVG-grafiikan käyttö (plotly.js)	3p
Kattava dokumentaatio	5p
Hyvä vertaisarviointi (tavoite)	5p

Yhteensä

42p

Tavoite harjoitustyöstä siis 40p