

Anton Pyrrö TTK20SD

CSC – Tieteen tietotekniikan keskus Oy:n sisäisen palveludokumentaation migraation suunnittelu

Tradenomi (AMK)

Tietojenkäsittely

Syky 2023



**KAMK • University
of Applied Sciences**

Tiivistelmä

Tekijä(t): Pyrrö Anton

Työn nimi: CSC – Tieteen tietotekniikan keskus Oy:n sisäisen palveludokumentaation migraation suunnittelu

Tutkintonimike: Tradenomi (AMK), Tietojenkäsittely

Asiasanat: migraatio, dokumentointi, ohjelmistot

Opinnäytetyön aihe oli CSC:n sisäisen palveludokumentaation migraatioon käytettävien työkalujen kartoittaminen ja migraation suunnittelu. Tarve tälle selvitystyölle tuli työharjoittelun aikana CSC:llä. Lähtökohdana oli etsiä ja verrata migraatioon sopivia työkaluja ja ohjelmistoja. Tiedonsiirtoa suunniteltiin Confluencesta GitLabiin, jotta pystyttäisiin vähentämään tiedonsäilytyspaikkoja ja tehostamaan työn tekoa.

Työssä keskityttiin työkalujen kartoittamiseen. CSC:llä oli asetetut kriteerit migraatioon käytettävälle työkalulle. Niitä olivat sisäverkossa tiedon säilyminen, Confluencen sivustorakenteen ja käyttöoikeuksien säilyminen, hyperlinkkien säilyminen ja mahdollisimman monen kosmeettisen ominaisuuden ehjänä pysyminen. Kriteerien nimeämisen jälkeen selvitettiin mahdollisuuksia ja päädyttiin kahteen vaihtoehtoon. Toinen näistä oli Pandoc-niminen ohjelma ja toinen erityisesti Confluencesta Gitlabiin siirtoa varten tehty skripti, joka on julkaistu GitHubiin.

Keskeisenä tuloksena oli migraatiotyökalujen löytäminen ja soveltuvuuden pohdinta. Opinnäytetyön perusteella päädyttiin suositteluun Pandoc-ohjelmaa ensisijaisesti. Toissijaiseksi vaihtoehdoksi jätettiin siirtoa varten tehty skripti. Molempia on tarpeen vielä kokeilla ennen lopullista päätöksentekoa. Testaaminen sekä itse siirto rajattiin työn ulkopuolelle ajan puutteen takia.

Johtopäätöksenä todettiin, että tarvitaan lisää tutkimusta ja käytännön kokeilua, mutta työllä luotiin pohja keväällä 2024 tapahtuvalle migraatiolle. Migraation jälkeen palveludokumentaatio käydään läpi ja sisältöä siistitään ja järjestellään niin, että varmistetaan tiedon eheys ja käytettävyys. Migraatio ei ole välttämättä helppo toteuttaa, joten toimeksiantajan tulee varautua ratkaisemaan mahdollisia siirtoon liittyviä haasteita. Työssä opittiin tekemisen aikana raportin kirjoittamista, tietoturvaa, verkostoitumista, ajanhallintaa ja suunnitelman luomista.

Abstract

Author(s): Pyrrö Anton

Title of the Publication: Planning of Documentation Migration for CSC Oy

Degree Title: Bachelor of Business Administration, Business Information Technology

Keywords: migration, documentation, software

The topic of the thesis was to map tools for the migration of internal service documentation at CSC Tieteen tietotekniikan keskus Oy and to plan the migration. The need for this study arose during an internship at CSC. The starting point was to search for and compare tools and software suitable for migration. Data transfer was planned from Confluence to GitLab to reduce data storage locations and streamline work.

The focus of the work was on mapping the tools. CSC had set criteria for the tool to be used for migration. These criteria included the preservation of information within the internal network, the preservation of Confluence's site structure and permissions, the preservation of hyperlinks, and the preservation of as many cosmetic features as possible. After naming the criteria, possibilities were explored, leading to the identification of two options. One was a program called Pandoc, and the other was a script specifically made for the transfer from Confluence to GitLab, which had been published on GitHub.

The key result was the identification of migration tools and the consideration of their suitability. Based on the thesis, Pandoc was recommended as the primary tool, with the script for transfer left as a secondary option. Both options need to be tested before making a final decision. Testing and the actual transfer were excluded from the scope of the thesis due to time constraints.

In conclusion, it was noted that more research and practical experimentation are needed, but the work laid the foundation for the migration planned for spring 2024. After the migration, the service documentation will be reviewed, and the content will be cleaned up and organized to ensure the integrity and usability of the information. Migration may not necessarily be easy to implement, so the client should be prepared to address potential challenges related to the transfer. The work involved learning report writing, cybersecurity, networking, time management, and plan creation during the process.

Sisällys

1	Johdanto	1
2	Migraatio	2
2.1	Datamigraatio	2
2.2	Ohjelmistomigraatio.....	2
2.3	Pilvimigraatio	3
2.4	Järjestelmämigraatio.....	3
3	Datamigraatio Confluencesta ja GitLabiin	4
3.1	Confluence	4
3.2	GitLab.....	5
4	CSC:n rajaukset/tavoitteet järjestelmälle.....	6
5	Vertailu ja arviointi.....	8
5.1	Pandoc	8
5.2	Confluence-to-GitHub	10
6	Yhteenveto.....	11
7	Päätäntö	12
8	Lähteet.....	13

1 Johdanto

Opinnäytetyön aihe on CSC:n sisäisen palveludokumentaation migraatioon käytettävien työkalujen kartoittaminen ja migraation suunnittelu. Suunnittelussa kartoitetaan migraatioon käytettävät työkalut, täyttävätkö ne CSC:n asettamat kriteerit ja miten migraatio kannattaa toteuttaa.

Työn ulkopuolelle on rajattu testaaminen ja alustojen kokeilu sekä siirto. Tarkoituksena on palveludokumentaation siirron suunnittelu vanhalta alustalta Confluencesta uudelle alustalle GitLabiin, jotta tiedonsäilytyspaikat vähenevät ja sitä kautta työ tehostuu uusilla ja vanhoilla työntekijöillä.

Tavoitteena on tehdä opinnäytetyö, jonka pohjalta keväällä 2024 tehtävä palveludokumentaation migraatio voidaan toteuttaa ilman jatkotutkimusta. Tämä tehdään tutkimalla migraatioon soveltuvat työkalut, migraatioon kuuluvat hyvät tavat ja huomioon otettavat asiat.

Migraation jälkeen palveludokumentaatio käydään läpi ja sisältöä siistitään ja järjestellään niin, ettei tieto toistu monessa paikassa. Laajasti tarkasteltuna aihe liittyy tietojohdantamiseen.

Aihe valikoitui tehtäväksi työharjoittelun aikana, koska se on ajankohtainen ja relevantti CSC:lle. Aiheessa kiinnostaa mahdollisuus selvittää migraatiota ja sen eri muotoja töiden ohella ja päästä jatkossa soveltamaan selvitettyä käytäntöön. Opinnäytetyö toimii ohjeena keväällä 2024 migraation toteuttamiseen.

2 Migraatio

Migraatio tarkoittaa tässä tapauksessa tietoteknistä migraatiota eli tiedon siirtämistä yhdestä järjestelmästä toiseen. Yleisin syy migraatioon on järjestelmän päivitys tai varmuuskopiointi. Migraatio voi tarkoittaa yhtä tai useammanlaista tiedon migraatiota, joita ovat datamigraatio (data migration), ohjelmistomigraatio (application migration), pilvipalvelunmigraatio (cloud migration) ja käyttöjärjestelmänmigraatio (operating system migration). [1.] Seuraavaksi kerrotaan tarkemmin erilaisista migraatiolajeista.

2.1 Datamigraatio

Datamigraatio on prosessi, jossa tietoa siirretään kahden eri tallennustilan, formaatin tai järjestelmän välillä. Tämä voi liittyä esimerkiksi ohjelmistojen päivityksiin, palvelinkeskuksien siirtämiseen, IT-infrastruktuurin ylläpitoon ja yritysten yhdistymiseen. Datamigraatio on kriittinen vaihe monissa IT-migraatioprojekteissa, ja sen on oltava huolellisesti suunniteltu, koska olemassa oleva data on usein yritykselle kriittistä. Datamigraation aikana on otettava huomioon mahdolliset ongelmat, kuten häiriöaika toiminnassa, datan mahdollinen korruptoituminen, menetetty tai huomaamatta jäänyt data sekä yhteensopivuusongelmat uuden ympäristön kanssa. [2.]

2.2 Ohjelmistomigraatio

Ohjelmistomigraatio tarkoittaa sovelluksen siirtämistä yhdestä IT-ympäristöstä toiseen. Ohjelmiston siirtämisessä on otettava huomioon useita tekijöitä. Tekniset vaatimukset ja yhteensopivuus määritellään yleensä jo suunnittelussa. Ne voivat sisältää muun muassa, pitääkö sovelluksen olla avoimesta lähdekoodista ja saako se siirtää tietoa ulos sisäisestä verkosta. Yhteensopivuudella tarkoitetaan, että ohjelmisto toimii uudessa ja vanhassa järjestelmässä. Organisaation tietoturva on tietysti myös siirtoa määrittävä tärkeä näkökulma. Ohjelmistomigraation vaiheet voivat vaihdella tapauskohtaisesti, ja ne voivat sisältää esimerkiksi sovelluksen siirtämisen pilvipalveluun, sovelluksen lähdekoodin muuttamista toiminnallisuus säilyttäen tai uuden sovelluksen käyttöönoton. [3.]

2.3 Pilvimigraatio

Pilvimigraatio on prosessi, jossa data tai ohjelmisto siirretään yrityksen järjestelmästä pilviympäristöön. Pilvimigraation tavoitteena on hyödyntää pilvipalveluiden tarjoamia etuja, kuten turvallisuutta ja kustannustehokkuutta. Pilvimigraation suunnittelussa on tärkeää määrittää, mitä tietoa tai sovelluksia siirretään pilveen ja kuinka siirto suoritetaan. Yrityksen on harkittava huolellisesti, onko pilvimigraatio tarpeellista ja kannattavaa. [4.]

2.4 Järjestelmämigraatio

Järjestelmämigraatio tarkoittaa IT-järjestelmän siirtämistä käyttöjärjestelmästä toiseen. Tämä voi sisältää esimerkiksi käyttöjärjestelmän päivittämisen uudempaan versioon tai jopa koko käyttöjärjestelmän vaihtamisen. Järjestelmämigraatiot ovat aikaa ja resursseja kuluttavia projekteja, ja ne voivat aiheuttaa katkoja päivittäisessä työnteossa. On tärkeää suunnitella ja valmistautua huolellisesti ennen järjestelmämigraatiota ja arvioida riskit sekä hyödyt. [1.]

Migraatioon sisältyy yleisesti monta liikkuvaa osaa ja tiukkoja vaatimuksia, jotka määritellään ja kartoitetaan suunnitteluvaihteessa. Näitä voivat olla tietoturvan ylläpito, tunnistautumismenetelmien säilytys, mahdolliset hajoamiskohdat ja miten niistä palaudutaan, milloin vanha järjestelmä otetaan pois käytöstä. Tarkka suunnittelu ja automaatio tekevät migraatiosta helpompaa. [1.]

3 Datamigraatio Confluencesta ja GitLabiin

Tässä opinnäytetyössä pääaiheena on datamigraatio. Migraatioon kuuluu CSC:n palveludokumentaatiossa olevan tiedon vieminen vanhasta järjestelmästä eli Confluencesta uuteen järjestelmään eli GitLabiin. Tiedon migraatiossa on kolme tärkeää vaihetta, jotka ovat siirron suunnittelu, toteutus ja validointi [5].

Suunnittelussa päätetään, mitä siirretään ja kuinka paljon pitää siirtää. Minkälaista tietoa se on, minkä laatuista ja missä kunnossa? Kun tiedon määrä, laatu ja kunto on tarkistettu, suunnitellaan, missä määrissä se siirretään. Siirretäänkö kaikki kerrallaan, pienissä erissä vai mitä tahansa siltä väliltä? Suunnittelussa määritellään myös, mitä kalustoa ja ohjelmistoa tiedon siirrossa käytetään. [5.]

Suunnittelun jälkeen tehdään itse migraatio. Siirron aikana valvotaan, että migraatio tapahtuu ilman ongelmia ja tiedon laadun muuttumista. Tiedon siirtyessä valvotaan myös, ettei tieto monistu eli kopioidu useamman kuin halutun kerran. Eli jos halutaan jonkin tiedoston kopioituvan vain kerran, niin tarkkaillaan, että näin myös käy. Siirtoa valvotaan edellä mainittujen asioiden takia sellaisessa ympäristössä, joka peilaa oikeaa työympäristöä ja mahdollistaa migraatiota toteuttavan tiimin nopean havainnoinnin ja reaktion, jos jotain ei-toivottua tapahtuu. [5.]

Tiedon siirron jälkeen tieto tarkistetaan ja varmistetaan uudessa järjestelmässä, ettei siirron aikana tapahtunut mitään ei-toivottua. Vanha tietojärjestelmä otetaan pois käytöstä vasta sitten, kun käyttäjät ovat varmistaneet uudelle alustalle siirretyn tiedon toimivaksi. [5.]

3.1 Confluence

Confluence on Atlassianin kehittämä tuote. Atlassianilla on töissä yli 10 000 työntekijää, joita sijaitsee 13 eri maassa. Atlassianin perusti vuonna 2002 Mike Cannon-Brookes ja Scott Farquhar. Atlassian erikoistuu tiimityöskentelyyn tarkoitettujen sovellusten tuottamiseen. [6.]

Confluence on tiimityötila, jossa tiimit voivat luoda, tallentaa ja tehdä yhteistyötä projekteissa. Työtilat auttavat organisoimaan ja jakamaan työtä. Confluence soveltuu kaikenkokoisille ja -tyyppisille tiimeille, olivatpa kyseessä kriittiset projektihankkeet tai tiimikulttuurin rakentaminen avoimella tavalla. [7.]

3.2 GitLab

GitLab on avoimen lähdekoodin koodivarasto ja yhteistyöhön perustuva ohjelmistokehitysalusta suurille DevOps- ja DevSecOps-projekteille ja on maksuton yksittäiskäyttäjille. GitLabissa on yli 30 miljoonaa rekisteröitynyttä käyttäjää ja GitLabia kehittävässä tiimissä on yli 2100 kehittäjää, joita sijaitsee yli 65 eri maassa. [8.]

GitLab tarjoaa paikan verkkokoodin säilyttämiseen ja mahdollistaa tehtävien seurannan sekä CI/CD-toiminnot. Tietovarasto mahdollistaa erilaisten kehitysketjujen ja versioiden hallinnoinnin ja antaa käyttäjille mahdollisuuden tarkastella aikaisempaa koodia ja palata siihen, jos ilmenee odottamattomia ongelmia. [9.]

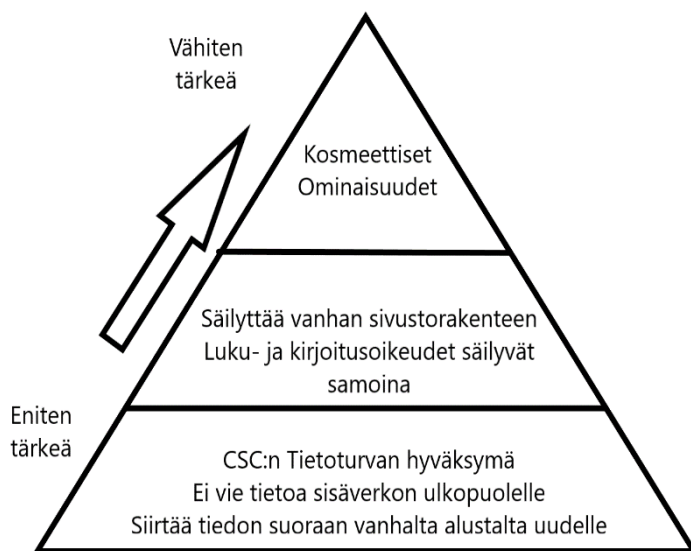
4 CSC:n rajaukset/tavoitteet järjestelmälle

Toimeksiantajan kanssa käytiin läpi migraation reunaehdoja. Käytiin läpi esimerkiksi teknisiä, yhteensopivuuden sekä tietoturvan vaatimuksia. Vaatimuksia olivat:

- Sovelluksen täytyy pystyä siirtämään tieto suoraan Confluencesta GitLabiin kulkematta kolmannen sovelluksen kautta. Tämä tarkoittaa, että käytettävä sovellus osaa hakea tiedon lähteestä ja integroida sen suoraan päämäärään tallentamatta sitä missään ulkoisessa tallennuspisteessä.
- Aiempaan liittyen sovellus ei myöskään saa siirtää tietoa sisäisestä verkosta ulkoverkkoon. Tämä tarkoittaa, että tiedon pitää kulkea vain turvallisesti yrityksen itse pystyttämässä verkossa ja sitä ei lähetetä kolmannelle osapuolelle tietoturvasyistä.
- Confluencessa olevan sivustorakenteen pitää pysyä mahdollisimman samanlaisena GitLabissa, että vältetään käsin työstämisestä. Mahdollista on, että jotkin rakenteet voivat muuttua, mutta tämä pyritään minimoimaan.
- Tavoitteena on saada Confluencesessa asetetut käyttöoikeudet, eli kenellä on lupa vain lukea ja kenellä myös oikeus muokata joitain sivuja, siirtymään GitLabiin. GitLabissa samoilla henkilöillä, joilla oli Confluencessa oikeus lukea ja kirjoittaa tiettyjä sivuja tai vain lukea eikä kirjoittaa, pitäisi myös pysyä oikeus samoihin toimintoihin GitLabissa.
- Hyperlinkkien pysyminen samana ja siirtyminen GitLabiin olisi toivottavaa, mutta sen ei oleteta toimivan ongelmitta. Tässä vaiheessa voi epäonnistua se, että Confluencessa olevat hyperlinkit, jotka ohjaavat muualle Confluencen sivulle päivittyisivät niin, että ne ohjaavat GitLabissa linkkiä painaessa takaisin Confluenceseen.
- Confluencessa on mahdollista muokata sivuille kosmeettisia elementtejä, joita ovat esimerkiksi korostukseen tarkoitettu laatikko, jonka sisälle voi laittaa tekstiä jonkin asian korostamiseksi. Mahdollista on myös piirtää erilaisia kuvia tai tehdä taulukoita. Tavoitteena on, että nämä elementit siirtyisivät sellaisinaan myös GitLabiin, mutta se ei ole pakollinen ominaisuus.

- Käytettävän migraatio-ohjelman täytyy olla CSC:n sisäisen tietoturvan hyväksymä. Järjestelmä ei myöskään saa olla uusi eli sen pitää olla jo olemassa oleva eikä uutta tehdä tätä tarkoitusta varten.

Kuvassa 1 on koostettu edellä mainitut vaatimukset. Alla on tärkeimmät ominaisuudet, joiden pitää toimia ja ylöspäin liikkuessa painoarvo vähenee.



Kuva 1. CSC:n rajauksien ja tavoitteiden painoarvo

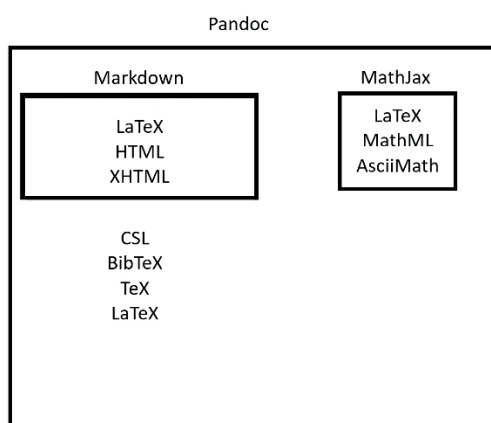
5 Vertailu ja arviointi

Migraatioon sopivia ohjelmia löytyi yksi kappale ja yksi skripti. Ensimmäisenä migraatiota toteuttaessa kokeillaan Pandoc-nimistä ohjelmaa ja jos tämä ei toimi tai ei täytä tietoturva-vaatimuksia, kokeillaan skriptiä.

5.1 Pandoc

Pandoc on tehokas avoimen lähdekoodin tekstinkäsittelytyökalu, joita ylläpitävät vapaaehtoiset ihmiset omalla ajallaan. Pandoc ymmärtää monia hyödyllisiä Markdownin syntaksin laajennuksia. Näihin kuuluvat muun muassa asiakirjan metatiedot, alaviitteet, taulukot ja määrittelyluettelot. Lisäksi Pandoc tukee ylä- ja alaindeksejä, yliviivausta, parannettuja järjestettyjä listoja, juoksevia esimerkkilistoja sekä rajattuja koodilohkoja syntaksin korostuksella. [10.]

Alla olevassa kuvassa näytetään eri ohjelmistot ja niiden relaatio toisiinsa. Kuvassa sisimmissä neliöissä ovat alimmat ohjelmistot ja niiden yläpuolella, mihin ne liittyvät. Kuvasta huomaa, että samoja ohjelmistoja pystyy käyttämään ilman välissä olevaa ohjelmistoa. Esimerkiksi Pandoc tukee suoraan LaTeXia, mutta Pandocin tukema ohjelmisto Markdown tukee sitä myös.



Kuva 2. Pandocin tukemat ohjelmistot ja niiden hierarkia.

Markdown on CSC:llä eniten käytetty Confluencen ominaisuus, jonka onnistunut siirto on tärkeää. Muut työkalut tai ohjelmat ja niiden käyttömäärä on epäselvää tässä vaiheessa ja niiden selvitykseen ei tätä työtä kirjoittaessa ole aikaa.

Markdown on tekstistä HTML-muotoon muuntava työkalu verkkokirjoittajille. Markdown mahdollistaa kirjoittamisen helposti luettavassa ja kirjoitettavassa pelkässä tekstistä koostuvassa muodossa, jonka voi sitten muuntaa rakenteelliseksi XHTML:ksi tai HTML:ksi. [11.]

Markdownin muotoilusyntaksin ohjaava suunnittelutavoite on tehdä siitä mahdollisimman luettavaa. Ideana on, että Markdowilla tehty asiakirja tulisi voida julkaista sellaisenaan, pelkkänä tekstinä, ilman että se näyttää tageilla tai muotoiluohjeilla merkityltä. Vaikka Markdownin syntaksiin on vaikuttanut useat olemassa olevat tekstistä HTML:ään muuntavat suodattimet, suurin yksittäinen inspiraation lähde Markdownin syntaksille on pelkästä tekstistä koostuva sähköpostin formaatti. [11.]

Pandoc osaa käsitellä älykkäitä lainausmerkkejä, viivoja, ellipseja ja mahdollistaa Markdownin HTML-lohkojen sisäisen käytön sekä sisäisen LaTeXin. LaTeX-matematiikka ja -makrot ovat käytävissä Markdown-dokumenteissa, ja Pandoc tarjoaa useita tapoja muuntaa matematiikka HTML-muotoon, kuten MathJaxin ja MathML:n avulla.

LaTeX on ladontajärjestelmä, joka on suunniteltu erityisesti teknisen ja tieteellisen dokumentaation tuottamiseen. Se tarjoaa monipuolisia ominaisuuksia, jotka mahdollistavat tarkan ja ammattimaisen dokumenttien muotoilun. LaTeX on vakiintunut de facto -standardiksi tieteellisten asiakirjojen viestinnässä ja julkaisussa, ja se on tunnustettu sen luotettavuudesta ja tehokkuudesta. [12.]

LaTeX on saatavilla ilmaisena ohjelmistona, mikä tekee siitä helposti saatavilla olevan työkalun tutkijoille, kirjoittajille ja muille ammattilaisille. Sen avulla käyttäjät voivat tuottaa laadukasta ja yhtenäistä dokumentaatiota ilman suurta taloudellista investointia. Tämä tekee siitä hyvän vaihtoehdon erilaisille tieteellisille ja teknisille yhteisöille, jotka tarvitsevat tehokasta tapaa tuottaa ja jakaa tietoa. [12.]

MathJax on avoimen lähdekoodin JavaScript-näyttömoottori, joka mahdollistaa LaTeX-, MathML- ja AsciiMath-merkintöjen visuaalisen esittämisen kaikissa nykyaikaisissa selaimissa. Lisäksi se on varustettu integroidulla tuella apuvälineteknologioille, kuten ruudunlukuohjelmille, mikä parantaa käytettävyyttä ja saavutettavuutta. [13.]

MathML:n tarkoituksena on helpottaa matemaattisen ja tieteellisen sisällön käyttöä ja uudelleen käyttöä verkossa sekä muissa sovelluksissa, kuten tietokoneavusteisissa algebralaskentajärjestelmissä, painettavassa ladonnassa ja puheen synteessissä. MathML mahdollistaa matemaattisen

merkinnän sekä visuaalisen esityksen korkealaatuisen näyttämisen että matemaattisen sisällön koodauksen sovelluksille, joissa semantiikalla on keskeinen rooli, kuten tieteellisessä ohjelmistossa tai puheen synteesissä. [14.]

Pandocissa on tehokas järjestelmä automaattisiin viittauksiin ja bibliografioihin. Tämä mahdollistaa viittausten kirjoittamisen yksinkertaisesti ja muuntaa ne asiakirjan loppuun muotoilluiksi viittauksiksi ja bibliografiaksi satojen CSL-tyylien avulla. Bibliografiset tiedot voivat olla BibTeXissä, BibLaTeXissa, CSL JSONissa tai CSL YAMLIssa, ja viittaukset toimivat kaikissa tulostusmuodoissa.

BibTeX on viitehallintaohjelma, jota käytetään yhdessä TeX:n ja LaTeX:n kanssa. BibTeX helpottaa edellä mainittujen ohjelmien kanssa kirjojen ja lehtiartikkeleiden bibliografian luomista. [15.]

Citation Style Language (CSL) on XML-pohjainen kieli, joka määrittelee viittausten ja kirjallisuusluettelon muodot. CSL:ssä on monia eri tyyliä, mukaan lukien laajasti käytetyt APA, Chicago ja Vancouver. Citation-Style-Language-paketti on suunniteltu tarjoamaan vaihtoehtoinen tapa muotoilla viitteitä LaTeXille, hyödyntäen CSL-tyylejä. [16.]

Pandocia voi mukauttaa monin tavoin, mukaan lukien mallijärjestelmä ja suodatinten kirjoitusjärjestelmä. Ohjelma sisältää sekä Haskell-kirjaston että itsenäisen komentoriviohjelman, ja uuden sisääntulo- tai tulostusmuodon lisääminen vaatii vain uuden moduulin lisäämisen.

5.2 Confluence-to-GitHub

Toinen löydetystä vaihtoehdoista on käyttäjän “sjones4” ja “xtaran” tekemä skripti, joka on julkaistu GitHubiin. Koska skriptin sivuilla on ohjeet suoritukseen ilman kuvia ja testaukseen mahdollisen ajan puutteen takia, on vaikea todeta, olisiko tämä vaihtoehto nopeampi vai jääkö siinä aiemmin mainittuja vaatimuksia täyttämättä.

Tällä vaihtoehdolla migraation toteuttaminen vaatii tietämystä skriptaamisesta, Gitistä ja tarvitsee lisäosan asennuksen migraatiota toteuttavalle koneelle, joka vaatii lisää tarkistuksia tietoturvalta. Tietämys kuitenkin kummastakin asiasta löytyy jo CSC:ltä, joten tämä vaihtoehto, jos se on todettu turvalliseksi, pitäisi onnistua kuitenkin suorittamaan.

6 Yhteenveto

Arvioinnin jälkeen ensimmäisenä kokeillaan Pandoc-sovellusta rajallisten vaihtoehtojen takia. Pandoc täyttää aiemmista vaatimuksista alustavasti kaikki, paitsi CSC:n sisäisen tietoturvan tarkistuksen.

Kokeilu toteutetaan keväällä 2024 ja samalla todetaan, täyttääkö se käytännössä kaikki kriteerit ja hyväksyykö CSC:n tietoturva sovelluksen. Jos kokeilu on onnistunut, aloitetaan varsinaisen migraation loppusuunnittelu ja toteutus. Jos toteutus ei onnistu Pandoc-sovelluksella, kokeillaan toista vaihtoehtoa ja toteutetaan edellä mainittu prosessi senkin suhteen.

Jos migraatio ei toteudu kummallakaan sovelluksella, aloitetaan uudelleen suunnittelu ja kartoitus. Tämä suoritetaan erikseen CSC:n sisällä.

7 Päätäntö

Opinnäytetyön keskeisenä tuloksena todettiin, että vaihtoehtoja oli rajallisesti. Onnistuttiin kuitenkin kartoittamaan kaksi vaihtoehtoa, joista toinen täyttää alustavasti hyvin kriteerejä ja siksi sitä aiotaan kokeilla ensimmäisenä migraation toteutustyökaluna. Kartoituksessa apuna olivat CSC:n henkilökunta ja ammattilaiset, verkossa ollut dokumentaatio ja verkosta löytyneet suositukset.

Vaihtoehtoiksi valittiin Pandoc-niminen universaali dokumentinmuuntaja ja GitHubiin julkaistu skripti, joka on kahden henkilön juuri tätä tilannetta varten luoma. Näistä kahdesta päädyttiin Pandoc-ohjelmaan, joka vaikuttaa tukevan laajemmin asetettuja vaatimuksia, joita olivat sisäverkossa tiedon säilyminen, Confluencen sivustorakenteen ja käyttöoikeuksien säilyminen, hyperlinkkien säilyminen ja mahdollisimman monen kosmeettisen ominaisuuden ehjänä pysyminen. Ainoa vaatimus, jota se ei ole täyttänyt, on CSC:n sisäisen tietoturvan hyväksyminen, joka toteutetaan migraatiota ennen keväällä 2024.

Opinnäytetyön tekemisen aikana opittiin, että varataan enemmän aikaa aiheen tutkimiselle ja vaatimukseen perehtymiselle. Tässä opinnäytetyössä lisäajalla olisi voinut esimerkiksi tutkia migraatioon käytettäviä työkaluja tarkemmin käytännön kokeilun kautta ja todeta, täyttäisivätkö ne kriteerit paremmin. Lisäksi tekemisen aikana opittiin raportin kirjoittamisesta, tietoturvasta, verkostoitumisesta, ajanhallinnasta ja suunnitelman luonnista.

Opinnäytetyötä seuraa keväällä 2024 suunnitelman toteutus CSC:llä käyttämällä näitä ohjeita. Siinä laitetaan opinnäytetyössä kirjoitettu teoria ja suunnittelu käytäntöön ja todetaan samalla, onko esimerkiksi migraatioon tutkitut vaihtoehdot sopivia vai tarvitaanko jatkotutkimusta ennen kuin migraatio voidaan suorittaa loppuun asti.

Onnistuneen migraation jälkeen palveludokumentaatio käydään läpi ja kirjoitetaan tarpeen mukaan uusiksi, että käytetään samoista asioista samoja termejä, linkataan tarpeen mukaan sivuja, joista löytyy jatkotietoa ja tiivistetään tekstiä.

8 Lähteet

1. RedHat. 2021. RedHat, What is IT migration? Viitattu 11.10.2023. Saatavissa: <https://www.redhat.com/en/topics/automation/what-is-it-migration>
2. Lelii, S. 2017. TechTarget, Data Migration. Viitattu 30.10.2023. Saatavissa: <https://www.techtarget.com/searchstorage/definition/data-migration>
3. IBM. 2019. IBM, Application migration. Viitattu 30.10.2023. Saatavissa: <https://www.ibm.com/cloud/learn/application-migration>
4. Semilof, M. 2021. TechTarget, What is cloud migration? An introduction to moving to the cloud. Viitattu 30.10.2023. Saatavissa: <https://www.techtarget.com/searchcloudcomputing/definition/cloud-migration>
5. Azure. What is data migration? Viitattu 11.10.2023. Saatavissa: <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-data-migration>
6. Atlassian, Company. Viitattu 27.11.2023. Saatavissa: <https://www.atlassian.com/company>
7. GitHub, Meet our team. Viitattu 27.11.2023. Saatavissa: <https://about.gitlab.com/company/team/>
8. Confluence basics, Atlassian. Viitattu 21.11.2023. Saatavissa: <https://www.atlassian.com/software/confluence/resources/guides/get-started/overview#about-confluence>
9. Definition Gitlab. TechTarget. Viitattu 21.11.2023. Saatavissa: <https://www.techtarget.com/whatis/definition/GitLab>
10. Pandoc a universal document converter. Viitattu 28.11.2023. Saatavissa:

<https://pandoc.org>

11. Daring Fireball, Markdown 2004. Viitattu 17.11.2023. Saatavissa:

<https://daringfireball.net/projects/markdown/>

12. LaTeX A document preparation system. Viitattu 19.11.2023. Saatavissa:

<https://www.latex-project.org>

13. MathJax Documentation. Viitattu 19.11.2023. Saatavissa:

<https://docs.mathjax.org/en/latest/>

14. What is MathML. Mathematical Markup Language (MathML). Viitattu 19.11.2023. Saatavissa:

<https://www.w3.org/Math/whatIsMathML.html>

15. Mikä on bibtex? 2023. Viitattu 17.11.2023. Saatavissa:

<https://fi.theastrologypage.com/bibtex>

16. Citation Style Language. Viitattu 19.11.2023. Saatavissa:

<https://citationstyles.org>