



JPCOAR スキーマと次期 IRDB

学術情報基盤オープンフォーラム 2018
2018年6月20日(水)

国立情報学研究所学術基盤推進部
学術コンテンツ課
新妻 聡

JPCOARスキーマとは

- 日本の機関リポジトリに登録されたコンテンツのメタデータを流通させるためのフォーマット（メタデータスキーマ）であり、プロトコルはOAI-PMH。
- 国際的な動向・技術の進展にあわせて、現在の機関リポジトリのメタデータスキーマであるjunii2を改訂した、新たなメタデータスキーマ。
詳細はJPCOARスキーマガイドラインのWEBサイト（<https://schema.irdb.nii.ac.jp>）を参照。
- JPCOAR（2016年度は機関リポジトリ推進委員会）と国立情報学研究所のメンバーから構成されるタスクフォースにより、2016年度から検討を開始して、2017年10月にJPCOARスキーマver1.0を策定。
 - 2016年度：メタデータ検討タスクフォース
 - 2017年度：メタデータ普及タスクフォース

JPCOARスキーマ策定（junii2改訂）の背景

1. 学術情報流通を取り巻く環境の変化や社会的な要請への対応
→ オープンサイエンス・オープンアクセス方針に対応した要素の拡充
2. フラットなメタデータ構造の限界
(複数人分の著者名・別名・ヨミ・所属機関の関連性が正確に記述できない)
→ 識別子の拡充・メタデータ構造の修正
3. 日本独自のメタデータ定義
(国際的な学術情報流通の進展とのギャップ)
→ 国際的に相互運用性の高いデータ交換のためのスキーマ定義

JPCOARスキーマ策定（junii2改訂）の基本方針

1. オープンサイエンス・オープンアクセス 方針に対応した要素の拡充

- オープンアクセス方針
 - 5 アクセス権
 - 6 APC
 - 著者ID（3.1 作成者識別子など）
 - 22 助成情報
- 研究データ対応
 - 21 位置情報（geoLocation）
 - contributorType属性（4 寄与者）

JPCOARスキーマ策定 (junii2改訂) の基本方針

2. 識別子の拡充・メタデータ構造の修正

junii2

- <creator lang="ja">国情, 太郎</creator>
- <creator lang="ja">情報, 次郎</creator>

氏名

- creator.alternative: コクジョウ, タロウ
- creator.alternative: ジョウホウ, ジロウ

氏名(ヨミ)
(独自項目、junii2
に出力されない)

- address: 東京大学
- address: 国立情報学研究所

所属
(独自項目、junii2
に出力されない)

creatorとaddressの対応関係がわからない
(たぶん順番通りのセットなのだろうけど)

JPCOARスキーマ策定（junii2改訂）の基本方針

2. 識別子の拡充・メタデータ構造の修正

JPCOARスキーマ

- `<jpcoar:creator>`
 - `<jpcoar:nameIdentifier nameIdentifierScheme="ORCID" nameIdentifierURI="https://orcid.org/0000-0001-0002-0003" />`
 - `<jpcoar:creatorName xml:lang="ja">国情, 太郎</jpcoar:creatorName>`
 - `<jpcoar:creatorName xml:lang="en">Kokujo, Taro</jpcoar:creatorName>`
 - `<jpcoar:creatorName xml:lang="ja-Kana">コクジョウ, タロウ</jpcoar:creatorName>`
 - `<jpcoar:affiliation>`
 - `<jpcoar:nameIdentifier nameIdentifierScheme="ISNI" nameIdentifierURI="http://www.isni.org/isni/0000000121691048" />`
 - `<jpcoar:affiliationName xml:lang="ja">東京大学</jpcoar:affiliationName>`
 - `<jpcoar:affiliationName xml:lang="en">University of`

氏名

氏名(ヨミ)

作成者ごとに情報がまとまる

所属

JPCOARスキーマ策定（junii2改訂）の基本方針

3. 国際的に相互運用性の高いデータ交換のためのスキーマ定義

定義	使用項目例
dc	title, rights, publisher, language, type
dcterms	alternative, accessRights, temporal
rioxterms	apc
datacite	description, date, version, identifier, geolocation, fundingReference
openaire	versionType
dcndl	dissertationNumber, degreeName, dateGranted
jpcoar	creator, contributor, subject, ...

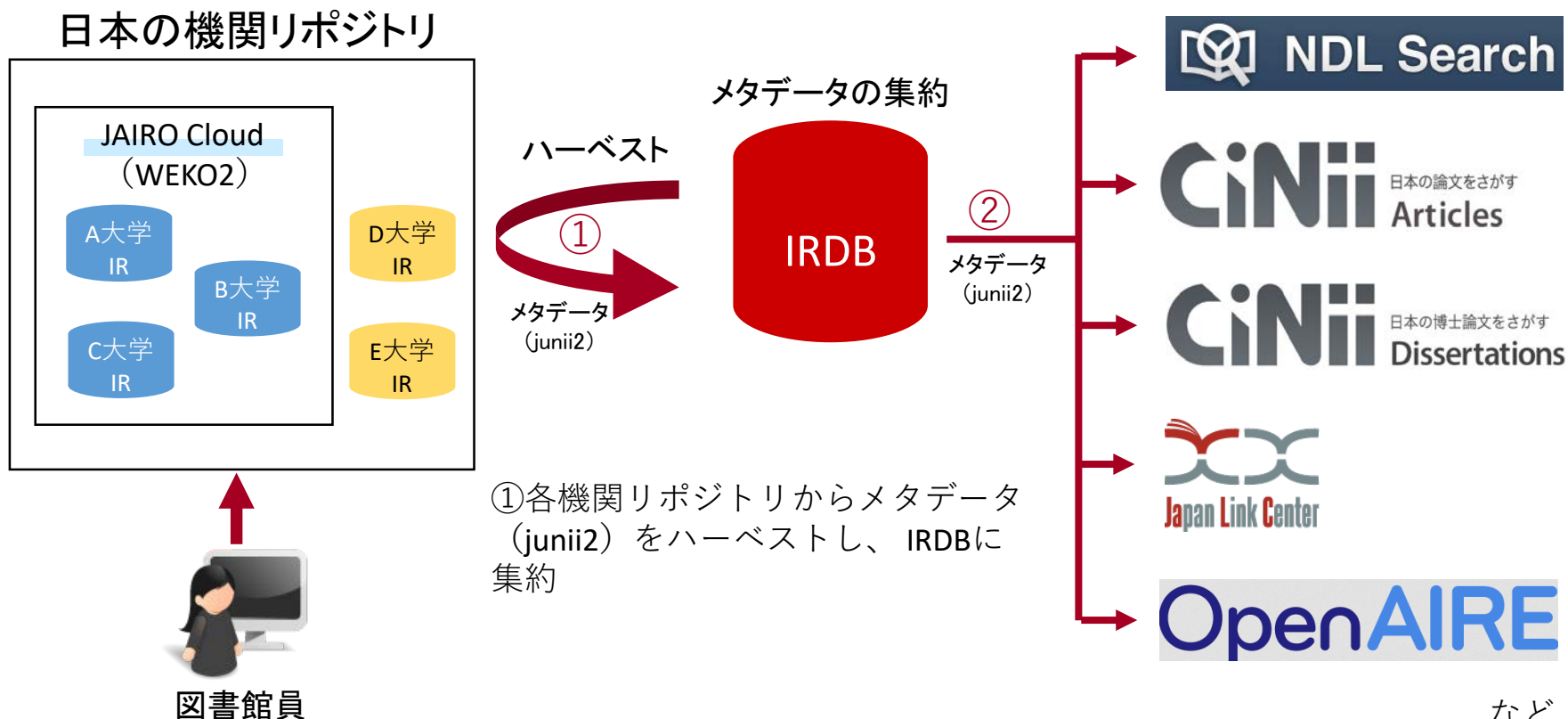
junii2/JPCOARスキーマ要素・属性まとめ

	junii2	JPCOARスキーマ
構造	1階層	3階層
要素数	64要素	79要素（68種類）
移行・内容改訂	21要素	21要素
統合・廃止	43要素	7要素
新規追加・拡張	—	51要素
属性の種類	4種類	15種類
言語属性（使用可能な要素数）	6要素	26要素
言語属性以外（使用可能な要素数）	14要素	29要素
識別子の種類	12種類	28種類
統制語彙（使用する要素・属性数）	5要素	22要素・属性

学術機関リポジトリデータベース (IRDB)

- 各機関リポジトリから収集（ハーベスト）したメタデータを集約し、各種のサービスに流通。

②各種のサービスにメタデータ (junii2) を流通

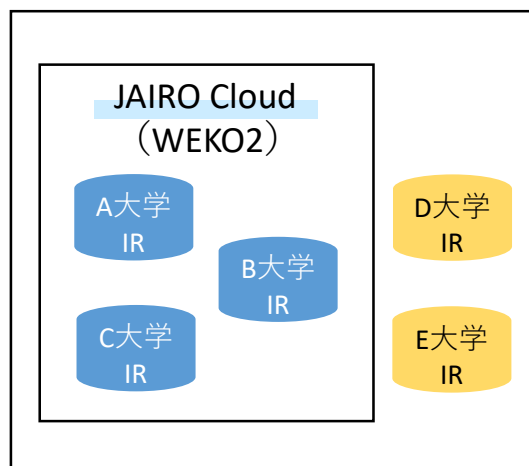


次期の学術機関リポジトリデータベース (IRDB)

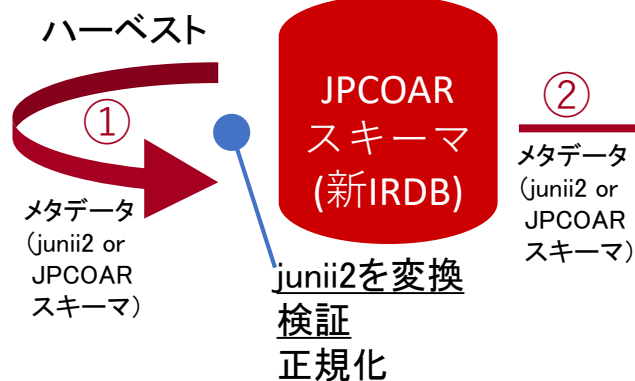
- JPCOARスキーマに対応することにより、質の高いメタデータが収集可能。
- 収集（ハーベスト）したメタデータの検証（バリデーション）や正規化を行うことにより、できる限り精緻かつ誤りのないデータを収集し提供することで、国内外の学術情報流通環境の向上に資する。

②各種のサービスにメタデータ（junii2またはJPCOARスキーマ）を流通

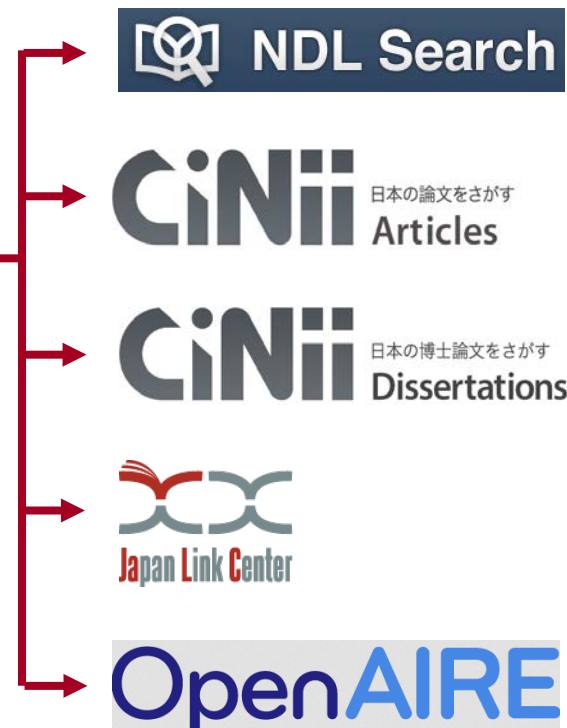
日本の機関リポジトリ



メタデータの集約

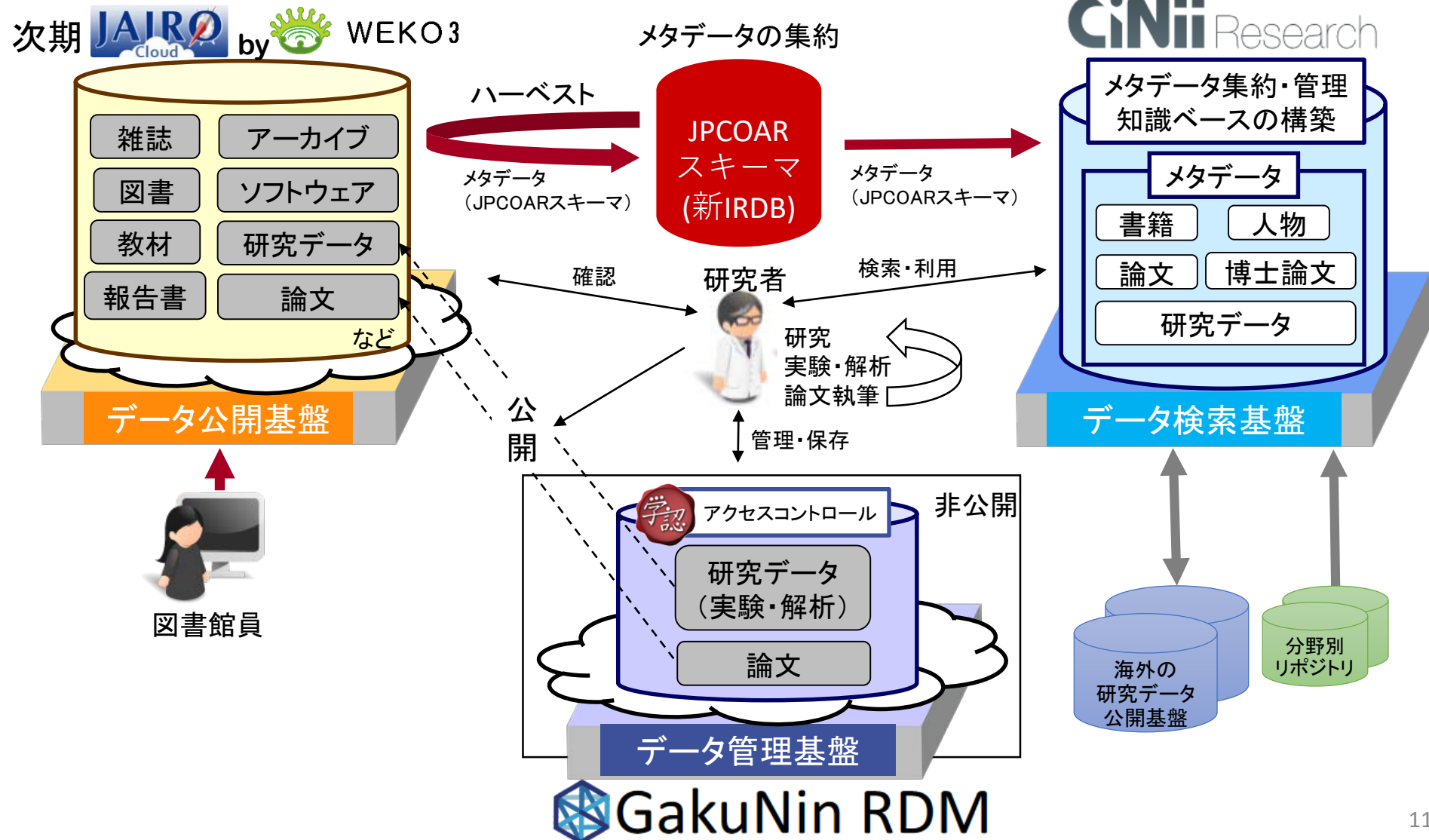


①各機関リポジトリからメタデータ（junii2またはJPCOARスキーマ）をハーベスト、検証（バリデーション）及び正規化した上でJPCOARスキーマでIRDBに集約



次期の学術機関リポジトリデータベース (IRDB)

新IRDBから見た研究データ基盤におけるデータの流れ



次期の学術機関リポジトリデータベース (IRDB)

- JAIROと現行IRDBを統合し、JAIROと現行IRDBは2019年3月末に終了。

JAIROの検索機能は新IRDBに継承



JAIRO Japanese Institutional Repositories Online

2018/05/31現在 690機関 2,908,330件

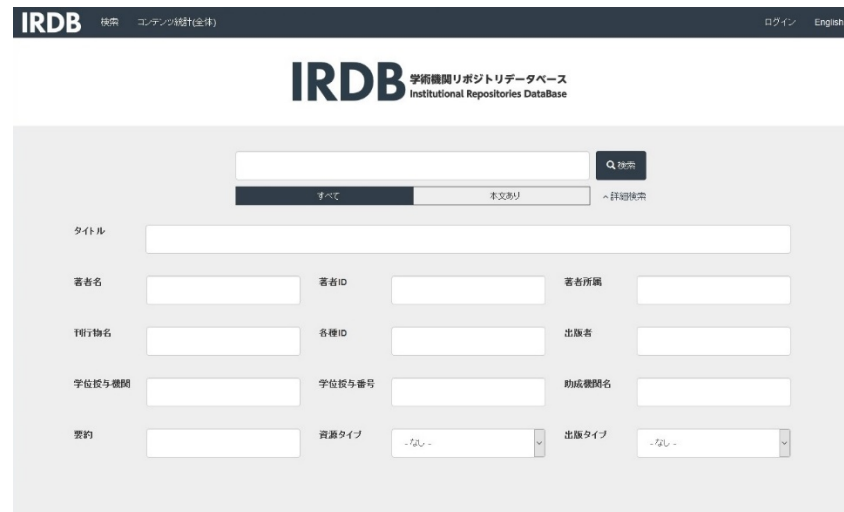
検索条件: すべて (1/1)

検索項目: すべて, キーワード, タイトル, 著者名(ID), 雑誌名, 資料種別, 機関名

検索結果: すべて, 本文あり

検索ボタン: 検索, クリア

JAIROの検索画面



IRDB 学術機関リポジトリデータベース Institutional Repositories DataBase

検索条件: すべて (1/1)

検索項目: タイトル, 著者名, 著者ID, 著者所属, 刊行物名, 各種ID, 出版者, 学位授与機関, 学位授与番号, 助成機関名, 要約, 資源タイプ, 出版タイプ

検索ボタン: 検索

新IRDBの検索画面

※画面は調整中のイメージです。

次期の学術機関リポジトリデータベース (IRDB)

- JAIROと現行IRDBを統合し、JAIROと現行IRDBは2019年3月末に終了。

現行のIRDBコンテンツ分析は新IRDBでリニューアル

IRDBコンテンツ分析

NII Institutional Repositories Database Contents Analysis

IRDBコンテンツ分析について、ヘルプ、お知らせ
お問い合わせ、English

機関名等から検索: 検索

HOME 機関名一覧から探す

ダウンロード

IRDB (Institutional Repositories DataBase: 学術機関リポジトリデータベース) コンテンツ分析システムでは、各所の詳細情報、コンテンツ内容の統計分析情報、機関リポジトリのクローラー(ロボット)情報を提供しています。

詳細

分析対象:
IRDB/ハーベスト機関リポジトリ数: 681、コンテンツ数 全体: 2,839,740件、本文あり: 2,151,826件
(2018/05/31)

コンテンツ増減

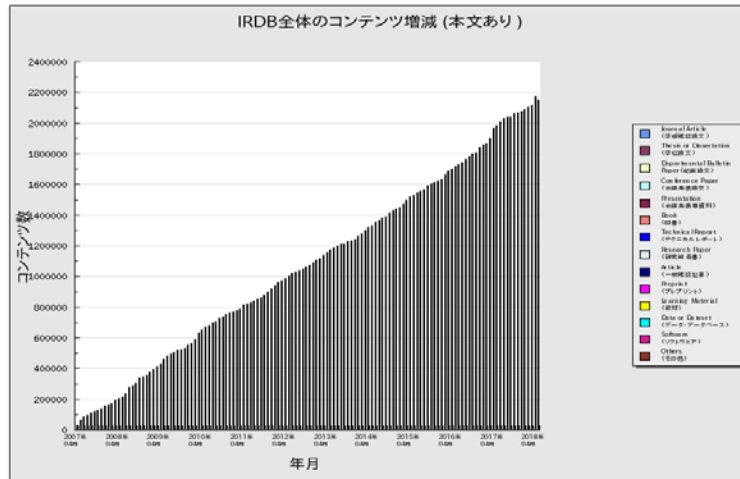
すべて 表示する年度を選択できます。 [本文あり] 本文ありと全体から選択できます。

最新月 [2018/5] のコンテンツ数 (本文あり)

TSVダウンロード

過去のコンテンツ数 (本文あり)

TSVダウンロード



IRDB

検索 コンテンツ統計(全体)

ログイン English

対象年:

2018

対象月:

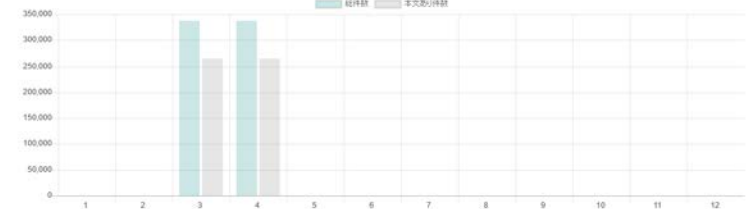
-すべて-

検索

CSV出力

月別件数

	1	2	3	4	5	6	7	8	9	10	11	12
総件数	0	0	337,346	337,346	0	0	0	0	0	0	0	0
本文あり件数	0	0	264,086	264,086	0	0	0	0	0	0	0	0



総件数	本文あり件数(率)	DOI付件数(率)
337,346	264,086 (78.28%)	11,334 (3.36%)

新IRDBのコンテンツ分析画面

※画面は調整中のイメージです。

現行IRDBのコンテンツ分析画面

次期の学術機関リポジトリデータベース (IRDB)

新たに追加した機能

- 機関の業務担当者向けのログイン機能を実装することにより、業務担当者が自機関のコンテンツを分析したり、ハーベストされたメタデータを検索、ハーベストエラーを確認できる。

IRDB 検索 コンテンツ統計(全体) マイコンテンツ コンテンツ統計(ユーザ) ユーザー情報 ログアウト English

大学

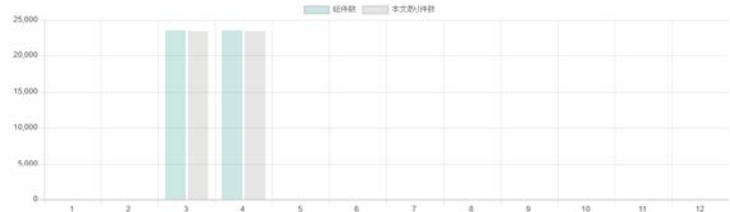
対象年: 2018 対象月: -すべて-

検索

CSV出力

月別件数

	1	2	3	4	5	6	7	8	9	10	11	12
総件数	0	0	23,527	23,527	0	0	0	0	0	0	0	0
本文あり件数	0	0	23,388	23,388	0	0	0	0	0	0	0	0



総件数	本文あり件数(率)	DOI付与件数(率)
23,527	23,388 (99.41%)	0 (0.00%)

自機関のコンテンツ分析画面

IRDB 検索 コンテンツ統計(全体) マイコンテンツ コンテンツ統計(ユーザ) ユーザー情報 ログアウト English

検索条件

機関ID: 機関IDのOAI-ID: OAI識別子:

更新日時: DOI:

項目・サブ項目(以上を検索): ワーキング件数(以上を検索): 項目実数件数(以上を検索):

メッセージ: 状態:

検索

23527件中1件目～50件目を表示

OAI-ID	表示/非表示	状態	OAI識別子	機関IDのOAI-ID	項目エラー件数	ワーキング件数	項目実数件数	メッセージ	DOI	更新日時
oai:ncic.jp:00024154	表示	---	oai:ncic.jp:00024154	0	0	0	0		04/21/2018 17:42	
oai:ncic.jp:00024153	表示	---	oai:ncic.jp:00024153	0	0	0	0		04/21/2018 17:40	
oai:ncic.jp:00024143	表示	---	oai:ncic.jp:00024143	0	0	0	0		04/21/2018 17:41	

※該当アイテムのURL: <http://hdl.handle.net/10003/1629>
<http://ncic.jp/00024154>
<http://ncic.jp/00024153>
<http://ncic.jp/00024143>

※項目登録エラー(1件)を確認しています。
項目登録エラー(1件)を確認しています。
(学内番号) (学内番号) (学内番号) (学内番号)

oai:ncic.jp:00024156	表示	---	oai:ncic.jp:00024156	1	0	0	0		04/21/2018 17:41
----------------------	----	-----	----------------------	---	---	---	---	--	---------------------

自機関のメタデータ検索画面

※画面は調整中のイメージです。 14

新IRDBへの移行スケジュール（予定）

2018年

2019年

	3	4	5	6	7	8	9	10	11	12	1	2	3	4
開発	→													
IRDB移行 アナウンス			★											
新規ハーベスト		→	→	→										
ハーベスト結果の 分析			→	→	→									
新規の 再ハーベスト						→	→	→						
レポート送付						→	→	→						
データ修正						→	→	→	→					
定期ハーベスト								→	→	→	→	→	→	→
外部データ提供先 機関との調整			→	→	→	→	→	→	→	→	→	→	→	→
システム切替									→	→	→	→	→	→

※NIIにおける予定

現行IRDB
終了

次期の学術機関リポジトリデータベース (IRDB)

今後の予定

- 新IRDBの機能詳細については、今年度のJPCOARスキーマ説明会で詳しく説明する予定です。

@東京会場

日時：2018年8月6日（月）13:00～16:15

会場：国立情報学研究所（NII）12F 1208会議室、1210会議室

@神戸会場

日時：2018年9月6日（木）13:00～16:15

会場：神戸大学瀧川記念学術交流会館会議室

詳しくは、JPCOARのWEBページをご覧ください。
興味のある方はご参加ください。

URL：https://jpcoar.repo.nii.ac.jp/?page_id=49#_href_229