

DH ・ 技術要素 ・ RO-Crate

RO-Crate入門

研究データを「箱」にまとめる

DH入門 / 技術要素シリーズ

中村

※実験的な取り組みです（構成・図・AI音声合成を含む）。内容をご確認・ご注意のうえご利用ください

この動画について

- ✓ **オープンに公開された仕様・資料**を参照し、独自に構成した解説です
- ✓ スライド・図は新規作成、ナレーションは**AI音声合成**（この回は本人のクローン声ではありません）
- ✓ これは**実験的な取り組み**です。内容は**ご確認・ご注意のうえ**ご利用ください
- ✓ 誤りに気づいたら概要欄からご指摘ください。出典・ライセンスは末尾と概要欄に記載しています

この回のゴール

研究データを、説明ごと「箱」にまとめて持ち運ぶ考え方をつかむ

- ✓ **RO-Crate** が、研究データとメタデータをまとめて梱包する軽量な方法だと説明できる
- ✓ crate=ディレクトリで、中心に **ro-crate-metadata.json** がある、と説明できる
- ✓ メタデータが **schema.org 語彙の JSON-LD** で書かれる、とイメージできる
- ✓ **3種のエンティティ**（箱・データ本体・関係するもの）の役割の見当がつく

リンクトデータ・JSON-LD の考え方を知っていると分かりやすいですが、必須ではありません。

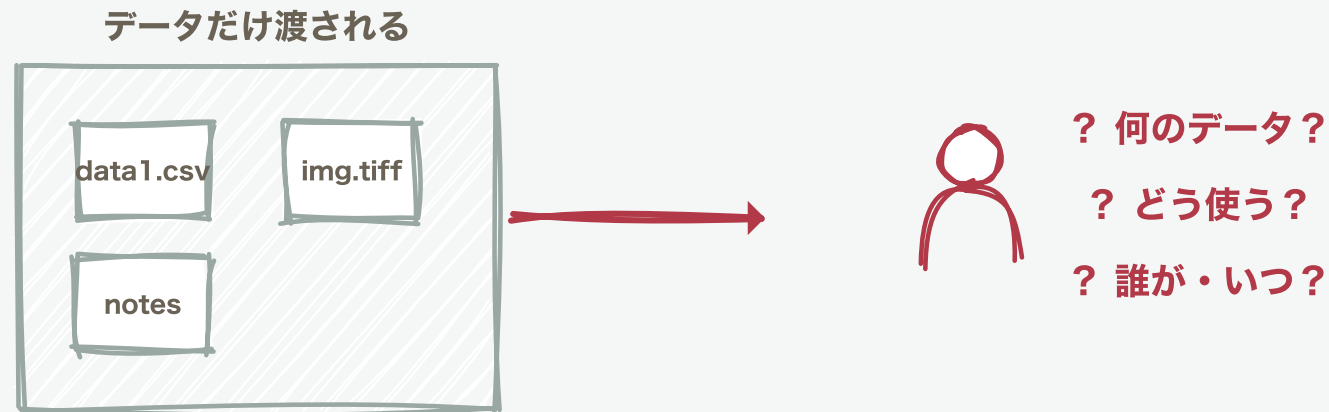
今日の流れ

- ✓ なぜデータを「箱」にまとめるのか
- ✓ crate のしくみ — ディレクトリと ro-crate-metadata.json
- ✓ メタデータの中身 — schema.org の JSON-LD と3種のエンティティ
- ✓ つないで広がる — 機械可読・FAIR・持ち運び
- ✓ 始め方・学ぶには

なぜ「箱」にまとめるのか

まずは、データだけ渡されたときの困りごとから

データだけでは、分からない



データだけでは、文脈が分からず、再利用しにくい

ファイルだけ渡されても、**何の・誰の・どう使う**データかが分からなければ、再利用は難しいものです

データと「説明」を、いっしょに

ひとつの「箱 (crate)」



データと、その「説明」を、いっしょに梱包する

データと、その**説明 (メタデータ)** を、ひとつの「**箱**」にまとめて持ち運べるようにする — それが
出発点です

RO-Crate — 研究データ梱包の軽量な方法

データ+メタデータを、まとめて持ち運ぶ

RO-Crate

研究データ梱包の
軽量な方法

コミュニティが策定 / 中身は schema.org の JSON-LD

RO-Crate (Research Object Crate) は、研究データとメタデータを**まとめて持ち運ぶ**軽量な方法。コミュニティが策定し、中身は **schema.org** の **JSON-LD** です

ここまでのポイント

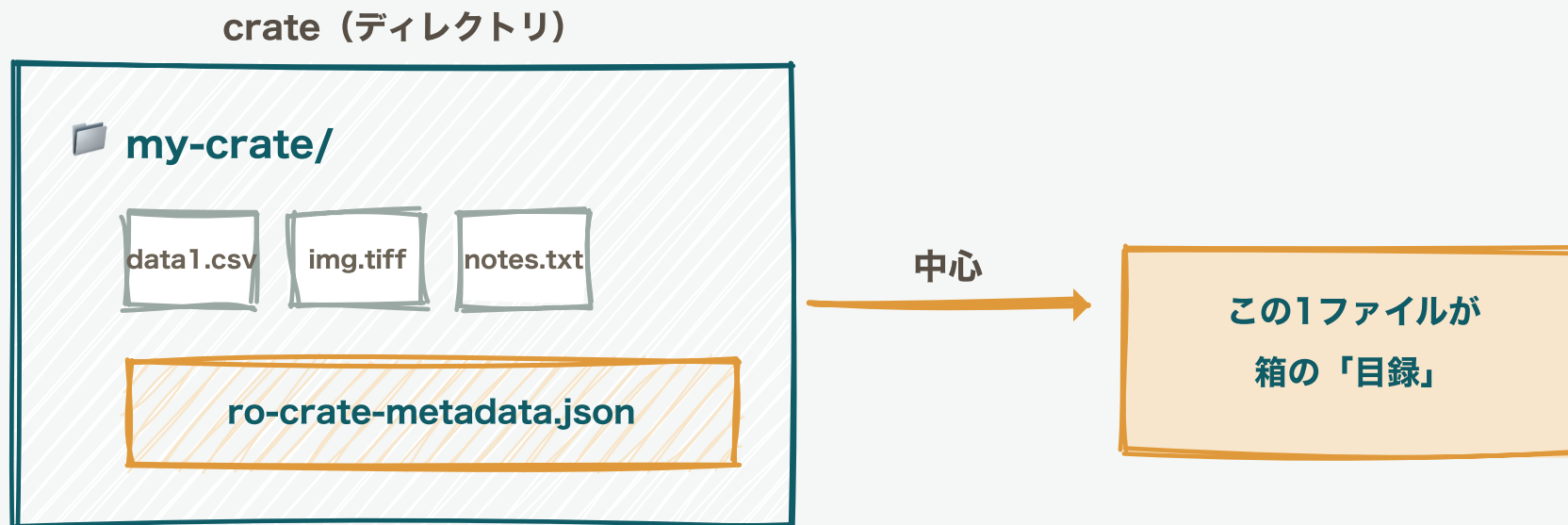
- ✓ データ**だけ**では、文脈が分からず再利用しにくい
- ✓ データと**説明（メタデータ）**を、ひとつの「箱」にまとめて持ち運ぶ
- ✓ それを軽量に実現する方法が **RO-Crate**（中身は schema.org の JSON-LD)

では、その「箱」は具体的にどんな形をしているのでしょうか。 → crate のしくみへ

crate のしくみ

「箱」の中身を見ていきます

crate = ディレクトリ + 目録ファイル



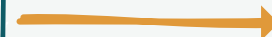
crateは、ふつうのフォルダ。中心に ro-crate-metadata.json を置く

crate は、ふつうの**フォルダ**。中心に **ro-crate-metadata.json** という「目録」を1枚置きます

目録は schema.org の JSON-LD

ro-crate-metadata.json

```
{
  "@context": "schema.org",
  "@graph": [
    {"@type": "Dataset", ...},
    {"@type": "Person", ...}
  ]
}
```



機械が読める
リンクトデータ

中身は schema.org 語彙の JSON-LD (リンクトデータの形)

目録の中身は **schema.org** 語彙の **JSON-LD**。リンクトデータの形なので、**機械が読めて**、つながりをたどれます

3種のエンティティ



箱・中身・関係者を、それぞれ記述する

目録は、**箱そのもの** (Root)、**データ本体** (Data)、**関係するもの** (Contextual : 人・組織・場所) を、それぞれ記述します

ここまでのポイント

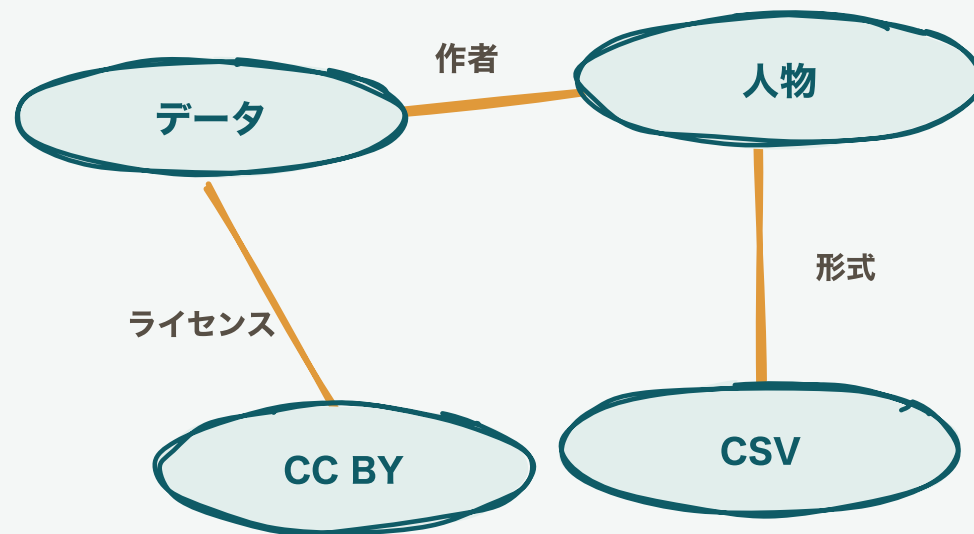
- ✓ crate は**ふつうのフォルダ**。中心に **ro-crate-metadata.json** を置く
- ✓ その目録は **schema.org** の **JSON-LD** = 機械が読めるリンクトデータ
- ✓ 目録は **箱・データ本体・関係するもの** (Root/Data/Contextual) を記述する
- ✓ 特別なDBもサーバも要らず、フォルダに1枚足すだけで始められる

この「箱」が、どう役立つのでしょうか。 → 機械可読・FAIR・持ち運びへ

つないで広がる

機械可読・FAIR・持ち運び

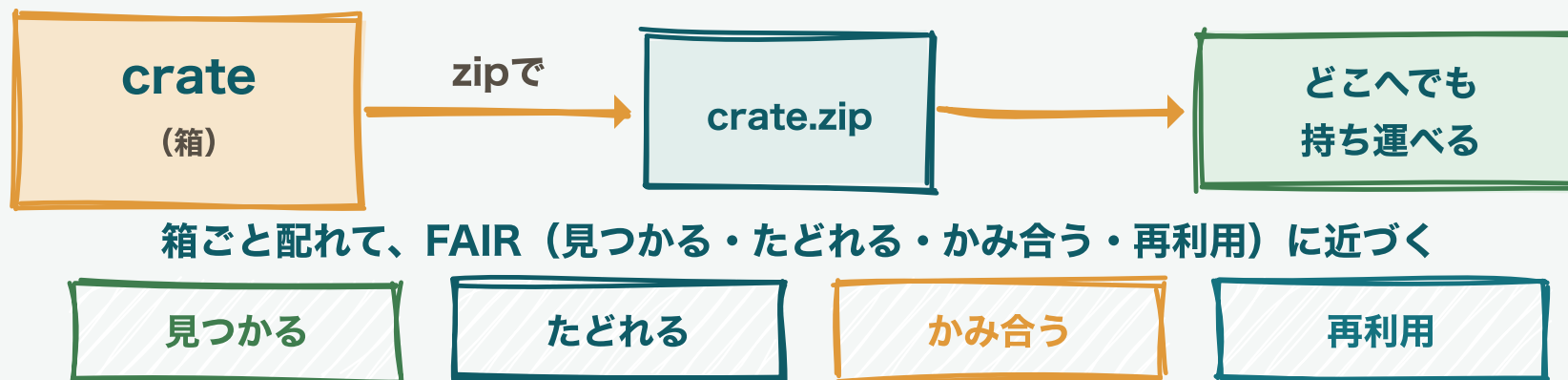
機械が「つながり」をたどれる



schema.org の語彙で、関係を機械が「つながり」としてたどれる

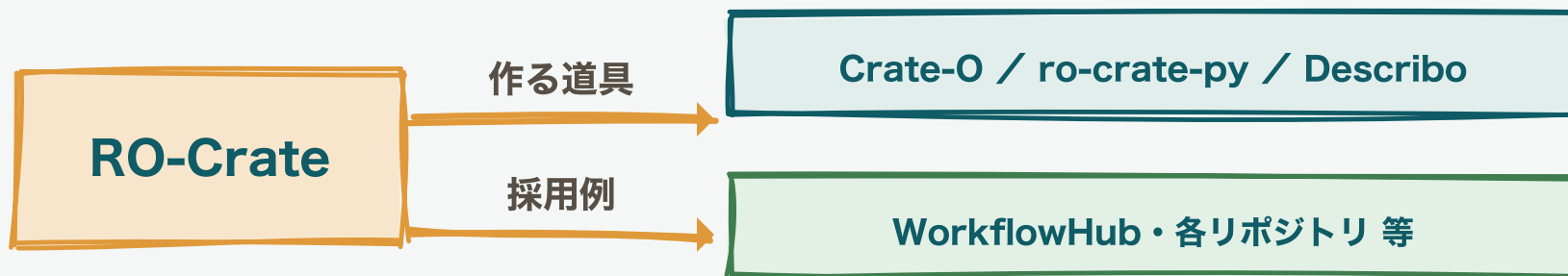
schema.org の語彙で書くから、データと、作者・ライセンス・形式などの**つながり**を、機械が、たどれます

箱ごと持ち運べて、FAIRに近づく



crate は **zip** で固めて、どこへでも持ち運べます。説明ごと運ぶので、**FAIR** (見つかる・たどれる・かみ合う・再利用) に近づきます

作る道具と、広がる採用



作る道具がそろい、研究基盤での採用も広がっている

手で書かなくても、**Crate-O** や **ro-crate-py** などの道具があります。**WorkflowHub** をはじめ、研究基盤での採用も広がっています

ここまでのポイント

- ✓ **schema.org** の **JSON-LD** なので、機械がつながりをたどれる
- ✓ crate は **zip** で固めて持ち運べ、説明ごと運ぶので **FAIR** に近づく
- ✓ **Crate-O**・**ro-crate-py** などの道具があり、**WorkflowHub** 等で採用が広がる

仕組みが見えたところで、少し立ち止まって考えてみましょう。

考えてみよう

あなたのデータを箱にするなら、**どんな説明**を添えますか？

何の・誰の・いつの・どう使うデータか。ここで少し、動画を止めて考えてみてください。

何を「説明」として残すか

- ✓ どんな文脈を、どこまで記述するかは、機械的には決まりません
- ✓ 将来の**誰が・何のために**使うかを想像して、残す説明を選びます
- ✓ 記述すべきことを見極める作業は、データを**どう理解してほしいか**を考えることでもあります

梱包は単なる作業ではなく、データの意味を**未来へ引き継ぐ**営みです

始め方・学ぶには

- ✓ まず作る：**Crate-O**（ブラウザのエディタ）で、手元のフォルダを crate にしてみる
- ✓ プログラムで：**ro-crate-py** などで、メタデータを生成する
- ✓ 体系的に：**RO-Crate 公式仕様**（researchobject.org）と、用例集
- ✓ 確かめる：できた **ro-crate-metadata.json** を開いて、3種のエンティティを眺めてみる

まずは「フォルダに目録を1枚足す」ところから。難しい道具立ては要りません

まとめ

- ✓ **RO-Crate** は、研究データとメタデータを**まとめて梱包**する軽量な方法
- ✓ crate=ディレクトリ、中心に **ro-crate-metadata.json** (schema.org の JSON-LD)
- ✓ 目録は **箱・データ本体・関係するもの** (Root/Data/Contextual) を記述
- ✓ 機械が読めて、**zip**で持ち運べ、**FAIR** に近づく。道具・採用も広がる

データに「説明」を添えて、ひとつの箱に。意味ごと未来へ手渡すための、軽い作法です

出典・ライセンス

本動画は、以下のオープンに公開された仕様・資料を参照して作成しました。

- ✓ RO-Crate Metadata Specification (researchobject.org/ro-crate) / RO-Crate community
- ✓ 事実確認 : schema.org / [JSON-LD \(W3C\)](https://www.w3.org/TR/json-ld/) / 採用事例 (WorkflowHub 等)

スライド・図は中村による新規作成 (概念を参照し、表現は新たに構成)。

ご清聴ありがとうございました