



KUNTIEN JA HUS:N ASIAKAS- JA POTILASTIETOJÄRJESTELMÄN HANKINTA

Tuotevertailun B pisteytysmalli

Versio: 1.0

20.10.2014



VERSIONHISTORIA:

Päivä	Versio	Sisältö	Tekijä
20.10.2014	1.0	<i>Tarjoajille lähetettävä versio</i>	<i>Apotti-hanketoimisto</i>

Sisällysluettelo

Yleistä	4
1 Painotukset	4
2 Vähimmäisvaatimukset	4
3 Osa-alueittainen pisteiden muodostus	4
3.1 Toiminnallinen laajuus ja laatu: Ryhmäarvioinnit	5
3.1.1 Toiminnallinen laajuus ja laatu: Ryhmäarvioinnit kotihoidon ja sosiaalityön sekä suun terveydenhuollon aihealueilla	8
3.2 Toiminnallinen laajuus ja laatu: Valmis sisältö	8
3.3 Käytettävyys	9
3.3.1 Tuloksellisuus	11
3.3.2 Virheettömyys	12
3.3.3 Tehokkuus ja opittavuus	14
3.3.4 Tyytyväisyys	16
3.3.5 Käyttöliittymäsuunnittelun onnistuneisuus	20
3.3.6 Esteettömyys	21

Liite 1: Käytettävyysarviointi: Tilastollisten merkitsevyyksien tarkastelu

Liite 2: Esimerkkilaskelma: Käytettävyyden kokonaisuuskohtaiset pisteet ja kokonaisarvion muodostuminen, tarjoaja X

Yleistä

Tässä dokumentissa kuvataan Tuotevertailu B:n pisteytysmalli ja sisäiset painoarvot.

1 Painotukset

Tuotevertailu B:ssä arvioinnin kohteina olevia osa-alueita on kaksi: toiminnallinen laajuus ja laatu sekä käytettävyys *Tuotevertailun B kuvaus ja toteutussuunnitelma –dokumentin luvussa 5* kuvatun mukaisesti. Kummallekin osa-alueelle on määritetty osa-aluekohtainen painotus, joka on kuvattu tarkennetun tarjouspyynnön liitteessä 5. Osa-alueista toiminnallinen laajuus ja laatu jakautuu edelleen toiminnallisten kokonaisuuksien toimivuuden ja soveltuvuuden (*jäljempänä ryhmäarvioinnit*) sekä tuotteen valmisisältöjen arviointiin.

Osa-aluekohtaiset painoarvot jakautuvat edelleen aihe-alueittain seuraavasti:

Tuotevertailun painoarvotaulukko												YHT	sisäinen paino
Ratkaisun toiminnallinen laajuus ja laatu: ryhmäarvioinnit													
Aihealue	Päivystystoiminta	Tehonäkymät	Nestehoito ja tehohoito	Leikkaus-sali	Avo-vastaanotto	Hoitotyö	Vuodeosasto	Lastensuojelu	Päätökset	Päihdehuollon alkukartoitus	Laitostoiminta		
Paino	10	10	10	10	10	10	10	10	10	5	5	100	30
Ratkaisun toiminnallinen laajuus ja laatu: valmisisältö													
Aihealue	Lastentaudit	Syöpätaudit	Sähköinen asiointi	Kansantautien hoidon ja seurannan tuki	Kardiologia ja neurologia	Infektio-seuranta	Naistentaudit ja synnytys	Ylätason johtamisen tuki	Päivystys, teho, leikkaussali				
Paino	10	10	10	10	10	10	10	20	10			100	10
Ratkaisun käytettävyys													
Kohde-alue	Terveystietohuolto	Sosiaalihuolto	Asiakas/potilas-portaali										
Paino	55	25	20									100	40
Kirjalliset vastaukset järjestelmän toiminnalliseen vaatimusmäärittelyyn: nämä arvioidaan lopullisen tarjouspyynnön vastausten perusteella													
Aihealue													
Paino												0	20

2 Vähimmäisvaatimukset

Tuotevertailu B:ssä ei ole vähimmäisvaatimuksia.

3 Osa-alueittainen pisteiden muodostus

Kaikkien osa-alueiden ja aihe-alueiden lopputulokset pyöristetään kahden desimaalin tarkkuudelle normaalien pyöristyssääntöjen mukaisesti.

3.1 Toiminnallinen laajuus ja laatu: Ryhmäarvioinnit

Arviointi toteutetaan *käyttäjätarinoihin perustuvina aihealuekohtaisina ryhmäarviointina* seuraavilla aihealueilla:

1. A1 Päivystystoiminta
2. A2 Tehonäkymät
3. A3 Nestehoito + tehohoito
4. A4 Leikkaussali
5. A5 Avovastaanotto
6. A6 Hoitotyö
7. A7 Vuodeosasto
8. A8 Kotihoito ja sosiaalityö (ks. jäljempänä kohta 3.1.1)
9. A9 Suun terveydenhuolto (ks. jäljempänä kohta 3.1.1)
10. SR1 Lastensuojelu
11. SR2 Päätökset
12. SR3 Päihdehuollon alkukartoitus
13. SR 4 Laitostoiminta

Tarjoajat esittelevät tuotteensa toiminnallisuuksia perustuen etukäteen heille lähetettyihin käyttäjätarinoihin ja niissä oleviin tehtäviin.

Kukin kolmestatoista ryhmäarvioinnin aihe-alueesta pisteytetään arviointitilanteessa nähtyjen tarinoiden ja tehtävien demonstraatioiden perusteella.

Ryhmäarvioinnin kokonaisuuden osapainoarvo koko tuotevertailun arvioinnissa on 30%. Se muodostuu kahdesta osa-alueesta eli ryhmäarvioinnin tehtäväkohtaisista arvosanoista ja ryhmäarvioinnin kokonaisarvioinnista.

1. Tehtäväkohtaiset arvosanat

Kukin ryhmäarviointi jakautuu useisiin käyttäjätarinan sisällä oleviin tehtäviin ja niiden toteutumisen arviointeihin. Arvioijat perustavat omat arvionsa siihen, miten heidän mielestään kukin käyttäjätarinan tehtävä onnistuu käyttämällä seuraavia vastausvaihtoehtoja:

- 3 (toteutuu hyvin)
- 2 (toteutuu kohtalaisen hyvin)
- 1 (toteutuu melko huonosti)
- 0 (toteutuu huonosti tai ei ollenkaan).

Mikäli tehtävää ei näytetä lainkaan, on tulos 0 pistettä. Lisäksi tehtäväkohta merkitään erikseen ja tarjoajalta pyydetään myöhemmin lisäselvitys syistä, miksi kyseessä olevaa tehtävää ei demonstroitu.

Kunkin arvioijan käyttäjätarinan tehtäville antamista arvosanoista lasketaan keskiarvo. Näistä arvioijakohtaisista keskiarvoista lasketaan edelleen keskiarvo.

Tehtäväkohtaisten arvosanojen osapainoarvo kunkin ryhmäarvioinnin kokonaisarviossa on 30%.



2. Ryhmäarvioinnin kokonaisarvio

Kunkin käyttäjätarinan demonstroidun toiminnallisuuden soveltuvuutta kokonaisuutena käyttötarkoitukseensa arvioidaan asteikolla 6-0 seuraavasti:

- 6 (erittäin hyvin)
- 5 (hyvin)
- 4 (melko hyvin)
- 3 (ei hyvin eikä huonosti)
- 2 (huonosti)
- 1 (erittäin huonosti)
- 0 (ei ollenkaan)

Arvioijien kokonaisarvosanoista lasketaan keskiarvo. Tämän arvosanan osapainoarvo kunkin ryhmäarvioinnin kokonaisarviossa on 70%.

Seuraavalla sivulla olevassa taulukossa on kuvattu laskentatapaa esimerkkitapauksessa, jossa tuotteen Z osalta aihealuetta "avovastaanotto" arvioi kolme arvioijaa ("Arv 1" jne.) kymmenen tehtävän ("teh 1" jne.) ja yhden kokonaisarviokysymyksen perusteella.

Tuote Z	Avovastaanotto			
	Arv 1	Arv 2	Arv 3	KA
teh 1	2	1	2	1,67
teh 2	1	3	3	2,33
teh 3	2	2	2	2,00
teh 4	3	3	3	3,00
teh 5	1	2	1	1,33
teh 6	0	0	0	0,00
teh 7	1	1	1	1,00
teh 8	3	2	3	2,67
teh 9	3	3	3	3,00
teh 10	2	1	2	1,67
Tehtävien KA	1,8	1,8	2	1,87
Maksimiarvosana				3,00
Tuloksen osuus maksimiarvosanasta tehtävienKA/maksimiarvosana				0,62
Sisäinen painoarvo				30,00
Loppupisteet1 = sis painoarvo x tuloksen osuus maksimiarvosanasta				18,67

Tuote Z	Avovastaanotto			
	Arv 1	Arv 2	Arv 3	KA
Kokonaisarviokysymys	4	5	5	4,67
Maksimiarvosana				6,00
Tuloksen osuus maksimiarvosanasta = kokonaisarviokysymysKA/maksimiarvosana				0,78
Sisäinen painoarvo				70,00
Loppupisteet2 = sis painoarvo x tuloksen osuus maksimiarvosanasta				54,44

Kokonaispisteet Tuote Z, avovastaanotto = Loppupisteet1 + Loppupisteet2	73,11
---	--------------

Joissakin aihe-alueissa on useampi käyttäjätarina tehtävineen. Niiden loppuarvosana lasketaan yllä olevan tavoin siten, että käyttäjätarinat muodostavat pistelaskennassa yhden käyttäjätarinan, jossa on yksi kokonaisarviokysymys.

3.1.1 Toiminnallinen laajuus ja laatu: Ryhmäarvioinnit kotihoidon ja sosiaalityön sekä suun terveydenhuollon aihealueilla.

Aihealueiden 8. Kotihoito ja sosiaalityö (THR A8) ja 9. Suun terveydenhuolto (THR A9) tuloksia ei lasketa mukaan Tuotevertailu B:n lopputulokseen. Näiden aihealueiden arviointi vaikuttaa hankinnan lopullisen laajuuden määrittelyyn.

3.2 Toiminnallinen laajuus ja laatu: Valmis sisältö

Arviointi tapahtuu seuraavilla aihealueilla

1. VS1 Lastentaudit
2. VS2 Syöpätaudit
3. VS3 Sähköinen asiointi
4. VS4 Kansantautien hoidon ja seurannan tuki
5. VS5 Kardiologia ja neurologia
6. VS6 Infektioseuranta
7. VS7 Naistentaudit ja synnytys
8. VS8 Ylätason johtamisen tuki
9. VS9 Päivystys, teho ja leikkaussali

Kukin näistä yhdeksästä aihe-alueesta arvioidaan demonstroitavien sisältöjen kuvauksiin pohjautuvien tehtävien perusteella kahden aihealuekohtaisen kokonaisarviokysymyksen avulla.

Kokonaisarviokysymykset kohdistuvat valmissisällön eri arviointialueille: 1) hyödynnettävyyteen ja 2) laatuun. Ne arvioidaan asteikolla 6-0 seuraavasti:

- 6 (erittäin hyvin)
- 5 (hyvin)
- 4 (melko hyvin)
- 3 (ei hyvin eikä huonosti)
- 2 (melko huonosti)
- 1 (huonosti)
- 0 (ei ollenkaan)

Kullekin arviointialueelle (hyödynnettävyys, laatu) lasketaan arvioijien kokonaisarvosanoista keskiarvo. Valmissisältöarvioinnin osapainoarvo toiminnallisen laajuuden ja laadun arvioinnissa on 10%. Keskiarvoista lasketaan keskiarvo, joka on kyseisen valmissisällön loppuarvosana.

Seuraavalla sivulla olevassa taulukossa on kuvattu laskentatapaa esimerkkitapauksessa, jossa tuotteen Z osalta aihealuetta "Lastentaudit" arvioi kolme arvioijaa ("Arv 1" jne.) kolmen kokonaisarviokysymyksen perusteella.

Tuote Z	Lastentaudit			
	Arv 1	Arv 2	Arv 3	KA
Hyödynnettävyys	6	6	5	5,67
Laatu	4	4	3	3,67
Arvosanojen KA				4,67
Maksimiarvosana				6,00
Arvosanojen KA osuus maksimiarvosanasta = arvosanojen KA/maksimiarvosana				0,78
Skaalaus				100,00
Loppupisteet = skaalaus x Arvosanojen KA:n osuus maksimiarvosanasta				77,78

3.3 Käytettävyys

Käytettävyyttä arvioidaan neljällä eri metodilla:

- Terveiden- ja sosiaalihuollon paritesteillä
- Käyttäjätarinoihin perustuvissa ryhmäarviointitilaisuuksissa
- Asiakas- ja potilasportaalien (jäljempänä portaali) käytettävyydesteillä
- Portaalien esteettömyyden asiantuntija-arviointina

Näiden metodien tuottamat arviointitulokset esitetään kolmena kokonaisuutena:

- Terveidenhuollon käytettävyys (paritestit, ryhmäarvioinnit)
- Sosiaalihuollon käytettävyys (paritestit, ryhmäarvioinnit)
- Portaalien käytettävyys (käytettävyytestit, esteettömyysarviointi)

Kokonaisuuksien painoarvot on esitetty jäljempänä taulukossa 1.

Terveiden- ja sosiaalihuollon paritestit: Paritestien aiheita ovat terveydenhuollon puolella "B1 Ajanvaraus", "B2 Lääkitys hoitaja" ja "B3 Lääkitys lääkäri" sekä sosiaalihuollon puolella "C1 Ilmoituksen/hakemuksen vastaanotto ja käsittely" ja "C2 Aikaisempien tietojen haku asiakkaasta". Paritestitilaisuudessa kaksi terveyden- ja sosiaalihuollon ammattilaisasiantuntijaa toimii testikäyttäjinä ja käytettävyydasiantuntija testiohjaajana. Käytettävyyttä arvioidaan testikäyttäjien testitehtävistä suoriutumiseen perustuen. Lisäksi käytettävyydasiantuntijat arvioivat paritestitehtäviin liittyviä optimaalisia suorituspolkuja (nk. tarjoajien toimittamat "oikeat polut").

Terveiden- ja sosiaalihuollon ryhmäarviointitilaisuudet: Käytettävyyttä arvioidaan osana toiminnallisen laajuuden ja laadun ryhmäarviointeja. Ammattilaisasiantuntijat arvioivat käytettävyyttä kaikissa tilaisuuksissa.

Käytettävyyssasiantuntijat arvioivat käytettävyyttä valituissa tilaisuuksissa (merkitty *:llä):

1. Terveydenhuolto:

- A1 Päivystystoiminta(*)
- A2 Tehonäkymät
- A3 Nestehoito + tehohoito(*)
- A4 Leikkaussali
- A5 Avovastaanotto(*)
- A6 Hoitotyö(*)
- A7 Vuodeosasto(*)

2. Sosiaalihuolto:

- SR1 Lastensuojelu
- SR2 Päätökset (osiot SR2A, SR2B, ja SR2C arvioidaan erillisinä *)
- SR3 Päihdehuollon alkukartoitus
- SR4 Laitostoiminta

Portaalin käytettävyyssitestit: Portaalin käytettävyyssitestitilaisuudessa kansalaiskäyttäjä toimii testikäyttäjänä ja käytettävyyssasiantuntija testiohjaajana. Testikäyttäjät ovat eri-ikäisiä ”tavallisia” kansalaisia, joilla ei ole sosiaali- tai terveydenhuollon koulutusta. Käytettävyyttä arvioidaan testikäyttäjien testitehtävistä suoriutumiseen perustuen. Lisäksi käytettävyyssasiantuntijat arvioivat käytettävyyssitestitehtäviin liittyviä optimaalisia suorituspolkuja (nk. tarjoajien toimittamat ”oikeat polut”).

Portaalin esteettömyysarviointi: Esteettömyysarvioinnissa esteettömyysarvioinnin asiantuntijat arvioivat portaalia kansalaiskäyttäjien näkökulmasta heuristista muistilistaa apuna käyttäen.

Käytettävyyssarvioinnin kokonaisarvo muodostuu terveydenhuollon, sosiaalihuollon sekä portaalin arviointipisteiden yhdistelmänä siten, että terveydenhuollon osuus on 55%, sosiaalihuollon osuus 25% ja portaalin osuus 20% käytettävyyden kokonaisarvioinnista. Kunkin kokonaisuuden pisteet jakautuvat edelleen tuloksellisuuden, virheettömyyden, tehokkuuden ja opittavuuden, tyytyväisyyden, käyttöliittymäsuunnittelun onnistuneisuuden sekä esteettömyyden pisteisiin, joiden painoarvo vaihtelee kussakin kokonaisuudessa seuraavalla sivulla olevan taulukon ”Taulukko 1” mukaisesti.

Taulukko 1: Käytettävyyssarvioinnin painoarvot.

	Kokonaisuudet		
	Terveysthuolto	Sosiaalihoito	Portaali
Kokonaisuuden painoarvo:	55%	25%	20%
Käytettävyyden aihealueet:	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Tuloksellisuus	30	30	25
Virheettömyys	20	20	10
Tehokkuus ja opittavuus	20	20	10
Tyytyväisyys	20	20	25
Käyttöliittymäsuunnittelun onnistuneisuus	10	10	-
Esteettömyys	-	-	30
<i>Yhteensä</i>	<i>100</i>	<i>100</i>	<i>100</i>

Käytettävyyssarvioinnin kokonaisarvo muodostuu kaikkien kokonaisuuksien käytettävyyden aihealueiden arviointipisteiden yhdistelmänä. Liitteessä 2 on esitetty laskelma käytettävyyden kokonaisuuskohtaisten pisteiden ja kokonaisarvion muodostumisesta.

Pisteiden laskenta ja kertyminen on kuvattu jäljempänä käytettävyyden aihealuekohtaisesti.

3.3.1 Tuloksellisuus

Tuloksellisuutta mitataan terveyden- ja sosiaalihoitoon paritesteissä ja portaalin käytettävyyssarvioinnissa onnistuneesti suoritettujen testitehtävien %-osuutena. Terveysthuoltoon paritesteiden aiheet B1-B3 käsitellään yhtenä kokonaisuutena, vastaavasti sosiaalihoitoon paritesteiden aiheet C1-C2.

Taulukko 2: Tuloksellisuuden mittari ja osapainoarvot.

Tuloksellisuus		Terveysthuolto (30%)	Sosiaalihoito (30%)	Portaali (25%)
Metodi	Mittari	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Paritesti / käytettävyyssarvio	Onnistuneesti suoritettujen testitehtävien %-osuus	100	100	100

Jokaisessa aiheessa on määritetty tehtäväsuorituksen maksimiaiika, joka on kaikille tehtäville sama. Jos testitehtävää ei ehditä suorittamaan sille varatussa ajassa, katsotaan tehtäväsuoritus epäonnistuneeksi. Testiohjaajana toimiva käytettävyyssarviointiaika merkitsee testin aikana tehtävän onnistuneeksi tai epäonnistuneeksi käyttäjien toimintaan perustuen verrattuna tavoiteltuun lopputulokseen, joka on määritetty testitehtäväkuvauksessa ja joka näkyy oikeassa polussa. Onnistumisprosentti kussakin kokonaisuudessa on kaikkien testitilaisuuksien kaikkien onnistuneesti suoritettujen tehtävien lukumäärän suhde suoritettavana olleiden tehtävien kokonaismäärään.

Kunkin kokonaisuuden tuloksellisuuspisteet määräytyvät onnistumisprosentin perusteella alla olevan taulukon mukaisesti:

Taulukko 3: Tuloksellisuuspisteiden määräytyminen.

Tuloksellisuuspisteet:	5	4	3	2	1
Onnistumis%	>90	$90 \geq \% > 80$	$80 \geq \% > 60$	$60 \geq \% > 30$	≤ 30

Mikäli tarjoajien saavuttamien pistemäärien ero on 1, varmistetaan laskemalla, että onnistumisprosenttien ero on tilastollisesti merkitsevä. Mikäli ero ei ole tilastollisesti merkitsevä, saavat molemmat tarjoajat samat pisteet, jotka määräytyvät korkeamman onnistumisprosentin mukaan. Tilastollisen merkitsevyyden laskeminen on kuvattu liitteessä 1.

Esimerkki: Kokonaisuus sisältää kaksi aihetta X ja Y. Aiheessa X on yhteensä 3 tehtävää ja toteutetaan 3 testiä. Aiheessa Y on yhteensä 4 tehtävää ja toteutetaan 3 testiä. Aiheessa X kaikissa kolmessa testissä suoritetaan onnistuneesti yhteensä 7/9 tehtävää. Vastaavasti aiheessa Y suoritetaan onnistuneesti 8/12 tehtävää. Onnistumisprosentti on siis: $(7+8) / (9+12) \rightarrow n. 71\%$ mikä vastaa 3 pistettä.

3.3.2 Virheettömyys

Virheettömyyttä arvioidaan terveyden- ja sosiaalihuollon paritesteissä ja portaalin käytettävyydesteissä onnistuneesti suoritettujen testitehtävien suorituksen aikaisten virheiden perusteella. Terveysthuollon paritestiä aiheet B1-B3 käsitellään yhtenä kokonaisuutena, vastaavasti sosiaalihuollon paritestiä aiheet C1-C2.

Taulukko 4: Virheettömyyden mittari ja osapainoarvot.

Virheettömyys		Terveysthuolto (20%)	Sosiaalihuolto (20%)	Portaali (10%)
Metodi	Mittari	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Paritesti / käytettävyydesti	Testitehtävien suorituksen aikaisten virheet	100	100	100

Virheellä tarkoitetaan testitehtävien suorituksen aikaisia käyttäjän ja järjestelmän vuorovaikutukseen liittyviä toimia, jotka poikkeavat "oikeasta polusta": tavoitteen kannalta ei-tarkoituksenmukaisia toimia (esim. vääriin näkymään siirtyminen), tahattomia ylimääräisiä toimia, erehdyksiä tai oleellisen asian huomioimatta jättämistä. Suorituksen aikaisten virheiden luokitellaan pieniin (0,5p.) ja suuriin (1p.) virheisiin niiden vaikuttavuuden perusteella. Näiden perusteella lasketaan kunkin onnistuneen tehtävän virhepisteet.

Virhepisteet määräytyvät seuraavalla tavalla:

1. Ensin lasketaan kunkin onnistuneesti suoritettujen tehtävien virhepisteet kussakin testitilaisuudessa.

2. Sen jälkeen lasketaan testitilaisuuskohtaisesti virhepisteiden keskiarvo. Mikäli yksittäisessä testitilaisuudessa ei suoriteta yhtään testitehtävää onnistuneesti virhepisteiden keskiarvo = 12p.
3. Tämän jälkeen lasketaan saman aiheen eri testitilaisuuksista virhepisteiden keskiarvojen keskiarvo. Näiden perusteella määräytyy kullekin aiheelle alla olevan taulukon mukaisesti virheettömyyspisteet välillä 1-5.
4. Nämä virheettömyyspisteet lasketaan kussakin kokonaisuudessa yhteen. Tämä on virheettömyyden arvio pisteinä välillä 1-15 terveydenhuollon kokonaisuudessa, 1-10 sosiaalihuollon kokonaisuudessa ja 1-5 portaalikokonaisuudessa.

Erytistapauksena on, mikäli missään aiheen testitilaisuudessa ei suoriteta yhtään tehtävää onnistuneesti. Tällöin virheettömyyspisteet aiheesta ovat 0.

Taulukko 5: Aihekohtaisten virheettömyyspisteiden määräytyminen.

Virheettömyyspisteet:	5	4	3	2	1	0
Virhepisteet:	0-1 kaksi pientä virhettä sallitaan	1 < pisteet ≤ 2	2 < pisteet ≤ 4	4 < pisteet ≤ 6	> 6 enemmän kuin neljä suurta virhettä	Erytistapaus: ei yhtään onnistuneesti suoritettua tehtävää yhdessäkään aiheen testissä.

Mikäli tarjoajien saavuttamien aihekohtaisten pistemäärien ero on 1, varmistetaan laskemalla, että virhepisteiden keskiarvojen ero on tilastollisesti merkitsevä. Mikäli ero ei ole tilastollisesti merkitsevä, saavat molemmat tarjoajat samat pisteet, jotka määräytyvät matalemmän virhepistekeskiarvon mukaan. Tilastollisen merkitsevyyden laskeminen on kuvattu liitteessä 1.

Esimerkki: Kokonaisuus sisältää kaksi aihetta X ja Y. Aiheessa X on yhteensä 3 tehtävää ja toteutetaan 3 testiä. Testissä X1 suoritetaan onnistuneesti 3/3 tehtävää, testissä X2 onnistuneesti 2/3 tehtävää ja testissä X3 onnistuneesti 0/3 tehtävää (kts. taulukko alla). Virhepisteiden keskiarvojen keskiarvo aiheessa X on: 4,444p., mikä vastaa virheettömyyspisteinä 1 pistettä. Aiheessa Y vastaava keskiarvo on 1,801p., mikä vastaa virheettömyyspisteinä 4 pistettä → Kokonaisuuden virheettömyyden pisteet ovat 5.

Aihe X	Testi X1	Testi X2	Testi X3
Tehtävä 1	Onnistui, 2p. virheitä	Onnistui, 1p. virheitä	Ei saatu suoritettua
Tehtävä 2	Onnistui, 1p. virheitä	Onnistui, 3p. virheitä	Ei saatu suoritettua
Tehtävä 3	Onnistui, 1p. virheitä	Ei saatu suoritettua	Ei saatu suoritettua
Virhepisteiden ka.	1,333p.	2,000p.	Ei onnistuneesti suoritettuja tehtäviä → 10,000p

3.3.3 Tehokkuus ja opittavuus

Tehokkuutta ja opittavuutta arvioidaan terveyden- ja sosiaalihuollon kokonaisuuksissa perustuen paritesteihin ja ryhmäarviointeihin viiden eri mittarin avulla ja portaalin kokonaisuudessa perustuen käytettävyydesteihin kahden eri mittarin avulla. Alla olevassa taulukossa on esitetty arviointimetodit ja niihin liittyvät mittarit.

Taulukko 6: Tehokkuuden ja opittavuuden mittarit ja osapainoarvot.

Tehokkuus ja opittavuus		Terveysthuolto (20%)	Sosiaalihuolto (20%)	Portaali (10%)
Metodi	Mittari	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Paritesti / käytettävyydesti	Tehokkuus: Oikeiden polkujen pituus askeleina	25	25	70
	Opittavuus: Käyttäjien vastaukset loppukyselyn opittavuuskysymyk-	15	15	30
Ryhmäarviointi	Tehokkuus: Käytettävyyssiantuntijan arvio käytön tehokkuudesta nähtyyn perustuen	20	20	-
	Opittavuus: Käytettävyyssiantuntijan arvio opittavuudesta ammattilaisiantuntijoiden vastauksiin perustuen	20	20	-
	Opittavuus ja tehokkuus: Käytettävyyssiantuntijan arvio mukautettavuuden opittavuudesta/tehokkuudesta nähtyyn perus-	20	20	-
Yhteensä		100%	100%	100%

Tehokkuus: Oikeiden polkujen pituus askeleina (paritesti, käytettävyydesti)

Kustakin kokonaisuudesta valitaan joukko testitehtäviä, joiden perusteella arvioidaan tehokkuutta. Arviointi tehdään laskemalla valittuihin testitehtäviin liittyvien oikeiden suorituspolkujen pituus askeleina. Askeleella tarkoitetaan yksittäistä vuorovaikutustapahtumaa käyttäjän ja järjestelmän välillä kuten valikon avaamista, painikkeen painamista tai tekstin syöttöä tekstikenttään. Laskenta tehdään suorittamalla tehtävät arvioitavalla järjestelmällä perustuen tarjoajan toimittamiin oikeisiin polkuihin, joissa on kuvattu optimaalinen suoritustapa. Mikäli askelten määrän laskeminen ei ole mahdollista toimitetun oikean polun perusteella tai oikea polku ei johda onnistuneeseen tehtäväsuoritukseen (polusta puuttuu osioita, tai se ei kuvaa tehtäväsuoritusta alusta loppuun), katsotaan polun pituuden olevan "ääretön ∞ ".

Oikeiden polkujen pisteet määräytyvät seuraavasti: Verrataan jokaisen valitun tehtävän suorituspolun pituutta tarjoajien välillä. Mikäli suorituspolut ovat molemmilla tarjoajilla kyseisen tehtävän kohdalla lähes saman pituisia (= askelten lukumäärän ero pienempi kuin 15 %) ovat molemmat suorituspolut 1 pisteen arvoisia. Mikäli toisen tarjoajan suorituspolku on toista enemmän kuin 15 % lyhyempi on tämä lyhyempi suorituspolku on 1 pisteen arvoinen. Pidemmän polun omaava tarjoaja saa 0 pistettä. Nämä pisteet lasketaan kokonaisuuskohtaisesti yhteen. Maksimipistemäärä on kussakin kokonaisuudessa valittujen tehtävi-

en lukumäärä. Erityistapauksena on se, että molempien tarjoajien saman tehtävän suorituspolun pituus on ∞ . Tällöin molemmat saavat 0 pistettä ko tehtävästä.

Opittavuus: Käyttäjien vastaukset loppukyselyn opittavuuskysymyksiin (paritesti, käytettävyydestä)

Opittavuutta arvioidaan pari- ja käytettävyydestien yhteydessä loppukyselyyn annettujen vastausten perusteella. Loppukysely sisältää opittavuuteen liittyviä väittämiä, joihin jokainen testiin osallistuva käyttäjä vastaa asteikolla 1-5. Kokonaisuuskohtaisesti lasketaan kaikista näihin väittämiin annetuista vastauksista keskiarvo, joka on suoraan pisteet tästä osuudesta.

Mikäli tarjoajien saavuttamien pistemäärien ero on 1, varmistetaan laskemalla, että keskiarvojen ero on tilastollisesti merkitsevä. Mikäli ero ei ole tilastollisesti merkittävä, saavat molemmat tarjoajat samat pisteet, jotka määräytyvät korkeamman keskiarvon mukaan. Tilastollisen merkitsevyyden laskeminen on kuvattu liitteessä 1.

Tehokkuus: Käytettävyyssasiantuntijan arvio käytön tehokkuudesta nähtyyn perustuen (ryhmäarviointi)

Tehokkuutta arvioidaan ryhmäarviointien käyttäjätarinoihin perustuen niissä arviointitilaisuuksissa, joihin osallistuu käytettävyyssasiantuntija (tilaisuudet A1, A3, A5, A6, A7, SR2): 5 = sujuva suorituspolku, joka sisältää tehtävän tavoitteeseen nähden mielekkään määrän vuorovaikutusaskleita; 1 = monimutkainen suorituspolku, joka sisältää paljon vuorovaikutusaskleita ja epäloogisia siirtymiä. Näistä arvioista lasketaan kokonaisuuskohtainen keskiarvo, joka on suoraan pisteet tästä osuudesta.

Opittavuus: Käytettävyyssasiantuntijan arvio opittavuudesta ammattilaisasiantuntijoiden vastauksiin perustuen (ryhmäarviointi)

Opittavuutta arvioidaan ryhmäarviointien käyttäjätarinoihin perustuen niissä arviointitilaisuuksissa, joihin osallistuu käytettävyyssasiantuntija (tilaisuudet A1, A3, A5, A6, A7, SR2). Tarinoiden yhteydessä käytettävyyssasiantuntija esittää osallistuville ammattilaisasiantuntijoille ennalta suunniteltuja ryhmäarviointitilaisuuden aikana nähtyihin asioihin perustuvia kysymyksiä liittyen käytön opittavuuteen ja intuitiivisuuteen. Käytettävyyssasiantuntija arvioi kunkin kysymyksen kohdalla vastausten ja järjestelmän suunnitellun käytötavan, jonka demonstroija näyttää, suhdetta toisiinsa asteikolla 5-1: 5 = ammattilaisasiantuntijoiden vastaukset vastaavat järjestelmän suunniteltua käyttötapaa; 1 = ammattilaisasiantuntijoiden vastaukset eroavat merkittävästi järjestelmän suunnitellusta käyttötavasta tai he eivät osaa antaa vastausta kysymykseen. Näistä arvioista lasketaan kokonaisuuskohtainen keskiarvo, joka on suoraan pisteet tästä osuudesta.

Opittavuus ja tehokkuus: Käytettävyyssasiantuntijan arvio mukautettavuuden opittavuudesta/tehokkuudesta nähtyyn perustuen (ryhmäarviointi)

Mukautettavuuden käytettävyyttä arvioidaan ryhmäarviointien käyttäjätarinoihin perustuen niissä arviointitilaisuuksissa, joihin osallistuu käytettävyyssasiantuntija (tilaisuudet A1, A3, A5, A6, A7, SR2). tarinat sisältävät ennalta suunniteltuja osioita, joiden avulla demonstroidaan järjestelmän mukautettavuutta ammattilaisasiantuntijoiden näkökulmasta. Kustakin osiosta käytettävyyssasiantuntija antaa nähtyyn perustuen arvionsa mukautettavuuden opittavuudesta ja tehokkuudesta asteikolla 5-1: 5 = looginen ja suju-

va suorituspoltu, joka sisältää tehtävän tavoitteeseen nähden mielekkään määrän vuorovaikutusaskelita; 1 = monimutkainen ja epäintuitiivinen suorituspoltu, joka sisältää paljon vuorovaikutusaskelita ja epäloogisia siirtymiä. Näistä arvioista lasketaan kokonaisuuskohtainen keskiarvo, joka on suoraan pisteet täästä osuudesta.

3.3.4 Tyytyväisyys

Tyytyväisyyttä arvioidaan terveyden- ja sosiaalihuollon kokonaisuuksissa perustuen paritesteihin ja ryhmäarviointeihin kolmen eri mittarin avulla ja portaalin kokonaisuudessa perustuen käytettävyydesteihin kolmen eri mittarin avulla. Alla olevassa taulukossa on esitetty arviointimetodit ja niihin liittyvät mittarit.

Taulukko 7: Tyytyväisyyden mittarit ja osapainoarvot.

Tyytyväisyys		Terveystenhuolto (20%)	Sosiaalihuolto (20%)	Portaali (25%)
Metodi	Mittari	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Paritesti / käytettävyydesti	Käytön aikaiset positiiviset / negatiiviset arviot	30	30	40
	Käyttäjien vastaukset loppukyselyyn	20	20	40
	Käyttäjien määrittämä paremmuusjärjestys	-	-	20
Ryhmäarviointi	Ammattilaisasiantuntijoiden vastaukset loppukyselyyn	50	50	-
Yhteensä		100%	100%	100%

Tyytyväisyys: Käytön aikaiset positiiviset / negatiiviset arviot (paritesti, käytettävyydesti)

Tyytyväisyyttä arvioidaan pari- ja käytettävyydestien yhteydessä testeihin osallistuvien käyttäjien antamiin käytön aikaisiin positiivisiin ja negatiivisiin arvioihin perustuen. Kaikissa pari- ja käytettävyydesteissä testikäyttäjät antavat tehtäväkohtaiset tyytyväisyysarviot jokaisen testitehtävän suorituksen päätteeksi. Lisäksi paritesteissä molemmat testikäyttäjät antavat tehtäväsuorituksen aikana kokemaansa perustuen henkilökohtaisia tyytyväisyysarvioita.

Positiivinen arvio ("*plus-arvio*") tarkoittaa testikäyttäjän kokemaa positiivista käytön aikaista asiaa ja testi-tehtävän suorittamiseen liittyvää tyytyväisyyttä. Vastaavasti negatiivinen arvio ("*miinus-arvio*") tarkoittaa käyttäjän kokemaa hankaluutta, epävarmuutta tai tyytymättömyyttä järjestelmän toimintaa kohtaan tehtäväsuorituksen yhteydessä.

Tulos lasketaan seuraavasti:

Portaalin käytettävyydestien aikana annetuista tehtäväkohtaisista arvioista lasketaan kaikkien käyttäjien plus-arvioimien tehtävien osuus kaikista tehtävistä. Tämän käytönaikaisen tyytyväisyysprosentin perusteella määrättyy seuraavalla sivulla olevan taulukon mukaisesti tyytyväisyyspisteet välillä 1-5. Mikäli testin aikana ei ehditä suorittamaan kaikkia suunniteltuja testitehtäviä, merkitään suorittamatta jääneet miinus-arvioiduiksi.

Terveyden- ja sosiaalihuollon paritestien yhteydessä lasketaan ensin testitilanne- ja tehtäväkohtaisesti kahden käyttäjän antamien plus- ja miinus-arvioiden lukumäärien summa. Tämän jälkeen lasketaan testitilannekohtaisesti niiden tehtävien osuus kaikista tehtävistä, jotka ovat saaneet enemmän plus- kuin miinus-arvioita. Mikäli testin aikana ei ehditä suorittamaan kaikkia suunniteltuja testitehtäviä, merkitään suorittamatta jääneet miinusarvioiksi. Mikäli tehtäväkohtaisten plus- ja miinus-arvioita on täsmälleen saman verran, merkitään tehtäväkohtainen tulos plus-arvioiksi. Tämän jälkeen lasketaan saman aiheen eri testitilaisuuksista plus-arvioitujen tehtävien osuuksien keskiarvo. Keskiarvon perusteella määräytyy kullekin aiheelle alla olevan taulukon mukaisesti käytönaikaiset tyytyväisyyspisteet välillä 1-5. Nämä tyytyväisyyspisteet lasketaan kussakin kokonaisuudessa yhteen. Tämä on käytön aikainen tyytyväisyysarvio pisteinä välillä 1-15 terveydenhuollon kokonaisuudessa ja 1-10 sosiaalihuollon kokonaisuudessa.

Tyytyväisyyspisteet:	5	4	3	2	1
Käytönaikainen tyytyväisyys%	>80	$80 \geq \% > 60$	$60 \geq \% > 40$	$40 \geq \% > 20$	≤ 20

Mikäli tarjoajien saavuttamien aihekohtaisten pistemäärien ero on 1, varmistetaan laskemalla, että käytönaikaisten tyytyväisyysprosenttien ero on tilastollisesti merkitsevä. Mikäli ero ei ole tilastollisesti merkitsevä, saavat molemmat tarjoajat samat pisteet, jotka määräytyvät korkeamman käytönaikaisen tyytyväisyysprosentin mukaan. Tilastollisen merkitsevyyden laskeminen on kuvattu liitteessä 1.

Esimerkki: Kokonaisuus sisältää kaksi aihetta X ja Y. Aiheessa X on yhteensä 3 tehtävää ja toteutetaan 3 testiä. Testissä X1 plus-arvioitujen tehtävien osuus on 66,6%, testissä X2 osuus on 33,3% ja X3 osuus on 66,6% (kts. taulukko alla). Aiheen X osalta käytön aikainen tyytyväisyyskeskiarvo on tällöin 55,6% mikä vastaa tyytyväisyyspisteinä 3 pistettä. Aiheessa Y vastaava keskiarvo on 71,5%, mikä vastaa tyytyväisyyspisteinä 4 pistettä → Kokonaisuuden käytönaikaiset tyytyväisyyspisteet ovat 7.

Aihe X	Testi X1	Testi X2	Testi X3
Tehtävä 1	plus-arviot = 3kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: plus	plus-arviot = 3kpl miinus-arviot = 4kpl → tehtäväkohtainen tulos: miinus	plus arviot = 3kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: plus
Tehtävä 2	plus-arviot = 2kpl miinus-arviot = 4kpl → tehtäväkohtainen tulos: miinus	plus-arviot = 0kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: miinus	plus-arviot = 3kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: plus
Tehtävä 3	plus-arviot = 3kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: plus	plus-arviot = 2kpl miinus-arviot = 2kpl → tehtäväkohtainen tulos: plus	Tehtävää ei suoritettu → tehtäväkohtainen tulos: miinus
Plus-arvioitujen tehtävien osuus kaikista	2/3 = 66,6%	1/3 = 33,3%	2/3 = 66,6%

Tyytyväisyys: Käyttäjien vastaukset loppukyselyyn (paritesti, käytettävyytesti)

Tyytyväisyyttä arvioidaan pari- ja käytettävyytestien yhteydessä loppukyselyyn annettujen vastausten perusteella. Loppukyselynä on standardoitu käytettävyysskysely SUS (System Usability Scale)¹. Kukin testi-käyttäjä vastaa testin päätteeksi kyselyn väittämiin asteikolla 1-5 (1 = täysin eri mieltä, 5 = täysin samaa mieltä).

Loppukyselyn tyytyväisyyspisteet määräytyvät seuraavasti: Ensin lasketaan kunkin käyttäjän SUS-kyselyn pistemäärän summa (maksimi 100 pistettä), jonka jälkeen lasketaan kokonaisuuskohtaisesti näiden keskiarvo. Näiden perusteella määräytyy kullekin kokonaisuudelle seuraavalla sivulla olevan taulukon mukaisesti loppukyselyn tyytyväisyyspisteet välillä 1-5.

¹ Brooke, J. (1996) SUS: a "quick and dirty" usability scale. Teoksessa: Jordan, P., Thomas, B, Weerdmeester, B., (toim.) Usability Evaluation in Industry. Taylor & Francis, Lontoo, s. 189-194.

Loppukyselyn tyytyväisyyspisteet:	5	4	3	2	1
SUS-kyselyn pisteet:	pisteet >78,9	78,8 ≥ pisteet > 72,6	72,5 ≥ pisteet > 62,7	62,6 ≥ pisteet > 51,7	pisteet ≤ 51,6

Mikäli tarjoajien saavuttamien kokonaisuuskohtaisten pistemäärien ero on 1, varmistetaan laskemalla, että SUS-kyselyn pisteiden keskiarvojen ero on tilastollisesti merkitsevä. Mikäli ero ei ole tilastollisesti merkitsevä, saavat molemmat tarjoajat samat pisteet, jotka määräytyvät korkeamman keskiarvon mukaan. Tilastollisen merkitsevyyden laskeminen on kuvattu liitteessä 1.

Tyytyväisyys: Käyttäjien määrittämä paremmuusjärjestys (käytettävyydestä)

Tyytyväisyyttä arvioidaan portaalin käytettävyydestien yhteydessä vertailukysymykseen perustuen. Jälkimmäisen portaalin testin lopuksi, loppukyselyn täyttämisen jälkeen, testikäyttäjä asettaa arvioimansa kaksi portaalia käyttäjäkokemuksensa perusteella paremmuusjärjestykseen. Paremmuusjärjestyksen tyytyväisyyspisteet määräytyvät seuraavasti: Lasketaan niiden käyttäjien %-osuus kaikista testikäyttäjistä, jotka ovat asettaneet ko. portaalin paremmuusjärjestyksessä ensimmäiseksi.

Tyytyväisyys: Ammattilaisasiantuntijoiden vastaukset loppukyselyyn (ryhmäarviointi)

Tyytyväisyyttä arvioidaan ammattilaisasiantuntijoiden näkökulmasta kaikissa ryhmäarviointitilaisuuksissa loppukyselyyn annettujen vastausten perusteella. Loppukysely sisältää 10 väittämää, joihin ammattilaisasiantuntijat vastaavat pohjautuen sekä käyttäjätarinan demonstraatioissa koettuun ja nähtyyn, että heidän omaan sosiaali- ja terveydenhuollon työn (substanssi) asiantuntemukseen, työprosessien ja käytäntöjen tuntemukseen ja aiempiin tietojärjestelmien käyttäjäkokemuksiin. Väittämiin vastataan asteikolla 3-0 (3 = vahvasti samaa mieltä, 0 = vahvasti eri mieltä).

Ryhmäarviointien loppukyselyn tyytyväisyyspisteet määräytyvät seuraavasti: Ensin lasketaan vastaajakohdaisesti vastauspisteiden summa (maksimi 30 pistettä), jonka jälkeen lasketaan ryhmäarviointikohtaisesti kaikkien vastaajien vastauspisteiden summan keskiarvo. Lopuksi näistä keskiarvoista lasketaan sosiaali- ja terveydenhuollon kokonaisuuskohtaiset keskiarvot.

3.3.5 Käyttöliittymäsuunnittelun onnistuneisuus

Käyttöliittymäsuunnittelun onnistuneisuutta arvioidaan terveyden- ja sosiaalihuollon kokonaisuuksissa niissä ryhmäarviointitilaisuuksissa, joihin osallistuu käytettävyyssiantuntija (tilaisuudet A1, A3, A5, A6, A7, SR2).

Taulukko 8: Käyttöliittymäsuunnittelun onnistuneisuuden mittari ja osapainoarvot.

Käyttöliittymäsuunnittelun onnistuneisuus		Terveystuolto (10%)	Sosiaalihuolto (10%)	Portaali (-)
Metodi	Mittari	Osapainoarvo %	Osapainoarvo %	Osapainoarvo %
Ryhmä-arviointi	Käytettävyyssiantuntijan arvio suunnitteluheuristiikkoihin ja käytyn keskusteluun perustuen	100	100	

Ryhmäarviointitilaisuuksissa käyttöliittymäsuunnittelun onnistuneisuutta arvioidaan käyttäjätarinan demonstraation aikana käytettävyyssiantuntijan toimesta. Arvioinnissa lähtökohtana käytetään ennalta valmisteltuja yleisiin suunnittelusääntöjä rikkoviin asiakas- ja potilastietojärjestelmien käytettävyysoongelmiin liittyviä kysymyksiä. Demonstaation tauoilla käytettävyyssiantuntija esittää näitä kysymyksiä ammattilaisiantuntijoille, jotka vastaavat kysymyksiin nähdyn ja koetun perusteella. Nämä kysymykset ovat samoja molemmille tarjoajille ja ne esitetään soveltuviin kohdin osana eri käyttäjätarinoita (molemmille tarjoajille samoissa ryhmäarviointitilaisuuksissa).

Käytettävyyssiantuntija antaa arvion käyttöliittymäsuunnittelun onnistuneisuudesta kysymyskohtaisesti asteikolla 5-1 perustuen ammattilaisiantuntijoiden vastauksiin: 5 = Ammattilaisiantuntijat eivät koe asian häiritsevän järjestelmän käyttöä tai työtehtävien suorittamista; 1 = Todennettu käytettävyysongelman, joka saattaa johtaa ammattilaisiantuntijakäyttäjän näkökulmasta vaikutukseltaan riski- tai haittapahtumatilanteeseen. Näistä arvioista lasketaan kokonaisuuskohtainen keskiarvo, joka on suoraan pisteet tästä osuudesta.

Esimerkki: Yleinen asiakas- ja potilastietojärjestelmien käytettävyysongelman on: "Yhteenvetönäkymästä on vaikeaa erottaa tärkeitä tietoja". Tähän liittyen esitetään ryhmäarviointitilaisuudessa seuraavat kysymykset: "Erottuvatko äsken nähdystä yhteenvetönäkymästä tärkeitä tiedot?", "Onko näkymä jäsennetty mielekkäästi?", "Ovatko käyttötilanteen näkökulmasta tarvittavat ja tärkeitä tiedot selkeästi esillä?".

3.3.6 Esteettömyys

Esteettömyyttä arvioidaan:

1. W3C:n² tuottamalla WCAG 2.0 -kriteeristöllä.³ WCAG 2.0 -kriteeristössä käytetään arviointitasona AA-luokitusta.
2. Kehitysvammaliiton (jäljempänä KVL) Papunet-verkkopalveluuyksikön⁴ selkokielisten verkkopalveluiden heuristiikoista⁵ kehitetyllä kriteeristöllä⁶.

Esteettömyysarvioinneissa hyödynnetään portaalin käytettävyydestin tehtäviä. Näistä valitaan arvioitavat tehtävät, joiden suorituspoluista arvioidaan kaikki niihin sisältyvät sivunäkymät esteettömyyden näkökulmasta. Näkymät määritetään tarjoajien toimittamista suorituspoluista (nk. "oikeat polut").

WCAG 2.0 sisältää A ja AA -tasoilla 38 erillistä kriteeriä. KVL:n heuristiikoissa kriteereitä on 19 eli yhteensä arvioitavia esteettömyyden kriteereitä on 57. Jokainen WCAG 2.0 -kriteeri pisteytetään niin, että se saa arvon 0, 1 tai 2. Vastaavasti KVL:n heuristiikoissa pisteytyksen lopputuloksena on arvo 0 / 0,5 / 1. Tällöin WCAG 2.0 -kriteerien maksimipistemäärä on 76 ja KVL:n heuristiikkojen 19.

Taulukko 9: Esteettömyysarvioinnin kriteerit ja pisteytys.

Kriteeristö	Kriteerien lkm	Mahdolliset pistearvot	Max. pisteet / tehtävä
WCAG 2.0	38	0 / 1 / 2	76
KVL:n heuristiikat	19	0 / 0,5 / 1	19

Kolme arvioijaa arvioi valittuihin testitehtäviin liittyvät sivunäkymät ja pisteyttää jokaisen esteettömyyskriteerin testitehtävittäin. Pistearvot määräytyvät niin, että 2 pistettä (KVL:n heuristiikat 1 piste) portaali saa testitehtävästä silloin, kun kriteeri täyttyy, 1 pistettä (KVL:n heuristiikat 0,5 pistettä) silloin, kun kriteeri täyttyy osittain ja 0 pistettä kun kriteeri ei täyty. Arvioinnin jälkeen lasketaan arvioijakohtaisesti annetut pisteet yhteen ja tämän jälkeen näiden keskiarvo. Tämä pistemäärä esitetään 0,5 pisteen tarkkuudella.

² <http://www.w3.org/>

³ <http://www.w3.org/Translations/WCAG20-fi/>

⁴ <http://papunet.net/yksikko>

⁵ http://papunet.net/saavutettavuus/wp-content/uploads/2013/05/WCAG2Papu-tarkistuslista_heuristiikat_15052013.pdf

⁶ <http://papunet.net/saavutettavuus/helppokayttoisen-verkkopalvelun-heuristiikat-wcag-2-0-arvion-tueksi/>



Arvioitavien järjestelmien toteutuksesta riippuen yksittäisiä kriteereitä ei välttämättä ole mahdollista arvioida, minkä johdosta arvioijat voivat jättää yksittäisen kriteerin arvioimatta. Arviointitulos muodostuu arvioituista kriteereistä saadun keskiarvopistemäärän prosenttiosuudesta suhteessa arvioitujen kriteerien maksimipistemäärään.

Esimerkki tuloslaskelmasta: Jos arvioitavan portaalin A pistemäärä on 310/480 pistettä, on suhteutettu pistemäärä $(310/480) \cdot 100$ eli 64.6. Vastaavasti jos arvioitavan portaalin B pistemäärä on 310/470 pistettä, on suhteutettu pistemäärä $(310/470) \cdot 100$ eli 66.

Esimerkki tilanteesta, jossa kriteeri jätetään arvioimatta: Järjestelmässä on esitetty ohje videona käyttäjille, jolloin sitä koskevat tietyt WCAG 2.0 –kriteeristön alakohdat (esimerkiksi. 1.2.1, 1.2.3). Jos taas ohje on esitetty tekstimuotoisena videon sijaan pätevät sen arviointiin eri alakohdat, jolloin edellä mainitut alakohdat 1.2.1 ja 1.2.3 jätetään arvioimatta.

Liite 1: Käytettävyyssarviointi: tilastollisten merkitsevyyksien tarkastelu

Tässä liitteessä on esitetty käytettävyyssarviointiin liittyvien tilastollisten merkitsevyyksien tarkastelu. Tarkastelussa hyödynnettävät kaavat on kuvattu kirjassa "*Quantifying the user experience – Practical statistics for user research*"⁷.

Binäärimuuttujiin perustuvat tulokset

Koskee käytettävyyden aihealueita:

- Tuloksellisuus: Onnistuneesti suoritettujen testitehtävien %-osuus (paritesti / käytettävyystesti)
- Tyytyväisyys: Käytön aikaiset positiiviset / negatiiviset arviot (paritesti / käytettävyystesti)

Ero testataan McNemarin testillä (*McNemar exact test*), jos samat käyttäjät arvioivat molemmat tuotteet. Vastaavasti mikäli samat käyttäjät eivät arvioi molempia tuotteita, käytetään N-1 otoksen suhteellisen osuuden testiä (*N-1 two-proportion test*).

Jos tilastollisen testin lopputuloksesta laskettava p-arvo on pienempi kuin 0,1 on tuloksilla tilastollisesti merkitsevä ero.

Jatkuviin muuttujiin perustuvat tulokset

Koskee käytettävyyden aihealueita:

- Virheettömyys: Testitehtävien suorituksen aikaiset virheet (paritesti / käytettävyystesti)
- Opittavuus: Käyttäjien vastaukset loppukyselyn opittavuuskysymyksiin (paritesti / käytettävyystesti)
- Tyytyväisyys: Käyttäjien vastaukset loppukyselyyn (paritesti / käytettävyystesti)

Ero testataan riippuvien otosten t-testillä (*paired t*), jos samat käyttäjät arvioivat molemmat tuotteet. Vastaavasti mikäli samat käyttäjät eivät arvioi molempia tuotteita, käytetään kahden riippumattoman otoksen t-testiä (*2-sample t*).

Jos tilastollisen testin lopputuloksesta laskettava p-arvo on pienempi kuin 0,1 on tuloksilla tilastollisesti merkitsevä ero.

⁷ Sauro, J. Lewis, J. R. (2012) Quantifying the user experience – Practical statistics for user research. Elsevier, Morgan Kaufman, MA, USA, s. 94-95.

Menetelmä		TERVEYDENHUOLTO						55	SOSIAALIHUOLTO					25	PORTAALI					20
Mittari		Tulos	Pisteet	Maks. pisteet	Skaalatut pisteet	Skaalatut maks. pisteet	Osapaino-arvo	Tulos	Pisteet	Maks. pisteet	Skaalatut pisteet	Skaalatut maks. pisteet	Osapaino-arvo	Tulos	Pisteet	Maks. pisteet	Skaalatut pisteet	Skaalatut maks. pisteet	Osapaino-arvo	
Tuloksellisuus							30						30						25	
Paritesti / käytettävyydesti	Onnistuneesti suoritettujen testitehtävien %-osuus	75 (maks. 100%)	3,00	5,00	9,90	16,50	100	29 (maks. 100%)	1,00	5,00	1,50	7,50	100	80,5 (maks. 100%)	4,00	5,00	4,00	5,00	100	
Virheettömyys							20						20						10	
Paritesti / käytettävyydesti	Testitehtävien suorituksen aikaiset virheet	4+4+3 (maks. 15p.)	11,00	15,00	8,07	11,00	100	2+1 (maks. 10p.)	3,00	10,00	1,50	5,00	100	4 (maks. 5p.)	4,00	5,00	1,60	2,00	100	
Tehokkuus ja opittavuus							20						20						10	
Paritesti / käytettävyydesti	Tehokkuus: Oikeiden polkujen pituus askeleina	12 (maks. 18p.)	12,00	18,00	1,83	2,75	25	4 (maks. 12p.)	4,00	12,00	0,42	1,25	25	3 (maks. 5p.)	3,00	5,00	0,84	1,40	70	
	Opittavuus: Käyttäjien vastaukset loppukyselyyn	3,2 (maks. 5p.)	3,20	5,00	1,06	1,65	15	2,0 (maks. 5p.)	2,00	5,00	0,30	0,75	15	4,1 (maks. 5p.)	4,10	5,00	0,49	0,60	30	
Ryhmäarviointi	Tehokkuus: Käytettävyyssiantuntijan arvio käytön tehokkuudesta nähtyyn	3,4 (maks. 5p.)	3,40	5,00	1,50	2,20	20	1,7 (maks. 5p.)	1,70	5,00	0,34	1,00	20							
	Opittavuus: Käytettävyyssiantuntijan arvio opittavuudesta ammattilaisiantuntijoiden	2,9 (maks. 5p.)	2,90	5,00	1,28	2,20	20	1,5 (maks. 5p.)	1,50	5,00	0,30	1,00	20							
	Opittavuus ja tehokkuus: Käytettävyyssiantuntijan arvio mukautettavuuden opittavuudesta/tehokkuudesta	4,0 (maks. 5p.)	4,00	5,00	1,76	2,20	20	2,1 (maks. 5p.)	2,10	5,00	0,42	1,00	20							
Tyytyväisyys							20						20						25	
Paritesti / käytettävyydesti	Käytön aikaiset positiiviset / negatiiviset arvot	51 (maks. 100%)	3,00	5,00	1,98	3,30	30	20 (maks. 100%)	1,00	5,00	0,30	1,50	30	70 (maks. 100%)	4,00	5,00	1,60	2,00	40	
	Käyttäjien vastaukset loppukyselyyn	65,2 (maks. 100p.)	3,00	5,00	1,32	2,20	20	53,0 (maks. 100p.)	2,00	5,00	0,40	1,00	20	75,0 (maks. 100p.)	4,00	5,00	1,60	2,00	40	
	Käyttäjien määrittämä paremmuusjärjestys													65 (maks. 100%)	65,00	100,00	0,65	1,00	20	
Ryhmäarviointi	Ammattilaisiantuntijoiden vastaukset loppukyselyyn	18,5 (maks. 30p.)	18,50	30,00	3,39	5,50	50	10 (maks. 30p.)	10,00	30,00	0,83	2,50	50							
Käyttööliittymäsuunnittelun onnistuneisuus							10						10						0	
Ryhmäarviointi	Käytettävyyssiantuntijan arvio suunniteluheuristikkoihin ja käyttöön keskusteluun perustuen	3,4 (maks. 5p.)	3,40	5,00	3,74	5,50	100	1,1 (maks. 5p.)	1,10	5,00	0,55	2,50	100							
Esteettömyys							0						0						30	
Esteettömyys-arviointi	Esteettömyyden asiantuntijoiden arvio esteettömyyskriteeristöihin perustuen													310/480	64,60	100,00	3,88	6,00	100	
Kokonaispisteet yhteensä:		52,81			Yhteensä	35,82	55			Yhteensä	6,86	25				Yhteensä	10,13	20,00		
Oikeiden polkujen maksimipisteet riippuvat arvioitavien tehtävien lukumäärästä. Tässä laskelmassa esitettyinä esimerkkinumerit.																				