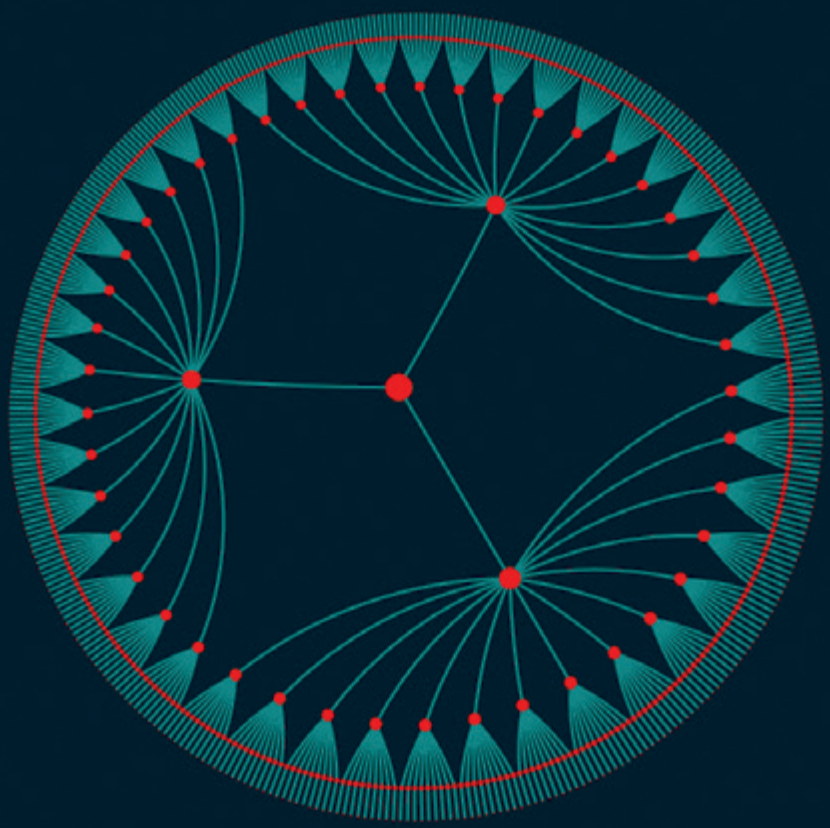
NRC·CNRC

Technologies for Indigenous languages in Canada

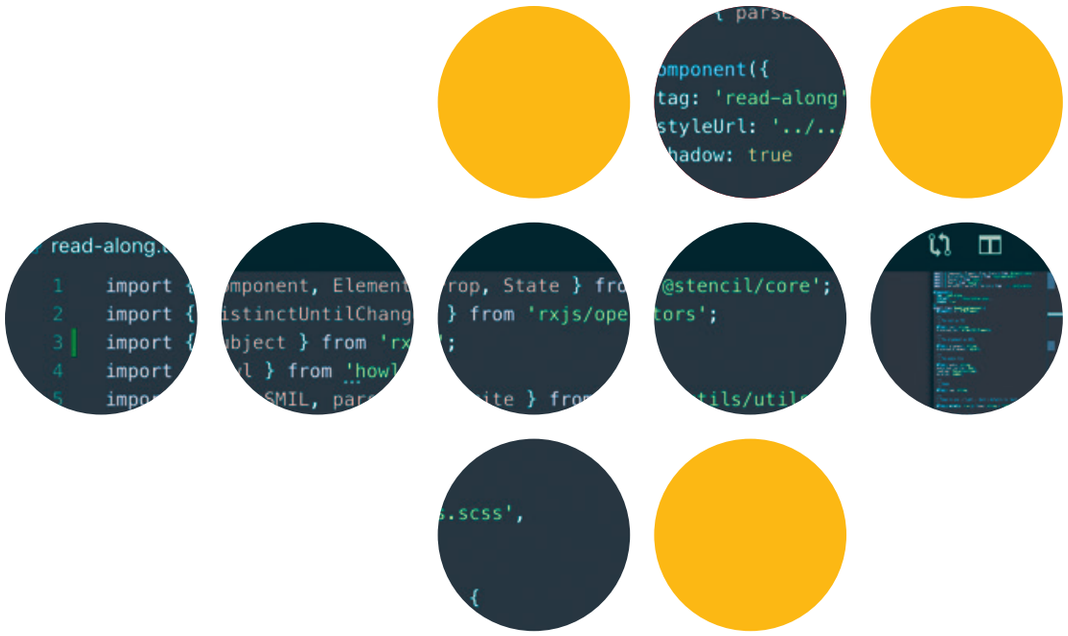
Working with Indigenous communities and language experts to develop technologies that support language reclamation, revitalization and stabilization.



National Research Council Canada
Conseil national de recherches Canada

Canada

In collaboration with Indigenous language experts, instructors, and communities, the National Research Council of Canada (NRC) is working on speech- and text-based technologies that contribute to the revitalization of Indigenous languages.



Researchers from the NRC's Indigenous Languages Technology Project are working together with partners to support them and their efforts on the future of Indigenous languages.

Featured projects

Verb conjugation software

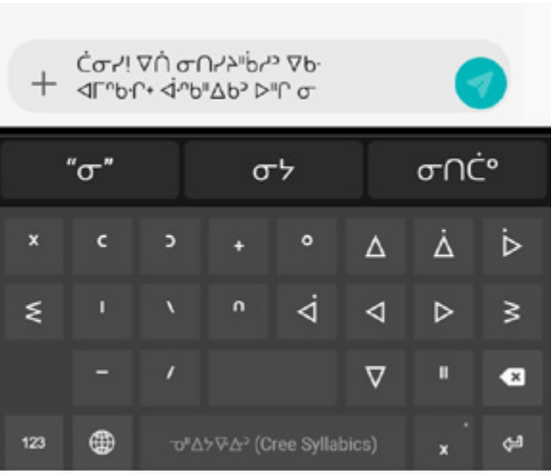
Collaborators: Onkwawenna Kentyohkwa Mohawk-language immersion school (Western Mohawk), Kanien'kehá:ka Onkwawén:na Raotitióhkwa Language and Cultural Center (Eastern Mohawk), Kitigan Zibi Cultural Centre (Algonquin), Prairies to Woodlands Indigenous Language Revitalization Circle (Michif)

- Most Indigenous languages in Canada are polysynthetic — making verb conjugation one of the hardest things to learn.
- Verb conjugation software has been developed for two dialects of Mohawk, as well as for Algonquin and Michif. One for Mi'kmaq is being prepared.

ReadAlong Studio

Collaborators: Carleton University's School of Linguistics and Language Studies, Michael Running Wolf (Northeastern University)

- The ReadAlong Studio software highlights words in Indigenous books as they are being read aloud. Listeners can click on any word to hear it spoken again.
- Educators and students find this tool very helpful in the classroom and for independent learning. It's been applied to 22 languages, with more to come.



Predictive text

Collaborators: First Peoples’ Cultural Council, Keyman, W̱SÁNEĆ School Board, Google Internationalization

- This software provides suggestions for texting on cell phones in ᓂᐢᐃᔭᐁᐃᓱ (Plains Cree) and Qʷi·qʷi·diččaq (Makah).
- The FirstVoices Language Tutor software has been updated and now includes predictive text software for FirstVoices keyboards in SENĆOŦEN.

Digital recordings

Collaborators: University of British Columbia (UBC)’s Indigitization Project, Council of Yukon First Nations, Tsilhqot’in National Government, Syilx Language House

- Older recording formats are degrading and becoming obsolete. Our project with UBC supports the digitization of audio and video recordings that are vital to each community.
- Language workers in the Yukon have created digital video recordings of Elders speaking in their Indigenous languages. These are being used in a new program designed to mentor the next generation of language leaders.
- The Tŝilhqot’in National Government has expanded its database of audio recordings and transcriptions, extended its suite of language applications, and built language technology and transcription capability in the community.
- There are fewer than 40 Elders remaining who are fluent in Nsyilxčn̓ (Interior Salish). New digital recordings have been made of Elders storytelling and sharing oral literature.

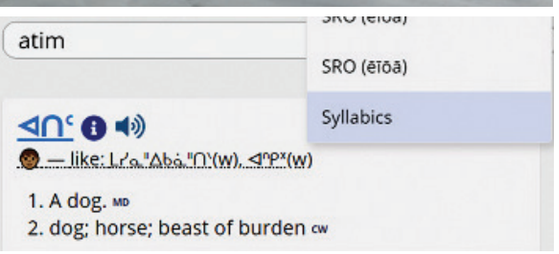
Online language courses

Collaborators: 7000 Languages, Prairies to Woodlands Indigenous Language Revitalization Circle

- Four Indigenous community teams have created online courses for Kwak’wala, Mi’kmaq, Naskapi and Southern Michif, with the NRC’s help.

More than 25 languages

Algonquin, Atikamekw, Cree (several dialects), Gitksan, Gwich’in, Hän, Innu, Inuktitut, Kaska, Kwak’wala, Lingít, Michif (the Métis language), Mi’kmaq, Mohawk (two dialects), Naskapi, Northern and Southern Tutchone, Nsyilxčn̓, SENĆOŦEN, Seneca, Tagish, Tŝilhqot’in, Tsuut’ina and Upper Tanana



Dictionaries

Collaborators: Prairies to Woodlands Indigenous Language Revitalization Circle, Turtle Mountain Community College, Carleton University's School of Linguistics and Language Studies, Alberta Language Technology Laboratory, Maskwacîs Education Schools Commission, Cree Literacy Network, Dr. Marie-Odile Junker (Carleton University)

- Michif first-language speakers have contributed to a new spoken dictionary app and website, making an out-of-print Michif dictionary accessible to educators and students.
- A new interface for itwêwina, the intelligent Plains Cree dictionary that analyses and assists with word formation, is in development.
- A common digital infrastructure helped improve learning tools for Algonquian languages, including dictionaries and a linguistic atlas.

What's next?

Reading and writing in Inuktitut

- We've recently created a suite of tools for reading and writing in Inuktitut (morphological analyzer, dictionaries, web search engine and reading assistant), which are at the user testing phase.

Automatic Speech Recognition (ASR)

- Beginning with Inuktitut and Cree, then expanding to Innu, Dénésiline, Tsuut'ina, Michif and more, we're developing language labelling and speech segmentation tools to enable ASR.

TTS technology

- In partnership with 3 Indigenous organizations, we will be working on a new project to develop text-to-speech (TTS) technology.

Machine translation

- Our latest project, still in the research stage, is machine translation for English to Inuktitut and vice versa.

Contact

Roland Kuhn, Project Leader
Indigenous Languages Technology Project
roland.kuhn@nrc-cnrc.gc.ca

canada.ca/nrc-indigenous-languages-technology

#NRCDigitalTech

Cat. No. NR16-396/2023E-PDF, ISBN 978-0-660-48086-2
An HTML version of this product is available on the NRC website. Également disponible en français. March 2023