|  |  |
| --- | --- |
| How to prepare an application to become a recognised CLARIN K-centre *(version 13.5, 26-03-2024)* | CLARINERIC_nosub |

1. **Background**

One of the building bricks of the CLARIN Knowledge Infrastructure are knowledge centres (abbreviated K-centres). Their role is to provide a place (physical or virtual) where people (both those using the infrastructure and those who help building or operating it) can get cross-border access to knowledge and expertise in specific areas.

K-centres are also expected to contribute to other CLARIN activities where appropriate.

The K-centres can be:

* *single-sited*: a group of people in one physical location or institute

or *distributed*: people spread over various locations in one or more countries;

* *specialised in one topic* (e.g. speech analysis, treebanking, Polish language technology) or

*covering a variety of topics* with a common denominator (e.g., the languages of Sweden);

* and are expected to offer a mix of proactive and reactive services.

The instruments they use may vary. Knowledge centres may or may not offer technical infrastructure or data services as well and may be located in or outside CLARIN ERIC member or observer countries. If your K-centre is located in a CLARIN country, it is recommended that you inform your National Coordinator of your plans to establish or join a K-centre as the activities of a K-centre are considered as contributions from your National Consortium to CLARIN.

An overview of all existing CLARIN K-centres can be found on <https://www.clarin.eu/content/list-clarin-k-centres>

1. **Application process**

Institutes that feel that they are sitting on specific knowledge and expertise that could be beneficial for current or future users of the CLARIN infrastructure, and that are willing to share them can submit an application to be recognised as a CLARIN K-centre. Please note that relevant expertise for CLARIN users is not necessarily linguistic in nature: IPR, ethics, Artificial Intelligence and linked open data are examples of potentially useful expertise for those who work with language.

Applications should make use of the form below and should be sent to   
[k-centre-application@clarin.eu](mailto:k-centre-application@clarin.eu)

Applications will be handled by the Knowledge Infrastructure Committee (abbreviated KIC), which will send feedback to the applicants. This could be questions for clarification, suggestions to improve the application or other comments.

The applicants submit their responses (normally in the form of an update of the application), and when the KIC is satisfied the application will be forwarded to the CLARIN ERIC Board of Directors (BoD), together with the KIC’s recommendations. The BoD will formally decide about the recognition as a K-centre for a period of 3 years. After approval by the BoD the applicants will be notified and they will receive the official K-centre certificate plus, if needed, guidelines for the landing page of the K-centre.

The K-centre will submit annual reports to the KIC on their activities, and after the 3 years period the K-centre and KIC will discuss the developments and if relevant prepare a proposal for a prolongation for another 3 years.

1. **Application form**

The application form below is a Word template, consisting of four main components:

* Administrative and contact data
* Information about what the K-centre plans to offer to users
* Confirmation of the K-centre’s willingness to comply with a number of requirements
* Information that should help KIC and BoD to judge the quality and sustainability of the K-centre

In section 4, after the blank application form, we provide a version of the application form with extra explanations and filled in with examples, mostly borrowed from descriptions of already existing K-centres. As there is no such thing as a “standard K-centre”, applicants should not hesitate to contact KIC at the email address below if they are not sure what to answer to the various questions. The process from initial application to (ideally) final approval by the KIC is normally iterative, and not a one step accept/reject process.

Please copy the form below to a separate Word document, complete it, and send it to   
[k-centre-application@clarin.eu](mailto:k-centre-application@clarin.eu)

|  |
| --- |
| K-centre application form (v13.5, March 2024) *Items marked with an asterisk (\*) are obligatory*  *Complete the form and send it as a Word document to* [k-centre-application@clarin.eu](mailto:k-centre-application@clarin.eu) |
| PART 1 Administrative data |
| 1. ***\*Proposed full name in English*** *(max 80 characters, the name must contain “CLARIN Knowledge Centre” or “CLARIN K-Centre”, and it should reflect the main theme(s) covered by the centre)* |
|  |
| 1. ***\*Proposed short name or acronym*** *(max 16 characters, no spaces, no special signs other than underscore or dash)* |
|  |
| 1. ***\*Host institution (only 1) and main contact(s) for this application*** |
| * 1. ***\*Name of the institution*** |
|  |
| * 1. ***\*Place and country*** |
|  |
| * 1. ***\*Website of the institution*** |
|  |
| * 1. ***\*Name of main contact(s)*** |
|  |
| * 1. ***\* Email of main contact(s)*** |
|  |
| 1. ***\*Draft landing page*** *of the K-centre* |
| * 1. ***URL*** *(if available, page must be in English)* |
|  |
| * 1. ***Planned content of the page*** *(If the URL is not yet available, max 500 char)* |
|  |
| PART 2 What the K-centre intends to offer and to whom *Please note*   * *that the information in Part 2 will be published on the CLARIN website and should help potential users finding the expertise they are looking for,* * *that this information will be used in searches on the CLARIN website,* * *that you can leave optional fields blank if they are not applicable to your specific situation* * *and that in section 4 of this document, below this table, you can find a version of this table with lots of explanations and examples to illustrate what sort of answers we expect* |
| 1. ***\*Non-technical prose description*** *of what the centre has to offer (max 500 char, no jargon)* |
|  |
| 1. ***\*Audience(s) targeted*** *by the centre (bulleted list of key words or phrases)* |
|  |
| 1. ***\*Types of services offered*** *(bulleted list of key words and phrases)* |
|  |
| 1. ***Portal (if applicable)****: List and/or descriptions of (groups of) languages for which the centre plans to serve as the main, broad portal covering all aspects of the language or its use (bulleted list languages and/or groups of languages)* |
|  |
| 1. ***Other languages (if applicable):*** *List and/or descriptions of (groups of) languages for which the centre has specific expertise or data, without serving as a main portal as indicated in point 4 above (bulleted list of languages and/or groups of languages)* |
|  |
| 1. *List of* ***modalities*** *covered by the expertise of the centre (bulleted list of key words and phrases)* |
|  |
| 1. *List of* ***linguistic topics*** *covered (bulleted list of key words and phrases)* |
|  |
| 1. *List of* ***language processing topics*** *covered (bulleted list of key words and phrases)* |
|  |
| 1. ***Data types*** *covered (bulleted list of key words and phrases)* |
|  |
| 1. ***CLARIN Resource Families supported, see https://www.clarin.eu/resource-families*** *(bulleted list of key words and phrases, to be selected from the* ***closed list*** *below; delete the ones that are not applicable, don’t add any new ones)* |
| ==CORPORA==   * Computer-mediated communication corpora * Corpora of academic texts * Historical corpora * L2 learner corpora * Legal corpora * Literary corpora * Manually annotated corpora * Multimodal corpora * Newspaper corpora * Oral history corpora * Parallel corpora * Parliamentary corpora * Reference corpora * Sign language resources * Spoken corpora   ==LEXICAL RESOURCES==   * Conceptual Resources * Dictionaries * Glossaries * Language models * Lexica * Wordlists   ==TOOLS==   * Corpus query tools * Normalization * Named entity recognition * Part-of-speech tagging and lemmatization * Tools for sentiment analysis |
| 1. ***Generic topics*** *covered, not connected with specific languages (bulleted list of key words and phrases)* |
|  |
| 1. ***Other key words and phrases*** *that describe the expertise of the centre (bulleted list of key words and phrases)* |
|  |
| PART 3 Compliance with CLARIN requirements for K-centres |
| 1. *\*Confirmation of willingness to register your centre in the CLARIN centre registry (yes/no):* |
| 1. *\*Confirmation that you will provide your services (at least) in English (yes/no):* |
| 1. *\*Confirmation that you will provide both pro-active and re-active services (yes/no):* |
| 1. *\*Confirmation that you will react to queries within 2 working days by means of a human agent (yes/no):* |
| 1. *\*Confirmation that you will keep track of the services provided and report on them on an annual basis following the guidelines provided (yes/no):* |
| PART 4 Quality, sustainability and organisation |
| 1. ***\*Publications*** *(please list at least 5 major recognized publications over the last 3 years covering all areas of expertise of the K-centre)* |
|  |
| 1. ***\*Critical mass*** *(please show that you have sufficient critical mass consisting of at least 3 permanent staff, which may be distributed over different physical locations, not necessarily in the same country); max 500 char* |
|  |
| 1. ***Organisation:*** *If the centre is distributed over different institutions please list all participating institutions (name, location, website), their specific contributions, and provide a description of the way the collaboration between the centres is organized to ensure that you can act as one knowledge centre; indicate which institution and which person will serve as our main contact if different for the main applicant* |
|  |
| PART 5 Any other relevant comments |
|  |

1. **Application form: examples and explanations**

|  |
| --- |
| **K-centre application form – Examples and explanations**  *Items marked with an asterisk (\*) are obligatory* |
| **PART 1 Administrative data** |
| 1. ***\*Proposed full name in English*** *(max 80 characters, the name must contain “CLARIN Knowledge Centre” or “CLARIN K-centre”, and it should reflect the main theme(s) covered by the centre)* |
| Example: CLARIN Knowledge Centre for linguistic diversity and language documentation |
| 1. ***\*Proposed short name or acronym*** *(max 16 characters, no spaces, no special signs other than underscore or dash)* |
| Example: CKLD |
| 1. ***\*Host institution (only 1) and main contact(s) for this application*** |
| * 1. ***\*Name of the institution*** |
| Example: Jozef Stefan Institute |
| * 1. ***\*Place and country*** |
| Example: Ljubljana, Slovenia |
| * 1. ***\*Website of the institution*** |
| Example: <https://www.ijs.si/ijsw/JSI> |
| * 1. ***\*Name of main contact(s)*** |
| Example: Darijo Srna |
| * 1. ***\* Email of main contact(s)*** |
| Example: [darijo.srna@ffzg.hr](mailto:darijo.srna@ffzg.hr) |
| 1. ***\*Draft landing page*** *of the K-centre* |
| * 1. ***URL*** *(if available, page must be in English)* |
| Example: http://www.clarin.si/info/k-centre/ |
| * 1. ***Planned content of the page*** *(If the URL is not yet available, max 500 char)* |
|  |
| **PART 2 What the K-centre intends to offer and to whom** *Please note*   * *that the information in Part 2 will be published on the CLARIN website and should help potential users finding the expertise they are looking for,* * *that this information will be used in searches on the CLARIN website,* * *and that you can leave optional fields blank if they are not applicable to your specific situation* |
| 1. ***\*Non-technical prose description*** *of what the centre has to offer (max 500 char, no jargon)* |
| Example: Information service offering advice on the use of digital language resources and tools for the Swedish language, minority languages in Sweden, the Swedish sign language, Swedish dialects, as well as other parts of the intangible cultural heritage of Sweden in text and speech, as well as language policy and plan |
| 1. ***\*Audience(s) targeted*** *by the centre (bulleted list of key words or phrases)* |
| Examples:  - Computational linguists  - Linguists  - Language teachers  - Historians  - Library staff  - Sociologists  - Citizen scientists  Please make sure that the description above makes clear to each of the audiences you list here why your offerings are relevant for them. |
| 1. ***\*Types of services offered*** *(bulleted list of key words and phrases, select one or more items from the closed list below and delete the others, or add additional ones not covered by this list)* |
| * Access to data * Access to document templates * Access to documentation * Access to tools * Data processing on demand * Depositing * FAQ * Helpdesk * How-to documents * Technical support * Technical support / user assistance * Training * User assistance * User forum * Webhosting |
| 1. ***Portal (if applicable)****: List the languages or groups of languages for which the centre plans to serve as the main, broad portal, covering all aspects of the language or its use (bulleted list of languages and/or groups of languages)* |
| Example 1:  - Belarusian  Example 2:  - Danish  - Danish sign language  Example 3:  - Spanish  - Basque  - Catalan  - Galician |
| 1. ***Other languages (if applicable):*** *List and/or describe languages or groups of languages for which the centre has specific expertise or data, without serving as a main portal as indicated in point 4 above (bulleted list of languages and/or groups of languages)* |
| The focus is not on the languages per se but rather on specific topics or technologies, and the centre has expertise in working with these languages. Examples:  - Under-researched languages and language families  - South Slavic languages  - English  - Dutch  - German |
| 1. *List of* ***modalities*** *covered by the expertise of the centre (bulleted list of key words and phrases, select none, one or more items from the closed list below and delete the others, or add additional ones not covered by this list)* |
| Examples:   * Audio: speech * Audio-visual * Multimodality * Sensor data * Sensor data: bio-signals * Text * Video: gestures * Video: sign language |
| 1. *List of* ***linguistic topics*** *covered (bulleted list of key words and phrases)* |
| By linguistic topics we refer to topics that are directly related to the study of language in a broad sense. Examples:  - Morphology  - Syntax  - Semantics  - Pragmatics  - Stylistics  - Language learning  - Translation studies  - Diachronic language studies  - Field linguistics  - Phonology  - Phonetics  - Language disorders |
| 1. *List of* ***language processing topics*** *covered (bulleted list of key words and phrases)* |
| Examples:  - Basic language processing  - Information extraction  - Machine translation  - Processing of morphologically rich languages  - Speech recognition  - Speech synthesis  - Text mining |
| 1. ***Data types*** *covered (bulleted list of key words and phrases)* |
| By data types we mean other data than collections of natural language in any of the modalities (such as text, or speech, video or sign language recordings). Examples:  - Language models  - Dictionaries  - Typological databases  - Treebanks  - Wordnets  - Linked open data  - Ontologies |
| 1. ***CLARIN Resource Families supported*** *(bulleted list of key words and phrases, select none, one or more items from the* ***closed list*** *below and delete the others, don’t add any new ones).* See for more information <https://www.clarin.eu/resource-families> |
| By a supported resource family we mean that your centre has specific expertise in e.g. the collection, annotation, analysis, use or other exploitation of members of the family.  ==CORPORA==   * Computer-mediated communication corpora * Corpora of academic texts * Historical corpora * L2 learner corpora * Legal corpora * Literary corpora * Manually annotated corpora * Multimodal corpora * Newspaper corpora * Oral history corpora * Parallel corpora * Parliamentary corpora * Reference corpora * Sign language resources * Spoken corpora   ==LEXICAL RESOURCES==   * Conceptual Resources * Dictionaries * Glossaries * Language models * Lexica * Wordlists   ==TOOLS==   * Corpus query tools * Normalization * Named entity recognition * Part-of-speech tagging and lemmatization * Tools for sentiment analysis |
| 1. ***Generic topics*** *covered, not connected with specific languages (bulleted list of key words and phrases)* |
| We use the term generic methods and issues to refer to topics of a very general nature, and that are not related to specific languages. Examples:  - Data management  - Legal issues  - Ethical issues  - Brain image recording  - Machine learning  - OCR  - Visualisation |
| 1. ***Other key words and phrases*** *that describe the expertise of the centre (bulleted list of key words and phrases)* |
| Please use this space for key words and phrases that are characteristic for your expertise, and that are not covered by the other questions in Part 2 of the form. Example:   * Animal language |
| **PART 3 Compliance with CLARIN requirements for K-centres**  (only yes answers expected) |
| 1. *\*Confirmation of willingness to register your centre in the CLARIN centre registry (yes/no):* |
| You will receive a form that you will have to complete for this registration |
| 1. *\*Confirmation that you will provide your services (at least) in English (yes/no):* |
| The landing page and all other pages that are part of the K-centre services should at least be available in English. |
| 1. *\*Confirmation that you will provide both pro-active and re-active services (yes/no):* |
| This means that your K-centre should contain material (or links to such material) that users can immediately use to get access to expertise they need, but that it should also provide a link to a help desk where users can submit queries. |
| 1. *\*Confirmation that you will react to queries within 2 working days by means of a human agent (yes/no):* |
| By this we mean a reply within 2 working days from a human agent, not with a promise that the question will be fully answered, but to make it clear that the question is really being taken care of and to offer the user an opportunity to clarify his/her request. A ticketing system or an automatic reply do not count as a human agent! |
| 1. *\*Confirmation that you will keep track of the services provided and report on them on an annual basis following the guidelines provided (yes/no):* |
| We will provide reporting guidelines. |
| **PART 4 Quality, sustainability and organisation** |
| 1. ***\*Publications*** *(please list at least 5 major recognized publications over the last 3 covering all areas of expertise of the K-centre)* |
| Please make sure that if you list several areas of competence, they should all be covered by one or more items in your list of publications. Please include links. |
| 1. ***\*Critical mass*** *(please show that you have sufficient critical mass consisting of at least 3 permanent staff, which may be distributed over different physical locations, not necessarily in the same country); max 500 char* |
| We want to make sure that the expertise you offer is anchored in your organisation and not just dependent on personal interests of individual members of your team. |
| 1. ***Organisation:*** *If the centre is distributed over different institutions please list all participating institutions (name, location, website), their specific contributions, and provide a description of the way the collaboration between the centres is organized to ensure that you can act as one knowledge centre; indicate which institution and which person will serve as our main contact if different for the main applicant* |
| This is only relevant for distributed K-centres. Example:  The CLARIN Knowledge Centre for SAFMORIL will operate as a distributed virtual centre supported by two different CLARIN member institutions. The Knowledge Centre will be coordinated by the University of Helsinki through its SAFMORIL helpdesk (safmoril@kielipankki.fi) and the main point of contact to CLARIN ERIC is Antti Niemi (antti.niemi@helsinki.fi).  SAFMORIL is supported by the following institutions:  University of Helsinki, Finland, www.helsinki.fi:  - Online Courses for Developing Morphologies, Helpdesk  University of Tromsø, Norway, www.uit.no:  - Proofing Tools, Morphology and Computational Linguistics, Language Repository  In addition, there are other institutions within the CLARIN infrastructure network contributing the work of individual experts to SAFMORIL. |
| **PART 5 Any other relevant comments** |
| Here you can provide additional information relevant for your application.  Examples: Your K-centre is part of or is a spin-off of a project, or it serves also as a centre of expertise in another research infrastructure (with opportunities for collaboration). |