

Sami Mansikkala

LÄHTÖAINEISTON KERÄÄMINEN JA HALLINTA TEHDASSIMULOINNISSA

Tekniikan ja luonnontieteiden tiedekunta
Kandidaatintyö
Toukokuu 2019

TIIVISTELMÄ

Sami Mansikkala: Lähtöaineiston kerääminen ja hallinta tehdassimuloinnissa
Kandidaatintyö
Tampereen yliopisto
Konetekniikan kandidaatin tutkinto-ohjelma
Toukokuu 2019

Tehdassimulointi on voimakas työkalu tuotantojärjestelmien suunnitteluun. Sen avulla voidaan arvioida nykyisen järjestelmän toimintaa ja uuden tuotantojärjestelmän tarpeita. Tehdassimuloinnissa hyödynnetään tapahtumapohjaista simulointia sen vaihtelevan aika-askeleen vuoksi. Simulointia voidaan hyödyntää kertaluontoisissa tuotantojärjestelmien suunnitteluprojekteissa tai jatkuvasti osana päivittäistä toimintaa, mikä luo omat vaatimuksensa lähtöaineiston keräykselle ja hallinnalle. Tässä työssä perehdyttiin kirjallisuuden avulla lähtöaineiston keräämiseen ja hallintaan liittyviin ongelmiin ajankäytön ja datan laadun kannalta. Työssä etsittiin kirjallisuudesta ratkaisuja näihin ongelmiin ja pohdittiin lähtöaineiston keräämisen ja hallinnan tulevaisuutta.

Tehdassimuloinnissa lähtöaineiston keräämiseen ja hallintaan voi kuluu jopa 30% koko simulointiprojektin ajasta. Tämä rajoittaa simuloinnin käyttöä tuotantojärjestelmän käyttäytymisen jatkuvassa ennustamisessa. Eniten datan keräys- ja hallintaprosessia hidastavat huono datan saatavuus ja mallinnettavien yksityiskohtien määrä. Yksityiskohtien mallintaminen lisää datan keräykseen tarvittavaa työtä, ja se korostuu etenkin, jos data ei ole saatavilla. Monimutkaisemmissa tuotantojärjestelmissä ongelmaksi saattaa tulla datan päällekkäisyys, jolloin samaa dataa on saatavilla useasta eri lähteestä. Luotettavamman lähteen selvittämiseen kuluu aikaa ja samalla lähteiden toisistaan eroavat arvot lisäävät epäluotettavuutta keräysprosessia kohden. Samat ongelmat vaikuttavat myös datan laatuun. Korkealaatuista dataa tarvitaan luotettavien simulointitulosten saavuttamiseksi.

Näitä ongelmia voidaan ratkaista mm. ottamalla käyttöön systemaattinen tapa kerätä dataa ja erilaisilla ohjelmistoilla, jotka keräävät dataa suoraan niiden lähteistä simulointimalliin. Ohjelmistot vähentävät merkittävästi datan keruuseen kuluvaan aikaan ja mahdollistavat simuloinnin jatkuvan käytön, mutta vaativat korkealaatuista dataa. Näitä ohjelmistoja varten on luotu myös standardi datan jäsentelylle ja rakenteelle. Standardi helpottaa simulointiohjelmistojen ja datan keräysohjelmistojen välistä integraatiota ja sitä voidaan myös hyödyntää datan varastoimisessa, jolloin tarvittava data on helposti löydettävissä. Ajan käyttöä voidaan myös vähentää yksinkertaisesti paremmalla kommunikoinnilla datan käyttäjien ja tuottajien välillä. Tällä voi olla myös positiivinen vaikutus datan laatuun.

Avainsanat: Tehdassimulointi, datan keräys, datan hallinta, datan laatu

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

SISÄLLYSLUETTELO

1.	JOHDANTO	1
1.1	Tehdassimulointi.....	1
1.2	Tutkimusongelma	1
1.3	Työn rakenne	2
2.	SIMULOINTI TUOTANNON SUUNNITTELUSSA	3
2.1	Tapahtumapohjainen simulointi	3
2.1.1	Toimintaperiaate	3
2.1.2	Käyttökohteet ja hyödyt.....	3
2.2	Projektiluontoinen tuotannon kehittäminen	4
2.3	Simulointi jatkuvana prosessina osana tuotantoa	4
3.	HAASTEITA DATAN KERÄÄMISESSÄ JA HALLINNASSA	6
3.1	Suurimpia haasteita.....	6
3.2	Ajankäyttö lähtöaineiston keräämisessä ja hallinnassa.....	8
3.3	Datan saatavuus.....	9
3.4	Datan laatu	10
4.	TAPOJA DATAN KERÄYKSEN JA HALLINNAN PARANTAMISEKSI	11
4.1	Datan keräyksen malleja	11
4.2	Datan hallinnan ja keräyksen tasoja	13
4.3	Henkilöstön rooli ongelmien poistamisessa	14
5.	DATAN HALLINNAN OHJELMISTOJA	16
5.1	Automaattinen datan keräys	16
5.2	Core Manufacturing Simulation Data	16
5.3	Generic Data Management -työkalu	18

5.4 Knowledge Extraction -työkalu.....	18
6. PÄÄTELMÄT	19
LÄHTEET	21
LIITE A: DATAN KERÄYSPROSESSI (MUOKATTU LÄHTEESTÄ ONGGO & HILL 2014).....	23

LYHENTEET JA MERKINNÄT

CMSD	Core Manufacturing Simulation Data
CR	Critical Ratio
ERP	Enterprise Resource Planning
FIFO	First In First Out
GDM -työkalu	Generic Data Management -työkalu
IIoT	Teollinen Internet (Industrial Internet of Things)
JSON	JavaScript Object Notation
KE -työkalu	Knowledge Extraction -työkalu
LPT	Longest Processing Time
SISO	Simulation Interoperability Standards Organization
SPT	Shortest Processing Time
XML	Extensive Markup Language

1. JOHDANTO

1.1 Tehdassimulointi

Tehdassimulointi on kriittinen osa tuotantojärjestelmien suunnittelua nykyaikana. Tuotantolaitoksista tehdään nykyään yhä monimutkaisempia, minkä vuoksi niiden kehittämisen avuksi tarvitaan ohjelmia, joilla voidaan testata erilaisia skenaarioita optimaalisen ratkaisun löytämiseksi.

Tehdassimulointi on voimakas työkalu, joka voi nopeuttaa ja tehostaa suunnitteluprosessia valtavasti. Sitä käytetään mm. perustelevaan suuria investointeja yrityksissä, layout-suunnittelussa ja ennustamaan tulevaa (Banks 1998, s. 519). Tehdassimulointi vaatii aina onnistuakseen luotettavaa dataa kehitettävästä järjestelmästä. Muutoksia tehdään pohjautuen simuloinnin tuloksiin, jotka muodostetaan kerätyn lähtöaineiston perusteella. Siksi onkin tärkeää, että kerätty data on hyvälaatuista ja se on hyödyllistä järjestelmän suunnittelun kannalta. Datan keräykseen ja hallintaan voi kuluja jopa n. 30 % koko projektin ajasta (Skoogh, Johansson 2007), joten on myös ajankäytön kannalta tärkeää suorittaa tämä vaihe huolellisesti.

1.2 Tutkimusongelma

Moni asia kuitenkin hankaloittaa lähtöaineiston keräämistä ja hallintaa, mikä vaikuttaa siihen kuluvaan aikaan ja itse simuloinnin tuloksiin. Yleisiä ongelmia ovat mm. datan tarkkuus, suhteellisuus ja saatavuus (Banks 1998, s. 60; Skoogh & Johansson 2007). Näitä ongelmia ratkaisemalla pystytään lyhentämään lähtöaineiston keräämiseen ja hallintaan kuluvaan aikaan ja poistamaan epävarmuuksia simulointituloksista.

Lähtöaineiston keräämisessä ja hallinnassa kuluvaan aikaan pitäisi pystyä lyhentämään, jotta simulointia voitaisiin hyödyntää paremmin päivittäisessä toiminnassa tuotantolaitoksissa. Datan keräykseen ja prosessointiin käytetyn ajan lyhentyessä datan laatu saattaa kuitenkin kärsiä. Tämän työn tarkoituksena on vastata kirjallisuuden avulla, kuinka lähtöaineiston keräämiseen ja hallintaan kuluvaan aikaan voidaan lyhentää tinkimättä datan laadusta. Työssä keskitytään vain lähtöaineiston keräämiseen ja hallintaan, vaikka simulointiprojektin muillakin vaiheilla voidaan vaikuttaa koko projektiin kuluvaan aikaan ja simulointituloksien laatuun.

1.3 Työn rakenne

Toisessa luvussa käydään läpi mitä tapahtumapohjainen simulointi on ja sen käyttökoh- teita ja hyötyjä. Lisäksi tarkastellaan simulointia työkaluna projektiluontoisessa tuotan- non kehittämisessä ja jatkuvassa käytössä osana päivittäistä toimintaa. Kolmannessa luvussa esitellään suurimpia haasteita ja ongelmia lähtöaineiston keräämisessä ja hal- linnassa. Neljäs luku käsittelee erilaisia ratkaisuja edellä esiteltyihin ongelmiin lähtöai- neiston keräämiseen ja hallintaan kuluvan ajan vähentämiseksi ja laadun paranta- miseksi. Viidennessä luvussa tarkastellaan, millaisia ohjelmistoja on luotu lähtöaineiston keräämisen ja hallinnan tueksi. Lopuksi pohditaan miltä datan keräyksen ja hallinnan tulevaisuus näyttää.

2. SIMULOINTI TUOTANNON SUUNNITTELUSSA

2.1 Tapahtumapohjainen simulointi

2.1.1 Toimintaperiaate

Yleensä simuloinnissa käytetään vakioaika-askelta ja tämä soveltuu hyvin, kun tarkastellaan järjestelmiä, jotka muuttuvat jatkuvasti, kuten kappaletta, johon tuodaan lämpöä. Tuotantoa simuloidessa järjestelmän tila muuttuu kuitenkin satunnaisesti aina kun tuote tai osa saapuu malliin tai se valmistuu työvaiheesta ja siirtyy seuraavaan vaiheeseen. Jos tuotantoa simuloitaisiin vakioaika-askeleella, se olisi hyvin raskasta tietokoneelle ja hidasta, sillä aika-askelen pitäisi olla hyvin pieni tarkkojen tulosten saavuttamiseksi. Tämän vuoksi tuotannosimuloinnissa hyödynnetään tapahtumapohjaista simulointia, jossa aika-askel vaihtelee järjestelmän muutosten mukaan. Näin ei turhaan simuloida hetkiä, jolloin järjestelmässä ei tapahdu mitään, ja simulaatiota voidaan nopeuttaa. (Banks 1998; Brailsford et al. 2014)

2.1.2 Käyttökohteet ja hyödyt

Tapahtumapohjainen simulointi on osoittautunut erinomaiseksi työkaluksi tuotantojärjestelmien suunnittelussa ja analysoinnissa. Sitä voidaan myös hyödyntää mm. sairaaloiden suunnittelussa ja sotilastarkoituksissa, mutta tässä työssä sitä käsitellään tuotannon näkökulmasta. Tapahtumapohjaisella simuloinnilla voidaan suunnitella, montako laitetta tarvitaan vastaamaan kysyntää. Lisäksi sen avulla voidaan tarkastella muutosten pitkäaikaista hyötyä, kouluttaa henkilöstöä, testata erilaisia häiriötilanteita tuotannossa ja myös perustella tehtävien muutosten taloudellista hyötyä. Simuloinnin avulla on myös mahdollista nähdä mihin tuotantojärjestelmän pullonkaula syntyy. (Allen 2011)

Simuloinnin edut verrattuna matemaattisiin ja analyttisiin metodeihin verrattuna tulevat esille monimutkaisia järjestelmiä tarkastellessa. Simuloinnilla pystytään paremmin kuvaamaan mallin objektien satunnaisuutta ja se tarjoaa myös visuaalisen esityksen tarkasteltavasta järjestelmästä. Monimutkaisen järjestelmän teoreettinen tarkastelu vaatii liikaa oletuksia ja yksinkertaistusta, jolloin myös tulokset kärsivät. Matemaattisten ja teoreettisten mallien tarkkuus ei ole riittävää ajatellen optimointia, joka on iso osa tehdassi-
mulointia. (Wang & Chatwin 2005)

2.2 Projektiluontoinen tuotannon kehittäminen

Simulointiprojekti on monitieteellinen ja monivaiheinen ryhmätehtävä, joka suoritetaan iteroiden (Kehris, Doulgeri 2002). Sillä voidaan pyrkiä eri tavoitteisiin: tehokkaampaan tuotantoon, osaavampaan henkilöstöön tai johdon vakuuttamiseen, kuten edellä mainittiin (Allen 2011). Simulointiprojekti voidaan karkeasti jakaa viiteen eri vaiheeseen (Allen 2011):

1. tehtävien määrittäminen
2. lähtöaineiston keräys ja prosessointi
3. simulointi
4. tulosten analysointi
5. päätöksen teon tuki.

Ensimmäisessä vaiheessa määritetään tehtävät. Päätetään mitä kukakin tekee ja kuinka karkealla tasolla. Myös projektin aikataulu sovitaan ensimmäisessä vaiheessa. Toisessa vaiheessa kerätään lähtöaineisto ja muutetaan se simulointimallille sopivaan muotoon. Kolmas vaihe on mallin simulointi. Neljännessä vaiheessa analysoidaan simuloinnin tuloksia, ja niitä vertaillaan ja validoidaan, jotta voidaan olla varmoja niiden täsmällisyydestä. Lopuksi tuloksista luodaan erilaisia kaavioita ja taulukoita tulosten visualisointia varten ja päätöksenteon helpottamiseksi. (Allen 2011)

Ennen näitä vaiheita täytyy kuitenkin määrittellä ratkaistava ongelma ja kerätä ymmärrys käsiteltävästä järjestelmästä, jotta voidaan asettaa projektille oikeat tavoitteet. Ongelman määrittely ja ymmärrys järjestelmästä auttavat myös selvittämään, tarvitaanko simulointia ollenkaan ongelman ratkaisemiseksi (Banks 1998).

2.3 Simulointi jatkuvana prosessina osana tuotantoa

Kilpailun kasvaessa ja asiakasvaatimusten lisääntyessä ja muuttuessa nopeasti, on myös tarve nopealle kehitykselle, kustannustehokkuudelle ja joustavuudelle lisääntynyt. Näihin tarpeisiin voidaan vastata tapahtumapohjaisella simuloinnilla. Simulointia pyritään käyttämään osana päivittäistä toimintaa, kuten laitehuoltojen suunnittelussa, jotta voidaan ennakoida laitteiden vikaantumista (Negahban, Smith 2014). Simuloinnin käyttö jatkuvana osana toimintaa muuttaa datan keräyksen tarvetta radikaalisti. Aiemmin simulointiprojekteissa datan keräykseen on varattu paljon aikaa, mutta jatkuvassa käytössä tähän ei ole varaa. Simuloinnin päivittäistä käyttöä varten datan tulee olla korkealaatuisia ja sen tulisi olla aina saatavilla. Nämä vaatimukset voidaan saavuttaa automatisoidulla

datan hallinnalla ja simulointiohjelmien ja datan lähteiden välisellä yhteensopivuudella. (Bokrantz et al. 2018)

Simulointia voidaan myös hyödyntää reaaliaikaisessa tuotannosuunnittelussa. Tuotantoa voidaan ohjata yksinkertaisten sääntöjen tai periaatteiden avulla, jotka kertovat, mikä työ tehdään seuraavaksi. Esimerkiksi valinta voidaan tehdä lyhimmän työajan mukaan, tuotteen määräajan perusteella tai sen mukaan kuinka kauan työ on ollut jonossa. Nämä ovat kuitenkin varsin karkeita tapoja ohjata tuotantoa eikä niillä voi huomioida muutostiloja tuotannossa. Ohjauksen optimoimiseksi voidaankin hyödyntää tapahtumapohjaista simulointia ja Markov-prosessia. Markov-prosessissa järjestelmän tilaa tarkastellaan jokaisen tapahtuneen toiminnon jälkeen, ja päätetään mitä tehdään seuraavaksi. Simuloinnilla voidaan arvioida näitä päätöksiä ja optimoida prosessia pidemmälle. (Rose et al. 2017)

Rose et al. (2017) testasivat tätä lähestymistapaa ja vertasivat sitä perinteisempiin ohjaustapoihin. Vertailtavia ohjaustapoja olivat FIFO (First In First Out), SPT (Shortest Processing Time), LPT (Longest Processing Time) ja CR (Critical Ratio). Vertailtavia lukuja olivat keskimääräinen tahtiaika, keskimääräinen keskeneräinen työ ja valmistuneiden töiden määrä. Hyödyntämällä simulointia ja Markov-prosessia saavutettiin parhaat tahtijat ja pienin keskeneräinen tuotanto, mutta eniten töitä valmistui käyttäen SPT-menetelmää, jossa seuraava työ valittiin lyhimmän työajan mukaan. Ero valmistuneiden töiden määrässä SPT:n ja Markov-prosessin välillä ei kuitenkaan ollut kovin suuri.

3. HAASTEITA DATAN KERÄÄMISESSÄ JA HALLINNASSA

3.1 Suurimpia haasteita

Perera ja Liyanage (2000) listasivat seitsemän suurinta ongelmaa lähtöaineistoon liittyen ja pyysivät vuoden 1997 Winter Simulation Conferencen osallistujia asettamaan ne tärkeysjärjestykseen. Lista on esitelty taulukossa 1. Osallistujat arvioivat suurimmaksi ongelmaksi datan huonon saatavuuden. Simulointimallien yksityiskohtien lisääntyessä tarvitaan myös enemmän dataa. Laadukasta dataa voi olla hankala löytää tai sitä ei ole olemassa ollenkaan. Täysin uutta tuotantojärjestelmää suunniteltaessa ei välttämättä ole edes mahdollista tuottaa haluttua dataa. Tätä tulosta tukee myös Skooghin ja Johansonin tekemä tutkimus (2007), jossa kartoitettiin lähtöaineiston keräämiseen ja hallintaan kulunutta aikaa ja datan saatavuuden vaikutuksia käytettyyn aikaan. Tämän tutkimuksen tuloksia käsitellään tarkemmin luvuissa 3.2 ja 3.3.

Taulukko 1: Suurimmat ongelmat lähtöaineistoon liittyen (Muokattu lähteestä Perera ja Liyanage 2000)

Suurimmat ongelmat	Sijointus
Huono datan saatavuus	1
Mallin tarkat yksityiskohdat	2
Datan saatavuuden tunnistamisen vaikeus	3
Simuloitavan järjestelmän monimutkaisuus	4
Selkeiden tavoitteiden puute	5
Rajoitettu datan käsittelykyky	6
Väärin määritellyt ongelmat	7

Toiseksi tärkeimmäksi ongelmaksi listattiin simulointimallien monimutkaisuus. Mahdollisimman yksityiskohtainen malli ei ole aina paras ratkaisu, sillä se lisää kerättävän datan määrää ja sitä kautta datan keräykseen kuluva aika kasvaa, ja pienten yksityiskohtien mallintaminen ei välttämättä edes paranna simuloinnin tulosten tarkkuutta merkittäväällä tavalla.

Seuraavaksi listalla on lähteiden tunnistamisen vaikeus. Dataa voi joutua keräämään useista eri lähteistä, ja ne voivat vaihdella yksinkertaisista manuaalisista järjestelmistä

monimutkaisiin tietokonejärjestelmiin. Joskus sama data voi löytyä kahdesta eri lähteestä, mutta niillä on eri arvot. Tämä lisää epäluotettavuutta, ja lisää aikaa kuluu, kun selvitetään, kumpi lähteistä on oikeassa. Lähteitä voi olla myös vaikea tunnistaa, koska tarvittava data on niin raa'assa muodossa muun datan seassa, että se vaatii paljon käsittelyä ennen kuin se on käyttökelpoista. Esimerkiksi koneen käyntiajasta kerätään dataa, mutta datassa esiintyviä käyttökatkoksia ei ole jaoteltu mitenkään. Jos datan joukosta tarvitsisi esimerkiksi etsiä ajanjaksot, jolloin kone on vikaantuneena, se olisi erittäin vaikeaa. Myös Skoogh ja Johansson (2007) huomasivat, että simulointiprojekteissa paljon aikaa kuluu saatavilla olevan datan kartoittamiseen.

Neljänneksi listalle asetettiin käsiteltävän järjestelmän monimutkaisuus. Suurista ja monimutkaisista järjestelmistä voidaan kerätä paljon erilaista dataa, mutta se vaatii kehittyneet menetelmät ja jatkuvaa vertailua, jotta saadaan eheää ja täydellistä dataa. Suurien järjestelmien ongelma on myös se, että yleensä dataa kerätään kuitenkin vain kertaluontoisesti tarpeen mukaan, vaikka datan keräys voisi olla jatkuva prosessi.

Viidentenä listalla on selkeiden tavoitteiden puute. Tavoitteet vaikuttavat simulointimallin yksityiskohtiin ja sitä kautta datan keräykseen kuluvaan aikaan. Epäselvät tavoitteet voivat johtaa sopimattomiin yksityiskohtiin mallissa, joita ei oikeasti tarvita.

Listan lopun täydentävät rajoitettu datan käsittelykyky ja väärin määritellyt ongelmat. Usein dataa tarvitsee järjestellä ja käsitellä, jotta sitä voidaan käyttää simulointimallissa, ja tähän kuluu ylimääräistä aikaa. Onnistunut simulointiprojekti vaatii syvän ymmärryksen järjestelmän toiminnasta ja ongelmista. Huono ymmärrys lisää epäonnistumisen riskiä ja datan keräykseen kuluvaan aikaan (Tye & Perera 1997, Perera & Liyanage mukaan 2000). Kun ei ymmärretä ongelmia ja järjestelmää, on mahdollista ratkaista väärin asioita.

Listan lisäksi on myös muita ongelmia: Tiedossa olevat ongelmat datan keräyksessä vaikuttavat datan luotettavuuteen ja sitä kautta laatuun. Simulointiasiantuntijat tiedostavat ongelmat datan keräyksessä eivätkä luota kerättyyn dataan täysin. Sen vuoksi he konsultoivat asiantuntijoita datan eri lähteistä, mihin kuluu paljon aikaa. (Bokrantz et al. 2018)

Useat tahot voivat myös tarvita samaa dataa eri tarkoituksiin. Tämän vuoksi datan tulisi olla tarpeeksi korkealaatuista, jotta kaikki tahot voivat käyttää dataa ongelmitta. Datan suhteellisuuden vuoksi jollekin data voi olla riittävää ja oikeassa muodossa, mutta riittä-

mätöntä toiselle. (Bokrantz et al. 2018) Esimerkiksi huoltopäällikölle voi riittää paljon karkeampi esitys koneen käynti- ja vika-ajoista, kun taas simulointiasiantuntija joutuisi tekemään vielä ylimääräistä työtä voidakseen mallintaa koneen käyttäytymisen oikein.

Dataa on usein saatavilla eri lähteistä, kuten yrityksen tietojärjestelmistä (esimerkiksi ERP, eli Enterprise Resource Planning), ulkoisista lähteistä, simulointi- tai järjestelmä-asiantuntijoilta tai sitten data täytyy kerätä itse. Datan hallinnan kannalta useat lähteet aiheuttavat haasteita mm. datan monistumisena, ja osa datasta voi vaatia jalostusta ennen kuin se on käyttökelpoista simulointimallissa. (Skoogh et al. 2012)

3.2 Ajankäyttö lähtöaineiston keräämisessä ja hallinnassa

Skoogh ja Johansson (2007) selvittivät, mihin lähtöaineiston keräämisessä ja hallinnassa kuluu eniten aikaa. He jakoivat lähtöaineistoon liittyvät toiminnot 9 eri vaiheeseen ja tutkivat 15 eri simulointitapausta yrityksissä. Vaiheet alkoivat tarvittavan datan tunnistamisesta ja päättyivät datan lopulliseen dokumentaatioon ennen kuin se viedään simulointimalliin. Tutkimuksessa selvisi, että ylivoimaisesti eniten aikaa kuluu lähtöaineiston keräämiseen. Keskimäärin 50% ajasta kului lähtöaineen keräämiseen yritysten tietojärjestelmistä ja niistä puuttuvan datan keräämiseen mittauksin tai muin keinoin. Seuraavaksi eniten aikaa kului saatavilla olevan datan kartoittamiseen ja kerätyn datan analysointiin. Näihin kumpaankin vaiheeseen kului 10% koko lähtöaineiston keräämiseen ja hallintaan kuluneesta ajasta. Kooste eri vaiheisiin kuluvista ajoista on esitetty taulukossa 2.

Taulukko 2: Ajan käytön jakautuminen lähtöaineiston keräämisen ja hallinnan eri vaiheille (muokattu lähteestä Skoogh & Johansson 2007)

Vaihe	Osuus kuluneesta ajasta
Tarvittavan datan tunnistaminen	8%
Datan tarkkuusvaatimusten määrittely	2%
Saatavilla olevan datan kartoitus	10%
Datan keräysmetodien valinta	4%
Dokumentoinnin perustaminen	4%
Datan keräys	50%
Datan analysointi ja valmistelu	10%
Datan validointi	7%
Loppudokumentointi	5%

Skoogh ja Johansson (2007) tutkivat myös, minkä tyyppisen datan keräämiseen kului eniten aikaa. 42% keräysajasta kului prosessiaikojen keräämiseen, ja siihen pidettiin haastattelujen perusteella syynä prosessien rajojen epäselvyyttä. Ei osattu sanoa selvästi, milloin prosessi alkaa ja milloin se loppuu.

3.3 Datan saatavuus

Lisäksi Skoogh ja Johansson (2007) tutkivat eri datatyyppien saatavuutta yrityksissä ja huomasivat siinä puutteita. Vain yhdessä tutkimuksen 15 yrityksestä oli kaikki tarvittava data saatavilla, kun simulointiprojekti aloitettiin ja kahdessa yrityksessä ei ollut lainkaan dataa saatavilla. Parhaiten dataa oli saatavilla koneiden vikaantumisista (64% yrityksistä) ja heikoiten materiaalin käsittelystä (0%). Muiden datatyyppien saatavuudet on esitetty taulukossa 3.

Taulukko 3: Datatyyppien saatavuus (muokattu lähteestä Skoogh & Johansson 2007)

Datatyyppi	Kaikki data saatavilla	Ei mitään dataa saatavilla	Osa datasta saatavilla
Prosessiajat	33%	27%	40%
Vikaantumiset	64%	9%	27%
Tuotannosuunnittelu	18%	55%	27%
Materiaalin käsittely	0%	62,5%	37,5%
Asetusajat	22%	44%	34%
Työkalun vaihtoajat	20%	80%	0%
Organisaation data	40%	40%	20%

Datan saatavuudella on suuri vaikutus lähtöaineiston keräämiseen ja hallintaan kuluvaan aikaan. Tutkimuksessa kului keskimäärin alle viikko prosessiaikojen keräämiseen, kun kaikki data oli saatavilla ja hieman yli kolme viikkoa, kun tarvittiin manuaalista keräystä. Yritys, jolla oli kaikki tarvittava data saatavilla, käytti vain 12% koko projektin ajasta lähtöaineiston keräämiseen ja hallintaan, kun taas keskimäärin tutkimuksessa siihen kului 31%.

3.4 Datan laatu

Bokrantz et al. (2018) kartoittivat laatuongelmia tapahtumapohjaisessa simuloinnissa käytännön simulointiprojektien avulla. He luokittelivat ongelmia 11 laadun eri ulottuvuuteen, jotka Balci et al. (2000) määrittivät. Laadun ulottuvuudet ovat:

- saatavuus
- virheettömyys
- yksityiskohtaisuus
- tarkkuus
- selkeys
- täydellisyys
- yhtenäisyys
- ajantasaisuus
- oleellisuus
- luotettavuus
- jäljitettävyys.

Tutkimuksessa ongelmia löytyi jokaiseen ulottuvuuteen liittyen. Esimerkiksi datan keräyksessä ei huomioida simulointia ja simulointiasiantuntija joutuu käyttämään paljon aikaa saadakseen datan sopivaan muotoon itselleen. Tahtiaikojen määrittäminen oli myös hankalaa manuaalisen kokoonpanon aiheuttaman vaihtelun vuoksi. Datan laatua kyseenalaistettiin korjaus- ja laskentavaiheessa tehtävien oletusten ja arvioiden vuoksi. Myös datan luokittelu ja nimeäminen oli vajavaista, mikä heikensi datan täydellisyyttä. Simulointiasiantuntijat tiedostivat näitä ongelmia eivätkä siksi luottaneet dataan. Näihin ja muihin esille tulleisiin ongelmiin saatiin ratkaisuja konsultoimalla eri asiantuntijoita yrityksen sisällä, mutta se ei kuitenkaan poistanut ongelmien juurisyytä. Yleisesti ottaen kaikki nämä ongelmat ja niiden selvittäminen lisäävät tiedon keräämiseen ja hallintaan kuluvaa aikaa ja epäluotettavuutta simulointituloksissa. Datan laatuun vaikuttavat myös itse keräysprosessin ja datan dokumentoinnin laatu. (Bokrantz et al. 2018)

4. TAPOJA DATAN KERÄYKSEN JA HALLINAN PARANTAMISEKSI

4.1 Datan keräyksen malleja

Simulointiprojekti alkaa tavoitteiden asettamisella, mikä tehdään asiakkaan tarpeiden perusteella. Kun projektin tavoitteet on määritelty, voidaan aloittaa simulointimallin rakentaminen ja datan keräys yhtäaikaista. (Rabe et al. 2008) Näin säästetään aikaa sillä prosessit ole riippuvaisia toisistaan. Onggo ja Hill (2014) esittelivät kaavion datankeräysprosessin helpottamiseksi (Liite A). Kaavio alkaa esittämällä asiakkaalle tarkan kuvauksen tarvittavasta datasta. Jos data on saatavilla ja vastaa kaikkia vaatimuksia, se voidaan kerätä talteen, käsitellä simulointia varten ja validoida. Jos data ei vastaa vaatimuksia, tulee pohtia, voiko sitä silti käyttää. Tällaisen datan käyttö voi vaatia joidenkin oletuksien tekemistä. Jos dataa ei voi käyttää, pitää päättää, onko sen uudelleenkeräämiseen kuluva aika siitä saatavan hyödyn arvoista. Jos dataa ei ole edes saatavilla, keskustellaan asiakkaan kanssa mahdollisuuksista kerätä kyseistä dataa. On tärkeää muistaa, että datan ei voi olettaa olevan saatavilla vain, koska sen luullaan kuuluvan asiakkaan prosessiin.

Kuhnt ja Wenzel (2010) esittelivät hyvin samanlaisen lähestymistavan datan keräykseen. Heidän mallinsa on hieman suoraviivaisempi, mutta yksityiskohtaisempi kuin Onggon ja Hillin (2014) kaavio. Malli on esitetty kuvassa 1.



Kuva 1: Datan keräyksen malli (Muokattu lähteestä Kuhnt & Wenzel 2010)

Malli alkaa tavoitteiden asettamisella. Selkeät tavoitteet helpottavat datankeräysprosessia, koska silloin tiedetään pääpiirteittäin, mitä dataa tarvitaan. (Kuhnt & Wenzel 2010)

Kun tavoitteet ovat selvillä, on tehtävä tarkennuksia kerättävään dataan. Tässä vaiheessa määritetään tarvittavan datan tarkkuus, sopivuus ja esitysmuoto. Määritellään ja arvioidaan myös sopivia tiedonlähteitä, sekä selvitetään, mitä dataa on saatavilla tietokannoista ja mitä dataa täytyy tuottaa. (Kuhnt & Wenzel 2010)

Seuraavassa vaiheessa valmistaudutaan keräämään dataa. Tämä vaihe sisältää lähteiden valitsemista ja karsimista, pohdintaa datan keräystavoista ja resurssien käytön suunnittelua. Datan keräystapoja pohtiessa pitää ottaa huomioon niiden tarvitsema työvoima, kustannukset ja häiriöalttius. Datan keräystä voi myös automatisoida, jos se on mahdollista. Datan keräyksen voi automatisoida keräämällä datan esimerkiksi ERP -järjestelmästä käyttöliittymän välityksellä. Tämä nopeuttaa datan keruuta ja vähentää virheitä, mutta riskinä on datan päällekkäisyys ja datan tarkkuus voi myös olla heikompaa. (Kuhnt & Wenzel 2010)

Seuraavaksi tulee kerätä data. Jos dataa on kerätty haastattelemalla asiantuntijoita, vastaukset tiivistetään, tehdään yhdenmukaisiksi ja jäsennellään käypään muotoon. Dataa verrataan tarvittuun tietoon ja validoidaan. (Kuhnt & Wenzel 2010)

Kun data on kerätty, se jäsennellään analyysijä varten. Datalle suoritetaan tilastollinen analyysi, jonka avulla pyritään tunnistamaan piirteitä ja trendejä datasta. Analyysillä vahvistetaan myös, että tulokset pitävät paikkansa ja ne vastaavat odotuksia. Lopuksi vielä tarkistetaan, että data vastaa aluksi määritettyjä tarpeita. (Kuhnt & Wenzel 2010)

Perera ja Liyanage (2000) esittivät myös mallin listaamiensa ongelmien ratkaisuksi. Mallin tarkoituksena on nopeuttaa tarvittavan datan tunnistamista ja keräystä. Se perustuu funktionaalisiin malleihin tuotantojärjestelmän eri osista ja näiden mallien välisiin suhteisiin. Ne helpottavat hahmottamaan mitä dataa tarvitaan ja kuinka se saadaan kerättyä.

Funktionaalisilla malleilla pyritään mallintamaan päätöksiä, toimintoja ja tapahtumia järjestelmässä. Esimerkiksi yksi malli voi kuvata yhden työstökoneen vaiheita, sen vaatimia syötteitä ja tuottamia tuotoksia. Suhteita mallintamalla selviää materiaalivirrat järjestelmässä ja mitä tuotteita voidaan valmistaa milläkin koneella. Kartoitustaulukoon syötetään funktionaaliset mallit, niiden ominaisuudet ja niiden väliset suhteet. Kartoitustaulukon avulla nähdään nopeasti ja helposti mitä dataa tarvitaan ja mistä sitä voidaan kerätä. (Perera & Liyanage 2000)

Mallin käyttö alkaa tutkimalla järjestelmää ja tutustumalla siihen. Tarkoituksena on selvittää sen toimintaperiaatteet. Nämä toimintaperiaatteet sitten muutetaan funktionaaliksi malleiksi, joiden avulla luodaan kartoitustaulukko. Seuraavaksi taulukosta tunnistetaan tarvittava data, joka kerätään talteen. Malli tuo järjestelmällisen tavan lähestyä datan tunnistamista ja keräystä ja sen avulla luotu tietokanta voidaan yhdistää suoraan simulointimalleihin. Valmiita kirjastoja voi hyödyntää funktionaalisia malleja luodessa, mikä nopeuttaa prosessia. (Perera & Liyanage 2000)

4.2 Datan hallinnan ja keräyksen tasoja

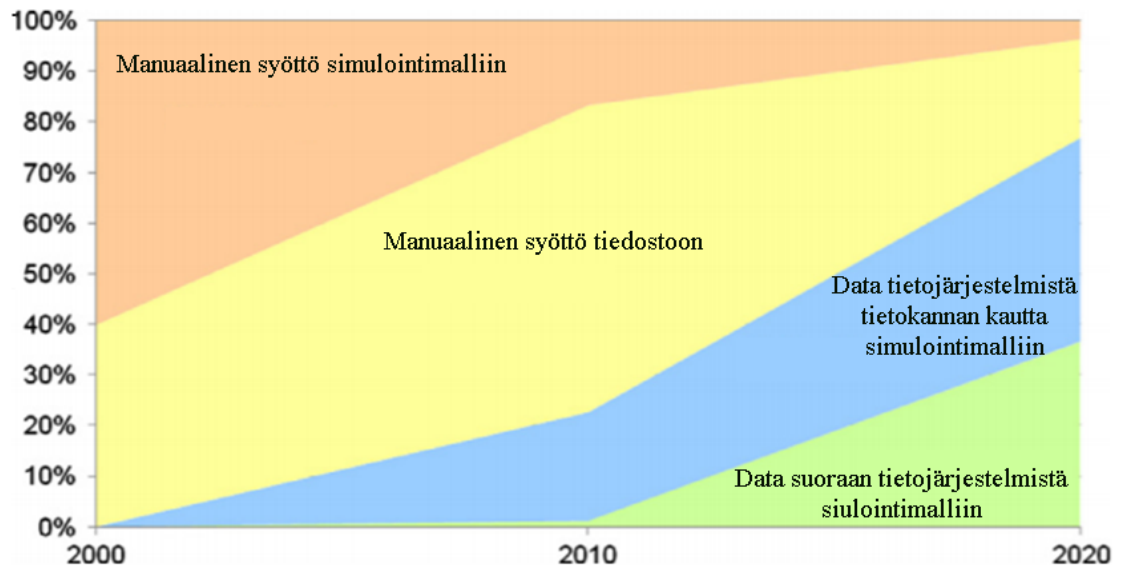
Kerätyn datan määrän kasvaessa on erittäin tärkeää hallinnoida ja varastoida dataa järkevästi, jotta oikea data löydetään helposti, kun sitä tarvitaan. Simulointimallia voidaan myös tarvita myöhemmin uudestaan, mitä varten tiedon ylläpitäminen ajan tasalla on hyödyllistä ja vaatii toimivaa datan hallintaa. (Skoogh et al. 2012)

Robertson ja Perera (2002) esittelivät neljä eri tapaa kerätä ja hallinnoida dataa simulointiprojektissa. Tutkimuksen tarkoituksena oli havainnollistaa silloista nykytilannetta datan keräyksen ja hallinnan suhteen ja esittää mahdollisia ratkaisuja niihin liittyviin ongelmiin. Vuosituhannen alussa oli yleistä kerätä tarvittava data manuaalisesti ja syöttää se joko suoraan simulointimalliin (tapa A) tai tiedostoon, esimerkiksi taulukkoon, josta data siirretään simulointimalliin (tapa B). Näiden tapojen etuna on niiden yksinkertaisuus, mutta ne ovat varsin alttiita inhimillisille virheille ja aikaa vieviä toimenpiteitä manuaalisen datan syötön vuoksi. Etenkin tapa A on joustamaton, koska data tallennetaan suoraan simulointimalliin. Tämä hankaloittaa mallin päivittämistä ja uudelleenkäyttöä. Tapa B on jo joustavampi ratkaisu, sillä data on tallennettu omaan tiedostoon, jossa se on organisoitu niin, että sitä on helppo lukea ja päivittää tarpeen mukaan.

Robertson ja Perera (2002) ennustivat tapojen A ja B suosion laskevan tapojen C ja D yleistyessä. Nämä tavat noutavat datan yrityksen tietojärjestelmästä, kuten ERP:stä. Tavassa C data kulkee simulointimalliin vielä erillisen tietokannan kautta, mutta tavassa D simulointimalli on linkitetty suoraan toiminnanohjausjärjestelmään. Tällä automaation asteella säästetään paljon aikaa ja välttyään inhimillisiltä näppäilyvirheiltä. Automaatio kuitenkin lisää toisenlaista epävarmuutta datan laadusta, sillä ohjelmat eivät osaa ottaa kantaa kerätyn datan arvoihin. Tällainen laaduntarkastus tulee suoritettua manuaalisessa menetelmässä samalla, kun arvoja syötetään taulukkoon tai simulointimalliin. Ohjelmat vain olettavat datan arvojen olevan oikein ja tarpeeksi korkealaatuista. Lisäksi sama data voi olla saatavilla useammasta eri paikasta, jolloin ei välttämättä ole varmuutta, käyttääkö simulointimalli niistä parhainta vaihtoehtoa. Muun muassa CMSD-standardi (Core Manufacturing Simulation Data) ja siihen pohjautuvat ohjelmat olettavat datan olevan tarpeeksi korkealaatuista käytettäväksi simulointimallissa (Bokrantz et al. 2018). Standardia käsitellään tarkemmin luvussa 5.2. Tapa D voi olla hankala ottaa käyttöön etenkin isommassa järjestelmässä sen monimutkaisuuden vuoksi.

Skoogh et al. (2012) palasivat Robertsonin ja Pereran (2002) tekemään tutkimukseen ja kartoittivat uudelleen datan keräyksessä ja hallinnassa käytettyjä tapoja kymmenen vuotta myöhemmin. Kymmenen vuoden aikana moni yritys oli siirtynyt keräämään dataa

erilliseen tietokantaan ja myös datan kerääminen automaattisesti toiminnanohjausjärjestelmistä oli yleistynyt. He myös tekivät myös vuodelle 2020 ennusteen, jonka mukaan dataa syötetään manuaalisesti simulointimalleihin enää 5%:ssa simulointiprojekteista. Ennusteen mukaan datan automaattinen kerääminen yritysten toiminnanohjausjärjestelmistä kasvaa 20%:sta noin 75%:iin. Tutkimusten tulokset ja ennuste vuodelle 2020 ovat nähtävissä kuvasta 2.



Kuva 2: *Datan keräyksen ja hallinnan kehitys (Skoogh et al. 2012)*

4.3 Henkilöstön rooli ongelmien poistamisessa

Lähtöaineiston keräämiseen ja hallintaan kuluva aika voidaan vähentää ja datan laatua parantaa myös yksinkertaisin keinoin. Kaikkien datan käyttäjien pitäisi yhdessä osallistua datan tuottoon ja kertoa mitä he vaativat datalta, jotta se olisi oikeanlaista jokaisen käyttötarkoitukseen. Tässä korostuu kommunikoinnin tärkeys eri osastojen välillä. Antamalla palautetta kerätystä datasta voidaan jatkuvasti parantaa datan keräysprosessia. Kommunikointi datan eri käyttäjien välillä auttaa muita hahmottamaan tarvittavan datan dimensioita ja tuottamaan sellaista dataa, jota kaikki voivat käyttää. Myös kouluttamalla henkilöstöä kerättävän datan tärkeydestä ja käyttötarkoituksista voidaan kehittää datan keräysprosessia. Kehittämällä datankeräysjärjestelmä, jota useampi taho voi käyttää ja joka on helppo ymmärtää ja muokata, voidaan myös parantaa datan laatua. (Bokrantz et al. 2018)

Simulointiasiantuntijoiden pitäisi ottaa aktiivinen rooli datan tuotannossa taatakseen luotettavuutta simulointituloksissa. Heidän tulisi painottaa raakadatan tärkeyttä yrityksessä ja jatkuvasti tunnistaa ja puuttua datan laatuongelmiin ja keskustella niistä datan tuottajien ja hallinnoitsijoiden kanssa. On tärkeää, että simulointiasiantuntijat tuovat esille datan tuottajille tarvitsemiaan asioita tuotettavasta datasta. Kerättyä dataa tulisi myös validoida erillään simulointimallin validoinnista ja siten, että kaikki tarvittava henkilöstön pätevyys otetaan huomioon. (Bokrantz et al. 2018)

5. DATAN HALLINNAN OHJELMISTOJA

5.1 Automaattinen datan keräys

Robertsonin ja Pereran (2002) esittelemät edistyneemmät tavat kerätä ja hallita dataa (tavat C ja D) vaativat hyvät ohjelmistot toimiakseen kunnolla. Tässä luvussa esitellään standardi, joka edistää eri ohjelmistojen yhteensopivuutta ja tarjoaa mallin kerättävän datan jäsentelylle. Tässä luvussa tutustutaan myös kahteen eri ohjelmistoon, jotka hyödyntävät tätä standardia ja yhdistävät yrityksen tietojärjestelmiä ja simulointiohjelmistoa (tapa C).

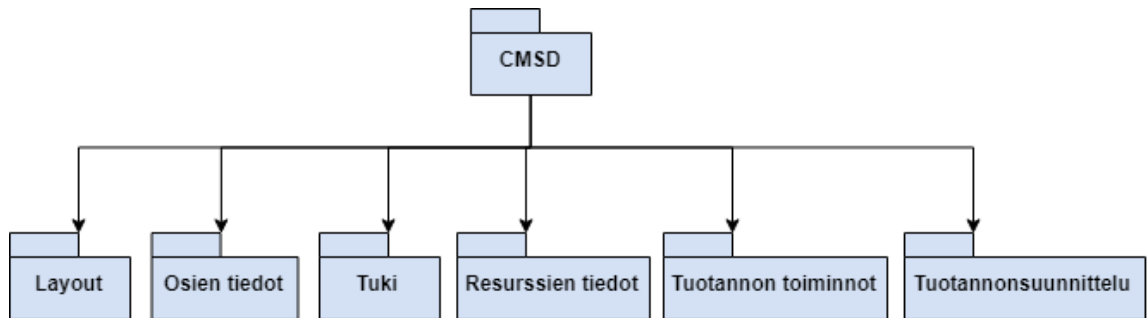
5.2 Core Manufacturing Simulation Data

Core Manufacturing Simulation data (CMSD) on standardi, jonka tarkoituksena on parantaa simulointiohjelmien ja muiden tuotannon tietojärjestelmien yhteentoimivuutta. Standardin on kehittänyt kansainvälinen Simulation Interoperability Standards Organization (SISO), joka pyrkii edistämään mallinnuksen ja simuloinnin yhteentoimivuutta ja uudelleenkäyttöä parantaakseen mallinnuksen ja simuloinnin laatua ja kustannustehokkuutta (SISO 2019).

Standardin avulla simuloinnissa tarvittava data voidaan kerätä sellaiseen muotoon, jota moni simulointiohjelma pystyy käyttämään, ja standardi toimii myös pohjana monelle tiedonhallintatyökalulle, jotka vähentävät tiedonkeruuseen ja -hallintaan kuluva aikaa. (KE -työkalu, GDM -työkalu). Nämä työkalut tallentavat tiedot standardin mukaiseen XML-tiedostoon, jota moni simulointiohjelma pystyy tulkitsemaan. Näitä työkaluja käsitellään lisää luvuissa 5.3 ja 5.4.

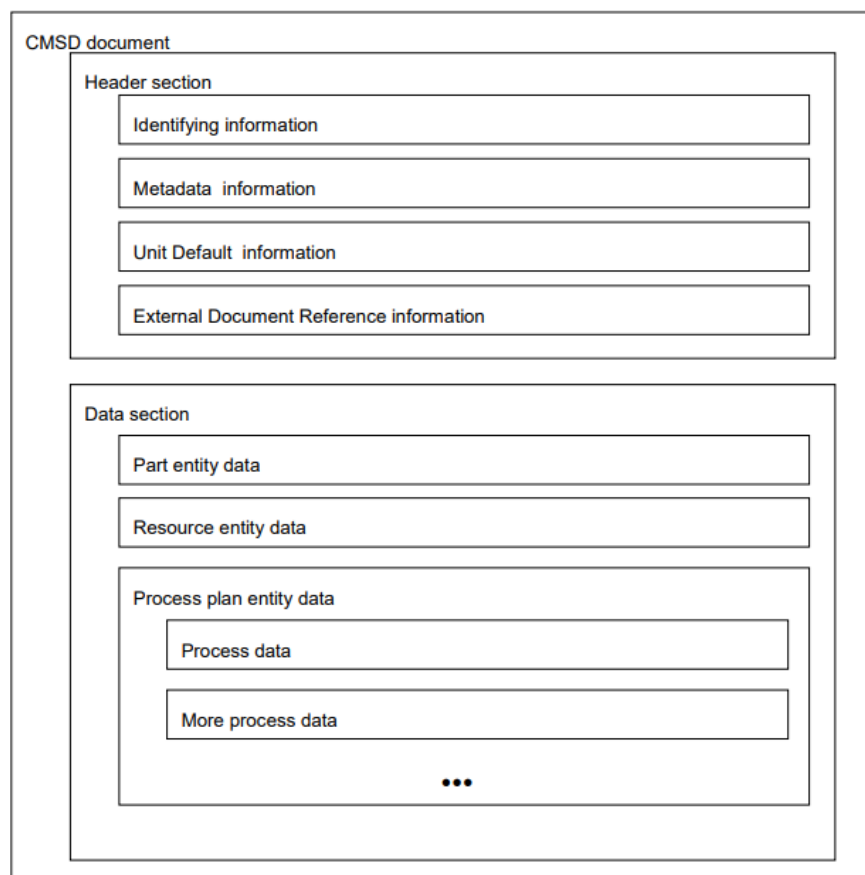
Standardin tietomalli on rakenne, joka sisältää välttämättömät objektit simulointiin, kuten tietoa osista, resursseista ja tuotannon toiminnoista. Rakenne myös kuvaa objektien välisiä suhteita, mikä tekee tiedonsiirrosta tehokkaampaa järjestelmien välillä. CMSD -malli

on jaettu tuotannon pääosa-alueiden perusteella kuuteen eri pakettiin, jotka sisältävät alipaketteja ja CMSD-dokumentteja (kuva 3). (SISO 2010)



Kuva 3: CMSD -malli (Muokattu lähteestä SISO 2010)

CMSD-dokumentti on kokoelma erilaisia objekteja, jotka liittyvät johonkin tuotannon vaiheeseen. Yhdessä dokumentissa voisi esimerkiksi olla yhteen laitteeseen liittyvää dataa, kuten asetusaikoja, työstöaikoja ja laitteen vikaantumisaikoja. Esimerkki CMSD-dokumentin rakenteesta on esitetty kuvassa 4. Dokumentti sisältää otsikko-osion (Header section), jossa on tietoa dokumentin sisällöstä ja rakenteesta. Dokumentin dataosiossa (Data section) on tuotantoprosessia koskevat tiedot. (SISO 2010)



Kuva 4: Esimerkki CMSD-dokumentin rakenteesta (SISO 2010)

5.3 Generic Data Management -työkalu

Generic Data Management -työkalu (GDM -työkalu) on kehitetty Chalmersin teknillisessä yliopistossa ja sen tarkoituksena on helpottaa lähtödatan hallintaa ja jäsentelyä, sekä vähentää aikaa, joka kuluu datan muokkaamiseen simulointimallille sopivaksi. Työkalu yhdistetään datan eri lähteisiin, joita se hallinnoi, ja käsittelee niiden tarjoamaa dataa. Koska työkalu on yhdistetty lähteisiin, sen avulla saadaan helposti tuorein data järjestelmästä simulointia varten. (Bengtsson et al. 2009)

GDM -Työkalu hyödyntää CMSD -standardia, jolloin dataa on helppo hyödyntää useassa eri simulointiohjelmassa, mikä tekee työkalusta sopivan vaihtoehdon moneen eri tapaukseen. Dataa voidaan tulkita useista erilaisista lähteistä erilaisten lisäosien avulla ja työkalu osaa tehdä datasta tilastollisia jakaumia. (Bengtsson et al. 2009)

5.4 Knowledge Extraction -työkalu

Knowledge Extraction -työkalu (KE -työkalu) on avoimen lähdekoodin ohjelma simulointin lähtöaineiston hallinnointiin. Se louhii raakadataa yrityksen lähteistä ja muuttaa sen simulointimallille käyttökelpoiseksi dataksi. Näin voidaan säästää datan käsittelyyn kuluva aikaa. Eräässä tapaustutkimuksessa lähtödatan keräämiseen ja hallintaan kului 81% vähemmän aikaa, kun käytettiin KE -työkalua. (Barlas ja Heavey 2016)

Data voidaan viedä työkalusta CMSD-standardin mukaiseen XML-tiedostoon, JSON-tiedostoon tai Excel-taulukkoon, josta on helppo tarkastella kerättyä dataa ja sen tilastollisia ominaisuuksia. Näin ollen työkalu on yhteensopiva monen simulointiohjelman kanssa. KE -työkalu on luotu Pythonilla hyödyntäen RPy2 -kirjastoa. Avoin lähdekoodi ja Python mahdollistavat työkalun kustomoinnin yrityskohtaisesti. KE -työkalu pohjautuu lähtödatan hallinnan rakenteeseen, jonka Skoogh ja Johansson määrittivät vuonna 2008. Rakennetta noudattaen työkalu koostuu kolmesta pääkomponentista, jotka ovat datan louhinta, datan käsittely ja outputin valmistelu. Rakenne on modulaarinen ja joustava, mikä edelleen helpottaa työkalun räätälöintiä omiin tarpeisiin. (Barlas ja Heavey 2016)

6. PÄÄTELMÄT

Tässä luvussa kerrataan työssä käsitellyjä asioita ja pohditaan, millaisia trendejä lähtöaineiston keräämiseen ja hallintaan liittyy. Luvussa käsitellään myös millaisia haasteita ja vaatimuksia nämä trendit luovat.

Tehdassimuloinnissa lähtöaineiston keräämiseen ja hallintaan kuluu n. kolmannes koko simulointiprojektiin kuluva ajasta. Eniten aikaa kuluu datan keräämiseen, koska tarvittavaa dataa ei ole valmiina saatavilla yrityksen tietokannasta. Datan saatuutta parantamalla voitaisiin merkittävästi lyhentää simulointiprojekteissa kuluva aikaa.

Teknologioiden kehittyessä on mahdollista kerätä yhä enemmän dataa. Teollinen internet (IIoT) mahdollistaa hyvin monenlaisen datan keräämisen tuotantojärjestelmistä, minkä vuoksi on tärkeää, että suuria datamääriä pystytään hallitsemaan ja varastomaan järkevästi ja tehokkaasti.

Saatavilla olevan datan määrän kasvaessa automaattinen datan keräys on kriittisessä roolissa, jotta simulointiprojekteissa käytetty aika datan keräämiseen ja hallintaan ei kasva liian suureksi. Kasvava datan määrä voi tuoda myös vaikeuksia tarvittavan datan tunnistamiseen. Tässä korostuu simulointiasiantuntijoiden ammattitaito kerätä tarpeeksi merkittävää dataa simulointitulosten kannalta. Simulointiprojektin alussa tehtävä ongelman ja tavoitteiden määrittäminen auttaa myös paremmin tunnistamaan projektin kannalta oleellista dataa.

Kun dataa on saatavilla yrityksen tietojärjestelmistä, se voidaan kerätä automaattisesti työkaluilla, kuten GDM -työkalu ja KE -työkalu, ja syöttää suoraan simulointimalliin, mikä nopeuttaa datan keräystä merkittävästi. Nämä ohjelmat kuitenkin olettavat kerätävän datan olevan korkealaatuista eivätkä ota kantaa datan laatuun. CMSD-standardiin pohjautuvat datan keräyksen ja hallinnan työkalut tulevat kehittymään ja yleistymään. Niitä mahdollisesti integroidaan osaksi simulointiohjelmiä, jotta dataa voidaan kerätä suoraan yrityksen tietojärjestelmistä simulointimalleihin.

Keräysprosessien kehittyessä simulointia voidaan paremmin hyödyntää esimerkiksi digitaalisessa kaksosessa ennustamaan ja arvioimaan oikean järjestelmän toimintaa ja suorituskykyä. Digitaalisessa kaksosessa suoritettavat simuloinnit tarvitsevat mahdoli-

simman tuoretta dataa, jotta tehdyt ennustukset ja arviot olisivat mahdollisimman tarkkoja. Teollinen internet mahdollistaa reaaliaikaisen datan käytön simuloinnissa. Se yhdistää koneita, ihmisiä ja muita tuotannon tekijöitä niin, että dataa on nopeasti saatavilla.

Yrityksissä on datan laatuun liittyen monenlaisia ongelmia, jotka yhdessä aiheuttavat epäluottamusta kaikesta kerätystä datasta. Luottamuksen saamiseksi simulointiasiantuntijat konsultoivat organisaation eri alojen asiantuntijoita saadakseen varmuuden datan laadusta, mihin kuluu paljon aikaa. Näiden ongelmien ehkäisemiseksi kommunikointi eri osastojen välillä ennen datan keräystä on tärkeää, jotta datan tuottajat tietävät mitä datan käyttäjät haluavat datalta ja miksi. Läpinäkyvyys datan tuottoprosessissa lisää luottamusta kerättyyn dataan ja parantaa datan laatua muillakin osa-alueilla. Myös ottamalla käyttöön erilaisia toimintamalleja datan keräykseen voidaan vaikuttaa käytettyyn aikaan ja kerätyn datan laatuun.

LÄHTEET

Allen, T., 2011. Introduction to Discrete Event Simulation and Agent-based Modeling. London: Springer.

Balci, O., Ormsby, W.F., Carr, J.T., III and Saadi, S.D., 2000. Planning for Verification, Validation, and Accreditation of Modeling and Simulation Applications, 2000, Society for Computer Simulation International, pp. 829–839.

Banks, J., 1998. Handbook of Simulation - Principles, Methodology, Advances, Applications, and Practice. John Wiley & Sons.

Barlas, P. and Heavey, C., 2016. KE tool, IEEE Press, pp. 472-483.

Bengtsson, N., Shao, G., Johansson, B., Lee, Y., Leong, S., Skoogh, A. and Mclean, C., 2009. Input data management methodology for discrete event simulation, Winter Simulation Conference, pp. 1335-1344.

Bokrantz, J., Skoogh, A., Lämkuill, D., Hanna, A. and Perera, T., 2018. Data quality problems in discrete event simulation of manufacturing operations. SIMULATION, **94**(11), pp. 1009-1025.

Brailsford, S., Churilov, L. and Dangerfield, B., 2014. Discrete-Event Simulation and System Dynamics for Management Decision Making. New York, United Kingdom: John Wiley & Sons, Incorporated.

Kehris, E. and Doulgeri, Z., 2002. Improving Simulation Project Efficiency Using Web Technology. SIMULATION, **78**(9), pp. 568-579.

Kuhnt, S. and Wenzel, S., 2010. Information acquisition for modelling and simulation of logistics networks. Journal of Simulation, **4**(2), pp. 109-115.

Negahban, A. and Smith, J.S., 2014. Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems, **33**(2), pp. 241-261

Onggo, B.S.S. and Hill, J., 2014. Data identification and data collection methods in simulation: a case study at ORH Ltd. Abingdon: Taylor & Francis Ltd.

Perera, T. and Liyanage, K., 2000. Methodology for rapid identification and collection of input data in the simulation of manufacturing systems. Simulation Practice and Theory, **7**(7), pp. 645-656.

Rabe, M., Spieckermann, S. and Wenzel, S., 2008. A new procedure model for verification and validation in production and logistics simulation, Winter Simulation Conference 2008, Winter Simulation Conference, pp. 1717-1726.

Robertson, N. and Perera, T., 2002. Automated data collection for simulation? Simulation Practice and Theory, **9**(6), pp. 349-364.

Rose, O., Shufang, X. and Tao, Z., 2017. Real-time job shop scheduling based on simulation and Markov decision processes, IEEE, pp. 3899-3907.

SISO – Simulation Interoperability Standards Organization, 2010, Standard for: core manufacturing simulation data - UML Model, CMSD Product Development Group.

SISO – Simulation Interoperability Standards Organization, Policies & Procedures, saatavilla: <https://www.sisostds.org/AboutSISO/PoliciesProcedures.aspx>, viitattu 3.4.2019

Skoogh, A. and Johansson, B., 2008. A methodology for input data management in discrete event simulation projects, Winter Simulation Conference, IEEE, pp. 1727-1735.

Skoogh, A. and Johansson, B., 2007. Time-consumption analysis of input data activities in discrete event simulation projects, Proceedings of the Swedish Production Symposium 2007, pp. 349-364.

Skoogh, A., Johansson, B. and Perera, T., 2012. Input data management in simulation – Industrial practices and future trends. Simulation Modelling Practice and Theory, **29**, pp. 181-192.

Wang, Q. and Chatwin, C.R., 2005. Key issues and developments in modelling and simulation-based methodologies for manufacturing systems analysis, design and performance evaluation. The International Journal of Advanced Manufacturing Technology, **25**(11), pp. 1254-1265.

LIITE A: DATAN KERÄYSPROSESSI (MUOKATTU LÄHTEESTÄ ONGGO & HILL 2014)

