



Text+ Consortium Progress Reports National Research Data Infrastructure (NFDI)

B-1 Progress Report Part 1, for publication

Version	1.0
Version date	2024-09-30
Project	Text+: Sprach- und textbasierte Forschungsdateninfrastruktur 2021-10-01 to 2024-09-30
Grant Number DFG	460033370
DFG Reference Number	NFDI 25/1

1 GENERAL INFORMATION	1
2 SUMMARY.....	2
3 COMPOSITION OF THE CONSORTIUM.....	4

1 General Information

- **Name of the consortium**

English: Text+: Language- and text-based Research Data Infrastructure

German: Text+: Sprach- und textbasierte Forschungsdateninfrastruktur

- **Research domains or research methods addressed by the consortium**

Language and text data are used in many disciplines addressed by Text+, especially the following:

RB No	DFG Review Board	DFG Subject Area	DFG Subject Area Number
1.11	Ancient Cultures	Greek and Latin Philology	1.11-02
1.14	Linguistics		
		General and Comparative Linguistics, Experimental Linguistics, Typology, Non-European Languages	1.14-01
		Individual Linguistics, Historical Linguistics	1.14-02
		Applied Linguistics, Computational Linguistics	1.14-03
1.15	Literary Studies		
		German Medieval Studies (Medieval German Literature)	1.15-01
		German Literary and Cultural Studies (Modern German Literature)	1.15-02
		European and American Literary and Cultural Studies	1.15-03
		General and Comparative Literary Studies; Cultural Studies	1.15-04
1.16	Social and Cultural Anthropology, Non-European Cultures, Jewish Studies and Religious Studies	All language and text data based research	

1.18	Philosophy		
		History of Philosophy	1.18-01
		Theoretical Philosophy	1.18-02
		Practical Philosophy	1.18-03

- URL of the consortium website and repositories used for publishing output
<https://www.text-plus.org>

2 Summary

Text+ follows the mission for the text- and language-oriented humanities and social sciences to extensively leverage the possibilities of digitisation in their research, teaching, and transfer, establishing a common data culture. Text+ empowers text and language data and access to such data. Access to digital sources should become a standard. Text+ enhances the digital literacy of researchers and encompasses the diversity of the research community and builds on their participation. To achieve this, we promote inter- and transdisciplinarity as well as innovation – for instance in the field of artificial intelligence – through the integration of infrastructure and research.

The consortium started with 30 partners, all making significant in-kind contributions. All the original partners remain, and three new partners have joined, expanding Text+ in terms of expertise, resources and community collaboration. Collaborative projects have been an effective means of disseminating Text+ services, resulting in a large number of applications and funded projects that continue to expand Text+.

Communities of interest play an active role in the governance of Text+ through community-based panels. The consortium benefits the communities by providing data, services and knowledge sharing, integrating diverse research efforts and developing new methodologies. Collaboration with other consortia and joint contributions to relevant conferences enhance these benefits.

Text+ is actively involved in the development and establishment of the National Research Data Infrastructure (NFDI) in Germany. We participate in various NFDI events, meetings and task forces to contribute to the mission and vision of the NFDI and cooperate with other consortia.

Text+ maintains close links with European Research Infrastructure Consortia (ERICs) such as CLARIN and DARIAH. We also collaborate with the European Open Science Cloud (EOSC) and other European cloud services. Text+ is involved in the exchange of RDM practices and standards through various international networks and organisations.

Text+ addresses three Grand Challenges: mastering the diversity of research data, supporting humanities scholars, and developing a governance model that supports shared responsibility and consensus building. The Text+ Data Space integrates data from different sources, improving the discoverability, searchability and reusability of humanities research data.

Success Stories

- Integration of new partners and data sets, such as the Deutscher Sprachatlas.
- Development of the Text+ Registry, Federated Content Search (FCS) and Text+ Consulting for improved data findability, access and interoperability.
- Successful cooperation projects that broaden the diversity of data types handled by the infrastructure.

Challenges to be Addressed

- Technological Advancements: Integrating large language models (LLMs) and addressing legal constraints related to data licensing and copyright laws.
- Structural Changes: Adapting to developments within the NFDI, such as the ones provided through Base4NFDI, developing basic services for the NFDI.
- Diversity and Inclusion: Ensuring gender balance and diversity across disciplines, career levels, and academic interests.

Text+ has made significant progress in expanding its consortium, integrating different data types and contributing to the development of the NFDI. The consortium's efforts in international networking, governance and community involvement have been instrumental in its success.

3 Composition of the consortium

▪ Applicant institution

Applicant institution	Location	GND-ID	ROR	Duration
Leibniz-Institut für Deutsche Sprache (IDS)	Mannheim	118318106X	https://ror.org/00hv-wkt50	10/21-09/26

▪ Spokesperson

Spokesperson	ORCID-ID	Institution, location	Duration
Professor Dr. Erhard W. Hinrichs ¹	0009-0006-4192-7779	Leibniz-Institut für Deutsche Sprache (IDS)	10/21-09/24
Professor Dr. Andreas Witt	0000-0002-0299-5713	Leibniz-Institut für Deutsche Sprache (IDS)	10/24-09/26

▪ Co-applicant institutions

Co-applicant institutions	Location	GND-ID	ROR	Duration
Berlin-Brandenburgische Akademie der Wissenschaften	Berlin	2131094-4	https://ror.org/05jgg9443	10/21-09/26
Deutsche Nationalbibliothek	Frankfurt am Main	10140798-1	https://ror.org/01n7gem85	10/21-09/26
Georg-August-Universität Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen	Göttingen	2024315-7	https://ror.org/01y9bpm73	10/21-09/26

¹ Prof. Erhard Hinrichs (Leibniz-Institute for the German Language) has been the spokesperson of the Text+ consortium since its inception. On October 1, 2024, Prof. Andreas Witt (Leibniz-Institute for the German Language), who was already serving as the consortium speaker in the NFDI association, succeeded him.

Nordrhein- Westfälische Akademie der Wissenschaften und der Künste	Düsseldorf	10370564-8	https://ror.org/03aym6q96	10/21-09/26
--	------------	------------	---	-------------

▪ **Co-spokespersons**

Co-spokespersons	ORCID-ID	Institution, location	Task area(s)	Duration
Dr. Alexander Geyken		Berlin- Brandenburgische Akademie der Wissenschaften, Berlin	Lexical Resources	10/21-09/26
Dr. Peter Leinen	0000-0002-3014-000X	Deutsche Nationalbibliothek, Frankfurt am Main	Collections	10/21-09/26
Professor Dr. Andreas Speer	0000-0002-7263-1405	Universität zu Köln, Köln	Editions	10/21-09/26
Regine Stein	0000-0003-3406-5104	Georg-August- Universität Göttingen, Göttingen	Infrastructure/ Operations	10/21-06/23
Professor Dr. Philipp Wieder ²	0000-0003-3406-5104	Georg-August- Universität Göttingen, Göttingen	Infrastructure/ Operations	06/23-09/26

² Until 7.6.2023 Regine Stein (Georg-August-Universität Göttingen, Göttingen State and University Library) was the Co-Spokesperson and Operations Speaker of Text+ and lead of the task area IO. From 7.6.2023 on, Philipp Wieder (GWDG and Georg-August-Universität Göttingen) took over her roles in Text+.

▪ **Participants**

Participating institutions	Location	GND-ID	ROR	Duration
Akademie der Wissenschaften in Hamburg	Hamburg	16057150-9	https://ror.org/01rd01t06	10/21-09/26
Akademie der Wissenschaften und der Literatur Mainz	Mainz	35487-9	https://ror.org/01kdxra28	10/21-09/26
Albert-Ludwigs-Universität Freiburg	Freiburg	2024338-8	https://ror.org/0245c6223	10/21-09/26
Bayerische Akademie der Wissenschaften	München	2005521-3	https://ror.org/001rdaz60	10/21-09/26
Bergische Universität Wuppertal	Wuppertal	64944-2	https://ror.org/00613ak93	01/24-09/26
Deutsche Akademie der Naturforscher Leopoldina - Nationale Akademie der Wissenschaften	Halle (Saale)	35672-4	https://ror.org/03v0k6341	10/21-09/26
Deutsches Literaturarchiv Marbach	Marbach am Neckar	10155895-8	https://ror.org/05x75gh12	10/21-09/26
Eberhard Karls Universität Tübingen	Tübingen	36187-2	https://ror.org/03a1kwz48	10/21-09/26
Forschungszentrum Jülich	Jülich	5008462-8	https://ror.org/02nv7yv05	10/21-09/26
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen	Göttingen	9415-8	https://ror.org/00cd95c65	10/21-09/26
Heidelberger Akademie der Wissenschaften	Heidelberg	2004744-7	https://ror.org/02d9f9b44	10/21-09/26
Herzog August Bibliothek Wolfenbüttel	Wolfenbüttel	8989-8	https://ror.org/02y5mkh60	10/21-09/26
Hochschule Darmstadt	Darmstadt	10133128-9	https://ror.org/047wbd030	10/21-09/26

Hochschule Wismar - University of Applied Sciences Technology, Business and Design	Wismar	5077582-0	https://ror.org/00s1ckt27	01/24-09/26
Julius-Maximilians- Universität Würzburg	Würzburg	36189-6	https://ror.org/00fbnyb24	10/21-09/26
Klassik Stiftung Weimar	Weimar	6045529-9	https://ror.org/02ad32n85	10/21-09/26
Ludwig-Maximilians- Universität München	München	2002585-3	https://ror.org/05591te55	10/21-09/26
Max Weber Stiftung - Deutsche Geisteswissenschaftliche Institute im Ausland	Bonn	1028661126	https://ror.org/02esxtn62	10/21-09/26
Niedersächsische Akademie der Wissenschaften zu Göttingen	Göttingen	16268620-1	https://ror.org/04hsa7a08	10/21-09/26
Otto-Friedrich- Universität Bamberg	Bamberg	2056127-1	https://ror.org/01c1w6d29	10/21-09/26
Philipps-Universität Marburg	Marburg	2001630-X	https://ror.org/01rdrb571	07/23-09/26
Salomon Ludwig Steinheim-Institut für Deutsch-jüdische Geschichte e.V.	Essen	212562-6	https://ror.org/01btqj997	10/21-09/26
Sächsische Akademie der Wissenschaften zu Leipzig	Leipzig	37863-X	https://ror.org/03zxjck07	10/21-09/26
Technische Universität Darmstadt	Darmstadt	2165342-2	https://ror.org/05n911h24	10/21-09/26
Technische Universität Dresden	Dresden	2024114-8	https://ror.org/042aaky30	10/21-09/26
Universität Duisburg- Essen	Duisburg	10050375-5	https://ror.org/04mz5ra38	10/21-09/26
Universität Hamburg	Hamburg	35534-3	https://ror.org/00g30e956	10/21-09/26

Universität Paderborn	Paderborn	64857-7	https://ror.org/058kzsd48	10/21-09/26
Universität Trier	Trier	26331-X	https://ror.org/02778hg05	10/21-09/26
Universität des Saarlandes	Saarbrücken	2024427-7	https://ror.org/01jdp yv68	10/21-09/26
Universität zu Köln	Köln	2024231-1	https://ror.org/00rcxh774	10/21-09/26

Participating individuals	ORCID-ID	Institution, location	Duration
Professor Dr. Andreas Henrich	0000-0002-5074-3254	Otto-Friedrich-Universität Bamberg, Bamberg	10/21-09/26
Professorin Dr. Verena Klemm	0009-0006-1780-770X	Universität Leipzig, Leipzig	10/21-09/26
Professorin Vivien Petras, Ph.D.	0000-0002-8113-1509	Humboldt-Universität zu Berlin, Berlin	10/21-09/26
Professorin Dr. Andrea Rapp	0000-0003-4933-4397	Technische Universität Darmstadt, Darmstadt	10/21-09/26
Professorin Dr. Sandra Richter	0000-0003-0641-5349	Deutsches Literaturarchiv Marbach	10/21-09/26
Professorin Dr. Ingrid Schröder	0000-0001-6855-9183	Universität Hamburg, Hamburg	10/21-09/26
Professorin Dr. Elke Teich	0000-0002-9936-7173	Universität des Saarlandes, Saarbrücken	10/21-09/26