Computing for Medicine

Google Classroom Code: dnd5qkt5

Monsoon 2025

Lecture 3

Data Sources

Open Discussion (5 Minutes)

Who Benefits from Health Data?

Computing for Medicine Starts With Data!

Patient -Source-Oriented Medical Record

Visits :

21 February 2006: dyspnea, coughing and fever. Dark defecation.
PE: BP 150/90, pulse 95/min, Fever: 39.3 °C.
Ronchi +, no abdominal tenderness.
Medications: 64 mg Aspirin/day.
Possible acute bronchitis and cardiac decompensation.
Possible bleeding due to Aspirin.
Rx: Amoxicilline 500 mg 2x1, Aspirin 32 mg/day.
4 March 2006: no cough, slight dyspnea, defecation normal.
PE: light rhonchi, BP 160/95, pulse 82/min.
Rx: Aspirin 32 mg/day.

Lab :

21 February 2006: ESR 25 mm, Hb 7.8, Fecal occult blood +.
4 March 2006: Hb 8.2, Fecal occult blood :-.

X-ray

21 February 2006: Chest x-ray: no atelectasis, light cardiac decompensation findings

Problem 1: Coughing

21 February 2006

S: dyspnea, coughing, fever.

O: pulse 95/min, Fever: 39.3 °C. Rhonchi+. ESR 25 mm.
Chest x-ray: no atelectasis, light cardiac decompensation findings.

A: Acute bronchitis.

P: Amoxicilline 500 mg 2x1.

4 March 2006

S: no coughing, slight dyspnea.

O: pulse 82/min. Slight rhonchi.

A: minimal bronchitis findings.

Problem 2: Dyspnea

21 February 2006

S: Dyspnea.

O: Rhonchi+, BP 150/90 mmHg.
Chest x-ray: no atelectasis, slight cardiac decompensation findings.

A: Slight decompensation findings.

4 March 2006

S: slight dyspnea.

O: BP: 160/95, pulse 82/min.

A: No decompensation.

Problem 3: Dark colored defecation

21 February 2006

S: Dark feces. Using Aspirin 64 mg/day.

O: No abdominal tenderness, rectal exam revealed no blood, Hb 7.8 mg/dl. Fecal occult blood +

A: Possible intestinal bleeding due to Aspirin.

P: Decrease Aspirin dose to 32 mg/day.

4 March 2006

S: Defecation normal.

O: Fecal occult blood -

A: No intestinal bleeding symptoms.

P: Continue Aspirin dosage 32 mg/day

Making Healthcare Data Science Reproducible and FAIR

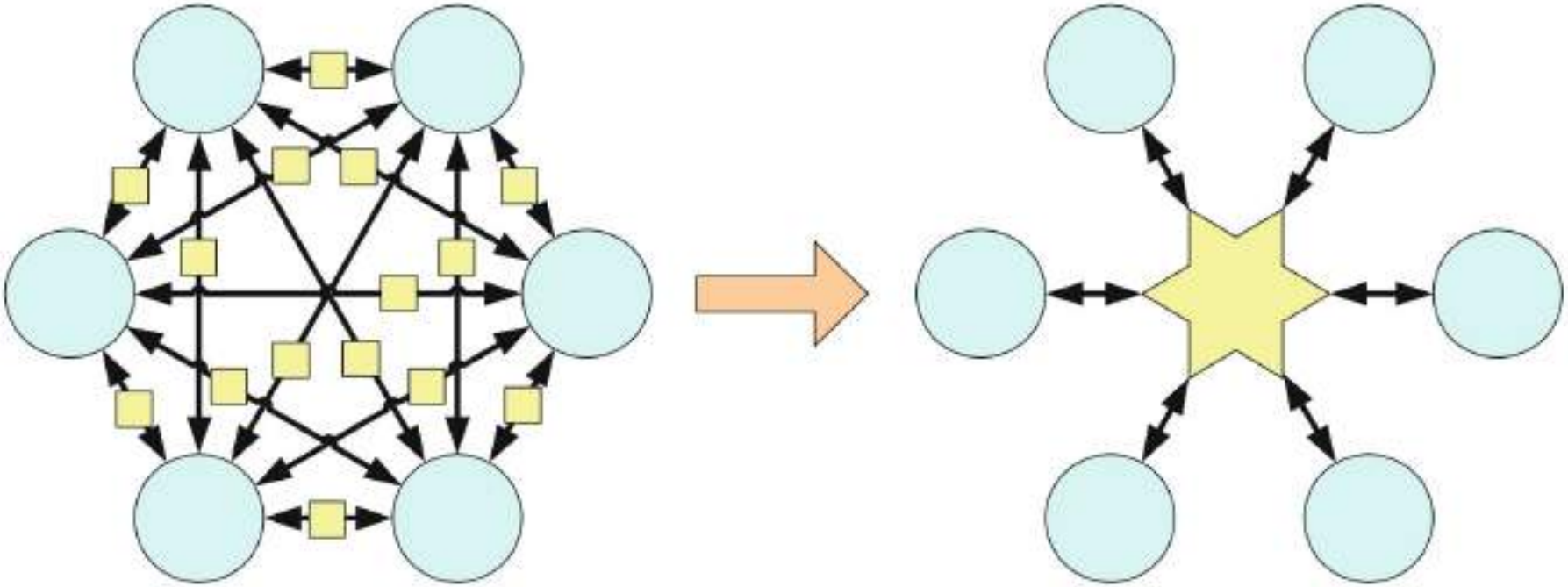
To be **Findable**:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be **Accessible**:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

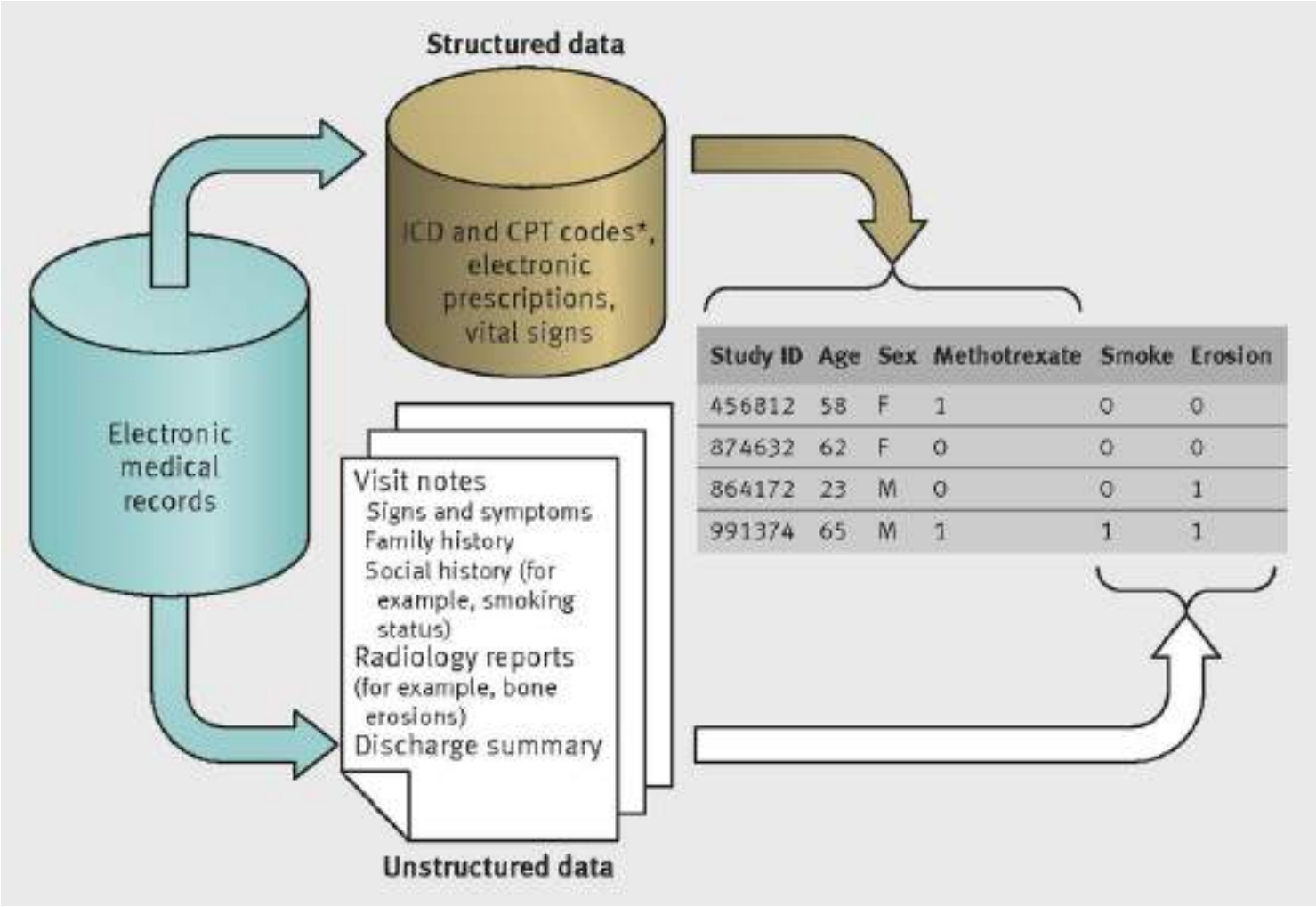
To be **Interoperable**:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be **Reusable**:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

Using Standards Allows the Same Vocabulary



How does a Computer Read a Medical Record?

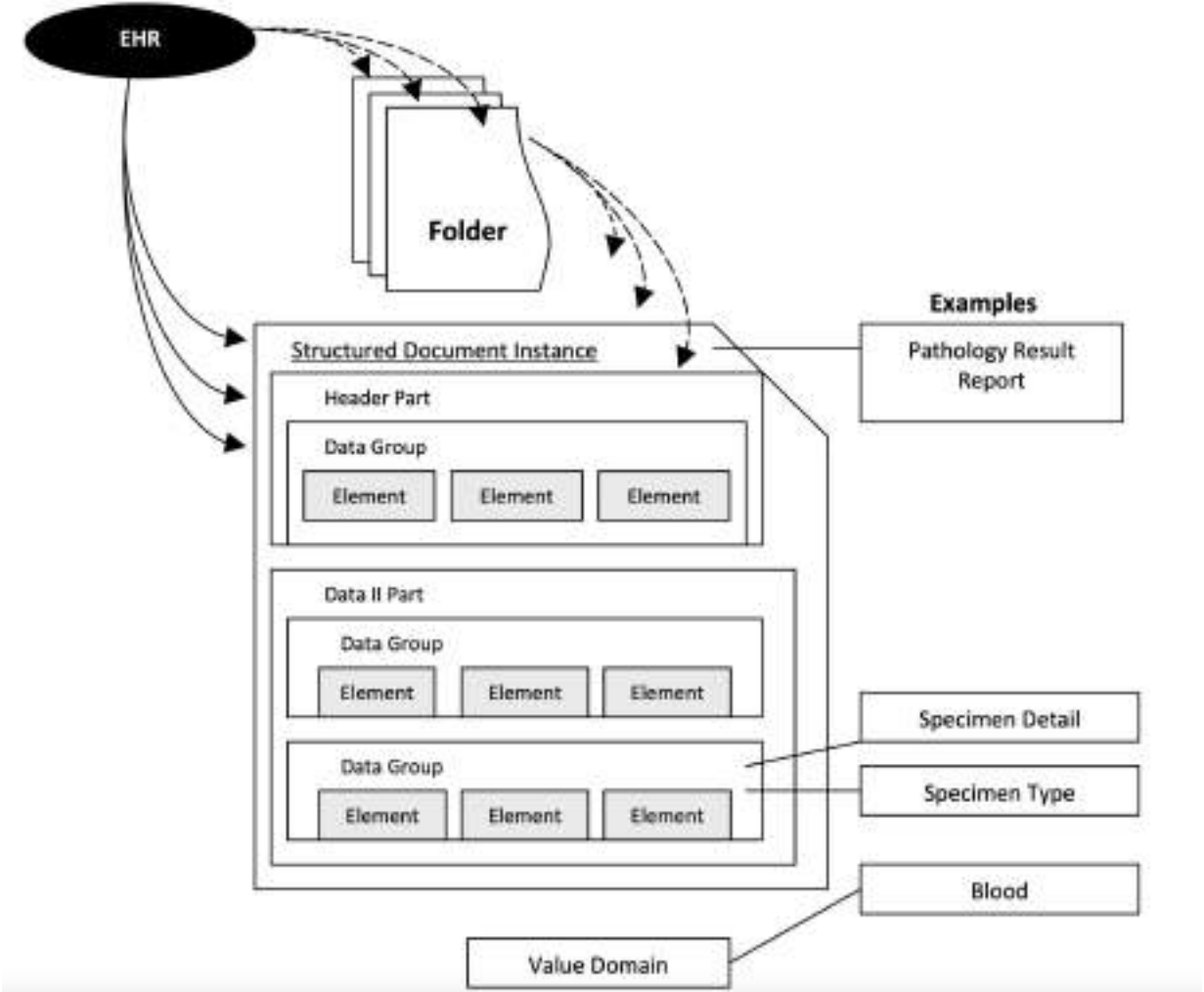

Structured data
ICD and CPT codes*, electronic prescriptions, vital signs
Electronic medical records
Visit notes
Signs and symptoms
Family history
Social history (for example, smoking status)
Radiology reports (for example, bone erosions)
Discharge summary
Unstructured data
Study ID Age Sex Methotrexate Smoke Erosion
456812 58 F 1 0 0
874632 62 F 0 0 0
864172 23 M 0 0 1
991374 65 M 1 1 1

What is an Electronic Health Record

ISO/DTR 20514: “A repository of information regarding the health of a subject of care in **computer processable** form, stored and transmitted **securely**, and accessible by multiple authorized users. It has a **standardized** information **model**, which is independent of EHR systems.

Its primary purpose is the support of continuing efficient and quality integrated healthcare and it contains information, which is **retrospective**, **concurrent**, and **prospective**”

EHR


EHR
Folder
Examples
Structured Document Instance
Pathology Result Report
Header Part
Data Group
Element
Element
Element
Data II Part
Data Group
Element
Element
Element
Specimen Detail
Data Group
Element
Element
Element
Specimen Type
Blood
Value Domain

Purpose of EHR

Australia (HealthConnect)

Initial Health Profile	Pharmacy Provision
Hospital Discharge—Emergency	Community-Based Health Consultation
Medical Consultation—General Practitioner	Allied Health Consultation
Medical Consultation—Specialist	Referral
Diagnostic Investigation—Imaging	Event Notification (for example, admission to hospital)
Diagnostic Investigation—Pathology	Pharmacy Provision

Austria (ELGA)

Primary Goals	**Secondary Goals**
Patient care	Quality management
Patient management	Automating healthcare activities
Streamlining the treatment process	Financial and Administrative processes
A lifelong EHR	Science and Research

What is Interoperability?

Interoperability is ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE 1990)

Levels of Interoperability

- Technical - exchange of information
- Semantic - capability of the recipient to “use” that information
- Process - actual use of information
- Human - clinical interoperability

Perspective | Open Access | Published: 20 August 2019

Why digital medicine depends on interoperability

Moritz Lehne ✉, Julian Sass, Andrea Essenwanger, Josef Schepers & Sylvia Thun

npj Digital Medicine **2**, Article number: 79 (2019) | Cite this article

Abstract

Digital data are anticipated to transform medicine. However, most of today's medical data lack interoperability: hidden in isolated databases, incompatible systems and proprietary software, the data are difficult to exchange, analyze, and interpret. This slows down medical progress, as technologies that rely on these data – artificial intelligence, big data or mobile applications – cannot be used to their full potential. In this article, we argue that interoperability is a prerequisite for the digital innovations envisioned for future medicine. We focus on four areas where interoperable data and IT systems are particularly important: (1) artificial intelligence and big data; (2) medical communication; (3) research; and (4) international cooperation. We discuss how interoperability can facilitate digital transformation in these areas to improve the health and well-being of patients worldwide.

https://www.nature.com/articles/s41746-019-0158-1

HSR Health Services Research

RESEARCH ARTICLE

Sharing information electronically with other hospitals is associated with increased sharing of patients

Jordan Everson PhD MPP ✉ Julia Adler-Milstein PhD

First published: 12 November 2019 | https://doi.org/10.1111/1475-6773.13240 | Citations: 8

Read the full text > PDF TOOLS SHARE

Abstract

Objective

One potential benefit of greater electronic health information exchange is a reduction in the effort required for patients to switch between providers. We therefore assessed whether hospital participation in health information organizations (HIOs) led to increased patient sharing.

Data Sources

Secondary data from 2010 to 2016.

Study Design

https://onlinelibrary.wiley.com/doi/10.1111/1475-6773.13240

Example of Syntactic Interoperability: HL7 Standard

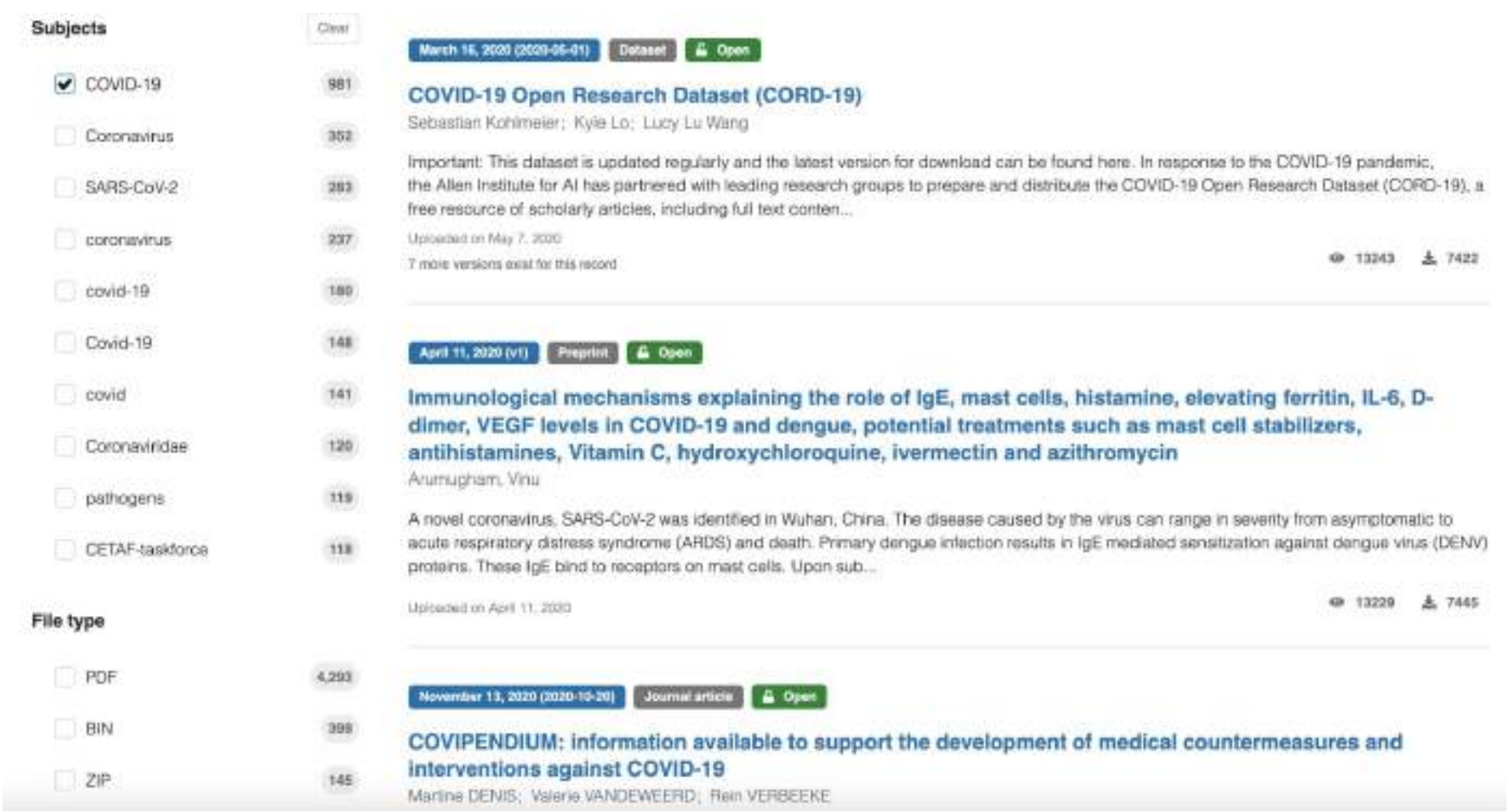
```
##### Translate this function from HL7 V2.8 ADT_A01 to HL7 V2.4 ADT_A05
### HL7 V2.8 ADT_A01
```

```
    MSH|^~\&|ADT1|GOOD HEALTH HOSPITAL|GHH LAB, INC.|GOOD HEALTH
HOSPITAL|198808181126|SECURITY|ADT^A01^ADT_A01|MSG00001|P|2.8||
    EVN|A01|200708181123||
    PID|1||PATID1234^5^M11^ADT1^MR^GOOD HEALTH
HOSPITAL~123456789^^^USSSA^SS||EVERYMAN^ADAM^A^III||19610615|M||C|222
2 HOME STREET^^GREENSBORO^NC^27401-1020|GL|(555) 555-2004|(555)555-
2004||S||PATID12345001^2^M10^ADT1^AN^A|444333333|987654^NC|
    NK1|1|NUCLEAR^NELDA^W|SPO^SPOUSE||||NK^NEXT OF KIN
    PV1|1|I|2000^2012^01||||004777^ATTEND^AARON^A|||SUR||||ADM|A0|

### HL7 V2.4 ADT_A05
```

Example of Syntactic Interoperability: FHIR Standard

```json
{
  "resourceType": "Bundle",
  "type": "message",
  "entry": [
    {
      "resource": {
        "resourceType": "MessageHeader",
        "eventCoding": {
          "system": "http://hl7.org/fhir/message-events",
          "code": "ADT_A01"
        },
        "destination": [
          {
            "endpoint": "http://localhost:8080/fhir/baseDstu3"
          }
        ],
        "source": {
          "name": "ADT1",
          "software": "GOOD HEALTH HOSPITAL",
          "endpoint": "http://localhost:8080/fhir/baseDstu3"
```



https://zenodo.org/search?q=coronavirus&f=subject%3ACOVID-19&l=list&p=1&s=10&sort=mostviewed

https://browse.welch.jhmi.edu/datasets/nih-data-repositories

Hopkins Initiatives

- COVID-19 Precision Medicine Analytics Platform Registry (JH-CROWN)
  The main data source is Johns Hopkins' electronic medical record, Epic. The registry is refreshed weekly with new and updated data and is available for Johns Hopkins investigators to analyze subsets of the COVID-19 patient population for retrospective analyses. CHSOR members Dr. Jodi Segal and Dr. Caleb Alexander have been using these data.
- COVID-19-specific Common Data Model
  PCORnet®, the National Patient-Centered Clinical Research Network, is creating a COVID-19-specific Common Data Model that will allow the use of information from patients across PCORnet's network At Johns Hopkins, Dr. Harold Lehmann leads these activities.
- Johns Hopkins COVID-19 Collaboration Platform
  Over 400 such trials have been registered on clinicaltrials.gov with dozens being added each day. Many of them are designed to answer similar questions and combining data or aggregating evidence could dramatically increase their efficiency and precision, getting answers to doctors faster and more reliably. *more...*
- National COVID Cohort Collaborative
  The National COVID Cohort Collaborative is the partnership among the NCATS-supported Clinical and Translational Science Awards (CTSA) Program hubs and the National Center for Data to Health (CD2H). At Johns Hopkins, Dr. Christopher Chute has taken the lead on this activity.

The Healthcare Data Revolution -> Transformation

- 4CE: Consortium for Clinical Characterization of COVID-19 by EHR
- Figshare: COVID-19 open data
- GitHub: COVID-19 Open Repo Data
- Harvard Dataverse: COVID-19 Data
- ICPSR: COVID-19 Data Repoasitory
- ImmPort: COVID 19
- Mendeley Data: Elsevier COVID-19 Research Environment
- National COVID Cohort Collaborative (N3C)
  OHDSI: Characterizing Health Associated Risks, and Your Baseline Disease In SARS-COV-2 (CHARYBDIS)
- Open-Access Data and Computational Resources to Address COVID-19.
- OpenSAFELY
  Tableau: COVID-19 Data Hub
- Vivli: Covid data
  Zenodo: Coronavirus Disease Research Community - COVID-19
- COVID-19 Research Database

MIMIC (Medical Information Mart for Intensive Care) Database

Medical Information Mart for Intensive Care III (MIMIC-III) is a large, freely-available database comprising de-identified health-related data associated with over 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012.

The database includes information such as demographics, vital sign measurements made at the bedside (~1 data point per hour), laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality (both in and out of hospital).

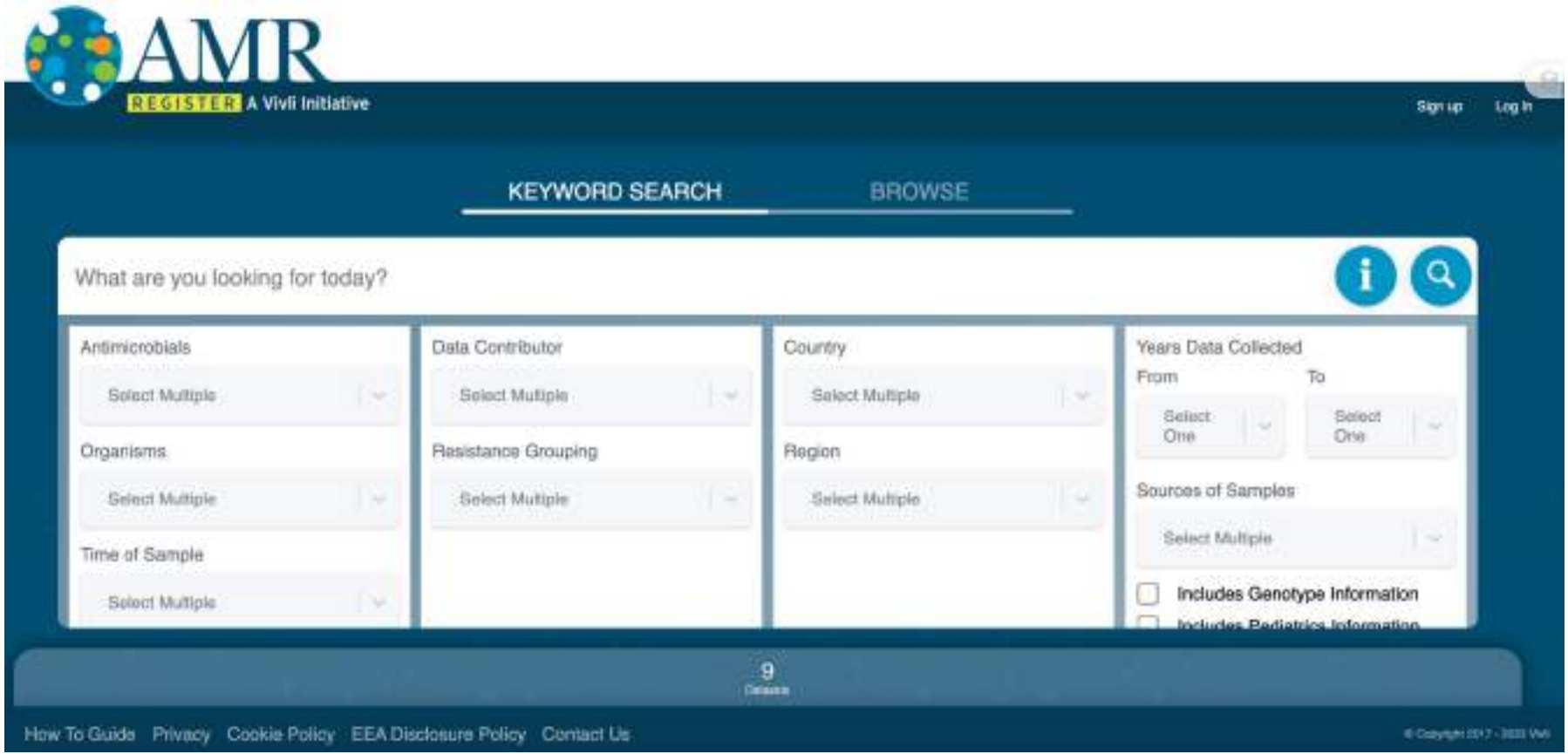
MIMIC-III supports a diverse range of analytic studies spanning epidemiology, clinical decision-rule improvement, and electronic tool development. It is notable for three factors:

- it is freely available to researchers worldwide
- it encompasses a diverse and very large population of ICU patients
- it contains high temporal resolution data including lab results, electronic documentation, and bedside monitor trends and waveforms

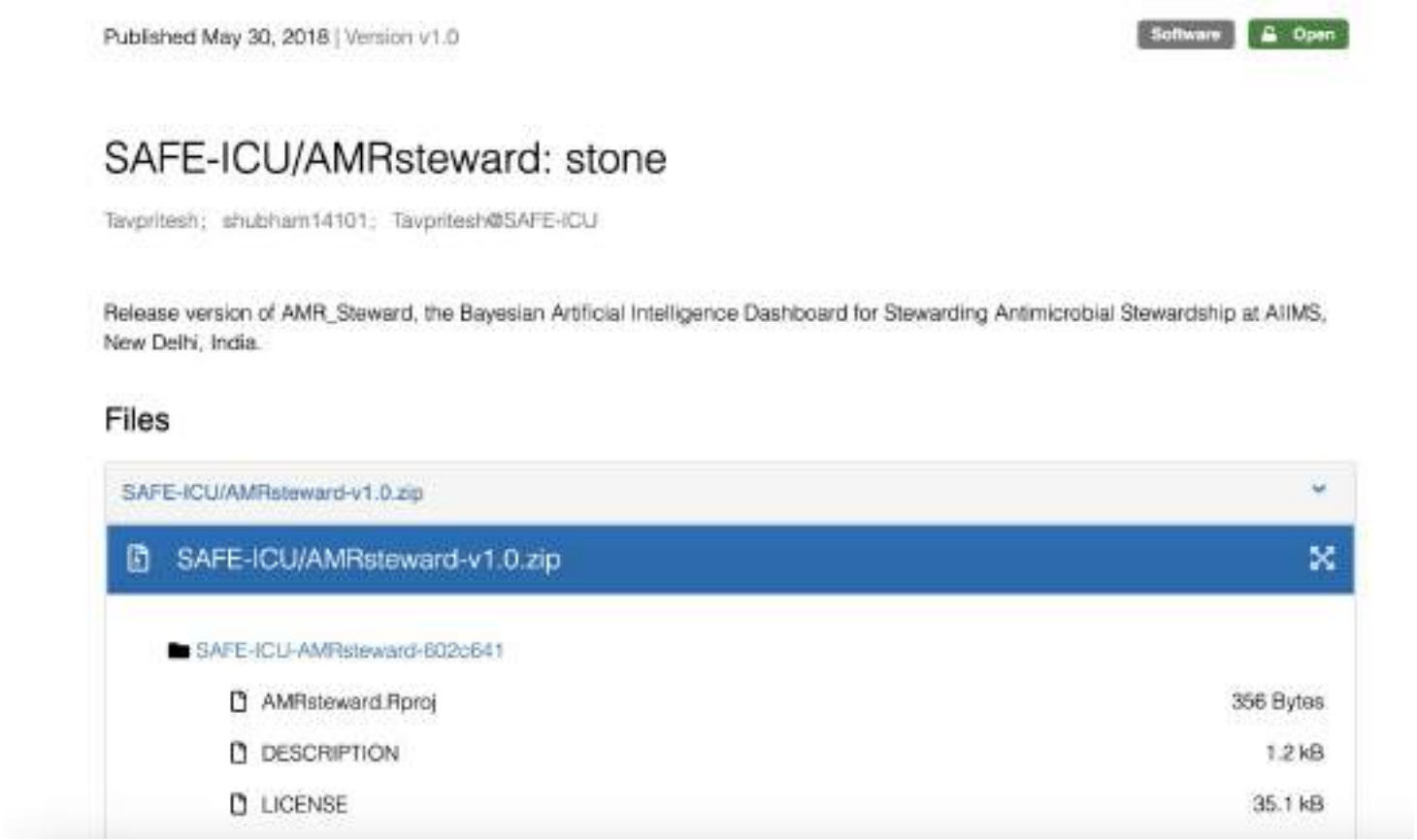
https://browse.welch.jhmi.edu/datasets/ehr-databases

Sources of Health Data

Sources of Health Data



Sources of Health Data



https://zenodo.org/records/1255584

Thanks for attending the class!