

Liite 2 (UTUR): Turun yliopiston FIN-CLARIN-panostus

<https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/FinClarinkonsortio>

Turun yliopisto (jäljempänä jäsen) toimii FIN-CLARIN -konsortiossa panostaen FIN-CLARINin ja CLARIN ERICin perussäännön mukaisten tavoitteiden saavuttamiseksi seuraavaa:

- Jäsen toimii CLARIN ERICin teknillisen ja tieteellisen kuvauksen (CLARIN ERIC Technical and Scientific Description, July 2011) mukaisena R-tason CLARIN-keskuksena (Recognized Centre).
- Jäsenellä on seuraavassa taulukossa kohdissa 5 ja 6 lueteltuja olemassa olevia ja tekeillä olevia kielivaroja ja jäsen pyrkii saamaan kyseiset kielivarat FIN-CLARINin ja CLARIN ERICin käyttöön tekijänoikeuksien ym. sallimissa puitteissa ja tarjoaa FIN-CLARINille asiantuntemusta mm. alla kohdissa 5 ja 6 mainituilla alueilla.

Alla on yhteenveto organisaation liittyessä FIN-CLARINille ja CLARIN ERICille tarjottavien palvelujen, kielivarojen ja -työkalujen ym. keruun ja valmistamisen käytetty työ- tai rahamäärä (2006-2015).

	Panostuksen laji	Kuvaus	Htkk
1	Tieto- ja palvelukeskus	-	-
2	Autentikoinnin ja auktorisoinnin yp.	-	-
3	Autentikoinnin ja auktorisoinnin kt.	HAKA-tunnistuspalvelun käyttö	-
4	Asiantuntemuksen jak. CLARINille	-	-
5	Olemassa olevat kielivarat, työkalut ja palvelut	1. Englanninkielisen biolääketieteellisen tekstin syntaktisesti ja semanttisesti annotoitu korpus 2. Suomen yleiskielen puupankki 3. Suomen yleiskielen Propbank 4. Puupankietsintä työkalu SETS 5. Suomenkielinen syntaktinen jäsenin 6. Biomolekulaaristen vuorovaikutusten resurssi EVEX 7. Terveystiedon korpus 8. Lauseopin arkiston murrekorpus 9. Mikael Agricolan teosten morfosyntaktinen tietokanta 10. Satakuntalaisuus puheessa (Sapu): aineiston keruu ja litteraattien tekeminen 11. Suomenkielisten arkikeskustelujen morfosyntaktinen tietokanta (Arkisyn) 12. Edistyneiden suomenoppijoiden korpus (LAS2)	576
6	Tekeillä olevat kielivarat, työkalut ja palvelut	1. Suomenkielinen semanttinen jäsenin 2. Internet-suomen korpus Internet Parsebank 3. Suomenkielisten arkikeskustelujen morfosyntaktinen tietokanta (Arkisyn) 4. Edistyneiden suomenoppijoiden korpus (LAS2)	288
7	CLARIN-standardien ja -käytänteiden omaksuminen ja levittäminen	-	2
	Yhteensä		866

Alla on yhteenveto organisaation arvioidusta vuosittaisesta panostuksesta liittymisen jälkeen.

	Panostuksen laji	Kuvaus	Htkk
1	Tieto- ja palvelukeskus	-	-
2	Autentikoinnin ja auktorisoinnin yp.	-	-
3	Autentikoinnin ja auktorisoinnin kt.	HAKA-tunnistuspalvelu ja Language Bank Rights käyttö	-
4	Asiantuntemuksen jak. CLARINille	-	-
5	Olemassa olevat kielivarat, työkalut ja palvelut	1. Englanninkielisen biolääketieteellisen tekstin syntaktisesti ja semanttisesti annotoitu korpus 2. Suomen yleiskielen puupankki 3. Suomen yleiskielen Propbank 4. Puupankietsintä työkalu SETS 5. Suomenkielinen syntaktinen jäsenin 6. Biomolekulaaristen vuorovaikutusten resurssi EVEX 7. Terveystiedon korpus 8. Lauseopin arkiston murrekorpus 9. Mikael Agricolan teosten morfosyntaktinen tietokanta 10. Satakuntalaisuus puheessa (Sapu): aineiston keruu ja litteraattien tekeminen 11. Arkisyn (suomenkielisten arkikeskustelujen morfosyntaktinen tietokanta) 12. Edistyneiden suomenoppijoiden korpus (LAS2): aineiston keruu ja annotointi	19
6	Tekeillä olevat kielivarat, työkalut ja palvelut	1. Suomenkielinen semanttinen jäsenin 2. Internet-suomen korpus Internet Parsebank 3. Suomenkielisten arkikeskustelujen morfosyntaktinen tietokanta (Arkisyn) 4. Edistyneiden suomenoppijoiden korpus (LAS2): lisäaineiston keruu ja annotointi 5. Prosodiakorpus	84
7	CLARIN-standardien ja -käytänteiden omaksuminen ja levittäminen	-	0
	Yhteensä		103