

TUTKA-RYHMÄ  
LOPPURAPORTTI

LUOVUTETTU REHTORILLE 3.6.2013



JYVÄSKYLÄN YLIOPISTO



## SISÄLLYS

Johdanto.....	1
1 Toimeksiannon täsmennys.....	2
1.1 Poisrajatut tiedot.....	3
1.2 Työryhmän toiminta .....	4
2 Esitys muutoksista yliopiston tutkimustietojen kirjaamisen prosesseiksi.....	5
3 Tietosisältö .....	6
3.1 Julkaisut .....	6
3.2 Kansainvälinen liikkuvuus.....	6
3.3 Täydentävän rahoituksen hankkeet.....	6
3.4 Merkittävät tieteelliset palkinnot .....	7
4 Tutkimustietojärjestelmän keskeiset vaatimukset .....	8
4.1 Järjestelmän ja prosessien yleiset vaatimukset.....	8
4.2 Julkaisut .....	9
4.3 Kansainvälinen liikkuvuus.....	10
4.4 Täydentävän rahoituksen hankkeet.....	10
4.5 Merkittävät tieteelliset palkinnot .....	11
5 Rajapinnat .....	13
5.1 TUTKAN nykyiset rajapinnat.....	13
5.2 Suunnitteilla olevia integrointeja .....	18
5.3 Uudessa järjestelmässä huomioitavia.....	20
6 Toteutusvaihtoehdot .....	26
6.1 Jatkokehitetään nykyistä TUTKAa.....	26
6.2 Uudistetaan TUTKA tutkimustietoihin liittyviltä osaltaan.....	27
6.3 Kehitetään uusi tutkimustietojärjestelmä.....	27
6.4 Ostetaan tutkimusjärjestelmä tai sen osia ulkopuolelta.....	28
7 Loppupäätelmät.....	29



## JOHDANTO

Rehtori Matti Manninen asetti TUTKA-ryhmän 10.11.2012. Työryhmän tehtävinä oli

1. selvittää, mitä tutkimustoimintaan liittyviä tietoja yliopiston tulee tuottaa ja tallentaa OKM-raportointia ja yliopiston sisäistä kehittämistä varten (tämä sisältää tietoihin liittyvien käsitteiden kuvaukset ja tietojen keruun täsmennetyn ohjeistuksen)
2. esittää kohdan 1 tulosten edellyttämät muutokset yliopiston tutkimustoiminnan prosesseihin
3. tehdä ehdotus edellä olevia toimintoja varten tarvittavista järjestelmäpalveluista ja niiden toteuttamisvaihtoehdoista sekä kytkennöistä yliopiston muihin järjestelmäpalveluihin ja järjestelmäarkkitehtuuriin.

Työryhmän jäseninä olivat vararehtori Kaisa Miettinen (puheenjohtaja), tietohallintopäällikkö Antti Auer, suunnittelija Jenni Kokko (sihteeri), tutkimuspäällikkö Sirkka-Liisa Korppi-Tommola, suunnittelupäällikkö Sakari Liimatainen, järjestelmäsuunnittelija Miika Nurminen, professori Jari Ojala ja johtaja Kimmo Tuominen.

Jyväskylässä 3.6.2013



## 1 TOIMEKSIANNON TÄSMENNYS

Tässä raportissa kuvataan Jyväskylän yliopiston tarpeisiin soveltuvan tutkimustietojärjestelmän toimintaa ja ominaisuuksia. Myöhemmin sanalla järjestelmä tarkoitetaan tällaista tutkimustietojärjestelmää. Keskeisiksi järjestelmään kirjattaviksi tutkimustiedoiksi on valittu

- julkaisut,
- kansainvälinen liikkuvuus,
- täydentävän rahoituksen hankkeet sekä
- merkittävät tieteelliset palkinnot.

Kokonaisarkkitehtuuri-hankkeessa määriteltyjen arkkitehtuuriperiaatteiden mukaan tietoarkkitehtuurin tavoitetilassa pyritään kohti yleisesti hyväksyttyjä kansallisia ja eurooppalaisia tutkimus- ja opetustoiminnan tietoarkkitehtuureja (XDW ja CERIF/euroCRIS), jotka ovat pohjana myös tietojen standardimuotoisessa raportoinnissa opetus- ja kulttuuriministeriölle (OKM). Arkkitehtuurityöhön kuuluu olennaisena osana organisaation ja sen yksiköiden toiminnan tavoitetilän kuvaaminen. Ydinprosessit ovat ensisijaisia ja tavoitetilän kuvausten perustana ovat toimintaprosessien kuvaukset. Tietojärjestelmien tulee tukea toimintaprosesseja ja kattaa määritellyt tietosisällöt.

Työryhmän valinta keskeisiksi tutkimustiedoiksi noudattaa yliopistojen rahoitusmallin linjauksia. Mallin tutkimusosion rahoituskriteereillä yliopistoja ohjataan tutkimustoiminnan perusedellytysten kehittämiseen. Tutkimustoiminnan keskeisin tuotos ovat julkaisut, ja ne muodostavat suurimman rahoituskokonaisuuden tutkimustoiminnan perusteella jaettavasta rahoituksesta ja toiseksi suurimman kokonaisuuden koko rahoitusmallissa. Toinen tärkeä tutkimustoiminnan kansainvälistymistä kuvaava kriteeri rahoitusmallissa on tutkijoiden kansainvälinen liikkuvuus. Sekä kansallinen että kansainvälinen täydentävä rahoitus ovat myös rahoitusmallin merkittäviä tekijöitä.

Rahoitusosuuden laskentaa ja toiminnan raportointia varten yliopistot toimittavat julkaisu- ja muut tutkimustoimintaa kuvaavat tiedot vuosittain opetus- ja kulttuuriministeriöön. On erittäin tärkeää välittää tutkimustiedot opetus- ja kulttuuriministeriölle oikein. Ohjeet kerättävistä sisältökentistä sekä tietojen muodosta on koottu ministeriön tiedonkeruukäsikirjaan, joka sisältää tarkat käsitelmäärittelyt niille tiedoille, jotka yliopistoilta edellytetään ministeriön valtakunnallisessa tiedonkeruussa. Jatkossa käsitteet tulevat sisällyttämään XDW-käsitelmalliin.

Yliopiston sisäisessä toiminnanohjauksessa, kehittämisessä ja raportoinnissa tutkimustiedoilla on myös merkittävä rooli. Yliopisto on siirtänyt valtakunnallisen rahoitusmallin ohjausvaikutuksen omaan sisäiseen rahoitusmalliinsa sisällyttämällä valtakunnallisen mallin osaksi sisäisen rahanjaon perusteita. Tiedoneuvosto laatii tiedebarometrin kootun tutkimustiedon perusteella. Tutkimustiedot ovat olennaisinta yliopiston toiminnan raportointia, ja niitä tarvitaan myös tieteellisten ansioitten arviointiin ja palkkauskeskusteluihin.

Keskeisiksi tutkimustiedoiksi on valittu myös merkittävät tieteelliset palkinnot. Palkintotiedot eivät sisälly ministeriön rahoitusmalliin, mutta ne ovat hyödyllisiä mm. tiedebarometria ja tutkimuksen näkyvyyttä ajatellen.



Tutkimustiedon käyttötarkoitukset on priorisoitu seuraavasti:

1. Viranomaisraportointi
  - Yliopistotasoiset julkaisu- ja vierailuraportit
2. Sisäinen raportointi
  - Tulosohtausraportointi
  - Barometrit
  - Yksikötasoiset raportit
3. Tutkimuksen näkyvyys
  - Hankkeiden ja partnerien seuranta
  - Yksiköiden www-sivujen päivittäminen, asiantuntijarekisteri
  - Julkaisujen avoin rinnakkaistallentaminen
4. Hanketoiminnan sisäinen seuranta
  - Apurahat, hankkeiden tuotokset
  - Rahoitushakemusten seuranta
5. Muu raportointi
  - Laadun arviointi
  - Kansainväliset arvioinnit

Työryhmä esittää, että tutkimustiedot kirjataan (pääsääntöisesti) keskitetysti yksiköissä. Tietojen tarkastusprosessi laajennetaan julkaisuista kaikkiin tutkimustietoihin. Näin toisaalta vastaan tiedoneuvoston tavoitteeseen antaa tutkijoille mahdollisimman paljon aikaa tutkimukselle vähentäen tukitoimiin käytettyä aikaa, ja toisaalta varmistetaan tietojen oikeellisuus. Esitys tarkoittaa, että kussakin yksikössä on yhdyshenkilö, jolle tutkijat toimittavat tutkimustiedot. Kirjaamisprosessiin liittyy lisäksi mahdollisesti muita, tietojen oikeellisuudesta eri osioissa vastaavia henkilöitä.

### 1.1 Poisraajatut tiedot

Rajaus keskeisiksi tutkimustiedoiksi tarkoittaa nykyiseen järjestelmään verrattuna tietosisällön rajaamista muusta kuin tutkimustiedosta. Poisraajatut tiedot ovat pääosin CV-tyyppisiä tietoja, jotka käyttötarkoituksensa kannalta kuuluvat muihin järjestelmiin. CV-tyyppisiä tietoja tarvitaan muissa yhteyksissä ja ne ovat hyödyllisiä esim. tutkimuksen näkyvyyden ja henkilöstön osaamisen seurannan kannalta, mutta ne eivät ole tutkimustietojärjestelmän pääasiallinen sisältö. Nykyisessä TUTKAssa kerättäviä, poisrajattavaksi esitettyjä CV-tyyppisiä tietoja ovat esitelmät ja esitykset, lausunnot sekä muu tieteellinen toiminta.

Nykyiseen TUTKAan on syötetty myös opinnäytteiden ohjauksia ja niitä on siirretty myös Korpista. Ohjaukset ovat tarpeellisia tietoja mm. YPJ-raporteissa ja laitosten työsuunnitelmien seurannassa, mutta opintoasioihin liittyvinä tietoina ne tulisi ensisijaisesti kerätä Korpissa, jossa on jo valmiudet ohjausten kirjaukseen ja raportointiin, sekä TUTKAN toimintoihin verrattuna myös perustoiminnallisuudet ohjausprosessin tukemiseen. Korppia käytetään nykyisellään jatko-opiskelijaksi ilmoittautumiseen koko yliopiston laajuisesti. Työryhmä esittää, että opinnäyteosiota kehitettäisiin niin, että se voidaan ottaa koko yliopiston käyttöön muissakin opinnäytteissä. Tutkimustietojärjestelmästä opinnäytteet tulee rajata pois.



Yhteiskunnallisen vuorovaikutuksen barometrissa arvioidaan useita sellaisia tietoja, joita ei nykyisellään kerätä missään tietojärjestelmässä. Nykyisen TUTKAN sisältöihin YVV-barometrin tietosisällöistä asiantuntijapalvelut ja kaupalliset oikeudet joiltakin osin soveltuvat, mutta esitetyn tietosisällön rajauksen mukaan edellä mainittuja tietoja ei tutkimustietojärjestelmässä kerätä.

Tilastokeskuksen tieteenalaluokituksen lisäksi on perusteltua harkita muidenkin tieteenalaluokitusten käyttöä, mutta työryhmällä ei ole kantaa mikä sen tulisi olla. Järjestelmän tulisi mahdollistaa luokkien muuntaminen eri luokitusten välillä niin, että yhdellä kirjauksella voidaan automaattisesti päätellä luokkia useamman luokituksen mukaisesti.

## **1.2 Työryhmän toiminta**

Työryhmä on kokoontunut 11 kertaa. Työryhmä kutsui kuultavaksi kolme asiantuntijavieraita. Ville Seppänen kertoi kokonaisarkkitehtuurihankkeen tuloksista tutkimustietojen osalta ja tutkimuspalvelujohtaja Tanja Grönlund Lappeenrannan teknillisestä yliopistosta kertoi taannosta tutkimustietojärjestelmäkilpailutuksesta ja sen tuloksista. Lisäksi tietoasiantuntija, projektipäällikkö Leena Huiku Tampereen teknillisen yliopiston kirjastosta kertoi tutkimusjärjestelmähankinnan tilasta.



## 2 ESITYS MUUTOKSISTA YLIOPISTON TUTKIMUSTIETOJEN KIRJAAMISEN PROSESSEIKSI

Työryhmä esittää, että tutkimustiedot tulee kirjata (pääsääntöisesti) keskitetysti esim. yksiköissä. Tämä edellyttää, että tutkija toimittaa asianmukaiset tiedot yhdyshenkilölle aina kun tutkimustietoja syntyy (ei esim. vasta vuoden lopussa). Toisin sanoen julkaisutietoja, liikkuvuustietoja ja merkittäviä palkintoja ei yksittäinen tutkija kirjaa itse tutkimustietojärjestelmään. Liikkuvuustiedot tulevat pääsääntöisesti SAP Travel -järjestelmästä. Luvussa 3.3 *Täydentävän rahoituksen hankkeet* kuvattavan hankeilmoituksen tiedot voi tutkijan syöttää itse, jos hankkeen rahoitushakemuksen lähettämisen kiireellisyys sitä edellyttää.

Erityyppisillä tiedoilla on erilaiset tarkistamis- ja hyväksymiskäytännöt. Tutkimustietojärjestelmän tulee tukea näitä prosesseja.

Järjestelmässä tulee olla mahdollisuus tallentaa tietojen hyväksymiseen liittyvä sähköinen viestintä, joka sisältää perustelut esim. julkaisujen luokittelulle. Kirjoittajan tulee tarvittaessa pystyä todistamaan, että esim. proceedings- ja kokoomateosartikkelit ovat käyneet läpi refereeprosessin.

Järjestelmään syötetty tieto on pääsääntöisesti julkista.

Nykyisin useimmat tutkijat kirjaavat tietonsa itse käytössä olevaan TUTKA-järjestelmään. Muutamat laitokset ovat organisoineet tietojen syötön keskitetysti. Jokaisella laitoksella on TUTKA-laitospääkäyttäjä ja julkaisujen hyväksyjä (vastuuhenkilö). Nämä voivat olla myös sama henkilö. Kun tutkija nykyisin syöttää julkaisun tiedot järjestelmään, tulee oman laitoksen vastuuhenkilön tarkastaa ja hyväksyä syötetyt tiedot ennen kuin julkaisu näkyy julkisissa raporteissa.

Esitetyssä kirjaamistavassa yhdyshenkilö kirjaa julkaisujen tiedot ja tutkija tarkistaa niiden oikeellisuuden. Tarvittaessa oikeellisuuden tarkistaa vielä laitoksen tai kirjaston vastuuhenkilö. Tutkimushankkeilla on hyväksymisprosessinsa, jossa laitoksen johtaja hyväksynnällään antaa luvan tutkimusrahoitushakemuksen lähettämiseen.



### 3 TIETOSISÄLTÖ

Tässä luvussa kuvattavissa tietosisältöjen määrittelyistä esitetään ensisijaisesti ne, jotka sisältävät muutoksia nykyiseen käytäntöön. Luvussa 4 *Tutkimustietojärjestelmän keskeiset vaatimukset* esitetään tarkemmin järjestelmän ja prosessien vaatimuksia.

#### 3.1 Julkaisut

Julkaistiedoilla tarkoitetaan ensisijaisesti opetus- ja kulttuuriministeriön keräämiä julkaisutietoja. Muutoksena nykyiseen käytäntöön työryhmä esittää, että henkilöstön kaikki julkaisut voidaan kirjata järjestelmään eikä vain niitä, jotka on tehty Jyväskylän yliopistossa työskennellessä. Nämä tiedot ovat olennaisia yliopiston tutkimuksen kokonaisarviointia tehtäessä. Koska ministeriöön raportoidaan vain ne julkaisut, joissa tekijällä on yhteys Jyväskylän yliopistoon työsuhteen tai ohjauksen kautta, tulee järjestelmässä voida erottaa Jyväskylän yliopistossa ja muualla tehdyt julkaisut.

Tutkija toimittaa julkaisun, mahdollisesti siitä rinnakkaistallennettavan version ja tekijän julkaisuluvan sekä muut olennaiset tiedot yhdyshenkilölle pääsääntöisesti sähköisesti.

Julkaistuista kerättävät tiedot perustuvat opetus- ja kulttuuriministeriön tiedonkeruukäsikirjaan ja korkeakoulujen XDW-tietomalliin mahdollisilla yliopistokohtaisilla tarkennuksilla täydennettyinä (esim. sisäisen raportoinnin tai tutkimuksen näkyvyyden tueksi).

#### 3.2 Kansainvälinen liikkuvuus

Kansainvälisellä liikkuvuudella (=matkatiedot) tarkoitetaan henkilöstön ja jatko-opiskelijoiden vierailuja ulkomaille sekä Jyväskylän yliopistoon saapuvia ulkomaisia vieraita opetus- ja kulttuuriministeriön ohjeistuksen mukaan. Matkatiedoista opetus- ja kulttuuriministeriö kerää opetus- ja tutkimusvierailut.

#### 3.3 Täydentävän rahoituksen hankkeet

Hankkeella tarkoitetaan täydentävän rahoituksen tutkimushanketta. Mikäli niin halutaan, vastaavalla tavalla voidaan käsitellä myös kehittämis- ja koulutushankkeita.

Työryhmä esittää periaatteeksi, että hankkeista syötetään järjestelmään tietoa vain hakuvaiheessa ja rahoituspäätöksen jälkeen kaikki olennainen tieto tulee suoraan taloustietojärjestelmästä. Työryhmä esittää, että nykyinen riskianalyysilomake korvataan hankeilmoituksella, joka on tutkimustietojärjestelmän osa. Hankehallintoon on lähemmin syventynyt hankehallinnon työryhmä.

Hankeilmoituksen sisältämät olennaiset tiedot ovat seuraavat: hankkeen nimi, kesto, vastuullinen johtaja ja mahdollinen varajohtaja, tiivistelmä (julkinen abstrakti), haettu rahoituksen kokonaismäärä rahoittajittain, rahoittajat, konsortio (ts. muut mukana olevat yliopistot, tutkimuslaitokset, yritykset tms.), mahdolliset riskit ja talousvastaava (budjetin tarkastaneen henkilön nimi). Kun rahoituspäätös tulee, yhdistetään rahoituspäätösnumero (SAP Talous - järjestelmästä) ja muut tarvittavat lisätiedot hankeilmoituksessa annettuihin tietoihin mahdollisimman automaattisesti, jolloin saadaan kokonaiskuva rahoitettujen hankkeiden perustiedoista.





Tällä hetkellä yliopistolla ei ole kattavaa tietoa säätiöiden myöntämästä ulkopuolisesta rahoituksesta. Tilannetta korjaa se, jos kaikki apurahatutkijat tekevät apurahansaajan sopimuksen ja myös säätiöille lähetettävistä hakemuksista tehdään hankeilmoitus.

Muu projektinhallinta ei olennaisesti kuulu tutkimustietojärjestelmään.

Tutkimustietojärjestelmää ei tulla käyttämään opetus- ja kulttuuriministeriön taloustiedonkeruuseen vaan talouspalvelut kerää tiedot suoraan SAP:sta.

### **3.4 Merkittävät tieteelliset palkinnot**

Merkittävä tieteellinen palkinto on palkinto, joka on myönnetty Jyväskylän yliopiston tutkijalle tunnustuksena korkeatasoisesta ja kansainvälisesti erityisen merkittävästä tieteellisestä tutkimuksesta tai tieteellisestä läpimurrosta (esimerkiksi Suomen tiedepalkinto, Suomen tiedeakatemian palkinto, muut merkittävät kotimaiset palkinnot ja kansainvälisten tieteellisten seurojen myöntämät palkinnot).

Järjestelmään kirjataan vain Jyväskylän yliopistossa työskennellessä saadut merkittävät tieteelliset palkinnot.

Merkittävistä tieteellisistä palkinnoista kerätään seuraavat tiedot: palkinnon saajan nimi, tieteellinen ala, palkinnon myöntäjän nimi, myöntövuosi ja palkinnon perustelu myöntöasiakirjasta.



## 4 TUTKIMUSTIETOJÄRJESTELMÄN KESKEISET VAATIMUKSET

Tässä luvussa kuvataan keskeiset vaatimukset tutkimustietojärjestelmälle ja sen edellyttämille prosesseille. Luku toimii pohjana varsinaiselle vaatimusmäärittelylle uutta tutkimustietojärjestelmää määriteltäessä.

### 4.1 Järjestelmän ja prosessien yleiset vaatimukset

- Tietomallin tulee olla muokattavissa vähintään kenttien ja eri tietokohteiden linkitysten osalta (käsittäen sekä tietosisällön määrityksen että validoinnin).
- Järjestelmän tuottamien raporttien tulee olla muokattavissa joustavasti (käsittäen hakuehdot, käytettävät kentät, tulosten muotoilun ja käytettävän formaatin). Varsinaisten raporttipohjien lisäksi myös ad hoc -tyylisten kyselyjen tulee olla mahdollisia.
- Tietojen tulee olla käsiteltävissä suoraan ohjelmointi- tai tietokantarajapinnan kautta. Tietojen siirto tutkimustietojärjestelmästä tietovarastoon ja toisaalta omista tietolähteistä tutkimustietojärjestelmään tulee olla automatisoitavissa.
- Järjestelmän tulee tukea työvuon hallintaa eri osioissa ja yksikkökohtaisesti käsittäen yleisesti tietojen keskitetyn kirjauksen ja tietojen ohjauksen hyväksyntästatuksen perusteella.
  - Hyväksymisprosessin vaiheen tulee olla näkyvillä järjestelmässä
  - Kirjaajien ja vastuuhenkilöiden tulee nähdä omaan yksikköönsä liittyvät, eri vaiheissa olevat tiedot. Prosessin kulku voi vaihdella eri osioissa (esim. tutkija tarkastaa julkaisunsa tiedot, laitoksen johtaja hyväksyy hankeilmoitukset), ja rooleihin liittyvät henkilöt voivat vaihdella yksikkökohtaisesti.
  - Järjestelmän tulee helpottaa tutkijoiden ja vastuuhenkilön välistä, yksittäiseen tietokohteeseen sidottua viestintää (esim. korjauspyynnöt, muistutukset tarkennusta vaativista tiedoista ja ilmoitukset prosessin etenemisestä).
  - Järjestelmässä tulee olla mahdollisuus tallentaa tietojen hyväksymiseen liittyvä (järjestelmän puitteissa käyty) viestintä liitetiedostoihin.
- Yksittäisen tietueen julkisuusasteen tulee olla määriteltävissä monipuolisesti.
  - Tiedon julkisuusasteen tulee olla muutettavissa kirjausprosessin eri vaiheissa. Esim. tietue näkyy julkisena vasta hyväksymisen jälkeen.
  - Yksittäisten kenttien tai tietojen osakohteiden (esim. hankehakemukset) näkyvyyden tulee olla rajattavissa.
  - Järjestelmään julkiseksi määriteltujen tietojen (oikeaksi todennetut julkaisut, kansainvälinen liikkuvuus, myönnetty rahoitus ja merkittävät tieteelliset palkinnot) tulee olla helposti selattavissa ja hakukoneiden indeksoitavissa.
- Järjestelmässä tulee olla toiminto tuplakirjausten havaitsemiseen ja tarvittaessa estämiseen mahdollisimman varhaisessa vaiheessa. Samoja tietoja sisältävien tietueiden tulee olla yhdistettävissä.



- Järjestelmän tulee tuottaa historiatietoa tehdyistä muutoksista niin, että järjestelmän tila voidaan tallentaa haluttuna ajanhetkenä esim. tietovarastoon<sup>1</sup>.
  - Valituissa tietokohteissa (esim. laitostiedot) historiatiedon tulee olla käytettävissä järjestelmän sisällä.
  - Historiatietojen lisäksi järjestelmän tulee tuottaa reaaliaikaista lokitietoa merkittävistä tapahtumista käyttäjän tarkkuudella (esim. tietueen muokkaus).
- Järjestelmän tulee noudattaa vähintään rajapintatasolla kansallisia ja (soveltuvien osin) kansainvälisiä tutkimustietoihin liittyviä standardeja (esim. CRIS/CERIF, XDW, OpenAIRE-suositukset).
  - Järjestelmän tulee tukea tietojen raportointia opetus- ja kulttuuriministeriön tiedonkeruun määritysten mukaisesti.
  - Järjestelmän tulee tukea Tilastokeskuksen tieteenalaluokitusta ja mahdollisesti muita tieteenalaluokituksia.
  - Järjestelmässä tulee olla valmius organisaation ja tutkijan tunnisteiden (esim. Haka, ORCID) kirjaamiseen.
- Järjestelmään tulee voida kirjautua Jyväskylän yliopiston käyttäjätunnuksilla. Käyttäjähallinnassa (esim. tutkijoiden yhdistäminen laitoksiin) tulee hyödyntää Jyväskylän yliopiston keskitettyä identiteettinhallintaa (IDM) ja roolipohjaista pääsynhallintaa (IAM).
- Käyttöliittymän tulee noudattaa saavutettavuuden (accessibility) periaatteita standardien ja kansallisten suositusten mukaisesti.

## 4.2 Julkaisut

- Julkaisun kirjaamisen tulee olla helppoa ja mahdollisimman vähän virhealtista.
  - Julkaisutietojen tulee olla poimittavissa valituista kansainvälisistä tietokannoista (Web of Science, Scopus).<sup>2</sup>
  - Julkaisutietojen tulee olla poimittavissa valituista kansallisista järjestelmistä (esim. ARTO).
  - Julkaisutietojen tulee olla siirrettävissä JYKDOKista.
  - Julkaisujen käyttöliittymän tulee mahdollistaa aiempien tietojen hyödyntäminen uusia tietoja kirjatessa (esim. samaan kokoomateokseen liittyvien artikkelien tulisi käyttää samoja tietoja kokoomateokseen yleisesti liittyvien kenttien osalta).
  - Järjestelmän tulee tukea Julkaisufoorumi-luokituksen laskentaa ja helpottaa kirjausvaiheessa julkaisun tietojen yhdistämistä Julkaisufoorumissa listattuihin julkaisukanaviin.
- Julkaisut tulee olla mahdollista kirjata keskitetysti työvuon hallintaa hyödyntäen. Tutkija vahvistaa julkaisun kirjauksen.

<sup>1</sup> Yliopiston järjestelmien historiointi tulee järjestää yhtenäisesti ja kestävästi osana yliopiston kokonaisarkkitehtuuria. Historiointi voidaan toteuttaa jokaisessa järjestelmässä erikseen tai esim. erillisessä järjestelmien yhteisessä historiointi- ja arkistointijärjestelmässä. Tietovarasto ei ole arkistojärjestelmä, mutta historiajärjestelmä voidaan toteuttaa tietovarastoinnin keinoin.

<sup>2</sup> Tietojen poiminta Web of Sciencesta ja Scopuksesta edellyttää lisenssiä, joka nykyisen FinELib-sopimuksen puitteissa sallii oman organisaation julkaisujen perustietojen siirtämisen omaan tietojärjestelmään. Neuvottelut uusista ehdoista ovat käynnissä kansallisella tasolla.



- Järjestelmän tulee mahdollistaa sekä opetus- ja kulttuuriministeriön että sisäisen raportoinnin vaatimien tietojen (ml. Julkaisufoorumi-luokitus) kerääminen Jyväskylän yliopiston omat erityisvaatimukset huomioiden.
  - Ministeriön raportointimallia laajentavien kenttien käyttö: esim. tutkijan affiliaatio yksikön tarkkuudella, koulutusalan päättely oppiaineesta, tieto julkaisun raportointivuodesta.
  - Kiinteät eräpäivät ja mahdollinen väliaikainen muutosten esto yksittäisiin osioihin (tukevat tietojen keräämistä mm. vuosittaisiin tilastoihin ja tiedebarometriin).
- Järjestelmän tulee tukea rinnakkaisjulkaisuprosessia ja julkaisujen avoimuuden määrittelyä.
  - Julkaisujen avointen rinnakkaisversioiden (yleensä post-print) tulee olla siirrettävissä automaattisesti (=muiden julkaisun metatietojen mukana) JYXiin niin, että tieto avoimen julkaisun tallennusstatuksesta ja pysyvä osoite avoimeen rinnakkaisjulkaisuun (jos rinnakkaisjulkaisuun saadaan lupa kustantajalta) palautuvat järjestelmään.
  - Järjestelmän tulee tukea lehtien ja sarjojen avoin saatavuus -tiedon selvitystä.
  - JYKDOKista siirrettyjen julkaisujen sähköiset versiot (jos tallennettu JYXiin) tulee merkitä järjestelmässä tiedonsiirron yhteydessä automaattisesti avoimiksi.
- Raportoitujen julkaisutietojen tulee olla tuotavissa takaisin järjestelmään esim. ministeriön tietovarastosta tai kansallisesta julkaisuportaalista erityisesti ministeriön tai CSC:n tekemien korjausten osalta.
  - Järjestelmässä tulee olla mahdollisuus kirjata henkilökohtaisia julkaisutietoja muissa organisaatioissa (tai esim. opiskelijana) tehdyt julkaisut, sekä
  - teokset, jotka eivät käytä ministeriön kriteerejä julkaisulle, mutta jotka ovat julkaisuja esim. tieteenalan tai tieteellisen CV:n käytäntöjen mukaisesti. Tällaisia ovat mm. tietyt toimitetut teokset ja ns. ei-traditionaalisissa julkaisukanavissa ilmestyneet julkaisut.
- Järjestelmässä tulee olla mahdollisuus kytkeä julkaisu hankkeeseen.

#### 4.3 Kansainvälinen liikkuvuus

- Järjestelmän tulee tukea matkatietojen keskitettyä kirjausta. Matkatietoja ovat henkilöstön ja jatko-opiskelijoiden vierailut ulkomaille ja ulkomaiset vieraat. Tiedoista tulee ilmetä matkan tarkoitus.
- Matkatietojen tulee ensisijaisesti siirtyä matkanhallintajärjestelmästä (SAP Travel).
- Järjestelmän tulee mahdollistaa myös tietojen manuaalinen syöttö ja täydentäminen. Manuaalisesti syötettäviä tietoja ovat esim. apurahatutkijoiden ja ulkopuolisten vierailijoiden matkat.
- Matkatiedot tulee voida raportoida opetus- ja kulttuuriministeriöön ministeriön tiedonkeruuohjeiden mukaan.

#### 4.4 Täydentävän rahoituksen hankkeet

- Järjestelmän tulee tukea hanketietojen keskitettyä kirjausta. Hankkeella ymmärretään täydentävän rahoituksen hanketta.



- Hanketiedoista ilmoittaminen koskee myös apurahatutkijoita, mikäli tutkimus tehdään Jyväskylän yliopistossa
  - Hankehallinnolla ei tarkoiteta esim. operatiivista talousseurantaa
- Järjestelmän tulee tukea rahoittajien hallintaa (esim. hierarkkisesti luokitellut rahoittajat, rahoittajien ja rahoitusohjelmien erottaminen, rahoittajakohtaisesti asetettavat vuosirajoitteet, rahoittajatietojen keskitetty muokkaus).
- Järjestelmässä tulee olla mahdollisuus syöttää hankeilmoitus, joka sisältää vähintään seuraavat tiedot:
  - hankkeen nimi
  - kesto
  - vastuullinen johtaja ja mahdollinen varajohtaja sekä näiden muutokset hankkeen kestoajana
  - tiivistelmä (julkinen abstrakti)
  - haettu rahoituksen kokonaismäärä rahoittajittain
  - rahoittajat
  - konsortio (ts. muut mukana olevat yliopistot, tutkimuslaitokset, yritykset tms.)
  - mahdolliset riskit
  - talousvastaava (budjetin tarkastaneen henkilön nimi)
  - mahdollisuus tallentaa liitteitä (esim. hankesuunnitelma)
- Järjestelmän tulee tukea hankeilmoituksen hyväksymisprosessia määritellyllä työvuolla.
  - Talousvastaava tarkistaa tarkistettavaksi määritellyt budjetit
  - Laitoksen johtaja tai vastaava antaa luvan rahoitushakemuksen lähettämiseksi hyväksymällä hankeilmoituksen
  - Hankeilmoituksen voi hyväksyä monessa eri roolissa
  - Järjestelmässä tulee olla mahdollisuus liitteiden tallentamiseen (esim. lähtenyt hakemus, sopimukset). Osa liitteistä voi olla pysyvästi (mutta ei välttämättä julkisesti) tallennettavaa tietoa.
- Järjestelmässä tulee olla mahdollisuus siirtää rahoituspäätöstiedot SAP Talous - järjestelmästä ja muista rahoittajien järjestelmistä (kuten Suomen Akatemia ja Tekes)
  - Muista järjestelmistä siirrettävien rahoituspäätöstietojen (esim. päätösnumero, myönnetty kokonaisrahoitus) tulee olla yhdistettävissä mahdollisimman automaattisesti kirjattuun hankeilmoitukseen.
- Muista järjestelmistä siirrettävien rahoittajatietojen tulee olla yhtenäistettävissä tutkimustietojärjestelmään merkittyjen rahoittajien kanssa.

#### 4.5 Merkittävät tieteelliset palkinnot

- Järjestelmän tulee tukea Jyväskylän yliopistossa työskennellessä saatujen merkittävien tieteellisten palkintojen keskitettyä kirjausta.
- Palkinnoista kerätään seuraavat tiedot:
  - palkinnon saajan nimi
  - tieteenala
  - palkinnon myöntäjän nimi
  - myöntövuosi



- palkinnon perustelu myöntöasiakirjasta

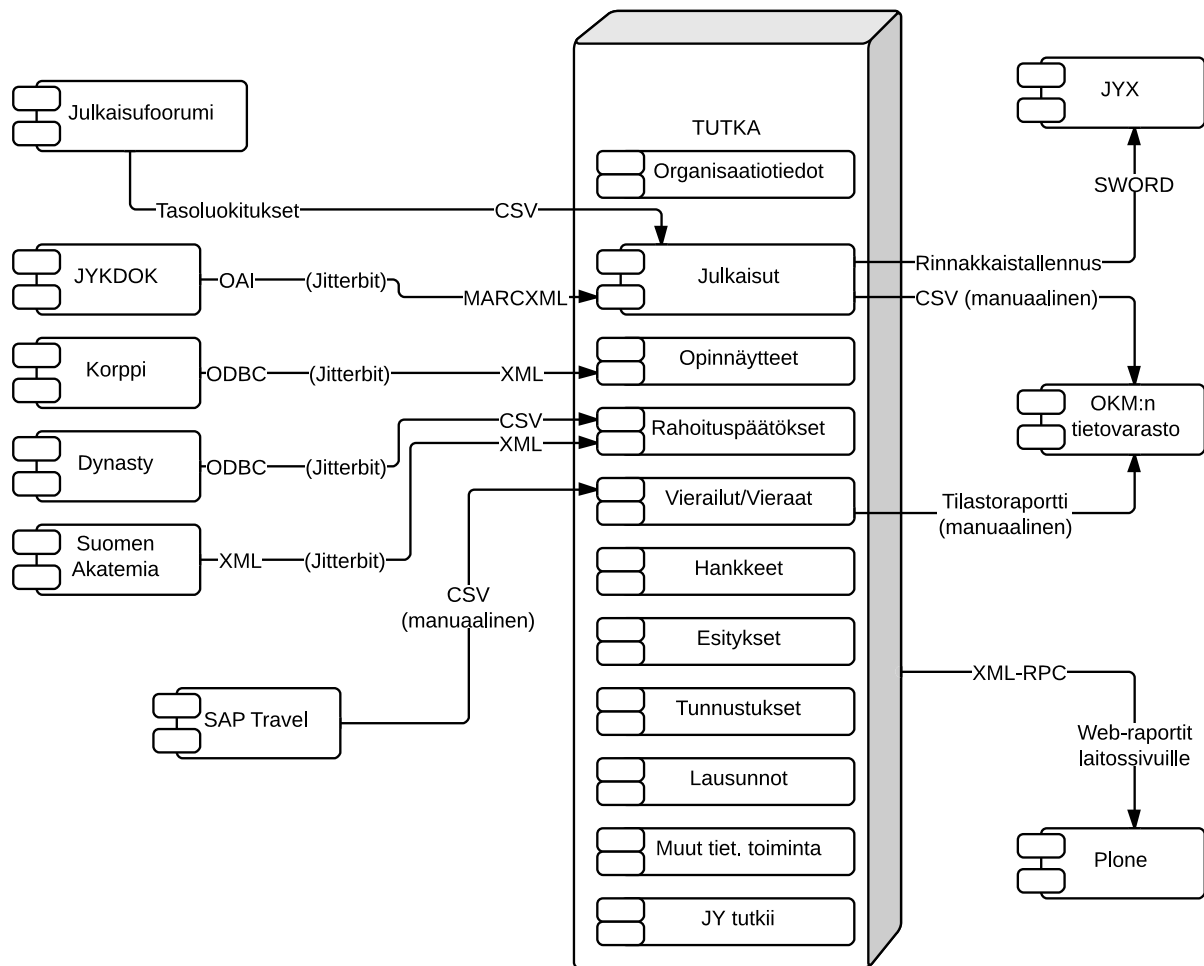


## 5 RAJAPINNAT

Tässä luvussa kuvataan tutkimustietojärjestelmän edellyttämiä rajapintoja nykyisessä ja ryhmän esittämässä järjestelmäympäristössä sekä tunnistettuja kehitystarpeita. Järjestelmän toteutustavasta riippumatta rajapintojen toimivuuden varmistaminen on tärkeää, koska ne mahdollistavat tietojen tuonnin automaattisesti järjestelmään ja toisaalta vaikuttavat tutkimuksen näkyyteen ja raportoinnin sujuvuuteen.

### 5.1 TUTKAn nykyiset rajapinnat

Aluksi kuvataan TUTKAan tällä hetkellä liittyviä tiedonsiirtoja (Kuva 1). Järjestelmään tuodaan tietoja pääosin eräajona kerran kuukaudessa. Samalla ylläpito tarkastaa tietojen oikeellisuutta ja muotoa. Erillisestä käyttäjätunnushallinnasta johtuen tarvittaessa luodaan uusia käyttäjiä, jos mukana on TUTKAlle tuntemattomia (mutta henkilökuntaan tai jatko-opiskelijoihin kuuluvia) henkilöitä. Ulkopuolelle suuntautuvissa tiedonsiirroissa voidaan käyttää enemmän automaatiikkaa. Esimerkiksi JYX-rinnakkaisjulkaisut ja julkaisulistat Plone-sivuille päivittyvät kerran vuorokaudessa. Nykyisessä siirtotekniikassa on tunnistettu ongelmia (viiveet ennen näkymistä käyttäjille, virheet lähtötiedoissa ja manuaalisen selvittelyn osuus), mutta vähintään vastaavat tiedot tulee pystyä siirtämään myös uuteen tutkimustietojärjestelmään.



Kuva 1 Tutkimustietojärjestelmän nykytila; tärkeimmät komponentit ja yhteydet muihin järjestelmiin.

### 5.1.1 Julkaisutietojen siirto JYKDOKista

Kirjaston JYKDOK-järjestelmästä on siirretty vuoden 2010 julkaisuista lähtien Jyväskylän yliopiston julkaisemat julkaisut (pl. lehtiartikkelit ja lehdet). Julkaisut määritellään ISSN- ja ISBN-tunnusten perusteella. Tiedot tuodaan TUTKAan MARCXML-tiedostoina Jitterbit-integrointipalvelimen kautta. TUTKA hakee tarvittaessa myös suoraan yksittäisten julkaisujen metatietoja OAI-haulla.

Tiedonsiirto ei ole täysin automaattista, vaan esim. henkilö- ja laitostietoja täytyy yksikäsittellää ja toisinaan korjailla. Muutokset kirjaston luettelointimäärityksissä ja opetus- ja kulttuuri-ministeriön määrityksissä vaativat sekä yksittäistapausten sisällöllistä tulkintaa että rajapinnan päivittämistä säännöllisesti. Esim. saman organisaation jäsenten kirjoittama tutkimusraportti on tulkittava yhtenä julkaisuna, vaikka se olisi JYKDOKissa luetteloitu osakohteisiin. Erityistä huomiota vaativat toimitetut teokset, koska julkaisuksi määrittelemine edellyttää laajaa johdantoa.

JYKDOK-siirron jälkeen tutkijat täydentävät tarvittaessa julkaisun tietoja. Lisäksi tiedot yhteinäistetään TUTKAan mahdollisesti (julkaisujen osalta ohjeiden vastaisesti) käsin merkittyihin





kopioihin. Kaksoiskappaleiden yhtenäistämiseksi on tarvetta myös käsin kirjattavilla julkaisuilla; erityisesti, jos julkaisujen kirjoittajat ovat eri laitoksilta.

### 5.1.2 Matkatietojen siirto SAP Travelista

SAP Travelista siirretään henkilökunnan ja ulkomaisten vieraiden matkatietoja CSV-tiedostoina. Siirtotiedostojen pohjana on SAP:in oma raportti, mutta tietojen täydentäminen ja tarkastaminen vaatii talouspalveluissa lisätyötä. Lisäksi TUTKAan siirtovaiheessa täytyy huomioida samat yksikäsitteistämistapahtumat kuin julkaisutietoja siirrettäessä. SAP-raportteja on kuitenkin tarkoitus parantaa järjestelmän tulevassa päivityksessä. Käsittelyn ulkopuolelle jäävät mm. apurahatutkijoiden matkat, jotka merkitään käsin. Toisinaan siirrossa ovat myös ongelmana tutkijoiden tarpeettomasti käsin kirjaamat matkat, joiden tiedot voivat olla ristiriidassa Traveliin merkittyjen kanssa. Todennäköinen syy käsinkirjauksille sekä julkaisujen että matkojen osalta on viive, joka voi olla useita kuukausia matkan tai julkaisun ilmestymisen jälkeen. Lisäksi joissakin tapauksissa virhekirjauksista johtuen tietue ei ilmesty ollenkaan siirtotiedostoon ennen kuin virhe huomataan (yleensä tutkijan itsensä toimesta).

### 5.1.3 Ohjaustietojen siirto Korpista

Ohjaustiedot siirretään Korppi-järjestelmästä Jitterbit-integrointipalvelimen muodostamassa XML-muodossa, joka huomioi myös muutokset siirtojen välillä (esim. Korpin päässä poistetut ohjaukset poistuvat myös TUTKAsta). Siirto toimii pääsääntöisesti hyvin, mutta lähtödatassa on tiettyjä eroja verrattuna TUTKAN tapaan tulkita opinnäytetietoja, mikä vaatii muunnoksia yksittäisiin kenttiin. Tulkintaa vaativat mm. oppiaineiden nimeäminen ja Jyväskylän yliopiston ulkopuolisten ohjaajien tunnistaminen.

### 5.1.4 Rahoituspäätösten siirto Suomen Akatemian järjestelmästä

Suomen Akatemia lähettää XML-muotoisia rahoitustietoja IT-palveluiden verkkolevylle, mistä Jitterbit-integrointipalvelin poimii ne TUTKAan. Yksittäisiä teknisiä katkoksia lukuun ottamatta siirto on toiminut pääsääntöisesti hyvin.

### 5.1.5 Rahoituspäätösten siirto Dynasty-asianhallintajärjestelmästä

Kirjaamon Dynasty-järjestelmän dokumenttienhallinnasta siirretään yliopiston kirjanpidon kautta kulkevien rahoituspäätösten tietoja TUTKAan, mikäli Dynastyssa on rahoituspäätöksen lisäksi myös projektinavauslomake (poikkeuksena Suomen Akatemian hankkeet, jotka siirretään omasta järjestelmästä). Tieto vastuullisesta tutkijasta saadaan, mikäli Dynastyyn on vietty myös rahoitushakemus. Rahoitustiedot siirretään Dynasty-järjestelmästä Jitterbit-integrointipalvelimen muodostamista CSV-tiedostoista, jotka huomioivat myös muutokset siirtojen välillä.

Rahoituspäätösten tiedonsiirto ei useista rajapinnan korjauksista huolimatta ole toiminut tyydyttävällä tavalla, koska Dynastyn oma käyttöliittymä ei tue tietojen kirjaamista riittävän täsmällisesti (esim. dokumenttien oikea liittämistapa ja rahoittajien kontrolloitu nimeäminen ovat käyttäjille työläitä ja vaativat tarkkaa suositeltujen käytäntöjen seuraamista). Lisäksi Dynastyn TUTKAa varten tarjoamassa rajapinnassa on vuoden 2012 Dynasty-päivityksen jälkeen ollut pysyvästi teknisiä ongelmia yksittäisten dokumenttien siirtymisessä. Nykyisellä siirtokäytännöllä ei ole myöskään saavutettu tavoitetta hakemusten järjestelmällistä keräämistä (erityisesti



kielteisten päätösten osalta), vaan lukuisat päätökset siirtyvät joko ilman hakemuksia tai epäyhtenäisesti hakemuksiin merkityillä rahoittajanimillä.

Rahoituspäätösten ja merkinnän erityisongelmana on myös tietojen päällekkäisyys monella tasolla. Hämmennystä aiheuttaa TUTKAN rahoituspäätösosion osittainen päällekkäisyys manuaalisesti merkittävien hankeosion tietojen kanssa. TUTKAN hanke- ja rahoituspäätöstietoja ei ole kuitenkaan linkitetty toisiinsa. Toisaalta myös Dynasty-järjestelmän tiedot ovat osittain päällekkäisiä SAP Talous -järjestelmän kanssa. Uudessa tutkimustietojärjestelmässä hanketietojen kerääminen on uudistettava.

#### **5.1.6 Tiedonsiirto Excel (CSV) -tiedostoista**

TUTKassa on kokeellinen Excel (CSV) -tiedostojen tuontitoiminto osalle järjestelmän osioista. Tähän mennessä tuontitoimintoa on käytetty joidenkin fysiikan laitoksen tietojen siirtoon sekä ylläpidon tekemiin massakorjauksiin. Excel-tiedonsiirto on yleistettävissä myös muiden laitosten käyttöön, mutta edellyttää tietojen muodon ja oikeellisuuden tarkistustoimintojen kehitystä edelleen.

#### **5.1.7 Julkaisufoorumin tietojen haku**

Tieteellisten seurain valtuuskunta tarjoaa julkaisuforumiluokat CSV-eroteltuina tiedostoina, joiden pohjalta Julkaisufoorumi-luokka voidaan päätellä paikallisesti julkaisuun kirjattujen tietojen perusteella. Julkaisufoorumi-luokka on laskettu TUTKassa vuoden 2012 julkaisuille syötetyistä ja kirjoituskäytännöiltään yhtenäistetyistä tiedoista. Jatkossa TUTKAan lisätään CSV-tietojen säännöllinen päivitys ja mahdollisuus valita TUTKAN omassa käyttöliittymässä lehti, sarja, konferenssin nimi tai julkaisija Julkaisufoorumissa listatuista tiedoista.

Julkaisufoorumi-luokan tarkassa laskennassa eri julkaisutypeille on vielä jonkin verran epäselvyyksiä ja tulkinnanvaraa, jotka saattavat vielä muuttaa toteutusta. Lisäksi luokkia tullaan uudelleenarviointien seurauksena päivittämään, ja näiden päivitysten suhdetta opetus- ja kulttuuriministeriön raportoinnissa tietyn vuoden julkaisuille vahvistettuihin luokkiin on vielä tarkennettava. Olisi toivottavaa, jos Julkaisuforumiluokituksen pystyisi jatkossa laskemaan yhteisesti käytössä olevan palvelun kautta tietyn vuoden määritysten mukaisesti annetuilla julkaisun metatiedoilla.

#### **5.1.8 Julkaisujen rinnakkaistallentaminen JYXiin**

JYX ja TUTKA muodostavat yhdessä rinnakkaistallentamisen mahdollistavan kokonaisuuden. TUTKassa on toiminnallisuus, jonka avulla mihin tahansa artikkeliin voidaan lisätä tiedostoja. TUTKA-palvelimella oleva skripti julkaisee artikkelin metatiedot METS-paketoinnilla ja tiedostot SWORD-rajapinnan kautta JYXiin ajastetusti. Viestintä TUTKAan tapahtuu XML-RPC-rajapinnan kautta. JYX lähettää julkaisun yhteydessä rinnakkaisjulkaisun tunnisteiden TUTKAan, jonka jälkeen julkaisun tietoihin ja Jyväskylän yliopisto tutkii -sivustolle lisätään linkki avoimeen julkaisuun.

Rinnakkaisjulkaisuprosessissa on jonkin verran kehitystarvetta, koska artikkelin siirron jälkeen mahdollisesti TUTKassa muuttuneet julkaisun metatiedot eivät päivitty JYXiin automaattisesti. Lupamenettelyt hidastavat prosessia jonkin verran. Nykyisillä suosituksilla jokaiselta julkaisun kirjoittajalta tarvitaan lupa rinnakkaisjulkaisuun, ja toisaalta osan kirjaston tekemästä lupaselvittelystä (=kustantajien linjausten esitarkastus SHERPA/ROMEO -tietokannasta) pystyisi au-



tomatisoimaan. Myöskään yleinen tietoisuus avoimesta julkaisemisesta ei ole vielä toivottavalla tasolla, vaikka tähän on yliopistotasoisesti kehoitettu vuonna 2010 (Toimenpiteet rinnakkaisjulkaisemisen edistämiseksi Jyväskylän yliopistossa).

### **5.1.9 Tiedonsiirto yliopiston www-sivuille**

Jyväskylän yliopiston Plone-sivustoille on toteutettu sisältötyyppi (Tutkasivu), joka hakee TUTKAsta yleisimpiä raportteja (esim. tutkijan julkaisut ja laitoksen ohjaukset) kaikista osioista. Siirto on toteutettu XML-RPC-kutsuilla ja ajetaan kerran vuorokaudessa. Siirtomekanismi toimii automaattisesti, mutta raporttien muokattavuus on TUTKAN omiin raportteihin verrattuna rajoitettua eivätkä useimmat tutkijat pääse niitä tekemään itse (tämä edellyttää oikeuksia laitossivujen päivitykseen). Suunnitteilla on kuitenkin TUTKA-tietojen integrointi ns. Roster-henkilökuntasivujen osaksi, mikä toimisi vakio pohjana laitosten henkilökuntalistoilta.

### **5.1.10 Julkaisujen ja vierailujen raportointi opetus- ja kulttuuriministeriölle**

TUTKAN keskeinen käyttötarkoitus on julkaisu- ja matkatietojen vuosittainen raportointi opetus- ja kulttuuriministeriölle. Tiedot vaikuttavat yliopistojen saamaan rahoitukseen. Aiemmin julkaisutiedot vietiin KOTA-tietokantaan (mistä syystä ministeriön tiedonkeruusta käytetään edelleen toisinaan nimitystä KOTA-raportointi) ja nykyisin ministeriön tietovarastoon. Raportoidut tiedot tulevat julkisesti selattavaksi lukumäärätietojen osalta CSC:n Vipunen-raportointiportaaliin ja viitetietojen osalta kansalliseen julkaisuportaaliin. CSC on yhdistänyt eri yliopistojen raportoidut tiedostot, tunnistanut yliopistojen väliset yhteisjulkaisut sekä julkaisukanavien julkaisufoorumiluokat ennen tietojen ajoa julkaisuportaaliin.

Julkaisujen osalta tiedot raportoidaan muotoiluvaatimuksiltaan tarkasti määritellyssä CSV-tiedostossa ja matkatiedot raportoidaan summatietoina (vähintään 5 työpäivää kestävä opetus- ja tutkimusvierailut). TUTKA tuottaa sellaisenaan vaadittavat raportit, mutta automaattista siirtoyhteyttä ministeriön tietovarastoon tai mekanismia ministeriön tai CSC:n tekemien muutosten palauttamisesta TUTKAan ei vielä ole. Julkaisuraportin generoinnin yhteydessä TUTKA tallentaa tiedon julkaisuista, jotka on raportoitu asetettujen aikarajojen puitteissa. Näin voidaan erottaa raportoinnin ulkopuolelle jääneet myöhässä lisätyt julkaisut, sekä julkaisut, joita ei hyväksytty raportointiin mukaan. Vastaava tietuekohtainen vuosikenttä on suunnitteilla myös matkatiedoille.

Opetus- ja kulttuuriministeriön raportoinnissa pakollisiksi vaaditut kentät vaikuttavat suoraan TUTKAan kirjattavien julkaisujen pakollisiin tietoihin. Määritykset ovat muuttuneet jonkin verran joka vuosi (joskus lähellä raportointivuoden loppua, jolloin suurin osa julkaisutiedoista on jo kirjattu), mistä johtuen ylläpito on joutunut tekemään julkaisutietojen kirjaukseen ja käsittelyyn liittyvää räätälöintiä ja tietojen korjausta jälkikäteen - joskus hyvinkin nopealla aikataululla. Kansallisen tutkimustietomallin kehityksestä huolimatta odotettavissa on, että tarvetta nopeille muutoksille lomakkeissa ja raporteissa on myös uudessa tutkimustietojärjestelmässä. Tulevaisuudessa myös julkaisu- ja vierailutiedot siirretään tietovarastoon, josta ne raportoidaan viranomaistietovarastoon (VIRTA).

### **5.1.11 Erilliset koodistot**

TUTKAan on tietojen kirjauksen tueksi integroitu useita listoja ja koodistoja, joista osa perustuu standardeihin (Tilastokeskuksen luokitukset kielet 2003, tieteenala 2010 sekä valtiot ja maat



2007; Jyväskylän yliopiston SAP-laitoskoodit, Jyväskylän yliopiston SAP-maakoodit ja opetus-hallinnon koulutuslaluokitus). Nämä ovat pääosin saatavilla koneellisesti luettavassa muodossa, mutta muutostilanteet saattavat vaatia erillistä kehitystä - erityisesti, jos luokitusta vaihdetaan eri vuosien välillä (esim. siirtymä tieteenala 2007 -luokitukselta tieteenala 2010 -luokitukselle vuoden 2011 julkaisuissa). Myös Jyväskylän yliopiston sisällä käytetyt nimet saatavat poiketa kansallisissa koodistoissa käytetyistä (esim. laitosten ja oppiaineiden yhdistämisen koulutuslaluokitukseseen).

## **5.2 Suunnitteilla olevia integrointeja**

TUTKAan (tai mahdolliseen uuteen tutkimustietojärjestelmään) on suunniteltu seuraavia integrointeja. Näiden toteutusaika riippuu käytettävissä olevista kehitysresursseista. Tutkijat ovat toivoneet erityisesti julkaisutietojen automaattisen poiminnan kehittämistä. IT-palveluissa erityinen paino on kertakirjautumisen ulottamisessa mahdollisimman moneen järjestelmään IDM:n myötä. Lisäksi on tärkeää osittaa laajoja järjestelmiä helpommin ylläpidettäviin ja vain omasta vastualueestaan huolehtiviin palveluihin, jotka viestivät keskenään käyttäjälle läpinäkyvästi yhteisen viestiväylän kautta.

### **5.2.1 Yhteys Jyväskylän yliopiston yhteiseen henkilötietojen hallintaan ja henkilötietojen hakuun**

IT-palveluiden ylläpitämissä järjestelmissä otetaan vuoden 2013 aikana käyttöön keskitetty identiteetin- ja pääsynhallinta (IDM, IAM). Tämä tulee mahdollistamaan myös TUTKAssa kirjautumisen yliopiston tunnuksilla ja suojatun HTTPS-yhteyden Access Gateway -välikerroksen avulla. Lisäksi laitospääkäyttäjien työ määrä uusien käyttäjien lisäämisessä vähenee huomattavasti: kun uusi (todennetusti henkilökuntaan tai jatko-opiskelijoihin) kuuluva henkilö kirjautuu TUTKAan, hänen tietonsa voidaan lisätä automaattisesti IDM-tiedoissa määritellyn laitoksen alle.

IDM ei kuitenkaan sellaisenaan ratkaise kaikkia TUTKAN erityisvaatimuksia käyttäjähallinnan suhteen: henkilökunnan ulkopuolisilla tutkijoilla ei välttämättä ole lainkaan yliopiston tunnusta. Lisäksi IDM säilyttää vain aktiiviset tunnukset ja tietojen tarkastus henkilökunnasta mahdollisesti jo poistuneista henkilöistä on hoidettava erikseen. Vanhojen henkilötietojen osalta IT-palveluissa on suunnitteilla LDAP-tietoja hyödyntävä erillinen palvelu henkilötietojen hakuun. Palveluun kytkeydytään IT-palvelujen viestiväylän kautta AMQP-protokollaa (Advanced Message Queuing Protocol) käyttäen.

IDM:n ulkopuolisten tutkijoiden tunnuksia hallinnoidaan jatkossakin TUTKAN sisäisillä tunnuksilla, joita voidaan käyttää yliopiston tunnusten rinnalla. Henkilökunnasta poistuneiden tutkijoiden tiedot on edelleen poistettava TUTKAsta manuaalisesti ja viivästetysti, koska esim. lehtiartikkelin hyväksyntäprosessi voi kestää yli vuoden, mutta julkaisun tietoihin kirjoittaja on merkittävä artikkelin kirjoitustilanteen mukaan henkilökunnan jäsenenä. TUTKAN sisällä voidaan kuitenkin kehittää automaattisia ilmoituksia laitospääkäyttäjille tutkijoista, joiden yliopiston tunnus on vanhentunut eikä aktiviteettia TUTKAassa ole ollut useampaan vuoteen.

### **5.2.2 Julkaisutietojen poiminta viitetietokannoista (Scopus, Web of Science)**

Uuden tutkimustietojärjestelmän keskeinen tavoite on vähentää työkuormaa julkaisutietojen kirjauksessa. Tärkeimpiä laajoja viitetietokantoja ovat Scopus ja Web of Science, jotka mm. tarjoavat mahdollisuuden julkaisutietojen hakuun tutkijan tai organisaation mukaan koneellisesti



käsiteltävässä muodossa. Poiminnan hyödyt vaihtelevat kuitenkin eri tieteenaloilla merkittävästi. Tietokannoissa olevat julkaisut ovat enimmäkseen lehtiartikkeleita, ja esim. luonnontiede on edustettuna yhteiskuntatieteitä laajemmin.

Nykyisenä esteenä julkaisutietojen poiminnan toteuttamiselle ovat olleet vähäiset kehitysresurssit. Toiminto edellyttää tutkimustietojärjestelmän räätälöintiä jokaista käyttöön otettavaa kirjastoa kohti (esim. tutkijoille tai vastuuhenkilöille suunnattu hakukäyttöliittymä, joka näyttää mahdollisesti lisättävissä olevat tiedot ja vertaa näitä jo kirjattuihin julkaisuihin – täysautomaattiseksi poimintaa ei todennäköisesti ole mahdollista tehdä), eikä importoinnin jälkeenkään julkaisua voida tyypillisesti hyväksyä sellaisenaan. Kuten JYKDOK-siirron yhteydessä, useita kenttiä joudutaan täydentämään (mikään viitetietokanta ei sellaisenaan kata kaikkia raportoinnissa vaadittuja pakollisia kenttiä) ja korjaamaan jälkikäteen (esim. yksittäisten kenttien kirjoitusasun osalta). Käytännössä poiminnan toteutus edellyttää muutaman kuukauden täysipäiväistä työpanosta.

### 5.2.3 Tiedonsiirto tietovarastoon

IT-palvelut on kehittänyt Tietovarastoa raportoinnin tueksi ja kokoamaan eri järjestelmien tietoa yhtenäiseen, pysyvään muotoon. TUTKasta on tarkoitus siirtää tietovarastoon vähintään opetus- ja kulttuuriministeriölle ilmoitetut julkaisut ja matkat vuosittain. Todennäköisenä formaattina ovat KOTA-raportin tyylliset CSV-tiedostot, joita siirretään Jitterbit-integrointipalvelimen kautta tietovarastoon. Tietojen yhdistäminen muista järjestelmistä tullessiin tietoihin kuitenkin edellyttää, että henkilö-, laitos-, ja oppiainetiedot kuvataan yhtenäisellä tavalla. Henkilötietojen osalta IDM tulee auttamaan yhtenäistämisessä, mutta Korpin ja TUTKAN esitystavat laitoksille ja oppiaineille poikkeavat toisistaan jonkin verran.

### 5.2.4 Julkaisutietojen siirto JYXistä

JYKDOK ei kata artikkelitasolla Jyväskylän yliopiston julkaisemia lehtiä, mutta useimpien yliopiston Open Access -lehtien artikkelit tallennetaan JYX-julkaisuarkistoon. IT-palvelut kehittää tekniikkaa, jolla lehtiartikkelien siirtoa JYXiin voidaan helpottaa entisestään Plone-pohjaisella käyttöliittymällä. JYXiin tallennettujen yliopiston henkilökunnan kirjoittamien lehtiartikkelien myötä on ilmennyt tarvetta kehittää tiedonsiirtoa myös JYXistä TUTKAan päin. Siirtoa ei ole mahdollista tehdä kokonaan automaattisesti, koska JYX ei sisällä suoraan tietoa artikkelien kirjoittajien affiliaatioista. Rajapinnan toteutusta priorisoidessa on huomattava, että integraation vaativuus tulee olemaan lähes samaa luokkaa kuin laajemmilla viitetietokannoilla, mutta potentiaalisesti poimittavissa olevien julkaisujen määrä on oleellisesti näitä pienempi.

### 5.2.5 Rahoituspäätösten siirto Tekesistä

Tekesin kanssa on neuvoteltu alustavasti rahoituspäätöstietojen koneellisesta siirrosta samaan tapaan kuin Suomen Akatemian päätöksille on tehty. Tarkempaa siirtotapaa tai aikataulua ei kuitenkaan ole vielä määritelty.

### 5.2.6 Jyväskylän yliopisto tutkii -asiantuntijarekisterin toiminnot

TUTKAN osana on alusta alkaen ollut Jyväskylän yliopisto tutkii -asiantuntijarekisteri. Rekisterissä näytettävät tiedot ovat vain osajoukko TUTKassa näytettävistä tiedoista (julkaisut ja hankkeet), ja tutkijoiden kannalta olisi toivottavaa, että sivulla näkyviä osia olisi mahdollista valita joustavammin. Teknisesti sivusto on suurimmaksi osaksi erotettu muusta TUTKasta,



mikä on mahdollistanut sivuston avaamisen hakukoneille vuoden 2012 lopussa. Tarvittaessa sivusto olisi mahdollista myös siirtää kokonaan toiselle palvelimelle.

Jyväskylän yliopisto tutkii -sivuston tarpeellisuutta on arvioitava uudessa tutkimustietojärjestelmässä, koska sivusto voisi sisältää myös selkeämmin CV-tyylistä tietoa (jolloin se kuuluisi ennemmin CV-järjestelmään). Toisaalta Plone-pohjaisille Roster-henkilökuntasivuille olisi mahdollista sisällyttää Jyväskylän yliopisto tutkii -sivustoa vastaavat tiedot (mikä kuitenkin edellyttäisi, että henkilökuntasivut olisivat järjestelmällisesti käytössä koko yliopiston laajuisesti). Toisaalta tutkimustietojärjestelmän yleisenä vaatimuksena on olla ”näyteikkunana” yliopistossa tehtävälle tutkimukselle, mikä puoltaisi järjestelmän tietojen avaamista ja esim. viimeimpien julkaisujen näyttämistä. Hakukonenäkyvyydessä tavoitetasoa voisi verrata esim. Helsingin yliopiston TUHAT-järjestelmään, jossa kaikki keskeiset tiedot yksittäisistä tutkijoista tutkimusryhmiin ja julkaisuihin löytyvät suoraan hakukoneilla.

### 5.2.7 Tutkimuksen näkyvyys

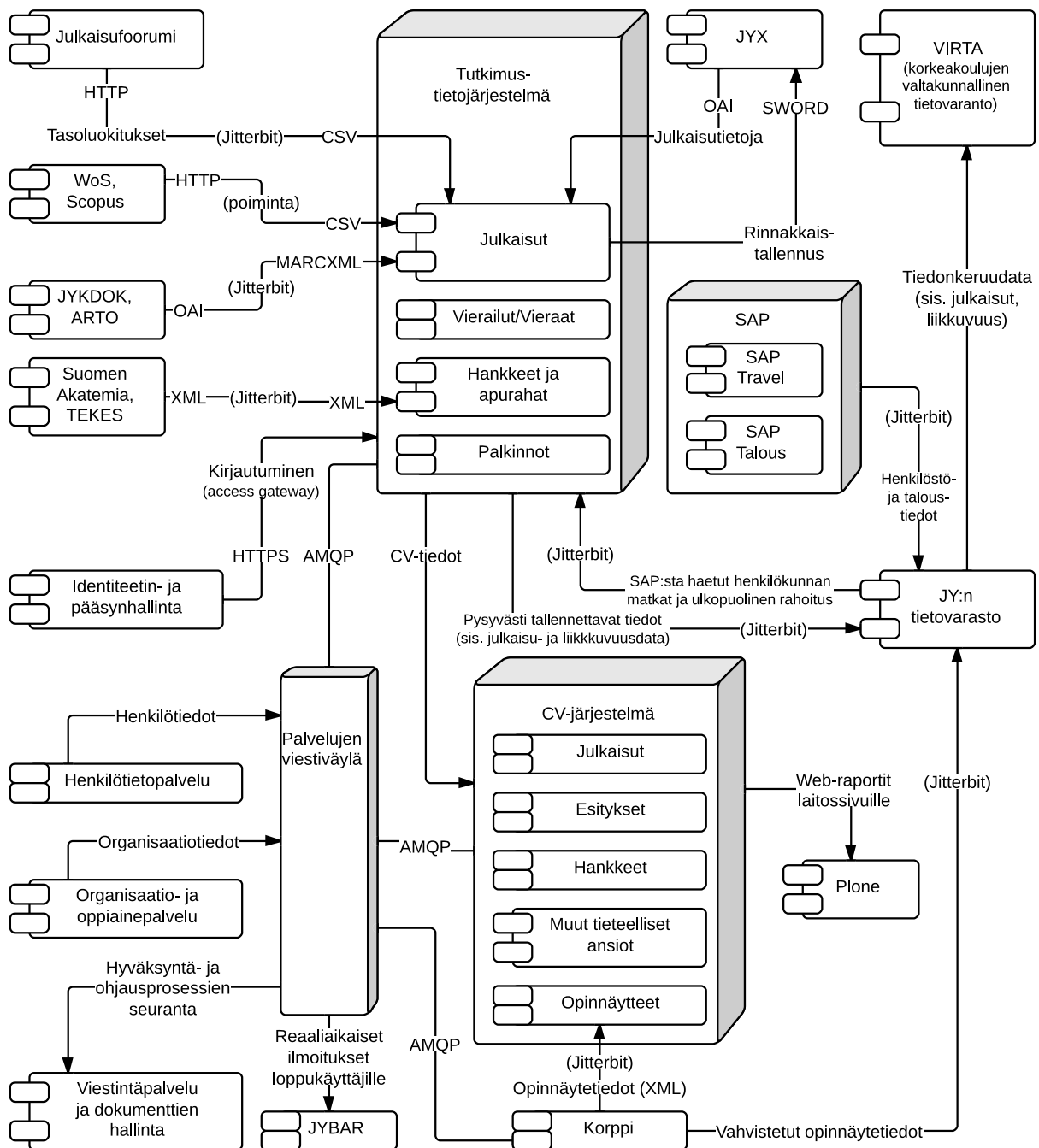
Tutkimuksen ulkoinen näkyvyys on tärkeää ja mm. lyhyet tiivistelmät rahoitetuista hankkeista ovat tämän tarpeen kannalta hyödyllisiä (mihin hankeilmoituksella pyritään). Vastaavasti voi esitellä hankkeita myös näihin liittyvien henkilöiden ja julkaisujen osalta. Henkilötietojen osalta tämä ei kuitenkaan ole yksiselitteistä, koska kaikkien hankkeisiin liittyvien henkilöiden ylläpito tutkimustietojärjestelmän sisällä aiheuttaisi kohtuuttomasti ylläpitoa, eikä edes kaikissa julkaisuissa ole aina selvää, minkä hankkeen nimissä eri tutkijat ovat osallistuneet julkaisun laatimiseen. Tutkimuksen näkyvyyttä voisi edistää joko Jyväskylän yliopisto tutkii -sivustoa uudistamalla, Plone-raportteja kehittämällä (esim. Jyväskylän yliopisto tutkii -sivuston voisi korvata vastaavat tiedot sisältävällä Plone-pohjaisella ”tutkimusportaalilla”) tai erillisellä CV-järjestelmällä.

Nykyiset TUTKAN tiedoista kootut Plone-raportit ovat riittävät laitossivuja ajatellen, mutta tutkijoiden henkilökohtaisen työn kokonaisvaltaisempaan esittelyyn ne ovat liian suppeita. Lisäksi tutkijoilla voi olla tarvetta esim. julkaisulistojen muodostukseen Plone-sivujen ulkopuolella tai kokonaan muussa formaatissa (esim. CSV, BibTeX) yliopiston ulkopuolisiin palveluihin (esim. tutkijoiden yhteisöpalvelut) vientiä varten.

### 5.3 Uudessa järjestelmässä huomioitavia

Seuraavat integrointi- ja tiedonsiirtotarpeet riippuvat keskeisesti siitä, mitä nykyisen TUTKAN osioita otetaan mukaan uuteen tutkimusjärjestelmään. Käytettävissä olevat palvelut riippuvat myös CV-, hanke-, ja ohjaustietojen tallentamiseen liittyvistä päätöksistä sekä kansallisten hankkeiden etenemisestä. Yliopiston omien järjestelmien välisen integroinnin tarkka toteutustapa riippuu IT-palveluiden kehitteillä olevasta viestipohjaisesta palveluarkkitehtuurista. Kuvauksissa on siis merkittävästi epävarmuutta verrattuna jo suunniteltuihin integrointeihin.

Kuva 2 on yleiskuvaus tutkimustietojärjestelmän tavoitetilasta olettaen, että CV:t ja opinnäyte-tiedot on hallittu omissa järjestelmissään, ja että useammassa järjestelmässä käytettyjä tietolähteitä on eriytetty omiin palveluihinsa. Toisaalta kuvassa on oletettu, että hanketoiminnan seuranta on sisällytetty (dokumenttien tallennusta lukuun ottamatta) tutkimustietojärjestelmän osaksi – käytännössä korvaten nykyisen Dynastyn roolin hankehallinnossa.



Kuva 2. Tutkimustietojärjestelmän tavoitetilän järjestelmäympäristö ja keskeisimmät tiedonsiirtorajapinnat.

### 5.3.1 Raportointirajapinta

Nykyinen TUTKA sisältää runsaasti sekä julkisia että ylläpidon erikseen räätälöimiä raportteja laitosten ja yliopiston hallinnon tarpeisiin. Tulevassa tutkimustietojärjestelmässä järjestelmän itse tarjoamien ja tietovarastosta saatavien raporttien (sekä mahdollisen CV-järjestelmän raporttien) suhdetta tulee selkeyttää niin, että tiettyä raporttimuotoa varten olisi yksi, ensisijainen tietolähde. Jos tietovaraston roolia julkisesti selattavien raporttien osalta päädytään kasvattamaan, näiden käyttöön on rakennettava erillinen www-palvelu, esim. nykyisten TUTKAN Plone-raporttien tapaan. Toisaalta nykyisen TUTKAN YPJ-raporttia vastaavalle palvelulle on mää-



riteltävä paikka uudessa järjestelmäkokonaisuudessa, koska tiedot joudutaan jatkossa kokoamaan useista eri järjestelmistä.

### 5.3.2 Laitostietojen ja oppiaineiden keskitetty kuvaus

Laitos- ja muut organisaatiotiedot sekä näihin liittyvät oppiaineet ovat useimmille yliopiston järjestelmille periaatteessa yhteistä tietoa, joka on kuitenkin tyypillisesti kuvattu jokaisessa järjestelmässä erikseen. Tämä aiheuttaa ongelmia erityisesti tilanteissa, jossa organisaatorakenne muuttuu. Laitos- ja oppiainetiedot (sisältäen historiatiedon aiemmista laitoksista) ovat TUTKAssa keskeistä tietoa, koska opetus- ja kulttuuriministeriön raportoinnissa koulutusala määritellään laitoksen tai monitieteisissä yksiköissä oppiaineen perusteella. Lisäksi useita raportteja tarkastellaan yliopiston sisällä laitosten tai tiedekuntien tasolla.

IT-palveluiden viestipohjainen palveluarkkitehtuuri mahdollistaisi erillisen palvelun, joka vastaisi laitos- ja oppiainetietojen hallinnasta ja tarjoaisi tämän tiedon edelleen muille järjestelmille (esim. Korppi, ROTI, tutkimustietojärjestelmä ja CV-järjestelmä) AMQP-protokollaa käyttäen. Käytännön ongelmana tällaisen tiedon määrittämisessä on harmonisoida eri järjestelmien tiedot. Esimerkiksi Korppi sisältää organisaatio- ja oppiainetiedot TUTKAa tarkemmassa muodossa (esim. oppiaineiden linjatiedot), mutta toisaalta TUTKA sisältää tiedot ajasta, jolloin laitos on ollut aktiivisena ennen mahdollista yhdistymistä toiseen laitokseen.

### 5.3.3 Hyväksymisprosessien viestinnän seuranta ja dokumenttien hallinta

Uudessa tutkimustietojärjestelmässä hyväksyntäprosessit ulotetaan julkaisuihin muihinkin osiin, mikä parantaa tietojen oikeellisuutta. Samalla varsinaiset hyväksymisprosessit tulee tehdä teknisesti mahdollisimman kevyiksi ja henkilöriippumattomiksi. Käytännössä tämä edellyttää ainakin hyväksymisprosessin aikana lähetettyjen viestien ja mahdollisten liitetiedostojen tallennusta joko tutkimustietojärjestelmään, tai erilliseen dokumenttien hallintajärjestelmään (vastaavalle mekanismille on tarve myös opinnäytteiden ohjausprosessissa). Tämä voidaan toteuttaa teknisesti useilla tavoilla (esim. Request Tracker -tukipyynnöjärjestelmä, Korpin ryhmät, Exchange-rooliosoitteet tai suorat sähköpostipalvelimella ajettavat skriptit). Dokumentit voisi tallentaa Ploneen erikseen määritellyksi ajaksi. Esim. kurssi- ja oppimateriaaliympäristö Kopan kehityksen pohjalta voidaan ottaa käyttöön palvelu dokumenttien tallennukseen sisältäen käyttöoikeuksien hallinnan. Viestinnässä käytetään AMQP-protokollaa.

Päätösten tueksi lähetettävien liitetiedostojen lisäksi uudessa tutkimustietojärjestelmässä saat-  
taa tulla tarvetta tallentaa pysyvästi yksittäisiä (mahdollisesti luottamuksellisia) dokumentteja, kuten hankeilmoituksiin liittyvät hankesuunnitelmat. Näiden arkistointi voidaan hoitaa muun hyväksyntään liittyvän viestinnän tapaan joko Plone-pohjaisella ratkaisulla, tai lähetyksellä erilliseen pitkäaikaissäilytykseen suunniteltuun järjestelmään, kuten JYX. Julkaisujen rinnakkaistallentamista JYXiin voidaan pitää erikoistapauksena tutkimustietojärjestelmässä käsitellyn dokumentin pysyvästä tallentamisesta.

### 5.3.4 Hanketietojen hallinta ja linkitys julkaisuihin

Hanketietojen hallinnan uudistaminen on keskeinen tutkimustietojärjestelmää sivuava kehitystarve. Tämä voidaan toteuttaa joko hanke- ja rahoituspäätöstietojen sisällyttämisellä tutkimustietojärjestelmään, tai 2-suuntaisella yhteydellä erilliseen hankejärjestelmään. Jälkimmäisessä tapauksessa toimivan rajapinnan määrittelyyn (ja ylipäänsä ulkoisen hankejärjestelmän sopi-





vuuden arviointiin toiminnallisuuden, käytettävyyden ja hankeprosessin tuen kannalta) on kiinnitettävä erityistä huomiota, jotta välttyttäisiin nykyisen Dynasty-käsittelyn ongelmilta. Automatisoitujen siirtojen Suomen Akatemiasta (ja mahdollisesti jatkossa Tekesistä) tulee toimia myös uudessa järjestelmässä. Yliopiston kautta kulkevan rahoituksen osalta tietoja on suotavaa jatkossa kerätä SAP Talous -järjestelmästä.

Mikäli hanketietojen kirjaus jätetään kokonaan varsinaisen tutkimustietojärjestelmän ulkopuolelle, on erillisen hankejärjestelmän silti tuettava mahdollisuutta välittää rahoitustietoja muihin järjestelmiin (kerätty tutkimusrahoitus on YPJ-raportissa tarvittavaa tietoa), sekä mahdollisuuteen yhdistää hanketietoja julkaisutietoihin. EU:n FP7-hankkeissa tuotetut julkaisut tulisi julkaista avoimesti OpenAIRE-suositusten mukaisesti, mikä edellyttää uusia metatietokenttiä julkaisuihin (ja edelleen JYXin julkaisurajapinnan laajentamista) – ja erityisesti julkaisuun liittyvän hankkeen tunnistetta. Julkaisujen linkittäminen hankkeisiin parantaa myös mahdollisuuksia tiedottaa hankkeiden tuotoksista web-raporteissa.

### 5.3.5 Tutkijan CV- ja YPJ-tiedot

Tutkimustietojärjestelmää sivuavana yleisenä vaatimuksena on koostaa tietoja erilaisten ansioluettelomallien pohjaksi. Esimerkki on Tieteellisten seurain valtuuskunnan CV-malli. CV on tutkimustietojärjestelmän lisäarvo, mutta ei ensisijainen vaatimus. YPJ-raportit ovat nykyisessä TUTKAssa olleet keskeisessä asemassa arviointiprosessin tukena, mutta raportin sovittaminen sellaisenaan uuteen tutkimustietojärjestelmään on ongelmallista, koska sen sisältämät tiedot ovat yhdistelmä varsinaisista tutkimustietojärjestelmän tiedoista ja CV-tiedoista (kutsutut suulliset esitelmät) ja opintoasioista (ohjaukset).

Koska uusi tutkimustietojärjestelmä tulee julkaisuja lukuun ottamatta sisältämään tietoja vain yliopiston nimissä tehdystä tutkimuksesta, ja toisaalta osa nykyisen TUTKAN osioista on päätetty rajata tutkimustietojärjestelmän ulkopuolelle, parhaat edellytykset sekä tutkijan CV:n koostamiselle että YPJ-raportin seuraajalle tarjoaisi erillisellä CV-järjestelmällä, johon siirretään tietoja muista relevanteista järjestelmistä. Nykyisessä TUTKAssa näitä ovat esitykset, lausunnot ja muu tieteellinen toiminta. CV-järjestelmän puitteissa lausunnot, tieteellinen toiminta ja (tutkimustietojärjestelmään edelleen sisältyvät) palkinnot ja muut tunnustukset olisi todennäköisesti selkeintä yhdistää saman käsitteen alle (TUTKAssa osioiden sisältämät tiedot on muodoltaan erittäin lähellä toisiaan), esim. "Tieteelliset ansiot".

Teknisiä vaihtoehtoja CV-järjestelmän toteutukseen on useita. Jyväskylän yliopistossa opiskelijoiden käyttöön kehitetty ePortfolio on yksi mahdollisuus tällaisen järjestelmän toteuttamiseksi. ePortfolion Plone-pohjaiset komponentit tarjoavat yliopiston www-arkkitehtuuriin sopivan, kohtuullisen helposti käyttöönotettavan ja huomattavasti nykyistä TUTKAA käyttäjäystävällisemmän teknisen alustan. ePortfolio mahdollistaa julkistettavien tietojen sijoittelun erilaisiin näyteportfolioihin sekä käsin kirjattujen ja muista järjestelmistä tuotujen "varmistettujen" tietojen sisällyttämisen samaan näkymään. Opiskelijan ePortfoliota varten kehitetyistä sisältötyypeistä osa on sellaisenaan sovellettavissa myös henkilökunnan CV:hen (esim. tutkinnot ja työkokemus).

Uudessa tutkimustietojärjestelmässä useista nykyisen TUTKAN osioista luovutaan. Tästä huolimatta jo kirjattujen tietojen säilytyksestä on huolehdittava. Kaikkien TUTKAN kirjattujen tietojen tulisi ainakin nykyisen henkilöstön osalta olla siirrettävissä korvaavaan järjestelmään, esimerkiksi osio kerrallaan XML-RPC-kutsuilla tai siirtotiedostoina. Erityisesti kutsutut suulli-



set esitelmät on syytä siirtää CV-järjestelmään, koska nämä ovat olleet YPJ-raportissa näytettävää tietoa. CV-järjestelmässä on mahdollista kirjata julkaisuiksi myös niitä jotka eivät täytä opetus- ja kulttuuriministeriön julkaisumääritelmää.

### 5.3.6 Ohjaustietojen tallennus yhdellä järjestelmällä

Tutkielmien ohjaukset ovat YPJ:ssä ja laitosten sisäisessä seurannassa keskeisesti tarvittavaa tietoa. Sisältönsä puolesta ne eivät kuitenkaan kuulu tutkimustietojärjestelmään. Yliopiston 19 opetuslaitoksesta kuuden (+Agora Centerin) ohjaustiedot siirretään Korpista automaattisesti TUTKAan. Lisäksi Korppia käytetään jatko-opiskelijaksi ilmoittautumiseen koko yliopistossa ja yksittäisten tutkijoiden osalta myös laitoksilla, joissa TUTKA-siirtoa ei tehdä. Näillä laitoksilla ohjaustiedot dokumentoidaan kuitenkin TUTKAan käsin. Nykyinen kahden kirjausjärjestelmän tilanne altistaa virheille, koska tietoja saattaa kirjautua siirron yhteydessä kahteen kertaan (erityisesti, jos automaattisesti siirrettävässä ohjauksessa on mukana henkilö laitokselta, jossa automaattisiirto ei ole käytössä). Toisaalta tietojen oikeellisuutta ei tarkisteta järjestelmällisesti myöskään Korpin puolella, ja eri laitosten käytännöt ohjausprosesseissa ja niiden merkinnässä vaihtelevat. Joillakin laitoksilla Korpin ohjausosio on jo otettu järjestelmällisesti opintoasioiden kehittämisen välineeksi, mutta toisaalta järjestelmässä on selkeitä kehitystarpeita. Ohjausosion kehittäminen (joko Korpin sisällä tai esim. ohjaajan työpöydän osana) ja sen käytön laajentaminen mahdollistaisi kuitenkin selkeämmin hallitun ohjausprosessin (sekä esim. yhteydet plagioinnin tunnistukseen ja tutkielman JYX-julkaisuun samalla käyttöliittymällä, mikä vähentäisi päällekkäistä työtä edelleen) ja kaikkien ohjaukseen liittyvien tietojen säilytyksen yhdessä paikassa.

Jos ohjausosio poistetaan TUTKAsta, Korppiin tai mahdollisesti tietovarastoon on rakennettava erillinen siirtomekanismi pelkästään TUTKAan kirjattujen ohjaustietojen tuontia varten. Riippumatta järjestelmästä jonne ohjauksia tullaan tallentamaan, tietojen oikeellisuus (samaa tapaan kuin tutkimustietojärjestelmän osioissa) sekä näkyvyys YPJ-raportissa (esim. tuomalla vahvistetut ohjaustiedot CV-järjestelmän osaksi) tulee varmistaa jatkossa.

### 5.3.7 Yhteys sähköiseen työpöytään

IT-palvelut kehittää sähköistä työpöytää, jonka tavoitteena on olla personoitu www-työpöytänäkymä opiskelijan ja henkilökunnan käyttöön. Näkymässä voidaan yhdellä kerralla nähdä omaan toimintaan liittyviä ajankohtaisia tietoja ja siirtyä joustavasti eri ohjelmistojen välillä. Myös tutkimustietojärjestelmä voisi hyödyntää sähköistä työpöytää ilmoitusten lähettämisen (esim. julkaisutietojen hyväksyntä) tai omien tietojen näyttämisen muodossa. Palveluita yhdistävänä käyttöliittymänä toimii JYBAR-navigointipalkki, joka voidaan tuoda eri soveluksiin näistä riippumatta. Sähköisen työpöydän kanssa kommunikoidaan IT-palvelujen viestiväylän kautta AMQP-protokollaa käyttäen.

### 5.3.8 Kotimaisten artikkelien poiminta (ARTO ja metatietovaranto)

ARTO on Kansalliskirjaston ylläpitämä kotimaisten artikkelien viitetietokanta, jonka tiedontuottajina toimivat useat eri kirjastot. Vuonna 2013 käynnistyneessä Kotimaisten tieteellisten artikkelien metatietovaranto -hankkeessa eli "artikkelitietovarantohankkeessa" tiedontuotantopohjaa pyritään laajentamaan. Tavoitteena on myös parantaa artikkelien ajantasaisuutta ja kattavuutta (erityisesti Julkaisuforumissa luokiteltujen lehtien osalta). Uuden tutkimustietojärjes-



telmän tulisi pystyä poimimaan viitetietoja kansainvälisten viitetietokantojen lisäksi myös ARTOsta.

### **5.3.9 Kansallinen tutkijarekisteri ja tutkijatunnisteet**

Tutkijoiden tunnistamisen helpottamiseksi on suunnitteilla kansallinen tutkijarekisteri, jonka vaatimukset on syytä ottaa huomioon uudessa tutkimustietojärjestelmässä. Lisäksi on olemassa erilaisia enemmän tai vähemmän vakiintuneita järjestelmiä tutkijoiden tunnistamiseen (esim. Scopusen ja Thomson Reutersin tutkijatunnisteet sekä ORCID). Kirjastokohtaiset tunnisteet ovat jo nykyisellään hyödyllisiä viitetietojen poiminnan apuna, mutta yleiskäyttöisiksi tunnisteiksi ne eivät sovellu, koska yksikään tunniste ei ole levinnyt riittävän laajalle. Opetus- ja kulttuuriministeriö tulee mahdollisesti edellyttämään jatkossa tutkijatunnisteita julkaisutietojen keruun osana.



## 6 TOTEUTUSVAIHTOEHDOT

Jyväskylän yliopistossa ei ole tällä hetkellä kokonaisvaltaista tutkimustietojärjestelmää. Keskeiset järjestelmät tällä hetkellä ovat TUTKA julkaisutiedon hallinnassa ja Dynasty asiakirjahallinnassa. Nykyinen TUTKA on tullut elinkaarensa päähän, koska käytetty tekniikka ei ole enää kaikilta osin tuettua. Lisäksi uudet tietotarpeet edellyttävät järjestelmän uudistamista. Dynasty ei tue hankehallinnon nykyisiä ja tulevia tarpeita.

Suomalaiset yliopistot ovat uudistamassa tutkimustietojärjestelmiään ja samalla tutkimustiedon käsittelyprosessejaan. Useimmiten pyritään sekä julkaisutiedon että hanketiedon käsitteilyyn samalla järjestelmällä mm. hakemusmenestyksen seuraamiseksi. Yleisenä vaatimuksena on eurooppalaisen CRIS/CERIF-mallin mukainen toteutus. Järjestelmänsä ovat jo uudistaneet Helsingin ja Turun yliopistot sekä Lappeenrannan teknillinen yliopisto. Lisää kilpailutuksia ja hankintoja on tulossa vuoden 2013 aikana. Käytännössä kilpailutus näyttää tapahtuvan tanskalaisen Puren (Helsinki) ja saksalaisen Converiksen (Lappeenranta ja Turku) kesken. Järjestelmien ja raportoinnin tukena käytetään yleensä myös omaa tietovarastoa.

Työryhmä on hahmotellut neljä vaihtoehtoista tutkimustietojärjestelmän toteuttamistapaa: 1) jatkokehitetään nykyistä TUTKAA, 2) uudistetaan TUTKA tutkimustietoihin liittyviltä osiltaan, 3) kehitetään uusi tutkimustietojärjestelmä tai 4) ostetaan tutkimusjärjestelmä tai sen osia ulkopuolelta. Sopivimman vaihtoehdon valinta riippuu myös tutkimustietojärjestelmään liittyvien muiden järjestelmien (esim. CV-tietojen hallinta) kehittämiseen liittyvistä päätöksistä.

Koska työryhmä esittää muutoksia nykyisiin tutkimustietojen kirjaamisprosesseihin, käyttöön-otto edellyttää koulutusta, mutta keskitetyn kirjauksen ansiosta koulutusta tarvitsevien henkilöiden määrä on huomattavasti nykyistä pienempi.

### 6.1 Jatkokehitetään nykyistä TUTKAA

Jatketaan nykyisen TUTKAN kehittämistä aiempaan tyyliin lisäten luvussa 5.21 kuvattuja rajapintoja. Vaihtoehdon etuna on periaatteessa saumaton siirtymä uuteen järjestelmään ja normaalin ylläpidon rinnalla tehtävä kehitys. Toisaalta TUTKAN perusrakenteessa olevien ongelmien (esim. suorituskyky-, ja käytettävyysongelmat) korjaus ja teknisen infrastruktuurin uudistaminen käy sitä hankalammaksi, mitä enemmän uusia ominaisuuksia nykyiseen järjestelmään lisätään ilman, että järjestelmän perustaa uudistetaan. Ylläpitoon tuovat myös rasitetta tutkimustietojärjestelmään kuulumattomien osioiden pitäminen TUTKAssa.

+ Aiempi TUTKAN kehitystyö voidaan hyödyntää ja toimintoprosessit voivat jatkua keskeytymättöminä.

+ Tekninen ympäristö (=Zope-sovelluspalvelin) on periaatteessa sama, mutta huomattavasti vanhempi versio, kuin yliopiston muissa Plone-pohjaisissa www-palveluissa.

+ Kehittäjät tuntevat toimintaympäristön ja prosessit.

- Mitä enemmän TUTKAA nykyisellään kehitetään sitä hankalammaksi ja työläämmäksi sen kehittäminen käy.

- Ylläpitokustannukset kasvavat.



## 6.2 Uudistetaan TUTKA tutkimustietoihin liittyviltä osaltaan

Jos TUTKAA uudistetaan keskeisimpien tutkimustietojärjestelmään liittyvien osioiden osalta, voidaan samalla poistaa vanhan TUTKAN ominaisuuksia, jotka eivät ole keskeisimpiä tutkimustietojärjestelmälle. Samalla uudistetaan TUTKAN teknistä ympäristöä ja tietomallia, jolloin järjestelmän elinkaari, suorituskyky ja ylläpidettävyys kasvavat (luonnollisesti alustavassa uudistamisessa tarvittavan kehitystyön kustannuksella). Esim. yhteys IT-palvelujen viestiväylään tai uusia käyttöliittymäkomponentteja voidaan ottaa käyttöön huomattavasti nykyistä TUTKAA helpommin. Tämän vaihtoehdon edellytyksenä on, että CV- ja hankejärjestelmistä sekä opinnäytteiden ohjaustiedoista tehdään linjaus.

- + Aiempi TUTKAN kehitystyö voidaan hyödyntää ja toimintoprosessit voivat jatkua keskeytymättöminä.
- + Tekninen ympäristö (=Zope-sovelluspalvelin) on sama, mutta uudistamisen jälkeen hieman vanhempi versio, kuin yliopiston muissa Plone-pohjaisissa www-palveluissa.
- + Voidaan poistaa vanhan TUTKAN ominaisuuksia, jotka eivät ole keskeisiä tutkimustietojärjestelmälle. Tämä keventää ylläpitoa.
- + Noudattaa IT-palveluiden näkemystä kokonaisarkkitehtuurista.
- + IT-palveluiden kehittämispalveluiden osaaminen hyödynnettävissä paremmin. Tämä pienentää henkilöriskiä nykyisestä.
- + Kehittäjät tuntevat toimintaympäristön ja prosessit.
- Kehityskustannukset ovat ensimmäistä vaihtoehtoa suuremmat, mutta ylläpitokustannukset pienemmät.
- Vaatii samanaikaisesti useisiin eri järjestelmiin liittyviä päätöksiä, joiden toimeenpano voi viedä aikaa.

## 6.3 Kehitetään uusi tutkimustietojärjestelmä

Uuden tutkimustietojärjestelmän kehittäminen edellyttää luonnollisesti merkittäviä suunnittelu- ja toteutuspanostuksia, mutta toisaalta tekee mahdolliseksi modernien tietoteknisten toteutustapojen hyödyntämisen. Vaihtoehto on suositeltava lähinnä tilanteessa, jossa nykyinen tutkimustietojärjestelmä toimii kokonaisuudessaan vanhentuneessa laitteisto- tai tietokantaarkkitehtuurissa (vrt. JORE ja ROTI). Ellei tämänkaltaisia painavia perusteita kokonaan uudelle järjestelmälle ole, ohjelmistotekniikan tutkimuksessa yleensä suhtaudutaan varauksellisesti "saman" järjestelmän uudelleenkirjoitukseen – varsinkin, jos "vanhaa" järjestelmää kehitetään samanaikaisesti uuden rinnalla. TUTKAN takana on 10-vuotinen kehitystyö ja järjestelmä sisältää merkittävästi "hiljaista" tietoa tietojen kirjaukseen ja siirtoon liittyvästä logiikasta. Tämän tiedon sisällyttäminen uuden järjestelmän vaatimusmäärittelyyn on hankalaa.

- + Parempi käyttöliittymä kuin nykyisessä TUTKAssa.
- + Kehittäjät tuntevat toimintaympäristön ja prosessit.
- + Noudattaa IT-palveluiden näkemystä kokonaisarkkitehtuurista ja mahdollistaa standardien tukemisen.
- + Järjestelmän ylläpito ja kehittäminen on joustavampaa verrattuna kaupalliseen järjestelmään.



- Aiemman TUTKA-koodin hyödyntäminen hankalaa.
- Edellyttää IT-palveluihin lisäresursseja ja erityisasiantuntemusta.
- Nykyistä TUTKAA on ylläpidettävä ja kehitettävä niin kauan kuin uusi järjestelmä on valmis.

#### 6.4 Ostetaan tutkimusjärjestelmä tai sen osia ulkopuolelta

Suomalaisten yliopistojen kilpailutuksissa on noussut esiin kaksi kaupallista järjestelmää, jotka sisältävät muitakin osia kuin tässä raportissa tutkimustietojärjestelmän ydintoiminnoiksi rajatut. Listattujen ominaisuuksiensa puolesta kaupalliset järjestelmät ovat monipuolisia ja sisältävät huomattavan määrän tutkimustietojärjestelmältä vaadittavia osioita. Järjestelmien soveltuminen Jyväskylän yliopiston järjestelmäympäristöön ja yliopiston erityisvaatimuksiin aiheuttaa kuitenkin merkittäviä epävarmuustekijöitä ja edellyttää laajaa testausta. Jos päätetään ostaa uusi tutkimustietojärjestelmä tai sen osio, tulee kilpailutusasiakirjat ja vaatimusmäärittelyt laatia huolella.

- + Kaupalliset järjestelmät ovat laajasti käytössä Euroopan yliopistossa (ja tulossa käyttöön myös suomalaisiin yliopistoihin).
- + Mahdollisuus yhteistyöhön muiden yliopistojen kanssa.
- + Tietomallit ovat yhtenäiset, joten voi olettaa, että kaupalliset järjestelmät pystyvät toteuttamaan Jyväskylän yliopiston tarpeet.
- + Pyrkivät ratkaisemaan yhdessä sovelluksessa merkittävän osan koko tutkimustietojärjestelmästä (hankehallinto, julkaisut ja tutkijan CV).
- + Valmiit julkaisutietojen poimintatoiminnot, todennäköisesti itse toteutettua helppokäyttöisempi käyttöliittymä.
- Aiemmin tehtyä työtä ei voida hyödyntää.
- Rajapintojen toimivuus muihin Jyväskylän yliopiston järjestelmiin epävarmaa ja vaatii räätälöintiä.
- Joustavuus ja reagointi ministeriön vaatimuksiin voivat olla haasteellisia.
- Nykyistä TUTKAA on ylläpidettävä ja kehitettävä niin kauan kuin uusi järjestelmä on valmis.
- Ylläpidon ja jatkokehityksen mahdolliset kustannukset (ja niiden työnjaon haastavuus).
- Riippuvuus ulkopuolisesta toimittajasta.



## 7 LOPPUPÄÄTELMÄT

Jyväskylän yliopisto tarvitsee nykyaikaiset vaatimukset täyttävän ja tulevaisuuden tarpeisiin joustavasti mukautuvan tutkimustietojärjestelmän. Työryhmä on hahmotellut neljä toteuttamisvaihtoehtoa uudeksi tutkimustietojärjestelmäksi. Uuteen järjestelmään siirtyminen kestää joka tapauksessa 2-3 vuotta jona aikana on ylläpidettävä ja kehitettävä nykyistä järjestelmää.

Yliopiston järjestelmiä tulee kehittää niin, että muiden järjestelmien toiminnot ja tietosisällöt huomioidaan. Päämääränä on se, että samaa tietoa ei kirjata useisiin järjestelmiin vaan tieto on siirrettävissä, ja eri järjestelmien kehittämistä ajatellaan kokonaisuutena.

Työryhmällä ei ole tällä hetkellä edellytyksiä arvioida eri vaihtoehtojen kustannuksia, sillä kokonaisuuteen liittyy päätöksiä määrityksistä, joita ei ole vielä tehty. Lisäksi kaupallisten järjestelmien soveltuvuuden arviointi edellyttää tätä raporttia laajempaa selvitystä. Työryhmä suosittelee, että tehdään kustannuslaskenta ja resurssiarvio vaihtoehtoista 2-4. Tähän liittyvät päätökset siitä, miten toteutetaan mm. tutkijan CV, tietojen kerääminen YPJ-keskusteluja varten sekä opinnäytteiden ohjausten kirjaus. (Toistaiseksi YVV-barometrin vaatimukset eivät ole vaikiintuneet kylliksi, jotta niihin voitaisiin ottaa kantaa.)

Tämä dokumentti ei riitä sellaisenaan kilpailutuksen vaatimusmäärittelyiksi. Jos vaihtoehto 4 valitaan, voidaan hyödyntää muiden yliopistojen tekemiä vaatimusmäärittelyjä ja mahdollisesti kilpailuttaa yhdessä.