

# Electronic data capture

Kristina Fišter (MD, MSc, DSc)

Assistant Professor, Head of the Division for Medical Informatics

University of Zagreb, School of Medicine, Andrija Štampar School of Public Health



BERLIN SCHOOL OF  
PUBLIC HEALTH

# Original research articles

DATA! – the defining element

good data management practices require transparency, integrity, and compliance with regulation



## Introducing: Project Open Data

Posted by Todd Park and Steven Vanhook on May 16, 2013 at 09:46 AM EDT



Technology evolves rapidly, and it can be challenging for policy, and its implementation to evolve at the same pace. Last week, President Obama launched the Administration's new Open Data Policy and Executive Order aimed at ensuring that data released by the government will be as accessible and useful as possible. To ensure this tech-focused policy can keep up with the speed of innovation, we created Project Open Data.

Project Open Data is an online, public repository intended to foster collaboration and promote the continuous improvement of the Open Data Policy. We wanted to foster a culture change in government where we are collaborative and where anyone can help us make open data work better. The project is published on an open source platform that allows communities of developers to collaboratively share and enhance code, resources and plug-and-play tools in Project Open Data can help accelerate the adoption of open data. For example, one tool instantly converts spreadsheets and databases into APIs for easier consumer developers. The idea is that anyone, from Federal agencies to state and local governments to private citizens can freely use and adapt these open source tools—and that's exactly what's happening.

## 113TH CONGRESS 1st Session H. R. 708

To provide for Federal agencies to develop public access to research conducted by employees of that agency or from interred by that agency.

### IN THE HOUSE OF REPRESENTATIVES

FEBRUARY 14, 2013

Mr. DOTY (for himself, Mr. YODER, and Ms. LOFGREN) introducing bill, which was referred to the Committee on Oversight and Reform

### A BILL

To provide for Federal agencies to develop public policies relating to research conducted by that agency or from funds administered by that

1 *Re it enacted by the Senate and House of Representatives,*

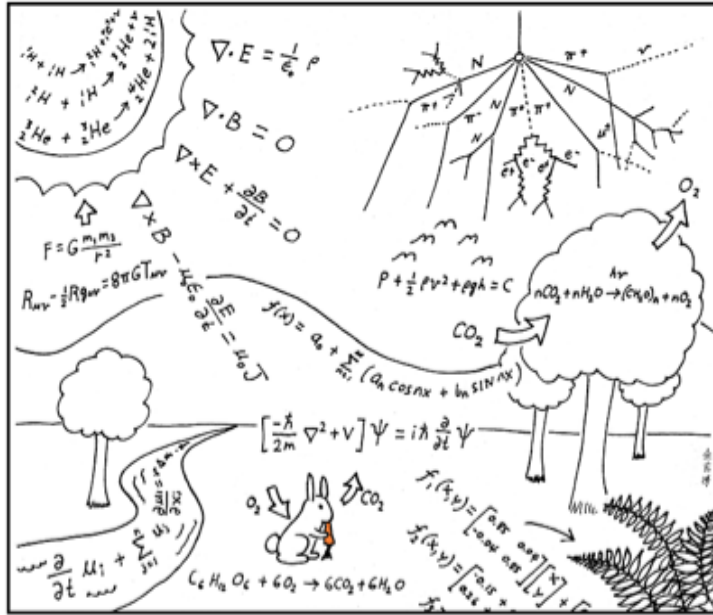
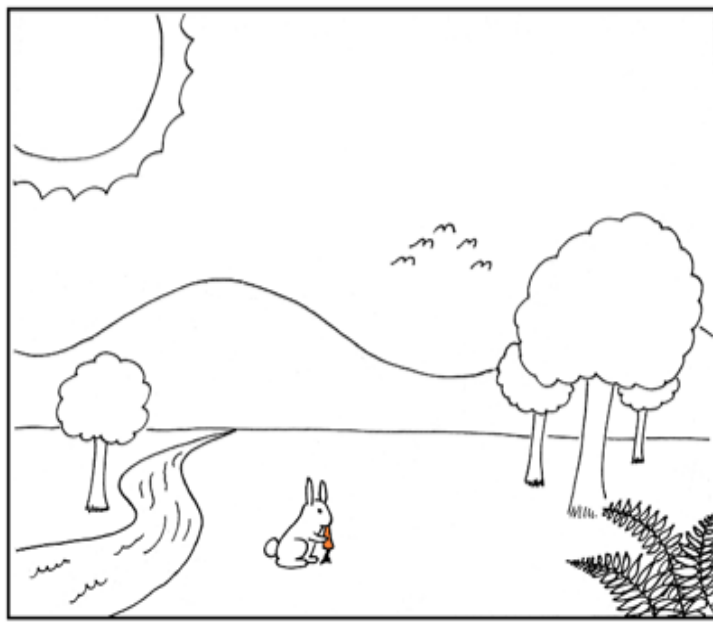


BERLIN SCHOOL OF PUBLIC HEALTH

# Electronic data capture solutions

- help meet regulation requirements in data management
- computerised systems that collect data in electronic format
- electronic data collection forms
- ability to provide remote and real-time access to research data
- automated checks and central revision - improved accuracy
- reduced cost





- **System** - a group of related parts that move or work together
  - clinic
- **Entity** - any object under consideration, such as a patient, visit, medical doctor, health institution, hospital, medication, etc.
  - patient
- **Attribute** - any characteristics of an entity, variable, parameter
  - blood pressure (this is two attributes!)
- **Attribute value (data)** - specific quality or quantity assigned to an attribute, for a specific entity instance (e.g. patient)
  - 160 mmHg (systolic)
  - 80 mmHg (diastolic)

# Two types of attribute values

- **Quantitative** - usually expressed in measurement units, but they can also result from counting
  - systolic pressure = 160 mmHg
  - number of children in a family = 5
- **Qualitative** - expressed by words, categories (mainly agreed in advance)
  - sex = female
  - name = John
  - patient's mobility = immobile



# Qualitative attribute values can be

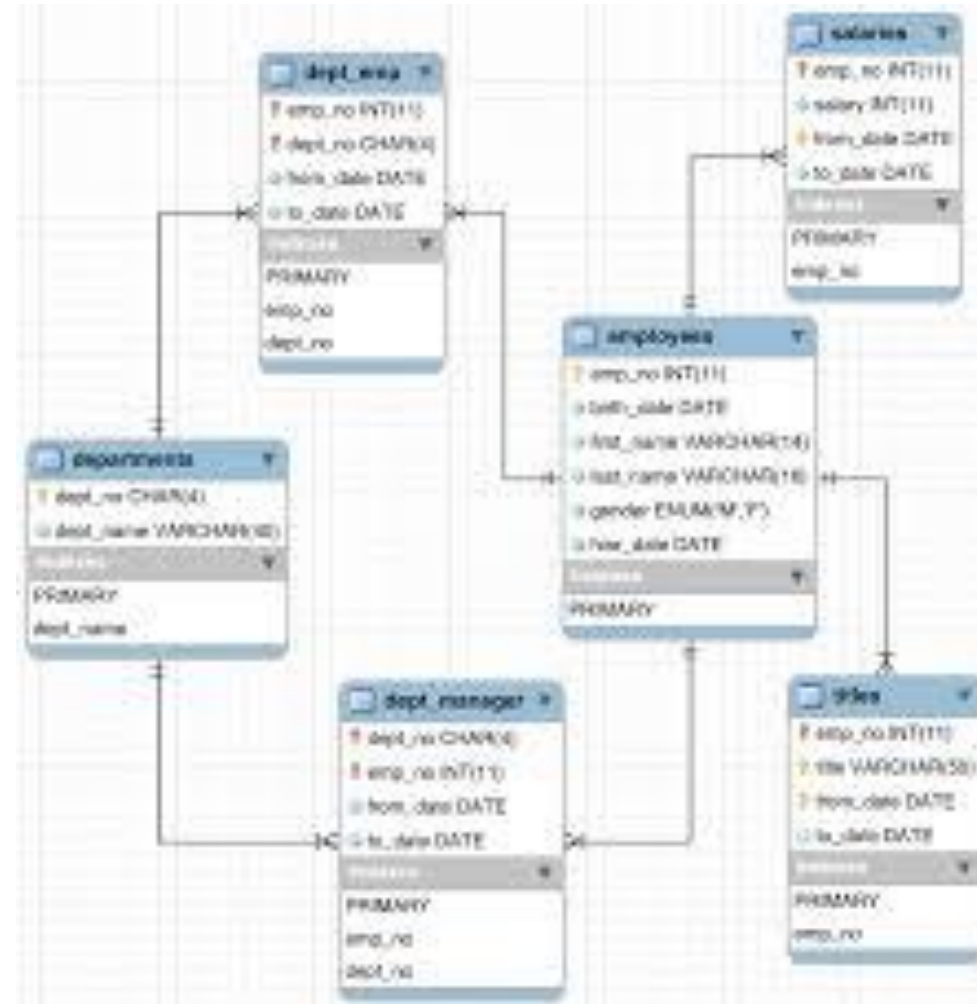
- **Free text** descriptions (medical history, discharge letter, notes, recommendations etc.)
- **Categories** or classes (sex: male, female; marital status: married, single, divorced, widower etc.)
  - structured data
  - code - one or more characters that replace a descriptive attribute value
    - e.g. M for men, W for women
    - ICD-10
    - SNOMED



BERLIN SCHOOL OF  
PUBLIC HEALTH

# Data structures

- a collection of data values, the relationships among them, and the functions or operations that can be applied to the data



<https://dev.mysql.com>



BERLIN SCHOOL OF PUBLIC HEALTH

# How do you collect your data?



BERLIN SCHOOL OF  
PUBLIC HEALTH





her  
Nor  
ree

**Somewhat  
Agree**

**Strongly  
Agree**

**Agree  
Completely**



CHARITÉ  
UNIVERSITÄTSMEDIZIN BERLIN

BERLIN SCHOOL OF  
PUBLIC HEALTH

<https://www.clearpointstrategy.com>

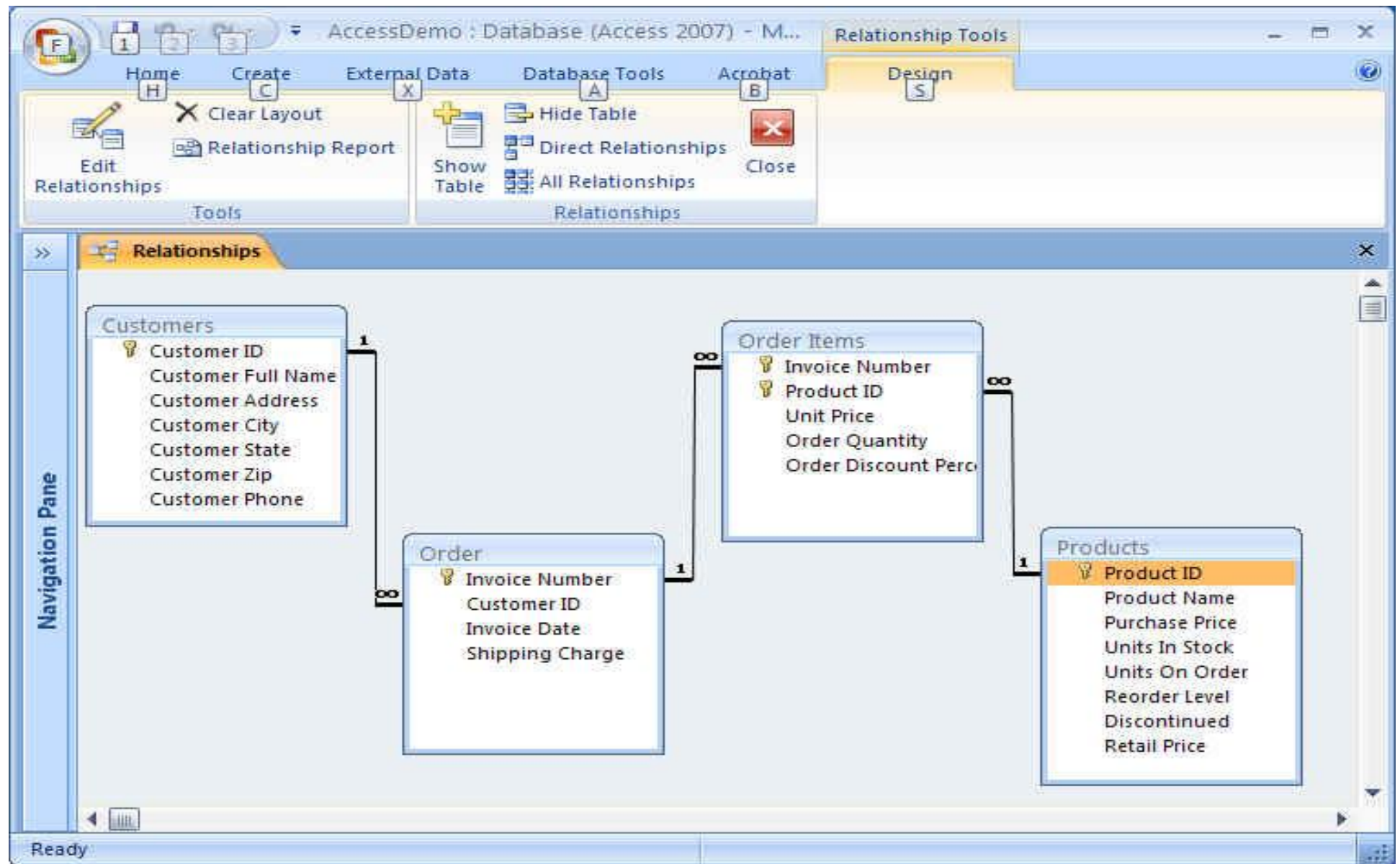
# MS Excel

The screenshot displays the Microsoft Excel interface with the following data table:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB				
2	3	3510	2	40	29	0	0	0	0	1	0				
3	16	3280	2	40	28	0	0	0	1	0	1				
4	30	4220	1	39	27	0	0	0	0	0	0				
5	31	3020	2	40	24	0	0	0	0	1	0				
6	42	3530	2	40	32	1	0	0	0	1	0				
7	51	3850	1	40	31	0	0	0	1	0	0				
8	54	4050	2	40	27	0	0	0	1	1	0				
9	63	3570	1	38	34	0	0	0	0	1	0				
10	71	3540	1	37	28	0	0	0	1	0	0				
11	78	4350	1	41	30	1	0	0	0	0	0				
12	79	2650	1	41	34	0	0	0	0	1	0				
13	81	3050	2	40	23	2	0	0	0	1	0				
14	103	3030	1	38	23	0	0	0	0	1	0				
15	108	3180	1	40	23	0	0	0	0	0	0				
16	124	4100	2	40	35	0	0	1	0	1	0				
17	129	2880	1	39	23	0	0	0	0	0	0				
18	130	4390	2	40	30	0	0	0	0	0	0				
19	131	3090	2	39	25	0	0	0	0	0	0				
20	132	3050	1	38	34	0	0	0	0	0	0				



# MS Access



csus.edu



BERLIN SCHOOL OF PUBLIC HEALTH

If you are doing a survey



BERLIN SCHOOL OF  
PUBLIC HEALTH



What are you looking for? 🔍



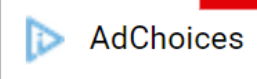
By Ross Rubin October 18, 2018 1:00PM EST



PCMag reviews products **independently**, but we may earn affiliate commissions from buying links on this page. [Terms of use.](#)

Product	Qualtrics	SurveyGizmo	SoGoSurvey	SurveyMonkey	WorldApp KeySurvey	Zoho Survey	Checkbox Survey	SurveyPlanet	Outside Software eSurveysPro	Toluna QuickSurveys
Lowest Price	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT
Editors' Rating	●●●●● EDITORS' CHOICE	●●●●●	●●●●○	●●●●○	●●●●○	●●●●○	●●●○○	●●●○○	●●●○○	●●●○○

Subscribe







Institutions **3481** Countries **131** Projects **706k** Articles **7061**

ABOUT PARTNERS RESOURCES SOFTWARE

### Basic Demography Form

Send custom surveys

Please complete the survey below.

Thank you!

#### Contact Information

1) First Name

2) Last Name

3) Street, City, State, ZIP

Progress indicator: 10 squares, 10th square highlighted.

Expand



CHARITÉ BERLIN SCHOOL OF PUBLIC HEALTH

# Education in biomedical informatics questionnaire (ID:533432)

Aktiviraj ovaj upitnik
⚙️ Preview survey ▾
✎ Postavke upitnika ▾
✂️ Alati ▾
📄 Prikaži / Izvezi ▾
👤 Survey participants
🔄 Odgovori ▾

- [🏠 Upitnik](#) ◀ ▶
- [📁 Questions and groups:](#) ▾
- [📁 Question explorer ▾](#)
- [☰ Sljedeća grupa pitanja](#)
- [☰ List questions](#)
- [🔗 Question organizer](#)
- [👤 Survey participants](#)

Education in biomedical informatics questionnaire

## Survey quick actions ➔

### Format:

Question by question
Group by group
  
Prikaži sva pitanja

### Select your template:

LimeSurvey ☰

## Flat and modern

Survey description : Flat and modern

---

### Single Choice Question

Add group	Dodaj pitanje
<span style="font-size: 2em; color: white;">+</span>	<span style="font-size: 2em; color: white;">+</span>

## Molimo recite nam nekoliko osnovnih podataka o sebi

Koje ste godine rođeni?

**!** U ovo polje mogu biti upisani samo brojevi.

Kojega ste spola?

♀  
ženskog

♂  
muškog



BERLIN SCHOOL OF  
PUBLIC HEALTH



**U kojoj mjeri podržavate sljedeće tvrdnje?**  
 ( 1 = nimalo ne podržavam, 5 = potpuno podržavam )

	1	2	3	4	5
Nastavni planovi sadrže dovoljan broj sati	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nastavne planove potrebno je osuvremeniti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Edukaciju iz biomedicinskoinformatičkih predmeta provodi nedovoljan broj nastavnika	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Visoka učilišta nedovoljno prepoznaju potrebu za obrazovanjem iz biomedicinske informatike	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Studenti već dolaze s dovoljnim znanjima i vještinama iz opće informatike	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sudenti ne pokazuju dovoljan interes za obrazovanjem iz biomedicinske informatike	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nastava iz biomedicinske informatike trebala bi se provoditi u završnim semestrima studija	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sveučilišne poslijediplomske stručne studije iz biomedicinske informatike trebalo bi uspostaviti za sve zdravstvene struke	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



BERLIN SCHOOL OF PUBLIC HEALTH

If you are collecting data in a trial  
or an observational study



BERLIN SCHOOL OF  
PUBLIC HEALTH

## Data Entry

### Features

- Repeating Forms and Visits
- Dynamic Skipping and Hiding of Forms/Visits
- Electronic Patient Reported Outcomes (ePRO)
- Multiple Subject Data Entry
- Automatic Subject Exclusion
- Attach Images and Files to eCRF Pages
- Double Data Entry
- Dictionary Coding (MedDRA, WHODrug, and More)
- Offline Data Entry with Excel
- External Data Loading and Mapping
- Complex Skip Logic / Dynamic Questions and Tables
- Calculated Variables (Even Across Forms/Visits)
- Carry-Over Data / Header Labels
- Embed Text or Images Anywhere on Forms
- Link Form Completion Guidelines
- Field-Specific Mouse-Over Help
- Tabular or Graphical Casebook Views
- AE & SAE Log Forms

## Data Cleaning

### Queries (Edit Checks/Data Discrepancies)

- Cross Form and Visit Custom Query Definitions
- Real-Time Hard and Soft Edit Checks at the Point

## eClinical Modules

### Electronic Patient Reported Outcomes (ePRO)

- Patients can Fill Out and Submit Surveys Online
- Submitted Data Instantly Captured in Unified Database

### Randomization

- Generate Subject IDs or Sample Numbers
- Assign Subjects Dynamically to Different Groups/Arms
- Unblind Notifications

### Dictionary Coding

- MedDRA, WHODrug, VedDRA, and More

### File Attachments

- Attach Any File Type up to 10mb per File

## Integration

- Integrate Medrio via our API
- Data Imports/Exports via API
- Sync with External Systems (IVRS, CTMS, etc.)
- Event Based Calls to API

## Paper/Hybrid Studies



BERLIN SCHOOL OF  
PUBLIC HEALTH

# A fully compliant, all-in-one platform for every type of study.

## Electronic Data Capture & Clinical Data Management

Easily design studies on a drag-and-drop interface

Create beautiful, ultra-capable eCRFs featuring real-time edit checks, skip logic, and auto-save

[Learn more](#)

## Randomization & Supply Management

Easily add randomization to your studies

Seamless, integrated user experience

Supports all common randomization methods

Track drug kits and monitory inventory

[Learn more](#)

## Patient reported outcomes



BERLIN SCHOOL OF  
PUBLIC HEALTH



- Alerts & Messages ▾
- Instructions ▾
- Info ▾
- Icon Key -

- Statuses**
- Not Started
  - Scheduled
  - Data Entry Started
  - Stopped
  - Skipped
  - Completed
  - signed
  - Locked
  - Invalid
- Actions**
- View
  - Edit
  - Remove
  - Restore
  - Reassign
  - Sign

## Subject Matrix for NIC5-15 in Subjects with AD (Alzheimer's Disease) ?

Study Subject ID	First Visit	1-Week F/U	2-Week F/U	Monthly F/U	EOS or LF/U	Logs	AE	registration visit	Actions
									Apply Filter Clear Filter
012									
TS10_101_10				x2					
TS10_102_10				x2					
TS10_103_10				x2					
TS10_104_10				x2					
TS10_105_10									
TS11_104_11									
TS11_105_11									
TS12_104_12									
TS12_105_12									
TS13_105_13				x2					
TS1_101_01									
TS1_102_01				x2					
TS1_103_01									
TS1_104_01									

Results 1 - 15 of 56.

github/OpenClinica

# Sample Physical Exam English

## ▼ CRF Header Info

Click the flag icon next to an input to enter/view discrepancy notes. Please note that you can only save the notes if CRF data entry has already started.

Exit

I Basic (0/9)


II Body...(0/35)

III Oth...(0/3)

-- Select to Jump --

### Title: Basic Information

#### Visit Information:

1 Date of Physical Examination:   2 Time of Physical Examination:  (HH:MM)

#### Physical Exam Information:

3 Height:  (in) 4 Weight:  (lb)

5 Temperature:  (F) 6 Pulse Rate:  (per min)

7 Respiration Rate:  (per min)

#### Blood pressure:

8 Systolic:  (mm) Diastolic:  (Hg)



CHARITÉ BERLIN SCHOOL OF  
PUBLIC HEALTH

If you are using routinely  
collected health care data



BERLIN SCHOOL OF  
PUBLIC HEALTH



## Software Demos

---

See for yourself, try any of these demos of open source analytics tools:

**ATLAS** – a web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis:



**ACHILLES** – a standardized database profiling tool for database characterization and data quality assessment:



---

## Open-Source Software







## Demo\_data\_1K\_synthetic\_patients

## Achilles Heel Report

## Data Quality Messages

Search: 

Message Type	Message
ERROR	400-Number of persons with at least one condition occurrence, by condition_concept_id; 12 concepts in data are not in correct vocabulary
ERROR	510-Number of death records outside valid observation period; count (n=1) should not be > 0
ERROR	600-Number of persons with at least one procedure occurrence, by procedure_concept_id; 8 concepts in data are not in correct vocabulary
ERROR	1,000-Number of persons with at least one condition era, by condition_concept_id; 12 concepts in data are not in correct vocabulary
NOTIFICATION	percentage of non-numerical measurement records exceeds general population threshold
NOTIFICATION	Percentage of patients with no visits exceeds threshold
NOTIFICATION	[GeneralPopulationOnly] Not all deciles represented at first observation
WARNING	4-Number of persons by race; data with unmapped concepts
WARNING	5-Number of persons by ethnicity; data with unmapped concepts
WARNING	200-Number of persons with at least one visit occurrence, by visit_concept_id; data with unmapped concepts
WARNING	402-Number of persons by condition occurrence start month, by condition_concept_id; 5 concepts have a 100% change in monthly count of events
WARNING	500-Number of persons with death, by cause_concept_id; data with unmapped concepts
WARNING	511-Distribution of time from death to last condition (count = 1); max value should not be positive, otherwise its a zombie with data >1mo after death
WARNING	512-Distribution of time from death to last drug (count = 1); max value should not be positive, otherwise its a zombie with data >1mo after death
WARNING	513-Distribution of time from death to last visit (count = 1); max value should not be positive, otherwise its a zombie with data >1mo after death

Showing 1 to 15 of 21 entries

Print

Previous

1

2

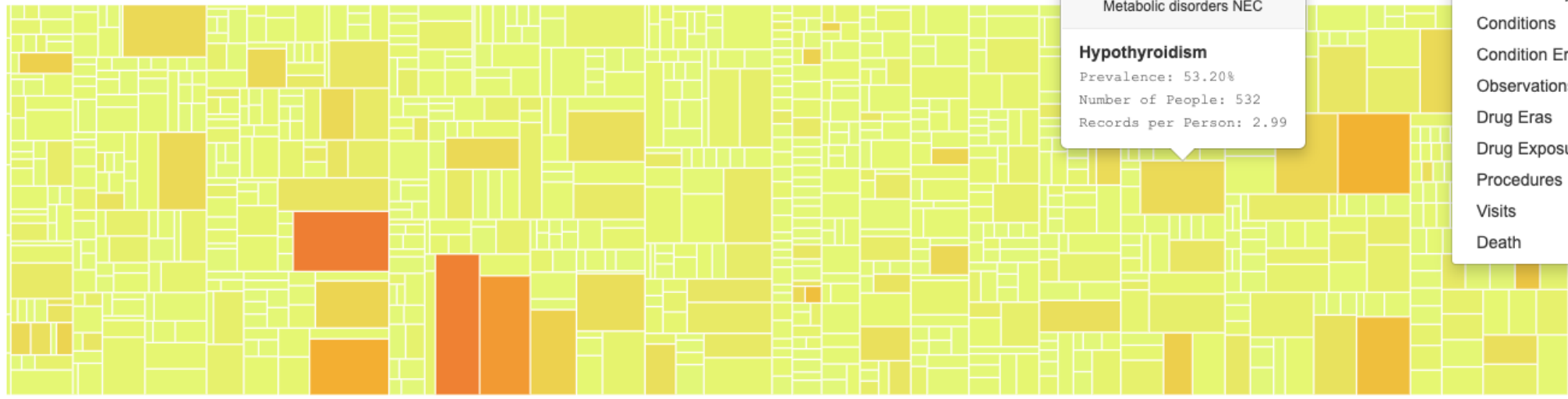
Next



### Conditions

#### Condition Prevalence

Treemap [Table](#)



NA  
NA  
Metabolism disorders NEC  
Metabolic disorders NEC

**Hypothyroidism**  
Prevalence: 53.20%  
Number of People: 532  
Records per Person: 2.99

- Dashboard
- Achilles Heel
- Person
- Observation Periods
- Data Density
- Conditions
- Condition Eras
- Observations
- Drug Eras
- Drug Exposures
- Procedures
- Visits
- Death

### Death

#### Death Prevalence by Age, Gender, Year

MALE FEMALE

#### Age Decile



## Cohort

Aspirin-Osteoporosis

Save

Close

Copy

Delete

Definition

Concept Sets

Generation

Reporting

Explore

Export

New Concept Set

Export All Concept Sets To CSV

Show 10 entries

Filter Repository Concept Sets:

Id	Title
0	Aspirin
1	Osteoporosis

Showing 1 to 2 of 2 entries

Previous 1 Next

Concept Set Expression

Included Concepts 2200

Included Source Codes

Export

Name:

Aspirin

Show 25 entries

Search:

Showing 1 to 1 of 1 entries

Previous 1 Next

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
1112807	1191	Aspirin	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Classification Non-Standard Standard

Delete Concept Set

Close Concept Set



CHARITÉ  
UNIVERSITÄTSMEDIZIN BERLIN

BERLIN SCHOOL OF  
PUBLIC HEALTH

# SNOMED

- Systematised Nomenclature of Medicine, from non-profit International Health Terminology Standards Development Organization (IHTSDO)
- General-purpose, comprehensive and computer-processable terminology suitable to code events and elements within the electronic health record
- Used in more than 40 countries, primarily in labs for coding of reports to generate statistics and facilitate data retrieval but increasingly also to code clinical data



# SNOMED (2)

- 1st level: multi-axial classification
  - 11 rules that define **modules**
- Remaining levels: hierarchical classification

## Module designator

Topography (T)

Morphology (M)

Function (F)

Diseases/Diagnoses (D)

Procedures (P)

Occupations (J)

Living Organisms (L)

Chemicals, Drugs and Biological Products (C)

Physical Agents, Forces and Activities (A)

Social Context (S)

General Linkage-Modifiers (G)



BERLIN SCHOOL OF  
PUBLIC HEALTH

# Allows the composition of complex terms from simpler terms

However exact rules for this haven't been developed so validity for clinical or research purposes is problematic

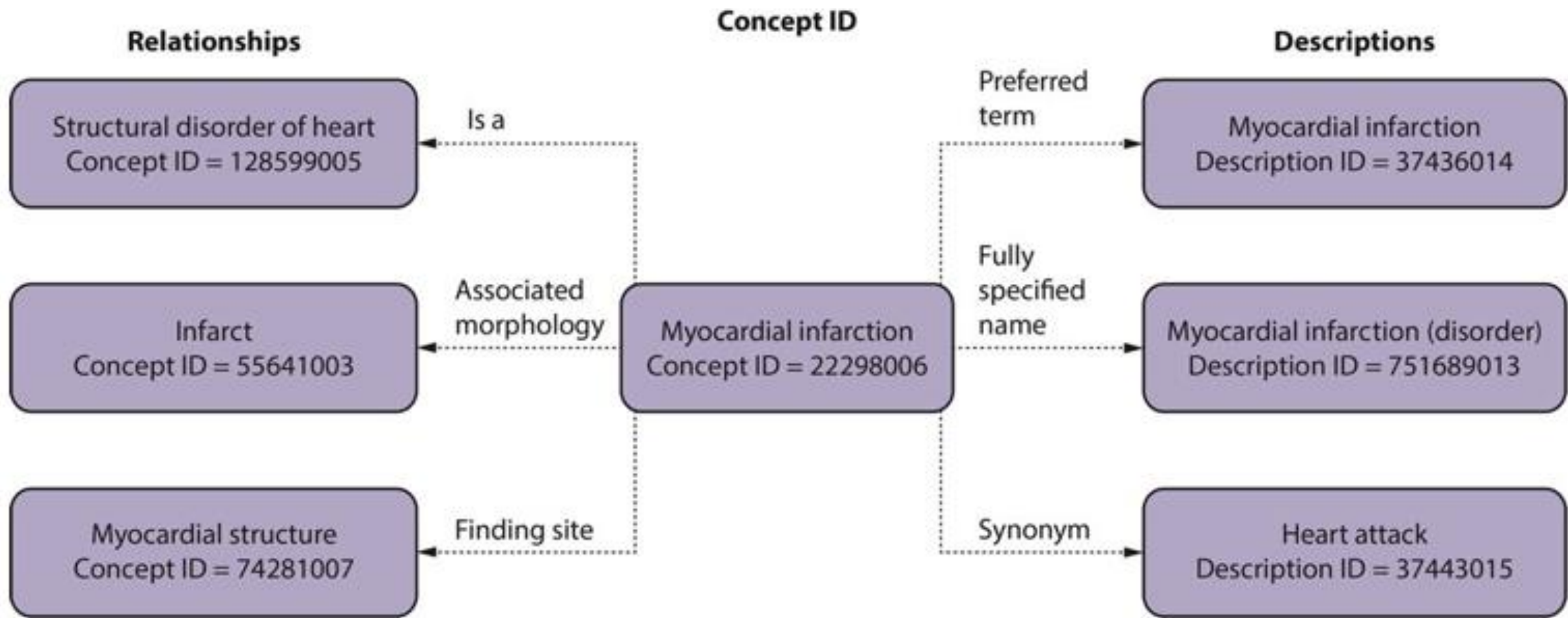
- Example:
  - Appendicitis (D5-46210)
  - Can also be coded as
    - Appendicitis, acute
    - Acute inflammation, in, Appendix
    - Acute, inflammation NOS (not otherwise specified), in, Appendix
- SNOMED CT was developed to solve this problem, and support the electronic storage, retrieval and analysis of clinical data



# SNOMED CT

- SNOMED CT (clinical terms) is taking over for use in electronic health records, decision support systems, and communication between applications
- Wider clinical scope, better context for the terms
- Supports multilingual terminological renderings of common concepts (aimed for international use)





*Enrico Coiera*



BERLIN SCHOOL OF  
PUBLIC HEALTH





OHDSI community



BERLIN SCHOOL OF PUBLIC HEALTH

## Day 1: Wednesday 22nd May 2019

10.00-11.30	What editors want (K. Fišter, Lecture)
11.45-13.15	Electronic data capture (K. Fišter, Lecture with demonstration)
14.15-15.45	Searching biomedical literature (D. Relić and P. Hrabač, Practical)
16.00-17.30	Managing citations and references (D. Relić and P. Hrabač, Practical)

## Day 2: Thursday 23rd May 2019

10.00-11.30	Common study designs (T. Kurth, Lecture)
11.45-13.15	Introduction to R/R Studio interface (J. Rohmann and M. Piccininni, Practical)
14.15-15.45	Causal inference (T. Kurth, Lecture)
16.00-17.30	Descriptive statistics using R (J. Rohmann and M. Piccininni, Practical)

## Day 3: Friday 24th May 2019

10.00-11.30	Analytical tools, epidemiological/statistical methods (T. Kurth, Lecture)
11.45-13.15	Data analysis in R: part 1 (J. Rohmann and M. Piccininni, Practical)
14.15-15.45	Data analysis in R: part 2 (J. Rohmann and M. Piccininni, Practical)
16.00-17.30	How to write: from protocol to research paper (K. Fišter, Lecture)
17.45-18.30	Exam



BERLIN SCHOOL OF  
PUBLIC HEALTH