Crossing Organizational Boundaries: Knowledge Management and Sharing to Advance Evidence Generating Medicine (EGM)

Philip R.O. Payne, Ph.D.

Associate Professor & Chair, Biomedical Informatics
Executive Director, Center for IT Innovation in Healthcare
Co-Director, Biomedical Informatics Program, Center for Clinical and Translational Science
Co-Director, Biomedical Informatics Shared Resource, Comprehensive Cancer Center

Improving People's Lives
through innovations in personalized health care

THE OHIO STATE UNIVERSITY Medical Center

Overview

1. **Motivation**
   - Realizing the promise of “Big Data”
   - Moving beyond traditional organizational boundaries
2. **Critical Approaches and Technologies**
   - Knowledge management
   - Integrative informatics platforms
3. **Challenges and Opportunities**
   - Reducing the distanced between data and knowledge generation
   - Enabling a systems-level approach to EGM
4. **Discussion**

The Ohio State University Medical Center

Overview

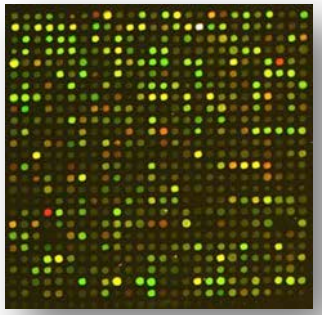
1. **Motivation**
   - Realizing the promise of “Big Data”
   - Moving beyond traditional organizational boundaries
2. **Critical Approaches and Technologies**
   - Knowledge management
   - Integrative informatics platforms
3. **Challenges and Opportunities**
   - Reducing the distanced between data and knowledge generation
   - Enabling a systems-level approach to EGM
4. **Discussion**

The Ohio State University Medical Center

Many Sources of Data!

Emergent Sources
PHR, Instruments, Etc.

Enterprise Systems and Data Repositories:
EHR, CTMS, Data Warehouses



Molecular Phenotype

Environment

The Ohio State University Medical Center

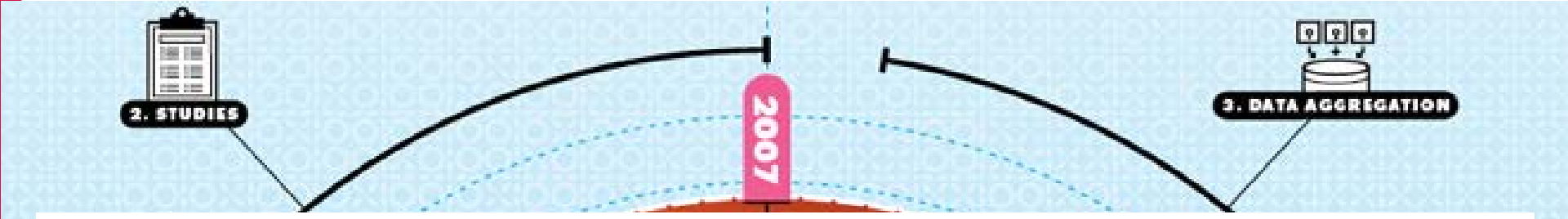
Big Data + Computing = Improved Health?

- **"Sergey Brin's Search for a Parkinson's Cure"**
  - Wired Magazine, July 2010
- **Leveraging Google's Computational Expertise To Mine Big Data**
  - Distributed computing
  - Reasoning across heterogeneous data types
  - Exchanging traditional measures of result validity for the predictive power of increasingly large data sets

THE OHIO STATE UNIVERSITY Medical Center

But Reasoning on Big Data Is Hard…

- **Unexpected problems**
  - Algorithms behave differently
  - Applicability of convention metrics
    - P-values don't mean allot peta-byte scale data sets
  - Signal vs. noise
    - Detection
    - Understanding of patterns
- **Physical computing**
  - Data storage
  - Computational performance



Traditional Model

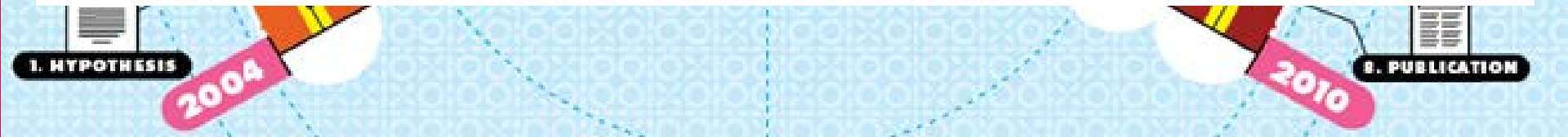
1. Hypothesis: An early study suggests that patients with Gaucher's disease (caused by a mutation to the GBA gene) might be at increased risk of Parkinson's.
2. Studies: Researchers conduct further studies, with varying statistical significance.
3. Data aggregation: Sixteen centers pool information on more than 5,500 Parkinson's patients.
4. Analysis: A statistician crunches the numbers.
5. Writing: A paper is drafted and approved by 64 authors.
6. Submission: The paper is submitted to *The New England Journal of Medicine*. Peer review ensues.
7. Acceptance: *NEJM* accepts the paper.
8. Publication: The paper notes that people with Parkinson's are 5.4 times more likely to carry the GBA mutation.

Total time elapsed: 6 years

Parkinson's Genetics initiative

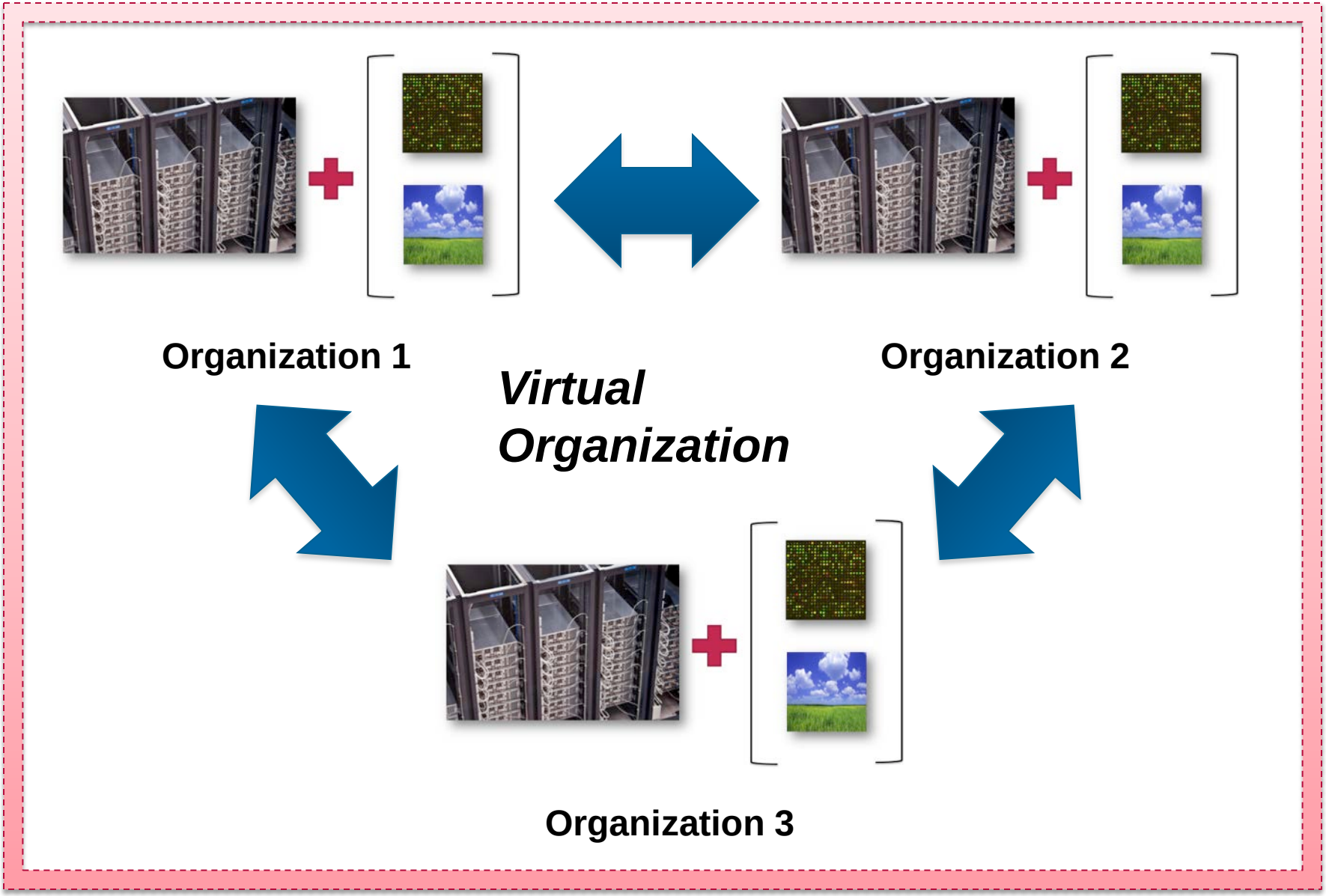
1. Tool Construction: Survey designers build the questionnaire that patients will use to report symptoms.
2. Recruitment: The community is announced, with a goal of recruiting 10,000 subjects with Parkinson's.
3. Data aggregation: Community members get their DNA analyzed. They also fill out surveys.
4. Analysis: Reacting to the *NEJM* paper, 23andMe researchers run a database query based on 3,200 subjects. The results are returned in 20 minutes.
5. Presentation: The results are reported at a Royal Society of Medicine meeting in London: People with GBA are 5 times more likely to have Parkinson's, which is squarely in line with the *NEJM* paper. The finding will possibly be published at a later date.

Total time elapsed: 8 months



Adapted From: "Sergey Brin's Search for a Parkinson's Cure", Wired (July, 2010)

Moving Beyond Organizational Boundaries



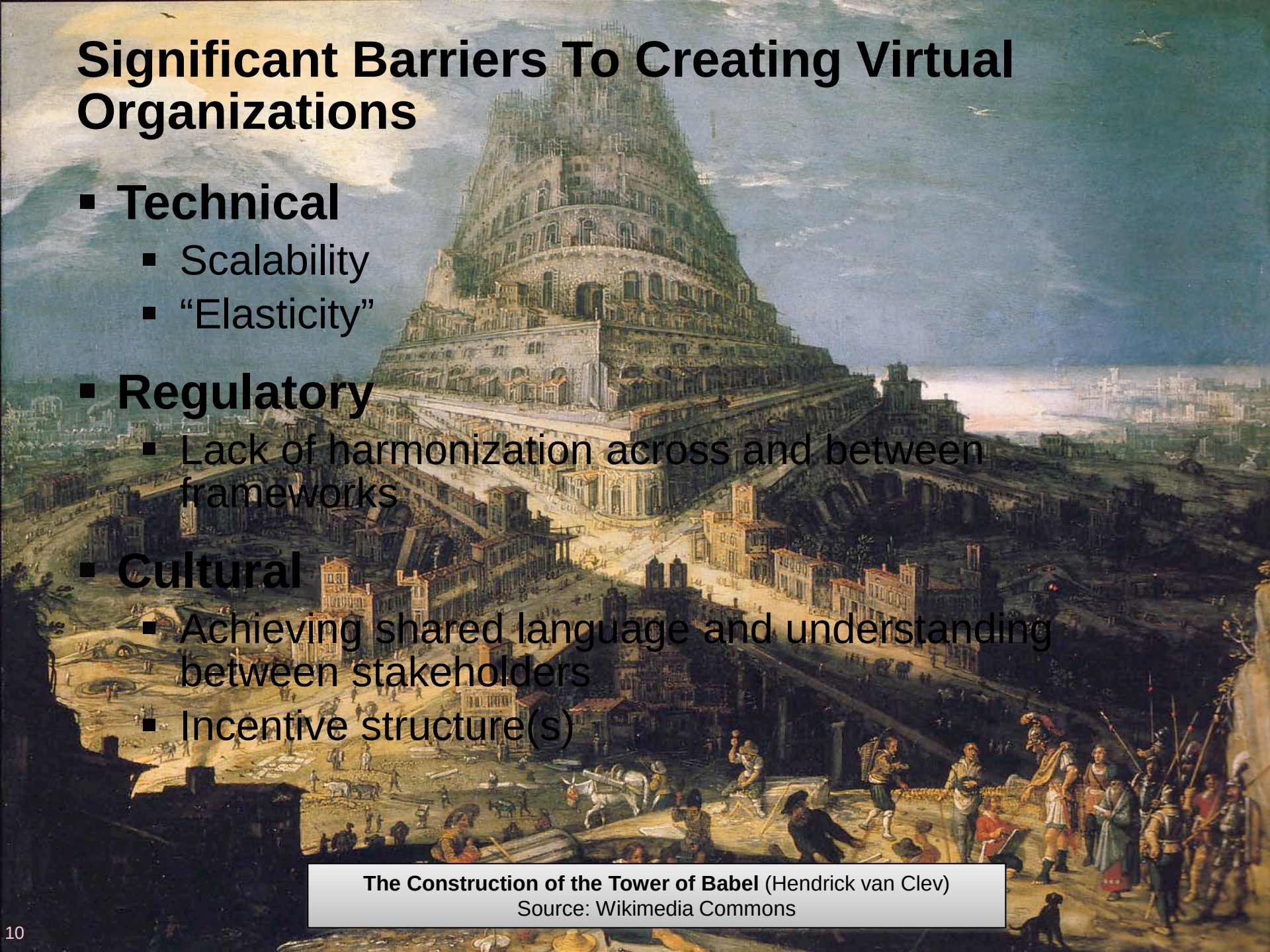
Benefits of Virtual Organizations

- **Larger patient populations**
  - Increased diversity
  - Ability to detect less common “signals”
- **Economies of scale**
  - Expertise
  - Resources
- **Extensibility of study outcomes**

THE OHIO STATE UNIVERSITY Medical Center

Significant Barriers To Creating Virtual Organizations

- **Technical**
  - Scalability
  - “Elasticity”
- **Regulatory**
  - Lack of harmonization across and between frameworks
- **Cultural**
  - Achieving shared language and understanding between stakeholders
  - Incentive structure(s)



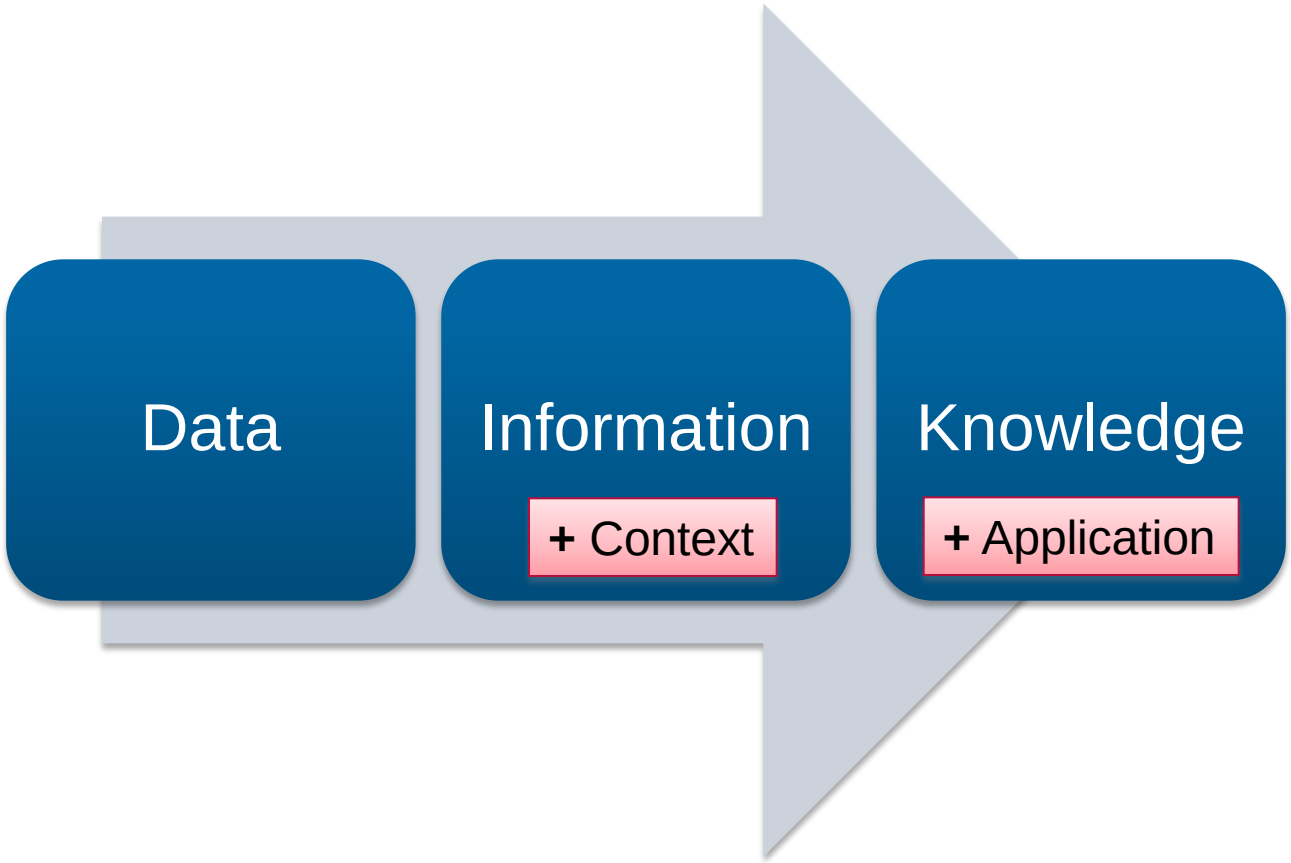
The Construction of the Tower of Babel (Hendrick van Clev)
Source: Wikimedia Commons

Overview

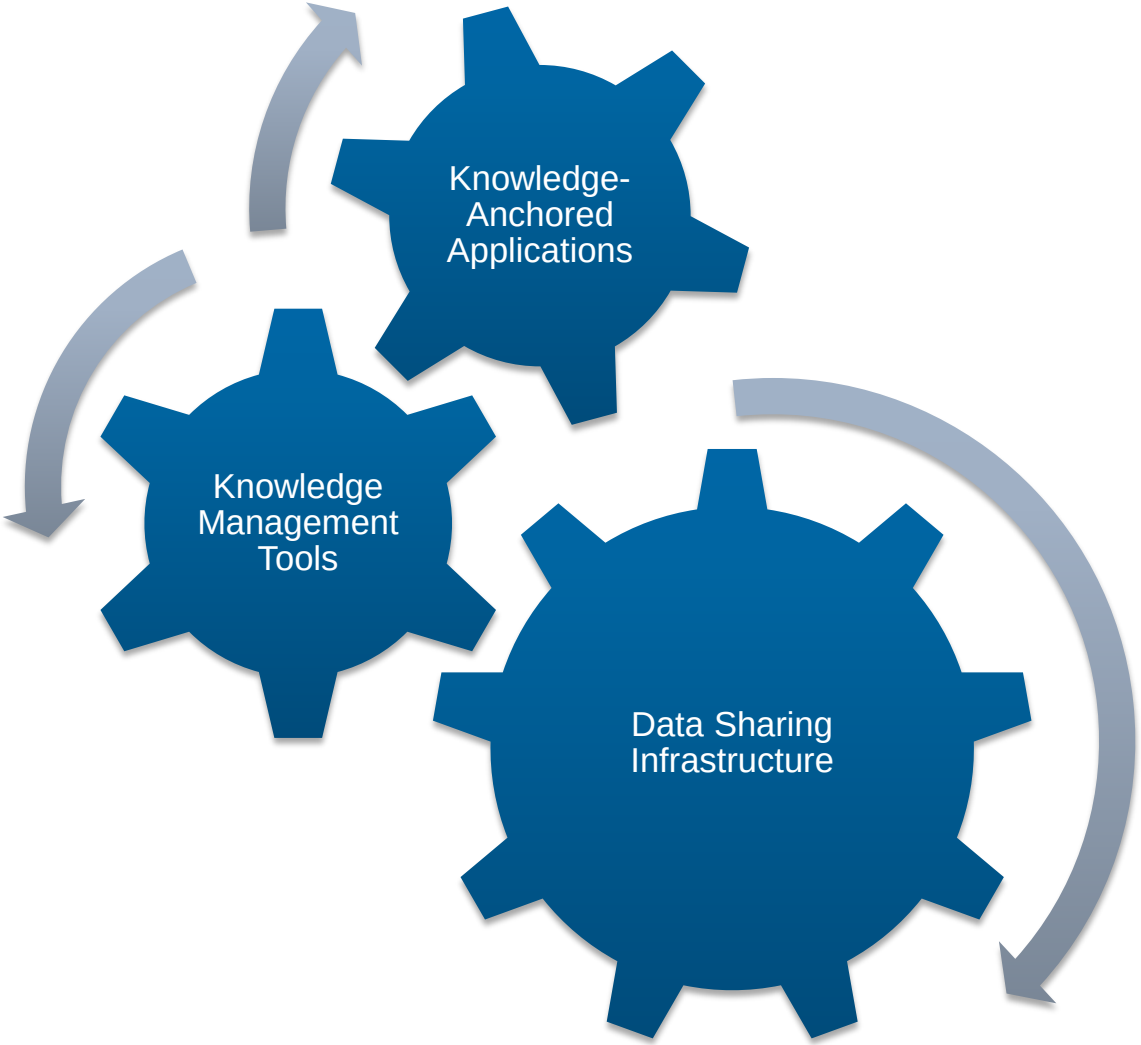
1. **Motivation**
   - Realizing the promise of “Big Data”
   - Moving beyond traditional organizational boundaries
2. **Critical Approaches and Technologies**
   - Knowledge management
   - Integrative informatics platforms
3. **Challenges and Opportunities**
   - Reducing the distanced between data and knowledge generation
   - Enabling a systems-level approach to EGM
4. **Discussion**

The Ohio State University Medical Center

The Role of Biomedical Informatics and HIT: Generating Information and Knowledge



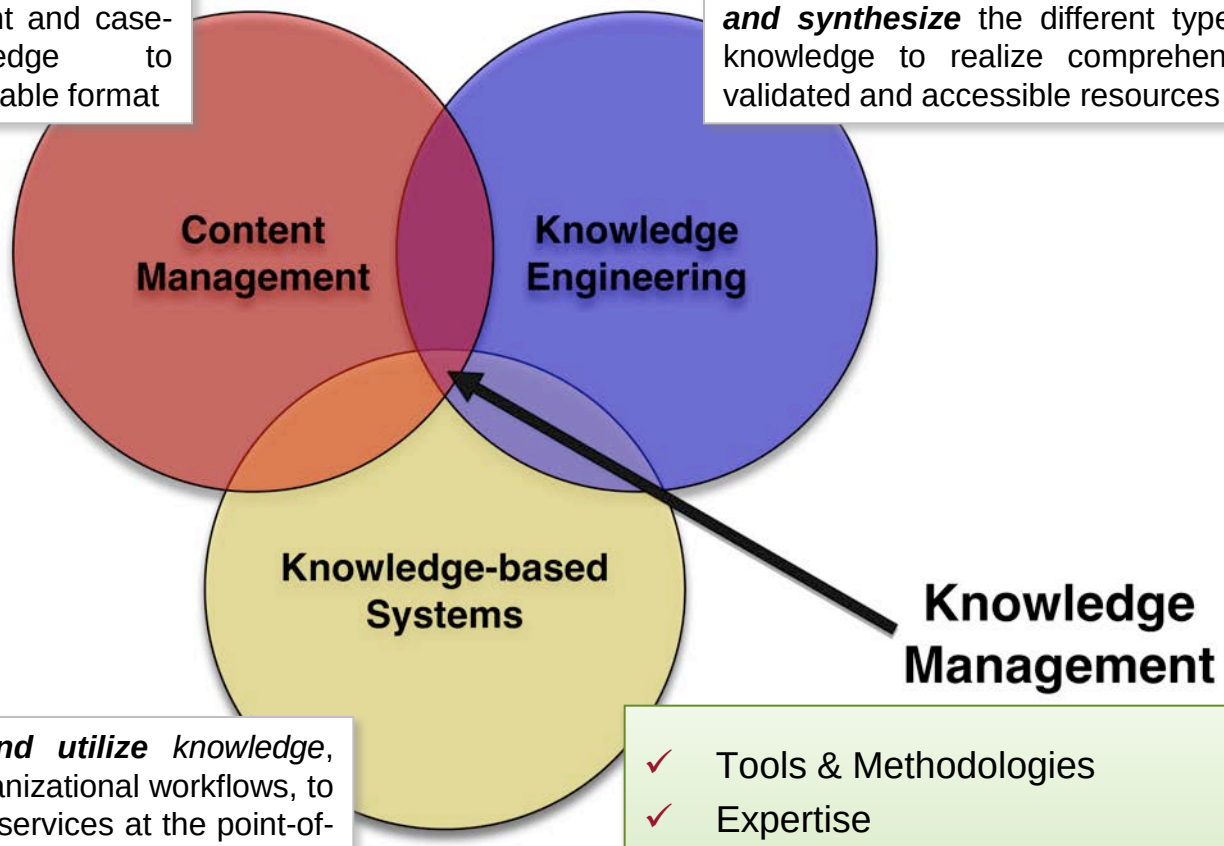
Core Platforms Supporting Virtual Organizations



Knowledge Management (KM):
A Core Competency

Access, share and disseminate current and case-specific knowledge to stakeholders in a usable format

Capture, represent, model, organize and synthesize the different types of knowledge to realize comprehensive, validated and accessible resources



Operationalize and utilize *knowledge*, within existent organizational workflows, to provide pragmatic services at the point-of-need (e.g., point-of-care decision support)

- ✓ Tools & Methodologies
- ✓ Expertise
- ✓ Focus on integration and dissemination of heterogeneous and multi-dimensional biomedical data sets

Abidi SSR. *Healthcare Knowledge Management: The Art of the Possible*. In: Knowledge Manage
Smaltz DH and RC Pinto. *Organizational Knowledge – Can You Really Manage It?* In: Proc HIM

Slide Source: Tara Payne, "*Knowledge Management for Research*"

The Importance of KM: Coping With Constant Evolution in Technology

1950-60's: Specialized computing facilities, programming languages, decision support, bibliographic databases, basic clinical documentation systems, first training programs

Today: Tele-health, mobile computing, widespread EHR adoption, service-oriented architectures, genomic and personalized medicine applications, translational research

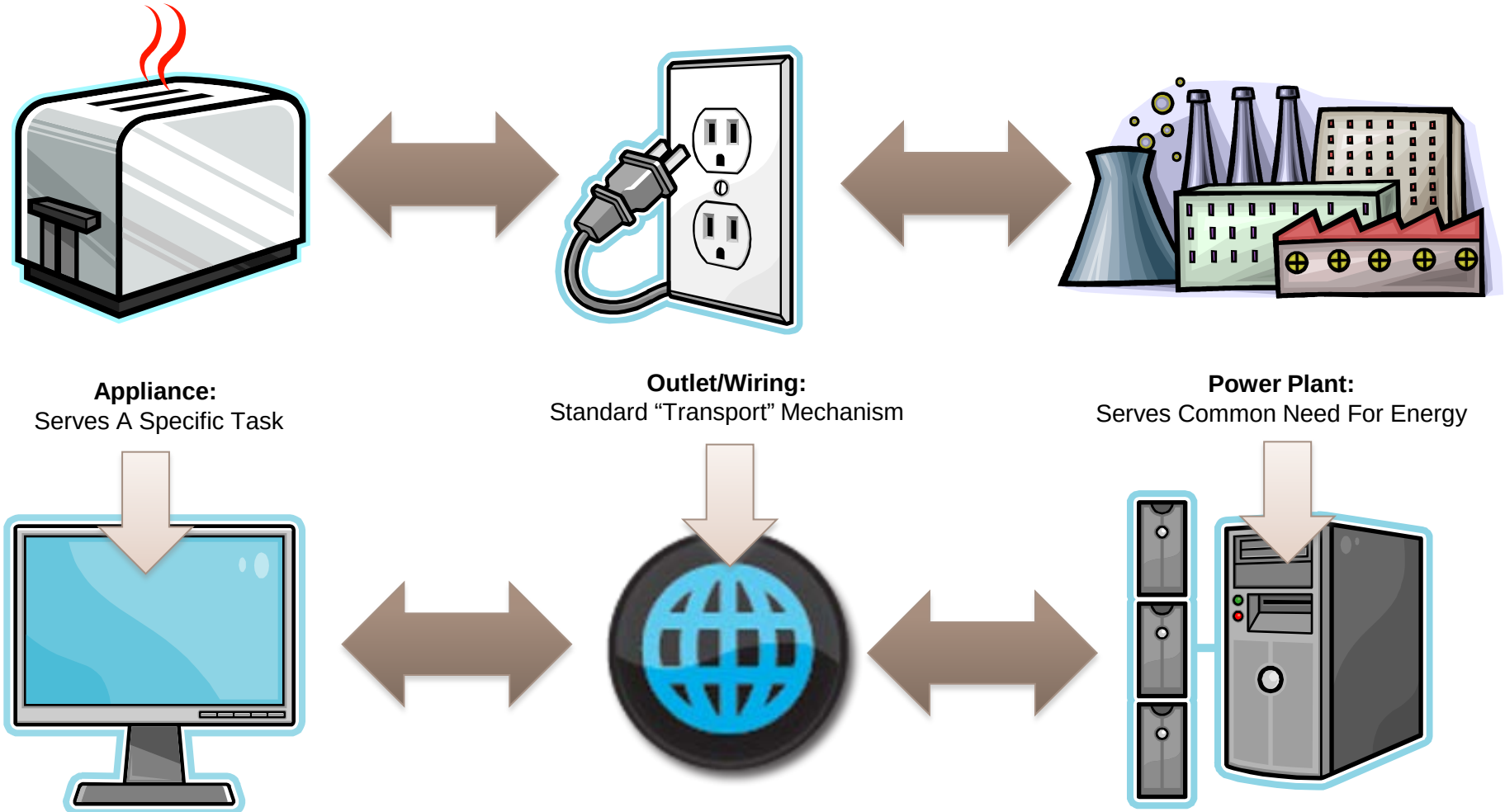
THE OHIO STATE UNIVERSITY Medical Center

Examples of Knowledge Management Tools

- **Terminology and Ontology Services**
  - Common data elements (CDEs)
  - Metadata and model repositories
- **Content Management Systems**
  - Document Management and Version Control
  - Wikis
- **Knowledge-bases**
  - Operational
  - Scientific
- **Social media**
  - Crowdsourcing
  - “Folksonomies”

The Ohio State University Medical Center

Bridging Organizational Boundaries: Service Oriented Architecture (SOA)



The Value Proposition for SOA-based Approaches to Data Federation

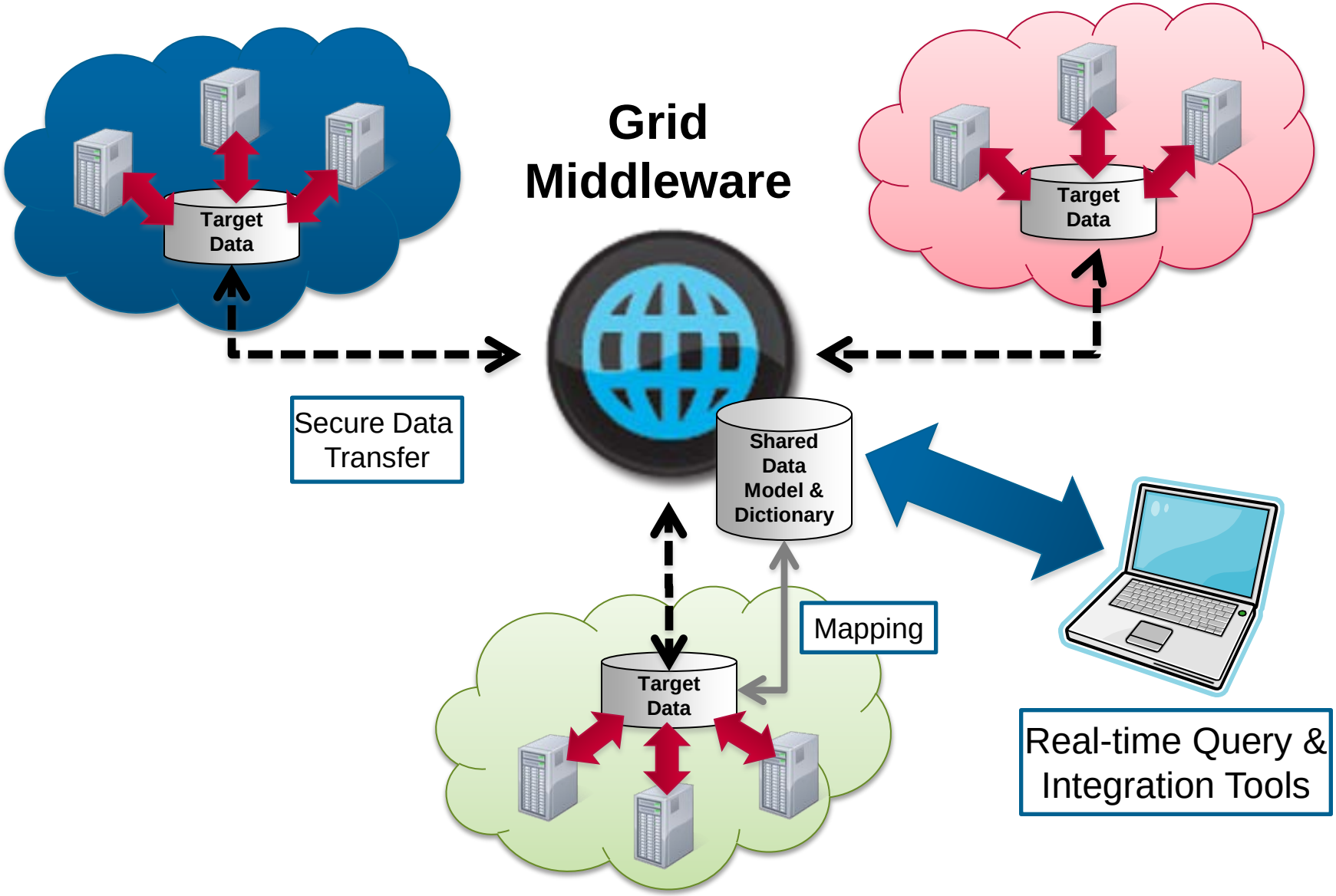
- **Reduced need to replicate data**
  - Data “lives” where it is initially generated or stored
  - Lowers infrastructure costs
- **Increased ability for data stewards to oversee access**
  - Fine-grained and policy-based access control
  - User-centered locus of control
- **“Elasticity”**
  - Ability to expand or contract resources based on current needs (e.g., plug and play)
- **Adaptability**
  - Platform-independent design allows for rapid evolution

THE OHIO STATE UNIVERSITY Medical Center

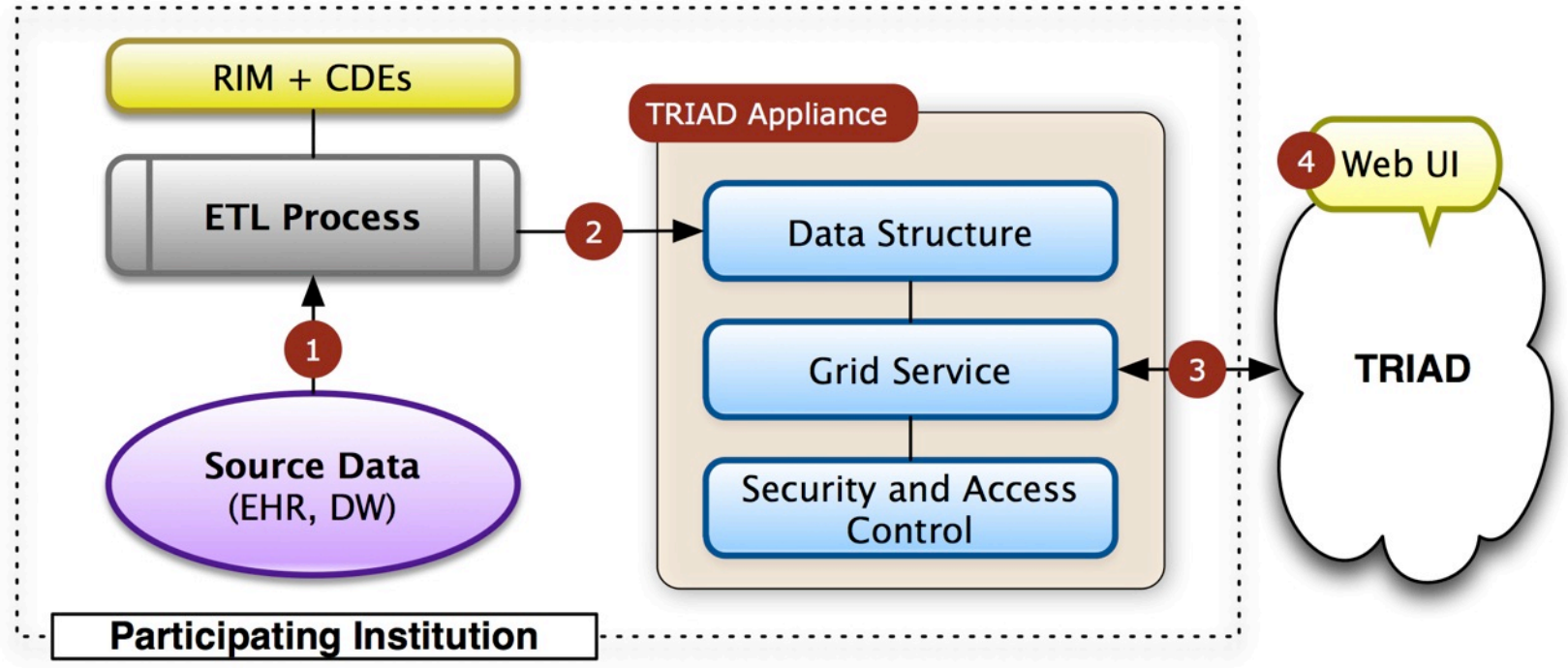
caGrid and TRIAD (Translational Research Informatics and Data Management Grid)

- ***caGrid and TRIAD are a generic and domain agnostic set of middleware and tools that enables service oriented science.***
  - Robust developer and adopter community
  - Developed and supported by the OSU Informatics Research and Development team
- caGrid and TRIAD aims to solve some of the basic challenges in research collaboration and data sharing across organizational boundaries

Use Case: Creating a Virtual Data Warehouse Using caGrid/TRIAD

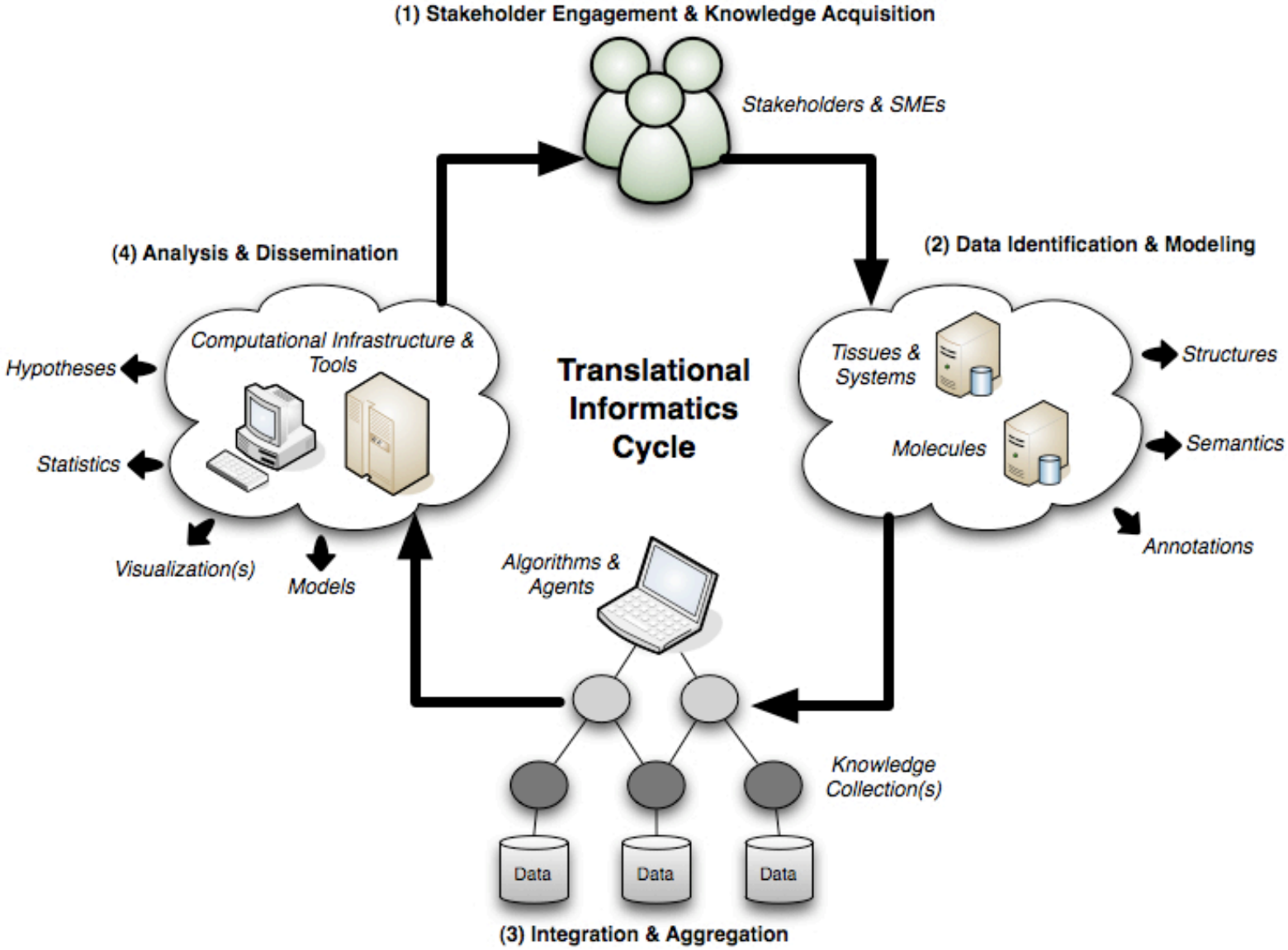


TRIAD Virtual "Appliance"



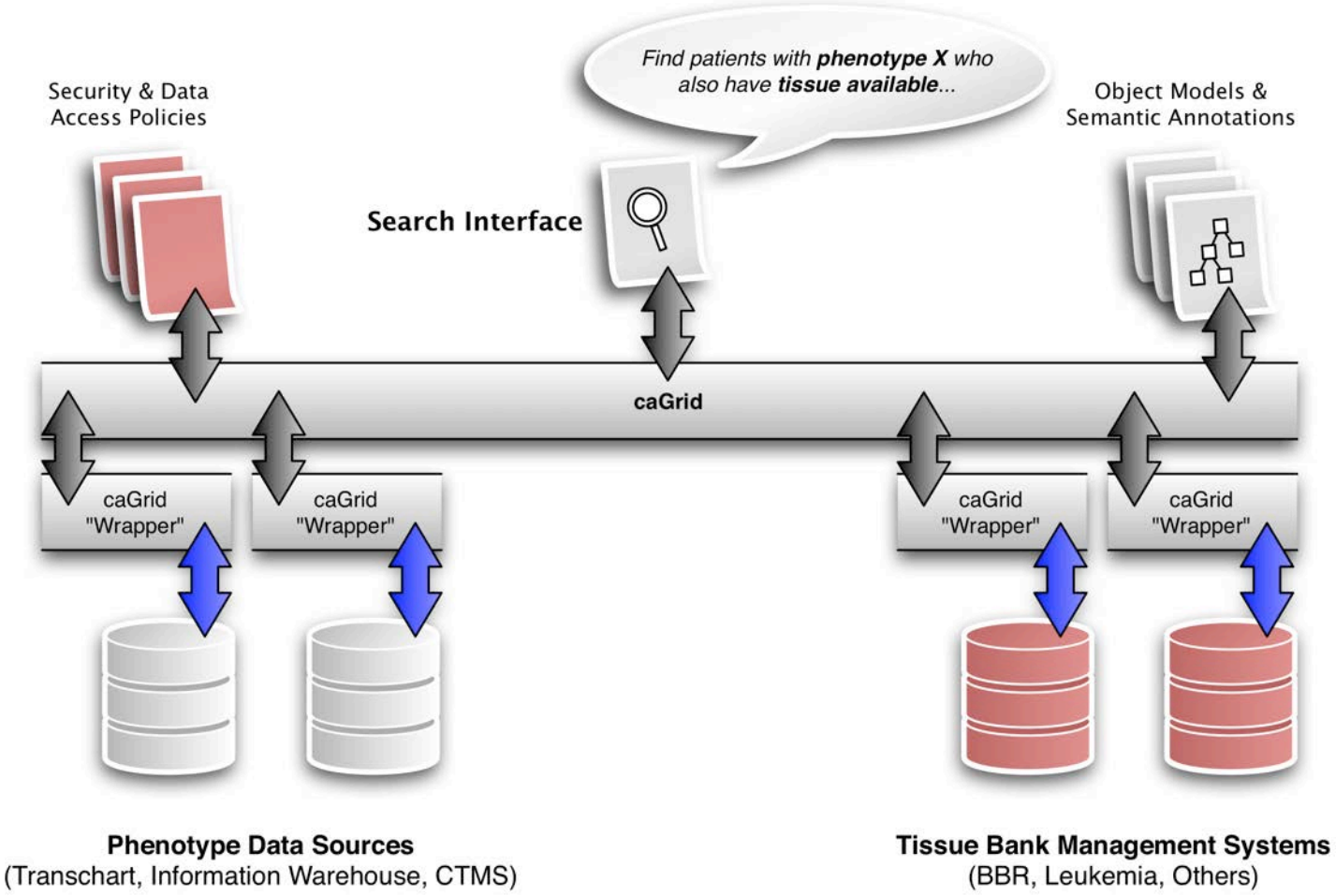
In this deployment model, a virtual server image containing the VA is installed at a participating site. Local source data that will be shared is subject to an Extract-Transform-Load (ETL) process (1) that is informed by a common reference information model (RIM) and common data elements (CDEs). Subsequently, conformant data is loaded into a data structure harmonized with the RIM (2) that is part of the VA, and securely exposed for discovery and distributed query purposes via TRIAD (3). End-users employ a simple, GWT-based user interface to construct and execute distributed queries spanning multiple VAs (4).

Designing Knowledge-Anchored Applications



Payne PR et al. *Translational informatics: enabling high-throughput research paradigms*. In: Physiol. Genomics 39: 131-140, 2009

Use Case: Distributed Cohort and Tissue Discovery



THE OHIO STATE UNIVERSITY Medical Center

CohortIQ Portal Interface

Diagnosis

Procedures

De-Identification

- History
- Results

Displaying 1 to 10 of 21 results. Export to Excel Export to XML

enterDateTime	outDateTime	postOperatingRoomDiagnosis	postOperatingRoomProcedure	preOperatingRoomDiagnosis	preOperatingRoomProcedure	scheduledOperatingRoomEnterDateTime	id
2008-05-31T07:19:00.000-04:00	2008-05-31T15:10:00.000-04:00	COLON CANCER	LAPAROSCOPIC SUBTOTAL COLECTOMY	COLON CANCER	LS VS OPEN HEMICOLECTOMY WITH ILEORECTAL ANASTAMOSIS	2008-05-31T07:30:00.000-04:00	AAF6p6AAKAAAM0kAAS
1899-12-10T00:00:00.000-05:00	1899-12-10T00:00:00.000-05:00			UNSTABLE ANGINA Iodine: N LHC CORS UA IPD IPD	LHC CORS UA IPD IPD	2009-01-20T09:00:00.000-05:00	AAF6p6AAKAAAM1HAAF
2008-10-19T13:16:00.000-04:00	2008-10-19T15:04:00.000-04:00	LARYNGEAL CANCER	LARYNGOSCOPY DIRECT OPERATIVE WITH BIOPSY;	LARYNGEAL CANCER	DLE WITH BX POSS SML WITH LASER EXCISION	2008-10-19T09:00:00.000-04:00	AAF6p6AAKAAAM4eAAH
2009-02-25T12:30:00.000-05:00	2009-02-25T20:46:00.000-05:00	NEUROGENIC BLADDER	ILEAL CONDUIT	NEUROGENTIC BLADDER	ROBOTIC CYSTECTOMY W/ILEAL CONDUIT URINARY DIVERSION	2009-02-25T10:00:00.000-05:00	AAF6p6AALAAAPmJAAB
1899-08-12T00:00:00.000-05:00	1899-08-12T00:00:00.000-05:00			MAL NEO BRONCH/LUNG NOS LUNG CANCER	AF BRONCH LEFT THOR LEFT UPPER LOBECTOMY RIB	2009-05-02T11:30:00.000-04:00	AAF6p6ABBAAANjHAAK
2008-11-22T11:58:00.000-05:00	2008-11-22T14:22:00.000-05:00	GOITER	TOTAL THYROIDECTOMY	GOITER	THYROIDECTOMY	2008-11-22T11:15:00.000-05:00	AAF6p6ABBAAANj2AAS
1899-12-10T00:00:00.000-05:00	1899-12-10T00:00:00.000-05:00			ANGINA PT SEEN IN ER9.14 Io OPD OUTPATIENT DIAGNOSTIC	ANGINA PT SEEN IN ER9.14 Io OPD OUTPATIENT DIAGNOSTIC	2009-08-24T06:00:00.000-04:00	AAF6p6ABBAAANqHAAD
2009-06-18T09:12:00.000-04:00	2009-06-18T11:53:00.000-04:00	PROSTATE CANCER	LAPAROSCOPY SURGICAL PROSTATECTOMY RETROPUBIC RADICAL INC	PROSTATE CANCER	ROBOTIC PROSTATECTOMY WITH BPLND	2009-06-18T09:00:00.000-04:00	AAF6p6ABCAAAYLNAAR
2002-08-16T07:05:00.000-04:00	2002-08-16T11:45:00.000-04:00	RECURRANT PERTIONEAL CA	EXP LAP PELVIC DISSECTION STENT PLACEMENT	OVARIAN CA	EXPL LAP REMOVAL OF PELVIC MASS	2002-08-16T07:00:00.000-04:00	AAF6p6ADTAAAADJAAW
2006-06-21T08:49:00.000-04:00	2006-06-21T09:53:00.000-04:00			BURN OF LEFT HAND	DEBRIDEMENT STSG OF LEFT HAND	2006-06-21T08:00:00.000-04:00	AAF6p6ADTAAADSVAAZ

Tissue Availability Filter

Cohort Discovery Tool

Specimen Available: True

Sex Genotype: UNKNOWN

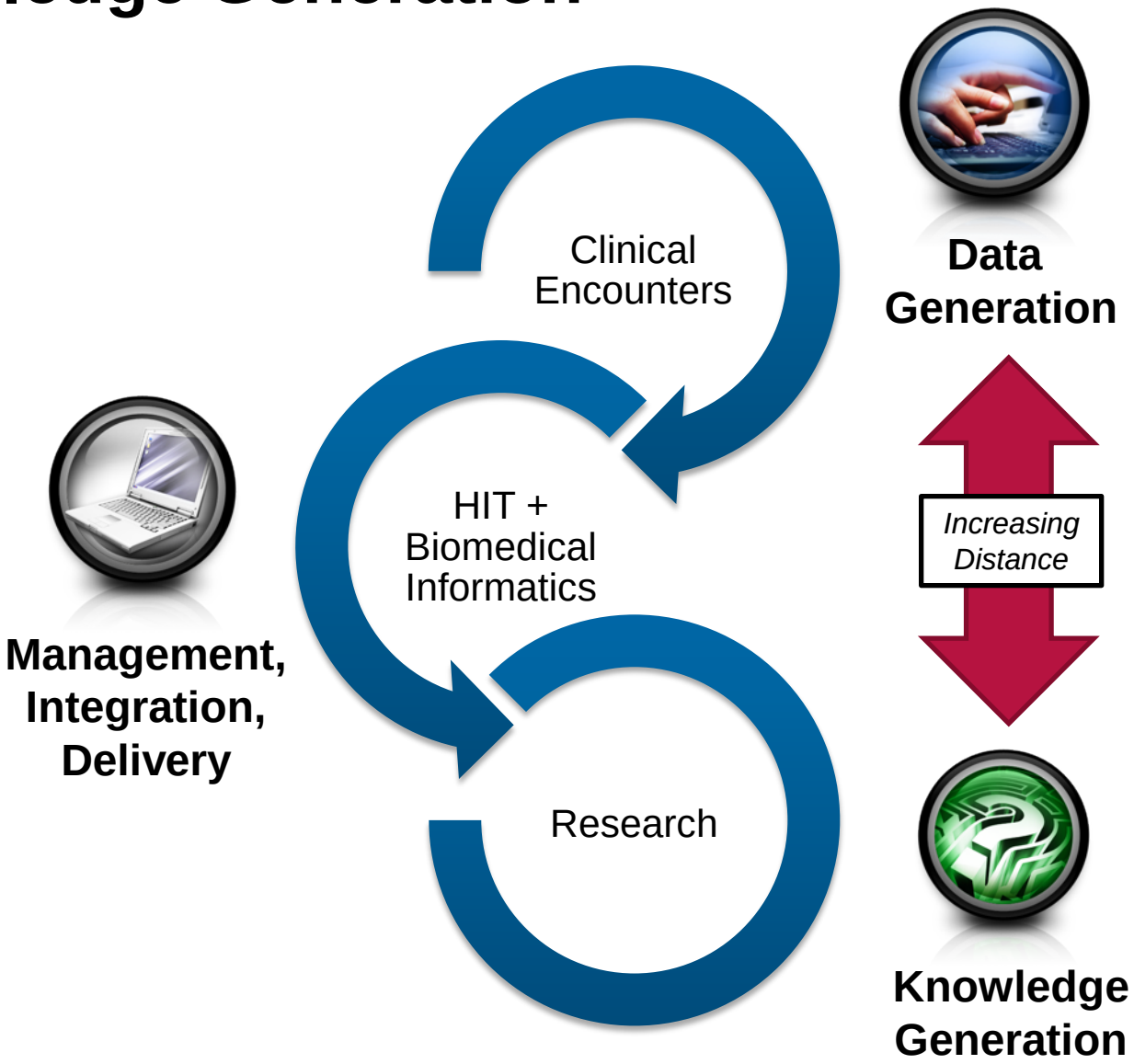
Submit Query

Overview

1. **Motivation**
   - Realizing the promise of “Big Data”
   - Moving beyond traditional organizational boundaries
2. **Critical Approaches and Technologies**
   - Knowledge management
   - Integrative informatics platforms
3. **Challenges and Opportunities**
   - Reducing the distanced between data and knowledge generation
   - Enabling a systems-level approach to EGM
4. **Discussion**

The Ohio State University Medical Center

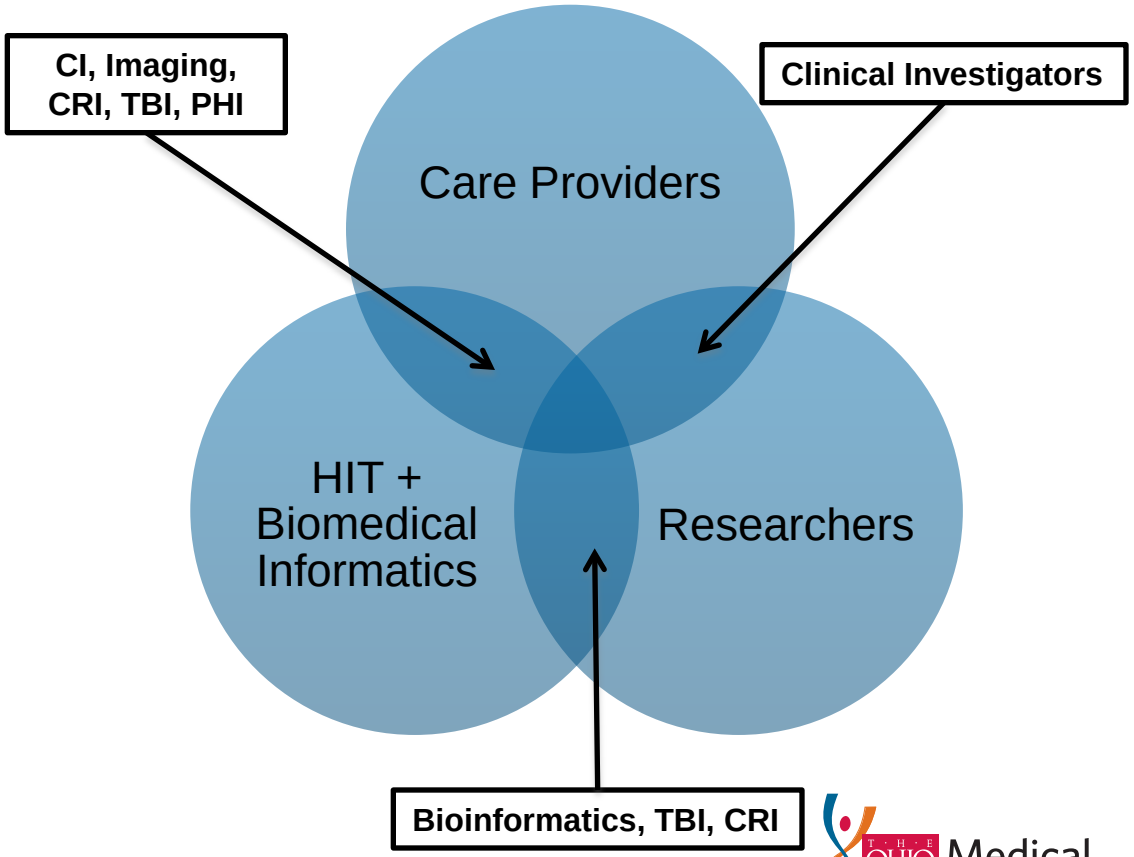
Increasing Distances Between Data and Knowledge Generation



Contributing Factors (1)

Regulatory, Technical, and Cultural Barriers Between Data and Knowledge Generation

- High performance systems require rapid adaptation
- Increasing demand for better, faster, safer, more cost effective therapies
- Simultaneous demand for increased controls over secondary use of clinical data
- Artificial partitioning of access to data for knowledge generation purposes
- **Critical overlaps and potential sources of conflict between these factors**



The Ohio State University Medical Center

Contributing Factors (2)

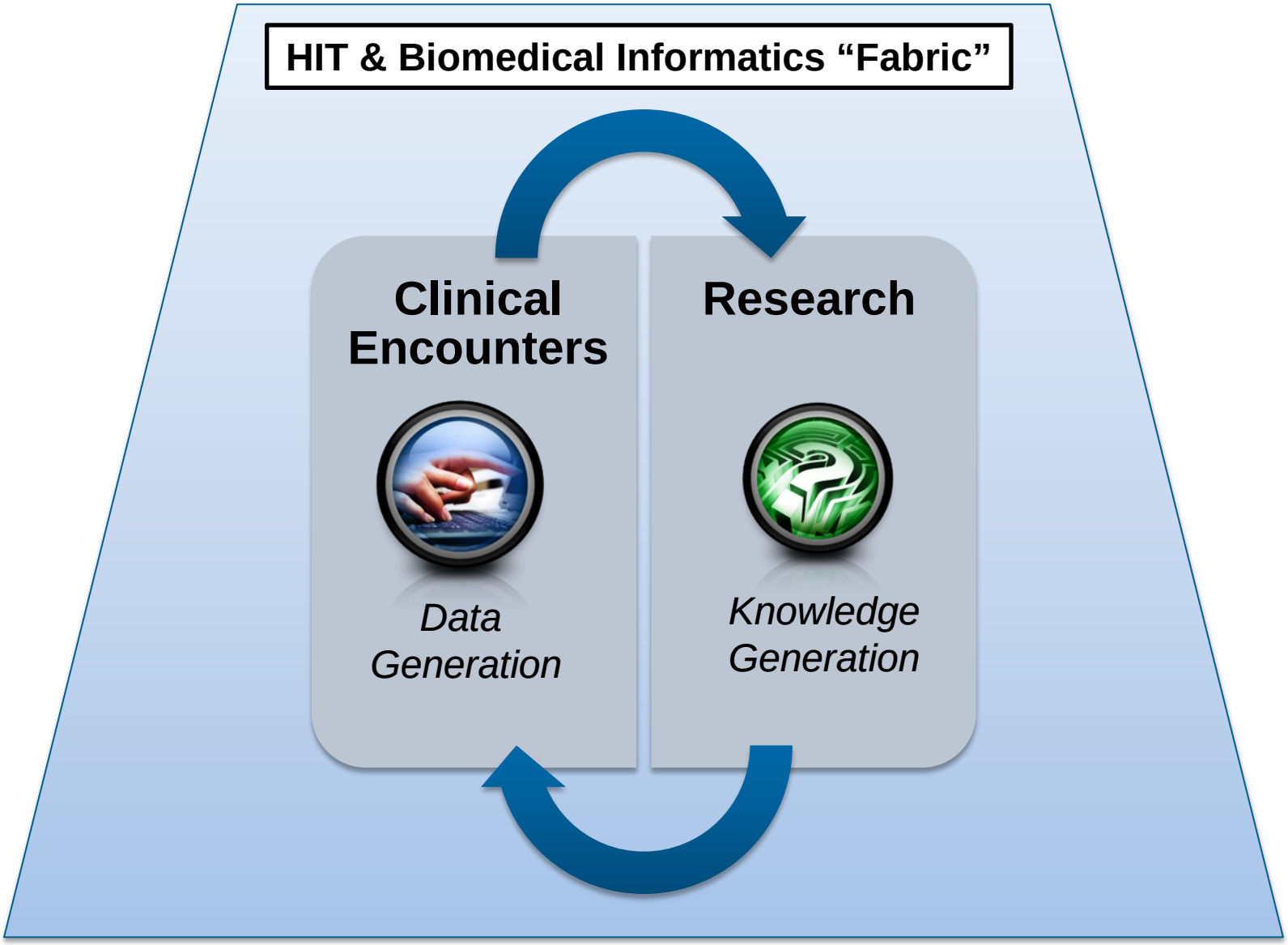
- **Historical precedence for reductionism in the biomedical and life sciences**
  - Break-down problems into fundamental units
  - Study units and generate knowledge
  - Reassemble knowledge into systems-level models
- **Influences policy, education, research, and practice**
- **Recent scientific paradigms have illustrated major problems with this type of approach**
  - Systems biology/medicine
- **Reductionist approach to data, information, and knowledge management is still prevalent**
  - HIT vs. Informatics
  - Informatics sub-disciplines

Overview

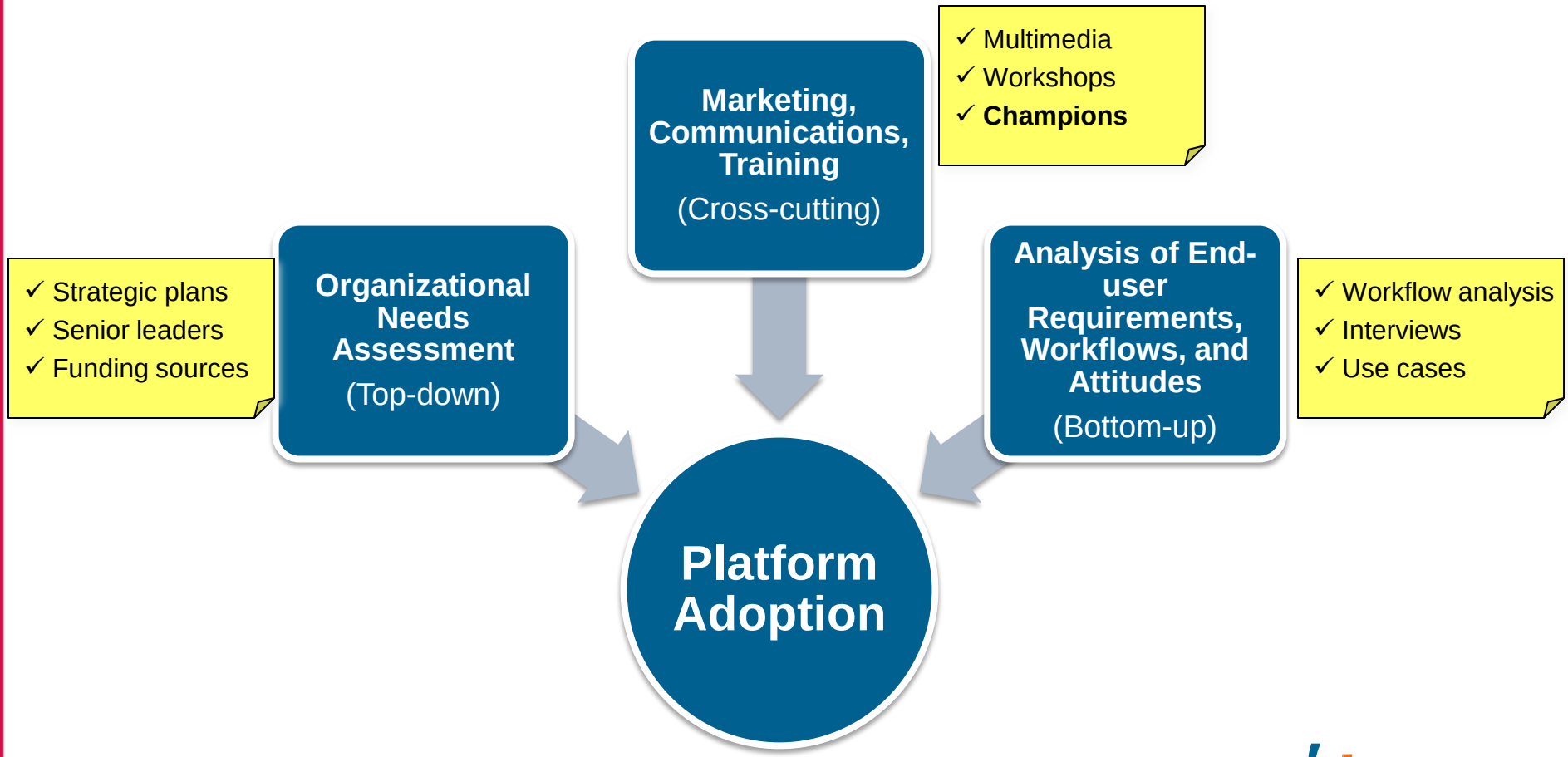
1. **Motivation**
   - Realizing the promise of “Big Data”
   - Moving beyond traditional organizational boundaries
2. **Critical Approaches and Technologies**
   - Knowledge management
   - Integrative informatics platforms
3. **Challenges and Opportunities**
   - Reducing the distanced between data and knowledge generation
   - Enabling a systems-level approach to EGM
4. **Discussion**

The Ohio State University Medical Center

Towards a Solution: A Systems Approach to Biomedicine



Overcoming Barriers: Socio-technical Approaches to Enabling Platform Adoption



The Ohio State University Medical Center

Realizing The Promise of BMI and HIT Requires Us To Build Robust and Innovative Infrastructures

Thank you for your time and attention!

- philip.payne@osumc.edu
- http://bmi.osu.edu/~payne