



(U) Human-Language Technology in Your Future

FROM: Anita H. Kulman
Mission Area Director, HLT PMO (S23)
Run Date: 07/20/2006

(U//FOUO) *New tools for analysts to use are on the horizon...*

(U) Would you like to look into the future? Some visionaries have been able to predict things to come. Throughout the years, science fiction writers created images of a technological world which now actually exists. A Robert Heinlein character of the 1940's carried a small, personal, portable telephone, and now cell phones have become so routine that it's hard to remember a time that we did not walk around cradling them to our ears.

(U//FOUO) But can you say what the future holds for the analysts' work environment?

(U//FOUO) The Human Language Technology Program Management Office (HLT PMO) and the Office of the Senior Language Authority (SLA) can predict that more and better technology services will be available to more analysts to help them pinpoint the most valuable data within the text, voice or image messages that could otherwise bury them. **First, let's look at the HLT PMO and then into the HLT crystal ball to see what the analysts' desktop might contain tomorrow or in a year or two, or more.**

(U//FOUO) The HLT PMO is a relatively new organization that is currently gathering strength and momentum. It was created to focus research, development and deployment of HLT services on one unified path toward agency mission goals. The PMO, for the most part, funds other organizations to do the work of building and refining and delivering HLT services, but some work is done within its organization as well.

(U//FOUO) The PMO may see into the crystal ball, but its work is not spectacular magic -- like the smoke and mirrors that you may have witnessed in some technology demos, particularly by outside companies hungry to sell their products to us, where presenters promise miracles. These products often turn out not to work well in our SIGINT environment. The PMO team works hard to help analysts do their daily tasks more easily. It is searching for the sturdy, long-lasting methods that will bring and maintain solid technology to the entire NSA/CSS Enterprise.

(S//SI//REL) In the PMO's crystal ball today, we can see **nine HLT capabilities** that will become more accurate and more useful as well as more available:

Information Extraction:

IE technologies organize unstructured data using information from their content and provide a way to retrieve important elements of those data.

Language Identification:



SERIES: (U) HLT

1. Human-Language Technology in Your Future
2. [For Media Mining, the Future Is Now!](#)
3. [For Media Mining, the Future Is Now! \(conclusion\)](#)
4. ['Knowledge Discovery': Finding the Best Material](#)
5. [Human-Language Technology -- Everywhere](#)
6. [Dealing With a 'Tsunami' of Intercept](#)
7. [Building Human-Language Technology](#)
8. [Strangers in a Strange Land?](#)

LID Labels written text or voice messages by language of interest.
(*Are they speaking in French or English? Is this document in Arabic?*)

Speech-to-Text:

STT provides automatic transcriptions of voice intercept in a written form of the foreign language. (*Chinese voice into Chinese text in characters.*)

Machine Translation:

MT provides automatic translation of foreign language text into English. (*Chinese text into English text.*)

Optical Character Recognition:

OCR transforms an image of a text, in the original language, to a text that can be edited and searched for pertinent words or information. (*Transforming a fax of a letter into a Word document.*)

Speaker Identification:

SID locates and labels voice messages where a speaker of interest is talking. (*Is that the terrorist we've been following? Is that Osama bin Laden?*)

Message Categorization:

This identifies the type of text or voice message, including the identification of the topic, of the genre, etc. (*Is this a diplomatic message or an email?*).

Information Retrieval / Question Answering:

Information Retrieval -- IR allows analysts to pull important bits of data from databases that contain information of potential intelligence value.

Question Answering -- Allows analysts to ask system natural language questions and to get specific answers, not a list of articles in which the answers may be found if you look for it. (*Where is Islamabad? Who is Musharaf?*)



(U//FOUO) Some of these capabilities are already available to analysts, while others are still in the first stages of research and may not reach the desktop for a few years. The HLT PMO is working on all of them, at differing levels of resources and funding, and with different predictions of delivery of the capabilities to operational offices, within its five major **Strategic Thrusts:**

- Media Mining;
- Knowledge Discovery;
- HLT throughout the Enterprise;
- High Speed-High Volume; and
- HLT for Information Sharing.

Look for articles on these Thrusts and other HLT components, how they relate to the capabilities just mentioned, and how they will help you.

(U//FOUO) For more information about these capabilities, please contact the HLT PMO office (" [go HLT](#) " or call [REDACTED]s).

"(U//FOUO) SIDtoday articles may not be republished or reposted outside NSANet

without the consent of S0121 ([DL sid comms](#))."

DYNAMIC PAGE -- HIGHEST POSSIBLE CLASSIFICATION IS
TOP SECRET // SI / TK // REL TO USA AUS CAN GBR NZL
DERIVED FROM: NSA/CSSM 1-52, DATED 08 JAN 2007 DECLASSIFY ON: 20320108