Assessment Information

CoreTrustSeal Requirements 2020–2026

Repository:	CLARIAH-IATEXT
Website:	https://clariah.iatext.ulpgc.es
Certification period:	July 31, 2026 - 30 July 2029
Requirements version:	CoreTrustSeal Requirements 2026-2028
This repository is owned by:	**Instituto Universitario de Análisis y Aplicaciones Textuales**

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background Information & Context

Re3data Identifier

Please fill your Re3data identifier from the website: https://www.re3data.org/

Response:

r3d100014720

Reviews

Reviewer 1:

Comments:

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Repository type

Select a repository type:

Response:

- Specialist repository

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Overview

Provide a short overview of key characteristics of the repository, reflecting the repository type selected. This should include information about the scope and size of data collections, data types and formats. Further contextual information may also be added.
When you have chosen Specialist repository in the former question (Repository type) please add your domain(s) and/or discipline(s) in this section.

Response:

The CLARIAH-IATEXT (1) repository is a specialized research data repository for Spanish linguistic resources, language technologies, and Digital Humanities applications, hosted at the Instituto Universitario de Análisis y Aplicaciones Textuales (IATEXT) (2) of the University of Las Palmas de Gran Canaria (3). The repository provides sustainable long-term access to diverse linguistic resources including Spanish textual corpora, linguistically annotated corpora, lexical databases, dictionaries, natural language processing tools, and Digital Humanities applications. It operates within CLARIAH-ES, Spain's node in the European research infrastructures CLARIN and DARIAH, which support digital research in the social sciences, arts, and humanities. 1 https://clariah.iatext.ulpgc.es 2 https://iatext.ulpgc.es 3 https://www.ulpgc.es

The repository covers 43 subject areas (https://clariah.iatext.ulpgc.es/browse/subject), grouped into the following thematic areas: (1) Linguistics and language analysis (18 resources): grammar, syntactic analysis, textual analysis, accentuation/stress marking, atonic and tonic syllables, syllabification, conjugation forms, lemma and word forms, prefix/suffix information for verb forms, verb, lexical relations, lexicon, Spanish lexicon, neologism, and written English. (2) Numerals and number systems (6 resources): cardinal, decimal, ordinal and Roman numerals, number, and number converter to text. (3)

Classical studies and Latin language (11 resources): Latin, classical philology, Dictionary of Latin Collocations, epigraphy, florilegium, Progymnasmata, roman-empire, 16th century, and 18th-century. (4) Digital Humanities, collections and repository infrastructure (26 resources): DARIAH Resource, database, literary and museum collections, repositories and collections, bibliographic information, galleries and museum studies, archeology, and morphology.

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Designated Community

Provide a clear definition of the Designated Community

Response:

As part of the CLARIAH-ES infrastructure, CLARIAH-IATEXT serves a broad and diverse community engaged with Spanish linguistic and textual resources. Its users include researchers, teachers, and students in linguistics, textual studies, history, social sciences, education, and Digital Humanities, including postgraduate and doctoral researchers. The repository also supports Spanish NLP developers, language educators, and professionals in the language industry, as well as employees involved in non-academic and commercial research. In addition, it extends cultural and educational institutions, such as museums and archives, and the wider public interested in Spanish linguistic resources and Digital Humanities. For further information about the repository, see the Evidence link: CLARIAH-IATEXT.

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Levels of Curation and Preservation

Select all relevant types from:

Response:

- A. Active preservation
- C. Initial Curation
- D. Deposit Compliance

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Levels of Curation and Preservation - Explanation

Please add the description for your Level(s) of Curation.

Response:

CLARIAH-IATEXT disseminates data in the form in which it is deposited. Prior to ingestion, each submission is subjected to a concise assessment of suitability and to the validation or completion of essential metadata elements. The repository employs the standard DSpace workflow, which systematically guides depositors through the sequential entry of metadata. Submissions are reviewed by curators, who may request amendments prior to publication. No conversion of file formats or advanced data enhancement is undertaken. Where appropriate, additional curation procedures may be arranged through specific agreements or dedicated funding.

The curation approach is based on the CLARIN-DSpace platform, a fork of DSpace optimised for language resource repositories and specifically designed to meet the requirements of CLARIN B-Centre certification. CLARIN-DSpace provides two levels of digital preservation: (1) bit-level preservation, which guarantees data integrity through checksums and regular backups; and (2) functional preservation, which includes format migration and metadata enrichment to ensure long-term accessibility. The platform is actively maintained by LINDAT/CLARIN and used by more than ten CLARIN B-Centres that run the same CLARIN-DSpace codebase, enabling collaborative development and mutual support.

Regarding active preservation, the repository currently performs regular backups to the SCAYLE supercomputing infrastructure (Supercomputación Castilla y León), ensuring long-term redundancy and data integrity. Active preservation measures are fully implemented and include monitoring of file integrity, maintenance of metadata standards, and planning for format migration as required by technological changes. These actions ensure that all deposited resources remain accessible and usable by the designated community over time.

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Cooperation and outsourcing to third parties, partners and host organisations

Please describe any cooperation and outsourcing to third parties, partners and host organisations.

Response:

Although CLARIAH-IATEXT is managed directly by the applicant institution (IATEXT), several external partners are essential to ensure long-term sustainability, technical robustness, and security.

Host organisation: The repository is hosted by the physical and network infrastructure of the University of Las Palmas de Gran Canaria (ULPGC), which provides network connectivity, IT services and technical support, institutional storage, and SSL certification (see Evidence links: ULPGC and ULPGC-IT). A Service Level Agreement is in place with ULPGC IT Services.

Federated identity: User authentication relies on federated identity infrastructures: the Federación de Identidad Académica Española (SIR/REDIRIS) for authentication (see Evidence link: REDIRIS), the ULPGC Identity Service Provider (IdP) for transnational access (see Evidence link: REDIRIS - ULPGC), and eduGAIN for ensuring international interoperability (see Evidence link: eduGAIN).

CLARIN infrastructure: The repository uses the CLARIN Concept Registry (CCR) as its controlled vocabulary system (see Evidence link: CLARIN-Registry), and Handle.net to assign persistent identifiers to resources (see Evidence link: HANDLE). CLARIAH-IATEXT participates in CLARIN ERIC as a certified K-Centre (Knowledge Centre providing specialised expertise and services in computational linguistics; see Evidence link: K-CENTRE and the K-Centre-Certificate.pdf included in the Attachments) and is a candidate for B-Centre certification (a certified data repository meeting requirements for long-term preservation and access to language resources, including CoreTrustSeal certification; see Evidence link: CLARIN ERIC).

National consortium: A Memorandum of Understanding is in place with CLARIAH-ES (see Evidence link: CLARIAH-ES and the Memorandum CLARIAH-ES.pdf included in the Attachments), the Spanish national research infrastructure consortium that brings together CLARIN and DARIAH institutions.

Long-term preservation: For off-site data backup and long-term preservation redundancy, the repository relies on SCAYLE (Supercomputación Castilla y León), a public supercomputing centre (see Evidence link: SCAYLE). Two formal agreements govern this cooperation: (1) the Memorandum of Understanding documenting SCAYLE's adhesion to the CLARIAH-ES infrastructure (Memorandum SCAYLE.pdf, included in the Attachments), and (2) the agreement establishing CLARIAH-IATEXT's responsibilities as a user of SCAYLE infrastructure (Memorandum SCAYLE 1.pdf, included in the Attachments).

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Applicants renewing their CoreTrustSeal certification: summary of significant changes since last application

CoreTrustSeal certification has an expectation of continuous improvement over time. Repositories undergoing recertification should summarise key changes since their last application, including any changes to their technical systems, Designated Community or governance/funding, as well as progress on Requirements. This summary should specify the area of change, provide a brief description, indicate the timing, and outline the impact on the repository.

Response:

N.A.

Reviews

Reviewer 2:

Comments:

Reviewer 1:

Comments:

Organisational Infrastructure

R1 Mission & Scope - R01

R01. The repository has an explicit mission to provide access to and actively preserve digital objects.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

CLARIAH-IATEXT repository has the mission of providing sustainable and long-term access to linguistic resources and language technologies for Spanish, supporting research in linguistics, textual analysis, and digital humanities. Its primary goal is to preserve, curate, and disseminate high-quality research data that serve the Spanish-speaking academic community and foster the advancement of knowledge in these disciplines. The mission statement is publicly available (see Evidence link: CLARIAH-IATEXT Mission Statement).

The repository promotes the submission of data in widely usable and interoperable formats suitable for language-related computational environments. Its scope covers diverse data types, including Spanish textual corpora of different varieties and registers, linguistically annotated corpora, dictionaries and lexical databases, NLP tools, discourse analysis data, and resources for Spanish language teaching, primarily in text-based formats. Although image-based, geolocation-based, or sound-based data may also be included where relevant.

To ensure scientific integrity and usability, only data that meet specific inclusion criteria are accepted: relevance to research in Hispanic linguistics or digital humanities, verifiable scientific quality, sufficient documentation to enable reuse, clearly established distribution rights, and, when applicable, compliance with CLARIN standards (see Evidence link: CLARIAH-IATEXT (DATA)).

The repository is firmly committed to the active preservation of its resources for a minimum period of ten years, implementing format migration when necessary to maintain long-term accessibility, ensuring the maintenance of persistent identifiers (Handles), and updating metadata in accordance with evolving standards. Furthermore, CLARIAH-IATEXT establishes succession plans to guarantee the continued preservation and accessibility of its resources in the event of institutional changes, thus ensuring the repository's sustainability and reliability over time (see Evidence link: CLARIAH-IATEXT (INFRASTRUCTURE).

Links:

- CLARIAH-IATEXT (INFRASTRUCTURE)
- CLARIAH-IATEXT (INFO)
- CLARIAH-IATEXT (DATA)
- CLARIAH-IATEXT Mission Statement

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R2 Rights Management - R02

R02. The repository maintains all applicable rights and monitors compliance.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

Upon first login or when submitting a new item, the depositor must accept a Distribution License Agreement (see Evidence link: CLARIAH-IATEXT (LICENCE)), which defines the repository's rights and obligations, confirms that the depositor holds the necessary rights to submit the data and authorizes the repository to distribute it on their behalf. All submissions are subsequently reviewed by a staff member.

Furthermore, each new resource or item must be associated with one of the available licences (see Evidence link: CLARIAH-IATEXT (LICENCE 2)), which can be either open or restricted. The available licence types cover a range of access levels: publicly available licences (e.g., Creative Commons, GNU GPL, MIT) allow free access and download by any user; academic licences restrict access and use to members of academic institutions and researchers, requiring institutional authentication; and restricted licences limit access to specific user groups as agreed with the depositor, with access granted upon request and verification.

By using the repository search and download functions, users agree to its Terms of Service (see Evidence link: CLARIAH-IATEXT (TERMS)), which set acceptable use, citation/attribution, and consequences of breach. Usage restrictions are made clear through visual indicators of licence openness (e.g., public, academic, restricted). Embargoes can be applied, so items are archived immediately and made public after a specified date. Where needed, custom end-user licences, including commercial use agreements, can be arranged directly with repository staff.

The author remains the owner of the data (the repository keeps a copy and takes good care of it under the agreed terms). In case of misuse or breach of licence/ToS terms, access may be suspended, and the research community informed as appropriate. Please note that the link to the commercial use licence agreement form within the Terms of Service page is currently under maintenance; commercial use requests should be directed to the repository team at clariah-iatext@ulpgc.es.

Links:

- CLARIAH-IATEXT (LICENCE 2)
- CLARIAH-IATEXT (LICENCE)
- CLARIAH-IATEXT (TERMS)

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R3 Continuity of Service - R03

R03. The Repository has a plan to ensure ongoing access to and preservation of its data and metadata.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The CLARIAH-IATEXT repository is firmly committed to the sustainable provision of long-term preservation and access services, ensuring the continued availability of its linguistic resources and language technologies for the research community (see Evidence link: CLARIAH-IATEXT (INFO)).

Funding and financial sustainability: The repository's financial sustainability is supported through a combination of institutional, national, and international funding sources (see also the Funding CLARIAH-IATEXT.pdf included in the Attachments). Current funding derives from the base budget of IATEXT (University of Las Palmas de Gran Canaria), participation in national and European research projects, financial support from CLARIAH-ES through the Ministry of Science and Innovation, and contributions from ULPGC for infrastructure and IT services. Looking ahead, the financial projection includes an institutional commitment from ULPGC for at least five years, the planned renewal of CLARIAH-ES funding in four-year cycles, and the diversification of funding sources through competitive projects. The main operational costs covered include salaries for technical and curation staff (equivalent to two full-time positions), maintenance of storage infrastructure and servers, bandwidth and network services, software licences, and staff training and professional development.

Off-site backup and preservation continuity: Long-term data redundancy and preservation continuity are ensured through a formal cooperation with SCAYLE (Supercomputación Castilla y León), a public supercomputing centre that is part of the CLARIAH-ES infrastructure. This cooperation is governed by two formal agreements included in the Attachments: (1) Memorandum SCAYLE.pdf, documenting SCAYLE's adhesion to the CLARIAH-ES infrastructure; and (2) Memorandum SCAYLE 1.pdf, establishing CLARIAH-IATEXT's responsibilities as a user of SCAYLE infrastructure for off-site backup.

Contingency plans: The repository has established contingency plans to ensure continuity under various circumstances. In the event of a funding reduction, priority will be given to core preservation functions over value-added services, collaboration with other CLARIAH-ES centres will be strengthened to share resources, and alternative funding opportunities will be sought. In case of organisational changes at IATEXT, responsibility for the repository would be transferred to another department within ULPGC, with key personnel retained through institutional contracts to maintain operational stability. In the unlikely event of cessation of operations, a succession plan is in place to ensure the transfer of data and metadata to another CLARIN centre—identified candidates include CLARIN-ES Barcelona and DIGITAL.CSIC—following a documented migration procedure that guarantees the preservation of digital objects and the maintenance of persistent identifiers through appropriate redirection.

Finally, CLARIAH-IATEXT is committed to a minimum of ten years of active preservation, with the long-term goal of indefinite preservation if there is community demand. The repository's sustainability and operational effectiveness are reviewed every three years to ensure that preservation strategies, infrastructure, and funding mechanisms remain aligned with evolving needs and best practices.

Written agreements with CLARIN centres: CLARIAH-IATEXT participates in the CLARIN infrastructure as a certified K-Centre, as documented in the K-Centre-Certificate.pdf included in the Attachments. Regarding the formal Centre Agreement governing policies, guidelines, procedures, and obligations between CLARIN centres: this agreement has not yet been signed, as CoreTrustSeal certification is an explicit prerequisite for its formalisation. The present application is precisely the required step to fulfil that condition. Once CoreTrustSeal certification is obtained, CLARIAH-IATEXT will proceed to sign the formal CLARIN Centre Agreement, which will establish the mutual obligations, service policies, and operational guidelines governing its participation in the CLARIN infrastructure as a certified B-Centre. This is endorsed by the Letter CLARIN 2026-01-13.pdf included in the Attachments (endorsement from the Spanish National Coordinator of CLARIN, Dr. Germán Rigau). The national consortium framework is further formalised through the Memorandum of Understanding with CLARIAH-ES (Memorandum CLARIAH-ES English.pdf, included in the Attachments).

A formal **Statement of Institutional Viability and Service Continuity** has been prepared in English and is provided as supporting evidence (Statement_Institutional_Viability_IATEXT_English.pdf). The document, signed digitally by the Director of the University Institute of Textual Analysis and Applications (IATEXT) and the Head of the CLARIAH-ES Canary Islands Node at the University of Las Palmas de Gran Canaria, demonstrates the institutional robustness of IATEXT and outlines the succession mechanisms in place within the CLARIAH-ES infrastructure.

The statement clarifies the continuity-relevant provisions of the CLARIAH-ES General Protocol of Action (Memorandum), including: the stable inter-institutional collaboration framework binding all 12 signatory nodes; the Monitoring Committee responsible for coordinated contingency management; the explicit provision that the Protocol remains in force for all remaining parties should any individual node withdraw; and the requirement that unfinished activities be concluded in an orderly manner under the Committee's supervision. Together, these mechanisms ensure the preservation of data, metadata, and continuity of access for users in the unlikely event that IATEXT ceases its activity.

Links:

- CLARIAH-IATEXT (INFO)

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments:

R4 Legal & Ethical - R04

R04. The repository ensures to the extent possible that data and metadata are created, curated, preserved, accessed and used in compliance with legal and ethical norms.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

CLARIAH-IATEXT aims to ensure that datasets deposited and made available comply with applicable legal and ethical norms. During the submission workflow, depositors must accept a Distribution License Agreement (see Evidence link: CLARIAH-IATEXT (LICENCE)) confirming they have the right to submit the data and granting the repository the permissions needed to store, preserve and distribute the materials on their behalf. No anonymous deposits are allowed.

Legal and ethical responsibilities are divided as follows: depositors are responsible for the legality, rights clearance and disciplinary ethics of their content; the repository verifies that appropriate documentation and licensing are in place, enforces access controls, and documents decisions in the curation workflow. In case of suspected misuse or rights violations, access may be suspended, items may be withdrawn pending resolution, and the community may be informed where appropriate.

The Centre follows relevant community guidance (see Evidence link: CLARIN (LEGAL)) and the EU GDPR framework. A Privacy Policy (see Evidence link: CLARIAH-IATEXT (PRIVACY)) describes what personal data are processed (e.g., identity attributes from the federation, licence signatures, logs), for what purposes (authentication/authorisation, audit, service operation), retention periods and user rights.

Regarding potentially disclosive content: the repository is aware that some submissions may include image-based, geolocation-based, or sound-based data, which could contain personally identifiable or otherwise sensitive information. During the curation workflow, submitted items are reviewed by staff to identify potentially disclosive content. Depositors are required to confirm, as part of the Distribution License Agreement, that their submissions comply with applicable legal and ethical norms, including data protection regulations (GDPR). Where disclosive content is identified, the repository works with the depositor to apply appropriate access restrictions (e.g., academic or restricted licences), apply anonymisation measures, or, if necessary, reject the submission. Licensing of disclosive data follows the same framework described in R02, with restricted or academic licences applied as appropriate to limit access to authorised users only.

Links:

- CLARIN (LEGAL)
- CLARIAH-IATEXT (LICENCE)
- CLARIAH-IATEXT (PRIVACY)

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R5 Governance & Resources - R05

R05. The repository has adequate funding and sufficient numbers of staff managed through a clear system of governance to effectively carry out the mission.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

Organizational structure: Instituto de Análisis y Aplicaciones Textuales is a research institute of the University of Las Palmas de Gran Canaria (ULPGC), established in 2014 with a solid organizational structure. It is organized into research divisions, including a specialized unit in computational linguistics and software development for Digital Humanities, with extensive experience in managing, preserving, and disseminating linguistic resources. IATEXT joined the national consortium CLARIAH-ES through the initiative of its researchers and directorate and, following its designation as a CLARIN K-Centre (as certified in the K-Centre-Certificate.pdf included in the Attachments), is now pursuing B-Centre certification. The institute has strengthened its team by hiring specialized full-time staff to support the development and maintenance of the repository, complementing its existing computational linguistics group and ensuring sustainable future operations. In addition, IATEXT benefits from the support of other CLARIAH-ES nodes in Spain and maintains close links with CLARIN and DARIAH institutions across Europe, ensuring access to expertise and collaborative support for sustainable repository operations.

Staff and FTE: The CLARIAH-IATEXT repository relies on a qualified and multidisciplinary team with solid experience in the management, preservation, and dissemination of linguistic and digital humanities resources. The repository is staffed by 2 dedicated FTE distributed among three key professional roles. The Repository Manager, holding a PhD in Computer Science, has extensive experience in research data management, strong knowledge of European research infrastructures (CLARIN and DARIAH) and of the Spanish national consortium (CLARIAH-ES), and proven skills in project coordination and management. The Data Curator has academic training in Data Science and Engineering and brings experience in the curation and annotation of linguistic corpora, familiarity with metadata standards such as CMDI and Dublin Core, and expertise in quality control and validation processes. The Technical Administrator is trained in Computer Engineering and possesses experience in managing digital repositories, implementing digital preservation technologies, and ensuring IT security and identity management in compliance with institutional and European standards.

Internal governance: The governance of the repository is supported by the CLARIAH-IATEXT Internal Commission, a body that oversees the strategic direction, operational decisions, and quality assurance of the repository. This commission is composed of senior IATEXT researchers and the repository management staff, ensuring that institutional knowledge and research community needs are embedded in the repository's governance. The Internal Commission coordinates with ULPGC institutional bodies, the CLARIAH-ES consortium, and external advisors as needed. (For further details on the Internal Commission and its role in repository workflows, see also R11.) The full list of commission members and team profiles is publicly available (see Evidence link: CLARIAH-IATEXT Team and Committee).

CLARIAH-IATEXT maintains an in-depth knowledge of its Designated Community, as many IATEXT researchers are active members of the same academic domains served by the repository. This proximity ensures a direct understanding of user needs through active participation in conferences, collaboration in digital humanities projects, and firsthand experience in the creation and use of linguistic corpora. Access to specialized expertise is reinforced through collaborations within and beyond the institution. The CLARIAH-ES network provides technical advice from other Spanish CLARIN centres, expertise in CMDI metadata from the Coordination Centre, and support from CLARIN ERIC for complex technical issues. Within ULPGC, the repository benefits from institutional IT services for infrastructure and security, and from the Legal Service for contractual and rights-related matters.

Funding: The sustainability of CLARIAH-IATEXT is supported through stable funding from the ULPGC and Government of the Canary Islands, complemented by funding and in-kind support from CLARIN, DARIAH, and the CLARIAH-ES consortium (see the Funding CLARIAH-IATEXT.pdf included in the Attachments for details). This combination of institutional backing, permanent staffing, and national and European funding frameworks provides a robust basis for the long-term operation, development, and preservation services of the repository.

Links:

- CLARIAH-IATEXT Team and Committee

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R6 Expertise & Guidance - R06

R06. The repository adopts mechanisms to secure ongoing expertise, guidance and feedback-either in-house, or external.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

As part of CLARIAH-ES, CLARIAH-IATEXT is a member of the CLARIN-ERIC infrastructure and maintains regular contact with its experts. Through national coordination within CLARIAH-ES, the repository participates in the Standing Committee for CLARIN Technical Centers and related user-involvement activities, and takes part in regular coordination meetings with other CLARIAH-ES national coordinators, where recent developments and necessary adjustments to repository practices are discussed. Staff also attend CLARIN conferences, workshops, and meetings to support knowledge exchange and remain aligned with community guidance, reflecting the repository's focus on language resources for research.

Day-to-day repository management, including submission workflows, metadata curation, and operational activities, is handled by a small in-house team with expertise in the creation and maintenance of linguistic resources and tools, as well as experience working with widely used formats and standards in the field (e.g. TEI, XML, CONLL-U). As the repository is based on CLARIN-DSpace, it benefits from the shared practices and technical support of the LINDAT/CLARIN community. Feedback from the Designated Community is collected through a lightweight helpdesk (email) and through occasional meetings, with requests reviewed by internal staff or domain experts and used to improve documentation, submission guidelines, and metadata templates. Information about the institute and its members, including the staff involved in repository activities, is publicly available on the IATEXT website (see evidence link).

Continuous training and professional development: The repository actively promotes continuous training and professional development for its staff. Team members regularly participate in the CLARIN workshops on metadata, data curation, and interoperability. They also undertake training in emerging digital preservation technologies and engage in exchange activities such as job shadowing with other CLARIN centers. Through this combination of internal expertise, institutional support, and ongoing professional growth, CLARIAH-IATEXT ensures the long-term excellence and resilience of its services.

Links:

- LINDAT
- CLARIN (CENTRES)
- CLARIN-DSPACE
- IATEXT Staff Website

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Digital Object Management

R7 Provenance and authenticity - R07

R07. The repository guarantees the authenticity of the digital objects and provides provenance information.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The repository ensures the integrity and authenticity of deposited data and records provenance information. CLARIAH-IATEXT ensures the integrity and authenticity of deposited digital data and provides provenance information through robust security measures. Specifically, it utilizes Shibboleth for institutional single sign-on (SSO), internal periodic backups that replicate sensitive data to prevent compromise, and records of an audit trail to track all significant actions.

Authenticity is further ensured through the authentication of depositors via institutional SSO (Shibboleth) or managed local accounts (see Evidence link: CLARIAH-IATEXT LOGIN). Identity attributes (e.g., name, email) are captured, and each dataset is linked to the Producer's account. Role-based permissions guarantee that only the Producer, or authorized curators, can edit or manage their submission. Once a dataset is published, the released version becomes immutable, and any subsequent changes create a new version.

Persistent identifiers: Persistent identifiers are assigned automatically by the CLARIN-DSpace platform upon publication. The repository uses the Handle System (handle.net), with the institutional prefix 20.500.14841, operated by the DONA Foundation through the CLARIN infrastructure. Each deposited resource receives a unique Handle PID in the form https://hdl.handle.net/20.500.14841/[ID], which is permanently linked to the resource and included in the metadata. This ensures long-term citability and discoverability of all published datasets.

Metadata follows community standards (CMDI, with Dublin Core exposed for interoperability), and DSpace provides standard dissemination interfaces (OAI-PMH) for efficient data sharing. Integrity is ensured through a combination of proactive data validation and robust storage practices. At ingest, file checksums are computed and stored. Fixity is verified at the time of writing, then re-checked at scheduled intervals, as well as during backup and restore operations. Data are stored on replicated storage with routine backups to an independent location and automated monitoring. In the event of discrepancies, affected items are quarantined, investigated, and restored from verified copies as needed.

Provenance is managed through DSpace's comprehensive audit trail, which records all significant actions, including submissions, edits, reviews, approvals, publication, versioning, and access-rights changes. Selected events are reflected in public provenance metadata, ensuring transparency, while detailed logs are securely stored and remain available to system administrators for further analysis or review.

Links:

- CLARIAH-IATEXT (LOGIN)
- CLARIN (CMDI)
- DublinCore
- OpenArchives
- Handle System (CLARIAH-IATEXT example)

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R8 Deposit, Appraisal & Accessibility - R08

R08. The repository accepts data and metadata based on defined criteria to ensure relevance and accessibility for users.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

Deposit process: CLARIAH-IATEXT evaluates submissions through the standard DSpace submission workflow. This documented process includes steps to ensure that data and metadata are sufficient for long-term preservation. The submission workflow guides depositors step-by-step through the following stages: (1) login and account identification using institutional SSO or local accounts; (2) selection of the appropriate collection; (3) provision of required descriptive metadata following CMDI and Dublin Core standards; (4) upload of data files and verification of file readability and format conformance; (5) acceptance of the licence agreement; and (6) review and curation by repository staff, who check completeness, correctness of metadata, legal/ethical readiness, and format compliance. If issues are found, submissions are returned to depositors for correction via the workflow. Once approved, the dataset is published and a persistent Handle identifier is assigned. Appraisal decisions and editor actions are tracked in the workflow and audit logs. The complete documented deposit and curation workflow is publicly available in the Data Management Policies of the repository (see Evidence link).

Collection development policy: The repository maintains a documented collection development policy that guides the selection and acceptance of digital objects. Submissions must fall within the repository's defined scope: language resources for research (corpora, tools, datasets in linguistics, digital humanities, and related fields). The policy specifies relevance criteria (alignment with the repository's scope and the needs of the Designated Community), legal/ethical readiness (clear rights, where relevant consent statements), documentation standards (sufficient metadata and file documentation), and format requirements (preferred formats following CLARIN recommendations; non-preferred formats accepted only if sufficiently documented and verifiable). Thorough content quality assessment and format conversion are outside the scope of the acceptance process, and depositors remain responsible for ensuring scientific and linguistic quality. The collection development policy is included in the Data Management Policies documentation.

Format guidance: Preferred, open, and well-documented formats are encouraged in line with CLARIN recommendations and the Standards Information System. CMDI-compliant metadata is supported and exposed alongside Dublin Core for interoperability.

Links:

- CLARIN (RECOMMENDATIONS)
- CLARIN (STANDARDS)
- CLARIAH-IATEXT Data Management Policies

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R9 Preservation plan - R09

R09. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

CLARIAH-IATEXT assumes responsibility for long-term preservation and manages this function in a planned and documented manner, as detailed in the repository's publicly available Data Management Policies (see Evidence link: CLARIAH-IATEXT Data Management Policies).

Bit-level preservation: The repository implements bit-level preservation for all items using DSpace. This mechanism guarantees scheduled backups at the primary server as well as replication to secondary servers. Off-site backup is ensured through a formal agreement with SCAYLE (Supercomputación Castilla y León), a public supercomputing centre that is part of the CLARIAH-ES infrastructure. This cooperation is documented in two agreements included in the Attachments: Memorandum SCAYLE.pdf (SCAYLE's adhesion to the CLARIAH-ES infrastructure) and Memorandum SCAYLE 1.pdf

(CLARIAH-IATEXT's responsibilities as a user of SCAYLE infrastructure). File checksums are computed at ingest and verified at regular intervals, as well as during backup and restore operations, to detect any corruption. When submitting a new version of any resource, an immutable and unique persistent Handle identifier (PID) is created and linked to the resource.

Functional preservation and format migration: In addition to bit-level preservation, the repository applies a documented functional preservation strategy. The approach distinguishes between:
- Format migration: For community-standard formats (e.g., TEI XML, CoNLL-U, plain text), the repository undertakes format migration when necessary to maintain readability and usability over time, while always retaining original bitstreams for reproducibility and authenticity.
- Normalization: At ingest, files are assessed for format conformance. Where feasible, files may be normalized to preferred open formats as part of the curation process.
- Emulation: Emulation is not currently within scope for the repository's content types (primarily text-based linguistic corpora and tools), given that the formats used are generally open and human-readable. However, the approach is kept under review as part of periodic preservation planning.
- Externally hosted tools and services: These are catalogued and described with sufficient metadata, rather than preserved at the bitstream level, as responsibility for preservation lies with the original providers.

Preservation planning and review: Published datasets are retained without a fixed end date and are periodically re-appraised to assess risk and usability. Possible outcomes include retaining, migrating, restricting, or withdrawing datasets. Withdrawals keep metadata and maintain PID resolution to a tombstone. The deposition agreement authorizes preservation actions, including copying, transforming, and storing data, and also covers the transfer of custody in case of hosting or funding changes. Content may be transferred to another CLARIN Centre or a designated institutional or national archive. Policies and procedures are documented and periodically reviewed.

Links:

- CLARIAH-IATEXT Data Management Policies

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments:

R10 Quality Assurance - R10

R10. The repository addresses technical quality and standards compliance, and ensures that sufficient information is available for end users to make quality-related evaluations.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

From the repository's perspective, quality assurance focuses on ensuring technical soundness and compliance with standards, rather than verifying the scientific accuracy of the content. Depositors authenticate through the DSpace submission workflow and are guided to complete mandatory metadata fields, such as type, title, description, creators, language, and license. Where applicable, values are encouraged from controlled vocabulary dictionaries. The metadata is then mapped to CMDI and exposed to Dublin Core to enhance interoperability and discovery. The repository combines automatic curation tasks (e.g., required-field checks, URL/handle validity, basic file/format sanity checks such as XML well-formedness) with editorial review for completeness and consistency. Curators may open sample files to confirm that bitstreams are accessible as described. Items that do not meet the baseline are returned to the depositor for correction via the workflow. Once accepted, a persistent identifier (PID) is assigned for citation and long-term reference, and records may include links to related publications or other digital objects to help end users assess fitness for use. By default, the repository operates on basic curation. Enhanced checks may be applied for specific collections by agreement. The repository monitors changes in community standards (metadata profiles and recommended formats) and updates submission templates and guidance accordingly. When such changes affect depositors or consumers, the required actions are communicated in the help and submission of materials. The repository help desk is available with help

and information about metadata descriptions and the use of data in the repository for both depositors and end users. The repository does not label data with quality scores, but end users are welcome to comment on the quality of data by contacting the help desk or directly reaching out to the repository staff.

Links:

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R11 Workflows - R11

R11. Digital object management takes place according to defined workflows from deposit to access.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The CLARIAH-IATEXT repository follows a controlled submission and curation workflow to ensure metadata quality, policy compliance, and long-term preservation. Depositors submit datasets and descriptive metadata through the DSpace interface. Curators then review the submission, verifying metadata accuracy, file integrity, and compliance with CLARIN and institutional policies. Upon approval, items are published, assigned as a persistent identifier (Handle), indexed by Solr, and exposed via OAI-PMH. Updates are managed through immutable versioning, with each version receiving separate PID and explicit links to earlier releases. Long-term reliability is supported through ongoing preservation monitoring, including fixity checks, backup verification, and periodic metadata audits. Preservation activities are performed according to a defined schedule with daily incremental backups and service monitoring, system updates, quarterly quality audits, and annual reviews of preservation strategies including format obsolescence evaluation and migration planning. Access is provided for both open and restricted resources. Open resources are directly downloadable, while access to restricted resources is controlled through Shibboleth authentication and attribute-based authorization. Versioning is managed through a controlled process covering request submission, verification, assessment, curation with provenance documentation, version assignment, linking between versions, and republication with updated metadata. All repository actions are logged in a central system, recording actor, action, time, resource identifier, and outcome. Access logs are retained for two years, while preservation and modification logs are kept permanently. Each deposit is tracked through all workflow stages using a unique identifier. After curation, data and metadata quality are independently verified by a designated curator or quality assurance officer. Automated processes include checksums, backups, integrity checks, OAI-PMH exposure, VLO indexing (29), and error monitoring, with logs reviewed daily. Policies are reviewed annually or when changes are required due to new CLARIN standards, technological developments, or community feedback, whichever occurs first. The annual review assesses efficiency metrics (processing times, error rates, user satisfaction), identifies bottlenecks, and incorporates best practices from other CLARIN centers. Regular certification processes with CoreTrustSeal and evaluation as a CLARIN B-Centre require CLARIAH-IATEXT to verify that digital preservation and curation processes conform to international best practices, update existing documentation, and create documentation for new processes. New processes or changes to existing workflows are discussed within the repository team and, when necessary, with the CLARIAH-IATEXT Internal Commission. Relevant policies and tracking methods are updated accordingly, with version control maintained for all procedure documents. Staff receive training on updated procedures, and changes are communicated to depositors and users. 29 https://vlo.clarin.eu/search

Links:

- CLARIN-VLO

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R12 Discovery and Identification - R12

R12. The repository enables users to discover the digital objects and refer to them in a persistent way through proper citation.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The repository can be searched using the DSpace interface backed by an Apache Solr index. Users can perform keyword search and browse/filter by facets (e.g., resource type, language, licence, date). Where applicable, text-based files are indexed for full-text search to enhance discovery.

To enhance the discoverability of content, the repository supports metadata harvesting via OAI-PMH protocol using DSpace's OAI v2 server. The CLARIN Virtual Language Observatory (CLARIN VLO) harvests metadata from our OAI endpoint into the CLARIN VLO, where language resources may be discovered using a facet browser (see Evidence link: CLARIN-VLO). Through VLO and the CLARIN Component Metadata Initiative, the potential user can search among larger collections of materials harvested from various data providers within CLARIN and related communities.

Each published dataset receives a persistent identifier (PID) using the Handle System (see Evidence link: HANDLE) that resolves to a stable landing page controlled by the repository. Published versions are immutable: new public versions receive a new PID (or a versioned PID) and link to prior versions on the landing page. A recommended citation is provided on each record (see Evidence link: CLARIAH-IATEXT (CITE)), and users are asked to acknowledge and cite data sources properly in all publications and outputs.

Links:

- CLARIN-VLO
- CLARIAH-IATEXT (CITE)
- HANDLE

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R13 Reuse - R13

R13. The repository enables reuse of the digital objects over time, ensuring that appropriate information is available to support understanding and use.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

CLARIAH-IATEXT adheres to the principles of Open Science and Open Access, fostering the dissemination and reuse of data generated by publicly funded research through trusted public repositories. By promoting data sharing, the repository contributes to increasing research visibility, avoiding redundant efforts, and enabling higher-quality research through verification and validation of existing data. In line with this commitment, CLARIAH-IATEXT fully embraces the FAIR principles (Findable, Accessible, Interoperable, and Reusable) as the cornerstone of its approach to sustainable data management, ensuring that research outputs remain usable, comprehensible, and valuable over the long term.

CLARIAH-IATEXT requires that a comprehensive set of metadata (both mandatory and recommended) providing information about the submitted resources are completed during deposit. The required metadata set is designed to support different metadata profiles and formats, including the CLARIAH-IATEXT CMDI (Component Metadata Infrastructure, see Evidence link: CLARIN CMDI) profile, and Dublin Core. Therefore, we support all these formats in our OAI-PMH endpoint. Because the other profiles and formats are dynamically constructed from the core CMDI metadata, the sustainability and future evolution of metadata formats can be easily supported. Users can view these descriptive metadata, together with licensing information covering intellectual property rights, conditions of use, and access restrictions, on the resource landing page.

Depositors are encouraged to upload files in standard formats for language resources suitable for long-term preservation that are maintained by the language resource community. Preferred formats include XML and plain text (UTF-8) for text corpora, CoNLL-U (a standard tab-separated format for annotated linguistic data, see Evidence link: CoNLL-U FORMAT) and XML for annotated corpora, and source code with Docker containers for tools. When depositors submit resources in other formats, curators require detailed documentation on how to process the data, which must be included in the resource itself.

Each resource includes documentation enabling independent reuse of CMDI metadata documents: creation methodology, annotation schemes, and quality control procedures; technical documentation provides XML schemas and annotation guidelines where applicable. All resources have clear licensing information displayed on landing pages and embedded in machine-readable CMDI metadata. Licenses used include Creative Commons licences (CC0, CC BY), and CLARIN-specific licences: CLARIN ACA (Academic/non-commercial use) and CLARIN RES (Restricted use with registration required). For further information on CLARIN ACA/RES licences, see the Evidence link: CLARIN ACA/RES LICENCES.

From a technical perspective, CLARIAH-IATEXT relies on a DSpace-based repository infrastructure (see Evidence link: WIKI-DSPACE), which provides robust mechanisms for long-term digital preservation. DSpace ensures the secure storage of bitstreams, the association of each resource with persistent identifiers, and the management of descriptive, administrative, and technical metadata. The platform supports integrity checking through checksums, versioning of resources, and controlled update workflows, ensuring that any modification to deposited materials is traceable and documented. Combined with regular backup procedures and controlled access policies, this infrastructure guarantees the durability, authenticity, and citability of the resources over time.

Links:

- WIKI-DSPACE
- CLARIN (CMDI)
- CLARIN ACA/RES LICENCES
- CoNLL-U FORMAT

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Information Technology & Security

R14 Storage & Integrity - R14

R14. The repository applies documented processes to ensure data and metadata storage and integrity.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The repository applies documented and auditable procedures across ingest, storage, and access to safeguard the integrity of both data and metadata (see link CLARIAH-IATEXT Data Managemente Policies). During ingest, the DSpace-based submission workflow requires depositors to complete mandatory metadata fields, registers file checksums (fixity information), and subjects submissions to baseline technical checks according to internal procedures: curators review deposits and may return them for correction when they do not meet acceptance criteria.

Fixity is generated at ingest and verified on write; scheduled jobs re-check fixity at defined intervals and during backup/restore operations. If discrepancies are detected, items are quarantined, investigated, and restored from verified copies. Storage integrity is supported by replicated service infrastructure and scheduled backups, with platform-level replication to a secondary node and periodic snapshots to establish recovery points: these measures enable restoration to a recent consistent state in the event of corruption or loss. Off-site backup redundancy is provided through a formal agreement with SCAYLE (Supercomputación Castilla y León), as documented in the Memorandum SCAYLE.pdf and Memorandum SCAYLE 1.pdf included in the Attachments.

Provenance and change history are maintained through an audit trail that records significant actions (submission, edits, approvals, publication, versioning, access-rights changes); selected events are exposed in public provenance metadata. Published versions are immutable: subsequent changes create a new version, and persistent identifiers are maintained to ensure reliable citation and traceability. To promote metadata integrity and discoverability, the repository exposes records via standards-based interfaces (OAI-PMH) and supports CMDI alongside Dublin Core: automated processes and periodic checks verify that dissemination endpoints remain consistent with curated records. Finally, preservation operations follow a documented preservation and security plan, including scheduled backups, regular integrity checks, and log-based monitoring, with results retained to support internal review and external audit.

Note: The Level of compliance dropdown does not display our selection. We have selected option 1 – Implemented: the requirement has been fully implemented by the repository, as documented in the CoreTrustSeal application report submitted on 2026-01-09, included in the Attachments.

Note: As mentioned, new versions receive a new PID, but the old PID of the old version is maintained.

Links:

- SCAYLE
- CLARIAH-IATEXT Data Management Policies

Reviews

Reviewer 2:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R15 Technical Infrastructure - R15

R15. The repository is managed on well-supported operating systems and other core infrastructural software and hardware appropriate to the services it provides to its Designated Community.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

The repository infrastructure is managed directly by IATEXT using its own dedicated hardware, with support from the ULPGC IT Services (Servicio de Informática de la ULPGC, SI-ULPGC; see Evidence link: ULPGC-IT) for general institutional matters such as network connectivity, SSL certification, and identity federation services.

The infrastructure itself consists of a Hyper-V cluster composed of two physical servers running Windows 11. The first server hosts the virtual machine (VM) running the CLARIN-DSpace (DSpace-based) repository stack, including the application layer, database, and associated services. The second server functions as a replication and high-availability (HA) node, supporting operational continuity in the event of hardware or system failure. Both nodes are synchronized via Hyper-V replication mechanisms to maintain data consistency and enable timely service recovery. The servers are operated within the institution's secure server facilities and connected through the institutional network, benefiting from centrally managed IT controls provided by SI-ULPGC. The infrastructure is provisioned with sufficient storage and compute capacity to support current repository operations and projected growth. Data protection is further supported through regular backups and tested recovery procedures, enabling restoration to a consistent state in case of accidental loss or corruption. To maximise the quality, interoperability, and sustainability of our repository service, the infrastructure adheres to a range of international and community standards and principles, including Handles, OAI-PMH, W3C standards, the Linked Data Platform, the Dublin Core Metadata Initiative, and the Component Metadata model of CLARIN.

Processes to monitor and manage technical change: The repository has implemented documented processes to monitor and manage the need for technical change, including in response to the evolving needs of preservation and reuse by the Designated Community. The repository management team holds regular operational reviews where infrastructure performance, service availability, and software versions are assessed. When changes in community standards, emerging preservation risks, or updated format recommendations from CLARIN are identified (e.g. through participation in CLARIN technical committees and workshops), these are evaluated against the current infrastructure and scheduled for implementation where appropriate. Change management follows an internal procedure: proposed changes are assessed for impact on data integrity, metadata interoperability, and service continuity before being applied. Significant infrastructure changes are logged and communicated to stakeholders. Feedback from the Designated Community, collected through the helpdesk and direct interactions, is also used as input to the review cycle. This ensures that the technical infrastructure evolves in alignment with the preservation strategy (R09) and the reuse needs of the Designated Community (R13).

Links:

- ULPGC-IT

Reviews

Reviewer 2:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

R16 Security - R16

R16. The repository protects the facility and its data, metadata, products, services, and users.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

Building on the infrastructure described in the previous section (see R15), which is designed with sufficient redundancy in mind so that outages caused by hardware failures should be rather rare, CLARIAH-IATEXT takes a practical, security-first approach to protecting research records and sensitive data.

We operate a replica of the service on a separate server to ensure service continuity, and we implement periodic snapshots to provide consistent restore points and meet a daily Recovery Point Objective: these are complemented by regular backups independent of the snapshots so we can recover from

data loss or—even in the worst case—the loss of an entire server. Off-site backup redundancy is provided through the SCAYLE infrastructure, as documented in the Memorandum SCAYLE.pdf and Memorandum SCAYLE 1.pdf included in the Attachments.

Physical security is provided by hosting the systems in secured facilities with controlled access for authorized personnel — specifically, IATEXT repository research staff (see Evidence link: IATEXT Authorized Personnel). At the network perimeter, the service is protected by a virtual-machine firewall, and communications are encrypted with a valid SSL/TLS certificate, exposing only the services strictly required for operation. Authentication relies on Shibboleth (SAML 2.0) single sign-on (see Evidence link: Shibboleth), allowing users to log in with institutional credentials while passwords remain with their identity providers rather than on our servers.

Operationally, we are adding active reviews of usage statistics and system logs to detect anomalous or abusive behaviour (e.g., inflated download counts or systematic scraping) and to respond promptly where necessary. Together, these measures provide a solid baseline for availability, integrity, and confidentiality as we continue to strengthen the service.

Links:

- Shibboleth
- IATEXT Authorized Personnel
- CLARIAH-IATEXT Personnel

Reviews

Reviewer 2:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments:

Reviewer 1:

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Comments:

Applicant Feedback

R17 Applicant Feedback

We welcome feedback on the CoreTrustSeal Requirements and the Certification procedure.

Compliance level:

Implemented: the requirement has been fully implemented by the repository. - 1

Response:

Regarding the Application Management Tool (AMT) bug notices: All notes previously included in the application that made reference to a bug in the CoreTrustSeal Application Management Tool have been removed.

Regarding R3 (Continuity of Access): In response to Reviewer 2's observation that no English translation of the relevant succession-related sections of the Memorandum of Understanding was provided, a formal Statement of Institutional Viability and Service Continuity has been prepared in English and included as a new attachment (Statement_Institutional_Viability_IATEXT_English.pdf). This document explicitly describes the succession and continuity mechanisms in place for the CLARIAH-IATEXT repository, including the distributed succession framework supported by the 11 remaining nodes of the CLARIAH-ES consortium. In addition, to address the concern that the succession commitments described therein rest on a unilateral declaration rather than a formalised inter-institutional agreement, the document references the formal membership of IATEXT-ULPGC in the CLARIAH-ES consortium as the legal and institutional basis for those commitments. Now, the CLARIAH-ES Memorandum of Understanding, provided in English, is also included as an attachment to this application (Memorandum CLARIAH-ES English.pdf).

Regarding R8 (Deposit, Appraisal & Accessibility): Sections identified as missing in R8 have now been incorporated into the project website and are publicly available at the Data Management page, including the documented deposit process for ensuring sufficient data and metadata preservation, as well as the collection development policy and selection procedures for digital objects.

Links:

Reviews

Reviewer 2:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments:

Reviewer 1:

Compliance level:

In Progress: the repository is in the implementation phase. - 0

Comments: