

典拠データの動向、 課題と展望

木村麻衣子（日本女子大学）

@2024年度書誌調整連絡会議



典拠データに含まれる要素

目録で扱われる特定の実体について

- ① 優先的に使用されるべき典拠形アクセス・ポイント（統一標目）
- ② 優先的に使用しないが、発見の手がかりとなる異形アクセス・ポイント（参照）
- ③ 他の実体との区別に資する属性など
- ④ 関連を有する別の実体の典拠形アクセス・ポイント
- ⑤ アクセス・ポイントの決定の根拠となった情報源

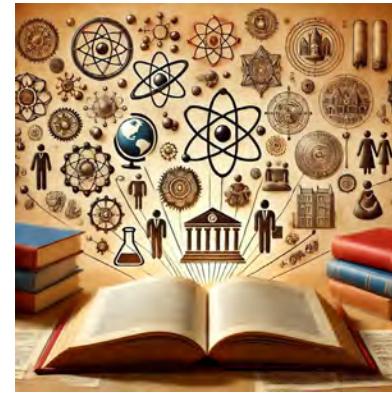
清少納言, 平安時代

ID	00271681	
典拠種別 skos:inScheme	個人名	
名称/タイトル xl:prefLabel	清少納言, 平安時代	①優先的に使用されるべき典拠形アクセス・ポイント
名称/タイトルのカナ読み ndl:transcription@ja-Kana	セイ ショウナゴン, ヘイアン ジダイ	
名称/タイトルのローマ字読み ndl:transcription@ja-Latn	Sei Shonagon, Heian Jidai	
別名/別タイトル xl:altLabel	清原 (姓) ; Sei Shonagon; Сей-Шонагон; Sei Shōnagon	②発見の手がかりとなる異形アクセス・ポイント
経歴 rda:biographicalInformation	女流文学者	③他の実体との区別に資する属性など
著作 is dct:creator of	枕草子	④関連を有する別の実体の典拠形アクセス・ポイント
関連リンク skos:exactMatch	n50009764 (LCNAME); 000000043360592X (ISNI); DA00360673 (NII); 285110 (KOKUSHO); NDL 00271681 (VIAF)	
出典 dct:source	枕草子 大人名事典	⑤アクセス・ポイントの決定の根拠となった情報源
作成日 dct:created	1993-06-25	
最終更新日 dct:modified	2024-03-08T14:35:39	
外部サイトへのリンク	Wikipediaで検索を行う ※Wikipediaの検索結果の一覧が表示されます。	出典: Web NDL Authorities, https://id.ndl.go.jp/auth/ndlna/00271681

日本の主な典拠データ



著者名
個人・家族・団体



件名
主題・個人・家族・団体・著作



著作

出版者, シリーズ名の典拠データも(一部には)あり

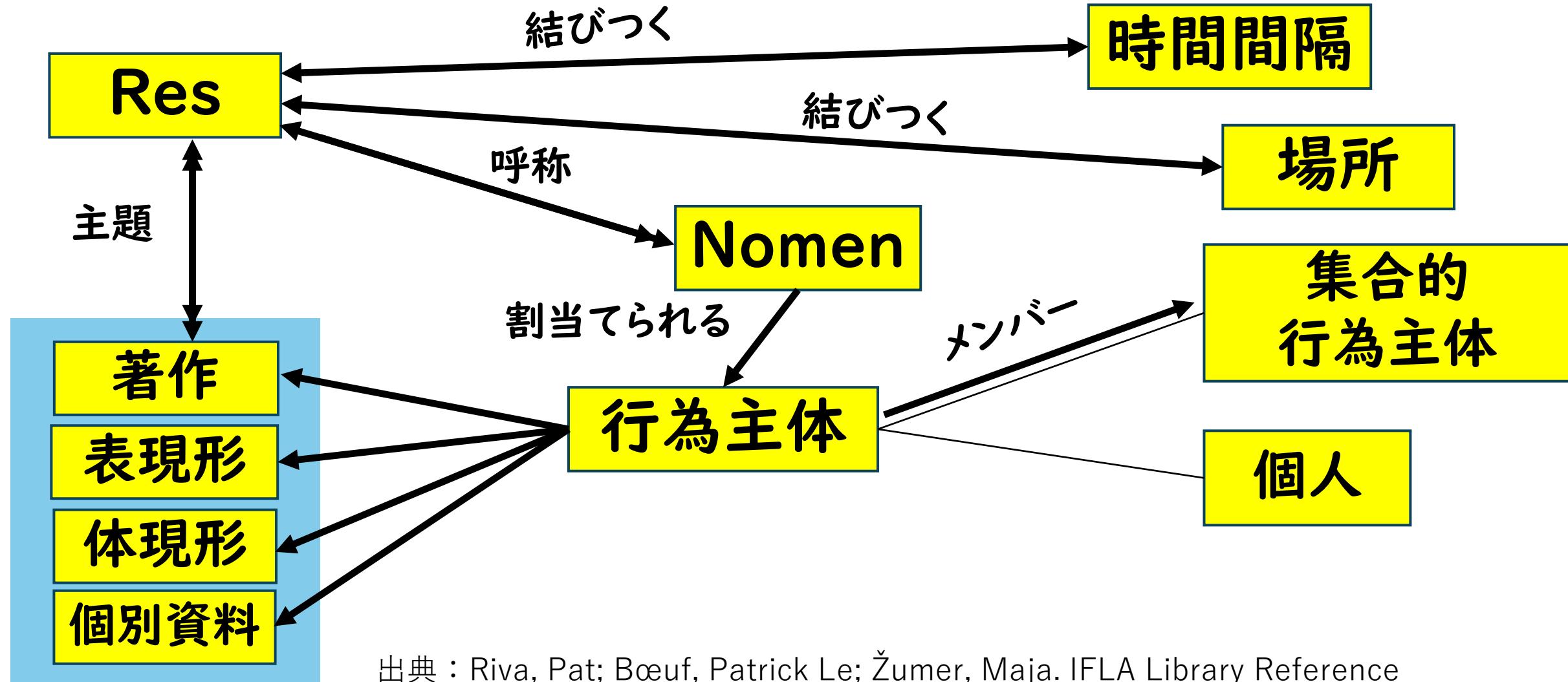
典拠データの作成単位

- FRBR以前：典拠形アクセス・ポイントごと
 - 1つの典拠形アクセス・ポイントにつき1件の典拠レコードを作成する
- 現在：実体ごと
 - 1つの実体につき1件の典拠データを作成する (ICP2009)

#典拠レコードを典拠データと呼ぶようになったのは「典拠データの機能要件(FRAD)」の頃から(2009~)か。

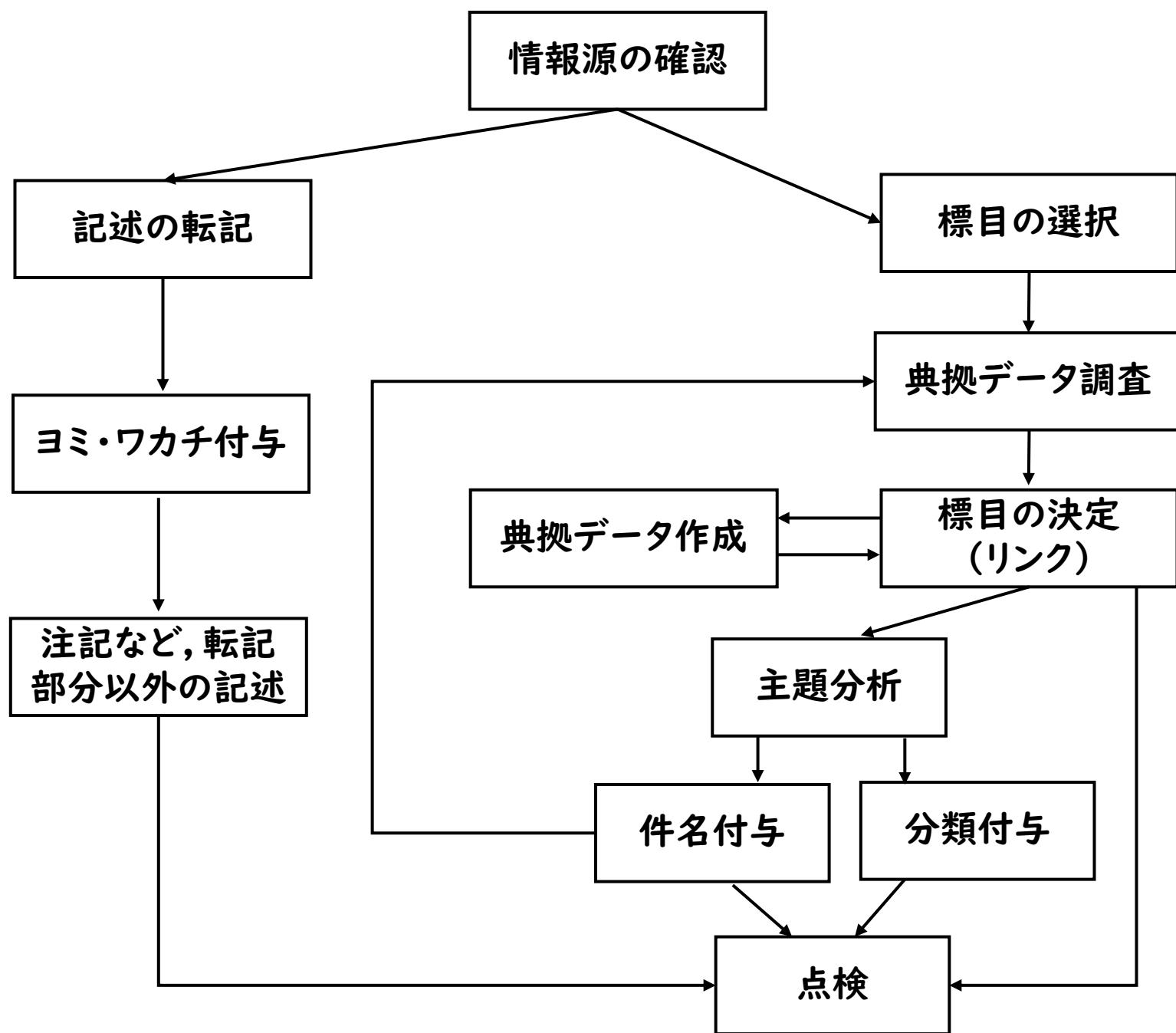
#典拠レコードの集積を典拠ファイルと呼んでいた名残りで、典拠データの集積も典拠ファイルと呼ばれる（典拠データベースと呼ぶべき？）

典拠データの作成対象となりうる実体 (IFLA LRM)

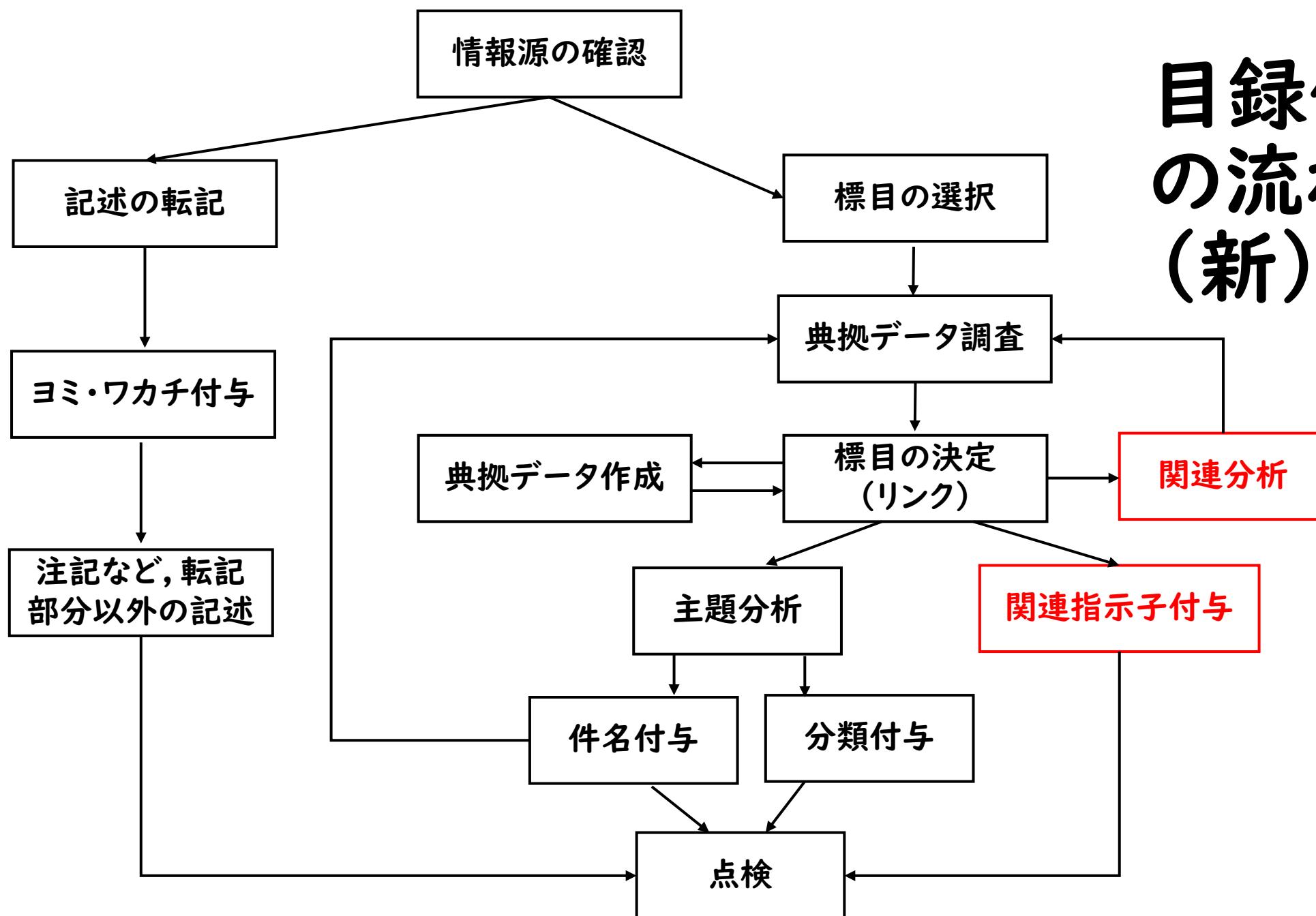


出典 : Riva, Pat; Bœuf, Patrick Le; Žumer, Maja. IFLA Library Reference Model : a Conceptual Model for Bibliographic Information. 2017-12. p. 86

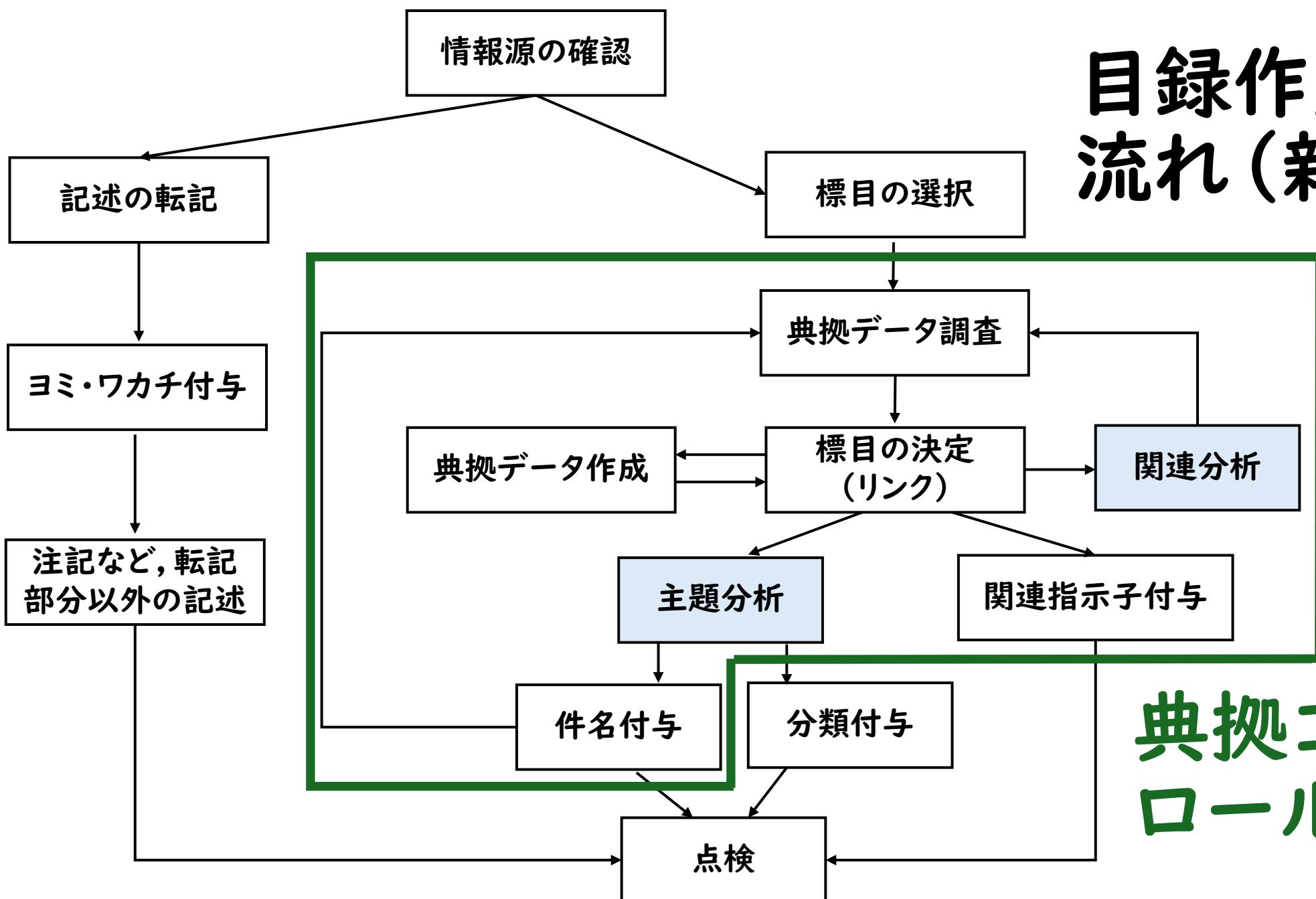
目録作業 の流れ (旧)



目録作業 の流れ (新)

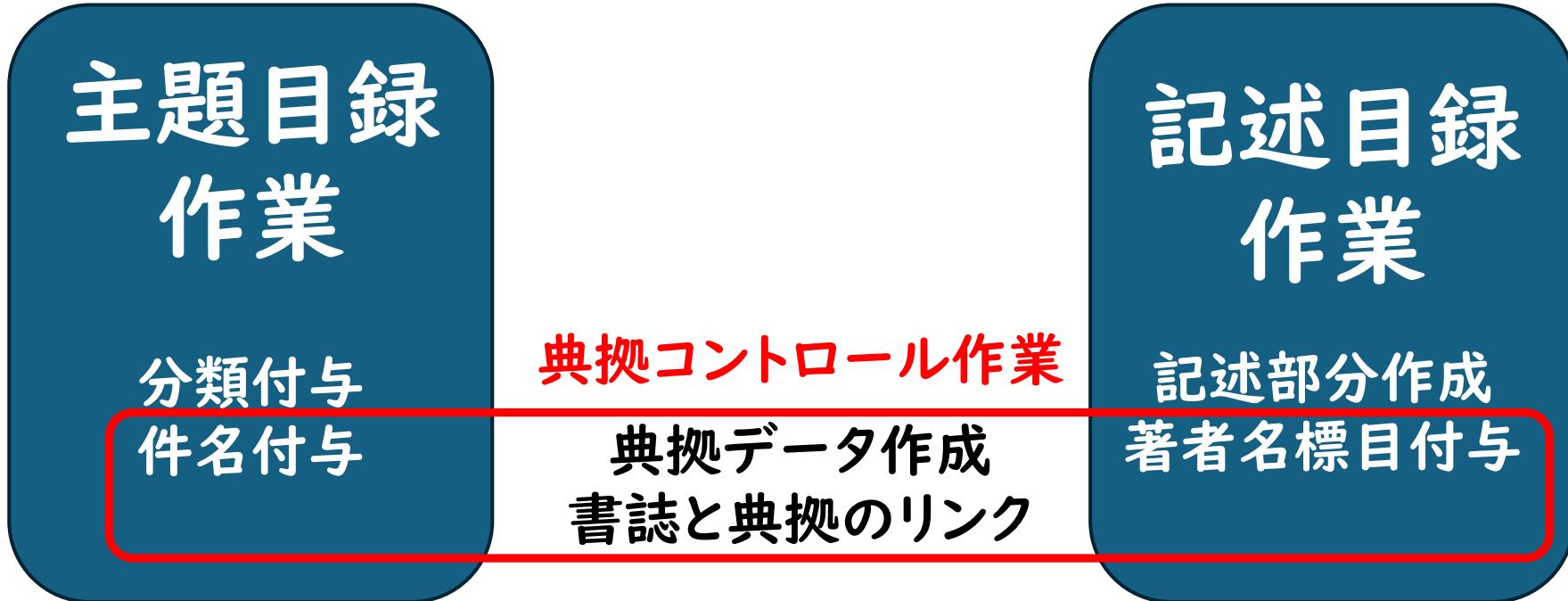


目録作業の流れ(新)



典拠コント
ロール作業

典拠コントロールの位置づけ (従来の型で説明)



- ・ 件名標目表から件名を選んで付与するだけなら典拠コントロール作業ではない
- ・ 著者名標目の形を別の書誌データの著者名標目の形に揃えるだけ(標目調整)なら典拠コントロール作業ではない

典拠コントロール作業

典拠データ作成

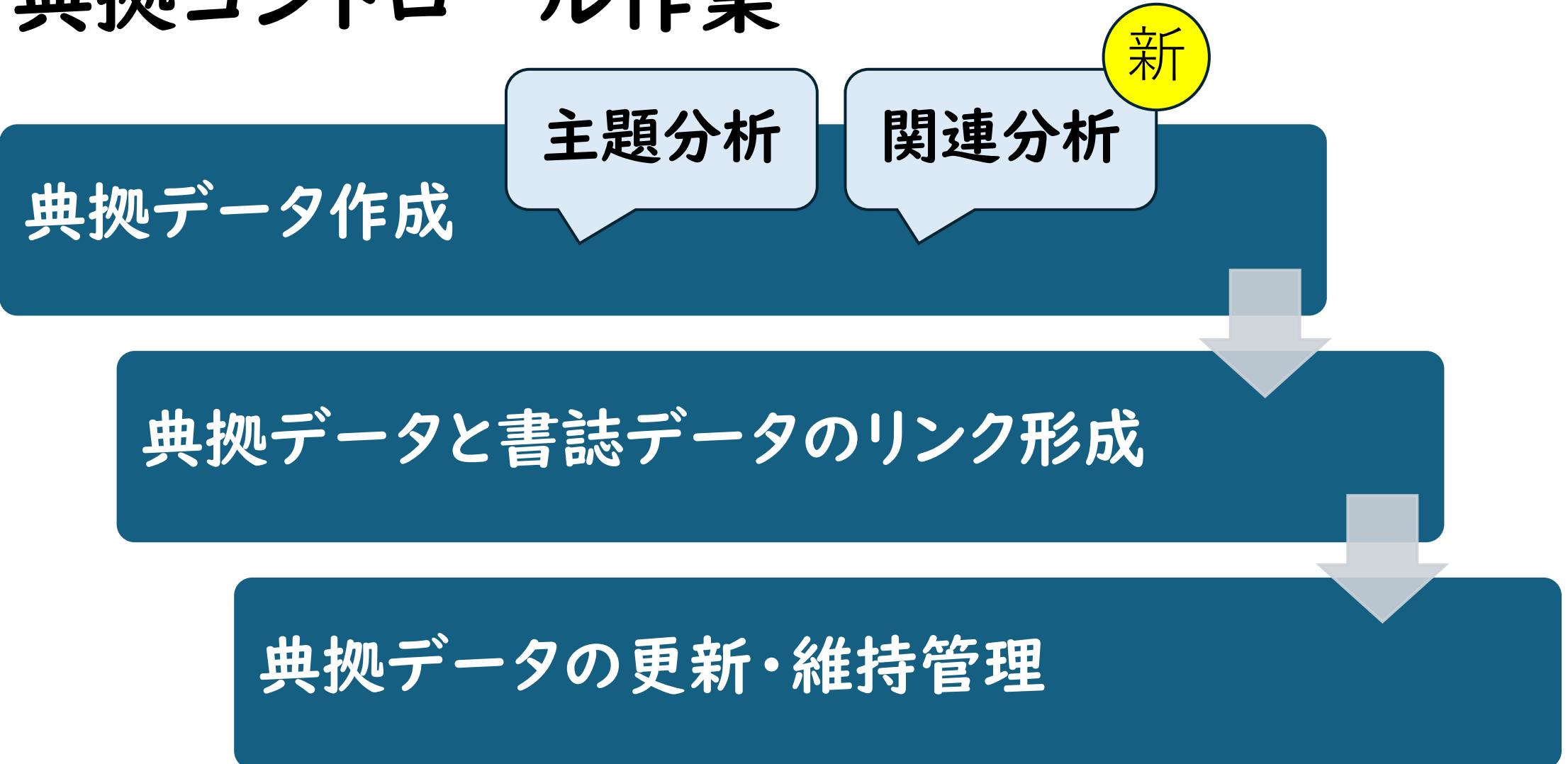


典拠データと書誌データのリンク形成



典拠データの更新・維持管理

典拠コントロール作業



典拠データ作成作業の核心

主題分析

主題目録作業の一環として行われる、情報資源の主題を考える作業

関連分析

目録中に現れる実体と他の実体との関連を考える作業



典拠ファイルの機能

1. 目録における典拠形アクセス・ポイントの使用に一貫性を持たせ、特定の実体が必ず特定の典拠形アクセス・ポイントによって識別されるようにする
2. 特定の実体を、異形アクセス・ポイントからも発見できるようにする
3. 書誌データにおける典拠形アクセス・ポイントの拠り所となる
4. 実体間の関連を通じて、ほかの実体をも発見できるようにする

新

附加価値

典拠コントロールのメリット

1 異名同実体の集中

異なる名称を持つ同一の
実体の検索漏れを防ぐ

2 同名異実体の識別

同じ名称を持つ異なる実体を
区別し、ノイズなく検索させる

3 実体の発見を支援

- ・さまざまな名称からの実体
の発見
- ・実体に関連する実体の発見

4 信頼のおける情報源の提供

- ・実体の異名、属性などが記録
されているため、著者や著作に
ついて調べる際の情報源として
典拠ファイルを利用できる
- ・AIの学習材料にも？

メリットを 最大限享受するには

- ・典拠データの作成を網羅的にする

典拠コントロールされていない標目もあるという状況は網羅的ではない

- ・典拠データの種類を増やす
- ・典拠データを幅広く共有する

皆で同じものを使えば相互運用・相互利用が容易





日本の現状

1. 館種毎に異なる典拠ファイルを使用
2. 件名の典拠コントロールは根付いていない
3. 著作の典拠コントロールは根付いていない
4. 著者名の典拠コントロールも必須と考えられていない
5. 典拠コントロールへの無理解？

I. 館種毎に異なる典拠ファイルを使用

典拠データ の種類	NDL ¹⁾	NACSIS-CAT ³⁾	TRC ⁵⁾ (単位：件)
個人	1,009,166		1,375,036
家族	2,766	1,865,190	
団体	234,957		305,182
地名	36,467		
出版者			210,987
件名	145,920		110,616
学習件名			24,245
シリーズ			126,309
全集			335,567
著作	6,193 ²⁾	40,165 ⁴⁾	42,832

注
 1)Web NDL Authorities提供件数, 2024-03-31現在
 2)統一書名+著作
 3)2025-02-16現在
 4)統一書名
 5)統一形+記述系+参照形の合計, 2025-01-31現在

出典：“統計からみた書誌データ（令和3年度以降）”. 国立国会図書館. https://www.ndl.go.jp/jp/data/basic_policy/statistics/FY2021-.html

“総合目録データベースの現況および参照MARCの収納状況”. 国立情報学研究所目録所在情報サービス. 2025-02-17, <https://contents.nii.ac.jp/catill/stats/cat/db>

“典拠ファイルT・累積件数表”. 2025-01-31. https://tooli.trc.co.jp/news/public/marc/routine_up/tenkyo/tenkyo_kensu.pdf

2. 件名の典拠コントロールは根付いていない

- ・複数の細目を事前結合した件名標目は付与が難しい
- ・細目つき件名がOPACでうまく検索されないことがある
- ・カード目録時代から件名目録を持っている図書館は20%以下¹⁾

⇒件名を付与しない図書館のほうが多い²⁾

⇒分類が決まれば件名が決まるような特定性の低い件名が多い³⁾

1)加藤宗厚. 件名入門. 理想社, 1972, p.41.

2) 2010年のJLA「目録の作成と提供に関する調査」では、目録作成業務を行っている館（公共は回答館の92.5%，大学は回答館の82.2%）のうち件名の典拠コントロールを行っていたのは公共36.2%，大学28.3%であった。

3) CiNii Booksの書誌データに付与された件名標目を見ての木村の感想。

3. 著作の典拠コントロールは根付いていない

- NCR1987では、「統一タイトル」標目は“無著者名古典、聖典および音楽作品の範囲内”でのみ作成という規定で、かつ、任意規定であったため、典拠ファイルが作成されることが稀であった



- NCR2018では、“全著作の典拠コントロール”が求められるが、その基礎となるべき「統一タイトル」典拠ファイルがほぼ存在しない
- NCR2018適用機関であっても、著作の典拠ファイルをゼロからすべて作成することは困難で、現在のところ部分的にしか作成されていない

4. 著者名の典拠コントロール も必須と考えられていない

- NCR 1965以来の標目に関する規則の一部は、NCR 1977にも引き継がれていたが、NCR 1987に至ってほぼ消滅（1987年9月）
 - 記述独立方式では優先的な標目を決めなくともよくなつたため？（それでも標目を記録しなくてよくなつたわけではないのだが）
- NASIS-CATの典拠リンク任意化（1987年11月）
- 著者名典拠コントロールを行っている図書館は、公共 56.0%，大学 63.7% (JLA, 2010¹⁾)



1) 2010年のJLA「目録の作成と提供に関する調査」での、目録作成業務を行っている館（公共は回答館の92.5%，大学は回答館の82.2%）のうちのパーセンテージ。「行っている」と回答していても、完全な典拠コントロールをおこなっているわけではないかもしれません。

5. 典拠コントロールへの無理解？①

以下すべて風聞

- ・ 民間MARCの購入に当たり、書誌データは契約するが典拠データは契約しない図書館もあるらしい
 - ・ 書誌データだけでは目録とは言えない
- ・ 典拠データを契約(orダウンロード)していても、更新には対応しない図書館もあるらしい
 - ・ 維持管理しないなら典拠コントロールとは言えない
- ・ 典拠データを契約(orダウンロード)していても、OPACでの検索には利用できない図書館もあるらしい
 - ・ 典拠データの役割が理解されていない

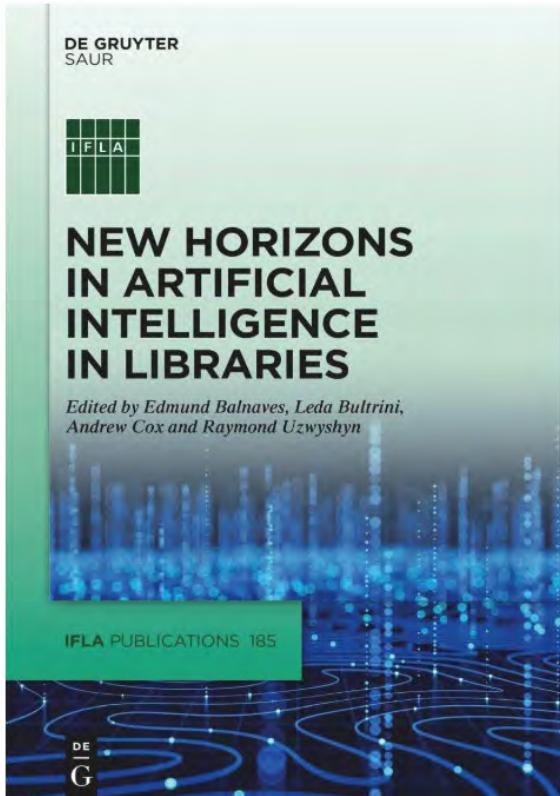
5. 典拠コントロールへの無理解？②

新

以下の声にどう答えるか？

- ①典拠データなぞ作らなくても、名寄せは機械学習でできるのでは
- ②AIが件名自動付与できるのでは
- ③件名なぞ不要
- ④著者名、書名の典拠データもAIで作ればいいのでは

『図書館における人工知能の新たな地平』



オープンアクセス

13章 スウェーデン国立図書館がテキストから主題語を抽出する試み

14章 ライプニツツ経済情報センターが件名自動付与を模索

16章 LCが電子書籍のテキストから実体を識別し名称典拠ファイルとの自動突合せを試行

17章 フランスの共同目録で責任表示記述から自動的にアクセス・ポイントを生成し典拠ファイルとリンクする仕組みを構築中

22章 Ex Librisが自社製品にAIを実装

Balnaves, E., Bultrini, L., Cox, A. & Uzwyshyn, R. (2025).
New Horizons in Artificial Intelligence in Libraries.
Berlin, Boston: De Gruyter Saur.
<https://doi.org/10.1515/978311336435>

5. 典拠コントロールへの無理解？②

新

以下の声にどう答えるか？

- ① 典拠データなぞ作らなくても、名寄せは機械学習ができるのでは
- ② AIが件名自動付与できるのでは
- ③ 件名なぞ不要
- ④ 著者名、書名の典拠データもAIで作ればいいのでは

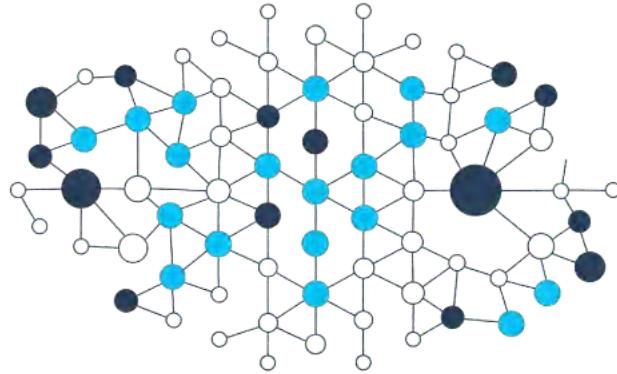
今のところの回答：

- ① 全部はできない。誤りへの許容度にもよる。異名からの検索はできない
- ② 全分野を網羅し、細目を含む特定性の高い件名を付与するのは無理
- ③ それに代わる検索手段が件名より優れているかはわからない
- ④ 誤りへの許容度によるが、誤りは最小限でないと機能を果たせない。人手で確認するのだとしたら、手間はあまり変わらない

今後に向けた課題



1.重要性・効能への無理解



2.関連分析の範囲



3.OPACの開発



I. 重要性・効能への無理解

30

空白

1987年から2018年まで



目録経験

図書館員の目録経験不足



研究者（教員）

情報資源組織化分野の研究者不足



2. 関連分析の範囲

1 時間とリソース

どこまでできるのか？

2 利用者のニーズ

どこまで求められているのか？

3 専門知識

本当にできるのか？

3. OPACの開発

- ・質の高い典拠データがあっても、それが検索に活用できないのではコストをかけるメリットがない
- ・図書館員（仕様を書く側）とベンダー（開発する側）の双方に啓蒙が必要



まとめ

- ・質の高い検索を提供するためには今後も典拠コントロールが必要
- ・典拠データを信頼のおける情報源として広く公開することは社会的に有益
- ・関連分析は典拠データに付加価値をつける、これからの典拠コントロールの核心部分（ただし、どこまでやるのかは要検討）
- ・日本では付加価値以前の現状である場合が多く啓蒙が必要
- ・質の高い典拠データを検索に活用できるようなOPACの開発が俟たれる