

samk



Satakunnan ammattikorkeakoulu  
Satakunta University of Applied Sciences

TUOMAS LAKKINEN

# **Datamigraation suunnittelu ja toteutus**

TIETOJENKÄSITTELYN TUTKINTO-OHJELMA  
2024

## TIIVISTELMÄ

Lakkinen, Tuomas: Datamigraation suunnittelu ja toteutus  
Opinnäytetyö, AMK  
Tietojenkäsittelyn tutkinto-ohjelma  
Huhtikuu 2024  
Sivumäärä: 32

Työn toimeksiantajana toimi Digia Finland Oy, joka on ohjelmisto- ja palveluyritys. Toimeksiantajalla oli tarve siirtää yrityksen sopimusarkisto vanhasta On-premise-varastosta uuteen pilvivarastoon, jotta voitiin ajaa vanha varasto alas.

Työn teoriaosuudessa on kuvattu datamigraation prosessit projektina, selitetty keskeiset käsitteet datamigraatiossa, esitelty datamigraation tyypit sekä tuotu esiin syitä, miksi datamigraatioita tehdään. Näiden lisäksi on myös käsitelty datamigraation eri vaiheita.

Toteutusvaihe aloitettiin suunnittelemalla datamigraation suoritustapa ja valitsemalla rakenne CSV-tiedostolle, jota datamigraatiossa tarvittiin. Tämän jälkeen data tuli valmistella, jotta voitiin viedä data lähdejärjestelmästä datan muokkausta varten. Valmistelun sekä muokkaamisen jälkeen aloitettiin datamigraation testausvaihe, jossa tuotiin muokattu data testiympäristöön toimeksiantajan työkalujen ja ohjeiden avulla. Testausvaiheen jälkeen tuli korjata kaikki ilmenneet virheet, jotta voitiin tuoda data virheettömästi kohdejärjestelmään.

Avainsanat: Datamigraatio, Microsoft Excel, M-Files, M-Files Importer

## ABSTRACT

Lakkinen, Tuomas: Data migration planning and implementation  
Bachelor's thesis  
Business Information Systems  
April 2024  
Number of pages: 32

The client of the work was Digia Finland Oy which is a software and service company. The client needed to move the company's contract archive from the old On-premise storage to the new cloud storage, so that the old storage could be brought down.

In the theoretical part of the work, the data migration processes were described as a project, key concepts in data migration were explained, all types of data migration were presented, and reasons why data migrations are carried out were highlighted. In addition to these, the different phases of data migration were also addressed.

The implementation phase began by planning the execution method of data migration and selecting the structure for the CSV file needed in the data migration. After this, the data needed to be prepared to export it from the source system for data modification. Following preparation and modification, the testing phase of data migration commenced, where the modified data was imported into the test environment using the tools and instructions provided by the client. After the testing phase, all identified errors had to be corrected to seamlessly import the data into the target system.

Keywords: Data migration, Microsoft Excel, M-Files, M-Files Importer

# SISÄLLYS

1 JOHDANTO .....	6
1.1 Toimeksiantaja .....	6
1.2 Työn tarkoitus ja tavoitteet.....	6
2 DATAMIGRAATIO JA SEN VAIHEET .....	7
2.1 Datamigraatio .....	7
2.2 Datamigraation vaiheet.....	8
2.2.1 Suunnittelu.....	10
2.2.2 Datan muokkaus sekä valmistelu .....	11
2.2.3 Datamigraation toteutus.....	12
2.2.4 Testaus.....	12
2.2.5 Jälkitoimet.....	13
3 TOTEUTUS.....	14
3.1 Taustatietoa.....	14
3.2 Suunnittelu .....	15
3.3 Datan valmistelu .....	17
3.3.1 Dokumenttien vienti (Export) .....	19
3.3.2 Datan muokkaaminen.....	20
3.4 Migraation suoritus testiympäristöön .....	23
3.5 Virheet.....	24
3.6 Importerin asetukset.....	26
3.6.1 M-Files Connection.....	26
3.6.2 File Format Options .....	27
3.6.3 Mappings .....	28
3.7 Migraation toteutus tuotantoympäristöön.....	29
3.8 Migraation viimeistely .....	30
4 LOPPUPÄÄTELMÄT .....	30
LÄHTEET.....	32

## SYMBOLI- JA LYHENNELUETTELO (EI PAKOLLINEN)

On-premise	Paikallinen järjestelmä, joka toimii yrityksen omassa IT-ympäristössä
NoSQL	Perinteisestä relaatiomallista poikkeava tietokanta
SQL	Kyselykieli, jolla relaatiotietokantaan voi tehdä erilaisia hakuja, muutoksia ja lisäyksiä
Database as a Service	Pilvipohjainen ohjelmistotyökalu, jota käytetään esimerkiksi datan hallintaan tietovarastossa
ETL	Englannin kielen sanoista Extract, Transform ja Load. Viittaa siis tietojen vientiin, muokkaukseen ja tuontiin.
CSV	Tiedostomuoto, jota käytetään taulukkomuotoisen datan tallentamiseen tekstimuodossa
XML	Merkintäkieli, joka mahdollistaa datan määrittelyn ja tallentamisen jaettavalla tavalla.
XLSX	Excel-tiedostomuoto, joka mahdollistaa monimutkaisempien taulukoiden ja kaavioiden laskentatiedon tallentamisen
Multifile	Voi viitata esimerkiksi tiedostopakettiin, joka koostuu useista eri tiedostoista

# 1 JOHDANTO

## 1.1 Toimeksiantaja

Opinnäytetyön toimeksiantajana toimi Digia Finland Oy, joka on ohjelmisto- ja palveluyritys. Yritys keskittyy yhdistämään teknologian mahdollisuudet ja ihmisten kyvyt älykkäämmän liiketoiminnan ja yhteiskunnan kehittämiseksi. Digia pyrkii rakentamaan kestävää tulevaisuutta hyödyntämällä teknologian tarjoamia mahdollisuuksia ja ihmisten osaamista. (Digia, n.d., kohta Digia yrityksenä.) Toimeksiantajalla oli tarve siirtää yrityksen sopimusarkisto vanhasta On-premise-varastosta uuteen pilvivarastoon.

Digia-nimi juontaa juurensa sanoista Digital Information Architects. Yrityksen historia ulottuu vuoteen 1990, jolloin se oli nimeltään SysOpen Oyj. Vuonna 2005 SysOpen Oyj yhdistyi Digia Oy:n kanssa, josta syntyi SysOpen Digia Oyj. Digia kuitenkin vaihtoi nimensä vuonna 2008 Digia Oyj:ksi. Vuonna 2016 Digia uudisti toimintaansa, ja yhtiö jakautui, minkä seurauksena Qt-liiketoiminta siirtyi Qt Group Oyj:lle. Näin ollen nykyisessä muodossaan Digia on toiminut vuodesta 2016 lähtien. (Digia, n.d., kohta Digian tarina.)

Pääkonttori sijaitsee Helsingissä, ja muut toimistot sijoittuvat Tampereelle, Jyväskylään, Turkuun, Joensuuhun, Raumalle, Ouluun, Vaasaan, Lahteen ja Kuopioon sekä Tukholmaan. Tytäryhtiöt toimivat Ruotsin ja Suomen lisäksi myös Alankomaissa. (Digia, n.d., kohta Globaalisti paikallinen.)

## 1.2 Työn tarkoitus ja tavoitteet

Opinnäytetyö toteutettiin toiminnallisena opinnäytetyönä ja sen toimeksiantajana toimi Digia Finland Oy. Opinnäytetyön tarkoitus oli luoda tuotos, joka käsittelee suorittamani datamigraation eri vaiheita sekä julkaista se Digian

Wikisivulla. Tuotos oli käytännössä opinnäytetyön toiminnallisen osan tiivistetty versio, sisältäen vähemmän yksityiskohtia. Tuotos säilytti kuitenkin pääpiirteissään saman sisällön ja viestin. Lisäksi tuotoksessa oli huomioitu luottamukselliset tiedot, joita ei voitu julkaista opinnäytetyössä. Tämä tarkoitti sitä, että joitakin tietoja tai yksityiskohtia oli jätetty pois tai niitä oli käsitelty yleisellä tasolla, suojellen samalla liikesalaisuuksia tai muita arkaluontoisia tietoja. Näin varmistettiin, että tuotos oli edelleen informatiivinen, mutta samalla kunnioitettiin ja säilytettiin yrityksen sisäisten tietojen luottamuksellisuutta.

Opinnäytetyön tavoite oli lisätä yleistä tietoisuutta datamigraatioprosessista ja sen merkityksestä liiketoiminnassa. Näin ollen opinnäytetyö voi auttaa muita organisaatioita ja ammattilaisia hahmottamaan, kuinka datamigraatio voidaan toteuttaa tehokkaasti käyttämällä Exceliä.

## 2 DATAMIGRAATIO JA SEN VAIHEET

### 2.1 Datamigraatio

Datamigraatiolla tarkoitetaan tiedonsiirtoprosessia tallennusjärjestelmästä, tietomuodosta tai tietokonejärjestelmästä sekä laitteesta toiseen (Dufasne ym., 2017, s.16; NetApp, n.d., kohta What is Data Migration; Right People Group, 2023, kohta 5 tärkeää vaihetta onnistuneeseen tiedonsiirtoon; Talend, n.d., kohta What is Data Migration). Siirrettävä data voi olla mm. yrityksen sopimusarkisto tai muita tärkeitä tiedostoja, joita ei saa hävittää.

Migraatiossa dataa ei kuitenkaan ole tarkoitus vain siirtää järjestelmästä toiseen, vaan ensin data puretaan vanhasta järjestelmästä, jonka jälkeen dataa valmistellaan sekä muokataan siirtoa varten. Toimenpiteet suoritetaan siksi, että tiedot olisivat yhteensopivia uuden järjestelmän kanssa. Kun tiedot on muokattu, ne siirretään uuteen järjestelmään. (Dufasne ym., 2017, s.16).

Datamigraatiot jaetaan yleensä kolmeen eri tyyppiin:

- **Tallennustilan siirto.** Olemassa olevan datan laajentamista nykyaikaisempaan rakenteeseen, jotta yhä useampi järjestelmä pystyy käsittelemään siirrettyä dataa. Tämä tuo mukanaan nopeamman ja tehokkaamman suorituskyvyn ja kustannustehokkaamman skaalauksen. (NetApp, n.d., kohta Types of data migration.)
- **Tietokannan siirto.** Usein on tarpeen siirtyä uuteen tietokantaan, jotta saadaan uusimmat liiketoimintavaatimukset täytettyä. Liiketoiminta- ja sovellustarpeistasi riippuen dataa voidaan siirtää tyyppillisestä On-premise-tietokannasta NoSQL-tietokantaan, pilvessä olevaan SQL-tietokantaan tai jopa Database as a Service -palveluun. (Talend, n.d., kohta Typical types of Data Migration.)
- **Sovellussiirto**-tyypissä sovellusohjelma siirretään järjestelmästä toiseen. Se voi sisältää koko sovelluksen siirron paikallisesta keskuksesta pilveen, pilvien välillä tehdyn siirron tai sovelluksen datan siirron kokonaan uuteen sovellukseen. (NetApp, n.d., kohta Types of data migration.)

Datamigraatio on erittäin tärkeä osa monen yrityksen tietojärjestelmien ylläpitoa, koska se mahdollistaa tietojen siirtämisen vanhoista varastoista uusiin vaivattomasti, jotta yritys voi jatkaa toimintaansa sujuvasti. Datamigraation voi tehdä seuraavista syistä:

- Tietojärjestelmän päivittäminen.
- Tietokannan uudelleenjärjestely.
- Pilvipalveluun siirtyminen.
- Vanhan laitteiston korvaaminen uudella.
- Yrityksien yhdistyminen.

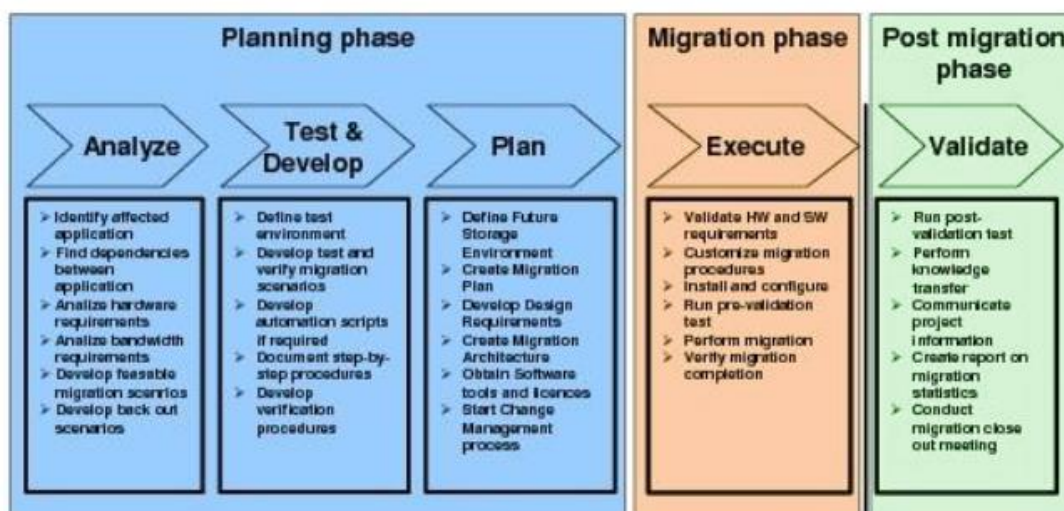
## 2.2 Datamigraation vaiheet

Voidaan sanoa, että lähes jokainen datamigraatioprosessi noudattaa kuvassa 1 näkyvää kolmivaiheista migraatorakennetta (Dufasne ym. 2017, s.17). Tietojen siirtäminen ei ole yksinkertainen tehtävä. Se vaatii huolellisen



suunnittelun, oikeat työkalut sekä hyvän strategian. Avain onnistuneeseen tiedonsiirtoon on huolellisessa suunnittelussa, oikean strategian valinnassa, tietojen oikeellisuuden ja laadun tarkistuksessa, perusteellisessa toteutuksessa ja siirron jälkeisissä toimissa. Oli kyseessä sitten tietokannan siirto, tallennustilan siirto tai vaikka pilvipalveluun siirtyminen. (Right People Group, 2023, kohta Johtopäätös.)

## Migration Process



Kuva 1. Datamigraation kolmivaiheinen prosessi (Dufrasne ym., 2017, s.17)

Suunnitteluvaihe alkaa nykytilanteen analysoinnilla ennen siirtoa. Tässä vaiheessa on tarkoitus määrittellä, mihin sovelluksiin siirto vaikuttaa ja mitkä ovat siirron riippuvuudet, vaatimukset ja rajoitukset. Suunnitteluvaiheen seuravassa osassa kehitetään ja testataan kaikki mahdolliset siirtoskenaariot, jonka jälkeen kootaan kaikki tulokset ja havainnot projektisuunnitelmaan. (Dufrasne ym., 2017, s.18.)

Siirtovaiheessa tietojen siirto suoritetaan käyttämällä työkaluja ja skenaariota, jotka on todettu sopiviksi edellisessä vaiheessa. Migraatiovaiheen tulos on tuotannon katkaisu uuteen ympäristöön. (Dufrasne ym., 2017, s.18.)

Validointivaiheessa tiedon siirron täydellisyys tarkistetaan sekä varmistetaan, että kaikki sovellukset toimivat oikein uudessa ympäristössä. Jos siirto

onnistuu, tuotanto siirtyy uuteen ympäristöön. Muussa tapauksessa siirto peruutetaan alkuperäiseen kokoonpanoon ja perimmäisen syyanalyysin on selvitettävä, mikä meni pieleen. Kun korjattava asia on löydetty, voidaan ongelmat korjata sekä aloittaa uusi siirtoyritys migraatiovaiheessa. (Dufrasne ym., 2017, s.18.)

### 2.2.1 Suunnittelu

Suunnitteluvaihe on yksi datamigraatioprojektin kriittisimmistä ja tärkeimmistä vaiheista. Sen tarkoitus on luoda selkeä suunta koko projektin ajaksi, tarjota oikea reitti sekä määritellä onnistumisen mittarit. Suunnitteluvaiheessa arvioidaan koko projektin laajuus, aika-arvio sekä mahdollinen budjetti. (Prohorchik, 2021, kohta Stage 1. Planning.) On myös tärkeää tutustua syvällisesti lähde- ja kohdejärjestelmään, jotta osataan valita tarpeisiin sopivat tiedonsiirtotyypit sekä tiedonsiirtotyökalut (Right People Group, 2023, kohta Vaihe 1: Perusteellinen suunnittelu ja arviointi).

Prohorchik (2021, kohta Choose a migration approach) määrittelee datamigraatiolle kaksi eri lähestymistapaa: "big bang" ja "trickle one". "Big bang" -tavalla tarkoitetaan lähestymistapaa, jossa koko migraatio suoritetaan yhtenä kokonaisuutena määritetyn aikaikkunan sisällä, jolloin siirrettävää dataa ei jaeta pienempiin osiin. Tämä tapa sopii paremmin pienemmille yrityksille, joilla data on suppeampaa. (Prohorchik, 2021, kohta Choose a migration approach). Jos "Big bang" -tapa on yrityksesi kannalta järkevin, on suositeltavaa suorittaa tiedonsiirto, esimerkiksi testiympäristöön, ennen varsinaista tapahtumaa (Tallend, n.d., kohta Data Migration Strategies).

"Trickle one" -lähestymistavalla tarkoitetaan tapaa, jossa tiedonsiirto tehdään järjestelmästä toiseen iteroimalla eli vaiheittain. Tarkemmin selitettynä siirrettävä data pilkotaan pienempiin kokonaisuuksiin, jolloin kaikki migraation eri työvaiheet toistetaan, kunnes kaikki haluttu data on siirretty kohdejärjestelmään. (Prohorchik, 2021, kohta Choose a migration approach).

Valtaosa yrityksistä käyttää migraatioissaan "Trickle one" -lähestymistapaa, vaikka se on huomattavasti monimutkaisempi prosessi kuin "Big bang" -lähestymistapa. Tämä johtuu siitä, että "Trickle one" -tapa vaatii todella huolellista suunnittelua. Suunnittelun lisäksi iterointi vaatii suuren määrän vaivaa sekä datan käyttäjiltä, että migraation suunnittelijoilta. Syy tämän lähestymistavan suosioon on se, että se ei aiheuta ylimääräisiä vakavia seisokkiaikoja eikä toimintahäiriöitä, jos tiedonsiirtoprosessi tehdään oikein. Tämän lisäksi tapa helpottaa huomattavasti virheiden sekä ongelmatilanteiden havaitsemisen ja helpottaa niiden ratkaisemista prosessin varhaisessa ajassa. (Prohorchik, 2021, kohta Choose a migration approach).

Jos datamigraatio toteutetaan yrityksen omalla datalla vanhasta järjestelmästä uuteen, tulee suunnitella myös vanhan järjestelmän käytön lopettaminen. Mikäli migraatio ja migraatiotiimi on ulkoistettu, vanhan järjestelmän lopettaminen on datamigraation tilaajan vastuulla. Vanhan järjestelmän lopettamisessa on otettava huomioon seuraavat asiat:

- Lakkautta kaikki aktiiviset sovelluslisenssit.
- Hävitä tarpeettomat laitteet.
- Peruuta vanhan järjestelmän ylläpito ja tuki.
- Suunnittele henkilöstön uudelleen koulutus, jotta uuden järjestelmän käyttö helpottuu. (Prohorchik, 2021, kohta Plan the old system retirement.)

### 2.2.2 Datan muokkaus sekä valmistelu

Suunnitteluvaiheen jälkeen voidaan aloittaa datan muokkaus ja valmistelu -vaihe. Se koostuu neljästä eri osa-alueesta:

- datan tarkastaminen
- validoiminen
- varmuuskopioiminen
- hallinnoiminen.

Vanha data tulee tarkastaa ja validoida huolellisesti, jotta itse migraatiossa ei tule ongelmia. On suositeltavaa käydä data läpi tietoaalueittain ja tutkia yksityiskohtaisesti mahdollisten epäjohdonmukaisuuksien, duplikaattien, virheellisten sekä ristiriitaisten tietojen varalta. Kaikki ongelmat tulee ratkaista ennen tiedonsiirtoa uuteen järjestelmään. (Prohorchik, 2021, kohta Plan the old system retirement). Hyvä vinkki tähän on tiedonsiirto testiympäristöön, josta kerrotaan tarkemmin kappaleessa 2.2.4.

Datan laadun tarkistamisen jälkeen on suositeltavaa varmuuskopioida data, koska ongelmatilanteet voivat johtaa tiedon häviämiseen tai vioittumiseen. Jos datan siirrossa tapahtuu virheitä, voidaan tarvittaessa palauttaa data varmuuskopiosta. Tämän takia varmuuskopiointi toimii ylimääräisenä ”suojakilpenä”. (Prohorchik, 2021, kohta Plan the old system retirement).

Datan hallinnointi -osiossa tulee selvittää kaikkien käyttäjien käyttöoikeudet datalle. Tarkemmin selitettynä tulee selvittää, kenellä on oikeus käyttää, muokata sekä poistaa mitään dataa. Kaikki nämä oikeudet pitää siirtää uuteen järjestelmään, jotta ne pysyvät samana myös uudessa järjestelmässä. (Prohorchik, 2021, kohta Stage 2. Preparing data.)

### 2.2.3 Datamigraation toteutus

Kun edelliset vaiheet on tehty, voidaan siirtyä datamigraation toteutusvaiheeseen ETL-prosessina. ETL viittaa datan poimimisen lähdejärjestelmästä (Extract), sen valmistamisen ja muokkaamisen (Transform) sekä lataamisen kohdejärjestelmään (Load). Tähän kuuluu myös tietojen muokkaaminen tiettyjen liiketoimintasääntöjen ja vaatimusten mukaisiksi. (DOMO, n.d., kohta What is ETL for data migration?)

### 2.2.4 Testaus

Datamigraation testaus tarkoittaa siirrettävän datan arviointiprosessia ennen datan tuontia kohdejärjestelmään datan tarkkuuden, täydellisyyden ja

johdonmukaisuuden varmistamiseksi. Testauksen päätavoite on varmistaa, että tiedot siirretään oikein vaarantamatta tietojen eheyttä tai häiritsemättä liiketoimintaa. Tätä varten testaus tulisi suorittaa ennen tuontia, sen aikana sekä sen jälkeen. (Trotino, 2023, kohta What is Data Migration Testing.)

Testaus tulee aloittaa heti, kun aloittaa tietojen käsittelyn, ja se jatkuu kaikissa tulevissa vaiheissa. Oli kyse sitten suunnittelusta, toteutuksesta tai siirron jälkeisistä toimista. Jos migraatiossa käytetään vaiheittaista siirtoa, tulee testata jokainen siirretty "tietoerä" havaitaksesi ajoissa kaikki mahdolliset virheet ja varmistaaksesi tietojen laadun. Virheiden korjaamisen jälkeen tulee testata data uudelleen. Kun siirto on valmis, pitää siirretyt tiedot tarkistaa käyttämällä yksikkö-, järjestelmä-, täyden volyymin tai osasovellustestausta. (Prohorchik, 2021, kohta Stage 5. Testing and going live.) Hyvä tapa datan testaukseen on suorittaa testaukset testiympäristöön. Vasta silloin, kun virheitä ei löydy ja kaikki osapuolet ovat tuloksiin tyytyväisiä, voidaan tietojen siirto aloittaa tuotantoympäristöön.

### 2.2.5 Jälkitoimet

Kun tiedonsiirto on tehty, vielä jää jäljelle yksi tärkeä vaihe, nimittäin siirtämisen jälkeiset toimet. Tässä vaiheessa tulee varmistaa, että kaikki toimii niin kuin pitää. Alapuolella käydään tarkemmin läpi asioita, joihin tulee jälkitoimissa keskittyä:

- **Tarkista datan oikeellisuus.** Tarkista kaikki kohdejärjestelmän tiedot varmistaaksesi, että ne ovat tismalleen samanlaiset kuin lähdejärjestelmässä. Tähän prosessin helpottamiseksi on suositeltavaa käyttää tiedonsiirtotyökaluja, jotka tarjoavat validointiominaisuuksia. (Right People Group, 2023, kohta Datan validointi ja liiketoimintaprosessien arviointi.)
  - Hyvä esimerkki datan tarkistamiseen on "**Ad hoc**" -tarkistus. Sen tarkoitus on varmistaa, että kohdejärjestelmä pystyy käsittelemään siirrettyä dataa oikein. Näillä tarkistuksilla voidaan

tarkistaa esimerkiksi tärkeiden taulujen data, jotta voidaan varmistaa sekä lähde- että kohdejärjestelmän tietojen täsmällisyys.

- **Tarkista liiketoimintaprosessit.** Varmistu siitä, että kaikki prosessit toimivat uudessa ympäristössä, ettei koidu ongelmia. Jos jokin on pielessä, nyt on viimeinen hetki korjata se. (Right People Group, 2023, kohta Datan validointi ja liiketoimintaprosessien arviointi.)

Tässä vaiheessa tulee myös muistaa poistaa käytöstä vanhat järjestelmät. Ennen poistamista tulee tarkistaa useampaan kertaan, että kaikki tiedot ovat siirretty onnistuneesti sekä kaikki prosessit toimivat oikein kohdejärjestelmässä. Käytöstä poistamisen ansiosta yritys voi kohdentaa enemmän resursseja kriittisempiin alueisiin, mikä voi johtaa huomattaviin kustannussäästöihin. (Right People Group, 2023, kohta Pois käytöstä vanhat järjestelmät.)

### 3 TOTEUTUS

Tässä kappaleessa käydään läpi suorittamaani datamigraatiota vaihe vaiheelta ja kerrotaan samalla, mitä työkaluja käytettiin migraatiossa.

#### 3.1 Taustatietoa

Ennen datamigraatiota käytössä on ollut vanha M-Filesin sopimusarkisto, joka ei enää vastaa nykyisiä toimintamalleja eikä sisällön puolesta ole tarpeellisia tietoja saatavilla (dokumenttien metadata). Tässä yhteydessä käytämme vanhasta sopimusarkistosta nimitystä M-Files On-premise, koska M-Files-dokumenttivarasto oli asennettu paikalliselle palvelimelle.

Vanhassa järjestelmässä oli useita ongelmakohtia, jotka korjaantuisivat pilvi-varastoon siirtyessä. Alapuolella on listattuna kaikki ongelmakohdat:

- Sopimuksien työnkulku ei enää vastaa nykyisiä toimintamalleja.
- Sopimuksien tarvittavat tiedot eivät ole pääsääntöisesti ajan tasalla.

- Asiakassopimusten ja toimitustietojen välillä ei ole kytkentää.
- Sähköinen allekirjoitus on erillinen manuaalinen prosessi.
- Uuden toiminnanohjausjärjestelmän käyttöönoton yhteydessä vanhat integraatiot eivät enää toimi.

Tavoitteena oli siirtyä M-Filesin pilviversioon, johon siirretään kaikki data. Uudessa varastossa otetaan tällöin käyttöön yksinkertaisempi hallintamalli. Alapuolella on listattuna pilvipalveluun siirtymisen hyvät puolet:

- Yksinkertaistetaan työnkulkua.
- Sähköinen allekirjoitus osaksi työnkulkua.
- Yhdistetään ratkaisu osaksi liiketoiminnan uudistettuja prosesseja tukevia järjestelmiä ja otetaan käyttöön tarvittavat integraatiot.

Tässä suunnitelmassa keskitytään datan siirtoon M-Filesin paikallisesta varastosta M-Filesin pilvivarastoon. Tarvittava arkkitehtuurimalli ja rakenne ovat mietitty jo valmiiksi.

### 3.2 Suunnittelu

Datamigraation suunnitteluvaiheessa tuli valita optimaalinen työkalu yrityksen sopimusten siirtämiseen uuteen järjestelmään. Vaihtoehtoja harkittaessa päätettiin käyttää **M-Files Importer** -työkalua, koska lähde- ja kohdejärjestelmän metadatarakenteet poikkesivat merkittävästi toisistaan. Tämä tarkoitti sitä, että metadatarakenteita oli muokattava huomattavasti migraation aikana.

Yllä mainittu M-Files Importer on monipuolinen tuontityökalu, jonka avulla data voidaan tuoda järjestelmästä dokumentteina M-Files-varastoihin metatietojen avulla. Importer tukee myös ei-asiakirjaobjektien, kuten asiakkaiden, tuontia CSV-, XML- ja XLSX-muodossa M-Filesiin. (M-Files, n.d., kohta M-Files Importer).

Kun selvitettiin datamigraation suoritustapaa, oli tarpeen myös päättää rakenteesta CSV-tiedostolle, jota tarvittaisiin migraatioon. Vaihtoehtoja rakenteelle oli kaksi:

- **Vaihtoehto 1**

- **Vaiheessa 1** luodaan uusi Sales Contract -luokan multifiledokumentti vanhan Sales Contract -kohteen ja siihen liittyvien metatietojen avulla.
- **Vaiheessa 2** siirretään vanhaan Sales Contract -kohteeseen liittyvät dokumentit uuteen Sales Contract -luokan multifiledokumenttiin.
  - Mikä on uusi vastaanottavan Sales Contract -luokan multifiledokumentti?
  - Vanhan Sales Contract -kohteen ja uuden Sales Contract -luokan dokumentti linkitetään Migration ID -ominaisuudella.

- **Vaihtoehto 2**

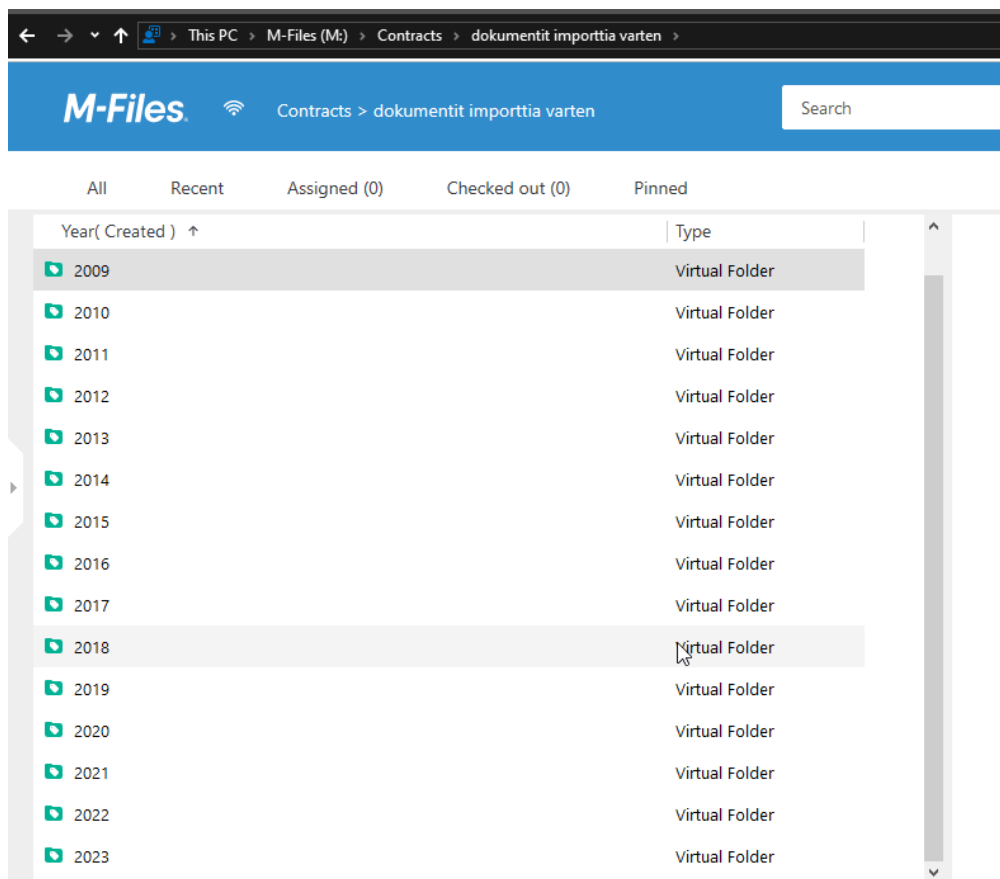
- Kootaan Excel-tiedostoon yksi rivi, jossa on metatiedot vanhalta Sales Contract -kohteelta, mutta tehdään se multifile-muodossa.
  - Sisältö listaa vanhaan Sales Contract -kohteeseen liittyvät dokumentit.

Kummassakin vaihtoehdossa lopputulos olisi ollut identtinen, mutta ensimmäinen vaihtoehto olisi vaatinut enemmän työvaiheita. Tässä vaihtoehdossa olisi myös ollut epäselvää, miten merkitä yhteys luodun (mutta dokumentteja sisältämättömän) Sales Contract -multifiledokumentin ja vanhan Sales Contract -kohteen välillä, jotta dokumentit voitiin liittää uuteen multifiledokumenttiin. Lopulta päätettiin hyödyntää vaihtoehtoa 2, koska se osoittautui paremmaksi tavaksi toteuttaa CSV-tiedoston rakenne yrityksen näkökulmasta.

Kun CSV-tiedoston rakenne oli päätetty, seuraavana vaiheena oli valita menetelmä, jolla datamigraatio suoritettaisiin. Päätettiin hyödyntää aiemmin mainittua "Trickle one" -menetelmää, jossa siirrettävä data jaettaisiin pienempiin palasiin. Tämä päätös tehtiin varmistaaksemme datamassojen sopiva koko sekä vienti- että tuontivaiheessa. Pienen tutkimuksen jälkeen päätettiin, että



jos yhden vuoden dokumenttien määrä ylittää 1000 dokumentin rajan, on järkevää ryhmitellä dokumentit kuukausittain. Kuitenkin huomattiin, ettei kuukausittain ryhmittelyä tarvinnut tehdä, koska yhdenkään vuoden dokumenttien määrä ei ylittänyt kyseistä rajaa. Lopullinen päätös oli siis ryhmitellä data vuosittain kuvan 2 mukaisesti.



Kuva 2. Näkymä lähdejärjestelmän sopimuksista ryhmiteltynä vuosittain.

### 3.3 Datat valmistelu

Ennen kuin päästiin viemään dataa lähdejärjestelmästä, tuli data valmistella oikeanlaiseksi. Tietorakenne tuli käydä läpi, jotta sopimusten sekä niiden liitteiden välinen riippuvuus löydettiin. Huomattiin, että liitteille piti erikseen luoda kenttä, johon haettiin ko. liitteelle liittyvän opportunityn ID. Sopimukselle liittyvät liitteet (eli sopimusdokumentit) kuuluvat M-Filesin luokkaan "Sales contract attachment". Ko. luokalle piti luoda erikseen uusi ominaisuus (property, ns. "kenttä") "Related Oppo ID", johon haettiin automatiikan avulla arvoksi sen

sopimuksen ID-arvo, johon kyseinen liite kuului. Kuvissa 3 ja 4 nähdään Sales Contract attachment -luokkaan liittyvät ominaisuudet sekä äsken mainittu Related Oppo ID -kenttä.

Class Properties - Sales Contract attachment

General | Templates | Permissions | Automatic Permissions | Advanced

Name: Sales Contract attachment

Object type: Document

Properties:

Name	Data Type	Required
Document Title *	Text	<input checked="" type="checkbox"/>
Sales Contract	Choose from list "Sales Contrac..."	<input checked="" type="checkbox"/>
Attachment Type	Choose from list "Attachment T..."	<input type="checkbox"/>
Oppo	Text	<input type="checkbox"/>
File Extension	Choose from list "Extensions"	<input type="checkbox"/>
Related Oppo ID	Text	<input type="checkbox"/>

Add... Remove Set As Name ↑ ↓

Kuva 3. Sales Contract attachment -luokkaan liittyvät ominaisuudet.

Property Definition Properties - Related Oppo ID

General | Automatic Values | Validation | Permissions | Advanced

Name: Related Oppo ID

Data type: Text

Content: Text

Show values from the following list:

Filter the list by using the value of the following property:

Sort values in the list in the following order:

Allow using this property with the following object type: Document

Enable automatic permissions via this property

Allow this property to be used as a grouping level in views

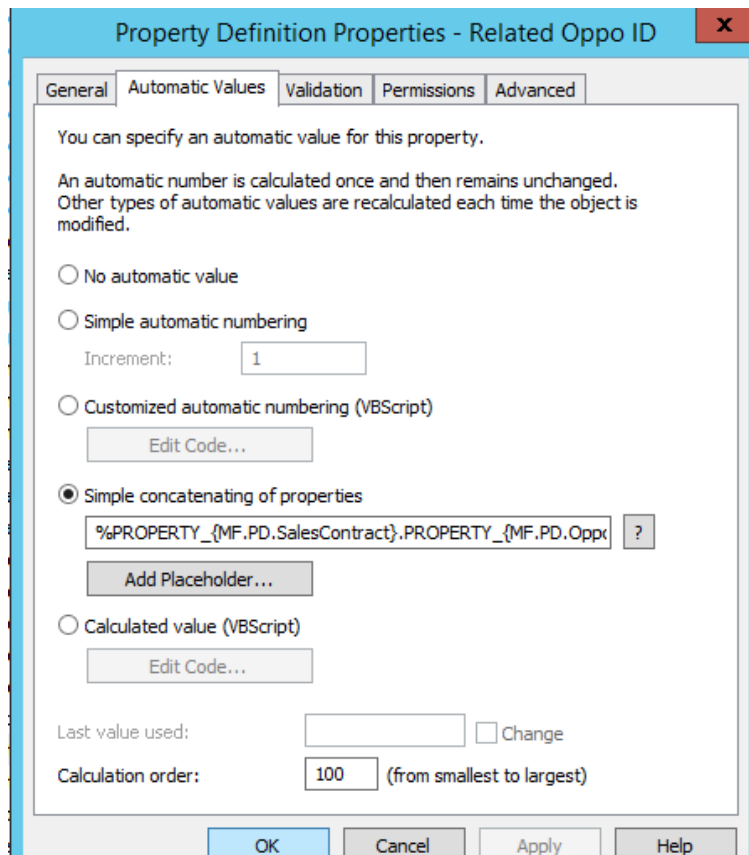
Allow searching for objects by this property

Do not search for old object versions

OK Cancel Apply Help

Kuva 4. Related Oppo ID -ominaisuuden perustiedot.

Kuvassa 5 näkyy kohta Simple concatenating of properties, jossa on komento, jonka avulla automaattikka hakee sopimuksen ID:n. Komento on kokonaisuudessaan `%PROPERTY_{MF.PD.SalesContract}.PROPERTY_{MF.PD.Opportunity}.PROPERTY_{MF.PD.OpportunityID}%`



Kuva 5. Related Oppo ID -ominaisuuden automaattitiettyttöön liittyvät tiedot.

Tämän tiedon avulla saatiin seuraavassa vaiheessa yhdistettyä sopimus opportunitylle ja liitteet sopimuksille opportunityn kautta.

### 3.3.1 Dokumenttien vienti (Export)

Datan valmistelun jälkeen oli aika viedä data lähdejärjestelmästä. Hyvin nopeasti tuli ilmi, ettei M-Filesissa pysty viemään dataa kuin 500 dokumenttia kerrallaan. Dokumentit piti siis viedä useammassa osassa, jos tämä raja ylittyi. Se ei kuitenkaan koitunut ongelmaksi, koska vain kolmen vuoden kohdalla kyseinen raja ylittyi.

Kun dokumentit saatiin vietyä lähdejärjestelmästä, piti viennissä luodut Excel-tiedostot yhdistää, mikäli 500 dokumentin raja ylittyi. Muussa tapauksessa päästiin suoraan datan muokkaukseen, josta puhutaan seuraavassa kappaleessa.

### 3.3.2 Datan muokkaaminen

Muokausvaiheessa erilliset Excel-välilehdet yhdistettiin käyttämällä Excelin omia käskyjä sekä luomalla tarvittavat koodit ja ehtolauseet eri toimintoja varten. Alapuolella käydään läpi kaikki migraatiossa käytetyt koodit ja ehtolauseet sekä niiden toiminnot.

- =IFERROR(VLOOKUP(A2;Sheet2!\$A\$2:\$B\$5;2;FALSE);" ")
  - Tämän ehtolauseen avulla onnistuttiin muuttamaan lähdevaraston Language-arvot vastaamaan kohdevaraston vastaavia arvoja. Ensimmäinen askel oli luoda Excel-tiedosto, jossa on kaksi välilehteä. Ensimmäiselle välilehdelle kerättiin vanhat (FI ja EN) ja uudet (FIN, SWE, ENG ja GER) Language-arvot omiin sarakkeisiinsa. Toisen välilehden ensimmäiseen sarakkeeseen sijoitettiin esimerkiksi vuoden 2017 dokumenttien Language-arvot, minkä jälkeen toiseen sarakkeeseen liitettiin yläpuolella oleva ehtolause. Näin varmistettiin, että uuden varaston arvot korvaavat vanhan varaston arvot. Kuvassa 6 näkyy selkeämmin yllä kuvattu tilanne.

old name	Name	ID	Ow	Language	Language
FI	FIN	1		FI	FIN
	SWE	2		FI	FIN
EN	ENG	3		FI	FIN
	GER	4		FI	FIN
				FI	FIN
				FI	FIN
				FI	FIN
				FI	FIN
				FI	FIN
				FI	FIN
				FI	FIN

Kuva 6. Language-arvojen muuntaminen.

- =B2&" "&A2
  - Tämän koodin avulla suoritettiin tarvittavat muutokset Owner-arvoihin. Luotiin Excel-tiedosto, jossa oli välilehti, johon syötettiin esimerkiksi vuoden 2017 dokumenttien Owner-arvot. Arvot olivat muodossa Sukunimi Etunimi. Tämän jälkeen suku- ja etunimet eroteltiin omiin sarakkeisiinsa, ja käytettiin yläpuolella olevaa koodia yhdistämään suku- ja etunimet uuteen sarakkeeseen. Lopullisten arvojen tuli siis olla muodossa Etunimi Sukunimi, kuten kuvassa 7 nähdään.

	A	B	C	D
1	Owner			Owner
2	Pouta	Pekka	(deleted)	Pekka Pouta
3	Myrsky	Jouni	(deleted)	Jouni Myrsky
4	Kuka	Muu	(deleted)	Muu Kuka
5	Sade	Emilia	(deleted)	Emilia Sade

Kuva 7. Owner-arvojen muuntaminen.

- =VLOOKUP(F2;Sheet5!\$A\$1:\$B\$2410;2;FALSE)
  - Tällä saatiin luettua oikeat Accountit Excel-välilehdestä, jossa kaikki Accountit oli listattuna. Kuvasta 8 saa paremman kuvan kyseisen komennon käytöstä.

	A	B	F	G	H
1	Related sales contract oppo ID	DBE Customer Name	Related sales contract oppo ID	Opportunity	Account (from Opportunity)
2	018772c8-c128-e711-80f9-5065f38aea21	FP OYJ	018772c8-c128-e711-80f9-5065f38aea21	21303 - FP palvelu	FP Oyj
3	dc840c89-c496-e711-815e-e0071b66d0e1	DFP	1dc8d675-da12-e711-80f6-5065f38aea21	21230 - FP konsultointi	PF OYJ
4	56b3917c-30ab-e711-8162-e0071b66d0e1	CA	b9745629-580a-e711-80f7-5065f38a4ad1	21198 - CRM ylläpito ja kehitys	FG OYJ
5	39c776fb-77ae-e711-8162-e0071b66d0e1	DG	987ba5be-df07-e711-80fa-5065f38bd502	21177 - Technologies	GPT OY

Kuva 8. Account-arvojen luku toiselta välilehdeeltä.

- =IFERROR(VLOOKUP(A2;Sheet4!\$A\$2:\$C\$88;3;FALSE);" ")
  - Tämän ehtolauseen avulla suoritettiin lähdejärjestelmän Cost Center -arvojen vaihto vastaaviin arvoihin kohdejärjestelmässä. Tarvittiin kaksi eri välilehteä Excelissä; ensimmäiselle välilehdelle kerättiin kaikki mahdolliset Cost Center -arvot

lähdejärjestelmästä, niiden ID:t sekä vastaavat arvot kohdejärjestelmästä, jokainen omalle sarakkeelleen. Toiselle välilehdelle koottiin esimerkiksi vuoden 2017 dokumenttien Cost Center -arvot. Näiden tietojen ja yläpuolella olevan ehtolauseen avulla toteutettiin lähdejärjestelmän arvojen korvaaminen kohdejärjestelmän arvoilla. Kuvassa 9 näkyy toteutus yllä kuvatusta tilanteesta.

Sheet2			Sheet1		
ID	Name		Name	ID	C
1200	Data Driven Solutions		OLD 1120 Servers	1120	
1210	Company CC 1		OLD 1140 Customers	1140	
1220	Company CC2		OLD 1150 Service	1150	
1230	Company CC3		1200 Data Driven	1200	Data Driven Solutions
1240	Climber Holding		2100 EP	2100	EPS
1250	Climber International		2120 MS D365	2120	MS Company

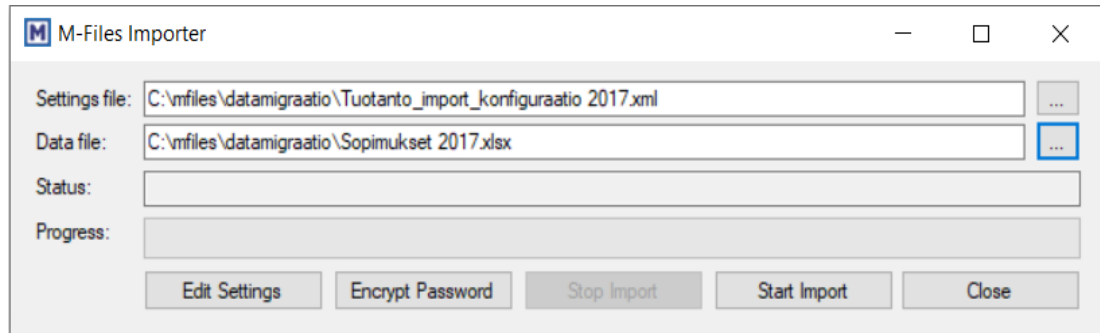
Kuva 9. Cost Center -arvojen muuntaminen.

- =IF(I2=I1;Q1&"|";P2), =IF(I3=I2;Q2&"|"&P3;P3), =IF(I2<>I3;"Last";"")
  - Näiden koodien avulla pystyttiin luomaan järjestys dokumenteille, jotta saadaan tiedostolistaus datamigraation vaatimaan muotoon.
- =XLOOKUP(F2;dokumentit!O122:O10420;dokumentit!Q122:Q10420;"";0;-1)
  - Tämän avulla pystyttiin kasaamaan kaikki luotavan multiframe-dokumentin tiedostot yhdelle riville Exceliin.

Huomioitavaa datan muokkauksessa näiden koodien ja ehtolauseiden avulla on se, että nämä vaiheet tuli toistaa jokaisen vuoden kohdalla. Pelkästään datan muokkauksessa meni siis hyvin paljon aikaa, varsinkin ensimmäisen vuoden kohdalla. Tämä vaihe helpottui, mitä pidemmälle vuosissa mentiin, koska asioita alkoi muistamaan jo ulkoa.

### 3.4 Migraation suoritus testiympäristöön

Tuonnit suoritettiin käyttäen aiemmin mainittua M-Files Importer -työkalua, jonka kuva näkyy alla. Kuvan alla on myös selitetty, mitä eri kohtiin tuli syöttää, jotta dokumenttien tuonti voitiin suorittaa onnistuneesti.



Kuva 10. M-Files Importer -työkalu.

- **Settings file** -kohtaan tuli liittää xml-tiedosto, jossa oli määritelty kaikki migraatiossa tarvittavat asetukset. Kyseisen asetukset käydään läpi seuraavassa luvussa.
  - Kuvassa esimerkkinä käytettiin tiedostoa "Tuotanto\_import\_konfiguraatio 2017.xml", joka sisälsi vuoden 2017 dokumenttien tuotantoympäristön konfiguraatitiedot. Käytännössä tämä tiedosto oli samanlainen kuin testiympäristön vastaava tiedosto.
- **Data file** -kohtaan piti liittää xlsx-tiedosto, josta dokumenttien tiedot luettiin.
  - Kuvan esimerkissä käytettiin vuoden 2017 sopimuksien Excel-tiedostoa.

Kaikki dokumentit tuotiin vuosittain erilliseen testiympäristöön ennen tuotantoympäristöön suoritettavaa tuontia. Tämä toimenpide oli suunniteltu välttämään mahdolliset ongelmat, joita käsitellään seuraavassa kappaleessa käsiteltävien virheiden osalta.

Testausvaiheessa kului odotettua enemmän aikaa, pääasiassa siksi, että virheitä ilmeni testituonnin jälkeen, ja ne oli korjattava ennen jatkamista. Jokaisen

vuoden kohdalla oli tarpeen suorittaa useita testituonteja, sillä uusia virheitä ilmeni, kun edelliset virheet oli korjattu. Syytä tälle ei keksitty. Testivaiheen hitauteen vaikutti myös nettiyhteyden hitaus. M-Files Importer edellyttää internet-yhteyttä toimiakseen, ja mitä hitaampi yhteys oli, sitä kauemmin testivaihe kesti.

### 3.5 Virheet

Testiympäristöön tehtyjen testituontien jälkeen tuli korjata edellisessä kappaleessa mainitut virheet. Alapuoolella on listattuna kaikki virheet ja niiden määrittymiset sekä ratkaisut virheiden korjaamiselle.

- **This object requires a source file.**
  - Viittaus olemassa olevaan kohdetiedostoon on puutteellinen/virheellinen.
    - Korjattiin käymällä läpi tiedostolistauksen ja xml-tiedoston väliset erot ja korjaamalla tiedoston nimi oikein.
- **Cannot find object with display ID 'x' for property 1054.**
  - Kyseiseen ID:hen liittyvää Opportunitya ei ole olemassa.
    - Lisättiin taustadataan puuttuva Opportunity tai mikäli kyseistä Opportunitya ei löytynyt, poistettiin kyseinen rivi xml-tiedostosta.
- **Source file not found.**
  - Viittaus olemassa olevaan kohdetiedostoon on puutteellinen/virheellinen.
    - Korjattiin käymällä läpi tiedostolistauksen ja xml-tiedoston väliset erot ja korjaamalla tiedoston nimi oikein. Oli myös tapauksia, joissa tiedostoja ei löytynyt, jolloin kyseinen rivi poistettiin xml-tiedostosta.
- **Assigning value 'x' to the M-Files property definition ID 1177 failed.** (ID 1177 viittaa Stagingissa Cost Centeriin, kun taas tuotannossa ID on 1167)
  - Staging-testauksen jälkeen piti huomata vaihtaa kohdekentän tuotantoa vastaava ID.



- Korjattiin vaihtamalla kenttien ID:t oikeiksi Importerin konfiguraatioon.
- **Assigning value 'x' to the M-Files property ID 1060 failed.** Duplicate values found. (ID 1060 viittaa Stagingissa Owneriin, kun taas tuotannossa ID on 1057)
  - Staging-testauksen jälkeen piti huomata vaihtaa kohdekentän tuotantoa vastaava ID.
    - Korjattiin vaihtamalla kenttien ID:t oikeiksi Importerin konfiguraatioon.
- **Assigning value 'x' to the M-Files property definition ID 1021 failed.**
  - Kyseistä accountia ei löydy. Yksi syy siihen voi olla se, että yrityksen nimi on muuttunut/muutettu.
    - Jotta virhe saatiin korjattua, piti käydä muuttamassa Accountin nimi oikeanlaiseksi. Oikea Account löytyi M-Files Online -palvelusta.
- **A file with the name "x" already exists in the document.**
  - Importer yrittää lisätä tiedoston samaan monitiedostoiseen dokumenttiin toisen kerran. Samanniminen tiedosto on jo siis olemassa.
    - Tämän virheen sai korjattua poistamalla duplikaatin Excelistä sekä dokumenttikansiosta.
- **The specified path, file name, or both are too long. The fully qualified file name must be less than 260 characters, and the directory name must be less than 248 characters.**
  - Yhden solun maksimi merkkimäärä on ylittynyt.
    - Lyhennettiin liian pitkien tiedostojen nimiviittaukset sekä itse tiedostonimeen että xml-tiedostoon. Tämä saattoi vaatia myös uuden rivin luomista xml-tiedostoon.
- **Network problems are preventing M-Files from communicating with the server. The remote procedure call failed.**
  - Importin aikana nettiyhteydessä on ilmennyt ongelmia, tai se on katkennut kokonaan.

- Tämän virheen korjaamiseen ei löydetty muita ratkaisuja, kuin tuonnin suorittaminen uudelleen.

### 3.6 Importerin asetukset

M-Files Importerin asetuksia voi käyttää joko valmiista tiedostosta (xml) tai luoda omia asetuksia graafista käyttöliittymää käyttäen. Asetukset koostuvat kymmenestä eri välilehdestä, jotka määrittävät haluttua käyttäytymistä. Seuraavaksi käsitellään ne välilehdet, jotka olivat käytössä datamigraatiossa.

#### 3.6.1 M-Files Connection

Kuvassa 11 näkyvällä M-Files Connection -välilehdellä on useampi asia, joita tuli ottaa huomioon datamigraatiossa:

- **Authentication type.** Määrittää tavan, jolla Importer muodostaa yhteyden M-Files varastoon.
  - Datamigraatiossa valittiin kirjautumistilin tyyppiä M-Files User eli vaihtoehto MFAuthTypeSpecificMFilesUser.
- **Username,** johon laitettiin Digian määrittelemä käyttäjätunnus.
- **Salasana,** joka on tallennettu Digian erilliseen salasankantaan.
- **Protocol sequence,** johon valittiin ncacn\_http.
- **Host,** joka tarkoittaa M-Files-yhteyden isäntäpalvelinta.
  - Datamigraatiossa Stagingin palvelimeksi laitettiin Digian määrittelemä testipalvelin ja tuotantoon puolestaan tuotantopalvelin.
- **Port,** jossa valitaan M-Files-yhteydelle portti.
  - Datamigraatiossa käytettiin porttia 4466.
- **Vault GUID,** joka tarkoittaa M-Files-dokumenttivaraston GUIDia.
  - Pilvivaraston GUID:in näkee kohteiden hyperlinkeistä M-Files Desktopissa. Mappings-tietojen pitäisi päivittyä automaattisesti, kunhan varastoyhteys toimii.

Settings Editor

General Errors Monitoring Transmittals Duplicates Timestamps **M-Files Connection** File Format Options Folders Mappings

Authentication type: MFAuthTypeLoggedOnWindowsUser

Username:

Password:

Domain:

Protocol sequence: ncacn\_ip\_tcp

Host: localhost

Port: 2266

Local computer name:

Vault GUID:

Time zone:

SPN:

Encrypted connection

Allow anonymous connection

Is M-Files client application

Maximum number of attempts to recreate lost connection:

Delay between attempts in milliseconds:

Retry the failed M-Files API call after recreating the lost connection

Test Connection

Kuva 11. M-Files Connection -asetukset. (M-Files)

### 3.6.2 File Format Options

Kuvassa 12 näkyvällä File Format Options -välilehdellä piti muuttaa vain kahta kohtaa:

- **CSV separator** -kohtaan laitettiin puolipiste.
- **Excel sheet name** -kohtaan piti laittaa sama nimi, kuin mikä on Excel tiedoston välilehden nimi, josta data luetaan. Muuten Importer ei pysty luomaan CSV-tiedostoa, jolloin tuonti ei onnistu.

Settings Editor

General Errors Monitoring Transmittals Duplicates Timestamps M-Files Connection File Format Options Folders Mappings

Data file type: XLSX

CSV

CSV encoding: windows-1252

CSV separator:

CSV list separator: |

CSV quotation: "

Excel

Excel connection string: Provider=Microsoft.ACE.OLEDB.12.0;Data Source={0};Extended Properties="Excel 12.0;HDR=NO;IMEX=1"

Excel sheet name: Sheet1\$

Overwrite any existing temporary CSV file during Excel import

XML

XML encoding: UTF-8

XML list separator: |

XML item selector:

XML validation file:

The XML file must contain at least one item

Each XML item must provide values for all mappings

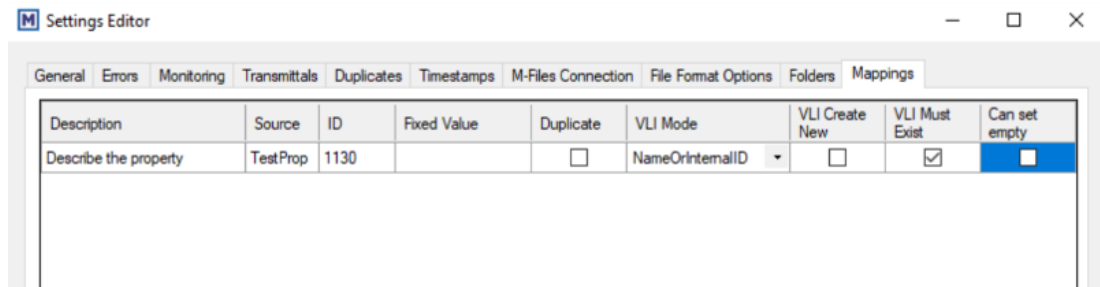
Kuva 12. File Format Options. (M-Files)

### 3.6.3 Mappings

Kuvassa 13 on oletusnäkyä Mappings-välilehdestä, jossa näkyy kaikki alapuolella olevat kohdat:

- **Description.** Vapaa muotoinen kuvaus sopimuksesta.
- **Source.** Mapattujen tietojen lähde. Tähän voi laittaa esimerkiksi CSV-tiedoston sarakkeen nimen.
- **ID.** M-Files ID, johon lähdetiedot tuodaan.
- **Fixed Value.** Voidaan pakottaa jollekin mappingille tietty arvo.
- **Duplicate.** Voidaan osoittaa, käytetäänkö kyseistä mappingia duplikaattien etsimiseen.
- **VLI Mode.** Tämä määrittää, kuinka lähdetiedot mapataan.
- **VLI Create New.** Kun tämä on laitettu päälle, Importer yrittää luoda uuden arvoluettelokohteen, jos se ei ole vielä olemassa.

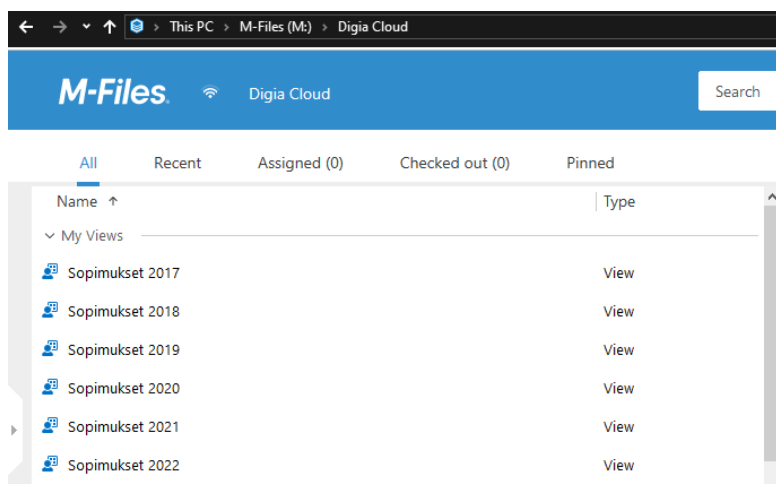
- **VLI Must Exist.** Kun tämä on laitettu päälle, Importer edellyttää, että arvolistakohde on jo olemassa. Kun tämä ei ole käytössä, tuonti onnistuu, vaikka arvoluettelon kohdetta ei löydy.
- **Can set empty.** Ilmaisee, voidaanko yhdistetty ominaisuus jättää tyhjäksi M-Filesissä, jos sillä ei ole arvoa datafilessa. Jos haluat tyhjentää ominaisuuden arvon M-Filesissä, ota tämä asetus käyttöön.



Kuva 13. Mappings. (M-Files)

### 3.7 Migraation toteutus tuotantoympäristöön

Kun kaikki edellä mainitut vaiheet oli käyty läpi, oli aika tuoda dokumentit tuotantoympäristöön. Tuonnit tuotantoympäristöön suoritettiin myös vuosittain. Tämä vaihe oli todella helppo suorittaa testausvaiheen ansiosta, koska ei tarvinnut huolehtia siitä, että virheitä tulisi vastaan. Alapuoella olevassa kuvassa 14 näkyy tuotantoon tuodut sopimukset ryhmiteltynä vuosittain.



Kuva 14. Näkymä M-Files Desktop -sovelluksesta tuotantoon tuoduista sopimusten vuosipaketeista.

### 3.8 Migraation viimeistely

Migraation viimeistelyvaiheessa tuli kerätä yhteen kaikki sopimukset, joita ei tuotu tuotantoon. Toisin sanoen ne sopimukset, joilla ei ollut liitteitä. Sopimusten yhteen kerääminen tehtiin siksi, että kaikki sopimukset saataisiin talteen, jotta lähdejärjestelmä voitiin ajaa alas. Tämä oli erittäin tärkeä vaihe siinä mielessä, että vanha palvelin oli jo tiensä päässä. Saatiin siirrettyä sopimustiedot turvaan, jolloin ei enää tarvitse panostaa yksittäisen palvelimen päivityksiin sekä ylläpitoon. Samalla sopimuksien hallinta ja pääsyoikeudet sekä sopimusten käsittelyn työkulku yksinkertaistui.

## 4 LOPPUPÄÄTELMÄT

Opinnäytetyön tavoitteena oli kasvattaa yleistä tietoisuutta datamigraatioprosessista ja sen merkityksestä liiketoiminnassa. Näin ollen opinnäytetyö voi tarjota apua muille organisaatioille ja ammattilaisille ymmärtämään, kuinka datamigraatio voidaan suorittaa tehokkaasti. Työn teoriaosuudessa käytiin läpi datamigraation prosessit projektina, esiteltiin datamigraation tyypit sekä tuotiin esiin syitä, miksi datamigraatioita tehdään. Näiden lisäksi käsiteltiin datamigraation eri vaiheita. Toteutusvaiheessa suunniteltiin datamigraation suoritustapaa sekä valittiin rakenne CSV-tiedostolle, jota datamigraatiossa tarvittiin. Näiden vaiheiden jälkeen tuli valmistella data, jotta voitiin viedä se lähdejärjestelmästä datan muokkausta varten. Valmistelun ja muokkauksen jälkeen aloitettiin testausvaihe, jossa tuotiin muokattu data testiympäristöön toimeksiantajan työkalujen ja ohjeiden avulla. Lopuksi tuli korjata kaikki ilmenneet virheet, jotta voitiin tuoda data virheettömästi kohdejärjestelmään.

Datamigraation lopputulos oli se, että saatiin siirrettyä Digian Sales Contract -luokan sopimukset omasta vanhasta M-Files On-premise -varastosta pilveen M-Files Cloudiin. Samalla tässä yhteydessä saatiin muutkin työn alussa määritellyt tavoitteet saavutettua.

Mielestäni opinnäytetyölle asetetut tavoitteet saavutettiin. Toiminnallisessa osassa ilmeni muutamia haasteita, erityisesti tiettyjen asioiden avaamisessa sekä kuvankaappausten tekemisessä M-Files-työkaluista ja Excel-tiedostoista. Tämä johtui siitä, että projektin päättyessä ei ollut enää pääsyä kyseisiin työkaluihin ja tiedostoihin. Onneksi toimeksiantajani avustuksella sain kuitenkin tarvittavaa apua, mikä mahdollisti asioiden tarkemman avaamisen kuvien avulla. Vaikka haasteita ilmeni, lopulta saatiin luotua toimiva lopputulos.

## LÄHTEET

Digia Finland Oy. (n.d.). Digia yrityksenä. Haettu 9.11.2023 osoitteesta <https://digia.com/digia-yrityksena>

Domo. (n.d.) What is ETL for data migration? Haettu 8.12.2023 osoitteesta <https://www.domo.com/learn/article/streamline-data-migration-and-business-processes-with-etl>

Dufrasne, B., Warmuth, A., Appel, J., Bauer, W., Douglass, S., Klee, P., Pura, M., Wells, M. & Wesselbaum, B. (2017). DS8870 Data Migration Techniques. IBM Redbooks. [https://play.google.com/books/reader?id=3fHDCAAAQBAJ&pg=GBS.PA3.w.23.0.8.0.1\\_33&hl=en](https://play.google.com/books/reader?id=3fHDCAAAQBAJ&pg=GBS.PA3.w.23.0.8.0.1_33&hl=en)

IBM. (n.d.). What is data migration? Haettu 8.11.2023 osoitteesta <https://www.ibm.com/topics/data-migration>

M-Files. (n.d.). M-Files Importer. Haettu 9.11.2023 osoitteesta <https://catalog.m-files.com/shop/m-files-document-importer/>

Microsoft. (n.d.). What is data migration? Haettu 8.11.2023 osoitteesta <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-data-migration>

NetApp. (n.d.). WHAT IS DATA MIGRATION? Haettu 9.11.2023 osoitteesta <https://www.netapp.com/data-management/what-is-data-migration/>

Prohorchik, K. (2021). A data migration strategy in six steps. <https://www.itransition.com/blog/data-migration-strategy>

Right People Group. (2023). 5 tärkeää vaihetta onnistuneeseen tiedonsiirtoon. <https://rightpeoplegroup.com/fi/5-vaihetta-onnistuneeseen-datamigraatioon/>

Talend. (n.d.). Understanding Data Migration: Strategy and Best Practices. Haettu 9.11.2023 osoitteesta <https://www.talend.com/resources/understanding-data-migration-strategies-best-practices/>

Trotino, G. (2023). Data Migration Testing: What You Need to Know. <https://www.k2view.com/blog/data-migration-testing/>