



Integrative and Translational Cancer Research Informatics Requirements

Tahsin Kurc, PhD
Center for Comprehensive Informatics
Emory University





Outline

- Research Pattern Templates
- Research Pattern Examples
- Informatics Requirements of Pattern Templates
- cancer Biomedical Informatics Grid (caBIG)



Research Pattern Templates

- Represent common sets of principles and processes employed by studies investigating similar problems.
- A research study may have elements of multiple research pattern templates
- Examples
 - Prospective clinical research study
 - Multiscale Investigations that encompass genomics, epigenetics, (micro)anatomic structure and function
 - Coordinated Systems Level Attack on Focused Problem
 - Integrative Clinical Analyses
 - Secondary Data Analysis
 - Adaptive Image Guided Intervention

Saltz J, Kurc T, Hastings S, Langella S, Oster S, Ervin D, Sharma A, Pan T, Gurcan M, Permar J *et al*: e-Science, caGrid, and Translational Biomedical Research. *IEEE Computer* 2008, 41(11):58-66.





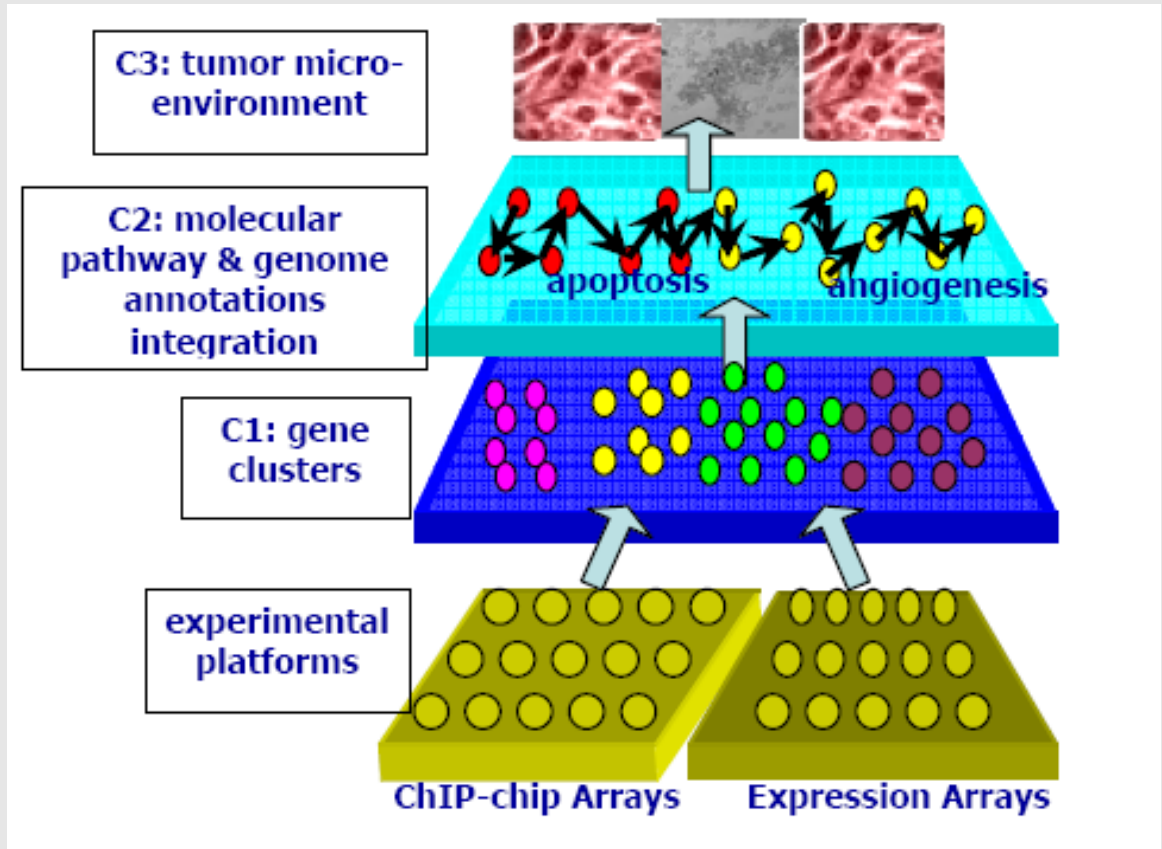
Research Pattern Template Examples

- Multiscale Investigations
 - Measure and quantify biomedical at multiple scales.
 - Experiments involving multiple types of microscopy imaging and high throughput genetic, genomic and epigenetic analyses.
 - Understand the interplay between anatomical structures, physiology and systems biology.
 - **Example: Tumor Microenvironment Characterization**
- Prospective Clinical Research Study
 - A group of patients are systematically tracked over a period of time.
 - Designed to explain risk factors for development and progression of disease and/or to assess the effects of various treatments.
 - **Example: Central review in Radiation Oncology**



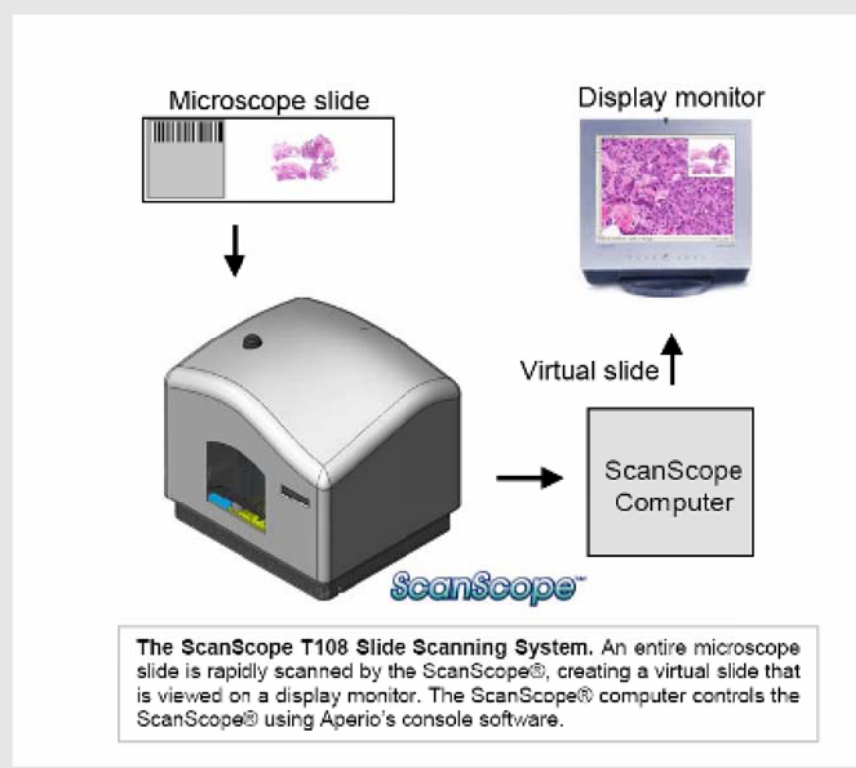
Tumor Microenvironment Research

- A tumor is an organ
- Structural and functional differentiation within tumor
- Molecular pathways are time and space dependent
- “Field effects” – gradient of genetic, epigenetic changes
- Anatomy, physiology, molecular biology of cancer

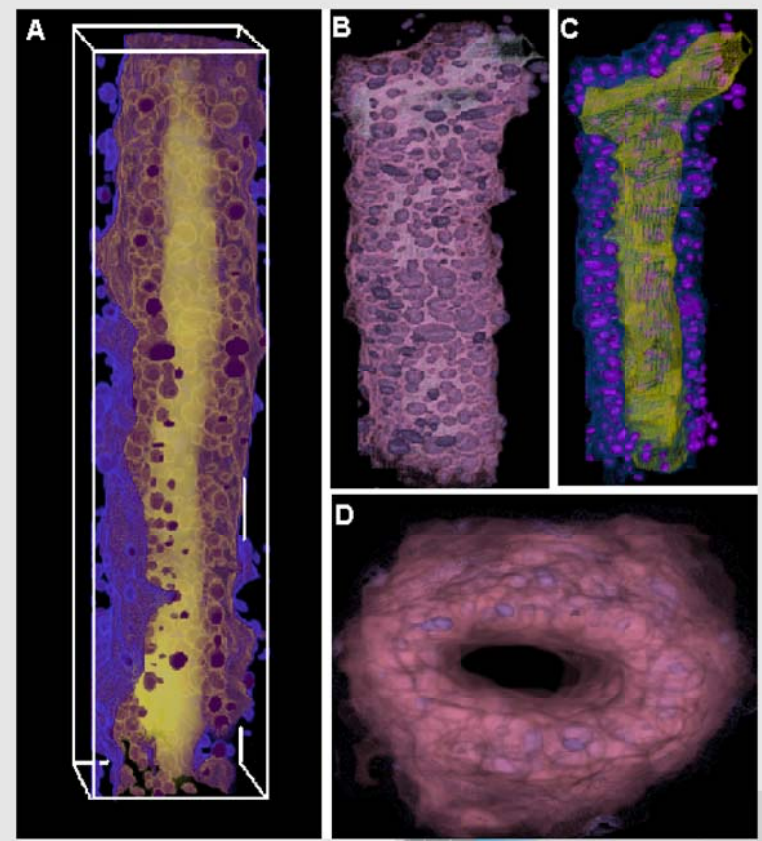


Tumor Microenvironment

Slide Scanning



Ducts



Imaging Team led by Raghu Marchiraju and Kun Huang OSU



Exploration of Datasets

- Molecular data, high-resolution imaging data
- Query data based on spatial characteristics
- Query data based on molecular characteristics
- Molecular and cellular information may be expressed in Ontologies
- Integrate spatial information with molecular information

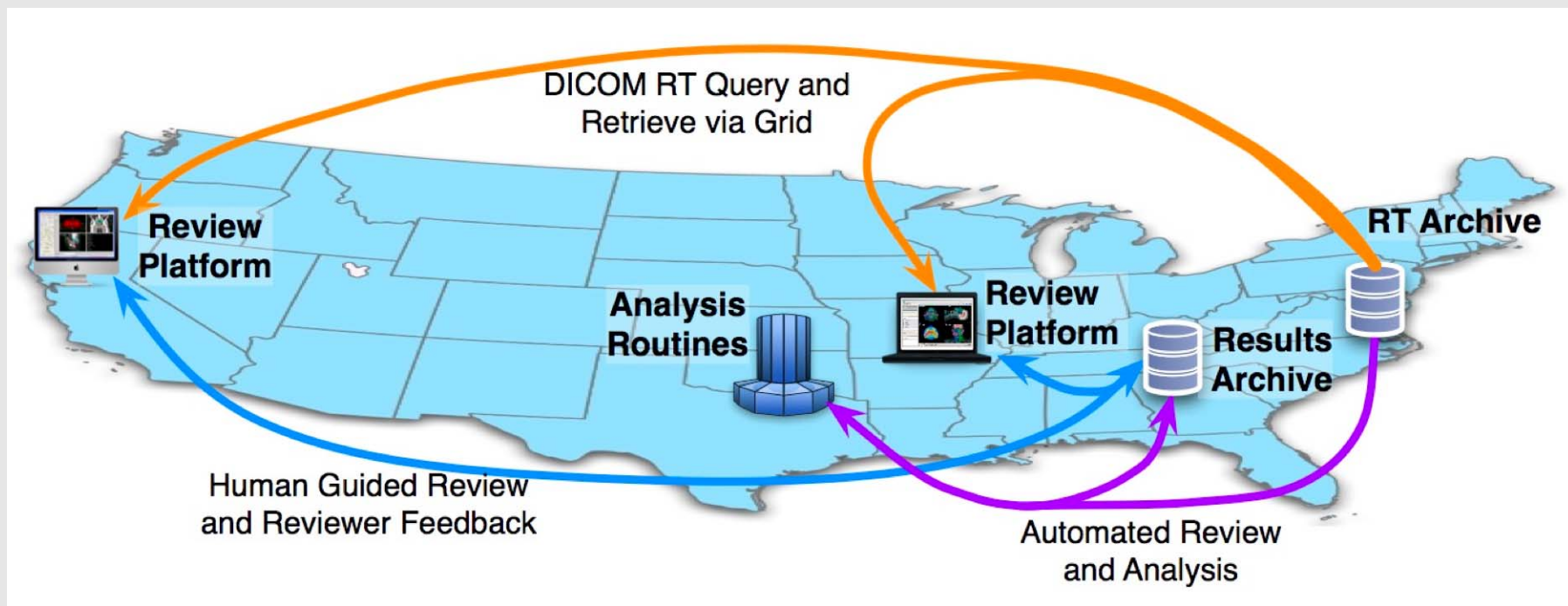


Central Review in Radiation Oncology

- Digital imagery
 - to define the tumor volume and outlining areas that receive radiation
 - to exclude the portions of healthy tissue that must be spared.
- High inter-observer variability amongst reviewers of image data.
- Central review of imaging objects.
 - Multiple expert radiologists at different institutions review image and clinical data, and an independent adjudicator incorporate their reviews into a consensus review.
- Expert reviewers at different institutions access images and annotate them.



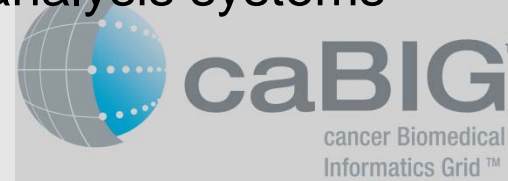
Prospective Clinical Research





Informatics Requirements

- Data management and analysis
 - Many different types of data
 - Support for clinical trials/studies such as scheduling of patient visits
- Semantics
 - Deep integration of many types of information to synthesize knowledge
 - Many formats, naming schemes developed in isolation
- Interoperability
 - Information drawn from home-brew and commercial/enterprise systems
 - Many different types of database systems and analysis systems





Informatics Requirements

- Large datasets and complex operations
 - Whole genome analyses, coordinated analysis of multiple types of molecular, high-resolution image data
 - Transferring large high-throughput or image datasets
- Resource Federation
 - Distributed resources
 - Discovery of what is available
 - Remote and coordinated use of resources
- Security
 - Authentication and access control
 - Distributed resources; crossing institutional boundaries
 - Compliance with regulatory and federal Guidelines



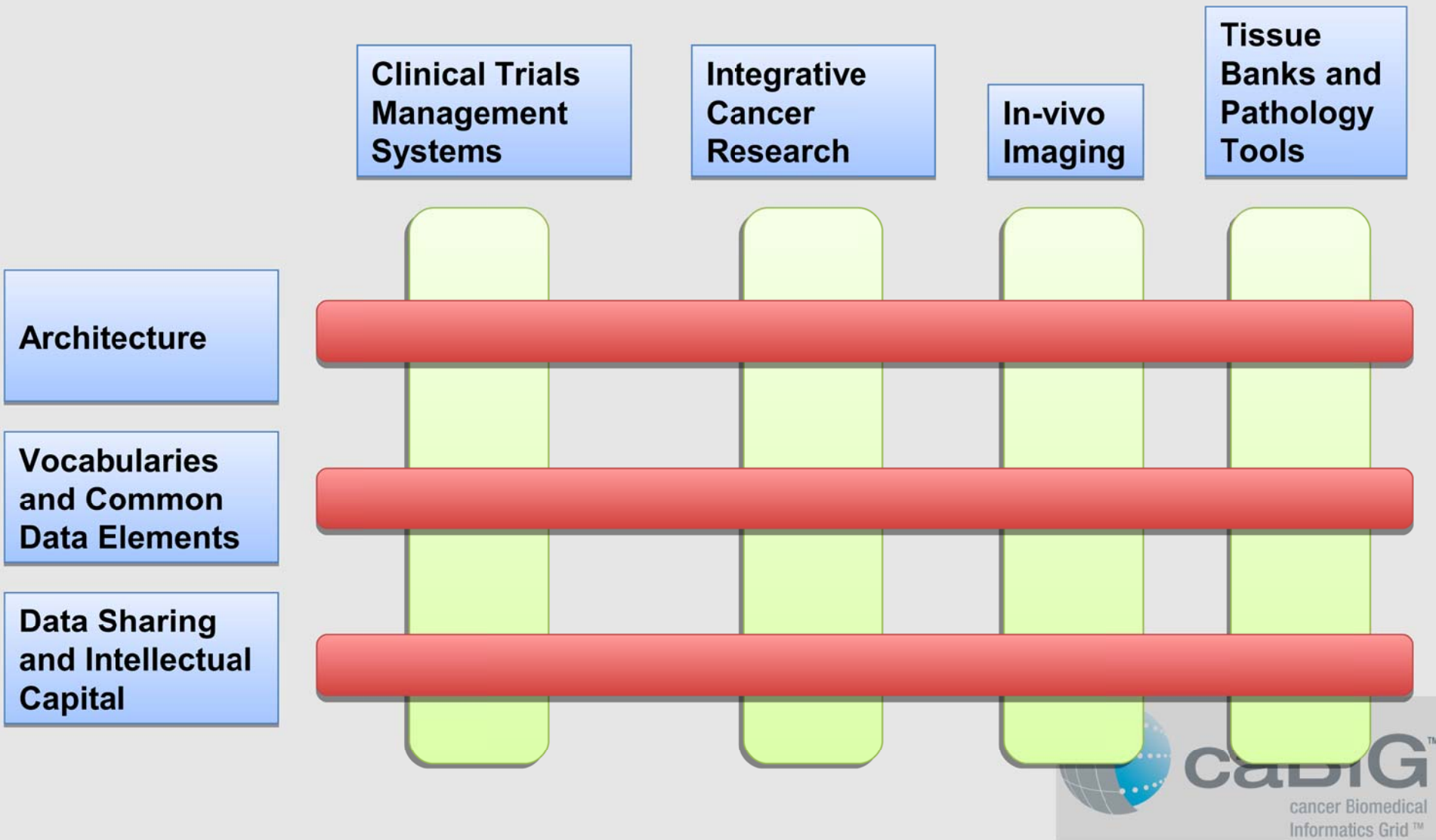


cancer Biomedical Informatics Grid

- Program funded by the NCI
 - Open to all, not just funded participants.
 - Over 190 institutions, 2000 people involved thus far.
- Overarching goal is to enhance cancer research significantly
- Create a voluntary network of research institutions, groups, and laboratories
 - Facilitate information, data, and tools sharing
- Develop common set of applications, tools, informatics standards, and infrastructure
 - Support for data management and data analysis
 - Interoperable systems



caBIG Organization





caBIG Tool and Application Examples

- Applications
 - caArray – microarray data management
 - caTissue – tissue information management
 - caAERS – adverse event reporting
 - PSC – patient study calendar
 - NBIA – MRI imaging archive
 - caGWAS – management of data in genome-wide association studies
- Infrastructure and Tools
 - caCORE SDK – application creation
 - caDSR and EVS (LexEVS) – common data elements and vocabularies
 - caGrid – Grid middleware for federation of resources





More Information

- Emory Center for Comprehensive Informatics
 - <http://www.cci.emory.edu>
- cancer Biomedical Informatics Grid
 - <https://cabig.nci.nih.gov/>

Funding

- The National Cancer Institute, National Institutes of Health under Contract No. HHSN261200800001E; by the NCI through Contracts 28X193, N01-CO-12400, 94995NBS23, and 85983CBS43; by Grant R01LM009239 from the National Library of Medicine; by Grant R24HL085343 from the National Heart, Lung, and Blood Institute; and by PHS Grant UL1RR025008 from the Clinical and Translational Science Award Program, NIH.
- The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.





Thank you.