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
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**
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**
Many Sources of Data!

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

Emergent Sources
PHR, Instruments, Etc.

Molecular Phenotype

Environment
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
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.
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

Virtual Organization

Organization 1 + Virtual Organization + Organization 2

Organization 3
Benefits of Virtual Organizations

- Larger patient populations
  - Increased diversity
  - Ability to detect less common “signals”
- Economies of scale
  - Expertise
  - Resources
- Extensibility of study outcomes
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 distance between data and knowledge generation
   • Enabling a systems-level approach to EGM

4. Discussion
The Role of Biomedical Informatics and HIT: Generating Information and Knowledge
Core Platforms Supporting Virtual Organizations

- Knowledge- Anchored Applications
- Data Sharing Infrastructure
- Knowledge Management Tools
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.


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
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”
Bridging Organizational Boundaries: Service Oriented Architecture (SOA)

Appliance: Serves A Specific Task
Outlet/Wiring: Standard “Transport” Mechanism
Power Plant: Serves Common Need For Energy
Application: Serves A Specific Task
Grid: Standard “Transport” Mechanism
Grid Services: Serves Common Need For Data & Analytical Platforms
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
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

Distributed Data & Knowledge

Syntactic & Semantic Interoperability

Security & Regulatory Frameworks

Socio-technical Factors

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

Target Data

Grid Middleware

Secure Data Transfer

Shared Data Model & Dictionary

Target Data

Mapping

Real-time Query & Integration Tools
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

Use Case: Distributed Cohort and Tissue Discovery
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 distance between data and knowledge generation
   • Enabling a systems-level approach to EGM

4. Discussion
Increasing Distances Between Data and Knowledge Generation

Clinical Encounters

HIT + Biomedical Informatics

Management, Integration, Delivery

Research

Data Generation

Increasing Distance

Knowledge Generation
Contributing Factors (1)

- 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

*Regulatory, Technical, and Cultural Barriers Between Data and Knowledge Generation*
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 distance between data and knowledge generation
   • Enabling a systems-level approach to EGM

4. Discussion
Towards a Solution: A Systems Approach to Biomedicine

HIT & Biomedical Informatics “Fabric”

Clinical Encounters
Data Generation

Research
Knowledge Generation
Overcoming Barriers: Socio-technical Approaches to Enabling Platform Adoption

Organizational Needs Assessment (Top-down)
- Strategic plans
- Senior leaders
- Funding sources

Marketing, Communications, Training (Cross-cutting)
- Multimedia
- Workshops
- Champions

Analysis of End-user Requirements, Workflows, and Attitudes (Bottom-up)
- Workflow analysis
- Interviews
- Use cases

Platform Adoption

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