

7-20-2009

Using caTIES as a case-finding tool in tissue repositories: system challenges and lessons learned

Micaiah Evans

B.W. Trevor Rohm

Frederick Schultz

Philip J. Kroth

Follow this and additional works at: <https://digitalrepository.unm.edu/hslic-posters-presentations>

Recommended Citation

Evans MH, Rohm BWT, Schultz FA, Kroth PJ. Using caTIES as a case-finding tool in tissue repositories: system challenges and lessons learned. Abstract and poster presented at the caBIG Annual Meeting; Washington, DC. May 20-22; 2009.

This Other is brought to you for free and open access by the Research and Scholarship at UNM Digital Repository. It has been accepted for inclusion in Posters and Presentations by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Using caTIES[®] as a case-finding tool in tissue repositories: system challenges and lessons learned



Micaiah H. Evans, MD, MPH,^{a,e} B. W. Trevor Rohm, MD, MS,^{c,e,f} Frederick A. Schultz, MA,^{b,c,e} and Philip J. Kroth, MD, MS^{c,d,e}

^aDepartment of Pathology, ^bHuman Tissue Repository, ^cHealth Sciences Library and Informatics Center, ^dDepartment of Medicine, ^eUniversity of New Mexico Health Sciences Center, Albuquerque, NM, ^fPresbyterian Medical Group, Albuquerque, NM

Introduction

- caTIES[®] = Cancer Text Information Extraction System
- caBIG[®]'s set of tools designed to streamline the tissue sample identification process
- Uses full text surgical pathology reports (SRPs) and natural language processing functionality
- De-identification is accomplished by software incorporated into caTIES[®] (De-ID[™])
- Facilitates researchers' access to tissue that is otherwise described only in uncoded text-based SPRs
- Reduces costs and delay associated with manual coding by human reviewers
- caTIES[®] is composed of multiple components
- Strict adherence to specifications of versions / configurations of these components is essential
- The installation and setup of caTIES[®] presents a number of unique challenges
- Support for installation and use of caTIES[®], as provided by caBIG[®] is inadequate
- Several key areas for improvement of caTIES[®] and caTIES[®] support are identified

Installation Process

- Four Phases of Installation:
 1. Introduction
 2. Information Gathering
 3. Installation
 4. Post Processing
- Ambiguous hardware requirements
- Further configuration of pre-installation software, Windows services management, and modification of configuration files required
- Working knowledge of IP addresses and ports is required
- Installation testing requires username and password provided only on installation instruction page.
- Installation is performed on "public" and "private" servers in parallel

Program Use / User's Manual

- Online User's Manual is for previous version of caTIES[®]
- Multiple levels of user accounts must be created, configured, and saved
- All work must be saved with "save" button despite lack of visual cues to the work flow and save processes
- Error messages are likely to appear
- Trouble-shooting the errors requires a working knowledge of the directory structure and log files
- No disaster / error recovery process: errors result in complete program failure

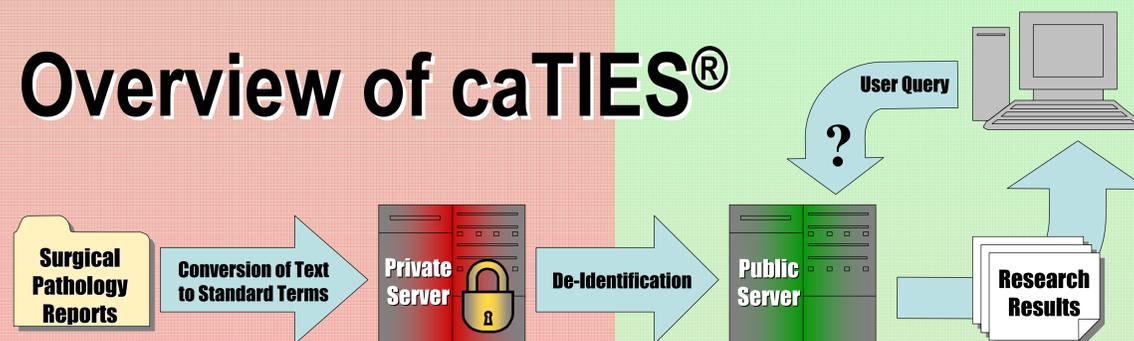
Help and Support

- Lack of help files
- Online user's manual is not updated to the current version and provides very few helpful specifics
- No explanation of HL7 encoding specifications or instructions for loading HL7 files provided in user's manual
 - Cerner Corporation has released *caBIG Data Extract Specifications* (2007) to facilitate use of caTIES[®] with Cerner's CoPathPlus[™]
 - No similar documents available to guide use of caTIES[®] with other surgical pathology information systems
- De-identification program (De-ID[™]) is not included with most the recent caBIG[®] distribution of caTIES[®]
- Sourceforge forum includes users' questions and user-posted answers, but is neither reliable nor exhaustive
- NCI and caBIG[®] forums do not address software incompatibilities, security issues, or HL7 specifications

Areas in Need of Further Development

- Preparation of single "installation package" that includes all required supporting components (especially De-ID[™])
- Updating of caTIES[®] to function with current versions of Java (JDK 1.6.14) and other supporting software
 - Most secure networks will require caTIES[®] to run only in a "demilitarized zone" due to dependence on outdated Java (JDK 1.5.1)
 - This isolation inhibits access of even the caTIES[®] public server by potentially interested researchers
- Development of a streamlined system for generation and transfer of security certificates
- Implementation of reliable disaster / error recovery measures and creation of a list of trouble shooting tips
- Periodically updated online user's manual corresponding to the current version of caTIES[®]
- Scheduled monitoring of, and response to the NCI and caBIG[®] forums by seasoned caTIES[®] experts
- Creation of an HL7 encoder that can adapt to the SPR format of any surgical pathology information system

Overview of caTIES[®]



Administrator

Researcher

Pre-Installation Process

- Testing and production environment must be configured prior to installation
- Pre-installation software requires working knowledge of supporting software:
 - Microsoft Windows Server
 - Java SDK 1.5+
 - Apache ANT 1.6+
 - Apache Tomcat Webserver 5.5+
 - MySQL 5.1+ Database Server
 - UMLS MMTX and NCI Metathesaurus
- Download and configuration of extensive pre-installation software is required
- Pre-installation programs must be of specific versions, without subsequent updates
 - The caBIG[®] distribution of caTIES[®] will not function with current version of Java (JDK 1.6.14); requires JDK 1.5.1
- File names created during pre-installation processes must be short (old MS-DOS type)
- Must use "/" rather than "\" in file paths
- Security certificate must be created, exported, and copied from one machine to the next