

VISTA Community Meeting
Day Three: January 16, 2011
Morning Session

Topic: PATIENT DEMOGRAPHICS AND PORTABILITY OF PATIENT DATA

Presenter: Chris Richardson

Building a test database out of a live database has traditionally been a serial process, doing one file after another. Mr. Richardson has found a way of turning it into a parallel process, so it takes less time to convert.

Names of patients are changed with a middle name of "pseudo." First and last names are changed to other names beginning with the same letter. Mr. Richardson's list of test names comes from lists of marathon results from various parts of the country, available on the internet.

In addition, Mr. Richardson is going through the TIU notes and removing names from that, and converting any numeric data into lowercase "n"s. Words not in the safe dictionary are replaced with X's. In this way, the volume of text remains roughly the same, but the file is redacted. It's been working well.

Internationalization: Mr. Richardson began working on this in 2004. The dialog file can hold multiple languages in the same structure. However, VISTA needs to be instrumented to reflect a call to the dialog file for each user's language. In this way, loading up the dialog file also updates the instrumented code.

It took about 45 minutes to build parallel directories of each VISTA routine. The dialog file was also expanded from 1900 entries to 165,000 entries. The goal was to load the English version into the dialog file, then extract the phrases from the file and put them into spreadsheets. Then, people interested in seeing their language used in VISTA need only to muster about 500 translators to translate the phrases in the spreadsheets. This is about an 85% solution, but that is probably good enough for most purposes.

Q: 500 translators for all the languages?

A: No, for each language. It depends on how fast you want it. If you want it real fast, you get 500. If you have 20, it's going to take longer.

Q: How long?

A: A couple of weeks. It's not a terribly onerous problem. And if you only want a subset, you could do it that way.

This ties into Meaningful Use (MU) with preferred language. Mr. Richardson posed an example where the doctor speaks French, but can also read German and Italian. If something has only been translated into Italian, the system can show it to him, if it has Italian as one of the doctor's secondary languages.

Mr. Richardson stressed that there isn't yet a solution such as Babel Fish for VISTA, although there may be one in future. But getting the idioms right is tricky. In English we miss the bus. In Spanish, the bus runs away from us. These things need to be taken into account for true internationalization.

The beauty of this [spreadsheet] approach is that there is always a directory containing the baseline

VISTA code. This is where any patches and upgrades can be applied; the updated system can then be run through the internationalization code again to generate the localized environment. This avoids what happened in Egypt and some other countries when they changed the code from English, and were not able to install subsequent patches.

Q: There's two things we're talking about, for clarity: Capturing the patient's preferred language, and capturing the preferred language for the provider as an enhancement.

A: It's not so much a preferred language, as a list of languages that the provider can accept. We can have a tree of different languages.

Q: But that second part is not needed for meaningful use?

A: No, it's not. Although it could still be useful; if you have a patient that only speaks a certain language, to know if there's somebody at the hospital who also speaks that language.

Mr. Richardson went on to discuss patient data. VA is running regional data centers. That can be a problem; when one single point of failure can shut down 18 hospitals. That kind of thing has happened, but not often, because VA has built redundancy into these systems. When such a shutdown happens, however, it interrupts a great many processes: doctors and nurses at the various hospitals all communicate with the server at the data center. VISTA wasn't really designed to be run this way; it was designed to be run locally.

A better design would be for the communication to happen locally, and the completed transactions be sent to the regional data center. That way, if there is a problem at the data center, the transactions can sit and wait before going to get processed; they are not time-sensitive. The patient data will eventually get updated. In addition, a national data center could sit atop the regional centers, also receiving transaction data and keeping the system updated nationally.

Mr. Richardson used the example of how to rebuild a hospital like New Orleans. Data from the New Orleans VA was pulled twice in response to hurricane Katrina. An initial extraction was put on tapes and brought to Houston before the actual hurricane hit. The VA facility in Baton Rouge was able to keep operating as long as the New Orleans facility had power, even though there was nobody there onsite in New Orleans. There was a second backup when the power failed and the system finally shut down, so that the Baton Rouge data was not lost.

A national repository as outlined above would enable VA to rebuild in a situation like that without having to physically go to the site.

Mr. Richardson has been working on VISTA for one, which is a way of generating a VISTA environment containing the data for a single patient. This is done using a recursive extraction of all files associated with that patient. That information can then be clustered into a single VISTA environment containing detailed information of all the patient's interactions with the hospital. This environment can be put on a live CD or DVD that the patient can carry with them. A doctor then only needs a PC to see the patient's history.

This technique can also be used to accession a patient into an existing environment. The only requirement is a PC linked to the server; it can negotiate the data dictionary. The operating system is on the CD; it will run regardless of the operating system on the machine. The data dictionary can determine which data can be brought in without modification and which data needs to go in an exceptions file. If there were a national data repository, it would be a way to rebuild a hospital from scratch, and be as up-to-date as the last transaction that came through.

Another important issue in moving patient data is the current time zone. If the computer is in one time zone, and the users are in three others, what time should VISTA display? It should be the time zone of the user; whoever is looking at the display. Sometimes there are users being transferred from DOD, they're in the Middle East, and they get transferred to Landstuhl, and then they get on a plane to Walter Reed. Middle East to Germany is not more than one or two time zones. But once they get on the plane to Walter Reed, the recorded time zone is GMT (per DOD policy). And then once he gets to Walter Reed, he's in a different time zone. So whether the patient can have another shot of morphine depends on when he was last given one somewhere over the Atlantic.

Mr. Richardson would like to see the time stamp expanded to contain more data, including time zone. VISTA could then calculate how to present the data to the user.

Q: Couldn't you include in the format, a time zone, so that everything would be in GMT?

A: You could, but I really think you want to do the conversion at the time the data is displayed. You're losing information if you've already converted it.

Q: If you have it in GMT plus time zone, it doesn't matter what time zone you're in.

A: Right now, it isn't being recorded. I'm saying let's add it to the time stamp so we always have it. Not only that. Right now, in areas that convert to Daylight Saving Time, in the winter they shut their system down for an hour. That is so dumb. All we have to do is record the time zone, and then one symbol gets changed at the hour.

Q: But that could be transparent. Nobody needs to know that you've been moving time zones.

A: Absolutely. As long as you can identify the time zone you're in, you're fine. You don't have to knock the system down for an hour. I think that's nuts. I suggested this five years ago, and got shot down. So it's one of the things we're working on now.

Other forthcoming changes include changes to Laboratory and its files. These files have evolved over time, and some of them are difficult to figure out. And as time goes on, the equipment changes. Some of the old Fileman code was written to conform to stack limits of 7, which led to a lot of limitations. Modern implementations can handle stacks of a thousand or more. The code has been working, but now trying to figure out how to improve it is a challenge.

Q: Again, this gets back to common conventions across disciplines. There are standards development organizations working on conventions that relate to the architecture, the data structure, and data representation dimensions. Clinical Laboratory Standards Institute, and the IVD Industry Connectivity Consortium (IICC). They are involved with the industry to develop the conventions for communication with devices. These are all part of what you're talking about. Folks in WorldVistA need to know those contacts and have some input into those conventions, because they address these issues..

A: Yes, and we're very sensitive to a lot of those. For example, we've just used some standards to increase the language file from the 14 entries that we had to 273.

Q: International Organization for Standardization (ISO) 639? Or locale codes?

Greg Woodhouse interjected at this point to explain that there are two issues being conflated. One is the patient language preferences, for example translators and discharge instructions in the appropriate language. There are two important standards: ISO 639 for languages and Federal Information Processing Standard (FIPS) 10-4 for locales. There are 250-odd languages out of 400 or so, if you want to include artificial and extinct languages and so on. But you also have to play

nice with software. For example if you are printing labels on drugs, you have to have software that will do that. Certainly Java and other platforms make use of locale to determine how strings are printed. Or, there are resource bundles like the dialog file that keep track of different versions of the same text, and they could use locale. You need to distinguish the two components (language and locale); these are composite entities. The language code is a Alpha-2 code from ISO 639 and they use the Alpha-2 code from ISO 3166, which is the country code, for locales. The point is, when we're dealing with languages, we need to maintain both bits of information. A language spoken in different countries isn't always spoken the same way.

Q: How many languages are you suggesting be available for the patient's preferred language?

A: Right now about 270.

Mr. Woodhouse added that 270 is the number of active languages in ISO 3166.

Q: How would you record that?

A: In VISTA, there are three places it is stored: the Kernel site parameter file (which is the default for the whole system), the New Person file (proficiency is also recorded there as a multiple), and the Patient file.

Q: Where would it be entered?

A: It's associated with the New Person file. It's also associated with the patient.

Nancy Anthracite added that there are local questions that can be entered at registration, and that's where a patient's preferred language could be recorded.