

コード（分類体系）導入実践ガイドブック

2019年（平成31年）3月28日

内閣官房情報通信技術（IT）総合戦略室

〔標準ガイドライン群 I D〕

1018

〔キーワード〕

識別、分類、有意コード、無意コード、標準コード、階層化、マッピング、コードの公開

〔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、個々の情報に割り振られた記号（英数字など）です。コードを使うことで、コンピュータ間でデータを交換する際の通信量の削減、データの一意性担保や分類等が可能となります。組織横断的に共通のコードを使うことで、組織間のデータ交換を効率的に行えるようになります。

本ガイドブックではコードの設計及び導入並びに運用について説明します。

改定履歴

改定年月日	改定箇所	改定内容
2019年3月28日	-	初版決定
2017年8月1日	-	β 版公開

目次

目次	i
1 はじめに	3
1.1 背景と目的	3
1.2 適用対象	3
1.3 位置づけ	3
1.4 用語	4
2 基本事項	5
2.1 コード（分類体系）の概要	5
2.2 標準化されたコード活用の意義	6
2.3 行政機関等が提供するコード	6
1) 国際標準のコード	6
2) 政府標準のコード	7
3) 業界全体の標準コード	7
3 コードの設計	7
3.1 コード設計の手順	8
1) コード化対象の明確化	8
2) 既存コードの調査	8
3) コードの意味の明確化	8
4) 有意コード、無意コードの検討	8
5) 桁数、文字種の検討	9
6) 階層化の検討（上位コード、下位コード）	9
7) 関連主要コードとの関係整理	9
8) チェックディジットの付加	10
9) コード表記の検討	10
3.2 既存コードの活用と拡張	10
3.3 コードマッピングの実施	11
1) 分類項目が異なる場合	11
2) 分類の粒度が違う場合	11
3.4 コード設計に当たっての留意事項	11
1) 直交性	11
2) 英語表記	12
4 コードの導入と運用	12
4.1 コード情報の提供	12

4.2	コード仕様と利用規約の提供	13
4.3	コードの変更手順	13
1)	コードの削除	13
2)	コードの更新	14
3)	コードの並行運用	14
5	参考情報	15
5.1	チェックディジット	15
5.2	行政機関保有コードの現状と課題	16

1 はじめに

1.1 背景と目的

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ました。

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ました。

平成28年12月に官民データ活用推進基本法（平成28年法律第103号）が制定され、世界最先端デジタル国家創造宣言・官民データ活用推進基本計画が平成30年6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ります。データの活用や流通にはデータの標準化などが重要になることから、データの標準化等に関するガイド類の整備が求められているところです。

現在、行政機関は400を超える分類体系コード（以下「コード」という。）を持っています。この中には、日本標準産業分類のようにJISに定義された標準コードだけでなく、独自に設計されたコードが存在します。

行政機関が設計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少なくありませんが、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おらず、類似するものがあるなど、行政機関を跨ってシステム連携する場合、コード間の整合性の確保が必要となります。

本ガイドブックは、行政機関において新規コードの設計及び現行コードの利用等に携わる方に向けて、コードの設計や運用の指針を示すことを目的として策定しました。

また、公開後も、利用者のご意見をいただきつつ、継続的な改定を行ってまいります。

1.2 適用対象

政府情報システムを対象とします。

1.3 位置づけ

本ガイドブックは、標準ガイドライン群の一つとして位置づけられます。

1.4 用語

本ガイドブックにおいて使用する用語は、本ガイドブックに別段の定めがある場合を除くほか、標準ガイドライン群用語集の例によります。

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）のことです。

コードは、他組織、他分野のシステムとの連携を容易にする観点から、既に標準化されているコードや、広く一般的に利用されているコードに準拠することを推奨します。独自の設計に基づくコードを運用している場合、システム間での連携において、相互の差異を埋める処理が必要となります。

コードは大きく「識別のためのコード」と「分類のためのコード」に分けられます。「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、製品コード等のように個別の対象物にコードがつけられたものです。

「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す。

また、コードの付与方式には、有意コードと無意コードがあります。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す。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して管理するような方式です。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す。各コードの特徴は下表のとおりです。

表 1 有意コード・無意コードの特徴

	区分
	有意コード
	無意コード

2.2 標準化されたコード活用の意義

ISO や JIS、業界団体等によって標準化されたコードを使うことで、他システムと連携する際に、同じコードが別のものを指すことを避けることができます。

例えば、契約や表彰等の政府が公開している法人関連情報を集約して提供する法人インフォメーションでは、国税庁が公開する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、ワンストップでの情報提供を可能としています。

また、公共調達で使われる調達物品の分類コードは、各調達機関で異なっており、同じ物品でも異なるカテゴリに分類されます。コードを標準化することで、各機関のコード定義に依存しない一元的な情報収集が可能になります。

上記のとおり、行政機関間や官民でデータを融合して活用するときキーとなるのがコードです。グローバルに情報交換する場合にも共通のコードを使って各種情報を連携することで効果的に情報交換が行われる等、標準化されたコードはデータを活用する上で欠かせません。

一方で、既にコード体系を保有している場合、標準コードへの見直しには多大なコストがかかるとともに、コードの変更に伴う周辺システムへの影響が大きい場合もあります。見直しコストに見合った効果が期待できない場合は、変換表（マッピング）等で整合性を確保しつつ、連携を行う方法もあります。

2.3 行政機関等が提供するコード

行政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く、データの整理や活用の鍵になることから、機械判読に適し、二次利用可能なオープンデータとしての公開が期待されています。国際標準やデファクトな標準があるときにはそれらの利用を優先する等、より広範に使えるコードを優先します。

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、主に以下のものが挙げられます。

表 2-1 国際標準機関が定めたコード体系
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
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

各府省が定めたコード体系には、主に以下のものが挙げられます。

表 2-2 各府省が定めたコード体系

	発行者	概要
日本標準産業分類	総務省	統計調査の結果を産業別に表示する産業分類コード
法人番号	国税庁	法人や団体に指定する一意の識別コード
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) 業界全体の標準コード

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す。例えば鉄道事業者間で使用するサイバネコードのようなものがあります。

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、設計時間や費用が無駄になるだけでなく、品質の劣化などにも繋がるおそれがあります。また、コード設計は、将来の追加・変更等を見据えた中長期的な視点で行う必要があります。

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します。

- ・ 既存のコードがあるときには、そのコードを利用します。
- ・ 既存のコードがあるがそのまま使えないときには、分類を追加したり、扱える件数を増やす等の拡張を行います。
- ・ 新しくコードを設計する場合には、中長期的な利用やデータ連携を前提

として設計します。

- ・ 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、コードの変更を行います。

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します。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問題ありません。

1) コード化対象の明確化

分類項目の整備や一意性の確保など、コード化するデータ項目を明確にします。

2) 既存コードの調査

コード化対象に適合する既存コードが存在しないか調査します。行政機関が保有するコードについては、政府CIOポータルサイトのコード一覧 (<https://cio.go.jp/codes>) ※で確認できます。利用目的に類似するコードが存在する場合は、3.2及び3.3も参照し、将来の追加・変更等の可能性も含め、活用を検討します。類似コードが存在しない場合は、3.4に示す考え方にに基づき新たなコードを設計します。

※ 新規に設計又は変更したコードは、コード情報の蓄積や、重複等の防止のため、内閣官房情報通信技術（IT）総合戦略室（当該Webページに記載の連絡先）に報告することを推奨します。

3) コードの意味の明確化

コードの各項目について説明します。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等を定義し、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作成します。

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います。

頻繁にコードの追加や変更が行われるような場合、固定的な設計になる有意コードでは拡張が難しいため、無意コードの採用を推奨します。

また、将来的にコードを付与する対象データ数がどの程度増加するか想定できない場合も、コード数の増加に柔軟に対応できる無意コードの採用を推奨します。

5) 桁数、文字種の検討

コードの使用範囲及び使用期間を検討し、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ります。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します。

6) 階層化の検討（上位コード、下位コード）

コード自体の管理やデータの管理及び検索を容易にするため、コードの階層化を検討します。例えば大分類、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できるコードもあります。

表 3-1 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ます。

7) 関連主要コードとの関係整理（マッピング）

関連主要コードとの関係を整理し、その対応表を公開します。また、1:nの対応方法等、関係整理の方法などを記述します。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以下の表で示します。

表 3-2 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手法として、チェックディジットがあります。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します（具体的な付与方法については、参考5.1を参照）。

9) コード表記の検討

111-1111、9999-9999-9999-9999 といったように、3、4桁ごと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します。

3.2 既存コードの活用と拡張

コード化対象である情報が、既にコード化されている場合、そのコードを利用します。目的の違いなどで既存コードを拡張する場合には、参照するコードとの整合性を取りながら拡張していくことを推奨します。

既存コードの拡張には、以下のような例があります。

- ・ 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

既存の分類に対して、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、参照する既存コードと拡張・変更した部分を明確化しておく必要があります。

複数のコードがある場合には、それぞれのコードについて活用性・拡張性を検証し、最適なコードを選択します。

3.3 コードマッピングの実施

コード設計後、既存のコードと、独自に設計したコードを示すマッピング表を作成します。これにより、他システムから連携するときに、コード間の差異を確認して、コード変換等の対応ができるようになります。

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」は同一ではありません。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、又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ません。

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、以下の表のように対応させます。ただし、コードAとコードBは1：1の関係ではないため、相互に正確なデータ交換が行えることを保証するものではないことに留意が必要です。

表 3-3 分類項目が異なる例

コードA	コードB
動植物園	動物園
	植物園

中長期的には共通のコード体系を使うことが理想的ですが、システム更改時などのタイミングで揃える、又はマッピングの精度を上げることが必要です。

3.4 コード設計に当たっての留意事項

その他、コード設計の留意事項は以下のとおりです。

1) 直交性

コードとコード付与対象の結びつきを一意にします。同じ対象を表現するコードが複数定義されてはいけませんし、逆に同じコードで複数の対象を示

すように設計してもいけません。割当先の情報（分類）に独立性を持たせません。

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ります。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るべきです。

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ます。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（単語の頭文字）だけを大文字で示すキャメルケースかの選択をします。

【スネークケースの例】

japanese_government

【キャメルケースの例】

japaneseGovernment

4 コードの導入と運用

コード設計等の後、運用管理するに当たっては、以下に留意します。

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こととし、可能な限りコード情報を提供する API も公開します。これにより、コードが指す内容の確認が容易になるほか、新規設計時のコードの重複設定を防止しやすくなります。API について、詳しくは「API 導入実践ガイドブック」を参照してください。

また、入力ミス防止や入力効率化のため、状況に応じてバーコードやQRコードでの提供を検討します。

コード体系（行政機関コード、全国地方公共団体コード等）のメタ情報として以下を明記します。これらの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ます。

- ・コードオーナー
- ・コード体系名

- ・コード体系バージョン
- ・発行日
- ・更新履歴

保有するコードを Microsoft Excel ファイル等の表作成ソフトで作成したファイル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す。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います（例：012345→12345）。これを防ぐため、先頭の文字を 0 以外の文字にするか、記号や文字列を入れ、数値として扱わないようにすることを推奨します（また、数字のみで構成されるコードについては、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ことで自動変換を避けることができます。)

4.2 コード仕様と利用規約の提供

コードを公開する場合には、以下の例のようにコードの各項目の定義を明示して公開します。

- (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します。

4.3 コードの変更手順

やむを得ずコードの変更（削除、更新）をする際には、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、コードを公開している Web サイトや、各行政機関 Web サイトの更新情報一覧等に、可能な限り早い段階で一定期間の告知を行います。その際、変更等することになった背景やその内容も付記することが望ましいです。

1) コードの削除

コードを割当てた項目の廃止（製品の生産中止等）等より削除したコードは、原則として再利用しないことを推奨します。ただし、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、利用しない期間を相当程度確保した上で、当該コードが再利用しているものであることを明確に示します。

2) コードの更新

時間の経過によって、当初想定していなかったコード化範囲の変化が発生し、これによるコード割当てが増加のため、現行のコード体系を維持できなくなる場合があります。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等が考えられますが、連携システム等への影響も勘案して対処方法を検討する必要があります。

また、コードの変更等を行った場合には、その内容を公開します。公開すべき項目は以下のとおりです。

- ・バージョン
- ・変更日
- ・変更理由
- ・変更点

コード変更等の際には、継続性の観点から、改定版と旧版との間で比較可能とすることが重要です。旧版のアーカイブを提供するとともに変換表や変換 API 等のソフトウェアの提供を行うことを推奨します。

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があります。

5 参考情報

5.1 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ものです。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ります。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ます。

チェックディジットの計算方法には様々なものがあります。例を以下の表で示します。

表 4-1 チェックディジットの計算方法

アルゴリズム名
Damm アルゴリズム (推奨)
モジュラス 10 ウェイト 2・1 (Luhn アルゴリズム)
モジュラス 10 ウェイト 3・1
モジュラス 11

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)
- ・ 行政手続における特定の個人を識別するための番号の利用等に関する法律の規定による通知カード及び個人番号カード並びに情報提供ネットワークシステムによる特定個人情報の提供等に関する省令（平成 26 年総務省令第 85 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ります。

5.2 行政機関が保有するコードの現状と課題

内閣官房のWebサイト「政府CIOポータル」では、各行政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ます。行政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す。

表 5 行政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
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」（<https://cio.go.jp/codes>）より作成