

付録

メタデータ比較検討用コンセプトペーパー

2021年8月

目次

1 背景と目的	3
1.1 背景	3
1.2 目的	3
1.3 メタデータとは何か	3
2 メタデータ項目の定義	4
2.1 DCAT の概要.....	4
2.2 DCAT シリーズ.....	6
1) BregDCAT	6
2) GeoDCAT	6
3) StatDCAT	7
2.3 ベース・レジストリのメタデータ項目	7
1) カタログ	8
2) カタログレコード	10
3) データセット	11
4) データサービス	15
5) ディストリビューション	16
2.4 メタデータの記載例	19

1 背景と目的

1.1 背景

令和2年12月25日に閣議決定された「デジタル・ガバメント実行計画」において、行政手続のワンスオンリーを実現するなど社会全体の効率性の向上を図るとともに、スマートシティなどの新しいサービスの創出を図るため、ベース・レジストリ¹の整備が喫緊の課題として挙げられています。また、「ベース・レジストリ・ロードマップ²」では、ベース・レジストリで管理するデータを行政機関等が検索、利用できるようにするため、レジストリ・カタログを構築することとしています。また、レジストリ・カタログの検索性や管理を高度化するため、ベース・レジストリに登録するデータセットにはメタデータ³を付与することが求められています。

これまでもオープンデータカタログや国立国会図書館の納本用メタデータ等のメタデータの活用事例はありましたが、今後は、一般のコンテンツ、オープンデータ、ベース・レジストリ等をシームレスにつなげるメタデータ管理が求められてきます。

1.2 目的

データセットに付与するメタデータは、ベース・レジストリやオープンデータ、コンテンツマーケットのカタログとの間でシームレスな検索が可能な仕組みとするため、国際標準に準拠することが重要になります。

本ペーパーは、メタデータの国際的な標準である DCAT (Data Catalog Vocabulary) に準拠したメタデータ項目を示すことで、コンテンツの発見性を高め、そのデータ連携や活用を容易にすることを実現するための基軸となる比較検証用メタデータを提示し、ディスカッションペーパーの作業を円滑に進めることを目的とします。

1.3 メタデータとは何か

メタデータとは、データに関するデータであり、データを検索する際などに使用します。データ関連データとしてデータに付帯して整備され、データカタログで検索に使用されます。対象地域、期間や品質情報などを含むことで、今後のデータ取引などで必須の情報になります。

¹ ベース・レジストリとは、公的機関等で登録・公開され、様々な場面で参照される、人、法人、土地、建物、資格等の社会の基本データであり、正確性や最新性が確保された社会の基盤となるデータベース群の総称。

² https://www.kantei.go.jp/jp/singi/it2/dgov/dail0/siryou_a.pdf

³ データに関するデータと呼ばれる見出し情報である。データ検索や管理に使われる。

2 メタデータ項目の定義

2.1 DCAT の概要

DCAT とは、国際標準化団体 W3C (World Wide Web Consortium) が勧告しているデータカタログのメタデータ標準であり、その構造は図 1 のようなクラスとプロパティで示されています。

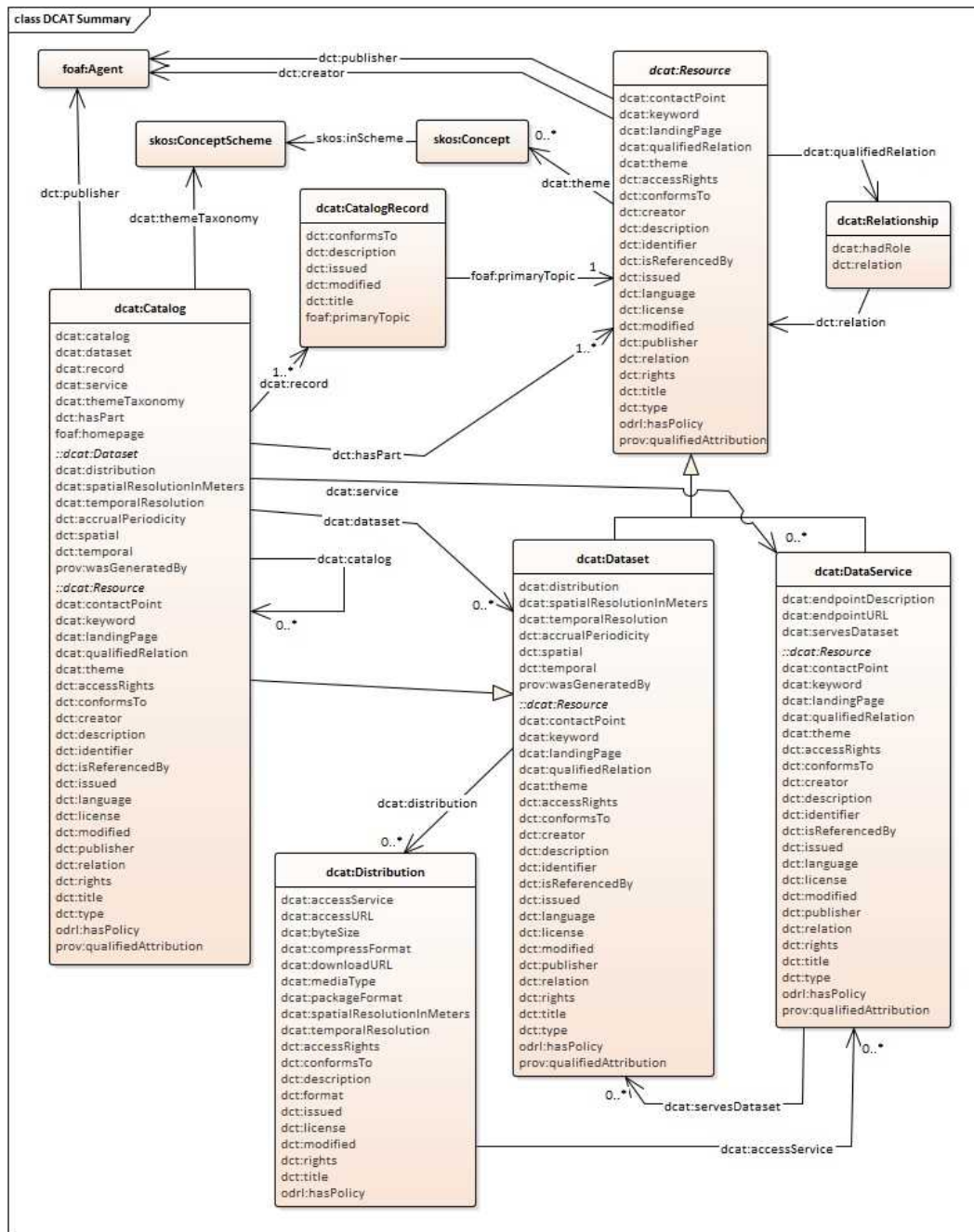


図 1 DCAT モデルの概要

(出典 : <https://www.w3.org/TR/2020/REC-vocab-dcat-2-20200204/>)

図1を概略化すると、以下のようにカタログ、カタログレコード、データセット、データサービス、ディストリビューションの階層構造で示すことができます。

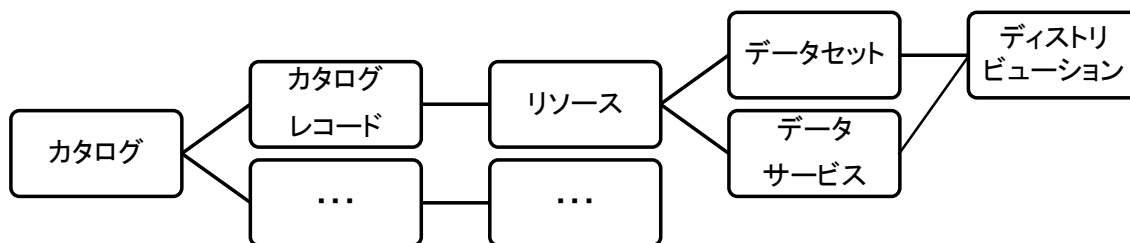


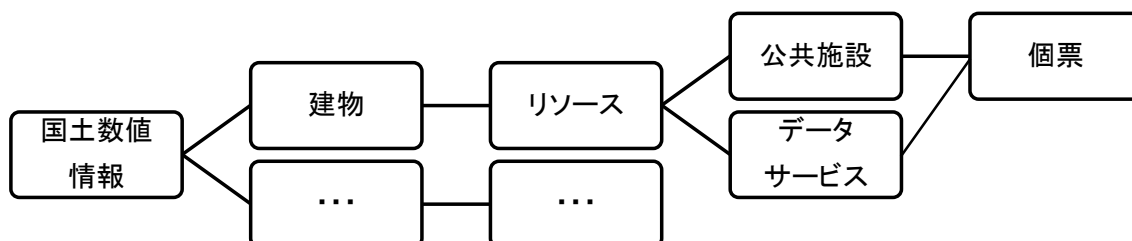
図2 DCATの構成概略

また、上記概略図で示す各階層（クラス）の定義は以下のとおりです。

カタログ	特定目的に対するデータの集合体のことをカタログと呼ぶ。カタログは、より上位のカタログを持つことや、カタログ内にサブのカタログを持つこともある。
カタログレコード	カタログの中で分野を設定したいときに持つ。カタログ冊子の章のようなものである。
リソース	カタログの中のデータに関する項目である。データ項目は持たず、後述するデータセットやデータサービスにより構成される。
データセット	公開される情報の束であり、複数のデータの集合体であることもある。品質やライセンス等、利用するための情報を含む。
データサービス	サブスクリプションやAPIでデータを入手するための詳細情報である。サービス形式などの情報を含む。
ディストリビューション (配信)	個々のデータ単位の情報である。配信するための技術的内容などを含む。

例えば、図2の構成を実際のデータセットに当てはめた場合、次の図3のように示すことができます。

例1：国土数値情報の場合



例2：会議体の場合

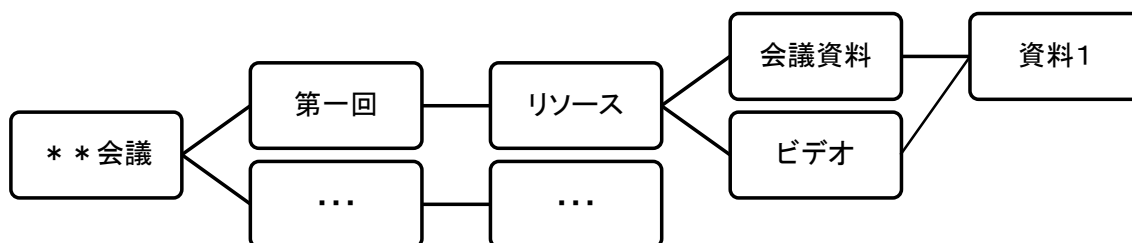


図3 DCATの使用例

2.2 DCAT シリーズ

DCATには様々な派生形があります。EU（欧州連合）の European Commission⁴では、DCATの仕様に基づき、対象別にそれぞれ以下のようなメタデータ標準を定めています。

1) BregDCAT

ベース・レジストリのメタデータ標準。行政機関の連絡先窓口やカタログの来歴情報などのメタデータ項目を定義⁵しています。

2) GeoDCAT

地理空間情報のメタデータ標準。住所や地理座標などのメタデータ項目を定義⁶しています。

⁴ 欧州連合の政策執行機関。

⁵ 詳細は <https://joinup.ec.europa.eu/collection/access-base-registries/solution/abr-bregdcap> を参照。

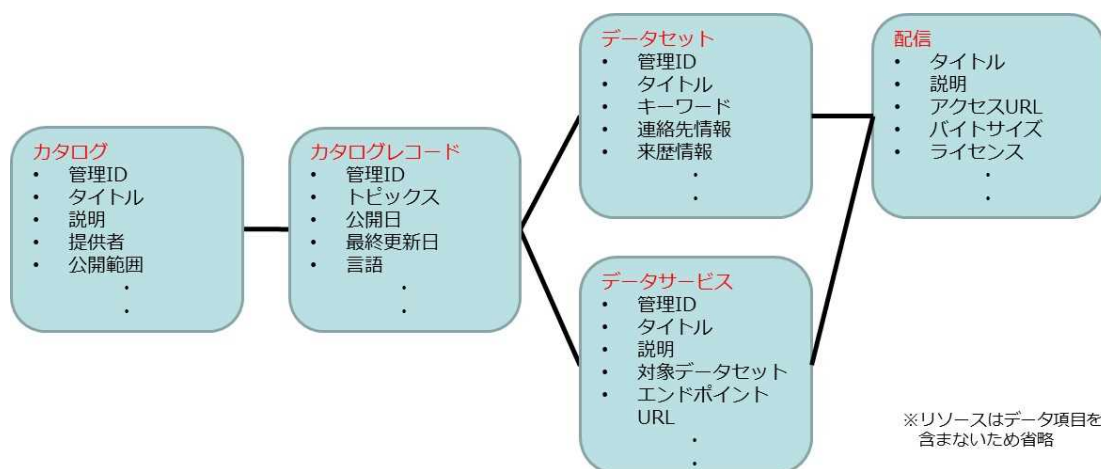
⁶ 詳細は <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/geodcat-application-profile-data-portals-europe> を参照。

3) StatDCAT

統計情報のメタデータ標準。統計データの系列や単位などのメタデータ項目を定義⁷しています。

2.3 ベース・レジストリのメタデータ項目

ベース・レジストリのメタデータ項目は、2.2で記したBregDCATを参考に、DCATの仕様にに基づき、それぞれのクラスで定めるべき項目を「DCAT-GoJ」という名称で定義します。DCAT-GoJのイメージは図3のとおりです。



DCAT-GoJで定めるメタデータ項目は次に示すとおりです。

なお、次に示す表の内容についても以下のとおりとなります。

- ・メタデータ項目：メタデータの項目名を示しています。
- ・分類：メタデータを付与するレベルを示しています。
 - ◎はメタデータの付与を必須とする項目、
 - はメタデータの付与を推奨する項目、
 - 空白はメタデータの付与を任意とする項目。
- ・項目説明：メタデータ項目の説明を示しています。
- ・データ形式：メタデータのデータ形式を示しています。
- ・記述方式：メタデータの記述方式を示しています。

指定された形式で記述とされている場合は、項目説明欄で指定された形式で記述してください。

⁷ 詳細は <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/statdcat-application-profile-data-portals-europe> を参照。

指定された選択肢から選択とされている場合は、項目説明欄で示されている選択肢から選択してください。

1) カタログ

メタデータ項目	分類	項目説明	データ形式	記述方式
管理 ID	◎	カタログをユニークに識別するための管理 ID を示す。	文字列型	自由記述
タイトル	◎	カタログのタイトルを示す。	文字列型	自由記述
サブタイトル		カタログのサブタイトルを示す。	文字列型	自由記述
説明	◎	カタログに収録されている情報の特徴を第三者に理解してもらうための説明を示す。また、関係性のあるカタログが存在する場合は、その名称も示す。	文字列型	自由記述
キーワード	○	カタログを容易に検索できるように、検索タグとして扱うキーワードを示す。	文字列型	自由記述
テーマ分類	○	カタログが扱う情報のテーマ分類を示す。	文字列型	自由記述
対象地域	○	カタログの対象地域を示す。	文字列型	自由記述
親カタログ ID	○	上位のカタログがある時にそのカタログ ID を示す。	文字列型	自由記述
提供者	◎	カタログを提供するエンティティ（組織又は個人）の名称を示す。 ※法人番号も検討	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
公開日	○	カタログの公開日を示す。 入力形式は以下のとおりとする。 ・ YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述
最終更新日	○	カタログを最後に更新・修正した日付を示す。 入力形式は以下のとおり。 ・ YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述
更新頻度	○	カタログが更新される頻度を示す。 定期的に提供する場合、年・月・週・日あたりの回数を示す。 不定期の場合は、不定期と示す。	文字列型	自由記述
言語	○	カタログがどの言語で記述されているかを示す。 なお、国際標準化機構（ISO）が発行する「ISO 639-1:2002」又は「ISO 639-2:1998」に準拠したアルファベット2文字又は3文字で表記する。 【例】日本語：ja、英語：en	言語型	指定された形式で記述
公開範囲		カタログの公開範囲を、以下のいずれかで示す。 ・ 公開 ・ 制限付き公開 ・ 非公開	文字列型	指定された選択肢から選択

メタデータ項目	分類	項目説明	データ形式	記述方式
公開条件		カタログの公開範囲が「制限付き公開」の場合に、制限条件及び制限解除するための条件を示す。	文字列型	自由記述
ライセンス	○	カタログに適用されるライセンスについて示す。 外部で定義されたライセンスへのリンク（URL）を示してもよい。	文字列型	自由記述
利用規約		カタログの利用規約を示す。 外部で定義された利用規約へのリンク（URL）を示してもよい。	文字列型	自由記述
ホームページ	○	カタログに関連するホームページを示す。	文字列型	自由記述

2) カタログレコード

メタデータ項目	分類	項目説明	データ形式	記述方式
トピックス	◎	カタログのトピックスを示す。	文字列型	自由記述
説明	◎	カタログレコードにどのようなデータセットが収録されているか、説明を示す。	文字列型	自由記述
公開日	○	カタログレコードの公開した日付を示す。 入力形式は以下のとおりとする。 ・ YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述

メタデータ項目	分類	項目説明	データ形式	記述方式
最終更新日	○	カタログレコードを最後に更新・修正した日付を示す。 入力形式は以下のとおりとする。 ・YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述
言語	○	カタログレコードがどの言語で記述されているかを示す。 なお、国際標準化機構（ISO）が発行する「ISO 639-1:2002」又は「ISO 639-2:1998」に準拠したアルファベット2文字又は3文字で表記する。 【例】日本語：ja、英語：en	言語型	指定された形式で記述

3) データセット

メタデータ項目	分類	項目説明	データ形式	記述方式
管理 ID	◎	データセットをユニークに識別するための管理 ID を示す。	文字列型	自由記述
タイトル	◎	データセットのタイトルを示す。	文字列型	自由記述
サブタイトル		データセットのサブタイトルを示す。	文字列型	自由記述
バージョン		データセットのバージョン情報を示す。	文字列型	自由記述
説明	◎	データセットの特徴を第三者に理解してもらうための説明を示す。	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
キーワード	○	データセットを容易に検索できるように、検索タグとして扱うキーワードを示す。	文字列型	自由記述
対象地域	○	データセットの対象地域を示す。	文字列型	自由記述
対象期間	○	データセットの対象期間を示す。 入力形式は以下のとおりとする。 ただし、明確に年月日で示せない場合は、自由記述とする。 ・開始年月日/終了年月日 : YYYY-MM-DD/YYYY-MM-DD	文字列型	指定された形式で記述 または 自由記述
分類	○	データセットが扱うデータの分類を日本標準産業分類に基づき示す。	文字列型	指定された形式で記述
提供者	○	データセットを提供するエンティティ（組織又は個人）の名称を示す。 ※法人番号も検討	文字列型	自由記述
作成者		データセットの作成に関わったエンティティ（組織又は個人）を示す。 ※法人番号も検討	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
連絡先情報	○	データセットについて問い合わせを行う際の連絡先情報として以下の項目を示す。 <ul style="list-style-type: none"> ・組織名 ・部署名 ・電話番号 ・メールアドレス ・フォーム URL ※法人番号も検討	文字列型	指定された形式で記述
タイプ		データセットの形式を示す。	文字列型	自由記述
来歴情報		データセットの出所や変更履歴等の来歴情報を示す	言語型	自由記述
品質評価		データセットの品質評価（評価方法は別途定義）を示す。	文字列型	自由記述
品質測定結果		データセットの品質測定結果（測定方法は別途定義）を示す。	文字列型	自由記述
公開日		データセットを公開した日付を示す。 入力形式は以下のとおりとする。 ・YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述
最終更新日	○	データセットを最後に更新・修正した日付を示す。 入力形式は以下のとおりとする。 ・YYYY-MM-DD（西暦-月-日）	日付型	指定された形式で記述

メタデータ項目	分類	項目説明	データ形式	記述方式
更新頻度		データセットが更新される頻度を示す。 定期的に提供する場合、年・月・週・日あたりの回数を示す。 不定期の場合は、不定期と示す。	文字列型	自由記述
言語		データセットがどの言語で記述されているかを記載する。 なお、国際標準化機構（ISO）が発行する「ISO 639-1:2002」又は「ISO 639-2:1998」に準拠したアルファベット2文字又は3文字で表記する。 【例】日本語：ja、英語：en	言語型	指定された形式で記述
公開範囲		データセットの公開範囲を、以下のいずれかで示す。 ・公開 ・制限付き公開 ・非公開	文字列型	指定された選択肢から選択
公開条件		カタログの公開範囲が「制限付き公開」の場合に、制限条件及び制限解除するための条件を示す。	文字列型	自由記述
準拠する標準		データセットが準拠する標準を示す。	文字列型	自由記述
関連ドキュメント		データセットに関する情報を持つドキュメントへのリンク（URL）を示す。	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
ランディングページ		データセットに関して、追加・補足できる情報が公開されている Web ページの URL を示す。	文字列型	自由記述

4) データサービス

メタデータ項目	分類	項目説明	データ形式	記述方式
管理 ID	◎	データサービスをユニークに識別するための管理 ID を示す。	文字列型	自由記述
タイトル	◎	データサービスのタイトルを示す。	文字列型	自由記述
説明	◎	データサービスの特徴を第三者に理解してもらうための説明を示す。	文字列型	自由記述
キーワード	○	データサービスを容易に検索できるように、検索タグとして扱うキーワードを示す。	文字列型	自由記述
対象データセット	○	データサービスが配信対象とするデータセットを示す。	文字列型	自由記述
提供者	○	データサービスを提供するエンティティ（組織又は個人）の名称を示す。	文字列型	自由記述
タイプ		データサービスの形式を示す。	文字列型	自由記述
公開範囲		データサービスの公開範囲を、以下のいずれかで示す。 <ul style="list-style-type: none"> ・公開 ・制限付き公開 ・非公開 	文字列型	指定された選択肢から選択

メタデータ項目	分類	項目説明	データ形式	記述方式
公開条件		カタログの公開範囲が「制限付き公開」の場合に、制限条件及び制限解除するための条件を示す。	文字列型	自由記述
ライセンス		データサービスを利用可能にするためのライセンスを示す。 外部で定義されたライセンスへのリンク (URL) を示してもよい。	文字列型	自由記述
準拠する標準		データサービスが準拠する標準を示す。	文字列型	自由記述
関連ドキュメント		データサービスに関する情報を持つドキュメントへのリンク (URL) を示す。	文字列型	自由記述
エンドポイント URL	○	データサービスのルート的位置又は主要エンドポイント (Web で解決可能な IRI (Internationalized Resource Identifier)) を示す。	文字列型	自由記述
ランディングページ		データサービスに関して、追加・補足できる情報が公開されている Web ページの URL を示す。	文字列型	自由記述

5) ディストリビューション

メタデータ項目	分類	項目説明	データ形式	記述方式
タイトル	◎	ディストリビューションのタイトルを示す。	文字列型	自由記述
説明	◎	ディストリビューションに関する説明を示す。	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
アクセスサービス		ディストリビューションにアクセスできるサービスを示す。	文字列型	自由記述
バイトサイズ		ディストリビューションのサイズをバイト単位で示す。	数値型	指定された形式で記述
圧縮形式		ディストリビューションのデータサイズを縮小するために圧縮されている場合、圧縮形式を示す。	文字列型	自由記述
メディアタイプ		ディストリビューションのメディアタイプを示す。 IANA(Internet Assigned Numbers Authority)管理語彙で定義されている配信データのメディアタイプから選択する。	文字列型	指定された形式で記述
公開日	○	ディストリビューションを公開した日付を示す。 入力形式は以下のとおりとする。 ・YYYY-MM-DD (西暦-月-日)	日付型	指定された形式で記述
最終更新日	◎	ディストリビューションを最後に更新・修正した日付を示す。 入力形式は以下のとおり。 ・YYYY-MM-DD (西暦-月-日)	日付型	指定された形式で記述

メタデータ項目	分類	項目説明	データ形式	記述方式
期間	○	ディストリビューションの配信期間を示す。 入力形式は以下のとおりとする。 配信期間：YYYY-MM-DD/YYYY-MM-DD	日付型	指定された形式で記述
ステータス		ディストリビューションのステータスを以下のいずれかで示す。 ・配信中 ・配信準備中 ・配信停止中	文字列型	指定された選択肢から選択
言語		ディストリビューションデータがどの言語で記述されているかを記載する。 なお、国際標準化機構（ISO）が発行する「ISO 639-1:2002」又は「ISO 639-2:1998」に準拠したアルファベット2文字又は3文字で表記する。 【例】日本語：ja、英語：en	言語型	指定された形式で記述
ライセンス	○	ディストリビューションを利用可能にするためのライセンスを示す。 外部で定義されたライセンスへのリンク（URL）を示してもよい。	文字列型	自由記述
利用規約		ディストリビューションの利用規約を示す。 外部で定義された利用規約へのリンク（URL）を示してもよい。	文字列型	自由記述

メタデータ項目	分類	項目説明	データ形式	記述方式
準拠する標準		ディストリビューションが準拠する標準を示す。	文字列型	自由記述
関連ドキュメント		ディストリビューションに関する情報を持つドキュメントへのリンク (URL) を示す。	文字列型	自由記述
アクセス URL	◎	ディストリビューションを取得できる Web ページの URL を示す。	文字列型	自由記述
ダウンロード URL		ディストリビューションがダウンロード可能な URL を示す。	文字列型	自由記述

2.4 メタデータの記載例

前記で示した BregDCAT に基づきメタデータ情報を記載した例は次のとおりです。

記載例	
カタログ	
管理 ID	CA00001
タイトル	文字情報基盤
サブタイトル	
説明	行政で用いられる人名漢字、法人用漢字、かな等の約 6 万文字の漢字です。
キーワード	漢字、戸籍統一文字、住民基本台帳ネットワーク統一文字、登記統一文字、変体仮名
テーマ分類	文字
対象地域	全国
提供者	文字情報技術促進協議会
公開日	2011-10-26
最終更新日	2020-8-26
更新頻度	不定期
言語	ja
公開範囲	公開

管理元で別途付番するため、記載不要です。

分類が必須(◎)でなく、未定または不明な項目は空欄可です。

指定された形式で記述してください。

指定された選択肢から選択してください。

記載例	
公開条件	
ライセンス	IPA フォントライセンス v1.0
利用規約	クリエイティブ・コモンズ 表示 - 継承 2.1 日本 ライセンス条件
ホームページ	https://moji.or.jp/mojikiban/
カタログレコード	
トピックス	文字
説明	行政で用いられる人名漢字、法人用漢字、かな等の約6万文字の漢字です。
公開日	2011-10-26
最終更新日	2020-8-26
言語	ja
データセット	
管理 ID	CA00001-DST001
タイトル	MJ 文字情報一覧表
サブタイトル	
バージョン	Ver. 006. 01
説明	各文字に関するコード、読み、字母、画数等をまとめた情報。
キーワード	フォント、ヨミガナ、画数
対象地域	全国
対象期間	
分類	全ての業務
提供者	文字情報技術促進協議会
作成者	情報処理推進機構
連絡先情報	組織名：一般社団法人 文字情報技術促進協議会 メールアドレス：info@moji.or.jp フォーム URL：https://moji.or.jp/about/contact/
タイプ	Strict Open XML
来歴情報	2020年10月に、情報処理推進機構から文字情報技術促進協議会に信託譲渡
品質評価	正確性、網羅性
品質測定結果	公務で使うのに十分な品質
公開日	2011-10-26

管理元で別途付番するため、記載不要です。

別途定義する基準に基づき記載してください。

記載例	
最終更新日	2020-8-26
更新頻度	不定期
言語	ja
公開範囲	公開
公開条件	
準拠する標準	
関連ドキュメント	
ランディングページ	
データサービス	
管理 ID	DSV00001
タイトル	文字情報基盤検索サービス
説明	
キーワード	
対象データセット	
提供者	
タイプ	web
公開範囲	公開
公開条件	
ライセンス	
準拠する標準	
関連ドキュメント	
エンドポイント URL	
ランディングページ	
配信	
タイトル	mji.00601-xlsx.zip
説明	
アクセスサービス	

管理元で別途付番するため、記載不要です。

記載例	
バイトサイズ	7.3M
圧縮形式	zip
メディアタイプ	
公開日	
最終更新日	2019-05-01
期間	
ステータス	配信中
言語	ja
ライセンス	
利用規約	
準拠する標準	
関連ドキュメント	
アクセス URL	https://moji.or.jp/mojikiban/mjlist/
ダウンロード URL	https://moji.or.jp/wp-content/mojikiban/oscdl/mji.00601-xlsx.zip