

Andmebaasidega eesti kirjanormide lätetel

Peeter Tinitš
Tekstipäev
Tallinn, 07.11.2022

Keel ja tekstide kogum

Kerge 2003. Eesti kirjakeele mõtteruum. ESA

- “[K]irjakeele mõistesse [kuulub] paratamatult kirjalike tekstide traditsioon, mille kujunemisel mängivad kaasa kirjaoskuse levik, tekstide levi ja kättesaadavus keelealal ning nende ühisosa plaanid: kui laialt tuntakse samu tekste, kui suur on tekstide keele ja kirjaviisi sarnasus jne.”
- “Kui tavaliselt seotakse normi spontaanset kujunemist suulise keelega, siis siinkirjutaja arvates on see ka kirjaliku keelevormi üks tunnuseid ja arengukomponente. Spontaansuse tekitavad tekstid, milles norm avaldub ja kujuneb ning mille kaudu ta omandatakse.”

> Kirjakeele kujundavad tekstid, mida kirjutatakse, millest keelt õpitakse. Küsimused: kes kirjutab, mida loetakse, ...?

EESTI KIRJAKEELE MÕTTERUUM

KRISTA KERGE

Kirjakeelena käsitab autor eesti keelt kogu vaadeldaval ajal, mil sellel on olemas olnud kiri. Niisuguse vaateviisi juures pole kirjakeel haritkeel ega kodukeel vastand vms, vaid EESTI KEEL ON KIRJAKEEL, üks Euroopa pika ajalooga kirjakeeli (vt ka Kerge 2000), mille kõnelejate arv kipub kohapeal jääma alla miljoni. Meie keelel pole samas isegi 20. sajandil alati olnud võimalust täita kõiki funktsioone, see aga seab keelele piire, ahendab mõtte- ja väljendusruumi. Parimailgi aegadel pole eesti keelt emakeelena kasutajate arv küllalt suur, et kõiki ülesandeid vaevata täita ja hooleta keelt koos hoida. Selle veendumuse taustal mõtiskletakse allpool kirjakeele olemuse üle, oskamata ja püüdmata kuidagi ennustada tema tulevikku ning õpetamata, kuidas teda kasutada. Pühendan oma mõtiskelu Henn Saarile, kelle sünnist möödus hiljuti 80 aastat.¹

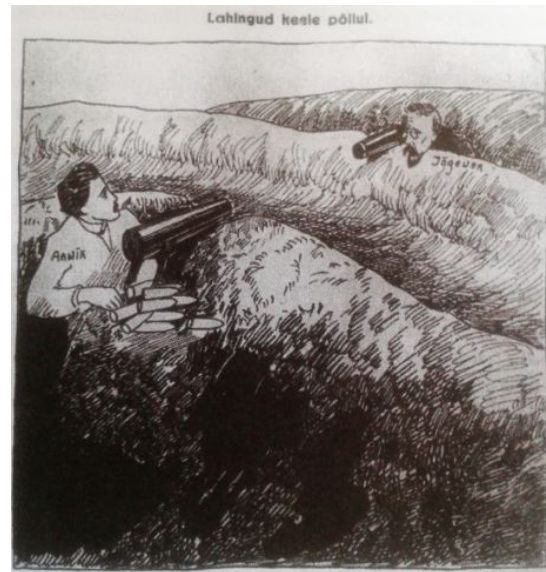
Tekstide ruum ja ajalooline sotsiolingvistika

Ajalooline sotsiolingvistika - püüab mõista keele ajalugu kasutajate kaudu.

- Kes kirjutas, kuidas, mis olukordades
- Sujuvad vs äkilised muutused
- Mitte-kesksed registrid jne

Küsimus: Aaviku põlvkonnas - kuidas nii äkitselt nii paljud muutused?

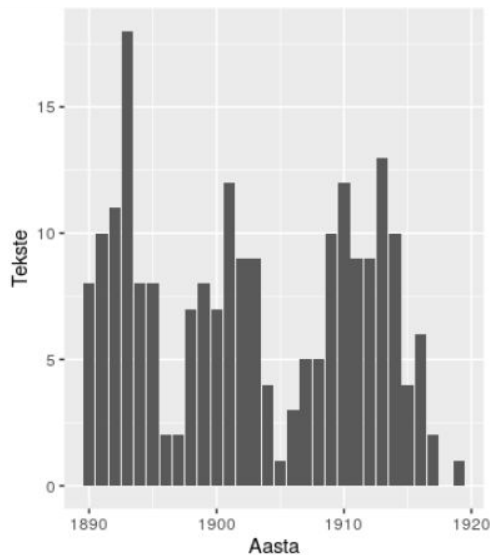
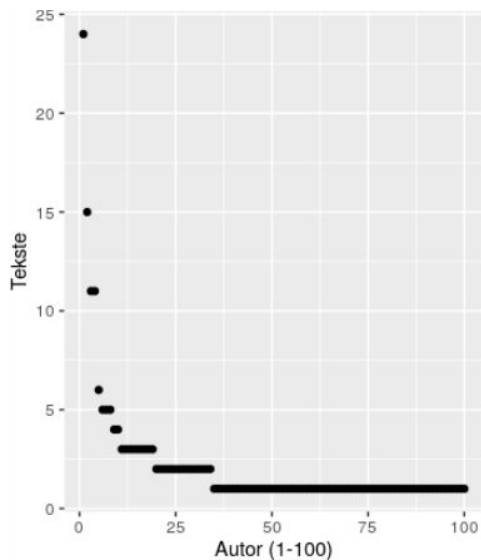
Kuidas uurida: tekstid, mida kirjutati, loeti ehk keelekorpused.



Lahingud keele põllul (Aavik ja Jõgeva) („Sädemed” nr 10, 1914.)

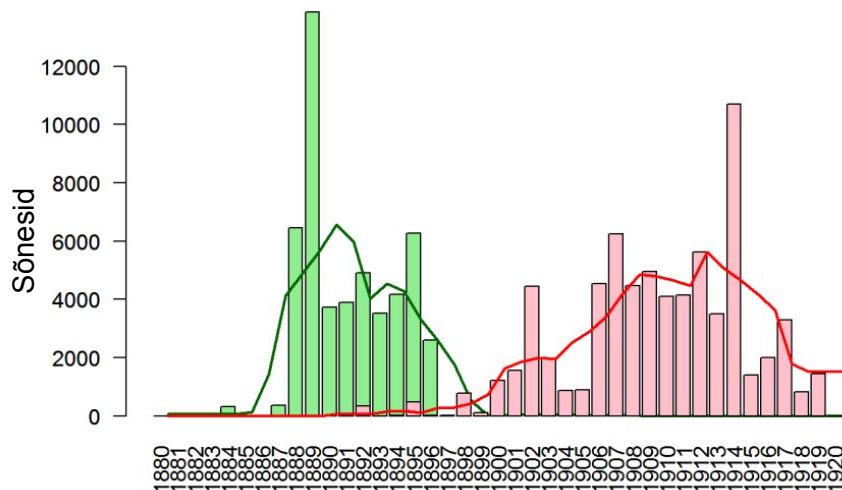
Korpuste loomine

Eesti kirjakeele korpus 1890-1990 (Hennoste et al. 1993)
100 autorit, 214 tekstikatkendit (enamikelt 1)



Uued korpused (Kirjandusmuuseumi arhiivist)

- 243 kirja 10lt autorilt (Aspe, Bergmann, Braks, Busch, Kruuspere, Küng, Lüüs, Oks, Speek, Suurkask)
- Hurda võrgustik: 269 autorit, 290 kirja



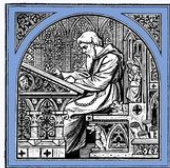
Digitud tekstid

Hulk tekste digiteeritud eri formaatides

- E-raamatud
- Isetehtud (Vikitekstid)
- Mäluasutuste kogud (KirMus kivike, TLÜAR ETERA jne)

Würst Gabriel ehk Piritä kloostri wiimsed päewad

Eduard Bornhöhe



Exported from Wikisource on 07/24/18

Eesti Raamatuportaal

Kõik kategooriad

Otsi raamatut, autorit, ISBN-i, kategooriat, toodet...

Raamatud E-Raamatud Mängud Kool ja kontor Kingitused Kodu ja aed Ilu ja stiil Muusika ja filmid Tehnika

Kultuuriministeeriumi toetusprogrammi EESTI KIRJANDUS e-raamatud

Kirjastuse raamatud

221 vastet

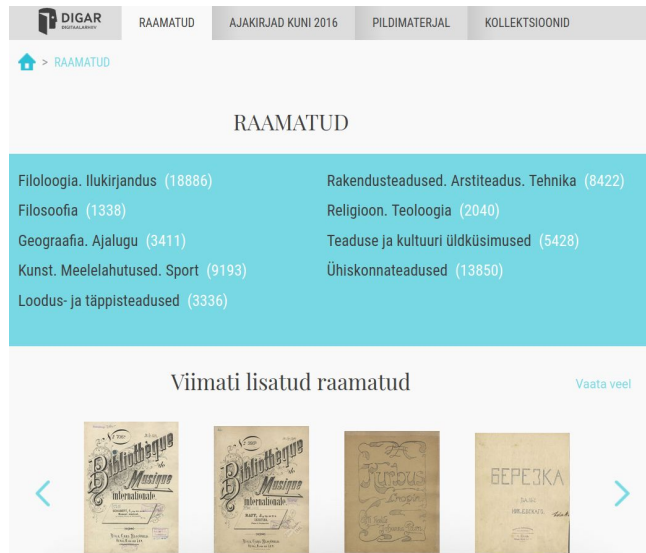
Populaarsed

Saadavus kauplustes	+				
Keel	-	E-RAAMAT: TASUJA: JUTUSTUS EESTIMA... 0,00 €	E-RAAMAT: HINGEDE OÖ 0,00 €	E-RAAMAT: KÜLALISTE LEIB 0,00 €	E-RAAMAT: LIBAHUNT 0,00 €
Hind	-				
Aed	+				
Ajalugu ja poliitika	+				
Aksesuaarid	+				
Albumid	+				
Arhitektuur ja sisekujundus	+				
Arvuti lisaseadmed	+				
Muusika	+				
Audioraamatud	+				

Digar ja rahvusraamatukogu aines

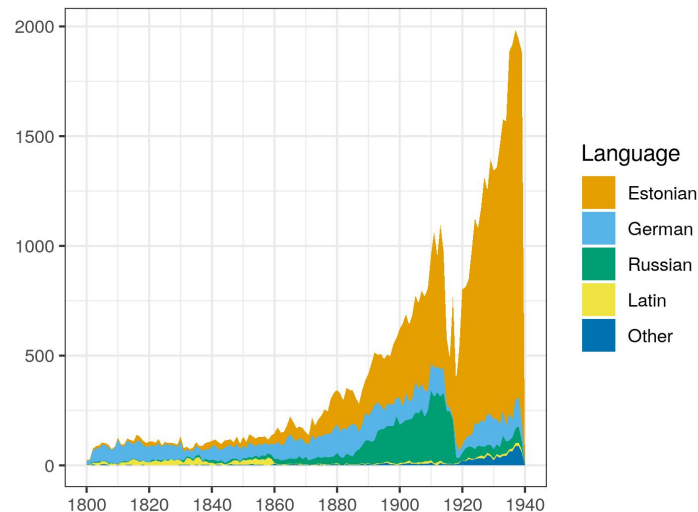
13k teost avatud teost 1800-2020

Eesti Rahvusbibliograafia 180k e.k. teost



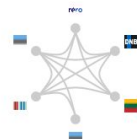
Digari teosed

Publications in Estonian National Bibliography



Rahvusbibliograafia aines

Kes kirjutas? (Sünnikoht, -aeg, amet jne)



[Artikkel](#)
[Arutelu](#)
[Vaata](#)
[Muuda](#)
[Muuda lähteteksti](#)
[Näita aialugu](#)
[Otsi Vikipeediast](#)

Julius Aamisepp

Julius Aamisepp (1. september (vkj) 20. august) 1883 Harjumaa, Kloostri vald, Karlepa, Karlepa-Tõnu talu – 19. jaanuar 1950 Jõgeva) oli eesti sordiaretaja.

Eesti biograafiline andmebaas ISIK

avalehekülg · nimekin · otsing · ametikohad [110] · asutused [470] · isikud [4486] · kaastöö väljaannetele [1133] · nime erikujud [925] · pseudonüümid [9179] · sugulased [309] · sünnikohad [625] · surmakohad [186] · tegevusala [599] · lühendid ja allikad · admin

AAMISEPP, JULIUS

Nimi			
Nime normitud kuju	Aamisepp, Julius		
Eludaatumid			
Sünniaeg	1883-09-01		
Sünnikoht	Karilepa-Tõnu t. Kloostri v. Harjumaa		
Surmaaeg	1950-01-19	100	1 2a Aam
Surmakoht	Jõgeva	375	2a 1 12
Markus	maetud Tartu Raadi kalmistule	500	1 20 (DE- Aamisepp /gnd#tar
Üldandmed			
Tegevusala	sordiaretaja		

[illegible]

001	ERRRja11124581 [VAF cluster]
003	ERRR
008	990322jadjnnnaabn ja
035	ja [ERRRja11124581]
040	ja ERRR ja est ic ERRR id ErTUR
043	Ic ee
100	1 ja Aamsepp, Julius, td 1883-1950.
400	1 ja Aamsepp, Julius.
500	1 ja Aamsepp, Jüny, td 1883-1950.
620	ja EE, 1 kd, 1985, EE, 14 kd, 2000.
680	1 Eesti põllumajandusteadlane, sõnaretaja, põllumajandusd (1947).

100	1	a	Aamissepp, Julius zw 1883-1960
375			
500	1	d	(DE-112) iso6128
548	1	i	(D-10101)1096201100 (D-588)1096201100 tw http://d-nb.info/gnd/1096201100 /gnd:Aamissepp, Iimar dz 1927-1994 d bezrt http://id-nb.info/standards/elementset/gnd#id-nb:infostandards/Relationship r v r1 Beziehung familiär Beziehung familiär v sO
548			1883-1960 d4 dat4 http://id-nb.info/standards/elementset/gnd#dateOfBirthAndDeath r v r1 Lebensdaten
548			
548	1	i	01.09.1883-19.01.1960 d4 dat4 http://id-nb.info/standards/elementset/gnd#dateOfBirthAndDeath r v r1 Exakte Lebensdaten
550	1	i	(D-10101)04203194X (D-588)4203194-1 tw http://id-nb.info/gnd/4203194-1 #Agnarwissenschaftler d4 berc http://id-nb.info/standards/elementset/gnd#professionOrOccupation r v r1 Charakteristischer Beruf
550	1	d	(D-10101)4779436 (D-588)4177943-6 tw http://id-nb.info/gnd/4177943-6 Revolutionär n4 beru http://id-nb.info/standards/elementset/gnd#professionOrOccupation r v r1 Beruf
561	1	i	Karielipa d4 ortq http://id-nb.info/standards/elementset/gnd#placeOfBirth r v r1 Geburtsort v sO Gernst. Läsane-Härrju

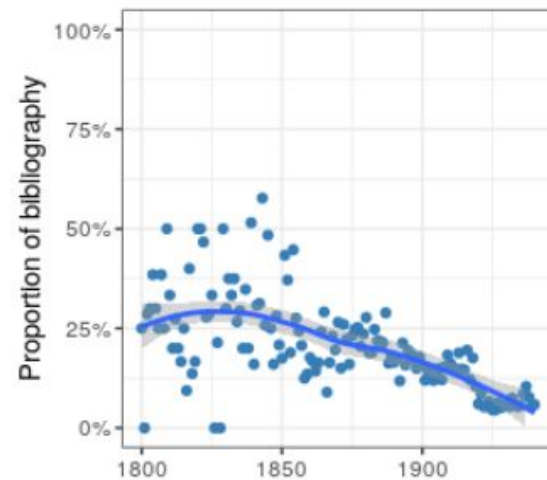
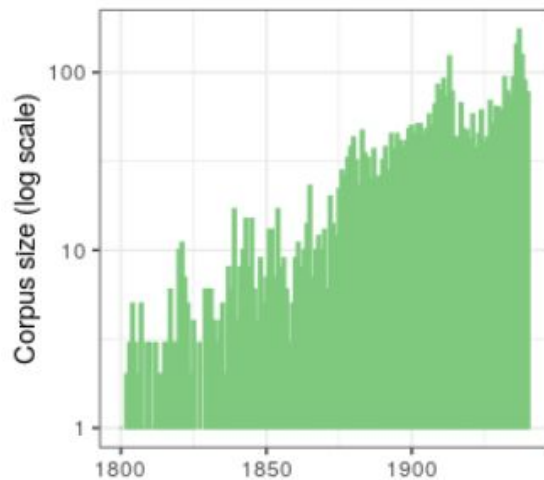


Korpus selle põhjal

Moodustatud tekstikogu 1800-1940 (4608 teksti)

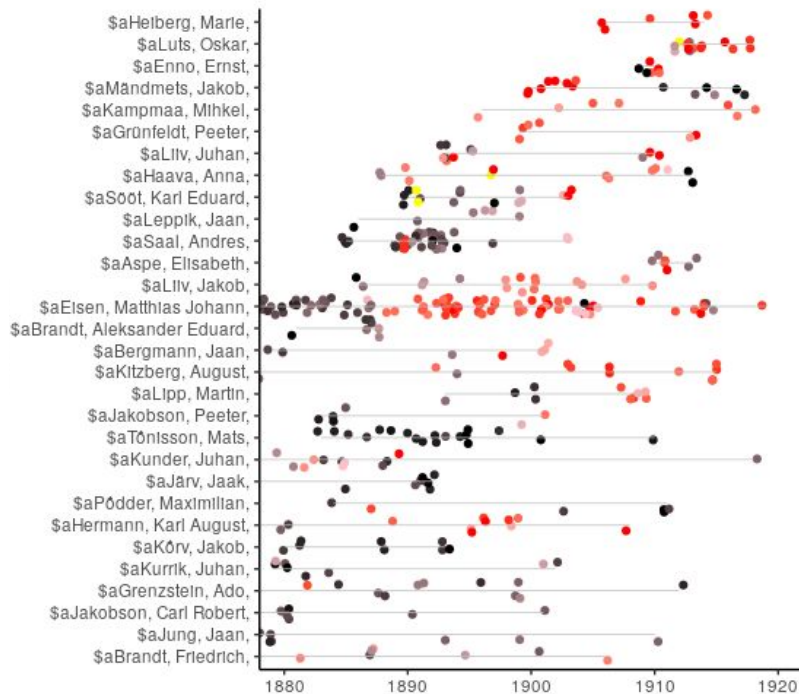
Metaandmed

- Eluloolised andmed (1k)
- Bibliograafia andmed (39k)
- Demograafia
- Murdeandmed

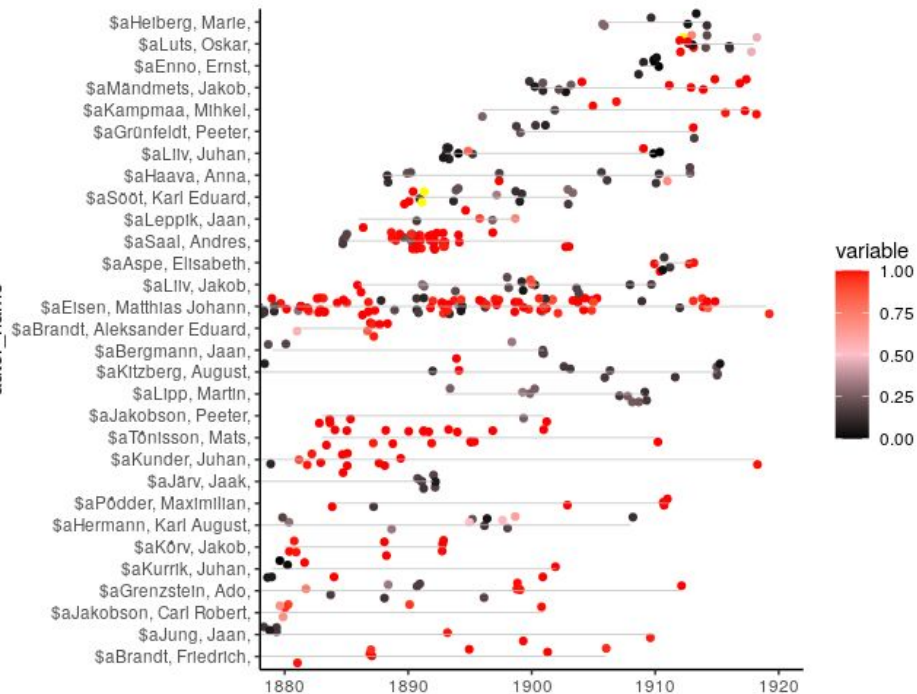


Autorite keelekasutus

Variation of ste_var



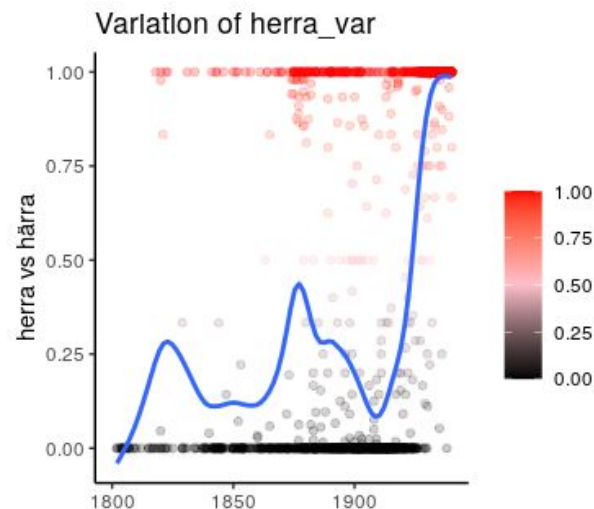
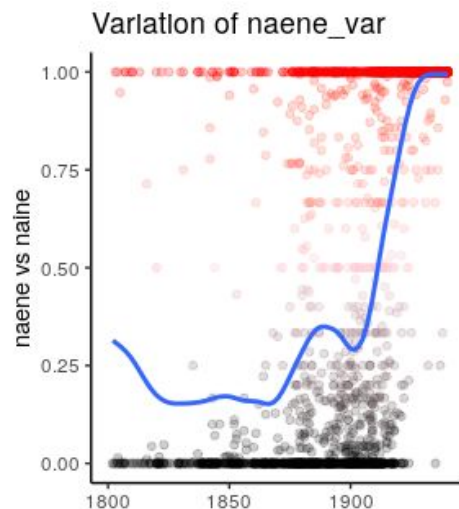
Variation of ää_var



Tekstide analüüs: keeleline muutuja (kirjakuju vs kirjakuju)

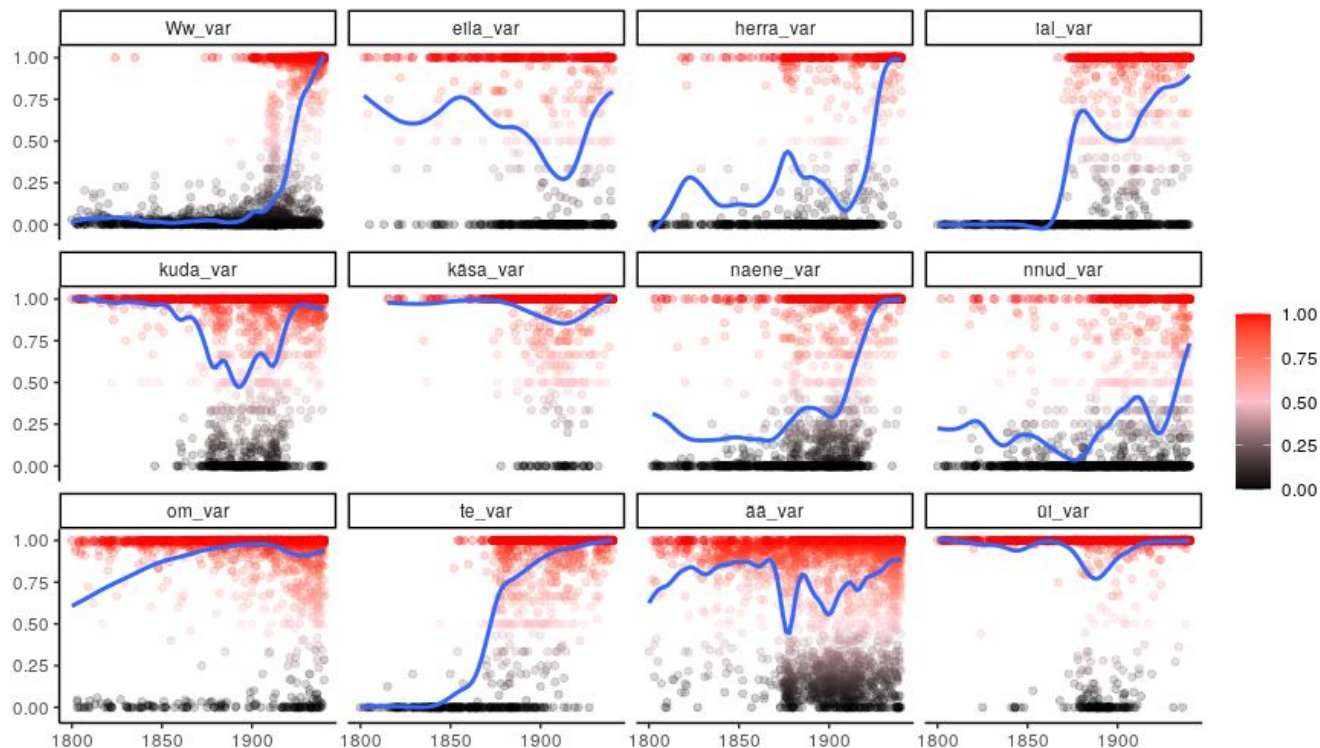
Keeleline muutuja:

- Konkureerivad keelekujud, mis tähendavad enamvähem sama (Labov 1966)
- naene - naine
- herra - härra
- hää/pää - hea/pea



Paljud keelelised muutujad

Mõned muutujad,
kus esines
huvitavat
variatiivsust
1900 a. paiku



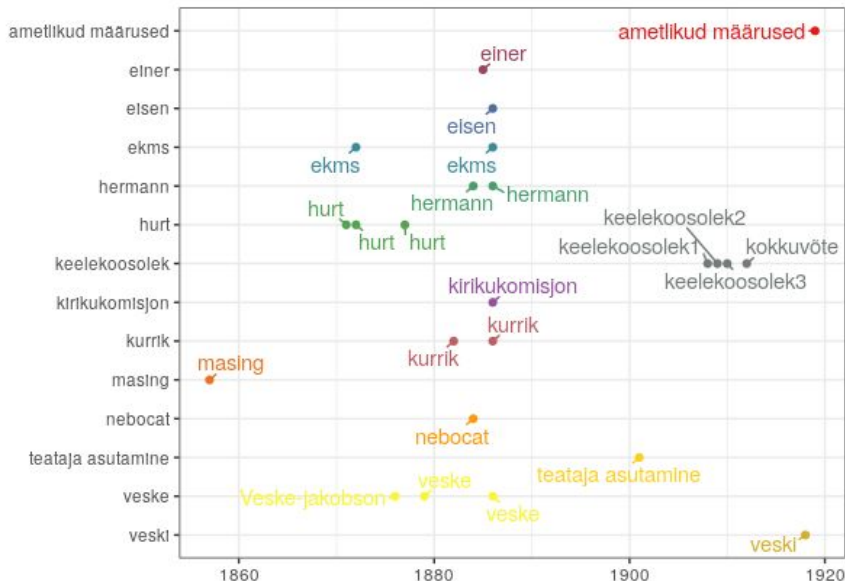
Soovitus keele kohta

Preskriptiivsed tööd (25 tööd)

- Sõnastikud ja komiteed
- Keelekonverentsid
- Ajaleheartiklid

Näiteks:

Palun kasutage “hää”, mitte “hea”.



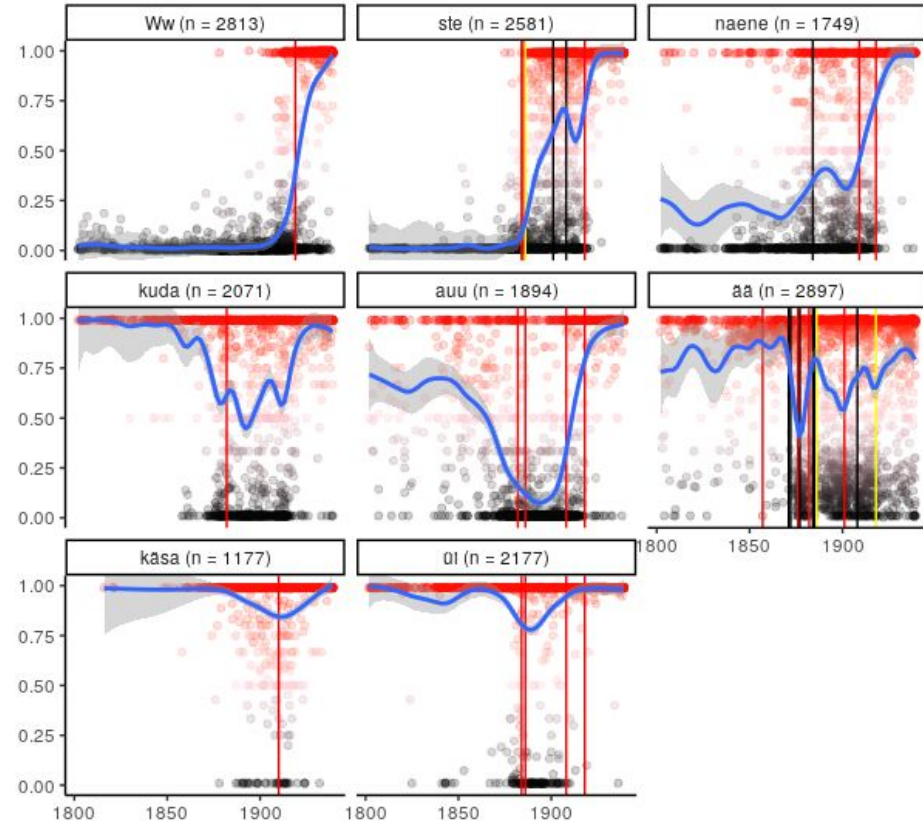
Ettekirjutuste mõju

Kombineerime andmeid variatiivsuse ja soovitude kohta

Kas ettekirjutused mõjutavad keele variatiivsust?

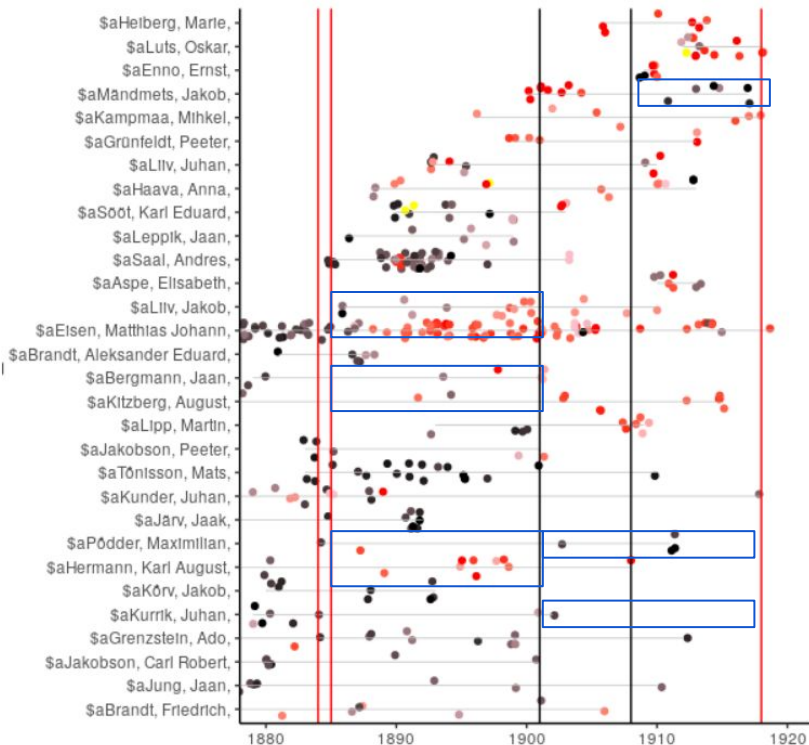
Kujutatud on populatsiooni keskmise mudel (sinine joon) ja ettekirjutuste ajastus (punased ja mustad vertikaaljooned).

Ettekirjutused tihti on täpselt samal hetkel kui populatsiooni trend muutub - jah, mõjutas kasutust.

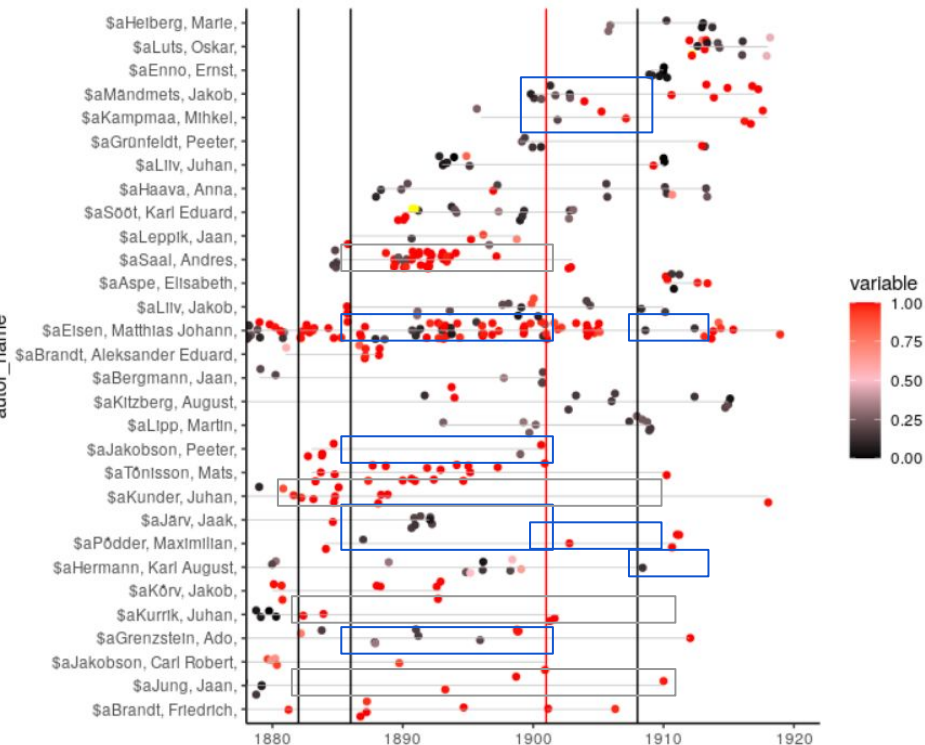


Autorite keelekasutus

Variation of ste_var

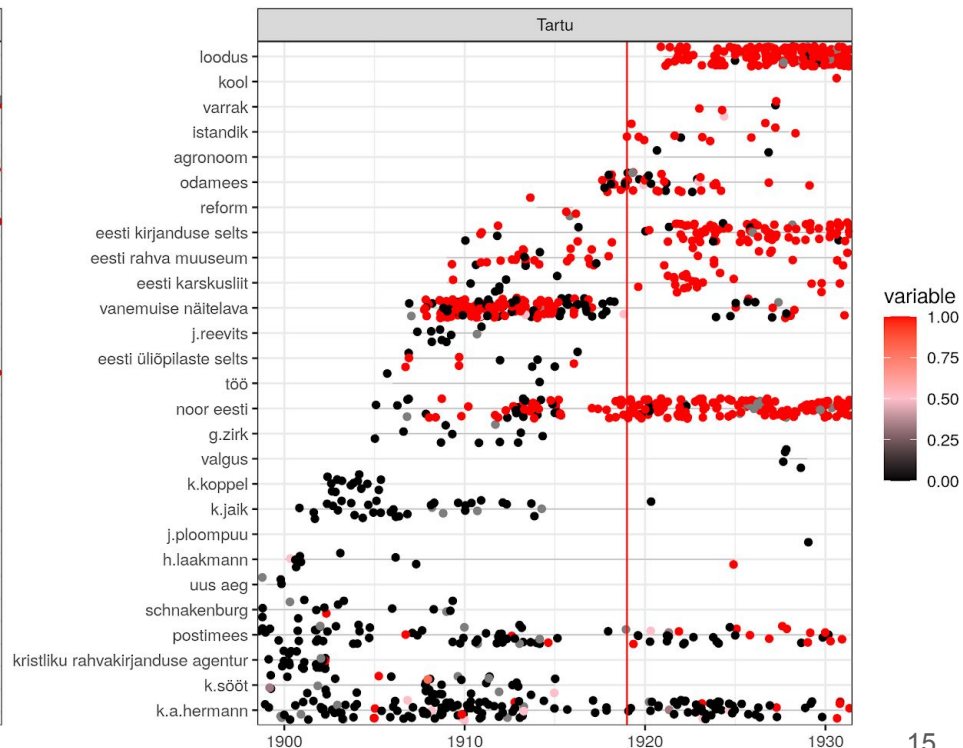
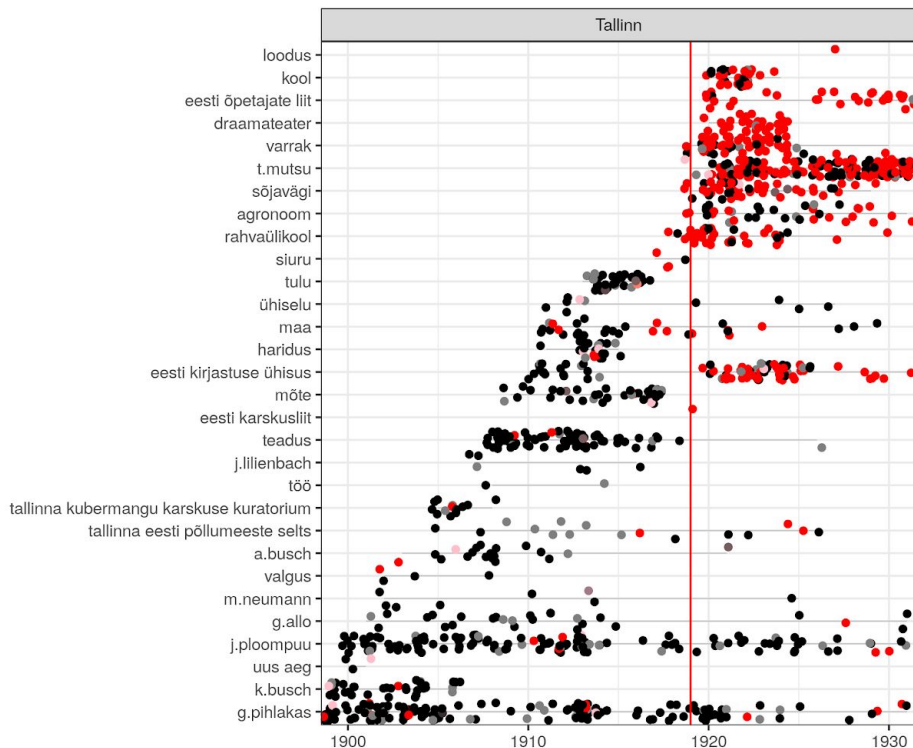


Variation of aa_var



Kirjastajate keelekasutus

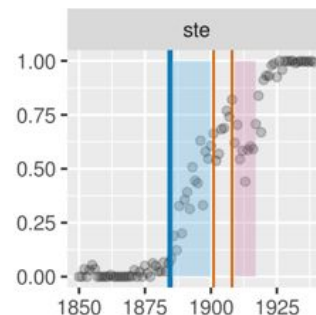
Variation of W vs V in print titles by publisher



Muud tunnused

Logistiline regressioon: kes olid juhid ja järeletulijad ettekirjutuste järgmisel

- Mõjud olid erinevad
- Edukate ettekirjutuste puhul pidid kas a) noored või b) suurlinnade autorid olema eesrinnas.



Phase 1

(n = 82)

Trend
Intended
tradition
dialect
accom
popsize
edu
youth

Phase 2

(n = 44)

Trend
Intended
tradition
dialect
accom
popsize
edu
youth

Phase 1

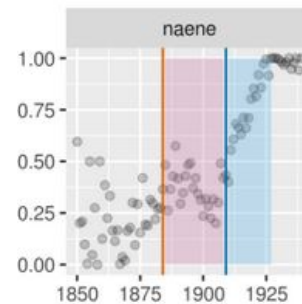
(n = 170)

Trend
Intended
tradition
popsize
edu
youth

Phase 2

(n = 86)

Trend
Intended
tradition
popsize
edu
youth



Phase 1

(n = 107)

Trend
Intended
tradition
dialect
accom
popsize
edu
youth

Phase 2

(n = 35)

Trend
Intended
tradition
dialect
accom
popsize
edu
youth

Phase 1

(n = 199)

Trend
Intended
tradition
popsize
edu
youth

Phase 2

(n = 87)

Trend
Intended
tradition
popsize
edu
youth

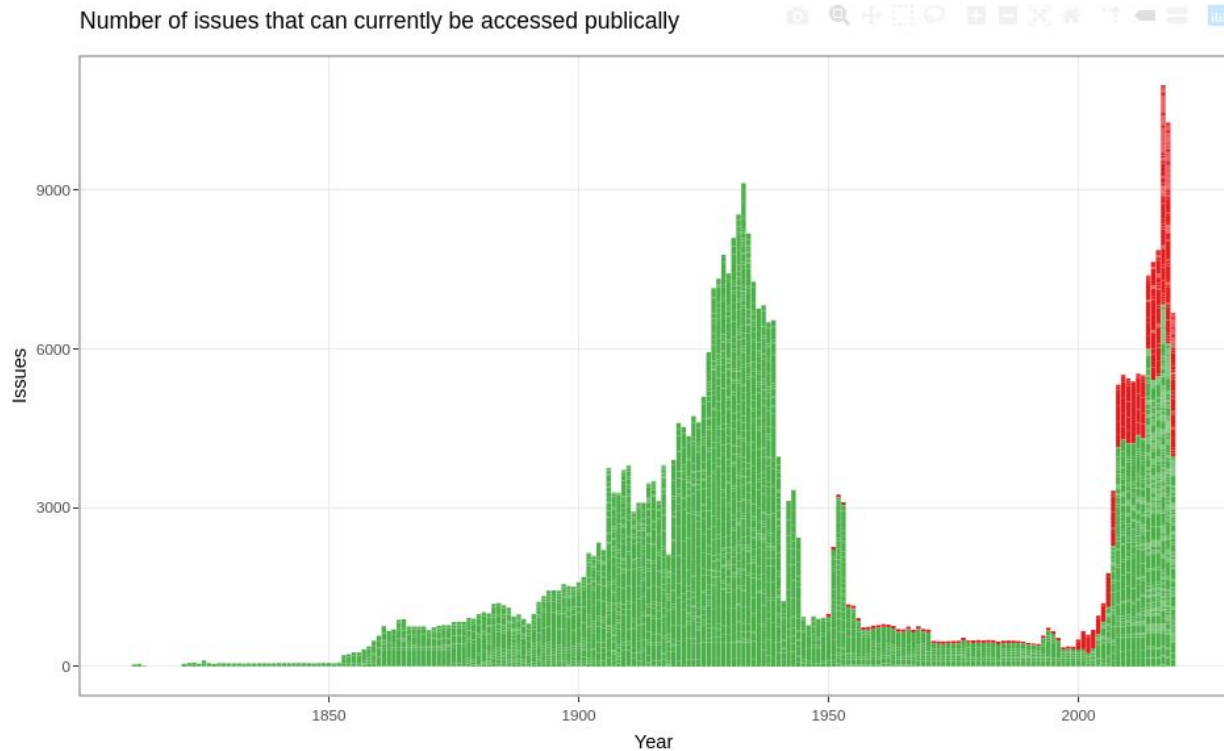
Teisi tulemeid

Koostöö

rahvusraamatukoguga:

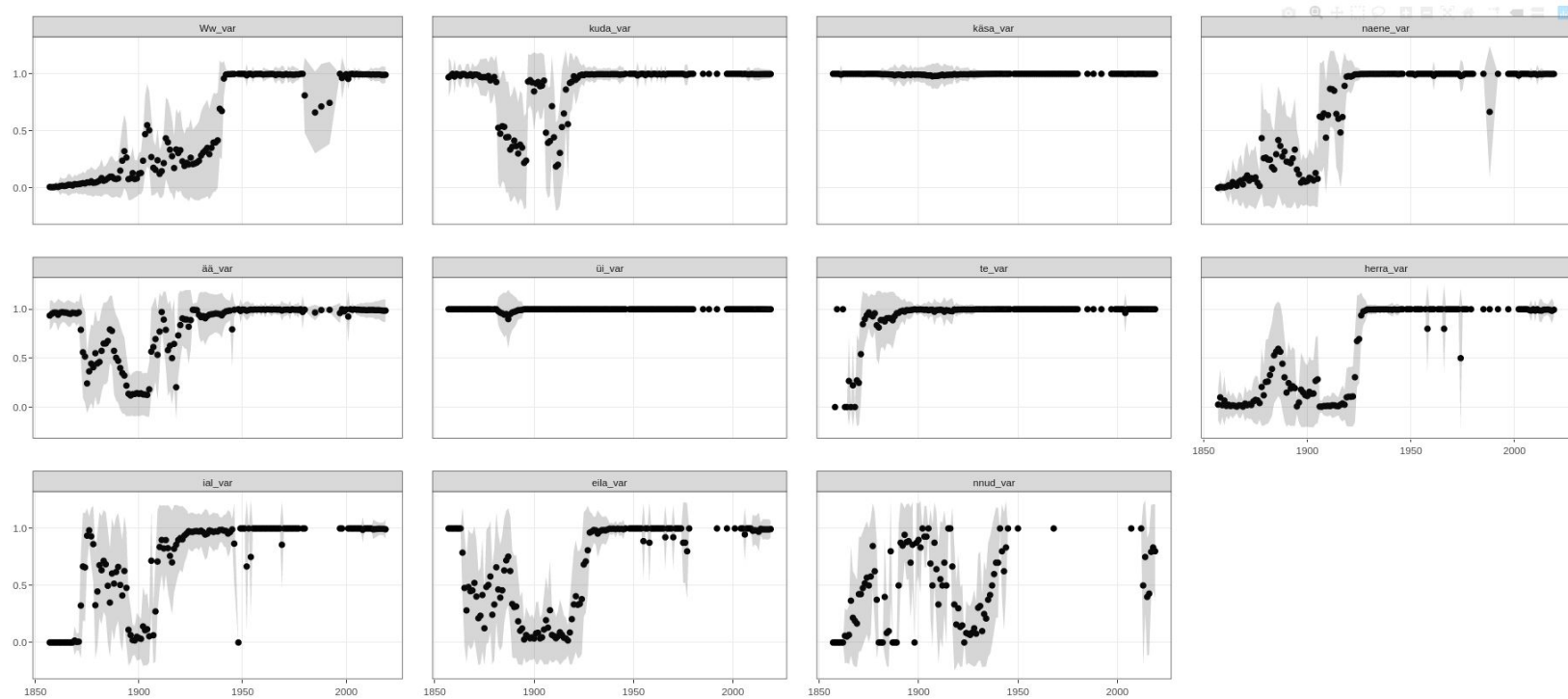
Avame RaRa digiainest
teksti- ja andmekaeveks
(digilabor, mis saab
avatud 2023 alguses).

Praegu 482k väljaannet,
~9,6 miljonit artiklit.



Teisi tulemeid

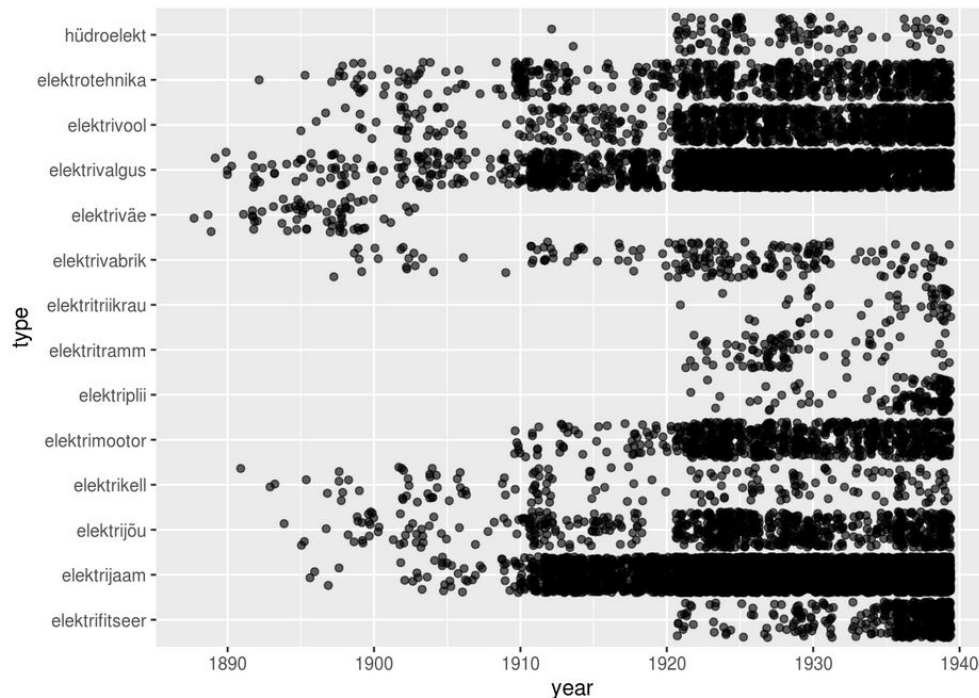
Samad otsingud ajalehetekstidel (pikemad aegread, suurem hulk tekste)



Teisi tulemeid

Tekste saab kasutada ka muul otstarbel - digihumanitaaria

- Paremal elektri ja elektrotehnika levik ajalehekogude põhjal
- Suured korpused on hea aines ühiskonna uurimiseks



Kokkuvõtteks

Suured andmekogud avavad võimalusi vanade küsimuste lahendamiseks uuel viisil või ka uute küsimuste lahendamiseks

Keelelised ettekirjutused 19. saj algul olid päris mõjukad

Kirjutajate elulooline taust ja kirjutiste metainfo avab tahke, kuidas ja mis viisil

Tekstikogu moodustab keeleruumi:

Tekstid, mida loeti, kirjutati kujundavad keele ajalugu ja seda mis vormid meieni jõuavad