

**Common Schema for Address**

1. This discussion paper analyses how we may proceed to develop a Common Schema (or Common Schemas) for “address”. The analysis tries to cater for practical problems that will be faced by B/Ds in data exchange based on data held in existing systems. This paper aims to stimulate discussion among members of the Task Force designated to work on the “address” data element.
2. First, we describe how addresses are stored and used in some government systems. The description is based on findings collected from a few B/Ds only. Then we discuss the considerations for developing the Address Common Schema.

**Types of “Addresses” Used in Government Services**

3. Addresses are stored in departmental systems either as an unstructured chunk of text or as a structured data element comprising sub-components like floor, flat, building, street, etc. Below is our observation.

***Unstructured addresses***

4. Unstructured addresses are mainly for interpretation by human (e.g. by a teacher who needs to do home visits for selected students, or by the postman when he delivers a letter). Unstructured addresses are often used for printing letters and mailing labels, or for display on screen. Because of the layout of letters and mailing labels, a maximum number of lines of a certain maximum length are often assumed.
5. Unstructured address can be stored according to a fixed dimension (e.g. max. 5 lines of max. 35 characters each). Alternatively, it can be stored as a single string of text with “carriage return” embedded in the text. In either case, a certain dimension is assumed.

***Structured addresses***

6. There may be several reasons driving a B/D to store the address as a structured

data element:

- the B/D needs to do analysis on the address (e.g. deriving which district council constituency a voter belongs to from the street name and street number of his address);
  - the B/D lets user input address in a structured way to facilitate data entry, and adopts the same structure in its database;
  - the B/D follows a recommendation on address format previously made by ITSD in the early 90s
7. In most cases, B/Ds capture addresses and store them in a structure that can be understood by layman. However, some B/Ds have the operational need to store addresses in a highly structured way that needs to be handled by trained personnel.
8. Some B/Ds that have adopted a structured address may sometimes encounter addresses (e.g. the address of some village houses) that cannot be mapped to the address structure in use. One way to deal with such exception is to store the information as free text, which is maintained as part of the address structure.
9. Some components of the address structure (e.g. District) may not carry an accurate value. It is purely auxiliary information to help the user locate the physical location identified by the address.

### **Considerations for Developing a Common Schema for “Address”**

(Note : The issues (in bold italics) below are open issues with no conclusive solution yet. The analysis we have put down are just one of the many possible viewpoints, they serve to stimulate discussion only.)

***Issue #1 : What principles should we observe when we develop a Common Schema for address ? (these principles should apply in most cases)***

10. The principles may include :
- B/Ds should be able to map address data from the common data exchange format to the format used in their backend systems and vice versa (This affects how we design the Common Schema and what tools we build to

facilitate the data mapping). If B/Ds currently need manual intervention to convert address data received from other parties to their internal format, then they should not need “more” human resources than what are required now to convert data from the common data exchange format

- Using a common data exchange format will have impact on B/Ds that are using a different format in their backend systems. We should ensure that the overall impact to all B/Ds is minimal

***Issue #2 : What potential benefits could we expect a Common Schema to bring about ?***

11. The benefits may include :

- If we adopt a Common Schema with some structured address information, then we can build tools based on the address structure (we will also need a comprehensive address database to make the tools work). These tools can include address data entry validation, address translation from Chinese to English and vice versa, etc. Data entry validation tools can enhance the integrity of the address data. These tools can be shared by all B/Ds. *(It should be noted that the accuracy of the data entry validation depends on a number of criteria such as the validation rules we define, and the address database we use. For example, without a proper database, we have no way to tell whether the 30<sup>th</sup> floor exists in “Guardian House” or not.)*
- If B/Ds share a common address format, then we can build and share common tools like mapping unstructured addresses to structured address components for data analysis, one-stop service for change of address, etc.

***Issue #3 : Should we have one single address structure, or should we have different structures for different purposes ? If we were to use a single structure, what should be the purpose of this address structure ?***

12. From the perspective of facilitating data exchange, we should use one common address structure for data exchange. This address structure is suggested to cater for **the addresses of persons and organizations**, as well as to serve as **postal**

**addresses** (including P.O. Box type). In contrast, it may not be specific enough for identifying a property where works projects need to be carried out.

13. B/Ds are free to use a different structure in their backend systems. If a B/D is maintaining addresses in a highly structured way in its backend system, it should be easy for this B/D to convert its addresses to the exchange format. If this B/D only acts as an address sender, then there is minimal impact on this B/D (apart from having to do the conversion before sending information to other B/Ds). If this B/D needs to receive address from other parties, then it may negotiate with its counterparts to see if a proprietary format should be used, otherwise, it should convert the address from the exchange format to its internal highly structured format.
14. As an example, a structure for data exchange can be something like this : (this is a composite structure comprising a free text portion as well as a structured portion)

<ol style="list-style-type: none"><li>1. Free text (a few lines of free text, see notes below)</li><li>2. Structured Components (with an 'AddressabilityVerified' indicator to show whether the structured components hold genuinely addressable objects or not)<ol style="list-style-type: none"><li>2.1. Flat</li><li>2.2. Floor</li><li>2.3. Block</li><li>2.4. Building</li><li>2.5. Estate</li><li>2.6. Street number</li><li>2.7. Street name</li><li>2.8. District</li><li>2.9. Area</li></ol></li><li>3. Country</li></ol> <p>Notes :</p> <p>There can be different approaches in using the free text component and the structured component. This is an important area to be considered.</p> <p>Approach #1 is that the free text portion is mandatory and it always holds the complete address. The structured components are optional. They are simply information for analysis purpose. The structured components can be derived</p>
--

from the free text. And in the opposite direction, if the structured components carry enough details, a complete free text address can be generated from these components (the IFCG SO can provide tools to generate the free text address if necessary)

Approach #2 is that at least one of either the free text part or the structured part should exist and the free text part should be considered as a complete address if it is not empty. If the free text component is empty, then a complete address can be generated from the structured components and the country.

Approach #3 is to use the free text component only for keeping information that cannot be mapped onto the structure (or for keeping information that we do not bother to map onto the structure because there is no business need to do so), and this free text component should be used together with other structured components. In this case, the way to generate a “print format” of the address is to (1) print the free text, then (2) print the non-empty structured components, and (3) print the country.

The structure above is for illustration only, the exact structure needs to be aligned among B/Ds so that the structure can address a majority of B/Ds’ data analysis requirements

Approach #1 appears to be more pragmatic and we will use this approach in the examples below for illustration purpose.

#### **Example 1 – local village house**

```
<FreeText>
  <Line>House No. 3A, Lot No. 1357, Ching Shan Tsuen</Line>
  <Line>Tuen Mun, New Territories</Line>
</FreeText>
<StructuredComponents AddressabilityVerified = ‘Y’>
  <District>Tuen Mun</District>
  <Area>NT</Area>
</StructuredComponents>
<Country>
  <Code>HK</Code>
</Country>
```

#### **Example 2 – local office apartment in urban area**

```

<FreeText>
  <Line>Suite 1301-03, 13/F, World Trade Centre</Line>
  <Line>280 Gloucester Road</Line>
  <Line>Causeway Bay, Hong Kong</Line>
</FreeText>
<StructuredComponents AddressabilityVerified = 'N'>
  <Flat>Suite 1301-03</Flat>
  <Floor>13</Floor>
  <Building>World Trade Centre</Building>
  <StreetNumber >280</StreetNumber>
  <StreetName>Gloucester Road</StreetName>
  <District>Causeway Bay</District>
  <Area>HK</Area>
</StructuredComponents>
<Country>
  <Code>HK</Code>
</Country>

```

**Example 3 – foreign address in the UK**

```

<FreeText>
  <Line>Cardinal Tower</Line>
  <Line>12 Farringdon Road</Line>
  <Line>London EC1M 3HS</Line>
  <Line>UK</Line>
</FreeText>
<Country>
  <Code>UK</Code>
</Country>

```

15. The considerations for using this type of address structure are as follows :

- this structure can cater for addresses held in existing systems either in a structured format or in an unstructured format
- if the recipient adopts this address structure in its backend systems, then it can accept both structured information (if any) and a complete unstructured address from different parties with minimal information loss while having the flexibility to choose whether to convert unstructured addresses to a structured

format based on business need (e.g. whether analysis needs to be done on the address). In the long run, there will be benefit if B/Ds adopt this common address structure in their backend systems, because they will be able to share data validation tools, data conversion tools, etc.

- the IFCG SO can develop tools to facilitate conversion from the free text component to the structured components, but B/Ds (or some designated parties) should verify and correct the generated structured components, and to handle exceptional cases
  
- say for example B/Ds have agreed that the common address structure should have max. 5 lines of free text with max. 35 characters each, but a B/D is maintaining addresses as 4 lines of 50 characters each, then the B/D should convert its address to /from the common format (the IFCG SO can develop tools to facilitate the data conversion but B/Ds may have to accept some limitations, otherwise, B/Ds will have to correct the converted data manually. An example of possible limitation is that “King’s Road” might appear as “King’s” on line 2 and “Road” on line 3. If the address is for human interpretation, then there should be no major problem. However, the address will not look elegant)

***Issue #4 : Should we have a common structure for both local and foreign address ?***

16. Some information systems need to maintain both local and foreign addresses. Therefore, preferably, we should use a common structure for both local and foreign address.

***Issue #5 : How do we cater for data exchange with existing systems where the address is stored in a different structure / format ?***

17. Below is an analysis of what the sender and receiver needs to do in some possible scenarios (assuming that we have agreed on using 5\*35 for the free text component and the ESD change of address data entry address structure for the structured component)

		receiver				
		Adopt exchange structure in backend. No need to do analysis on data	Insist using only 5*35 free text in backend	Insist using only 4*50 free text in backend	Insist using its own structured address structure in backend	Insist using only ESD structure in backend
sender	Using 5*35 free text	Neither sender nor receiver needs to do any conversion	Neither sender nor receiver needs to do any conversion	Receiver Converts from 5 lines to 4 lines	Receiver converts from free text to its own structure	Receiver uses IFCG SO supplied tools to convert free text to structured components
	Using 4*50 free text	Sender converts from 4 lines to 5 lines	Sender converts from 4 lines to 5 lines	Sender and receiver should agree among themselves whether to use their proprietary format for exchanging data	Sender converts from 4 lines to 5 lines. Receiver converts from free text to its own structure	Sender converts from 4 lines to 5 lines. Receiver uses IFCG SO supplied tools to convert free text to structured components
	Using Own structured address structure	Sender converts from its own structure to free text and provides structured components if possible.	Sender converts from its own structure to free text and provides structured components if possible.	Sender converts from its own structure to free text and provides structured components if possible. Receiver converts from 5 lines to 4 lines	Sender converts from its own structure to free text and provides structured components if possible. Receiver converts from free text to its own structure.	Sender converts from its own structure to free text and provides structured components if possible. Receiver uses IFCG SO supplied tools to convert free text to structured components
	Using ESD structure	Sender generates free text using the IFCG SO supplied tools and supplies the structured components.	Sender generates free text using the IFCG SO supplied tools and supplies the structured components.	Sender generates free text using the IFCG SO supplied tools and supplies the structured components. Receiver converts from 5 lines to 4 lines	Sender generates free text using the IFCG SO supplied tools and supplies the structured components. Receiver converts from free text to its own structure.	Sender generates free text using the IFCG SO supplied tools and supplies the structured components.



***Issue #6 : If we were to use a structure comprising both an unstructured component (the free text part) plus some structured components ,how structured should the structured part be ? like the address structure used by the Rating and Valuation Department (RVD) or like how the public input addresses through the Electronic Service Delivery (ESD) system ?***

18. This is debatable. Below are the pros and cons of using RVD's format :

Pros:

- n RVD's format is highly structured, therefore it can cater for most types of analysis needed by B/Ds
- n as a number of B/Ds are already receiving addresses from RVD, they have already prepared mapping tables to facilitate conversion from RVD's address format
- n as the address information maintained by RVD are comprehensive<sup>1</sup> and error free, the IFCG SO can develop address data entry validation, address data conversion, and address translation (Chinese to English and vice versa) tools by making use of RVD's address database (if RVD allows us to do so). If B/Ds adopt RVD's structure, then it may be easier to develop the tools. Alternatively, ESDLife also has an address database based on which we can develop validation / conversion / translation tools.

Cons:

- n RVD's address format is highly structured, we cannot expect an untrained layman to be able to input an address according to RVD's format. If we were to adopt RVD's format, we need to (1) develop user friendly tools to facilitate the general public to input their address and then the tool should intelligently convert the address to RVD's format; (2) establish a central pool of trained personnel to help B/Ds verify and correct batch address conversion as necessary
- n Since RVD's address structure is different from most B/Ds' address structure, almost all B/Ds that need a structured format for analysis will need to do data conversion and software program modification

---

<sup>1</sup> RVD is maintaining the address of all ratable objects (which include all properties) at floor level (i.e. at 3-D level). This does not include the address of (1) public housing where the information is maintained by the Housing Department, and (2) government properties where the information is maintained by the Government Property Agency.

- n Since RVD's format is highly structured, although it provides greater flexibility in data analysis, the algorithm to manipulate the address may get more complicated

19. The choice of the structure should depend on impact analysis across all B/Ds. Statistics may also help the Task Force make its decision, these statistics can be collected via surveys if deemed necessary.

***Issue #7 : How should address in Chinese be handled ?***

20. RVD's or ESDLife's address database can help us develop tools to translate the structured address components from English to Chinese and vice versa. With regard to data exchange, one possible way is to do nothing to the free-text part and exchange the address as is (it will be the recipient's responsibility to translate the free text address manually), but perform automatic translation for the structured components as necessary, as far as possible (it will be the recipient's responsibility to verify and correct the translation).

***Issue #8 : Should we cater for postcode which appears in mainland and some foreign addresses ?***

21. Probably not, there is no postcode in local addresses. For foreign address, B/Ds seldom need to do analysis on the address (apart from the Country information), hence foreign addresses are often maintained as unstructured information which can easily embed the postcode / zipcode.

***Issue #9 : What are the industry standards we can make reference to ?***

22. Please refer to Appendix A

***Issue #10 : What are the address formats currently in use by B/Ds that we can make reference to ?***

23. Please refer to Appendix B

***Issue #11 : What are the address formats adopted by other economies that we can make reference to ?***

24. Please refer to Appendix C

**Way Forward**

25. The Task Force is recommended to proceed as follows :

- discuss the issues raised in this paper and agree on the higher level issues first, then drill down to the issues at the lower level, e.g.
  - n what other principles should we observe ?
  - n should we have one address structure or multiple ?
  - n should we have a hybrid structure comprising both free text and structured components ?
  - n how to use the free text component and the structured component ?
  - n how should we structure the structured component ?
  - n what is the format (e.g. field length) of individual components ?
- conduct survey if necessary to collect statistics to support decision making
- align the address structure and format (including field length), plus any necessary controlled vocabularies such as Street Name and Building Name
- discuss details of tools required, including conversion rules and data validation rules. These tools may include :
  - n address data entry validation tools;
  - n convert from free text to structured components and vice versa;
  - n translate from Chinese to English and vice versa;
  - n convert free text address from one dimension to another;
  - n abbreviate addresses to fit a certain dimension (e.g. abbreviate 'Street' as 'St'); etc.
- align exception handling procedures for various types of data exchange / conversion scenarios, e.g.

- n on the fly data exchange / conversion for transactions which normally do not involve manual intervention
- n batch data conversion
  
- publish the candidate Common Schema and announce when the tools will be available
  
- request B/Ds (through their Common Schema Liaison Officers) for comment and request them to conduct impact analysis on how much effort they will take to adopt this schema for exchanging addresses
  
- agree on the Common Schema
  
- develop and test the tools
  
- draw plans to facilitate data conversion as necessary
  
- B/Ds to prepare their adoption plan

**Information Technology Services Department**  
**October 2003**

## Appendix A - Industry Standards Related to Address

		URL
Address	UN/EDIFACT	<a href="http://www.unece.org/trade/untdid/d00a/trsd/trsdadr.htm">http://www.unece.org/trade/untdid/d00a/trsd/trsdadr.htm</a>
UN/PROLST	UN/EDIFACT	Draft copy (issued in May 2001) - A global Address for direct mail marketing and e-commerce <a href="http://xml.coverpages.org/IAEC-V10-5-31-01.pdf">http://xml.coverpages.org/IAEC-V10-5-31-01.pdf</a>
ISO 11180:1993 Standard for Postal Addressing	ISO	Provides mailing address structure requirements (not read yet, need to pay for the standard)  A reference example can be found in: <a href="http://www.jtc1sc32.org/sc32/jtc1sc32.nsf/Attachments/B17566EF26D3A77788256A530061148F/\$FILE/32N0648T.PDF">http://www.jtc1sc32.org/sc32/jtc1sc32.nsf/Attachments/B17566EF26D3A77788256A530061148F/\$FILE/32N0648T.PDF</a>  Page 4 of the example shows the Mailing Address Group which comprises: <ul style="list-style-type: none"> <li>- Mailing Address Building Number</li> <li>- Mailing Address Street Name</li> <li>- Mailing Address State Code</li> <li>- Mailing Address Country Name</li> </ul>
xNAL (includes xNL: eXtensible Name Language and xAL- eXtensible Address Language)	OASIS Customer Information Quality (CIQ) Technical Committee	<a href="http://www.oasis-open.org/committees/ciq/download.shtml">http://www.oasis-open.org/committees/ciq/download.shtml</a>  select <a href="#">Download the HTML formatted documents for CIQ SCHEMAS only</a>
British Standard 7666	BSI IST/36	Part 3 Spatial datasets for geographic referencing: Specification for Addresses <a href="http://www.govtalk.gov.uk/documents/IA_APD_1_1.doc">http://www.govtalk.gov.uk/documents/IA_APD_1_1.doc</a>
CEN TC/331 N231	European Committee for	Draft issued for Postal Services – Address data bases – Part 1 – Components of Postal Addresses

	Standardization	<a href="http://www.nen.nl/cen331/n284.pdf">http://www.nen.nl/cen331/n284.pdf</a>
HR-XML	HR-XML Consortium	The HR-XML Consortium will develop a common HR vocabulary and model, and will develop schemas for common HR objects used in Recruiting and Staffing, Benefits Enrollment, Payroll, etc. It is working on the many attributes of the <b>Person</b> object which must be handled consistently by different HR processes. The first two attributes defined are <b>PersonName</b> and <b>PostalAddress</b> <a href="http://xml.coverpages.org/HRXML-PostalAddressv12.pdf">http://xml.coverpages.org/HRXML-PostalAddressv12.pdf</a>
Address Data Content Standard	Federal Geographic Data Committee (FGDC)	<a href="http://www.census.gov/geo/www/standards/scdd/AddressStandardV2_April%2017_2003.htm">http://www.census.gov/geo/www/standards/scdd/AddressStandardV2_April%2017_2003.htm</a>  FGDC-STD-001-1998 Content Standard for Digital Geospatial Metadata ( <a href="http://www.fgdc.gov/standards/documents/standards/metadata/v2_0698.pdf">http://www.fgdc.gov/standards/documents/standards/metadata/v2_0698.pdf</a> ). It defines addresses for contacts (persons and organizations) with metadata such as Address type + Address (0..n occurrences) + City + State_or_province + Postal_Code + Country (optional)
International Address Standard UPU S42-1	Universal Postal Union (UPU)	Not read yet, need to pay for the standard, an overview is available at : <a href="http://xml.coverpages.org/Lubenow-UPUS42.html">http://xml.coverpages.org/Lubenow-UPUS42.html</a>

## Appendix B - Address Formats In use by B/Ds

B.1 Building Address defined in the draft Version 0.3 report of the Data Alignment Measure (DAM) project (this address is at building-level, i.e. it is a 2-dimensional address)

Table B.1.1

Data Item	Description	Format	Mandatory
Record ID	Identifier of the concerned Building Geographical Info record	N(9)	Y
Address Source	Provider of this address, where R = Recorded address provided by RVD S = Surveyed address provided by LandsD When Address Source = 'S', only the Bldg No Num, Bldg No Alpha, Bldg No Ext and St Loc ID data items are applicable and all the other non-mandatory data items will be left as Null.	X(1)	Y
Address ID	Identifier of address assigned by Data Owner. When Address Source = 'R', it references to the concerned Block ID as in RVD's PMS	N(9)	Y
Eng Bldg Name 1	Line 1 of building name in English e.g. 'KAM FAI BUILDING', 'BLOCK 40'	X(35)	
Eng Bldg Name 2	Line 2 of building name in English	X(35)	
Eng Bldg Name 3	Line 3 of building name in English	X(35)	
Chi Bldg Name 1	Line 1 of building name in Chinese e.g. '金輝樓', '第 40 座'	CX(14)	
Chi Bldg Name 2	Line 2 of building name in Chinese	CX(14)	
Chi Bldg Name 3	Line 3 of building name in Chinese	CX(14)	
Eng Block Desc	Block description in English e.g. 'BLDG', 'APT', 'HSE', 'BLK'	X(35)	
Chi Block Desc	Block description in Chinese e.g. '大廈', '洋房', '座'	CX(14)	
Block No Num	Numeric part of block number e.g. '1' in 'TOWER 1'	N(5)	
Block No Alpha	Alphabetic part of block number	X(10)	

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
	e.g. 'A' in 'BLK A'		
Block No Alpha Prec Ind	Determine ordering of numeric and alphabetical parts of block number when both parts are not Null, where Y = Block No Alpha precedes Block No Num N = Block No Num precedes Block No Alpha	X(1)	
Block Desc Prec Ind	Determine ordering of block description and block number in full address when both are not Null, where Y = block description precedes block number (e.g. in the cases of 'BLK A' and 'TOWER 1') N = block number precedes block description (e.g. in the cases of 'NORTH BLK' and 'WEST TOWER')	X(1)	
Bldg No Num	Numeric part of building number e.g. '18' in '18A/20A', and '1' in '1B/1C'	N(4)	
Bldg No Alpha	Alphabetical part of building number e.g. 'A' in '18A/20A', and 'B' in '1B/1C'	X(2)	
Bldg No Ext	Extension of building number e.g. '/20A' in '18A/20A', and '/1C' in '1B/1C'	X(8)	
St Loc ID	Identifier of the concerned street-location ( <i>see table B.1.2</i> )	N(9)	
Unofficial Address Code	Indicate that the address unofficial, where 1 = Building number to be bracketed/No building number 2 = Building number and street name to be bracketed 3 = Dummy address i.e. building number within the range occupied by the building and recorded for search purpose, but there is no tenement/unit in the building using this building number in full address	X(1)	
Bldg No Confirm Code	Indicate whether the assigned building number is confirmed, where C = Confirmed building number P = Provisional building number	X(1)	



<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
	<p>A = Not yet allocated but action to be taken</p> <p>B = Use existing old number but action to be taken</p> <p>X = Not applicable</p> <p>When this code is C or P, at least one of the three Bldg No fields must not be Null.</p>		
Eng Full Addr 1	<p>Line 1 of the full postal address in English. Concerned lot number(s) is generated as part of the full address (line 1-5) for unofficial addresses, where unofficial addresses refer to those with non-blank Unofficial Address Code.</p> <p>The full address is mandatory for records with Address Source = 'R' and Unofficial Address Code = 1, 2 or Null.</p>	X(35)	
Eng Full Addr 2	Line 2 of the full postal address in English	X(35)	
Eng Full Addr 3	Line 3 of the full postal address in English	X(35)	
Eng Full Addr 4	Line 4 of the full postal address in English	X(35)	
Eng Full Addr 5	Line 5 of the full postal address in English	X(35)	
Eng Loc Addr 1	<p>Line 1 of location name part for postal addresses with location name.</p> <p>Location name part (line 1-3) is applicable normally for buildings in the New Territories only, and is appended to the last non-blank line of English full address (line 1-5) to form the full postal address.</p>	X(35)	
Eng Loc Addr 2	Line 2 of location name part for postal addresses with location name	X(35)	
Eng Loc Addr 3	Line 3 of location name part for postal addresses with location name	X(35)	
Eng Addr Verify Code	<p>Indicate whether the English full postal address is verified, where</p> <p>N = Not yet verified</p> <p>V = Verified, and the full address is generated by the structural address</p> <p>S = Verified, but the full address is manually input by user</p>	X(1)	

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
Chi Full Addr 1	Line 1 of the full postal address in Chinese. Concerned lot number(s) is generated as part of the full address (line 1-5) for unofficial addresses, where unofficial addresses refer to those with non-blank Unofficial Address Code. The full address is mandatory for records with Address Source = 'R' and Unofficial Address Code = 1, 2 or Null.	CX(14)	
Chi Full Addr 2	Line 2 of the full postal address in Chinese	CX(14)	
Chi Full Addr 3	Line 3 of the full postal address in Chinese	CX(14)	
Chi Full Addr 4	Line 4 of the full postal address in Chinese	CX(14)	
Chi Full Addr 5	Line 5 of the full address in Chinese	CX(14)	
Chi Addr Verify Code	Indicate whether the Chinese full postal address is verified, where N = Not yet verified V = Verified, and the full address is generated by the structural address S = Verified, but the full address is manually input by user	X(1)	

Table B.1.2

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
St Loc ID	Unique identifier of a street name-location name combination	N(9)	Y
Street ID	Identifier of the concerned street name e.g. references to 'CASTLE PEAK RD' for 'CASTLE PEAK RD-TSUEN WAN' e.g. references to 'TAI MONG TSAI RD' for 'TAI MONG TSAI RD, TSAM CHUK WAN, SAI KUNG' (see table B.1.3)	N(9)	Y
Location ID 1	Identifier of the first concerned location name, if any. Normally not available for streets in Hong	N(9)	

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
	Kong Island and Kowloon. e.g. references to 'TSUEN WAN' for 'CASTLE PEAK RD-TSUEN WAN' e.g. references to 'TSAM CHUK WAN' for 'TAI MONG TSAI RD, TSAM CHUK WAN, SAI KUNG' <i>(see table B.1.4)</i>		
Location ID 2	Identifier of the second concerned location name, if any. Normally not available for streets in Hong Kong Island and Kowloon. e.g. references to 'SAI KUNG' for 'TAI MONG TSAI RD, TSAM CHUK WAN, SAI KUNG' <i>(see table B.1.4)</i>	N(9)	
Location ID 3	Identifier of the third concerned location name, if any. Normally not available for streets in Hong Kong Island and Kowloon. <i>(see table B.1.4)</i>	N(9)	
Location Nature	Nature of the street-location combination, where 1 = Street-location is equal to gazetted street name 2 = Street name part alone is equal to gazetted street name 3 = Village 4 = Others	X(1)	Y
Eng Full Name	Full street-location name in English generated according to the concerned street name and location name(s) e.g. 'CASTLE PEAK RD-TSUEN WAN'	X(57)	Y
Chi Full Name	Full street-location name in Chinese generated according to the concerned street name and location name(s) e.g. '青山公路－荃灣段'	CX(24)	Y

Table B.1.3

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
Street ID	Unique identifier of a street name	N(9)	Y
Eng Name	Street name in English e.g. 'CONNAUGHT' in 'CONNAUGHT RD C'	X(35)	Y
Eng Type	Street type in English e.g. 'RD' in 'CONNAUGHT RD C'	X(10)	
Eng Direction	Street direction in English e.g. 'C' in 'CONNAUGHT RD C'	X(10)	
Eng Type Prec Ind	Determine ordering of type and direction parts of the English street name when both parts are not Null, where Y = Eng Type precedes Eng Direction N = Eng Direction precedes Eng Type (Normally 'Y' for all records)	X(1)	
Chi Name	Street name in Chinese e.g. '干諾' in '干諾道中'	CX(14)	Y
Chi Type	Street type in Chinese e.g. '道' in '干諾道中'	CX(5)	
Chi Direction	Street direction in Chinese e.g. '中' in '干諾道中'	CX(5)	
Chi Type Prec Ind	Determine ordering of type and direction parts of the Chinese street name when both parts are not Null, where Y = Chi Type precedes Chi Direction N = Chi Direction precedes Chi Type e.g. 'Y' for '干諾道中', and 'N' for '西洋菜南街'	X(1)	

Table B.1.4

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
Location ID	Unique identifier of a location name	N(9)	Y
Eng Name	Location name in English	X(35)	Y

<b>Data Item</b>	<b>Description</b>	<b>Format</b>	<b>Mandatory</b>
Chi Name	Location name in Chinese	CX(14)	Y

B.2 Address format used to capture user input in ESD Change of Address transaction

Field Description	length
Room/Flat	20
Floor	20
Block	20
Building/Phase	50
Estate/ Village No & Name	50
Street No & Name	50
Area (Select from a predefined list)	
District (Select from a predefined list)	
Other Information (e.g. Overseas Address, PO Box No)	3 x 500

B.3 Existing databases in B/Ds

B.3.1 Property Master System (PMS) of Rating and Valuation Department

B.3.2 Register of Quarter System (RQS) of Census & Statistics Department

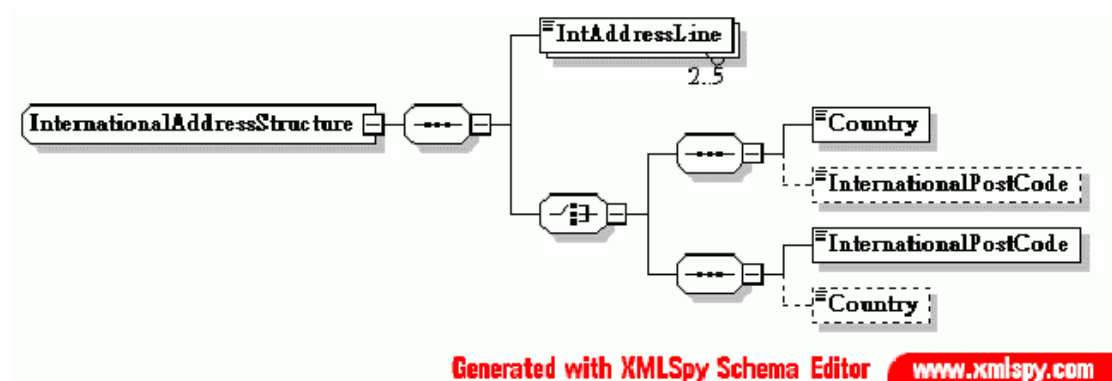
## Appendix C – Addresses Used in Other Economies

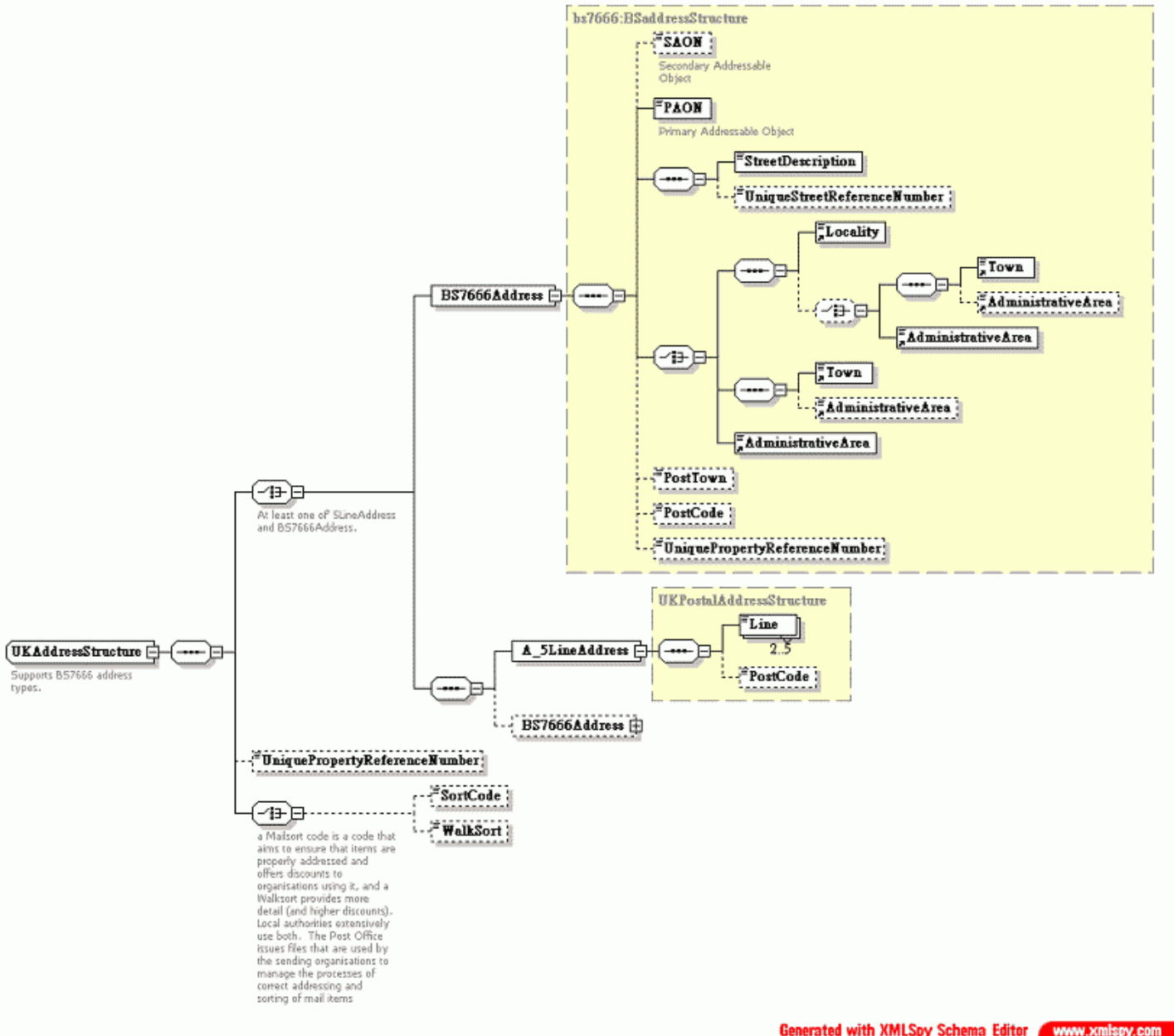
### C.1 UK address as stipulated in “UK Online – Information Architecture – Address and Personal Details Fragment”

[http://www.govtalk.gov.uk/documents/IA\\_APD\\_1\\_1.doc](http://www.govtalk.gov.uk/documents/IA_APD_1_1.doc)

The following examples on International Address and UK Address are compiled according to data specifications published on the UK Govtalk homepage (Schemas & Standards - Agreed Schemas - Address & Personal Details)

[http://www.govtalk.gov.uk/schemasstandards/agreedschema\\_schema.asp?schemaid=182](http://www.govtalk.gov.uk/schemasstandards/agreedschema_schema.asp?schemaid=182)





## C.2 US

The US Postal Service (USPS) maintains a standard and technical guideline for mailing type address – Postal Addressing Standard.

<http://pe.usps.gov/cpim/ftp/pubs/Pub28/pub28.pdf>

The Address data elements for Postal service are described on page 42 of the guideline:

- Street number
- Predirectional,

- Street name
- Street Suffix,
- Postdirectional,
- Secondary Unit Indicator
- Secondary Number
- Company Name
- PO Box Number
- City
- State
- ZIP Code
- ZIP+4 Code
- Carrier Route Code...

In addition, they had provided a ZIP+4 look-up system on the web, to avail the public of a simple means of obtaining ZIP+4 Code information.