

Clinical Data

<u>Click here for: Book</u>

B

С

Color index:

🔆 🛛 Golden note

- Doctors notes

Editing File

0

- **Data:** are symbols or observations reflecting differences in the world. (ex: 5)
- **Information:** is data with meaning. (ex: 5 fingers)
- **Knowledge:** is information that is justifiably believed to be true. (ex: elevation in FBG level, the likelihood of diabetes is increased, smokers are likely to develop lung cancer more than non-smokers)

Data

• **Wisdom:** is the critical use of knowledge to make intelligent decisions.

Each zero or one is a bit, each 8 bits equal a byte



Is important to recognize that neither data types nor file formats define the meaning of the data.

Types of clinical data					
Narrative	Numerical measurements	Coded data			
recording by clinician- maternity history.	blood pressure, temperature, glucose level.	selection from a controlled terminology system, anything that has a drop down menu (example being the term MI that may mean myocardial infarction or mitral insufficiency).			
Textual data	Recorded signals	Pictures			
other results reported as text.	EKG, EEG.	radiographs, photographs, and other images.			

Clinical Data

Clinical data: are a collection of observations about a patient. **A datum:** is a single observation of a patient.

General categories of data entry:

- **Free-form** entry by historical methods:
 - Writing
 - Dictation
 - Typing
- Structured (menu-driven) data entry by mouse or pen.
- **Speech recognition** for either of above.

Informatics and Computer Science

- The domain of:
 - Computer scientists is **Data**
 - Informaticians is **information**
- Information retrieval involves both computer science (data) and informatics (information)
 - Retrieving all the documents related to Aspirin & heart attacks.

Artificial Intelligence (AI)

- AI is concerned with the development of systems that can do something that previously required human intelligence. (chess, driving vehicles)
- **Example in medicine:** derma skin lesions categorization
 - A system is given large sets of images which are labeled (benign, malignant,...)
 - The system tries to learn to be able to identify lesions in the features
- Limitations:
 - Require large sets of Data
 - Cannot justify judgment (why did it classify this picture as benign?)



Clinical Data

Data to Information:



ICD: International Classification of Disease, 10 is the version.

Interoperability: Ability of two or more systems or components to exchange the information and use the information that has been exchanged. (ex: medical record at hospital A, sends information to a medical record at hospital B without losing this information).

Information to Knowledge:

Two types of data:

- **Structured data** (from *drop down menus* such as; billing codes, lab results, problem nests and medication nests).
- Unstructured (free text) (simple human language ex: history).

OB TEST DOB 2	8/2/2008 64 26+2	BW 1070				Linemal 217
IDN 122422	Day 22 Compared 64	100 1280	104			Civer poor 227
INIT ALOTOL	Day 55 * Corrected OX 5	1140 1250g en 0.	1/04		Lop Filec (C) 1	nages (2) Calculator
arrest Status	Admissions Respiratory N	strition Other	Ireatments Lost Re	alts		
espiratory Support	Adhiesian	Ranning	Discharge			
04#75,H0224	Liverpool Hospital		Admitted: 28/0	2/08 at 4 hours		
	Admission Age 1	Corr.GA 25 We	ght 10/0 75% HD	255 71% Lengt	h[35 55%	
uids / Feeds	Date & Time 20102/20	OD 15/30 IN Hospit	al Deepool Housital		MRN[29432
N 10% Fat 3g	Bed D1	M To	Nacu 💌	Reason(s) for	Prenaturity	×
andice	Consultant Ban Calan	der 💌 Ireusr	ce Hospital 💌	Admission	Respiratory Distress	
A/03 SBY 135 Billblarket	MATERNAL HISTORY					
ther	Ann is a 28 year old G2 P1 under the care of Kaisher	. (now) woman whose however delivered at t	blood group is 0 positiv iverpool Hospital under	e. She was booked the care of Dr Feter	to deliver at Campbellto Hannell, She had a hist	wn Hospital tory of essential
JO3 Mod PDA DOSTBLE NEC	hypertension. This pregna	ncy was complicated b	y hypertension of pregn	ancy, fetal growth n	estriction, Gilateral Penal class. clinically in error to	Pelvis dilatation
	choricannionitis. Ann was	treated with antenata	steroids, tocolytics, an	d anthypertensive d	rugs. Following the sport	taneous onset of
	Moor, she proceeded to	a naginal belowry. Art	boocs were given bero	re centery.		
reatments						
ectments ntavte, Folic Acid nalice.	PERIDUKTAL HISTORY Jan was born at 13:00 hou	rs with a birth weight	of 1070 grams (76th ce	ntile). Apgars were	3 at 1 minute and 7 at 5	minutes
verteents ntavte, Folic Acid ngline,	PERBUKTAL HESTORY Jan was born at 13:00 hou respectively treated with a	rs with a birth weight	of 1070 grams (76th ce m. The arterial cord pH	ntile). Apgars were was 7.24 and the b	3 at 1 minute and 7 at 5 see excess -5. Jan was f	minutes ten retrieved to
eetments etavta, Folc Acid ngline, ast Results	PERDUATAL HISTORY Ian was born at 13:00 hou respectively treated with a Added to Worksheet 0 This is freetext.	rs with a birth weight tabation and ventilate 1/03 Orders on W Prostext order	of 1070 grams (76th or on. The arterial cord pH orksheet: 01/03	ntile). Apgans were was 7.24 and the bi le click on	3 at 1 minute and 7 at 5 see excess -5. Jan was f	minutes nen rethieved to
votments navna, Polic Acid ngline, sat Results (03 Ale 136	PERINATIA, HESTORY Jan was born at 12:00 hou respectively traded with a Added to Warksheet 0 This is freetest	rs with a birth weight tubation and ventilate U/03 Orders on W Preetoxt orders	of 1070 grams (76th ce on, The arterial cord pH erisheet: 01/03 5 100.0 100.0	ntile). Apgars were was 7.24 and the ba le click on a claiete)	3 at 1 minute and 7 at 5 see encess -5. Jan was f	minutes en retrieved to
ectments ntavka, Folic Acid galine, 103 Alas 136 (03 Hai 136 (03 Hai 135	PERINATAL HISTORY Jan was born at Job0 hou respectively traded with a Added to Warksheet 0 This is freeboxt	rs with a birth weight tubation and ventilate 1/03 Orders en W Prestout order	of 3070 grams (76th ce on. The arterial cord pH erksheet: 01/05 bast 5	ntile), Apgars were was 7.24 and the bi le dick on a diates)	3 at 1 minute and 7 at 5 see access -5. Jan was f	minutes ten retrieved to
ectaents ntwite, Folic Acid ngline, 03 No 136 03 Ho 136 03 Ho 135 03 Ho 135 03 Ho 135 03 Ho 135 03 Ho 135 04 Hows B/P I	PERINATAL HISTORY Jan was born at 15:00 hou respectively tosted with a Added to Worksheet 0 This is freetext Hospital Episodes Exerptiol Hospital	rs with a birth weight tabation and vertilate 1/00 Onders on W Freetoxt order 1/29422	of 3070 grams (76th ce n. The arterial cord pH erksheet: 01/03 8 Jacksteet: 01/03 8 Admitted: D 28 Feb 2008 16:30	ntile), Apgans were was 7.24 and the ba le click on a claims) achanged	3 at 1 minute and 7 at 5 see encoses -5. Jan was fi Add Another	minutes ten nationed to Admission
vertwents navva, Folic Actd ngline, (Cl Na 126 (Cl Hb 125 (Cl Hb 125 (Cl Hb 125 (Cl Hb 125 (Cl Hb 125 (Cl Hb 127) (Cl Hb 127) (Cl Hb 127) (Cl Hb 127) (Cl Hb 127) (Cl Hb 127) (Cl Hb 127)	PREDUCTAL HEATORY Jan was born at 1300 hot respectively tradied with in Added to Worksheet O This is frontext Hospital Episodes Liverpool Hospital NETS Campbellown Hospital	rs with a birth weight industrian and worklows I/00 Orders on W Freetowit orden I/01 I/02	of 1070 grams (764 oc n. The articrisi cord pH erisisheet: 01/03 5 100,05 28 Feb 2008 15:00 28 Feb 2008 15:00 2 28 Feb 2008 15:00 2	ntile), Apgars were was 7.24 and the bi e click on s cliliona) scharged 3 Feb 2008 35:30 3 Feb 2008 35:30	a at 1 minute and 7 at 5 an access -5. Jan was 5 Add Another Delete Han . the	minutes en retrieved to Admission

Dr. Notes:

Each has its own **advantages** and **disadvantages**, they both are needed and depended on. The **structured data** is easier to manage and faster in analysis while **unstructured** gives the ability to express thoughts in natural language but it's harder to manage. (ex: retrieval of all patients medical records that are in ICD-10-CM system with the same disease such as breast cancer).

Natural Language Processing: studied the analysis of unstructured data. (takes the natural human language as an input, process it (low level processing: assigns the part of speech to words "verb, noun" while high level processing such as Siri).

What Makes Biomedical Informatics Difficult?

- Incomplete information.
- Uncertain information.
- Imprecise information.
- Vague information.
- Inconsistent information.
 - Ex: Birthdate: 8/29/66 and 9/17/66.

Clinical Data Warehouse

- A modern **way to convert medical information to knowledge** is to use a clinical data warehouse (CDW).
- **Function:** a database system that collects, integrates and stores clinical data from a variety of sources including electronic health records (**EHR**), radiology and other information systems.
- **Staging:** the process of cleaning the data from any incorrect or missing values.
- Meta-Data: "data about data" example: ICD-10-CM
- Difference between EHR & CDW:
 - EHR: are systems implemented in hospitals that focus on
 - individual data for one patient.
 - Address real time updates regarding patient data
 - **CDW:** support many EHR, focus on group query.
 - Information update depends on how often it is updated (monthly, weekly,...)

Uses:

- Monitor Quality by allowing users to query for specific quality measures
- Identify trends
- Comparative effective research: practice based research, answers very specific questions (is the treatment A better than treatment B? We can answer this question by looking for the prognosis of people who were treated with both in the system)

i2b2¹ platform (an example of CDW)

- A harvard project to integrate biology and the bedside used by other institutes
- Its an open source and modular and incorporate genomic and clinical information for research purposes
- Database consists of **facts** (diagnoses, lab results, etc.) queried by users and **dimensions** that describe the fact
- **Design:** star scheme
 - \circ In the middle there is a fact table (quantitative, information)
 - \circ On the side dimension tables (more information about these facts.

Use of Aggregated Clinical Data

- **Concept extraction:** the problem of identifying concepts within unstructured data, such as discharge summaries or pathology reports.
 - Usually, these concepts are mapped to a controlled vocabulary.
- **Classification:** the problem of categorizing data into two or more categories.
 - Supervised machine learning.

1- I2B2 stands for "informatics for integrating Biology & Bedside





Big Data

Definition: collecting large collections of data from various healthcare foundations followed by storing, managing, analyzing, visualizing, and delivering information for effective decision making.

The Six <u>V's</u> of big data:

Everything after this line is Dr. notes

- Volume: large quantity of data produced by organization.
 - Variety: a lot of sources in different formats, contribute in big data.
 - **Organized** (structured data: laboratory data, sensors data, data from databases).
 - **Semi-organized** (ex: website data).
 - **Unorganized** "*most common*" (ex: social media data, medical records that doesn't use medical terminology).
- **Velocity:** massive frequency during the current details created, supplied and managed.
- **Veracity:** refers to accuracy of this data if it's accurate or valid (ex: in big data there is usually low veracity compared to other databases).
- **Variability:** data is fluctuating, changeable data, guidelines are changeable as well and might me updated so data might change during life.
- Value: methods of extracting valuable information from these data (big data analytics).



Analysis: process of transforming raw data into information.

Stage 1 Data acquisition: there are different sources (Electronic record, social media, smart phone, apps, websites... etc) to collect the data.

Stage 2 Storage: store in cloud chambers.

Stage 3 Management: organizing, cleaning and data mining to make sure that the data is valid, no missing data or values.

Stage 4 Analytics:

- Descriptive analysis: collects historical data, what happened in the health care?
- Diagnostic Analysis: route cause of the problem, why did it happen?
- Predictive Analysis: use both real time data and historical data, it has probabilities, what will happen? what is the future trends?
- Prescriptive Analysis: number of different outcomes before the decision, what should we do? And what is the best outcome? And how we can make it happen?

Stage 5 Data visualization and reporting: how I represent these results? (ex: in graphs).



QUIZ

1. which of the followings is picture data?

- A. Temperature.
- B. Blood test.
- C. X-ray.
- D. EMG.

2. Which of the following is a major limitation in using abbreviation in narrative data?

- A. Produce ambiguity.
- B. Hard to read.
- C. Needs skills.
- D. Against patient's privacy.

3. "Pick from a list" allows wrong selection, compliance concerns, cloning and limitations are issues for?

- A. Coded data.
- B. Voice recognition.
- C.Language processing.
- D.Manual data entry.

4. The veracity of big data is due to which of the following?

- A. Uncertainty of data
- B. Velocity of data
- C. Multiple sources of data
- D. Different forms of data

5. What of these describe narrative?

- A. Business Documents.
- B. Choose from a list.
- C.Report as text.

6. "140/90 mmhg" is considered which one of data elements?

- A. attribute
- B. value of the attribute
- C. time of the observation
- D. method obtained

7. "Collection of records" is definition of what?

- A. Field
- B. File
- C. Database
- D. Record

8. The type of data that makes the data (understandable) for the computer:

1.C 2.A 3.A 4.A 7.B 6.B 8.A 8.A

- A. Coded data
- B. Free form data
- C. Unstructured data
- D. Dictated data





team leader KHALID ALKHANI

DONE BY OUR AMAZING MEMBER: Omar Aldosari

NOTE TAKEN BY OUR SHARP MEMBER:

