Academic Software Development

User Perspectives

Carmelo Gaudioso, MD, PhD

ROSWELL PARK CANCER INSTITUTE

ROSWELL PARK CANCER INSTITUTE

Agenda

- Background
- Reasons and considerations
- Technical support
- Software quality
- Derived benefits
- Desired improvements

ROSWELL PARK CANCER INSTITUTE

TIES: The Software

- An **NLP and IR system** for de-identifying, annotating, storing and retrieving pathology and radiology documents
- A system for **indexing research resources** (FFPE, FF, images) with document annotations
- A **GUI for querying** large repositories of annotated documents
- A platform to support **data and biospecimen sharing** among networks of cancer centers and other institutions
- Supported by NCI Grant R01 CA132672 and NCATS Grant 2ULRR024153

TIES Cancer Research Network

Member Institutions

- U Pitt Cancer Institute
- U Penn Abramson Cancer Center
- Roswell Park Cancer Institute
- Georgia Regents Cancer Center

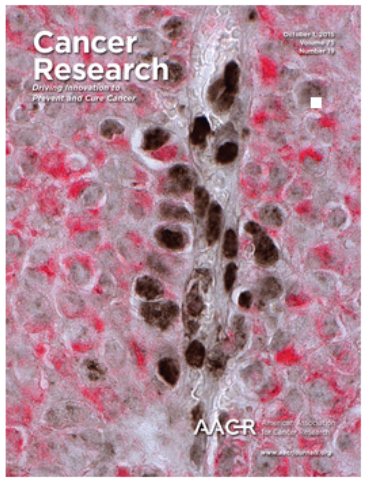
Network Trust Agreements

- Predetermined NHSR data
- Established common SOPs

Total Pathology Cases

- 5.8 million reports

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)



• View the new Impact Factor
• View the Most-Cited Articles of *Cancer Research*

Resource

Cancer Research

A Federated Network for Translational Cancer Research Using Clinical Data and Biospecimens

Rebecca S. Jacobson[1], Michael J. Becich[1], Roni J. Bollag[2], Girish Chavan[1], Julia Corrigan[1], Rajiv Dhir[1], Michael D. Feldman[3], Carmelo Gaudioso[4], Elizabeth Legowski[1], Nita J. Maihle[2], Kevin Mitchell[1], Monica Murphy[4], Mayurapriyan Sakthivel[4], Eugene Tseytlin[1], and JoEllen Weaver[3]

Abstract

Advances in cancer research and personalized medicine will require significant new bridging infrastructures, including more robust biorepositories that link human tissue to clinical phenotypes and outcomes. In order to meet that challenge, four cancer centers formed the Text Information Extraction System (TIES) policies, and procedures, enable regulatory compliance. The TIES Cancer Research Network now provides integrated access to investigators at all member institutions, where multiple investigator-driven pilot projects are underway. Examples of federated search across the network illustrate the potential impact on

Reasons and Considerations for Using TIES

Reasons

- Automatically extracts information from reports to identify eligible patients
- Create, store, retrieve, and share case sets
- Automatically link identified patients to available biospecimens
- Access to a larger data pool

Considerations

- Ease of deployment/use
- Data security
- Performance: accuracy & completeness
- Cost and required resources
- Value added to translational research efforts
- Supported, maintained

ROSWELL PARK CANCER INSTITUTE

Support and Help

- Extensive supporting materials
- Annual face 2 face meeting
- Monthly Executive Committee meetings
- Weekly teleconferences throughout the installation, configuration, and QA testing phase
- 2 - 3 releases per year
- Online support after every major release
- Sourceforge User Forum, responses within 24 hours

ROSWELL PARK CANCER INSTITUTE

Customizations

- Created additional de-ID scripts to loading process
- Added user regulatory admin role
- Added audits reporting feature to monitor and report TIES user activity
- Removed tissue ordering feature

ROSWELL PARK CANCER INSTITUTE

Software Quality

- Information
  - Precision
  - Recall
  - Structured and easily viewable
- System
  - Accessible
  - Flexible
  - Reliable
  - Querying across networks
  - Response time

ROSWELL PARK CANCER INSTITUTE

Software Benefits

- User friendly graphical interface
- Ability to do concept searching AND text searching in the same query
- Ability to formulate temporal queries
- Highlighting features for report review
- Direct access to de-identified clinical data to investigators
  - Predetermined NHSR status of the data
  - User interface: create, store, retrieve, and share case sets
- Access to a larger, interinstitutional data set, enhancing the ability to study rare diseases and promote collaboration

Desired Improvements

- TIES – LIMS Interface
  - Search for corresponding tissue samples in LIMS.
    - Get the case sets data from TIES central hub
    - Create an API
    - Write the case set information into the RPCI-TIES
- Admin's ability to search for cases using de-ID # and surgical pathology #
- Annotate pathology reports with molecular information

ROSWELL PARK CANCER INSTITUTE

Summary

- Outstanding user experience
- Extraordinary leadership and project management
- Great opportunity to participate in a global effort in support of translational research including rare diseases
- Significant capability to electronically identify eligible patients interrogating pathology reports across TCRN
- TIES is up and running with minimal local resources
- NHSR database with directed access to investigators
- Conducting pilot projects using TIES and TCRN
- Looking forward to future enhancements and extended ways of using TIES and the TCRN

ROSWELL PARK CANCER INSTITUTE

Thanks!

Questions?