

## 2. A Guide to Translating Native Data to SDP

The Guide to Translating Native Data to SDP describes the Guidance documents and tools, as well as a straight-forward process to map native schedule data to the SDP data concepts and formats.

This manual is written for a technical reader who has XML, programming or database experience, and who has knowledge of the procedures used to translate one format to another.

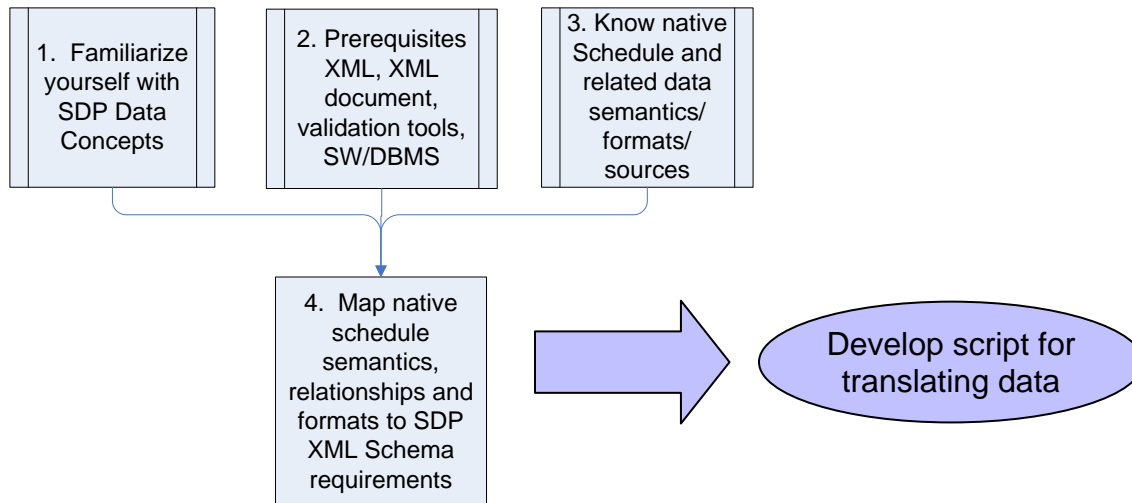
### 2.1. Templates, Quick Start and User Requirements Guidance

Several Guidance documents and tools were developed to support the translation of native data to SDP. The most useful documents for the “uninitiated” SDP translator are the following:

- ***SDP Quick Start Guide*** ([http://www.consystem.com/tsdea/rstwg/SDP\\_QS\\_web.htm](http://www.consystem.com/tsdea/rstwg/SDP_QS_web.htm)). This web document provides a high level overview of the requirements for translating native data to the mandatory elements of the SDP XML Document. Each mandatory element is described, along with its format and an example of its usage. In addition, conventions and code values are included in the document.
- ***SDP Template Spreadsheet***. This MS Excel spreadsheet contains several pages that describe the entire SDP XML Schema (mandatory and optional). The template includes a comprehensive data concept glossary, as well as a description of all the codes and their values. The template will be described in Section 2.3 below.
- ***SDP Guidance Documents***. As described in Section 1.2, the Guidance Documents consist of four parts (Ed. Note: part 4 will be implemented when the TSDEA is operating). Part 2: User Requirements Manual may be used as a comprehensive reference.
  - ***SDP Guidance Document Part 2 User Requirements***: Part 2 may be used as a reference manual to view detailed examples of how different organizations map their native data to the SDP. In addition, the context and requirements related to the SDP data concepts are explained in this document. Chapters 4 through 10 describe how to map native data to the SDP data concepts by each branch in the SDP XML Schema. Appendix A: Considerations for Rail Transit discusses how to apply the SDP concepts to rail concepts, particularly those that differ from bus transit. A glossary is included in Appendix B of this document.

### 2.2. A Simple Process for Translating Native to SDP

The process described in this section is similar to one used by many software developers to map one data format to another. As depicted in Figure 1, there are four major steps needed to map native data to SDP.



**Figure 1: SDP Translation Process**

Steps to build a native data to SDP XML document translator:

1. Familiarize yourself with SDP data concepts, codes and schema elements. The *SDP Quick Start* web document will provide the basic overview. To see more detail on non-mandatory element, see the *User Requirements Manual (Part 2 Guidance document)*.
2. Check your skills and experience with key technical areas to develop script and code:
  - Do you understand XML and SDP Schema organization?
  - Do you have experience programming or developing modules for DBMS?
  - Do you understand how to validate SDP XML documents?
3. Review your understanding of the semantics, relationships and formats of your organization’s native schedule and related data set(s) that cover the SDP data requirements.
4. Map Native Data to SDP concept

Based on your knowledge of your schedule and related data meanings, identify the native table and field that maps to each SDP mandatory and optional element descriptions.

Using the *SDP Template Spreadsheet* (to record your findings) or other tool:

- Map native types, flags and codes to SDP codes
  - Map and document the obvious native data (and transformations) to SDP
  - Identify missing identifiers; strategize on approaches for completing them while ensuring identifier uniqueness constraints.
  - Identify missing mandatory data fields; strategize on approaches for completing them.
  - Identify native data that is not represented (and should be) in the SDP.
5. Develop scripts and modules to translate codes, data types and extract data from native sources, and load the data into the SDP format.

- You can use the Data Mapper software (csv2xml) application to generate the XML Document from comma-separate value (csv) files. For information on applying this method, see Chapter 5.
- You can develop embedded procedures in your database to generate an XML Document. You can reuse procedures from any of the applications developed for this project. See Chapter 5 for user manuals and design documents.

### 2.3. Using the SDP Template Spreadsheet

The SDP Template Spreadsheet provides a useful “cheat sheet” for developing translations between native data and SDP data concepts. The spreadsheet includes 9 worksheets:

- Notes on How to Use the Template: includes notes on how to use the worksheets.
- CodeList: includes the SDP element name, description, XML code fragment and guidance on how to use the code values. Guidance includes how to classify special real-world objects. For example, a subway is a “heavy rail” – HR mode.
- Glossary: defines the data concepts and branch where it is placed.
- Each SDP Branch has its own worksheet. Each worksheet is organized similarly. The worksheet contains five Columns:
  - Required: describes whether the high level element is mandatory [M] or optional [O]
  - Element Name: lists the SDP data concept (high level element)
  - Type: lists the XML type or referenced type. This field also includes referential integrity rules such as unique, identifier (identifying key). A nested element will list the name of another Element Name (either on the same worksheet or on another one).
  - Questions to Ask: describes the guidance related to a high level, embedded or nested element.
  - Native Format: a placeholder for recording notes on which native data match the SDP element.
- AgencyRegistration (branch): describes the elements, nested and embedded elements in the AgencyRegistration branch of the SDP XML Schema. This page also includes the SDP Document attribute group for the ScheduleVersion data concept.
- Service (branch): describes the elements, nested and embedded elements in the Service branch of the SDP XML Schema.
- Transit Network (branch): describes the elements, nested and embedded elements in the Service branch of the SDP XML Schema.
- Transit Gazetteer (branch): describes the elements, nested and embedded elements in the Transit Gazetteer branch of the SDP XML Schema.
- Transit Facility (branch): describes the elements, nested and embedded elements in the Transit Facility branch of the SDP XML Schema.
- AddressStructure: describes the data structure for an address. The AddressStructure is nested in the Agency and TransitStop elements.

The Excel Template may be found at the following link:

[http://www.consystec.com/tsdea/rstwg/documents/SDPxml\\_template\\_v1\\_0.xls](http://www.consystec.com/tsdea/rstwg/documents/SDPxml_template_v1_0.xls)