

**Big Data:
What Keeps Me Up At
Night**

January 7, 2014

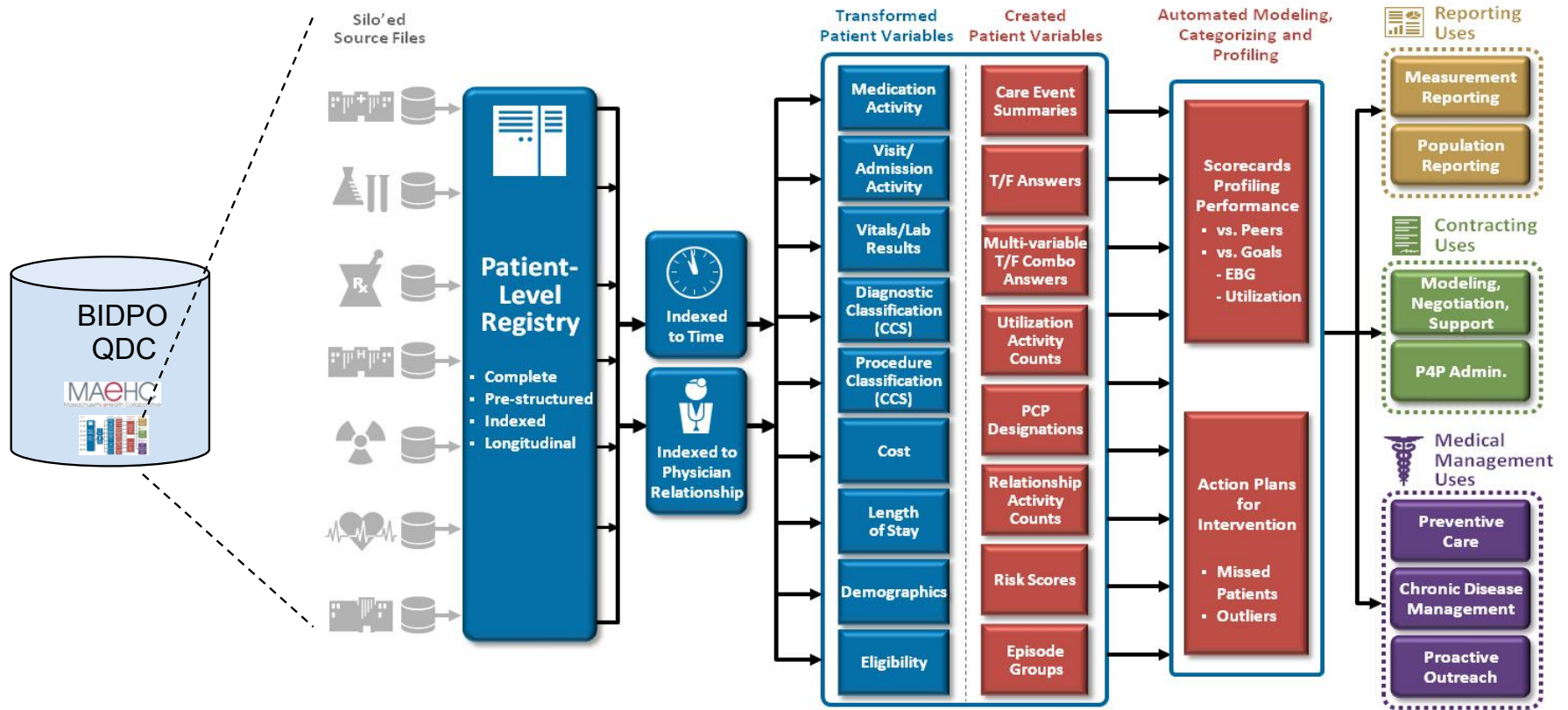
What is Big Data?

- From May 2012 by danah boyd and Kate Crawford's "Critical Questions for Big Data" in Information, Communications, and Society.
- "a cultural, technological, and scholarly phenomenon that rests on the interplay of: (1) Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets. (2) Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims. (3) Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy."

Centralized Data Mining

- Affordable Care Act requires ACOs to report quality measures across all sites of care
- Many EHRs, many vocabularies, many data quality issues
- Multi-year effort to create a Quality Data Center of continuous, longitudinal, normalized patient-centric data

Quality Data Center Components



Quality Measures derived from Big Data

Scorecard: Measure Results				
Measures				
	Population	Numerator	Rate	Missed
A1C testing - DM				
Adolescent Well-Care Visits (AWC)				
Antidepressant Medication Mgmt (AMM) - Effective Acute Phase Treatment				
Antidepressant Medication Mgmt (AMM) - Effective Continuation Phase Treatment				
Avoidance of Antibiotic Treatment for Adults With Acute Bronchitis (AAB)				
Breast Cancer Screening (BCS)				
Cervical Cancer Screening (CCS)				
Chlamydia Screening in Women 21-24 (CHL)				
Cholesterol Management for Patients With Cardiovascular Conditions (CMC) - LDL-C Control				
Cholesterol Management for Patients With Cardiovascular Conditions (CMC) - LDL-C Screening				
Colorectal Cancer Screening (COL)				
Comprehensive Adult Diabetes Care (CDC) - Blood Pressure Control				
Comprehensive Adult Diabetes Care (CDC) - Eye Exam (Retinal) Performed				
Comprehensive Adult Diabetes Care (CDC) - HbA1c Poor Control (>9.0%)				
Comprehensive Adult Diabetes Care (CDC) - Hemoglobin A1c (HbA1c) Testing				
Comprehensive Adult Diabetes Care (CDC) - LDL-C Control				
Comprehensive Adult Diabetes Care (CDC) - LDL-C Screening Performed				
Comprehensive Adult Diabetes Care (CDC) - Medical Attention for Nephropathy				
Comprehensive Ischemic Vascular Disease Care (IVD) - Use of Aspirin or Another Antithrombotic				
Controlling High Blood Pressure (CBP)				
Use of Appropriate Medications for People With Asthma (12-50) (ASM)				

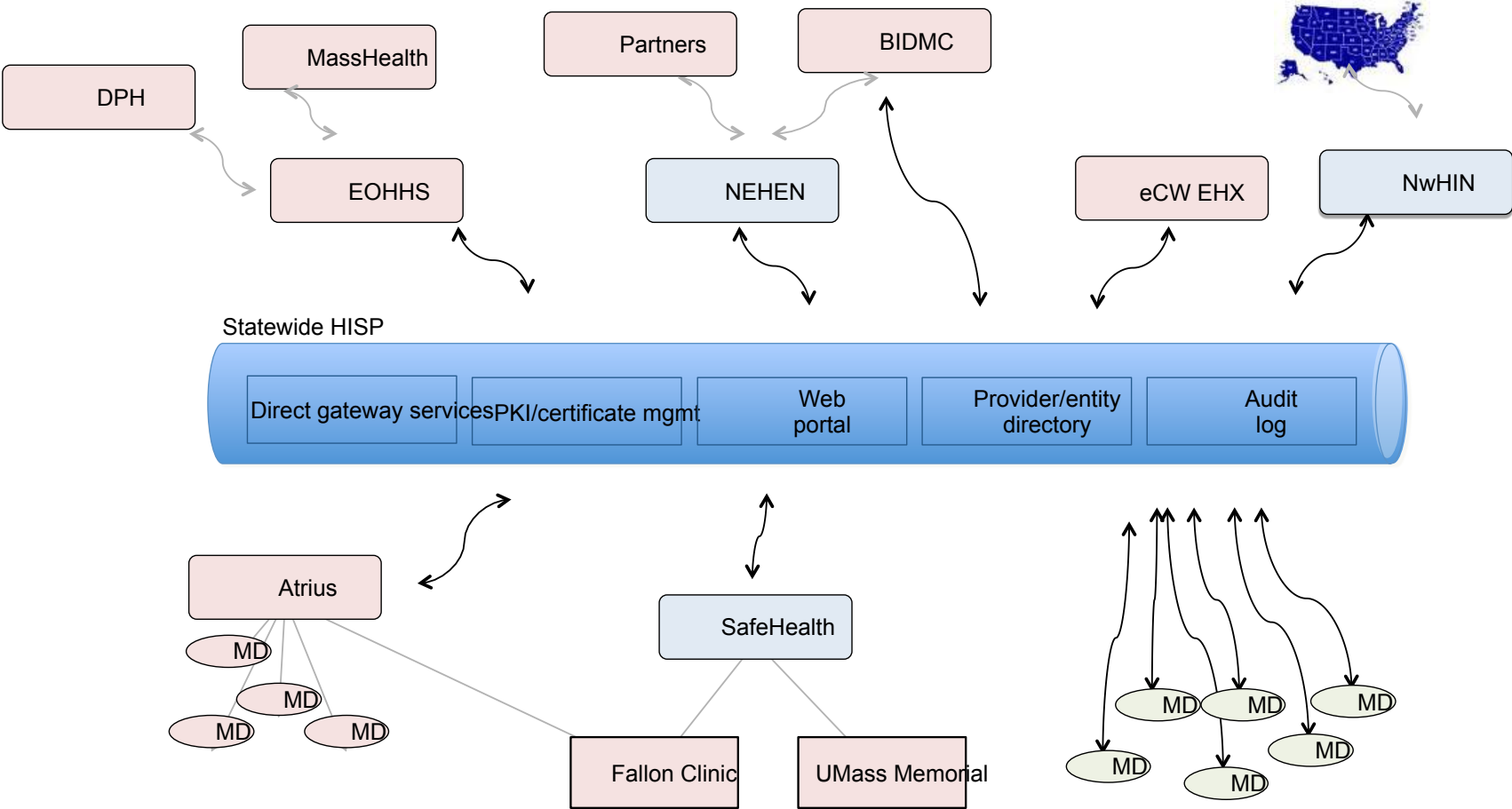
Federated Data Mining

The screenshot shows a web application titled "Clinical Query 2" running in a Windows Internet Explorer browser. The interface is divided into several sections:

- Header:** "Clinical Query 2" on the left and navigation links "Find Patients | Message Log | Help | Logout" on the right.
- Left Panel:** "Navigate Terms" and "Find Terms" tabs. A tree view shows categories: Demographics (Age, Gender (Female, Male, Unknown), Race), Diagnoses, Laboratory Tests, and Medications.
- Query Tool:** "Query Name: Hyper-BETA -Male@02:28:25". "Temporal Constraint: Treat all groups independently".
- Query Groups:** Three groups are defined:
 - Group 1: "Hypertensive disease" with a green box labeled "one or more of these".
 - Group 2: "BETA BLOCKERS/RELATE" and "ACE INHIBITORS" with a red box labeled "none of these".
 - Group 3: "Male" with a green box labeled "one or more of these".The groups are connected by "AND" operators.
- Buttons:** "Run Query", "Clear", "Print Query", "New Group", and "3 Groups".
- Previous Queries:** A list showing "Hyper-BETA -Male@02:28:25 [3-12-2012] [g]".
- Query Status:** "Finished Query: 'Hyper-BETA -Male@02:28:25' [12.1 secs]". "Compute Time: 10.8 secs". "Number of patients for 'Hyper-BETA -Male@02:28:25' patient_count: 39000±3".

The browser's taskbar at the bottom shows the "Internet" icon and a 100% zoom level.

Health Information Exchange is an Enabler for Big Data Registries



- Dr. Famous and his Laptop
- The Chest X-rays in China
- The Nurse and Angry Birds

Security of Big Data is a major issue



Questions?

- jhalamka@bidmc.harvard.edu
- <http://geekdoctor.blogspot.com>