

PTX資料建模與XMLSpy使用技巧

創代科技股份有限公司 Datarget Inc.

執行長 關嘉宏 博士

<http://www.datarget.com.tw>

臺北市基隆路二段189號17-4F

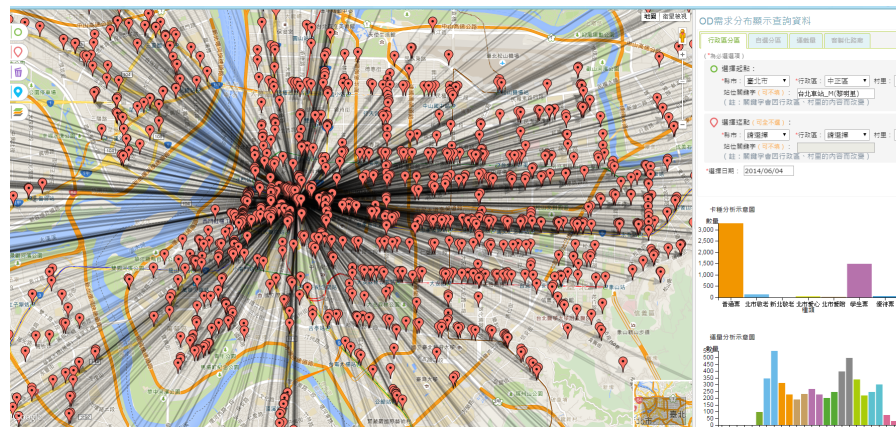
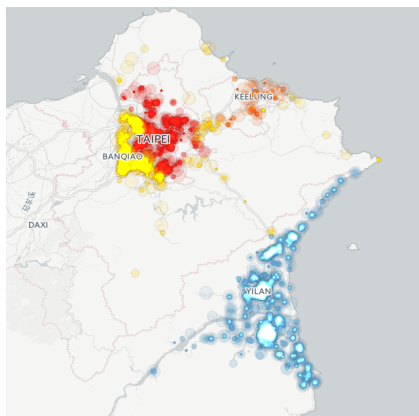
+886-2-8732-0511

Who am I ?

闕嘉宏

<https://tw.linkedin.com/in/charleschueh>

- 2013, 「創代科技股份有限公司」 創辦人兼執行長
- 2012, 雲端Telematics&3G/4G物聯網資通訊服務
- 2010, GIS&ITS聯合應用
- 2006, 台大碩士保送台大博士班
- 2004, 逢甲GIS&交通管理學系



- GIS & ITS 雙系統技術整合
- 車載資通訊平台整合技術
- 近場即時交通資訊技術
- 3G/4G 多元資訊服務設計

- 多模車機整合及先進安全駕駛輔助系統
- 票證增值服務應用研究
- 公共運輸整合系統設計與分析
- 異質性時間序列融合模型

Who am I ?

闕嘉宏

<https://tw.linkedin.com/in/charleschueh>

- 交通部-PTX資料模型設計師
- 台灣大學先進公共運輸研究中心-技術總顧問
- 淡江大學巨量資料分析與應用-授課教授
- 臺北捷運公司 TRTS4-數據顧問
- 國道高速公路局-數據分析顧問
- 2015,2016 台灣資料科學年會-智慧交通授課教授
- 2015,2016 台北市公務人員訓練中心-資料科學授課教授
- 2016 世界ITS大會 墨爾本 資料科學特聘教授

1. 國際性資料模型發展關鍵-Google GTFS為例

2. PTX結構設計特性

3. 繼承技術與特例

4. 重覆封裝引用

5. 資料清單與特例

6. XSD輸出UML相關技術

1. 國際性資料模型發展關鍵

General Transit Feed Specification (GTFS)

<https://developers.google.com/transit/gtfs/>

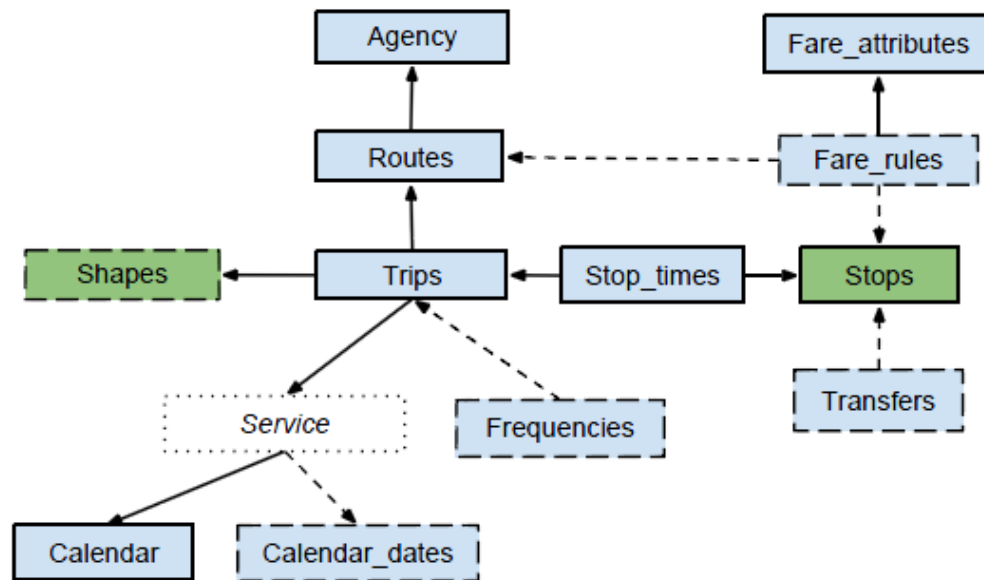
- 利用Google Map圖台進行資訊服務
- Static Public Transit data
 - [Trip Updates](#)
 - [Service Alerts](#)
 - [Vehicle Positions](#)
- GTFS RealTime

Overview



The GTFS specification defines the following files along with their associated content:

Filename	Required	Defines
agency.txt	Required	One or more transit agencies that provide the data in this feed.
stops.txt	Required	Individual locations where vehicles pick up or drop off passengers.
routes.txt	Required	Transit routes. A route is a group of trips that are displayed to riders as a single service.
trips.txt	Required	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
stop_times.txt	Required	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	Required	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.
calendar_dates.txt	Optional	Exceptions for the service IDs defined in the calendar.txt file. If calendar_dates.txt includes ALL dates of service, this file may be specified instead of calendar.txt



Key

Entity (File)

Spatial entity

Virtual entity

————> Foreign key

- - - -> Foreign key - optional

提供開放式資料交換平台 <http://www.gtfs-data-exchange.com/>

- 可以下載到全世界提供的GTFS資料，並提供檢核工具
- 判定是否已正式被Google Map所使用
- 地區分類，提供單位，提供表格

All Agencies | Recent Updates

Transit Agencies Providing GTFS Data

View: by Agency | [by Location](#) | [by Last Update](#) | [as Table](#)

Filter: All Sources | [Official Sources](#)

Data is Available for 1000 Transit Agencies.

[RSS of All Updates](#)

A

- [A. Reich GmbH Busbetrieb](#)
- [ABCBus Service](#)
- [ABOUT ME s.r.o.](#)
- [ABQ Ride](#)
- [AC Transit](#)
- [ACAD](#)
- [ACTION](#)
- [ACTP S.r.l.](#)
- [ACTV S.p.a](#)
- [AMAT](#)
- [AMAT Palermo S.p.A.](#)
- [AMTRAK](#)
- [ARILIX OÜ](#)

Upload By [trimet-archiver](#) 2 months ago
pertains to [Portland Aerial Tram](#) [Portland Streetcar](#) [TriMet](#)

Archived from <http://developer.trimet.org/schedule/gtfs.zip>

[trimet-archiver_20160505_0144.zip](#) 32047933 [more info »](#)

Upload By [bkv-zrt-archiver](#) 2 months ago
pertains to [BKK](#)

Archived from http://www.bkk.hu/gtfs/budapest_gtfs.zip

[bkk_20160505_0141.zip](#) 29543662 [more info »](#)

Upload By [charlottesville-archiver](#) 2 months ago
pertains to [Charlottesville Transit Service](#)

Archived from <http://avlweb.charlottesville.org/rtt/public/utility/GTFS.aspx>

[charlottesville-transit-service_20160505_0140.zip](#) 164651 [more info »](#)

定義FeedInfo進行資料版本控制

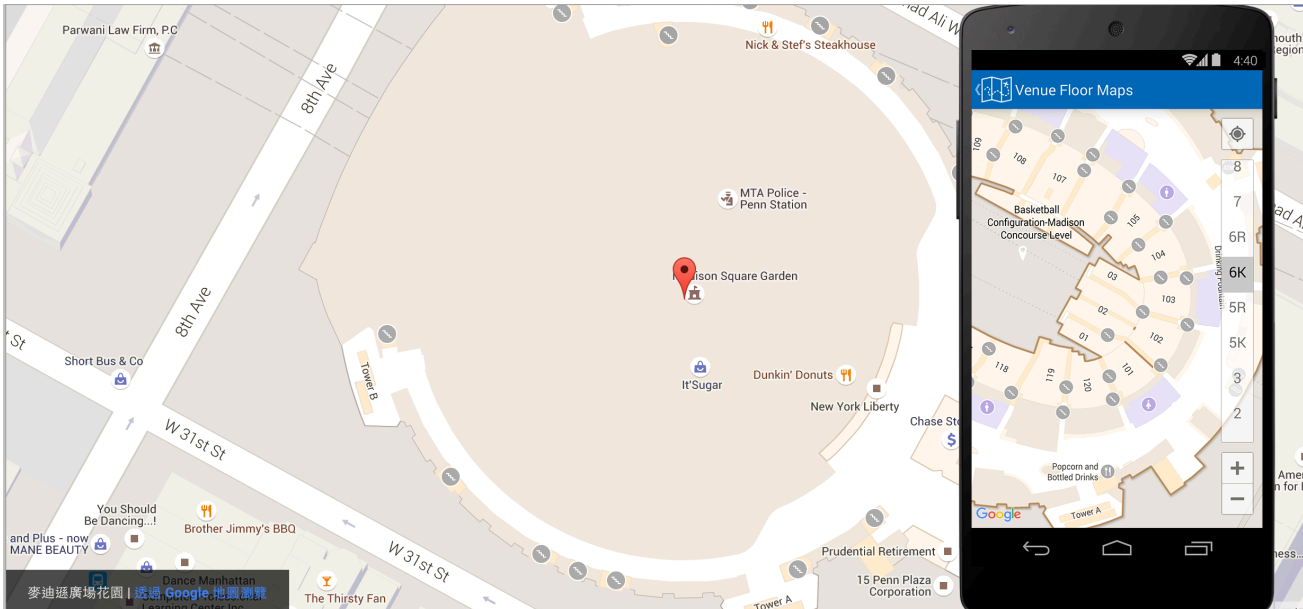
- 公共運輸資料一定會遇到路線，站牌調度的問題
- 鮮少人能夠真正的對資料版本進行控管

Field Name	Required	Details
feed_publisher_name	Required	Contains the full name of the organization that publishes the feed. (This may be the same as one of the agency_name values in agency.txt .) GTFS-consuming applications can display this name when giving attribution for a particular feed's data.
feed_publisher_url	Required	Contains the URL of the feed publishing organization's website. (This may be the same as one of the agency_url values in agency.txt .) The value must be a fully qualified URL that includes http:// or https:// , and any special characters in the URL must be correctly escaped. For a description of how to create fully-qualified URL values, see http://www.w3.org/Addressing/URL/4_URI_Recommendations.html .
feed_lang	Required	Contains a IETF BCP 47 language code specifying the default language used for the text in this feed. This setting helps GTFS consumers choose capitalization rules and other language-specific settings for the feed. For an introduction to IETF BCP 47, please refer to http://www.rfc-editor.org/rfc/bcp/bcp47.txt and http://www.w3.org/International/articles/language-tags/ .

Field Name	Required	Details
feed_start_date	Optional	The feed provides complete and reliable schedule information for service in the period from the beginning of the feed_start_date day to the end of the feed_end_date day. Both days are given as dates in YYYYMMDD format as for calendar.txt , or left empty if unavailable.
feed_end_date		The feed_end_date date must not precede the feed_start_date date if both are given. Feed providers are encouraged to give schedule data outside this period to advise of likely future service, but feed consumers should treat it mindful of its non-authoritative status. If feed_start_date or feed_end_date extend beyond the active calendar dates defined in calendar.txt and calendar_dates.txt , the feed is making an explicit assertion that there is no service for dates within the feed_start_date or feed_end_date range but not included in the active calendar dates.
feed_version	Optional	The feed publisher can specify a string here that indicates the current version of their GTFS feed. GTFS-consuming applications can display this value to help feed publishers determine whether the latest version of their feed has been incorporated.

我想讓全世界都看到台灣

- GTFS-TW
- Google Map Indoor
- GTFS RealTime



What is Live Transit Updates for Google Maps?

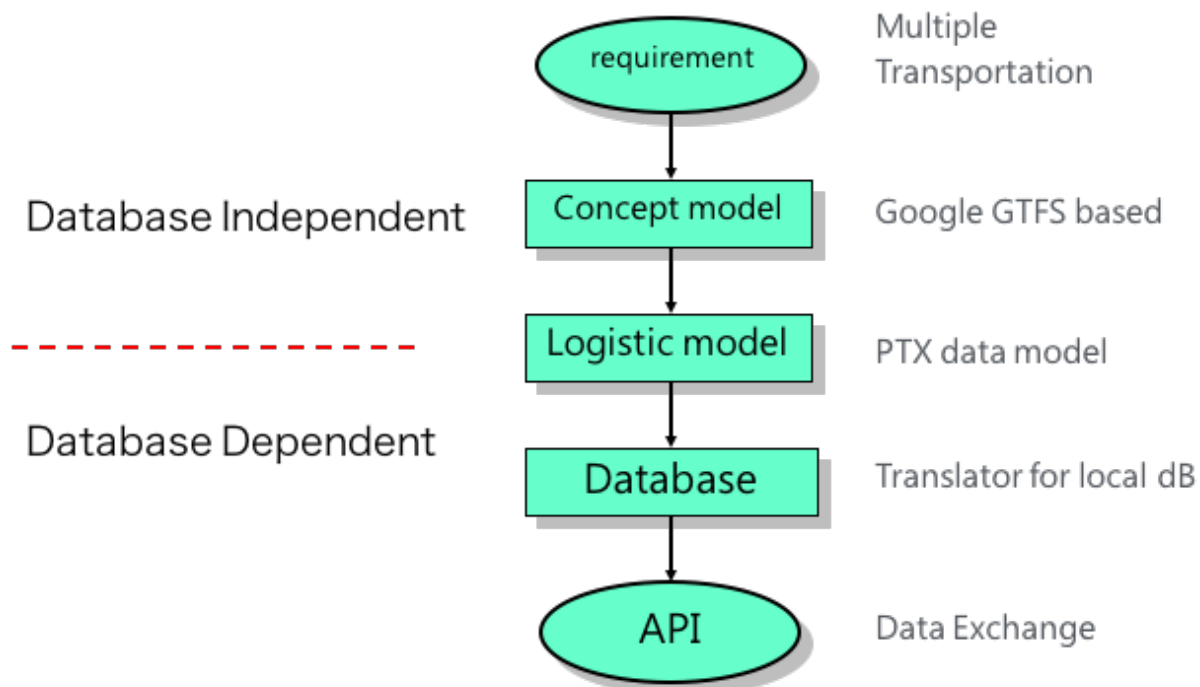
Live Transit Updates is a service providing real-time transit updates to users of Google Maps for mobile. These updates include live departure and arrival times as well as service alerts.



2. PTX資料結構設計特性

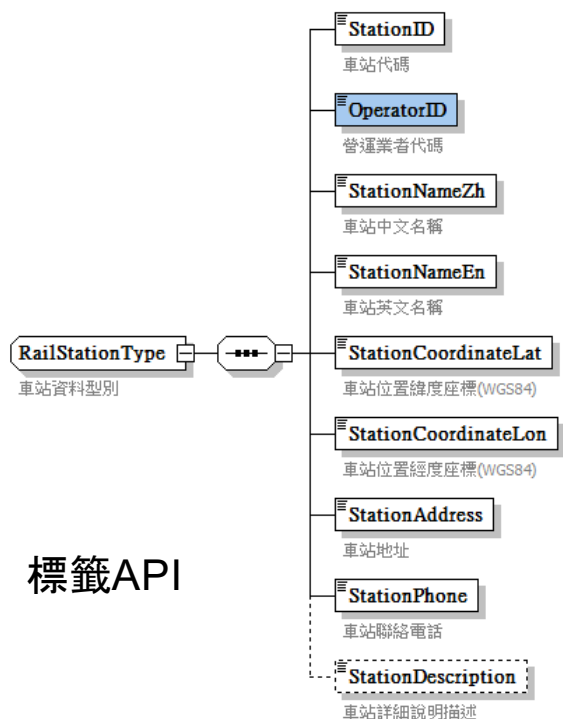
重新認識資料模型的定義

- ★採用開放授權，讓手上的資料（任何資料格式）可以在網路上取得¹
- ★★讓這份資料能以結構化的方式取得（例如用 Excel 取代掃描的表格）²
- ★★★使用開放格式取代專屬格式（例如用 CSV 取代 Excel）³
- ★★★★★使用 固定網址 來表示資料，使其它人可以連結到資料在資料網絡中的位置⁴
- ★★★★★★鏈結你的資料到其它資料，以提供資料之間的脈絡關係，例如兩份資料間的相等關係（owl:sameAs）

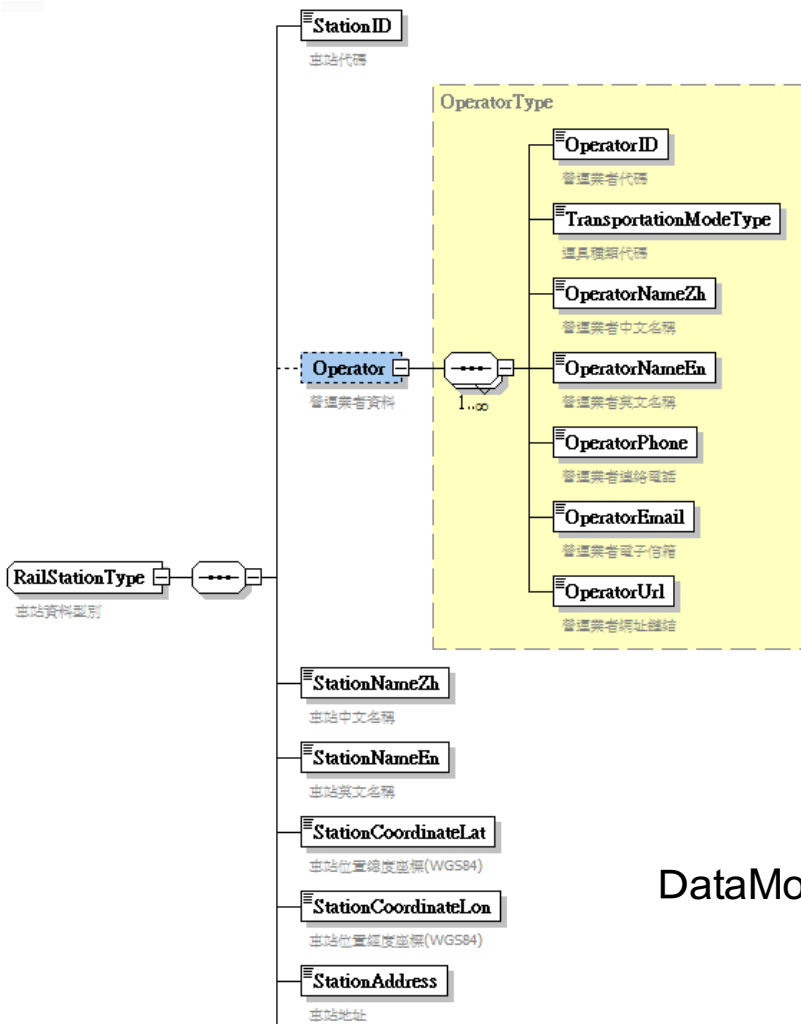


資料模型 vs 標籤API

- 資料模型保留完整的物件屬性
- 標籤API使用簡潔的ID
- 完整性與效率地權衡
- OData的興起



標籤API



DataModel

The screenshot shows the details for the **Operator** type in an OData metadata editor. The **Details** tab is selected, showing the following information:

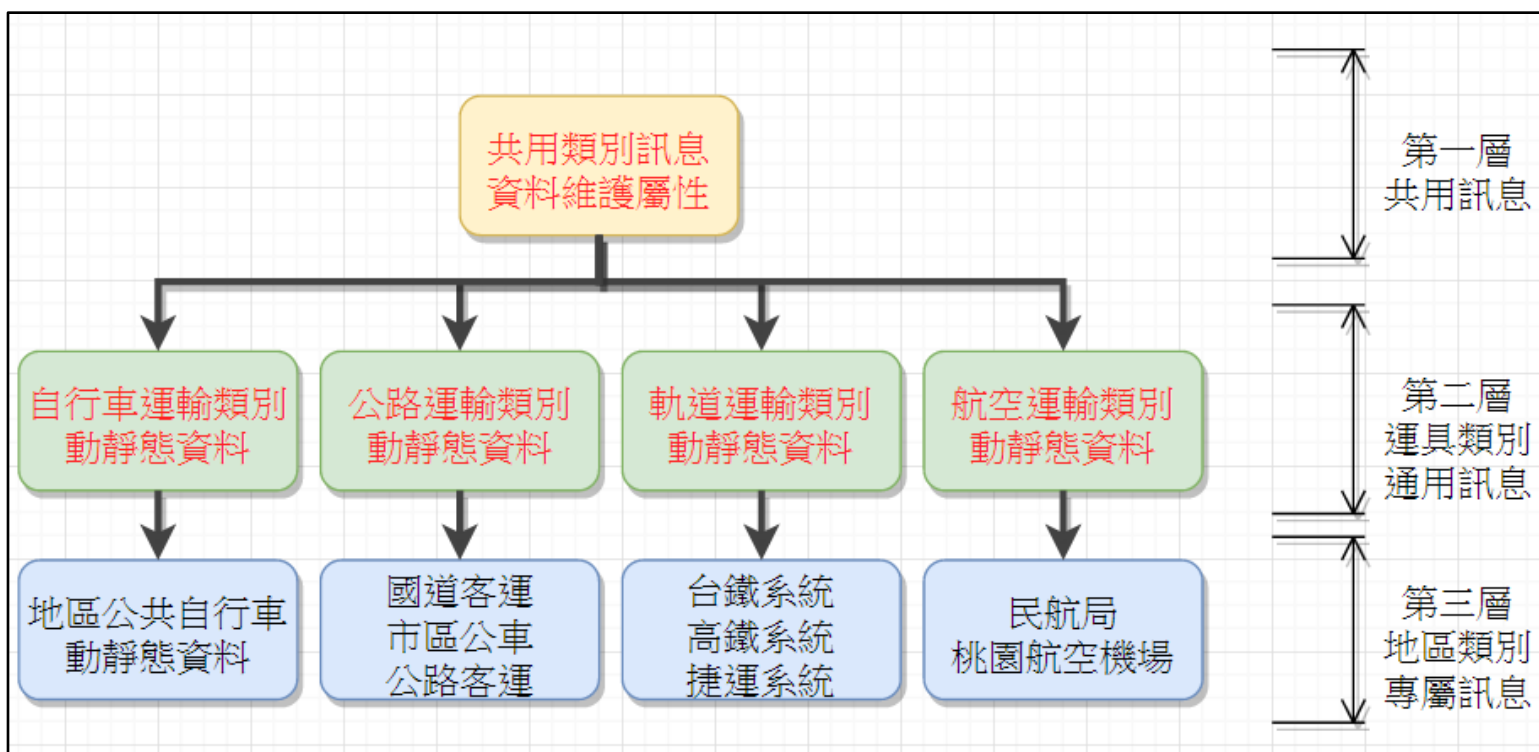
Property	Value
name	Operator
isRef	<input type="checkbox"/>
minOcc	0
maxOcc	1
type	OperatorType
content	complex
derivedBy	
mixed	
nullable	
block	
form	
id	

On the right side, there are tabs for **Globals** and **Namespaces**. Below the details, there is a **Facets** section.



開發出屬於台灣的三層式資料模型架構

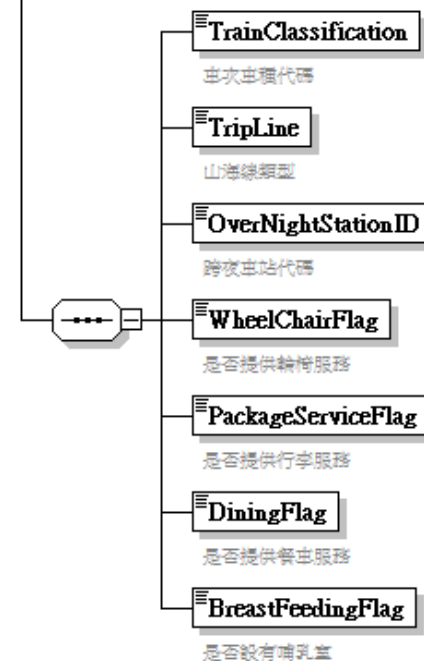
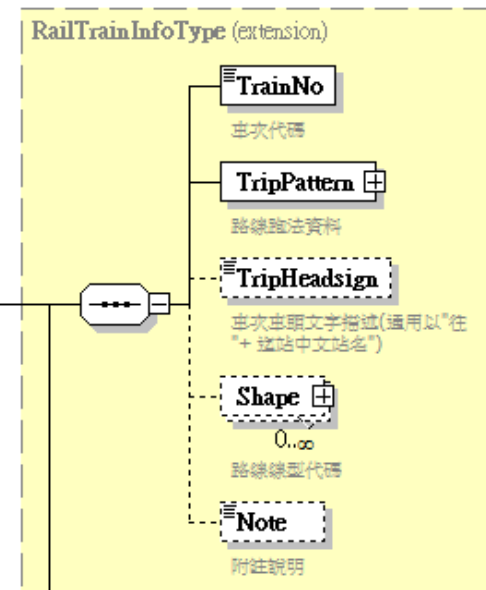
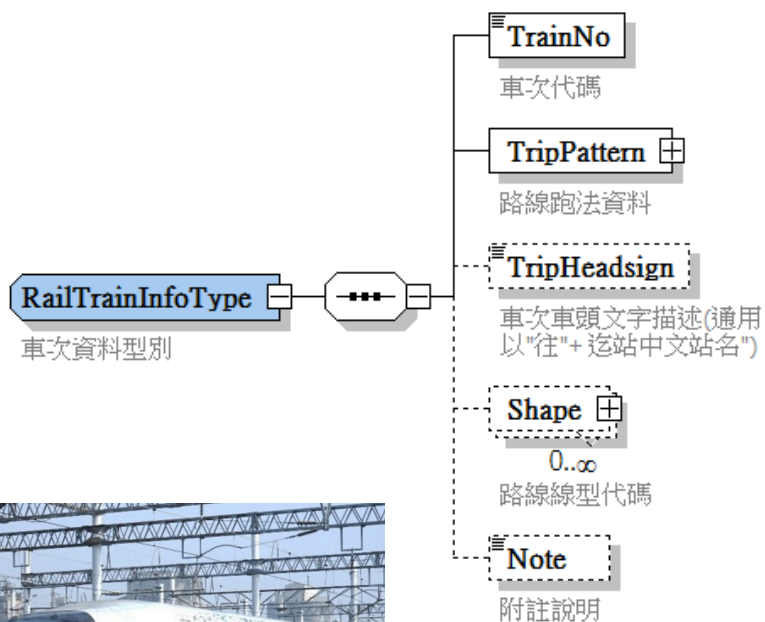
- 以GTFS技術為基底
- 引入物件導向之父類別概念
- 地方管理者可自行擴充子類別屬性



3.繼承技術與特例

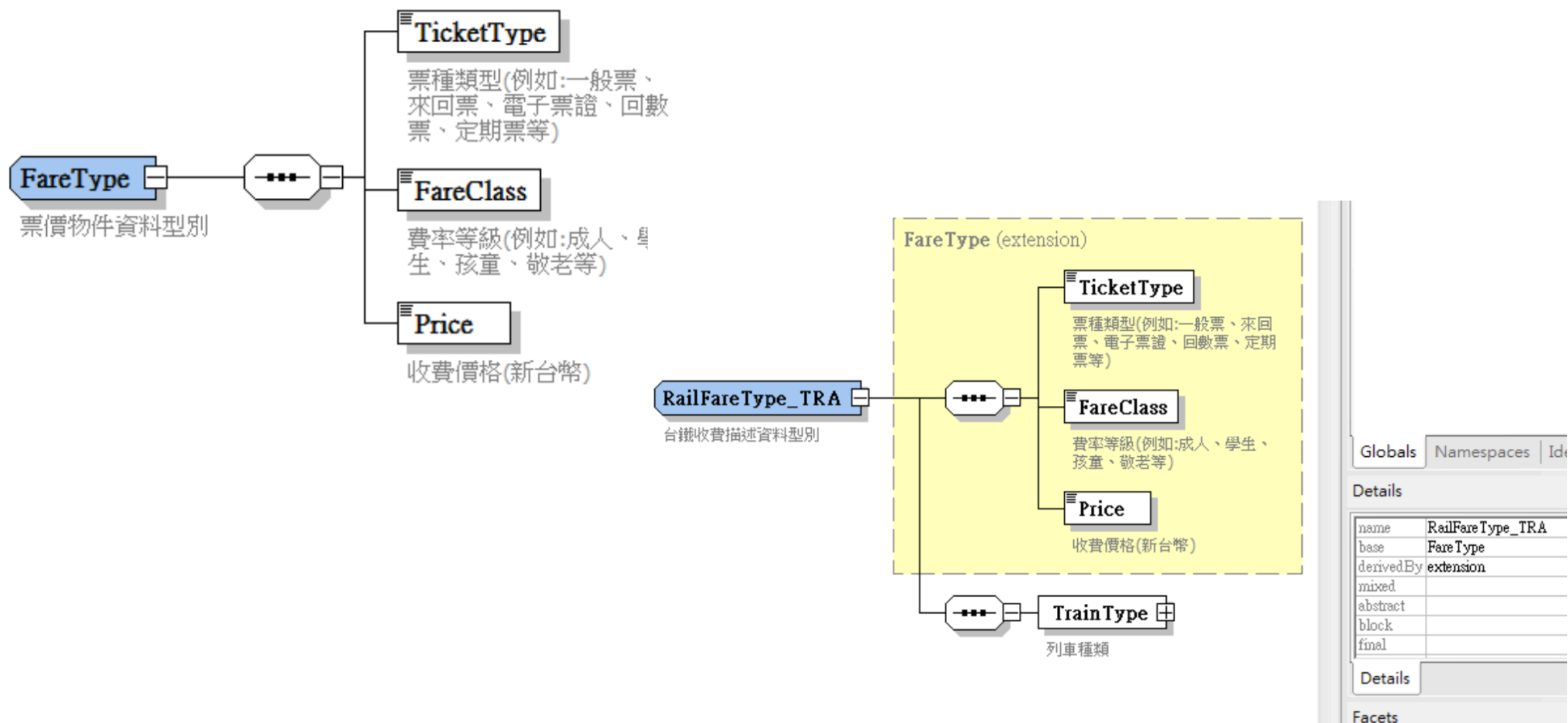
Extension for Local Attributes

- 台鐵與高鐵都屬於軌道列車
- 台鐵獨有車種，哺乳車廂，輪椅車廂，或是過夜車站等欄位



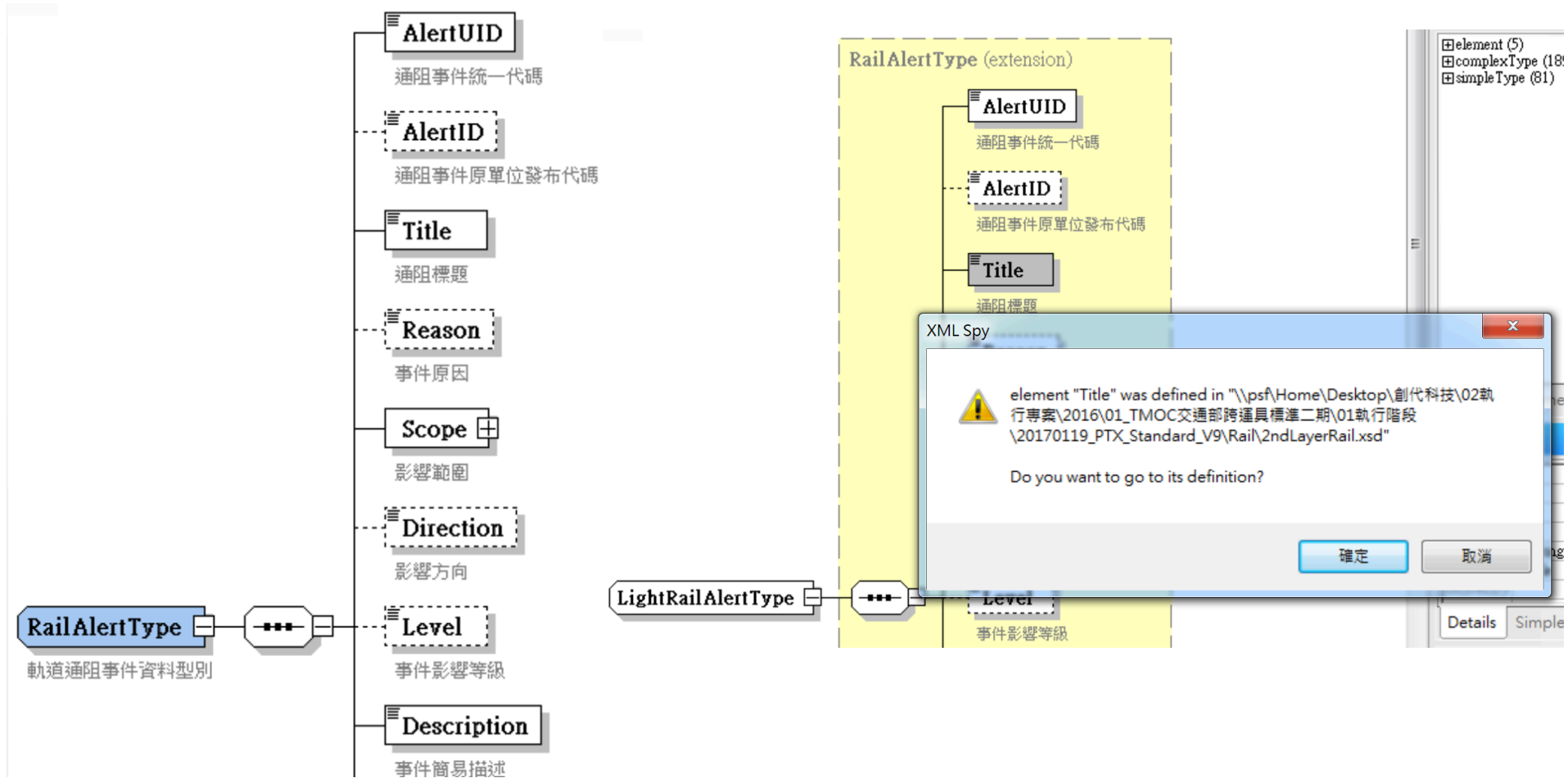
Extension for Local Attributes

■ 實際應用與教學情境說明



繼承技術因為確保了父類別型別，因此造成逆向填值產生特例

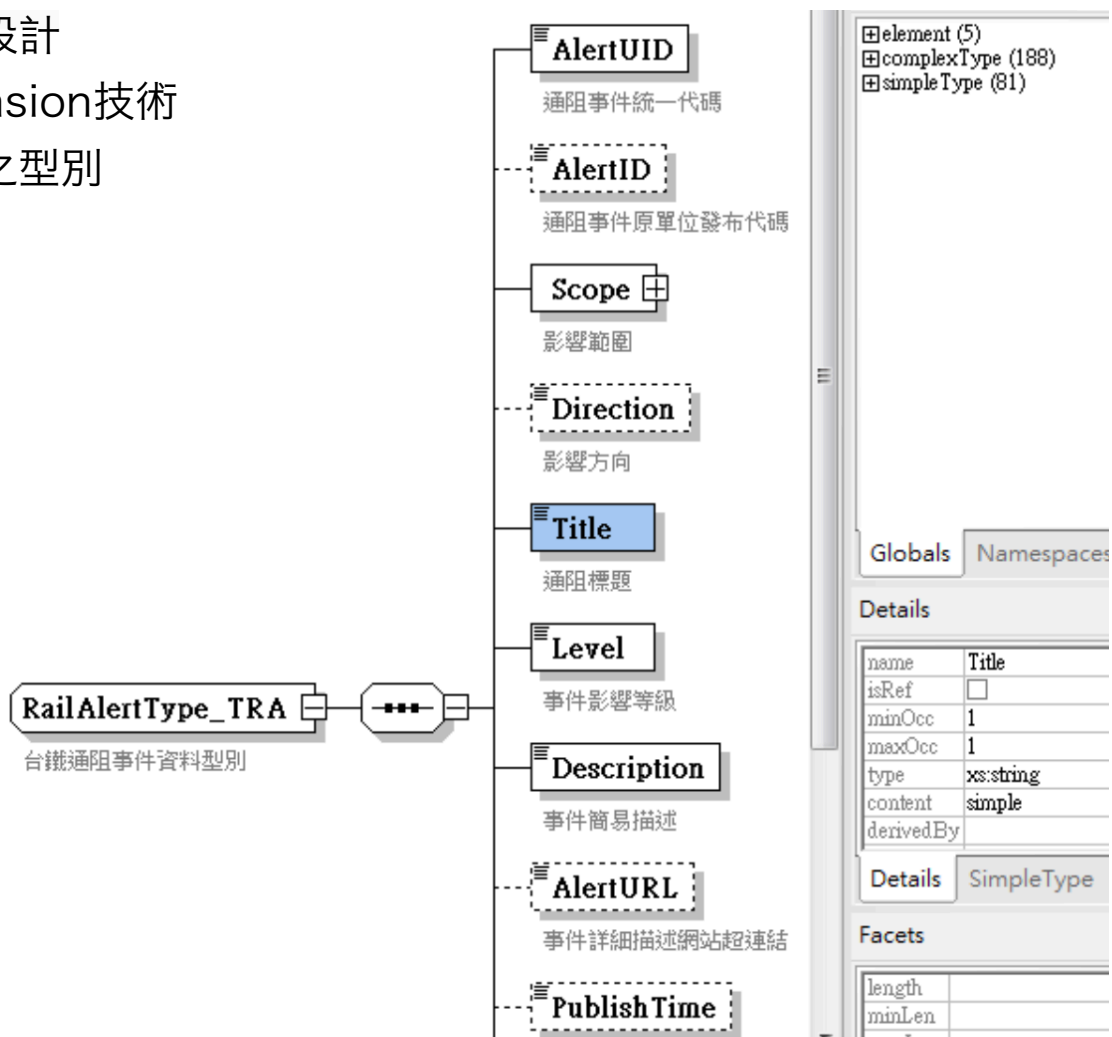
子類別無法逆向改變父類別「既有的值域」



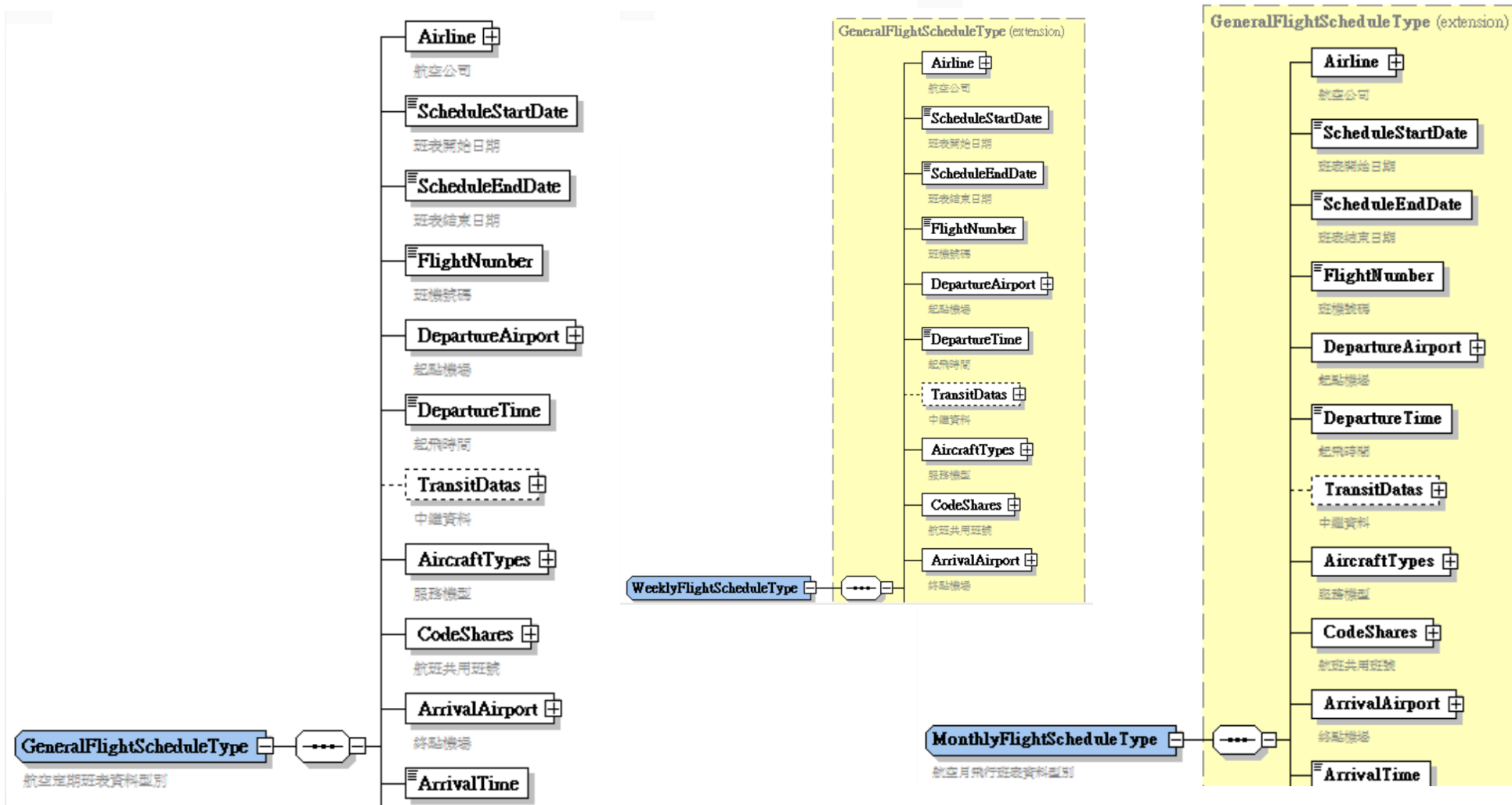
繼承技術因為確保了父類別型別，因此造成逆向填值產生特例

為了確保各個運具仍然「內部客製化欄位」的設計解套

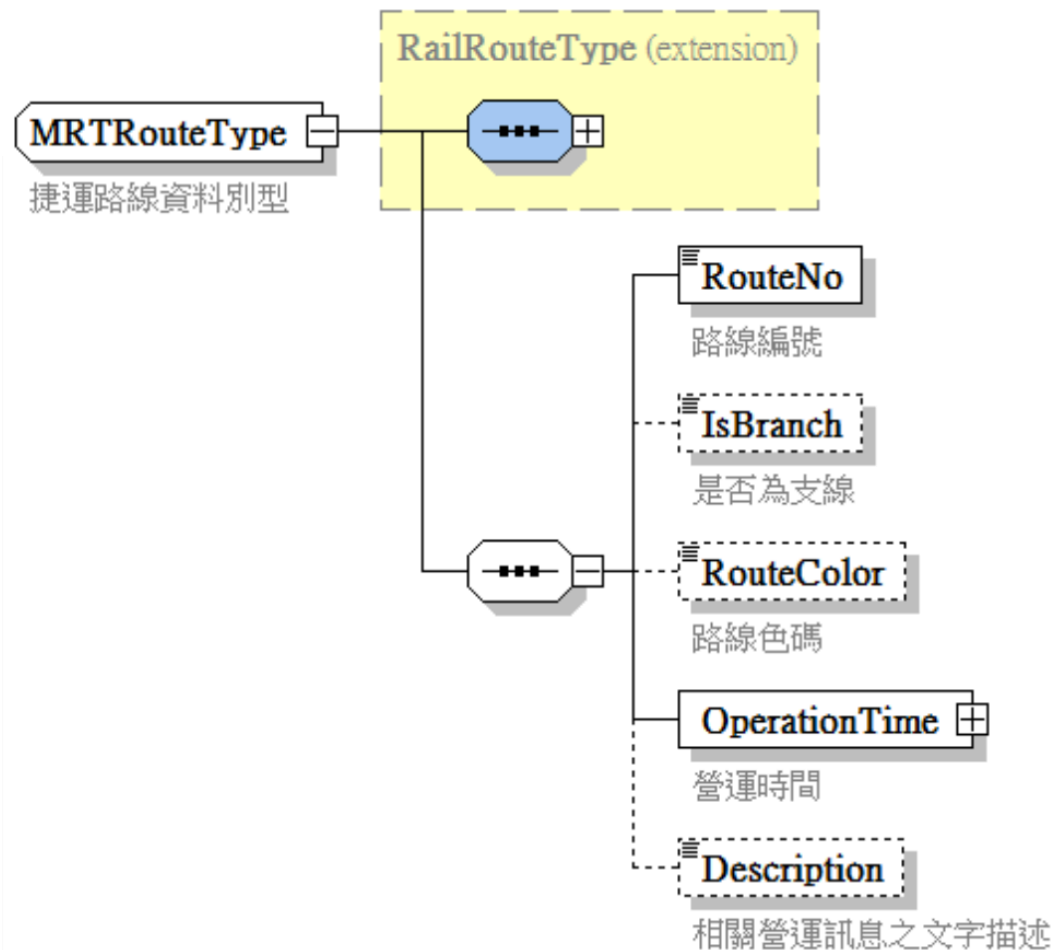
- 仿造父類別之結構設計
- 但不直接使用extension技術
- 可改變各欄位封裝之型別



航空時刻表，可透過定期時刻表進行收編



捷運路線資訊，具有更細緻的資料需求



4.重覆封裝引用

公車系統問題範例

- 台北市系統獨有的主副路線
- 新北市路線沒有定義方向性，去程ID1, 返程ID2
- GTFS透過Trip_id來組合出Route_id的集合

PathDetail.aspx	附屬路線與路線對應資訊
屬性	說明
Id	主索引
PathAttributeId	附屬路線 ID
StopId	站牌 ID
SequenceNo	站序
Type	類別：0：正常 CPT：控制點)
TypeAttribute	類別屬性描述 (例如控制點時間之描述)

11111		藍27
11111	155955	藍27(不進三總)
11111	157692	藍27(僅往程)
11111	157693	藍27(僅返程)
11111	11111	藍27(進三總)

Field Name	Required	Details
route_id	Required	Contains an ID that uniquely identifies a route. This value is referenced from the routes.txt file.
service_id	Required	The service_id contains an ID that uniquely identifies a set of dates when service is available for one or more routes. This value is referenced from the calendar.txt or calendar_dates.txt file.
trip_id	Required	Contains an ID that identifies a trip. The trip_id is dataset unique.
trip_headsign	Optional	Contains the text that appears on a sign that identifies the trip's destination to passengers. Use this field to distinguish between different patterns of service in the same route. If the headsign changes during a trip, you can override the trip_headsign by specifying values for the stop_headsign field in stop_times.txt .
trip_short_name	Optional	Contains the text that appears in schedules and sign boards to identify the trip to passengers, for example, to identify train numbers for commuter rail trips. If riders do not commonly rely on trip names, please leave this field blank. A trip_short_name value, if provided, should uniquely identify a trip within a service day; it should not be used for destination names or

路線類別也容易因為這樣的狀況，產生問題

公車系統問題範例

- 公車路線類別定義全部都不一樣
- 高鐵接駁，臺灣好行，觀光巴士，免費巴士沒有交集定義

快線公車

幹線公車

紅線接駁

橘線接駁

一般公車

其他

✓ 全部

- 一般公車
- 捷運先導公車
- 捷運紅線接駁公車
- 捷運藍線接駁公車
- 捷運綠線接駁公車
- 捷運棕線接駁公車
- 捷運橘線接駁公車
- 山區公車
- 通勤公車
- 市民小巴
- 新巴士
- 幹線專車
- 特定路線專車
- 活動專車

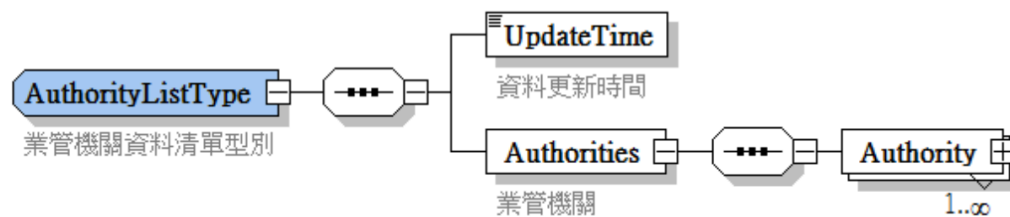
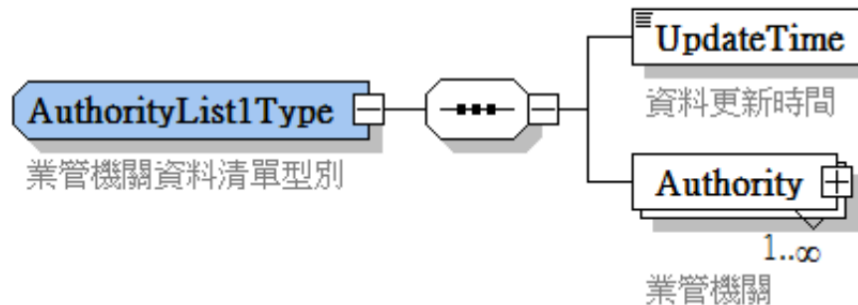
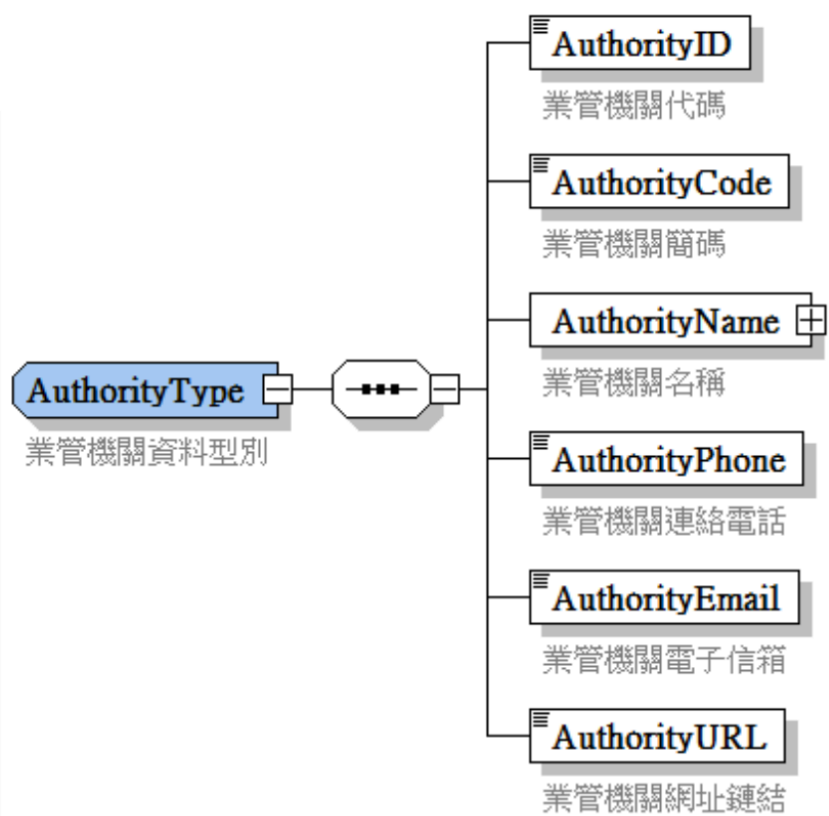
✓ 路線類別

- 市區公車
- 觀光公車
- 高鐵接駁車
- 綠線
- 藍線
- 棕線
- 橘線
- 黃線
- 紅線

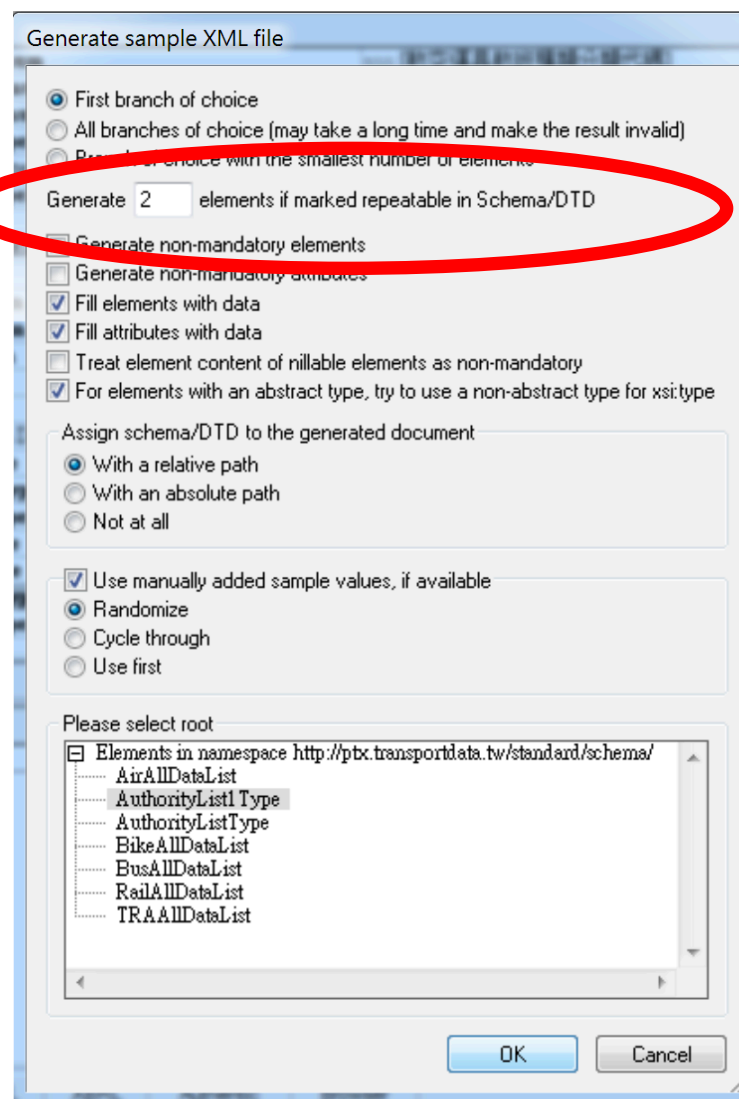
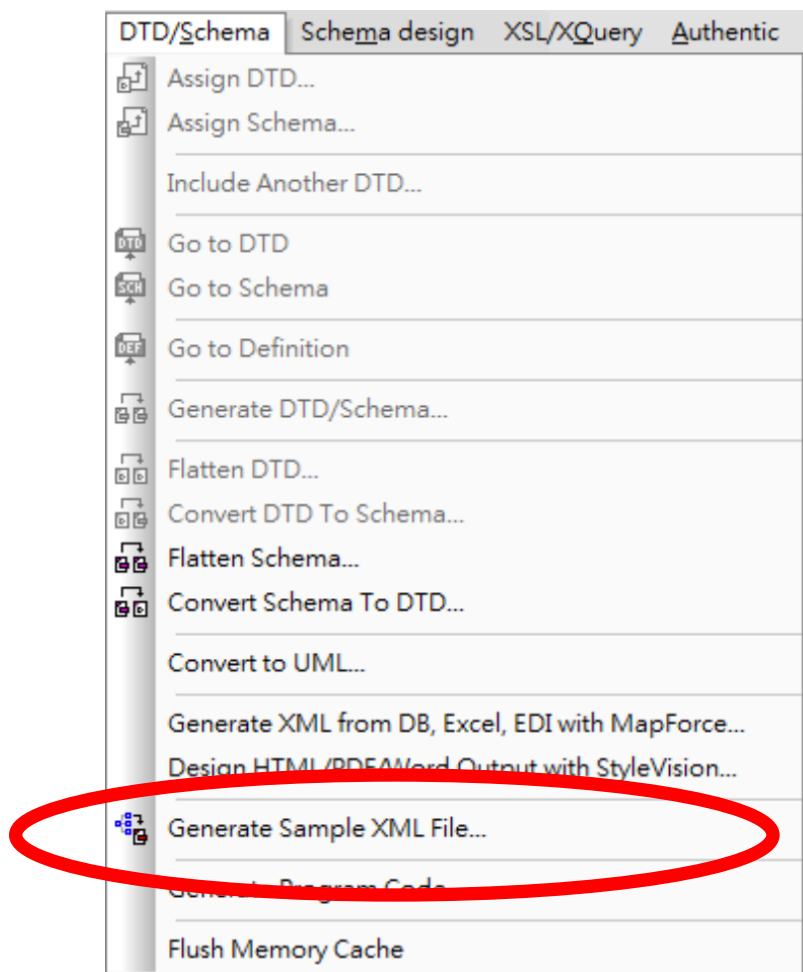
市區公車	免費公車	公路客運
✓ 請選擇(市區公車)		
全部		
桃園客運		
中壢客運		
統聯客運		
新竹客運		
亞通客運		

5. 資料清單與特例

以業管單位為範例，資料清單型別利用基本型別，產生出批次檔（配合更新時間）



利用elementType,產生出XML範例說明



有無巢狀的利與弊

無巢狀結構

```

<UpdateTime>2001-12-17T09:30:47Z</UpdateTime>
<Authority>
  <AuthorityID>String</AuthorityID>
  <AuthorityCode>String</AuthorityCode>
  <AuthorityName>
    <Zh_tw>String</Zh_tw>
    <En>String</En>
  </AuthorityName>
  <AuthorityPhone>String</AuthorityPhone>
  <AuthorityEmail>String</AuthorityEmail>
  <AuthorityURL>http://www.altova.com/</AuthorityURL>
</Authority>
<Authority>
  <AuthorityID>String</AuthorityID>
  <AuthorityCode>String</AuthorityCode>
  <AuthorityName>
    <Zh_tw>String</Zh_tw>
    <En>String</En>
  </AuthorityName>
  <AuthorityPhone>String</AuthorityPhone>
  <AuthorityEmail>String</AuthorityEmail>
  <AuthorityURL>http://www.altova.com/</AuthorityURL>
</Authority>
</nl:AuthorityListType>

```

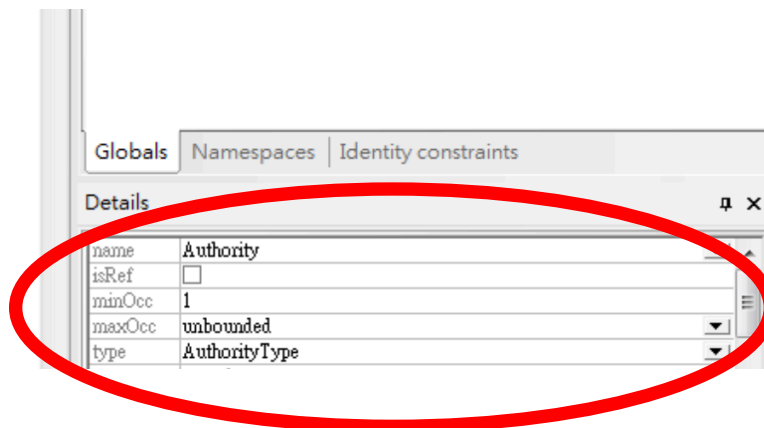
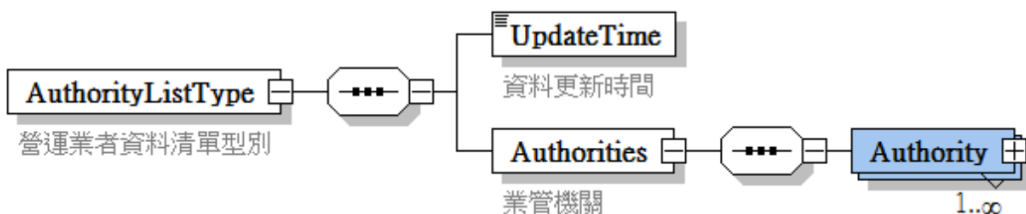
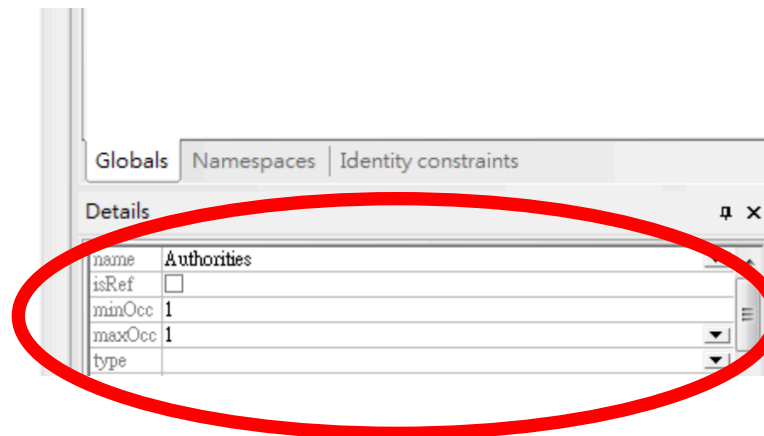
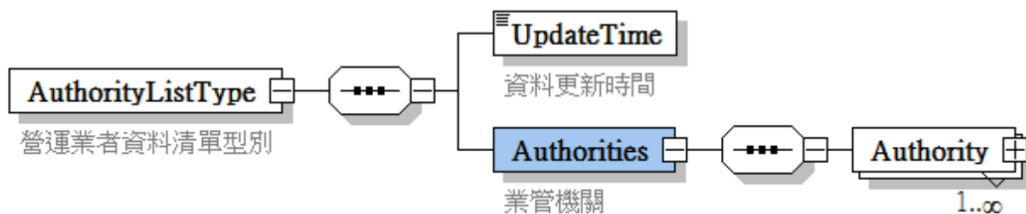
有巢狀結構

```

<UpdateTime>2001-12-17T09:30:47Z</UpdateTime>
<Authorities>
  <Authority>
    <AuthorityID>String</AuthorityID>
    <AuthorityCode>String</AuthorityCode>
    <AuthorityName>
      <Zh_tw>String</Zh_tw>
      <En>String</En>
    </AuthorityName>
    <AuthorityPhone>String</AuthorityPhone>
    <AuthorityEmail>String</AuthorityEmail>
    <AuthorityURL>http://www.altova.com/</AuthorityURL>
  </Authority>
  <Authority>
    <AuthorityID>String</AuthorityID>
    <AuthorityCode>String</AuthorityCode>
    <AuthorityName>
      <Zh_tw>String</Zh_tw>
      <En>String</En>
    </AuthorityName>
    <AuthorityPhone>String</AuthorityPhone>
    <AuthorityEmail>String</AuthorityEmail>
    <AuthorityURL>http://www.altova.com/</AuthorityURL>
  </Authority>
</Authorities>
</nl:AuthorityListType>

```

有無巢狀的利與弊

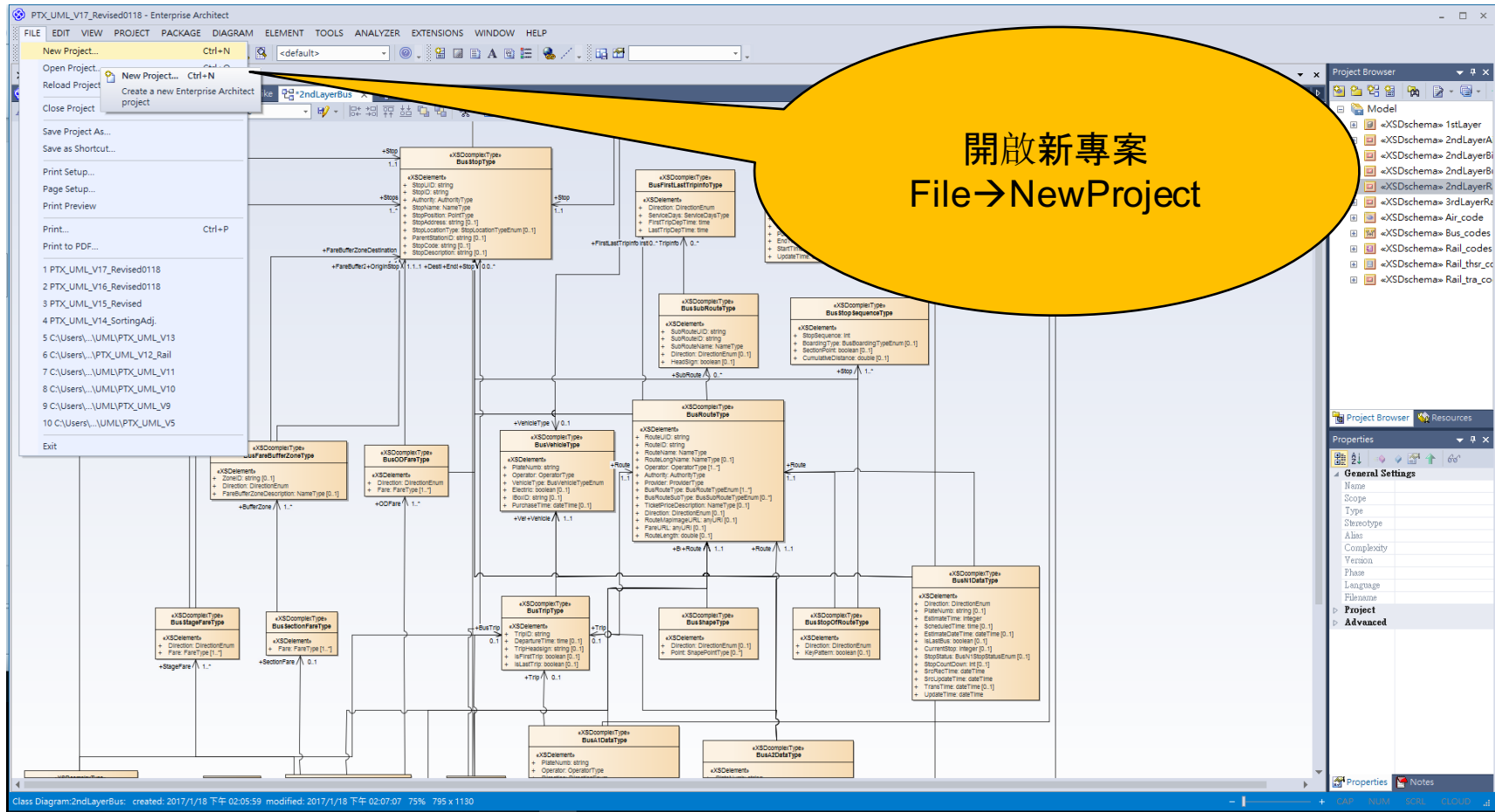


交通部目前購買之軟體技術

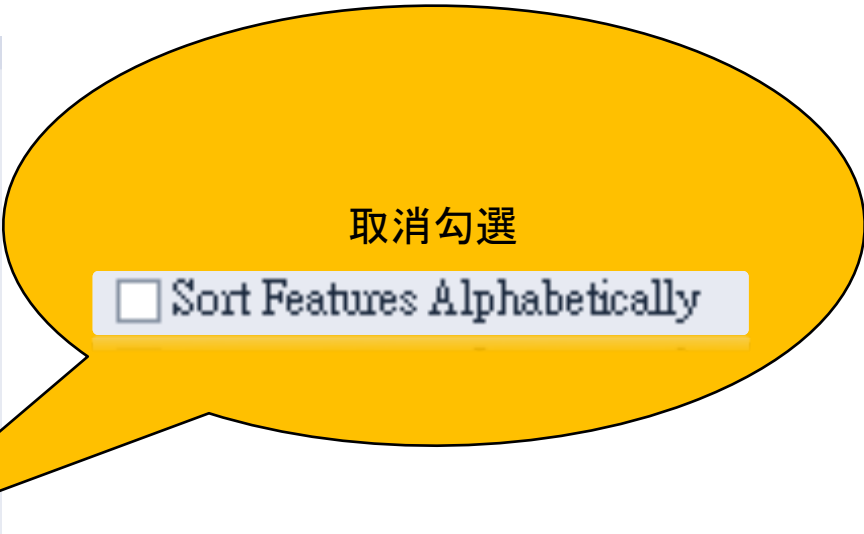
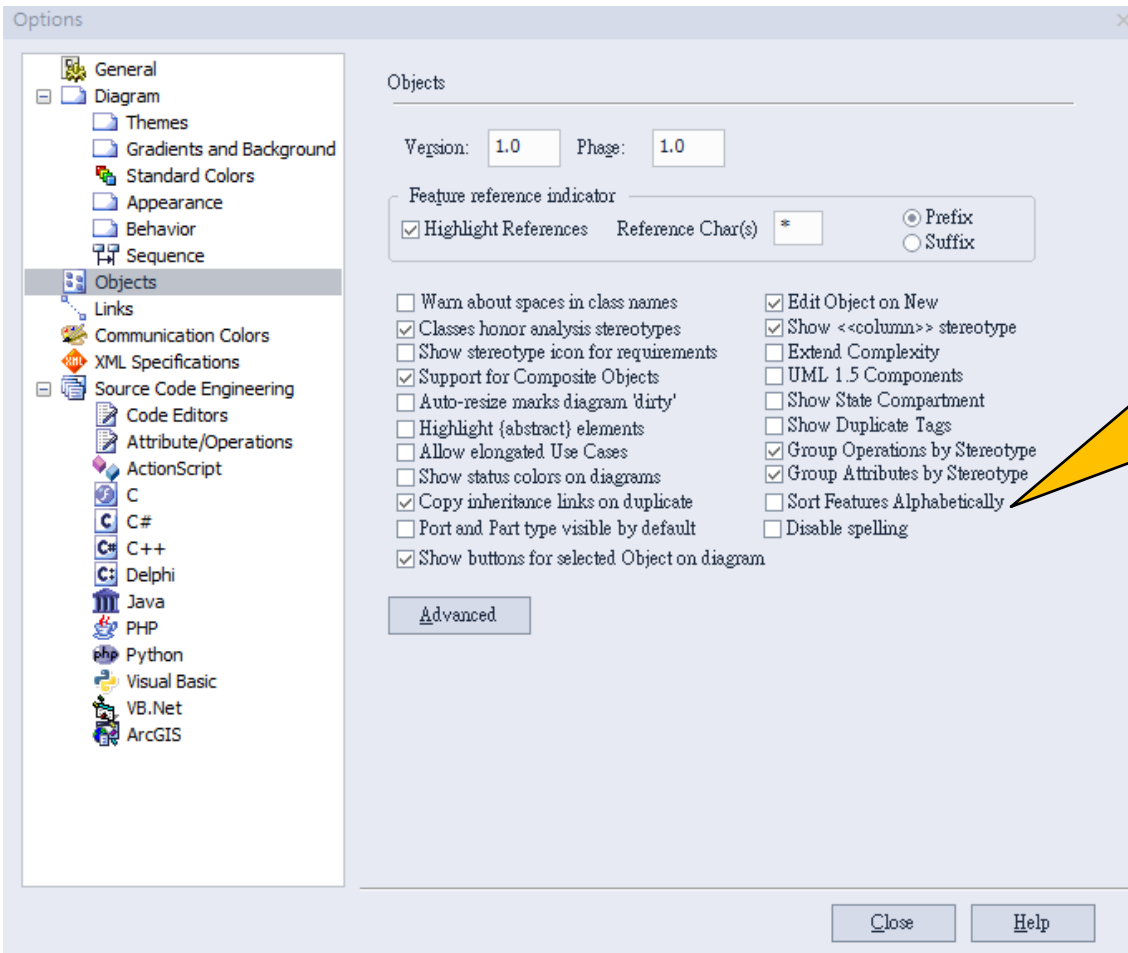
軟體工具	特性說明	UML DM上限	巢狀結構是否適用
Altova	UML採用各個資料型別，單獨繪製	50個資料型別	可克服，無型別結構也能判斷前後順序
EA	以XSD為單位，將資料型別彙整為單一UML	無上限	無法克服，必須給訂明確的資料型別

6. XSD輸出UML相關技術

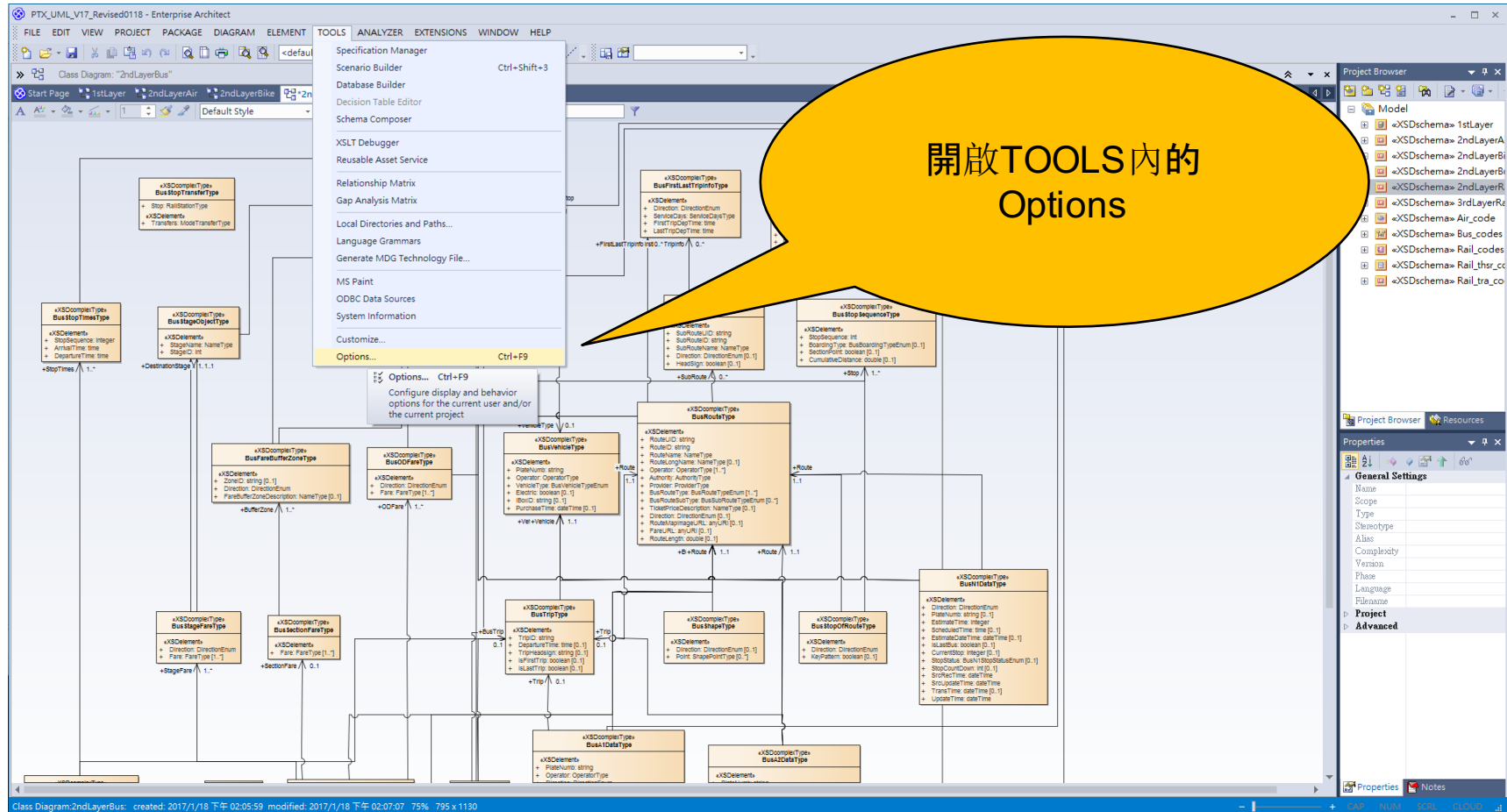
匯入XSD檔



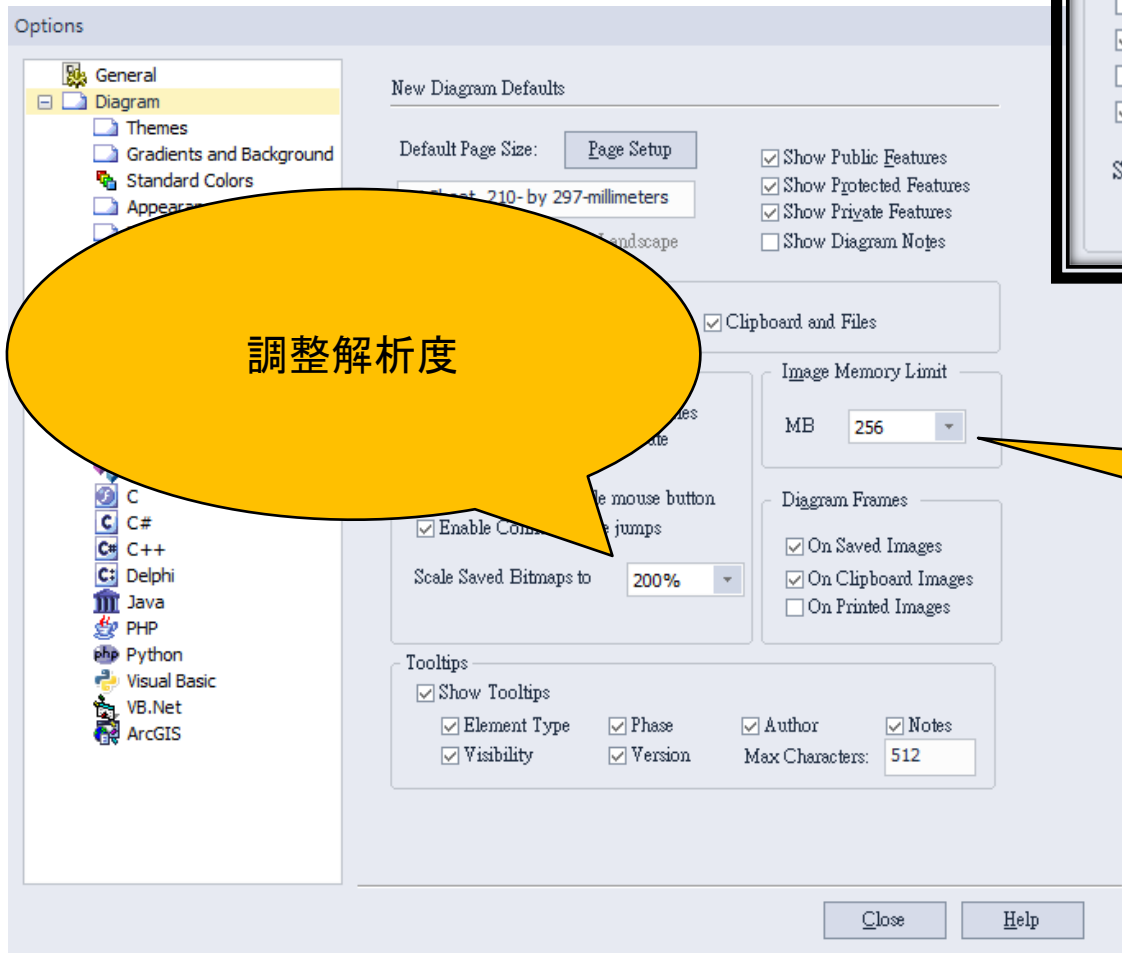
UML資料型別的排列次序調整



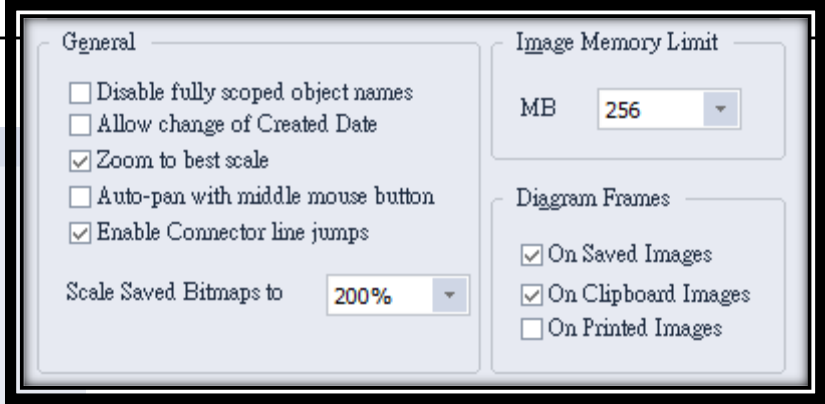
調高UML圖檔輸出解析度-方法1



調高UML圖檔輸出解析度-方法1



調整解析度



調整圖檔輸出容量大小

調高UML圖檔輸出解析度-方法2

The screenshot shows the Enterprise Architect interface with a UML Class Diagram titled '2ndLayerAir'. The diagram features several classes, each represented as a UML class box with its name and a list of attributes. The classes include:

- AirportType**: Attributes include Departure, Arrival, AirportID, Authority, AirportName, AirportIATA, AirportICAO, AirportLocation, AirportLatitude, AirportLongitude, AirportName, AirportAddress, and AirportPhone.
- AirlineType**: Attributes include AirlineID, AirlineName, AirlineIATA, AirlineICAO, AirlineMnemonic, AirlineAddress, AirlinePhone, and AirlineLocation.
- TransitType**: Attributes include TransitSequence, TransitBase, TransitAirportArrival, and TransitAirportDeparture.
- FlightType**: Attributes include FlightDate, FlightNumber, AirFlightType, ScheduleAirline, ActualArrivalTime, ActualDepartureTime, FlightStatus, FlightStatusC, Terminal, Gate, CoShare, IsCargo, ACType, and CapacityCounter.
- FlightScheduleType**: Attributes include FlightDate, FlightNumber, AirFlightType, ScheduleAirline, ActualArrivalTime, ActualDepartureTime, FlightStatus, FlightStatusC, Terminal, Gate, CoShare, IsCargo, ACType, and CapacityCounter.
- FlightScheduleType**: Attributes include FlightDate, FlightNumber, AirFlightType, ScheduleAirline, ActualArrivalTime, ActualDepartureTime, FlightStatus, FlightStatusC, Terminal, Gate, CoShare, IsCargo, ACType, and CapacityCounter.
- FlightScheduleType**: Attributes include FlightDate, FlightNumber, AirFlightType, ScheduleAirline, ActualArrivalTime, ActualDepartureTime, FlightStatus, FlightStatusC, Terminal, Gate, CoShare, IsCargo, ACType, and CapacityCounter.
- FlightScheduleType**: Attributes include FlightDate, FlightNumber, AirFlightType, ScheduleAirline, ActualArrivalTime, ActualDepartureTime, FlightStatus, FlightStatusC, Terminal, Gate, CoShare, IsCargo, ACType, and CapacityCounter.

A yellow callout bubble with a black border points to the diagram, containing the text: **Ctrl+A全選後, 直接複製到Word檔案**

創代 科技

用數據科學，提升生活品質