

「研究データリポジトリ整備・運用ガイドライン（案）」の検討

リポジトリの信頼性

国立研究開発法人科学技術振興機構
バイオサイエンスデータベースセンター(NBDC)
八塚 茂



「コンテナ」としてのリポジトリの信頼性



リポジトリの**コンテナ**としての信頼性に焦点を当てる

- 安定した運営
- 高度なキュレーション（収集・分析・再構成）
- 持続可能性
- 明確なライセンス・ポリシー
- 堅牢なデータ管理
- データ利用の実績
- 高い専門性

出典：絹谷弘子,北本 朝展「データリポジトリの信頼性と要求要件に関する調査検討状況報告」
国際的動向を踏まえたオープンサイエンスの推進に関する検討会（第7回）資料1 平成30年10月26日(金)

CoreTrustSeal(CTS)

- 「信頼できるデータリポジトリを認定するための中核的な統一要件」を満たすリポジトリに対して付与される認証
- Data Seal of Approval (DSA) と国際科学会議 (ICSU) 世界科学データシステム (WDS) がRDAのワーキンググループで策定
- 各リポジトリに対して16項目の自己評価を実施
- 各項目は5段階(0~4)で評価
- 概ね3以上のレベルであれば認証される(らしい)

出典：南山泰之「信頼できるデータリポジトリの中核的な統一要件」
<http://current.ndl.go.jp/e1888>



CTSの要求要件

- 組織的な基盤 (Organizational Infrastructure)
 - R1. ミッション/スコープ (Mission/Scope)
 - R2. ライセンス (Licenses)
 - R3. アクセスの継続性 (Continuity of access)
 - R4. 機密保持/倫理 (Confidentiality/Ethics)
 - R5. 組織的な基盤 (Organizational infrastructure)
 - R6. 専門家によるガイダンス (Expert guidance)
- デジタルオブジェクト管理 (Digital Object Management)
 - R7. データの完全性・信頼性 (Data integrity and authenticity)
 - R8. 評価 (Appraisal)
 - R9. 保管手続きの文書化 (Documented storage procedures)
 - R10. 保存計画 (Preservation plan)
 - R11. データの品質 (Data quality)
 - R12. ワークフロー (Workflows)
 - R13. データの発見と識別 (Data discovery and identification)
 - R14. データの再利用 (Data reuse)
- 技術 (Technology)
 - R15. 技術的な基盤 (Technical infrastructure)
 - R16. セキュリティ (Security)

参照元 : https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf

TRUST原則

RDA 13th plenary (2019.4)で発表

ホワイトペーパー作成中

- T: Transparency
 - リポジトリが何を含んでいる（いない）をオープンなエビデンスで示すこと
- R: Responsibility
 - リポジトリは高品質なデータサービスを提供すること
- U: User community
 - 現在あるいは潜在的なリポジトリの利用に着目すること
- S: Sustainability
 - リポジトリはデータの長期間にわたる保管と利用を可能にすること
- T: Technology
 - リポジトリのオペレーションを可能にするインフラを整備すること

参照元 : <https://bit.ly/2Ih7g8F>