CLARIN
AGREEMENT

CLARIN Agreement with the National Consortium of Finland

(1)   CLARIN ERIC, a European Research Infrastructure Consortium established by Commission Implementing Decision (EU) 2012/136 of 29 February 2012 (hereinafter "**CLARIN ERIC**"), and,

(2)   the National Consortium of Finland, represented by the University of Helsinki, FI03134717 (hereinafter "**National Consortium**"),

hereinafter collectively referred to as the "Parties", and individually as "Party",

HAVE AGREED on the following specific terms and conditions for contribution of resources to CLARIN ERIC during the period

January 1, 2022 until further notice according to Article 3.4.

Article 1 – Scope

The Parties have agreed to conclude an agreement in accordance with Article 6.5 of the Statutes, including the specific terms and conditions described in this Agreement which relate to Art. 2, 6.1 and 6.2 (d)-(m) of the Statutes.

Article 2 – Definitions

"National Coordinator" means the person appointed as a national coordinator by a CLARIN ERIC member in accordance with Article 6(2)(e) of the Statutes and which is responsible for the National Consortium.

"Statutes" means the Statutes of CLARIN ERIC including its Annexes.

Article 3 – Obligations

3.1 Contribution to CLARIN ERIC overall and specific obligations listed in Articles 2 and 6 of the statutes

In addition to the specific contributions mentioned below in Article 3.2 below, the National Consortium shall (i) promote the adoption of relevant standards in national resource and tool creation projects, as well as the uptake of CLARIN services among researchers in their country, (ii) gather user feedback and requirements, and (iii) support CLARIN centres in the member country by facilitating integration into national and other relevant infrastructures.

3.2 Specific National Contributions

This section describes the national level in-kind contribution to CLARIN; part of this contribution will be coordinated at the transnational level through CLARIN ERIC, and part will be coordinated at the national level. This agreement lists the components of the offering at the time of signature. For each component, details are provided in technical annexes where relevant.

3.2.1. Contributions to be coordinated by CLARIN ERIC

(a) **Data and service centres**: In accordance with Art. 6.2 (f) of the Statutes, the National Consortium shall contribute at least one centre of type B (cf. the check list https://office.clarin.eu/v/CE-2013-0095-B-centre-checklist-v6.pdf).

Detailed descriptions of the centres and the services they offer are included in Annex 1-B. CLARIN centers are expected to obtain a deep level of integration with the infrastructure by connecting a substantial portion of their language resources and tools via:

· harvestable metadata for the VLO

· endpoints for the Federated Content Search

· web applications for the LR Switchboard

(b) **Access**: The National Consortium shall comply with statutes Article 19 of the Statutes on access policies for users. Any restrictions or exceptions are specified in Annex 1-F, to the extent it is applicable.

(c) **User Authentication and Authorization System**: The National Consortium shall comply with the specification given in Annex 1-C.

(d) **Knowledge infrastructure (KI)**: The National Consortium shall contribute to the KI by providing access to its knowledge and expertise to the CLARIN community at large in accordance with Annex 1-D.

(e) **Membership of CLARIN ERIC committees and any other contribution** being offered to CLARIN and lending itself for coordination by CLARIN ERIC at the European level. Detailed descriptions are included in Annex 1-E.

3.2.2. Contributions coordinated at the national level, but compliant with CLARIN

The National Consortium shall provide for:

(a) Enhancements and improvements to existing resources, tools and services in accordance with the detailed descriptions included in Annex 2-A.

(b) Creation of new resources, tools and services in accordance with the detailed descriptions included in Annex 2-B.

(c) Measures to promote the adoption of and adherence to CLARIN standards in national resources and tools creation programmes and projects in accordance with the detailed descriptions included in Annex 2-C.

(d) Measures to create a knowledge infrastructure at the national level, with a view to promoting uptake of CLARIN services by the target audience, to exchanging knowledge and expertise, and to advancing Humanities and Social Sciences research through the support by CLARIN. A brief description of those items that are not already covered by 3.2.1-(d) is provided in Annex 2-D.

(e) Any other contribution relevant to the goals of CLARIN, including collaboration with third parties that CLARIN as a whole could contribute to or from which CLARIN could benefit. Detailed descriptions are included in Annex 2-E (if applicable).

3.3 National Contributions made jointly with other Parties

In accordance with Article 6.3 of the Statutes, members may decide to provide some of the national contributions in cooperation with other members, observers or third parties. In such cases, a separate multi-party agreement may be made between the CLARIN ERIC and those members / observers / third parties concerned.

3.4 Termination and Amendment of the Agreement

The Agreement may be terminated pursuant to a withdrawal of Finland from CLARIN ERIC, and the Agreement may be amended if there is a change in the national consortium or the commitments of its members. Otherwise the Agreement may be revisited for Amendments or Termination every five years upon request by either party.

Signatures

Date: **February 28, 2022**   Date:

Name: Krister Lindén   **Name: Franciska de Jong**

Function: National Coordinator   **Function: Executive Director**

Annex 1 National Contributions to be coordinated by CLARIN ERIC

Annex 1-A

The National Consortium of Finland, FIN-CLARIN, consists of the following partner(s): https://www.kielipankki.fi/organization/fin-clarin/

1. The **University of Helsinki** (UHEL) offers expertise, e.g., in the following areas: European and world languages, corpus linguistics, Variation of English (VARIENG), language universals, language technology including but not limited to morphological, surface-syntactic parsing, speech synthesis as well as lexical data including synonymy and terminology databases, and annotated multimodal and multilingual corpora.

2. **CSC – IT Center for Science Ltd**. offers services for language-based research including an extensive number of text and speech data collections as well as software suitable for research in an efficient computing environment. CSC offers its expertise, e.g., in the following areas:
   - Evaluation, installation, tailoring, development and maintenance of language-based software and tools.
   - Web-based collaboration environments for research communities as well as web interfaces for language-based tools, e.g. as part of puhti.csc.fi, korp.csc.fi, sanat.csc.fi.
   - For the European CLARIN infrastructure and as needed for implementing functionalities for international use, e.g., technical definitions, PID compatibility for resources, authorization service extensions, installation and operation of new resources and their required computing and archiving servers, agreements concerning international resource use, research tool compatibility and integration on an international level.

3. The **Institute for the Languages of Finland (KOTUS)** offers expertise, e.g., in the following areas: internet access to scientific dictionaries, other dictionaries, lexicons and wordlists, resources for linguistic and onomasiological maintenance, text, speech and video corpora as well as other audio-visual digital data collections. Resources are available in particular for Finnish and Swedish but also for languages related to Finnish.

4. The **Aalto University** offers expertise, e.g., in the following areas: automatic speech recognition, speaker adaptation for speech recognition and synthesis, alignment of text and speech, indexing, retrieval and browsing of audio and audio-visual material, methods and tools for the analysis and parameterization of the glottal source, vocoders in HMM-based speech synthesis, methods for robust parameterization of speech, as well as resources with simultaneous recordings of speech and EGG data.

5. The **University of Eastern Finland** offers expertise, e.g., in the following areas: planning, recording and annotation of speech corpora and creation of speech and text corpora, multi-lingual resources, in particular for translation and contrastive language studies.

6. The **University of Jyväskylä** offers expertise, e.g., in the following areas: language skills assessment corpora, multi-lingual learners' corpora, corpora of English as a teaching language and as a foreign language, as well as corpora of child and student language use.

7. The **University of Oulu** offers expertise, e.g., in the following areas: regional literary Finnish, Finnic minority and regional languages, Northern varieties of Finnish, learners' corpora, corpus linguistics, corpus methodology, Saami linguistics, Saami language technology, training researchers and other professionals who have a profound knowledge of Saami language and culture.

8. The **University of Tampere** offers expertise, e.g., in the following areas: search tools, text corpora and parallel text corpora.

9. The **University of Turku** offers expertise, e.g., in the following areas: syntactic and semantic annotation of text and speech, text retrieval and analysis tools, the dialect corpus of the syntax archive, and the morpho-syntactic database of the works of Michael Agricola, health data corpora, bio-molecular interaction resources, search and analysis services, and collections of Satakunta regional data, Finnish language learner data, and old literary Finnish.

10. The **University of Vaasa** offers expertise, e.g., in the following areas: terminology, especially conceptual analysis, research in language for special purposes as well as discourse analysis.

The National Coordinator is:
Dr. Krister Lindén
Research Director
Faculty of Arts
Department of Digital Humanities
P.O.Box 24, Unioninkatu 40B
FI-00014 University of Helsinki, Finland
Phone: +358 2941 29303
Email: krister.linden@helsinki.fi

The invoice for Annual fee is to be sent to:
Academy of Finland EDI identifier: 003702458939
Opus Capital Group Oy (operator): 003710948874
Academy of Finland Business ID: FI0245893-9
Contact information:
Ms. Merja Särkioja
Academy of Finland
Hakaniemenranta 6
POB 131
FI-00531 Helsinki
Phone: +358 295 335 000
Registrar's Office: +358 295 335 049
Emails: merja.sarikoja@aka.fi, kirjaamo@aka.fi

Annex 1-B Data and Computing Centres

The following data and computing centres will be providing services to CLARIN:

1. The **University of Helsinki** (http://www.helsinki.fi/fin-clarin) is responsible for the contents of the FIN-CLARIN Language Bank (http://www.kielipankki.fi).

2. **CSC - IT Center for Science Ltd,** (http://www.csc.fi/) offers data processing services for the FIN-CLARIN Language Bank, including storage capacity and computing facilities for storing and analyzing language resources, provision of tools that can access language resources, implementing and maintaining a web-based interfaces to language bank tools, user management and authorization infrastructure for the language resources, support of persistent identifiers for digital objects, participation in further development of IT infrastructure for CLARIN in accordance with European CLARIN partners and developments, and maintenance of language bank services.

B centres and certification status:

· **CSC - IT Center for Science Ltd** has been certified as a CLARIN B centre since 2015.

Annex 1-C Authentication and Authorisation Infrastructure

Specification of the implementation of user authentication and authorisation system: All national centres (or the universities they belong to) have joined the AAI federation of the Finnish National Research and Education Network (HAKA).

Annex 1-D Knowledge Infrastructure

The National Consortium FIN-CLARIN will (i) publish the relevant Knowledge Infrastructure activities and facilities it offers on the CLARIN website according to guidelines provided by CLARIN ERIC, (ii) will keep this list up to date, and (iii) will participate in the coordination of these activities with other National Consortia in order to maximize their efficiency and effectiveness. In particular, FIN-CLARIN provides the K-Centre SAFMORIL:
https://www.kielipankki.fi/safmoril/

Annex 1-E Committee Participation

FIN-CLARIN will contribute to the permanent CLARIN ERIC committees where applicable, and will participate in ad hoc committees as required and within the limits of the national budget:
· National Coordinators' Forum: Krister Lindén (UHEL)
· Standing Committee for CLARIN Technical Centres: Martin Matthiesen (CSC)
· Legal Issues Committee: Krister Lindén (UHEL)
· User Involvement Committee: Mietta Lennes (UHEL)
· Standards Committee: Jussi Piitulainen (UHEL)
Any other contribution:
· Ad-hoc or sub-committees: Krister Lindén (UHEL) unless delegated.

Annex 1-F   FAIR Access to Data and Tools

Resources, tools, and services provided by FIN-CLARIN will comply with statutes article 19 on Access Policies for Users. Any restrictions or exceptions are specified in the Language Bank Rights authorization tool (https://www.kielipankki.fi/data-protection-statement/user-of-the-lbr-service/) or by similar means for getting access to the resources according to the FAIR principles[1].

Annex 2 National Contributions to be coordinated at the national level

Annex 2-A National Work Programme

The FIN-CLARIN partners provide the following resources, tools, and services through the Language Bank of Finland.

SERVICES AND TOOLS
- Name: https://www.kielipankki.fi/tools/
  Contact: finclarin@helsinki.fi

CORPORA
- Name: https://www.kielipankki.fi/corpora/
  Contact: finclarin@helsinki.fi

ON-GOING ENHANCEMENT WORK

1. The **University of Helsinki** will improve and extend Finnish and Finnish-Swedish language corpora, such as the National Library Newspaper collection, the social media corpus Suomi24, Finnish bilingual text and speech corpora as well as the Helsinki Term Bank of Arts and Sciences.

2. The **Institute for the Languages of Finland (KOTUS)** will improve their existing resources and services.

3. The **Aalto University** will improve and develop their speech recognition software and other language processing tools.

4. The **University of Eastern Finland** will improve and develop their corpora for cross-linguistic studies.

5. The **University of Jyväskylä** will improve their corpora of language proficiency assessments.

6. The **University of Oulu** will improve and maintain the Saami Culture Archive as well as develop Finnish multimodal and multilingual corpora.

7. The **University of Tampere** will improve and maintain their parallel corpora of literary texts, multilingual parallel and comparable corpora of juridical texts, a resource collection of human-computer dialogues, and the https://mustikka.uta.fi web server.

8. The **University of Turku** will improve and maintain syntactically and semantically annotated text and speech corpora, text retrieval and analysis tools, the syntax archive dialect corpus and related user guidance, as well as an annotated corpus of historical Finnish and its documentation.

9. The **University of Vaasa** will maintain WasaTerm - Terminologisia sanastoja, Terminology Forum

Annex 2-B New Resources

National Work Programme:

FIN-CLARIN is one of the more than 30 infrastructures on the Finnish national roadmap of research infrastructures. The FIN-CLARIN plan is to create new language resources and tools for language-based processing, to collect and upgrade existing language resources to make them widely usable, and to include new language-based resources created in other projects as they become available. For additional information on new resources, see the table of forthcoming new FIN-CLARIN resources:
https://www.kielipankki.fi/corpora/forthcoming/

Annex 2-C CLARIN Standards

Measures to promote the adoption of and adherence to CLARIN standards in national resources and tools creation programmes and projects:

FIN-CLARIN cooperates with the Academy of Finland in including appropriate funding requirements in research project as part of their data management plan to produce language resources. The goal is to have research proposals include funding for making resources available in standard formats with proper metadata so that the resources can be seamlessly integrated in the FIN-CLARIN repository, preserved and re-used by other scholars. In addition, instructions and documentation are maintained for creating metadata in CLARIN compatible format. For additional information on how FIN-CLARIN implements CLARIN ERIC standards, see the FIN-CLARIN Data Management Policy:
https://www.kielipankki.fi/support/data-management-policy/

Annex 2-D Knowledge Infrastructure

Contributions to the Knowledge Infrastructure mainly targeting national audiences and not already covered by Annex 1-D.

Members of the national consortium have their areas of expertise as mentioned in the description of each member. Each member creates and maintains resources, tools and services related to its area of expertise in FIN-CLARIN and will advise students and researchers of other FIN-CLARIN members on the creation and usage of such resources, tools and services. In particular, we foresee the following areas of activity:
- Maintaining a national web portal hosting or linking to web services of all Finnish CLARIN members
- Close cooperation with associations of the humanities and social sciences
- Specific workshops, training events, and summer schools for researchers from the humanities and social sciences
- Hosting and training of junior researchers
- Development of teaching modules for student curricula
- Support and help desk for users of FIN-CLARIN web applications

Annex 2-E Other

Any other contribution:
- Close cooperation with DARIAH-FI as a component of FIN-CLARIAH on the joint national RI roadmap.
- A co-operation agreement with the other RIs in the area of SSH in Finland (CESSDA, ESS, SHARE).

---

[1] https://www.fairdata.fi/en/about-fairdata/fair-principles/