

Digitaalisten dokumenttien hallinta - Lisääkö (digitaalinen) tieto tuskaa?

Tietoa käsitellään entistä enemmän digitaalisessa muodossa ja tallennetun tiedon määrä kasvaa kiihtyvällä vauhdilla. Valtaosa organisaatioiden tallennetusta tiedosta on dokumenteissa. Toisaalta oppivat organisaatiot ja hiljainen tieto ja tietämys ovat olleet vahvasti esillä viime vuosina. Mistä ja miten merkittävistä asioista oikein on kysymys, miten organisaatioiden informaation hallinnasta saa otteen, mikä on tärkeää ja mitä pitäisi tehdä? Lisääkö tieto ja tiedon digitaalisuus tuskaa vai vähentääkö se sitä?

1 Tietotulva

Mikä merkitys dokumenteilla ja sisällöllä on tietointensiivisille organisaatioille? Eri lähteiden [1] [2] mukaan tietointensiivisissä organisaatioissa:

- tiedon hakuun käytetään 16 % työajasta,
- 10 - 15 % organisaatioiden tuloista käytetään dokumenttien luomiseen, hallitsemiseen ja jakeluun, ja
- dokumenttien käsittelyyn käytetään 60 % työajasta.

Erään oman tapaustutkimuksemme [3] mukaan organisaation ihmiset lähettivät ja vastaanottivat keskimäärin 92 sivullista tietoa työpäivässä. Toisessa organisaatiossa [4] luku oli 74 sivullista. Joissakin puhtaasti tietoa käsittelevissä organisaatioissa luku saattaa nousta yli 200:n sivun.

Vastaavasti, 1900-luvun puolivälin jälkeen arvioitiin, että viimeisen 50 vuoden aikaan tietoa oli tuotettu enemmän kuin maailmanhistoriassa siihen saakka. 1985 arvioitiin tiedon määrän kaksinkertaistuvan organisaatioissa viidessä vuodessa. 1994 tämä aika-arvio oli enää 9 kuukautta. Myöhemmin 6 kuukautta. Tiedon määrän kaksinkertaistumiseen menevä aika on siis lyhentynyt radikaalisti näissä arvioissa.

Tietotekniikan kehittymisen myötä tietokoneiden kapasiteetti on noussut jatkuvasti. Mooren laiksi kutsutun nyrkkisäännön mukaan teollisessa massatuotannossa olevien integroitujen piirien pakkaustiheys kaksinkertaistuu noin puolentoista - parin vuoden

välein. Käytännössä tämä tarkoittaa sitä, että tietokoneiden teho ja tallennuskapasiteetti ovat satakertaistuneet kymmenen vuoden välein. Kun tiedon massatallennusvälineiden kapasiteetti on noin satakertainen tietokoneen keskusmuistiin nähden, kaikki se tieto, jota tallennettiin tietokoneen massamuisteille kymmenen vuotta sitten mahtuu nyt tietokoneen keskusmuistiin.

Kun tietotekniikka kehittyy näin huimalla vauhdilla, ratkaiseeko se tietotulvan? Riittääkö, että saamme käyttöömmme uudet koneet, joiden kapasiteetti riittää? Itse asiassa tilanne on ainakin osittain juuri päinvastainen. Tietotekniikan kehittyminen on johtanut ihmiset toimimaan tavalla, jota ihmiset itse eivät hallitse. Tietojenkäsittelymenetelmien puute lisää ihmisten tuskaa, vaikka digitaalisuus voisi sitä vähentää.

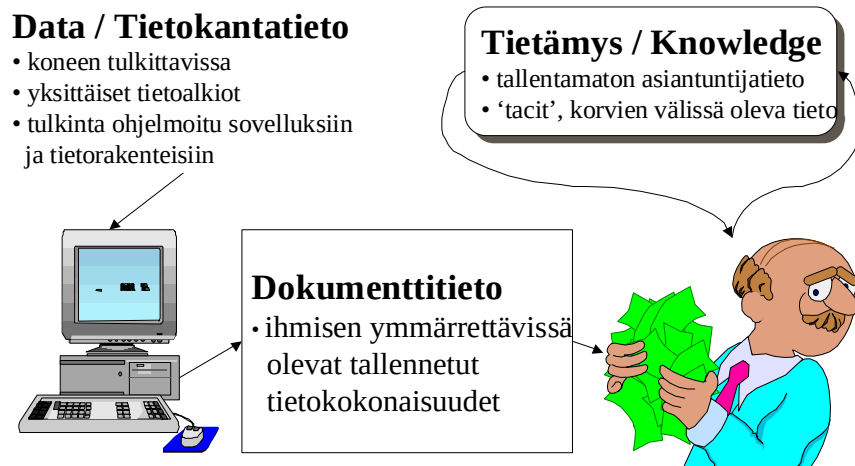
Tämä esitys purkaa auki tietotulvan problematiikka hieman tarkemmalla tasolla. Aluksi käydään käsiksi terminologiaan, sitten tilannetta tarkastellaan esimerkkien, havaintojen ja mittaustulosten valossa ja lopuksi esitellään lähestymistapa ja suuntaviivoja organisaation tietojen hallinnan kehittämiseen.

2 Tieto, tietämys ja informaatio organisaatiossa

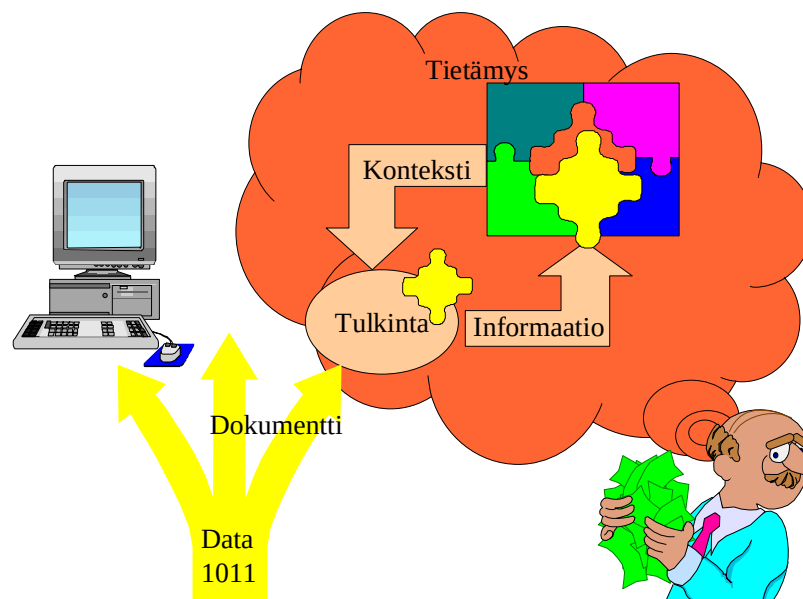
Tietoon liittyviä termejä käytetään usein eri merkityksissä. Tässä käytämme oheisia määritelmiä, jotka perustuvat lähinnä Shanonin informaatioteorian näkökulmaan:

- **Data** tarkoittaa tässä symboleja, joilla on tietty merkitys.
- **Tietämys** tarkoittaa ihmisen kognitiivisia malleja ja kokemusta maailmasta.
- Data sisältää **informaatiota** jollekin ihmiselle, jos hänen tietämyksensä kasvaa datan tulkinnan perustella - siis jos hän kykenee sidostamaan sen olemassa oleviin tietämysrakenteihinsa ja se sisälsi hänelle uutta informaatiota. Ihminen kykenee siis dataa tulkitessaan antamaan sille kontekstin ja tuo asiayhteys eli **konteksti** antaa datalle oikean tulkinnan.
- **Tieto** tulkitaan tässä huonosti määritellyksi yleiskäsitteiksi, joka kattaa tietämyksen sekä usein myös datan ja informaation.
- **Dokumentti** (määritellään jäljempänä tarkemmin) on väljästi määritellen pysyvä, ensisijaisesti ihmiselle tarkoitettu tietokokonaisuus, jossa data on kontekstissaan eli asiayhteydessään.
- **Sisältö** on väljästi määritellen dokumenttitiedon synonyymi, ihmisille julkaistavaksi tarkoitettua tekstiä ja muuta mediaa, joskaan ei välttämättä asiayhteydessään.

Kuva 1 havainnollistaa näiden käsitteiden suhdetta organisaation tietovarannon näkökulmasta. Kuva 2 taas havainnollistaa tarkemmin datan, tietämyksen ja informaation suhdetta.



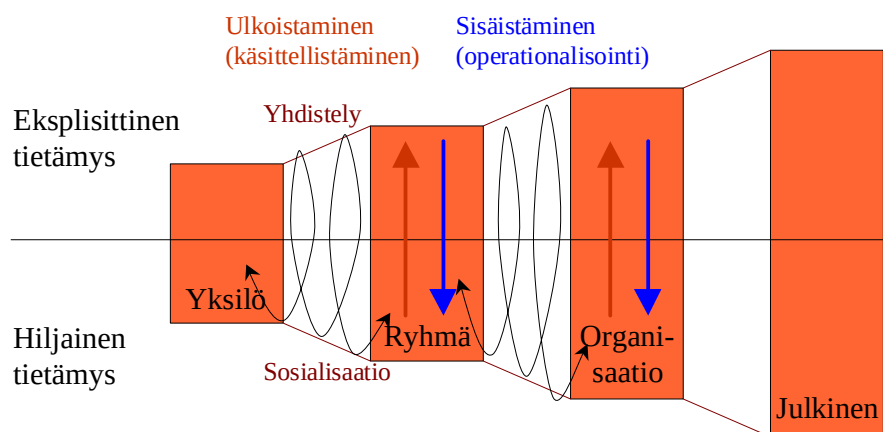
Kuva 1. Data, dokumentit ja tietämys [5].



Kuva 2. Tieto, data, informaatio ja tietämys [6] [7], .

Eksplisiittisellä tietämyksellä tarkoitetaan symbolisessa muodossa esitettyä tietämystä - dokumentteja, kaavioita, ohjeita - joka on siis monistettavissa ja jaettavissa. Hiljaisella tietämyksellä (tacit knowledge) eli hiljaisella tiedolla tarkoitetaan ihmisten kykyä toimia menestyksellisesti ilman, että he tiedostavat toimintaperiaatteitaan. Esimerkkinä voisi olla vaikkapa lääkärin tekemä oikea diagnoosi, jota hän ei välttämättä kykene perustelemaan sanallisesti.

Pelkistetysti sanoen Japanilaiset yritykset luovat uutta tietämystä ulkoistamalla, eli muuntamalla hiljaista tietämystä eksplisiittiseksi tietämykseksi ([8], s. 11, kuvan 3 nuoli ylöspäin). Hiljaisen tietämyksen käsitteellistämistä voi seurata uuden tiedon luominen tietoja yhdistelemällä. Eksplisiittistä tietoa otetaan käyttöön organisaatiossa sisäistämällä se eli muuntamalla se operationaaliseen muotoon, toimintatavoiksi (kuva 3, nuoli alas). Toinen tapa siirtää hiljaista tietämystä on sosialisatio, eli toimintamallien ja muun hiljaisen tietämyksen oppiminen yhdessä toimimalla.



Kuva 3. Tietämyksen muodostuminen ([6] mukaan [8]).

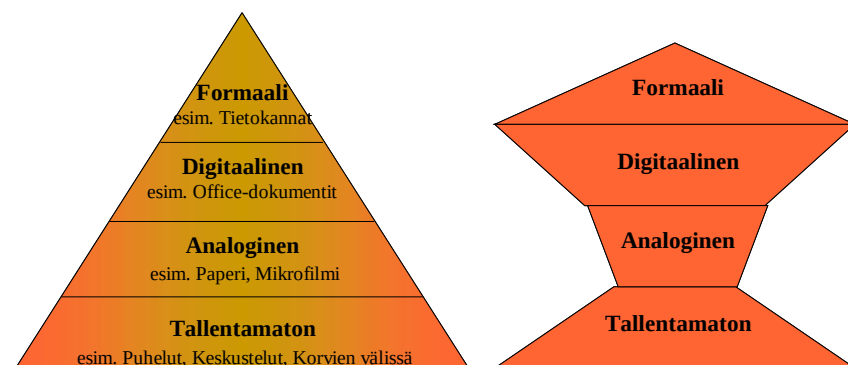
Länsimaaisessa kulttuurissa korostuu Japania enemmän eksplisiittisen tiedon käyttö [8] eli myös tietämyksen siirto dokumenttien kautta.

Organisaatioissa käsiteltävän tiedon määrä vaihtelee varsin paljon. Eräässä omassa tutkimuksessamme arvioimme tyypillisen tietointensiivisen organisaation jokaisen ihmisen saavan ja lähettävän päivittäin tietoa määrän, joka vastaisi noin 90 sivullista tekstiä – mukaan lukien suullisesti, dokumentteina ja digitaalisesti välitetty tieto. Kuinka suuri osa esimerkiksi sähköposteista todella luetaan, onkin sitten jo toinen kysymys. Toisessa organisaatiossa, jossa ihmiset lähinnä vain käsittelevät tietoa, määrä oli vielä kaksin-kolminkertainen edelliseen verrattuna.

Missä muodossa organisaatioiden tieto ja tietämys ovat?

- Vain pieni osa (arviolta 20 %) organisaatioiden tietämyksestä on tallennettu - loppu tietämys liikkuu verbaalisesti käytäväkeskusteluissa tai on yksilöiden käytössä hiljaisen tietämyksen muodossa - ilman, että sitä esitetään sanallisesti [9] [1].
- Tallennetuista tiedoista 80 % on dokumenteissa (1995) ja vain 15 % tietokannoissa [2] [10].

- Dokumenteista valtaosa on ollut analogisessa muodossa, yleisimmin paperilla tai mikrofilmillä, mutta nyt on siirrytty digitaaliseen tiedon tuottamiseen (joskin eri lähteet mittaavat hieman eri asioita):
 - 1987 [9]: 99 % analoginen (95 % paperi), 1 % elektroninen
 - 1991 [11]: 97 % analoginen (90 % paperi), 3 % digitaalinen
 - 2000 [12]: 0.003 % paperia (kaikesta tuotetusta tiedosta)
- Esimerkiksi USA:ssa tuotettiin yli 1'000'000'000 sivua paperia päivässä 1996 [13]. Maailmassa tuotettiin kaikkiaan vuosittain 1-2 exabyteä (2'000'000'000'000'000'000 byteä) tietoa 1997 [12] eli noin 250 megatavua henkilöä kohti.
- Dokumenttien määrä kaksinkertaistuu organisaatioissa [2] [1]
 - 1985: 5 vuodessa
 - 1994: 9 kuukaudessa
- Digitaalisten dokumenttien osuus on kuitenkin muutamassa vuodessa noussut oleellisesti.
- Valtaosa (80 %) digitaalisessa muodossa olevasta tiedosta on henkilökohtaisissa tietovarastoissa, PC:n kiintolevyillä [2] [1].



Kuva 4. Organisaatioiden tiedon pyramidin muuttuminen.

Näiden lukujen perusteella kuva 4 esittää tiivistetysti tietojenkäsittelyn tilannetta organisaatioissa 1990-luvulla (vasemmalla) ja 2000-luvun ensimmäisen vuosikymmenen lopulla (oikealla). 90-luvulla tieto oli karkeasti ottaen jaettavissa neljään kerrokseen, joista kunkin tason sisältämä tietomäärä oli noin nelinkertainen ylempien tasojen yhteiseen tietomäärään. Siis tallentamatonta tietoa oli nelinkertaisesti tallennettuun verrattuna, analogisesti tallennettua nelinkertaisesti digitaaliseen tallennettuun verrattuna ja tekstimuotoista digitaalista tietoa tietokannanhallintajärjestelmillä hallittuun formaaliin tietokantatietoon verrattuna. Digitaalisen tiedon määrän ylittäessä analogisen tallennetun tiedon määrän tilanne muuttuu suunnilleen oikean puoleisen kuvan mukaiseksi.

Tiedon pyramidia voidaan käyttää myös kuvaamaan eri tieteen- ja teknologian alueiden painotuksia:

- Informaatioteknologia (IT) painottaa ylintä ja osin toiseksi ylintä osaa pyramidista – miten tietokoneita ja verkkoja käytetään näihin tarkoituksiin.
- Dokumenttien hallinta painottaa kahta keskimmäistä kerrosta.
- Sisällön hallinta (content management) painottaa kahta ylimmäistä kerrosta.
- Tietämyksen hallinta (knowledge management) painottaa alinta kerrosta ja muiden kerrosten käyttöä sen apukeinoina.
- Informaation hallinta (IM, IS) kattaa periaatteessa kaikki tasot. Käytännössä organisaatioissa on kuitenkin vain IT-ryhmiä ja osastoja, ei IS-organisaatiota.

Kuvassa 5 esitetään organisaatioissa käsiteltävän tiedon ja tietämyksen kokonaisuutta.



Kuva 5. Tiedon varastointi ja kommunikointi organisaatioissa [6].

3 Dokumentit ja dokumenttien hallinta

Dokumenteilla on useita rooleja organisaatioissa. Dokumentti on

- **operatiivisen toiminnan välikappale**, esimerkkeinä tilaukset, projektisuunnitelmat, kirjeenvaihto, laskutus...
- **yrityksen tiedon ja osaamisen (tietämyksen) varasto**, esimerkkeinä tuotetieto, asiakastiedot, ohjeet, jne. eli tiedot, jotka yleensä ovat muutoin henkilöriippuvia tietoja, joiden saatavuus on kiinni asian tietävän henkilön paikallaolosta.
- **tuote tai osa tuotetta**, kuten on-line -dokumentit, käyttöohjeet, markkinatutkimukset, suunnitelmat, ...

Mutta mikä on "dokumentti"? Tässä käytämme seuraavaa määritelmää:

Dokumentti on

- **ensisijaisesti ihmisen havainnoitavaksi tarkoitettu**
vaikkakin usein tietokoneella käsiteltävä tai tietokoneen käyttämä
- **ihmisen kannalta merkityksellinen tietokokonaisuus,**
kokonaisuuden laajuus vaihtelee, mutta tietojen pitää ihmisen mielestä liittyä toisiinsa
- **josta voidaan esittää toistuvasti**
dokumenttia ei välttämättä tallenneta yhtenä kokonaisuutena, vaan se voidaan koostaa osista tai tuottaa dynaamisesti muista tiedoista useita kertoja samanlaisena, jos tuottamiseen tarvittavat tiedot ja mekanismi ovat käytettävissä
- **yksi tai useampia näkymiä.**
samaa tietokokonaisuuteen voidaan usein tuottaa erilaisia suodatettuja ja vaihtoehtoisia “näkymiä”, joissa on visuaalisten elementtien lisäksi muitakin ihmisen aistittavia medioita.

Eri organisaatioissa ja eri ihmisille dokumentti merkitsee eri asioita. Esimerkiksi joku voi pitää teknisiä piirrustuksia dokumentteina, toinen ei. Samoin on www-sivujen, pankkiautomaattikuittien, sähköposti-viestien tai automaatiojärjestelmän näyttöjen ja näyttökopioiden laita.

Dokumenttien hallinnan ja tietokantatiedon hallinnan välillä on oleellisia eroja, jotka aiheutuvat dokumenttitiedon perusominaisuuksista. Näitä ovat seuraavat:

Kokonaisuus. Dokumentti on laaja kokonaisuus, jossa data on kontekstissaan eli ihminen kykenee tulkitsemaan tietoa sen aihepiirin perusteella. Kokonaisuus voidaan myös pilkkoa osiin esimerkiksi eri medioiden käytön vuoksi, mutta liiketoimintaprosessien näkökulmasta osia halutaan yleensä hallita yhtenä loogisena kokonaisuutena.

Elinkaari. Dokumentin tietosisältö, tallennusmedia ja esitysmuoto voivat vaihtua moneen kertaan ennen dokumentin hävittämistä. Sisällölliset muutokset pakottavat yleensä versioimaan dokumentteja ja käsittelemään niiden tilaa suhteessa organisaation liiketoimintaprosesseihin.

Konfiguraation hallinta. Kun elinkaareen yhdistetään kokonaisuus - osa -suhde, päädytään jo konfiguraation hallintaan. Täytyy kyetä hallitsemaan versiot rinnakkaiset variantit, replikaatit jne. Konfiguraatioiden olemassaolo vaikeuttaa myös tietoturvan hallintaa dokumenttia pienempien osien suhteen.

Dokumenttien hallinta (DH) tukee yrityksen liiketoimintaa, prosesseja ja tuotteita hallitsemalla dokumenttitietoa. Dokumenttien hallinnan sisältö vaihtelee sen mukaan, mitä dokumentilla ja sisällöllä ymmärretään tiettyssä organisaatiossa. Dokumenttien hallinta on siis yksi osa yrityksen informaation hallintaa.

Dokumenttien hallinnan kehittäminen ei rajaudu olemassa olevien dokumenttityyppien hallinnan kehittämiseen, vaan ottaa kantaa myös siihen:

- mitä tallentamatonta tietoa ja tietämystä kannattaisi koota ja tallentaa dokumenttien muodossa,
- mitä dokumentteja kannattaisi muuntaa formaaliksi tiedoksi esimerkiksi siirtämällä se tietokantoihin,
- mitä dokumentteja ei enää kannata käyttää.

DH:n kehittämisessä lähtökohtana ovat organisaation tavoitteet ja sen tiedon hallinnan tavoitteet. Resurssien oikeaan kohdentamiseen voidaan käyttää myös tiedon käytettävyyden analysointia.

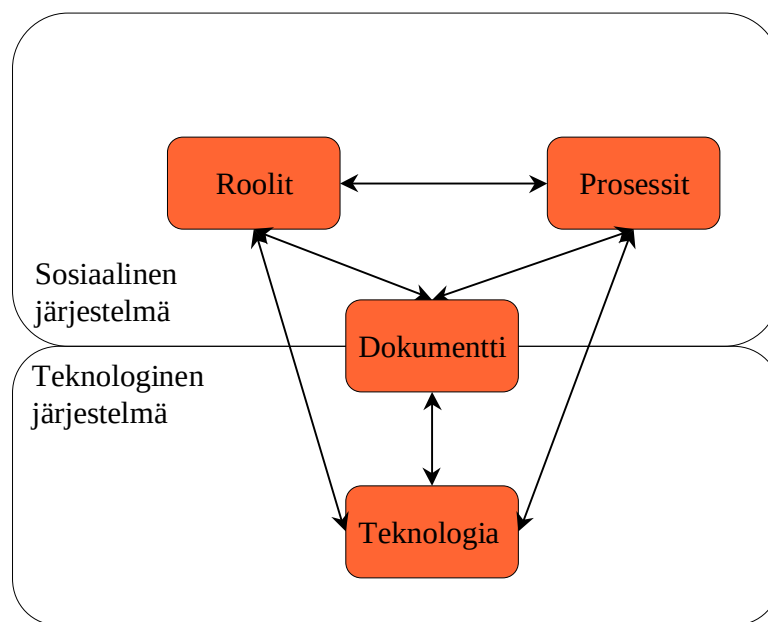
4 Timanttimalli

Informaation hallinnan kenttää voidaan lähestyä useasta eri näkökulmasta. Kun haluamme kehittää organisaation dokumentti-muotoisen informaation hallintaa, voimme käyttää apuna timanttimalliä (kuva 6 [14] [6] [15]). Organisaation informaation hallinta on jaettavissa karkeasti kahteen osajärjestelmään:

- **sosiaaliseen osajärjestelmään**, joka koostuu organisaation työprosesseista, niiden käsittelemistä tiedoista sekä rooleista, joiden kautta ihmiset osallistuvat prosessien toimintaan
- **teknologiseen osajärjestelmään**, joka koostuu teknisistä laitteista, ohjelmista, verkoista ja niiden avulla käsiteltävistä tiedoista.

Dokumentit sijaitsevat näiden kahden järjestelmän rajapinnassa. Dokumenteissa näkyvät sekä sosiaalisen järjestelmän muutokset - mitä tietoja organisaatio käsittelee - että teknologisen järjestelmän muutokset - millä tekniikoilla tietoja käsitellään. Yleensä näitä kahta järjestelmää käsitellään organisaatiossa erillisinä. Eri ihmiset miettivät liiketoimintaprosessien muutoksia kuin IT ja IS asioita. Näiden kahden tekijän yhteensovittaminen onkin keskeinen tekijä DH:n tehokkaan toiminnan kannalta.

Timanttimalliä voidaan käyttää myös dokumenttien hallinnan kehittämisen näkökulmina. Kehittäminen voi lähteä liikkeelle rooleista, prosesseista, tiedoista tai teknologiasta käsin.



Kuva 6. Timanttimalli [6] [15].

Kun informaation hallintaa aletaan kehittää **dokumenteista** käsin, analysoidaan ensin organisaation käsittelemät dokumentit, niiden tietosisältö, rakenteet, säilytysvaatimukset jne. Analyysi jatkuu kohti dokumentteja käyttäviä prosesseja, tekniikoita ja tietojen käyttäjiä.

Prosessilähtöinen kehittäminen määrittelee ensin organisaation liiketoimintaprosessit, sekä pää- että tukiprosessit. Prosessien tietotarpeet ja tietojen vaihto analysoidaan ja muunnetaan vaatimuksiksi dokumenteille ja muulle informaation käsittelylle.

Roolilähtöisessä kehittämisessä analysoidaan eri roolit, joita organisaatiossa on - sekä prosessien että toimintojen kuin myös yksilöiden näkökulmasta. Rooleissa kiinnitetään erityisesti huomiota niiden tietotarpeisiin ja siihen, miten eri dokumentit ja muu informaatio vastaavat näitä tarpeita. Kannattaa vielä muistaa, että sosiaalisen järjestelmän keskeinen yhteinen tekijä on **ihminen**, joka roolien kautta toimiessaan saa prosessit toimimaan.

Teknologiaalähtöisessä kehittämisessä analysoidaan tarpeita organisaation DH:n ja muun informaatiotekniikan standardeille, järjestelmille, verkoille ja muulle infrastruktuurille. Painopiste on sopivien teknologioiden löytämisessä, joita pyritään soveltamaan liiketoiminnan eri tarpeisiin.

Rajautuminen yhden näkökulman käyttöön ei tuota hyvää tulosta organisaation kokonaisuuden kannalta - kaikki osa-alueet täytyy ottaa huomioon. Kannattaa myös huomata, että liiketoiminnan eri osa-alueiden tietämys on jakautunut eri puolille organisaatiota. Riittävän

tiedon kokoaminen edellyttää kohtuullisen suuren osallistujajoukon saamista mukaan DH ja muiden informaation hallinnan kehittämishankkeiden valmisteluun.

5 Genret - organisaation kommunikaation lajityypit

Genre [16] on organisaation sisällä yleisesti tunnettu ja ymmärretty kommunikaatiomuoto, jolla on tietty tietosisältö ja tietty muoto. Genreen liittyy aina tietty käyttötarkoitus, jonka sen käyttäjät tuntevat. Esimerkkeinä genreistä voidaan mainita varaosatilaus, osaluettelo, projektipalaverimuistio, henkilötiedot, ruokalista, asiakaspalaute. Käytämme tässä termejä genre ja kommunikaatiogenre, jotka kattavat sekä dokumenttigenret että muut muilla medioilla välitetyt kommunikaatiomuodot. Kuvassa 7 on esimerkkejä organisaatiokohtaisista genreistä, joilla on tarkasti rajattu merkitys organisaatiossa. Niistä viimeinen – Tilaus - voisi olla myös geneerinen genre, joka tunnetaan yleisesti eri organisaatioissa, kun taas kahdella ensimmäisellä on selvästi organisaatiokohtainen sisältö

Tuottaja	Genre	Käyttäjä
Tuotehallinta	Lista valituista laitteista	Varaosa (-osasto/ryhmä)
Myynti	Tukipyyntö	Projektiorganisaatio
Asiakas	Tilaus	Myynti

Kuva 7. Esimerkkejä organisaatiokohtaisista kommunikaatiogenreistä.

Dokumenttigenre poikkeaa hieman käsitteestä **dokumenttityyppi**, sillä samaa teknistä dokumenttityyppiä (kuten ”sisäinen muistio”) voidaan käyttää useamman genre tietosisällön välittämiseen ja toisaalta samaa tietosisältöä (genreä) voidaan myös välittää useamman dokumenttityypin, puhelimen tai tietokantasovelluksen kautta.

Genresysteemi [17] on joukko genrejä, jotka ovat osa samaa toimintaprosessia. Esimerkiksi työpaikkailmoitus, työhakemus ja haastatteluraportti voivat muodostaa genresysteemin.

Genrerepertuaarilla tarkoitetaan tietyn organisaation käyttämiä genrejä. Kun useampi organisaatio toimii yhteistyössä, niillä on yleensä omia, sisäisiä genrerepertuaareja sekä yhteistyössä käytettävistä, jaetuista genreistä muodostuva yhteinen genrerepertuaari.

Tehdyissä organisaatioanalyysissä yhdestä organisaatioyksiköstä on yleensä löytynyt kohtuullisella työmäärällä 400-1000 genreä.

Vaikuttaisi sille, että tämän verran niitä on löydettävissä myös muista vastaavista organisaatioista. Genrejä yhdistelemällä ja yleistämällä niiden määrä olisi luultavasti pudotettavissa noin 150-400 kappaleeseen, jotka toteutetaan ehkä noin sadan teknisen dokumenttityypin/-rakenteen avulla. Organisaatioissa on esimerkiksi useita "tilaus" ja "lasku" genrejä, joista ainakin osa olisi yhdistettävissä pariin yleiseen genreen, jolla on yhtenäinen tietosisältö ja esitystapa sekä media. Yleensä yksi ihminen ei tunne tarkasti kuin murto-osan kaikista organisaation genreistä. Tämän takia dokumentteihin kohdistuvat muutokset vaativat yleensä varsin kattavan osaan oton tiedon keruuseen organisaation eri henkilöstöryhmistä.

Organisaatioiden informaation hallinnasta löytyy aina suuri määrä vaihtoehtoisia kehityskohteita. Parhaan kehityskohteen valinta tai saavutetun hyödyn mittaaminen jälkikäteen saati arviointi etukäteen ei ole helppoa. IT-järjestelmiin investoidaan ilman varmuutta siitä, tuottavatko investoinnit (ns. IT investment paradox).

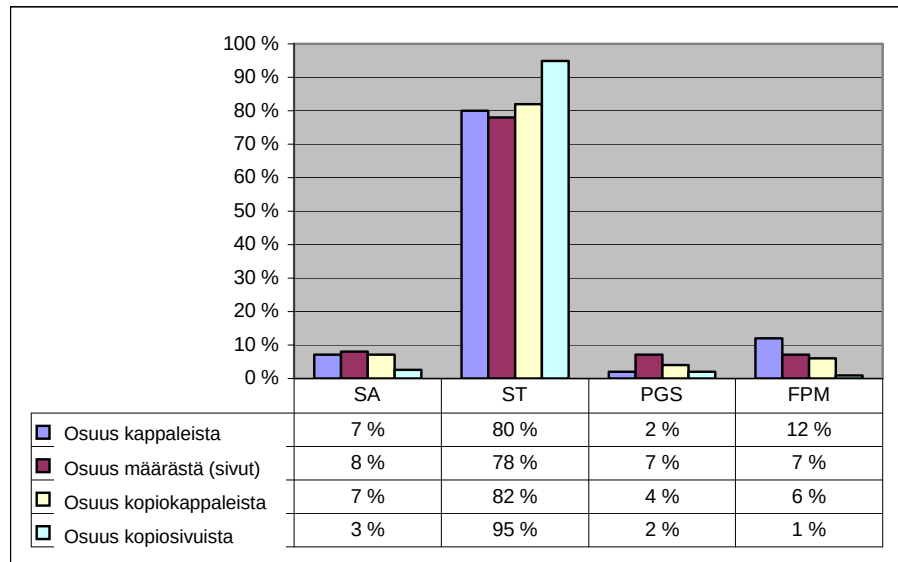
Yksi keskeinen pohjatieto näiden ongelmien ratkaisemiseksi on tieto informaation määrästä. Eri formaateissa käsiteltävien dokumenttigenrejen lukumäärä on yksi mittari. Sen lisäksi voidaan mitata [18] [19] [20]:

- dokumenttien, sähköpostien, tai muiden kommunikaatiotapahtumien lukumäärää,
- edellistä pois lukien kopiot (esimerkiksi sama sähköposti usealle vastaanottajalle, eli ”uniikit instanssit”, UI, kopioiden sijasta),
- kommunikoidun tiedon määrä, tai
- edellistä ilman kopioita (”uniikin” eli ainutkertaisen tiedon määrää)

Tiedon määrän mittaamista vaikeuttaa se, että tallentamaton tieto, kuten puhe, sisältää myös sanattomia viestejä, joiden määrää on hankala arvioida. Samoin esimerkiksi yksi tekstisivu voi bitteinä mitattuna kasvaa moni sataa kertaiseksi vain vaihtamalla tallennusmuoto tekstitiedostosta bittikartaksi. Tältä pohjalta käytännöllisimmäksi mittayksiköksi nousee ”sivu” eli se määrä tietoa, joka voidaan esittää ihmiselle yhtenä visuaalisena A4-paperisivuna [18].

Kuvassa 8 esitetään erään organisaation käsittelemien tietojen jakautuminen neljän pääprosessialueen (SA, ST, PGS ja FPM) mukaisesti [20]. Jos tarkastellaan kommunikaatiotapahtumien määrää, yksi prosessialue (ST) muodosti 80 % organisaation kommunikaatiosta. Jos tarkastellaan tiedon määrää (ml. kaikki kopiot), kyseinen prosessialue käsitti noin 95 % kyseisen organisaation kommunikoinnista. Tämä ilmiö onkin varsin yleinen – organisaation ydinprosessissa liikkuu yleensä vähintään puolet käsiteltävistä tietomäärästä. Tukiprosesseissa voidaan kyllä käsitellä hyvinkin

monenlaisia tietoja – eli suurta lukumäärää genrejä – mutta niissä liikkuvan tiedon määrä jää vähäisemmäksi.



Kuva 8. Erään organisaation tietojen jakautuminen neljän osa-alueen välillä [20].

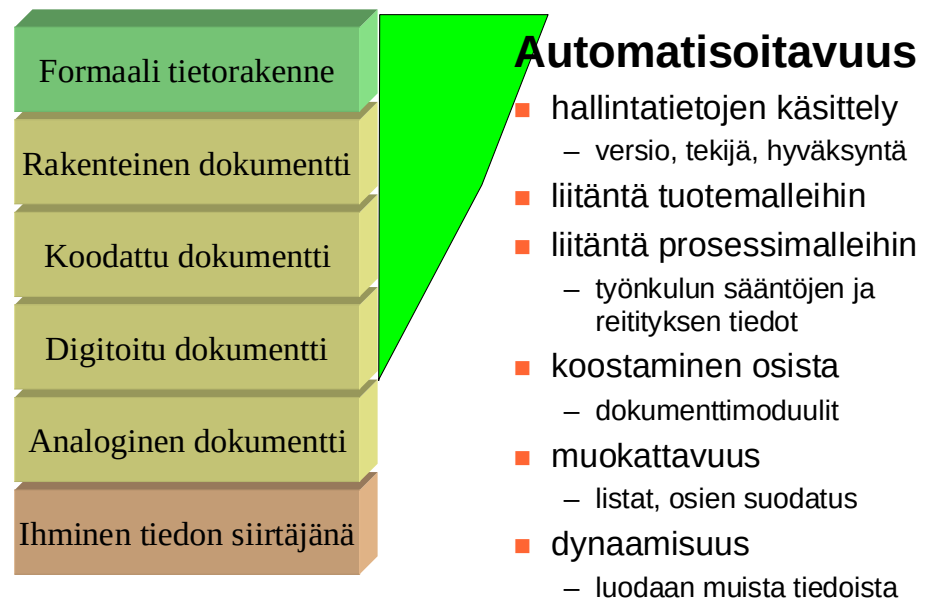
Tältä pohjalta panostaminen keskeisen ydinprosessin tehokkuuteen on useissa tapauksissa perusteltua, vaikka se edellyttäisikin juuri tätä tarkoitusta tai toimialaa varten räätälöityjen järjestelmien käyttöä. Toisaalta esimerkiksi tilaus – laskutus – prosessin käsittely muodostaa useissa organisaatioissa suhteellisen suuren tietomäärän ja sen automatisointiin kannattaa panostaa, sillä tähän alueeseen sopivia yleiskäyttöisiä ohjelmistoja on saatavana varsin taloudellisesti.

6 Ydinprosessien tehostaminen

Aiemmin mainitut dokumenttien erityispiirteet tulevat näkyviin myös teknologiavaihtoehtoissa, joilla dokumentteja käsitteleviä ydinprosesseja tehostetaan. Tietokoneiden kannalta käsiteltävä tieto pitäisi voida määritellä formaalisti, jotta kaikkia tietokoneen tarjoamia mahdollisuuksia voitaisiin hyödyntää.

Tässä käytämme termiä **dokumenttitiedon formaalisuus** [6] viittaamaan tiedon tallennusmuodon ulottuvuuteen, joka vaikuttaa tietokoneiden kykyyn hallita ja prosessoida sisältökomponentteja. Kuvassa 9 esitetään jatkumo, jossa dokumenttitiedon formaalisuus kasvaa alhaalta ylöspäin. Kun tietoa ei ole tallennettu, sen hallittavuus ja käsiteltävyys on luonnollisesti hyvin heikko - voidaan enintään tallentaa viitetietoa tiedosta, jota tiedetään organisaatiossa käsiteltävän (esimerkiksi laatujärjestelmä voi olla kiinnostunut siitä, miten strategia on kommunikoitu työntekijöille, vaikka tämä tieto välitetään suullisesti

osastokokouksessa). Edettäessä ylemmille tasoille tietokoneen ominaisuudet ja tehtävien automatisoitavuus kasvavat, kunnes formaalisti määritellyssä tiedossa tietokone kykenee manipuloimaan metatiedon lisäksi tehokkaasti itse tietoa.



Kuva 9. Dokumenttitiedon formaalisuus [6].

Sisällönhallinnan tarpeita vastatakseen tiedon täytyy olla tietokantatiedon, rakenteisten dokumenttien tai vähintäänkin puolirakenteisen tiedon tasolla, jotta prosessoinnin kannalta riittävä metatieto olisi olemassa.

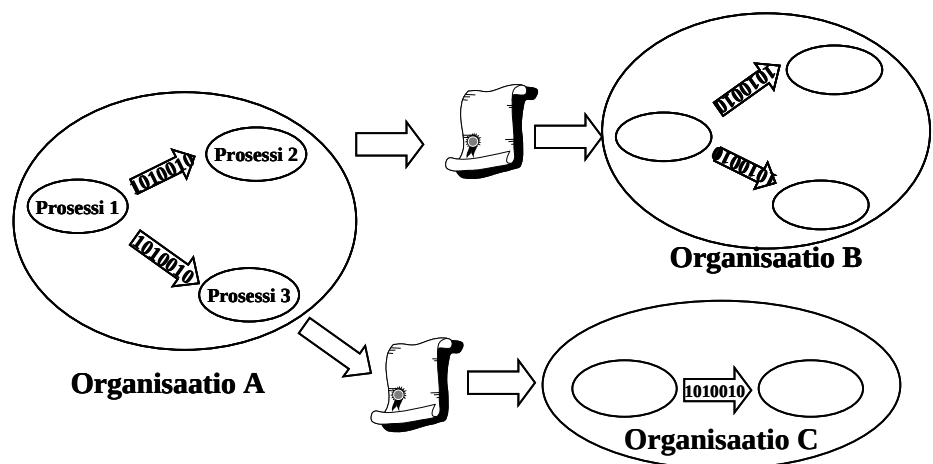
Kannatta kuitenkin huomata, ettei tiedon formaalisuus ole itseisarvo sinällään. Formaali esitysmuodot sopivat suhteellisen vakiintuneille kommunikaatiomuodoille (genreille) ja selkeisiin tilanteisiin. Vähemmän formaalit ja kommunikaatiokanavaltaan rikkaammat viestintämuodot (median richness) sopivat taas vuorovaikutteisempina monitulkintaisempiin kommunikaatiotilanteisiin, kuten uuteen yrityskumppaniin tutustumiseen tai palautteen antoon alaiselle.

Dokumenttitiedon formaalisuuteen liittyy kiinteästi **granulariteetti** eli tiedon käsittelyn rakeisuus. Mitä suurempi käsittelyn rakeisuus (granulariteetti) on sitä pienempinä kokonaisuuksina tietoa käsitellään, eli sitä pienempi tiedon raekoko. Jos dokumenttitietoa käsitellään suurina kokonaisuuksina, kuten suurina dokumentteina tai kokonaisina dokumentaatioina, hallittavien kohteiden lukumäärä on pieni ja hallintatyö siinä mielessä helpompaa. Tällöin joudutaan kuitenkin usein ylläpitämään rinnakkaisia dokumentteja, joilla on yhteisiä osia ja ongelmaksi tulee kaikkien tarvittavien dokumenttien päivittäminen muutostilanteissa.

Jos taas dokumentit jaetaan pienempiin osiin, voidaan paljon välttää saman tiedon esittäminen kahdessa dokumentissa, eikä rinnakkaisten versioiden ylläpito tule ongelmaksi. Tällöin kuitenkin joudutaan ylläpitämään tietoa siitä, mitkä osat muodostavat minkäkin kokonaisuuden, eli hallittavien kohteiden määrä kasvaa. Lisäksi yhtä dokumenttia muutettaessa ei välttämättä tiedetä, missä kaikissa yhteyksissä (dokumenttikoosteissa) kyseinen tietopalanen tulee esiintymään eli lukija ei välttämättä kykene tulkitsemaan tietoa oikein ilman riittävää asiayhteyttä.

Äärimmilleen vietyinä dokumentin raekoko voidaan pienentää tasolle, jolla jokainen tekstinpala tallennetaan tietokannan kenttänä tai lyhyenä XML-merkattuna elementtinä, jota käytetään dynaamisesti luoduissa dokumenttikokonaisuuksissa tarpeen mukaan. Tämä edellyttää kuitenkin dokumenttien koostamisen tarkkaa formalisointia ja on sidoksissa dokumenttitiedon formaalisuuteen. Tässä lähestymistavassa dokumenttikäsite alkaa hämärtyä ja yleensä puhutaankin mieluummin **sisällönhallinnasta** (content management). Dokumenttien hallinnan ja sisällönhallinnan käsitteet ovat kuitenkin paljon päällekkäisiä.

Dokumenttitiedon formaalisuuden nostaminen helpottaa siis sisäisten prosessien automatisointia. Kuva 10 esittää yleistä tilannetta, jossa organisaatioiden välinen tiedonsiirto hoidetaan tulostamalla digitaalinen tieto paperille ja kirjoittamalla se taas sisään seuraavassa organisaatiossa.



Kuva 10. Informaatiosaaristo, jossa organisaatioiden sisäinen tietojen välitys on digitaalista mutta organisaatioiden välinen tietojen siirto tapahtuu paperimuodossa. [21]

Käytännössä organisaation sisäisen integroinnin jälkeen vuorossa ovat yleensä prosessien uudelleenorganisointi ja organisaatioiden välisten toimintaketjujen automatisointi käyttäen elektronisen liiketoiminnan sovelluksia.

Näihin tarkoituksiin käytetään yleensä XML-pohjaisia rakenteisia dokumentteja (XML-pohjaisia EDI-”standardeja”), joiden tietomäärittäjiä on vakioitu useilla foorumeilla:

- Common Business Library (CBL, www.commerceone.com)
- BizTalk (Framework, skeemat, server, Microsoft server www.biztalk.org, www.microsoft.com/biztalk)
- RosettaNet (American Express, Microsoft, Netscape, IBM, www.rosettanet.org)
- Commerce XML resources (www.cxml.org www.ariba.com)
- ebXML (www.ebxml.org, www.ebxml.org, OASIS)

Esimerkiksi BizTalk-määrittäksen mukaan kirjatilaus (kuva 11) voidaan lähettää oheisen kaltaisena elektronisena kirjeenä, jossa kirjekuoren (<Envelope>.... </Envelope>) sisällä ovat lähettäjän ja vastaanottajan tiedot sekä kirjeessä itsessään kirjatilaus (<PurchaseOrder>). Kirjatilauksen tiedoissa on myös mainittu osoite Internetissä, josta löytyy kirjatilauksen käyttämän tietomallin kuvaus. Näin eri organisaatioiden välisen tiedon siirron yhdenmukaisuus pyritään turvaamaan. Tilauksen käsittelijöinä ovat siis tietokonesovellukset ja se välitetään esimerkiksi Internetin välityksellä. Kirjatilaus BizTalk-muodossa voisi siis näyttää (yksinkertaistettusti) seuraavalle:

```
<Envelope>
  <Header>
    <to> <address type=department>
      Book Orders</address>
    </to>
    <from> <address type=organization">
      Book Lovers</address>
    </from>
  </Header>
  <Body>
    <PurchaseOrder
      po="http://electrocommerce.org/purchase_order/">
      <po:Title>Essential BizTalk</po:Title>
    </po:PurchaseOrder>
  </Body>
</Envelope>
```

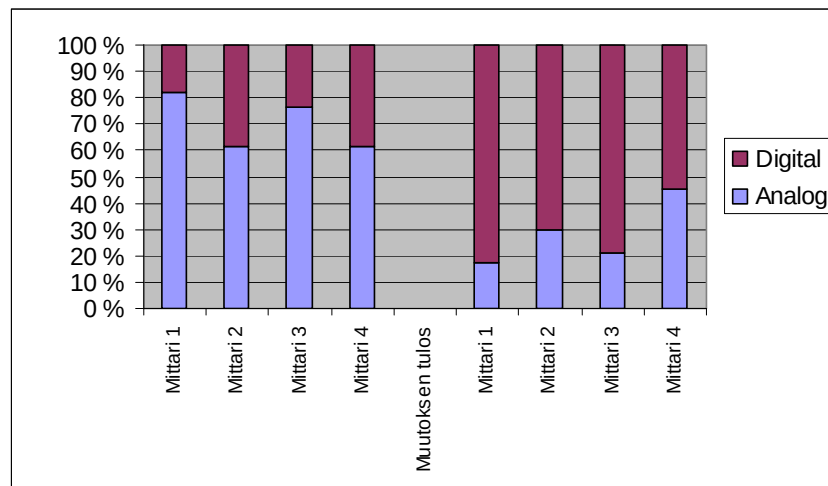
Kuva 11. Kirjatilaus digitaalisessa kirjekuorisanomassa.

Esimerkkinä elektronisen liiketoiminnan eduista on verkkolaitevalmistaja Cisco, joka kertoi saavuttaneensa webin käytön avulla 6,5 Mrd Mk säästöt, eli 6% liikevaihdostaan. 80% säästöistä eli 4400 MMk syntyi asiakaspalvelun siirtämisestä valtaosin itsepalveluperiaatteelle. Nyt 400 insinööriä vastaa puhelinkyselyihin, ilman itsepalvelujärjestelmää olisi tarvittu 1000 henkilöä enemmän.

Cisco pystyi siis tarjoamaan enemmän palvelua ja nostamaan myyntiään radikaalisti olemassaolevalla henkilömäärällä. Muita Cison saavuttamia säästöjä tuottivat muunmoassa elektroninen kaupankäynti 413 MMk ja toimitusketjun hallinta 2000 MMk. Kaiken kaikkiaan tuloksena oli 2,5 Mrd markan lisämyynti, toimitusajan lyheneminen 8 viikosta kahteen ja varastojen pieneneminen 45%. Jatkossa suurin potentiaali on verkko-opetuksen soveltaminen 80% koulutuksesta, jolla tullaan arvion mukaan saavuttamaan 40-60% säästöt. [22]

7 Suuruusluokkia organisaatioiden informaatiomääristä

Miten helppoa organisaation tietoja on digitoida? Eräässä organisaatiossa toteutetussa, keskeisten tietojen digitalisoinnissa saavutettiin kuvan 12 mukaiset tulokset [18].

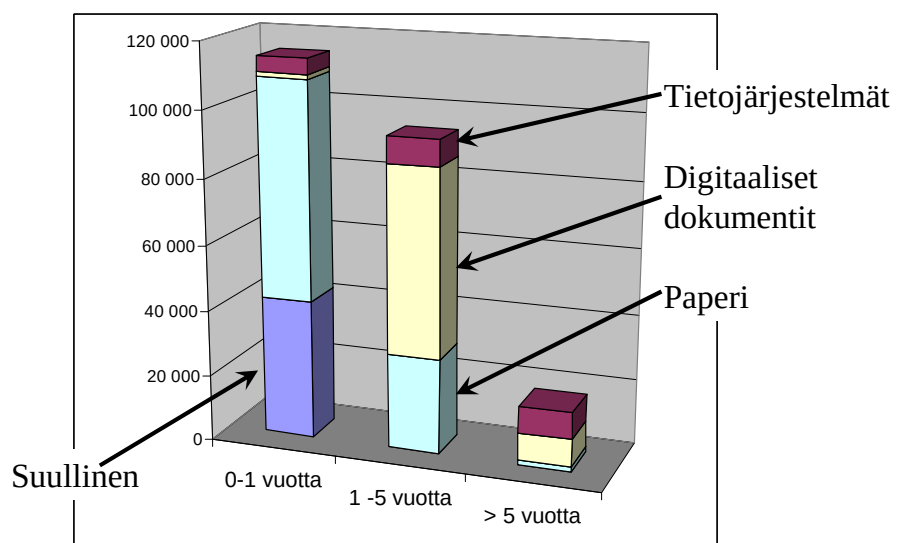


Kuva 12. Yhden keskeisen genren digitalisoin vaikutus otannan digitaalisuuteen [18].

Kuva 12 esittää otantaa organisaation tiedoista vuonna 1999 (10% genreistä analysoitu). Vasempien neljän palkin mukaisesti tämän otannan tiedoista noin 60-80% prosenttia organisaation tallennetuista tiedoista oli analogisessa muodossa – riippuen käytettävästä määrällisestä mittarista (mitataanko tapahtumia, kopioita vai määrää, kts. luvun alku edellisellä sivulla). Kun yksi tämän otannan keskeinen genre digitalisoitiin, muuttui tilanne kaikilla mittareilla kohtuullisen dramaattisesti. Nyt 50-80% otannan tiedoista oli digitaalisessa muodossa. Yksittäisilläkin, hyvin kohdennetuilla muutoksilla voi olla selvä vaikutus organisaation tietojenkäsittelyn kokonaisuuteen ja toiminnan tehokkuuteen.

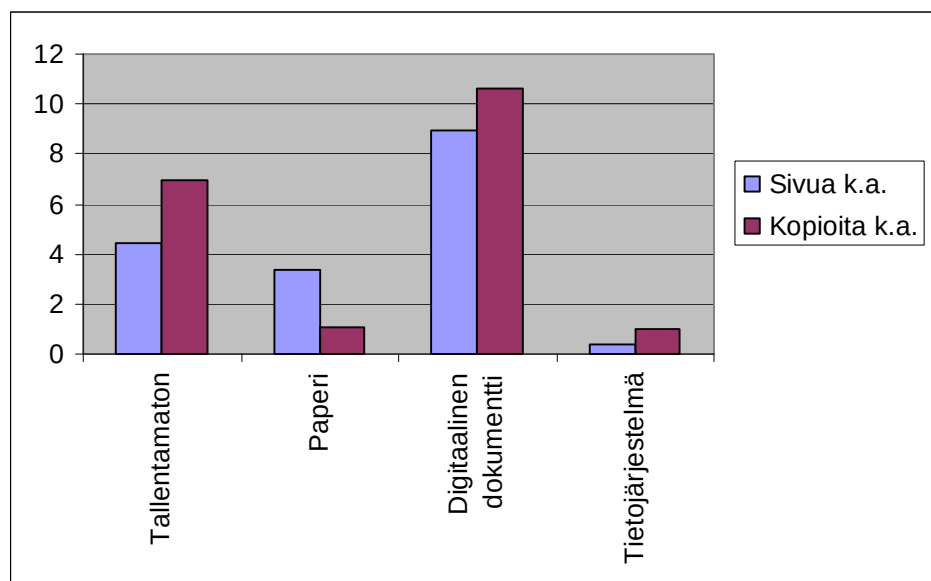
Kuva 13 esittää erään organisaation käsittelemien tietojen jakautumista tallennusajan (x-akseli) ja tiedonsiirtomuotojen mukaan (suullinen tieto, paperi, digitaaliset dokumentit ja tietojärjestelmät). Tiedon määrä

on mitattu sivuina, ilman kopioita. Noin puolet tiedosta välitetään medially, jota säilytetään yli vuoden, mutta vain 7-8% säilytetään yli viisi vuotta. Toisin kuin yleensä oletetaan, tietoa ei ainakaan tässä organisaatiossa arkistoida paperille, vaan valtaosa pitkäaikaisvarastoinnista hoidetaan tietojärjestelmissä tai digitaalisina dokumentteina. 1-5 vuotta säilytettävistä tiedoista suurin osa oli digitaalisina dokumentteina. Kyseisen organisaation prosesseihin liittyvästä lyhytikäisestä tiedosta kohtuullisen suuri osa välitetään suullisesti, vaikkakin suurin osuus onkin paperilla. Näyttää sille, että tässä tapauksessa valtaosa paperitiedosta on tilapäistä käyttöä varten tulostettua digitaalista tietoa.



Kuva 13. Erään organisaation vuotuinen tietojenkäsittely jaoteltuna säilytysajan ja tiedonsiirtotavan mukaan, pois lukien kopiot. Värät: sininen = suullinen, vihreä = paperi, keltainen = digitaaliset dokumentit, punainen = tietojärjestelmät.

Kuvassa 14 on esitetty kyseisen saman organisaation dokumenttien ja muiden tietokokonaisuuksien keskimääräinen koko sivuina sekä yhdestä dokumentista jaeltujen dokumenttien kopioiden keskiarvo. Kuten kuvasta näkyy, tietojärjestelmissä käsiteltävät tietokokonaisuudet ovat hyvin pieniä – tyypillisesti niissä on tietoa vain muutamia rivejä, eli reilusti alle yhden sivullisen. Myös valtaosa paperidokumenteista on hyvin pieniä. Sen sijaan digitaalisen dokumenttien keskikoko on noin kymmenen sivua. Vastaavasti digitalisista dokumenteista jaetaan keskimäärin kymmenkunta kopiota, kun paperidokumenteista jaetaan vain muutamia kopioita.



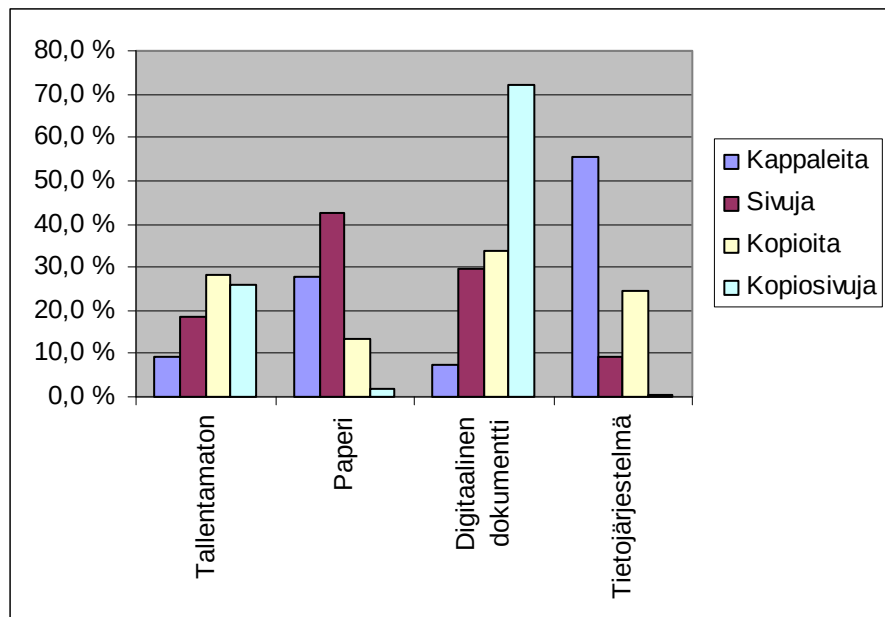
Kuva 14. Erään organisaation tietokokonaisuuksien keskikoko ja keskimääräinen kopiomäärä kommunikaatiomuotojen mukaan ryhmiteltynä.

Kyseisen organisaation tietojen kokonaisvolymi näyttää kuvan 15 mukaiselta. Suurin lukumäärä erilaisista tietokokonaisuuksista (siniset pylväät) tallennetaan tietokantoihin (yli 50%), erilaisia paperidokumentteja on noin kolminkertainen määrä digitaalisiin dokumentteihin verrattuna (vajaa 30% vs. alle 10%).

Tietokantatieto koostuu pienistä tietokokonaisuuksista, joten tietojärjestelmissä on sivuina mitattuna vain vajaa 10% erilaisesta tiedosta (punaiset pylväät). Paperidokumentit nouseekin tässä organisaatiossa suurimmaksi erilaisen tiedon sivumäärällä mitattuna (yli 40%) digitaalisten dokumenttien osuuden noustessa 30% tasolle.

Jos taas tarkastellaan tietokokonaisuuksien lukumäärää mukaanlukien myös kopiot (keltaiset pylväät) nousee digitaalisten dokumenttien ryhmä suurimmaksi noin kolmanneksen osuudellaan. Edelleen, jos tarkastellaan kopioiden sisältämää tietomäärää mitattuna sivuina (vaalean turkoosit pylväät) nousee digitaalisten dokumenttien osuus yli 70% tämän organisaation käsittelemästä tiedosta. Edellisen kuvan mukainen suuri kopiomäärä yhdistettynä tietokokonaisuuksien suurempaan keskikokoon tuottaa valtaosan tämän organisaation tietomäärästä – ja siis myös tietotulvasta.

Miksi näin? Onko digitaalisia dokumentteja hyödynnetty käyttötarkoituksiin, joissa vaaditaan suuri määrä kopioita suurista tietokokonaisuuksista? Vai onko dokumenttien digitalisointi aiheuttanut turhaa kointia? Jos on, niin onko se haitallista?



Kuva 15. Esimerkkiorganisaation käsittelemien tietojen määrät jaoteltuina neljällä eri kriteerillä:

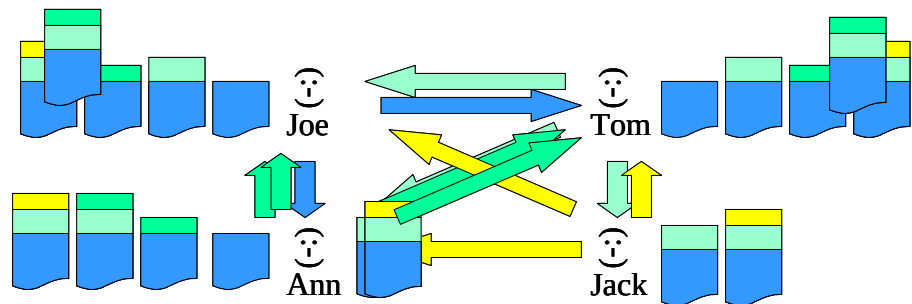
- Kappaleita = erilaisten dokumenttien tai muiden kommunikaatiotapahtumien määrä
- Sivuja = kommunikaation määrä mitattuna sivuina
- Kopioita = kappalemäärä mukaanlukien kopiot
- Kopiosivuja = kommunikaatiomäärä mukaanlukien kopioiden sivumäärä.

Digitaalisia dokumentteja hyödynnetään kyllä sellaisten tietojen tallentamiseen, joita on tarve jaella suurelle määrälle ihmisiä. Mutta näyttää myös sille, että helppo tietojen jakelu sähköpostin kautta aiheuttaa myös paljon turhaa, huonosti harkittua ja myös huonosti valmisteltua kommunikointia. Kopioita lähetetään varmuuden vuoksi useammille ihmisille – ”tiedoksi”. Tämä lisääntynyt tiedon lähettäminen myös lisää tiedon vastaanottamista, tiedon prosessointitarvetta vastaanottajalta ja siten tietotulvaa.

Edelleen, vastaanotettua digitaalista dokumenttia saatetaan muokata ja lähettää se takaisin korjattuna ja/tai päivitettyinä. Jos alkuperäistieto on lähetetty useammalle ihmiselle, joista pari-kolme muokkaa omaa versiotaan ja jakelee edelleen korjauksensa useammalle henkilölle, päädytään helposti tilanteeseen, jossa muutaman hengen ryhmällä on kenties kymmenestä erilaisesta versiosta kymmeniä kopioita – eikä kukaan ole varma, mikä tieto on ajan tasalla. Kuva 16 havainnollistaa tällaista tilannetta.

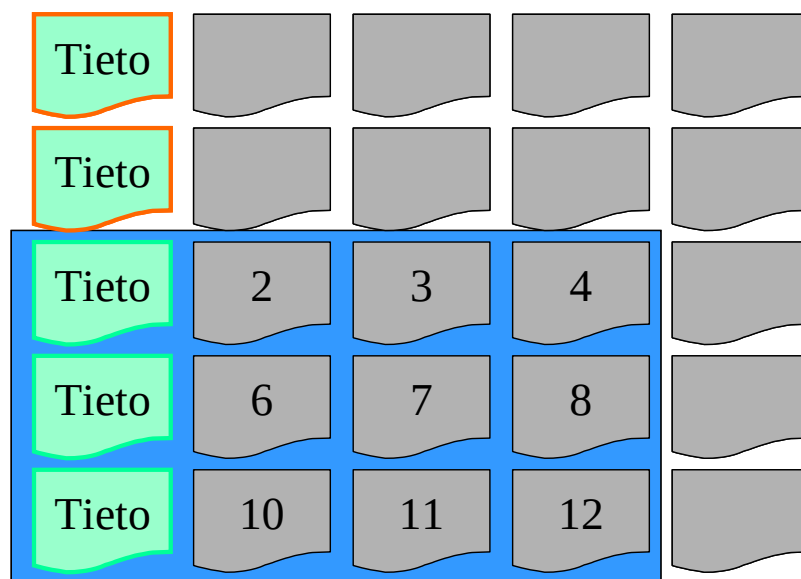
Mooren lain perusteella voisi sanoa, ettei tämä ole ongelma: ostetaan lisää levytilaa, tallennetaan varmuuden vuoksi kaikki versiot ja käytetään hakukoneita tiedon hakemiseen. Tässä lähestymistavassa

tulee kuitenkin ongelmia rinnakkaisten versioiden hallinnan lisäksi muunmoassa tiedon etsinnän suhteen.



Kuva 16. Esimerkki tilanteesta, jossa useampi henkilö tekee lisäyksiä (eri väriset) yhden henkilön (Joe) tekemään dokumenttiin (tummansininen). Rinnakkaisten versioiden kaaos kasvaa nopeasti [23].

Jos esimerkiksi joku koettaa koostaa puhelinluetteloa kandestakymmenestä viidestä eri luetteloversiosta, joissa on viisi eri lisäystä, tilanne näyttää suunnilleen kuvan 17 mukaiselle.



Kuva 17. Esimerkki tilanteesta, jossa kahdestakymmenestä viidestä dokumentista halutaan löytää viisi ”Tieto”-tekstillä merkittyä, mutta kysely palauttaa sinisellä pohjalla olevat 12 dokumenttia mukaanlukien kolme haluttua. Haun saanto on 3/6 eli 60 % ja tarkkuus 3/12 eli 25% [23].

Käyttämällä hakukonetta saadaan kolme halutuista viidestä dokumentista (saanto 60 %) ja yhdeksä turhaa dokumenttia (tarkkuus 25%). Tiedonhaun tutkimuksessa näiden kahden suureen välistä suhdetta pyritään optimoimaan tilanteen tarpeita vastaavaksi muokkaamalla kyselyjä optimaalisiksi. Jos esimerkiksi haluttaisiinkin vain yksi sellainen puhelinluettelo-versio, jossa on haluttu lisäys, tarkempi haku voisi palauttaa kaksi dokumenttia, joista toisesta löytyisi haluttu numero. Jos taas halutaan kaikki viisi lisäystä, pitää luultavasti

tehdä kysely, joka palauttaa (lähes) kaikki dokumentit ja lukija joutuu selaamaan lähes kaikki dokumentit läpi. Saannon ja tarkkuuden samanaikainen optimointi on erittäin vaikeaa ilman lisätietoja. Jokainen turha dokumentti – siis kopiot – tuottaa tietotulvaa ja maksaa työaikana merkittäviä summia.

Meidän on syytä tarkastella joitakin suuruusluokkia, jotta organisaatioiden tiedon hallinnan problematiikka saisi oikeat konkreettiset mittasuteet. Kuvan 18 taulukossa esitetään joitakin suuruusluokkia organisaatioiden informaation käsittelystä. Vasemmalla esitetään mittayksiköt, keskellä suuruusluokat yhden henkilön työpäivää kohden ja oikealla samat skaalattuna 100 henkilön vuodessa tekemään työmäärään nähden. Yksi henkilö voi siis lukea noin 300 sivua tekstiä päivässä. Landaurin tutkimusten mukaan muistiin jäävä osuus on tästä prosenttien – parin suuruusluokkaa. Näihin reunaehtoihin suhteutettuina havaintomme noin sadan sivullisen informaatiomäärän vastaanottamisesta tai lähettämisestä työpäivää kohden tuntuvat realistisille: Jos vastaanottaa kolmisivuisen dokumentin ja lähettää sen edelleen neljälle kolleegalleen, käsittelee jo $(1 + 4) * 3$ eli 15 sivullista informaatiota. Sata sivullista henkilötyöpäivää kohden skaalautuu kahdeksi miljoonaksi sivuksi vuodessa sadan hengen organisaatiossa.

Mittayksikkö	1 henkilö päivässä	100 henkilöä vuodessa
Lukunopeus	~300 sivua	
Muistiin jää	~10 sivua (1-2 %)	(lähde: Landauer)
Tiedon käsittely	~100 sivua	2'000'000 sivua
Tietoa tuotetaan ja kopioidaan	1-10 sivua * kopiot (3-4-..)	20-200'000 sivua * kopiot (3-4-..)
Josta säilytetään pitkäaikaisesti	0,2 – 0,5 dokumenttia	5'000- (20'000-) dokum. (sivua)

Kuva 18. Suuruusluokkia organisaatioiden informaation käsittelystä.

Kannattaa huomata, että päivittäin käsitellystä sadan sivun tietomäärästä itse tuotettua uutta tietoa on vain suhteellisen pieni osa: yleensä alle kymmenen sivullista – lisäksi siitä tuotetut kopiot. Edelleen, pitkäaikaissäilytettävää tietoa on suhteellisen vähän, joitakin prosenteja. Tosin organisaation toiminnan luonne vaikuttaa oleellisesti erityisesti tähän osuuteen.

Mikä edellä kuvatun tiedonhaun problematiikan merkitys on organisaatioiden informaation käsittelylle? Esimerkiksi: Jos organisaatiossa on 100'000 dokumenttia, haku voi palauttaa 1'000 dokumenttia, joiden läpikäyminen veisi 7 työpäivää ja maksaisi palkkakustannuksina sivukuluineen kenties 2'000 euroa. Jos haku voidaan rajata yhteen genreen / dokumenttilajiin (esimerkiksi: vastaus korvaushakemukseen), haun kohteena oleva dokumenttijoukko voi pudota 1'000 dokumenttiin ja haku palauttaakin kenties 50 dokumenttia, joiden selaaminen vie kolme tuntia ja maksaa 100 euroa. Edelleen, jos dokumenteissa on organisaation toiminnan kannalta mielekästä metatietoa (mihin asiakkaasee, projektiin, tuotteeseen, asiaan... dokumentti liittyy) haun kohteena oleva dokumenttijoukko voi rajautua 20 dokumenttiin, joista haku palauttaa 10, joiden selaamiseen menee 35 minuuttia, mikä maksaa organisaatiolle palkka yms. kuluina 20 euroa.

8 Ihminen, tieto, tiedon puute ja tuska

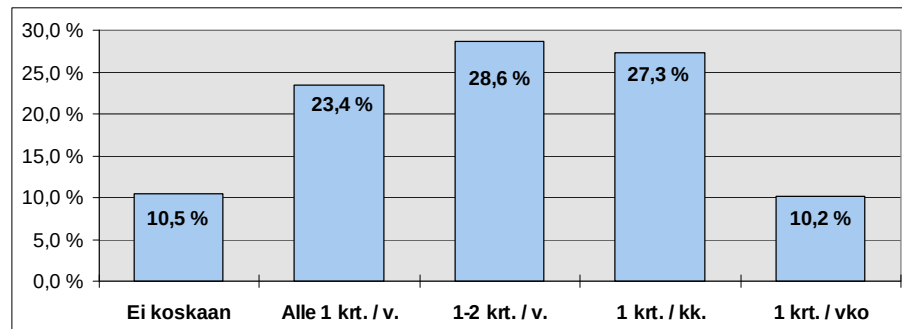
Usein käytetyn tiedon organisoiminen tiedon hakemista varten voi siis periaatteessa pudottaa tiedonhankintakustannuksia sadasosaan. Organisaation käyttämä tieto, genrerepertuaari, kartoitetaan ja keskeiset tietokokonaisuudet organisoidaan hakua varten esimerkiksi dokumenttienhallintajärjestelmään, jossa tiedosta on vain yksi kopio, haettavissa metatietojen avulla ja/tai tekstihaulla. Edelleen, keskeinen tieto voidaan integroida tiiviisti keskeisiin toimintaprosesseihin sovelluksilla, joiden avulla niiden käsittely on tehokasta.

Vähemmän käytetty tieto voi jäädä muiden tiedonhankintamekanismien varaan – hajanaisempiin paperiläjiin, ihmisten korvien väliin ja niin edelleen. Samoin organisaation oppimisen kannalta keskeistä uuden tiedon muodostamista voidaan vain osin tukea keskitetyillä tietovarastoilla ja niiden analysoinnilla – osa tietämyksen luomisesta tapahtuu koostamalla tietoa satunnaisemmin hankituista tietolähteistä. Samoin monet prosessit vaativat monipuolisempaa ihmisten välistä vuorovaikutusta, jota ei voi hoitaa digitaalisilla lomakkeilla: esimiehen kanssa hoidettavaa vuotuista kehityskeskustelua tai potilaan diagnoosia ei kannata yrittää kovin tarkasti määritellyillä proseduureilla ja lomakkeilla.

Edellisestä tiedon hankintaan liittyvästä esimerkistä kannattaa myös huomata, että tehokkaamman tiedon hallinnan ja tiedon saatavuuden puuttuessa ihmiset eivät yleensä käytä seitsemää päivää yhden tiedon hankkimiseen. He etsivät tietoa jonkin aikaa ja luopuvat. Tiedon puuttuessa päätökset tehdään arvailujen, ei tiedon varassa. Lopputuloksena on yleensä korkeampi riskitaso ja keskimäärin huonompi päätösten tuottama lopputulos. Ja tuskaa, kun päätöksen

tekijä tiedostaa tämän. Tällaisessa tilanteessa tiedon puute siis lisää tuskaa, ei tieto.

Kuvassa 19 on esimerkki tämän tuskan määrästä erään organisaation kyselyssä saatujen vastausten perusteella. Kysymys oli: "Kuinka usein et löydä etsimääsi tietoa, mutta se tuleeikin sattumalta vastaan myöhemmin?" Vastausten perusteella tämä tapahtuu useimmille vain pari kertaa vuodessa, mikä on yleisesti ottaen varsin hyvä tulos.



Kuva 19. Kuinka usein vastaan tulee tietoa, jota ei etsittäessä löydetty? [24].

Tiedon hankinnan ongelmien rinnalla kuvan 18 taulukosta kannattaa kiinnittää huomiota myös tiedon käsittelyn rajoituksiin.

Organisaatioissa ihmisten päivittäin kommunikoiman tiedon määrä - noin sata sivullista - pysyy suunnilleen samana, vaikka tietokoneiden kapasiteetti tuhatkertaistuukin. Organisaatio voi siis helposti tukkia sisäisen kommunikointinsa ja puolittaa työskentelytehonsa yksinkertaisesti lähettämällä jokaisesta tiedosta kopio kaksinkertaiselle määrälle vastaanottajia.

Kaiken tämän perusteella voidaan perustellusti sanoa, että dokumenttien digitalisointi johtaa helposti hallitsemattomaan informaatiotulvaan, joka lisää tiedon vastaanottajan tuskaa tiedon käsittelyn viemän ajan takia. Syynä on siis tilanne, jossa tietoa ei hallita. Toisaalta, vaihtoehtona on aiemmin ollut tilanne, jossa tiedon puute on lisännyt tuskaa.

Tietotulvan tai tiedon puutteen aiheuttaman epävarmuuden lisäksi tietojenkäsittelyn digitalisoimattomuus tai digitalisointi voivat aiheuttaa muitakin inhimilliseen toimintaan kuuluvia oleellisia ilmiöitä. Ihmisten, yksilöiden vaikutusta ei voi vähätellä missään organisaation toimintojen kehittämisessä. Muutosvastarinta ja muutosjohtaminen ovat keskeisessä roolissa. Yleensä muutos onnistuu parhaiten saamalla ihmiset mukaan jo varhaisessa vaiheessa kertomaan mielipiteensä. Tämä pätee erityisesti dokumenttien hallinnassa, missä kenelläkään yksittäisellä henkilöllä tai ryhmällä ei ole kaikkea muutoksen suunnitteluun tarvittavaa tietoa.

Dokumenttien hallinnan kehittämisessä yksi keskeinen kulttuurimuutos liittyy tietojen jakamiseen muiden kanssa. Omien tietovarastojen käyttämisen sijasta omalla tietokoneella (PC - personal computer) puhutaankin organisaation tiedoista organisaation yhteisissä, jaetuissa tietovarastoissa.

Millainen on organisaation tiedon jakamisen kulttuuri? Tätä voi selvittää esimerkiksi aiemmin esitellyllä tiedon käytettävyyden kartoituksella [24]. Yhden organisaation tuloksia:

- Varahenkilöttömiä 65 %
- Halu/mahdollisuudet luopua henkilökohtaisista tietovarastoista:
"Voisitko luopua henkilökohtaisista tietovarastoista, mikäli tietotekniset ratkaisut tarjoaisivat siihen mahdollisuuden?"
 - En ollenkaan 9,0 %
 - Osittain 59,3 %
 - Lähes kaikista 22,0 %
 - Kaikista 9,0 %

Tiedon käsittelyn kulttuurimuutoksella organisaatiossa voi olla monia vaikutuksia:

- **Tehostuminen**
Kun informaation hallintaa kehitetään, organisaation tehokkuus paranee erityisesti tiedon uudelleenkäytön avulla ja parhaiden käytäntöjen leviämisen kautta.
- **Toiminnan läpinäkyvyys**
Ihmiset saattavat kokea yksilöllisyytensä tulevan loukatuksi. Kun muut näkevät "minun tietoni" he näkevät, miten toimin... he näkevät "minut" - vai "meidän tietomme" ja "meidän toimintatapamme"?
- **Tieto on valtaa**
Kun tieto on teknisesti jaettavissa kaikille, joudutaan julkisesti määrittelemään, mitä tietoja kukin voi käyttää. Ennen tämän hoiti jokainen itse jakamalla tietoaan muille halunsa mukaan. Nousee esille uusi kysymys: Kuka määrittelee oikeudet tietoon?
- **Yhteistyön tarve korostuu**
Muutos vie aikaa, vaikka tekniikkaa olisi.

Kannattaa muistaa, että muutos organisaation toiminnassa tapahtuu vain, jos se on työntekijöille kokonaisuudessaan edullista. Esimerkiksi myyntimiehet eivät jaa asiakaskontaktitietojaan toisilleen, jos he samalla menettävät valtaosan myyntipalkkioistaan. Yleensä tekniikan käyttöönotto on helpompi osa muutosta ja asenteiden ja toimintatapojen muutos vie enemmän aikaa.

9 Organisaation informaation hallinnan kehittäminen

Millä keinoilla organisaation kannattaa kehittää dokumenttien hallintaa ja miksi? Oheinen taulukko 1 tiivistää keskeiset päämäärät ja keinot.

Taulukko 1. Organisaation dokumenttien hallinnan kehittämisen päämääriä ja keinoja.

Päämäärä	Keinot
Tiedon ja osaamisen muodostuminen	Tiedostaminen, oppiminen
Tiedon jaettavuus ja varmuus tiedon saatavuudesta	Tiedon tallentaminen ja siirtyminen PC:stä organisaation dokumenttien ja informaation hallintaan
Oppiva organisaatio	Tiedon jakamisen kulttuurin edistäminen
Tiedon tulkittavuus	Organisaation yhteiset genret määriteltä
Tietokoneavusteisuus ja hallittavuus	Tiedon formaalisuuden asteen nostaminen ydinprosesseissa ja sopivan granulariteetin valinta
Verkostoitumisen tukeminen	Elektronisen liiketoiminnan standardien mukaisten prosessien ja genrejen käyttö.
Tietotulvan vähentäminen	Keskittetty tietovarasto. Tarpeettomat kopiot pois jakelusta.
Tiedon löytyminen	Tiedon hankinnasta tiedon hakuun organisoimalla ydinprosessien tieto.
Tuskan vähentäminen	Parempia päätöksiä, kun tieto löytyy ja oikea tieto on käytettävissä. Tietotulvan vähentäminen.

Miten asiassa kannattaisi edetä? Yksinkertaistettu lähestymistapa perustuu seuraavien kysymysten vastaamiseen (ja tulosten kirjaamiseen ja jakamiseen):

- Mikä on organisaation ja sen eri prosessien tavoite?
- Mikä on kokonaisuuden laajuus - mitä prosesseja käsitellään?
- Mitä tietoja (genrejä) prosessien sisällä käsitellään ja niiden välillä vaihdetaan?
- Mitä sisältöjä halutaan julkaista, mitä kanavia käyttäen?
- Personoidaanko tai lokalisoidaanko julkaisuja henkilö- tai ryhmäprofiilien avulla tai muuten?
- Mitkä genret tulisi tallentaa ja mitkä voivat olla tilapäisiä? Miten hoidetaan tallentamattoman ja hiljaisen tietämyksen siirto?
- Mitkä genret käsitellään dokumentteina, mitkä tietokannoissa, mitkä sisällönhallintajärjestelmällä?

- Mitkä ovat näiden dokumenttityyppien tai sisältökomponenttiluokkien tietosisällöt, metatiedot, formaatit ja mediat?
- Mitä genrejä siirretään organisaatioiden välillä, mitä eBusiness-formaatteja niille käytetään?
- Mikä on prosessin kannalta tarkoituksenmukainen dokumenttitiedon formaalisuuden taso?
- Minkä kokoisina rakenneyksikköinä tietoja käsitellään (granulariteetti)?
- Mitä metatietoja tarvitaan?
- Mitä hallintajärjestelmiä, työnkulkujärjestelmiä, extra/intranet, e-business yms. sovelluksia kannattaa käyttää näiden tietojen käsittelyyn? Integrointi muihin järjestelmiin?
- Dokumentointi ja koulutus kaikille asianosaisille!

Lisätietoja löytyy esimerkiksi METODI-projektin tuottamasta CD-ROM-muotoisesta menetelmäkirjastosta [14], joka sisältää dokumentoidut menetelmät (prosessikuvaukset, roolit, dokumenttipohjat jne.) selainsovelluksena sekä raportteja niiden käyttökokemuksista pilottiorganisaatioissa.

10 LÄHTEET JA LISÄTIETOJA

Lisätietoja: Prof. Pasi Tyrväinen, Jyväskylän yliopisto, Informaatioteknologian tiedekunta, Digitaalinen media, (Agora C515.3) PL 35, 40351 Jyväskylä, puh. (014) 260 3093, fax. (014) 260 3068, Pasi.Tyrvainen@jyu.fi, <http://www.jyu.fi/~pttyrvai>

1. DMA, Document Management Alliance.
2. IDC, International Data Corporation.
3. Tyrväinen, P. *Estimating Applicability of New Mobile Content Formats to Organizational Use*. in *36th Hawaii International Conference on Systems Sciences*. 2003. Hawaii, USA: IEEE Computer Society Press.
4. Tyrväinen, P., Järvenpää, M., and Sievänen, A. *On Estimating the Amount of Learning Materials, A Case Study*. in *5th International Conference on Enterprise Information Systems*. 2003. Angers, France: ICEIS Press.
5. Tyrväinen, P., *Millä menetelmillä dokumenttien hallintaa kehitetään teollisuudessa*, in *Multimedian teolliset sovellukset -seminaari*. 1998: Helsinki.
6. Tyrväinen, P., *Dokumenttien hallinta yrityksen tietämyksen hallinnassa*, in *METODI-päätösseminaari*. 2000, Jyväskylän yliopisto: Vantaa.
7. Tyrväinen, P. *Mikä on dokumenttien hallinnan konkreettinen merkitys liiketoiminnassa?* in *Dokumenttien hallinta -seminaari*. 2001. Helsinki: Kontakti.net.
8. Nonaka, T., *The Knowledge-Creating Company*. 1995.
9. C&L, Coopers & Lybrand.

10. Sprague, R.H., *Electronic Document Management: Challenges and Opportunities for Information Systems Manager*. MIS Quarterly, 1995. **19**(1): p. 29-49.
11. Delphi, Delphi Consulting Group.
12. Lesk, M., *How much information is there in the world?* 1997.
13. Sutton, M.J.D., *Document Management for the Enterprise - Principles, Techniques, and Applications*. 1996: John Wiley & Sons, Inc. 369.
14. METODI-projekti, *METODI menetelmäkirjasto CD-ROM*. 2000, Tietotekniikan tutkimusinstituutti, Jyväskylän yliopisto.
15. Päivärinta, T. and P. Tyrväinen, *Structuring information by genres to bridge the social and technological in information resources management - Leavitt's framework revis(IT)ed*, in *Proceedings of the 24th Information Systems Research Seminar in Scandinavia, Norway, 11-14 August 2001.*, A. Mørch, et al., Editors. 2001, University of Bergen, Department of Information Science: Ulvik in Hardanger. p. 16.
16. Yates, J. and W.J. Orlikowski, *Genres of organizational communication: A structurational approach to studying communication and media*. Academy of Management Review, 1992. **17**(2): p. 299-326.
17. Yates, J., W.J. Orlikowski, and J. Rennecker, *Collaborative genres for collaboration: Genre systems in digital media*, in *Proceedings of the 30th Annual Hawaii International Conference on System Sciences. Digital Documents Track.*, R.H. Sprague, Editor. 1997, IEEE Computer Society Press: Maui (HA). p. 50 - 59.
18. Tyrväinen, P., *Estimating applicability of new mobile content formats to organizational use*, in *Proceedings of the 36th Hawaii International Conference on Systems Sciences*. 2003, IEEE: Los Alamitos, CA. p. CD ROM, 10 p.
19. Tyrväinen, P. and T. Päivärinta, *How Digital is Communication in Your Organization? - A Metrics and an Analysis Method*, in *The Proceedings of the 5th Information Conference on Enterprise Information Systems, ICEIS*. 23-25.4. 2003, ESEO: Angers, France. p. 11.
20. Tyrväinen, P., M. Järvenpää, and A. Sievänen, *On Estimating the Amount of Learning Materials - A Case Study*, in *The Proceedings of the Information Conference on Enterprise Information Systems, ICEIS*. 23-25.4. 2003, ESEO: Angers, France. p. 9.
21. Kouloupoulos, T., M. and C. Frappaolo, *Electronic Document Management Systems. A Portable Consultant*. 1 ed. Computing. 1995, New York: McGraw-Hill, Inc., U.S.A. 273(297).
22. Talouselämä, in *Talouselämä*. 2001.
23. Tyrväinen, P., *Dokumenttien hallinnan lainalaisuuksia - sosiaalisen ja teknisen maailman rajoituksia*, in *Dokumenttien hallinta -konferenssi*. 2003, Kontakti.net: Espoo.
24. Tyynysniemi, J., *Tiedon käytettävyyden merkitys ja mittaaminen liiketoiminnassa*, in *METODI-seminaari*. 1999, Jyväskylän yliopisto: Helsinki.