

コード（分類体系）設計・運用実践 ガイドブック

β版

平成29年8月1日

内閣官房情報通信技術（IT）総合戦略室

【目次】

1	経緯及び目的	3
2	基本事項	4
2. 1	コード（分類体系）の概要	4
2. 2	コード活用の意義	5
2. 3	行政機関等が提供するコード	5
3	コードの設計	6
3. 1	コード設計の手順	6
3. 2	標準や既存コードの活用と拡張	8
3. 3	データ交換用コードマッピングの実施	8
3. 4	コード設計にあたっての留意事項	9
4	コードの導入と運用	10
4. 1	コード情報の提供	10
4. 2	コード仕様と利用規約の提供	10
4. 3	コードの変更手順	10

【更新履歴】

版数	更新年月日	更新内容	備考
β	2017年8月1日	β 版作成	

1 経緯及び目的

各府省では、「電子行政分野におけるオープンな利用環境整備に向けたアクションプラン」（平成 26 年 4 月 25 日各府省情報統括責任者（CIO）連絡会議決定）に基づき、自府省内の諸活動に関する透明性を高め、開かれた行政の実現を図ってきた。

また、内閣官房では、国民、企業等の社会・経済活動に有益な情報資源の充実に資する観点から、行政情報を有効活用し、より利便性の高い公共サービスの実現に向けて各種ガイドの整備と普及を推進してきた。

平成 28 年 12 月に官民データ活用推進基本法が制定され、世界最先端 IT 国家創造宣言・官民データ活用推進基本計画が平成 29 年 5 月に閣議決定されたことから、今後は、地方公共団体も含め、データ流通環境の整備が加速されることとなる。データの活用や流通にはデータの標準などが重要になることから、データに関するガイド類の整備が求められているところである。

現在、政府機関は 400 を超える分類体系コード（以下「コード」という。）を持っている。この中には、日本標準産業分類のように JIS に定義されたマスターコード、自サービスで使うために設計された独自のコードが存在する。

行政機関で制定したコードは社会のマスターコードとして扱われる場合が多いが、これらのコードは、政府共通のマスターコードとして定義されていなかったり、類似するものが複数あるなどにより、業務連携が煩雑になるとともに、システム連携時に無駄なコストが発生する等の様々な問題点が指摘されている。

また、コード作成時に手引きとする文書が存在していなかったこともあり、コード設計や管理に改善の余地があるものも見受けられる。

本ガイドブックは、政府機関において新規コードの作成及び現行コードの変更に携わる行政職員に向けて、適切な設計方法や運用手順を示すことを目的として策定した。

なお、本ガイドブックは、遵守を必須としたものではなく、データを効果的に活用するために参照し、効率的に作業することを目的としたものとして整理した。

また、行政機関において、これらデータのガイドブックの提供は初の試みとなることから、β版での公表を行い、利用者、開発者の声を聴き、順次改善を行っていく。

2 基本事項

2. 1 コード（分類体系）の概要

コードとは、情報サービスや事務処理を効率的に行うことを目的として、個々の情報を簡易に管理するためにつけた記号（英数字など）のことである。

組織や分野毎に独立してコードが作られ利用されている場合もあるが、異なるコードを使用していると組織及び分野間で電子的な情報の交換をすることができないため、社会的に標準コードを設けている分野が多い。

行政機関が付与するコード体系は、広く社会で使われるマスターコードとして扱われることが多いため、拡張性や運用性を考えた上で、コードの整備を行う必要がある。

コードは、識別・分類・配列が三大機能と呼ばれ、大きくは「識別のためのコード」と「分類のためのコード」に分けられる。「識別のためのコード」は、マイナンバー、法人番号、顧客コード、製品コード等にみられる個別の対象物にコードがつけられたものである。「分類のためのコード」は標準産業分類や性別コード等のように情報の分類に使うためのコードである。

また、コードの付与方式は、有意コードと無意コードがある。有意コードとは、コードの記号に意味を持たせるコードである。例えば、JP-CAS-2017-003のように、日本のCASという組織が2017年に3番目に付番したものと管理するような方式である。無意コードは、コードの記号に乱数などを使い、コードに意味を持たせない方式である。コードの意味付けを下表にて、容易に管理することができる。

表 有意コード・無意コードの特徴

区分	特徴
有意コード	<ul style="list-style-type: none">・ 利用者がコードの指す内容を理解することができる（人にやさしい）。・ 分散して運用すると、末端の通番管理が難しい。・ コードのパターンが固定されていることから、コードの運用方法を変えることが難しい（組織再編等）。・ データ項目の増加により、コード体系が破綻しやすい。
無意コード	<ul style="list-style-type: none">・ データ項目数の変更に対応することができる（機械にやさしい）。・ コードと意味を関連付けるテーブルが必要である。・ コードの運用変更があっても、システム改修への影響が小さい。・ コード番号に上下関係が示されない。・ uuid等、コードの重複を避ける技術的な解決策が提案されている。

コードの配列機能は有意コードの時に有効な方法である。コードに意味があるので、組織等、コードの中にある分類ごとにデータを抽出して処理をしたり、配列をすることが容易にできる。

2. 2 コード活用の意義

共通の識別コードを使うことにより、複数のデータを一体として活用することが可能になる。契約や表彰等の政府の法人活動情報を集約して提供する法人インフォメーションでは、法人番号を識別コードとして、各府省の保有する情報を連携させ、ワンストップで情報提供をしている。データ交換時に、コードを送信することで各組織が共通で持つマスターデータの関連情報の送信を省略する等、データ交換量を減らしたり、データの連携を効率化することが可能である。

また、分類コードを共通化することで組織横断の情報の集約が効果的に可能になる。公共調達で使われる調達物品のコードは、各調達機関で異なっているが、分類コードを共通化することで、一元的な情報収集が可能になる。

このように行政機関間や官民でデータを融合して活用するときにキーとなるのがコードである。グローバルに情報交換する場合にもマスターコードを使って各種情報をつなぐことで効果的に情報交換が行われる等、データの活用には必須の仕組みである。

ただし、各システムでは既存のコード体系を保有している場合が多い。コード変換は多大なコストがかかるとともに影響範囲が大きいことから、既存のコードをそのまま使いたいという要望が大きい。そうした場合には、識別コード、分類コードともにコードの変換表（マッピング）等で対応を行うこととなる。

2. 3 行政機関等が提供するコード

政府機関などの公的機関が公開するコードは、社会的なマスターコードとして扱われることが多い。データの整理や活用の鍵になることから、オープンデータ化の要望が大きいデータである。国際標準があるときには国際標準を優先する等、より広範に使えるコードが優先されることが多い。

(1) 国際標準のコード

国際標準機関で決定されたコード体系には、主に以下のものがある。

	国際標準	概要
性別コード	ISO 5218	性別に関する国際標準
国名コード	ISO 3166-1	国名に関する国際標準
言語コード	ISO 639	言語に関する国際標準。2文字と3文字の表記がある。
日時	ISO 8601	日時の表記に関する国際標準

※日時はコードではないが、定型的な記述が求められる基本体系であるため記載

(2) 政府標準のコード

政府内の各府省が決定したコード体系には、主に以下のものがある。

	発行者	概要
標準産業分類	総務省	統計を始め様々な場面で使われる産業コード
法人番号	国税庁	登記された法人に一意に振られるコード
全国地方公共 団体コード	総務省	地方公共団体に付与されたコード

(3) 業界標準のコード

業界のデータ流通を高度にするために業界団体等が定めたコードである。行政機関の制定したコードではないが社会全体で使われるコードも多い。

3 コードの設計

サービスを構築するたびにコード再設計を行うことは、設計時間や費用の無駄なだけでなく、品質の劣化などにも繋がるものである。中長期的視点を持ってコード設計を行う必要がある。

以下に、本ガイドブックが示すコード設計の基本的方針を示す。

- (1) 標準コードがあるときには、その標準コードが使えないか考える。
- (2) 標準コードがあるがそのまま使えないときには、拡張を考える。
- (3) 新しくコードを作るときには、中長期的な利用やデータ連携を前提とする。
- (4) 既存のコードを利用しているが改善の必要がある場合には、コード変更による効果と影響を見極めた上で、コードの変更を行う。

3. 1 コード設計の手順

コードの設計は以下の手順で実施する。ただし、(4)以降は同時並行で進められることが多く、実施順序は状況に応じて変えてもよい。

- (1) 目的の明確化
なぜコードが必要か利用者と利用イメージを明確にする。
- (2) 標準コードの調査
目的に適ったコードが既に存在しないか調査する。政府機関関係のコードについては、政府 CIO ポータルのコード一覧 (<https://cio.go.jp/codes>) で確認することができる。類似のコードが存在する場合は、追加・改変の必要性も含め、活用可能性を検討する。類似コードが存在しない場合は、以下に示す考え方に基づき新たなコードを作成する。作成した新しいコードは、前述の政府 CIO ポータルなどに追加・公開する。
- (3) コード内容の意味の明確化
コードの各項目について意味を明確に定義する。例えば、「博物館」は何を意味するのか、「複合文化施設」は何を意味するのかを定義し、コード利用者が迷わず

に選択することができるように解説をする。

(4) 有意コード、無意コードの検討

コード化対象の特性（コードの使用範囲、コード運営者・利用者の視点、コードの増加や再設計の可能性、データの使用頻度・更新頻度など）を分析し、有意コード・無意コードの選択を行う。

なお、コード化対象データ数が将来どれだけ増加するか見積もりにくい場合は、無意である方が柔軟に対応することができるので、無意コードを推奨する。

(5) 桁数の検討

コードの使用範囲・使用期間を検討し、現在のコード化対象数を考慮した上で、将来必要となるコード数を見積もる。その上で、コードの必要桁数、コードに使う数字、英数字などの文字種を検討する。

(6) 階層化の検討（上位コード、下位コード）

管理を容易にするため、また、データの管理や検索を容易にするため、コードの階層化を検討する。大分類、中分類、小分類のように、用途に応じて粒度を選択して活用することができるコードも多い。

郵便番号の階層の例

10	0	0001
東京都	銀座郵便局（東京都千代田区）	千代田

郵便番号では上2桁が都道府県を表し、上3桁又は5桁で管轄郵便局が、下4桁又は2桁で町域が、それぞれ分かるようになっている。

無意コードを使っている場合にも、その上位コードで有意コードを使い、有意コードと、無意コードを組み合わせ使用することもある。

(7) 関連主要コードとの関係整理

関連主要コードとの関係を整理し、その対応表を公開する。また、1:nの対応方法等、関係整理の方法などを記述する。

例として、日本標準産業分類と国際標準産業分類の対応関係を下記の表で示す。

表 日本標準産業分類と国際標準産業分類の対応関係（一部抜粋）

分類番号 (JSIC)	日本標準産業分類 (JSIC) 第13版	分類番号 (ISIC)	国際標準産業分類 (ISIC) 第4版（仮訳）
0121	酪農業	0141	畜牛及び水牛の飼育
		0144	羊及び山羊の飼育
0122	肉用牛生産業	0141	畜牛及び水牛の飼育
0123	養豚業	0145	豚の飼育
0124	養鶏業	0146	家禽類の飼育
0126	養蚕農業	0149	その他の動物の飼育

出所：総務省「日本標準産業分類と国際標準産業分類の対応関係」

(http://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/)

(8) チェックディジットの付加

入力ミスによる誤った処理を避けるため、チェックディジットを入れることがある。チェックディジットを使う場合には、コード末尾に付加することを推奨する。

(9) コード表記の検討

111-1111、9999-9999-9999-9999 といったように、3、4桁毎にハイフンを記載するなどして、視認性の高い表記を検討する。

3. 2 標準や既存コードの活用と拡張

標準コード調査により、コード化対象である情報が既に標準コードとして体系化されていると分かった場合、そのコードをベースとして活用することができないか検討する。目的の違いなどで新しいコードを作る場合には、オリジナルとなるコードとの整合性を取りながら拡張をしていくことが望ましい。他に選択肢がない場合には新規にコードを起こすこととなる。

既存コードの拡張には、以下のような例がある。

- ・ 詳細化したい項目に枝番をつけるケース

法人番号はあるが事業所管理をするために枝番を独自に付加

- ・ 既存分類コードに分類コードを追加するケース

古い分類に対して、最新の分類項目を付加

いずれの場合においても、将来のデータ活用を容易にするために、オリジナルのコードと変更した部分を明確化しておく必要がある。

複数の既存コードがあり標準コードが存在しない場合には、既存のコードのマッピングにより検証することが望ましい。多くのコードをマッピングすることで、より利用者のニーズに合った分類などが見つけやすくなり、比較を通じて相互運用性の面から必要な分類体系が分かってくることもある。また、主要コードをマッピングすることで既存コードの保有者も全体像を把握することができ、納得性の高いコードを整備することが可能になる。

3. 3 データ交換用コードマッピングの実施

最終的なコードを決定した時に、既存コードとコード間のマッピングを行う必要がある。前項でのコード項目検討時の比較以上に、詳細な意味定義も含め連携するサービスの双方の視点からデータを見るなど、具体的に検討を進めて行く必要がある。

(1) 分類項目が異なる場合

「1：会社員、2：公務員、3：自営業、4：学生 5：無職」と「1：会社員、

2：公務員、3：自営業、4：学生 5：その他」の場合、5番の「無職」と「その他」は同一ではない。そうした場合には、「その他」の内の無職割合を換算しデータをつなぐ、もしくは連携させない等のデータ連携の扱いを考えなければいけない。

(2) 分類の粒度が違う場合

「動植物園」と「動物園」「植物園」のようにデータ項目の粒度が違う場合には、以下のように対応させるが、あくまでも変換であり、相互に正確なデータが交換することができることを保証するものではないことに留意が必要である。

コードA	コードB
動植物園	動物園
	植物園

中長期的には共通化したコード体系を使うことが理想的であるが、システム更改時などタイミングを見て揃えていく、若しくは、マッピングの精度を上げていくことが必要である。

3. 4 コード設計にあたっての留意事項

その他、コード設計の留意事項は以下のとおりである。

(1) 直交性

コードとコード付与対象の結びつきを一意にする。同じ情報を表現するコードが複数定義されてはならないし、逆に同じコードで複数の情報を示すように設計してはならない。割当先の情報（分類）に独立性を持たせる。

例えば、企業の合併などにより、企業が2つのコードを持ってしまう場合があるが、この場合には1つの継承コードでその後の処理をしていくように運用で対処していく必要がある。また、「30101:屋外イベント」「30102:屋内イベント」のように状態を示すものと、「30105:音楽イベント」等のイベント種別を示すものが同じコード上に混在して、「屋外コンサート」のように1つのコードを選択することが困難となるコード設計は避けなければならない。

(2) 英語表記

国際的運用を見据え、データ項目に適切な英語表記を加える。複数単語による英語項目名をつける必要がある場合には、単語の区切りを「_」とするスネークケースか区切りの文字だけを大文字で示すキャメルケース化の選択をする。

(3) バーコード等の利用

入力ミス防止や入力効率化のため、状況に合わせてバーコードやQRコードを活用する。

4 コードの導入と運用

コード作成後、運用管理するにあたって以下に留意する。

4. 1 コード情報の提供

作成したコードは原則として公開する。可能であれば、コード情報を提供する API を公開する。これにより、コードが指す内容の確認のほか、新規発行における重複発行防止などが可能になる。API について、詳しくは API ガイドブック・API テクニカルガイドブックを参照する。

コードにはメタ情報として以下の 4 点を明記する。これら 4 情報と該当コードを記載することで、対象物の一意性が確保される。

- コードオーナー
- コード名
- コードバージョン
- 発行日

保有するコードを Microsoft Excel ファイル等の表作成ソフトで公開している場合には、入力形式を自動で判別して変換することがあるので注意が必要である。例えば、数値のみを入力した際には標準の設定で形式が「数値」となり、“0”が先頭列にある場合は削除されてしまう（例：012345→12345）。これを防ぐため、Excel 等の表作成ソフト内で形式をあらかじめ「文字列」に指定するか、OpenRefine

(<http://openrefine.org/>) などのデータ自動変換がないエディタで一覧データを扱う必要がある。

4. 2 コード仕様と利用規約の提供

コードを提供するときには、以下のようにコードの各項目の定義を明示した仕様を公開する。

(例) 項目名：都道府県名

コード：2 桁 Integer (整数)

解説：日本の都道府県名

コードに利用制限などがある場合には、コードの利用規約の中でそれを明記する。

4. 3 コードの変更手順

コードの廃止/移行/統合/分割の際には、現行のコードを利用している個人や組織の対応準備期間として、一定期間の告知を行う。また、早めの周知を心がける。その際、廃止/移行/統合/分割することになった背景やその内容を記すことが望ましい。

(1) コードの削除

コード対象の廃止などにより削除したコードは、再利用しないことが望ましい。コー

ドの桁数不足などで、やむを得ず再利用せざるを得ない場合には、相当期間再利用しない期間を空けた上で、再利用しているコードがあることを周知した上で再利用を実施する。

(2) コードの改定

十分に検討され設計されたコードであっても、時間の経過につれて想定していなかったコード化範囲の変化や割当ての増加によって、現状の割当てを維持することができなくなることがある。コードの欠番の利用やコード桁数の追加、サブコードの追加で対応するケースが多いが、接続先システムへの影響も勘案して対処方法を検討する必要がある。

また、コードの改定を行った場合には、改定内容をまとめた文書を公開する。記載すべき項目は以下のとおりである。

- ・ バージョン
- ・ 改定日
- ・ 改定理由
- ・ 改定点

改定の際には改定版と旧版との間で比較可能とすることが、継続性という観点から重要である。旧コードのアーカイブを提供するとともに変換表の提供を行うことが望ましい。

(3) コードの並行運用

複数システムでコードを運用している場合には、変換表で対応するのではなく、関係システムのコード変更が終わるまで並行運用する場合もある。

参考1 チェックディジット

コードは記号化されているため、入力ミスに気づきにくい。そこで、入力エラーを検知するための検査用数字をコードに付加することがある。これをチェックディジットと呼び、一定の計算式でコードから算出し、先頭や末尾に付加される。

チェックディジットの計算方法には様々なものがある。例を以下の表で示す。

表 チェックディジットの計算方法

アルゴリズム名	例
モジュラス 10 ウェイト 2・1 (Luhn アルゴリズム)	・ クレジットカード番号 ・ IMEI (International Mobile Equipment Identity)
モジュラス 10 ウェイト 3・1	・ ISBN-13 (国際標準図書番号) ・ JAN/EAN などのバーコード
モジュラス 11	・ 個人番号 (マイナンバー) ・ 住民票コード ・ 運転免許証番号

出所：下記資料を参考に作成

- ・ ISO/IEC 7812-1 (<https://www.iso.org/standard/70484.html>)
- ・ IMEI (<https://imeidata.net/tool/calculator>)
- ・ ISBN-13 (<https://isbn-information.com/index.html>)
- ・ 流通システム開発センター (http://www.dsri.jp/jan/check_digit.html)
- ・ 総務省令第八十五号

なお、上記のような代表的計算式の変形版を使うこともある。

参考 2 行政機関保有コードの現状と課題

内閣官房管轄の Web サイト「政府 CIO ポータル」では、各政府機関の保有するコードを区分別に整理し、表形式で公開している。政府機関が保有するコードは合計 447（2015 年 7 月時点）あり、区分別のコード数は以下のとおりである。

表 政府機関が保有する各区分のコード数及び例

区分	コード数	例
行政	82	政令指定都市番号、都道府県コード
地域	26	土地利用分類コード、地域番号一覧、郵便番号
自然	11	河川コード、島コード、海域区分コード
産業	25	日本標準産業分類、産業分類コード
交通	3	港符号表、港コード
医療	178	医療機関番号、二次医療圏
教育	6	大学番号一覧、高等学校等コード表
統計・調査	5	刊行頻度コード、自治体中間標準レイアウト
その他	111	地震コード、品種コード、住所例外コード
合計	447	-

出所：内閣官房「政府 CIO ポータル」の公開資料（<https://cio.go.jp/codes>）より作成

これらのコードの中には、類似したものが存在していたり、マスターコードとの関係性が不明確なものがあるなど、更なる整理が必要である。