West Valley Information Sharing Enterprise
Data Dictionary Product Summary

Suzanne W. Dietrich
Arizona State University

The goal of the data dictionary component of the project was to gather a description of the data for four out of the ten agencies participating in the West Valley Information Sharing Enterprise (WISE). These data descriptions lay the groundwork for developing future data exchanges. To provide added value to the project, all agencies were contacted to provide their data description or metadata. The process of gathering the data dictionary included various email and voice communications with WISE agency representatives and vendor contacts. Some telephone conferences were held with vendor contacts and agency representatives to discuss the requirements for the data descriptions. Most metadata were described using an XML representation, which was not compliant with the Global Justice XML Data Model (GJXDM), or in an Excel worksheet. The ASU team took this raw data and transformed the metadata descriptions into a collection of Excel spreadsheets that document the data maintained by each agency.

The accomplishment of the data dictionary was greatly facilitated by the cooperation and effort of individuals dedicated to the project from each of the participating agencies in the west valley. These representatives were responsible for many simultaneous projects at their own agency. It should also be noted that this project had very short timelines, which impacted requests for each of the participating agencies. This document summarizes the deliverable for the data dictionary component of the project, and reports on the challenges and recommendations for data exchange.

Table 1 provides a summary for each agency of the record management system used, whether a nondisclosure agreement was required, and the degree of information provided for the data description. The Spillman RMS is used by four agencies: Avondale, Buckeye, Surprise, and Tolleson. Although they all use Spillman, the modules used and the specific data in the corresponding code tables differ except for Buckeye and Tolleson, which share the same server and therefore, have the same modules and code tables. El Mirage's vendor, Cisco, provided only a sample of the data that they maintain in XML, which used tags located in a data dictionary for the Criminal Justice Information System (CJIS) Electronic Fingerprint Transmission Specification. Goodyear's vendor, New World Systems, provided a metadata description that emphasized person data. HTE would not provide their data directly. However, Peoria facilitated the gathering of the data description from their vendor. Wickenburg was not an active participant in the WISE committee. Youngtown's CAD data is collected through El Mirage, and its RMS vendor, Xpediter, did not have a description of their data but provided a copy of their product so that we could provide some representation of the data.

Table 1 Record Management System and Data Description by Agency

Agency	Record Management System (RMS)	Nondisclosure Agreement	Data Description
Avondale	Spillman	Yes	Complete
Buckeye	Spillman	Yes	Complete
El Mirage	Cisco	No	Sample
Glendale	CHIPS written by Glendale in Oracle DB	No	Complete
Goodyear	New World Systems	No	Person Data
Peoria	HTE	No	Complete
Surprise	Spillman	Yes	Complete
Tolleson	Spillman	Yes	Complete
Wickenburg	None	NA	NA
Youngtown	Cisco for CAD via El Mirage & Xpediter for RMS	No	Screen Shots

The deliverable for the data dictionary component of the project is a CD that includes detailed Excel spreadsheets for each agency providing metadata that documents the description of the data and the data in the code tables. The deliverable media, which was submitted to the sponsor of the project, contains the following files for the various agencies:

- AvondaleMetadata.xls, AvondaleCodeTables.xls
- BuckeyeMetadata.xls, BuckeyeCodeTables.xls
- ElMirageMetadataCodeTables.xls
- GlendaleMetadata.xls, GlendaleCodeTables.xls
- GoodyearMetadataCodeTables.xls
- PeoriaMetadata.xls, PeoriaCodeTables.xls
- SurpriseMetadata.xls, SurpriseCodeTables.xls
- TollesonMetadata.xls, TollesonCodeTables.xls
- YoungtownXpediterScreenshots.doc

In addition, there is an executive summary Excel spreadsheet (ExecutiveSummaryWorkbook.xls) that provides a high-level overview of the data maintained by the agencies. One worksheet within the executive summary spreadsheet compares the agencies information with respect to person data, which was identified by the needs assessment as the primary type of information that the agencies needed to share. This overview along with the detailed metadata provide the foundation of knowledge necessary for future data sharing efforts.

Challenges and Recommendations

The gathering of the data dictionary was a challenging task. Vendors of the various record management systems (RMS) were reluctant or unwilling to provide the metadata of their product because it discloses the design of their system:

> "We do not normally provide our proprietary record system database schema or tables to third parties. The potential damage to the business should the information end up in the wrong hands is extreme. This could either be competitors, or "hackers" which could then pose a threat to all our customers. Because of the above reasons, we politely decline to provide the information you requested."

Some vendors required a nondisclosure agreement to be signed, such as Spillman. The detailed data description for the Spillman agencies can only be disclosed to members of the WISE committee. As an alternative to disclosing their system design, vendors were asked whether they

could provide a GJXDM description of their data. To the best of our knowledge, none of the vendors have a complete mapping to the Justice XML:

> *"We have mapped our data to the GJXDM for a few specific projects but I do not have a mapping document/specification for the entire database. We've chosen to do this on an as needed basis because the effort to map everything would be quite time consuming."*

Most vendors were also reluctant to provide a metadata description because of the time required to do so:

> *"I don't want to appear difficult, but at the same time, I'm sure that you can understand that we are hesitant, as a private company, to spend too many resources on unfunded projects. We walk a fine line between trying to provide all of the services that our customers (and their other vendors) want, and attempting to be a profitable undertaking."*

One vendor provided only a sample of the data that their RMS maintains.

Based on the above challenges with vendors, a recommendation for proceeding with future data exchange would be to explore the integration of data sharing solutions provided by various vendors (e.g., Spillman's Insight and HTE's Police-to-Police [P2P]) and agencies (e.g., the Arizona Criminal Justice Commission [ACJC] and the Integrated Criminal Justice Information System [ICJIS]). The focus of integrating data sharing solutions will emphasize the use of standards and will provide a necessary level of abstraction above the vendor's proprietary data solution.

All agencies within the criminal justice system are being challenged to increase the sharing of information across jurisdictions and to share that information in a more effective and efficient manner. Any solution obviously relies on the cooperation of the vendors of the record management systems. Since most RMS vendors probably have a data sharing solution for person data, which was identified as the primary type of information that the agencies needed to share by the needs assessment, a recommendation for a first step would be to develop a first tier of data sharing for person data. The CJIS has already identified standards for representing most of the person data, which appears to be consistent with the GJXDM. A simple strategy that shares person data based on the known standards will provide an initial step towards data sharing for the west valley agencies.