Electronic data capture

Kristina Fišter (MD, MSc, DSc)

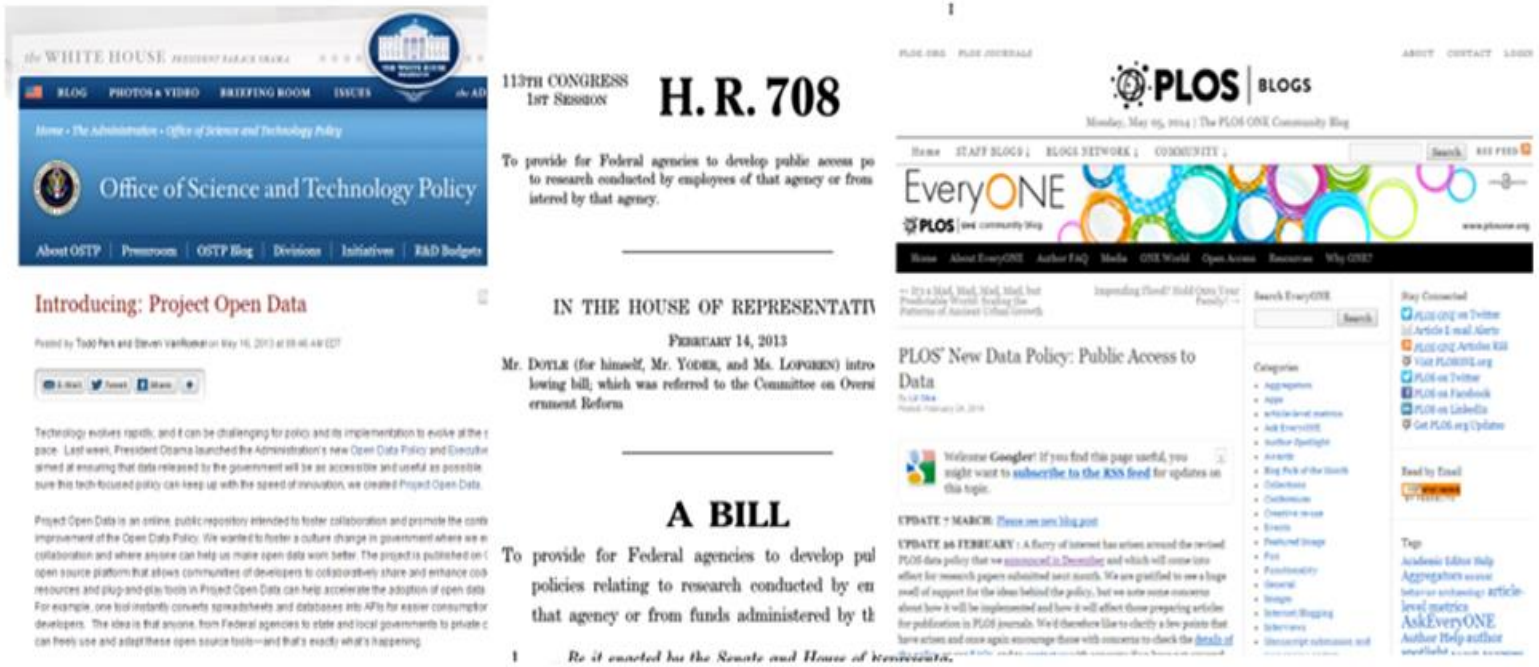
Assistant Professor, Head of the Division for Medical Informatics

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

Original research articles

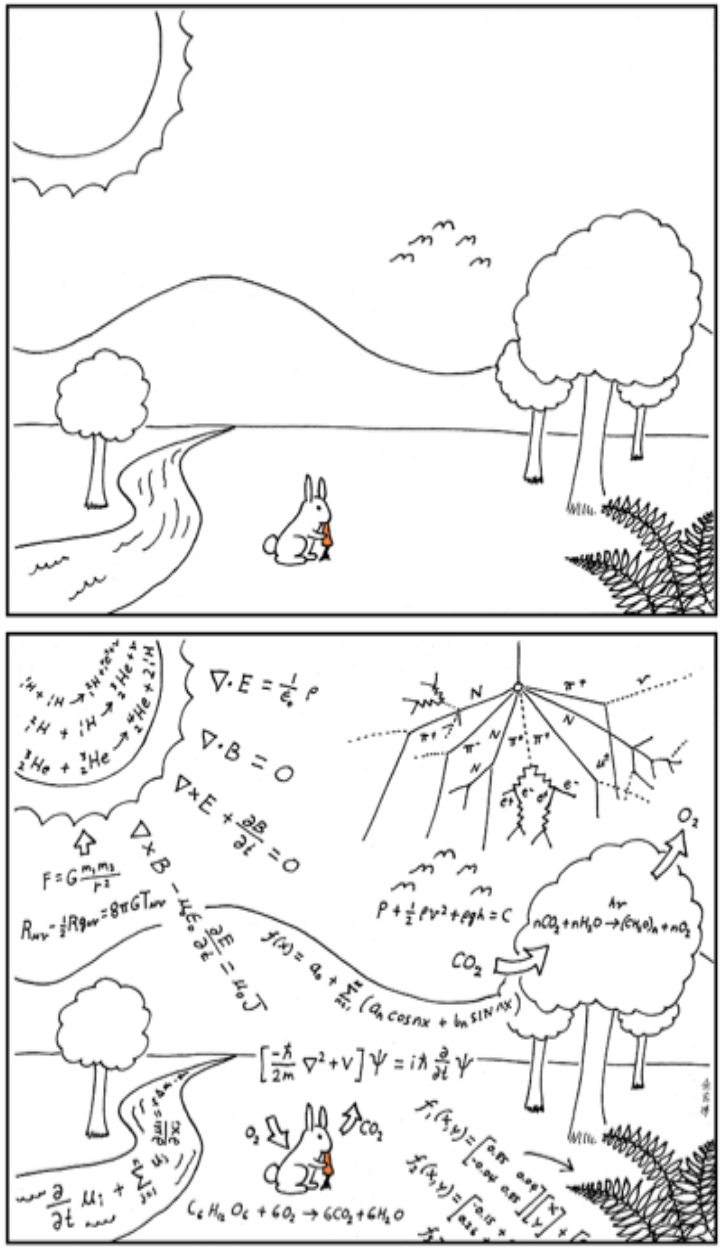
DATA! – the defining element

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



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

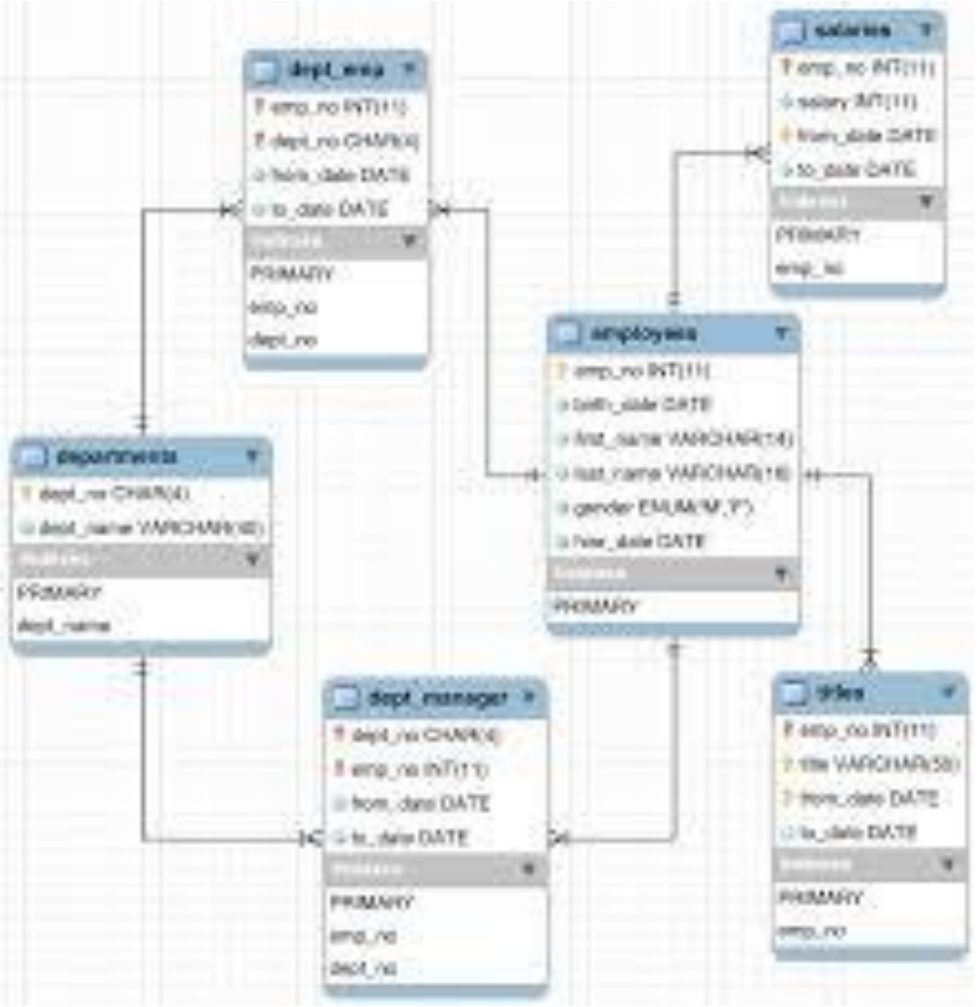
- **Quantitative** - usually expressed in measurement units, but they can also result from counting
    - systolic pressure = 160 mmHg
    - number of children in a family = 5
- **Qualitative** - expressed by words, categories (mainly agreed in advance)
    - sex = female
    - name = John
    - patient's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

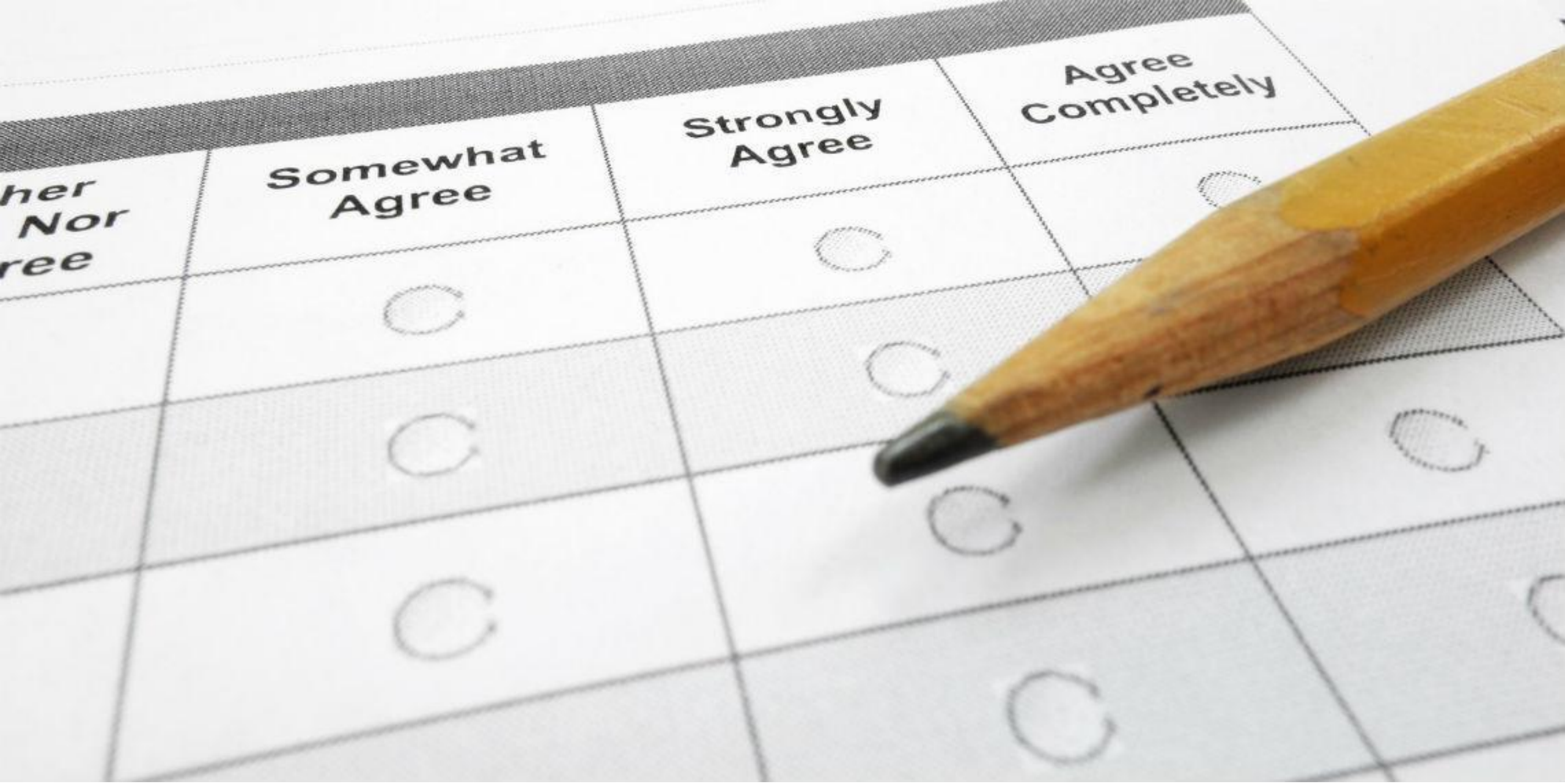
Data structures

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



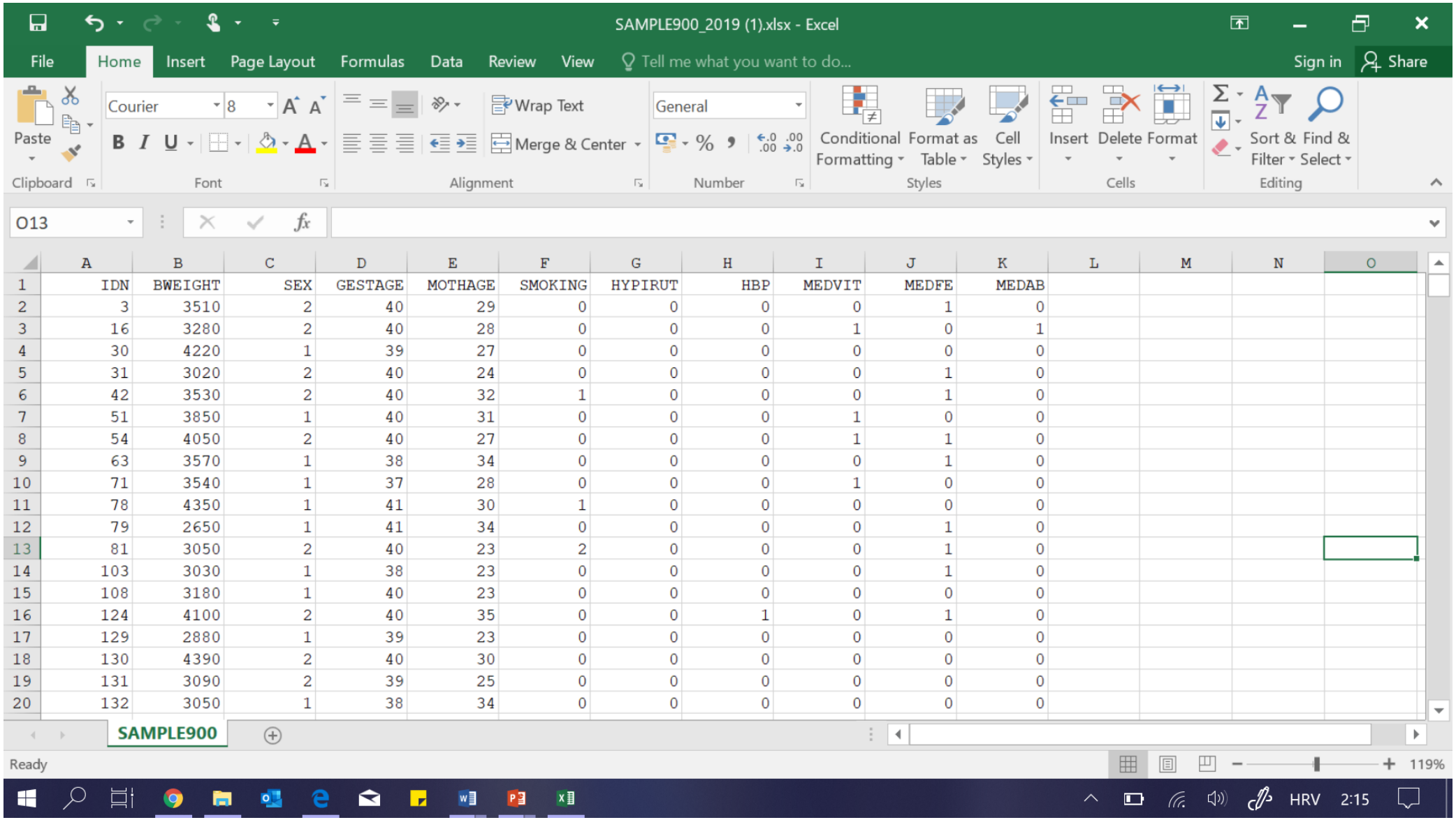
https://dev.mysql.com

How do you collect your data?

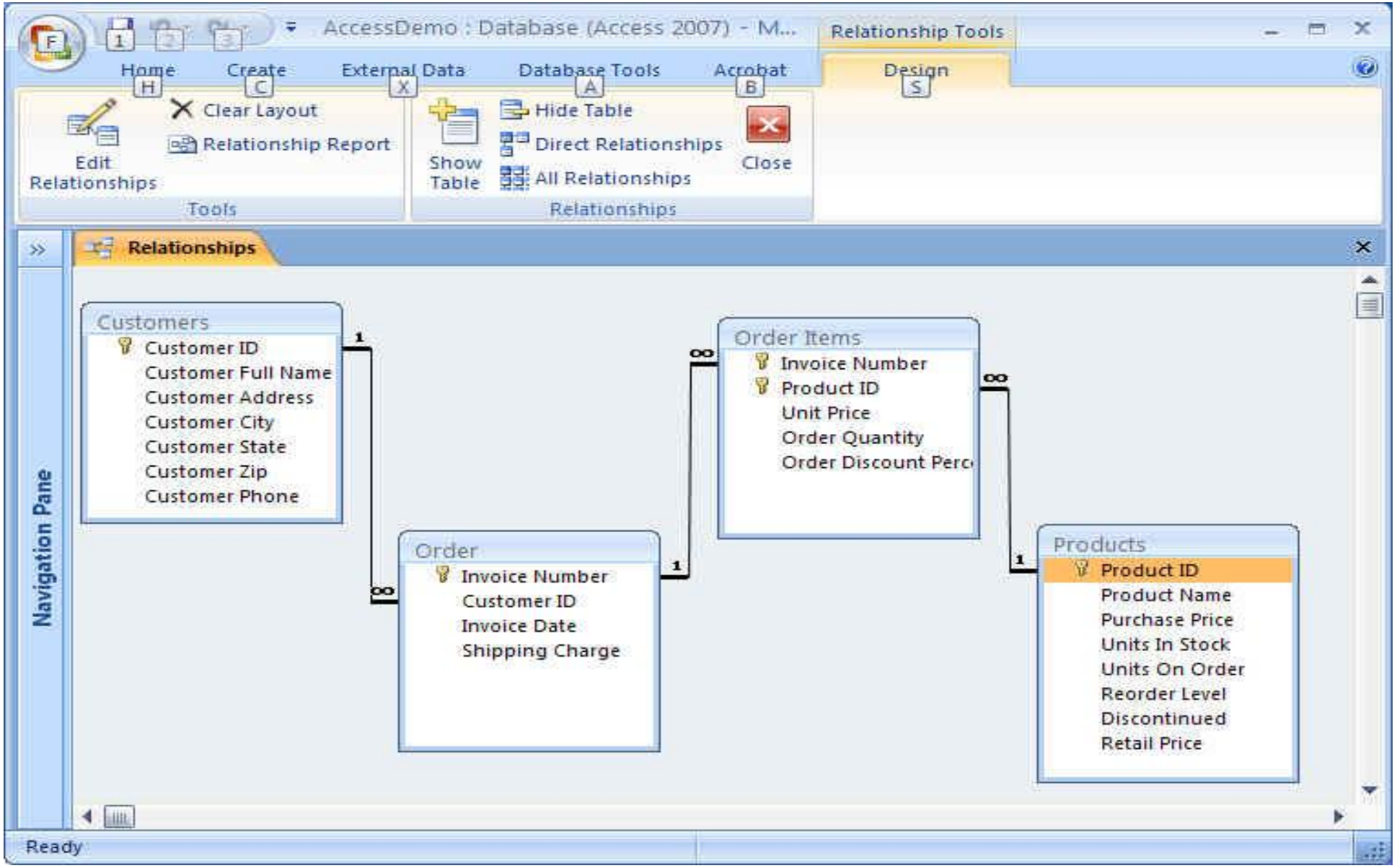
https://www.clearpointstrategy.com

MS Excel



	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

If you are doing a survey


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

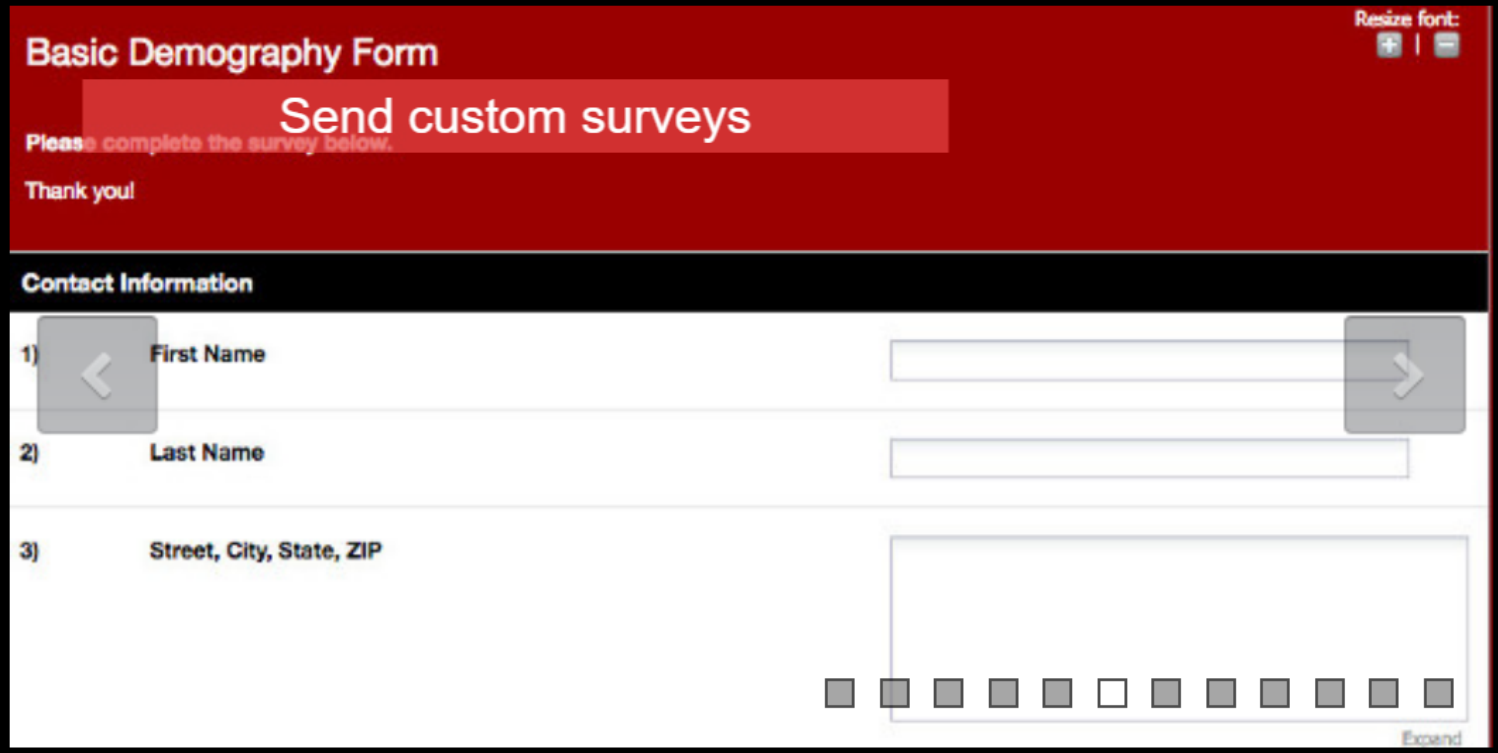
Product	Qualtrics	SurveyGizmo	SoGoSurvey	SurveyMonkey	WorldApp KeySurvey	Zoho Survey	Checkbox Survey	SurveyPlanet	Outside Software eSurveysPro	Toluna QuickSurveys
	qualtrics	surveygizmo	SoGoSurvey	SurveyMonkey	KeySurvey	ZOHO	Checkbox	surveyplanet	eSurveysPro	Toluna
Lowest Price	SEE IT	SEE IT	SEE IT	SEE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

https://www.project-redcap.org

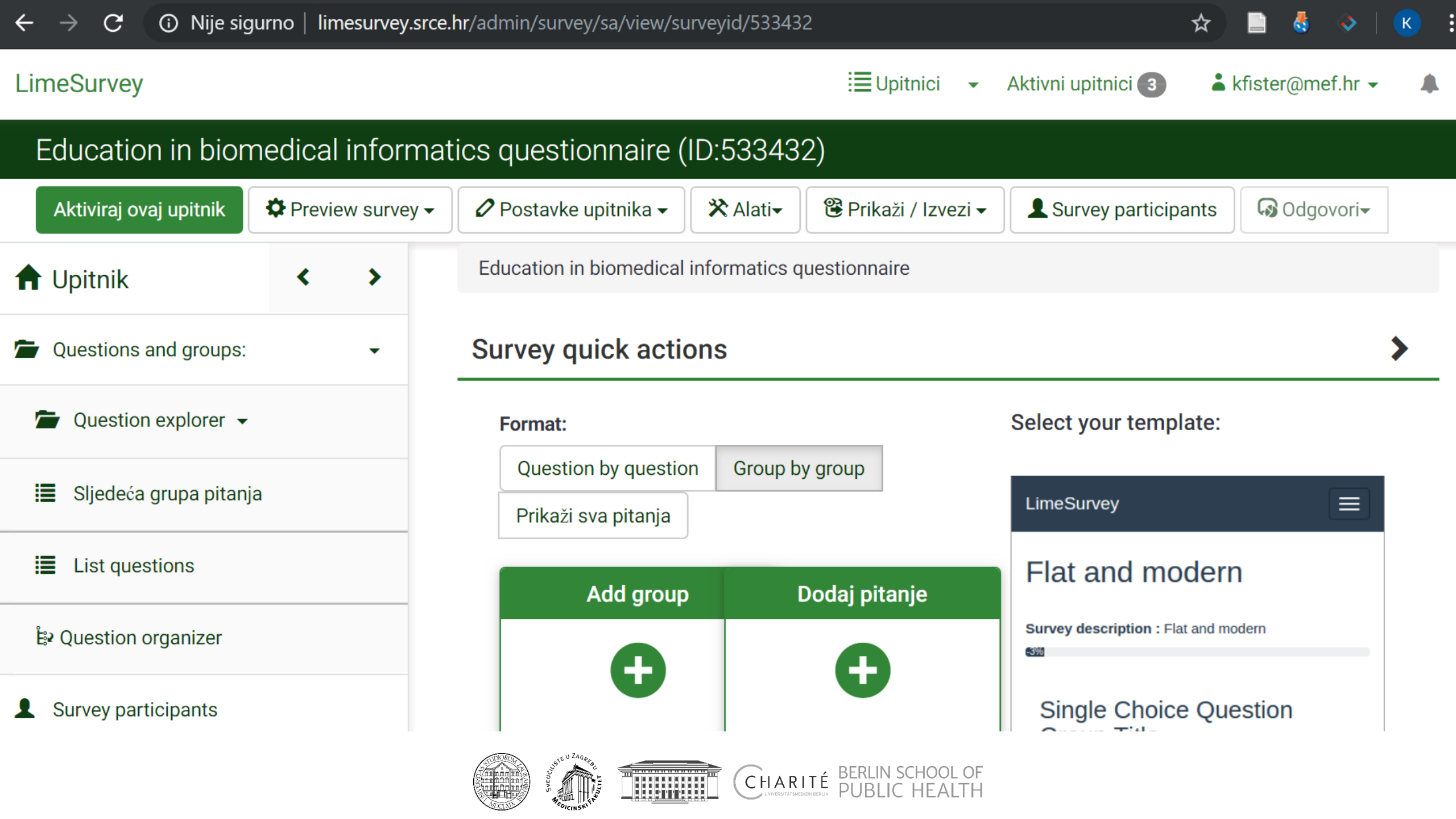
REDCap
Research Electronic Data Capture

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

ABOUT | PARTNERS | RESOURCES | SOFTWARE



STANFORD UNIVERSITY | SVEUČILIŠTE U ZAGREBU MEDICINSKI FAKULTET | CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN | BERLIN SCHOOL OF PUBLIC HEALTH


Nije sigurno | limesurvey.srce.hr/admin/survey/sa/view/surveyid/533432
LimeSurvey
Upitnici
Aktivni upitnici 3
kfister@mef.hr
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
Add group
Dodaj pitanje
Select your template:
LimeSurvey
Flat and modern
Survey description : Flat and modern
Single Choice Question
CHARITÉ
BERLIN SCHOOL OF PUBLIC HEALTH

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

U kojoj mjeri podržavate sljedeće tvrdnje?

(1 = nimalo ne podržavam, 5 = potpuno podržavam)

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

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

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 Typeup 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

medrio

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN 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

github/OpenClinica

CHARITÉ BERLIN SCHOOL OF PUBLIC HEALTH

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)

If you are using routinely collected health care data




BERLIN SCHOOL OF PUBLIC HEALTH

https://www.ohdsi.org/analytic-tools/

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 browing, cohort definition, and population-level analysis:

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

Open-Source Software

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN BERLIN SCHOOL OF PUBLIC HEALTH

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

Software – OHDSI | Achilles | ATLAS | Inbox (328) - kristina.fishter@

Not Secure | www.ohdsi.org/web/achilles/#/Demo_data_1K_synthetic_patients/death

Error

Achilles

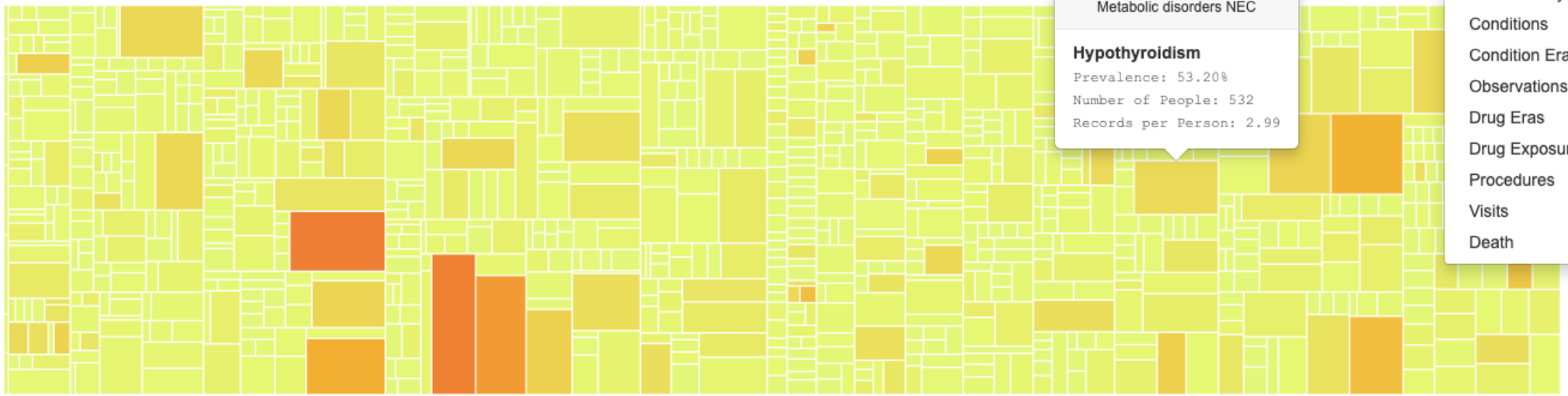
Data Sources | Reports

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

Conditions

Condition Prevalence

Treemap | Table

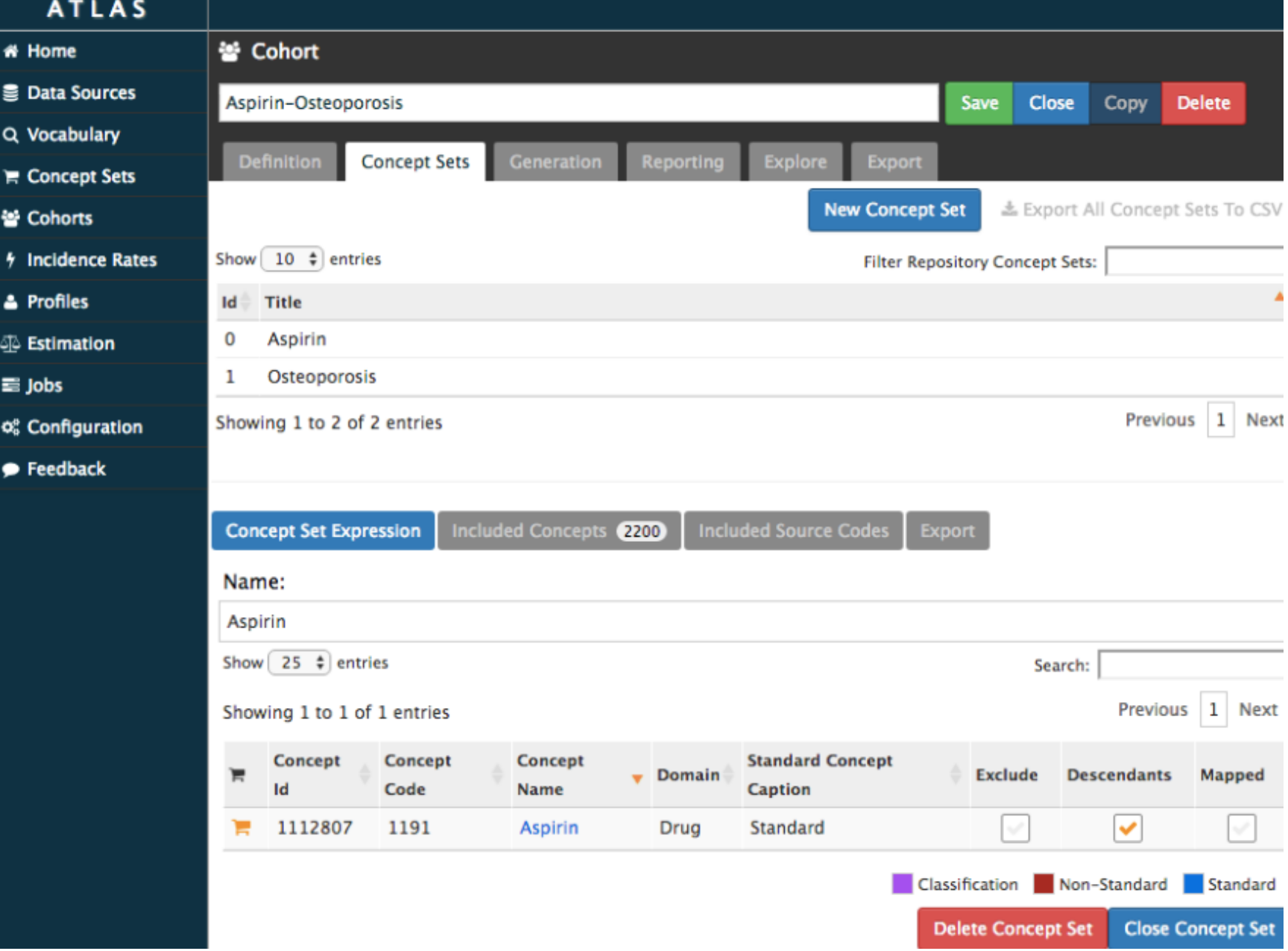


Box Size: Prevalence, Color: Records per Person (Blue to Orange = Low to High), Use Ctrl-Click to Zoom, Alt-Click to Reset Zoom

Death

Death Prevalence by Age, Gender, Year

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN | BERLIN SCHOOL OF PUBLIC HEALTH



CHARITÉ 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)

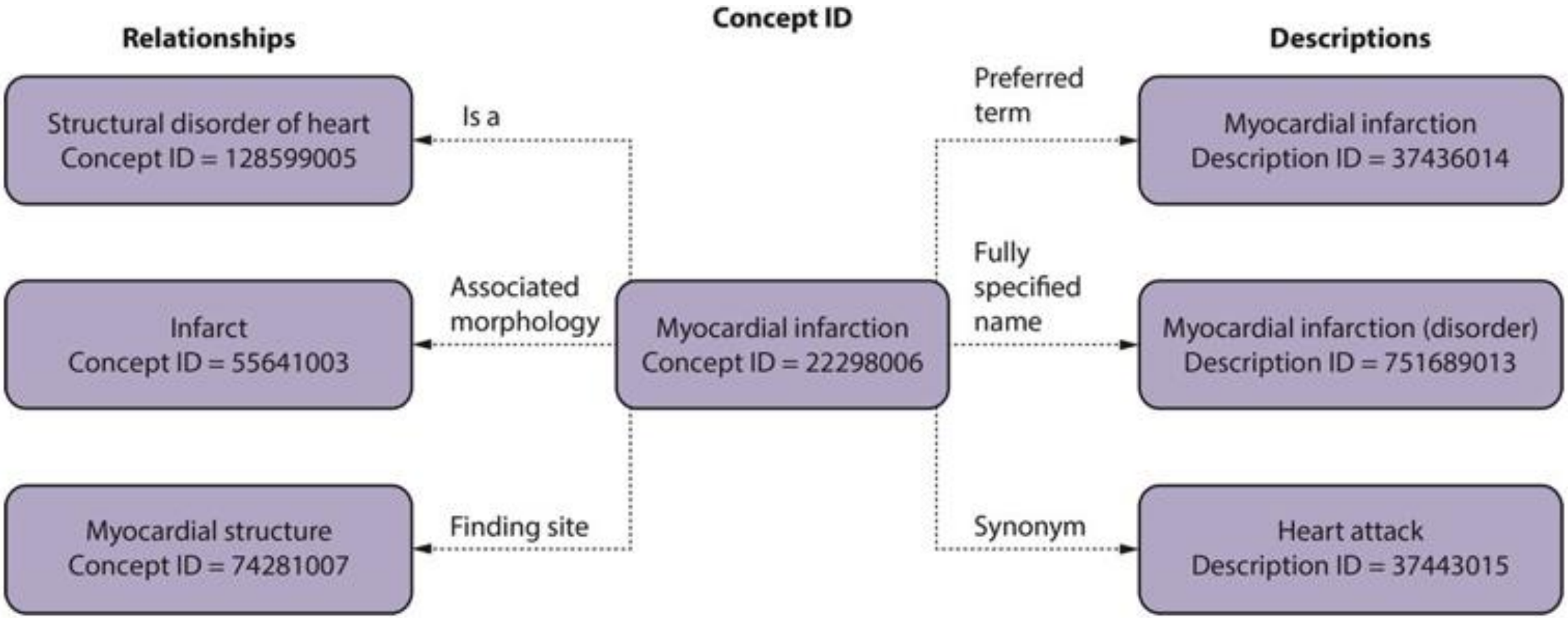
Allows the composition of complex terms from simpler terms

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

OHDSI community

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN
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