

TWiki > KitWiki Web > FinnishActivities > FinnishLanguageResourceCentre > FinClarin20080527Suunnittelukokous (2011-11-14, KimmoKoskenniemi)

FiLRC: FIN-CLARIN suunnittelukokous 27.5.2008

- Kimmo Koskenniemi, Martti Vainio, Antti Arppe (HY)
- Pirkko Nuolijärvi (Kotus)
- (Pirjo-Leena Forsström ei pääse, eikä sijainenkaan) (CSC)
- Jussi Niemi ([JoY](#))
- Jaana Kekäläinen ([TaY](#))
- Marketta Harju-Autti (OY)
- Ari Huhta (JY)

Kokouksen päätökset ja omaksutut tehtävät

1. FIN-CLARIN Organisaation vakiinnuttaminen ja täydentäminen, johtoryhmän perustaminen

- Vahvistettava konsortion kunkin nykyisen organisaation edustaja (yliopistoissa esim. neuvottelemalla asianomaisten tutkijoiden ja laitosten kanssa sekä informoimalla niiden tiedekuntien dekaaneja, joille kieliresursseilla on merkitystä) -- **HY, TaY, JoY, OY**
- Jyväskylän yliopiston liittäminen konsortioon. Konsortiosopimuksen virallinen rehtorin allekirjoitus sekä edustajan nimeäminen kuten yllä. Nykyisille konsortion jäsenten keskushallinnoille tämä tiedoksi. **JY**
- Mahdolliset yhteistyötahot, olennaiset toimijat saatava mukaan joko konsortion jäseniksi tai muuten:
 - TY suomen kielen laitos, SKS -- **Pirkko Nuolijärvi**
 - Svenska Litteratursällskapet/Fred Karlsson -- **Kimmo Koskenniemi**
 - TSK ja VY (terminologiat)
- Johtoryhmän (tai yhteistyötahon) edustajan tehtävänä on tiedottaa organisaationsa sisällä FIN-CLARINin toiminnasta organisaationsa sisällä sekä välittää organisaationsa toiveita ja tarpeita johtoryhmälle. -- **kaikki**
- Johtoryhmän ja yhteistyötahojen edustajien videoneuvottelukokous tiistaina 17.6.2008 -- **kaikki**

Aikataulu: ennen juhannusta

2. Organisaatioiden ja toiminnan yleiskuvaus ja konsortion vuosikertomus

- Tiivis kuvailu organisaation kieliresursseista ja toimijoista tarvitaan kaikista konsortion jäsenistä ja yhteistyötahoista. Laaditaan ensin malli. -- **HY** 6.6. mennessä - ks. luonnos [FinClarinSiteHY](#)
- Mallin pohjalta muiden organisaatioiden edustajat selvittävät oman organisaationsa osalta vastaavat seikat ja laativat KITWIKlin tiiviin englanninkielisen selostuksen omasta organisaatioistaan. -- **kaikki** juhannukseen mennessä
- Konsortion on varauduttava tekemään vuosittain vuosikertomus. -- **HY**

3. FIN-CLARINin toimintasuunnitelma ja BLARK

- Toimintasuunnitelma kansallisen kieliresurssi-infrastruktuurin rakentamiseksi vuonna 2009 ja sitä seuraavina kahtena viisivuotiskautena. Tarvitaan tiedot nykyisistä kieliresursseista (aineistot, kielimoduulit, sovellukset) konsortion osapuolilta (ks. ed.). Infrastruktuurin toteuttamissuunnitelman alustava versio -- **HY** elokuun loppuun mennessä
- Suunnitelman edellyttämät haastattelut. -- **HY** syksy 2008

4. CLARINin jäsenyys

- Niiden konsortion jäsenten, jotka eivät ole liittyneet CLARINin jäseniksi, tulisi liittyä sellaiseksi ottamalla yhteyttä CLARINIin, esim. steven.krauwer (ät) let.uu.nl. -- **OY, JY** ensi tilassa

5. Kokoustekniikka

- a. Verkkokokousten harjoittelu, Henna-Riikka Laitinen ottaa yhteyttä testausta varten ja ohjeistaa henna-riikka.laitinen(ät)helsinki.fi ja järjestää testitilaisuuksia. -- **HY**
- b. Johtoryhmän seuraava kokous 17.6. klo 9.00 alkaen verkkokokouksena

6. Muut yhteistyötahot

- a. Lista mahdollisista muista yhteistyötahoista yhteydenottoa varten Yle, muut mediayhtiöt

Kokouksen taustamateriaali

CLARIN lyhyesti

EU:n CLARIN, FIN-CLARIN ja paikalliset CLARIN-hankkeet

- EU:n CLARIN on komission rahoittama kolmivuotinen valmisteluvaiheen hanke, jonka tehtävänä on luoda perusta rakennettavalle kieliresurssi-infrastruktuurille. CLARIN on yksi 34:stä ESFRIn tiekartalla olevasta infrastruktuurista.
- OPM:n HY:lle myöntämä CLARINia tukeva rahoitus (200 k[€] v. 2008, ehkä 300 k[€] v. 2009) sekä kansallinen Kieliresurssikonsortio (HY, CSC, Kotus, [TaY](#), [JoY](#), OY).
- Yliopistojen ja organisaatioiden omat CLARINIin liittyvät hankkeet

Taustamateriaalia

- [Tiivis esitys CLARINista ja Suomen osuudesta siinä](#) - EU:n CLARIN, Suomen FIN-CLARIN (Kieliresurssikonsortio)
- [EU:n laajuisen CLARINin sivusto](#)
- [CLARIN Newsletter](#): **CLARINin tiedotuslehden** ensimmäinen numero
- Paikalliset hankkeet, kuten HY:ssa [HFST](#) ja [OMOR](#).

Mitä CLARIN on

- Nykyisellään Euroopan kieliresurssit ovat pirstoutuneena eri puolille, eikä yhteiskäyttö juurikaan ole mahdollista. CLARINin tavoitteena on luoda pohja toimivalle yhteiselle käytölle. Toimimalla Suomessa osana CLARINia varmistetaan Euroopan-laajuinen resurssien saatavuus ja tekninen yhteensopivuus.
- CLARINin FIN-CLARINin avulla pyritään varmistamaan tuleva kansallinen ja kansainvälinen rahoitus mm. pyrkimällä valtakunnalliselle infrastruktuurien tiekartalle

Tehtävä: Suunnitelma FIN-CLARIN -konsortion nimissä

- **Organisoituminen:** Kieliresurssikonsortion tulisi järjestäytyä siten, että sillä olisi ainakin johtoryhmä, joka voi toimia sen nimissä, sillä muuten FIN-CLARIN ei kelpaa kansallisten tutkimusinfrastruktuurien tiekartalle. FIN-CLARINin tulee järjestäytyä siten, että kukin osapuoli ottaa tiettyä vastuuta oman organisaationsa piirissä tapahtuvasta relevantista toiminnasta, erityisesti sen raportoinnista. Osapuolten tulisi olla myös CLARINin jäseniä (kuten CSC, HY, [JoY](#), Kotus ja [TaY](#) jo ovat). Myös OPM:n ja EU:n komission rahalla palkattujen henkilöiden keskinäinen työnjako tehtävä selväksi ja tiedotettava tarvittaville tahoille.
- Laadittava kattava **selvitys kieliresurssien nykytilanteesta, puutteista ja tarpeista**
- Laadittava konsortion yhteisesti hyväksymä **suunnitelma kieliresurssi-infrastruktuurin toteuttamiseksi** sekä siihen liittyvät prioriteetit, aikataulut, ja kustannusarviot (tavoitteena myös varmistaa FIN-CLARINin pääsy kansalliselle tiekartalle).
- Vaikutettava rahoittajiin (Akademian, Tekes jne), esim. säännöiksi rahoituksen ja tuotettavien resurssien mahdollisimman avoimen saatavuuden takaamiseksi.
- Huolehdittava siitä, että käyttäjille järjestetään tarvittavaa koulutusta.

Suunnitelmanmukaiset kieliresurssit

- **Löydettävyy**s ja saavutettavuus, s.o. käyttäjän pitää pystyä tunnistamaan, onko toivotunlaisia aineistoja tai ohjelmia saatavilla ja pystyä paikallistamaan, missä sellaiset ovat julkisten metatietojen perusteella (CLARIN WP2).
- **Oikeudet ja lisenssit** (IPR), s.o. käyttäjän pitää pystyä hankkimaan oikeus aineistojen tai ohjelmien käyttöön selkeästi määritellyllä tavalla ja vastavuoroisesti oikeuksien myöntämisen pitää sujua mahdollisimman vähällä vaivalla, mutta silti luotettavasti (CLARIN WP7).
- **Käytettävyy**s, s.o. aineistojen tulee olla sovittujen normien ja standardien muodossa ja ohjelmien pitää olla yhteensopivia tällaisten standardien kanssa, jotta niitä voidaan soveltaa aineistoihin ja yhdistellä keskenään (CLARIN WP5).
- **Autentikointi** ja **auktorisointi** luvan saaneiden käyttäjien ja vain heidän tulle saada käyttää resursseja (CLARIN WP2 ja WP7).
- **Tekniikka** eli tietokonekeskukset, jotka toimivat yhteen eräänlaisena federaationa, jolloin käyttäjälle ei ole väliä, missä keskuksessa aineistot tai ohjelmat (alunperin) ovat tai olivat (CLARIN WP2).
- **Käyttöliittymät** eli tietokoneohjelmia, joilla aineistoja voi esim. selata, valikoida niistä esimerkkejä ym. (Vrt. esim. [venäläinen Integrum -järjestelmä](#))

BLARK ja kieliresurssisuunnitelma

Basic Language Resource Kit (BLARK) pyrkii määrittelemään mitä tarvitaan, jotta paikallista kieltä voitaisiin käyttää nykyisessä ja tulevassa maailmassa tasavertaisena englannin tai muiden ylikansallisten kielten kanssa. Ks. artikkeli [A Field Survey for Establishing Priorities in the Development of HLT Resources](#) hollannin kielen BLARKista sekä muuta aiempaa materiaalia [BLARK](#).

BLARK perustuu osiin:

- **Sovellukset** joita tavalliset kansalaiset käyttävät ja joihin moduuleja tarvitaan
- **Moduulit** eli komponentteina toimivat tietokoneohjelmat, joiden rakentamiseksi puolestaan aineistoja tarvitaan.
- **Aineistot** (tekstikorpuksset, puhekorpuksset, leksikot, annotoidut aineistot, jne.)
- Sovellusten, moduulien ja aineistojen väliset **riippuvuudet**, jotka vaikuttavat siihen, missä järjestyksessä ja tahdissa aineistoja, moduuleja ja sovelluksia ylipäättänsä on tarkoituksenmukaista koota ja rakentaa.
- **Toimintasuunnitelma**, jossa vaiheistetaan, rajataan ja mitoitetaan tehtävät osa sekä varustetaan ne hinta-arviolla.

CLARIN näkee BLARKin **laajemmin ja humanistisemmin** kuin pelkästään kutakin kieltä tukevinä kieliteknologisina resursseina.

Aikataulu vuodelle 2008

- Kesä - elokuu: verkoston luonti
- Heinä - elokuu: resurssiselvitys yhdessä CLARINin WP5:n (Tübingen) kanssa
- Syyskuu: asiantuntijahaastattelut
- Loka - joulukuu: suunnitelman laadinta (projektisuunnitelma); 1-2 iteraatiota FIN-CLARINin sisällä
- Tammikuu 2009: suunnitelman julkistus

FIN-CLARIN tuleva toiminta

- Kokoukset, videoneuvottelut (matkustamisen vähentämiseksi)
- Tiiviit kuvailut kustakin yliopistosta ja osapuolesta: mitä resursseja, mitä toimintaa ...
- Seminaari vuoden vaihteessa (suunnitelman julkistus)
- Konsortion täydentäminen: JY, SKS, TY, Svenska litteratursällskapet, VY

Liitännät muihin infrastruktuureihin tai suuriin hankkeisiin

- Kirjastot
- YLEn ja muiden media-yhtiöiden arkistot
- Muiden alojen teknologiset ja yhteiskuntatieteelliset resurssit

-- [Martti Vainio](#) - 26 May 2008

Liitteitä

- [tavHanke-2.pdf](#): OPM:lle toimitettu suunnitelma CLARINin täydentävästä rahoituksesta 2009
- [a-field-survey-for.pdf](#): Artikkelin hollanninkielinen BLARKista: A Field Survey for Establishing Priorities in the Development of HLT Resources

I	Attachment	Action	Size	Date	Who	Comment
	a-field-survey-for.pdf	manage	70.2 K	2008-05-26 - 14:19	KimmoKoskenniemi	Artikkeli hollanninkielinen BLARKista: A Field Survey for Establishing Priorities in the Development of HLT Resources
	tavHanke-2.pdf	manage	22.5 K	2008-05-26 - 13:54	KimmoKoskenniemi	OPM:lle toimitettu suunnitelma CLARINin täydentävästä rahoituksesta 2009

Topic revision: r11 - 2011-11-14 - [KimmoKoskenniemi](#)

Copyright © 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideat, pyynnöt, ongelmat TWiki? [Send feedback](#)

