

■企画工程用

・本チェックリストは、職員がデータを収集する際に実施すべき項目をまとめたものです。

No	分類	チェック項目	NG例	チェック欄
1	正確性	アンケートや調査票を作成する際に、書式を指定できるところについて、入力欄の書式を指定しているか	<ul style="list-style-type: none"> ・生年月日の入力欄に全角ひらがなや記号を入力できてしまう ・郵便番号を入力する欄で、ハイフンで区切ったものと区切らないものが混在してしまう 	
2		アンケートや調査票を作成する際に、入力する内容があらかじめ決まっているものについて、入力できる内容をプルダウン形式などで指定しているか	<ul style="list-style-type: none"> ・施設名を入力する欄に文字列を自由に入力できてしまい、誤った施設名のまま入力を完了できてしまう 	
3		アンケートや調査票を作成する際に、明らかな誤入力を防ぐ設計にしているか	<ul style="list-style-type: none"> ・日付を入力する欄に、「平成35年」などと明らかな誤りを入力できてしまう ・郵便番号を8桁で入力できてしまう 	
4	完全性	アンケートや調査票を作成する際に、利用目的と照らして必要なデータ項目が網羅されているか	<ul style="list-style-type: none"> ・アンケートを通してあるサービスについての利用者満足度を年齢や住まいなどの属性別に分析することを想定しているものの、利用者の属性に関する設問が設定されていない 	
5		必須項目に入力漏れがある場合、入力漏れがわかるような仕組みになっているか	<ul style="list-style-type: none"> ・必須項目の入力漏れがあるにもかかわらず、データの入力を完了できてしまう ・選択式の入力欄で、3つ以上の選択が必須であるにもかかわらず、1つしか選択されていない状態で入力を完了できてしまう 	
6	一貫性	同一ファイル内でデータの矛盾はないか	<ul style="list-style-type: none"> ・同じ施設に対して施設コードが2つ存在している ・あるセルでは部署の正式名称が記載されている一方、別のセルでは略称が記載されている 	
7		関連するファイル間でデータに矛盾はないか	<ul style="list-style-type: none"> ・施設の敷地面積が、あるファイルでは10,000㎡で、別のファイルでは9,000㎡となっている ・あるファイルでは、データの名称が「面積」、別のファイルでは名称が「施設面積」となっており、同じデータを指しているか分からない ・あるファイルのセルが別ファイルのセルを参照しているとき、参照先が間違っている 	
8	可用性	必要な時にいつでもデータにアクセスできるようになっているか	<ul style="list-style-type: none"> ・ファイルの名前や格納場所が頻繁に変わり、必要なファイルに辿り着けない ・Excelファイルを庁内のポータルサイト上に公開しているが、所有者がファイルを編集している時間はファイルがポータルサイト上から消えており、アクセスできない 	
9	最新性	公開データの更新サイクルが元データの更新サイクルに対して適切に設定されているか	<ul style="list-style-type: none"> ・元データは月次で更新されている一方で、公開データは半年に1回しか更新されない ・運用の際の更新頻度について方針が定められておらず、担当者が変わると更新されなくなってしまう 	
10		データの特徴を踏まえ、データが収集時から十分に短い期間で公開されているか	<ul style="list-style-type: none"> ・データを収集してから公開するまでに半年近くかかる 	
11		ファイル等で提供される場合は、最終更新日時及び最新版の所在が明記されているなど、更新版の有無が確認できるようになっているか	<ul style="list-style-type: none"> ・アンケートや調査票の結果がいつ時点のデータか確認できない 	

■要件定義工程用

・本チェックリストは設計・開発工程の仕様書を作成する際に、データ品質の要件として考慮すべき観点をまとめたものです。

No	分類	チェック項目	要件が設計・開発で反映された際の機能・設計の具体例	チェック欄
1	正確性	利用者が入力するデータを間違えないよう、バリデーションチェック等の仕組みを設定すること	<ul style="list-style-type: none"> ・書式を指定できるものについては、入力欄の書式を制御すること ・入力する内容があらかじめ決まっているものについては、入力できる内容をプルダウン形式等で制御すること ・明らかな誤入力は制御すること 	
2	完全性	データの入力漏れを制御する等、利用目的の達成に必要なデータが漏れなく入力されるようにすること	<ul style="list-style-type: none"> ・利用目的と照らした際に必要なデータ項目は網羅すること ・必須項目に入力漏れがある場合、エラーが表示されるような設計にすること 	
3	一貫性	入力データに矛盾がないよう制御する等、システム内及び東京都の既存システムとの連携においてデータの一貫性を保てるようにすること	<ul style="list-style-type: none"> ・原則、システム内で一意に主キーを設定すること ・システム内では入力データに矛盾がないように制御すること ・システム間でデータに矛盾がないように制御すること 	
4	可用性	データを常に利用可能な状態にすること	<ul style="list-style-type: none"> ・必要な時にいつでもデータにアクセスできるようにすること ・データを庁内又は庁外に公開するシステムは常時稼働させること 	
5	最新性	データの利用目的に適した更新頻度で庁内又は庁外に公開するデータを更新すること。また、更新時には最終更新日時や、ソースファイル等があればその所在をシステム内に明記すること。	<ul style="list-style-type: none"> ・庁内又は庁外に公開するデータの更新サイクルは元データの更新サイクルに対して適切に設定すること ・データの特徴を踏まえ、データが収集時から十分に短い期間(※)で公開されていること ・ファイル等で提供される場合は、最終更新日時及び最新版の所在が明記されている等、更新版の有無が確認できるようになっていること 	

■設計・開発工程用

・本チェックリストはシステムの設計や開発工程において、仕様書で定めたデータ品質の要件が適切に設計・開発に組み込んでいるかを確認する項目をまとめたものです。

No	分類	チェック項目	NG例	チェック欄
1	正確性	書式を指定できるものについて、入力欄の書式を制御しているか	<ul style="list-style-type: none"> ・生年月日の入力欄に、全角ひらがなや記号を入力できてしまう ・郵便番号を入力する欄で、ハイフンで区切ったものと区切らないものが混在してしまう ・パーセントによる表記と小数による表記が混在してしまう 	
2		入力する内容があらかじめ決まっているものについて、入力できる内容をプルダウン形式などで制約しているか	<ul style="list-style-type: none"> ・施設名を入力する欄に文字列を自由に入力できてしまい、存在しない施設名を入力できてしまう 	
3		明らかな誤入力を制御しているか	<ul style="list-style-type: none"> ・日付を入力する欄に、「平成35年」などと明らかな誤りを入力できてしまう ・郵便番号を8桁で入力できてしまう 	
4	完全性	利用目的と照らした時に必要なデータ項目が網羅されているか	<ul style="list-style-type: none"> ・経費精算システムで「経費の金額」がデータ項目に含まれていない ・個人を識別する必要があるにもかかわらず、「従業員番号」がデータ項目に含まれていない 	
5		必須項目に入力漏れがある場合、入力漏れがわかるような仕組みになっているか	<ul style="list-style-type: none"> ・必須項目の入力漏れがあるにもかかわらず、データの入力を完了できてしまう ・選択式の入力欄で、3つ以上の選択が必須であるにもかかわらず、1つしか選択されていない状態で入力を完了できてしまう 	
6	一貫性	システム内で一意に主キーを設定しているか	<ul style="list-style-type: none"> ・事業者ごとに主キーが設定されていないため、類似した名称の事業者が存在する場合、それが同じ事業者なのか、異なる事業者なのか分からない 	
7		システム内で入力データに矛盾がないように制御しているか	<ul style="list-style-type: none"> ・入力が不要な欄にデータを入力できてしまう ・あるデータがシステム内の別データを参照しているとき、参照先が間違っている 	
8		システム間でデータに矛盾がないように制御しているか	<ul style="list-style-type: none"> ・施設の敷地面積が、あるシステムでは10,000㎡で、別のシステムでは9,000㎡となっている ・あるシステムでは、「1」と整数で登録される仕様となっている一方で、別のシステムでは「1.00」と小数点第2位の表記で登録されている仕様となっている。両者とも「1」だが、システム上はエラーとなる可能性がある ・あるシステムでは、文字コードが「UTF-8」である一方で、別のシステムでは「シフトJIS」となっている。文字コードが異なるため、文字化けしてしまう 	
9	可用性	必要な時にいつでもデータにアクセスできるようになっているか	<ul style="list-style-type: none"> ・データ公開システムにアクセス可能な時間帯が限定されている ・同じデータを意味しているにもかかわらず、データの名称が「面積」になったり「施設面積」になったりと、頻繁に変わってしまい、必要なデータに辿り着けない 	
10		データを公開するシステムは常時稼働しているか	<ul style="list-style-type: none"> ・頻繁にシステムが停止する 	
11	最新性	公開データの更新サイクルが元データの更新サイクルに対して適切に設定されているか	<ul style="list-style-type: none"> ・元データは月次で更新されている一方で、公開データは半年に1回しか更新されない ・運用の際の更新頻度について方針が定められておらず、担当者が変わると更新されなくなってしまう 	
12		データの特徴を踏まえ、データが収集時から十分に短い期間で公開されているか	<ul style="list-style-type: none"> ・データを収集してから公開するまでに半年近くかかる 	
13		ファイル等で提供される場合は、最終更新日時及び最新版の所在が明記されているなど、更新版の有無が確認できるようになっているか	<ul style="list-style-type: none"> ・システムからダウンロードしたファイルがいつ時点のデータか確認できない 	