


政府資料標準訂定介紹

一起開始

111年9月版



政府資料標準訂定目的及方法

PART 1



政府資料標準訂定目的及方法

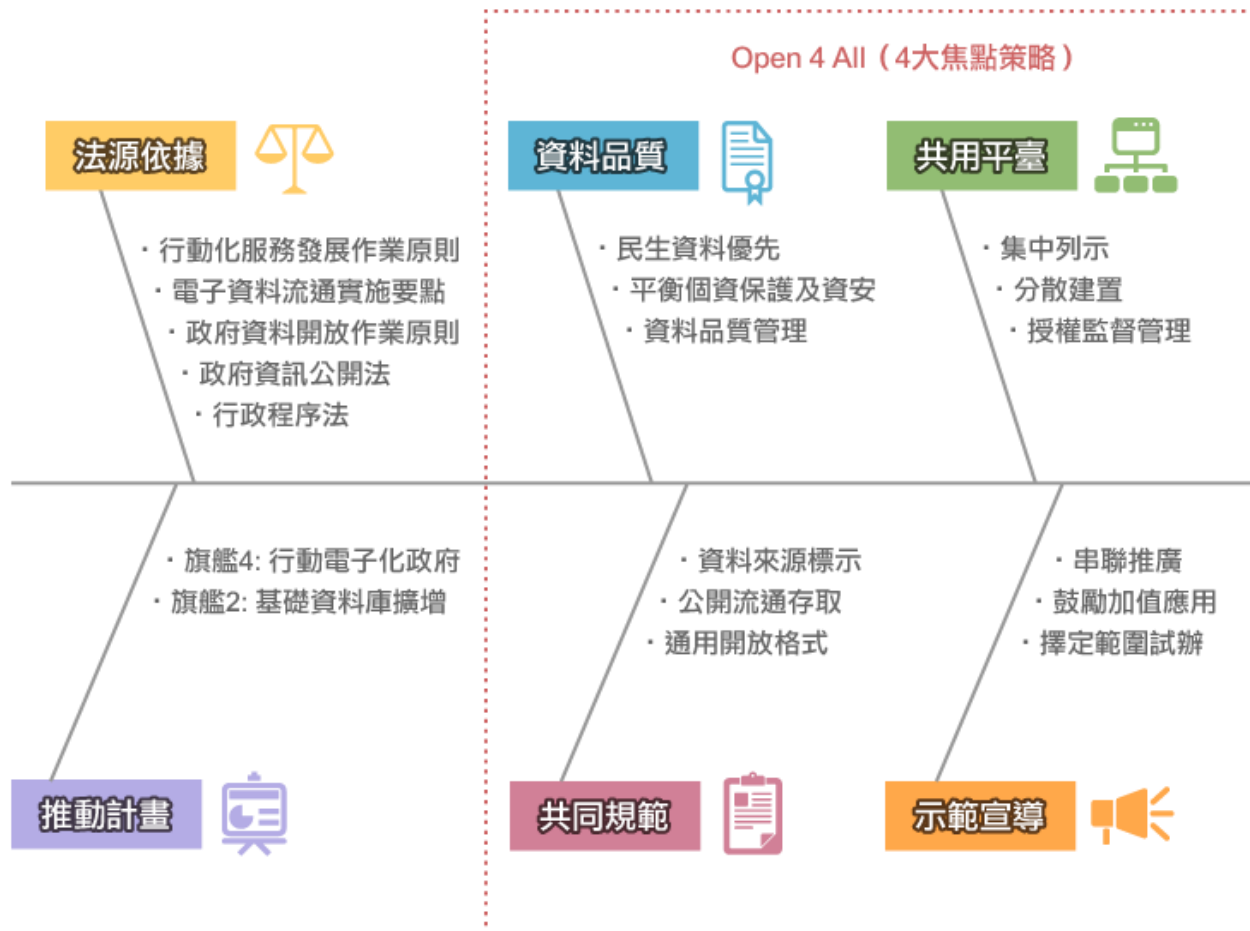
1) 資料開放後的挑戰

資料標準推動緣起-從開放資料談起

將政府蒐集的相關數據，供民間使用，發展更多食衣住行育樂

- 2015年及2016-2017年，「英國開放知識基金會」公布開放資料指標(Open Data Index)評比，臺灣連續2屆獲得第1名，在完整度、深度、廣度獲得國際肯定。

- 行政院宣示啟動OpenData深化應用元年，中央政府各機關設立政府資料開放諮詢小組，盼望促進資料品質提升與強化民間應用鏈結。



資料開放後的挑戰

各地方機關開放資料提供



民間團體接手開放資料



自102年4月政府推動資料開放以來，截至111年9月已累計開放逾5萬8,068項，並開放多項與公共利益及經濟發展密切相關之高應用價值資料集。

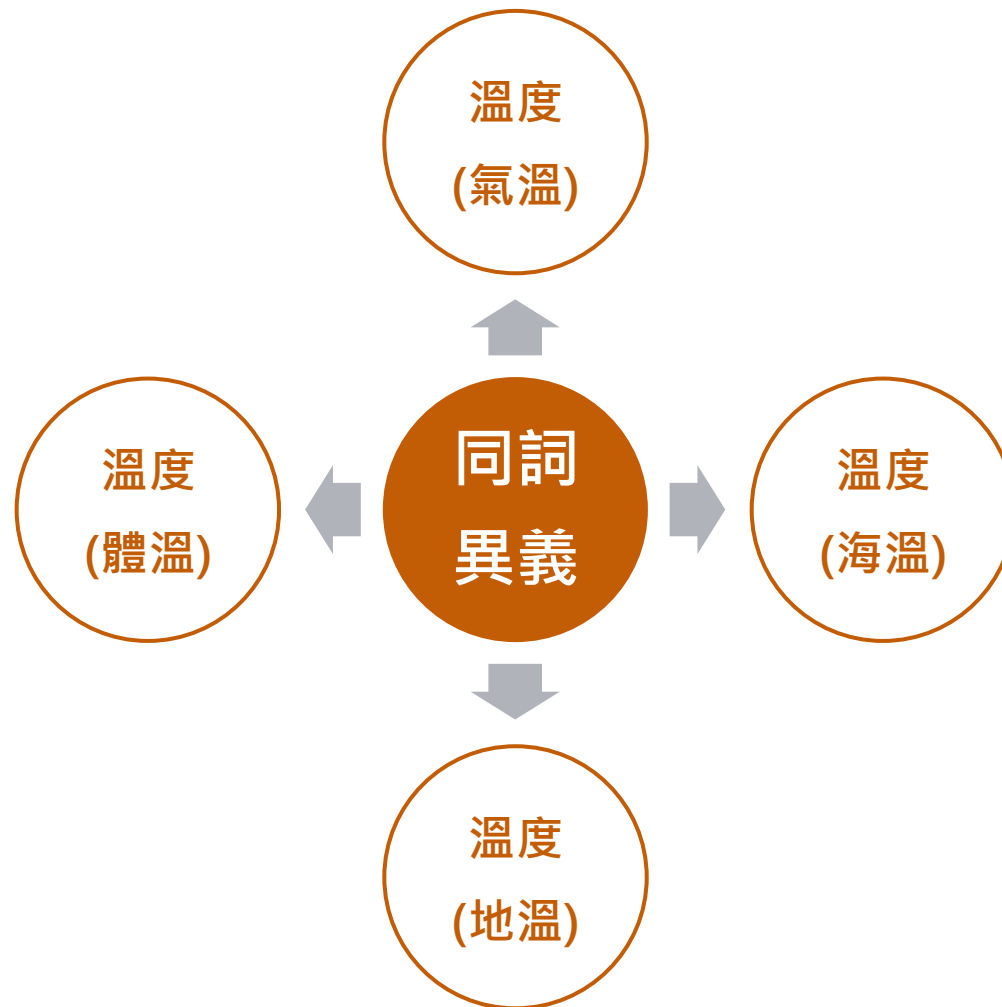
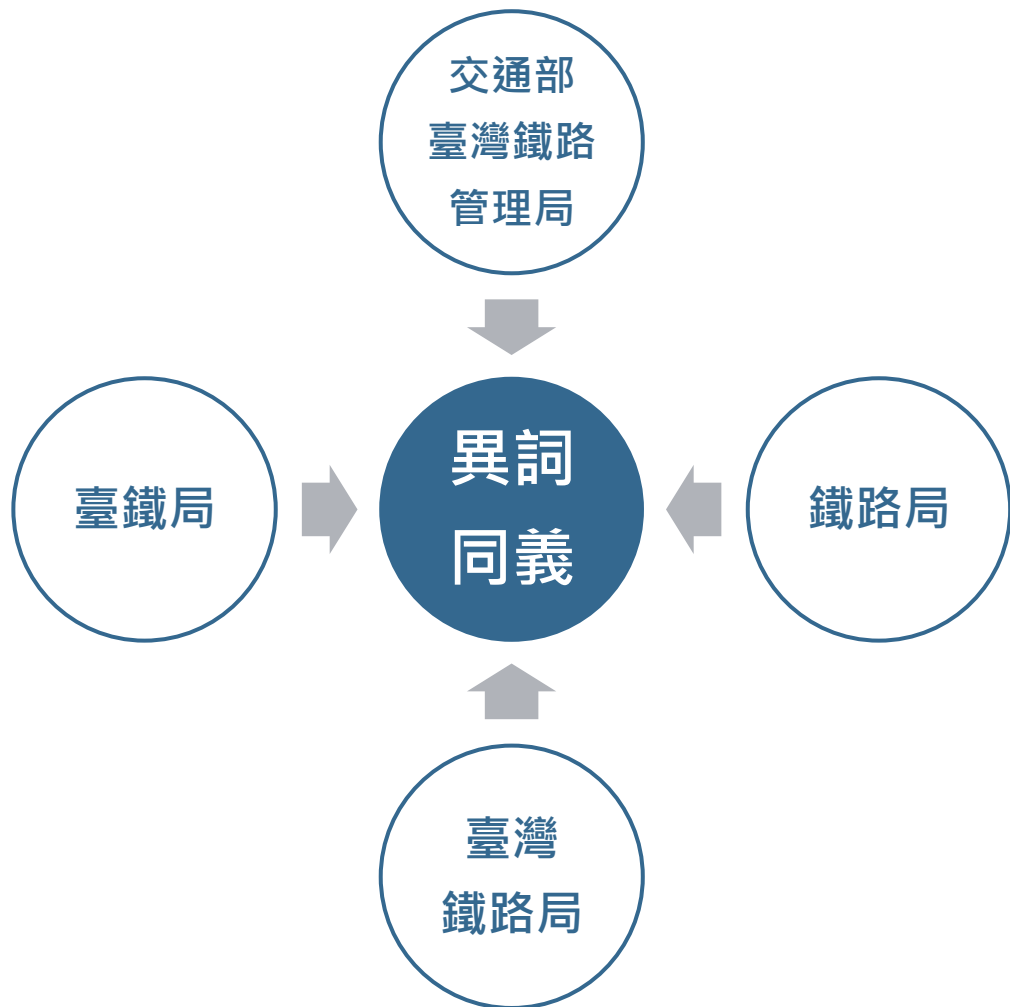
資料來源多樣、資料標準不同
使用門檻大幅提升
降低開放資料使用意願

資料標準不同產生的問題(1/5)

1

語意問題

例如異詞同義、同詞異義



資料標準不同產生的問題(2/5)

2

格式問題

例如全形/半形、資料型態、長度

12345 半形 A機關

1 2 3 4 5 全形 B機關

日期
2019/1/1
2019/1/31
2019/12/1
2019/12/31

VS

日期
2019/01/01
2019/01/31
2019/12/01
2019/12/31

↑ 長度不一致

↑ 長度一致

資料型態不同

中文名稱	英文名稱	說明	型別	值域	資料範例
西元年份	AD year	YYYY	String	[0000-9999]	2020
西元年	CE year	西元年(4位)	Number	[0000-9999]	2020
工作時數	Work hours	以小時為計算單位之工作時間	Number	整數	2、8
加班時數	Overtime hours	以小時為計算單位之加班時間	Number	浮點數	2.0、1.5

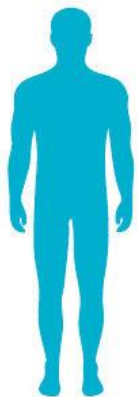
資料型態不同，
勾稽時會出錯

資料標準不同產生的問題(3/5)

3

值域問題

例如值域範圍不一致、不合理



正常人體體溫

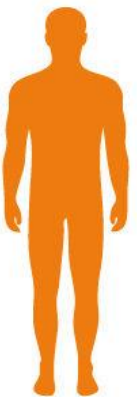
直腸(肛溫)	97.9°F (36.11°C) ~ 100°F (37.94°C)
耳溫	96.4°F (35.78°C) ~ 100.4°F (38°C)
舌下	95.9°F (35.5°C) ~ 99.5°F (37.5°C)
腋下	94.5°F (34.72°C) ~ 99.1°F (37.27°C)

A校

A幼兒園採肛溫判別發燒與否

B校

B幼兒園採腋溫判別發燒與否



發燒徵兆體溫

直腸(肛溫)	100.4°F (38°C) 或更高
舌下	100°F (37.8°C) 或更高
腋下	99°F (37.2°C) 或更高

收到溫度回報37.5度，有沒有發燒？是A校或B校的小朋友？

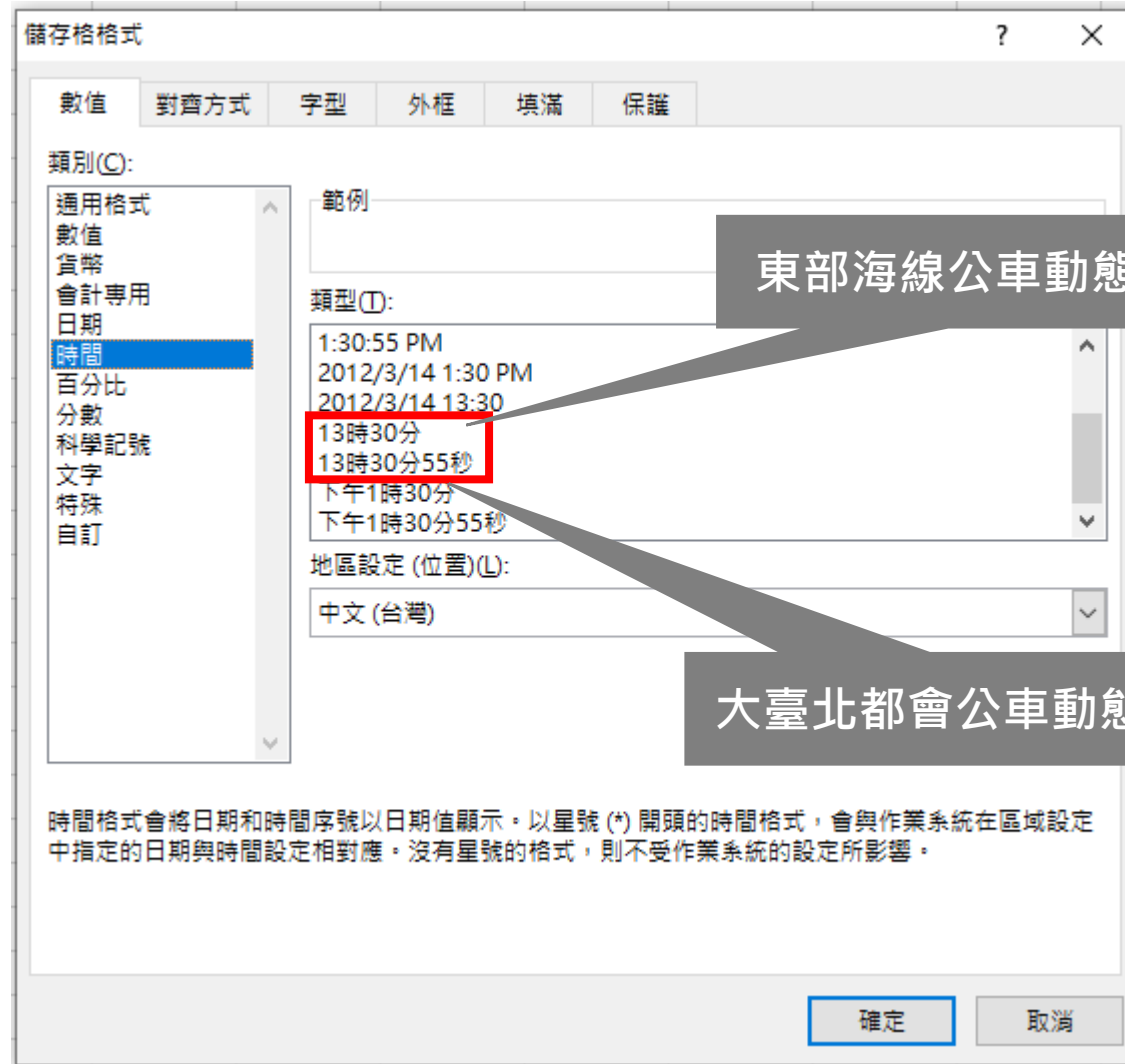
收到溫度回報96度，單位是攝氏°C或華氏°F？是測到高溫水還是美國小朋友的體溫？

資料標準不同產生的問題(4/5)

4

顆粒度問題

例如時間粗細程度不一致



資料標準不同產生的問題(5/5)

5

更新頻率問題

例如代碼表更新、動態資料產生頻率



A機關
更新時間

10:30



B機關
更新未知

?

13:30

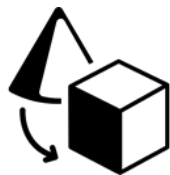
上述問題交換資料每次都需要進行 費工的資料處理

政府流通資料困難



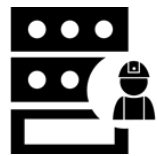
資料勾稽串接前要溝通資料內容

多次多方溝通，好費時



不同單位拋來的資料格式不一

要轉成一致，好費工



不同系統資料維護方式不一

多樣系統要維運多樣標準資料

上述問題使用資料每次都需要進行 客製的資料處理



民間應用資料困難



每個資料集都要寫一隻程式去處理

想想就覺得頭痛



有些欄位根本看不懂做什麼用

種類：A/B/C 到底是什麼意思



資料寫錯了，還要手動清理

大寫字填成小寫字，還要額外處理

面對各樣資料問題
積極迎接資料開放後挑戰

推動資料標準
達成資料品質、跨域資料整合

便利各界應用資料，發展創新服務





政府資料標準訂定目的及方法

2)做了資料標準有什麼好處呢?

資料標準是什麼

1

描述資料標準欄位

說明欄位作用，使用方式

2

設定資料型別

如字串、數字、代碼表規則

3

設立正規表示式

轉成機讀格式，方便開放資料檢核

該平臺將資料標準依據『領域』→『主題』→『類別』進行存放。

資料標準 [共通性](#) > [共通性](#) > [共通資料](#)

資料標準URI下載 [XML](#) [JSON](#)。本網站使用之正規表示式以 PHP 版本為主。

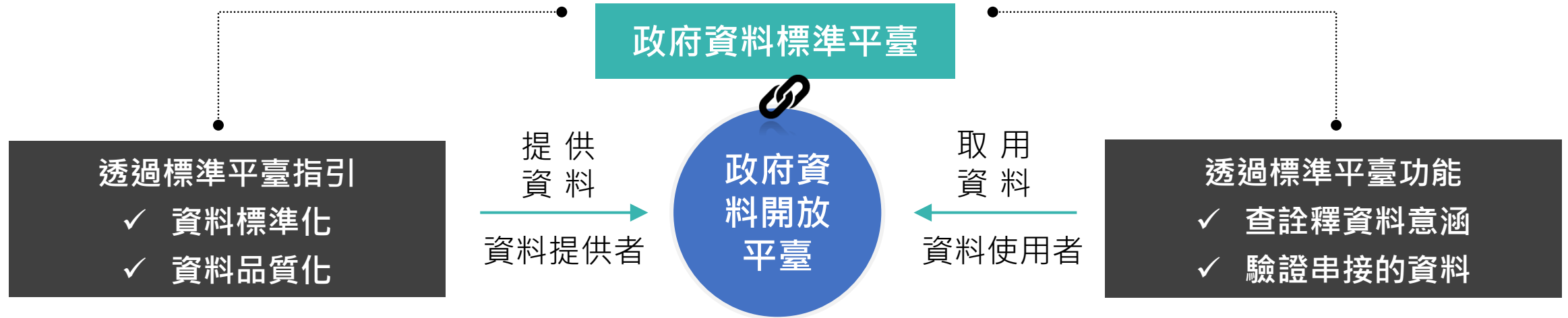
範例1：查找身分證字號的標準，可進入『共通性』→『共通性』→『共通資料』進行查找

Properties from 共通資料

中文名稱	英文名稱	說明	型別	值域	附註	附件連結	驗證
姓名	Name	姓氏與名字	String				
國民身分證統一編號	Id	辨別個人身分之專屬代碼	String	[A-Z][1-2][00000000-99999999]	第1碼編碼：縣市別 第2碼編碼：1為男性、2為女性		驗證
性別	Gender	依國民身分證統一編號第2碼編訂	String	[1-2]	1為男性 2為女性		驗證

資料標準有哪些優點

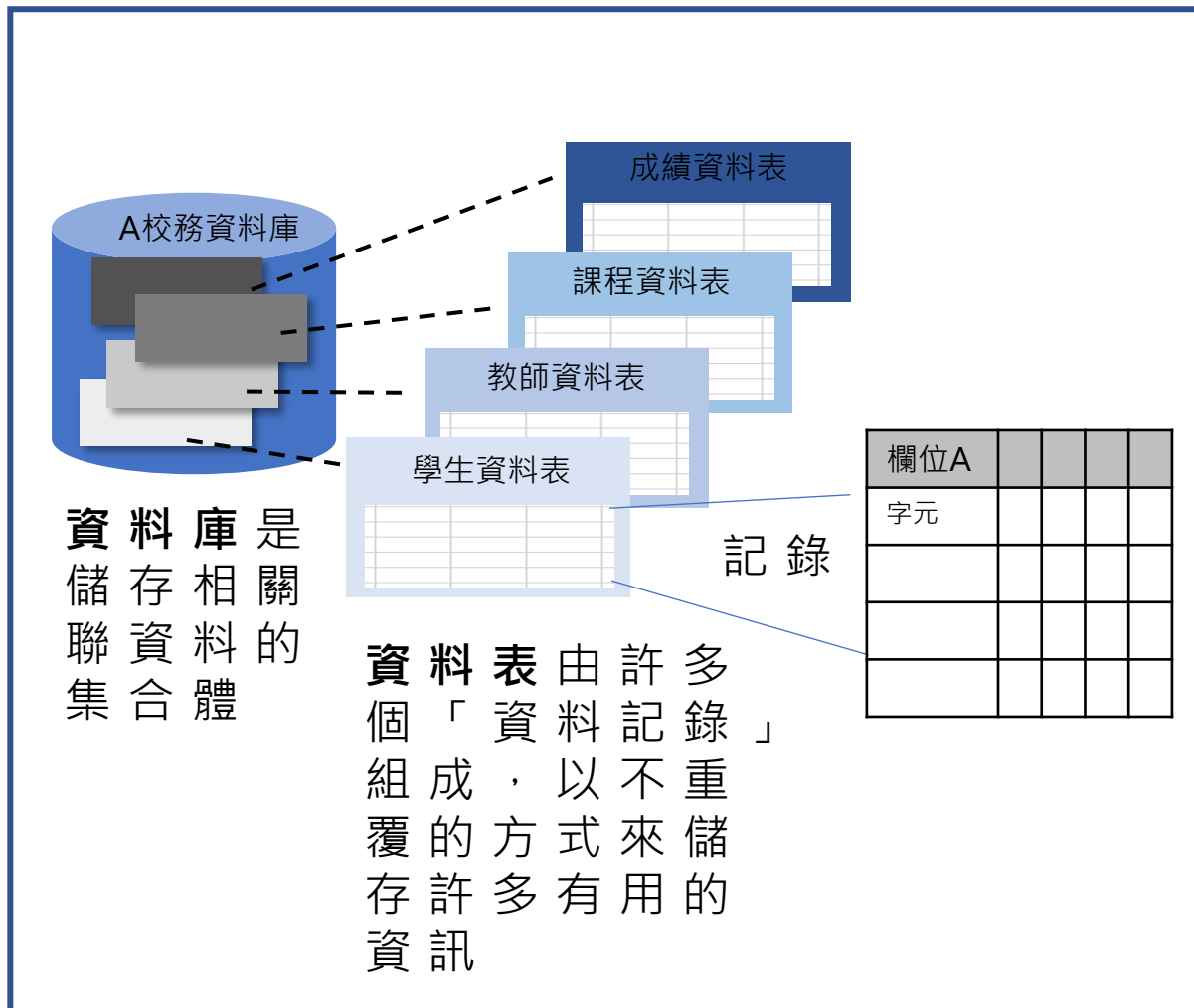
讓民間團體更快速瞭解資料樣貌



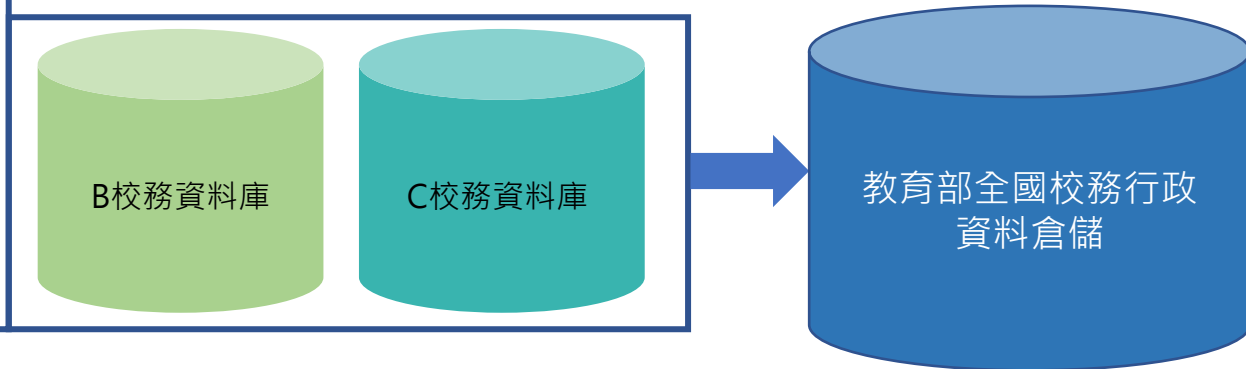
1. **資料容易溝通**：使不同系統或單位在取用資料時，有所依據及規範，有利於溝通與檢索資料意涵。
2. **資料容易串接**：使不同系統或單位在取用資料時，有一致的資料標準串接資料，彼此資料容易交換、整合及分析。
3. **系統容易維運**：使用資料標準化建置系統，促進資料和軟體的重複使用性，減少多樣系統要維運多樣標準不一致資料，所產生的維運成本。
4. **創造三贏**：綜上所述，初期各單位需投入較多資源及時間，但長期規劃其經濟效益將更大，亦使資料開放品質上升，有效達成政府、民眾、企業資料交換流通。

資料標準如何提升資料品質及促進流通(1/3)

資料層級的基本觀念 - 由簡到繁

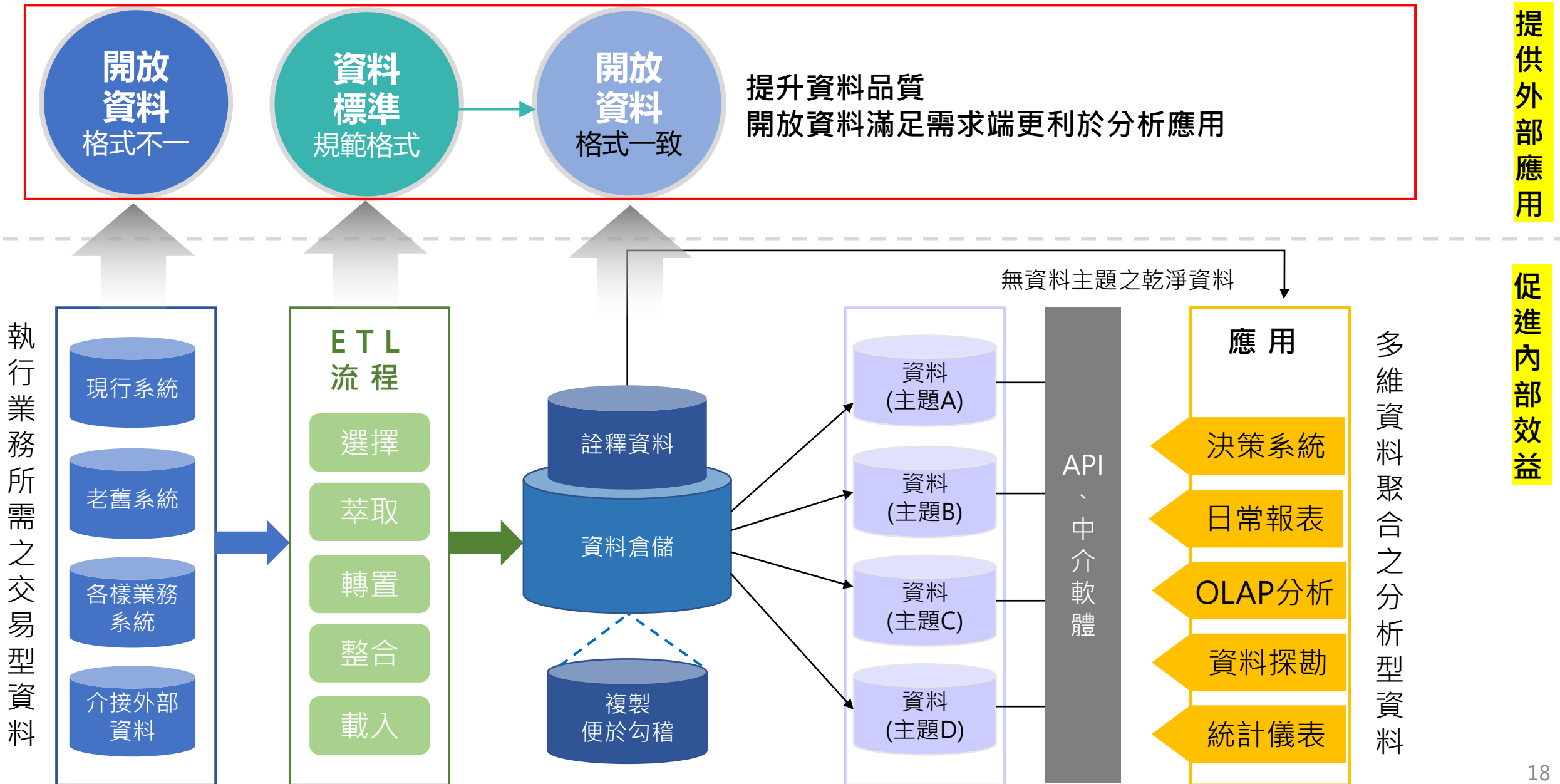


資料層級	層級描述	資料範例
欄位	由字元所組成	學生學號
資料記錄	數個欄位組合形成一筆記錄	學籍資料
資料表	由相同格式定義之紀錄所組成	全班學籍資料
資料庫	由多個相關資料表所組成	校務行政資料庫，包括：成績資料表、學籍資料表、選課資料表...等
資料倉儲	<ol style="list-style-type: none"> 1. 整合性的資料儲存體 2. 內含各種與主題相關的大量資料來源 3. 可提供企業決策性資訊 	教育部的全國校務行政資料倉儲，可進行彙整分析提供決策資訊



全國總共10,931所學校

資料標準如何提升資料品質及促進流通(2/3)



資料標準如何提升資料品質及促進流通(3/3)



臺北捷運廣場各場地活動日期及名稱

臺北捷運廣場各場地活動日期及名稱

主要欄位說明: **日期**、地點、活動名稱、辦理單位

臺北大眾捷運股份有限公司 / 詮釋資料更新時間: 2020/07/31 12:09

17 3 0

CSV

上傳驗證
通過資料

[5-7,A-Z]

驗證正規表示式

中文名稱: 日期

正規表示式:

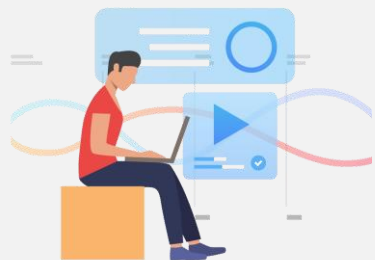
送出

中文名稱	英文名稱	說明	型別	值域	附註	附件連結	驗證
日期	Date	描述特定日子	Date	[0000-9999][01-12][01-31]	引用標準名稱: CNS 7648 「資料元件及交換格式-資訊交換-日期及時間表示法」 (ISO 8601 Data elements and interchange formats - Information interchange - Representation of dates and times)		<input type="button" value="驗證"/>

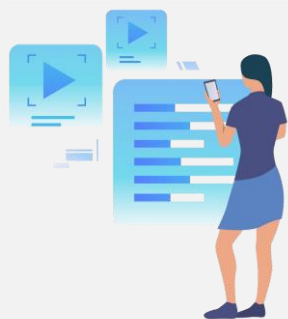
資料標準
日期定義及驗證功能

運用資料標準驗證提升資料開放品質

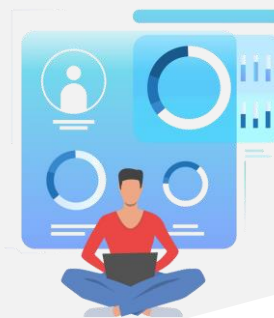
各機關上傳開放資料
到開放資料平臺



針對資料欄位
選擇驗證的資料標準



按照選擇的資料標準
機器自行驗證



通過資料標準驗證
民間團體更好取得資料



反覆
驗證

這個欄位可以對到...共通性>共通性>地址欄位
這個欄位可以對到...戶政>戶籍資料>個人戶籍資料>役別代碼
這個欄位好像找不到可以對應的資料標準...先跳過

對應這個欄位的資料有沒有符合資料標準
若未通過驗證...
機關資料上傳者則須依資料標準修改資料內容

推動政府資料標準三大重點策略

政府資料標準

各領域
建立資料
標準模型

各領域建立資料標準模型

各領域應以資料流通、資料治理目的為主，白金標章機制為輔，建立一致的資料標準模型。

建置
資料典

建置資料典

建置規範資料品質的資料典，讓使用者知道如何使用資料、清理資料。

協調
利害關係
流通對象

協調利害關係流通對象

資料能夠流通的對象越多，越能創造資料的加乘效果。

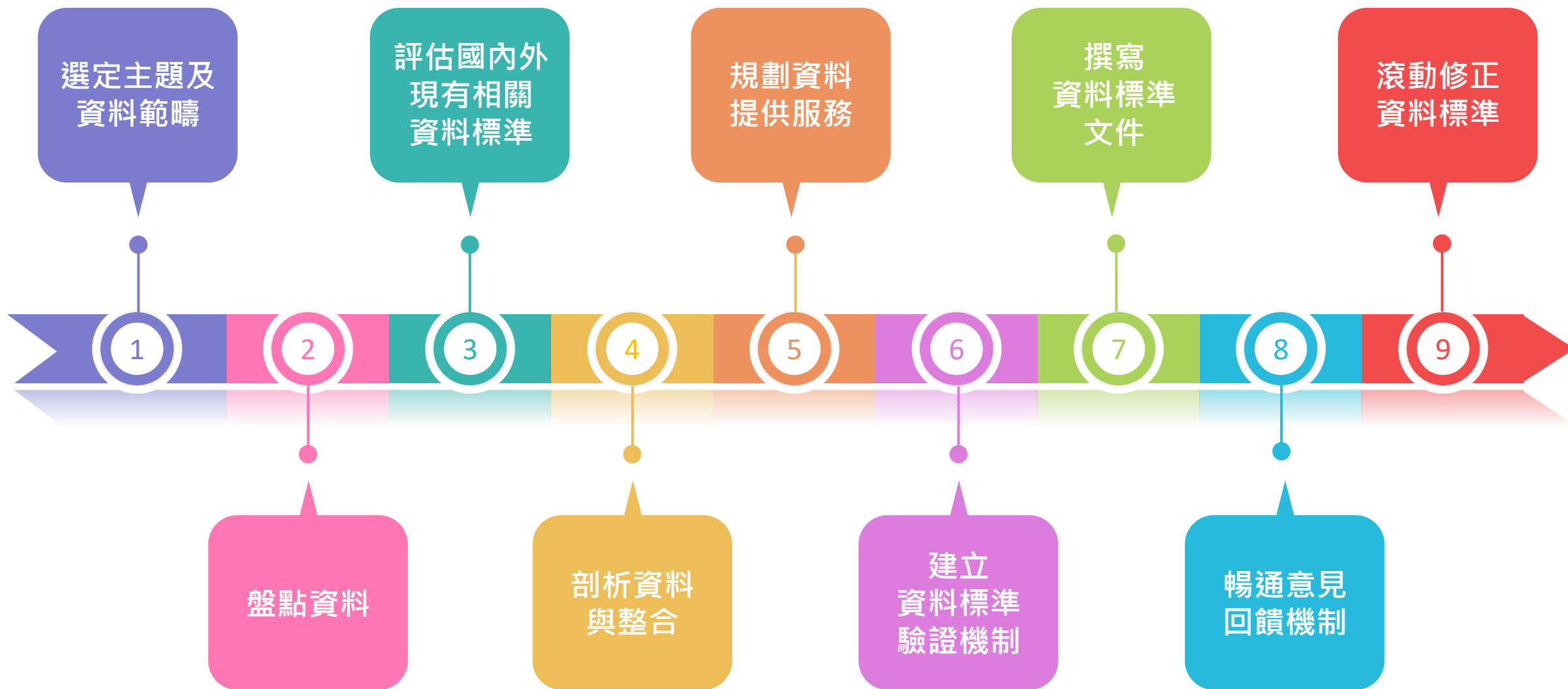


政府資料標準訂定目的及方法

3) 怎麼訂定資料標準呢?

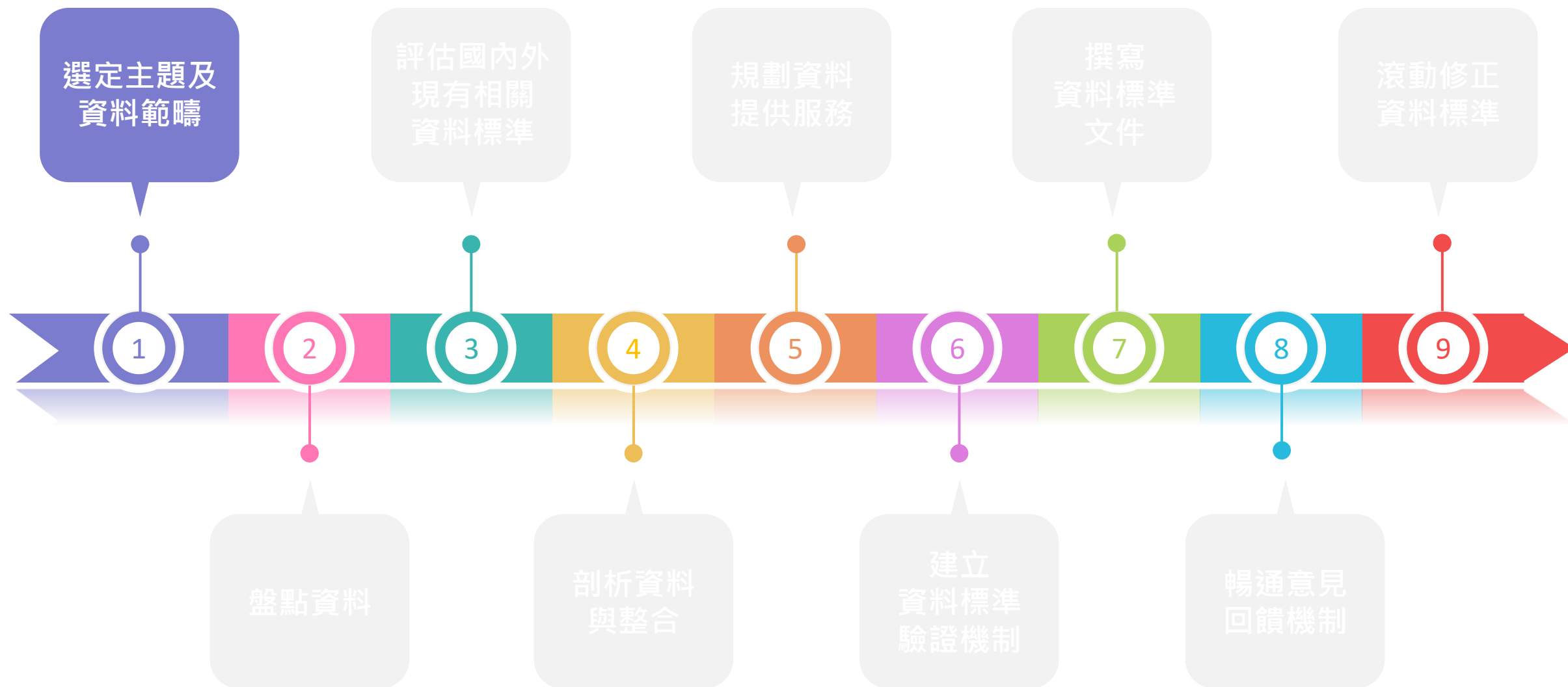
機關訂定資料標準方法流程(9大步驟)

詳參數位發展部「領域資料標準訂定流程參考指引」



機關訂定資料標準方法流程(步驟1)

詳參數位發展部「領域資料標準訂定流程參考指引」



1 資料儲存系統化

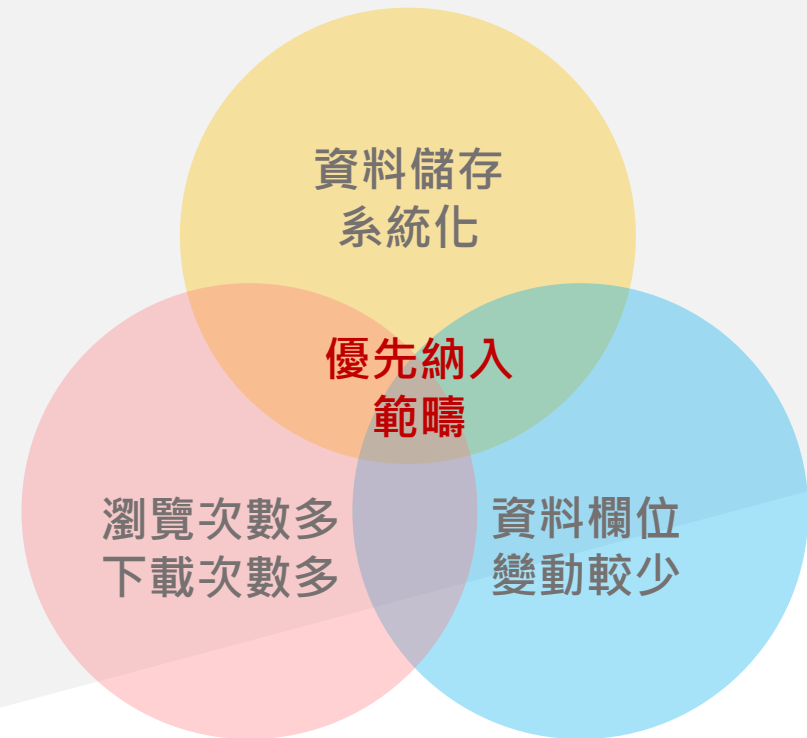
若資料更新係由系統產生及儲存
則將來資料調整與取用通用基準較容易

2 資料集下載瀏覽次數多

1. 開放/申請資料集下載量越多，代表民間使用者較有興趣之資料
2. 資料交換頻率及交換量較大者，代表機關流通需求之資料

3 該資料欄位較為固定不會經常變動

許多資料集欄位已經使用多時，欄位和輸入格式已經達到比較穩定的狀態



選定主題及資料範疇

在政府資料開放平臺，瀏覽開放資料集

範例：民眾感興趣的資料集 如：景氣指標及燈號

排序： 瀏覽次數多至少 10項/頁

公文電子交換系統地址簿

收錄公文電子交換系統之使用機關，以供電子交換系統使用

主要欄位說明: ORGID、ORGNAME、STATUSCODE、UPDATETIME

國家發展委員會檔案管理局 2022-02-09 10:45:37 更新 70753 16381 21

景氣指標及燈號

景氣指標、景氣對策信號(燈號及綜合判斷分數)

主要欄位說明: Date、工業生產指數(Index2016=100)、電力(企業)總用電量(十億度)、製造業銷售量指數(Index2016=100)、批發、零售及餐飲業營業額(...詳內)

國家發展委員會 2022-08-03 14:48:58 更新 67424 3116 11

中央機關一般職稱雙語詞彙

本資料集主要提供中央機關一般職稱之中英對照詞彙等資訊。

主要欄位說明: 一般職稱類別、一般職稱類別細目、一般職稱中文、職等、一般職稱英譯

國家教育研究院 2021-06-25 12:12:23 更新 27595 7177 1

CSV JSON

ZIP

CSV

1

2

3

每月固定產生，應該是系統產生

每月自動產生CSV檔案，動態生成

景氣指標查詢系統

https://index.ndc.gov.tw/n/zh_tw#/

互動良好，民間需求高

瀏覽人次破萬、下載千次，互動意見頻繁

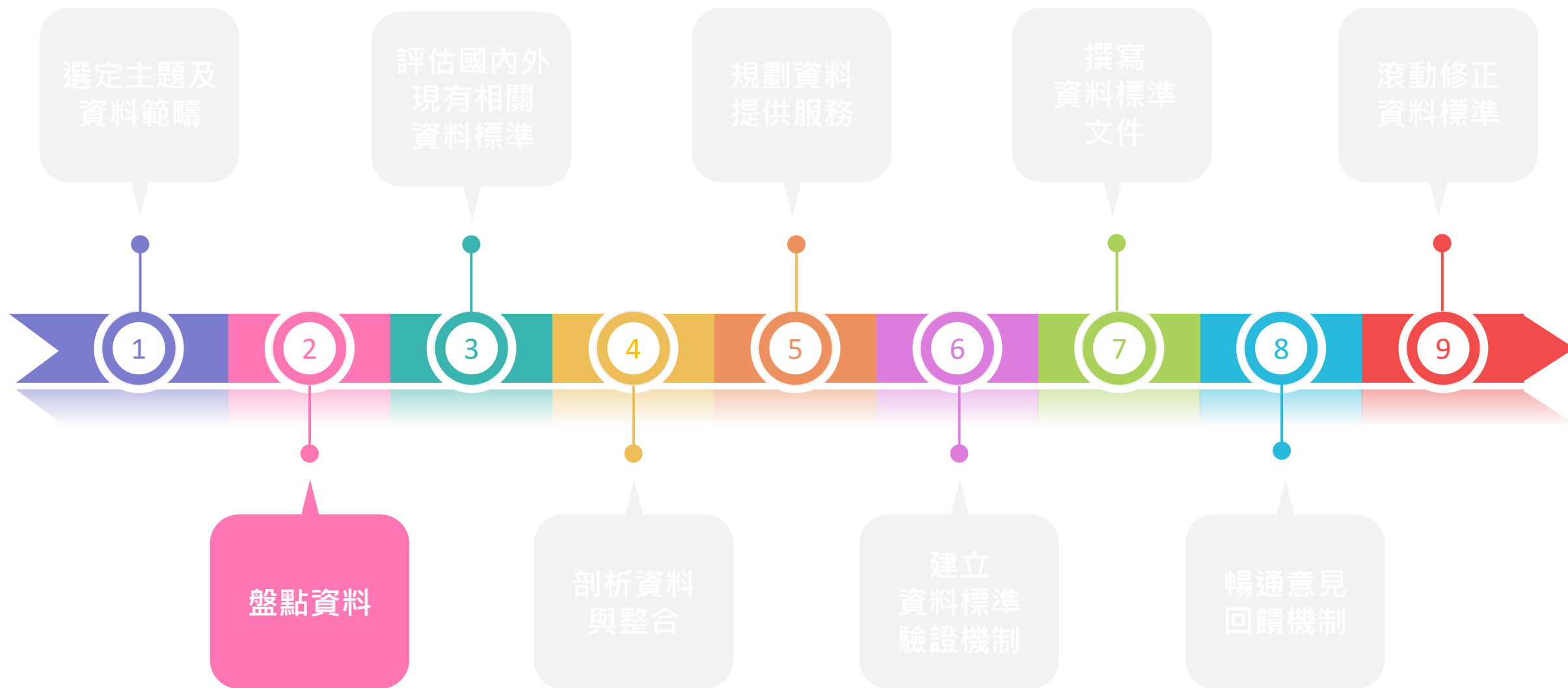
資料欄位固定，穩定性高。

1982年至今，資料欄位包含「領先指標」、「同時指標」及「落後指標」...等等構成項目

欄位穩定，資料品質也會較好

機關訂定資料標準方法流程(步驟2)

詳參數位發展部「領域資料標準訂定流程參考指引」



盤點資料：鎖定預計的資料集進行分析

前往 schema.gov.tw 查看是否有適合的領域/主題/類別

- 若有，可進入查看是否有類似的可以使用
- 若無，可預備規劃自己的新領域放入新主題

政府資料標準平臺
schema.gov.tw

網站導覽 | 問題回報 | English

首頁 全部領域 資料標準列表 政府資料代碼專區 API服務 關於平臺

全部領域
下載全部資料標準。 [XML](#) [JSON](#)。本網站使用之正規表示式以 PHP 版本為主。

<p>共通性</p> <p>近30日點閱數：669 累計點閱數：14,774</p> <p>XML JSON</p>	<p>戶政</p> <p>近30日點閱數：118 累計點閱數：3,588</p> <p>XML JSON</p>	<p>財稅</p> <p>近30日點閱數：148 累計點閱數：3,530</p> <p>XML JSON</p>
<p>商工</p> <p>近30日點閱數：397 累計點閱數：10,149</p> <p>XML JSON</p>	<p>公路監理</p> <p>近30日點閱數：33 累計點閱數：1,240</p> <p>XML JSON</p>	<p>勞保</p> <p>近30日點閱數：305 累計點閱數：3,750</p> <p>XML JSON</p>
<p>健保</p> <p>近30日點閱數：93 累計點閱數：1,728</p> <p>XML JSON</p>	<p>長照</p> <p>近30日點閱數：107 累計點閱數：1,767</p> <p>XML JSON</p>	<p>金融</p> <p>近30日點閱數：764 累計點閱數：4,217</p> <p>XML JSON</p>
<p>政府主計</p> <p>近30日點閱數：261 累計點閱數：645</p> <p>XML JSON</p>	<p>貨幣與金融</p> <p>近30日點閱數：1,127 累計點閱數：1,696</p> <p>XML JSON</p>	<p>新領域名稱</p>

盤點機關資料...確認沒有可放入的領域
就可以用機關業務性質或資料性質命名新領域

盤點資料：盤點原始資料文件

點開資料集「景氣指標及燈號」看一下資料內容，詳細包含「景氣指標與燈號」、「領先指標構成項目」、「同時指標構成項目」、「落後指標構成項目」、「景氣對策信號構成項目」

構成項目	原始資料來源	季調資料來源
領先指標		
- 外銷訂單動向指數(以家數計)	經濟部	國發會
- 實質貨幣總計數 M1B ¹	中央銀行	同左
- 股價指數	台灣證券交易所	不季節調整
- 工業及服務業受僱員工淨進入率 ²	主計總處	國發會
- 建築物開工樓地板面積 (住宅、商辦、工業倉儲)	內政部	國發會
- 實質半導體設備進口值 ³	財政部	採3個月移動平均值
- 製造業營業氣候測驗點	台灣經濟研究院	不季節調整
同時指標		
- 工業生產指數	經濟部	同左
- 電力(企業)總用電量	台灣電力公司	國發會
- 製造業銷售量指數	經濟部	國發會
- 批發、零售及餐飲業營業額	經濟部	同左
- 非農業部門就業人數	主計總處	國發會
- 實質海關出口值 ³	財政部	同左
- 實質機械及電機設備進口值 ⁴	財政部	國發會
落後指標		
- 失業率 ⁵	主計總處	同左
- 製造業單位產出勞動成本指數	主計總處	國發會
- 金融業隔夜拆款利率	中央銀行	不季節調整
- 全體金融機構放款與投資	中央銀行	不季節調整
- 製造業存貨價值	經濟部	不季節調整

「景氣指標及燈號」，看起來原始資料來源與各機關有關好像需要邀請各業務資料之相關者討論資料標準

Date	領先指標綜合指數	領先指標不含趨勢指數	同時指標綜合指數	同時指標不含趨勢指數	落後指標綜合指數	落後指標不含趨勢指數	景氣對策信號綜合分數	景氣對策信號
201901	108.89	100.03	108.2	99.4	111.85	102.75	20	黃藍
201902	109.56	100.48	107.89	98.94	111.54	102.29	17	黃藍
201903	110.26	100.95	107.83	98.72	111.03	101.65	20	黃藍
201904	110.82	101.3	108	98.72	110.49	101	21	黃藍
201905	111.22	101.5	108.37	98.9	110.07	100.45	18	黃藍
201906	111.53	101.61	108.89	99.21	109.74	99.98	21	黃藍
201907	111.81	101.7	109.43	99.54	109.51	99.61	22	黃藍
201908	112.12	101.81	109.89	99.79	109.32	99.28	20	黃藍
201909	112.48	101.97	110.25	99.95	109.15	98.96	19	黃藍
201910	112.85	102.14	110.55	100.07	108.94	98.96	18	黃藍

- 看一下有沒有共通性欄位(可前往schema.gov.tw上查詢)
 - Date** 不可對應到「日期」YYYYMMDD 為String用以記錄年份月報
- 看一下有沒有自由欄位(沒有任何限制的字串欄位)
 - 景氣對策信號**非自由欄位 為有限選擇:綠燈、紅燈、藍燈、黃紅燈、黃藍燈
- 看一下有沒有數字統計欄位(純數字)
 - 綜合指數**是數字統計欄位
 - 趨勢指數**是數字統計欄位
 - 綜合分數**是數字統計欄位

了解一下「景氣指標及燈號」資料注入的方式

Date	領先指標綜合指數	領先指標不含趨勢指數	同時指標綜合指數	同時指標不含趨勢指數	落後指標綜合指數	落後指標不含趨勢指數	景氣對策信號綜合分數	景氣對策信號
201901	108.89	100.03	108.2	99.4	111.85	102.75	20	黃藍
201902	109.56	100.48	107.89	98.94	111.54	102.29	17	黃藍
201903	110.26	100.95	107.83	98.72	111.03	101.65	20	黃藍
201904	110.82	101.3	108	98.72	110.49	101	21	黃藍
201905	111.22	101.5	108.37	98.9	110.07	100.45	18	黃藍
201906	111.53	101.61	108.89	99.21	109.74	99.98	21	黃藍
201907	111.81	101.7	109.43	99.54	109.51	99.61	22	黃藍
201908	112.12	101.81	109.89	99.79	109.32	99.28	20	黃藍
201909	112.48	101.97	110.25	99.95	109.15	98.96	19	黃藍
201910	112.85	102.14	110.55	100.07	108.94			黃藍

1. 哪些欄位是系統自動產生
 - ① **Date** 是由系統自動產生嗎
2. 哪些欄位在填寫過程中有選單或規則(不會出現額外雜訊)
 - ① **景氣對策信號** 是由系統選單選擇嗎
3. 哪些欄位資料型別是字串string
 - ① 填入是空白欄位，由人員自由填打嗎 (此範例無，另舉例:姓名)
 - ② 填入是英數編碼，有填打格式規則嗎 (此範例無，另舉例:身分證號)
4. 哪些欄位資料型別是數值number
 - ① **景氣對策信號綜合分數** 有規定只能填入整數嗎
 - ② **綜合指數、趨勢指數** 有規定必填成浮點數嗎

盤點資料：盤點資料欄位範圍

確認一下「景氣指標及燈號」資料可能範圍及格式、相關資料

The screenshot shows the Microsoft Excel ribbon with the 'Data' tab selected. The 'Sort & Filter' group is highlighted with a red box, containing icons for 'Sort' (升序, 降序, 降序, 升序), 'Filter' (篩選), 'Clear' (清除), 'Refresh All' (重新套用), and 'Advanced Filter' (進階...). The 'Data Analysis' group is also highlighted with a red box, containing icons for 'Data Analysis Tools' (資料剖析), 'Quick Fill' (快速填入), 'Remove Duplicates' (移除重複項), 'Data Validation' (資料驗證), 'Merge Cells' (合併彙算), 'PivotTable' (關聯), 'Manage Data Model' (管理資料模型), 'Modeling' (模擬分析), 'Forecast' (預測), 'Group' (組成群組), 'Ungroup' (取消群組), and 'Subtotal' (小計).

打開excel，運用工作表之[資料]功能頁籤中「排序與篩選」和「資料剖析」工具進行盤點

Date	領先指標綜合指數	領先指標不含趨勢指數	同時指標綜合指數	同時指標不含趨勢指數	落後指標綜合指數	落後指標不含趨勢指數	景氣對策信號綜合分數	景氣對策信號
198401	22.04	103.98	22.34	105.41	21.06	99.33	42	紅
198402	22.36	104.68	22.6	105.8	21.51	100.71	39	紅
198403	22.64	105.2	22.78	105.85	21.91	101.77	37	黃紅
198404	22.83	105.26	22.87	105.48	22.22	102.45	34	黃紅
198405	22.87	104.68	22.92	104.91	22.48	102.92	37	黃紅
198406	22.77	103.51	22.96	104.36	22.76	103.46	30	綠
198407	22.61	102.04	22.97	103.69	23.06	104.1	30	綠
198408	22.42	100.51	22.97	102.98	23.38	104.82	27	綠
198409	22.28	99.17	22.96	102.2	23.73	105.65	25	綠
198410	22.21	98.2	22.92	101.35	24.08	106.48	28	綠

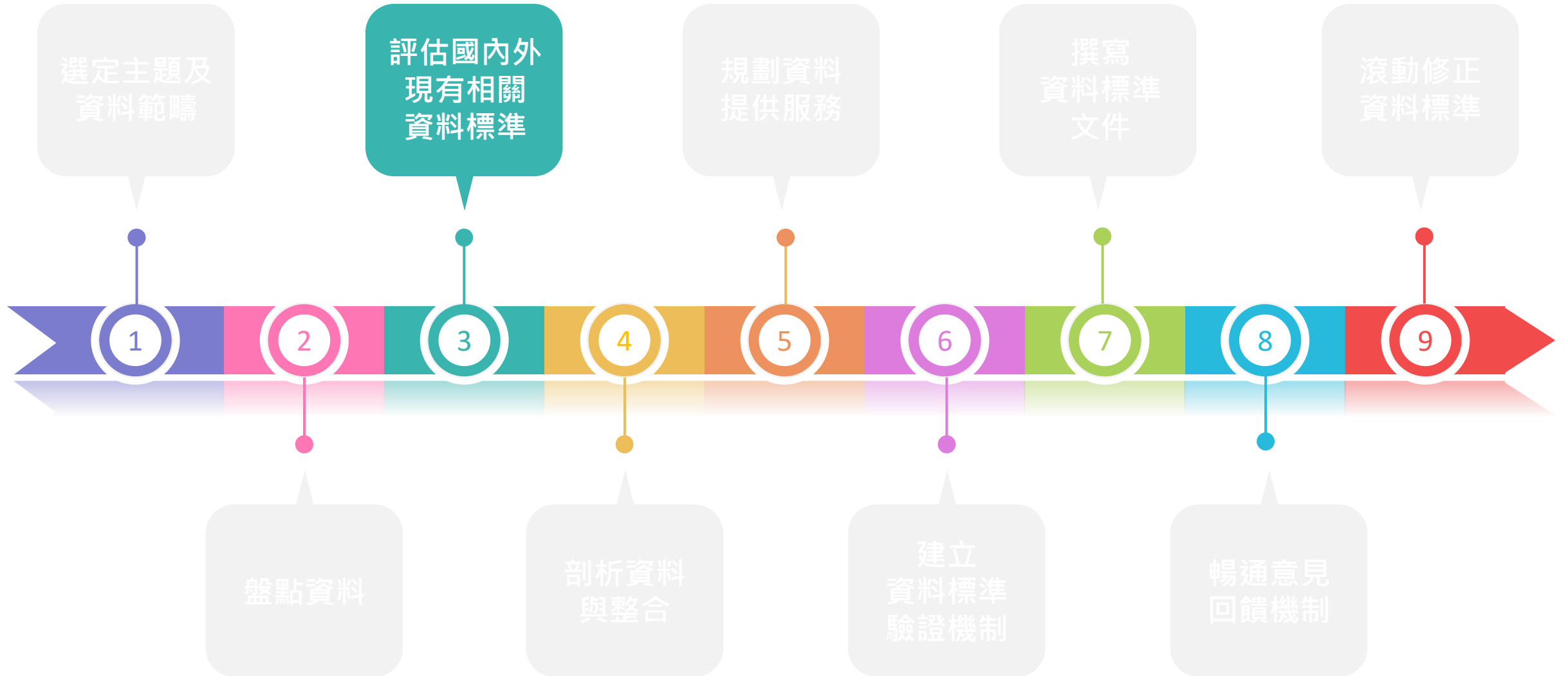
Date	領先指標綜合指數	領先指標不含趨勢指數	同時指標綜合指數	同時指標不含趨勢指數	落後指標綜合指數	落後指標不含趨勢指數	景氣對策信號綜合分數	景氣對策信號
201901	108.89	100.03	108.2	99.4	111.85	102.75	20	黃藍
201902	109.56	100.48	107.89	98.94	111.54	102.29	17	黃藍
201903	110.26	100.95	107.83	98.72	111.03	101.65	20	黃藍
201904	110.82	101.3	108	98.72	110.49	101	21	黃藍
201905	111.22	101.5	108.37	98.9	110.07	100.45	18	黃藍
201906	111.53	101.61	108.89	99.21	109.74	99.98	21	黃藍
201907	111.81	101.7	109.43	99.54	109.51	99.61	22	黃藍
201908	112.12	101.81	109.89	99.79	109.32	99.28	20	黃藍
201909	112.48	101.97	110.25	99.95	109.15	98.96	19	黃藍
201910	112.85	102.14	110.55	100.07	108.94	98.6	18	黃藍

1. Date：看起來都是西元年搭配月份之YYYYMM格式
2. 綜合指數：看起來都是浮點數(為2-3位正整數與0-2位小數)
3. 趨勢指數：看起來都是浮點數(為2-3位正整數與0-2位小數)
4. 景氣對策信號綜合分數：看起來都是正整數，範圍為[9-45]
5. 景氣對策信號：看起來都是5種燈號[綠燈,紅燈,藍燈,黃紅燈,黃藍燈]

※建議業務單位與資訊單位一起討論

機關訂定資料標準方法流程(步驟3)

詳參數位發展部「領域資料標準訂定流程參考指引」



評估國內外現有相關資料標準

優先查看 schema.gov.tw 是否已經有可用的資料標準
避免重複訂定



資料標準列表

領域-主題-類別
資料標準分類篩選

- 共通性 (1)
 - 共通性 (1)
- 戶政 (3)
 - 戶籍資料 (3)
- 公路監理 (2)
 - 監理資料 (2)
- 勞保 (3)
 - 勞保納保 (1)
 - 勞保給付 (1)
 - 國保業務 (1)
- 財稅 (1)
 - 房屋稅 (1)
- 金融 (14)
 - 證券期貨 (3)
 - 共通 (7)
 - 銀行 (2)
 - 保險 (2)
- 商工 (33)
 - 公司 (10)
 - 商業 (5)
 - 工廠 (6)
 - 有限合夥 (7)
 - 共通 (5)
- 健保 (1)
 - 納保 (1)
- 長照 (1)
 - 長照 (1)
- 政府主計 (2)
 - 歲計會計 (2)
- 貨幣與金融 (3)
 - 經濟研究與統計 (1)
 - 共通性資料 (2)

1

查看Schema.gov是否有類似標準

若有可直接參照引用，可減少分歧

2

參考國外是否有相關資料標準

旁徵博引，更了解訂定的標準是否有疏漏

3

查看目前法規有沒有特定規範

數位發展部「領域資料標準訂定參考指引」
數位發展部「共通性資料標準」
其他特定領域的法規規範

評估國內外現有相關資料標準

■ 共通標準

所有領域適用之標準，例如：ISO標準、國民身分證統一編號、郵遞區號，參照引用共通領域即可。

■ 主要標準

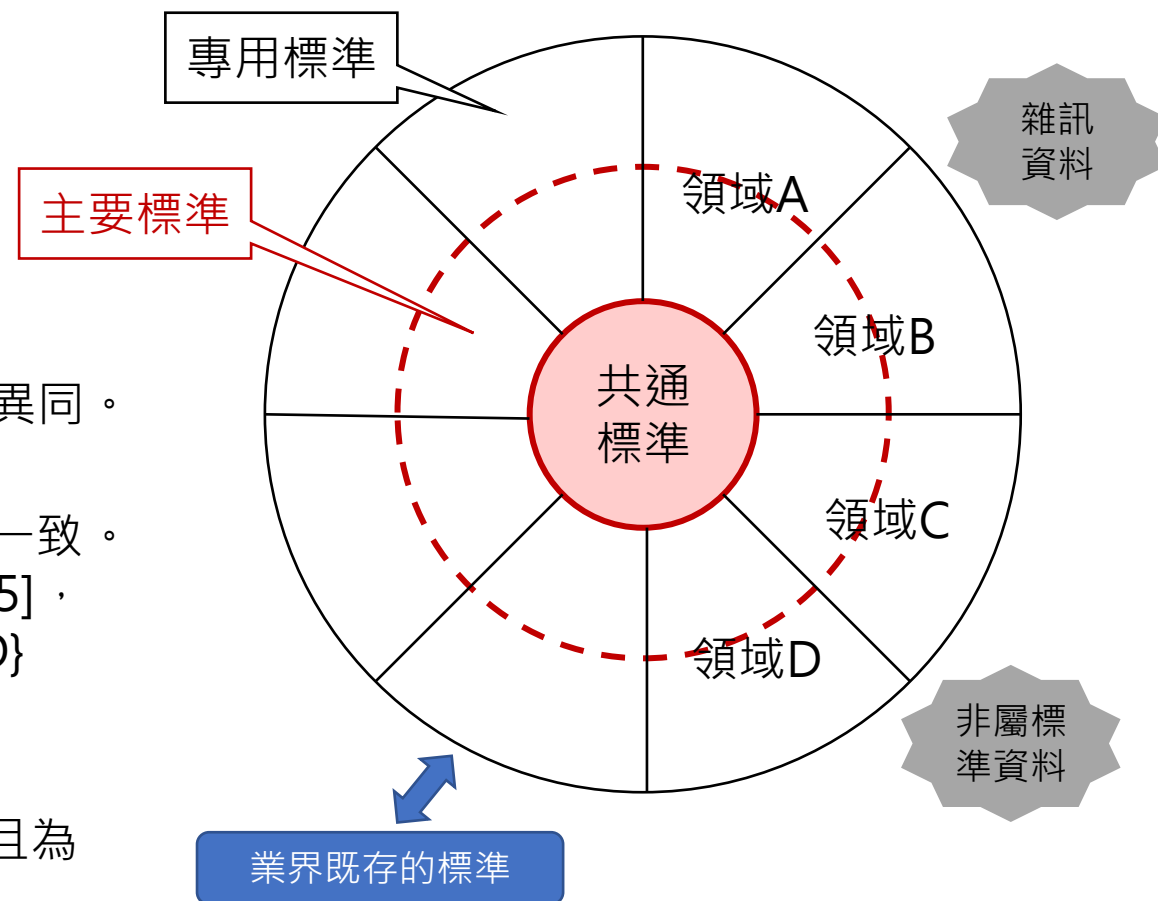
特定領域之主要標準，例如：氣象領域-溫度(意指氣象溫度)，公衛領域-溫度(意指人類體溫)，動保領域-溫度(意指生物體溫)

- ① 跨領域的勾稽：標準意涵可互通，但值域範圍可能有所異同。例如人體恆溫[36.1-37.5]，恆溫動物[35.5-43.2] °C
- ② 同領域跨機關的勾稽：若涉及跨機關標準不同則須協調一致。例如自行車站點代號新北市[1001-1731]、臺南市[2-85]，交通部協調編列一致之UID {業管機關代碼} + {StationID}

■ 專用標準

該領域專用標準，例如：氣象領域之颱風命名與編號，且為該機關完全主責之標準。

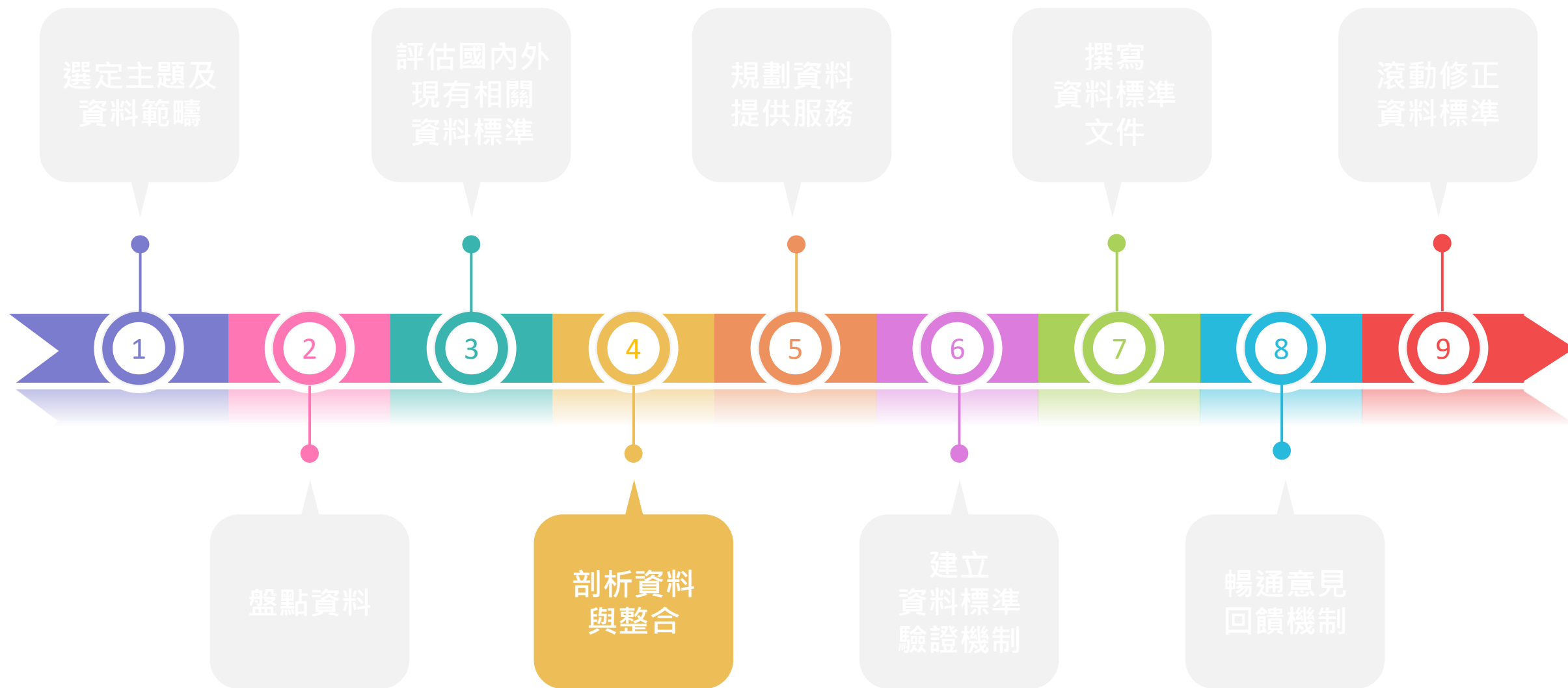
建議參考共通語彙基盤概念
將盤點後的資料與國內外領域資料標準進行初步比對、分類



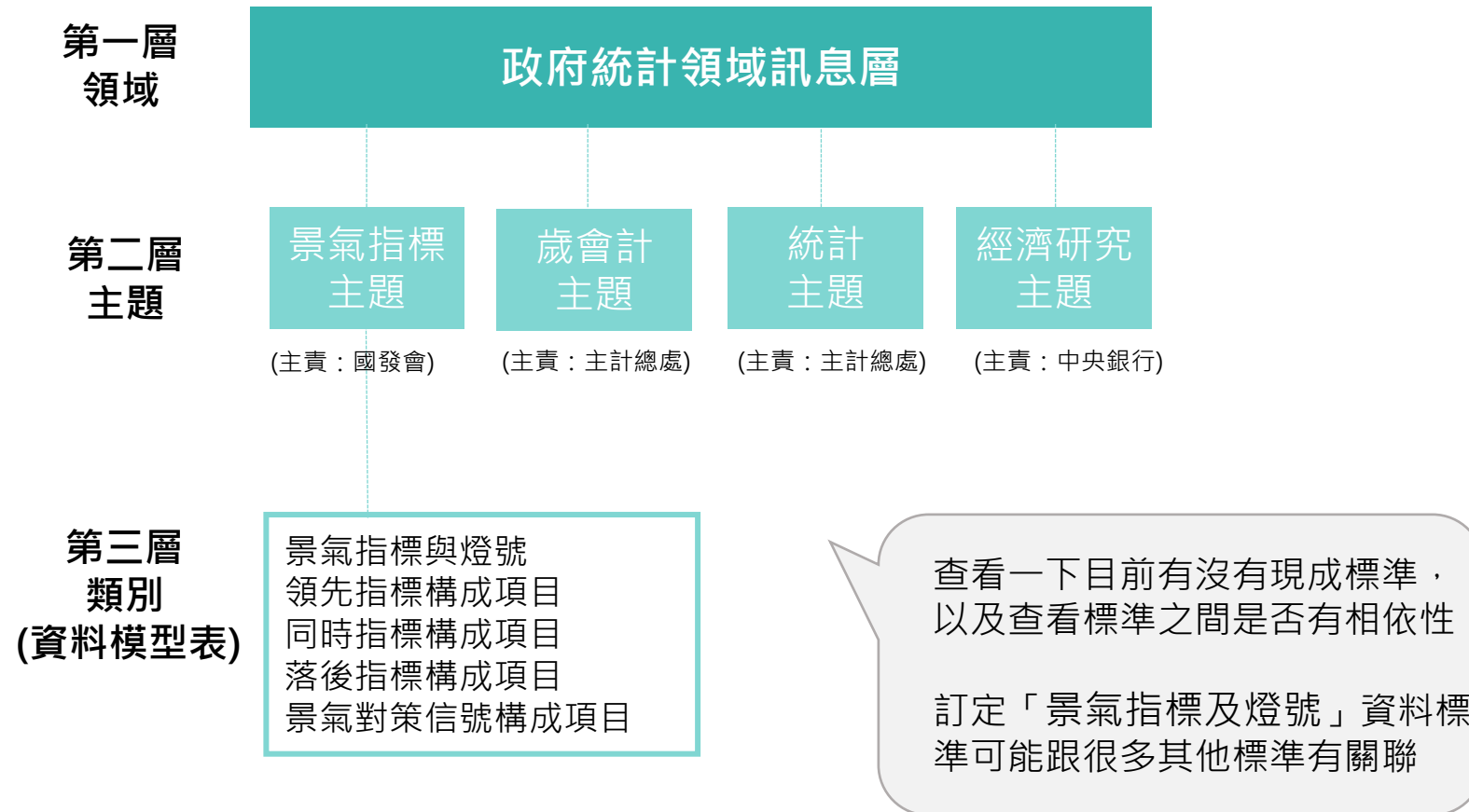
基本上可透過對應表來互轉資料

機關訂定資料標準方法流程(步驟4)

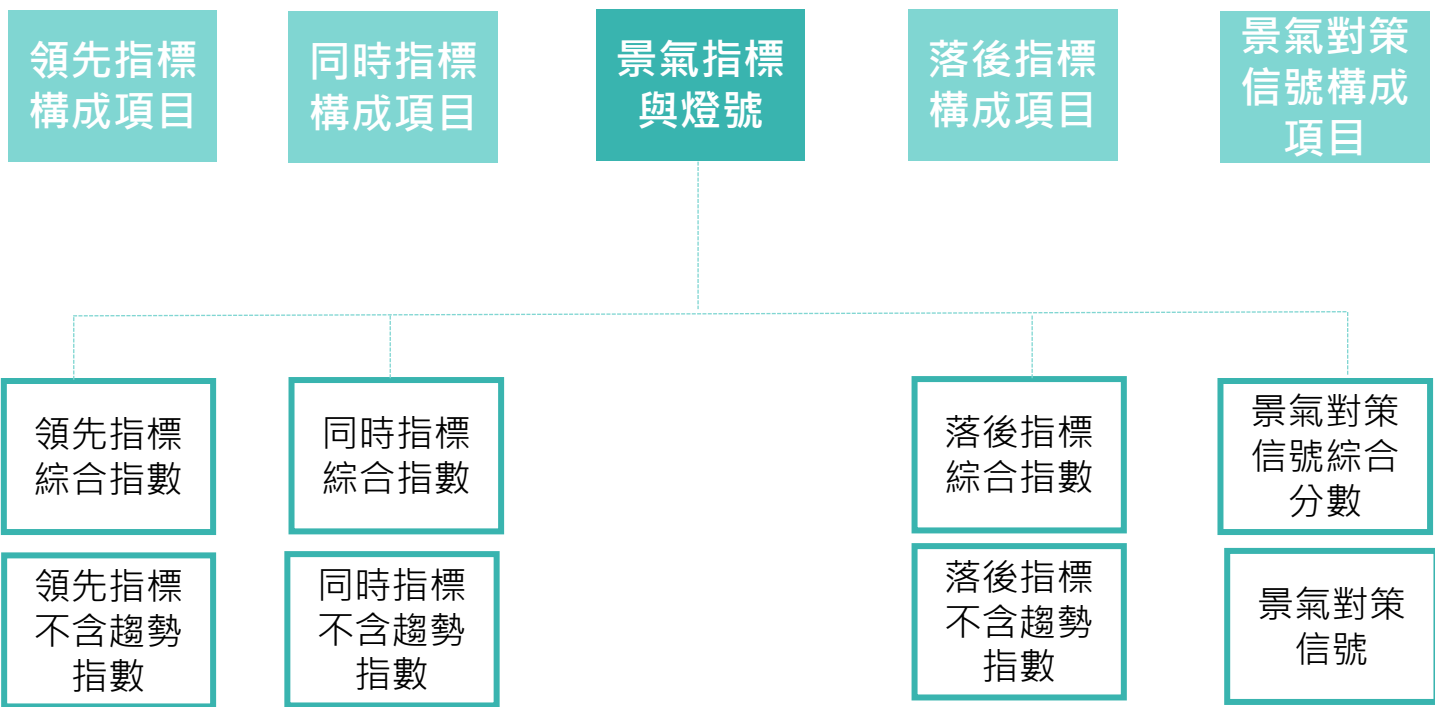
詳參數位發展部「領域資料標準訂定流程參考指引」



資料盤點結果所設定之各資料種類未必來自同一資料庫，各資料擁有者、主責單位亦不同，因此須藉由特性分析確認相互之關係，意即確認資料階層與架構



釐清資料特性分析，透過資料模型建置，將各種基礎資料(業務型資料)、應用資料(分析型資料)結合以呈現完整領域資料標準

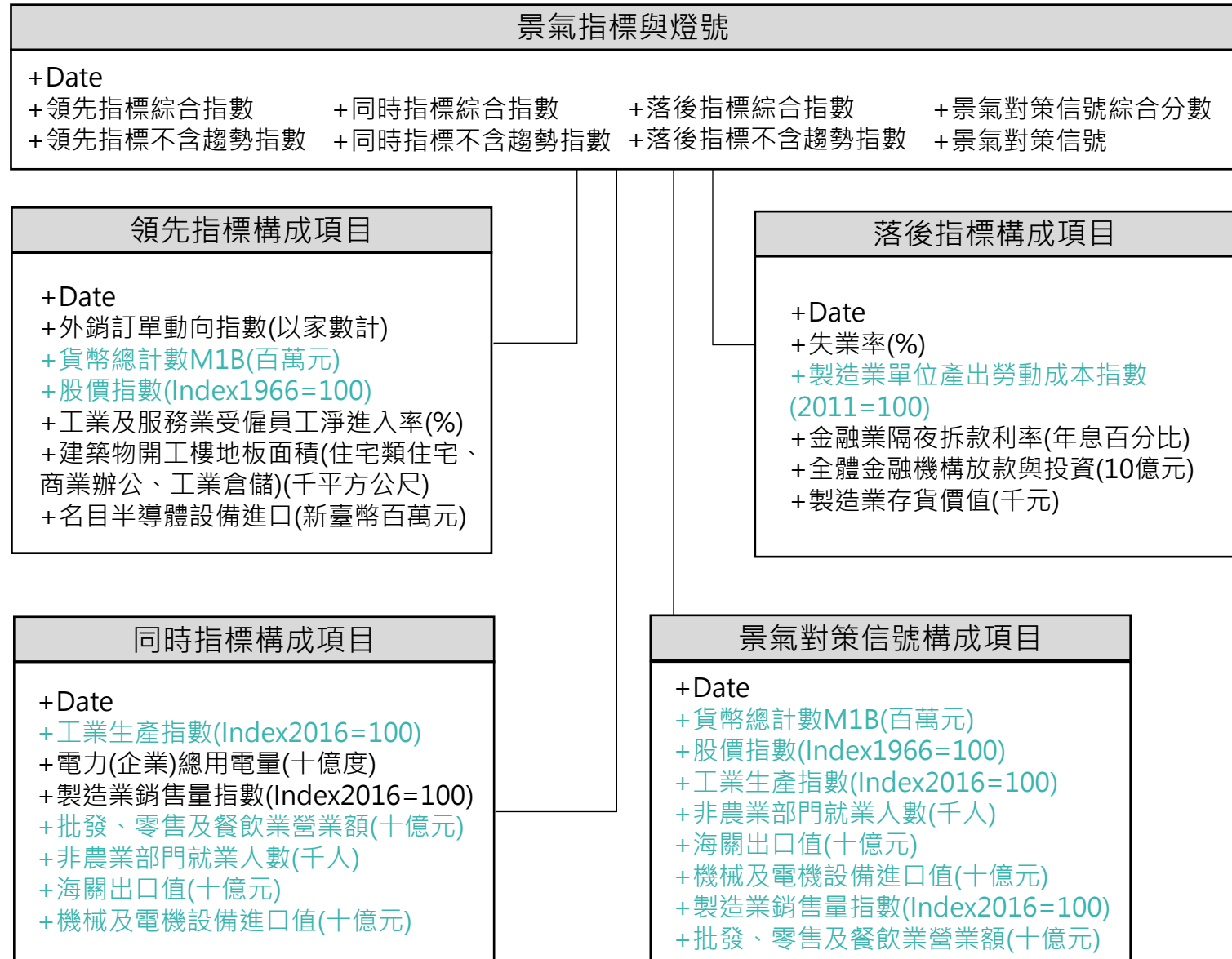


建議參考現行資料內容與實際資料供需狀況，進行必要性屬性分析後，納入應用綱要設計。

※實際依各領域專業進行描述資料的類別集合、資料屬性之間的關係

了解「景氣指標及燈號」，可跟其他資料表**應用的方式和範圍**看起來...領先指標構成項目、同時指標構成項目、落後指標構成項目、景氣對策信號構成項目，關聯「景氣指標及燈號」之指標指數與信號分數

將特定領域之資料內容文件化，以達成一般且可正確解讀資料之目的，建議將特性分析及類別設計成果以正式的模組化語言(例如UML結構類別圖)來描述應用綱要。



可以將系統內的資料表進行串聯

※建議業務單位與資訊單位一起討論

撰寫資料標準時須以資料典形式提供各欄位之詳細定義及必要說明，以利使用者解讀所取得之資料內容，資料典定義說明如下表所示

名稱	定義
項次	資料典之序號。
中英文名稱	資料屬性列舉於資料典中。
說明	提供屬性之意義，以淺顯易懂的文字介紹該項資料，協助使用者了解屬性之內容，避免誤解使用。
型別	說明屬性之資料型態填列String、Number、Boolean、Date、Time、DateTime。
值域	說明該欄位之資料型別值域，為資料內容是否正確的判斷方式。
必要性	說明該屬性描述之必要性，區分為「必要屬性」(Mandatory)、及「選擇屬性」(Optional)，本文說明方式將以簡寫之(M)與(O)表示之。
最多發生次數	依實際需求，至多出現一次(1)、只出現特定次數(特定次數)或出現多次(N)。
附註	針對屬性或關係提供前述項目無法提供之額外說明，例如代碼表、編碼規則。

- 根據不同項目制定資料欄位內容，為了讓一般使用者能夠更了解資料特性，資料說明以越詳細越好。
- 若有代碼，則額外以代碼表方式說明

項次	類別	中文屬性名稱	英文屬性名稱	資料說明	必要性	發生次數	資料型別	值域	附註
1	景氣指標與燈號	年月份	Monthly report	描述台灣景氣指標年月份	M	1	String	[0000-9999][01-12]	
2		領先指標綜合指數	Composite Leading Index	將領先指標不含趨勢指數進行趨勢還原後所得出	M	1	Number	浮點數	
3		領先指標不含趨勢指數	Trend adjusted Leading Index	合成景氣領先指標後進行振幅調整所得出	M	1	Number	浮點數	
4		同時指標綜合指數	Composite Coincident Index	將同時指標不含趨勢指數進行趨勢還原後所得出	M	1	Number	浮點數	
5		同時指標不含趨勢指數	Trend adjusted Coincident Index	合成景氣同時指標後進行振幅調整所得出	M	1	Number	浮點數	
6		落後指標綜合指數	Composite Lagging Index	將落後指標不含趨勢指數進行趨勢還原後所得出	M	1	Number	浮點數	
7		落後指標不含趨勢指數	Trend adjusted Lagging Index	合成景氣落後指標後進行振幅調整所得出	M	1	Number	浮點數	
8		景氣對策信號綜合分數	Scores of Monitoring Indicators	由貨幣總計數M1B變動率等9項指標構成分數總和	O	1	Number	[9-45]	45-38分為熱絡 37-32分為轉向 31-23分為穩定 22-17分為轉向 16-9分為低迷
9		景氣對策信號	Monitoring Indicators	以景氣燈號標誌景氣波動的指標	O	1	String	字串	詳閱「景氣對策信號」代碼表

名稱(中文)	代碼
景氣穩定	綠燈
景氣熱絡	紅燈
景氣低迷	藍燈
景氣由穩定轉向熱絡	黃紅燈
景氣由穩定轉向低迷	黃藍燈

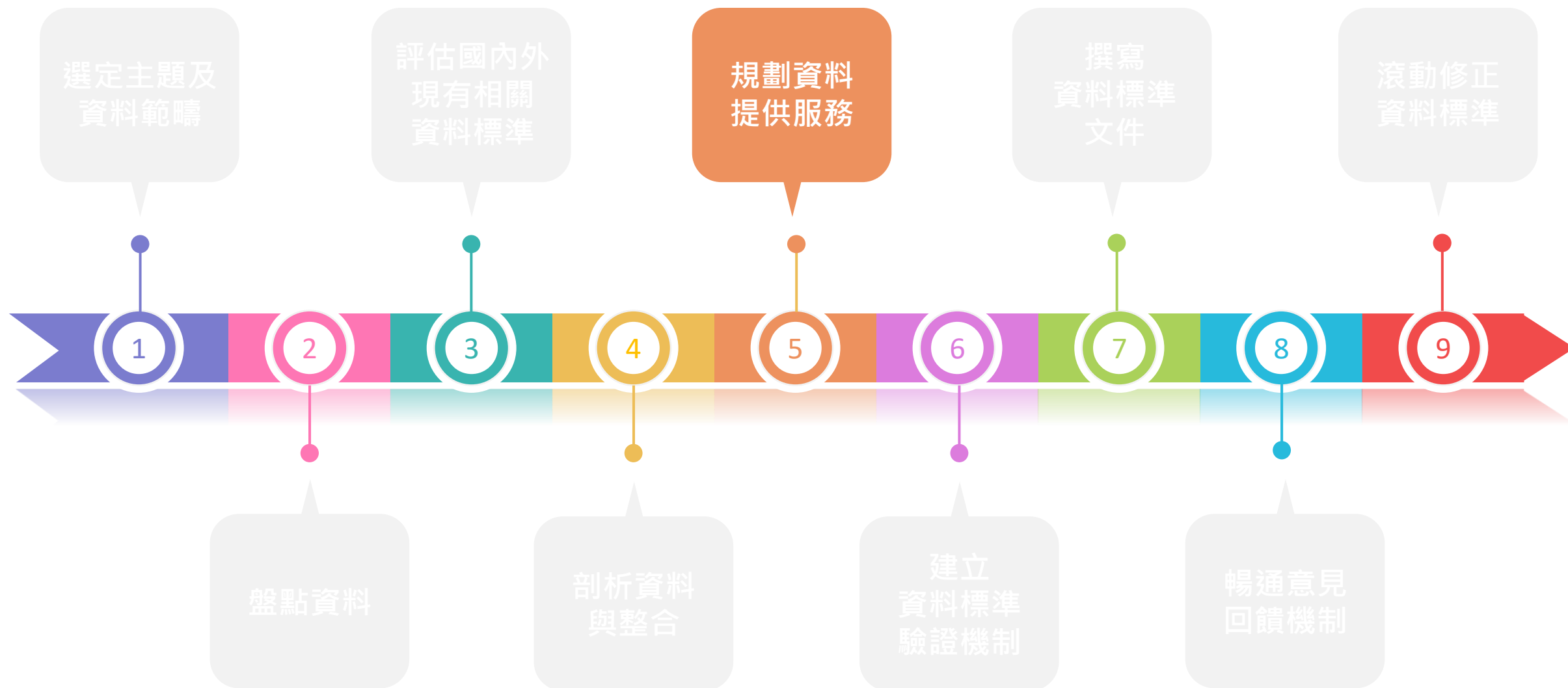
為達資料實質流通交換，須將應用綱要轉換為合適的資料編碼架構，例如XML、JSON、KML、GML等開放格式

房屋稅籍中文主檔Schema

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <!-- 中文主檔: HOUT100 -->
  <xs:element name="ROWSET">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="ROW">
          <xs:complexType>
            <xs:sequence>
              <xs:element type="xs:int" name="HSN_CD"/>
              <xs:element type="xs:int" name="HOU_LOSN"/>
              <xs:element type="xs:string" name="OWNER_TP"/>
              <xs:element type="xs:string" name="HOU_BL_TP"/>
              <xs:element type="xs:string" name="USE_LIC_NO"/>
              <xs:element type="xs:string" name="HOU_YN_SPRMAG"/>
              <xs:element type="xs:string" name="LOCATION"/>
              <xs:element type="xs:integer" name="LOCATION_CD"/>
              <xs:element type="xs:string" name="SPC_ADDR_MK"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

機關訂定資料標準方法流程(步驟5)

詳參數位發展部「領域資料標準訂定流程參考指引」



規劃資料提供服務

規劃資料提供欄位、資料服務提供模式

以便利民眾及機器於未取得原始程式碼、文件等情況下，即可探索及瞭解資料內容

項次	類別	欄位名稱	英文欄位名稱	內容格式	選填條件
1	目錄 資訊 Catalog	分類編號	categoryCode	文字	必須填寫
		資料集編號	identifier	文字	必須填寫
3	資料集 資訊 Dataset	資料集名稱	title	文字	必須填寫
4		資料集描述	description	文字	必須填寫
5		主要欄位說明	fieldDescription	文字	必須填寫
6		資料集類型	type	文字	必須填寫
7		授權方式	license	文字	必須填寫
8		授權說明網址	licenseURL	網址	必須填寫
9		計費方式	cost	文字	必須填寫
10		計費說明網址	costURL	網址	選擇性填寫
11		計費法令依據	costLaw	文字或 網址	選擇性填寫
12		資料集提供機關之上級機關名稱	organization	文字	必須填寫
13	資料集提供機關之上級機關聯絡人姓名	organizationContactName	文字	必須填寫	

項次	類別	欄位名稱	英文欄位名稱	內容格式	選填條件
31	資源供應 資訊	資料資源編號	resourceID	文字	必須填寫
32		資料資源描述	resourceDescription	文字	必須填寫
33	Distribution	檔案格式	format	文字	必須填寫
34		編碼格式	characterSetCode	文字	必須填寫
35		資料資源最後更新時間	resourceModified	日期或 時間	建議填寫
36		資料存取網址	accessURL	網址	建議填寫
37		資料下載網址	downloadURL	網址	建議填寫
38		領域別詮釋資料參考網址	metadataSourceOfData	網址	建議填寫

配合政府資料開放平臺之推動，建議採用數位發展部頒布之「資料集詮釋資料標準規範」，資料供應單位落實詮釋資料內容之建置。

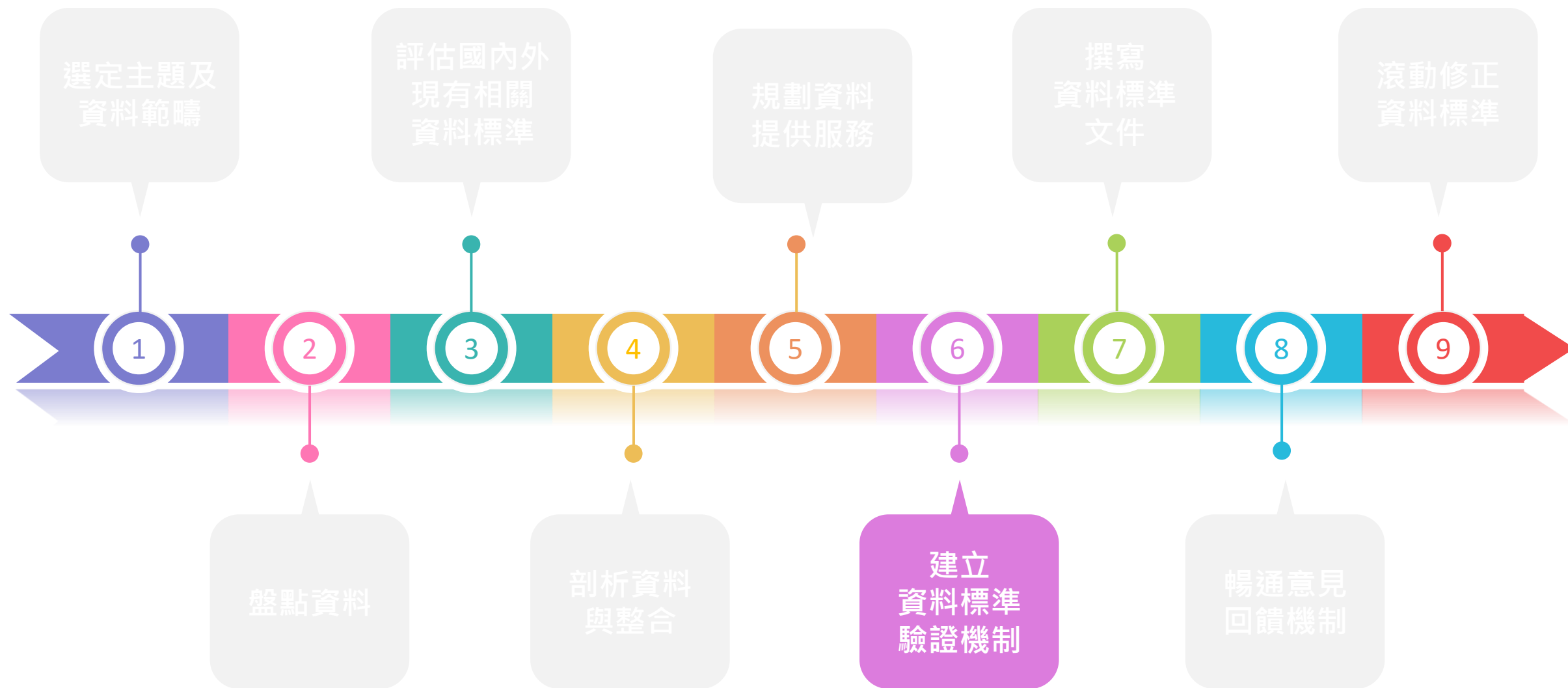
視實際需要，擴充項目之格式與型別，並須遵循該標準中之相關規定，且應伴隨詮釋資料提供資料取得者參考。

※建議業務單位與資訊單位一起討論

範例：公共運輸旅運資料標準-通用性資料集標準框架表(擷取部分示意)

機關訂定資料標準方法流程(步驟6)

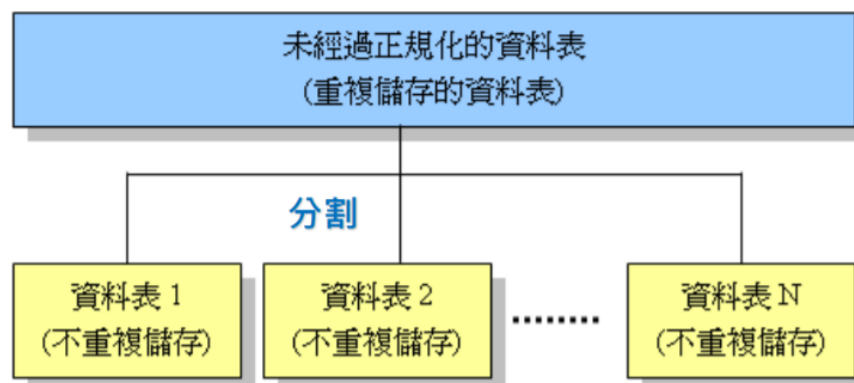
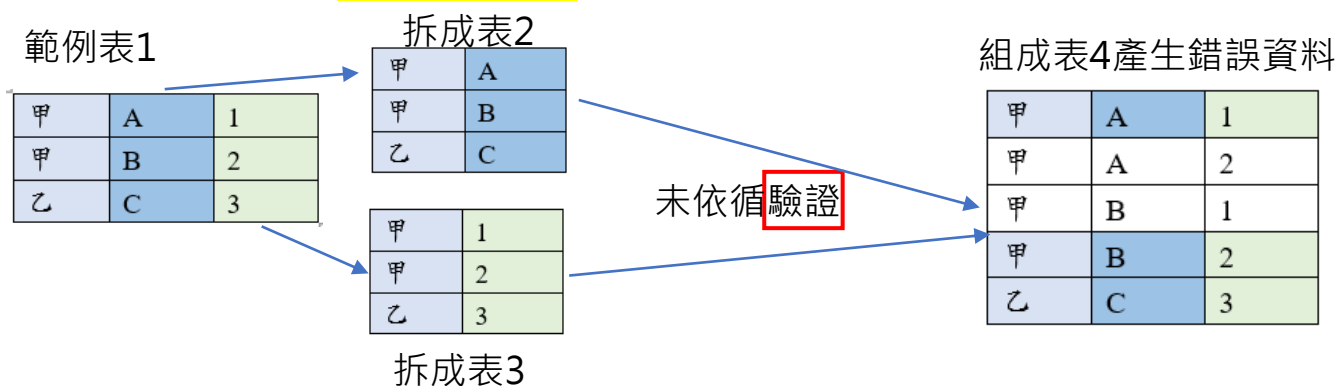
詳參數位發展部「領域資料標準訂定流程參考指引」



建立資料標準驗證機制(1/3)

設計資料標準及資料服務之驗證機制，以驗證所盤點之資料可通過此標準，確保對外提供資料服務之正確性、易用性、即時性及採適當格式開放。

方法1：產出資料前，將資料庫正規化以正確儲存資料結構



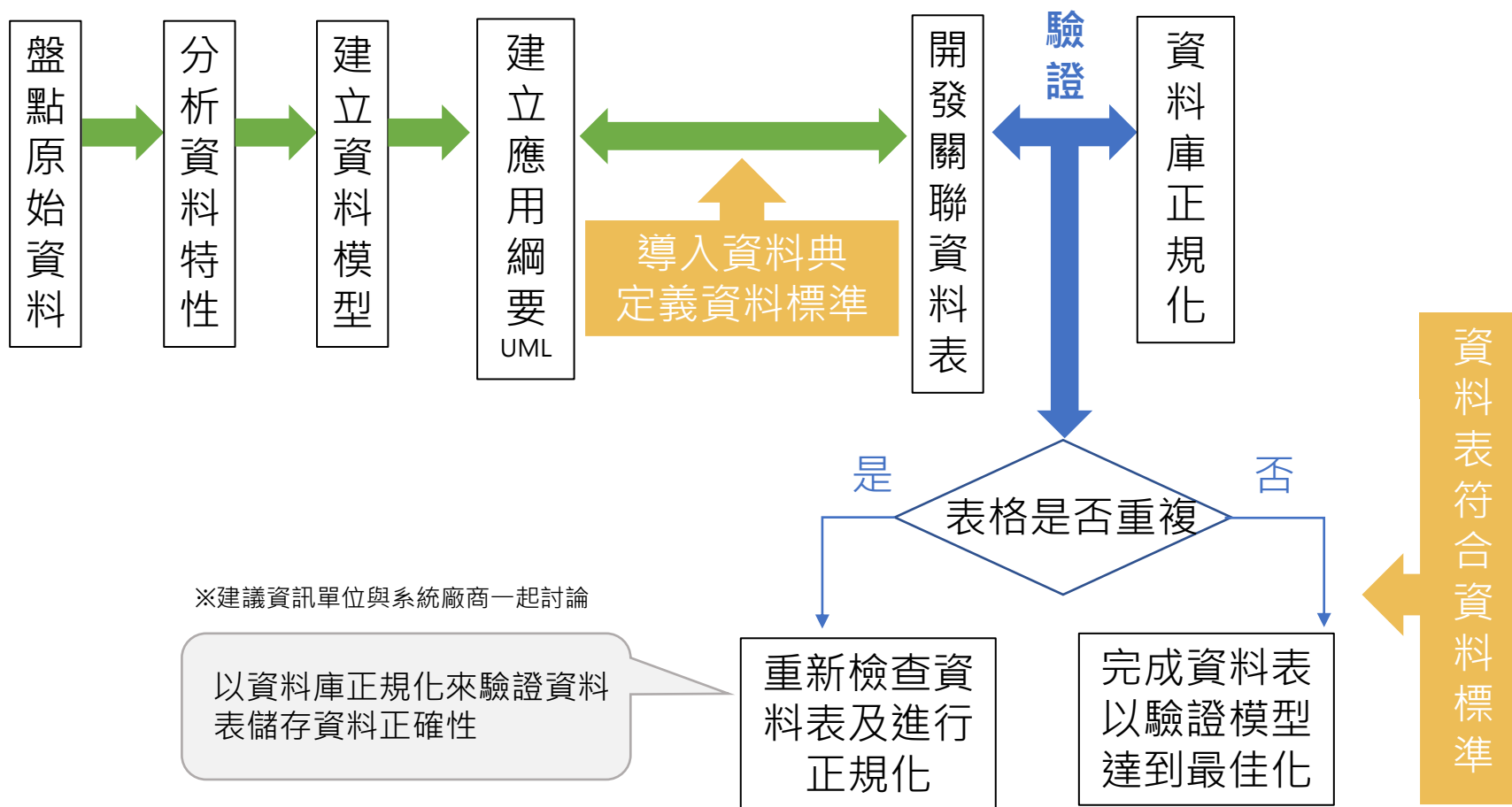
資料應用綱要的設計必須配合實務上的需要，因此，使用資料庫正規化(Normalization)，藉此檢視設計是否良好

以欄位的「相依性」作為分割資料表的依據，將整個資料表中重複的資料剔除，避免在資料關聯表產生新增異常、刪除異常、修改異常的狀況發生。

建立資料標準驗證機制(2/3)

設計資料標準及資料服務之驗證機制，以驗證所盤點之資料可通過此標準，確保對外提供資料服務之正確性、易用性、即時性及採適當格式開放。

方法1：**產出資料前**，將資料庫正規化以正確儲存資料結構



建立資料標準驗證機制(3/3)

設計資料標準及資料服務之驗證機制，以驗證所盤點之資料可通過此標準，確保對外提供資料服務之正確性、易用性、即時性及採適當格式開放。

方法2：產出資料後，資料開放前，使用資料標準平臺驗證功能

資料標準

共通性 > 共通性 > 共通資料

資料標準URI下載 [XML](#) [JSON](#)。本網站使用之正規表示式以 PHP 版本為主。

Properties from 共通資料

中文名稱	英文名稱	說明	型別	值域	附註	附件連結	驗證
姓名	Name	姓氏與名字	String				
國民身分證統一編號	Id	辨別個人身分之專屬代碼	String	[A-Z][1-2] [00000000-99999999]	第1碼編碼：縣市別 第2碼編碼：1為男性、2為女性		驗證
性別	Gender	依國民身分證統一編號第2碼編訂	String	[1-2]	1為男性 2為女性		驗證

驗證正規表示式

中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證失敗

驗證正規表示式

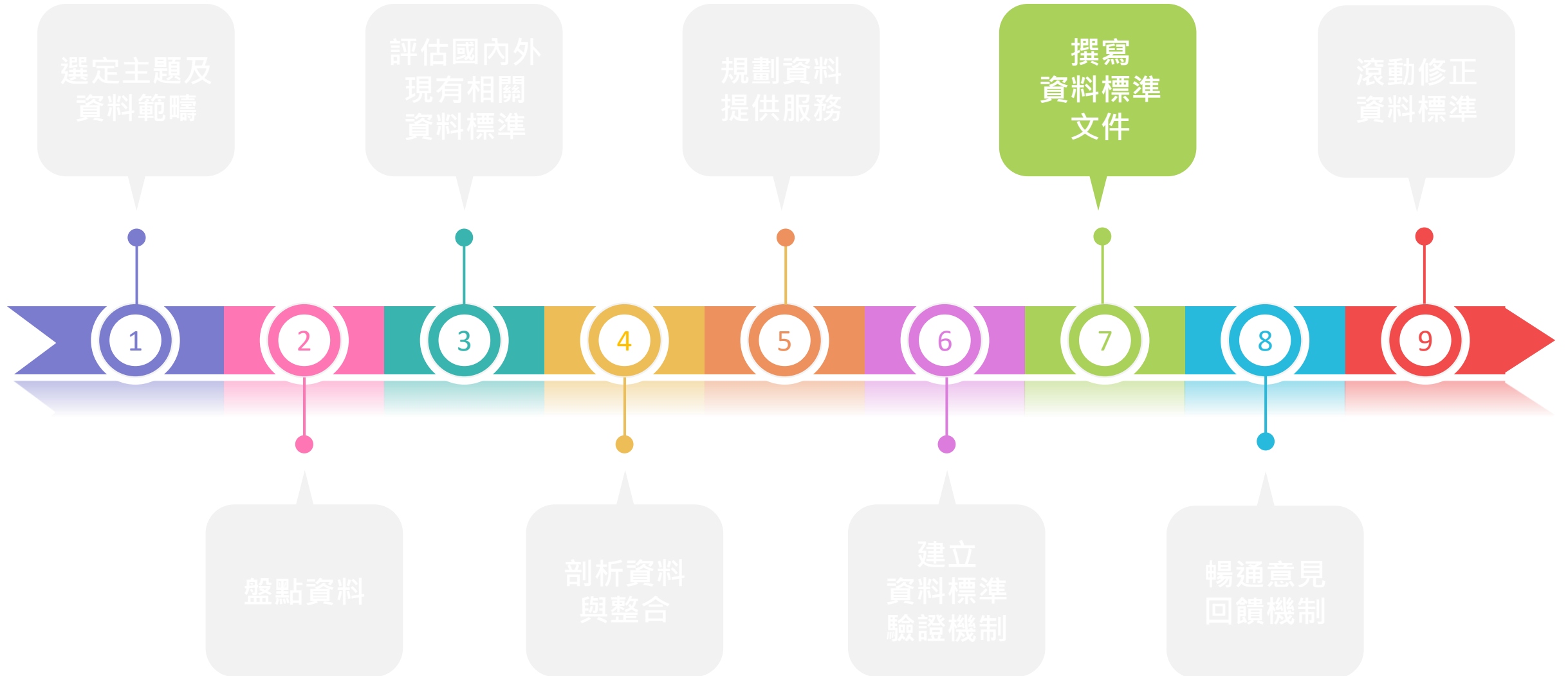
中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證成功

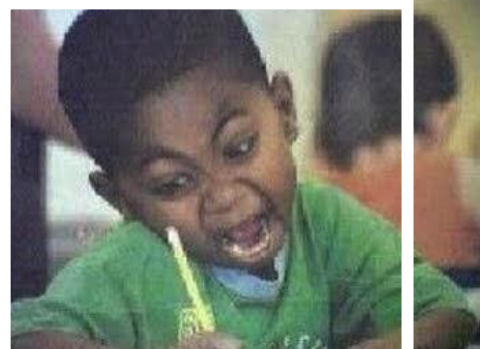
政府資料標準平臺提供將資料標準之值域正規化表示(Regular Expression)的驗證功能，方便機關運用以檢查開放或不開放之各類資料，是否符合資料標準格式，若未符合則需重新修改原始資料。

機關訂定資料標準方法流程(步驟7)

詳參數位發展部「領域資料標準訂定流程參考指引」



機關同仁須撰寫資料標準文件
設計供應機制，並提供資料使用者作為解讀資料之參考

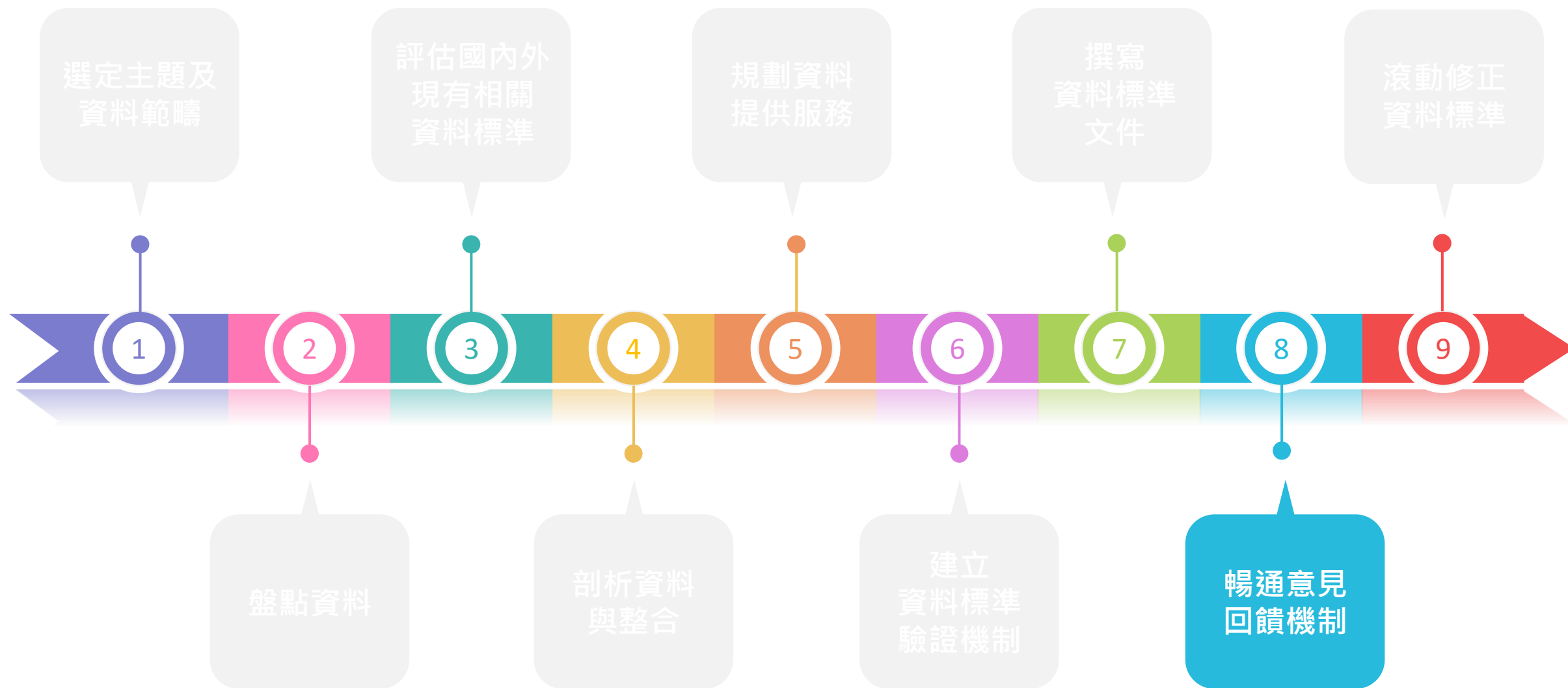


撰寫資料標準文件

應有章節內容	章節內容說明
目的	說明資料標準之設計目的、作業模式、運用流程架構等
範圍	說明資料標準之領域-主題-類別描述範疇、未來擴充內容
應用場合及使用限制	說明資料標準之應用場合與適用對象、資料生產單位，參考引用國內外相關標準或法規
特性分析	整體類別資料屬性結構層級分析
應用綱要	資料標準之資料屬性依據特性分析以結構類別圖說明應用
資料典	以表列說明資料標準之內容包含(項次、中文名稱、英文名稱、說明、型別、值域、必要性、最多發生次數、附註)
編碼規則	編碼轉換規定依循XML Schema或其他
標準訂定權責維護	說明相關聯絡資訊
附錄	資料相關說明、代碼表、系統資料檔屬性、初審意見等

機關訂定資料標準方法流程(步驟8)

詳參數位發展部「領域資料標準訂定流程參考指引」





暢通意見回饋機制

機關提出申請、學者專家給予意見
持續修正資料標準



撰寫資料標準文件

機關人員釐清資料標準後，進行資料標準訂定



提報數位發展部初審

數位發展部進行初審意見回覆，並於修改後進行複審



舉辦工作圈複審

因應機關需求或民間需求，調整資料標準規則

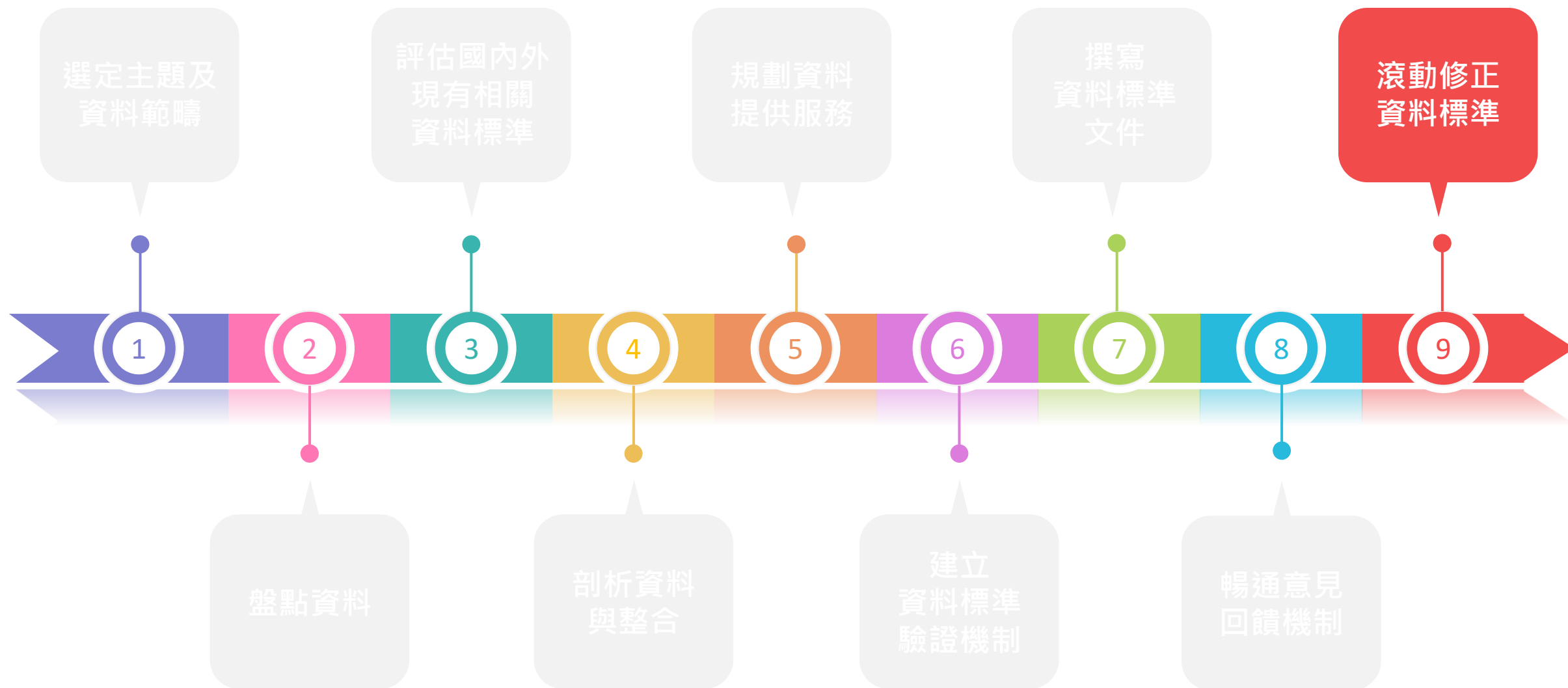


持續滾動式修正

達到更好的資料標準品質

機關訂定資料標準方法流程(步驟9)

詳參數位發展部「領域資料標準訂定流程參考指引」



滾動修正資料標準帶來開放資料的良性循環

回饋開放資料需求

民間得到回饋，願意提出更多資料需求，讓機關得以更了解社會大眾的渴望

更多服務到民間生活

民間團體有更多意願發展開放資料產品，促進更多食衣住行的發展



提升資料品質

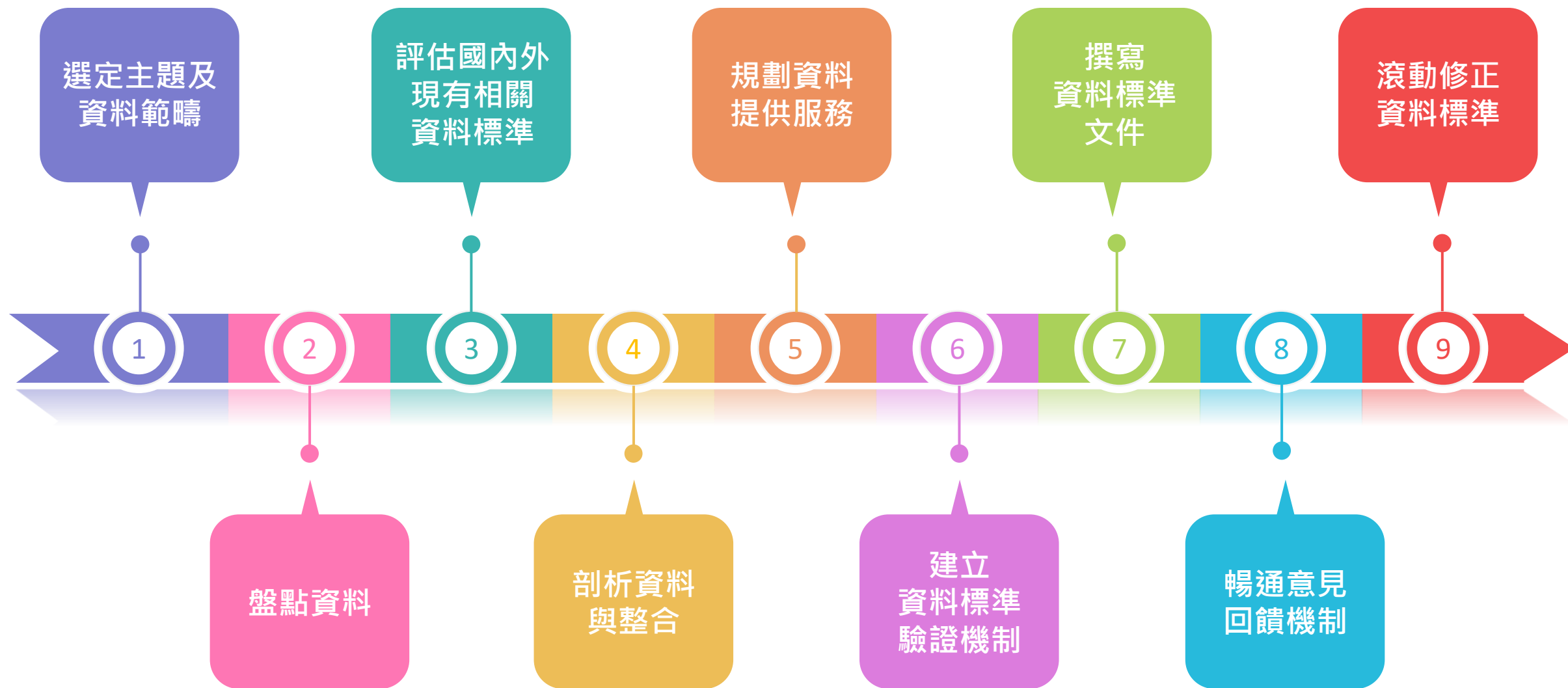
由國發會開發機器檢核機制，有效讓機關內部資料品質提升

促進開放資料應用

資料品質提升後，讓更多民間機關可以應用政府累積的資料，達到開放資料的目的

機關訂定資料標準方法流程(9大步驟)

詳參數位發展部「領域資料標準訂定流程參考指引」



訂定資料標準流程小複習

1. 選定主題及資料範疇



- ① 整理出目前流通性高、民眾使用率高之資料
- ② 選定訂定資料標準之主題名稱、資料範疇

2. 盤點資料



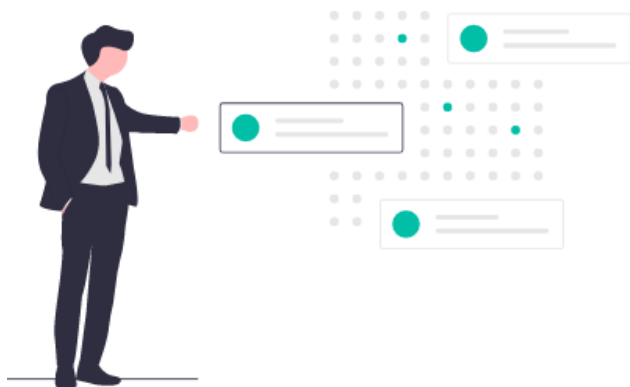
- ① 盤點相關資料之原始文件
- ② 盤點系統資料概況
- ③ 盤點資料欄位範圍

3. 評估國內外資料標準



- ① 評估國內外相關資料標準
- ② 將資料做初步分類，共通標準、主要標準專用標準等

4. 剖析資料與整合



- ① 資料特性分析確認資料領域、主題、類別階層與架構
- ② 建置資料模型
- ③ 設計應用綱要UML圖
- ④ 定義資料典
- ⑤ 進行資料編碼

5. 規劃資料提供服務



- ① 規劃資料提供欄位
- ② 規劃資料服務提供模式

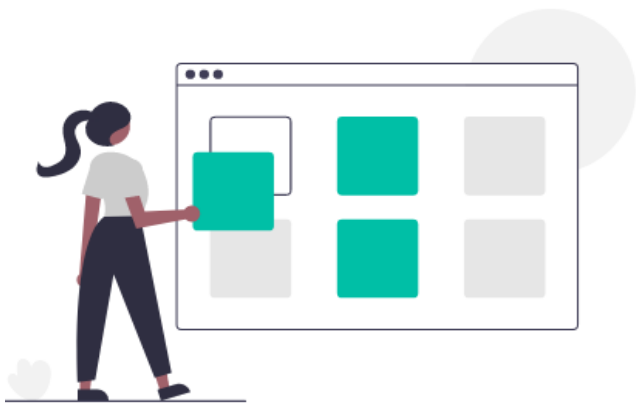
6. 建立資料標準驗證機制



- ① 設計資料標準及資料服務之驗證機制
- ② 建立資料標準驗證機制

訂定資料標準流程小複習

7. 撰寫資料標準文件



- ① 目的
- ② 範圍
- ③ 應用場合及使用限制
- ④ 特性分析
- ⑤ 應用綱要
- ⑥ 資料典
- ⑦ 編碼規則
- ⑧ 標準訂定權責維護
- ⑨ 附錄

8. 暢通意見回饋機制



- ① 完成草案後提報國發會進行初審
- ② 依初審意見修正後提送工作圈複審

9. 滾動修正資料標準



- ① 依複審通過之資料標準內容、資料標準驗證機制，逐步調整資料庫欄位正規化、產出標準化資料
- ② 依實際情形修訂及維護資料標準

資料標準是什麼(1/5)

1

是參考國際標準

結合國內實際資料情形，以適合我國國情與需求

現行標準

中文名稱	英文名稱	說明	型別	值域	附註
日期時間	DateTime	描述特定日子及時刻	DateT ime	[0000-9999][01-12][01-31][00-24][00-59][00-59]	引用標準名稱：CNS 7648 「資料元 件及交換格式－資訊交換－日期及 時間表示法」 (ISO 8601 Data elements and interchange formats - Information interchange - Representation of dates and times)
時間	Time	描述特定時刻	Time	[00-24][00-59][00-59]	引用標準名稱：CNS 7648 「資料元 件及交換格式－資訊交換－日期及 時間表示法」 (ISO 8601 Data elements and interchange formats - Information interchange - Representation of dates and times)
日期	Date	描述特定日子	Date	[0000-9999][01-12][01-31]	引用標準名稱：CNS 7648 「資料元 件及交換格式－資訊交換－日期及 時間表示法」 (ISO 8601 Data elements and interchange formats - Information interchange - Representation of dates and times)

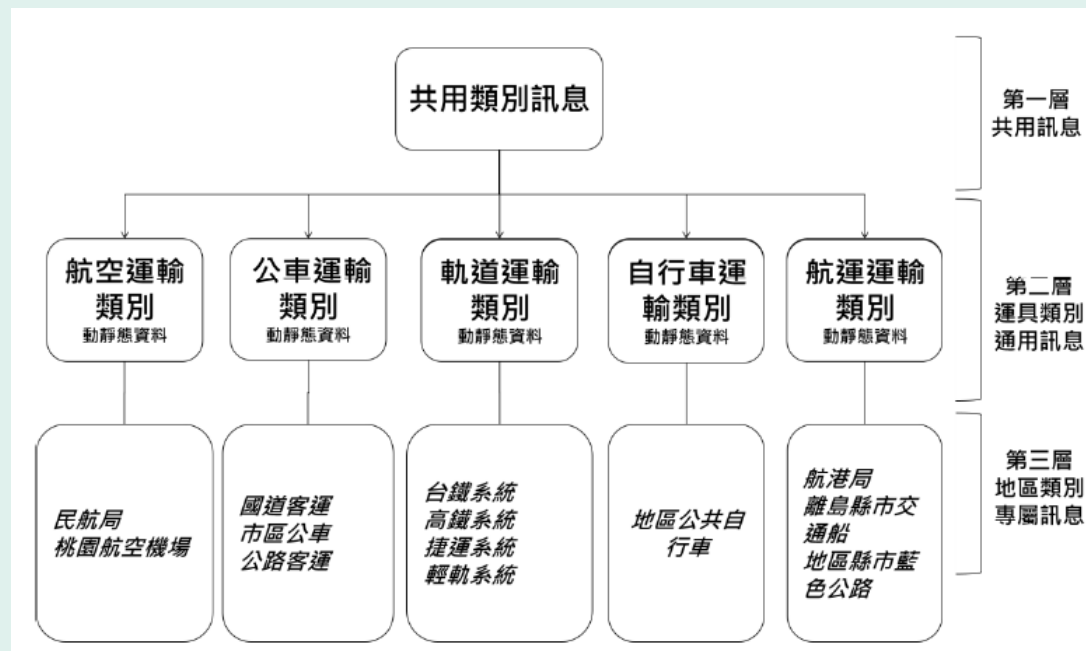
本國國情需要民國年、季、旬

2

能資料特性分析

可透過資料特性分析確立階層與架構

交通部範例



公共運輸旅運標準

資料標準是什麼(2/5)

3

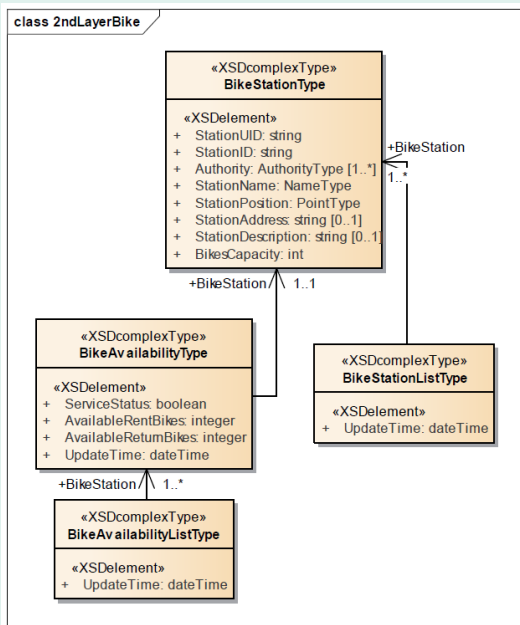
能建立資料模型

設計應用綱要(UML)符合資料建模(Data Modeling)

交通部範例-公共自行車



公共自行車資料模型圖



公共自行車應用綱要

交通部範例-公共自行車服務應用

區域	租賃站點查詢	可借車輛	可停空位
中正區	仁愛林森路口	1	41
中正區	捷運善導寺站(1號出口)	7	40
中正區	南昌公園	12	22
中正區	國家圖書館	23	27
中正區	捷運臺大醫院(4號出口)	29	18
中正區	信義連雲街口	25	14

公共自行車站點動態

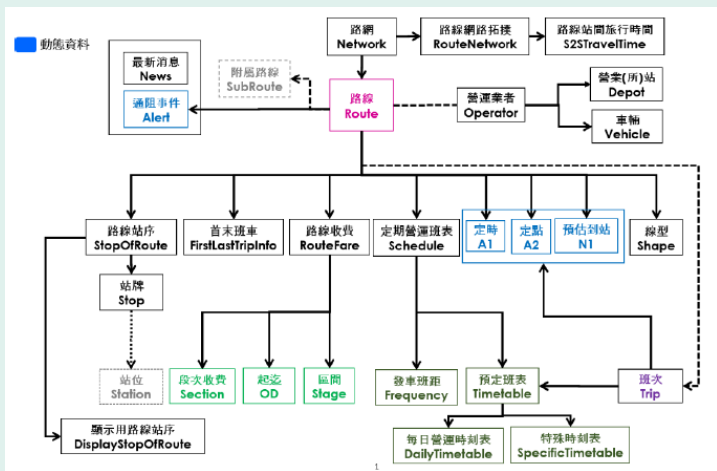
資料標準是什麼(3/5)

3

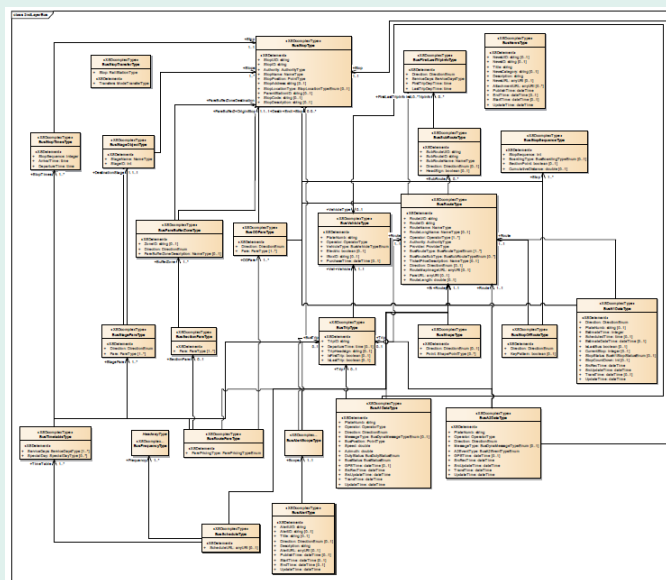
能建立資料模型

設計應用綱要(UML)符合資料建模(Data Modeling)

交通部範例-公路運輸(市區公車、公路客運)



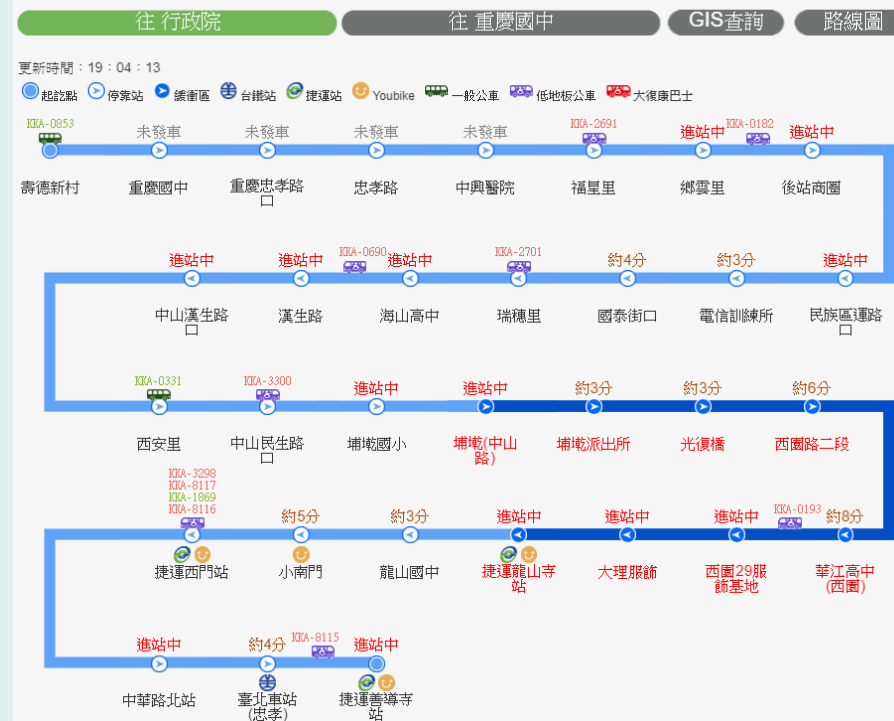
公路運輸資料模型圖



公路運輸應用綱要

交通部範例-公路運輸服務應用

265區



市區公車路線動態

資料標準是什麼(4/5)

4

能歸納資料典

可歸納資料典以明確定義資料意涵與格式，讓機關單位、民間團體更快速瞭解資料樣貌

資料標準平臺

政府資料標準平臺 schema.gov.tw 網站導覽 | 問題回報 | English

首頁 / 資料標準列表 / 內容

資料標準 共通性 > 共通性 > 共通資料

資料標準URI下載 [XML](#) [JSON](#)。本網站使用之正規表示式以 PHP 版本為主。

Properties from 共通資料

中文名稱	英文名稱	說明	型別	值域	附註	附件連結	驗證
姓名	Name	姓氏與名字	String				
國民身分證統一編號	Id	辨別個人身分之專屬代碼	String	[A-Z][1-2][00000000-99999999]	第1碼編碼：縣市別 第2碼編碼：1為男性、2為女性		驗證
性別	Gender	依國民身分證統一編號第2碼編訂	String	[1-2]	1為男性 2為女性		驗證
地址-村里	Village	鄉以內之編組為村；鎮、縣轄市及區以內之編組為里	String				驗證

1

描述資料標準欄位

說明欄位內容，使用方式

2

設定資料型別

如字串、數字、代碼表規則

3

設立正規表示式

轉成機讀格式
方便開放資料檢核

資料標準是什麼(5/5)

5

使用統一代碼

資料使用主管機關之統一代碼

範例-性別

中文名稱	英文名稱	說明	型別	值域	附註	附件連結
性別	Gender	依國民身分證統一編號第2碼編訂	String	[1-2]	1為男性 2為女性	

驗證正規表示式

中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證失敗

送出

驗證正規表示式

中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證失敗

送出

驗證正規表示式

中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證成功

送出

驗證正規表示式

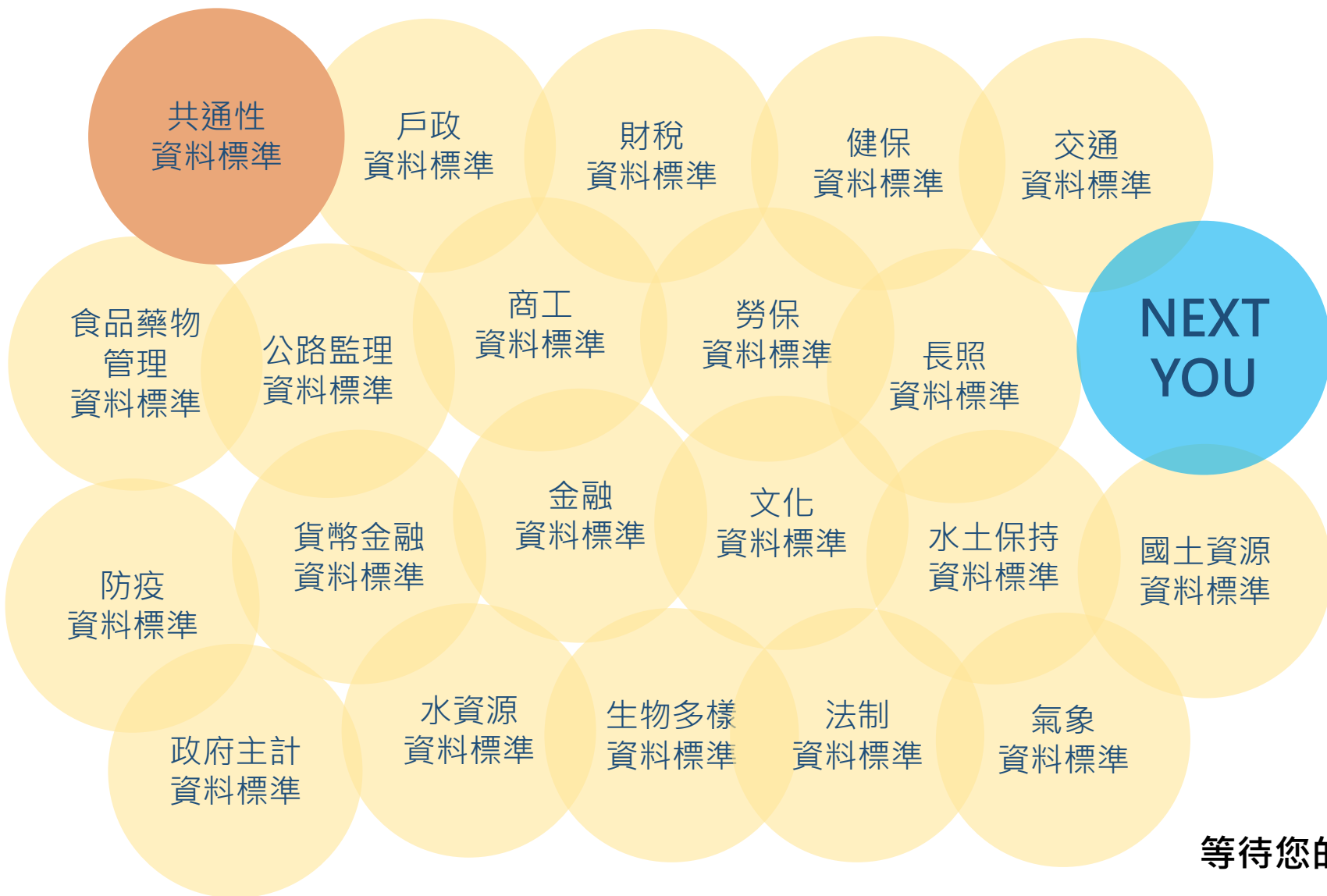
中文名稱: 性別
正規表示式: /^[1-2]\$/

驗證成功

送出

目前資料標準成果

自2018年10月建立「政府資料標準平臺」持續納入新領域資料標準



等待您的一起加入

感謝聆聽

有任何政府資料標準問題歡迎聯絡客服
專線(02) 89787198 #289、#281
Email : schema@moda.gov.tw

一起努力

A person in a red jacket stands on a grassy mountain peak, looking out over a vast sea of white clouds. The sky is clear and blue. The foreground shows the rocky and grassy terrain of the mountain.