Why Do We Need A Standard Way To Synchronize Human Services Data?

Introduction:

Human services agencies share client records to perform both day-to-day program as well as disaster relief operations.[1] They do this by sending and receiving series of data files formatted to one of many human services data standards. Although a great deal of effort has gone into creating and using these client record file formats, the formats are limited to only holding client information, with no context of what exactly should be done with the data (i.e. should the sent records be added or removed from the receiving system?).

To elaborate, a given information system may need to send client records to another (e.g. for reporting to a regional data warehouse, or as a contributing partner to a larger local tracking system). Usually, the goal of the data exchange is to make sure the client records added to the first system are also added to the second system, so they contain the same records. Optionally, the second system can also send back its own records to the first system, so that in the end, both systems have the exact same client records for a given time period. This task of matching, adding, removing, and updating records so two different information systems finally have matching sets of records, is called "synchronization". There is no standard defining exactly how to go about synchronizing records over the Internet.

Status Quo*: Lack of Standards for Human Services Data Synchronization and Resulting Consequences.*

Even though client record synchronization is the goal of most human services data integration projects, there is no human services or disaster response standard which defines how synchronization should take place exactly. We currently possess adequate client record file formats, but there is no sequence of expected procedures for:

- using the Internet to start a new synchronization,
- sending and receiving the specific client information which changed since the last time the two different information systems synchronized,
- comparing the client records with those already in the receiving system to see if there are any differences between the two information systems,
- and "marking" certain client records for addition/removal to/from each information system being synchronized.

Because there is no human services synchronization standard, nowadays each system typically sends all its client records for a given time period to another information system for comparison. This full comparison takes significant effort, as it can involve large quantities of data and data checking. More importantly, each system must come up with its own method to compare the records and provide its own set of instructions describing how the other system must go about sending the records. Usually, an old method of sending data files called the File Transfer Protocol (FTP) is used. Unfortunately, FTP is usually blocked by human services agency firewalls. A more efficient way of sending client records would be to use Hypertext Transfer Protocol (HTTP). HTTP is the method currently used to send web

1 Examples of client data formats are the Disaster Client Data Standard (http://clientdatastandard.org), a pending Patient Information Tracking format, HUD CARS (www.hud.gov/offices/hsg/sfh/hcc/CARS23_Quick_Interface_Reference_Guide_for_CMS_16.pdf), and HUD HMIS (http:/hmis.info).

pages over the Internet . HTTP passes easily through most human services firewalls.

Because the current way of sending client records is so customized and difficult, the sending system repeatedly sends large batches of client records, each covering a subsequent time period, and the receiving system may delete or add to those same received records, if it chooses. Eventually, the receiving system becomes significantly different ("out of synch") with the sending system, and reconciling the two systems grows increasing difficult over time.

A Vision for the Future

Instead of this tedious and customized process for human services client data synchronization, a new standard is needed. It would standardize a way for two systems to:

- recognize and sign on with each other,
- ask for and begin a synchronization for a certain time period,
- exchange records in a predefined order to carry out the actual synchronization which determines which records to add or delete.
- exchange the actual added or deleted records,
- keep track of what was marked as previously sent during a prior synchronization,
- perform the synchronization "over the web", transferring small amounts of data at a time, as opposed to large files for batch processing.

An Existing, Proven Solution: the OMA Data Synchronization Protocol

Fortunately, most of the work of defining human services synchronization protocol has already been completed by the Open Mobile Alliance Data Synchronization Working Group. While originally intended for synchronizing cell phone contacts and the like, this group has standardized a set of rules for two systems to synchronize records of any type, including client service records. This set of rules is entitled the "SyncML Representation Data Synchronization", and it is currently at version 1.2.1, and version 2.0 is nearing completion.

A Call for Action

We must begin a concerted public/private effort to develop, test, and implement, the OMA Data Synchronization Protocol for use in a human services context. This will require the development of a usage guide tailored to human services, sample code, testing schedules, and compliance oversight. A large number of partners will be required for this effort, which could be spearheaded by the Human Services XML group within the National Human Services Data Consortium, given its broad interagency focus.