

CoreTrustSeal Requirements: Periodic Review 2019

Report of the CoreTrustSeal Review 2019 Working Group

(Version 2.0)



Contents

| | |
|---|-----------|
| Attending: CoreTrustSeal Review 2019 Working Group | 3 |
| Artefacts | 3 |
| Version History | 3 |
| Introduction | 4 |
| Goals and Scope | 4 |
| Outcomes | 4 |
| Review Process | 5 |
| Feedback | 5 |
| Redundancy and Overlap | 5 |
| Requirements Text Edit | 6 |
| Requirements Revision Overview | 7 |
| General Extended Guidance | 7 |
| Context (R0) | 7 |
| Organizational Infrastructure | 9 |
| Digital Object Management | 10 |
| Technology | 12 |
| Appendix: Community Review Communications | 14 |
| Schedule | 14 |
| Webpage | 14 |
| Email | 14 |

Attending: CoreTrustSeal Review 2019 Working Group

Rorie Edmunds (Working Group Lead; World Data System)

Jonas Recker (CoreTrustSeal Chair; GESIS-Leibniz Institute for the Social Sciences, Germany)

Hervé L'Hours (CoreTrustSeal Vice-chair; UK Data Archive, University of Essex, UK)

Jonathan Crabtree (CoreTrustSeal Board; Odum Institute Data Archive, USA)

Mustapha Mokrane (CoreTrustSeal Board; Data Archiving and Networked Services, The Netherlands)

Ilona von Stein (CoreTrustSeal Board ex officio; Data Archiving and Networked Services, The Netherlands)

Artefacts

The following information artefacts were integrated into the review process:

- CoreTrustSeal Extended Guidance: Community Comment version
- CoreTrustSeal Extended Guidance: Board Comment version
- Consolidated feedback spreadsheet containing direct comments received by the Board and applicant comments received during the review process

Version History

| | | |
|------------|--|--------|
| 2019-07-25 | Interim edit during the review process | v00.01 |
| 2019-07-25 | Restructured and revised to include the review process and a change log for each Requirement | v00.02 |
| 2019-07-29 | As released for review and approval by the CoreTrustSeal Board | v00.03 |
| 2019-09-03 | Release with the revised Requirements for public comment | v01.00 |
| 2020-05-27 | Edited to include changes to the Extended Guidance 2020–2022 | v02.00 |

Introduction

This Report supports the final public release of the revisions to the Core Trustworthy Data Repositories Requirements 2020–2022. The CoreTrustSeal Board would like to thank contributors from the CoreTrustSeal community and beyond for their expert, detailed, and thoughtful feedback during this process.

We have integrated a large number of feedback items into this revision while seeking to keep the structure of the Requirements fundamentally stable for the 2020–2022 period. Feedback that has not been integrated now will be used as inputs for the next review. At the time of the next review, we expect to be able to make more substantive changes in response to the evolving role of CoreTrustSeal and related efforts, including those around FAIR (Findable, Accessible, Interoperable, Reusable) data.

The Working Group met to review each item of feedback, and the recommendations made by past applicants and CoreTrustSeal reviewers.

An interim release of the Core Trustworthy Data Repository Requirements (and associated Guidance) was released to allow responses based on errors or omissions in the draft version.

Once a stable revision of the Requirements was agreed, the Board progressed to reviewing and revising the CoreTrustSeal Glossary and Extended Guidance. These were also available for a period of public comment before all the documents were formally released as the 2020–2022 Requirements, Extended Guidance, and Glossary.

Goals and Scope

- To adapt and improve the Core Trustworthy Data Repositories Requirements in response to feedback from the CoreTrustSeal and wider data management communities, including the experiences of applicants, reviewers, and the Board.
- To maintain the generically applicable and ‘core’ level of the CoreTrustSeal certification.
- To maintain the fundamentals of the previous version such that the impact is minimized for those repositories still transitioning to CoreTrustSeal from its predecessor certification standards.
- To develop the processes and knowledge necessary to support a more wide-ranging review and revision in 2021–2022 for the 2023–2025 Requirements.

Outcomes

The following items were released to the community:

- Core Trustworthy Data Repositories Requirements 2020–2022 (<https://doi.org/10.5281/zenodo.3638211>).
 - Change file: <https://doi.org/10.5281/zenodo.3828622>
- CoreTrustSeal Extended Guidance (<https://doi.org/10.5281/zenodo.3632533>).
 - Change file: <https://doi.org/10.5281/zenodo.3828636>
- CoreTrustSeal Glossary (<https://doi.org/10.5281/zenodo.3632563>).
- Final Report from the 2019 Review (this document).

Review Process

The review and revision process was undertaken in several distinct phases. Feedback received from applicants during past reviews was included for cross-check purposes, even if the issues raised were addressed some time ago.

In addition to comments that were directly relevant to the Requirements text, the Working Group considered where feedback had implications for changes to Board processes, communications, or the online CoreTrustSeal Application Management Tool.

Feedback

The Working Group undertook an initial review of all feedback received. Each suggestion was triaged into the following categories.

- Accepted, with a brief explanation of how this would be addressed.
- Rejected, with a brief explanation as to why.
- Deferred, with a brief explanation as to why this would be reconsidered in the time leading up to, and during, the next review, and who should take the lead in doing so (CoreTrustSeal Board, or the community).

Suggestions were typically rejected when considered too granular for core-level requirements, or were overly disciplinary or domain-specific.

Redundancy and Overlap

There were a number of comments that indicated overlap between Requirements, or a level of redundancy whereby the same evidence could potentially be repeated across multiple Requirements.

Each feedback item in this category was identified, collated, and cross-checked against the Requirements. All instances of actual redundancy were removed where this did not detract from the internal consistency of the Requirement. In other cases, further guidance was provided to clarify the different expectations of each Requirement.

Requirements Text Edit

After examining all comments and potential cases of redundancy/overlap, the outcomes were then included in a full-text review that also included proposed edits from the public feedback and from the Board commented version of the Extended Guidance.

This process used a 'clean' copy of the Requirements text with all changes tracked. The text edit was conducted in the following Requirements order.

Technology Requirements: R15 and R16

The Technology Requirements, R15 (Technical Infrastructure) and R16 (Security), were addressed first as feedback indicated that they present a challenge to respondents. Such challenges may have arisen because those leading the application process tend to be from a repository management rather than a technical background, or because expectations and norms around technology and security issues are continuing to evolve.

Addressing these Requirements first also supported a review of R1 to R14 with a view to moving or removing guidance with a technology element.

Requirements R1 to R14

Requirements R1 to R14 were then fully reviewed, including integrating feedback, addressing redundancy, and refining technical aspects.

R0: Context

The Context section was left until this stage so that the Working Group could identify the full range of contextual information implied by R1 to R16.

Cross-references among Requirements

The last step of the structure and content process was to review, validate, and simplify the cross-references among requirements. Text used in cross-references was also standardized to reduce ambiguity.

Final Editorial Edit

The final stage of the process was to ensure consistent descriptions and terminology were used throughout the text.

Requirements Revision Overview

This part of the Report acts as a change log for the proposed revisions to the Core Trustworthy Data Repositories Requirements. Specific responses to feedback are covered in a separate feedback spreadsheet. The focus here is on identifying additions, deletions, or amendments. Some clarifications have been made to address the knowledge base of applicants—for example, providing further detail on the Designated Community—but the Requirements and Guidance cannot replace a working knowledge of data management practices, including the [OAIS \(Open Archival Information System\) reference model](#).

Changes to the Requirements and to expectations during review both will be communicated to the community. CoreTrustSeal reviews will become more rigorous over time as community norms are established and as technical solutions are developed. For example, there may be a higher expectation that globally Persistent Identifiers (PIDs) are used because these services have become more widely available and affordable.

The Requirements Guidance remains non-exhaustive, and applicants should respond based on their local practices.

General Extended Guidance

General Points: A reference to the 2014 Digital Preservation Coalition Technology Watch Report, 'The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)' by Brian Lavoie (<https://doi.org/10.7207/twr14-02>), was added in response to a request for further detail on the OAIS model.

Missing Information/Evidence: It was clarified that reviewers and peer repositories should be able to understand the final application as a standalone document including clear statements about what information was being provided in evidence and where. A short statement on cross-referencing was also included.

Application Structure and Length: A clarifying statement was added to explain that 'Applications should not respond to each item of guidance in a question-and-answer format. Applications should include prose responses to each Requirement.'

Context (R0)

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community inputs on the best approach to describing repository collections. These 'collection profile' descriptions will support improved reviews.

Renewals

A Renewals section is added to allow applicants renewing their CoreTrustSeal to provide an overview of relevant changes since their last application in terms of funding, governance, data

collections, Designated Community, or technology. This includes information that addresses any reviewer comments from the final assessment of their previous application; for instance, an expectation that documentation would be made public.

Repository Type

Repository types were agreed with the [RDA/WDS Publishing Data Cost Recovery for Data Centres IG](#) during the development of CoreTrustSeal. These remain open for revision over time. Library/Museum/Archives are now split into separate headings. The definition of Publication Repository will be re-examined in the Glossary and potentially refined.

Brief Description of the Repository

Further guidance is provided. It is highlighted that organization structure diagrams and descriptions are very helpful in the review process.

Designated Community

‘Designated Community’ is used instead of Repository Designated Community in response to a query over the use of Repository versus Archive (in the OAIS sense). A sentence from the Extended Guidance is moved to the Guidance to stimulate more formalized descriptions of the Designated Community.

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community inputs on the best approach to defining the Designated Community.

EXTENDED GUIDANCE

The Extended Guidance now addresses the Repository versus Archive definitions as they apply to CoreTrustSeal. In response to feedback, additional details about the expectations of designated community information are provided; in particular, concerning repositories serving multiple user-communities.

Levels of Curation

EXTENDED GUIDANCE

Language to encourage applicants to add further details on the curation levels performed has been included and questions added concerning information about cases where different parts of the collection are under different levels of curation.

Insource/Outsource Partners

‘Outsource Partners’ has been replaced by ‘Insource/Outsource Partners’, and the explanation provided in the Guidance revised accordingly.

Organizational Infrastructure

I. Mission/Scope

The ideal for a CoreTrustSeal applicant remains that they have a public mission statement specifically referencing research data management, and that this statement is approved and periodically reviewed at the highest level of the organization through a standard governance procedure. However, the text is revised to make it clear that evidence implying a clear mission is also acceptable.

II. Licenses

The text is revised to clarify that ethical and privacy elements of licences and rights management should be covered under R4 (Confidentiality/Ethics).

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community inputs on consolidating all rights management under R2, including a proposal to move the ‘right to preserve’ under R10 (Preservation Plan) into R2. The Requirement will be renamed appropriately to reflect these changes.

III. Continuity of access

‘Continuity of access’ in some areas of data management and preservation may be synonymous with ‘preservation’. The text is updated and the cross-references simplified to clarify that the scope here is business continuity, and disaster and succession planning, whereas technology aspects are covered under R15 (Technical infrastructure).

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community inputs on additional revisions to this Requirement, as well as renaming it, to further increase the clarity of its scope.

IV. Confidentiality/Ethics

The text is revised to clarify that evidence demonstrating good governance of data with disclosure risk should include guidance for data depositors and users.

The Requirement cross-references are simplified.

EXTENDED GUIDANCE

Language has been added to emphasize the repository’s responsibility to put in place measures that ensure no personal data are uploaded to the repository unless explicitly permitted.

V. Organizational infrastructure

The Requirement cross-references make it clear that expert advice beyond internal staff expertise is covered under R6 (Expert guidance).

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community input on renaming this Requirement to avoid duplication with the section heading.

EXTENDED GUIDANCE

In response to feedback asking about periodic renewal, it has been clarified that this refers to the renewal of funding.

Text now includes that evidence should describe organization structure, governance, and skills.

VI. Expert guidance

The types of expertise in scope are slightly revised.

Digital Object Management

VII. Data integrity and authenticity

The text is amended to better differentiate between integrity and authenticity, including the need to document intended changes and to mitigate and recover from unintended changes.

The Requirement cross-references are simplified. It is made explicit that details of the technical management of integrity and authenticity are covered under R15 (Technical infrastructure)

VIII. Appraisal

The terms 'evaluation' and 'selection' are added for applicants from communities in which 'appraisal' is not commonly used.

The text across R8 (Appraisal), R11 (Data quality) and R14 (Data reuse) is revised to clarify perceived overlaps. Appraisal reflects interactions at the point of deposit; curation actions post-deposit are moved to R11. References to automated validation of metadata against schema are also added to R8.

The notion of 'reappraisal' is added to reinforce that re-evaluation and disposition of data over time is permitted and expected. In addition, a bullet point concerning the process for removing items from the data collection is added.

The Requirement cross-references are refined.

IX. Documented storage procedures

The text is revised to clarify that this Requirement refers to data flows and storage procedural information from a repository manager perspective.

The Requirement cross-references are edited to note that technical details of data storage management are covered under R15 (Technical infrastructure) with elements of physical and logical security of storage systems covered under R16 (Security).

EXTENDED GUIDANCE

Additional examples of possible evidence have been provided.

X. Preservation plan

A specific question about plans for future migrations or other measures to address the risk of obsolescence is added. To avoid ambiguous interpretations of preservation policies, strategies, plans, and implementation plans, the text states the need for a ‘documented approach to preservation’ rather than for a specifically named item of evidence.

The Requirement cross-references clarify that rights management of data access and use are covered under R2 (Licences).

EXTENDED GUIDANCE

Further details on what is expected from a preservation approach have been provided.

XI. Data quality

The text now increases the focus on quality improvement and assurance during the curation process, including the need to meet Designated Community expectations.

The Requirement cross-references note that selection criteria at the point of deposit are covered under R8 (Appraisal).

EXTENDED GUIDANCE

In response to user feedback, text has been clarified on the expected approach in cases where quality thresholds cannot be reached.

XII. Workflows

This Requirement identifies whether all relevant processes have been sufficiently documented; for example, to ensure consistent service delivery and to mitigate risk. The text has been updated to make this clearer and to indicate that processes should map to repository activities, such as the functions identified by the OAIS reference model.

The Requirement cross-references have been removed to avoid any suggestion that the applicant is expected to repeat information provided elsewhere.

XIII. Data discovery and identification

In response to feedback and to the increased availability of PID minting and resolution services, the text now asks specifically which PID system is in use by the applicant.

Redundant text related to the provision of search interfaces and to the benefits of curation, both of which fall outside the remit of the core certification process, has been removed from the Requirement and added to the Extended Guidance.

EXTENDED GUIDANCE

In response to feedback, the text on citation, credit, and attribution has been adjusted such that it is applicable beyond the purely academic scope.

XIV. Data reuse

Specific mention is made to the knowledge base of the Designated Community.

References to metadata provision are now more generic, with exact examples moved to the Extended Guidance.

Redundant requirements cross-references are removed.

EXTENDED GUIDANCE

Additional guidance is provided for cases where general rather than domain-specific schema and community standards are used.

Technology

By the 2021–2022 review of the Requirements, the CoreTrustSeal Board will seek community inputs on further defining community norms for documentation of technical infrastructure and security. The process will attempt to define both repository minimal/best practices and the provision of evidence from third-party technical service providers.

XV. Technical infrastructure

References to community standards are made more generic, with exact examples moved to the Extended Guidance.

A specific reference to data recovery provisions is added.

The cross-references are revised to address perceived and actual overlaps between R15 and other Requirements. They make it explicit that governance is covered in R3 (Continuity of access), governance of storage processes is covered in R9 (Documented storage procedures), and security arrangements are covered in R16 (Security).

EXTENDED GUIDANCE

Additional examples of relevant standards are provided. Detailed reference to data streams have been moved here from Guidance.

XVI. Security

Questions are added asking what levels of security are required and how they are supported, and what authentication and authorization procedures are used.

The Requirement cross-references now clarify that responses to R16 should focus on the measures taken to ensure premises, systems, and network security, in contrast to the overall technology focus of R15 (Technical infrastructure) and to the governance and application of data storage workflows of R9 (Documented storage procedures).

EXTENDED GUIDANCE

More specific questions have been added, including on security policies and access/Authentication and Authorization Infrastructure methods.

Appendix: Community Review Communications

Schedule

01 March 2019–30 April 2019: Open review period.

01–31 May 2019: Extension of open review period.

31 July 2019: Draft revision of the Core Trustworthy Data Repositories Requirements available for comment.

31 October 2019: Final version of the Requirements released.

01 November 2019: Suspension of application submissions against the Requirements for 2017–2019.

01 January 2020: Application submissions accepted against the Requirements for 2020–2022.

Webpage

'Review of Requirements' page on the CoreTrustSeal website announcing the 2019 Review and the resulting draft documents:

- <https://www.coretrustseal.org/why-certification/review-of-requirements/>
[Accessed 02-09-2019]

Email

Mailmerge 1

Email sent on 13–14 March 2019 to a mailing list of 541 CoreTrustSeal community members:

Subject

Review of CoreTrustSeal Requirements - Inputs Invited

Body Text

[You are receiving this email as a member of the CoreTrustSeal Community. We will only ever contact you in this way with information of importance. Please notify info@coretrustseal.org if you wish to be removed from our list.]

Dear Firstname Lastname,

A review of the [CoreTrustSeal](#) will take place in 2019 to define the Requirements for the period 2020–2023. This has no impact on the certifications of current CoreTrustSeal-certified repositories, which continue to run for three years from the date awarded.

The 2019 review process will focus on applicant feedback received during past reviews, other feedback received during communications and outreach activities, and an **open review period to run from 1 March 2019 to 30 April 2019**. Given the feedback received to date and the fact that a number of past WDS and DSA repositories are continuing to transition to the CoreTrustSeal Requirements, we expect the number, structure, and content of the Requirements to remain fundamentally stable for 2020–2023. This stability will enable us to work to support a wider range of repository types and repository technical infrastructure providers in the future. It will also provide an established baseline from which to consider other external demands on the CoreTrustSeal, including the development of the European Open Science Cloud and the further clarification of [FAIR Data Principles](#).

Feedback will be incorporated into revisions of the Requirements, Extended Guidance, and Glossary as necessary. **Please contact 2019review@coretrustseal.org or submit your feedback using the form at the bottom of [this page](#).**

A draft revision of the CoreTrustSeal Requirements will be made available by 31 July 2019 for comment and a final version will be released by 31 October 2019.

Applications and renewals against the current version of the Requirements are asked to submit by 1 November 2019 to minimize the number of reviews that overlap different versions of the Requirements. Applications against the revised requirements will be accepted from 1 January 2020.

Thank you in advance for your inputs,

Mailmerge 2

Email sent on 02–03 May 2019 to a mailing list of 541 CoreTrustSeal community members:

Subject

Review of CoreTrustSeal Requirements - review period extended till 31 May

Body Text

[You are receiving this email as a member of the CoreTrustSeal Community. We will only ever contact you in this way with information of importance. Please notify info@coretrustseal.org if you wish to be removed from our list.]

Dear Firstname Lastname,

Thank you to everyone who has sent us their valuable comments thus far. We are pleased to inform you that owing to popular demand, we are extending the deadline for community feedback until 31 May. We look forward to receiving your thoughts soon!

--

A review of the [CoreTrustSeal](#) will take place in 2019 to define the Requirements for the period 2020–2023. This has no impact on the certifications of current CoreTrustSeal-certified repositories, which continue to run for three years from the date awarded.

The 2019 review process will focus on applicant feedback received during past reviews, other feedback received during communications and outreach activities, and an **open review period to run from 1 March 2019 to ~~30 April 2019~~ 31 May 2019**...(Remainder of email as Mailmerge 1.)