

京都大学の研究データ保存環境の これまで

京都大学

情報環境機構 / 学術情報メディアセンター

青木 学聡 (あおき たかあき)

<https://kyouindb.iimc.kyoto-u.ac.jp/j/zH8mD>

<https://orcid.org/0000-0002-5926-4903>

<https://researchmap.jp/takaakiaoki>

なぜ研究データ保存？ (政府等の対応)

- 2013-14年頃: 多くの研究不正事案が取りざたされる
- 2014.8: 文部科学省「研究活動における不正行為への対応等に関するガイドライン」
"研究者に対して一定期間研究データを保存し、必要な場合に開示することを義務付ける規程を整備し、その適切かつ実効的な運用を行うこと"
http://www.mext.go.jp/b_menu/houdou/26/08/1351568.htm
- 2015.3: 日本学術会議「科学研究における健全性の向上について」
"論文等の形で発表された研究成果のもととなった実験データ等の研究資料は、当該論文等の発表から10年間の保存を原則とする。試料や標本などの有体物については5年間の保存を原則とする。"
http://www.mext.go.jp/a_menu/jinzai/fusei/1355898.htm

なぜ研究データ保存？

(大学での対応)

京都大学での対応

http://www.kyoto-u.ac.jp/ja/research/ethic/research_guide

- 2014(H26)年度

- 8月: 文部科学省「研究活動における不正行為への対応等に関するガイドライン」
- 2月: 「京都大学における公正な研究活動の推進等に関する規程(※)」
 - 教職員, 部局, 全学の役割, 責任の確認. 研究データ一定保存義務
- 3月: 日本学術会議「科学研究における健全性の向上について」

- 2015年度(H27)年度

- 7月: 「(※)の研究データの保存、開示等について定める件」
年度内に管理方針, データ保存計画を, 部局毎に制定

- 2016年度(H28)年度より

- 部局ごとに決定した管理方針, データ保存計画を元に行動

部局の対応

- 京都大学工学研究科の場合(学内専用ページなので非公開)
 - 工学研究科における基本的なルール(2016.3)
 1. 公表成果の根拠資料として保存すべきデータの範囲と保存方法は、研究者・監督者が判断する。
 2. 監督者は、研究室の特性に応じた保存計画を作成し、研究科に提出する。
 3. 研究科は、必要最小限の研究データを所定の期間保存するための支援を行う。【工学研究科研究データ保存システム】
必要最小限のデータ…外部に発表した原著論文及び博士学位論文の根拠となる研究データ
 4. 退職、転出の際は、研究データの追跡を可能にしておく。
学生・研究者…監督者に研究データを引き継ぐ、または、自身で研究データを保存する場合は、監督者に新たな連絡先を報告する。
監督者…研究科のデータ保存システムに追加する、または、自身で研究データを保存する場合は、部局長に新たな連絡先を報告する。
 - 「データ保存計画ひな形」「研究室におけるデータ保全の心得」等のテンプレート、ガイドラインを提供
 - 組織としてのデータ保存の必要性に言及

研究室の対応

- 京都大学工学研究科 掛谷研究室
 - http://sk.kuee.kyoto-u.ac.jp/ja/internal/research_data/
- 論文になるまでのデータも含め、研究室のNASにデータを集約するように指導

組織的な研究データ保存サービスの構築

- 2014年度: 政府、全学レベルでの方針策定
- 2015年度: 部局レベルでの方針策定
 - 一部の部局、機構でプロトタイプを構築
 - 工学研究科附属情報センターが部局内向けサービス提供
 - 情報環境機構キャンパスICTラボが全学向けサービス提供
- 2016年度: 情報環境機構にてデータ保存のためのミドルウェア、ハードウェアを調達
- 2017年度: 新しいデータ保存システムの構築
- 2018年度: 情報環境機構のサービス切り替え

プロトタイプシステム例

- 情報環境機構, 工学研究科等で2015年度末より運用
 - 電子化されたデータを安心して保存できる仕組みが必要との判断
 - 既に運用, 開発経験のあるWebフレームワークや, CMSを改良しデータ保存システムを構築

ResearchData (beta) Takaaki Aoki Logout

Logged in. Last login: 2016/08/18 15:43:52

My ResearchData

Title	Year / Month	Filename	Filesize	
Molecular Dynamics Simulations Study of Nano Particle Migration by Cluster Impact	2016/02	SMMIB2015-submit-v1.zip	3.4 MB	Detail

Total filesize: 3.4 MB

[+ Add](#)

(情報環境機構のシステムの例)

ResearchData Upload

(*): required

Research Data File(*)
[ファイルを選択](#) 選択されていません

Title(*)

Authors

Type

Year / Month
 /

Bibliography

Digital Object Identifier (DOI)

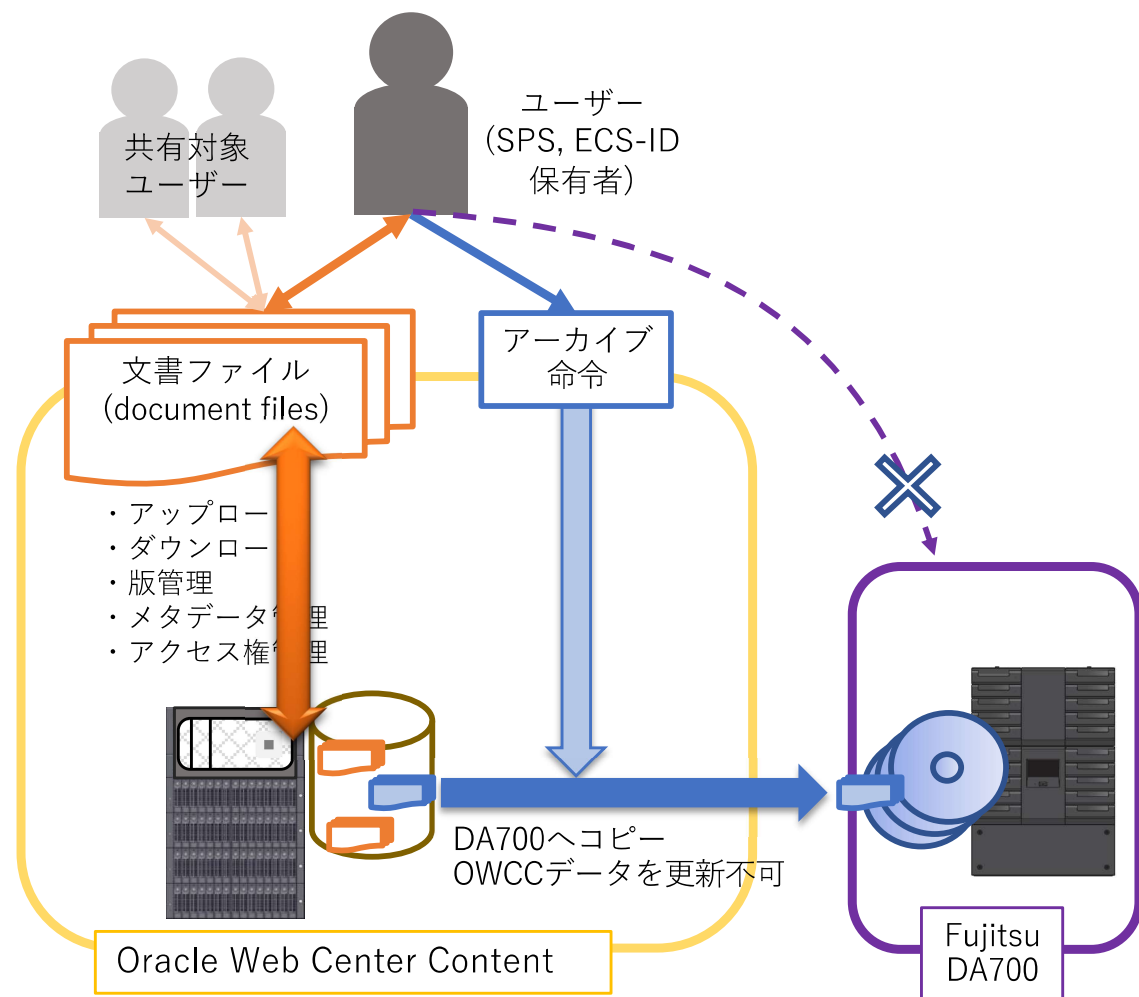
Memo

Share with: (mail addresses of KUMail or KUMOI)

[Save](#) [Cancel](#)

ダークアーカイブ型データ保存システム

- エンタープライズコンテンツ
マネジメントシステム(ECM)
と大容量光ディスクの組み合
わせ
 - 前者をライトアーカイブ
(ユーザーが随時利用できる)
 - 後者をダークアーカイブ
(ユーザーは直接利用できない)
として異なる保存ポリシーを適用



ライトアーカイブ

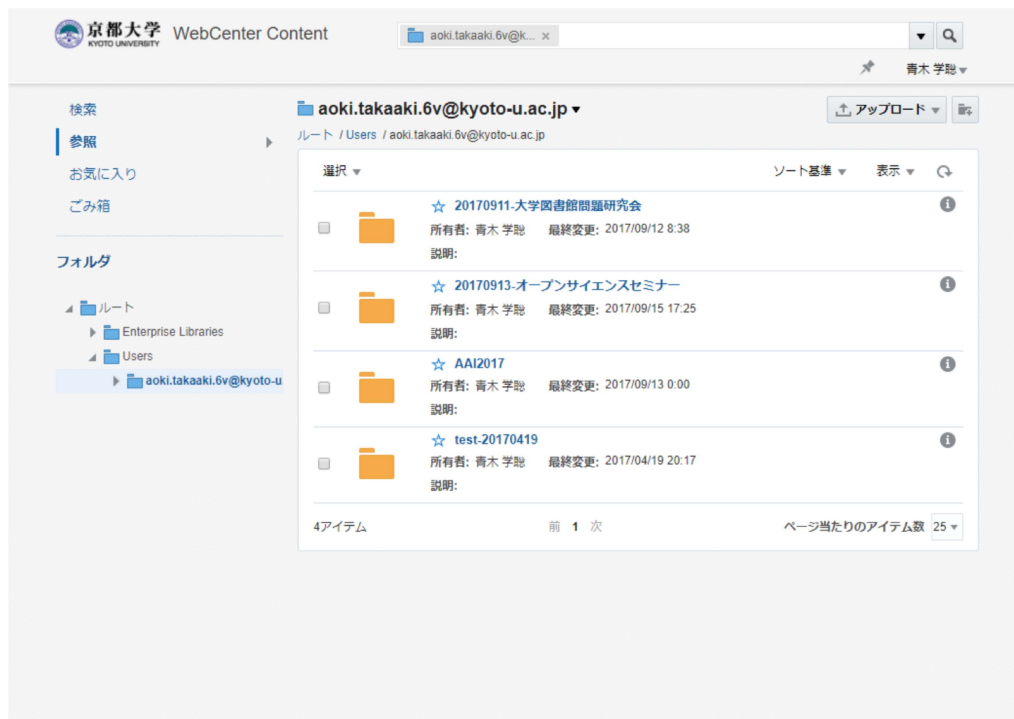
ダークアーカイブ

ECMでのライトアーカイブ(1)

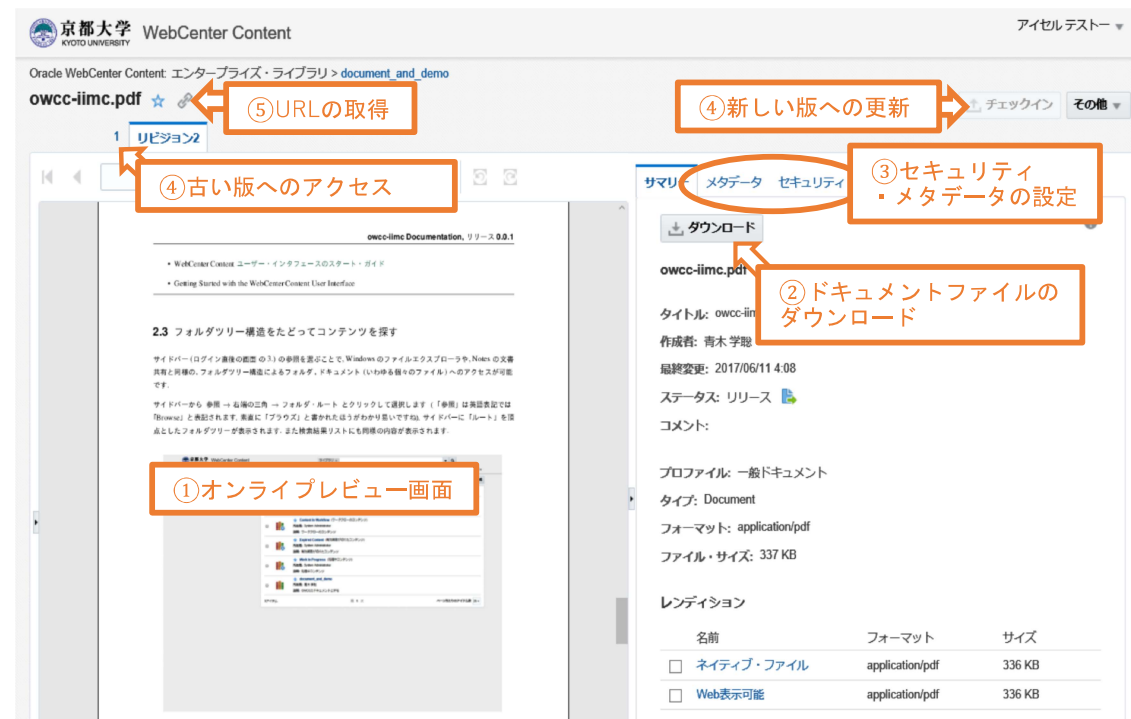
- WebUIによるファイルのアップロード、ダウンロード機能を提供
 - ツリー形式のファイル、フォルダのブラウズ
 - コンテンツプレビュー
 - メタデータ付与
 - リビジョン管理
 - ファイル共有
 - WebDAVでの一括up/download
 - (クラウド上でのデータ公開)
- 「オンプレ電子メールシステム」並みのデータ保護
 - システム二重化
 - 遠隔地バックアップ

ECMでのライトアーカイブ(2)

コンテンツのブラウジング



コンテンツのプレビュー



ダークアーカイブ(1)

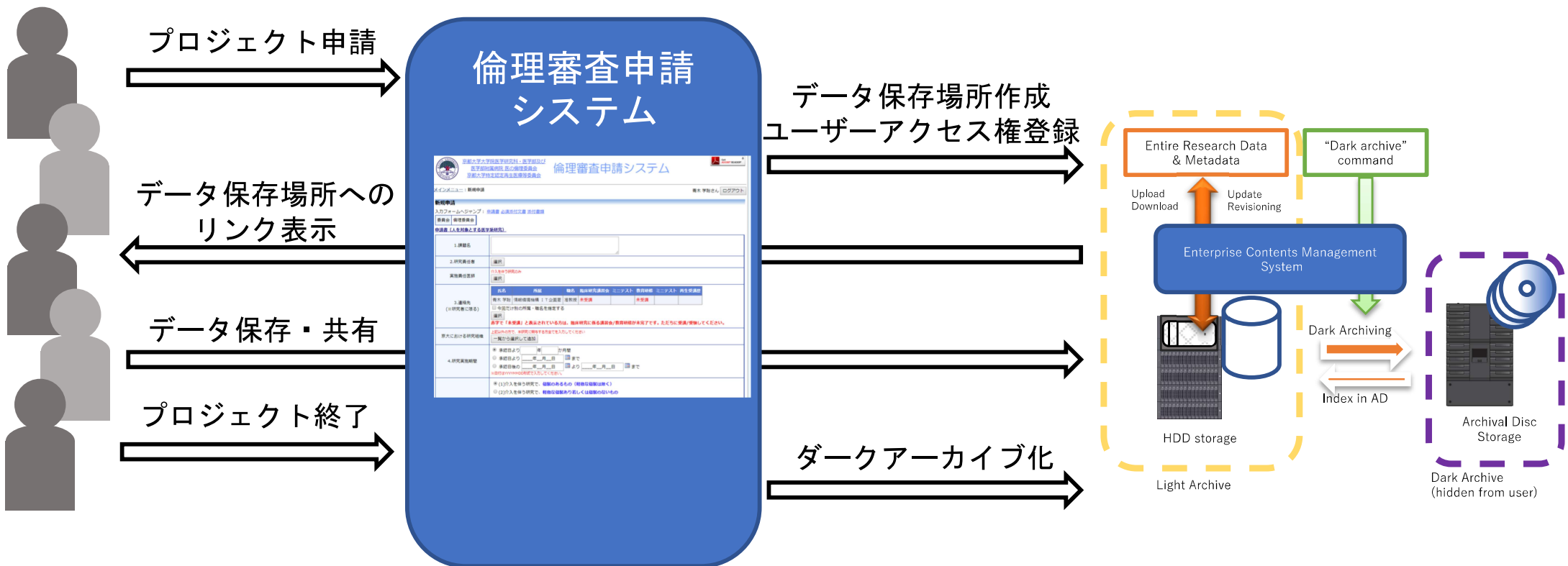
- ECMによるライトアーカイブ
 - 元来, ECMシステム並びにこれが使用するHDDそのものは長期間のデータ保存を保証しない. 冗長化, バックアップ等副次的対策が必要.
 - 要求される保存期間に比べ, ECMシステムの寿命が短い為, システムの更新, あるいはシステムの変更毎にデータ移行作業, そして移行内容の正当性の確認作業が必要.
- ダークアーカイブ側に求められる機能
 - 長期間保存が保証されたメディアの利用
 - 古典的, また汎用的なデータフォーマットでの保存

ダークアーカイブ(2)

- 利用者がフォルダ単位でダークアーカイブ命令をECMに発行
- ECMがバッチ処理として、フォルダの内容をダークアーカイブにコピー。ライトアーカイブ側にはダークアーカイブでの管理番号を保存
- コピー完了後、ライトアーカイブ側の対象ファイルを「更新不可」に、読み取りは引き続き可能
 - ライトアーカイブ・ダークアーカイブ間のコンテンツの一致を保証
 - 利用者が明示的な作業により、ライトアーカイブも許可。これにより、ダークアーカイブ側での世代管理も可能

部局の研究プロジェクト管理システムとの連携

京大病院、医学研究科の臨床研究データ保存(2019.1.22-)



京大医学研究科、病院の 倫理審査申請システム

京都大学大学院医学研究科・医学部及び
医学部附属病院 医の倫理委員会
京都大学持病生医学生区域委員会
京都大学臨床研究委員会

倫理審査申請システム

メインメニュー

お知らせ・新着情報

2015年09月09日 医の倫理委員会ホームページはこちら<http://www.e...> (続きあり) [表示](#)

2015年09月09日 特定部北再生医療等委員会ホームページはこちら<http://...> (続きあり) [表示](#)

2015年02月04日 <http://www.ec.med.kyoto-u.ac.jp...> (続きあり) [表示](#)

申請者用メニュー

新規申請 [➡](#)

担当するプロジェクト一覧 [➡](#)

利用ガイド [➡](#)

プロジェクト一覧

☐ 終了済みも表示

受付番号	課題名	研究期間	審査状況	表示
C9999	〇〇〇と△△△に関する研究 その1	2018/1/1~2025/12/31	表示	表示 ④

共通メニュー

マイページ [➡](#)

メールログ [➡](#)

さん ログアウト

プロジェクトの申請者

削除 ①

追加

https://archive.iimc.kyoto-u.ac.jp/wcc/faces/wccmain2?_af562A44E18F2F55B64561AF798209A54 ⑤

OWCC

研究責任者、連絡先

研究責任者、連絡先

研究責任者、連絡先

データ・マネージャ (更新・削除) ※②

データ・マネージャ (更新)

その他

プロジェクトの研究組織	OWCC 権限名	OWCC 操作
研究責任者 実施責任医師 連絡先 連絡窓口	研究責任者、 連絡先	- 新規フォルダの作成 - ファイルのアップロード、移動、閲覧 - フォルダ/ファイルの削除 (trashbox フォルダへの移動) - data フォルダ内での上記 3 項目の作業
データ・マネジメント (更新・削除)	データ・マネー ジャ (更新・削 除)	- 新規フォルダの作成 - ファイルのアップロード、移動、閲覧 - フォルダ/ファイルの削除 (trashbox フォルダへの移動)
データ・マネジメント (更新)	データ・マネー ジャ (更新)	ファイルのアップロード、移動、閲覧
統計 モニタリング CRC その他	その他	ファイルの閲覧

利用状況と課題

- まだユーザー数、コンテンツ数ともに少ない
- 運用側、利用側ともにノウハウが少なく、認知度が低い
 - 運用側: 本当に「学内すべての」データを預かれるのか?
 - 利用側: 本当にデータを預けて安心できるのか?
- 安定的な運用モデルの模索
 - 現状は無償での提供
 - いずれは安定的な財源の確保や、従量課金モデルを導入する必要がある
- 研究者がすでに採用・利用している手法からの漸進的な移行プロセスの提示

利用状況と課題

- ダークアーカイブへのアクセス方針
 - 「当人または、所属している部局の部局長からの申し出により開示する。」という内部ルールを作っているが、ユーザーへの利用条件の明示や実際の承認プロセスは未整備
- データの引継ぎ
 - ライトアーカイブ、ダークアーカイブともに4年間レンタルシステムなので、長期間の安定したサービスとして定着させられない

「研究公正とコンプライアンス」 のための研究データ管理の視点

- データの長期保存ルール(policy)が先行, 現実的な実施手順(procedure)の策定と実施(deployment)が後追い
- 長期保存に対する客観的な質保証は研究者個人の能力をはるかに超える。組織的な対応が後回し
- 残念ながら「ほとんどが死蔵されるデータ」
 - 研究者個人, 限られた研究グループ内で保持するバックアップ最終手段
 - 最悪の場合(研究公正にかかわるインシデント)への備え、というリスク管理の問題

研究データ保存の「コスト」から「ベネフィット」へ

- 心理的、金銭的障壁をできる限り抑え、適切なアクセス制御の下、データ共有を低コストで実現する
 - 研究パフォーマンスの向上
 - 研究データに新たな価値を付加する
- どのようなデータが生成され、これがどのように共有、保存されるか？
 - フレームワークでの管理モデルの提示
 - モダンなRDMのイニシエーション
- 研究支援、知財管理、オープンサイエンス対応等、大学の学術コンテンツ戦略の根幹と位置付けられるか？