

機関リポジトリ収録文献の ビジビリティ向上

平成21年度CSI委託事業報告交流会(コンテンツ系)

平成22年6月22日

「機関リポジトリ上の情報資源の発見及び

アクセス性の向上のための調査研究開発(AIRway)」

北海道大学附属図書館学術システム課

野中雄司

1

AIRwayプロジェクト発足のきっかけ

とある学内での研究者向け機関リポジトリ説明会でのこと

- **説明)図書館**

- 機関リポジトリに論文を登録すると、Google などからアクセスできるようになります。

- **質問)研究者**

- 私たちは文献を探すときは普通Web of Scienceとか、PubMedを使うのですが、それらからはアクセスできないのでしょうか？

どうすればよい？

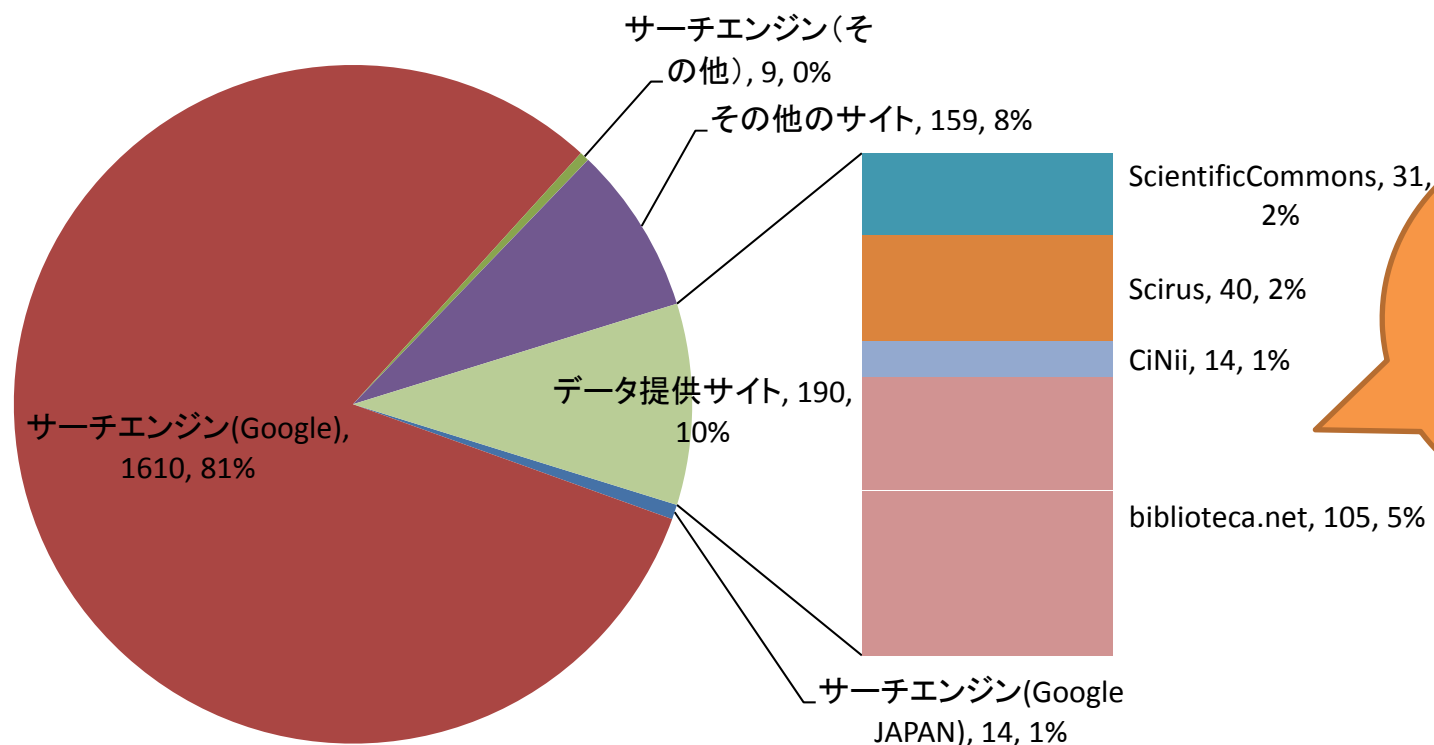
- OAI サービス・プロバイダ(OAIsterなど)の認知度は未だ低く、一般的なインターネット検索エンジンは学術文献の探索という用途には**通常用いられない**。
- 研究者の**自然な情報探索行動**の中に、機関リポジトリ上の文献発見を位置づけていくためにはどのような方法があるか？

HUSCAPへのアクセス経路

- 学術雑誌論文(英語)のみを対象
- 国外からのアクセスのみを対象
- 教育機関からのアクセスのみを対象

平成21年
1月～8月

もちろん
フィルタリン
グ後



ほとんど
Google
から!

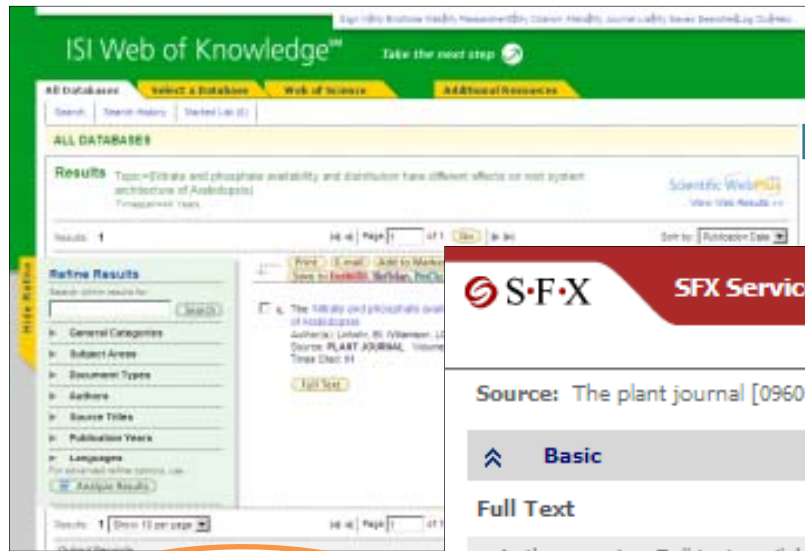
2

AIRwayとは？

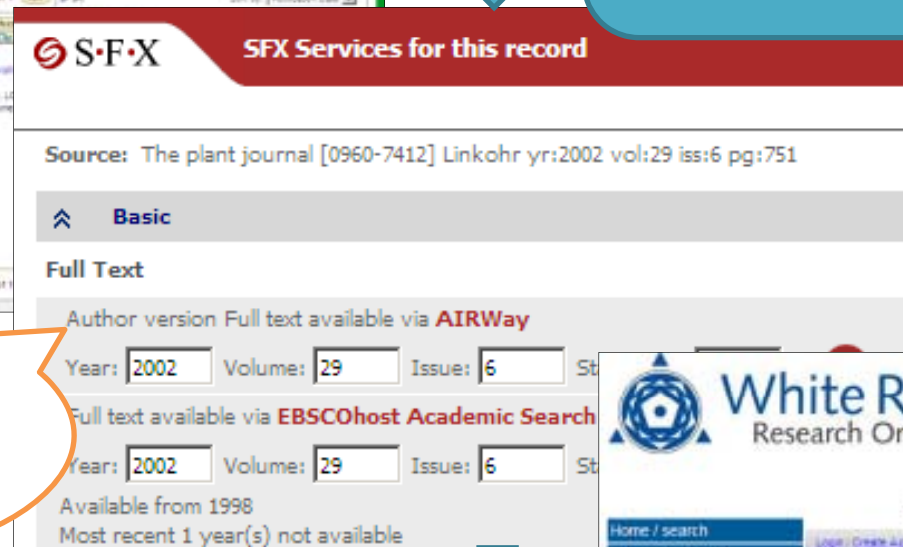
リンクリゾルバを使用し，文献データベースから電子ジャーナルだけでなく，機関リポジトリに搭載された文献へ利用者を導く

AIRway

文献データベースの検索結果からお目当ての論文についているリゾルバアイコンをクリック



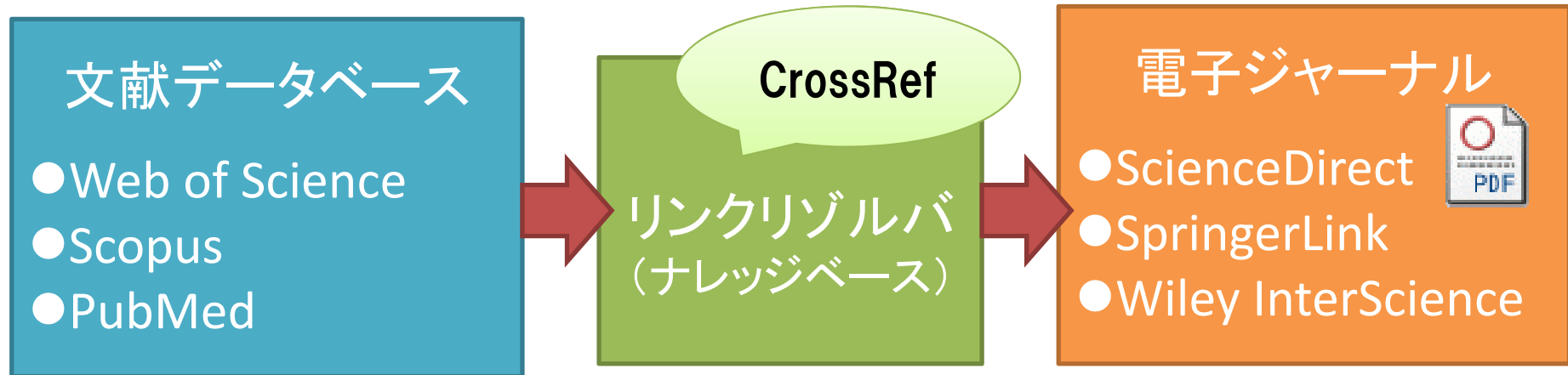
AIRwayでヒットしたらAIRwayリンクが出現



リゾルバ中間窓のAIRwayリンクをクリック。
機関リポジトリ上の該当論文へ。



AIRway概要

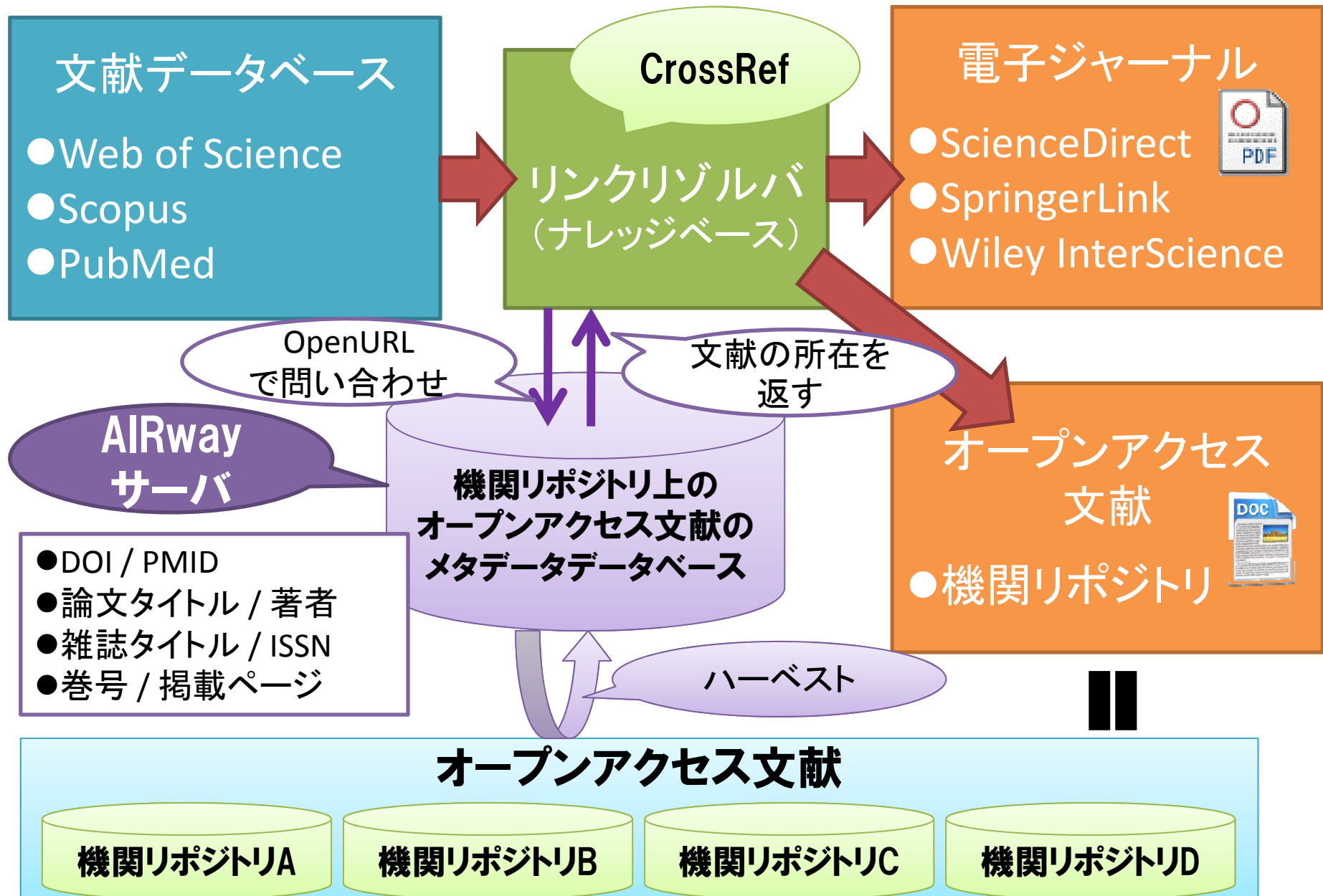


もし雑誌の購読ライセンスがあるならば、リンクリゾルバからその論文へ到達することができるが、ライセンスがなければ、リンクリゾルバは論文へナビゲートしてくれないし、当然読むこともできない。
その論文が**世界のどこかの機関リポジトリ**にありうとも。

オープンアクセス
文献

- 機関リポジトリ

AIRway概要



3

AIRway現況

現況

- **対応リンクリゾルバ**
 - SFX(Ex Libris社)
 - グローバルバージョンは平成21年度から
 - WorldCat Link Manager(OCLC)
 - 360 Link(Serials Solutions社)
 - 平成21年度から
- **収録レコード数**
 - 約35万レコード
 - (平成20年度末は27万レコード)

現況

ハーベスト対象機関リポジトリ

(国内:JAIRO、国外:30リポジトリ(増加作業中))

- Japan
 - JAIRO : Japanese Institutional Repositories Online
- Australia
 - Queensland University of Technology ePrints Archive (QUT ePrints Archive)
- Germany
 - Max Planck Society eDoc Server
- India
 - Dspace@NITR
- United Kingdom
 - White Rose Research Online
- United States
 - Deep Blue at the University of Michigan
 - PubMed Central (PMC) #only author manuscripts

など

メタデータの収集は容易ではない

- 各機関リポジトリごとに、OAI-PMHでどのようなメタデータフォーマットがあり、どのような値が取得できるかを調査し、かつそれに適合するハーベスタ(データ取得用のプログラム)を作らなければならない。
- OAI-PMHで必須であるoai_dcで充足できればベストではあるが、そううまくはいかない。
 - 機械可読な掲載誌書誌情報取得不可
 - 識別子(DOIなど)の記述方法がいろいろ。
 - 本文のあり/なしや、アクセス制限の有無、その本文の版、などは確認できない。
- junii2はOK。

現状のAIRwayで使用しているハーベスタを作るのも一苦労だった！

AIRwayハーベスタ

- **ハーベスタ1 (junii2)**
 - IRDB(JAIRO)からハーベスト
- **ハーベスタ2(oai_dc)**
 - [dc:format=application/pdf] がセットされているかつ [dc:identifier=http://dx.doi.org] もしくは [dc:relation=http://dx.doi.org] が含まれているもの
- **ハーベスタ3(zim_export)**
 - [genre=Article] のみ, [fturl viewftext=PUBLIC] のみ, [fturl filename=pdf] が含まれているもののみ
- **ハーベスタ4(oai_dc+context_object)**
 - [dc:format=application/pdf] or [dc:format=pdf] があり, かつ [dc:type=Article] があること。
 - 上記条件をクリアしたらcontext_objectから書誌事項取得
- **ハーベスタ5(PubMed Entrez)**
 - 「author manuscript」のみを取得

4

国外機関リポジトリの メタデータフォーマット調査

メタデータフォーマット調査

- OpenDOARで調査
- 収録コンテンツに論文を含む、機関リポジリ
に限定して調査
- コンテンツ数上位300機関リポジリ(日本除
く)
- 調査項目
 - 本文があることを確認できるか？
 - 出版社版のDOIは含まれているか？
 - 機械可読な掲載誌書誌情報(掲載雑誌のISSNや
巻号、ページ)を取得可能か？

調査結果

	機関 リポジトリ数	メタデータ フォーマット数
対象	256	676
本文あり確認可	205	447
DOI出力例あり	78	132
書誌情報取得可	48	65

- エラーを除く256機関リポジトリを調査対象
- メタデータフォーマットは**61種類**あった

**世界中の機関リポジトリから
AIRway対応メタデータを全て収集するのは現状困難**

5

AIRwayの今後

AIRwayの今後

- **本文の有無に関する情報や版情報、機械可読の掲載誌書誌情報をOAI-PMHにより取得可能な状態にしておくことによって、オープンアクセス文献の視認性は向上することを確認できた。**
- **ただし全てを網羅するにはグローバルなレベルでのメタデータの充実とそのメタデータフォーマットの標準化とが必要。**

今後

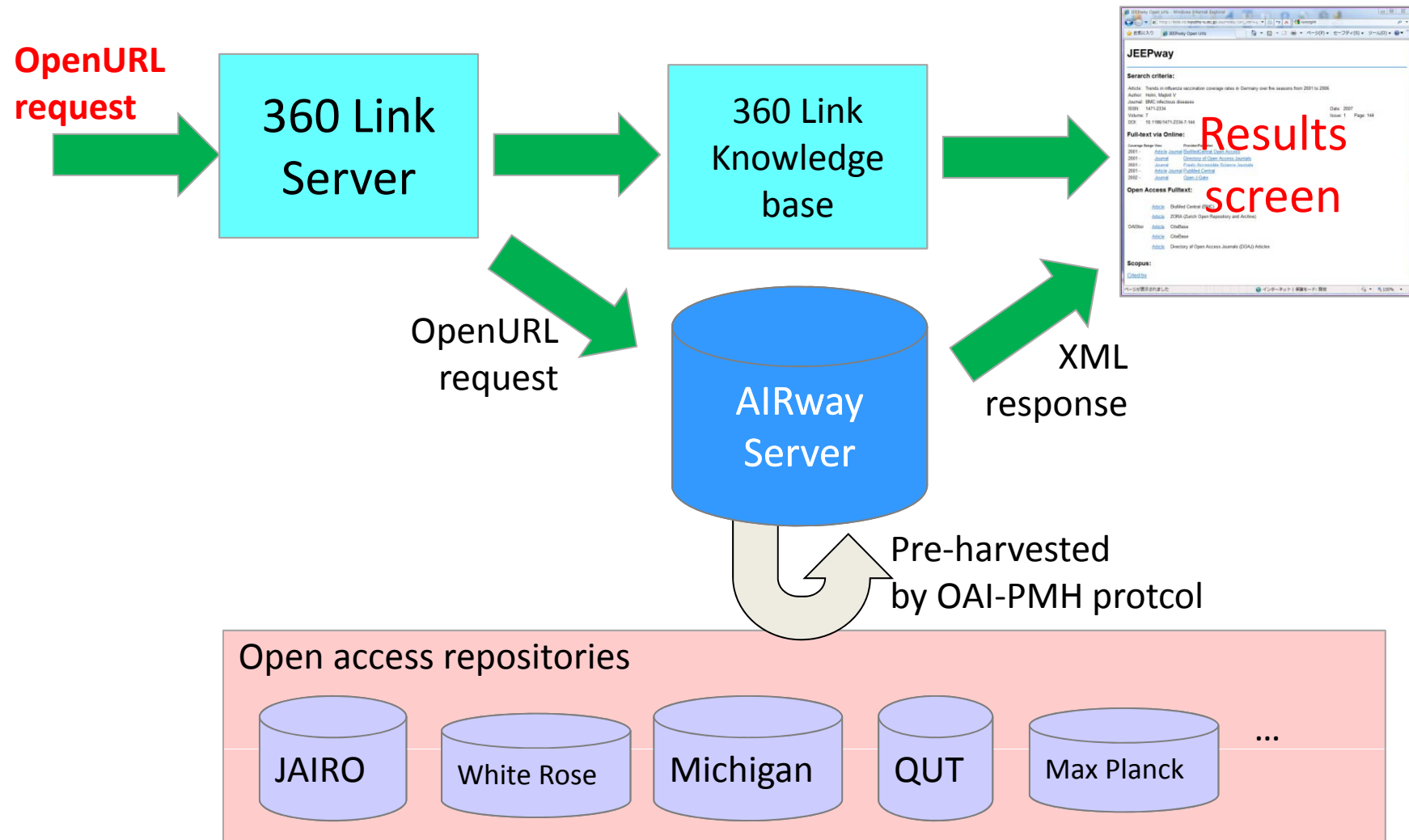
- 機関リポジトリ上のメタデータの充実、フォーマットの統一はかなり大変だし、これにコストをかけるのは正しいか？
- AIRwayの成果をステップとして別のアイデアが必要か？
 - JEEPway(九州大学で開発)
 - 識別子を使用した異版(著者版オープンアクセス論文)へのアクセス
 - 出版社版とは異なった識別子を付与する？
 - さらに版の情報も標準化し、各版との関係の記述方法も標準化され、それを集約したデータベースがあれば・・・？

JEEPway

- 同じくリンクリゾルバを使用し、機関リポジトリ上のオープンアクセス文献に対してより曖昧な情報から文献を特定するクエリー解決モジュール。
- 具体的にはリンクリゾルバが受け取った論文タイトル、雑誌タイトル、著者を使用し、オープンアクセス文献のサービスプロバイダ(JAIRO、OAlster、Scientific Commons)を検索、結果を取得、整形してリンクリゾルバ中間窓に表示する。

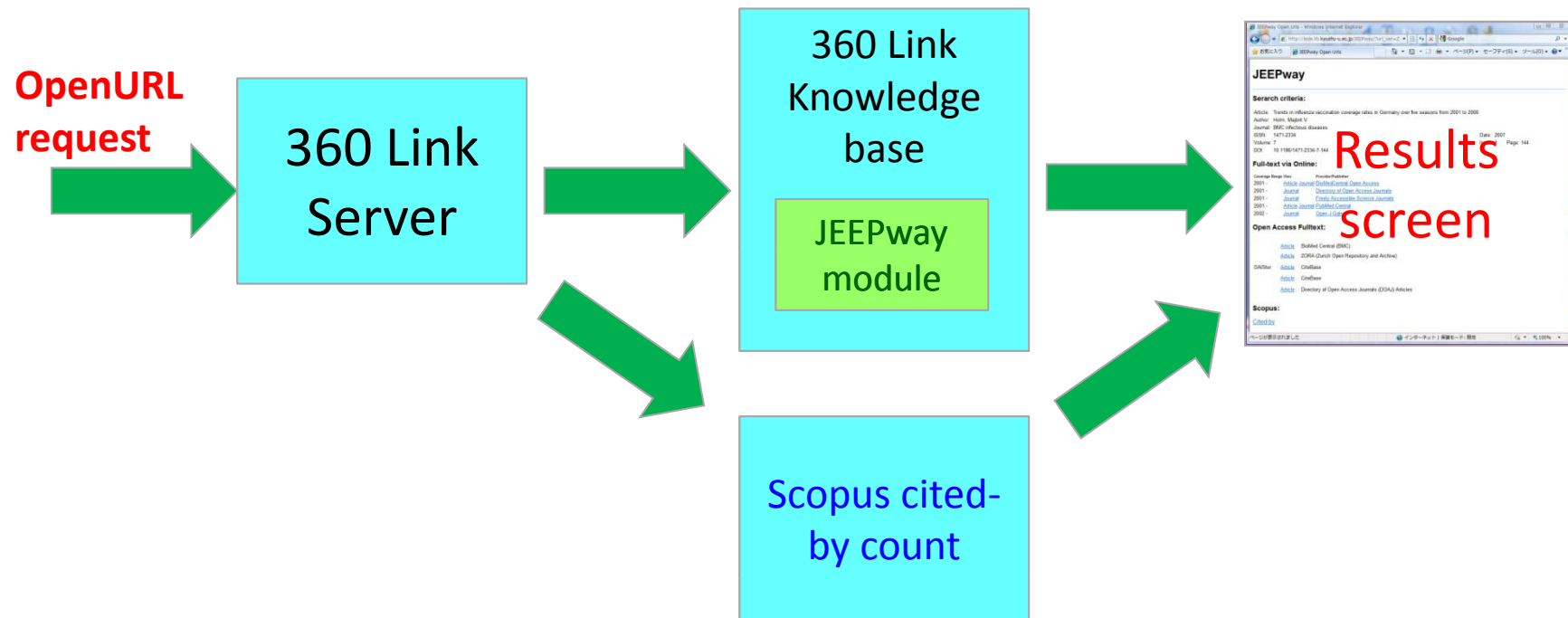
AIRwayとの比較(1/2)

AIRwayモデル



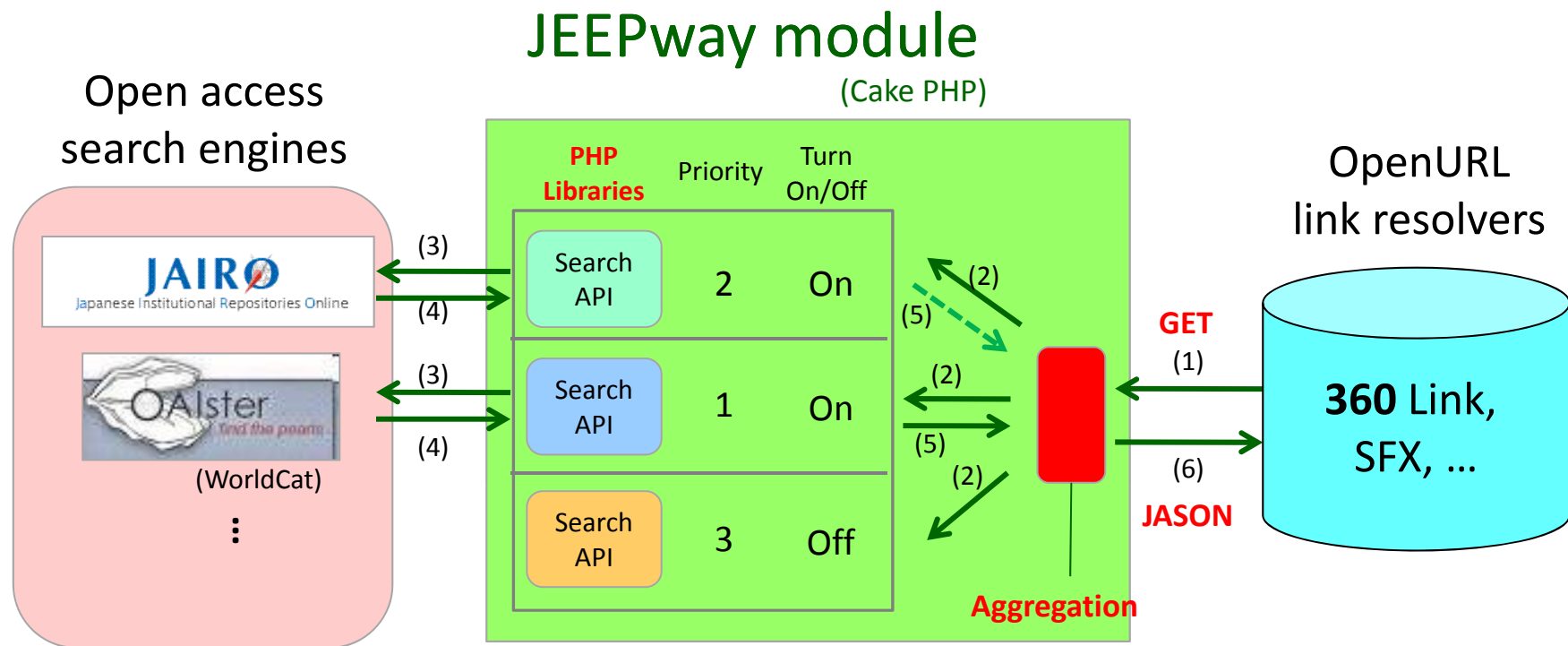
AIRwayとの比較(2/2)

JEEPwayモデル



- ハーベストは必要ない
- その都度サービスプロバイダを検索する。

JEEPwayモジュール



研究者の自然な情報探索行動の中に、機関リポジトリ上の文献発見を位置づけていくためにはどのような方法があるか？

**AIRwayは今後もこれを
追い求めていきたいと
思います！**