



UiT Norgga árktaš universitehta

Borealium

Verktøy for små språk i Norden

Sjur Moshagen (UiT)
Kristine Eide (Språkrådet)

Språk og AI i Norden, 5. og 6. november 2024

Oversikt:

- bakgrunn, historikk (Kristine)
- kva vil brukarmiljøa ha? (Sjur)
 - alt på ein plass
 - enkelt å installera og ta i bruk
- innhald på plattformen (Kristine))
 - oversikt over verktøy og ressursar
 - nyhende
 - instruksjonar - manglar!! (Men ting bør vera utforma så bra at det ikkje trengst)
- lokalisering og språktilpassing (Sjur)
 - arkitektur
 - pontoon
- bruksstatistikk (Sjur)
- målsetjing (Kristine)
- finansiering (Kristine) (takke Nordisk Ministerråd, (Divvun: spesifikasjoner og gjennomføring))
 - drift (Kristine)

- demo & vising i levande live (begge)

Bakgrunn

- Låg språkteknologisk dekning for mange språk
 - skriveverktøy, digitale ordbøker, automatisk opplesing, teksting, maskinomsetjing ...
- Tilbakemeldingar: eksisterande teknologi for små språk kravde større teknologisk kompetanse enn mange brukarar hadde
- Manglande tilgang til digital infrastruktur som språkteknologien garanterer

Historikk

Initiativ frå prosjektet små språk i Norden (2019-2022) i samarbeid med ASTIN

1. *Nordisk insats for att öka kunskapen om språkteknologins betydelse både for ökad digital inkludering och for att stärka och skydda de små språken i Norden.*
2. *Gemensam nordisk hållning och samarbete for att inkludera de små språken i EU:s arbete med språklig jämlikhet i Europa 2030.*
3. *Gemensam nordisk dialog med och strategi for att driva på tech-giganternas arbete med och inkludering av små språk i Norden.*

Resultat (finansiert av nordisk ministerråd)

1. *Kartläggning av behoven av språkteknologiska resurser for de små språken i Norden [European Language Equality – Report on the Nordic Minority Languages](#).*
2. *Utbildning av språkambassadörer for att sprida användningen av befintliga språkteknologiska verktyg*
3. *Utveckling av en webbportal for de små språken i Norden*
 - a. **=> Isländska och neråt** (Nordisk ministerråd)

Kva vil brukarmiljøa ha

- færrest moglege hinder
- lett å bruka språket
- alt samla på ein eller få stader
- verktøya skal vera så like som mogleg dei for majoritetsspråka
- oversikt over dei ressursane som finst

Borealium er tenkt som svar på fleire av desse punkta

Browser address bar: borealium.org

Borealium 🔍 🇸🇦 NORSK, NYNORSK ▾

VELKOMEN

Borealium — verktøy for små språk i Norden

Borealium samlar alle verktøy for språkstøtte for små nordiske språk på ein plass. Her kan ein søkja i alt som finst, velja eit språk nedanfor, eller berre bla gjennom alt som finst: stavekontrollar, grammatikkontrollar, ordbøker, tastatur med meire. Nesten alt er gratis og bygt med open kjeldekode.

🇸🇦 Sjå alle verktøy for eit språk ved å klikka på språket nedanfor:

Map labels: Kalaallisut, Davvisámegiella, Anaråškielá, Nuorttsáá'mkiõll, Meänkieli, Bidumsámegiella, Yiddisch, Farøyskt, Íslenska, Aarjelsaemien giele, Ubmejesámegiella, Julevsámegiella, Kväänin kieli.

TASTATUR

Tastatur for mobiltelefon og nettbrett

Mobilappen Divvun Keyboards inneheld tastatur for samiske og andre språk. For dei fleste språka finst det òg ein innebygd stavekontroll. Appen er tilgjengeleg for iOS og Android.

Divvun Keyboards
på iOS og Android

PROGRAMINSTALLERING

Ei nedlasting — resten blir handtert automatisk

Med Divvun Manager kan ein installera retteprogram, tastatur og andre verktøy. Divvun Manager hentar automatisk nye versjonar når dei er tilgjengelege.

Divvun Manager
på Windows og macOS

NYHENDE

🇸🇦 Nå kan datamaskiner for første gang snakke sørsamisk — på 100-årsdagen til stemmegiveren

UiTs samiske språkteknologimiljø har laget en sørsamisk talesyntese basert på arkivmateriale av taleren, i samarbeid med Samisk arkiv - Arkivverket. Dette kan bidra til å bevare det unike språket.

LES MEIR

🇸🇦 Nå kan datamaskiner for første gang snakke lulesamisk

UiTs samiske språkteknologiutviklere har nemlig laget en lulesamisk talesyntese med tre kunstige stemmer. Dette kan bidra til å bevare det unike språket.

LES MEIR

🇸🇦 Vad är språkteknologi för små språk?

Vad är språkteknologi och vad används det till? Och varför saknas det teknologi för små språk? Trond Trosterud, professor i samiska språk vid Giellatekno på universitetet i Tromsø, Norge berättar varför språkteknologiska verktyg är viktiga för att kunna läsa, skriva, förstå och prata små språk i Norden.

LES MEIR

Utviklingsarbeidet

- prosjektet leidd av Divvun-gruppa
- ASTIN har vore referansegruppe
- arbeidet gjort av Necessary Innovations (Gøteborg)
- ein fyrste versjon ferdig i desember 2023
- offentleg lansering 6.11.2024

Målsetjing

Minst mogleg arbeid med:

- vedlikehald
- utvikling
- lokalisering
- drift
- lett å leggja til nye ressursar

Bruksstatistikk:

- samla inn statistikk over bruken av portalen
- i samsvar med GDPR
- vi nyttar plausible.io

Lokalisering og språktilpassing

- fleirspråkleg
- minst mogleg tekst å omsetja
- enkelt å omsetja
- omsetjingsbehov tilpassa kvart språk
- Omsetjing i Pontoon (Mozilla)
- eit kjerneprosjekt for all felles info
- eitt prosjekt pr ressurspråk
 - lokalisering av info om t.d. alle islandske ressursar i eit separat prosjekt
- fall tilbake på nærmaste relevante språk om omsett info ikkje finst, i siste hand engelsk
- alle med GitHub-konto kan omsetja (med kontroll)

Pontoon

The screenshot displays the Pontoon web interface. At the top, there's a navigation bar with 'Teams', 'Projects', 'Contributors', and 'Machinery'. The main heading is 'Projects'. Below this, there are statistics: 16 projects, with 'boreallium core' having the most translations (Terminology) and 'sátni.org' having the most missing strings. A large circular gauge shows 70% completion. To the right, a legend indicates 2,893 translated strings, 0 pretranslated, 0 warnings, 0 errors, and 1,201 missing strings. Below the statistics is a search bar for projects and a table listing various projects with their target dates, priorities, latest activity, and progress bars.

PROJECT	TARGET DATE	PRIORITY	LATEST ACTIVITY	PROGRESS
boreallium core	Nov 4, 2024	★★★★★	2 minutes ago	78%
boreallium fao	Nov 4, 2024	★★★★★	6 days, 13 hours ago	53%
boreallium fit	Nov 4, 2024	★★★★★	6 days, 13 hours ago	44%
boreallium fkv	Nov 4, 2024	★★★★★	6 days, 13 hours ago	75%
boreallium isl	Nov 4, 2024	★★★★★	1 week, 1 day ago	58%
boreallium kal	Nov 4, 2024	★★★★★	6 days, 12 hours ago	54%
boreallium multi lang	Nov 4, 2024	★★★★★	13 seconds ago	48%
boreallium sjd	Nov 4, 2024	★★★★★	6 days, 12 hours ago	93%
boreallium sje	Nov 4, 2024	★★★★★	6 days, 12 hours ago	100%
boreallium sma	Nov 4, 2024	★★★★★	6 days, 12 hours ago	42%
boreallium sme	Nov 4, 2024	★★★★★	6 days, 12 hours ago	35%
boreallium smj	Nov 4, 2024	★★★★★	6 days, 12 hours ago	80%
boreallium smn	Nov 4, 2024	★★★★★	—	Not synced yet
boreallium sms	Nov 4, 2024	★★★★★	—	23%

Finansiering og drift

Prosjektet er finansiert med:

- midlar frå Nordisk Ministerråd
- eigeninnsats til innhald, lokalisering og omsetjing frå ulike språkrøktorganisasjonar i Norden:
 - Sámi Giellagáldu
 - Divvun
 - Oqaasileriffik / Grønlands språksekretariat
 - Fróðskaparsetur Føroya
 - Árnastofnun (Island)
 - ISOF (Sverige)
 - Dansk Sprognævn
 - Språkrådet (Norge)

Drift:

- Divvun driv portalen over eige budsjett
- Eigeninnsats frå alle samarbeidspartar



UiT Norgga árktaš universitehta

Borealium

Verktøy for små språk i Norden

Demo - takk for merksemda!

Har du ein eller fleire ressursar til portalen? Ta kontakt med Divvun-gruppa: feedback@divvun.no