

Guidage étendu version 2.0



**Exigences relatives aux référentiels de données fiables CoreTrustSeal:
Orientation étendue 2020–2022**

Contenu

Introduction	3
Contexte et conseils généraux	3
Glossaire des termes	4
Conseils généraux étendus	4
Introduction : Généralités	4
Informations/preuves manquantes	4
Compréhensibilité de la documentation	4
Documentation en langue autre que l'anglais....	5
Documents sensibles et autres documents internes.....	5
Structure et longueur de l'application	5
Exigences	6
Informations générales	6
Le contexte	6
Infrastructure organisationnelle	10
1. Mission/ Périmètre	dix
2. Licences	11
3. Continuité d'accès	12
4. Confidentialité/ Ethique	13
5. Infrastructure organisationnelle	14
6. Conseils d'experts	15
7. Gestion des objets numériques	16
8. Intégrité et authenticité des données.....	16
9. Évaluation	17
10. Procédures de stockage documentées	18
11. Plan de conservation	19
12. Qualité des données	20
13. Flux de travail	21
14. Découverte et identification des données.....	22
15. Réutilisation des données.....	23
16. Technologie.....	24
17. Infrastructures techniques	24
18. Sécurité.....	25
19. Rétroaction du candidat	26
20. Commentaires/réactions	26

Introduction

Ce document contient le texte intégral des **exigences CoreTrustSeal Trustworthy Data Repositories** pour 2020-2022 avec des paragraphes d'introduction sur le contexte et les orientations générales.

En plus des exigences CoreTrustSeal, qui restent stables pour la période 2020-2022, ce document fournit les **directives étendues** pour les examinateurs et les candidats CoreTrustSeal. Le texte des directives étendues peut être mis à jour au cours de la période 2020-2022 sous réserve de l'approbation du conseil d'administration de CoreTrustSeal. Le document contient également une référence au glossaire des termes.

Le document est destiné à maximiser la cohérence des examens à travers le large éventail de candidats CoreTrustSeal. Le public principal est les évaluateurs, mais il est également utile pour les candidats lors de la préparation d'une auto-évaluation de candidature.

Contexte et conseils généraux

Les exigences relatives aux *référentiels de données fiables CoreTrustSeal* décrivent les caractéristiques des référentiels fiables. Toutes les exigences sont obligatoires et évaluées en tant qu'éléments autonomes.

Bien qu'un certain chevauchement soit inévitable, la duplication des preuves recherchées pour chaque exigence a été réduite au minimum. Les options dans les listes de contrôle (par exemple, le type de dépôt et le niveau de conservation) ne sont pas considérées comme exhaustives et peuvent être affinées à l'avenir. Les candidats sont encouragés à ajouter d'autres options.

Chaque exigence est accompagnée d'un texte d'orientation décrivant les informations et les preuves que les candidats doivent fournir pour permettre un examen objectif.

Le demandeur doit indiquer un niveau de conformité pour chacune des exigences :

- 0 – Non applicable
- 1 – Le référentiel n'a pas encore pris cela en compte
- 2 – Le référentiel a un concept théorique
- 3 – Le référentiel est en phase de déploiement
- 4 – La ligne directrice a été entièrement mise en œuvre dans le référentiel

Les niveaux de conformité sont un indicateur des progrès auto-évalués du candidat, mais les examinateurs jugent la conformité par rapport aux déclarations de réponse et aux preuves à l'appui. Si un candidat estime qu'une exigence n'est pas applicable (0), cela doit être justifié en détail. Les niveaux de conformité de 1 ou 2 ne sont pas suffisants pour une candidature réussie. La certification peut être accordée si certaines exigences sont en phase de mise en œuvre (3).

Les déclarations de réponse fournies par les candidats doivent inclure des liens vers des preuves à l'appui en ligne. Étant donné que le processus de certification de base n'inclut pas de visite du site par un auditeur, ces preuves accessibles au public fournissent une assurance transparente de bonnes pratiques. Les liens URL doivent être vérifiés immédiatement avant de soumettre les candidatures.

Toutes les réponses doivent être en anglais. Bien que des tentatives soient faites pour jumeler les évaluateurs aux candidats en termes de langue et de discipline, cela n'est pas toujours possible. Des traductions complètes des preuves ne sont pas requises, mais si des preuves non anglaises sont fournies, un résumé en anglais doit être inclus dans la déclaration de réponse.

Aucune divulgation d'informations sensibles n'est requise pour acquérir le CoreTrustSeal, mais des dispositions sont prises dans le cadre du processus de certification pour les référentiels qui souhaitent partager des éléments de preuve contenant également des informations confidentielles.

Le CoreTrustSeal est valable trois ans à compter de la date d'attribution. Bien que les systèmes et les capacités du référentiel évoluent en permanence en fonction de la technologie et des besoins des utilisateurs, ils pourraient ne pas subir de changements majeurs dans ce laps de temps. Une organisation avec des processus opérationnels et des dossiers bien gérés devrait être en mesure de présenter une nouvelle demande avec des révisions minimales après trois ans, sauf si :

- L'organisation, sa collecte de données ou la communauté désignée a considérablement changé.

- Les exigences CoreTrustSeal ont été mises à jour de manière à avoir un impact sur le candidat.

Les exigences CoreTrustSeal sont soumises à un examen et à une révision tous les trois ans. Cela n'affecte pas un candidat retenu jusqu'à ce qu'il demande un renouvellement.

Glossaire des termes

Veuillez vous référer au glossaire des exigences relatives aux principaux référentiels de données dignes de confiance : <https://doi.org/10.5281/zenodo.3632563>.

Orientation générale étendue

Introduction : Généralités

La révision des informations sur les preuves tous les trois ans à des fins de certification uniquement n'est ni efficace ni efficace. La gestion continue des informations commerciales nécessaires à l'exécution des services d'un référentiel doit être suffisante pour demander et conserver la certification. Plus précisément, les référentiels qui documentent suffisamment bien leurs politiques et procédures pour garantir que la qualité reste constante, que le risque de départ du personnel est atténué, etc., devraient uniquement préparer des réponses aux applications et gérer des versions publiques de leurs preuves.

Les examinateurs peuvent proposer un niveau de conformité différent de celui sélectionné par le demandeur. Si les examinateurs changent le niveau sélectionné pour une *valeur inférieure*, la raison sera expliquée. Il y a une attente de progrès pour une exigence avec un niveau de conformité inférieur à 4 lorsque la certification est renouvelée.

Il n'est pas possible de couvrir tous les scénarios de stockage possibles dans le Guide ou le Guide étendu. De même, toutes les puces de toutes les exigences ne sont pas obligatoires. Les réponses des candidats doivent faire référence aux problèmes soulevés dans le texte du Guide et fournir des réponses en fonction de leur contexte local. L'évaluation finale d'une exigence dépend de l'exhaustivité et de la qualité de la réponse. Les évaluateurs recherchent des déclarations de preuve claires et ouvertes spécifiques au candidat.

Les concepts et la terminologie utilisés dans les Exigences sont informés par le Modèle de référence OAIS. L'utilisation de la terminologie de l'OAIS peut aider à assurer la compréhensibilité et la clarté de la candidature, et les candidats sont donc fortement encouragés à se familiariser avec l'OAIS avant de préparer les réponses aux exigences. Le rapport de veille technologique 2014 de DPC 'The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)' par Brian Lavoie (<https://doi.org/10.7207/twr14-02>) fournit une introduction utile au modèle de référence OAIS.

Informations/preuves manquantes

Le processus de certification CoreTrustSeal dépend des réponses étayées par des preuves. La qualité des preuves publiques à l'appui devrait s'améliorer avec le temps. Les demandes sont plus difficiles à évaluer si les informations sont manquantes, insuffisantes ou peu claires; si les liens URL sont rompus; ou si les déclarations de preuves se recoupent continuellement.

La familiarité ou la connaissance interne du référentiel par l'examineur ne doit pas jouer un rôle dans l'évaluation des preuves disponibles. La déclaration de preuve publique finale doit également être claire pour que les référentiels homologues puissent la comprendre.

On ne s'attend pas à ce qu'un examinateur effectue une recherche sur le site Web du candidat pour trouver des preuves. Si les informations fournies sont insuffisantes pour permettre à l'examineur de prendre une décision, la demande sera renvoyée avec une explication des raisons pour lesquelles les preuves sont insuffisantes. Aucun niveau de conformité ne sera donné à ce stade par l'examineur.

Compréhensibilité de la documentation

Les examinateurs et les lecteurs de l'examen public final devraient être en mesure de comprendre les réponses aux exigences sans avoir à lire en détail les preuves à l'appui. Lorsque des documents plus longs sont présentés

comme preuve, ou un document est utilisé comme preuve pour plus d'une exigence, le demandeur doit se référer spécifiquement aux sections pertinentes et citer/résumer les informations dans sa réponse.

Documentation en langue non anglaise

La documentation dans des langues autres que l'anglais est acceptable si son contenu est suffisamment et clairement expliqué dans un résumé en anglais. Ce résumé peut être assez bref pour certains types de documents (par exemple, une liste de formats préférés), mais devrait être plus long pour d'autres (par exemple, un document de politique de préservation).

Documents sensibles et autres documents internes

La certification CoreTrustSeal n'exige pas la publication d'informations complémentaires confidentielles, commercialement sensibles ou présentant un risque pour la sécurité. Cela s'applique également aux documents qui ne sont disponibles que sur l'intranet d'un référentiel. Les candidats peuvent disposer d'informations commerciales contenant à la fois des informations sensibles et des preuves pertinentes pour le CoreTrustSeal. Ces preuves peuvent être soumises de manière confidentielle aux examinateurs et aux documents nommés et décrits dans la demande¹. Au fil du temps, nous nous attendons à ce que les candidats séparent les preuves pertinentes des documents confidentiels et s'assurent qu'une version publique est mise à disposition pour le prochain examen.

Si la documentation n'existe pas encore, est en cours de réalisation ou est actuellement destinée à un usage interne uniquement (par exemple, un wiki), une date de disponibilité publique doit être indiquée dans la candidature. La certification peut être approuvée sur la base de ces assurances. Les candidats doivent fournir la documentation publique lorsqu'ils renouvellent leur certification.

Structure et longueur de l'application

Les candidatures en cours sont confidentielles pour les examinateurs CoreTrustSeal et le Conseil, mais les candidatures retenues sont rendues publiques. Les candidats doivent donc garder à l'esprit tous ces publics. Les candidatures ne doivent pas répondre à chaque élément d'orientation sous forme de questions-réponses.

Les candidatures doivent inclure des réponses en texte à chaque exigence, incorporant les éléments pertinents des directives et des directives détaillées fournies.

Le conseil d'administration de CoreTrustSeal comprend que les candidats proviennent d'un large éventail d'organisations dont la mission, la taille et la complexité varient à la fois en termes de structure organisationnelle et de variété de collecte de données. Même les directives étendues ne peuvent pas couvrir tous les sujets et types de preuves qui pourraient être pertinents pour la demande. Nous comprenons également qu'il faut de l'espace pour expliquer la pertinence des preuves fournies; surtout s'il n'est pas disponible en anglais. La Commission ne fixe pas de longueurs minimales ou maximales pour les réponses, mais d'après son expérience, même les déclarations de preuves les plus complexes se situent à l'extrémité inférieure de la plage de 500 à 800 mots. Dans la mesure du possible, les déclarations de preuves doivent être étayées par des liens publics vers la documentation utilisée pour régir votre organisation et gérer vos objets numériques. C'est cette preuve publique qui offre le plus d'assurance qu'une organisation gère ses collections en tant que référentiel de données digne de confiance.

Les candidats ne devraient pas avoir besoin de répéter de longues portions de texte dans différentes réponses aux exigences. Dans les cas où les preuves s'appliquent à plus d'une exigence, un bref résumé des informations pertinentes doit être ajouté, puis un renvoi à la réponse appropriée pour plus de détails.

¹ Les documents confidentiels doivent être soumis au Secrétariat CoreTrustSeal via info@coretrustseal.org

Conditions

Informations d'arrière-plan

Le contexte

R0. Veuillez fournir le contexte de votre référentiel.

– **Type de référentiel. Sélectionnez tous les types pertinents parmi :**

- Référentiel basé sur un domaine ou un sujet
- Référentiel institutionnel
- Système de référentiel national, y compris gouvernemental
- Référentiel de publications
- Bibliothèque
- Musée
- Archiver
- Référentiel de projets de recherche
- Autre (Veuillez décrire)

- *Brève description du référentiel*

- *Brève description de la communauté désignée*

– **Niveau de conservation effectué. Sélectionnez tous les types pertinents parmi :**

- A. Contenu distribué tel que déposé
- B. Conservation de base - par exemple, vérification brève, ajout de métadonnées de base ou de documentation
- C. Amélioration de la conservation - par exemple, conversion vers de nouveaux formats, amélioration de la documentation
- D. Conservation au niveau des données - comme dans C ci-dessus, mais avec une édition supplémentaire des données déposées pour précision

commentaires

– *Partenaires internes/externalisés. Le cas échéant, veuillez les énumérer.*

- *Résumé des changements importants depuis la dernière demande (le cas échéant)*

– *Autres informations pertinentes*

Réponse

Conseils:

Les informations contenues dans cette section fournissent le contexte et le contexte dont les examinateurs ont besoin pour évaluer pleinement les réponses aux autres exigences. Il est donc d'une importance vitale pour l'ensemble de la demande que des réponses détaillées soient données à chaque question. Veuillez sélectionner parmi les options et fournir des détails pour les éléments qui apparaissent dans l'exigence contextuelle.

(1) Type de référentiel. Cet élément aidera les réviseurs à comprendre quelle fonction exécute votre référentiel. Choisissez la meilleure correspondance pour votre type de référentiel (sélectionnez tout ce qui s'applique). Si aucune des catégories n'est appropriée, n'hésitez pas à fournir un autre type descriptif. Vous pouvez également fournir des détails supplémentaires pour aider l'examineur à comprendre votre type de référentiel.

(2) Brève description du référentiel. Fournir un bref aperçu du référentiel ; en particulier, veuillez ajouter des informations sur le type de données acceptées par le référentiel (c'est-à-dire l'étendue de sa collecte). Si le référentiel a des partenaires externes, fait partie d'un réseau ou d'une organisation mère, la réponse doit idéalement inclure un schéma et une description de la structure organisationnelle globale.

(3) Communauté désignée. Une définition claire de la communauté désignée démontre que

le candidat comprend la portée, la base de connaissances et les méthodologies, y compris les logiciels/formats préférés, de la ou des communautés d'utilisateurs qu'il cible. Veuillez vous assurer que la réponse est suffisamment précise pour permettre aux évaluateurs d'évaluer l'adéquation des mesures de conservation et de préservation décrites tout au long de la candidature.

(4) Niveau de conservation. Cet élément est destiné à déterminer si le référentiel distribue son contenu aux consommateurs de données sans aucun changement, ou si le référentiel ajoute de la valeur en améliorant le contenu d'une manière ou d'une autre. Tous les niveaux de conservation supposent (1) que les dépôts initiaux sont conservés tels quels et que les modifications ne sont effectuées que sur des copies de ces originaux, et (2) des métadonnées qui permettent à la communauté désignée de comprendre et d'utiliser les données de manière indépendante (c'est-à-dire sans avoir à consulter le créateur original) est présent au dépôt ou ajouté par le référentiel. Les annotations/modifications doivent respecter les termes de la licence convenue avec le producteur de données et être clairement dans les compétences de ceux qui se chargent de la conservation. Ainsi, le référentiel devra démontrer que ces annotations/modifications sont entreprises et documentées par des experts appropriés et que l'intégrité de toutes les copies originales est maintenue. Le savoir aidera les examinateurs à évaluer d'autres exigences de certification. D'autres détails peuvent être ajoutés pour aider à comprendre les niveaux de conservation que vous entreprenez.

(5) Partenaires internes/externalisés. Veuillez fournir une liste des partenaires avec lesquels votre organisation travaille, décrivant la nature de la relation (organisationnelle, contractuelle, etc.) et indiquant si le partenaire a entrepris une évaluation fiable du référentiel. Si une fonction ou une preuve à l'appui n'est pas sous le contrôle direct du demandeur, elle entre alors dans cette catégorie. Il peut s'agir d'une organisation hôte ou d'une autre relation « d'internalisation », ou d'une externalisation ou d'une autre dépendance vis-à-vis d'un tiers. Ces relations peuvent inclure, mais sans s'y limiter : tout service fourni par une institution dont vous faites partie, le stockage fourni par d'autres dans le cadre de la redondance multicopie ou l'adhésion à des organisations qui peuvent entreprendre la gestion de votre collecte de données lorsqu'un problème de continuité des activités survient. De plus, veuillez énumérer les exigences de certification pour lesquelles le partenaire fournit tout ou partie de la fonctionnalité/du service pertinent, y compris les contrats ou accords de niveau de service en place. Étant donné que l'externalisation sera presque toujours partielle, vous devrez toujours fournir des preuves appropriées pour les exigences de certification qui ne sont pas externalisées et pour les parties du cycle de vie des données que vous contrôlez.

Les qualifications/certifications, y compris, mais sans s'y limiter, la certification CoreTrustSeal (et ses prédécesseurs) sont préférées pour les partenaires d'externalisation. Cependant, il n'est pas nécessaire qu'ils soient certifiés. Nous comprenons qu'il peut s'agir d'un domaine complexe à définir et à décrire, mais ces détails sont essentiels pour garantir un processus d'examen complet.

(6) Résumé des changements importants depuis la dernière demande. La certification CoreTrustSeal s'attend à une amélioration continue. Les référentiels en cours de recertification doivent souligner brièvement aux examinateurs tout changement significatif dans les systèmes techniques, la communauté désignée, le financement, etc. au cours des trois années précédentes. Ce faisant, veuillez vous référer à tous les commentaires qui vous ont été donnés par les examinateurs de votre précédente application CoreTrustSeal. Des informations détaillées sur un changement doivent être ajoutées à l'exigence appropriée.

(7) Autres informations pertinentes. Le référentiel peut souhaiter ajouter des informations contextuelles supplémentaires qui ne sont pas couvertes par les exigences, mais qui peuvent être utiles aux examinateurs pour effectuer leur évaluation. Par exemple, vous pourriez décrire :

- L'utilisation et l'impact des fonds de données du référentiel (citations, utilisation par d'autres projets, etc.).
- Un rôle national, régional ou mondial que le référentiel remplit.
- Tout cluster global ou organisation réseau auquel appartient le référentiel.

Guidage étendu R0.

Type de référentiel et brève description du référentiel

La sélection de plus d'un type de référentiel doit être étayée par une explication dans la *brève description du référentiel* de la manière dont ces multiples rôles sont remplis. Cette explication pourrait faire référence

les collections de données pertinentes, les types de données, les formats et les disciplines avec lesquels le référentiel travaille.

Brève description de la communauté désignée

Comme indiqué dans la définition (voir Glossaire), il est possible pour un référentiel d'avoir une Communauté Désignée composée de différentes « sous-communautés » ; par exemple, pour différentes collections. Si tel est le cas pour un candidat, il doit fournir une définition et une description suffisamment détaillée de chacune de ces sous-communautés. En outre, il est important de noter que la communauté désignée peut être plus petite que le groupe global d'utilisateurs d'un référentiel. Les collections numériques d'un musée d'histoire naturelle peuvent intéresser un large groupe d'utilisateurs intéressés, y compris le grand public.

Néanmoins, le musée peut définir sa communauté désignée comme étant plus restreinte que cela (par exemple, *des biologistes et des anthropologues effectuant des recherches sur des sujets du domaine de l'histoire naturelle*).

Pour bien servir sa communauté désignée, un référentiel doit avoir une compréhension approfondie de la composition, des compétences, de la base de connaissances et des besoins de la communauté désignée, et de la manière dont ceux-ci peuvent évoluer au fil du temps. Tout au long de la candidature, les preuves doivent démontrer une compréhension des exigences de conservation (contexte supplémentaire, formats préférés, etc.) pour mieux servir la communauté désignée (y compris les sous-communautés respectives, le cas échéant), ainsi qu'une démonstration que le candidat surveille et répond aux changements des besoins de la communauté désignée.

Un référentiel avec une communauté désignée très spécifique et étroite pourrait facilement indiquer la base de connaissances attendue (par exemple, le degré de compréhension de la génétique ou le niveau d'expertise dans l'utilisation de logiciels statistiques). En revanche, une large communauté désignée (c'est-à-dire composée de plusieurs communautés d'utilisateurs) signifie que le référentiel doit avoir une compréhension suffisante de toutes leurs bases de connaissances et offrir un large éventail de documentation contextuelle pour garantir que ses données peuvent être comprises par tout le monde au sein de la communauté désignée. Communauté. En ce qui concerne la définition de la base de connaissances de la communauté désignée, les candidats doivent indiquer explicitement toute hypothèse tacite, telle que les compétences en langues (étrangères), la capacité à accéder à des systèmes d'exploitation ou à des navigateurs Internet spécifiques, à utiliser certains logiciels, etc.

Niveau de conservation effectué

Plusieurs options (A, B, C ou D) du niveau (ou de l'étendue) de conservation peuvent être sélectionnées, en fonction du type de données et des conditions de conservation convenues avec le déposant. Lorsqu'un référentiel effectue une conservation à plusieurs niveaux, des informations supplémentaires doivent être ajoutées sur la proportion des données de la collection conservées aux niveaux respectifs. Les réponses aux exigences doivent ensuite indiquer comment les flux de travail de conservation reflètent les différents niveaux de conservation et comment les niveaux de conservation sont liés aux niveaux de préservation. Par exemple, les objectifs et les actions de préservation sont-ils les mêmes pour toutes les données, quel que soit le niveau de conservation.

En plus d'indiquer les niveaux de conservation, un référentiel doit donc démontrer qu'il assure l'accessibilité à long terme des données à mesure que les besoins de la communauté désignée évoluent. Cela est moins susceptible d'être possible aux niveaux de conservation A ou B, car sans normalisation des formats de fichiers soumis à un format de conservation commun, il peut être difficile d'effectuer des migrations de format à l'avenir en fonction de l'hétérogénéité de la collection. De même, le manque de métadonnées et de documentation riches peut poser un risque concernant l'utilisation continue des données.

Les examinateurs s'attendent à un niveau plus élevé de provenance formelle, d'intégrité et de gestion des versions (journaux des modifications, etc.) à mesure que les niveaux de curation progressent de A à D.

Partenaires internes/externalisés

Si plus d'un partenaire est impliqué, un diagramme indiquant toute l'échelle du processus d'internalisation/d'externalisation est utile pour aider les examinateurs. Avoir plusieurs partenaires d'internalisation/d'externalisation (par exemple, un pour le stockage, un pour la maintenance des sites Web) est acceptable tant que toutes les relations sont clairement indiquées. Les examinateurs demanderont que la réponse dans cette section soit révisée si les déclarations de preuve plus loin dans la demande font référence à des entités non mentionnées ici.

Autres informations pertinentes

Un référentiel peut faire référence ici à son enregistrement re3data (<http://www.re3data.org/>), nombre d'employés, taille de la collection, nombre moyen de téléchargements, son évolution dans le temps, modèle commercial ou de financement, etc. Une description claire et cohérente de l'approche organisationnelle dans son ensemble est généralement utile.

Infrastructure organisationnelle

1. Mission/portée R1.

Le référentiel a pour mission explicite de fournir un accès aux données de son domaine et de les conserver.

Niveau de conformité:

Réponse

Conseils:

Les référentiels assument la responsabilité de la gestion des objets numériques et veillent à ce que les matériaux soient conservés dans l'environnement approprié pendant des périodes de temps appropriées. Les déposants et les utilisateurs doivent être clairs sur le fait que la préservation et l'accès continu aux données est un rôle explicite du référentiel.

Pour cette exigence, veuillez décrire:

- La mission de votre organisation en matière de préservation et de fourniture d'accès aux données, et inclure des liens vers déclarations explicites de cette mission.
- Le niveau d'approbation que la mission a reçu au sein de l'organisation.

La preuve de cette exigence pourrait prendre la forme d'un énoncé de mission public approuvé, de rôles mandatés par les bailleurs de fonds, d'un énoncé de politique signé par le conseil d'administration.

Guidage étendu R1.

Si la préservation des données n'est pas mentionnée dans la mission du référentiel, cette exigence ne peut pas avoir un niveau de conformité de 3 ou plus.

2. Licences

R2. Le référentiel conserve toutes les licences applicables couvrant l'accès et l'utilisation des données et surveille la conformité.

Niveau de conformité:

Réponse

Conseils:

Les référentiels doivent disposer d'un modèle de droits approprié couvrant l'accès et l'utilisation des données, communiquer à leur sujet avec les utilisateurs et contrôler la conformité. Cette exigence concerne les réglementations d'accès et les licences applicables fixées par le référentiel de données lui-même, ainsi que tous les codes de conduite généralement acceptés dans le secteur concerné pour l'échange et l'utilisation appropriée des connaissances et des informations. Les preuves doivent démontrer que le référentiel a mis en place des contrôles suffisants selon les critères d'accès de ses fonds de données, ainsi que des preuves que toutes les licences ou tous les processus pertinents sont bien gérés.

Pour cette exigence, veuillez décrire:

- Contrats de licence en cours d'utilisation.
- Conditions d'utilisation (Droits de propriété intellectuelle, diffusion, utilisation prévue, protection des données sensibles, etc.).
- Documentation sur les mesures en cas de non-respect des conditions d'accès et utilisation.

Notez que si toutes les banques de données sont entièrement publiques et sans conditions imposées aux utilisateurs, telles que des exigences d'attribution ou un accord pour rendre l'analyse secondaire ouvertement disponible, alors cela peut simplement être indiqué.

Les dispositions en matière d'éthique et de confidentialité qui ont un impact sur les licences sont traitées dans R4 (Confidentialité/Éthique). L'assurance que les licences de dépôt accordent des droits suffisants au référentiel pour maintenir, préserver et offrir l'accès aux données doit être couverte par la R10 (Plan de préservation).

Guidage étendu R2.

Les stipulations relatives à l'accès aux données et à leur utilisation pourraient être définies dans un ensemble de conditions générales standard, ou différenciées par déposant ou ensemble de données. Pour les données sensibles, en particulier, les licences peuvent préciser les limites d'utilisation, l'environnement d'utilisation (salle sécurisée, accès distant sécurisé), et les types d'utilisateurs (chercheur agréé, ayant reçu une formation, etc.). Les options de licence populaires incluent, mais sans s'y limiter, celles proposées par Creative Commons (<https://creativecommons.org/>) telles que les licences «[CC 0 Waiver](#)» et «[données du domaine public](#)».

Bien qu'il puisse être difficile d'identifier les cas de non-conformité, il convient de tenir compte des conséquences si une non-conformité est détectée (par exemple, des sanctions sur l'accès/l'utilisation actuels ou futurs des données). Dans le cas de la divulgation de données personnelles sensibles, des sanctions légales sévères peuvent être appliquées à la fois à l'utilisateur et au référentiel. Idéalement, les référentiels devraient avoir une politique publique en place en cas de non-conformité.

Un niveau de conformité de 4 est nécessaire si le demandeur donne actuellement accès à des données personnelles.

3. Continuité d'accès R3.

Le dépôt dispose d'un plan de continuité pour assurer l'accès continu et la préservation de ses fonds.

Niveau de conformité:

Réponse

Conseils:

Cette exigence couvre la gouvernance liée à l'exploitation continue du dépôt dans le temps et pendant les catastrophes, ainsi que les preuves relatives à la planification de la relève; à savoir, les mesures en place pour assurer l'accès et la disponibilité des fonds de données, à la fois actuellement et à l'avenir.

Les examinateurs recherchent des preuves que des préparatifs sont en place pour faire face aux risques inhérents à l'évolution des circonstances, y compris dans la mission et/ou la portée.

Pour cette exigence, veuillez décrire:

- Le niveau de responsabilité assumé pour les banques de données, y compris toute garantie périodes de conservation.
- Les plans à moyen terme (trois à cinq ans) et à long terme (> cinq ans) en place pour assurer la disponibilité et l'accessibilité continues des données. En particulier, la réponse aux changements rapides de circonstances et la planification à long terme doivent être décrites, en indiquant les options de relocalisation ou de transition de l'activité vers un autre organisme ou le retour des fonds de données à leurs propriétaires (c'est-à-dire les producteurs de données). Par exemple, que se passera-t-il en cas d'arrêt du financement, qui pourrait être dû à un retrait inattendu du financement, à une fin planifiée du financement d'un référentiel de projets limité dans le temps ou à un changement d'intérêts de l'institution hôte?

Les preuves pour cette exigence doivent se rapporter spécifiquement à la gouvernance. Les aspects techniques de la continuité des activités et de la planification des sinistres et de la succession doivent être couverts dans R15 (Infrastructure technique).

Guidage étendu R3.

L'examineur recherche des informations pour comprendre le niveau de responsabilité assumé pour les données, le niveau de risque pour l'organisation actuelle et le niveau de planification de la relève pour l'avenir de la collecte de données. Par exemple, le demandeur est-il le gardien principal ou unique? Le déposant partage-t-il la responsabilité du devenir des données? Le référentiel fournit-il l'accès, la préservation et/ou le stockage des données à un niveau de qualité minimum pendant une période de temps minimum? Ces informations aident l'examineur à juger si le référentiel est durable en termes de finances et de processus; notamment la continuité de ses encaissements et de ses responsabilités en cas d'interruption temporaire ou définitive de service.

La responsabilité de la durabilité peut ne pas incomber au référentiel lui-même, mais à un hôte supérieur ou à une organisation mère. Si tel est le cas, cela doit être clairement indiqué. De plus, si le dépôt fait partie d'une organisation plus grande, cette organisation ou toute autre organisation (par exemple, une archive nationale) a-t-elle garanti qu'elle assumerait la responsabilité en cas de discontinuité de service? S'il n'y a pas d'accord écrit formel entre le référentiel et une telle organisation, alors le niveau de conformité peut être au maximum de 3 seulement.

4. Confidentialité/Ethique

R4. Le référentiel garantit, dans la mesure du possible, que les données sont créées, conservées, consultées et utilisées conformément aux normes disciplinaires et éthiques.

Niveau de conformité:

Réponse

Conseils:

Le respect des normes éthiques est essentiel à une science responsable. Le risque de divulgation - par exemple, le risque qu'un individu qui a participé à une enquête puisse être identifié ou que l'emplacement précis d'une espèce en voie de disparition puisse être identifié - est une préoccupation à laquelle de nombreux référentiels doivent répondre.

Les preuves doivent démontrer que le référentiel a de bonnes pratiques pour les données présentant des risques de divulgation, y compris des conseils pour les déposants et les utilisateurs. Cela est nécessaire pour maintenir la confiance de ceux qui acceptent que des données personnelles/sensibles soient stockées dans le référentiel.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Comment le référentiel respecte-t-il les normes disciplinaires applicables ?
- Le référentiel demande-t-il la confirmation que la collecte ou la création de données a été effectuée conformément aux critères juridiques et éthiques prévalant dans la zone géographique ou la discipline du producteur de données (par exemple, comité d'examen éthique/comité d'examen institutionnel ou législation sur la protection des données)?
- Des procédures spéciales sont-elles appliquées pour gérer les données présentant un risque de divulgation ?
- Les données présentant un risque de divulgation sont-elles gérées de manière appropriée pour limiter l'accès ?
- Les données présentant un risque de divulgation sont-elles diffusées dans des conditions appropriées ?
- Des procédures sont-elles en place pour examiner le risque de divulgation des données et prendre les mesures nécessaires pour anonymiser les fichiers ou pour fournir un accès sécurisé?
- Le personnel est-il formé à la gestion des données à risque de divulgation ? • Existe-t-il des mesures en place si les conditions ne sont pas respectées ? • Le référentiel fournit-il des conseils sur le dépôt, le téléchargement et l'utilisation responsables des données divulguées ou potentiellement divulguées?

Cette exigence concerne les dispositions éthiques et de confidentialité qui ont un impact sur la création, la conservation et l'utilisation des données. Les détails sur toutes les licences conformes à ces dispositions éthiques et de confidentialité doivent être couverts dans R2 (Licences).

Guidage étendu R4.

Toute organisation responsable de données a le devoir éthique de les gérer au niveau attendu par la pratique scientifique de sa Communauté Désignée. Pour les référentiels contenant des données sur des individus, des organisations ou des aires et espèces protégées, il existe des attentes juridiques et éthiques supplémentaires selon lesquelles les droits des personnes concernées seront protégés.

La divulgation de ces données pourrait également présenter un risque de préjudice personnel, une violation de la confidentialité commerciale ou la divulgation d'informations critiques (par exemple, la localisation d'espèces menacées ou d'un site archéologique). S'il existe un risque que des données identifiables soient déposées, par exemple par accident, le référentiel doit prendre les mesures appropriées pour traiter ces données et s'assurer qu'elles sont traitées (éliminées) conformément aux réglementations légales.

Le niveau de conformité doit être à 4 si le référentiel fournit actuellement un accès à des données personnelles ou à d'autres données sensibles.

Les preuves doivent démontrer que le demandeur comprend son environnement juridique et les pratiques éthiques pertinentes, et qu'il a mis en place des procédures documentées pour assurer la conformité.

5. Infrastructure organisationnelle

R5. Le référentiel dispose d'un financement adéquat et d'un nombre suffisant de personnel qualifié géré à travers un système de gouvernance clair pour mener à bien sa mission.

Niveau de conformité:

Réponse

Conseils:

Les dépôts ont besoin de financement pour s'acquitter de leurs responsabilités, ainsi que d'un personnel compétent ayant une expertise dans l'archivage des données. Cependant, il est également entendu que la continuité du financement est rarement garantie, et cela doit être mis en balance avec le besoin de stabilité.

Pour cette exigence, les réponses doivent inclure des preuves liées aux éléments suivants:

- Le référentiel est hébergé par une institution reconnue (garantissant une stabilité à long terme et durabilité) appropriée à sa communauté désignée.
- Le référentiel dispose d'un financement suffisant, y compris des ressources humaines, des ressources informatiques et un budget pour assister aux réunions si nécessaire. Idéalement, cela devrait être pour une période de trois à cinq ans.
- Le référentiel veille à ce que son personnel ait accès à une formation continue et à des développements.
- L'éventail et la profondeur de l'expertise de l'organisation et de son personnel, y compris affiliations pertinentes (par exemple, des organismes nationaux ou internationaux), est appropriée à la mission.

Des descriptions complètes des tâches effectuées par le référentiel et des compétences nécessaires pour les exécuter peuvent être fournies, si elles sont disponibles. Ces descriptions ne sont toutefois pas obligatoires, car ce niveau de détail dépasse le cadre de la certification de base.

L'accès à des conseils d'experts objectifs au-delà de ceux fournis par un personnel qualifié est couvert dans R6 (Conseils d'experts).

Guidage étendu R5.

La réponse à cette exigence doit contenir des preuves décrivant les processus décisionnels de gouvernance/gestion de l'organisation et les entités impliquées. Le personnel doit avoir une formation appropriée en gestion des données pour garantir des normes de qualité cohérentes. Il est également important de savoir quelle proportion du personnel est employée sur une base permanente ou temporaire et comment cela pourrait affecter la qualité professionnelle du dépôt, en particulier pour la conservation à long terme.

Dans quelle mesure le financement est-il structurel ou basé sur des projets ? Cela peut-il être exprimé en nombre d'ETP?

À quelle fréquence le renouvellement périodique du financement a-t-il lieu?

6. Conseils d'experts

R6. Le référentiel adopte un ou plusieurs mécanismes pour garantir des conseils et des retours d'experts continus (soit internes, soit externes, y compris des conseils scientifiques, le cas échéant).

Niveau de conformité:

Réponse

Conseils:

Un référentiel efficace s'efforce de s'adapter aux évolutions des types de données, des volumes de données et des débits de données, ainsi que d'adopter les nouvelles technologies les plus efficaces afin de rester précieux pour sa communauté désignée. Compte tenu du rythme rapide des changements, il est donc conseillé pour un référentiel de s'assurer régulièrement des conseils et retours d'utilisateurs experts afin d'assurer sa pertinence et son amélioration continue.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Le référentiel dispose-t-il de conseillers internes ou d'un comité consultatif externe pouvant être composé d'experts techniques, de curation, de science des données et disciplinaires?
- Comment le référentiel communique-t-il avec les experts pour obtenir des conseils ?
- Comment le référentiel communique-t-il avec sa communauté désignée pour obtenir des commentaires?

Cette exigence vise à confirmer que le référentiel a accès à des conseils d'experts objectifs au-delà de ceux fournis par le personnel qualifié mentionné dans E5 (Infrastructure organisationnelle).

Guidage étendu R6.

L'examineur recherche des preuves que le référentiel est lié à un réseau d'expertise plus large afin de démontrer l'accès à des conseils et à une orientation à la fois pour ses activités quotidiennes et le suivi des nouveaux défis potentiels à l'horizon (communauté et veille technologique). Si une partie de ces informations a déjà été fournie sous 'R0. Brève description de la communauté désignée du référentiel » et « Autres informations pertinentes », le demandeur doit y faire référence.

Gestion des objets numériques

7. Intégrité et authenticité des

données **R7. Le référentiel garantit l'intégrité et l'authenticité des données.**

Niveau de conformité:

Réponse

Conseils:

Le référentiel doit fournir des preuves pour montrer qu'il exploite un système de gestion des données et des métadonnées adapté pour garantir l'intégrité et l'authenticité pendant les processus d'ingestion, de stockage d'archives et d'accès aux données. Cette exigence couvre l'ensemble du cycle de vie des données dans le référentiel.

Pour protéger l'intégrité des données et des métadonnées, toute modification intentionnelle des données et des métadonnées doit être documentée, y compris la justification et l'origine de la modification. Des mesures doivent être en place pour garantir que les modifications non intentionnelles ou non autorisées peuvent être détectées et que des versions correctes des données et des métadonnées peuvent être récupérées.

L'authenticité couvre le degré de fiabilité des données originales déposées et leur provenance, y compris la relation entre les données originales et celles diffusées, et si oui ou non les relations existantes entre les ensembles de données et/ou les métadonnées sont maintenues.

Pour cette exigence, les réponses sur l'intégrité des données doivent inclure des preuves liées aux éléments suivants:

- Description des vérifications pour vérifier qu'un objet numérique n'a pas été altéré ou corrompu (c'est-à-dire, contrôles de fixité) du dépôt à l'utilisation.
- Documentation de l'exhaustivité des données et des métadonnées.
- Détails sur la manière dont toutes les modifications apportées aux données et aux métadonnées sont enregistrées.
- Description de la stratégie de contrôle de version.
- Utilisation des normes et conventions internationales appropriées (qui doivent être précisées).

La gestion des preuves de l'authenticité doit porter sur les questions suivantes:

- Le référentiel a-t-il une stratégie pour les changements de données ? Les producteurs de données sont-ils informés cette stratégie ?
- Le référentiel conserve-t-il les données de provenance et les pistes d'audit associées ?
- Le référentiel maintient-il des liens vers les métadonnées et vers d'autres ensembles de données? Si c'est le cas, comment?
- Le référentiel compare-t-il les propriétés essentielles des différentes versions du même fichier ?
Comment?
- Le référentiel vérifie-t-il l'identité des déposants ?

Guidage étendu R7.

L'examineur bénéficiera d'un aperçu clair des processus et des outils utilisés pour garantir que l'authenticité et l'intégrité des données sont protégées tout au long du cycle de vie de la conservation, y compris le niveau de pratique manuelle et automatisée, et de la manière dont ces processus, outils et pratiques sont documentés. Comme défini dans le Guide de l'exigence, le candidat peut trouver utile de répondre à chaque puce séparément et d'aborder l'intégrité et l'authenticité de manière indépendante (notez que la réponse doit être rédigée en prose complète).

Les pistes d'audit, qui sont des enregistrements écrits des actions effectuées sur les données, doivent être décrites dans les preuves fournies.

8. Évaluation

R8. Le référentiel accepte les données et les métadonnées en fonction de critères définis pour garantir la pertinence et la compréhensibilité des utilisateurs de données.

Niveau de conformité:

Réponse

Conseils:

La fonction d'évaluation est essentielle pour évaluer si les données répondent à tous les critères de sélection et pour assurer une gestion appropriée pour leur préservation. L'évaluation et la réévaluation au fil du temps garantissent que les données restent pertinentes et compréhensibles pour la communauté désignée.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Le référentiel utilise-t-il une politique de développement des collections pour guider la sélection des données à archiver ?
- Quelle approche est utilisée pour les données qui ne relèvent pas du profil de mission/collecte ?
- Le référentiel a-t-il mis en place des procédures pour déterminer que les métadonnées requises pour interpréter et utiliser les données sont fournies ? • Existe-t-il une évaluation automatisée de l'adhésion des métadonnées aux schémas pertinents ?
- Quelle est l'approche du référentiel si les métadonnées fournies sont insuffisantes pour préservation ?
- Le référentiel publie-t-il une liste de formats préférés ?
- Des contrôles sont-ils en place pour s'assurer que les producteurs de données respectent les formats préférés ?
- Quelle est l'approche vis-à-vis des données déposées dans des formats non préférés ?
- Quel est le processus de suppression des éléments de votre collection, en gardant également à l'esprit l'impact sur identifiants persistants existants ?

Cette exigence couvre les critères de sélection appliqués au point de dépôt. La qualité des données et l'amélioration au cours du processus de curation doivent être couvertes par R11 (Qualité des données).

Guidage étendu R8.

Le demandeur doit démontrer que des procédures sont en place pour s'assurer que seules les données appropriées à la politique de collecte sont acceptées. Le personnel du dépôt doit disposer de toutes les informations, procédures et connaissances spécialisées nécessaires pour assurer la conservation et l'utilisation à long terme applicables à la communauté désignée.

Pour que la collection reste pertinente et utilisable par la communauté désignée - en particulier à la lumière des changements technologiques, culturels ou législatifs (par exemple, la protection des données ou les droits de propriété intellectuelle) - les critères de sélection peuvent devoir être révisés au fil du temps et les actifs numériques réévalués en conséquence. Des politiques et des procédures documentées doivent être en place pour le retrait d'articles d'une collection.

9. Procédures de stockage documentées

R9. Le référentiel applique des processus et des procédures documentés dans la gestion de l'archivage des données.

Niveau de conformité:

Réponse

Conseils:

Les référentiels doivent stocker des données et des métadonnées depuis le point de dépôt, via le processus d'ingestion, jusqu'au point d'accès. Les dépôts qui effectuent la préservation numérique doivent offrir un «ystockage d'archives» selon les termes de l'OAIS.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Comment les processus et procédures pertinents sont-ils documentés et gérés ?
- Le référentiel a-t-il une compréhension claire de tous les emplacements de stockage et de la manière dont ils sont gérés ?
- Le référentiel a-t-il une stratégie pour les copies multiples ? Si oui, qu'est-ce que c'est ?
- Des techniques de gestion des risques sont-elles utilisées pour éclairer la stratégie ?
- Quels contrôles sont en place pour assurer la cohérence entre les copies d'archives ?
- Comment la détérioration des supports de stockage est-elle gérée et surveillée ?

Les détails sur la mise en œuvre technique du stockage doivent être couverts en R15 (Infrastructure technique), et les dispositions spécifiques pour la sécurité physique et logique en R16 (Sécurité).

Guidage étendu R9.

L'examineur cherche à comprendre chacun des emplacements de stockage qui prennent en charge les processus de curation, comment les données sont gérées de manière appropriée dans chaque environnement et que les processus sont en place pour surveiller et gérer les modifications apportées à la documentation de stockage. Les procédures sont-elles documentées et standardisées de telle manière que différents gestionnaires de données, tout en effectuant les mêmes tâches séparément, arriveront sensiblement au même résultat? Des exemples de preuves incluent des diagrammes de flux de données couvrant les lieux de dépôt, de conservation et d'accès (ainsi que toute restriction d'accès). Pour le stockage des archives, les preuves peuvent comprendre des descriptions des dispositions multisites (sur site, à proximité du site, hors site), la combinaison de supports de stockage et toute redondance (y compris l'intégrité par des sommes de contrôle).

10. Plan de préservation

R10. Le dépôt assume la responsabilité de la préservation à long terme et gère cette fonction de manière planifiée et documentée.

Niveau de conformité:

Réponse

Conseils:

Le référentiel, les déposants de données et la communauté désignée doivent comprendre le niveau de responsabilité assumé pour chaque élément déposé dans le référentiel. Le référentiel doit avoir le droit d'assumer ces responsabilités. Les procédures doivent être documentées et leur exécution assurée.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Le référentiel a-t-il une approche documentée de la préservation ?
- Le niveau de responsabilité pour la préservation de chaque élément est-il compris ? Comment est-ce défini ?
- Des plans liés aux migrations futures ou des mesures similaires pour faire face à la menace d'obsolescence sont-ils en place ?
- Le contrat entre le déposant et le dépositaire prévoit-il toutes les actions nécessaires pour assumer les responsabilités ?
- Le transfert de garde et le transfert de responsabilité sont-ils clairs pour le déposant et le dépositaire ?
- Le référentiel a-t-il le droit de copier, transformer et stocker les éléments, ainsi que d'y donner accès ?
- Les actions relatives à la préservation sont-elles spécifiées dans la documentation, y compris le transfert de garde, des normes d'information sur les soumissions et des normes d'information sur les archives ?
- Existe-t-il des mesures pour s'assurer que ces actions sont entreprises ?

Les droits concernant l'accès et l'utilisation des données, et le contrôle de leur conformité doivent être couverts par R2 (Licences).

Guidage étendu R10.

Le terme plan de préservation fait référence à une approche documentée pour définir et mettre en œuvre des actions de préservation. Les Exigences ne définissent ni ne font de différence entre une politique, un plan, une stratégie ou un plan d'action de préservation.

L'examineur recherchera une documentation claire et gérée pour garantir: (1) une approche organisée de la conservation à long terme, (2) un accès continu aux types de données malgré les changements de format, et (3) une documentation suffisante pour prendre en charge l'utilisabilité par la Communauté désignée. La réponse doit indiquer si le dépôt a défini des niveaux de préservation et, si oui, comment ceux-ci sont appliqués. Le plan de préservation doit être géré de manière à garantir que les changements apportés à la technologie des données et aux exigences des utilisateurs sont traités de manière stable et en temps opportun.

Si les niveaux de préservation diffèrent entre les classes ou les collections d'articles, le demandeur doit expliquer les différences dans l'approche de préservation, ainsi que les critères appliqués pour déterminer le niveau de préservation. Cela peut être pertinent si, par exemple, la taille du fichier d'un objet ou la sensibilité des données qu'il contient détermine le nombre de copies redondantes effectuées; ou, seuls les éléments déposés dans les formats préférés sont convertis en formats de conservation standard et seront migrés à l'avenir.

Si le demandeur n'est pas lié à une approche de préservation documentée, il ne peut être qu'au niveau de conformité 3 maximum et il doit en avoir une en place au moment du prochain examen.

11. Qualité des données

R11. Le référentiel possède l'expertise appropriée pour traiter la qualité des données techniques et des métadonnées et garantit que suffisamment d'informations sont disponibles pour que les utilisateurs finaux puissent effectuer des évaluations liées à la qualité.

Niveau de conformité:

Réponse

Conseils:

Les référentiels doivent s'assurer qu'il existe suffisamment d'informations sur les données pour que la communauté désignée puisse évaluer la qualité des données. L'évaluation de la qualité devient de plus en plus pertinente lorsque la communauté désignée est multidisciplinaire, où les utilisateurs peuvent ne pas avoir l'expérience personnelle nécessaire pour effectuer une évaluation de la qualité à partir des seules données. Les référentiels doivent être en mesure d'évaluer l'exhaustivité et la qualité des données et des métadonnées.

Les données, ou les métadonnées associées, peuvent présenter des problèmes de qualité liés à leur valeur de recherche, mais cela n'empêche pas leur utilisation si un utilisateur peut prendre une décision éclairée sur leur pertinence grâce à la documentation fournie.

Pour cette exigence, veuillez décrire:

- L'approche de la qualité des données et des métadonnées adoptée par le référentiel.
- Le référentiel dispose-t-il de contrôles de qualité pour garantir l'exhaustivité et l'intelligibilité des données déposées ? Si tel est le cas, veuillez fournir des références aux normes de contrôle de la qualité et aux mécanismes de signalement acceptés par la communauté de pratique concernée, et inclure des détails sur la manière dont les problèmes sont résolus (par exemple, les données sont-elles renvoyées au fournisseur de données pour rectification, corrigées par le référentiel, notées par des drapeaux de qualité dans le fichier de données, et/ou inclus dans les métadonnées qui l'accompagnent?).
- La capacité de la communauté désignée à commenter et/ou évaluer les données et les métadonnées.
- Si des citations vers des ouvrages connexes ou des liens vers des index de citations sont fournis.

Cette exigence fait référence aux normes de qualité des données et à l'assurance pendant la conservation. Les critères de sélection sont couverts dans R8 (Évaluation).

Guidage étendu R11.

Le demandeur doit indiquer clairement dans sa réponse qu'il comprend les niveaux de qualité que l'on peut raisonnablement attendre des déposants. Les preuves doivent décrire comment la qualité sera assurée pendant la conservation et les attentes de qualité de la communauté désignée. Le référentiel et ses déposants sont tenus de documenter tous les domaines dans lesquels la qualité des données ou des métadonnées est inférieure à la norme attendue.

12. Flux de travail

R12. L'archivage s'effectue selon des flux de travail définis, de l'ingestion à la diffusion.

Niveau de conformité:

Réponse

Conseils:

Pour assurer la cohérence des pratiques entre les ensembles de données et les services et pour éviter les actions ad hoc, les flux de travail doivent être définis en fonction des activités du référentiel et clairement documentés.

Des dispositions pour le changement géré doivent être en place. Le modèle de référence OAIS peut aider à spécifier les fonctions de workflow d'un référentiel.

Pour cette exigence, les réponses doivent inclure des preuves liées aux éléments suivants:

- Descriptions des workflows/processus métier.
- Communication claire aux déposants et aux utilisateurs sur le traitement des données.
- Niveaux de sécurité et impact sur les flux de travail (protection de la confidentialité des sujets, etc.).
- Contrôle qualitatif et quantitatif des sorties.
- Les types de données gérées et tout impact sur le workflow.
- Gestion des décisions au sein des workflows (par exemple, transformation des données d'archives).
- Changer la gestion des workflows.

Cette exigence confirme que tous les workflows sont documentés.

Guidage étendu R12.

L'examineur recherche des preuves que le candidat adopte une approche cohérente, rigoureuse et documentée pour gérer toutes les activités tout au long de ses processus et que les modifications apportées à ces processus sont correctement mises en œuvre, évaluées, enregistrées et administrées.

L'Exigence n'exige pas de descriptions détaillées des flux de travail, mais cherche à savoir comment et où ces flux de travail sont documentés.

13. Découverte et identification des données

R13. Le référentiel permet aux utilisateurs de découvrir les données et de s'y référer de manière persistante grâce à une citation appropriée.

Niveau de conformité:

Réponse

Conseils:

La découverte efficace des données est la clé du partage des données. Une fois découverts, les ensembles de données doivent pouvoir être référencés par des citations complètes, y compris des identifiants persistants pour garantir que les données seront accessibles à l'avenir.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Le référentiel offre-t-il des fonctionnalités de recherche ?
- Le référentiel maintient-il un catalogue de métadonnées consultable selon les normes appropriées (convenues au niveau international) ?
- Quels systèmes d'identifiants persistants le référentiel utilise-t-il ?
- Le référentiel facilite-t-il la collecte automatique des métadonnées ?
- Le référentiel est-il inclus dans un ou plusieurs registres disciplinaires ou génériques de ressources ?
- Le référentiel propose-t-il des citations de données recommandées ?

Guidage étendu R13.

La réponse doit contenir la preuve que toute conservation de données et de métadonnées soutient la découverte d'objets numériques clairement définis et identifiés, et permet leur lien avec des objets numériques connexes conformément aux normes du domaine. Il doit être clair pour la communauté désignée comment les données sont citées de manière à ce que le crédit et l'attribution appropriés soient attribués aux individus/organisations qui ont contribué à leur création.

14. Réutilisation des données

R14. Le référentiel permet la réutilisation des données au fil du temps, garantissant que les métadonnées appropriées sont disponibles pour soutenir la compréhension et l'utilisation des données.

Niveau de conformité:

Réponse

Conseils:

Les référentiels doivent s'assurer que les données continuent d'être comprises et utilisées efficacement à l'avenir malgré les changements technologiques et la base de connaissances de la communauté désignée. Cette exigence évalue les mesures prises pour s'assurer que les données sont réutilisables.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Quelles métadonnées sont fournies par le référentiel lors de l'accès aux données ?
- Comment le référentiel assure-t-il la compréhensibilité continue des données ?
- Les données sont-elles fournies dans les formats utilisés par la communauté désignée? Quels formats ?
- Des mesures sont-elles prises pour tenir compte de l'évolution possible des formats ?

Le concept de «réutilisation» est essentiel dans les environnements où les résultats d'analyse secondaires sont redéposés dans un référentiel aux côtés des données primaires, car la chaîne de provenance et les problèmes de droits associés peuvent alors devenir de plus en plus compliqués.

Guidage étendu R14.

Pour répondre à cette exigence, le candidat doit démontrer à la fois une connaissance approfondie des scénarios de réutilisation et des besoins de la communauté désignée en termes de pratiques, d'environnement technique et (de respect) des normes applicables. L'évolution de la technologie et des méthodologies et normes employées par la communauté désignée peut conduire à la nécessité de reconsidérer le format dans lequel les données sont diffusées. De même, des métadonnées appropriées et de qualité, conformes à un schéma généralisé et/ou disciplinaire, jouent un rôle essentiel et doivent être mentionnées dans les preuves fournies. Ces dernières informations sont essentielles pour concevoir des processus de conservation qui garantissent que les objets numériques restent au fil du temps compréhensibles et utilisables par la communauté désignée. En particulier, si seul un schéma de métadonnées généralisé (tel que Dublin Core ou DataCite) est utilisé, le demandeur doit fournir la preuve que cela est suffisant pour que la communauté désignée puisse continuer à comprendre le contenu préservé.

Technologie

15. Infrastructures techniques

R15. Le référentiel fonctionne sur des systèmes d'exploitation et d'autres logiciels d'infrastructure de base bien pris en charge et utilise des technologies matérielles et logicielles appropriées aux services qu'il fournit à sa communauté désignée.

Niveau de conformité:

Réponse

Conseils:

Les référentiels doivent fonctionner sur des infrastructures centrales fiables et stables qui maximisent la disponibilité des services. En outre, le matériel et les logiciels utilisés doivent être pertinents et adaptés à la communauté désignée et aux fonctions qu'un référentiel remplit. Le modèle de référence OAIIS précise les fonctions d'un référentiel pour répondre aux besoins des utilisateurs.

Pour cette exigence, les réponses doivent inclure des preuves liées aux questions suivantes:

- Quelles normes le référentiel utilise-t-il comme référence ? S'agit-il de normes internationales et/ou communautaires ? À quelle fréquence sont-ils révisés ?
- Comment les normes sont-elles mises en œuvre ? Y a-t-il des écarts importants par rapport à la norme ? Si oui, veuillez expliquer.
- Le référentiel dispose-t-il d'un plan de développement de l'infrastructure ? Si oui, qu'est-ce que c'est ?
- Un inventaire des logiciels est-il maintenu et une documentation système est-elle disponible ?
- Des logiciels pris en charge par la communauté sont-ils utilisés ? Décrivez s'il vous plaît.
- La disponibilité, la bande passante et la connectivité sont-elles suffisantes pour répondre aux besoins de la Communauté ?
- Le référentiel dispose-t-il d'un plan d'urgence et d'un plan de continuité d'activité ? En particulier, des procédures et des dispositions sont-elles en place pour assurer une récupération ou une sauvegarde rapide des services essentiels en cas de panne ? Que sont-ils ?

Les aspects de gouvernance de la continuité des activités, de la planification des catastrophes et de la planification de la relève doivent être couverts dans R3 (Continuité d'accès). Les détails sur le processus de stockage doivent être traités dans R9 (Procédures de stockage documentées). Les dispositions de sécurité sont couvertes par R16 (Sécurité).

Guidage étendu R15.

Les flux de travail et les acteurs humains fournissant des services de stockage doivent être soutenus par une infrastructure technologique adaptée qui répond aux besoins de la communauté désignée et permet au stockage de se remettre des catastrophes à court terme. L'examineur recherche des preuves que le candidat comprend l'écosystème plus large de normes, d'outils et de technologies disponibles pour la gestion et la conservation des données (de recherche), et a sélectionné des options qui correspondent aux exigences locales. Si possible, cela devrait être démontré en utilisant un modèle de référence.

Des exemples de normes pertinentes comprennent les normes d'infrastructure de données spatiales (SDI), l'Open Geospatial Consortium (OGC), le W3C ou les normes ISO.

Pour les flux de données en temps réel ou quasi réel, la fourniture d'une connectivité 24 heures sur 24 aux réseaux publics et privés à une bande passante suffisante pour répondre aux responsabilités mondiales et/ou régionales du référentiel ?

16. Sécurité

R16. L'infrastructure technique du référentiel assure la protection de l'installation et de ses données, produits, services et utilisateurs.

Niveau de conformité:

Réponse

Conseils:

Le référentiel doit analyser les menaces potentielles, évaluer les risques et créer un système de sécurité cohérent. Il doit décrire les scénarios de dommages basés sur des actions malveillantes, des erreurs humaines ou des défaillances techniques qui constituent une menace pour le référentiel et ses données, produits, services et utilisateurs. Il devrait mesurer la probabilité et l'impact de tels scénarios, décider des niveaux de risque acceptables et déterminer les mesures à prendre pour contrer les menaces pesant sur le dépôt et sa communauté désignée. Cela devrait être un processus continu.

Pour cette exigence, veuillez décrire:

- Votre système de sécurité informatique, les employés ayant des rôles liés à la sécurité (par exemple, les responsables de la sécurité) et tous les outils d'analyse des risques (par exemple, DRAMBORA2) que vous utilisez.
- Quels niveaux de sécurité sont requis et comment ceux-ci sont pris en charge.
- Toute procédure d'authentification et d'autorisation utilisée pour gérer en toute sécurité l'accès à systèmes en cours d'utilisation (par exemple, Shibboleth, OpenAthens).

Les processus de stockage et l'infrastructure technique qui utilisent ces mesures de sécurité doivent être couverts respectivement par R9 (Procédures de stockage documentées) et R15 (Infrastructure technique).

Guidage étendu R16.

L'examineur recherche des preuves que le demandeur comprend tous les risques techniques applicables au service fourni à la communauté désignée, ainsi qu'à l'environnement physique. Le demandeur doit démontrer qu'il a mis en place des mécanismes pour prévenir, détecter et répondre à un incident de sécurité.

De quelle manière la sécurité de l'infrastructure technique est-elle maîtrisée par le référentiel ou par son établissement hébergeur/externalisateur ? Qui est en charge?

Les procédures d'authentification et d'autorisation en place sont-elles suffisantes pour garantir la sécurité des fonds de données à chaque étape du flux de travail (par exemple, en exigeant une authentification à deux facteurs pour les données sensibles)?

Quelles politiques de sécurité de l'entreprise sont en place pour régir la sécurité de tous les systèmes, y compris la sécurité du réseau, les contrôles d'intrusion, la sécurité des installations physiques et la politique de mot de passe?

² <https://www.repositoryaudit.eu/>

Rétroaction du candidat

Commentaires/réactions

Ces exigences ne sont pas considérées comme définitives et nous apprécions votre contribution pour améliorer la procédure de certification CoreTrustSeal. Tout commentaire sur la qualité des exigences, leur pertinence pour votre organisation ou toute autre contribution sera pris en compte dans les itérations futures.

Réponse