

FiLRC: FIN-CLARIN-johtoryhmän videoneuvottelukokous

ma 26.10.2009 klo 10~11.30

PÖYTÄKIRJA

Kokous alkoi klo 10.11.

Osanottajat:

HY: Kimmo Koskenniemi (pj), Antti Arppe, Mirja Nissinen (siht.), Jussi Piitulainen

CSC:

JY: Ari Huhta

JoY:

Kotus: Pirkko Nuolijärvi

OY: Marketta Harju-Autti

TaY: Jaana Kekäläinen, Mikhail Mikhailov

1. Yleistä - ilmoitusasioita

Kuulumisia CLARIN-projektista (KK):

[NEERI](#): Networking Event for European Research Infrastructures + Standards Workshop pidettiin Helsingissä 30.9.-2.11.2009.

Lisätietoa, konferenssiesitykset ja kuvia konferenssista löytyy [NEERI-nettisivulta](#).

EU-CLARINin johtoryhmän kokouksia: Kööpenhamina 22.-23.10.2009

[ERIC](#) (European Research Infrastructure Consortium) - EU-CLARIN pyrkii ERICiksi; ERICin osapuolia ovat valtiot; ERICin tehtävänä on pitää huolta siitä, että keskuksat (kuten CSC, MPI) ovat yhteensopivia; pyritään siihen, että saataisiin Suomi mukaan tähän infrastruktuuriin > aiesopimus; ERICin aikataulu: CLARIN haluaa edetä niin nopeasti kuin mahdollista, keskustelut aloitettu, ensi vuoden aikana (kesällä) lopullisiin ratkaisuihin?

2. FIN-CLARINin kansallisen rahoituksen tilannekatsaus ja lähitulevaisuus (KK)

a. Kieliaineistojen tiekartan toteuttaminen (KK:n pohjustus)

Helsingin yliopisto (yleisen kielitieteen laitos) olisi toteuttamista koordinoiva instanssi < HY anoo valtakunnalliseen tehtävään rahoitusta OPM:ltä - ei muuta tapaa anoa rahaa; FIN-CLARINin valmisteluvaiheen loppuunsaattamiseen (palkkarahaa ja loput nk. ulkoa hankittaviin töihin eli kaikkea ei käytetä HY:ssa)

tiekartan mukaisia resursseja kehitettäisiin niissä yliopistoissa, joilla on siihen parhaat edellytykset

FIN-CLARIN-johtoryhmä priorisoisi ja ohjaisi tiekartan komponenttien toteuttamista

kaikkien kannattaisi omissa tulosneuvotteluissaan mainita aina FIN-CLARIN ja viitata myös HY:n tilanteeseen

b. FIN-CLARINin virallisempi järjestäytyminen

kirjallinen johtosääntö konsortion toiminnalle (mistä HY/koordinoiva taho vs. FIN-CLARIN-johtoryhmä vs. osasia suorittavat yliopistot vastaavat; kuinka päätöksiä tehdään)

kohdan 1 periaatteet

vastuut, toimintasuunnitelma, toimintakertomukset vuosittain

Periaatteet kannattaa lyödä lukkoon, vaikka pitempiaikainen jatkorahoitus vielä epävarma

Järjestäytymisestä hyötyä myös tulevassa rahanjaossa

AA ehdottaa, että HY:ssa tehdään johtosääntö luonnos, joka tulee johtoryhmän käsittelyyn ja hyväksyttäväksi

Kotus voi tehdä tässä yhteistyötä HY:n kanssa

Johtoryhmän seuraavaan kokoukseen mennessä jonkinlainen luonnos (parin kk:n kuluessa) (AA ja MN)

3. FIN-CLARIN-autentikointi- ja auktorisointi-infrastruktuurin kehitystyön eteneminen CSC:llä (AA)

työ päästy aloittamaan hieman myöhässä mutta ollaan aikataulussa

koodaustyö on alkanut kesällä

järjestelmän betatestaus vuodenvaihteessa; johtoryhmän jäsenet voisivat osallistua testaukseen

järjestelmän ensimmäisen version pitäisi olla valmis Suomen osalta maaliskuussa

järjestelmää pääsee käyttämään oman kotiyliopiston tunnuksilla

EU-tasolla vireillä yhteistyöprojekti: pari keskusta Hollannista ja pari Saksasta + CSC - tarkoitus tarjota näiden kolmen maan

kielentutkijoille pääsy toisen maan aineistoihin > laajennetaan myöhemmin, nyt tekeillä prototyyppi

HY osallistuu kustannuksiin

4. FinWordNetin ja suomenkielisen puupankin prototyypin kehittämisprojektit Helsingin yliopistolla Krister Lindénin johdolla loppuvuodesta 2009 (KK/AA)

[WordNet](#): elektroninen synonyymisanakirja, käytetty myös esim. semanttiseen luokitteluun

ensin käännetään englanninkielinen versio suomeksi, sitten tehdään kakkosversio, jossa otetaan huomioon myös mm.

kulttuurisidonaisuus > lopputulos pyrkii olemaan oikeasti suomenkielinen [WordNet](#)

KL valmistellut asiaa

1. vaihe noin kuudesosa koko työstä

tutkimuskäyttöön (kaupallisia versioita mm. Lingsoftilla ja Kielikoneella)

monikielinen tiedonhaku

jatkokehittämissä lisää yhteistyötä - nyt tekeillä prototyyppi

puupankki

prototyyppi

ISK:sta esimerkkivirkkeet

selvitetään ongelmia, mitä puupankin luomisessa voi tulla

jatkokehittämissä tässäkin lisätään yhteistyötä, vain prototyyppiä aloiteltu HY:ssa

5. Mahdollinen FIN-CLARIN-työpaja

kun kansallinen jatkorahoitus selviää

esim. Kotuksen tarpeita kieliaineistojen talletuksen suhteen voisi työstää yhteisessä työpajassa, missä voisi miettiä, miten CSC:llä luotavaa AA-infraa käytetään oikeasti; tämä voisi kiinnostaa muitakin

CLARIN-tiedotusta ja -markkinointia lisättävä: kokousten kautta, road show, muuta?

road show? < kutsuttaisiin humanististen aineiden edustajia katsomaan miten infrastruktuuri toimii; ensi keväänä; esim. kielten jatkokoulutusseminaarien yhteydessä; Kielitieteen päivät Helsingissä + Fonetiikan päivät Joensuussa (AA ja MN selvittävät asiaa), [AFinLAn syysseminarium](#) marraskuussa Tampereella (AH menossa, esittelee FIN-CLARINia; selvitetään posteritilanne ym.; AH ja AA)

työskentelyyn tähtäävä johtoryhmän kokous ja johtoryhmä levittäisi tätä tietoa myös ympärilleen

työpaja kevättalvella Helsingissä?

tiekarttasuunnitelman läpikäynti marras-joulukuun vaihteessa, viimeistään helmikuussa; matkoja voidaan maksaa OPM:n rahasta;

(AA ja MN selvittävät päivämäärää Doodlella); samalla voitaisiin katsoa johtosäännön luonnosta ja vilkaista hieman CSC:n AA-infraa

6. Konsortion mahdolliset uudet jäsenet tai yhteistyötahot

tilannekatsaus ja mahdolliset uutiset yhteydenottojen suhteen (PN, KK, AA)

Tilanne 22.9.2008 pidetyn kokouksen pöytäkirjan mukaan (tämän jälkeen pöytäkirjoissa ei merkintää uusista tiedoista; yhteydenottoja on jatkettu):

TY < Pirkko Nuolijärvi: yhteydet Kaisa Häkkiseen (Suomen kielen oppiaine) ja Anneli Pajuseen (TS-korpus); mutta ei vielä responssia alkukiinnostuksen jälkeen; seuraavaksi konsortio/aiesopimuksen lähettäminen asioiden edistämiseksi.

humanisteista ei tietoa; Tapio Salakoski ja Filip Ginter tietotekniikan puolella; lähetetään muistutus ja aiesopimus (MN)

VY < PN: yhteydet Merja Koskelaan (vararehtori) kielikylypaineistoista ja Nykysuomen laitokseen (Anita Nuopponen, Birgitta Romppainen & Esa Lehtinen); seuraavaksi aiesopimuksen lähettäminen.

lähetetään muistutus ja aiesopimus (MN)

SKS < PN; puhe statuksesta yhteistyökumppanina, jolloin aiesopimuksen allekirjoittamista ei tarvita.

Svenska Litteratursällskapet i Finland < Kimmo Koskenniemi: virallinen kirje yhteistyökumppaniksi kutsumisesta; tilanne edelleen kesken.

AA: Urpo Nikanne < Antti Arppe: yhteydenotto ruotsin kielen laitoksen (Svenska språket) johtajaan Saara Haapamäkeen.

OY: Giellagas-instituutti ja saamenkielisten aineistot mukana Oulun kautta FIN-CLARINissa; neuvottelevat YLE:n kanssa saamelaisradion aineistojen sisällyttämisestä

MH-A: Giellagas-instituutti tehnyt oman arkiston

jäsenet ja liitännäisjäsenet erikseen; kaikkien jäsenten ei tarvitse olla FIN-CLARINin jäseniä

Lisäksi:

TKK < Antti Arppe: TKK:lta Timo Honkela on ilmoittanut, että heidän liittymisensä on etenemässä.

TKK:n myötä myös Aalto tulossa mukaan

7. Muut esille tulevat asiat

Muita asioita ei tullut esille.

8. Seuraava kokous

työpaja marras-joulukuussa? (AA ja MN selvittävät ajankohtaa; ks. kohta 5.)

Tehtävälista

Mitä / kuka / milloin

1) Johtosäännön luonnos / AA ja MN / parin kk:n kuluessa eli seuraavaan kokoukseen mennessä

2) Road show'n osallistuminen Kielitieteen päiville toukokuussa Helsingissä + Fonetiikan päiville 25.-26.2.2010 Joensuussa / AA ja MN selvittävät asiaa / seuraavaan kokoukseen mennessä

3) Road show'n osallistuminen [AFinLAn syysseminariumiin](#) marraskuussa Tampereella / AH menossa, AH ja AA selvittävät posteritilanteen / pian

4) Konsortion mahdolliset jäsenet ja yhteistyötahot - muistutukset ja aiesopimukset / MN / lähiaikoina

5) Seuraavan kokouksen (työpajan) järjestäminen marras-joulukuussa / AA ja MN selvittävät päivämäärää Doodlella / pian

Kokous päättyi klo 12.07.

Liite:

Tiekarttasuunnitelma, joka aikanaan jätettiin:

Timetable

Years for construction/major upgrade: 2009-2020

Estimation of distribution of the costs/years:

2009: 425,000 €, 2010: 525,000 €, 2011: 500,000 €, 2012: 550,000 €, 2013: 450,000 €, 2014: 450,000 €, 2015-2020: 350,000 € per year

1. Technical construction of the CLARIN-compatible Authentication and Authorization Infrastructure for language resources in

Finland: 200,000 € (2009-2010),

2.

2a. web-based search and browsing interface: 100,000 € (2009-2010),

2b. linguistic analysis tools: 300,000 € (2009-2011), (aloitettu, ei käytetty FIN-CLARIN-rahaa)

2c. adaptation and development of annotation and editing tools: 100,000 € (2011-2012),

3. Basic Linguistic resources for the languages of Finland (missing components of BLARK):

3a. grammatically annotated text corpus (tree bank) of Finnish text: 1,000,000 € (2010-2020), (aloitettu alustavasti Kotuksen kanssa > annotoinnin mallikorpus, mukana kaikki kieliopilliset ilmiöt)

3b. comprehensive lexical database with synonym sets ([WordNet](#)) for Finnish: 300,000 € (2010-2013), (käynnistetty, pyydetty tarjouksia, muutamia kymmeniä tuhansia euroja ensimmäiseen vaiheeseen)

3c. conversion and upgrading of existing linguistic resources to conform CLARIN standards: 100,000 € (2009-2010),

3d. digitizing and integrating nationally important text collections, audio and video recordings into CLARIN: 1,900,000 € (2010-2020) and

4. project administration and contractual affairs: 1,000,000 € (2009-2015)

-- [MirjaNissinen](#) - 2009-10-26