Research Data Infrastructure for
Text- and Language-based Research Data

The Text+ Data Space as Curated and FAIR Access Point to Text- and Language-based Research Data

Text+

FAIR Access to Text- and Language-based Research Data

The Text+ Data Space enhances the findability, searchability, and reusability of humanities research data. It integrates data from various sources using existing mapping and matching techniques, but crucially, it leaves data sovereignty with the original sources. The Text+ Data Space is composed of:

- Text+ Centres: Data and competence centres that provide Research Data Management (RDM) services and form a distributed environment for data, certified according to CTS and/or Nestor.
- Text+ Registry and the Federated Content Search (FCS).
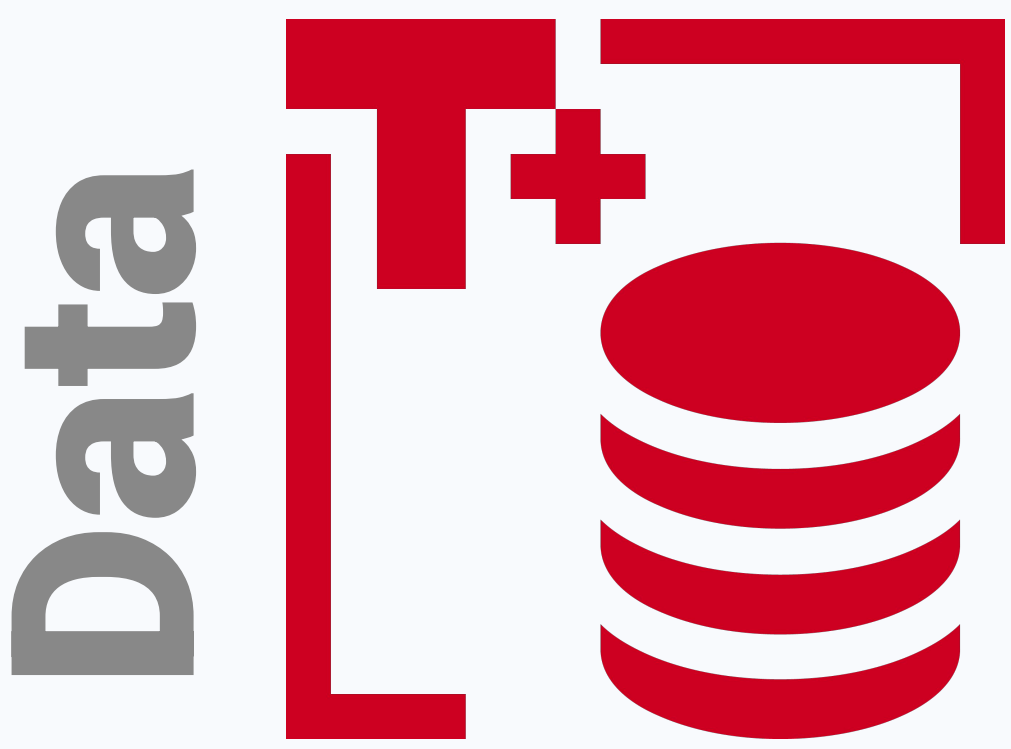
Search & Retrieval in the Text+ Data Space

The distributed data management approach addresses the complexities of data types, disciplines, and the legal status of the data.

The Text+ Registry is a versatile system for describing and cataloguing diverse entities and resources. The Federated Content Search (FCS) is the central Text+ application for directly searching across distributed research content, including the often extensively annotated datasets of the Text+ Centres. The modular concept of the Text+ Centres is scalable and open to both new partners and new data domains.
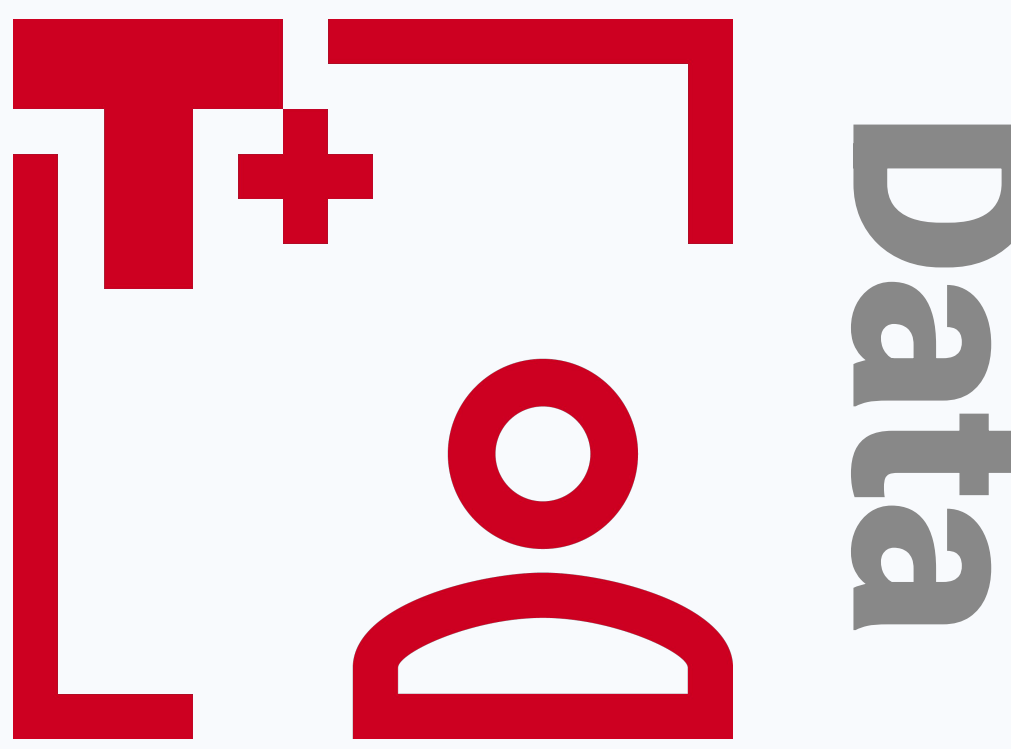
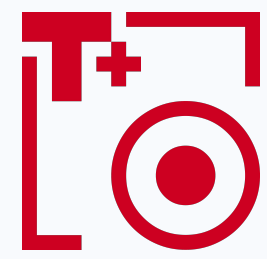## Data sharing
in the Text+ Data Space

Data — Data Owner

Data exchange and data processing along the data value chain (examples)
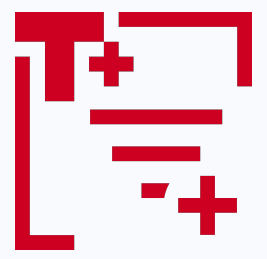
Data — Data User

Centers
Data, Competence

2025 N=34 Text+ forms a distributed infrastructure that is supported by Text+ centers.
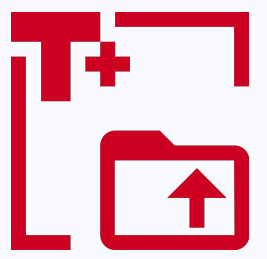
Enrichment
GND, Authority Files

Research data within the Text+ Data Space is intended for authority file enrichment, using GND entities. entityXML and the Text+ GND Agency support this effort.

Data Depositing

Researchers and research projects may hand over their data to Text+ for longterm preservation and FAIRification.
The research data is stored in one of the Text+ Centers.

Search & Retrieval
FCS, Registry

With the Text+ Registry for metadata search over various domains and collections and the Federated Content Search (FCS) for search in structured and full text data, Text+ offers a powerful Search & Retrieval bench.

Interoperability Layer

**LLM-Hub
Knowledge Graphs
Ontology Harmonization
Certificates (CTS, nestor)
Standards (DIN)**

T+ Portal

Federated Content Search
Search in full- and structured data. Various extensions are under development as part of Text+. These include the extended support of Lexical Resources, entity search, as well as the simplified integration of FCS in new application contexts. The search platform can be used both as a stand-alone web application (FCS Aggregator) and via the integration into the Text+ website.

Registry
Metadata-based directory of data and services. The Text+ Registry is a comprehensive information and reference system that provides easy access to resources relevant for research and teaching in the field of language and text-based disciplines.

Training- and educational activities
Ranges from consulting services to events in the field of "data and tool literacy" – these can be conferences, workshops or summer schools, but also regular events such as the Text+ IO Lectures or the Research Rendezvous – to the publication of guidelines and training material or application examples, to the provision of detailed documentation on the use of resources.

Text+ maintains numerous collaborations with other infrastructures and communities. They serve either to adapt its own offerings, for example, regarding standards or interoperability, or to expand the user base of its own offerings.

This is technically implemented, for instance, in a joint helpdesk service with the Specialized Information Services (Fachinformationsdienste), collaboration within the group of humanities and cultural studies consortia (Humanities@NFDI), for example, in ontology harmonization, or intensive participation in the GND community concerning authority data.

Text+ also maintains close ties with the European Research Infrastructure Consortial (ERIC) CLARIN and DARIAH. These proven relations date back to the national CLARIN, DARIAH and CLARIAH projects, which today form a part of the Text+ infrastructure and service offerings.

Utilize and Preserve Text- and Language-based Research Data

Text+ provides an extensive collection of text collections and corpora, editions and lexical resources for re-use as part of the NFDI. In addition to its own data and services, the service portfolio also includes the option of integrating and preserving research data from the community in the long term.

Data Search | Data Depositing | Services | Consulting/Helpdesk

Text+ and the NFDI
As part of the National Research Data Infrastructure (NFDI), the DFG is funding a total of 26 consortia from the four fields of humanities and social sciences, engineering sciences, life sciences and natural sciences. The Text+ consortium is one of six consortia from the humanities, including Berd@NFDI, KonsortSWD, NFDI4Culture, NFDI4Memory, NFDI4Objects. While there are many cross-consortia activities via the NFDI sections, the Text+ consortium is in particularly close dialogue with the humanities and social sciences consortia.

Showcase
Using the example of the transfer of research data stored on discs, the outlined workflow provides an impression of data integration in Text+. It ranges from initial contact and consulting to the integration, re-utilisation and long-term preservation of the data (registry, FCS, MonaPipe, GND agency, TextGrid repository).
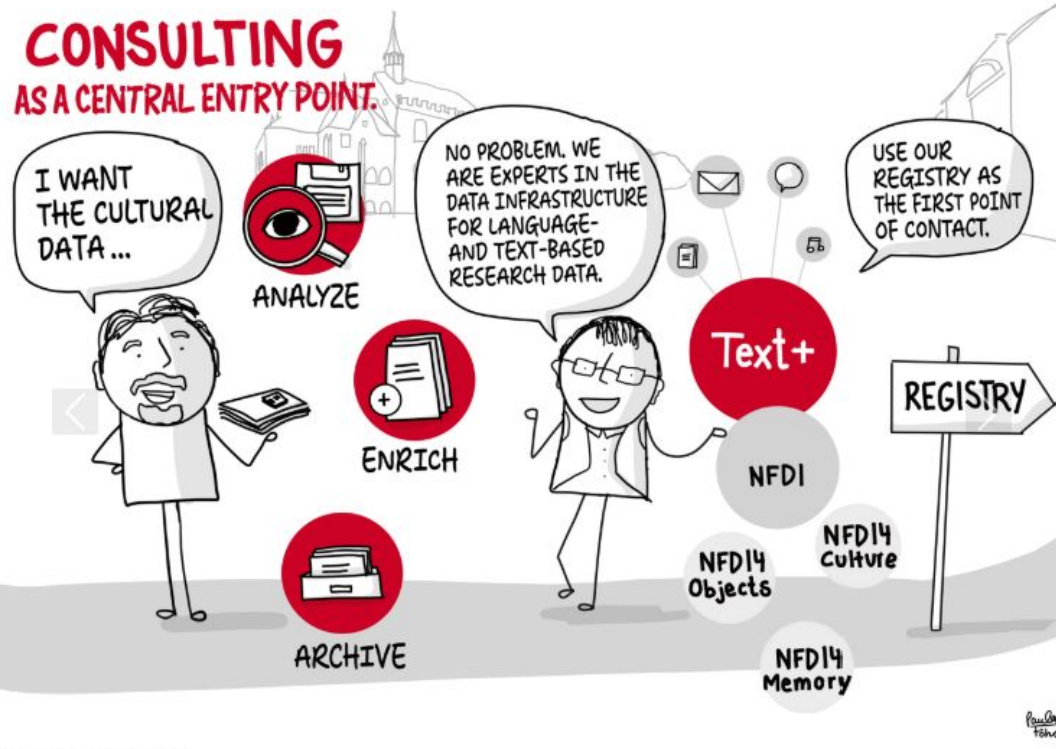
CONSULTING AS A CENTRAL ENTRY POINT

Illustration: Paula Fohr

Text+ Blog
Bekanntgabe der neuen Kooperationsprojekte von Text+ für 2026
Antonina Werthmann
21. October 2025 09:58

Text+ fördert jährlich ausgewählte Kooperationsprojekte – mit dem Ziel, das Angebot an Daten und Services kontinuierlich zu erweitern und der wissenschaftlichen Gemeinschaft langfristig zugänglich zu machen. Wir freuen uns, für 2026 die Förderung folgender fünf Projekte bekanntzugeben:

open blog

current events

| Fri–Fri | 14:00–16:00 |
| 24.–21. October – November 2025 | Digitale Editionen und … Virtuell |

| Thu | 10:00–11:30 |
| 20. November 2025 | Werkstattreihe Standardisierung: BERIA Collection Virtuell |

| Tue | 09:00–10:00 |
| 25. November 2025 | Text+ Research Rendezvous Virtuell |

more events

Data Depositing Service
As part of the NFDI, Text+ serves as a hub for text- and language-based research data. The Text+ Centers offer the academic community a data depositing service. These datasets are not only archived for the long term but are also made available for reuse in accordance with legal regulations. Access is open to all researchers with an academic affiliation.

Cooperation within NFDI
Governance, Sections, Working Groups
Workshops and Series with Other Consortia

International Cooperation
ERICs
EOSC
DINI
RDA
ISO, DIN

Consulting/Helpdesk
Events, Blog, Workshops

Base4NFDI is an integral part of the Text+ architecture. Text+ uses and integrates - wherever feasible - individual services into the architecture and service portfolio. Currently (2025) these services are expecting integration in our infrastructure or are offered to our user community through our portal: DMP4NFDI is supported with a Text+ incubator, IAM4NFDI (identity and access management) as extension of the AcademicCloud AAI, which is used by Text+, Jupyter4NFDI, a service with roots in Text+ and as data processing tool in addition to Text+ provided tool chains/pipelines, PID4NFDI for persistent identification of resources, TS4NFDI as terminology hub to utilize other domains' terminologies

The European Open Science Cloud (EOSC) provides an important (potential) infrastructure context for Text+. Many services and first and foremost the discoverability of research data are intended to be used by European researchers beyond the national NFDI domain.

Text+ employs the SSH Open Marketplace as a central catalogue for the dissemination of its services and training materials. Resources are gathered by the Text+ portal and subsequently harvested through nightly builds, to be published as browsable catalogues.

This cooperation yields benefits not only for Text+, but also for the wider European humanities community, by enabling the use of Text+ resources.

DARIAH-EU — Digital Research Infrastructure for the Arts and Humanities | CLARIN | CLARIAH-DE

SSH Open Marketplace — Social Sciences and Humanities Open Marketplace | eosc

Text+ is part of
nfdi Nationale Forschungsdaten Infrastruktur

Contact: office@text-plus.org
Title: The Text+ Data Space as Curated FAIR Access Point to Text- and Language-based Research Data
Authors: Text+ Task Area Coordinators
Date: 27. November 2025, Bonn