

RO-Crateを用いた材料研究データの データリポジトリへの登録



国立研究開発法人
物質・材料研究機構
National Institute for Materials Science



田辺浩介 <https://orcid.org/0000-0002-9986-7223> TANABE.Kosuke@nims.go.jp

松田朝彦 <https://orcid.org/0000-0001-5989-027X> MATSUDA.Asahiko@nims.go.jp

物質・材料研究機構 統合型材料開発・情報基盤部門 材料データプラットフォームセンター

材料データのデータリポジトリ登録の課題

- 複数の分野にまたがっており、データの形式や内容が多種多様
- データに付与するメタデータスキーマの設計や記述が困難
- 複数のデータファイルを、データ生成時のディレクトリ構造をそのまま保存・共有する必要がある
- データ駆動型研究を実現するために、登録・公開処理の自動化が必要

研究データパッケージング
フォーマット

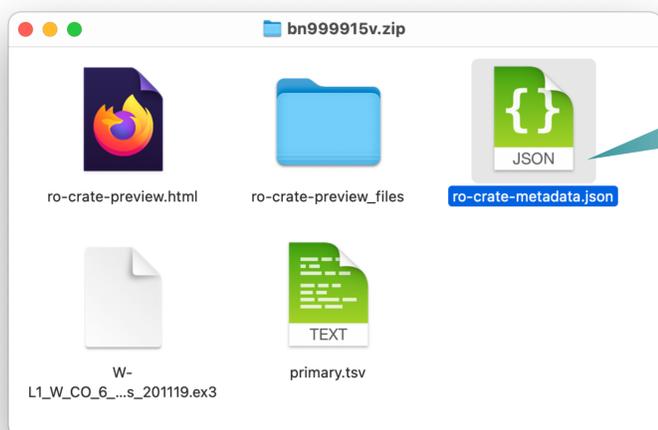
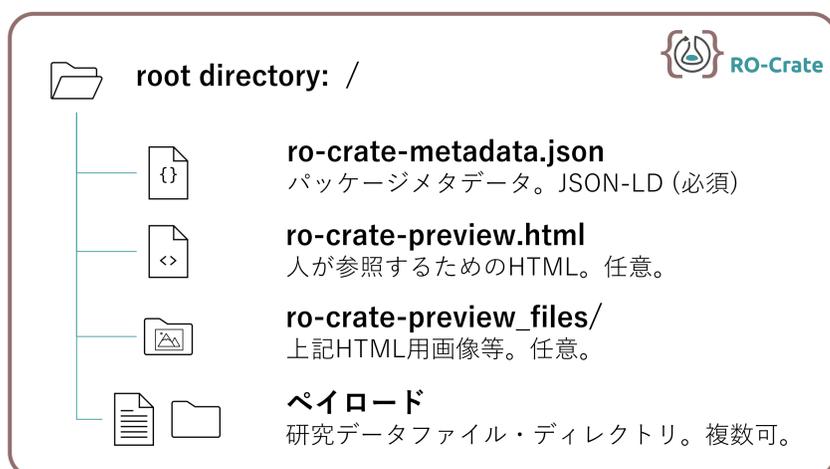


RO-Crate を選択

- <https://www.researchobject.org/ro-crate/>
- Research Object communityによる開発
- 2019年にv1.0, 2021年にv1.1 リリース

パッケージの中身と実例

(ZIPファイルでもディレクトリでも可)



JSON-LD で書き、Bioschemas と Material Schemas の語彙を利用

```
{
  "@context": [ "https://w3id.org/ro/crate/1.1/context", { "bio": "http://schema.org" } ],
  "@graph": [
    {
      "@type": "CreativeWork",
      "@id": "ro-crate-metadata.json",
      "conformsTo": { "@id": "https://w3id.org/ro/crate/1.1" },
      "about": { "@id": "." }
    },
    {
      "@id": ".",
      "@type": "Dataset",
      "url": "https://mdr.nims.go.jp/concern/datasets/bn999915v",
      "identifier": "https://doi.org/10.48505/nims.2563",
      "name": "XAFS spectrum of Tungsten carbonyl",
      "datePublished": "2021",
      "description": "This dataset consists of X-ray absorption fine structure (XAFS) spectra at W L21-edge of Tungsten carbonyl measured at SPring-8 BL14B2",
      "hasPart": [
        { "@id": "W-L1_W(CO)6_Si311_50ms_201119.ex3" },
        { "@id": "primary.tsv" }
      ],
      "material": [ { "@id": "#material-0" } ]
    },
    {
      "@id": "#material-0",
      "@type": "Material",
      "name": "Tungsten carbonyl",
      "description": "Standard Sample",
      "bio:ChemicalSubstance": {
        "bio:chemicalComposition": "W(CO)6",
        "identifier": "https://matvoc.nims.go.jp/wiki/Item:Q1765"
      }
    },
    {
      "@id": "W-L1_W(CO)6_Si311_50ms_201119.ex3",
      "@type": "File",
      "encodingFormat": "text/plain",
      "url": "https://mdr.nims.go.jp/downloads/73666690b",
      "mainEntityOfPage": { "@id": "https://mdr.nims.go.jp/concern/parent/bn999915v/file_sets/73666690b" }
    },
    {
      "@id": "primary.tsv",
      "@type": "File",
      "encodingFormat": "text/tab-separated-values",
      "url": "https://mdr.nims.go.jp/downloads/k930c032p",
      "mainEntityOfPage": { "@id": "https://mdr.nims.go.jp/concern/parent/bn999915v/file_sets/k930c032p" }
    }
  ]
}
```

パッケージの情報

データセット全体の情報

試料・対象物質

材料科学固有のメタデータ

ペイロード情報 1

実験データファイル

ペイロード情報 2

実験独自のメタデータ

RO-Crateによる研究データパッケージングの利点

- データ作成時の文脈（ディレクトリ構造等）をそのまま保存
- RDFによる柔軟なメタデータ記述
- HTMLプレビューファイルで人間向けの可読性も確保
- メタデータのファイル名が“ro-crate-metadata.json”固定なので、アプリケーションでのメタデータ読み取り実装が容易
- 研究コミュニティによるオープンな規格、かつ静的ファイルなのでアーカイブ化・長期保存にも有利

簡潔な自動データ登録と公開による
データ駆動型研究への貢献

材料科学分野での適用を推進中
あなたの分野でもぜひ！

