

研究データの早期のオープン化に 資する「逐次公開」型運用モデルと 運用支援環境の検討

金沢大学総合メディア基盤センター

高田良宏

目次

- 自己紹介
- 地域学術資料／研究室データ
- 現状／考え方の転換
- 「逐次公開」型運用モデルの検討
 - 発生源入力とオープン化の段階
 - 整理途中における複数の時点における公開
- 運用支援環境検討
 - 求められる公開基盤(リポジトリ)
 - 運用支援環境の構築事例
- まとめ

自己紹介

• 所属



- 金沢大学総合メディア基盤センター
 - 大学情報システム開発・大学のデータの利活用
 - 学術資源リポジトリ(データリポジトリ)

- 学術資源リポジトリ協議会
 - 学術資源リポジトリの確立と普及



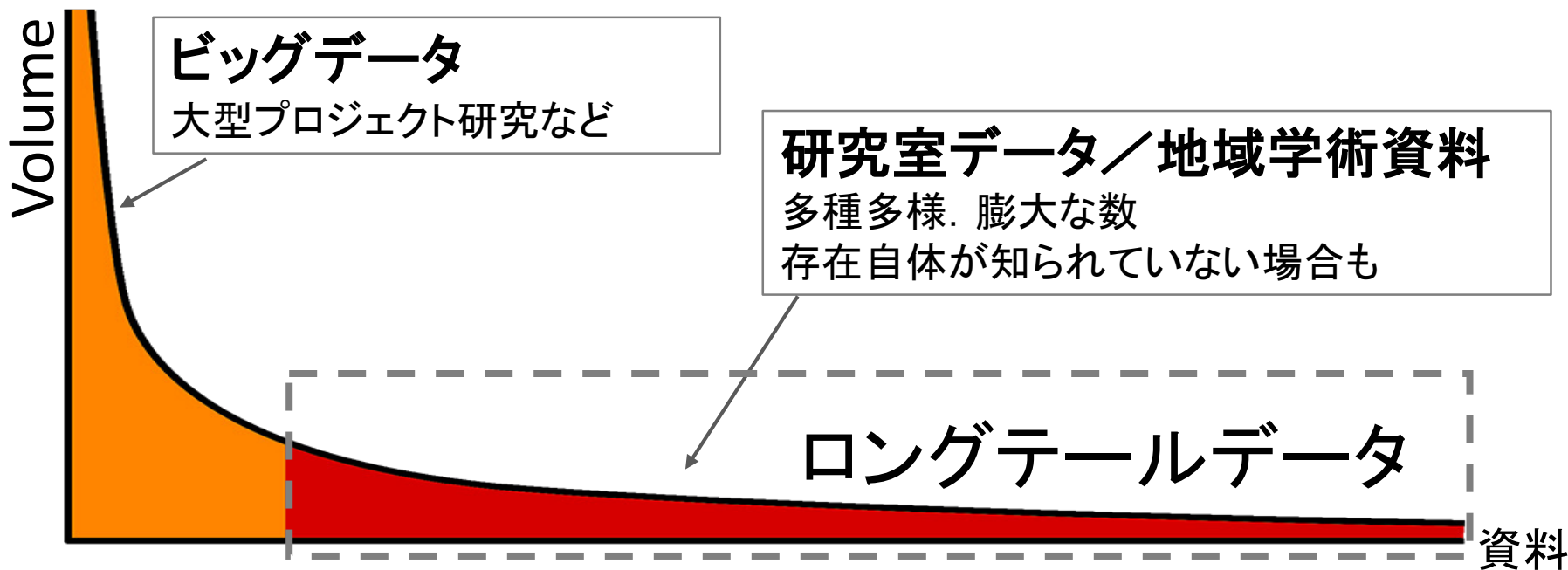
• 研究データリポジトリとの関わり

- 2006年にインドの宗教美術(仏像, 遺跡等々)の画像の公開依頼される.
- 論文などの文献資料に対して非文献資料と称して, DSpaceを使って非文献資料のためのリポジトリ構築

研究室データ／地域学術資料

研究室データと地域学術資料、おかれている環境は非常に似ている(以降**研究データ**)

研究室データでの早期公開の試みが地域学術資料にも適用できるのでは？



現状

- 公開・活用される研究データは一握りで、多くの研究データは死蔵されており、さらには喪失の危機に瀕している
- 公開・活用は(重要)課題であるという認識は共通
- 進まない理由
 - 資金・人材不足
 - 公開しにくい(公開がふさわしくない)データの混在
 - 研究データ公開に対する考え方(社会の要請との不一致)



- 国や社会は、オープンサイエンスに代表されるように広い範囲での早期の公開を求めている

考え方の転換

- 研究データを分類・整理・解釈し、きちんとしたメタデータを付して研究データリポジトリ(DB)に登録、公開しようしている

- 永遠に登録・公開されない可能性大(資金・人材不足)
- エビデンスデータならばそこまでする必要はあるが、ほとんどのデータは手を抜いてもよいのではないか... ?
- 社会は、分類・整理・解釈をそこまで望んでいないのではないか... ?



- 早期の段階で公開し、その後の状況により逐次充実させていけばよい。状況: 己の興味, 資金・人材, 社会の要請, etc.

「逐次公開」型運用モデルの検討

「逐次公開」型運用モデル 発生源入力とオープン化の段階

時間・情報量

原則すべてのデータがオープン化候補

従来のデータリポジトリ

最小限の
メタデータ

追加の
メタデータ

付加価値
メタデータ

研究
データ

デジタル化
された

推奨メタ
データが
揃った

メタデータが
充実した

所在が明らかになった

所在がわかる

オープンデータの
候補となる

ニーズがわかる

データがわかる

再利用が可能となる

信頼性が向上

再利用が進む

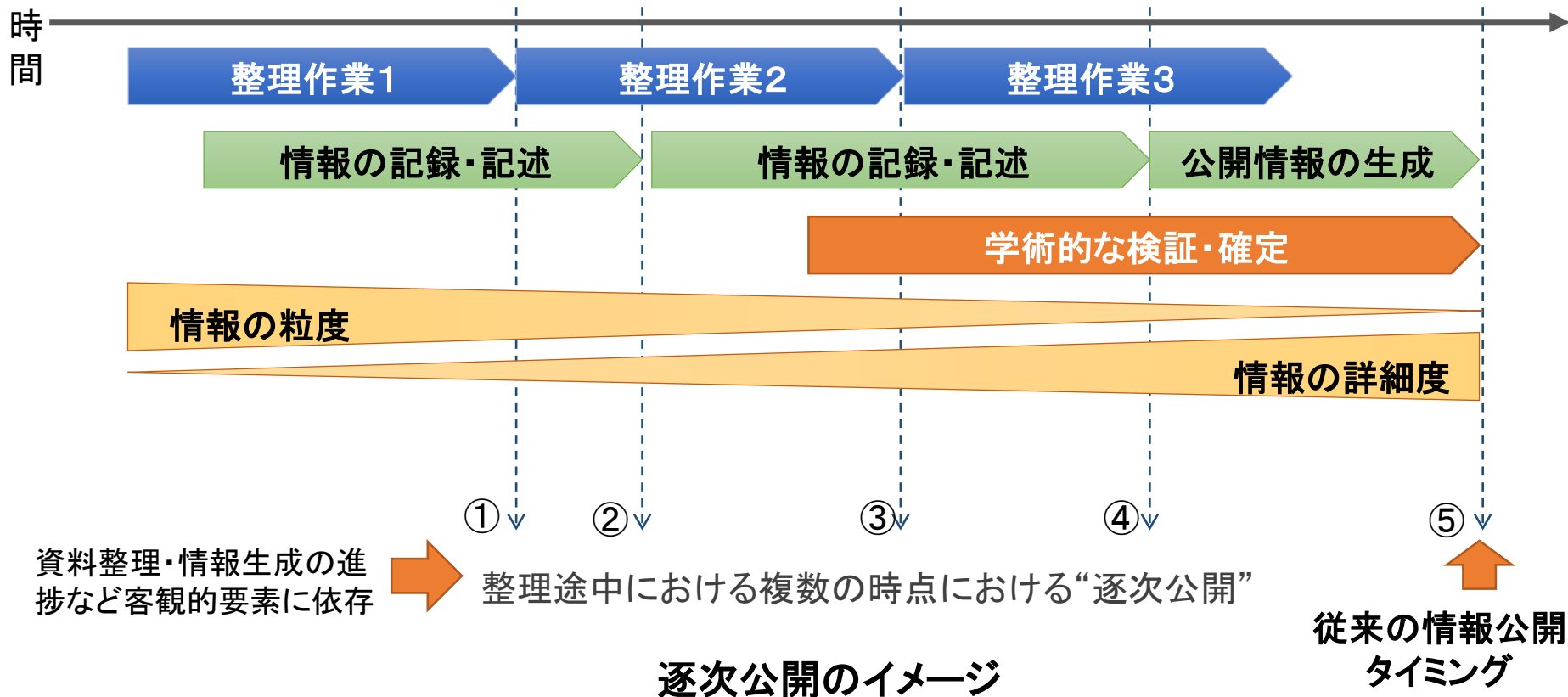
「逐次公開」型運用モデル 発生源入力とオープン化の段階

- オープン化の段階を従来より十分広く解釈する
 - ①所在が明らかになった段階(オープン化候補)
 - ②データ本体と最小限のメタデータが登録された段階・再利用が可能な段階(とりあえず利用できる)
 - ③再利用が進む段階(解釈が進みメタデータ充実)
 - ④完全な段階(従来はこの段階で公開していた)



- 問題: 誰が段階を切り分けるのか? その段階に達したかの基準・判断は? それをどのように表すのか? その段階の切り方を利用者が望んでいるか?

「逐次公開」型運用モデル 整理途中における複数の時点における公開



準備ができた任意の時点で公開(逐次公開)する。その後、作業の進捗に合わせて任意の時点で情報の追加・修正(逐次修正)繰り返す

運用支援環境の検討

運用支援環境の検討 求められる公開基盤(リポジトリ)

• データ所有者の声

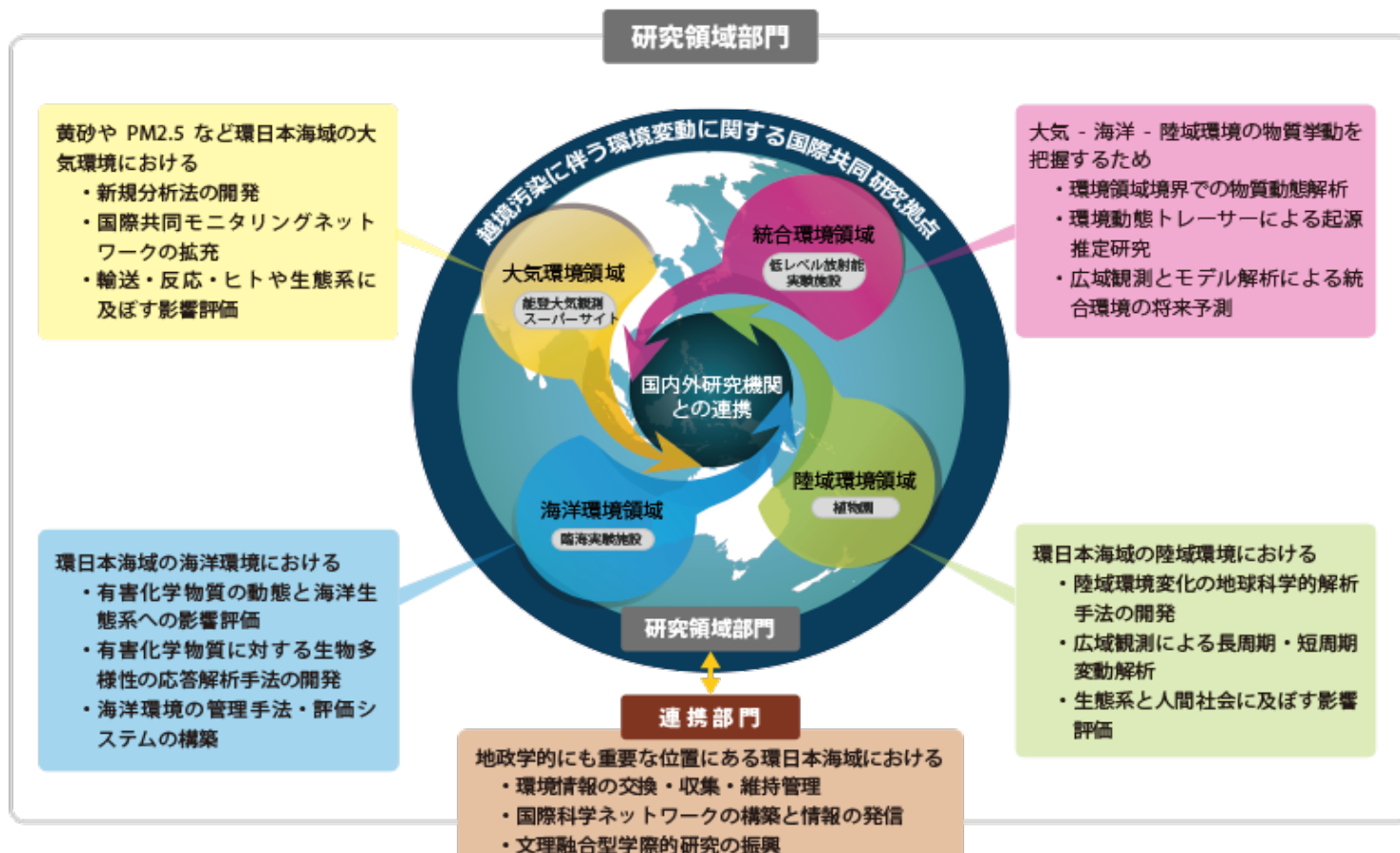
- 整理途中でも随時公開したい(逐次公開)
⇒整理の完了はいつになるか見当もつかない
- 整理が進めばメタデータは随時更新したい
⇒簡単な手順で行いたい
- データ自身も更新されるので、バージョン管理が必要
⇒写真の撮り直し、データの校正法の変更等々
- 研究用DBと公開用のDBをできれば同じにしたい
⇒2種類のデータは管理する余裕はない
⇒研究用としてはすべて登録したい、公開用として問題のあるものは非公開にしたい

求められる公開基盤(リポジトリ) 運用支援環境の仕様(概要)

- 方針: ICTおよびリポジトリに精通しているとは限らないデータ所有者がリポジトリ上のデータ管理を行えるような機能⇒引継ぎを行える逐次公開を支援するデータ管理機能
- 仕様(概要)
 - ✓ OSに非依存(Webアプリ)
 - ✓ データを逐次一括で登録可能(逐次公開)
 - ✓ Excelなどの一般的なソフトでデータの整理が可能
 - ✓ 一度登録したデータに対して、逐次更新が可能
 - ✓ 実データファイル更新時、過去のバージョンを履歴として残す
- いくつかの研究センターで実証

求められる公開基盤(リポジトリ) 運用支援環境の構築事例

- 金沢大学環日本海域環境研究センター 公開中
- 多種多様なデータを所有: 観測データ, 標本, 画像, 文書



求められる公開基盤(リポジトリ) 運用支援環境の構築事例

金沢大学国際文化資源学研究中心 — 公開準備中



特徴

- ・非常に対象が広い
 - ・分野・領域, 地域, 年代
- ・資料の状態が多様
 - ・画像, 動画, 音声, 文書
- ・保存されている媒体多用
- ・権利等の問題
 - ・著作権, 肖像権, 個人情報, その他



形態文化資源部門
有形文化資源に関する調査・研究



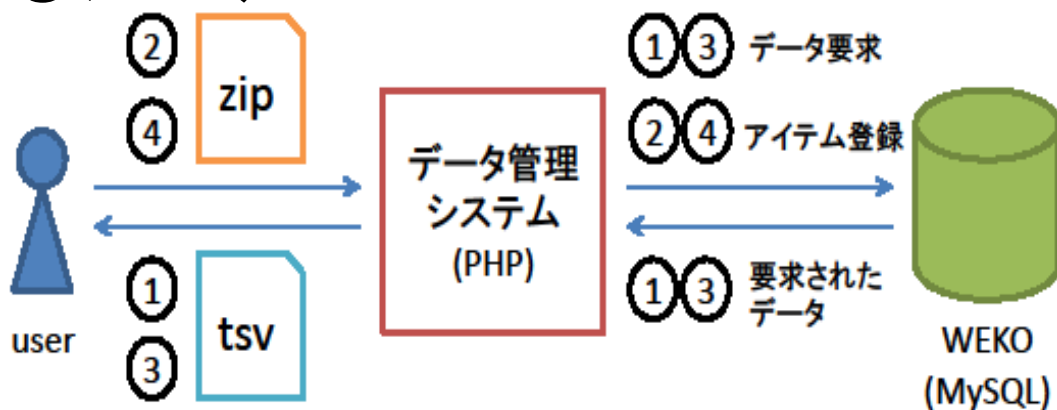
伝承文化資源部門
無形文化資源に関する調査・研究



文化資源情報部門
文化資源情報の収集・社会還元

データ管理システム概要

- ブラウザベースで逐次公開／逐次更新の管理ができるシステム



- ①メタデータテンプレート作成機能
- ②アイテム登録機能
- ③登録済みメタデータエクスポート機能
- ④アイテム更新機能

登録用機能

更新用機能

運用支援環境の構築事例

登録手順

• ユーザーが登録する際の手順

1. テンプレート作成



①メタデータテンプレート
作成機能

2. メタデータファイル作成

3. 作成したメタデータファイルと実データファイルをフォルダにまとめ、ZIPファイル化

- 実データファイル(生データや解析後のデータ、写真、論文など)
- メタデータファイル(実データに関する情報)

4. 登録先インデックスを選択し、アップロード

②アイテム登録機能

運用支援環境の構築事例

更新手順

- ユーザーの更新時の手順

1. 登録済データをエクスポート

2. メタデータファイルを編集

3. 追加・更新する実データファイルとメタデータファイルをフォルダにまとめてZIPファイル化

4. アップロード

④アイテム更新機能



③登録済みメタデータ
エクスポート機能


運用支援環境の構築事例

更新フラグ

m	メタデータのみ
rh	実データファイルを更新し、履歴を 公開 する
ph	実データファイルを更新し、履歴を 非公開 にする
replace	実データファイルを更新し、履歴を 残さない *公開前限定


m|rhのように記述するとメタデータと実データ
ファイルを同時更新可能

運用支援環境の構築事例 更新結果（ハリサンショウウニ）

File / Name
P1232641.JPG
 P1232641.JPG (662.94KB) [0 downloads]
更新履歴



タイトル(英)	00001
アイテムタイプ	海洋生物標本
言語	日本語
標本No.	00001
棚No.	A
段No.	1
ラベルの有無	あり
容器の種類	シャーレ
保存状態	乾燥
和名	ハリサンショウウニ
学名	Temnopleurus reevesii (Gray, 1855)
動物門	棘皮動物
サムネイル	P1232641-t.JPG 

File / Name
更新後
 更新後 (13.09KB) [0 downloads]
更新履歴



タイトル(英)	00001
アイテムタイプ	海洋生物標本
言語	日本語
標本No.	11111
棚No.	B
段No.	1
ラベルの有無	あり
容器の種類	シャーレ
保存状態	乾燥
和名	ハリサンショウウニ
学名	Temnopleurus reevesii (Gray, 1855)
動物門	棘皮動物
サムネイル	P1232641-t.JPG 

運用支援環境の構築事例

更新結果：バージョン管理

WEKO

バージョン	更新日時	ファイル名	ユーザ名	管理
Current	2019-01-30 16:38:06	update.docx	miyamoto	
1	2018-05-09 16:18:12	P1232641.JPG	miyamoto	<input checked="" type="radio"/> 公開 <input type="radio"/> 非公開

環日本海域環境研究センター データベース

メニュー

[トップページ](#)

カウンタ

COUNTER 000000

お知らせ

Welcome to KINET DataBase!

当サイトは金沢大学環日本海域環境研究センターの各研究領域（大気環境、海洋環境、陸域環境、統合環境）で随時収集、あるいは保管されている夕、環境試料、標本、出版物などの学術成果を広く共有、公開することを目的として運用されています。公開試料、データの取り扱いにあたってはとに定めるTerms of useに従ってご利用くださいますようお願いいたします。

データ利用規約

1. 当施設が測定したデータは、教育目的もしくは研究目的でのみ利用して下さい。
2. 測定データを利用した学会発表や論文発表等では、金沢大学環日本海域環境研究センターの測定データを利用した旨を発表内で明記して下さい。
3. 学会発表や論文発表等で測定データを利用される際は、当センターまでその旨ご連絡下さい。
4. 測定データの改ざんや公序良俗に反する使用は認めません。
5. 測定データを転載したり、直接リンクを張りつけたりしないで下さい。リンクを張る場合はこのページを張りつけて下さい。
6. データを利用したことにより損害を生じても当センターは一切の責任を負いません。

WEKO

[トップ](#) [ランキング](#)

検索

[▶ 詳細検索](#) 全文検索 キーワード検索 ?

Language

日本語

インデックスツリー ?

- 表面錯体モデリングユーザーイン
- 大気環境領域
- 海洋環境領域
- 陸域環境領域
- 統合環境領域
- Person

インデックスリスト

- 表面錯体モデリングユーザーインターフェース [3件]
- 大気環境領域 [61件]
- 海洋環境領域 [2066件]
- 陸域環境領域 [3件]
- 統合環境領域 [7件]

まとめ

- 逐次公開型の運用モデルを提案した
 - データの所有者にも一定の理解を得られている
- 逐次公開(逐次登録・修正)を支援するシステムを構築した
 - ICTに精通していない場合でも、データ管理が行えるシステムであり、自立して管理を行うことができるようになった
 - 地域学術資料に対しても適用可能ではないか
- 課題
 - 今後、事例を増やして検証をすすめる
 - 所有者や作業者のメリット・デメリット考える
 - 公開インセンティブやリスクについても考える必要あり

ご清聴ありがとうございました

科研費: 18K18525, 挑戦的研究(萌芽)
歴史的資料喪失等を考慮した発生源入力によるオープン化モデルと学術資源基盤の確立