



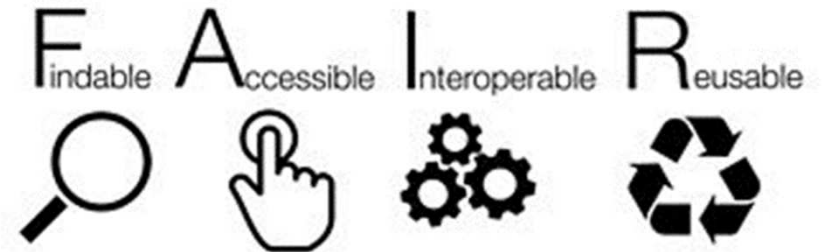
J-STAGE Dataでの FAIR原則に沿った 研究データ公開

2024年1月

FAIR原則とはなにか？

研究データの公開にあたり、国際的に広く認められ
求められている **データ共有の原則**

- **F**indable (見つけられる)
- **A**ccessible (アクセスできる)
- **I**nteroperable (相互運用できる)
- **R**eusable (再利用できる)



FORCE11: THE FAIR DATA PRINCIPLES (2016).
<https://www.force11.org/group/fairgroup/fairprinciples>, NBDC研究チーム(訳), "FAIR原則 (「THE FAIR DATA PRINCIPLES」和訳)" (2019).
<https://doi.org/10.18908/a.2019112601>

G7広島首脳コミュニケ (2023年5月) でも言及

「**G7は、FAIR原則に沿って**、科学的知識並びに研究データ及び学術出版物を含む公的資金による研究成果の公平な普及による、**オープン・サイエンスを推進**する」

J-STAGE DataはFAIR原則を達成可能

- データ自体が検索できる
- 元論文と関連づいている
- データ自体に独自のDOIがある
- データのメタデータがある
- DOIでデータにアクセスできる
- 利用ライセンスが明示されている

電子付録（Supplementary Material）ではFAIR原則は達成できない

FAIR 15の原則

To be Findable: (見つけられるために)	To be Accessible: (アクセスできるために)
F1. (メタ) データが、グローバルに一意で永続的な識別子 (ID) を有すること。	A1. 標準化された通信プロトコルを使って、(メタ) データを識別子 (ID) により入手できること。
F2. データがメタデータによって十分に記述されていること。	A1.1 そのプロトコルは公開されており、無料で、実装に制限が無いこと。
F3. (メタ) データが検索可能なリソースとして、登録もしくはインデックス化されていること。	A1.2 そのプロトコルは必要な場合は、認証や権限付与の方法を提供できること。
F4. メタデータが、データの識別子 (ID) を明記していること。	A2. データが利用不可能となったとしても、メタデータにはアクセスできること。
To be Interoperable: (相互運用できるために)	To be Re-usable: (再利用できるために)
I1. (メタ) データの知識表現のため、形式が定まっていて、到達可能であり、共有されていて、広く適用可能な記述言語を使うこと。	R1. メタ (データ) が、正確な関連属性を豊富に持つこと。
I2. (メタ) データがFAIR原則に従う語彙を使っていること。	R1.1 (メタ) データが、明確でアクセス可能なデータ利用ライセンスと共に公開されていること。
I3. (メタ) データは、他の (メタ) データへの特定可能な参照情報を含んでいること。	R1.2 (メタ) データが、その来歴と繋がっていること。
	R1.3 (メタ) データが、分野ごとのコミュニティの標準を満たすこと。

J-STAGE DataでのFAIR原則対応

	FAIR原則に沿ったデータ	FAIR原則に沿わないデータ
ファイル	数値データのファイルは、人間の可読性だけでなく、機械判読可能になっており、閲覧者がそのままダウンロードして、自身の環境で再検証したり、別の解析に用いたりできる。	<ul style="list-style-type: none">・ 二次利用できない、しにくい・ 数値データが機械判読不能
メタデータ	検索されやすい端的なTitleで、Descriptionによって、どんなデータであるのか、各ファイルの見方、二次利用方法が詳細に説明されている。	<ul style="list-style-type: none">・ 記述が不十分でデータの内容が理解できない。・ 元論文に依存した内容が記載されている。・ データそのものだけでなく、解釈、結論、考察が含まれる。

ファイル

- ふさわしくない例
 - 数値データをPDFファイルで公開
 - 特定のソフトウェアやハードウェアに依存した形式
- 望ましい
 - 特定のソフトウェアに依存せずに再利用できる形式
 - CSV、TSV、JSON、XMLなど
 - 機械判読可能な表記（後述）
 - READMEファイルの配置
 - 複数ファイルの場合ファイル命名規則で違いを明確に

ファイル –機械判読可能な表記–

- 表タイトル、注釈は含めない（Descriptionに記載する）
- 1セル1データ
- 数値データは数値属性にし、単位・区切り・注釈記号などの文字列を入れない
- セル結合しない
- スペースや改行等で体裁を整えない

参考：統計表における機械判読可能なデータ作成に関する表記方法

https://www.soumu.go.jp/main_content/000723626.pdf

表データの公開にあたっては、かならず参照し従ってください

FAIR原則に沿ったメタデータ

Title*	検索されやすく、このデータを端的に示すタイトル
Authors*	このデータの著者（元論文著者と異なる場合もある）
Categories*	分野
Journal*	ジャーナル名
Item type*	データ種類
Keywords*	このデータのキーワード
Description*	元論文を見なくてもわかる詳細な説明や再利用に関する内容
Funding	このデータが関連する助成金情報
Resource title*	元論文のタイトル
Resource DOI*	元論文のDOI
License*	このデータの二次利用に関するライセンス
References	このデータが関連する文献のURL、DOI
Corresponding author email address*	このデータの連絡先著者メールアドレス
Title(in Japanese)	Titleの日本語表記
Description(in Japanese)	Descriptionの日本語表記
Manuscript title(in Japanese)	Resource titleの日本語表記
Authors(in Japanese)	Authorsの日本語表記
Copyright*	発行年、著作権者を明示

**元論文に依存しない
そのデータに関する内容
であることが重要**

メタデータの付与にて考えるべきこと

データは、論文とは独立して流通する

- データだけを見て、どんなデータであるか、その内容がわかるようにメタデータを付与する

たとえば、Google検索において適切なクエリを投げればそのデータがヒットするように考える。また、そのデータがヒットしたとき検索をした人にどんなデータであるかすぐわかるように考える

(論文を参照しないとその内容がわからないようでは不可)

- データも論文と同様・同等と考える

(論文は独立に流通しておりデータを参照せずともその内容はわかる。ここで、論文とデータを入れ替えても同じことが成り立つように)

- データの解釈や考察は「論文」で、データのメタデータには含めない

メタデータのふさわしくない例①

- Title
 - 「(ジャーナル名、巻号ページ) のデータ」
→ 見つけれにくく独立して流通する意味をなさない。
 - 論文と同じタイトル
→ 同じ場合もあり得るが、データを端的に示す内容を書く。
- Description
 - 論文の抄録と同じ内容
→ 論文の説明ではなくデータ自体の詳細説明、再利用に関する内容を書く。
- Funding
 - 論文には記載されているのに未入力
→ Dimensionsと連携すればアクセスされやすくなる。

メタデータのふさわしくない例②

- (in Japanese)項目
 - 和文誌なのに未入力
 - 英語だけより見つけられやすくなる。
 - 特に和文誌ではManuscript title(in Japanese)は、英語のResource titleより正式な記事タイトルに相当する。
- Title (in Japanese)
 - Title (英語) と内容が一致していない
- Description (in Japanese)
 - Description (英語) と内容が一致していない

メタデータのふさわしくない例サマリ

メタデータ項目	ふさわしくない例
Title	「(ジャーナル名、巻号ページ)のデータ」
	論文と同じタイトル
Description	論文の抄録と同じ内容
Funding	論文には記載されているのに未入力
Title (in Japanese)	和文誌なのに未入力
	Title (英語) と内容が一致していない
Description (in Japanese)	和文誌なのに未入力
	Description (英語) と内容が一致していない
Manuscript title (in Japanese)	和文誌なのに未入力 (Resource titleよりこちらが正式記事タイトルに相当)
Authors (in Japanese)	和文誌なのに未入力
Copyright	「© 公開年 著作権者」の形式になっていない

J-STAGE Dataの目指すところ

- J-STAGE掲載誌の論文付随データ公開の場
研究の透明性また論文の信頼性を担保するため、論文に付随するデータ、とくに論文根拠データを公開する
ジャーナル編集委員会によって査読・確認されたデータのみを公開する
- 研究データの二次利用推進
二次利用が期待される・想定される研究データを、**第三者による二次利用が可能**なよう、**FAIR原則に沿って公開**する
たんなるデータ置き場にはしない
- 研究データ公開およびデータ産生者の評価
今後は、研究データを公開すること、それが引用・二次利用されることが研究者としての業績となっていく
論文の著者だけでなく、データの著者、つまりデータ産生者の評価にもつながっていくことを期待