Database Interoperability: Technology and Process for Sharing Resource Information[1]

Nancy Shank,
University of Nebraska Public Policy Center

Information and Referral (I&R) programs provide a critical connector in the health and human service delivery system. I&R programs include those that offer specialized information about services (such as respite services) or are targeted at a specific population (such as seniors), and those that offer comprehensive information about a broad range of health and human services. I&R services enable individuals, families, and staff from helping agencies to identify, understand, and access programs and services. However, I&Rs have struggled to easily share the information in these rich resource databases. Technical and process barriers have frustrated the exchange and merging of resource information. I&Rs who are committed to sharing data have been forced to use the same software product or to apply less than optimal technological solutions that result in lost data elements or time-consuming and awkward workarounds each time services or programs change

Because I&Rs are unable to easily share resource information with one another, multiple I&Rs may find themselves allocating scarce resources to identify and maintain the same resource information about services in their communities. Multiple, isolated databases may exacerbate the already complex system of health and human services, requiring persons (particularly those with complex access or service needs) to make numerous contacts before finding an appropriate referral. Data sharing between I&Rs would enable seamless provision of information. This article will explore the technical and process solutions that are being applied to enable sharing across software products.

Information and Referral (I&R) programs are found in nearly every community in the United States, and they are critical connectors in the health and human service delivery system. I&R services enable individuals, families, and staff from helping agencies to identify, understand, and access programs and services. Families and individuals looking for appropriate medical, educational, social service, and public health services find it difficult to locate and access the multiple programs that may be available (Agosta & Melda, 1995; Koenning, Benjamin, Todaro, Warren, & Burns, 1995; Levinson, 2002). National economic conditions and the profound changes in welfare over the past decade are increasing the need for services (Mathematica, 2002).

[1] This information was made possible, in part, by a grant (#31-60-I03013) from the Technology Opportunities Program, National Telecommunications and Information Administration, U.S. Department of Commerce.

Professionals working with families also lack information about the ever-changing programs offered by other organizations. To assist clients, many agencies create their own databases of services but find that maintaining current, comprehensive information is time-consuming, costly, and duplicative of others' efforts. In Nebraska, 141 agencies reported committing nearly 31,000 hours to tracking services offered by other organizations, and the 57 agencies that disclosed expenses reported they spent well over *$4 million annually* (University of Nebraska Public Policy Center, 2000).

I&R resource databases and calling statistics also allow planners and policymakers to determine whether there are duplications or gaps in programs and services, thus promoting for strategic decision-making based on demonstrated need. Increasingly, I&Rs are also seen as a vital component of homeland security/ bioterrorism response. After 9/11/2001, Connecticut's 2-1-1 system managed donor and volunteer information, received crisis calls, provided technical assistance to other agencies, and was able to quickly include resource information about newly-available services established for the crisis. During Toronto's August 2003 blackout, 2-1-1 assisted callers essential needs as food and water, and keeping cool (Even more recently, 2-1-1 was a critical component in the disaster response during Florida's hurricanes of 2004 (see, e.g., "California Gets New 211 Service," 2004, para. 5; M. Botero, personal communication, February 24, 2005; United Way of America, 2004). For example, the coordination between 2-1-1s in Atlanta and throughout Florida enabled over 1,500 callers to receive seamless information about such critical needs as where obtain supplies such as sandbags and dry ice, and where to find shelter (Lee, 2004).

Access to More Comprehensive Information

No single I&R is able to meet the needs of all possible consumers at all times, because consumers have needs that transcend individual I&R missions. I&R programs include those that offer *specialized* information about services (e.g., elderly population, respite services, disability- or disease-specific, child care) and those that offer *comprehensive* information about a broad range of health and human services, usually for a specific geographic area. Information sharing would enable individual I&Rs to better meet clients' needs and eliminate the need to divert staff time to maintaining data outside agencies' missions. For example, a client of the Area Agency on Aging may need child care information for a grandchild, or a First Call for Help© caller may need resources for a sibling in another county. Data sharing promises more comprehensive information will be at the fingertips of those already working with a client.

Data sharing reduces duplication between databases. For example, a disease-specific I&R and a homeless I&R may both maintain information about a state's disability programs, because disability programs are relevant to both organization's clients. Working together, the I&Rs could jointly determine who would be responsible for gathering and maintaining state disability program information. The other I&R, then, could simply get updated information through the routine sharing process. I&R information sharing may also reduce costs for agencies listed in the databases.

Agencies need only to maintain updated information with one of the I&Rs within an information sharing network.

Indeed, data sharing is not *just* a good idea, it is a requirement for all I&Rs who are accredited through the Alliance of Information and Referral Systems (Alliance of Information and Referral Systems, p. 18):

> I&R services within the system shall endeavor to participate in local database collaboratives as a means of avoiding duplication of database maintenance activities and achieving broader coverage of different types of community resources. The I&R services within the system shall: Appropriately divide information gathering tasks and oversee the regular and systematic exchange of resource database information; and Maintain comprehensive, accurate, and up-to-date information on the community resources for which they have maintenance responsibility.

Database Interoperability

Interoperability is one of the cutting edges of information technology. According to the Institute of Electrical and Electronics Engineers (1990), interoperability is "the ability of two or more systems or components to exchange information and to use the information that has been exchanged." Interoperability goes far beyond sharing information as stand-alone databases, rather, it means that information is readily usable by participating organizations to integrate into systems and manipulate and use for their purposes.

As discussed earlier, many I&Rs possess both the interest and the desire to share information and to be able to actively use it to deliver services. However, most information sharing has not been at the interoperability level. That is, information may have had to be keyed in again or simply stand alone, alongside, but not part of another I&R's database. It has been time-consuming, expensive, and frequently impossible to combine data from different I&R programs to create larger, integrated, and more useful data sets or to join I&R data with information from other sources.

In the past, the most straightforward solution has been for organizations to agree to use the same software product. Yet, using the same software is not a viable solution for many organizations that wish to share information. Organizations may be unwilling or unable to use the same product because of capital and training investments, because of the need for different features, because of organizational concerns, and a variety of other reasons.

This article provides an overview of data sharing standards, discusses sharing processes that are needed to supplement meaningful data sharing, and presents future directions for improving interoperability.

Step One: Creating Translation Standards

Across many fields, interoperability has been enhanced by data exchange standards that guide users and vendors in design and implementation. Without data standards that are recognized nationally, it is difficult to efficiently share resource information—particularly when I&Rs use different software products. Several years ago, the Michigan Association of United Ways produced a standards document that identified the most significant information (e.g., agency name, contact, taxonomy code, etc) that would be crucial for I&Rs to share. In 2002, the Alliance of Information and Referral Systems (AIRS) Board of Directors approved the Michigan document as the standard for interoperability of database information.

The standard uses a technology called XML (eXtensible Markup Language). While it's not necessary to understand what XML technology is and does in order to appreciate its benefits, it can be helpful to have a sense of how it works in order to understand what it does with data from I&R databases.

XML is a internationally-recognized tagging structure lets users create their own standards (or *language*) for how information is categorized with a document or database. XML has a solid foundation in computer technology as it is written in SGML, the international standard metalanguage for text markup systems. A good resource about XML is the *XML Schema Primer* at www.w3c.org/TR/xmlschema-0/

Unlike its better-known cousin HTML (HyperText Markup Language), that uses fixed tagging conventions to specify how Web pages are rendered by browsers, XML does not have fixed tagging conventions. Instead, users create their own tagging names and definitions. This makes XML usable for a vast array of applications (from medical records to libraries to I&R). The "dictionary and grammar" of any particular group's tagging conventions is called their XML Schema. In essence, XML Schema (*Schema* is both the singular and plural form of the noun) create "buckets" for information. Each bucket has specific information that should be included in it and may even have constraints, such as length or pick lists. The way information is tagged identifies which bucket or buckets that particular piece of information belongs to in a database.[2]

Ideally, the mapping of information to various buckets is done behind the scenes by the software being used by an I&R. Vendors use the schema to ensure that users can export data from their databases to fill the buckets. In the example of I&Rs, one of the XML "buckets" is reserved for free text information about a service's proximity to public transportation. Vendors then make sure that they have a field in their databases that is available for users to include this information. Another bucket is reserved for the state that a service agency is located. Vendors make sure that state information is separable from ZIP code, for example.

---

[2] One way of grasping a fundamental difference between HTML and XML is understanding that the purpose of HTML is to control the *display of data* (such as font face, size, and special effects or the placement of graphic images or the playing of audio files) while the purpose of XML is *to describe relationships between elements of the data*. HTML has to rely on highly standardized conventions so that individuals can use a spectrum of browsers to properly render Web pages, but XML has to allow authors the latitude to describe all sorts of relationships on the fly.

It soon became clear that, although the AIRS-approved standard provided a foundation for sharing, further refinement was needed. The standard needed greater specificity in order to guide vendor compliance. During 2003 and 2004, AIRS convened a workgroup to:

1) further refine and define additional "fields" needed to capture a greater breadth and depth of information that I&Rs would likely want to share; and
2) provide greater guidance to vendors by taking advantage of all the functionality offered through current programming technology.

The workgroup included participants from across the United States and Canada, including representatives from the initial Michigan project. The group reviewed the schema and recommended a number of changes. The changes included adding new information as well as removing information, re-titling information to more precisely reflect content and practice, and creating specific criteria for information when possible. For example, one change created designations for phone types (e.g. "voice" or "pager").

The AIRS Board approved the workgroup recommendations to upgrade the AIRSXML standards and asked the workgroup to continue to refine the standard. The AIRS Board established the workgroup as a subcommittee of the Technology Committee. Additionally, the Board determined that “generational” changes to the standard would be made no more frequently than annually, but that minor versioning could occur throughout the year. By late Summer 2004, AIRSXML 2.00 was released to vendors.

Since that version, several minor changes have been made to the versioning. These minor changes have been in response to vendor identification of implementation issues and minor error in the standard. The current version, AIRSXML 2.03, is available in the Library section of the AIRS website: www.airs.org/pub/pub_library.asp.

Vendors have been involved in the development of the AIRSXML standards since the inception of the idea several years ago. They were quick to recognize interoperability as a selling point that could change traditional models of customer retention. Vendors who incorporate AIRS XML-compliant exporting and importing capability into their products will be at a competitive advantage because that capability gives I&Rs more independence in choosing a software product that best meets their need, regardless of the product used by other I&Rs with whom they collaborate. Vendors have played an important role in providing information and feedback about the initial versions of the XML schema and, most recently, in identifying fixes to the current version (2.03) as they have been working to develop XML compliance functionality to their products. At the time of this writing, XML functionality is not yet commercially available as a component of any of the major I&R vendors’ products. However, it is expected that several vendors will have these tools available in 2005. When this functionality is available in vendor products, I&Rs will be able to easily create data exports (sometimes called *data dumps*) that may be shared with other I&Rs.

Step Two: Standardizing Processes

Although the ability to easily create export files will be an important technological innovation, that technology is only one step in a three step process. As described, the XML Schema creates "buckets," however, standardizing *processes* between collaborating I&Rs ensures that the information that fills each of the buckets may be meaningfully used. Although most I&R programs maintain roughly the same types of information, their databases are dissimilar in important ways. Different I&R programs may have established unique conventions for inputting and organizing data.

One of the most important conventions for inputting and organizing resource information is through the adoption of a taxonomy for services. A taxonomy is a system of classifying information. For I&Rs it would be a system for classifying the kinds of services that are offered. Many I&Rs have adopted the AIRS/INFO Line Taxonomy of Health and Human Services (Information and Referral Federation of Los Angeles County, Inc., 2004). This taxonomy (see www.211taxonomy.org/) is a hierarchically-organized structure of over 7,100 terms that define a wide range of health and human services. In order for I&Rs to most effectively share information about resources, there should be agreement, not only on the coding or taxonomy to be used, but also on the minimum depth and use of that taxonomy. For example, one I&R may choose to keep very precise information about emergency food providers and has internally decided to code to the fourth level (terms like *Food Pantries* or *Soup Kitchens*). If its collaborating I&R also maintains non-duplicative information about emergency food providers resources, but indexes it at the third level term *Emergency Food*, the indexing will be too general for use by the former organization. Thus, levels of specificity and the means for coding them should be set agreed upon by collaborating I&Rs. This does not preclude I&R from keeping more precise information, if they choose to because many software products enable more precise taxonomy coding to be "rolled-up" to more general categories).

Partnering I&Rs should also make decisions about how target populations for services, modes of delivery for services, and facilities are reflected in their databases. The AIRS/INFO LINE Taxonomy, for example, contains not just a structure for classifying services, but also gives organizations the ability to use target population codes (for services that are available to populations of specific interest to an I&R—rather than indexing a service as merely *Disease/Disability Information*, for instance, it could be indexed as *Disease/Disability Information* targeted at *Muscular Dystrophy*),[3] modalities (whether a specific service like *Divorce Assistance* is provided via *Advocacy*, *Legal Representation*, or some other means), and facilities (using common facility-type terms like *Hospitals* and *Public Libraries* often circumvents the need to index the specific services provided by an organization). See Bruni (2000) and Sales (2003) for excellent overviews on decision making strategies in using the AIRS/INFO LINE Taxonomy.

[3] For an in-depth examination of issues involved in using target terms, see Diane Gatto and Cathleen Kelly, "Indexing Using Target Population Terms in the AIRS/INFOLINE Taxonomy" elsewhere in this volume.

Even when collaborating I&Rs observe the AIRS XML, there are many fields in I&R databases that are fairly unrestricted. For example, geographic catchments of services may be coded by as ZIP codes, counties, neighborhoods, census tracks and so on. Other examples are ages, hours of operation, and physical accessibility, to name a few. Collaborating databases should also examine their inclusion/exclusion policies, so that collaborating partners are receiving only the kind of resources they want. Standards should be set for how much information about any resource will be shared. Finally, partners will want to review issues such as the frequency and means of maintaining resource information.

Exporting and Managing the Merged Data

After the technological ability to export data and standards for indexing are established, the final step to interoperability is the exporting and merging of data. Some I&Rs may be exporting data to a collective database on a central repository (i.e., computer server) for coordination, others may be importing data into their own organizational database on a regular basis. Whatever its form or scope, the most successful data set merging takes advantage of automated processes and minimizes manual interventions. Automated data conversion rules and routines should enable databases to define how missing or mismatched data should be mapped into a centralized database. There will be missing data because different software programs contain different data elements, indexing systems, and ways of structuring their data. Data will be mismatched because vendors, even with compliant exports, may continue to offer divergent selection choices within fields.

In October 2003, the University of Nebraska Public Policy Center along with numerous collaborating partners, was awarded a two-year U.S. Department of Commerce Technology Opportunities Program grant (#31-60-I03013) to develop a technological tool that will prepare data and create a new consolidated data set. For example, the tool will: perform quality checks (compliance with the XSD) and de-duplication (creates hierarchies for accepting changes, in the way a personal digital assistant synchronizes with a computer); confirm addresses and create precise geographical coding (through a Web service) to enable geographically-based searching; and perform basic formatting and spelling checks. The new data set is scrubbed and ready for export into a centralized database and/or available for web serving.

When completed, the technological tool will be made available for others across the nation who are interested in creating databases from diverse data sets. To ensure the product will be affordable to others, the tool employs open source software whenever possible to meet the needs of the system architecture. The tool, therefore, will be usable by entities with limited capital resources. Some of the open source solutions being used include packages from the Apache Project (www.apache.org): the Apache Web server, Xalan and Xerces XML packages, and the Jakarta Struts application framework. The JBOSS (www.jboss.org) application server will provide Java J2EE compliant services. The relational database being used is PostgreSQL (www.postgresql.org). All of this software runs on the FreeBSD and Linux open source operating systems. The one

exception to the open source approach is the de-duplication process. A Web-based service will be used to de-duplicate and geo-code agencies through the implementation of address standardization.

Future Directions

There are some promising efforts underway to achieve database interoperability in I&R; however, the field still is really only in the beginning stages of achieving stable and mature systems for sharing information. Although there are no processes underway for certification at this time, certification of products to ensure compliance with the AIRS XML will be a future important advance. Certification is necessary to make product choice more transparent and of lower risk for I&Rs. Certification would enable vendors to represent their product as credibly compliant with the standard.

Vendors and their products' features will be a key element in ensuring that data interoperability is accessible to the typical I&R. Products should enable easy, user-friendly tools for sharing. For example, software products should allow I&Rs to flag the desired taxonomy level or even prevent indexing to "unapproved" levels. Software products should enable I&Rs to "flag" those resources that they do not want included in data exports. Products should provide an easy and user-friendly means of creating compliant data exports. Automated processes for "roundtripping" the information (i.e., enabling I&Rs to send updates and receive back and synchronize the entire coordinated database system to their local systems) would enable I&Rs more comprehensive and customizable access to the entire shared data.

As interoperable databases are developed, they may naturally serve as the information backbone for 2-1-1 call centers and other regional collaborations. Publicly accessible interoperable databases also will serve as ideal mechanisms for Web-based searching for community resources. Attention should be given to creating a Web site that may be intuitively searchable by someone with minimal knowledge of the health and human services field, as well as efficiently searchable by health and human services professionals. Several different means of searching may be offered including: searching by key word, searching by drilling down into the taxonomy, searching via alphabetical listings, and searching by finding matching words. Innovations such as enabling users to leave breadcrumb paths of where they have been to reach a result and the ability to maintain "favorites" lists should also be explored. Of course, all searching should allow users to create geographic filters to specific service locations. Geographic filters may employ ZIP codes, mileage distances from a specified location (e.g., home address), counties, or regions. In Nebraska, a new website is attempting to incorporate these type of features (see www.ne211.org).

As I&Rs have the tools to share information more easily more people will know about the services that are available to them in a more accessible and efficient manner. The field stands at the cusp of realizing exciting new opportunities for true collaborative processes.

REFERENCES

211 Assists Toronto Through Blackout. (2003, August 15). Canada NewsWire Group. Retrieved February 23, 2005 from http://www.newswire.ca/en/releases/archive/August2003/15/c2012.html

Agosta, J., & Melda, K. (1995). Supporting families who provide care at home for children with disabilities. Exceptional Children, 62(3), 271-282.

Alliance of Information and Referral Systems (October 2002). Standards for professional information and referral (4th ed.). Retrieved February 11, 2005 from www.airs.org/downloads/NewStandardsforweb10-02.pdf

Bruni, M.G. (2000). Indexing with the AIRS/INFO LINE Taxonomy of Human Services. Information and Referral, 22, 83-109.

California gets new 211 service. (2005, February 11). TheDigest.Com. Retrieved February 24, 2005 from http://www.thedigest.com/articles/173/10.html

Lee, E. C. (2004). "Hurricane Chronicles From 2-1-1 Community Resources in Orlando: August 13 - September 09, 2004." (Available from El Cabrel Lee, Director 2-1-1, United Way Community Services, 1212 Griswold Street, Detroit, MI 48226)

Flora, C. B., Flora, J. L., Spears, J. D., & Swanson, L. E. (1992). Rural communities: Legacy & change. Boulder, CO: Westview Press.

Information and Referral Federation of Los Angeles County, Inc. (2004). A taxonomy of human services: a conceptual framework with standardized terminology and definitions for the field (Rev. ed.). San Gabriel, CA:Author.

Institute of Electrical and Electronics Engineers. (1990). IEEE standard computer dictionary: A compilation of IEEE standard computer glossaries. New York:Author.

Koenning, G. M., Benjamin, J. E., Todaro, A. W., Warren, R. W., & Burns, M. L. (1995). Bridging the 'med-ed gap' for students with special health care needs: A model school liaison program. Journal of School Health, 65(6), 207-212.

Levinson, R. W. (2002). New routes to human services: Information and referral. New York: Springer Press.

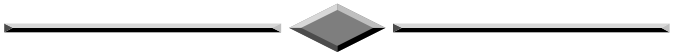
Mathematica Policy Research. (2002). The evaluation of welfare reform in Iowa: Final impact report (MPR Reference No. 8217-125 & 530). Washington, DC: Author.

Sales, G. (2003). An orientation to the structure and contents of the AIRS/INFO LINE Taxonomy. Information and Referral, 25, 1-41.[4]

United Way of America. (2004). Florida United Ways, Florida 2-1-1 network launch coordinated community action plan for Hurricane Charley response. Retrieved February 23, 2005, from http://national.unitedway.org/files/pdf/press_releases/Hurricane_Charley_Relief_Fund_Presrel.pdf

United Way of Metropolitan Atlanta. (1999). Who can help me? How can I get involved? [Brochure]. Atlanta: Author.

University of Nebraska Public Policy Center. 2000). Final 211 report: Survey of existing I & R services and a Nebraska 211 system cost/benefit analysis. Lincoln: Author.



Nancy Shank is the Associate Director of the University of Nebraska Public Policy Center. She is the principal investigator for a U.S. Department of Commerce Technologies Opportunities Program grant (#31-60-I03013) that joins the Center with national and state partners to develop the technological link between propriety software and XML standards to enable interoperable information and referral resource sharing. She co-authored a Nebraska study that surveyed existing information and referral services and created a cost/benefit analysis that has served as a model for other states. Her research interests include database interoperability for information sharing between organizations, information and referral, organizational systems and behavior, social capital, and human services systems

University of Nebraska Public Policy Center
121 S. 13th Street, Suite 303
Lincoln, NE 68588-0228
402.472.5687 *(voice)*
402.472.5679 *(fax)*
nshank@nebraska.edu *(e-mail)*

[4] A revised version of this article appears in this volume.