



La RDA – Certification des entrepôts de données

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

La dix-septième plénière

- La prochaine plénière aura lieu du 20 au 22 avril
- RDA France peut prendre en charge les frais d'inscription. Pour demander un soutien:
 - Modalités: <https://www.rd-alliance.org/group/rda-france/wiki/participation-la-dix-septieme-session-pleniere-de-la-rda>
 - Formulaire: <https://framaforms.org/rda-france-rda-p17-1614242381>
 - **Date limite: 1^{er} avril**
- Atelier en français « La RDA pour les Nouveaux Venus » le 19 avril à 15h heure de Paris
 - <https://www.rd-alliance.org/group/rda-france/wiki/la-rda-pour-les-nouveaux-venus-19-avril-2021>

Parmi les activités

- Appels à commentaires en cours
 - [Sensitive Data Interest Group Charter](#)
 - [RDA Data Usage Metrics WG Recommendations](#)
- WG FAIR Data Maturity Model
 - Recommandation
 - <https://www.rd-alliance.org/group/fair-data-maturity-model-wg/outcomes/fair-data-maturity-model-specification-and-guidelines-0>
 - Continue en Maintenance Group
 - <https://www.rd-alliance.org/groups/fair-data-maturity-model-wg>
 - Retours sur l'implémentation
 - Mise à jour de la recommandation
- WG RDA/FORCE11/ReSA FAIR for Research Software
 - <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>
 - Travail sur la définition de FAIR pour les logiciels
 - Ils vont proposer un Groupe de Travail pour définir des métriques

Pour recevoir les informations

- › Liste RDA France
 - › <https://listes.services.cnrs.fr/wws/subscribe/rda-france>
- › Liste Certification
 - › <https://listes.services.cnrs.fr/wws/subscribe/rda-france-certification>
- › Liste Entrepôts
 - › <https://listes.services.cnrs.fr/wws/subscribe/rda-france-entrepots>
- › Liste FAIR
 - › <https://listes.services.cnrs.fr/wws/subscribe/rda-france-fair>
 - › Pour proposer un contenu: rda-france-fair-request _à_ services.cnrs.fr
- › Mail de contact pour RDA France
contact_rdafrance _à_ services.cnrs.fr

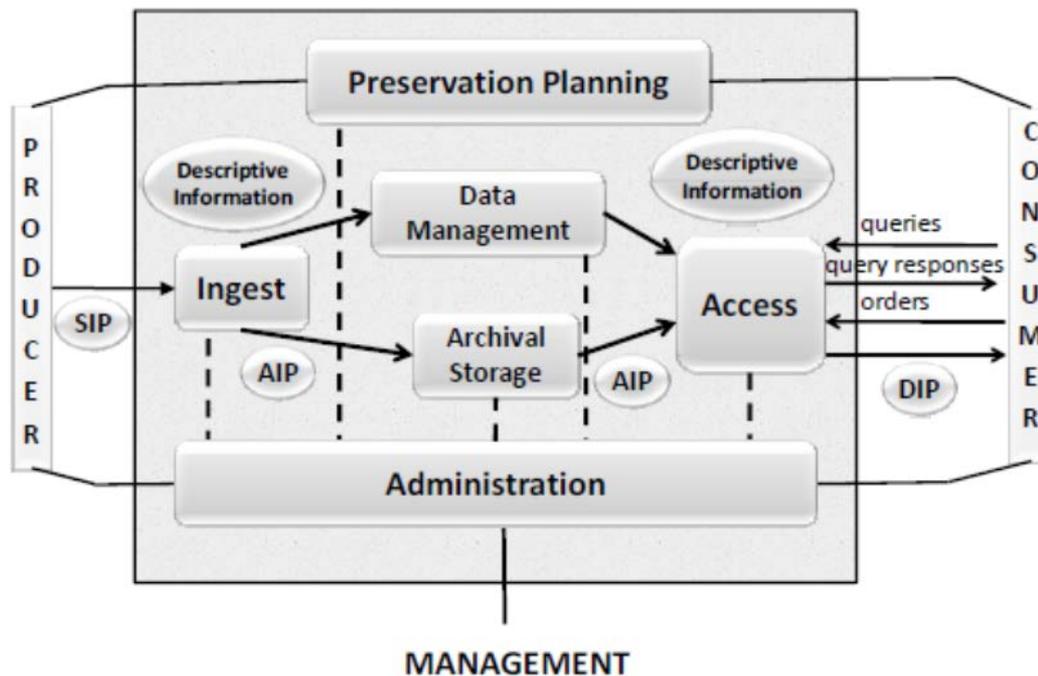
La certification des entrepôts de données

- Une des mesures du Plan National pour la Science Ouverte: Engager un processus de certification des infrastructures de données
- Cité dans la Feuille de Route Science Ouverte du CNRS (nov. 2019) pour les infrastructures de recherche qui produisent des données
- Plan Données de la Recherche du CNRS (nov. 2020): objectif d'aller vers la certification des entrepôts et services de données
- Une des priorités de RDA France
 - ~200 personnes ont suivi un des ateliers « débutants »
 - Groupe de Travail commun à RDA France et au Collège Données du Comité pour la Science Ouverte pour mettre en place des activités en soutien
 - Enquête sur le paysage de la certification
 - <https://framaforms.org/entrepots-et-services-de-donnees-interesses-par-la-demarche-de-certification-1590503416>
 - Enquête sur les critères de certification
 - <https://framaforms.org/criteres-de-certification-1600678445>

Pourquoi la certification?

- Répondre à l'injonction « politique »
- Obtenir un label de « centre de confiance » attribué après une évaluation par des experts internationaux
- Même si on ne va pas jusqu'à la certification, examen détaillé et identification d'améliorations à apporter aux pratiques dans les différentes « cases » de l'activité (organisation, gestions des données, aspects techniques)
- Gilles hier: « Un cercle vertueux »
- Il faut construire la documentation qui doit être publique (sauf exception justifiée)

Pour la gestion des données: le modèle OAIS



SIP: Submission Information Package

AIP: Archival Information Package

DIP: Dissemination Information Package

La certification CTS

- 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
 - En cours, Extended Guidance V0.2
 - <https://doi.org/10.5281/zenodo.3632533>
 - Traduction française par RDA France **de la version précédente**
 - <https://www.rd-alliance.org/coretrustseal-criteres-de-conformite>

- « Administrative fee » 1000€

Auto-évaluation

- Questionnaire à remplir
- Il faut les compétences
 - De la direction (mission, organisation, ...)
 - Des personnes en charge du contenu
 - Des personnes en charge de l'informatique
- Pour le CDS: un travail d'équipe qui a impliqué la direction, les documentalistes, l'informaticien en charge du service et l'ingénieur système

La procédure de certification

- › 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

- › Les éléments du dossier sont publics (sauf exception légitime)
- › Les évaluateurs sont membres de la communauté CoretrustSeal

A l'INSU

- L'IR Data Terra est engagée vers la certification de ses Centres de Données et de Services
 - Deux ANR (ANR Flash, Terre Solide et Océan)
- Priorité de l'ASOV
 - Participation soutenue aux formations
 - Retours apportés à l'enquête sur les critères– MERCI!
- Une atelier « avancé » pour l'astronomie?
- Aide de collègues du GT Certification + Voir si une participation du Board CoreTrustSeal est possible

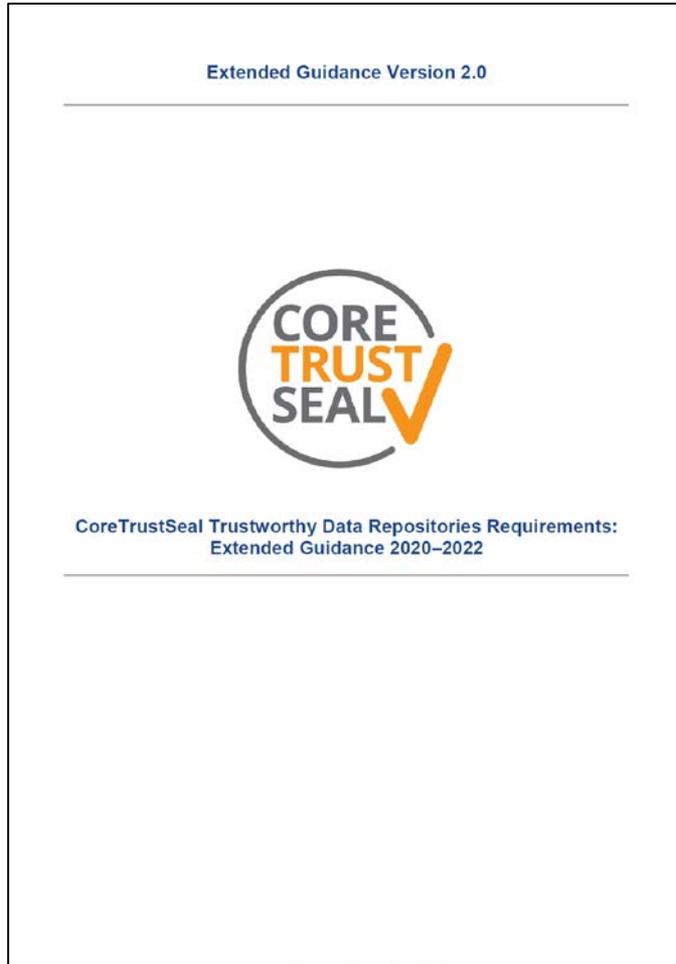
Le modèle de certification européen



Centres de données de confiance - Trustworthy Data Repositories

thanks to Mustapha Mokrane (DANS)- NestorSeal and ISO numbers updated 22 January 2021

Les critères de certification CTS



- R0 - Le contexte
- 16 critères, 3 thèmes:
 - Infrastructure organisationnelle
 - Gestion des objets numériques (données et des métadonnées)
 - Technologie

Le contexte

- › Type d'entrepôt
- › Brève description de l'entrepôt
- › Brève description de la communauté concernée
- › Niveau de curation
 - › Contenu en accès tel que déposé
 - › Curation de base (p. ex. vérification rapide, ajout de métadonnées de base ou de documentation)
 - › Curation avancée (p. ex. conversion vers de nouveaux formats, amélioration de la qualité de la documentation)
 - › Curation au niveau des données
- › Partenaires
- › *Résumé des modifications depuis la candidature précédente (s'il y a lieu)*
- › Autres informations pertinentes

Infrastructure organisationnelle

- R1 – Mission/périmètre
- R2 – Licenses
- R3 – Continuité de l'accès
- R4 – Confidentialité/éthique
- R5 – Infrastructure organisationnelle
- R6 – Conseils d'experts

Gestion des objets numériques

- › R7 – Intégrité et authenticité des données
- › R8 – Appréciation et sélection des données
- › R9 – Procédures d'archivage documentées
- › R10 – Plan de préservation
- › R11 – Qualité des données
- › R12 – Processus de traitement (Workflows)
- › R13 – Découverte et identification des données
- › R14 – Réutilisation des données

Technologie

- R15 – Infrastructure technique
- R16 – Sécurité