1-1-2008

A Case Study in Data Sharing: 211 Helplink and the AIRS XSD

Ian Cottingham
University of Nebraska-Lincoln

Ben Kutsch
University of Nebraska-Lincoln

Nancy C. Shank
University of Nebraska Public Policy Center, nshank@nebraska.edu

A Case Study in Data Sharing:
211 Helplink and the AIRS XSD

Prepared by:

Ian Cottingham, Chief Software Architect
Department of Computer Science and
Engineering
University of Nebraska - Lincoln

Ben Kutsch, Development Lead
Department of Computer Science and
Engineering
University of Nebraska – Lincoln

Nancy Shank, Associate Director
University of Nebraska
Public Policy Center

June 6, 2008

UNIVERSITY OF
Nebraska

Table of Contents

Acknowledgements

The authors wish to thank the following individuals whose participation made this project possible:

Augustine "Tino" Paz
Manager, 2-1-1 / CAN 2-1-1 Liaison
United Way of America

Craig Quincy
Software Development Manager
VisionLink

Zac Sackrison
Senior Web Developer
VisionLink

Ed Toomey
Software Systems Manager
RTM Designs

Cristina Vetrano
Project Manager - Implementation
Disaster Services
American Red Cross

Tony Zhou
Assistant Database System Manager
211 Helplink, United Way of the Bay Area

EXECUTIVE SUMMARY

The Coordinated Assistance Network (CAN) helps communities prepare for and respond to disasters. The ability to exchange up-to-date information about resources and clients is central to effective response. To facilitate the exchange of resource data, CAN has been working with information & referral organizations, namely 2-1-1s, to ensure that their existing data about community services may be easily and quickly shared with CAN in the event of disaster. CAN has been working with 211 Helplink (San Francisco, California) to develop an exchange using the data standard developed by the Alliance of Information and Referral Systems, AIRS XSD 2.07. The data exchange has failed. The University of Nebraska Public Policy Center and the University of Nebraska-Lincoln Department of Computer Science and Engineering agreed to analyze the exports, determine the failure points, and make recommendations for this and future data exchanges.

Based on the analysis, the following recommendations are suggested to improve data interoperability both in the specific instance of 211 Helplink and CAN, as well as between information and referrals and CAN throughout the United States:

THE 211 HELPLINK AND CAN DATA EXCHANGE

1. Both vendors should come into full compliance with the AIRS XSD 2.07 data specification.

2. Vendors and clients who wish to define elements and attributes beyond the standard should use the extensibility available within the structure of the AIRS XSD specification.

3. Vendors should develop and enforce proper data forms by disallowing non-compliant data to be accidently entered to ensure data quality.

4. Non-compliant vendors should provide the language that will transform their non-compliant data sets into an AIRS XSD 2.07 compliant format.

5. In order to accommodate the most streamlined data transfer process, thereby enabling the fastest response time during a disaster, CAN should require partner compliance with the AIRS data specification.

2-1-1 AND CAN EXCHANGE ACROSS THE UNITED STATES

6. Human service organizations should be encouraged to use software products that are able to produce AIRS XSD compliant data sets.

7. The Information and Referral community should be equipped with information about software product's ability to create AIRS XSD compliant data sets.

8. The AIRS XSD Workgroup should continue evaluating and recommending improvements to the AIRS XSD specification.

9. An omnibus tool for data normalization should be developed.

10. Solutions to interoperability challenges that extend beyond the standard should continue to be developed.

INTRODUCTION

The Coordinated Assistance Network (CAN) (Appendix A) works with communities across the United States: helping caseworkers improve assistance to individuals impacted by disaster. By creating the capability to data exchange *prior* to a disaster, CAN has been able to improve community readiness and ensure that information will quickly be made available to the range of human service organizations involved in disaster response.

Information and referral organizations (I&Rs) identify, code, and disseminate information about human services available in their communities. Many I&Rs, such as 2-1-1s (Appendix B), also operate call centers where the public and professionals contact trained specialists to identify appropriate services. I&Rs have served as information hubs in past disaster responses, including responses to September 11, 2001 and Hurricanes Katrina and Rita (see Shank, 2007; United Way of America, n.d.-a; United Way of America, n.d.-b). More recently I&Rs have provided assistance during the California wildfires of 2007 (2-1-1 California & United Way of America, 2008; Graham, 2007; Phelps, 2007) and other local disaster events throughout the United States.

To improve the ability of I&Rs to exchange data, the national I&R association, AIRS[1], has developed and promulgated data exchange standards. The standard defines the structure and form of data that I&Rs should be able to exchange. At this time, several major I&R software vendors have created utilities that allow their clients to create AIRS XSD-compliant exports. Observance of the standard is voluntary and there is no certification process in place that assures that vendors observe the standard. A new version of the standard, named AIRS XSD 3.0[2], is currently under consideration for adoption.

San Francisco's major I&R, 211 Helplink (http://www.211sf.org/), has been working with CAN for approximately a year to prepare for disaster data exchange. However, the exchange of data from the 2-1-1 system to the CAN system has failed: This, despite the fact that both the 211 Helplink software product (Refer, a product of RTM Designs) and the CAN software product (Tapestry, a product of VisionLink) purport to be observing the AIRS XSD 2.07. Further, there is a dispute between the vendors over which vendor is responsible for the failure and which vendor is responsible to intervene and correct the problems.

---

[1] The Alliance of Information & Referral Systems (AIRS) is the professional association for over 1,200 I&Rs, primarily in the United States and Canada (http://www.airs.org/).
[2] By the time of final submission of the report the status of AIRS 3.0 has changed. See the addendum to this document for more information.

The American Red Cross on behalf of the Coordinated Assistance Network contracted with the University of Nebraska Public Policy Center and the University of Nebraska–Lincoln Computer Science Laboratory for Advanced Research Computing to assess the compliance of the 211 Helplink and CAN data exports with AIRS XSD 2.07, identify areas of non–compliance, suggest remedies, and provide recommendations for future interoperability between 2–1–1s and CAN.

APPROACH

Representatives from the American Red Cross, the United Way of America, and the University of Nebraska approached the failure of the data exchange based upon the following understandings:

- CAN desires an overarching approach to data exchange that leverages communities' existing I&R resources;

- 2-1-1s are positioned to provide valuable information in disaster planning, response and recovery;

- CAN requires that data be rapidly available to their partners and others in times of disaster;

- Vendor-neutral standards eliminate the need for customized mapping of data between I&Rs and CAN and provide means for efficient exchange of information;

- Without certification processes, it is difficult for vendors or their customers to assess compliance with standards, or to assess other challenges to data exchange;

- The 211 Helplink and CAN data exchange difficulties provide a case study to analyze the challenges to data exchange; and,

- The analysis of the data exchange failure will be undertaken with the participation of both vendors (RTM Designs-the purveyor of Refer, and VisionLink-the purveyor of Tapestry), and both data sources (211 Helplink and CAN).

Key Activities and Timeframe

The key activities and timeline for the case study were as follows:

Key Activity	Timeline (2008)
Orientation Conference Calls with Both Vendors and Both Data Sources	March 20 - 25
VisionLink (CAN)	March 20
RTM Designs (211 Helplink)	March 25
University of Nebraska receives data exports	March 24 – April 3
VisionLink (CAN)	March 24
RTM Designs (211 Helplink)	April 3
University of Nebraska analyzes data exports	March 24-April 17
University of Nebraska sends to each vendor and data source an analysis of compliance, suggested fixes, and a copy of the validation output.	April 18-25
Vendor and data source deadline for response to University of Nebraska analysis	April 25
University of Nebraska issues preliminary report to American Red Cross	May 2
University of Nebraska issues final report to American Red Cross	June 6

Data Analysis

The vendors agreed to provide data sets to the University of Nebraska for evaluation. Neither vendor provided a copy of their software product to the researchers. Thus, it was not possible to verify that the exports were actually created using export utilities available in the vendor products.

The analysis of the data included validation against the AIRS XSD 2.07 specification release with an automated track of all validation errors and a manual examination of elements and attributes containing errors.

The first stage involved a standard validating parse of the XML file against the AIRS XSD 2.07. The parser was implemented using C# and did not terminate on errors. This way, the parser made a full pass of the XML document, logging validation

events for each non-compliant element and/or attribute encountered. The full results from the parse for both data sets are provided as separate documents.

The second stage of analysis involved a manual evaluation of elements or attributes in error to determine how the validation errors could be corrected to allow specification compliant data sets to be transferred. The results of the individual element or attribute analysis are contained in the following section.

The result of the analysis provided a complete understanding of those aspects of the provided data sets which do not meet specification compliance. This understanding allows for formal recommendations to be made to allow these data sets to be shared as part of CAN.

The analysis was provided to both participating vendors. Vendor comments are reported in the Results section.

RESULTS

Data sets were provided to University of Nebraska-Lincoln Computer Science and Engineering by both VisionLink (the CAN data) and RTM Designs (the 211 Helplink data). The validation against the AIRS XSD 2.07 specification resulted in errors in both vendors' data sets. The VisionLink (CAN) data set fared significantly better than the RTM Designs (211 Helplink) data set. The VisionLink data set had no attribute errors and had a very small rate of element errors, all of the same type. The RTM Designs data set had high rates of attribute and element errors, of a number of different types.

The results are summarized in the following table:

	VisionLink (CAN)	RTM Designs (211 Helplink)
File Name	airsxmlexport.xml	UWBAData.xml
File Size	100.3MB	1.5MB
Elements Processed	2,171,094	32,721
Attributes Processed	1,167,086	20,116
Elements in Error	35	754
Attributes in Error	0	327
Element Error Rate	0.000016	0.023043
Attribute Error Rate	0.000000	0.016256

Attributes and Elements

Analyses performed on the data sets provided by the vendors included data encapsulated as both attributes and elements. *Elements* are generally used to hold specific pieces of data. *Attributes* are used to provide additional information about an element.

For example, the AIRS XSD uses *elements* for such items as:
- Defining what constitutes a valid address ("Address"), as well as defining specific subparts of the address, for example, City ("City")

The AIRS XSD uses *attributes* for further definition of elements. For example,
- Defining whether an address is confidential ("Confidential")

VisionLink Data

There were a total of 35 errors found in the automated analysis of the data set provided by VisionLink. These 35 errors were the result of a single error type:

1. The addition of data this is not in the list of supplied choices listed in AIRS 2.07

The following table displays the error, classification and a work-around to bring the data into compliance with AIRS 2.07.

Problem	Reason	Type
Inclusion of Urdu as a language offering.	The language code 'urd' is not a recognized code.	1
Work-around		
The following change can be made by VisionLink to allow provision of data with Urdu as a language in a specification compliant way. <Languages><Name>OTHER</Name><Notes>Urdu</Notes></Languages>		

VisionLink believes that the omission of Urdu is an error in the AIRS XSD 2.07 because it is an ISO 639-2 language. However, not all ISO 639-2 languages are included in the AIRS specification, so VisionLink's inclusion of the language, in the manner it was included, is a violation of the specification.

RTM Designs Data

There were a total of 1,081 errors found in the automated analysis of the data set provided by 211 Helplink. These 1,081 errors are summarized as follows:

1. The addition of non-AIRS 2.07 elements into the XML
2. Misuse of AIRS 2.07 elements or attributes
3. The addition of data that is not in the list of supplied choices listed in AIRS 2.07

The number of errors in the data set provided by RTM Designs prevented a complete analysis of all issues from being conducted in the time provided. The following table summarizes selected errors, classifications and work-arounds to bring the data into compliance with AIRS XSD 2.07.

Problem	Reason	Type
Agency elements contain: <DisabilitiesAccess> </DisabilitiesAccess> <OVERVIEW> </OVERVIEW> <TRANSPORTATION> </TRANSPORTATION>	These are not valid Agency elements	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
Site elements contain : <AREASERVED> </AREASERVED> <CROSSSTREET> </CROSSSTREET> <DISASTERSURVEYCOORDINATORNAME> </DISASTERSURVEYCOORDINATORNAME> <DISASTERSURVEYCOORDINATORNAMEANDNUMBER> </DISASTERSURVEYCOORDINATORNAMEANDNUMBER> <DISASTERSURVEYCOORDINATORNUMBER></DISASTERSURVEYCOORDINATORNUMBER> <DISASTERSURVEYDATEUPDATED></DISASTERSURVEYDATEUPDATED> <DISASTERSURVEYSECONDARYCONTACTNAME> </DISASTERSURVEYSECONDARYCONTACTNAME> <DISASTERSURVEYSECONDARYCONTACTNAMEANDNUMBER> </DISASTERSURVEYSECONDARYCONTACTNAMEANDNUMBER> <DISASTERSURVEYSECONDARYCONTACTNUMBER></DISASTERSURVEYSECONDARYCONTACTNUMBER> <TRANSPORTATION> </TRANSPORTATION>	These are not valid Site elements	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
Service elements contain: <Program> </Program> <AREASERVED> </AREASERVED>	These are not valid Service elements	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
Taxonomy elements contain: <Name/> <URL/>	These are not valid Taxonomy elements	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
The Source element has attribute 'OriginTimeStamp'	Attributes are case sensitive	2
Work-around		
Change 'OriginTimeStamp' to 'OriginTimestamp'		

Problem	Reason	Type
The Agency and Site elements have attribute 'ExcludeFromWebSite'	Attributes are case sensitive	2
Work-around		
Change 'ExcludeFromWebSite' to 'ExcludeFromWebsite'		

Problem	Reason	Type
AKA element contains invalid Name elements	AKA cannot have any child elements	2
Work-around		

Instead of
<AKA Confidential="false">
    <Name>AIRRC</Name>
</AKA>
Use
<AKA Confidential="false">
    AIRRC
</AKA>		

Problem	Reason	Type
Contact elements are invalid	Contact elements must contain at least one Phone or at least one Email element	2
Work-around		
Include the required element(s)		

Problem	Reason	Type
There are Longitude values greater than 0	Longitude must be in the form -180<=x<=0	2
Work-around		
This should include all longitudes in the U.S. There appears to be a data issue where values of in 211 Helplink of -122 are being entered as 122		

Problem	Reason	Type
Sites are missing SiteDescription	SiteDescription is a required element	2
Work-around		
Include required element <SiteDescription/>		

Problem	Reason	Type
Sites are missing IRSStatus	IRSStatus is a required element	2
Work-around		
Include required element < IRSStatus />		

Problem	Reason	Type
URL elements are missing Address and Note elements	Address and Note are required elements	2
Work-around		
Include required <Address/> and <Note/> elements		

Problem	Reason	Type
Urls are written without "http://" or "https://"	For the url string to be valid it must include the "http://" or "https://"	2
Work-around		
Make sure that all URL are in the following example format: <URL> <Address>http://www.google.com</Address> <Note></Note> </URL>		

Problem	Reason	Type
Service elements have an element ServiceDescription	Service elements require a Description element	2
Work-around		
Replace <SERVICEDESCRIPTION/> with <Description/>		

Problem	Reason	Type
Service elements are missing Hours elements	Services must have at least one valid Hours element	2
Work-around		
Include at least one <Hours> element		

Problem	Reason	Type
ServiceSiteLink elements contain <Key/> elements	Key is not a valid element for ServiceSiteLinks	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
Taxonomy elements contain <Name/> and <URL/> elements	Name and URL are not valid Taxonomy elements	1
Work-around		
Do not include invalid elements		

Problem	Reason	Type
SourceOfFunds elements contain the following invalid values: <SourceOfFunds>City</SourceOfFunds> <SourceOfFunds>County</SourceOfFunds> <SourceOfFunds>Donations</SourceOfFunds> <SourceOfFunds>Federal</SourceOfFunds> <SourceOfFunds>Fees</SourceOfFunds> <SourceOfFunds>Grants</SourceOfFunds> <SourceOfFunds>Independent Fund Raising</SourceOfFunds> <SourceOfFunds>State</SourceOfFunds>	These are the only valid SourceOfFunds: UNKNOWN OTHER Contributions Foundations/Grantmaking Public Charities (other than United Way) United Way FundRaising Government - County/Regional Government - Federal Government - Municipal Government - State/Provincial Government - Unspecified In-Kind Support Only Investment Income Program Fees or Membership Dues PrivateGrants Special Events	3
Work-around		
Only use data in the valid list.		

Problem	Reason	Type
The PhoneType attribute contain 72 different invalid entries	These are the only valid PhoneTypes: Voice Fax TTY/TDD Pager Modem Cell Voice/Fax Voice/TTY/TDD Unspecified	3
Work-around		
Only use data in the valid list.		

DISCUSSION OF RESULTS

The primary issue at hand with regard to sharing data between vendors is a lack of compliance with the AIRS XSD 2.07 specification. The purpose of the specification is to provide a common interface for organizations sharing data. Deviation from the specified data structures creates a degree of ambiguity requiring some form of intervention to consume data. This issue develops anytime the specification is deviated from, even in instances where the specification is deficient in some way. For this reason, the analysis performed does not take into account the rationale for deviation; adherence to the specification was the sole factor of analysis.

The purpose of a data specification is to alleviate the need to determine the structural meaning of some set of elements or attributes. By providing a common set of well-defined structures, the specification exists to allow free flow of data between systems based solely on the agreed-upon definition of the structures (elements or attributes). Were each data set to be fully specification compliant, each could be easily consumed by a receiving party as the data therein would be equivalent in terms of the structural representation.

In the case of the two data sets in question, both deviated from the specification in some form. Deviation from the specification requires a mapping of non-equivalent elements or attributes to the receiving party's data structure and leads to an inability to quickly consume data from many vendors. By deviating from the specification, a data provider might conceivably gain the ability to include a wider and more varied breadth of data. However, this perceived benefit of including a greater number of data elements or attributes is detrimental to data sharing. By providing data that is not clearly defined, the vendor is inhibiting the data sharing process through obfuscation. When a significant portion of a data set is non-compliant, it cannot be easily digested into other systems, thereby rendering the data problematic until it can be defined.

The analysis performed by the University of Nebraska demonstrated that neither the VisionLink nor the 211 Helplink data sets where compliant with the AIRS XSD 2.07 specification, resulting in some form of translation to digest either. In the case of VisionLink, the inclusion of an unrecognized language type would require additional processing to consume those records. Beyond that inclusion, the VisionLink data set is fully AIRS XSD 2.07 compliant. Exclusion of the non-compliant data or the recommended correction would bring the data set into compliance and allow for data sharing. The case of the 211 Helplink data set is considerably more complex. The data set is, at best, loosely based off the AIRS XSD 2.07 specification. To imply that this data set is AIRS XSD 2.07 compliant would be fundamentally false.

The analysis performed by the University of Nebraska found: misused elements, ignored structural relationships, use of unrecognized elements, and ignored attribute restrictions.

Significant effort and a more detailed understanding of the data structures generated by the RTM Designs software product would be required to translate this data set into an AIRS compliant data set or to allow import into an existing database.

Beyond simply bringing the data sets into compliance with the AIRS XSD 2.07 specification, the issue can be resolved in one of two ways:

1. Each vendor can provide an Extensible Stylesheet Language Transformation (XSLT) for their data set that will transform the non-compliant AIRS XML into AIRS XSD 2.07 compliant XML.
2. The receiving party can create an XSLT for each non-compliant XML data set they wish to consume that will translate said data set into AIRS 2.07 compliant XML.

The first option places the responsibility of compliance on the party in violation and represents the least overall effort. This is due primarily to the fact that the provider of the data set is aware of the non-compliant elements provided and can much more effectively provide proper translations for them, thereby minimizing data loss and misrepresentations. The second option places responsibility for the translation of the non-compliant data on the party wishing to consume the data (e.g., aggregate the data), rather than on the party in violation. This option is considerably less effective (in correctly translating data), less timely (in the ability to quickly and easily exchange data), and not sustainable (a small change in the sending party's data structure may invalidate the translation). The multiplicity created by forcing each data consumer to create a translation document for a single data provider results in potential data loss through differing degrees to which the translation takes place. By creating and distributing a single translation document, the provider is able to ensure that multiple consumers each translate the non-compliant data in the same way.

Therefore, if a vendor is unable or unwilling to comply with the AIRS XSD and instead expects a data transformation to be made, it is appropriate to place the responsibility for developing the translation specification on the vendor providing non-compliant data sets. To facilitate broad data sharing activities with minimal effort, vendors should produce AIRS XSD 2.07 compliant XML. Those vendors who choose, for whatever reason, not to be compliant with the specification should bear the additional burden of facilitating compliance in order to participate in data sharing activities. This also gives the vendor an increased level of control in preventing data loss through the translation of ambiguous records and provides an increase in data integrity. In this way, a system is created allowing the vendors the freedom to deviate from the specification while not inhibiting the process of data

sharing in a disaster scenario. It cannot, however, be stressed enough that compliance with AIRS standards is the most effective and preferred way of sharing data.

Data exchange during and immediately following disasters requires the ability to rapidly exchange data from a multitude of data sources. The following recommendations are offered to ensure that this type of data exchange is possible. The recommendations respond both to the specific exchange of data between 211 Helplink and CAN as well as to other exchanges that will be needed across the United States.

THE 211 HELPLINK AND CAN DATA EXCHANGE

1. **Both vendors should come into full compliance with the AIRS XSD 2.07 data specification.**

   The AIRS XSD was created to facilitate data sharing. The standard is vendor-neutral and has been created over the past several years with the participation of I&Rs, vendors, and computer scientists. Standards provide the foundation for straightforward data exchange. It prevents costly and time-consuming customized mapping exercises. Both vendors should fully-observe the AIRS XSD. In VisionLink's case, only a minor workaround is required. RTM Design's workarounds are more extensive, but achievable.

   AIRS XSD 3.0 removes specificity of content (e.g., pick lists) in the standard. Thus, for example, VisionLink's attribute error (i.e., the addition of the language Urdu in the language list) will no longer result in non-compliance.

2. **Vendors and clients who wish to define elements and attributes beyond the standard should use the extensibility available within the structure of the AIRS XSD specification.**

   The AIRS XSD was created to define the basic data set that I&R would be interested in exchanging. However, it was recognized, that no national standard could (or should) represent all the data that specific I&Rs may ever want to exchange. For example, the AIRS XSD does not attempt to represent locally-unique information about services (e.g., whether an agency is a member of a community planning collaborative, or whether services are funded by a particular local funding organization). So that I&Rs could use the AIRS XSD, but add to it when needed, the standard was structured to accommodate extensions. Extensive use of this functionality, however, is strongly discouraged as it circumvents many of the standard structures defined by the XSD; the goal of extensibility is to accommodate limited special cases only.

3. **Vendors should develop and enforce proper data forms by disallowing non-compliant data to be accidently entered.**

   Improper content may cause invalid data. For example, a user may incorrectly type an email address without the standard "@" within the alphanumeric string, or may omit the "http://" or "https://" at the beginning of a website address. Vendors should develop editors, guides, or alerts that identify proper and improper data forms. Although vendors cannot be held responsible for all content errors, vendors could enable customers to input data that not only will be valid in the AIRS XSD, but also will be less likely to have errors in their everyday I&R use.

4. **Non-compliant vendors should provide a translation mechanism that will transform their non-compliant data sets into an AIRS XSD 2.07 compliant format.**

   If a vendor chooses not to create exports compliant with the AIRS XSD, the vendor should be responsible for translating the data structure so that it may be consumed in an AIRS XSD 2.07 normalized way. An XSLT document provides the language needed to translate an XML document into another XML form. Non-compliant vendors should create the XSLT that "maps" the non-compliant data into an AIRS XSD compliant form. Non-compliant vendors are in the best position to create the XSLT because they are most familiar with their data structures and would be able to minimize unnecessary data loss. Due to the uniqueness of data structures maintained by vendors in their data base systems, as well as the varied level of non-compliance with the AIRS specification, providing a generalized estimate concerning the amount of effort this process would take is extremely difficult. Such an estimate would have to be provided on a vendor-by-vendor basis with an analysis of: 1) the database schema employed by the vendor; and 2) a full analysis of the XML produced by the vendor's software.

5. **In order to accommodate the most streamlined data transfer process, thereby enabling the fastest response time during a disaster, CAN should require partner compliance with the AIRS data specification.**

   The most efficient means of improving data sharing between partners, particularly when rapid sharing is required, is to require all partnering organizations to use software that is able to product AIRS compliant exports. When organizations use products that create AIRS compliant exports, then "aggregators" (e.g., CAN) of the information are able to much more efficiently combine data sets.

6. **Human service organizations should be encouraged to use software products that are able to produce AIRS XSD compliant data sets.**

   Compliance with the AIRS XSD provides a foundation for the more timely exchange of data. Without widespread adoption of the standard by vendors, data exchange will continue to be a vexing, time-consuming, and costly endeavor. Additional education is needed to promote standards within the I&R and vendor communities. I&Rs, and other human service organizations, should be encouraged to use software products that are able to export AIRS XSD compliant data sets. The use of software products which do not easily export AIRS XSD compliant data sets should be avoided.

7. **The Information and Referral community should be equipped with information about software product's ability to create AIRS XSD compliant data sets.**

   I&Rs and other human service providers are currently fairly dependent upon vendors to confirm that their products are AIRS XSD compliant. The field lacks any unbiased review or certification of products that could inform users. Two options that should be considered to assist I&Rs are:

   - Independent product review
   - Certification process

   An independent product review could be conducted of all major I&R software products. This process could be done with or without active participation of the vendors. The assessment would be done with the same data set entered into each product and XML data dump would be generated. The resulting XML data would be validated against the AIRS data specification in a fashion similar to that used for this review. Based on the results, a rating could be applied to the software package. Such a rating would allow vendors and users to determine suitability of a given software package for data sharing activities in this space. Such a review would need to be conducted by an independent organization who themselves do not participate in data exchange in this space.

   A certification process could further formalize vendor compliance with the AIRS XSD. Vendors could submit their products for testing and certification to a neutral third-party certifying organization. Ideally, this certifying organization would operate as an independent consultant to AIRS (as the copyright holder of the AIRS XSD) or another public or private organization (such as a federal governmental agency or CAN).

8. **The AIRS XSD Workgroup should continue evaluating and recommending improvements to the AIRS XSD specification.**

   As with any standard, the AIRS XSD will benefit from continued refinement and evolution. The AIRS Board has committed to no more frequent than annual major versioning. Limiting versions helps to ensure that vendors are not overwhelmed by frequent, significant changes to the standard. Since the establishment of the AIRS XSD, only three major versions have been released. The I&R community and vendors benefit from a standard that is responsive to needs of the field.

9. **An omnibus tool for data normalization should be developed.**

   A data-translation language and accompanying software package should be developed that would allow vendors to define rule-sets for their data to facilitate translation of independent data sets into AIRS XSD compliant data dumps. By utilizing such a translation tool, value-add and non-specification compliant data structures can be transparently embedded in AIRS compliant data dumps and extracted on-demand by participating vendors recognizing such structures. Development of such a tool through CAN or some other organization, while not a substitute for vendors simply providing compliant data, would be beneficial to achieving a transparent data transfer mechanism, in the absence of vendor compliance. Having such a mechanism in place would alleviate much of the cumbersome translation activities recommended in this report to bring vendor data into AIRS compliance.

10. **Solutions to interoperability challenges that extend beyond the standard should continue to be developed.**

    The standard defines the data set for exchange of resource data. However, the standard, itself, does not ensure interoperability beyond the data. Additional frameworks are necessary to address the issues of:

    a) Taxonomy coding levels
    b) Structure for identifying organizational data ownership
    c) Matching and deduplication strategies (for client information)
    d) Privacy and security

REFERENCES

2-1-1 California and United Way of America. (February 2008). Trial by Fire: How 2-1-1's regional response to the 2007 Southern California wildfires underscored the need for a statewide network. Retrieved on April 17, 2008 from http://www.211.org/documents/TriabyFireUW.pdf

Graham, M. (2007, November 8) Disaster Gets Its Own Phone Number…Some Places. *Wired*. Retrieved November 30, 2007, from http://www.wired.com/politics/security/news/2007/11/two_one_one

Phelps, A. (2007, November 28) Wildfires gave reverse 911 its first test. National Public Radio. Retrieved November 30, 2007, from http://www.npr.org/templates/story/story.php?storyId=16703144

Shank, N. (2007) 2-1-1 in the Mid South Delta: A Collective Case Study, University of Nebraska Public Policy Center, Lincoln, NE. Available from http://ppc.nebraska.edu/program_areas/documents/2-1-1MSDI_Case_Study.pdf

United Way of America (n.d.-a) Shelter From The Storm: The Monroe Experience. Alexandria, VA.

United Way of America (n.d.-b) 2-1-1:A critical tool in times of disaster, Alexandria, VA. Retrieved April 17, 2008 from http://www.211.org/documents/211CriticalTool.pdf

ADDENDUM

At the time of writing, the AIRS 3.0 data specification was under final review by the AIRS Board. The specification was approved for adoption on May 18, 2008. The new version of the standard strips some of the content-specific elements (e.g., pick lists) from the standard. This enables vendors and users to more easily respond to unique content needs thus minimizing changes to the standard, except for substantial structural issues. The changes found in AIRS 3.0 will result in fewer compliance issues for both of the vendors who provided data for this analysis. While this is not a specific endorsement of either data set as AIRS 3.0 compliant, many of the attribute errors found in the data when validating against AIRS 2.07 will no longer exist in a 3.0 validation. At a minimum, the effort required to bring data into AIRS compliance will be significantly minimized with version 3.0 of the specification.

AUTHORS

Ian Cottingham is the Chief Software Architect for the Laboratory for Advanced Research Computing in the Department of Computer Science and Engineering (CSCE) at the University of Nebraska – Lincoln. The Laboratory for Advanced Research Computing (LARC) provides industry-level software support for funded intra-disciplinary research projects. Mr. Cottingham has directed operations of LARC since its formation in 2004. In this capacity he has been involved in an ongoing partnership with the United States Department of Agriculture Risk Management Agency to develop tools for agriculture producers and researchers to study and quantify the effects of climate risk events, specifically drought. In conjunction with CSCE faculty and partners from the National Drought Mitigation Center this work has resulted in over $10 million in funding across 6 years. Mr. Cottingham has also served as a software development lead in IBM Global Services, President of OrangePeel inc., and Technology Committee Chair for the Lancaster County Republican Party. He continues to focus on business driven computing research and the commercialization of research-oriented intellectual property. He is currently the principal investigator on a cooperative partnership with Region V Services of Nebraska: *An Interactive Education System Proof of Concept,* as well as a partnership with Los Angeles County 211: *AIRS Specification Support.* Mr. Cottingham is pursuing a PhD in computer science with a focus on data modeling and data-driven software architecture.

Ben Kutsch, graduated with a degree in Computer Science from the University of Nebraska - Lincoln in 2002. After graduation he spent two years working at a local publishing company as integral part of the database administration/development team. He is currently a lead software developer with the University of Nebraska – Lincoln, Department of Computer Science Laboratory for Advance Research Projects (LARC). He has over 6 years of programming experience and has worked on numerous projects and has managed projects during their complete lifecycle from planning to implementation to deployment. Most recently he oversaw and developed the redesign of the Nebraska 211 website to facilitate search and data retrieval of health and human services. His focus is on databases as well as web programming. In his spare time he is working on completing his MBA with an expected graduation of spring 2009.

Nancy Shank, MBA, is the Associate Director of the University of Nebraska Public Policy Center (http://ppc.nebraska.edu/). The Center contributes to public policy by developing and making available objective analysis of public issues for the public, elected officials, executive agencies, and others who represent the diverse interests of Nebraskans. Ms. Shank's work has focused on information technology and health and human services delivery systems. In 2000, Shank, Dr. David Rosenbaum (College of Business Administration, University of Nebraska-Lincoln), and Dr. Gregg Wright (Center on Children Families and the Law, University of Nebraska-Lincoln), co-authored Survey Of Existing I&R Services And A Nebraska 211 System Cost/Benefit Analysis (University of Nebraska Public Policy Center). This report described the existing information and

referral environment in Nebraska, presented operational steps and alternatives, and calculated cost/benefit ratios. This report has generated considerable interest and has served as the basis for analysis and policy discussion in a number of other states. It has also been cited in supportive documentation for the Calling for 2-1-1 Act. Ms. Shank was the principal investigator for a three-year U.S. Department of Commerce Technologies Opportunities Program grant that brought partners together across the U.S. to develop the technological link between propriety software and XML standards to enable interoperable information and referral resource sharing. She has published several articles about information and referral services. Ms. Shank is currently the principal investigator for a three-year (2005-08) Transforming Healthcare Quality Through Information Technology-Implementation Grant through the Agency for Healthcare Research and Quality to create a sustainable electronic system for health information exchange among western Nebraska's hospitals, public health, behavioral health, and other healthcare stakeholders. Nancy serves on the Nebraska Information Technology Commission eHealth Council, the National Human Services Data Consortium Board and its Human Services XML Workgroup, and is a member of the Alliance of Information and Referral Services XML Technology Subcommittee. She has consulted with numerous organizations throughout Nebraska and the U.S. on organizational development, program evaluation, and information sharing.

Appendix A

The Coordinated Assistance Network (CAN)

Origins and Background

The origins of the Coordinated Assistance Network (CAN) grew out of the terrorist attacks of September 11, 2001. The tragedy showed clearly that making clients find their way through a web of service providers caused added confusion in an already trying time. Several disaster clients were lost within the improvised system; others were shuttled from appointment to appointment, having to tell their painful story time and time again. The lessons learned in New York had a significant impact on both national and local disaster relief agencies, leading to an intensive search for not just a better way to deliver services to those in need, but to also help communities improve planning for natural and man-made disasters.

Galvanized by the September 11th experience, in 2003 seven leading relief organizations joined together to form CAN. The founding members are the Alliance of Information and Referral Systems (AIRS), the American Red Cross, National Voluntary Agencies Active in Disaster (National VOAD), The Salvation Army, and United Way of America – and two New York organizations – Safe Horizon and the 9/11 United Services Group. The American Red Cross serves as the fiscal agent of CAN.

CAN's Charge

CAN's mission is clear: To support disaster relief organizations to develop and implement a sustainable system and proactive approach to integrated service delivery and to appropriately and effectively share client and resource information following large scale disasters.

CAN's vision is ambitious: Disaster relief organizations support one another for the common good through a coordinated assistance network. The network promotes and locally adopts integrated service delivery models and a standards-based system of client and resource information exchange to ensure clients receive the most efficient, effective assistance following disasters.

Understanding there is no market share on human suffering, the founding agencies have created a fully functional, user driven technology platform now in place to ease the burden on relief agency works and disaster victims, alike. Caseworkers from over 200 non governmental agencies log onto the secure, web based CAN database to review client and resource information and assistance provided by other agencies. The capability enables caseworkers to

provide assistance quickly reducing paperwork so disaster victims are not burdened by repeatedly retelling their stories.

In addition, because caseworkers can see the details of each client's case, including what services and resources have been provided by other agencies, they are better able to craft a more comprehensive, holistic recovery plan for that client or family, without unnecessarily wasting resources.

CAN by the Numbers:
- 19 national, 4 regional and 230 local of the nation's leading relief agencies (253 total) have become CAN Participating Agencies
- 4,688 users have registered on the system since Hurricane Katrina's landfall; 2,553 are currently active users.
- The CAN database contains 369,544 active client records
- The CAN database contains 14,379 resource agencies
- The CAN database contains 32,504 resource services
- The CAN database contains over 90,000 client records provided by FEMA to eliminate duplication of benefits.

Community Based Disaster Planning
In early 2005, CAN began a pilot communities program to introduce the CAN model of collaboration operational coast-to-coast in the U.S. The six pilot communities are: New York City; Washington, DC, New Orleans, Oklahoma City, San Francisco and Seattle

The pilot communities are working to develop the scope, magnitude and complexity of plausible major natural disasters, terrorist attacks or other major emergencies that pose the greatest risk to their communities. Working with the CAN staff, the communities are working to identify the tasks, capabilities, resources and service delivery modifications required for a coordinated response and recovery to such disasters. Each community continues to show significant interest in CAN, particularly in the use of the technology to enhance existing systems.

In a 2006 pilot community survey completed, 93% of the participants responded affirmatively when asked whether the activities associated with CAN helped their community improve the capacity to respond to a disaster within the community. In the same survey, 100% of respondents affirmed that the activities associated with CAN helped their community improve the capacity to recover from a disaster within the community.

Appendix B

2-1-1 Fact Sheet

2-1-1 is an easy to remember phone number that connects individuals with resources in their community.

With approximately 1.5 million nonprofit organizations in the United States plus scores of government agencies, finding help can be confusing and intimidating. 2-1-1 centers are staffed by trained specialists who quickly assess the callers' needs and refer them to the help they seek. It's simple to remember, accessible to everyone at no cost, and available 24/7 with multilingual capabilities.

2-1-1 enables people to get help or give help.

2-1-1 offers information on a broad range of services, including rent assistance, food banks, affordable housing, health resources, child care, after-school programs, elderly care, financial literacy, and job training programs. Specialists at 2-1-1 centers facilitate thousands of volunteer hours and direct donors to locations where their gifts may be most needed and appropriate. One call center's referrals facilitated nearly 65,000 volunteer staff hours worth over $1 million.

2-1-1 benefits the nation.

The human services system in many of our cities and states is not only inefficient and costly, but is confusing and time consuming for consumers seeking to give or get help. It need not be this way. A 2004 University of Texas at Austin cost-benefit analysis of 2-1-1 estimates a net value to society approaching $130 million in the first year alone, and a conservative estimate of $1.1 billion over ten years. A national 2-1-1 system produces cost savings for tax payers, employers and government; and 2-1-1 in any community saves time and enhances the human services experience for those needing assistance.

Supporting 2-1-1 benefits your community.

2-1-1 is locally designed by community stakeholders who are aware of their local and state needs and resources. 2-1-1 is part of the community fabric, employs local citizens, and serves the local community. Businesses, nonprofit organizations, and government officials support 2-1-1 as a way to improve the lives of the residents in their communities. As of January 1, 2007, 2-1-1 was available to over 190 million Americans – approximately 65% of the U.S. population – with 209 active 2-1-1 systems operating in 41 states, the District of Columbia and Puerto Rico.

2-1-1 enhances public safety and crisis recovery efforts.

From hurricanes and floods to bio-terrorism, 2-1-1 call centers that already exist in communities, operating 24/7, are the logical platform for building emergency response communication capacity. In the wake of Hurricanes Katrina and Rita, hundreds of thousands of Louisiana, Texas and Alabama residents called 2-1-1 with a multitude of needs, including shelter, transportation, medical, food and water, construction materials, mental health, and questions about the availability of and application process for federal, state, and nonprofit assistance. 9-1-1 referred non-emergency calls to 2-1-1, freeing up 9-1-1 operators for life-and-death situations.

> "2-1-1 offers information on a broad range of services, including rent assistance, food banks, affordable housing, health resources, child care, after-school programs, elderly care, financial literacy, and job training programs."

2-1-1 is a private-public partnership.

2-1-1 is funded through local and state sources including local United Ways and other nonprofits, foundations, businesses, and state and local government. The *Calling for 2-1-1 Act* seeks to authorize federal funding of $150 million for years one and two, and $100 million for years three through five through the U.S. Department of Health and Human Services (HHS) to help implement and sustain 2-1-1 nationwide. States would be required to provide a 50% match in order to draw down the federal dollars.