



La certification des entrepôts de données

Françoise Genova, CDS/Observatoire Astronomique de Strasbourg, RDA France, GT Certification

Aude Chambodut, EOST, CNRS INSU, Board CTS

Olivier Rouchon, CNRS, Board CTS

Gilles Ohanessian, GT Certification

La certification des entrepôts de données

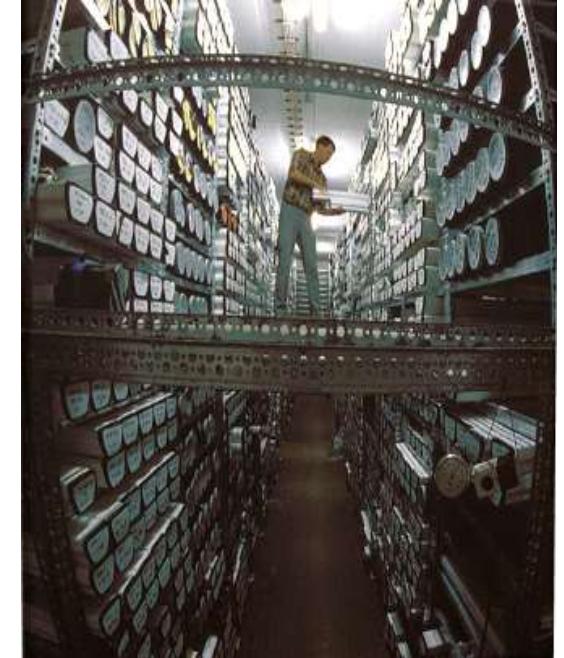
- › Pourquoi?
- › La procédure de certification
- › Les critères du CoreTrustSeal
- › Conclusions
- › Annexes
 - › Historique du CoreTrustSeal. Le rôle de la RDA
 - › Le modèle OAIS
 - › DataVerse et CoreTrustSeal

La confiance (Trustworthiness)

- › La confiance est au cœur de la Science Ouverte
- › Les entrepôts sont une source majeure de données
- › Il leur faut avoir la confiance
 - › de leurs utilisateurs
 - › Les personnes qui produisent et déposent les données
 - › Les utilisateurs des données
 - › de leurs autorités et de leurs financeurs

Qu'est-ce qu'un entrepôt de confiance?

- Mission de fournir un accès fiable, long-terme à des ressources numériques contrôlées pour sa communauté cible, maintenant et dans le futur
 - Supervision, planification et maintenance constantes
 - Compréhension des menaces et risques dans les systèmes
- Cycle d'audit et/ou de certification régulier



Credits: Ingrid Dillo & Hervé L'Hours

Qu'est-ce qu'un entrepôt de confiance?

- Mission de fournir un accès fiable, long-terme à des ressources numériques contrôlées pour sa communauté cible, maintenant et dans le futur
 - Supervision, planification et maintenance constantes
 - Compréhension des menaces et risques dans les systèmes
- Cycle d'audit et/ou de certification régulier



Credits: Ingrid Dillo & Hervé L'Hours



Pourquoi une certification formelle?

- › Assurer que le centre est « de confiance »
- › Mais... il a peut-être déjà la confiance de ses utilisateurs...
- › L'exemple du CDS
 - › Crée en 1972
 - › Centre de données de référence pour la communauté astronomique internationale
 - › Infrastructure de Recherche sur la Feuille de Route nationale
 - › 2 000 000+ requêtes/jour sur les services

Pourquoi une certification formelle?

- › Critères établis par des personnes compétentes et applicables quel que soit le cadre disciplinaire
- › Au préalable, **auto-évaluation** selon les critères, qui permet de vérifier l'organisation et les process et d'identifier des améliorations possibles
- › Evaluation externe par des personnes compétentes
- › Le dépôt dans un centre de données certifié est **un point important dans les Plans de Gestion des Données (PGD)**

La stratégie nationale

Plan National pour la Science Ouverte

2018

Structurer

- ➔ Généraliser la mise en place de plans de gestion des données dans les appels à projets de recherche
- ➔ Développer des centres de données thématiques et disciplinaires.
- ➔ Développer un service générique d'accueil et de diffusion des données simples.
- ➔ Engager un processus de certification des infrastructures de données.



Organiser

- ➔ Soutenir la **Research data alliance** (RDA) et créer le chapitre français de l'alliance (RDA France).
- ➔ Soutenir **Software heritage**, la bibliothèque des codes sources

2021

- ➔ Poursuivre le processus de **certification** (*Core trust seal*) des entrepôts de données français.

Les entrepôts français actuellement certifiés CoreTrustSeal

- › [CDS](#), Strasbourg
- › [CDSP](#), Paris
- › [HAL](#), Paris
- › [IDOC/IDOC-DATA](#), Orsay
- › [IFREMER SISMER](#), Plouzané
- › [IPSL](#), Paris
- › [ORTOLANG](#), Nancy
- › Epos-France SI-Sismo ([RESIF](#)),
Grenoble
- › [SHOM](#), Brest
- › [SPAR/BNF](#), Paris
- › [ESRF](#), Grenoble



Les cadres de certification



Centres de données de confiance - Trustworthy Data
Repositories

Certification formelle ISO 16363:
2012 > 2025

(2 Centres ont été certifiés, 100+ critères)
<https://www.iso.org/standard/87472.html>

Certification étendue DIN 31644
(6 Centres certifiés, 34 critères)

www.langzeitarchivierung.de

Certification CoreTrustSeal

(200+ centres, 16 critères)
www.CoreTrustSeal.org

La procédure de certification

La procédure d'évaluation (1)

- › Auto-évaluation de l'entrepôt
- › Soumission d'un formulaire en ligne
- › Examen par deux évaluateurs
- › Examen des évaluations par le Board
 - › Renvoi des évaluations des évaluateurs et des commentaires du Board aux candidats
 - › Acceptation du dossier



La procédure d'évaluation (2)

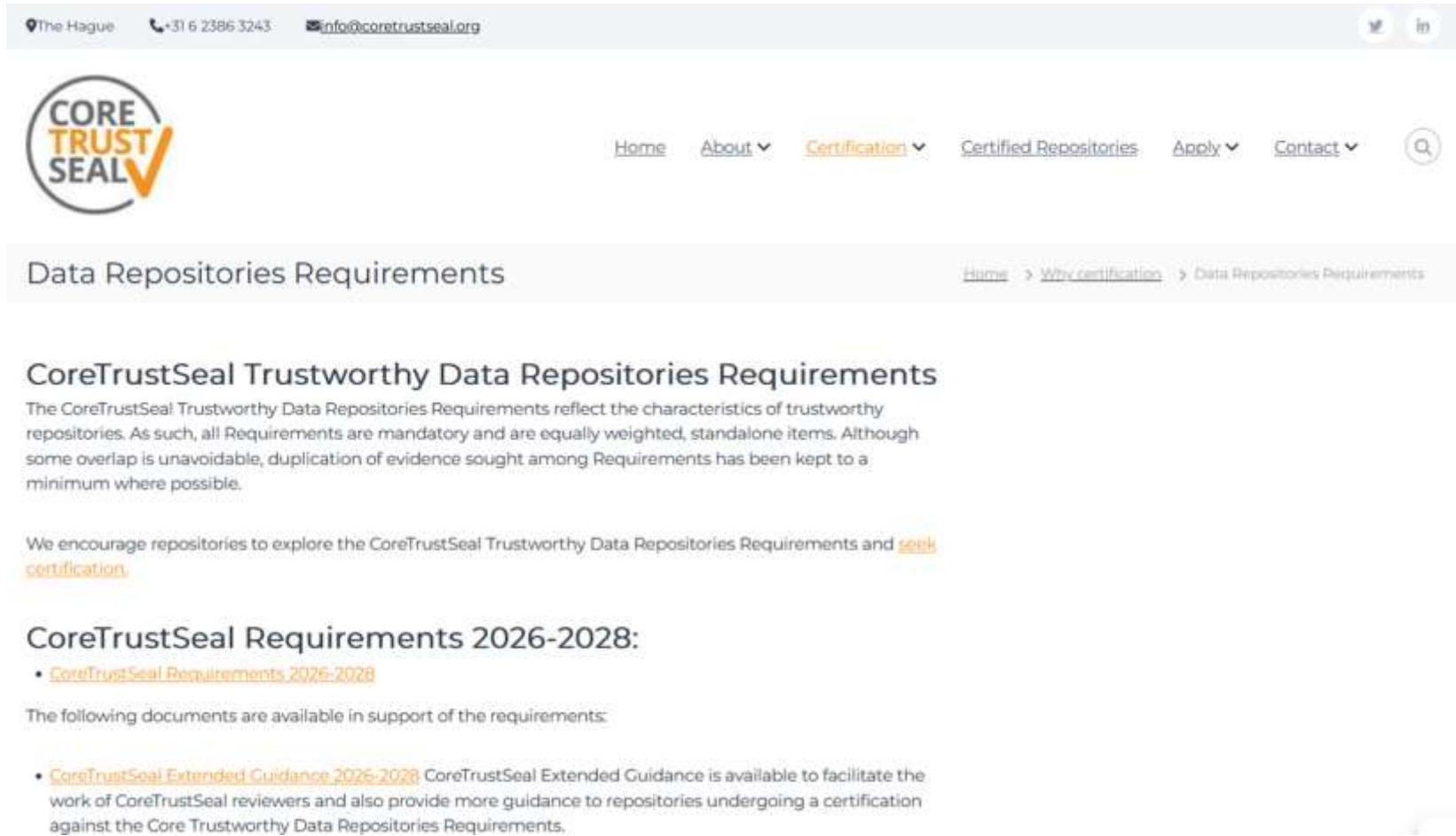
- › Les éléments du dossier sont publics (sauf exception légitime)
- › Les évaluateurs sont membres de la communauté CoreTrustSeal
 - « Assemblée des évaluateurs »
- › Pour devenir membre du World Data System (WDS)
 - › Il faut obtenir la certification CTS...
 - › et remplir quelques conditions supplémentaires
- › Renouvellement de la certification tous les 3 ans

Les critères du CoreTrustSeal

La certification CoreTrustSeal

- Toute l'information est sur le site de CoreTrustSeal
 - <https://www.coretrustseal.org/>
- Contexte + 16 critères
 - <https://www.coretrustseal.org/why-certification/requirements/>
- Document pour guider les évaluateurs et les candidats
 - Extended Guidance 2020-2022 (V2.0) > 2023-2025 (V3.0) > 2026-2028 (V01.00)
<https://doi.org/10.5281/zenodo.17659853>
 - Glossaire
<https://doi.org/10.5281/zenodo.17660010>
 - Changements
<https://doi.org/10.5281/zenodo.17661225>
- Lorsque les critères évoluent, les entrepôts qui ont soumis un dossier selon les anciens critères restent dans ce cadre
- « Administrative fee » 3000€

Les critères sur le site CoreTrustSeal



The screenshot shows the CoreTrustSeal website's "Data Repositories Requirements" page. At the top, there is a header with links for "The Hague", "+31 6 2386 3243", and "info@coretrustseal.org". Social media icons for YouTube and LinkedIn are also present. The main navigation menu includes "Home", "About", "Certification", "Certified Repositories", "Apply", and "Contact". A search icon is located on the right side of the menu.

The page title is "Data Repositories Requirements". The breadcrumb navigation shows "Home > Why certification > Data Repositories Requirements".

CoreTrustSeal Trustworthy Data Repositories Requirements

The CoreTrustSeal Trustworthy Data Repositories Requirements reflect the characteristics of trustworthy repositories. As such, all Requirements are mandatory and are equally weighted, standalone items. Although some overlap is unavoidable, duplication of evidence sought among Requirements has been kept to a minimum where possible.

We encourage repositories to explore the CoreTrustSeal Trustworthy Data Repositories Requirements and [seek certification](#).

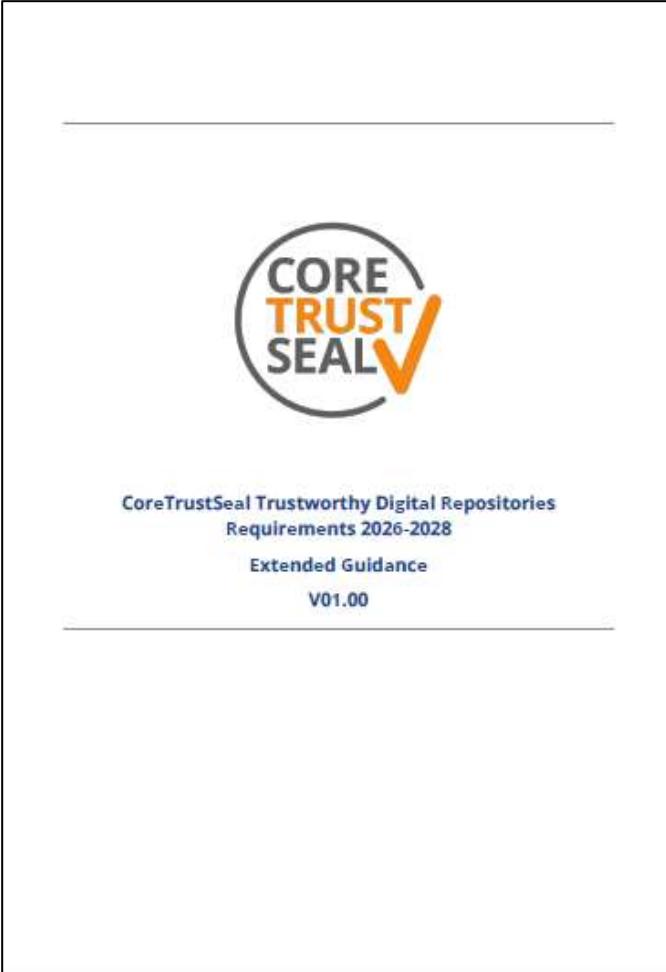
CoreTrustSeal Requirements 2026-2028:

- [CoreTrustSeal Requirements 2026-2028](#)

The following documents are available in support of the requirements:

- [CoreTrustSeal Extended Guidance 2026-2028](#) CoreTrustSeal Extended Guidance is available to facilitate the work of CoreTrustSeal reviewers and also provide more guidance to repositories undergoing a certification against the Core Trustworthy Data Repositories Requirements.

Les critères



- › R0 - Le contexte
- › 16 critères, 3 thèmes:
 - › Infrastructure organisationnelle
 - › Gestion des objets numériques (données et métadonnées)
 - › Technologie de l'information et sécurité

Le contexte

- › Identifiant Re3data
- › Type d'entrepôt
- › Vue d'ensemble de l'entrepôt
- › Description de la communauté concernée
- › Niveau de curation et préservation
 - Z. Level Zero. Content distributed as deposited. Unattended deposit-storage-access.
 - D. Deposit Compliance
 - C. Initial Curation
 - A. Active preservationPartenaires
- › Coopération et externalisation auprès de tiers, partenaires et organisation d'accueil
- › *Résumé des modifications depuis la candidature précédente (s'il y a lieu)*

Infrastructure organisationnelle

- › R01 – Mission/périmètre
- › R02 – Gestion des droits
- › R03 – Continuité de service
- › R04 – Légalité/éthique
- › R05 – Gouvernance et ressources
- › R06 – Compétences et accompagnement



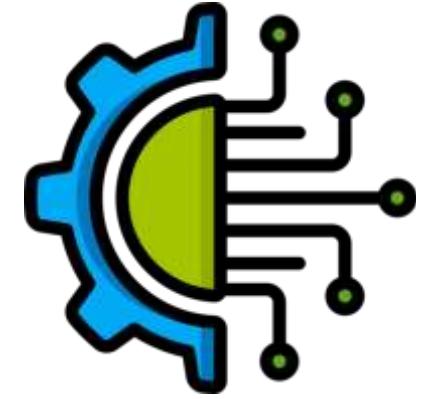
Gestion des objets numériques

- › R07 – Provenance et authenticité
- › R08 – Dépôt, appréciation et accessibilité
- › R09 – Plan de préservation
- › R10 – Assurance qualité
- › R11 – Processus de traitement (Workflows)
- › R12 – Découverte et identification
- › R13 – Réutilisation



Technologies de l'information et sécurité

- › R14 – Stockage et intégrité
- › R15 – Infrastructure technique
- › R16 – Sécurité



Révision des critères 2026-2028



- Revue des dossiers de candidatures des trois dernières années
 - Proposition de changements et additions
-
- Sondage de la communauté
 - Revue des retours
 - Incorporation des retours au projet de révision
 - Revue par le Board et approbation de la nouvelle version
-
- Publication de la version finale
 - Correction des dernières erreurs par la communauté et le Board
 - Publication des critères 2026-2028

Conclusions

Pourquoi la certification

- › Un travail d'équipe dans la plupart des cas (tout compris)
 - › De quelques semaines à beaucoup plus mais ça vaut le coup!
- › Evaluation – amélioration des process
 - › Auto-évaluation
 - › Evaluation externe
- › Importance croissante pour les financeurs des centres de données et des projets (PGD)
- › Priorité au niveau politique en France, intérêt des organismes (CNRS, Universités)

L'impact de la RDA

- › Fusion des deux cadres de certification ‘de base’, DSA (Data Seal of Approval) et WDS (World Data System)
- › Clarification du paysage pour les centres de données et les agences de financement
- › Deux cadres complémentaires au départ: le résultat est meilleur que chacun des originaux!
- › Création de CoreTrustSeal
- › Nombreux nouveaux candidats à la certification

Le rôle de RDA France

- La certification est une priorité – aussi parce que la réflexion sur les critères permet d'améliorer la manière dont on gère ses données, même si on ne va pas jusqu'à candidater à CoreTrustSeal
- Groupe de Travail commun avec le collège Données du Comité pour la Science Ouverte depuis juillet 2021
 - Prise en charge des frais administratifs de certification
 - Outil de visualisation et de partage des dossiers de certification
<https://crusoe.ouvrirlascience.fr/>
- Ateliers, présentations à la demande, etc.
 - Plus de 300 personnes ont suivi les ateliers
- Démarrage de la mise en réseau des entrepôts certifiés et « sur le chemin de la certification »
- Liste de diffusion
<https://listes.services.cnrs.fr/wws/subscribe/rda-france-certification>

Informations complémentaires

- › Le [support de l'atelier de septembre 2025](#) contient en Annexe (à partir de la diapositive n°30)
 - › L'historique de CoretrustSeal
 - › Des informations sur le modèle OAIS et sa mise à jour de décembre 2024 (V3)
 - › Modèle de référence sous-jacent dans CoreTrustSeal OAIS – Open Archive Information System Version 2
 - › La version actuelle: Version 3, décembre 2024
 - <https://public.ccsds.org/Pubs/650x0m3.pdf>
 - › Site en français
 - <https://www.cines.fr/archivage/un-concept-des-problematiques/le-modele-de-reference-loais/>
- › Des informations sur DataVerse et CoreTrustSeal

Liens utiles

RDA Global

Contact - <https://www.rd-alliance.org/contact/>

Web - www.rd-alliance.org

Twitter - [@resdatall](https://twitter.com/resdatall)

LinkedIn - www.linkedin.com/in/ResearchDataAlliance

Slideshare -

[http://www.slideshare.net/ResearchDataAlliance](https://www.slideshare.net/ResearchDataAlliance)

RDA FRANCE

<https://www.rd-alliance.org/groups/rda-france/activity/>

Email - contact-rdafrance@services.cnrs.fr