

Minnesota HMIS XML Data Integration Community Project Profile

THE PARTNERS

Bowman Systems (<http://www.bowmansystems.com>) created ServicePoint, a web database application for tracking client information used in over 200 community and regional installations involving 30,000 users and 2,500 agencies across 43 states. ServicePoint was selected as Minnesota's Homeless Management Information System (HMIS) software in 2002.

Wilder Research (<http://www.hmismn.org/>) administers Minnesota's HMIS, including a state-wide ServicePoint database in use by more than 170 organizations serving people experiencing homelessness in Minnesota. Wilder Research also took lead responsibility for developing, documenting and refining Minnesota's platform-agnostic XML data standard, based on the federal XML HMIS data standard.

The State of Minnesota (<http://www.headinghomeminnesota.org/>) opted to align with and extend the federal HMIS requirements to meet its own needs, rather than creating separate reporting systems for state-funded programs dedicated to ending homelessness. In addition, the State authorized the creation of a portable data standard that would allow agencies to meet their data and reporting requirements without using Minnesota's HMIS system directly.

Hearth Connection (<http://hearthconnection.org/>) is a nonprofit intermediary organization working with government and dozens of community social service agencies serving people with long histories of homelessness in supportive housing. Hearth Connection's own web-based data system, Co-Pilot, was designed to capture the required data elements for the federal and Minnesota HMIS data standards. Hearth Connection volunteered to be Minnesota's first provider to transform its own data to the Minnesota HMIS data standard for uploading to ServicePoint.

THE PROJECT

The partners began meeting in the summer of 2005 to design a process for uploading data from an arbitrary source repository to Minnesota's implementation of ServicePoint, including the many extensions and customizations made to support Minnesota's state-funded homeless programs. The first phase of this work was to explicate the attributes and permitted values for the combined data elements. The thorough review needed as part of this comprehensive documentation led to several changes in the specific data elements being collected, and also helped motivate resolution of specific, XML-friendly and, where possible, industry standard key codes to represent values that have applicability in many domains (e.g., state or county abbreviations). Finally, the review of data elements also helped ensure a common and consistent semantic understanding of the values intended to be contained by those data elements.

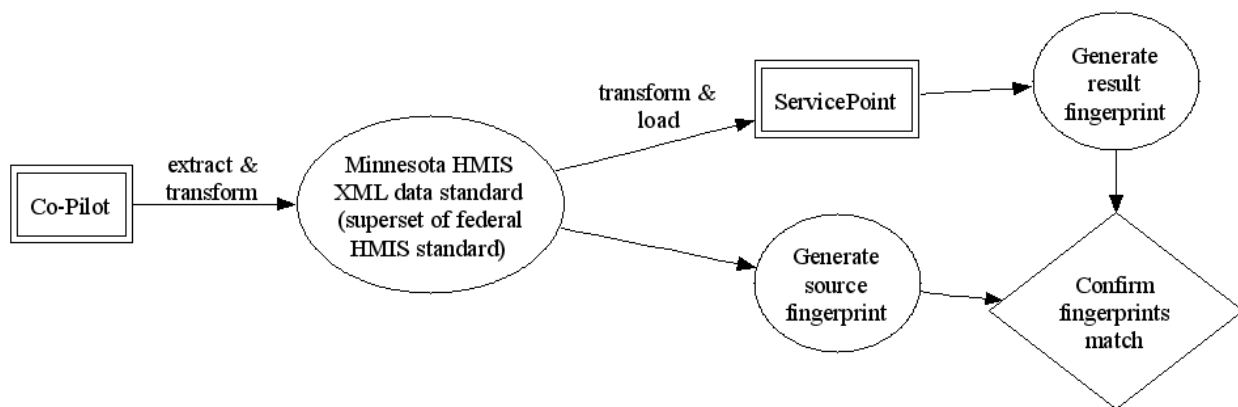
Once the data requirements were described in English, stored in an Excel spreadsheet, Wilder Research initiated the development on an XML schema, the formal data definition document that would encode the English-language semantic understanding in a format amenable to computer

processing. With support from HUD technical assistance funds and contributions of Eric Jahn from Alexandria Consulting, a draft XML schema was developed and went through many rounds of revision. The schema was intended to be an extension, or superset, of the federal HMIS XML standards that were under development at the same time. (This was not always possible, given the relationship between some state and federal data requirements, and the limited support for extensibility built into early drafts of the federal schema.) A draft version of Minnesota's resulting schema is available online: <http://www.hmismn.org/MNmodel>

In developing the schema, the partners also agreed to policies regarding its development, including a commitment to limiting changes to Minnesota-specific data elements (and corresponding changes to the schema) once per year. In practice, the schema has evolved more rapidly than that, as new requirements or problems have been identified and addressed. Once it has stabilized through additional use, the once-per-year timetable for changes to the schema may be more realistic. The partners all recognize the need for an ongoing commitment to the schema's development to keep it current with the changing data landscape.

In addition to the partners listed above, this initial planning and documentation phase involved several other agencies (including the association of Minnesota's community action agencies, Catholic Charities and Hennepin County). Once development of a draft XML schema was complete, Hearth Connection volunteered to be the first agency to use the schema for live data, working with Wilder Research and Bowman Systems to refine the schema as subsequent issues were identified. Hearth Connection initiated development of its transformation of its Co-Pilot's data to the schema, and produced the first such file in December 2006. This file, containing representative sample data extracted from Co-Pilot, helped surface technical hurdles with the remaining steps of the import process, and the need for verification that the data imported into ServicePoint matched the data provided. In 2007, the partners worked to tackle these challenges, to solidify the legal agreements governing the transmission and protection of individually identified participant data, and to develop a fingerprint report for verification purposes.

The fingerprints were designed to analyze the source and imported data in ways that would be unlikely to match by chance (e.g., a sum of the ASCII codes for the characters in every uploaded participant's first name). Elements of the fingerprint were chosen that favored XML-friendly calculations, so that the fingerprint itself could be generated as an XML transformation of the source data (this was only partially achieved). These steps are shown schematically below.



After many intermediate steps and partially successful uploads (each of which helped identify technical issues with the schema or transformation steps), Hearth Connection's source data from Co-Pilot was imported into ServicePoint successfully in March 2008, and again in August 2008 and February 2009. The partners continue improving the schema and have identified possible future refinements including automation, handling for anonymous data, supporting differential (or "delta") uploads and ultimately, synchronization between data systems. A number of other agencies have also expressed interest in using the upload process for their data.

For more information about this project, please contact:

Josh Johnson

Bowman Systems
318-213-8780 x131
josh@bowmansystems.com

Craig Helmstetter

Minnesota HMIS
Wilder Research
651-280-2700
cdh@wilder.org

Vicki Farden

Minnesota Housing
651-296-8125
vicki.farden@state.mn.us

Eric Grumdahl

Hearth Connection
651-645-0676
eric@hearthconnection.org