National Cancer Institute

caBIG™ cancer Biomedical Informatics Grid™

Delivering caBIG™ Compatible Solutions

caBIG™ Industry Partners Meeting

Sue Dubman

National Cancer Institute

September 30, 2005

Natcher Conference Center
National Institutes of Health

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Approaches Taken by NCI and caBIG™ in Delivering caBIG™ Compatible Solutions

1) Modify/enhance existing vendor solution

2) Custom develop solution leveraging caCORE/caBIG™ tools where appropriate

3) Open standards middleware layer for integration

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: C3D – Oracle Clinical

Vendor solution made caBIG™ compatible

- NCI made C3D “bronze level” compatible early on
  - Integrated C3D with caDSR metadata via open source scripts and procedures
  - Provided external interfaces necessary for clinical trials management
  - Developed open source solution for CDUS and Theradex/CTMS transmissions
- NCI is working with Oracle to make application “silver-level” compliant
  - Oracle has committed to accept/process a standards-based (HL7 eDCI) CRF message
  - We are talking about provision of complete metadata APIs

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: C3D Made caBIG™ Compatible

C3D Extensions for “Bronze” Level Compatibility

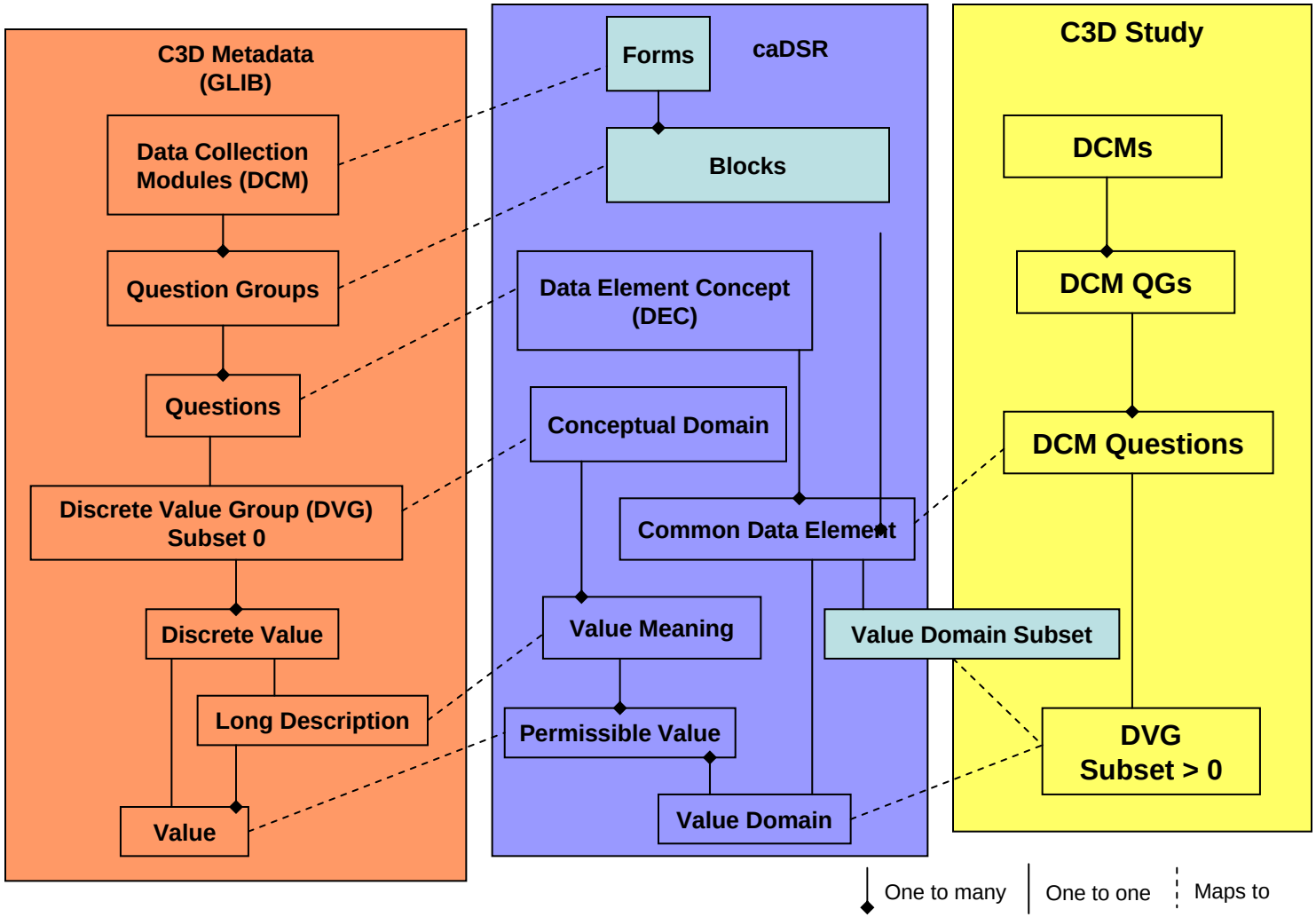
- CDE Loader
- Lab Loader
- CTMS Submission Extraction
- CDUS Submission Extraction
- Validation and Derivation Procedures
- CDE Compliance Tools
- Library of Template eCRFs
- Cross Study Reports
- New OC API for central patient registry

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: C3D Made caBIG™ Compatible

C3D Mapping to caDSR



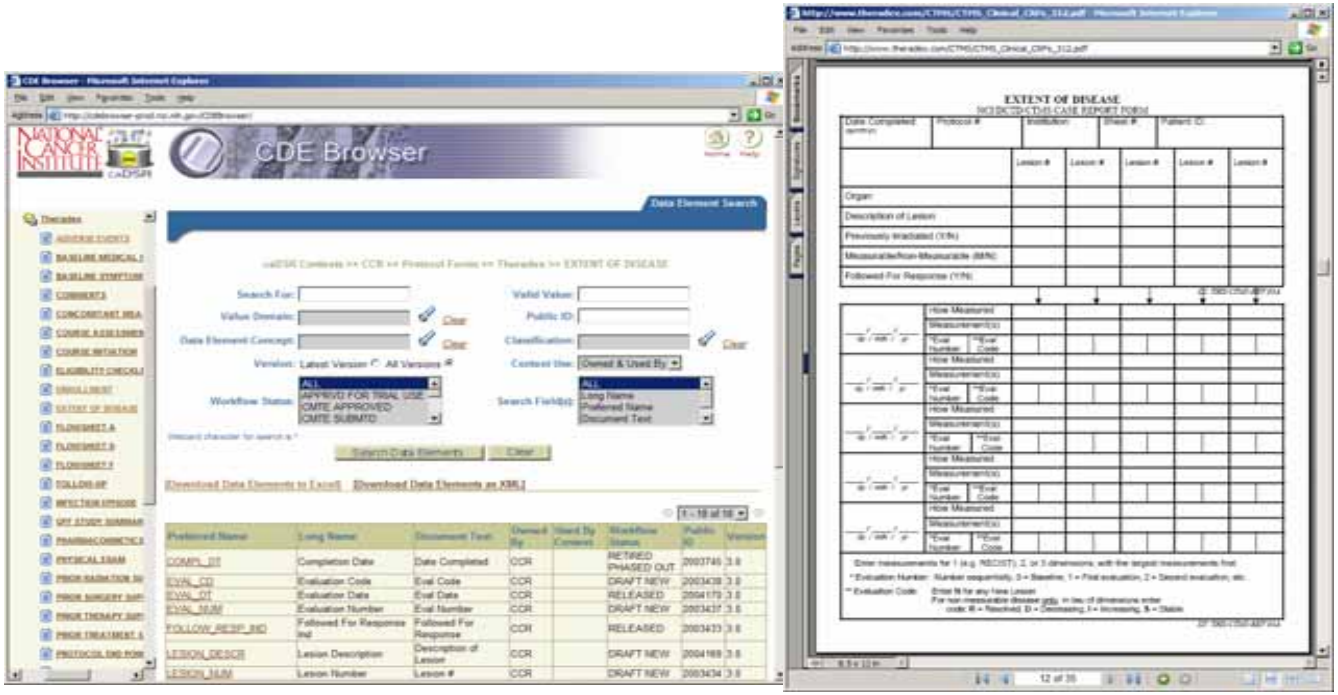
National Cancer Institute

caBIG cancer Biomedical Informatics Grid™

Case Study: C3D Made caBIG™ Compatible

Library of Template CRFs Built with CDEs

- Forms and CDEs defined in caDSR
- Template CRF layouts defined as PDFs



National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: C3D Made caBIG™ Compatible

Moving towards "Silver" Compatibility, Later "Gold"

- The "Silver" and Later "Gold" Vision for C3D includes:
  - C3D would receive HL7 v3 eDCI messages in define mood to load its metadata.
  - The HL7 v3 messages would, of course, be made up of common data elements and use standard terminology.
  - C3D would render screens for the collection of the data in a way that enables it to generate the same message in an event mood.
  - For electronic data collection C3D would accept HL7v3 event messages from any source to which it has registered interest and consent to receive consented patient clinical data.
  - C3D would also be able to natively generate HL7v3 messages for all its data and metadata.
  - C3D's repository would be based on HL7 v3 RIM
  - Expose C3D data through CTOM APIs

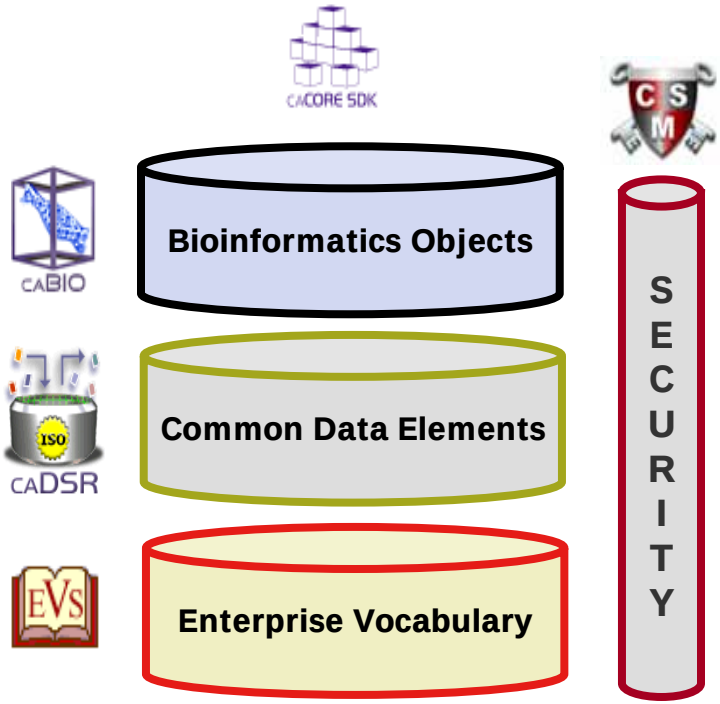
National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Custom caBIG™ Development

Leveraging Core NCI/caBIG™ Infrastructure

- Many NCI applications built on top of caCORE Infrastructure today including:
  - caArray
  - I3
  - Firebird
  - C3PR
  - CIAF
  - HL7 SDK
  - among others

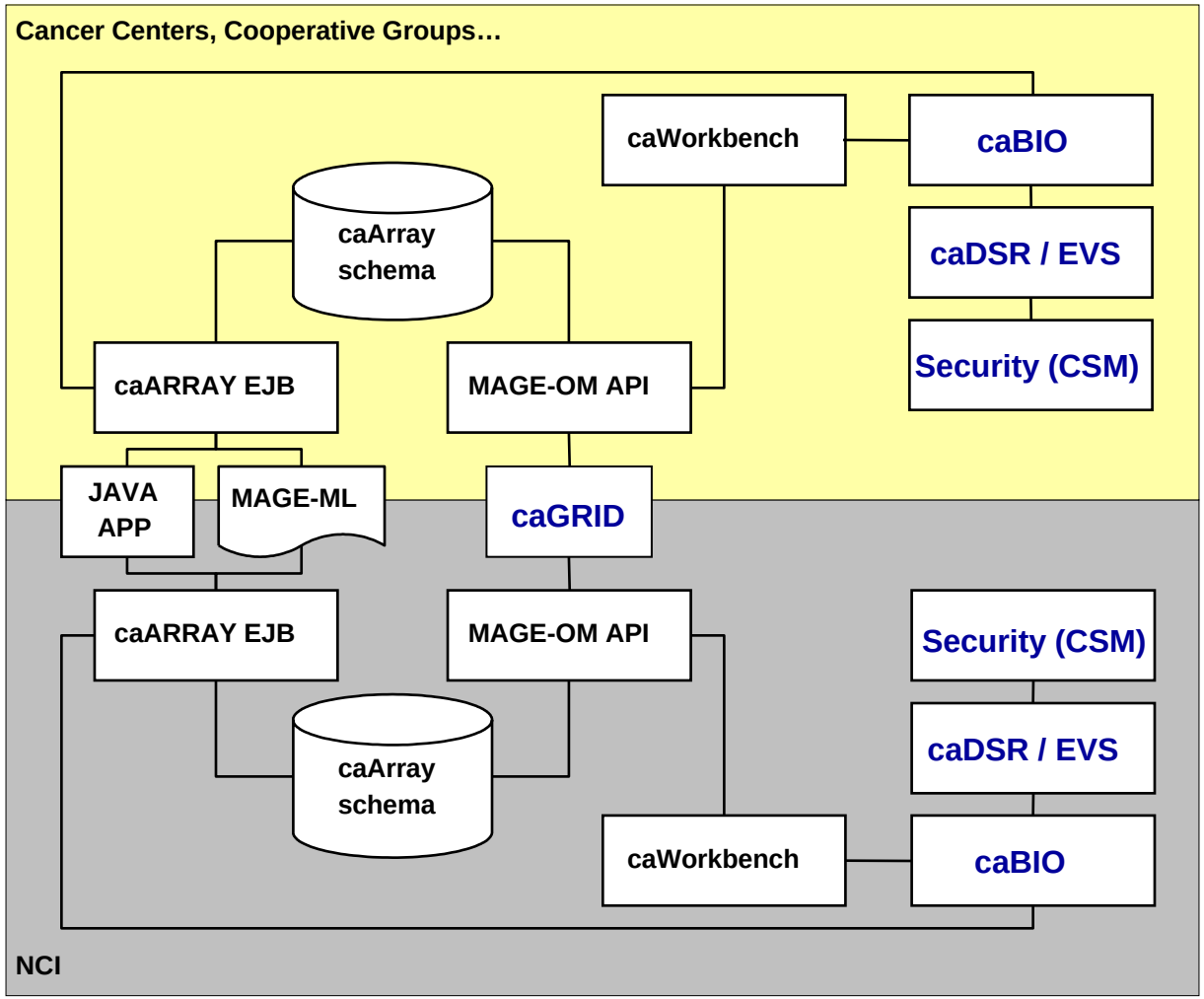


National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Custom caBIG™ Development

"Gold Level" caArray Configuration



National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Custom caBIG™ Development

caArray Compliance with caBIG™ Principles

- Open source
  - caArray source code and all other artifacts are available under an open source license
  - caArray utilizes only open source technologies such as MAGE-stk, JBoss, Xerces, Struts, Ant, OJB, etc
- Open development
  - All caArray development plans, artifacts are shared with the community
- Open Access
  - caArray allows submission and retrieval of all public data. The protection group/element concept in caArray allows researchers to preclude sharing of sensitive data
- Federated
  - caArray is one of several caGrid reference implementations

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Custom caBIG™ Development

caArray Compliance with caBIG™ Architecture

- caArray allows programmatic interface to its data via the EJB Managers, MAGE-OM API, and the MAGE-ML document.
- caArray is built upon MAGE-OM object model, MIAME and MGED-Ontology standards
- caArray APIs and messages support the delivery of data and also of accompanying metadata, in order to ensure that aggregated data sets are comparable
- caArray supports and extends MAGE-OM which allows for the deep annotation of microarray experiments according to MIAME
- caArray is built to utilize and complement the caCORE infrastructure

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Custom caBIG™ Development

caArray Compliance with caBIG™ Data Standards

- caArray metadata resides in caDSR, a ISO/IEC 11179 derived repository
- caArray is built to support MAGE OM (an OMG specification), as described in Uniform Modeling Language (UML)
- caArray utilizes the MGED Ontology, a set of open and standard controlled vocabularies and ontologies built to support annotation of microarray data.
- caArray is built upon MAGE-OM object model, MIAME 1.1, and MGED-Ontology standards

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Open Standards Integration Layer

The BRIDG Model for Clinical Research

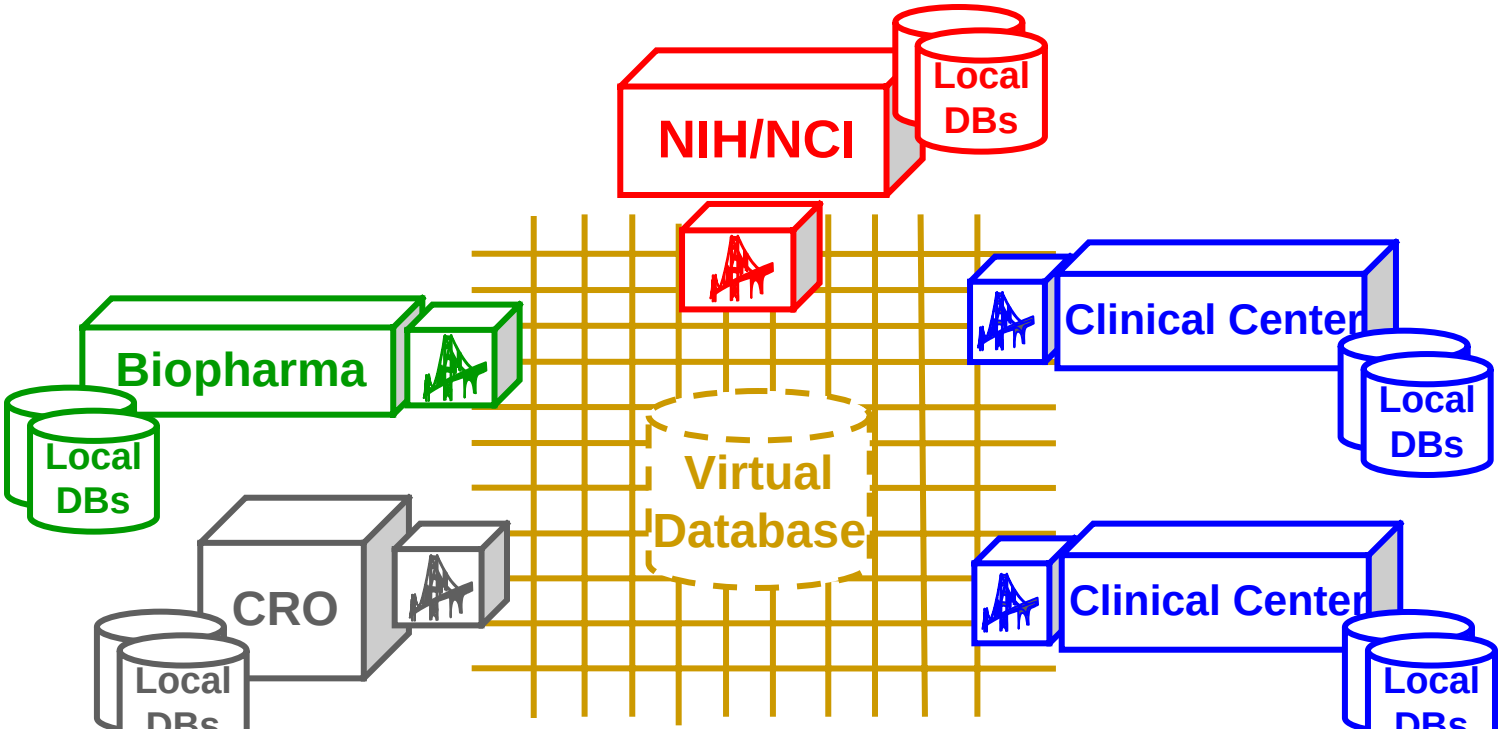
- The BRIDG (Biomedical Research Integrated Data Group) model is an open standards model based on a collaboration between HL7, CDISC, NCI/caBIG, FDA and industry
- Several early BRIDG implementations represent an open standards approach to achieving interoperability:
  - Focus is on defining the APIs that a clinical trials application must implement to interoperate with other clinical trials applications
  - Approach does not dictate application functionality or design

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Open Standards Integration Layer

CTOM (Clinical Trials Object Model): A reference implementation of BRIDG to support collaborative research



caGRID – Grid Services:

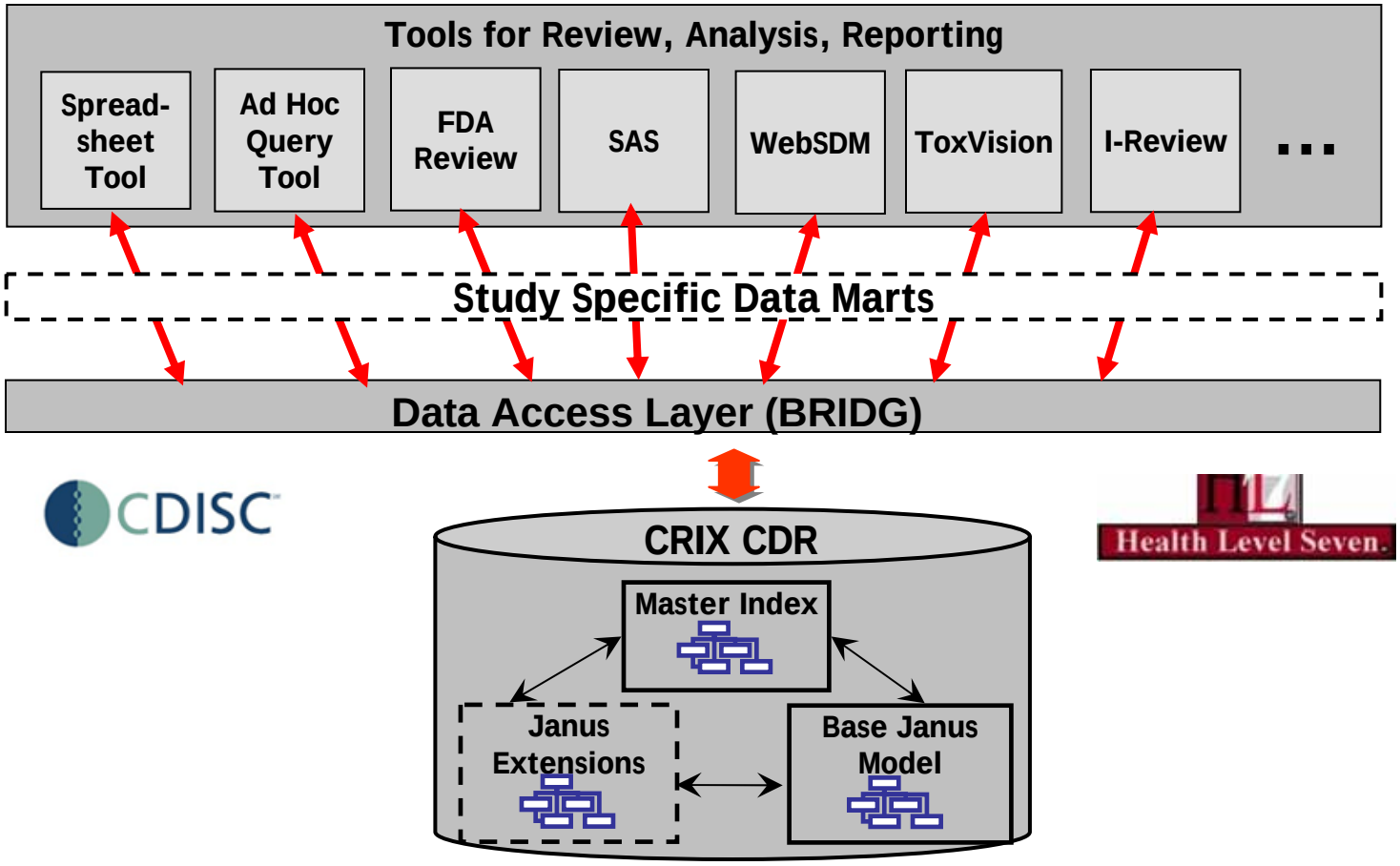
- Data Services (caBIO/CTOM/BRIDG)
- Security Services (CSM)
- Semantic Services (caDSR)
- Vocabulary Services (EVS)
- etc.

National Cancer Institute

caBIG cancer Biomedical Informatics Grid

Case Study: Open Standards Integration Layer

CRIX (Clinical Research Information Exchange) data access layer based on BRIDG



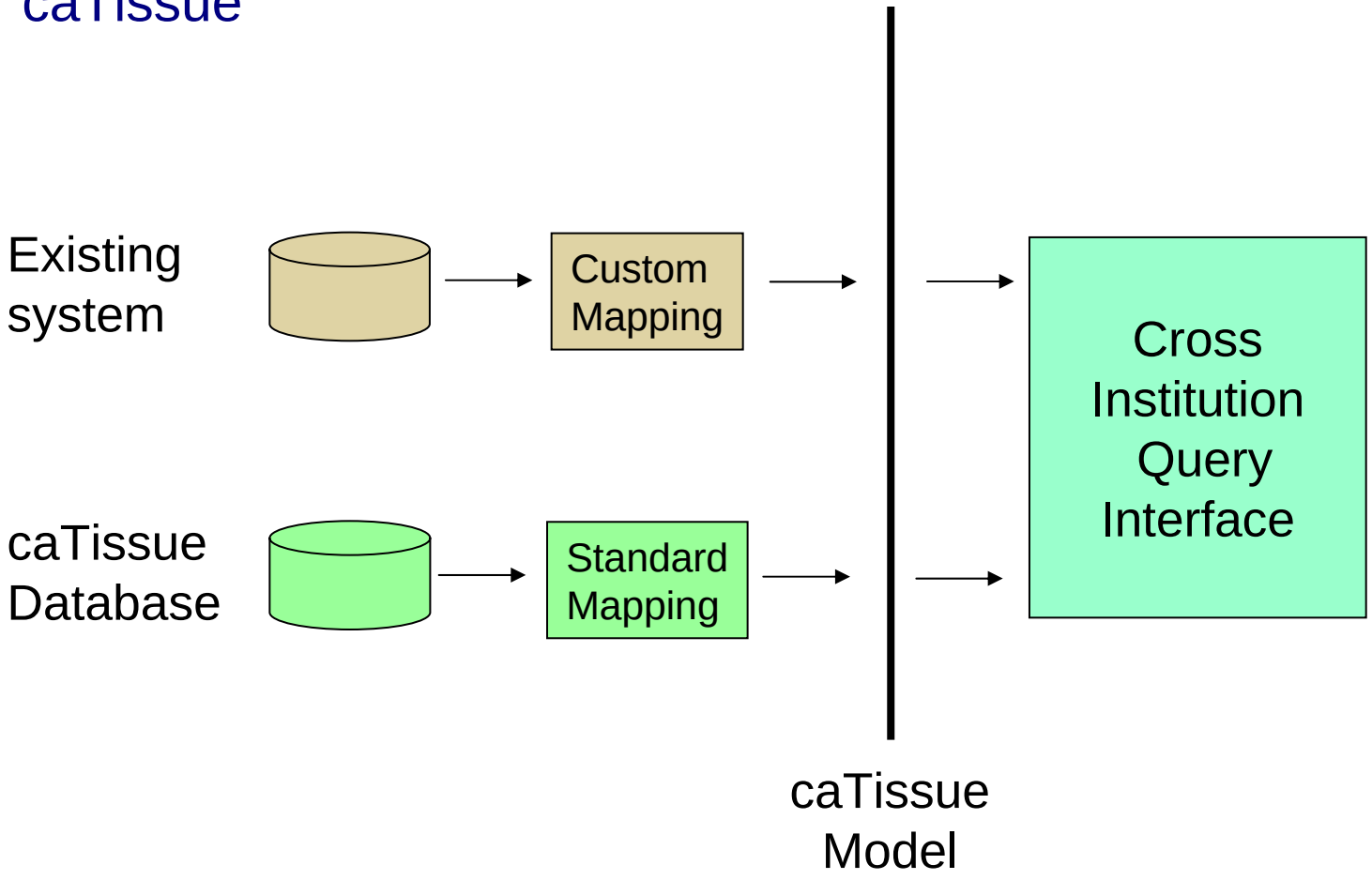
Case Study: Open Standards Integration Layer

caTISSUE Object Model

- Model to expose existing biospecimen informatics systems
- Plays well with others - via caBIG compatibility
- Modular design so that additional functionality (biological annotations, billing and financial modules, etc.) can be added without architectural redesign

caTISSUE Object Model

Allows sites to use existing/vendor solutions or adopt caTissue



National Cancer Institute

caBIG™ cancer Biomedical Informatics Grid™

Summary

- There is no single approach to caBIG compatibility
- Existing vendor solutions have been made caBIG™ compatible relatively easily
- Open standards models for interoperability are progressing
- CDEs are being harmonized across the caBIG community
- Tools to support development of caBIG™ compatible solutions are improving and these types of implementations are getting easier