Goethe-Universität | 60325 Frankfurt am Main
Universitätsbibliothek Johann Christian Senckenberg

Prof. Dr. Erhard Hinrichs
Spokesperson Text+
Leibniz-Institut für Deutsche Sprache
R-5, 6-13
68161 Mannheim

Letter of intent in support of the planned consortium Text+

18.09.2020

**Universitätsbibliothek
Johann Christian Senckenberg**

Direktorin

Daniela Poth

Campus Bockenheim
Bockenheimer Landstr. 134-138
60325 Frankfurt am Main

Telefon +49 (0)69 798 39229/ 39230
d.poth@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de

The Specialised Information Service *Fachinformationsdienst (FID) Linguistik* is a joint project of the University Library Frankfurt and the Applied Computational Linguistics lab at the Goethe University Frankfurt. *FID Linguistik* welcomes the plans for a national infrastructure for language- and text-based research data. *FID Linguistik* supports the proposed consortium Text+ and is interested in collaborating on the development and establishment of the relevant infrastructure.

FID Linguistik is a service facility for the studies of General and Comparative Linguistics and the linguistics of single languages at German universities and non-university research institutions. The main objective of the FID is the nationwide information supply for study and research. It plans and develops services in compliance with the needs of the linguistic community taking into consideration various domain-specific resources such as print media, digital publications, scientific online resources as well as research data. The current development focuses on the research data that constitute the empirical basis of Linguistics. We take measures to improve the search for research data and enhance their visibility and accessibility.

The online platform of the FID is the *Lin/gu/is/tik* portal (www.linguistik.de). The search functionality provided on the platform has been continuously enhanced in order to integrate more research data. A recent development and one of the portal's outstanding features is its interlinking with Linked Open Data (LOD). To this end, *FID Linguistik* re-modelled its authority data according to Semantic Web standards and linked them to relevant repositories. In addition, the FID developed methods for the enrichment of language resources with content-related metadata using LOD technologies.

Another step towards better visibility of linguistically relevant research data is the ongoing linking of language corpora with associated secondary literature. For this purpose, we create authority records for mono- or multilingual language corpora given they have a persistent identifier or a website and subsequently link these records to the bibliographic references of the relevant publications.

Gradually, the interlinking will be expanded to include other resource types such as lexical databases and dictionaries.

The FID plans additional technical measures in order to optimise the search for relevant information, its export, and further machine processing. A joint, semantically enriched search index for all modules of the *Lin|gu|is|tik* portal will be established based on a uniform Linked Data format. Thus, we will promote interoperability and enable an easy re-use of the aggregated and generated information.

The efforts of *FID Linguistik* with regard to research data comply very well with the measures planned in the context of the national research data infrastructure (NFDI). The proposed consortium Text+ focuses on research data types that build the foundation of modern linguistic research, i.e. language- and text-based collections, dictionaries and other lexical resources as well as editions. The FID supports the planned actions towards the interlinking of these research data types in a common infrastructure that conforms to the FAIR principles (Findability, Accessibility, Interoperability and Reusability).

Building on our expertise and practical experience in the areas of metadata standards and metadata interoperability, we are ready to collaborate with Text+ on the establishment of a federated metadata infrastructure and the development of metadata and research data search services. *FID Linguistik* intends to participate in the knowledge exchange focusing on solutions for search and data modelling that enhance findability across potentially heterogeneous metadata and data sources. Another starting point for mutually beneficial cooperation can be the interlinking of research data and related publications.
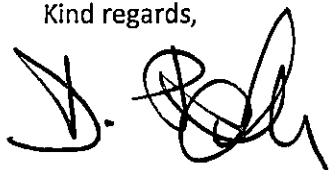
FID Linguistik appreciates the efforts of the initiative Text+ to implement a balance between scientific and operational aspects when developing its data and service portfolio. As a representative of the linguistic community on one side and a provider of information infrastructure on the other, *FID Linguistik* took part in a series of workshops in preparation of the NFDI. Consequently, we are ready to participate in the governance structure of Text+ and thus directly contribute to the establishment of its service portfolio. *FID Linguistik* can join the initiative as a participant in one of the planned Scientific Coordination Committees that will encompass representatives of the research communities. We consider the areas Collections (i.e. language- and text-based collections) and Dictionaries (i.e. dictionaries and other lexical resources) of particular importance to the linguistic research. The FID could also play a consultative role, e.g. by appointing a representative in an advisory board.

FID Linguistik intends to contribute to the dissemination of the services provided by Text+ within the linguistic community by offering NFDI-related information services via the *Lin|gu|is|tik* portal. Furthermore, we can coordinate our activities at the annual conferences of the leading German professional associations in the field of Linguistics and offer joint presentations, information desks or training events. Adopting a needs-oriented approach, Text+ and *FID Linguistik* can collaborate on the development of appropriate methods for training and the organisation of workshops, summer schools, etc.

FID Linguistik regards the development of the NFDI as essential for the future of Linguistics in Germany. We welcome the plans of the proposed consortium Text+ and are confident that combining expertise, exchanging knowledge, and collaborating will lead to synergies and mutual benefits for both platforms and, above all, will be beneficial for the linguistic community.

We are very much looking forward to future collaborations.

Kind regards,

(Daniela Poth)
Library Director