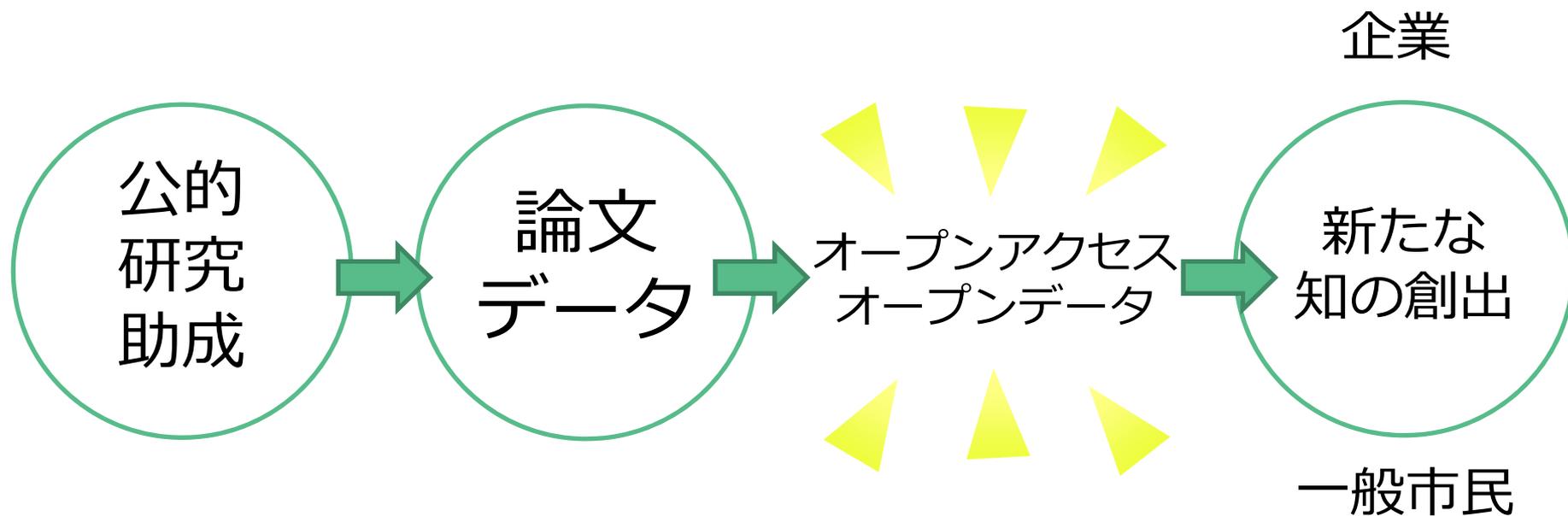


消えゆくデータを供養する

人文系研究データの ケーススタディ

京都大学学術研究支援室
天野絵里子

オープンサイエンス政策



過去の研究データはどうなったのだろうか？
人文系の研究データはどうなったのだろうか？

なくなる前に機関リポジトリに保存しよう

データベースの探索先

国立国会図書館

Dnavi

22,196 件

公開 17,470 件

公開停止 4,726 件

データベース科研

データベース科研とは？

- 科学研究費助成事業・研究成果公開促進費
「データベース」
- 「研究成果の公開発表、重要な学術研究の成果の発信及び、データベースの作成・公開について助成することによって、我が国の学術の振興と普及に資するとともに、学術の国際交流に寄与することを目的とするものであり、優れた研究成果の公的流通の促進を図るものです。」
- データ作成について助成
- 平成18年～ 公開されたデータベースの一覧あり
http://www.jsps.go.jp/j-grantsinaid/13_seika/#data

人文系の研究データとは

- ○○新聞記事索引
- ○○文書目録
- ○○文庫貴重資料画像
- テレビCMデータベース
- ○○史文献データベース
- ○○地図デジタルアーカイブ

過去に採択されたDBの現状

データ
ベース

URLを
クリックすると

現状

A

異なるサイトを
表示

別のサイトへ
移動



OK

B

DBを表示

2010年の更新で
止まっている



管理に不安

C

Not found

Not found



行方不明

生きているデータベースの条件

- 運営組織が現在もしっかりしている
- 更新日が最近
- ウェブで検索可能、閲覧可能
- オープンである
- 相応の理由がないかぎり誰でもアクセスできる
- 無料

これらにあてはまらないデータベースを
機関リポジトリに登録したい

データベース科研 調査

対象：

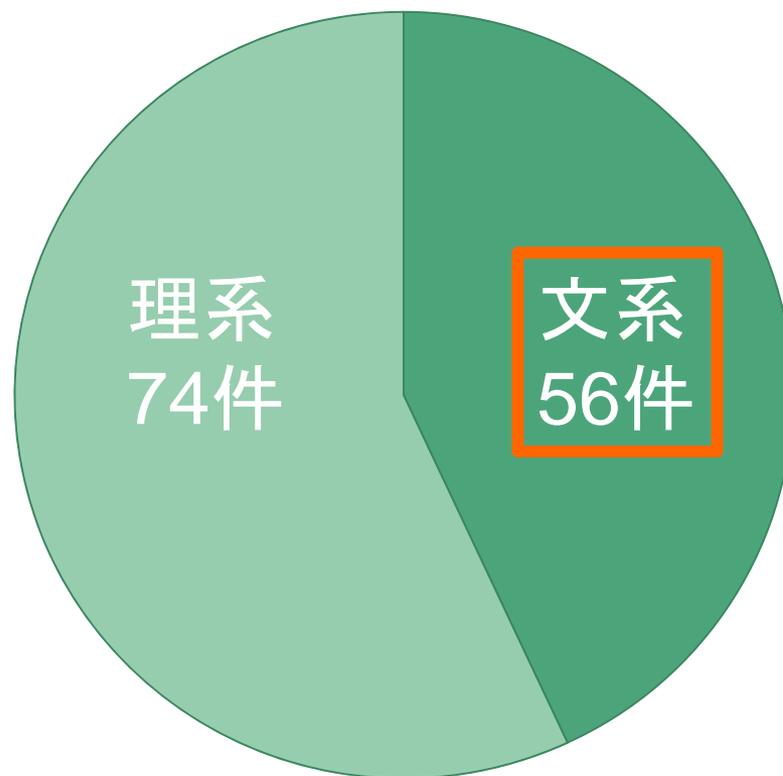
平成21～23年度採択

252件（約80件/年）

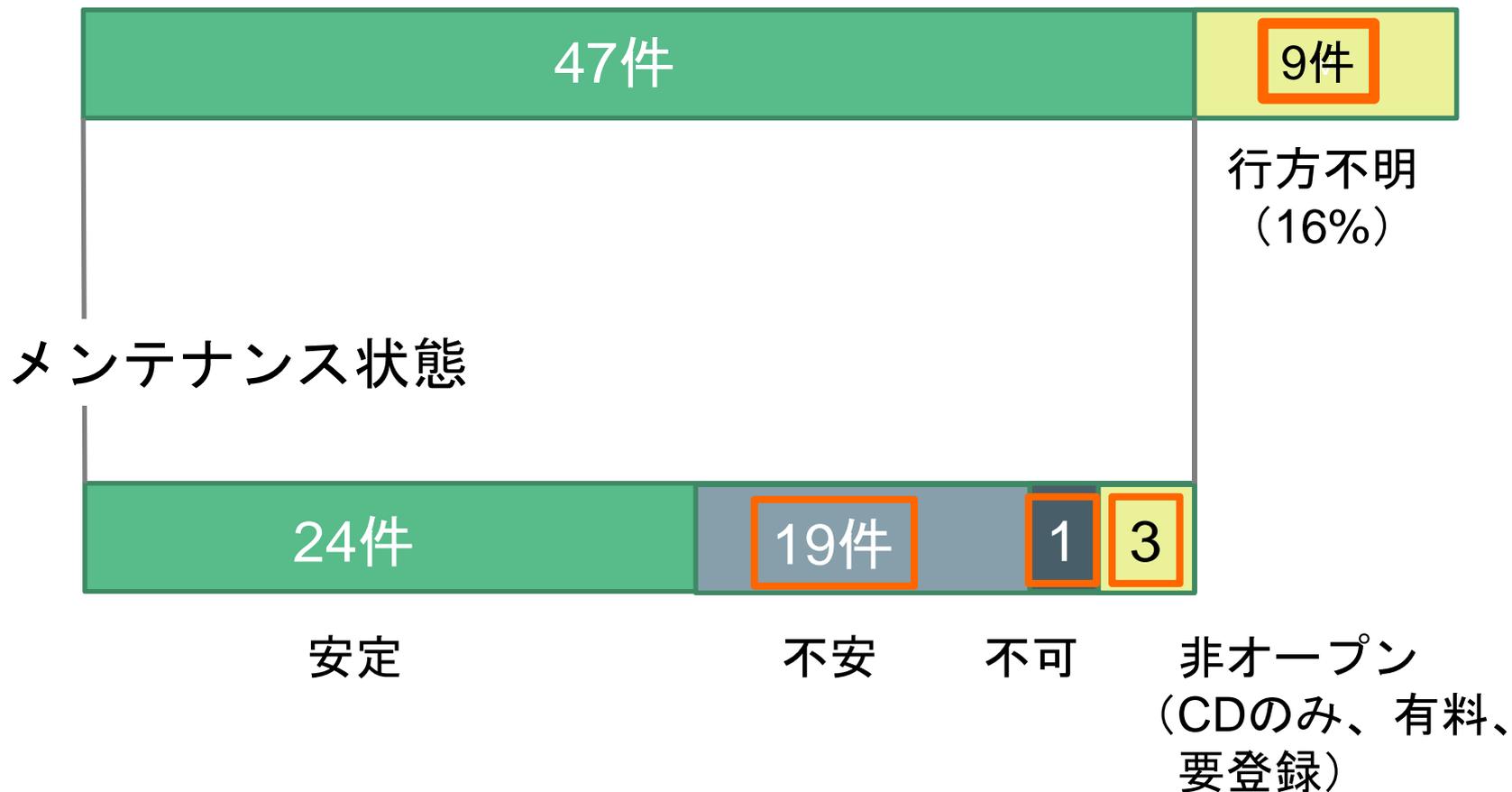
重複を除き、

「公開」済みのもの

129件



URLが活着ている／移動している



科研費の助成を受けて作成されたデータベースの半分が適切に継続管理されていない

データをどう移行する？

1. データー一つ一つにメタデータを振り直して登録

一件ずつ登録するため、再利用可能で綺麗なメタデータを付与できる

⇔ 工数がかかる。独自スキーマへの対応は難しい場合も。

2. データセット単位にメタデータを振って登録

期間、範囲、取得者などで階層を分け、階層ごとアーカイブ

⇔ 作業量は減少するが、再利用しづらい。

データをどう移行する？（続き）

3. 一括登録

DBのメタデータスキーマを機関リポジトリ側のメタデータスキーマにマッピングし、一括登録

⇔ DBのメタデータスキーマは（恐らく）バラバラ。
事案ごとにマッピングする必要がある・・・

4. メタデータ検索サイト（データベース）を別途構築する

機関リポジトリ側にはデータ+既存のスキーマに合致するメタデータのみを登録し、固有のメタデータについては別途メタデータ検索サイトを整備する

→ 利用と保存を分けられるため、検索サイト側の変更（あるいは寿命）に対応できる

⇔ 機関リポジトリ側はトラフィック増加に対応できるか？ 識別子の管理は？

今後の課題

- 管理している研究者や機関との交渉を進める
- 著作権に関する手続きについて検討する
- 機関リポジトリへの移行のベストプラクティスをいくつか実現させる