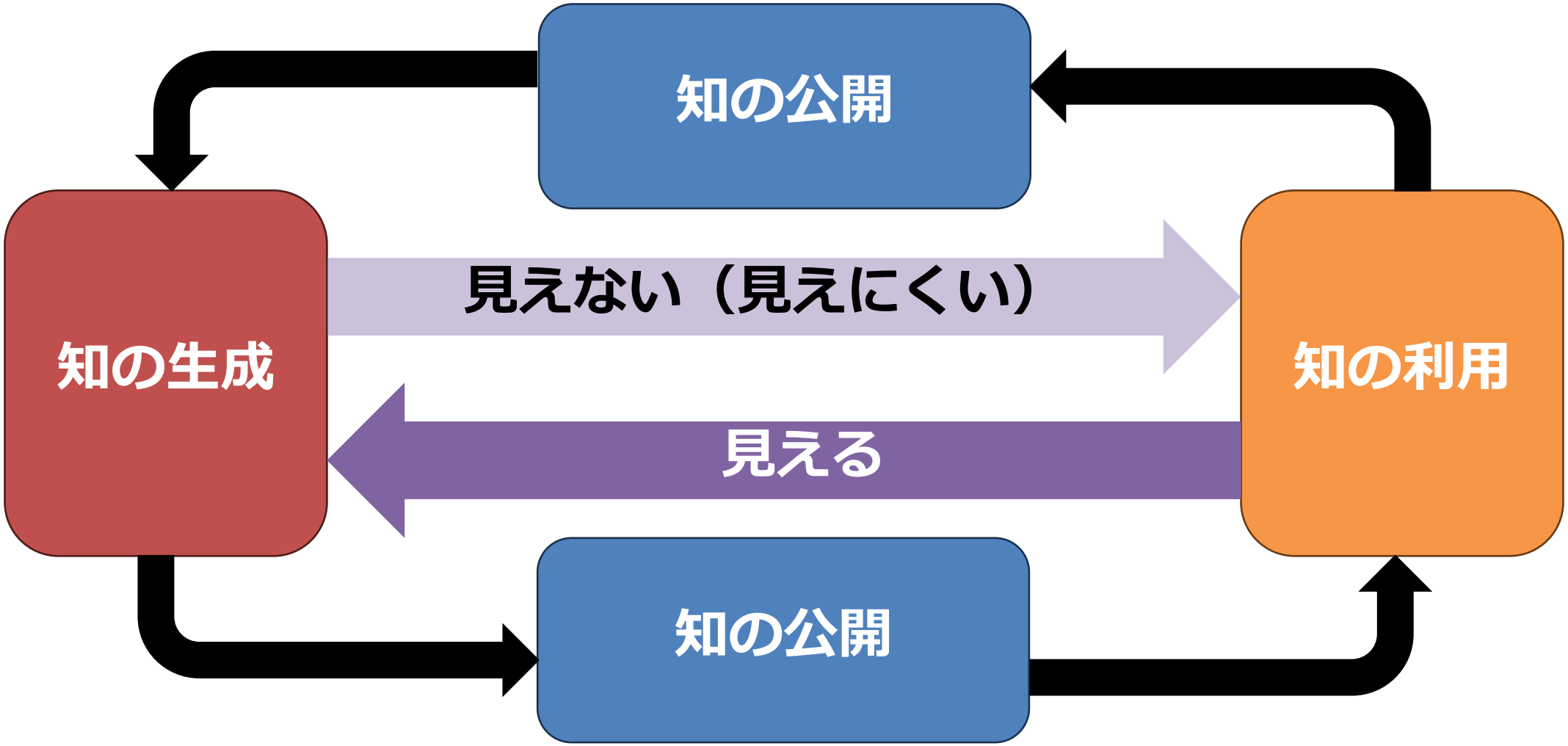
Mahalo Button：研究データ利用事例共有プラットフォーム

北本 朝展
国立情報学研究所

https://dias.ex.nii.ac.jp/

https://mahalo.ex.nii.ac.jp/

オープンサイエンスと知の循環



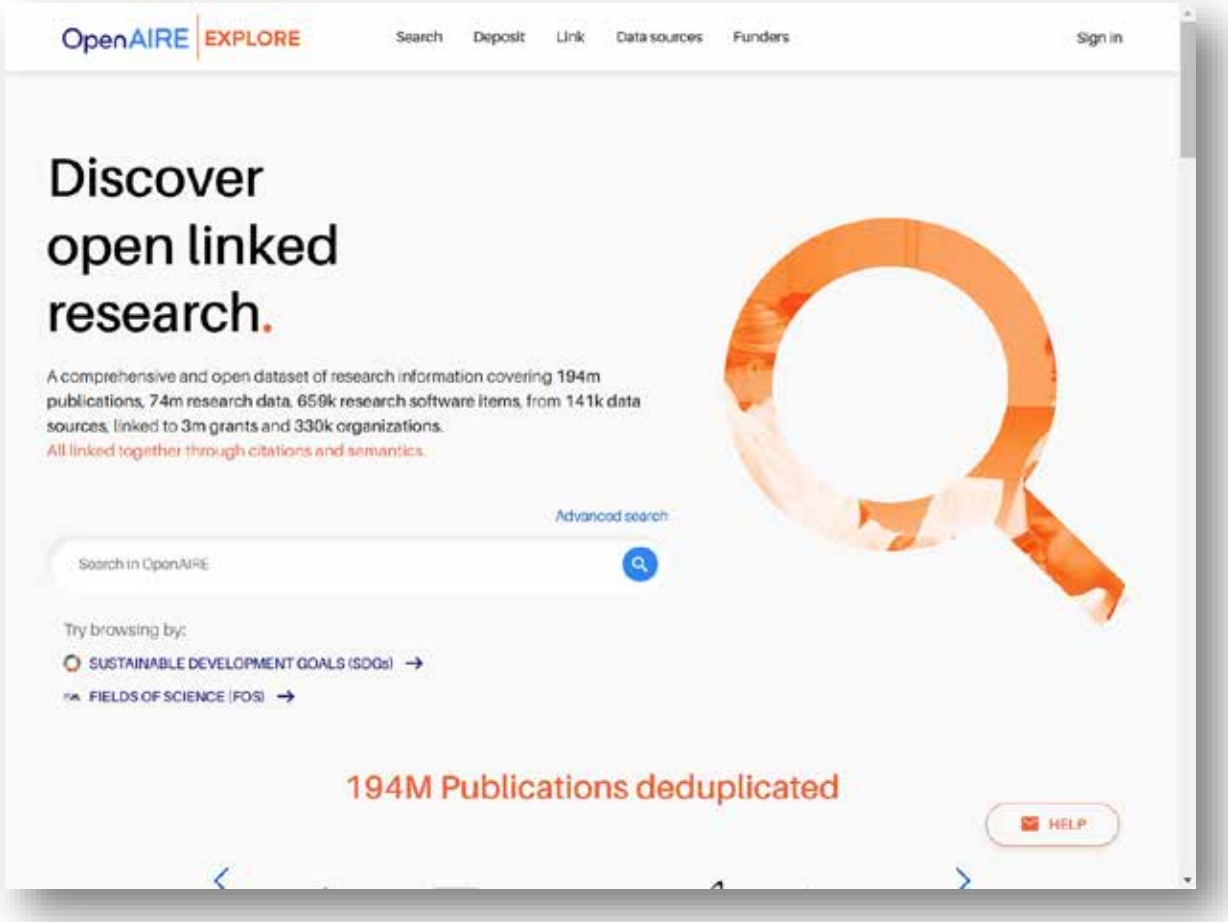
公開と利用のミッシングリンク

1. **知の生成⇒公開**：オープンサイエンスでよく議論される課題であり、インフラも充実
2. **知の公開⇒利用**：公開された知は、利用（FAIRのR）されてこそ価値が上がる
3. **知の利用⇒公開⇒生成**：知を利用する側から知を生成する側は見えているが、逆は見えない（見えにくい）
4. **ミッシングリンク**：生成側から利用側が見えれば、生成の価値を可視化し、利用方法を汎用化できる

知の循環の課題

1. 知を生成しても、適正に評価されない（特に**引用プラットフォームに乗せにくいデータセット**など）
2. 知を公開しても、どう利用されているかわからない（**生成者だけでなく公開者にとっても悩み**）
3. 知を利用しても、生成者や公開者にフィードバックできない（**改善の提案や感謝の伝達への障壁**）
4. 知の循環がつながらないと、**知のフライホイールを加速することができない**

Scientific Knowledge Graph



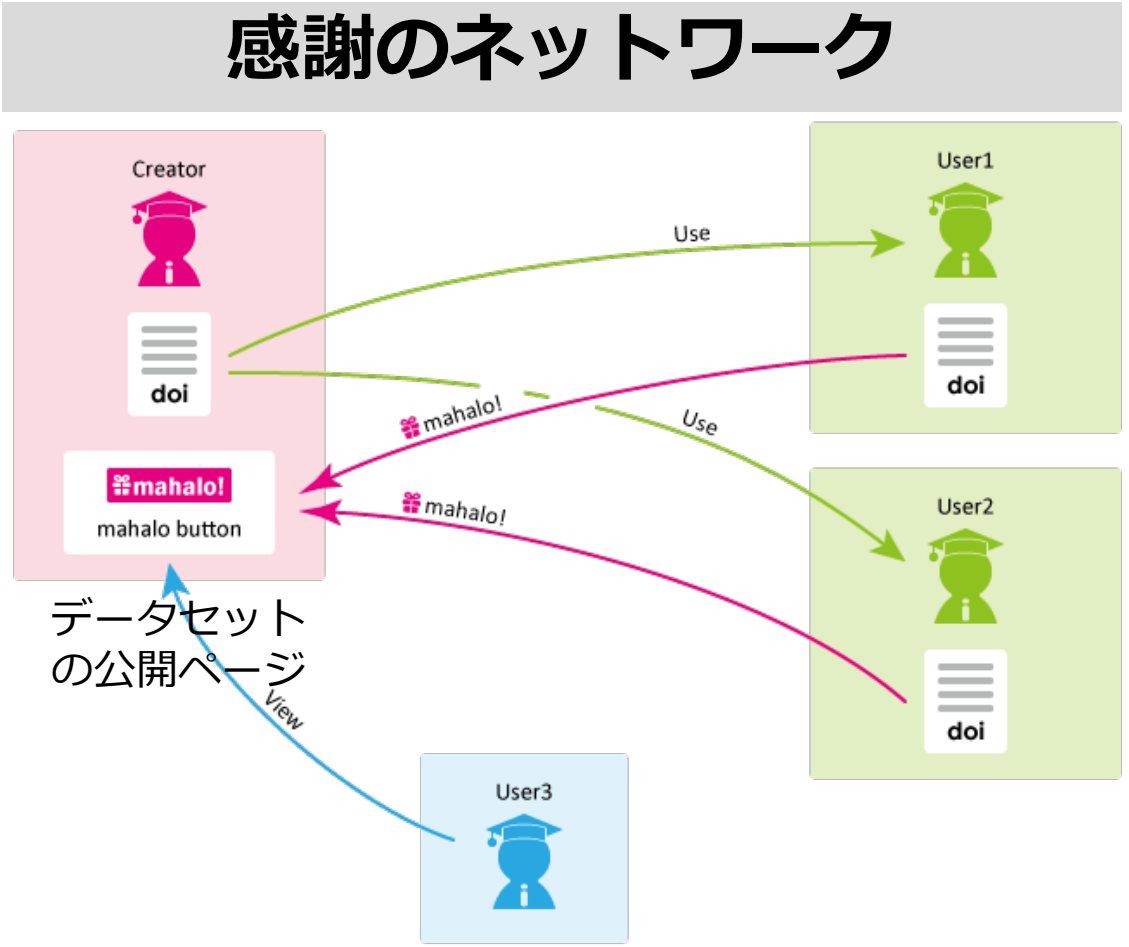
https://explore.openaire.eu/

1. 出版、データ、ソフトウェア、プロジェクト、組織など、**エンティティ**の関係をグラフ化
2. リソースの依存関係を見える化できるため、**研究評価**には有用
3. **利用事例**に関する詳細な情報は得られない

Mahalo Button

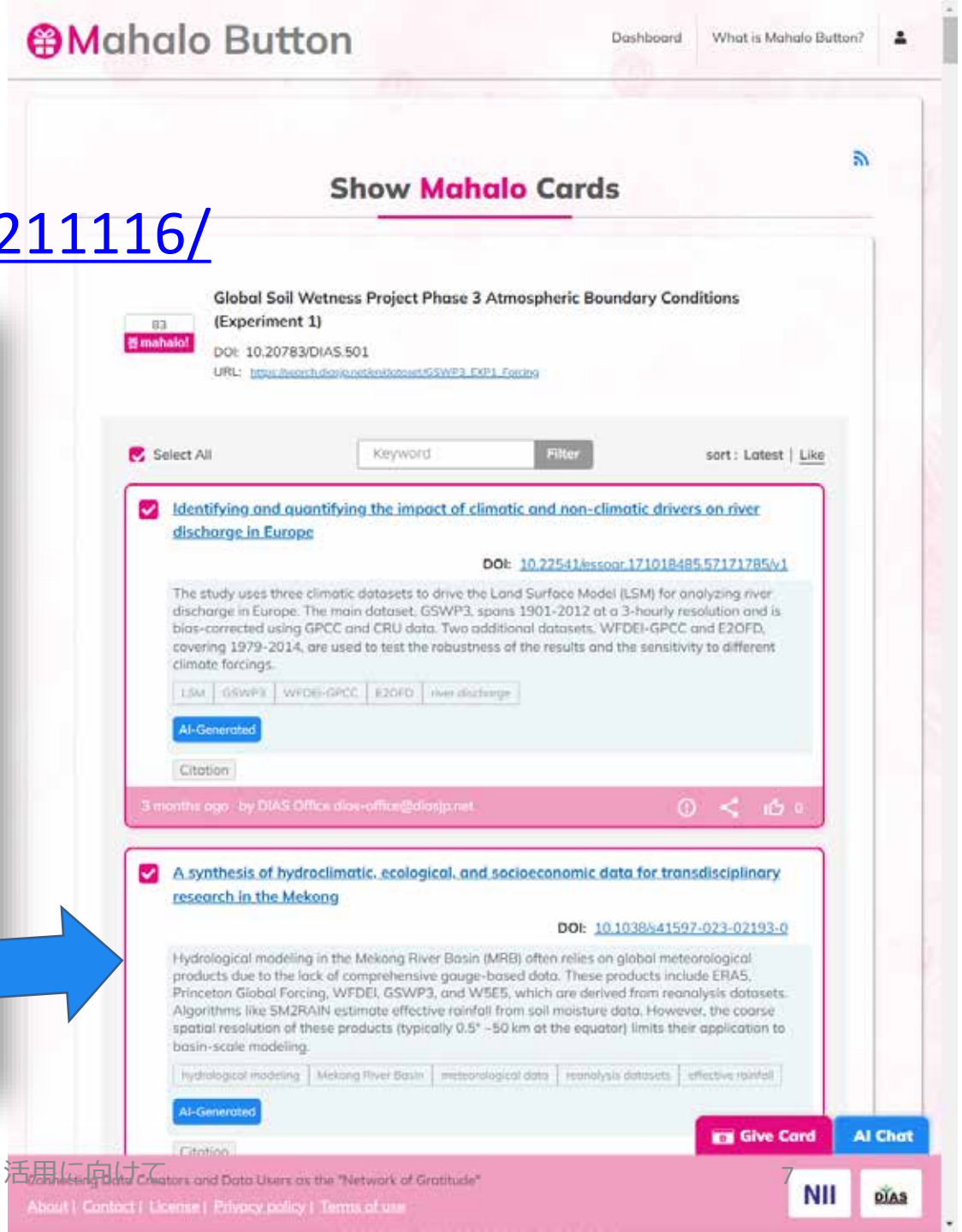
https://mahalo.ex.nii.ac.jp/

1. **Mahalo Button**＝データセット公開ページに設置するボタン
2. **データ利用者**や**データ公開（管理）者**：データ利用事例を登録
3. **潜在的データ利用者**：データ利用事例から学ぶことで、新たなデータ利用者となる
4. **ボタンで利用事例数を集計し、データ作成者の業績とする**



DIASとMahalo Button

https://diasjp.net/information/topics/20211116/



基本的な構成要素

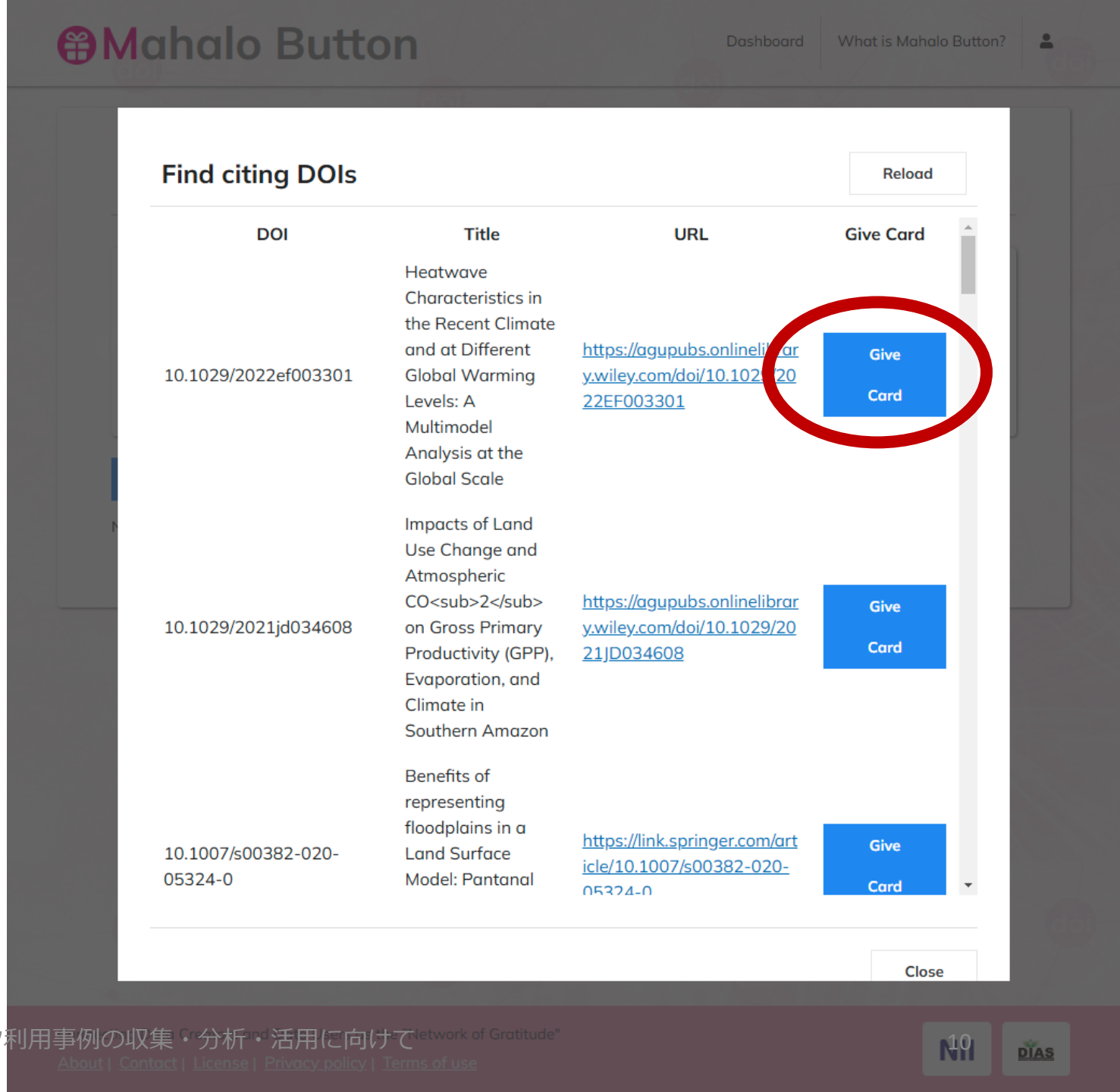
1. **Mahalo Button**＝**データに紐づく**：データ作成者（公開者）が作成し、データ公開ページに設置
2. **Mahalo Card**＝**データ利用事例に紐づく**：データ利用者が、Mahalo Buttonに紐づけて登録
3. **Mahalo Report**=**Mahalo Cardに紐づく**：Mahalo Cardに問題があればMahalo Button管理者に連絡
4. **アカウント**：一部の閲覧機能以外は、すべてログインして操作し、ダッシュボードで管理

Mahalo Card＝利用事例

1. データセットをどのように使ったかという詳細な記述（引用）を登録する
2. 単に文書（論文）へリンクするだけでなく、「引用の文脈」を収集し、より深い分析を可能とする
3. 生成AI機能を導入し、詳細な記述（テキスト）の管理・活用を支援する
4. Mahalo Cardの記述はオリジナル文書（論文）の発見支援であり、正確な内容は原典で確認する必要あり

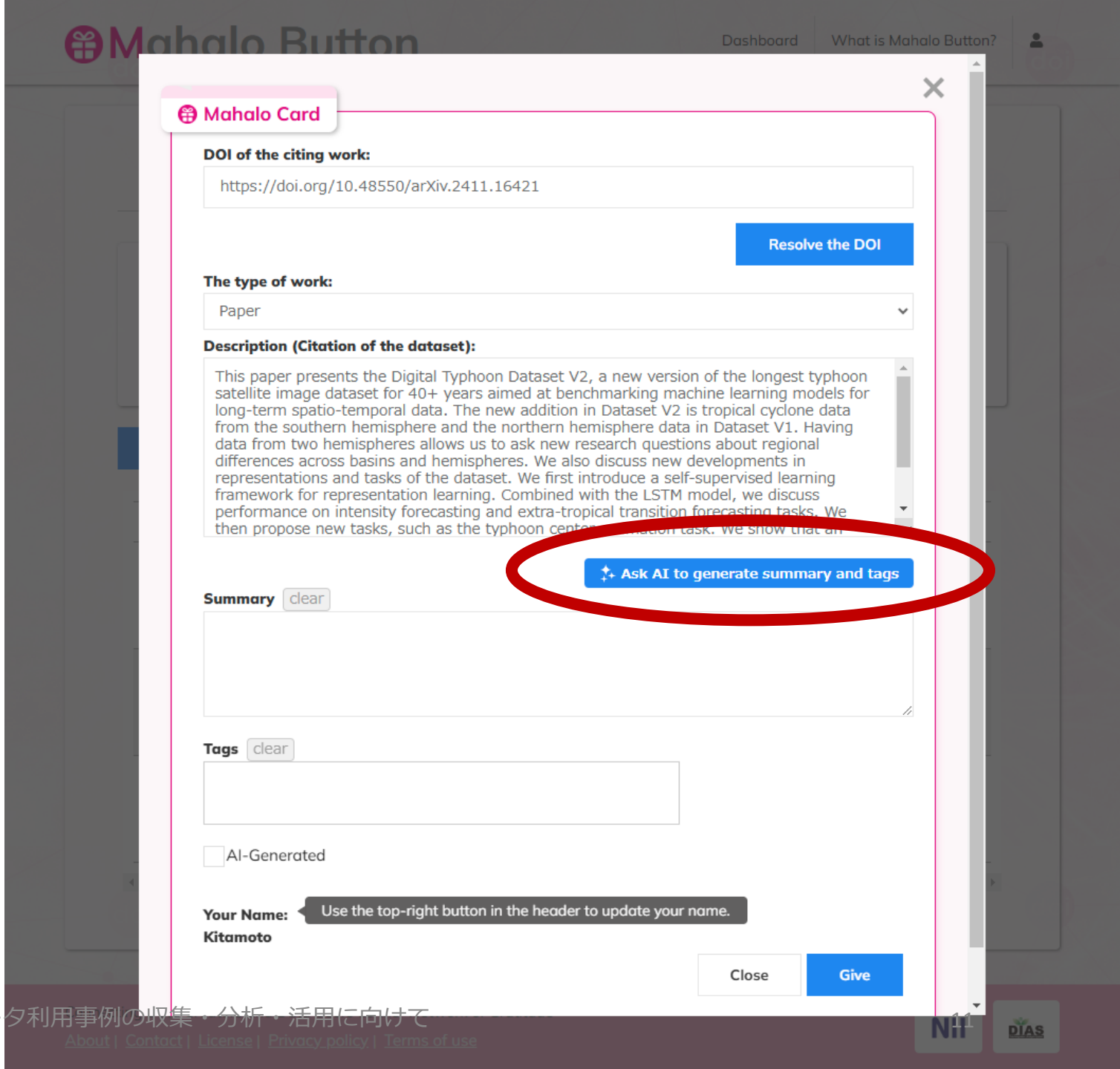
1. データ引用論文の検索支援

1. DataCite APIやOpenCitations APIを利用し、DOIをキーとしてデータ引用論文を検索
2. 「Give Card」をクリックすると、スムーズにデータ引用論文を登録可能



2. 生成AIによるデータ登録支援

1. TitleとDescriptionを元の論文から抜き出して、SummaryとTagsを自動生成
2. メタデータを、人手で整理せず、自動的に生成するため、作業負荷が大幅に減る



3. 生成AIによるデータ利用支援

1. Mahalo Cardに登録した利用事例に生成AIチャットが使える
2. プロンプトを変えると、様々な内容の情報が取得できる
3. データセットの潜在的利用者を支援する

データリポジトリの価値向上

1. 「感謝のネットワーク」のコンセプトは、データ利用者が自発的に利用事例を入力（感謝）する想定
2. 現在のDIASでは、データキュレーターが利用事例を検索して入力している
3. **データ公開者に対するサービスの一つ**＝リポジトリの付加価値機能と位置付ける
4. 将来的には、リポジトリの利用報告を高度化し、リポジトリの貢献度評価にも利用できないか？

オープンサイエンスと持続可能性

1. Mahalo Buttonは、**オープンサイエンスの実験プロジェクト**として、どなたでも利用できる
2. 利用規約やガイドラインを守って利用する
3. 不適切利用への対応方法も用意しているが、限られたリソースでの運用には限界もある
4. プロジェクト予算に依存しているため、長期的な運用は保証できていない
5. 入力したデータはダウンロードすれば長期保存可能

Mahalo Buttonの利点

https://dias.ex.nii.ac.jp/mahalo/

1. データ生成者：データの利用事例を把握し、貢献を可視化するサービスとして利用
2. データ公開（管理）者：利用事例に基づきデータに付加価値を与え、利用を促進するサービスとして利用
3. データ利用者：生成AIチャットの支援によりデータの利用方法を把握し、自らの研究に利用
4. システム開発者（DIAS）：Mahalo Buttonの開発を通じて、利用事例共有サービスのモデル化と普及