FINDING DATA

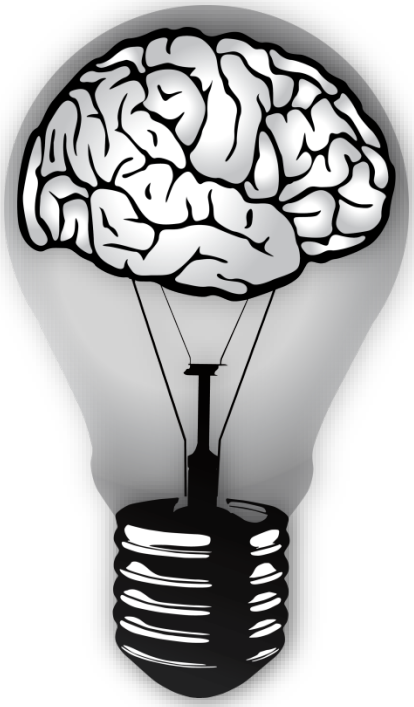
Literature | Data | The Meaning of Life

Data After Dark

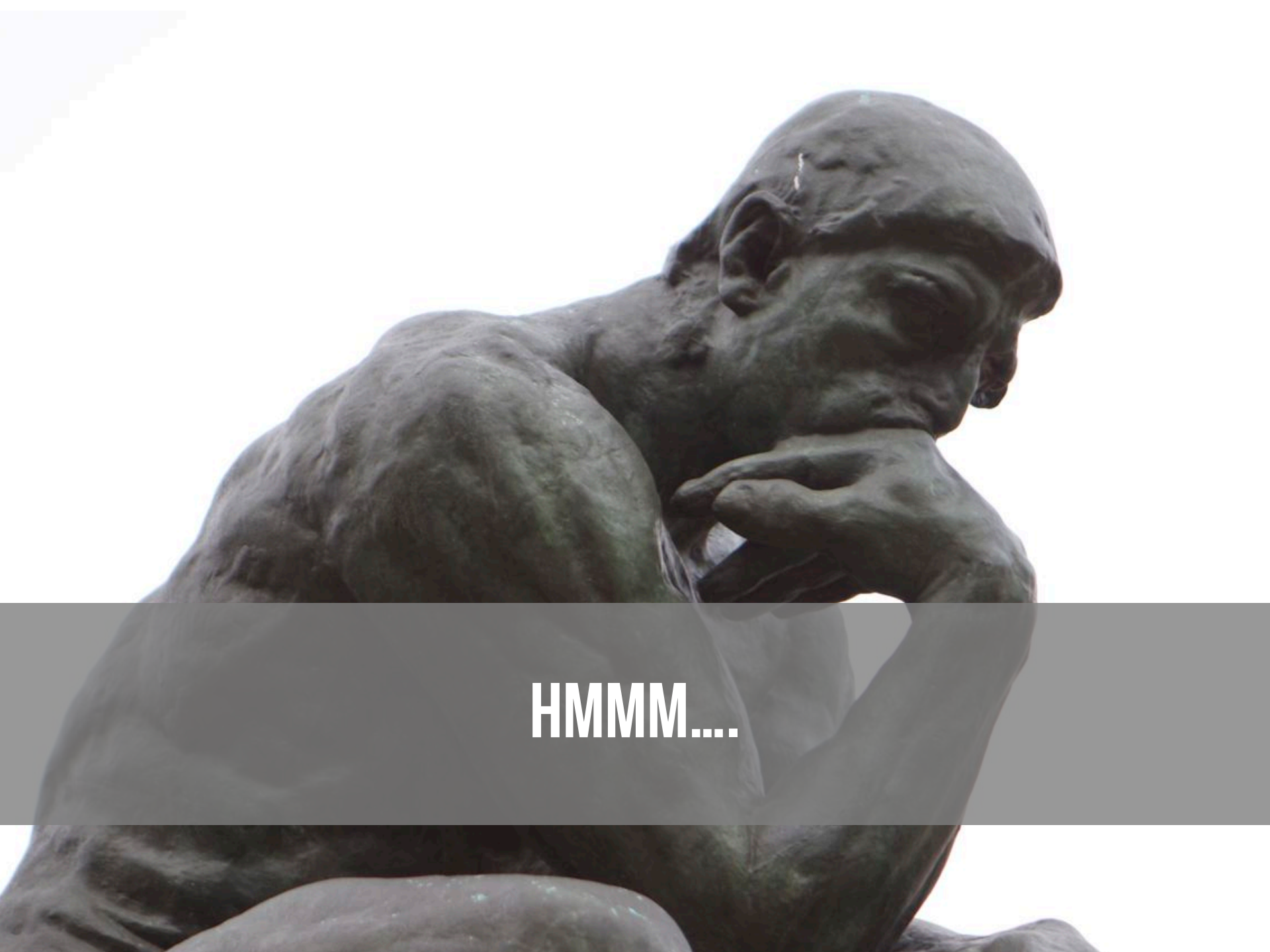
OHSU BD2K Data Science Workshop

Jackie Wirz, PhD

13th January 2016



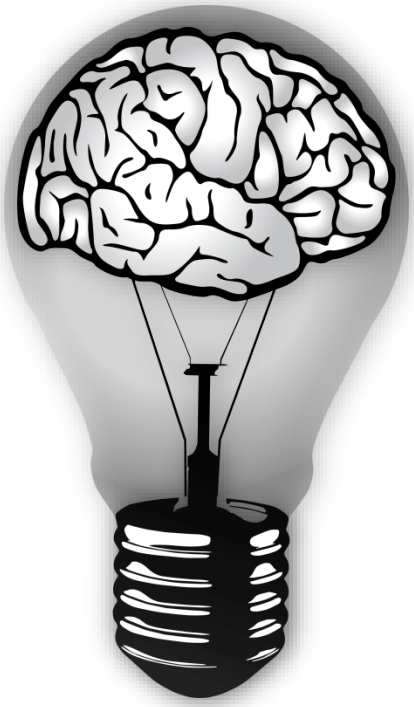
WHERE DOES GOOD DATA COME FROM?


HMMM....

YOUR EXPERIMENTS

LITERATURE

DATASETS OUT IN THE ETHER

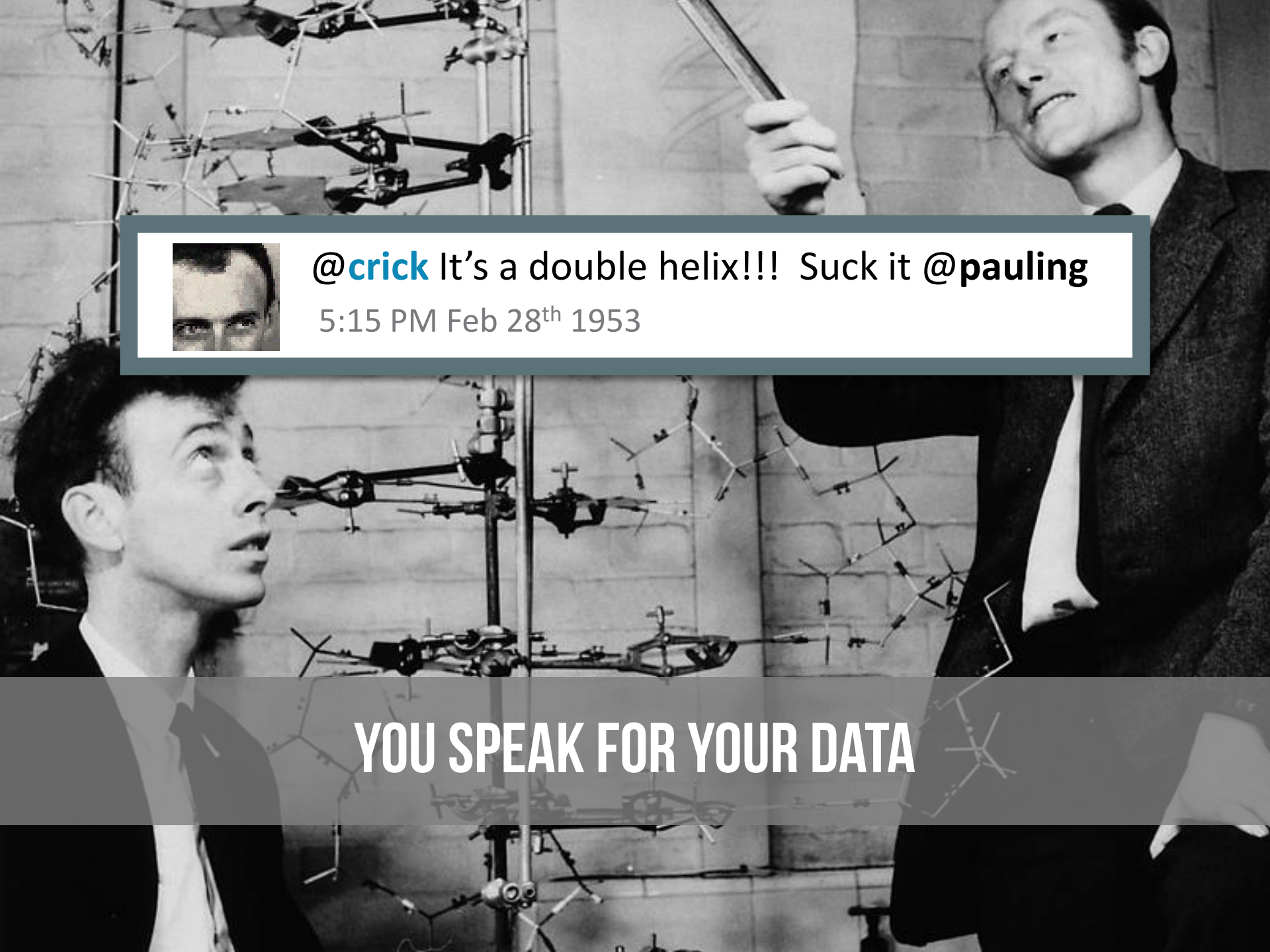


HOW DO I FIND GOOD DATA?

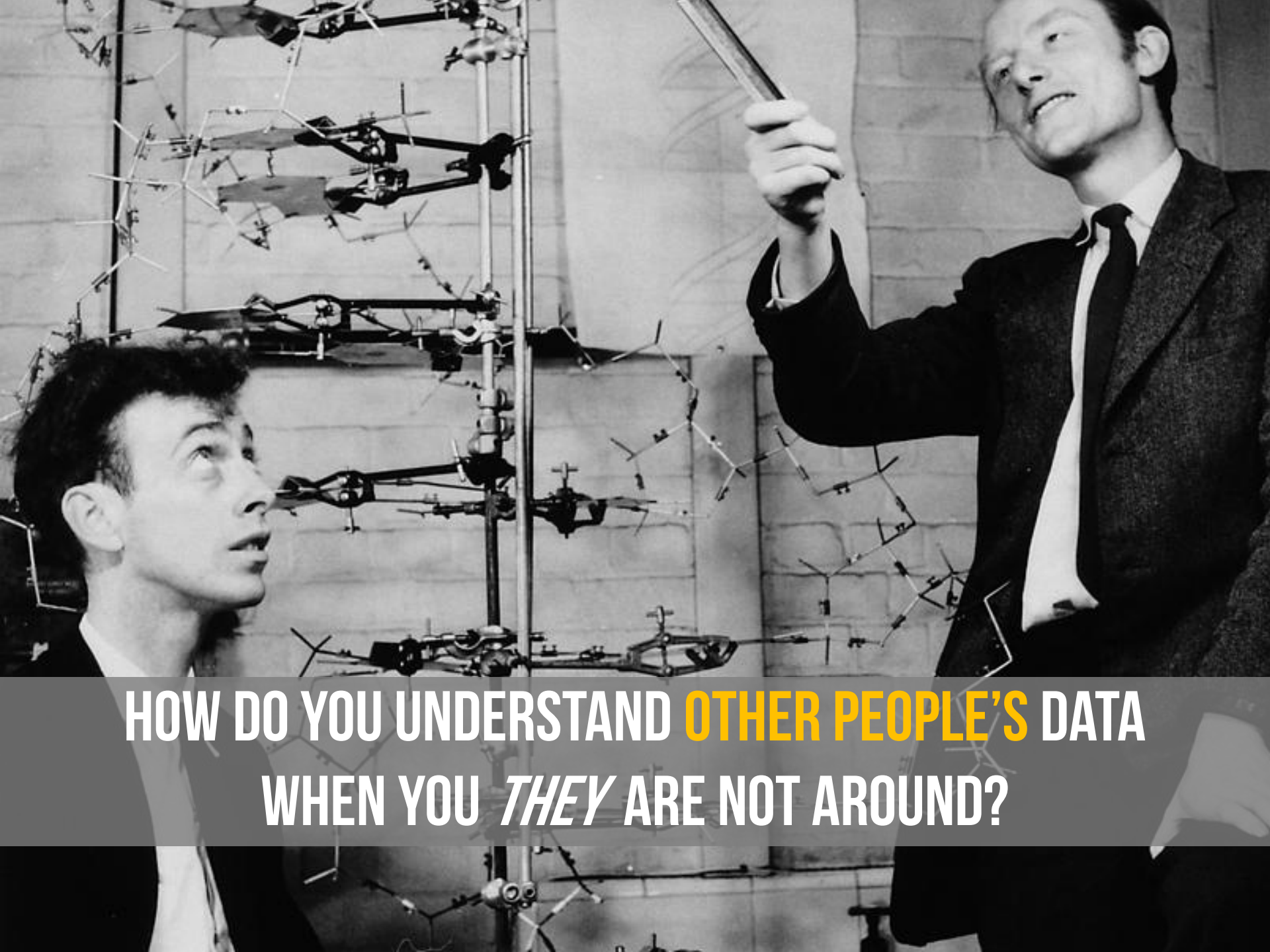
IF YOU PRODUCE IT YOURSELF...

DROP MIC AND WALK AWAY

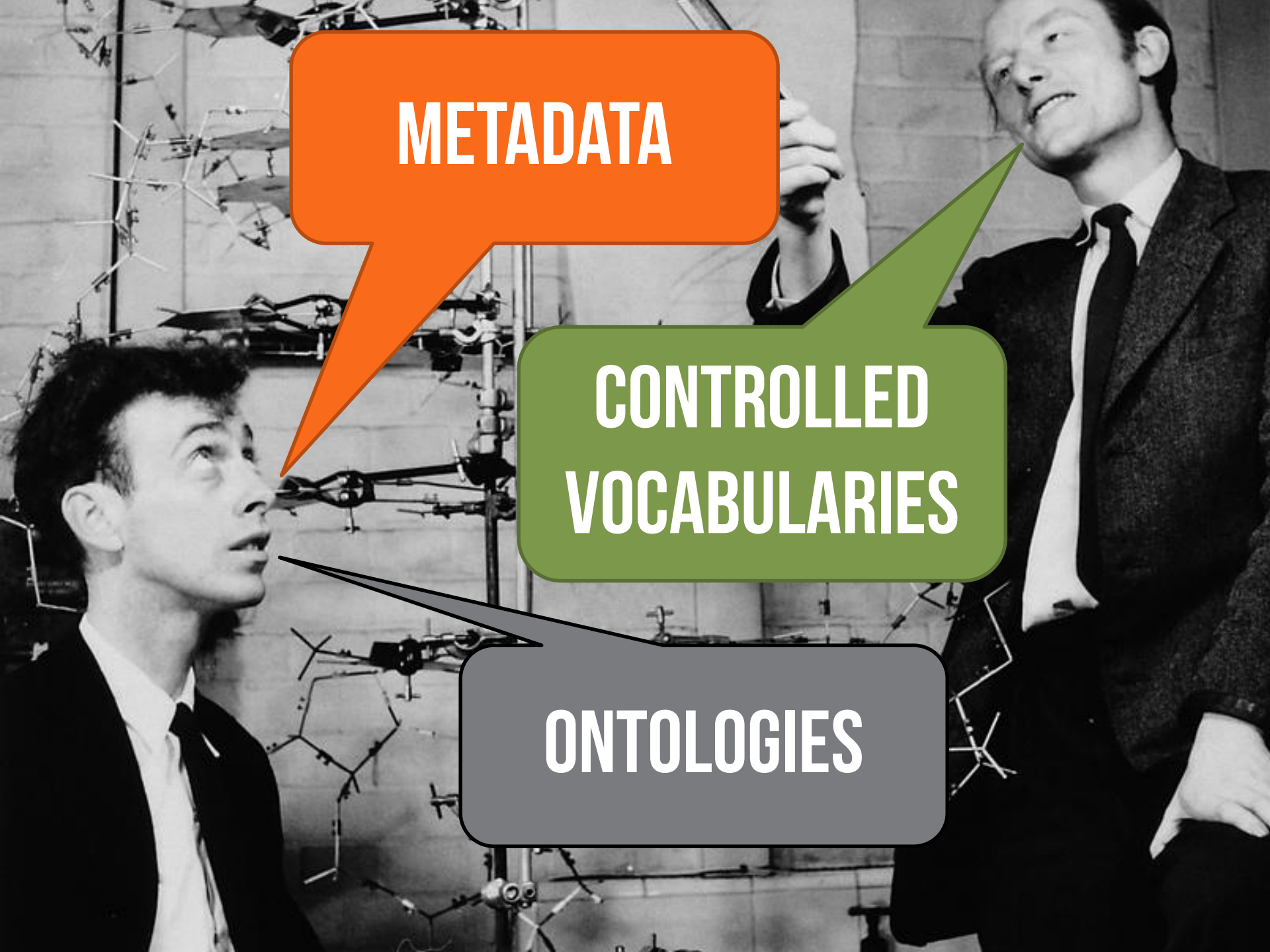
JUST KIDDING...


@crick It's a double helix!!! Suck it @pauling
5:15 PM Feb 28th 1953
YOU SPEAK FOR YOUR DATA

@watson @crick You guys both suck…
5:17 PM Feb 28th 1953
HOW DO YOU SPEAK FOR YOUR DATA
WHEN YOU ARE NOT AROUND?



HOW DO YOU UNDERSTAND OTHER PEOPLE'S DATA WHEN YOU *THEY* ARE NOT AROUND?


METADATA
CONTROLLED VOCABULARIES
ONTOLOGIES

METADATA

CONTROLLED VOCAB

ONTOLOGIES

WHAT IT IS

WHAT IT TAKES TO DO IT

RELEVANT VARIABLES

DEFINITIONS

GROUPING

CLASSIFICATION

CONNECTION

WHAT IS METADATA, REALLY?

METADATA
IS A LOVE NOTE
TO THE FUTURE
Behappy.me

YOUR METADATA SHOULD MAKE YOUR DATA UNDERSTANDABLE TO OTHERS WITHOUT YOUR INVOLVEMENT

- ANNE GILLILAND

Person 1	Control	blue
Person 2	Control	brown
Person 3	Control	amber
Person 4	Control	brown
Person 5	Treatment	black
Person 6	Treatment	grey
Person 7	Treatment	brown
Person 8	Treatment	brown

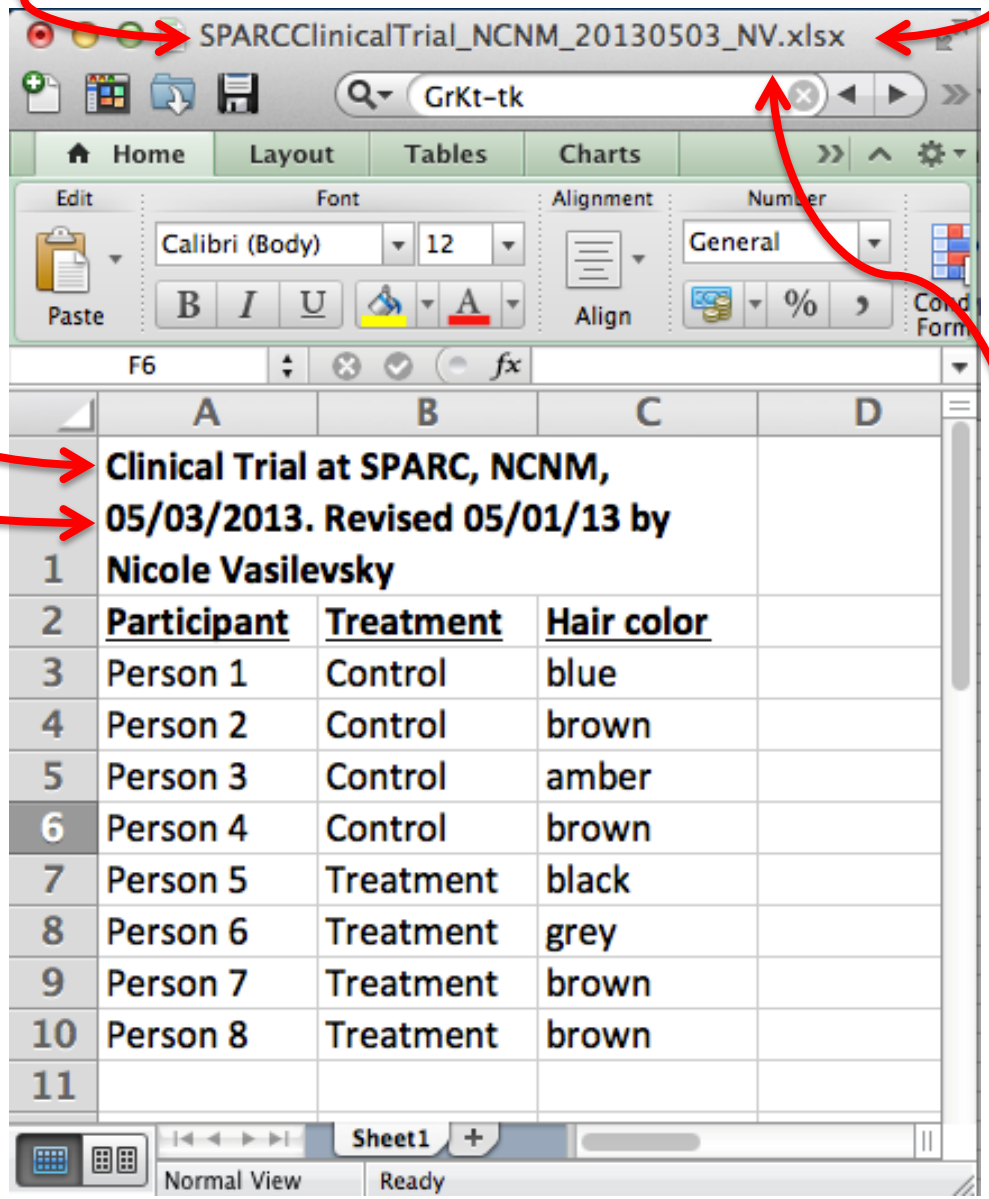
FILE NAME

FILE TYPE

TITLE

DATE CREATED

WHO CREATED THE DATA



	A	B	C	D
1	Clinical Trial at SPARC, NCNM, 05/03/2013. Revised 05/01/13 by Nicole Vasilevsky			
2	Participant	Treatment	Hair color	
3	Person 1	Control	blue	
4	Person 2	Control	brown	
5	Person 3	Control	amber	
6	Person 4	Control	brown	
7	Person 5	Treatment	black	
8	Person 6	Treatment	grey	
9	Person 7	Treatment	brown	
10	Person 8	Treatment	brown	
11				

Data Field	Name	Definition	Data Type	Format	Field Size	Values	Source System	Date First Entered	Why Item Is Included
Inpatient Daily Census	IP_DAY_CENSUS	The number of inpatients present at census-taking time each day, plus any inpatients who were both admitted and discharged after the previous day's census-taking time	numeric	x to xx	3	Any whole number from 0 to 999	Patient Census	2/23/2008	Provides analysis of budget variances, aids future budgetary decisions, and allows quicker response to negative trends
Medical Record Number	MR_NUM	The unique number assigned to a patient's medical record The medical record is filed under this number	alphanumeric	xxxxxx: requires leading zeros	6	000001 to 999999	Patient Census; Practice Mgt		Provides analysis of services, resource utilization, and patient outcomes at the physician level
Patient Age	PT_AGE	Age of patient calculated by using most recent birthday attained before or on same day as discharge	numeric or alphanumeric	Age in days = xD to xxD OR Age in months = xM to xxM OR Age in years = x to xxx	3	Age must be > 0, and < OR = 124 years; children less than 1 year must be > 0 M AND < 1 year	Patient Census; Practice Mgt	2/23/2008	Patient age impacts the services utilized and payer sources

BEST BOOK OF 2016: THE DATA DICTIONARY

IF YOU LIKE IT
PUT SOME METADATA ON IT
memegenerator.net

WHAT IS CONTROLLED VOCAB, REALLY?

Like New Sofa with Chaise
$375 (Milwaukie)
furniture - by owner

~~~~****CHAISE LOUNGE AND ACCEN
$125 (SE Portland)
furniture - by owner

Very nice modern Chaise
$200 (Clackamas)
furniture - by owner

victorian fainting couch
$600 (Oregon City)
antiques - by owner

Custom Mid Century Upholstered Coffee Table
$110 (I deliver)
furniture - by owner

Victorian Fainting Couch
$225 (Beaverton)
furniture - by owner
~~~~

NCBI Resources How To Sign in to NCBI

MeSH

MeSH Search

Limits Advanced Help

MeSH

MeSH (Medical Subject Headings) is the NLM controlled vocabulary thesaurus used for indexing articles for PubMed.

Using MeSH

- Help
- Tutorials

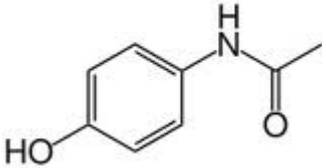
More Resources

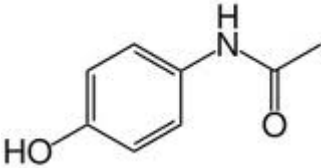
- E-Utilities
- NLM MeSH Homepage

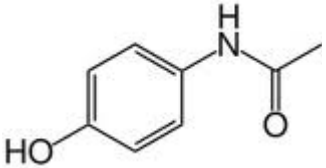
NEW VALUE SIZE
Extra Strength
TYLENOL
Gelcaps

Acetaminophen
NON-ASPIRIN

PARACETAMOL
TABLETS


H
N
O
HO


H
N
O
HO


H
N
O
HO

- Acetominophen
- APAP
- Hydroxyacetanilide
- Paracetamol
- N-Acetyl-p-aminophenol
- p-Acetamidophenol
- p-Hydroxyacetanilide
- Acetamidophenol
- N-(4-Hydroxyphenyl)acetanilide
- Acephen
- Acetaco
- Tylenol
- Anacin-3
- Anacin 3
- Anacin3
- Datril
- Panadol
- Acamol
- Algotropyl

acetominophen

PLoS Biol. 2012;10(5):e1001339. doi: 10.1371/journal.pbio.1001339. Epub 2012 May 29.

Dealing with data: a case study on information and data management literacy.

Haendel MA, Vasilevsky NA, Wirz JA.

Library, Oregon Health & Science University, Portland, Oregon, USA. haendel@ohsu.edu

PMID: 22666180 [PubMed - indexed for MEDLINE] PMCID: PMC3362643 **Free PMC Article**

Publication Types, MeSH Terms, Grant Support

Publication Types

Research Support, American Recovery and Reinvestment Act

Research Support, N.I.H., Extramural

MeSH Terms

Cooperative Behavior

Databases, Factual*

Humans

Information Dissemination

Information Literacy*

Information Management/methods

Information Storage and Retrieval/methods*

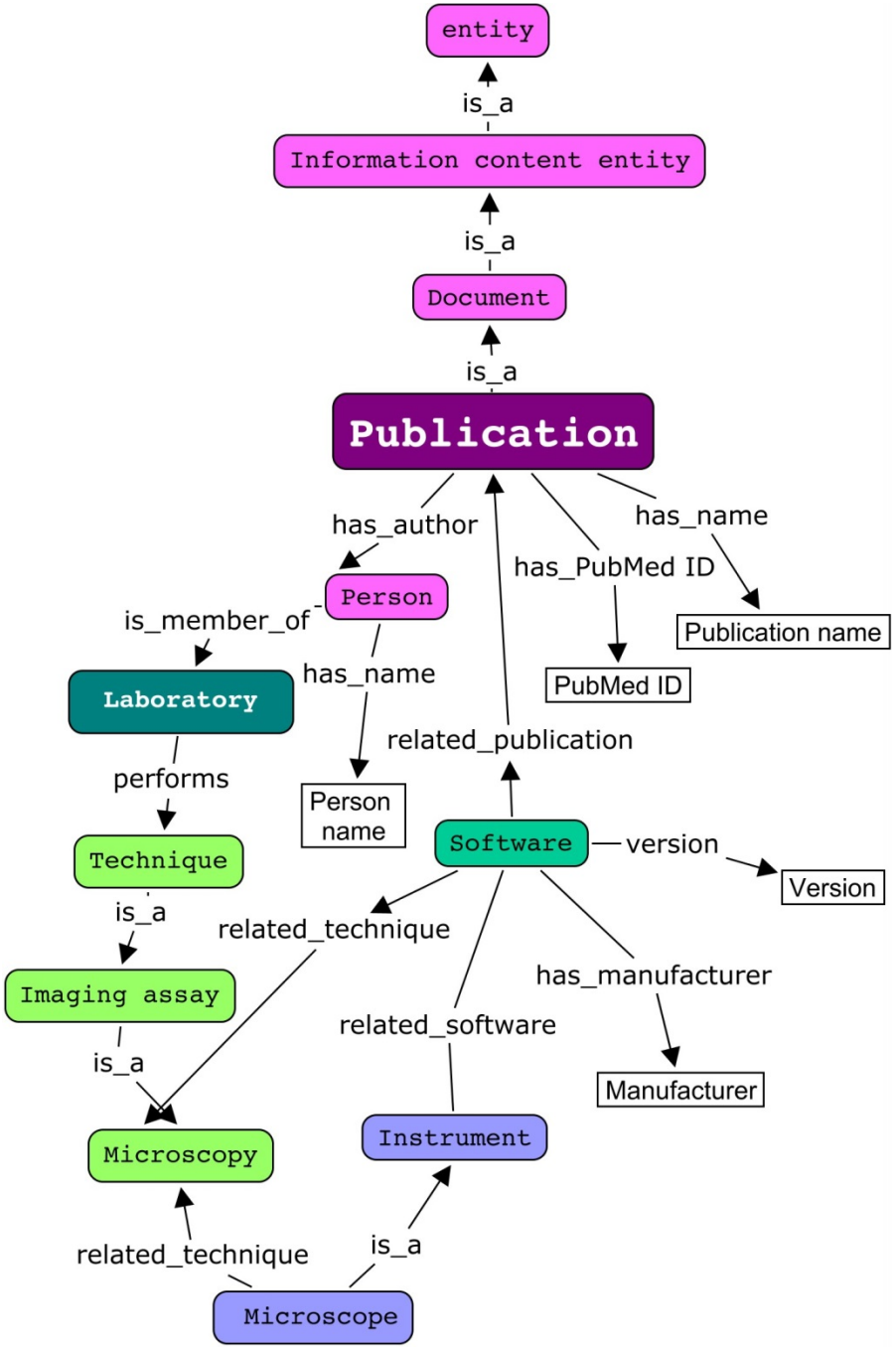
Interdisciplinary Communication

Professional Role

PUBMED INDEXES ARTICLES WITH MESH TERMS

WHAT IS AN ONTOLOGY, REALLY?

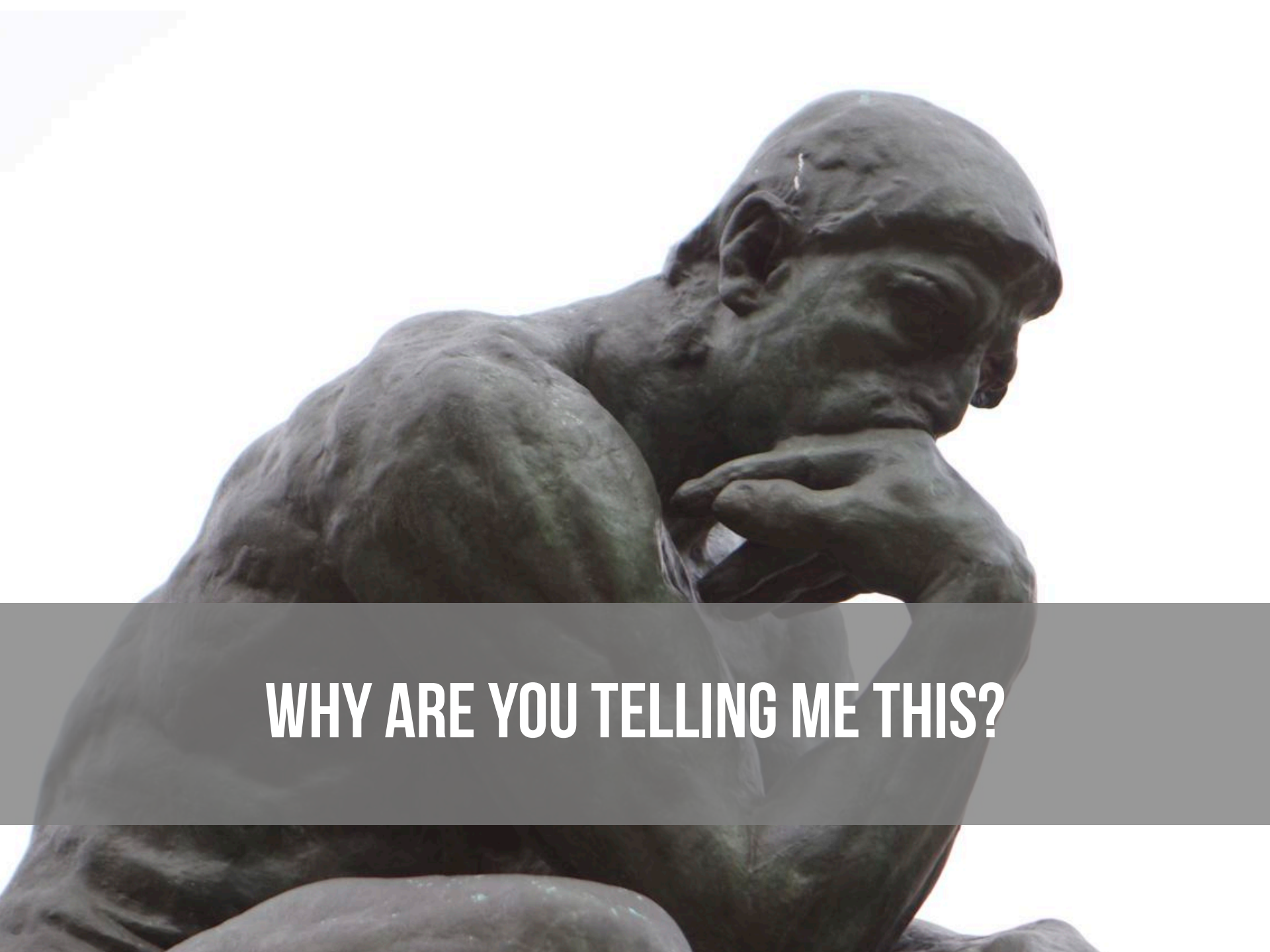
Mac OS X

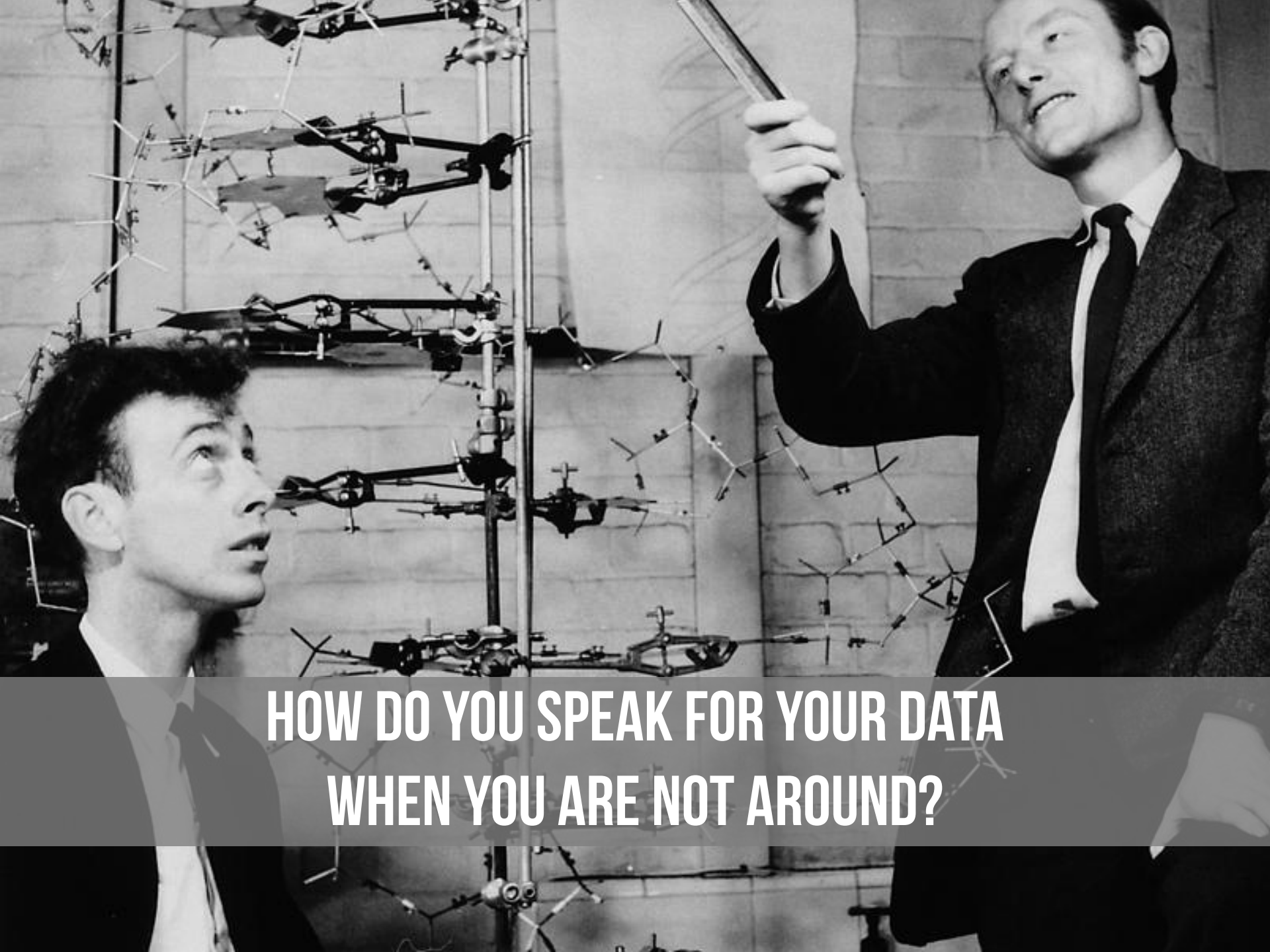

entity
is_a
Information content entity
is_a
Document
is_a
Publication
has_author
has_name
has_PubMed ID
Person
Publication name
is_member_of
PubMed ID
Laboratory
has_name
related_publication
performs
Person name
Software
version
Technique
Version
is_a
related_technique
Imaging assay
has_manufacturer
related_software
is_a
Manufacturer
Microscopy
Instrument
related_technique
is_a
Microscope

NETFLIX

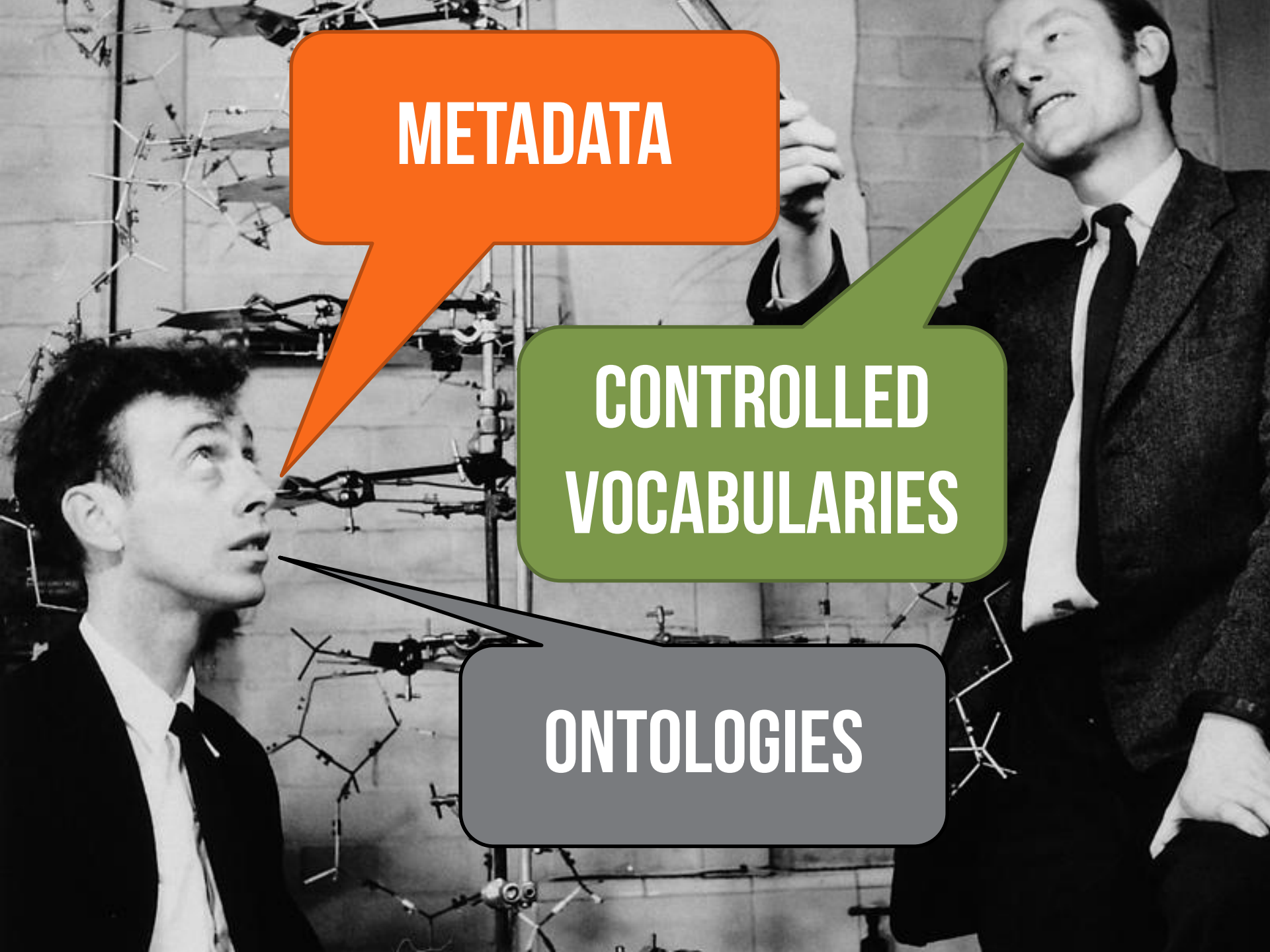
amazon

monarch
INITIATIVE

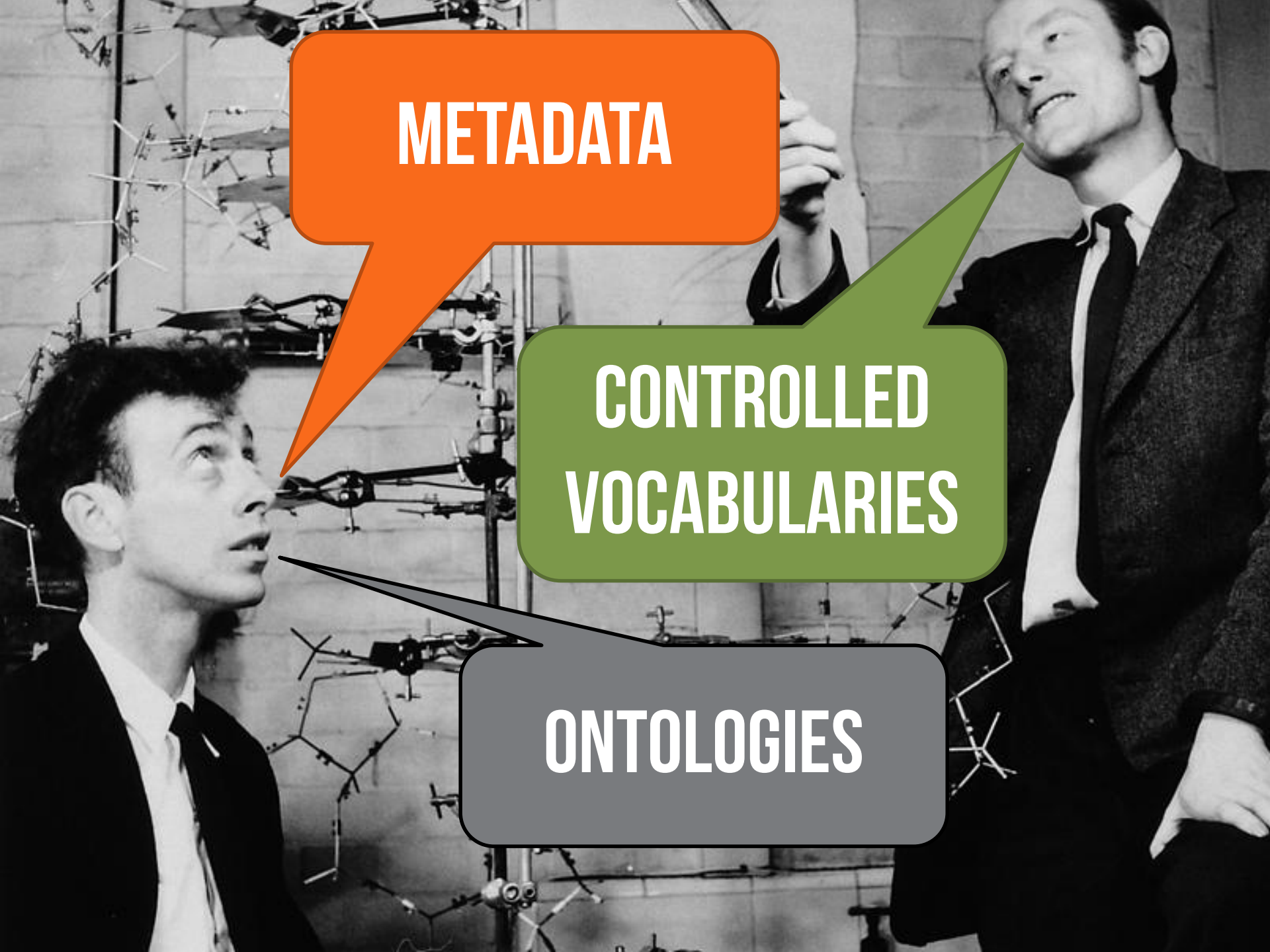

WHY ARE YOU TELLING ME THIS?

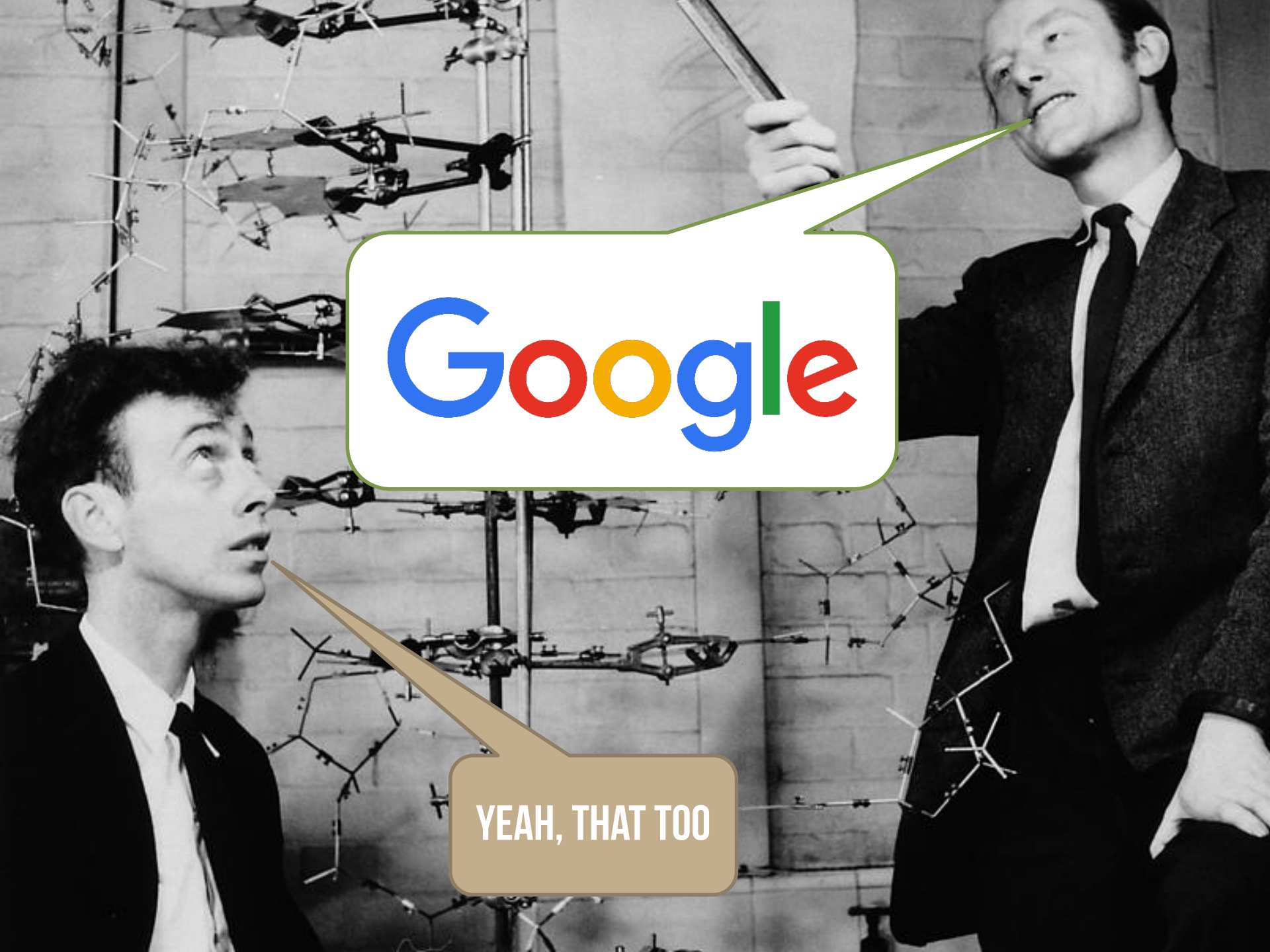


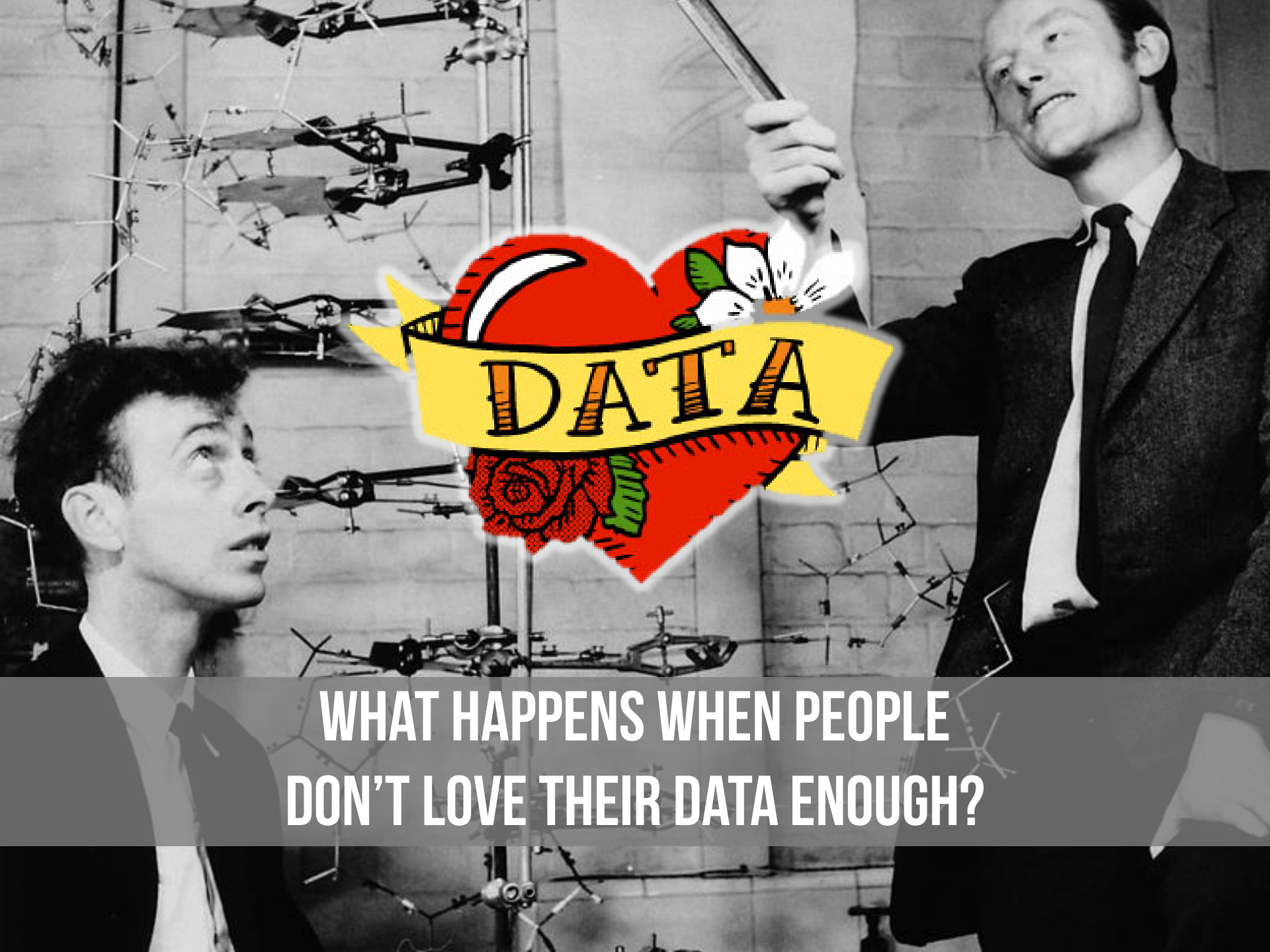
HOW DO YOU SPEAK FOR YOUR DATA WHEN YOU ARE NOT AROUND?


METADATA
CONTROLLED VOCABULARIES
ONTOLOGIES

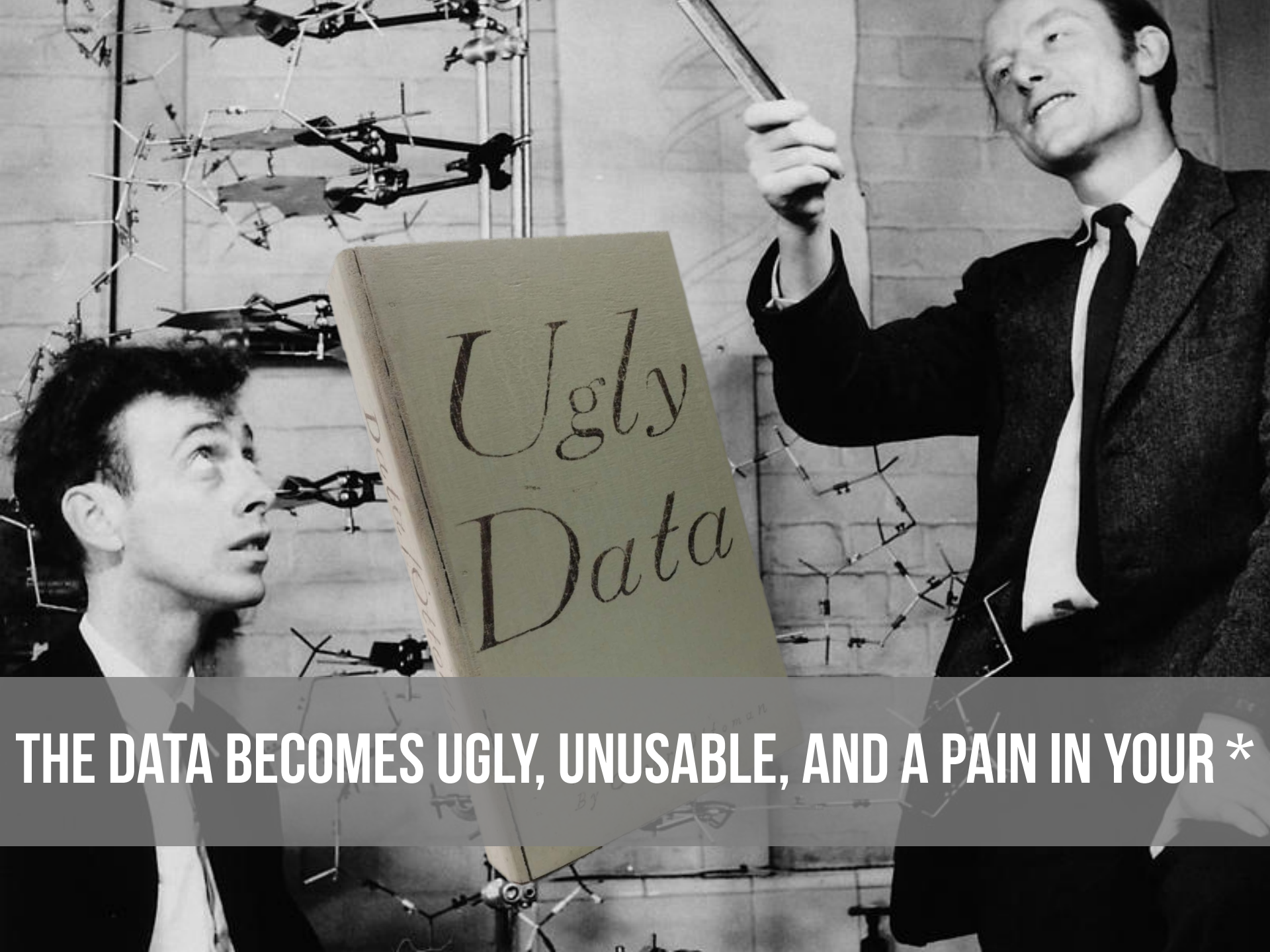
HOW DO YOU FIND DATA?


METADATA
CONTROLLED VOCABULARIES
ONTOLOGIES


Google
YEAH, THAT TOO



WHAT HAPPENS WHEN PEOPLE DON’T LOVE THEIR DATA ENOUGH?



THE DATA BECOMES UGLY, UNUSABLE, AND A PAIN IN YOUR *

LET'S TAKE A LOOK AT A SMALL EXAMPLE

MEET THE URBAN LAB

THE URBAN LAB ANTIBODIES

Primary Antibodies:												
Antibody Target	**Ok?**	**Rec'd**	**Manufacturer**	**Stock Number**	**Species**	**Diluted 2:1?**	**Storage**	**State**	**Final Concentation**	**Incubation Time**	**Unmasking?**	**Notes**
BrdU (for IdU)	✓	3/23/12	BD Biosystems	347580	mouse	no	4C	Liquid	1:50	overnight	Acid Wash	Use secondary at 1:300 for best detection.
BrdU or CldU	✓	0/0/2010	Serotec	MCA2060	rat	no	4C	Liquid	1:1000	1 hr	Acid Wash	
Calbindin	✓	1/11/08	Chemicon	AB1778	rabbit	no	-20C	Frozen Aliquots	1:1000	1 hr		
Calretinin (N-18)	✓	8/15/06	Santa Cruz	SC-11644	goat	no	4C	Liquid	1:1500	1 hr		
CamKII		6/2/05	Affinity Bioreagents	MA1-048	mouse	no	-20C	Frozen Aliquots	nd	nd		
Caspase-3	✓	10/19/04	Cell Signaling Technology	9661	rabbit	no	-20C	Liquid	1:250	overnight		
cFos		3/26/03	Genosys	OA-11-824A	sheep	yes	-20C	Liquid	nd	nd		
c-Fos (Ab-3)		10/29/09	Calbiochem	OP53	mouse	no	-20C	Liquid	1:1000	1 hr		
cFos (Ab-5)	✓	2/24/11	Calbiochem	PC38	rabbit	no	-20C	Frozen Aliquots	1:1000/1:5000	1 hr/overnight		
Connexin 43		unk	Invitrogen (Zymed)	71-0700	rabbit	no	-20C	Frozen Aliquots	1:500?	1 hr?		
doc2b		4/1/10	NeuroMab/Antibodies, Inc.	73-214	mouse	no	4C	Liquid	nd	nd		Ascites fluid
doc2b	✓	5/1/10	Novus Biologicals	NBP1-03473	rabbit	no	-20C	Frozen Aliquots	1:25	1 hr		Use DNAse1 for BrdU labeling(?)
Doublecortin (C-18)	✓	8/15/06	Santa Cruz	SC-8066	goat	no	4C	Liquid	1:250	1 hr		
EGR-1 (S-25)		10/29/09	Santa Cruz	SC-101033	mouse	no	4C	Liquid	1:2000	1 hr		
EGR-1 (C-19)	✓	2/24/11	Santa Cruz	SC-189	rabbit	no	4C	Liquid	1:500	1 hr		
Frequenin (NCS-1)		1/11/05	Chemicon	AB5906	Chicken	no	-20C	Frozen Aliquots	nd	nd		
GABA	✓	6/22/06	Chemicon	AB175	Guinea Pig	no	-20C	Frozen Aliquots	1:2000	1 hr		
GABA	✓	4/23/09	Sigma (Preferred)	A0310	mouse	no	-20C	Frozen Aliquots	1:2000	1 hr		
GAP-43	✓	4/1/08	Sigma	G9264	mouse	no	-20C	Frozen Aliquots	1:16000	overnight		
GFAP	✓	8/7/12	Abcam	ab7260	rabbit	no	4C	Liquid	1:1000	1 hr		
GFAP	✓	9/1/06	Promega	G5601	rabbit	no	4C	Liquid	1:1000	1 hr		
GFP	✓	3/5/13	Abcam	ab6673	goat	no	4C	Liquid	1:10000	1 hr		
GFP		4/08	Chemicon (for EM)	AB3080	rabbit	no	-20C	Frozen Aliquots	1:600	overnight		
GFP	✓	5/3/12	Molecular Probes (Preferred)	A11122	rabbit	no	4C	Liquid	1:1000	1 hr		
Growth Cone		5/6/08	Serotec	MCA-1712	mouse	no	4C	Liquid	1:500 or 1:1000	1 hr		
HCN1		3/9/11	NeuroMab/Antibodies, Inc.	75-110	mouse	no	-20C	Frozen Aliquots	1:100	1 hr	Pepsin	
HCN2	✓	9/12/12	Alomone Labs	APC-030	rabbit	no	4C	Liquid	1:500	1 hr	Pepsin?	Titrated in p9, older animals probably need pepsin unmaski
HCN2		12/20/10	NeuroMab/Antibodies, Inc.	75-111	mouse	no	-20C	Frozen Aliquots	nd	nd		GML18, p9--no labeling with pepsin digestion.
HCN3		12/20/10	NeuroMab/Antibodies, Inc.	75-175	mouse	no	-20C	Frozen Aliquots	nd	nd		
HCN4		12/20/10	NeuroMab/Antibodies, Inc.	75-150	mouse	no	-20C	Frozen Aliquots	1:50	1 hr	Pepsin	
HCN4		8/9/11	Abcam	ab69054	rabbit	no	-20C	Frozen Aliquots	nd	nd	Pepsin	
HSV-1	✓	2/26/07	Dako	B0114	rabbit	no	4C	Liquid	1:10000	2-3 days		
Ki67	✓	11/12/08	Abcam	ab15580	rabbit	no	4C	Liquid	1:1000	1 hr		
Kv1.1	✓	10/28/09	NeuroMab/Antibodies, Inc.	75-007	mouse	no	-20C	Frozen Aliquots	1:1000	1 hr	Pepsin	
Kv1.2		3/9/11	NeuroMab/Antibodies, Inc.	75-008	mouse	no	-20C	Frozen Aliquots	1:1000	1 hr	Pepsin	
Kv1.2		8/23/11	Abcam	ab55987	rabbit	no	-20C	Frozen Aliquots	nd	nd	Pepsin	GML19, pp1-5--no labeling with FL or DAB
Kv1.2	✓	10/16/12	Santa Cruz	SC-11188	goat	no	4C	Liquid	1:50	overnight	Pepsin?	Titrated in p9, older animals probably need pepsin unmasking. La
Kv1.3		3/9/11	NeuroMab/Antibodies, Inc.	75-009	mouse	no	-20C	Frozen Aliquots	1:1000	1 hr	Pepsin	
Kv3.1b		7/1/11	NeuroMab/Antibodies, Inc.	75-041	mouse	no	-20C	Frozen Aliquots	nd	nd	nd	
Kv3.2	✓	9/9/11	Abcam	ab101787	rabbit	no	-20C	Frozen Aliquots	1:1000	overnight	Pepsin	Better labeling than Santa Cruz.
Kv3.2 (L-14)	✓	6/21/11	Santa Cruz	SC-54394	goat	no	4C	Liquid	1:100	overnight	Pepsin	
Kv3.3		6/23/11	Abnova	H00003748-M01	mouse	no	-20C	Frozen Aliquots	nd	nd	nd	
Kv4.2		4/8/11	NeuroMab/Antibodies, Inc.	75-016	mouse	no	-20C	Frozen Aliquots	1:250	1 hr	Pepsin	
Kv4.2 (C-20)	✓	6/16/10	Santa Cruz	sc-11683	goat	no	4C	Liquid	1:250/1:2500	1 hr/overnight	Pepsin	
Map2	✓	5/13/08	Abcam	ab11267	mouse	no	-20C	Frozen Aliquots	1:4000	1 hr		Use 2% TX-100 during permeabilization for good Map2 penet
mGluR1a	✓	10/07	Abcam	ab6439	rabbit	no	-20C	Liquid	1:1000	1 hr		

A+ ORGANIZATION

“BAX ANTIBODY (SANTA CRUZ)”

Bax Antibodies

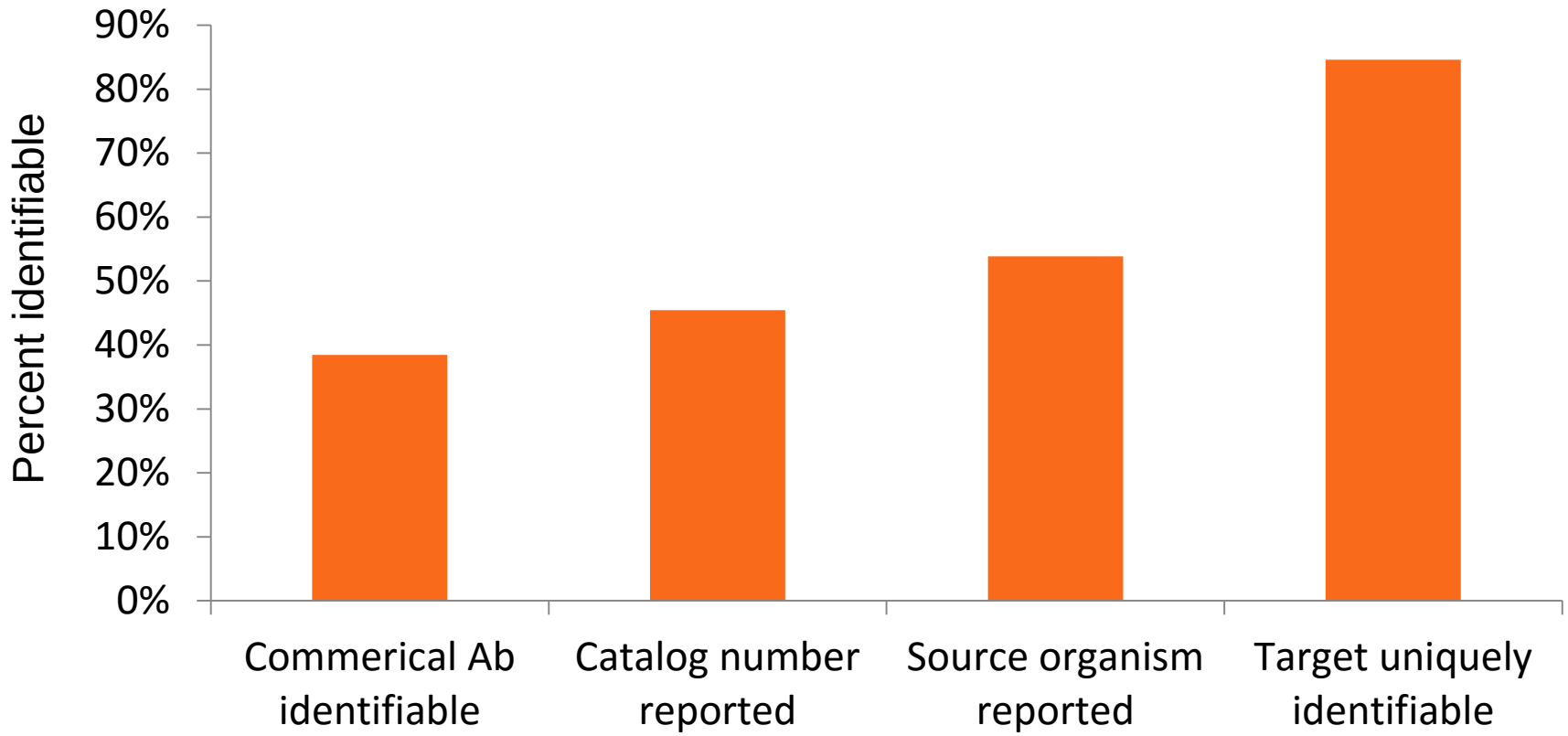
Santa Cruz Biotechnology, Inc. offers a broad range of Bax antibodies. Select Bax antibodies from several monoclonal and/or polyclonal Bax antibodies listed below. View detailed Bax antibody specifications by linking to the specific product blocks. Select appropriate Bax antibodies for your research by isotype, epitope, applications and species reactivity. Bax gene silencer products in siRNA, shRNA Plasmid, shRNA Lentiviral Particle, CRISPR/Cas9 Knockout and CRISPR Double Nickase Plasmid formats are also available as well as CRISPR/dCas9 Activation Plasmids and CRISPR Lenti Activation Systems for gene activation.

Introducing HOVERcruz™, a unique system for rapid identification of Bax Antibodies. Hover over the icon next to the product names in the table to see representative data for that product. **Live chat is now available! (click here)** Our live chat system is available to answer your product and ordering questions.

See chemical Bax Activators for functional analysis of cellular responses to Bax.

Also see additional Bax Antibodies including, Bax beta

PRODUCT NAME	CAT. #	ISOTYPE	EPITOPE	APPLICATIONS	SPECIES	CITATIONS	RATING
Bax (2D2)	sc-20067	mouse IgG_1	3-16 (h)	WB, IP, IF, IHC(P), FCM	human	69	★★★★★
Bax (Δ 21)	sc-6236	rabbit IgG	1-171 (m)	WB, IP, IF, IHC(P), ELISA	m, r, h, c, b, p	62	★★★★★
Bax (6A7)	sc-23959	mouse IgG_1	N-terminal (h)	WB, IP	m, r, h	58	★★★★★
Bax (B-9)	sc-7480	mouse IgG_{2b}	1-171 (m)	WB, IP, IF, IHC(P), FCM	m, r, h	222	★★★★★
Bax (N-20)	sc-493	rabbit IgG	N-terminal (h)	WB, IP, IF, ELISA	m, r, h, e, c, b, p	625	★★★★★
Bax (P-19)	sc-526	rabbit IgG	N-terminal (m)	WB, IP, IF, IHC(P), ELISA	m, r, h, c, b, p	280	★★★★★
Bax (YTH6A7)	sc-80658	mouse IgG_1	clone 5B7	WB, IP	m, r, h	2	★★★★☆
Bax (6D149)	sc-70405	mouse IgG_1	3-16 (h)	WB, IP, FCM	human	2	★★☆☆☆
Bax (4H32)	sc-70407	mouse IgG_1	N-terminal (h)	WB, IP	m, r, h	9	★★★★★
Bax (6D150)	sc-70408	mouse IgG_1	N-terminal (h)	WB, IP	m, r, h	9	★★★★★
Bax (SPM336)	sc-65532	mouse IgG_1	3-16 (h)	WB, IP, IF, IHC(P)	human	5	★★★★★
Baxβ (K-17)	sc-20287	goat IgG	internal (h)	WB, IF, ELISA	human		★★☆☆☆
Baxβ (S-18)	sc-20288	goat IgG	C-terminal (h)	WB, IF, ELISA	human		★★☆☆☆

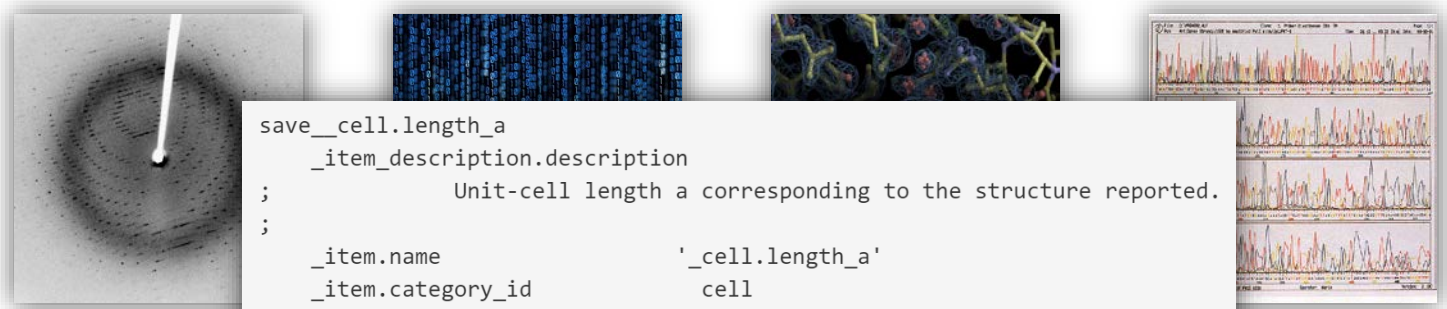


OF 14 ANTIBODIES PUBLISHED IN 45 ARTICLES, ONLY 38% WERE IDENTIFIABLE

Data standards are the rules by which data are described and recorded. In order to share, exchange, and understand data, we must standardize the **format** as well as the **meaning**.

www.usgs.gov/datamanagement/plan/datastandards.php

STANDARDS IN ACTION



```
save__cell.length_a
    _item_description.description
;             Unit-cell length a corresponding to the structure reported.
;
    _item.name                  '_cell.length_a'
    _item.category_id             cell
    _item.mandatory_code          no
    _item_aliases.alias_name    '_cell_length_a'
    _item_aliases.dictionary      cif_core.dic
    _item_aliases.version         2.0.1
     loop_
    _item_dependent.dependent_name
                                '_cell.length_b'
                                '_cell.length_c'
     loop_
    _item_range.maximum
    _item_range.minimum             .    0.0
                                  0.0   0.0
    _item_related.related_name  '_cell.length_a_esd'
    _item_related.function_code   associated_esd
    _item_sub_category.id         cell_length
    _item_type.code               float
    _item_type_conditions.code    esd
    _item_units.code              angstroms
     save_
```

Structural changes linked to proton translocation by subunit *c* of the ATP synthase

Vinit K. Rastogi & Mark E. Girvin

Biochemistry Department, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA

F_1F_0 ATP synthases use a transmembrane proton gradient to drive the synthesis of cellular ATP. The structure of the cytosolic F_1 portion of the enzyme and the basic mechanism of ATP hydrolysis by F_1 are now well established, but how proton translocation through the transmembrane F_0 portion drives these catalytic changes is less clear. Here we describe the structural changes in the proton-translocating F_0 subunit *c* that are induced by deprotonating the specific aspartic acid involved in proton transport. Conformational changes between the protonated and deprotonated forms of subunit *c* provide the structural basis for an explicit mechanism to explain coupling of proton translocation by F_0 to the rotation of subunits within the core of F_1. Rotation of these subunits within F_1 causes the catalytic conformational changes in the active sites of F_1 that result in ATP synthesis.

(e-mail: girvin@aecom.yu.edu). Atomic coordinates are available at the Protein Data Bank for the structure of subunit *c* at pH 8 (1c99), the refined structure of subunit *c* at pH 5 (1c0v), and the model for the ac_{12} complex (1c17). Chemical shifts are available at the BioMagResBank for subunit *c* at pH 8 (entry 4151) and pH 5 (entry 4146).

Structure Summary | 3D View | Annotations | Sequence | Sequence Similarity | Structure Similarity | Experiment

NMR Ensemble

View in 3D JSmol or PV (in Browser)

1C99

ASP61 DEPROTONATED FORM OF SU ATP SYNTHASE OF ESCHERICHIA COLI

DOI: 10.2210/pdb1c99/pdb

Classification: PROTON TRANSPORT HYDROLAS

Deposited: 1999-04-30 **Released**: 1999-11-19

Deposition author(s): Rastogi, V.K., Girvin, M.E.

Organism: Escherichia coli

Expression System: Escherichia coli

Structural Biology Knowledgebase: 1C99 (1 model >23 annotations) SBKB.org

Display Files | Download Files

- FASTA Sequence
- PDB File
- PDB File (Header)
- mmCIF File
- mmCIF File (Header)
- PDBML/XML File
- PDBML/XML File (Header)

Experimental Data Snapshot

Method: SOLUTION NMR

wwPDB Validation reports are not available for NMR entries.

Contact U

2 Structure Hits | 1 Citation

Query Parameters: Query Details | Save Query to MyPDB

Simple query for a list of PubMed IDs: 10580496

Query Refinements: ***Select an item or pie chart***

Organism
- Escherichia coli only (2)

UniProt Molecule Name
- ATP synthase subunit c (2)
- ATP synthase subunit a (1)
- refine query

Taxonomy
- Bacteria only (2)

Experimental Method
- Solution NMR (2)

Release Date
- Nov 1999 (2)
- refine query

Polymer Type
- Protein (2)

Enzyme Classification
- 3.6.3.14: H(+)-transporting two ... (1)

SCOP Classification
- F1F0 ATP synthase subunit C (2)
- F1F0 ATP synthase subunit A (1)

Protein Symmetry
- Asymmetric (2)
- refine query

Protein Stoichiometry
- Monomer (1)
- Homomer (1)
- refine query

Membrane Proteins
- Subunit C of the F_1F ... (2)

Refine Query with Advanced Search

Show only representatives at Select ▼ sequence identity

RESOURCES

FORCE11
The Future of Research Communications and e-Scholarship

www.force11.org/node/4463

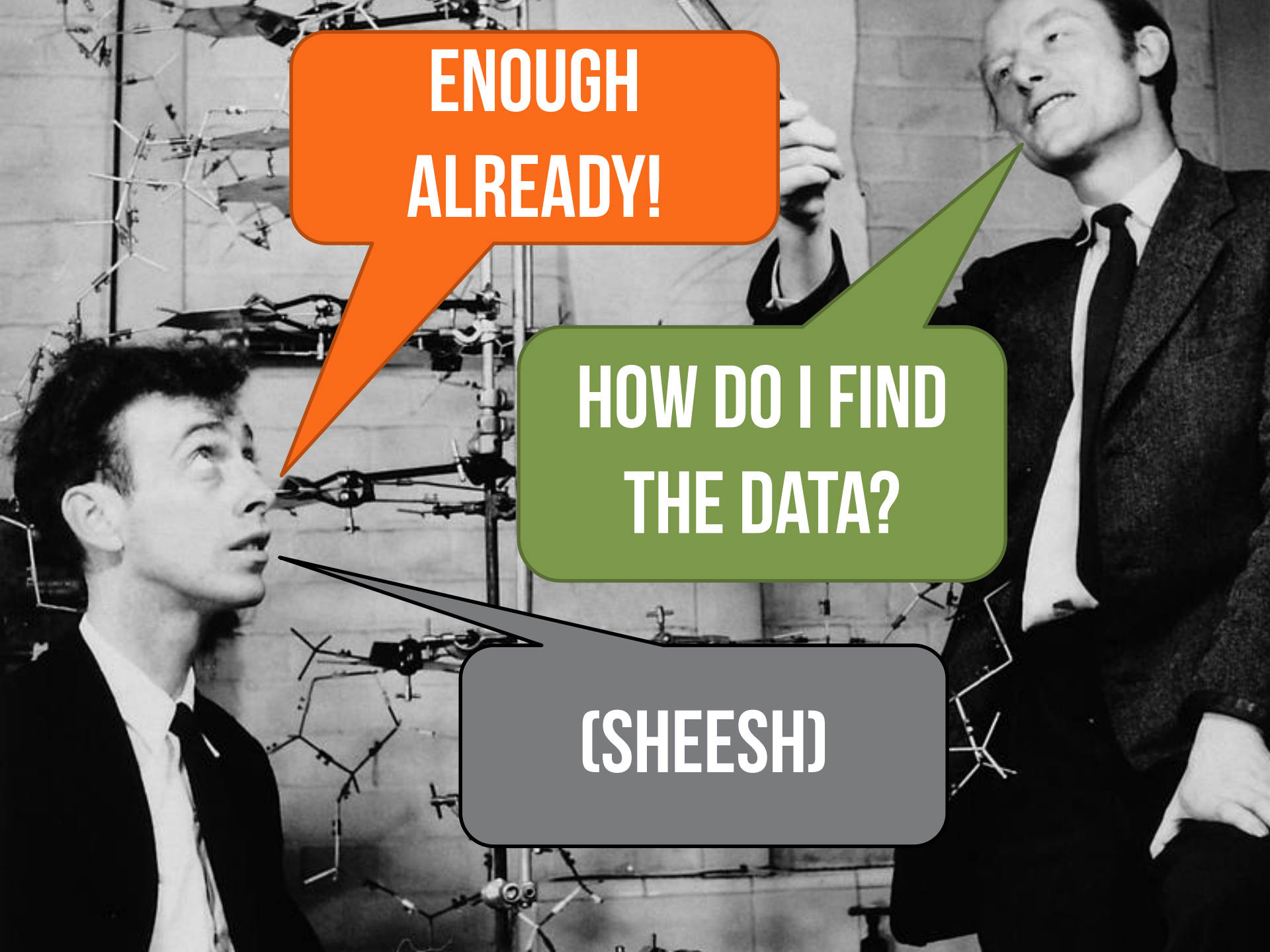
biosharing

biosharing.org/bsg-000532

mibbi

http://www.biosharing.org/standards/mibbi

REPORTING STANDARDS


ENOUGH ALREADY!
HOW DO I FIND THE DATA?
(SHEESH)

YOUR DATA I HOPE YOU CAN FIND...

LITERATURE & DATA

LITERATURE -> DATA

PubMed

25,628,538

3,851,969

1,779 | 320 | 3,501

1

Percept Mot Skills. 2010 Apr;110(2):580-92.

Nonconscious activation of an elderly stereotype and speed of driving.

Branaghan RJ, Gray R.

[1au]	1st Author
[au]	Author
[lastau]	Last Author
[ti]	Article Title
[ta]	Journal Title Abbreviation
[journal]	Journal Title
[dp]	Date of Publication

[FIELD TAGS]

PubMed.gov
U.S. National Library of Medicine
National Institutes of Health

Search: PubMed
RSS Save search Limits Advanced search Help
smith[1au] AND journal of biological chemistry[journal] AND proteoglycans[M
Search Clear

J Biol Chem. 1985 Sep 5;260(19):10761-7.

Interaction of proteoglycans with the pericellular (1 alpha, 2 alpha, 3 alpha) collagens of cartilage.

Smith GN Jr, Williams JM, Brandt KD.

Abstract

Interaction between cartilage proteoglycan and the collagen(s) composed of 1 alpha, 2 alpha, and 3 alpha chains was studied in vitro. Most of the collagen was insoluble under the conditions of assay (0.15 M NaCl, 0.008 M phosphate buffer, pH 7.4; 4 degrees C) and was in the form of fibrils 20 nm in diameter or thinner. The larger fibrils had 60-70 nm periodicity, characteristic of native collagens. Proteoglycan monomers which had been labeled by incubating cartilage slices in vitro with Na2 35SO4 were used to assay the interaction. The insoluble collagen fraction bound proteoglycan from solution. At proteoglycan:collagen ratios lower than 1:2, binding was rapid and linear, and the dissociation constant was 1.7 X 10(-9) M. At higher proteoglycan:collagen ratios, more proteoglycan was bound, but at a slower rate. Binding of proteoglycan to collagen did not require fibrils, since soluble 1 alpha, 2 alpha, and 3 alpha containing collagen also bound to proteoglycan and formed an insoluble complex. Denatured collagens did not bind proteoglycan or compete for binding with normal collagen. Optimum binding occurred with intact proteoglycan, but proteoglycan which had been treated with protease was also bound at low levels. Both protease-treated proteoglycan and free chondroitin sulfate competed with intact proteoglycan in the binding assays, but neither chondroitinase ABC-treated proteoglycan nor the oligosaccharides produced by digestion of chondroitin sulfate with testicular hyaluronidase altered the binding of proteoglycan to collagen. Hyaluronic acid did not compete with radioactive proteoglycan, but heparin and dextran sulfate were extremely effective inhibitors of binding. These data suggest a relatively nonspecific interaction between sulfated polyanions and 1 alpha, 2 alpha, and 3 alpha containing collagens. However, given the location of these collagens near the chondrocyte surface, the interaction of fibrillar 1 alpha, 2 alpha, 3 alpha collagen with proteoglycan is likely to occur and to be of biological importance.

PMID: 4030769 [PubMed - indexed for MEDLINE]

Google scholar smith journal of biological chemistry proteogly Search Advanced Scholar Search

Scholar Articles and patents | anytime | include citations | Create email alert Results **1** - **10** of about **7,980**. (**0.26** sec)

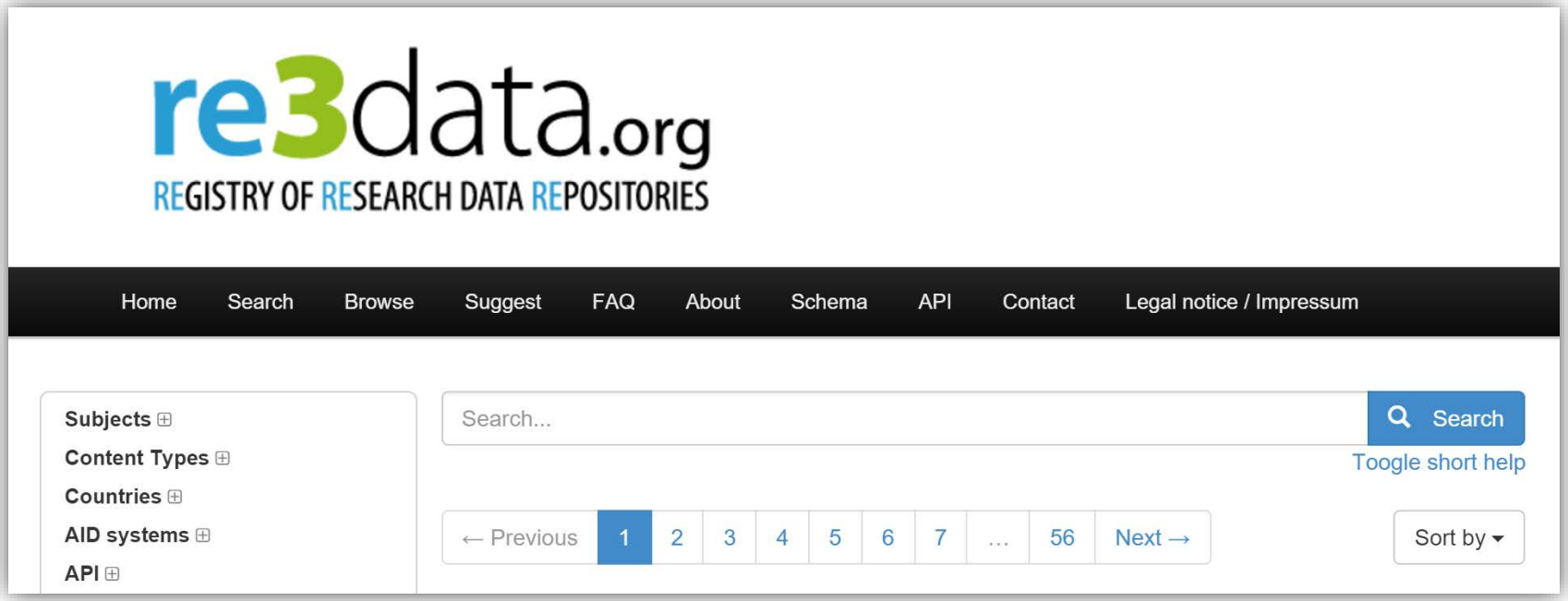
Cell surface heparan sulfate **proteoglycans**
..., LS Ostedgaard, C O'Riordan, AE **Smith**... - Cell, 1992 - ASBMB
Table of Contents. May 15, 1992; 267 (14). Cell surface heparan sulfate **proteoglycans**. M Yanagishita and; VC Hascall. J. Biol. Chem. 1992 267: 9451-9454. **...** BC Tilly,; MC Winter,; LS Ostedgaard,; C O'Riordan,; AE **Smith**,; and MJ Welsh. J. Biol. Chem. 1992 267: 9470-9473. **...**
Cited by 317 - Related articles - All 6 versions

The potential role of **proteoglycans** in cationic lipid-mediated gene delivery [HTML] from jbc.org
CM Wiethoff, JG **Smith**, GS Koe... - ... of **Biological Chemistry**, 2001 - ASBMB
Recent evidence supports a role for **proteoglycans** in polycation-mediated gene delivery. Therefore, the interaction of glycosaminoglycans with cationic lipid-DNA complexes (CLDCs) has been characterized using a combination of biophysical approaches. At low **...**
Cited by 67 - Related articles - BL Direct - All 3 versions

Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer [HTML] from jbc.org
..., JL Slightom, MJ Bienkowski, CW **Smith**... - ... of **Biological Chemistry**, 1999 - ASBMB
... the Protein Science, § Genomics Research, ‡ Structural, Analytical, and Medicinal **Chemistry**, and ¶ Research **...** several proteins may be retained in complexation with the **proteoglycans**, or released **...** of HSPG function impacts a wide spectrum of **biological** phenomena underlying **...**
Cited by 133 - Related articles - BL Direct - All 5 versions

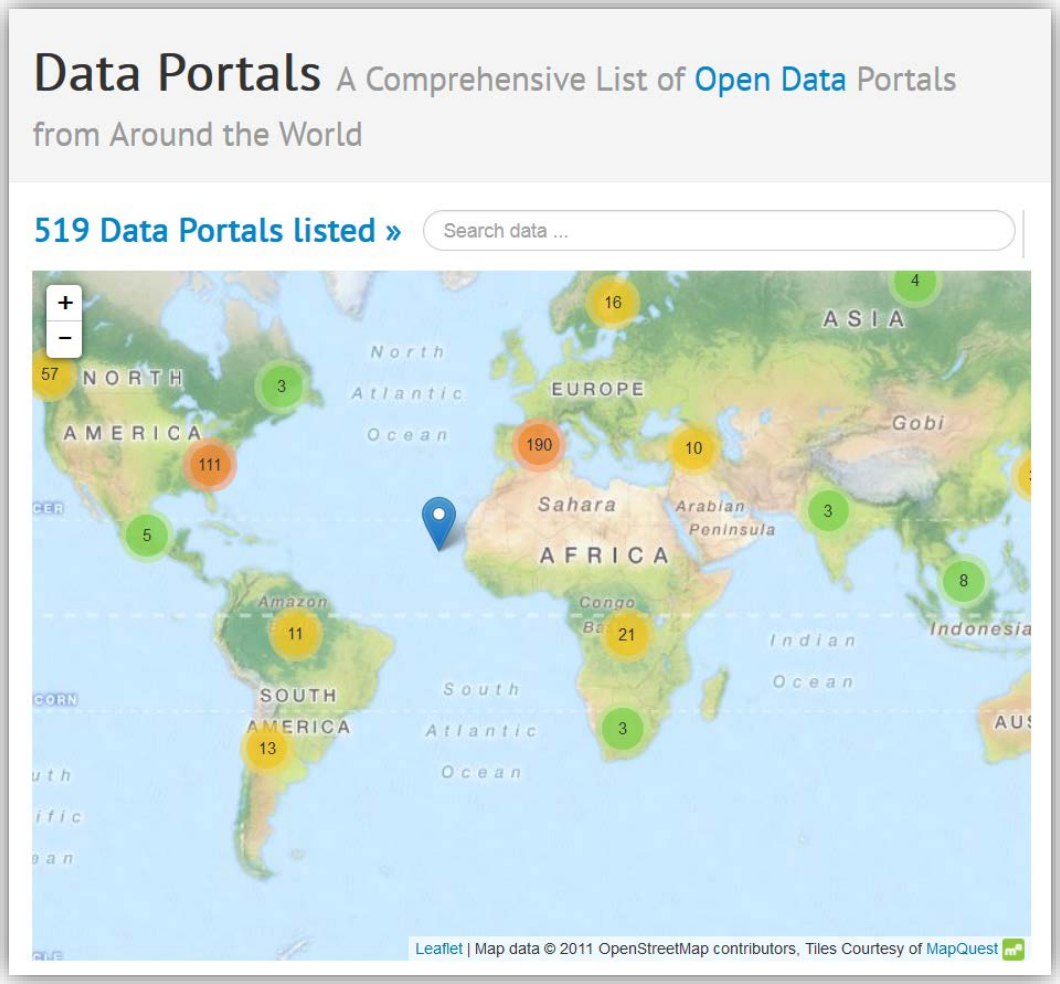
Structural analysis of sequences O-linked to mannose reveals a novel Lewis X structure in cranin (dystroglycan) purified from sheep brain [HTML] from jbc.org
NR Smalheiser, SM Haslam, M Sutton-**Smith**... - ... of **Biological Chemistry**, 1998 - ASBMB
The **Journal** of **Biological Chemistry** Skip to main page content. **...** Footnotes. ↵* This work was supported by the Biotechnology and **Biological** Sciences Research **...** Abstract/FREE Full Text. ↵: Haslam SM,; Coles GC,; Munn EA,; **Smith** TS,; **Smith** HF,; Morris HR,; Dell A. (1996) J. Biol. **...**
Cited by 87 - Related articles - BL Direct - All 4 versions

DATA SOURCES



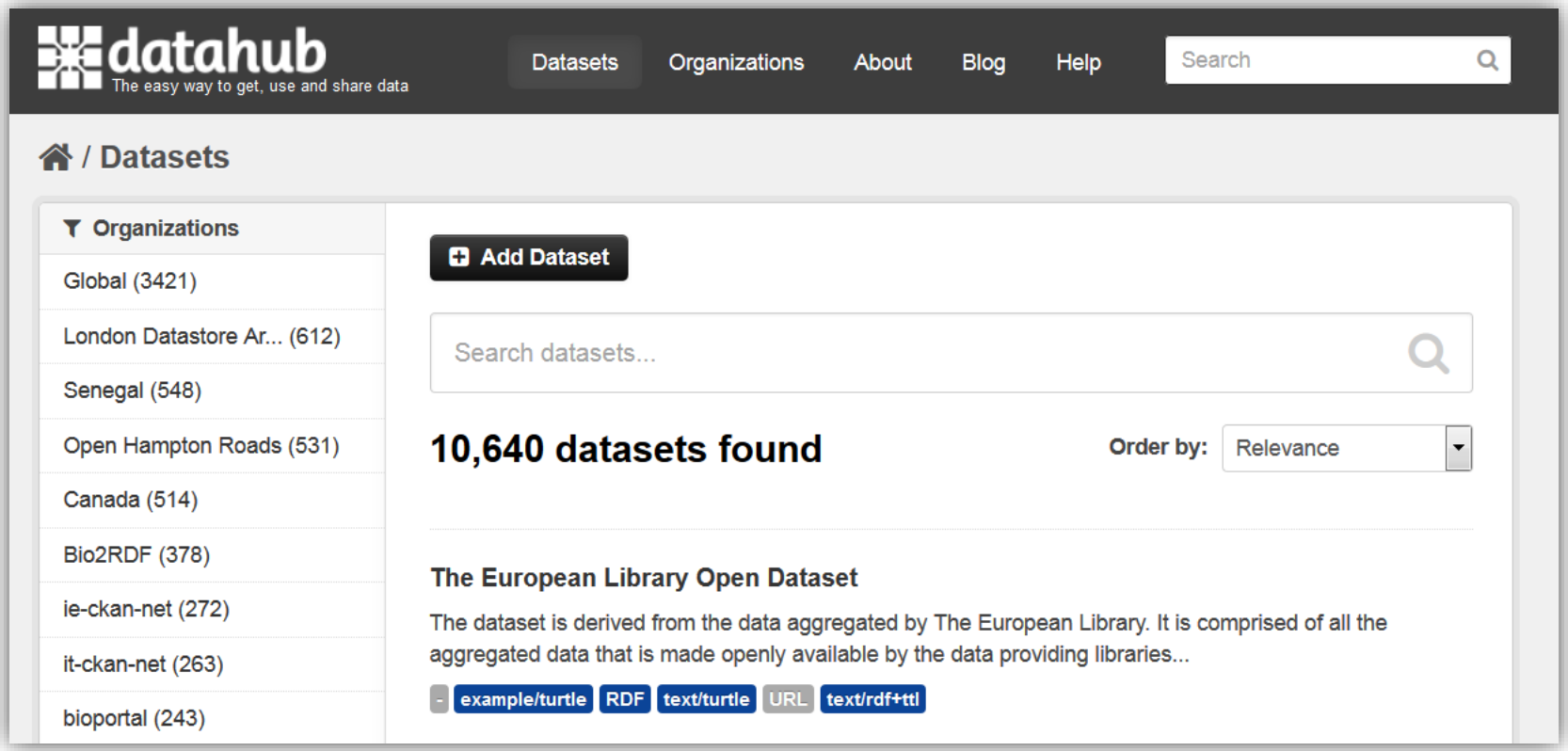
REGISTRY OF RESEARCH DATA REPOSITORIES

http://re3data.org



DATA PORTALS

http://dataportals.org



DATAHUB

http://datahub.io/dataset

ZANRAN

http://zanran.com

DATA.GOV

http://data.gov

U.S. National Library of Medicine
National Institutes of Health

Contact NLM

Search

The World's Largest Medical Library

Databases | Find, Read, Learn | Explore NLM | Research at NLM | NLM for You

NIH Trans-NIH BioMedical Informatics Coordinating Committee (BMIC)

BMIC Home | CDE Resource Portal

Home

NIH Data Sharing Repositories

This table lists NIH-supported data repositories that accept submissions of appropriate data from NIH-funded investigators (and others). Also included are resources that aggregate information about biomedical data and information sharing systems. The table can be sorted according by name and by NIH Institute or Center and may be searched using keywords so that you can find repositories more relevant to your data. Links are provided to information about submitting data to and accessing data from the listed repositories. Additional information about the repositories and points-of-contact for further information or inquiries can be found on the websites of the individual repositories.

Show 50 entries

Search:

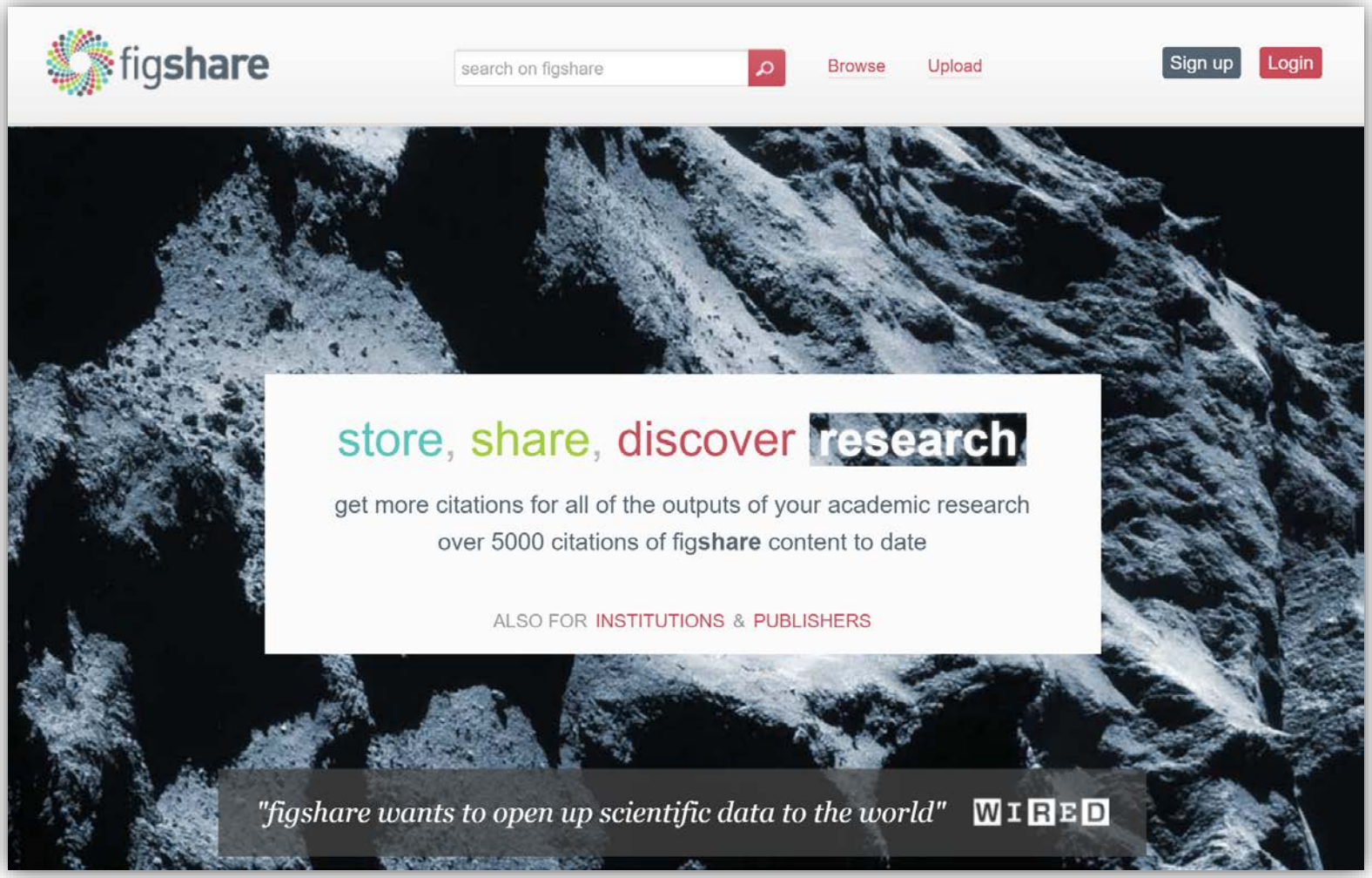
IC	Repository Name	Repository Description	Data Submission Policy	Access to Data
NCI	The Cancer Imaging Archive (TCIA)	The Cancer Imaging Archive (TCIA) is a large archive of medical images of cancer accessible for public download. All images are stored in DICOM file format. The images are organized as "Collections", typically patients related by a common disease (e.g. lung cancer), image modality (MRI, CT, etc) or research focus.	How to Submit Data to TCIA	How to Access TCIA Data
NCI (NHGRI, NIGMS)	PeptideAtlas	PeptideAtlas is a multi-organism, publicly accessible compendium of peptides identified in a large set of tandem mass spectrometry proteomics experiments. Mass spectrometer output files are collected for human, mouse, yeast, and several other organisms, and searched using the latest search engines and protein sequences.	How to Submit Data to PeptideAtlas	How to Access PeptideAtlas Data
NHGRI	FlyBase: A Drosophila Genomic and Genetic Database	Drosophila Genomic and Genetic database that includes proteomics data, microarrays and Tiling BAC's.	How to Submit Data to FlyBase	How to Access FlyBase Data
NHGRI	The Zebrafish Model Organism Database (ZFIN)	ZFIN serves as the zebrafish model organism database. It aims to: a) be the community database resource for the laboratory use of zebrafish, b) develop and support integrated zebrafish genetic, genomic and developmental information, c) maintain the definitive reference data sets of zebrafish research information, d) to link this information extensively to corresponding data in other model organism and human databases, e) facilitate the use of zebrafish as a model for human biology, and f) serve the needs of the research community.	How to Submit Data to ZFIN	How to Access ZFIN Data
NHGRI	WormBase	WormBase is an international consortium of biologists and computer scientists dedicated to providing the research community with accurate, current, accessible information concerning the genetics, genomics and biology of C. elegans and related nematodes.	How to Submit Data to WormBase	How to Access WormBase Data
NHLBI	Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC)	The goal of BioLINCC is to facilitate and coordinate the existing activities of the NHLBI Biorepository and the Data Repository and to expand their scope and usability to the scientific community through a single web-based user interface.	How to Submit Data to BioLINCC	How to Access BioLINCC Data

DATA SHARING REPOSITORIES

https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html

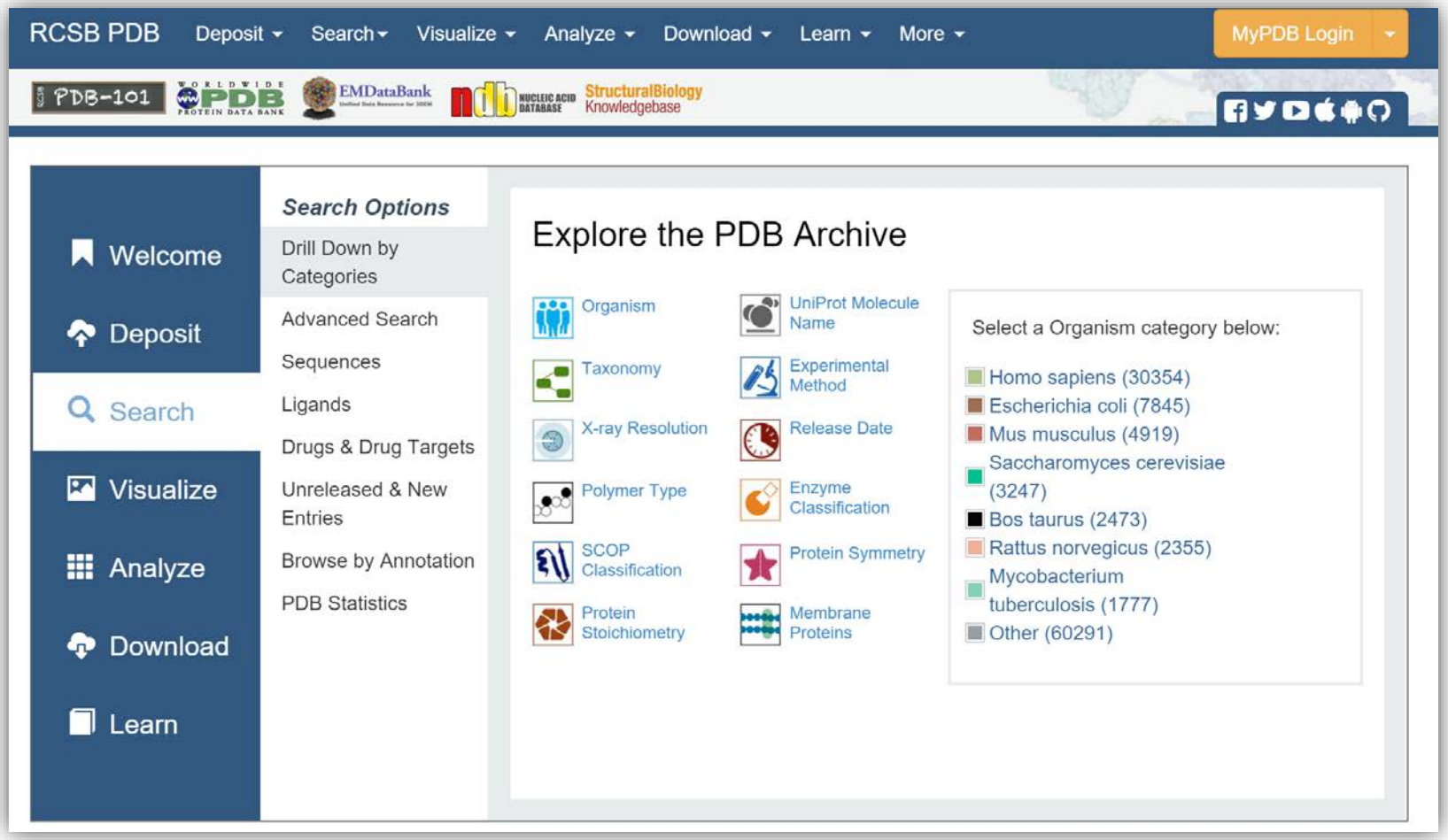
DATA DRYAD

http://datadryad.org



FIGSHARE

www.figshare.com



PDB – SHINING STAR OF SUBJECT SPECIFIC DATASETS

www.rcsb.pdb.org

Stanford Large Network Dataset Collection

- Social networks : online social networks, edges represent interactions between people
- Networks with ground-truth communities : ground-truth network communities in social and information networks
- Communication networks : email communication networks with edges representing communication
- Citation networks : nodes represent papers, edges represent citations
- Collaboration networks : nodes represent scientists, edges represent collaborations (co-authoring a paper)
- Web graphs : nodes represent webpages and edges are hyperlinks
- Amazon networks : nodes represent products and edges link commonly co-purchased products
- Internet networks : nodes represent computers and edges communication
- Road networks : nodes represent intersections and edges roads connecting the intersections
- Autonomous systems : graphs of the internet
- Signed networks : networks with positive and negative edges (friend/foe, trust/distrust)
- Location-based online social networks : Social networks with geographic check-ins
- Wikipedia networks and metadata : Talk, editing and voting data from Wikipedia
- Twitter and Memetracker : Memetracker phrases, links and 467 million Tweets
- Online communities : Data from online communities such as Reddit and Flickr
- Online reviews : Data from online review systems such as BeerAdvocate and Amazon

SNAP networks are also availalbe from UF Sparse Matrix collection. **Visualizations of SNAP networks** by Tim Davis.

SNAP

https://snap.stanford.edu/data/index.html

Cornell University

ROPER CENTER

Polls Membership Blog Support About Us Search iPOLL Search Datasets

Polls

Search ...

Explore the Roper Center Archives

The Roper Center holds polling data ranging from the 1930s, when survey research was in its infancy, to the present. Over 650,000 questions of U.S. samples can be searched at the question-level in **iPOLL**, while **Roper*Express*** offers downloads of over 20,000 datasets from over 100 countries. As part of its mission to maximize use and access for the data in its collection, the Center also provides multiple resources on politics, elections, U.S. policy, and social issues.

Politics

The Roper Center archives provide essential data for understanding U.S. politics. Resources on politics and elections are drawn from the collection's national public opinion polls, voter exit polls at the state and national

Issues

The Roper Center collection contains public opinion data on an extremely wide range of topics. Go to **Topics at a Glance** for overviews of popular subject areas or browse recent **Issue Briefs** below.

Data

The Roper Center archive consists of **over 20,000 datasets** from public opinion surveys dating back to the 1930s. By collecting and preserving the collection's U.S. and international datasets, the Center provides

POLLS – THE ROPER CENTER

http://ropercenter.cornell.edu/polls/

CIVICAPPS FOR GREATER PORTLAND [BETA]

DATASETS IDEAS APPS CONTRIBUTE ABOUT

SEARCH

Datasets

Welcome to the CivicApps Data Catalog!

The purpose of the Catalog is to make public data easier to access, and in the process make government data more useful for everyone. The vision is that diverse and rich regional data will facilitate an unprecedented range of new information and options, enabling others to derive solutions to citizens needs.

The Catalog is unique because it includes public datasets from a wide array of local government jurisdictions. It is the only inter-jurisdictional repository of local public data of its kind in the United States, at least as far as we know. This would not be possible without data owners across the greater Portland area joining hands to make government an open data platform.

So check out the Catalog for yourself, beginning with the most recent datasets displayed first. Tell us which datasets are most important to you. If you don't see data that you think should be included, suggest new datasets via the Contact form.

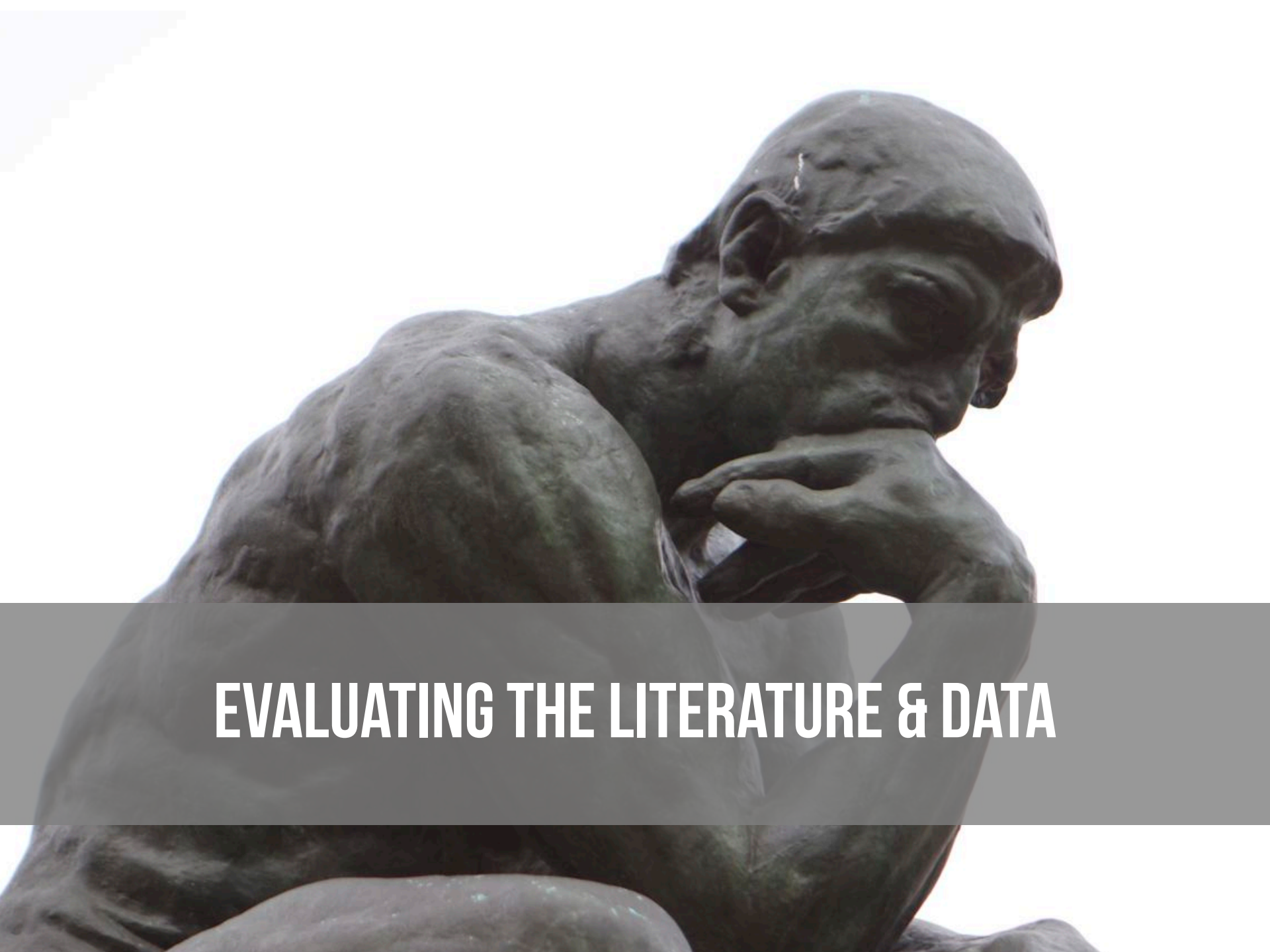
Name	Rating	Agency	Category	Format
Crime Incidents - 2014 This static file reflects crimes reported to the City of Portland Police Bureau in 2014. Classification of the crime type is based on the Uniform Crime...		City of Portland Police	Public Safety	CSV/Text
Garbage Collection Schedule Areas Residential garbage / recycling / composting curbside collection days and schedules by hauler.		City of Portland BPS	Community	Shapefile

CIVICAPPS

www.civicapps.org/datasets

GOOGLE

HERE'S A TIP...
DON'T ASK FOR WHAT YOU WANT
DESCRIBE WHAT YOU EXPECT TO FIND

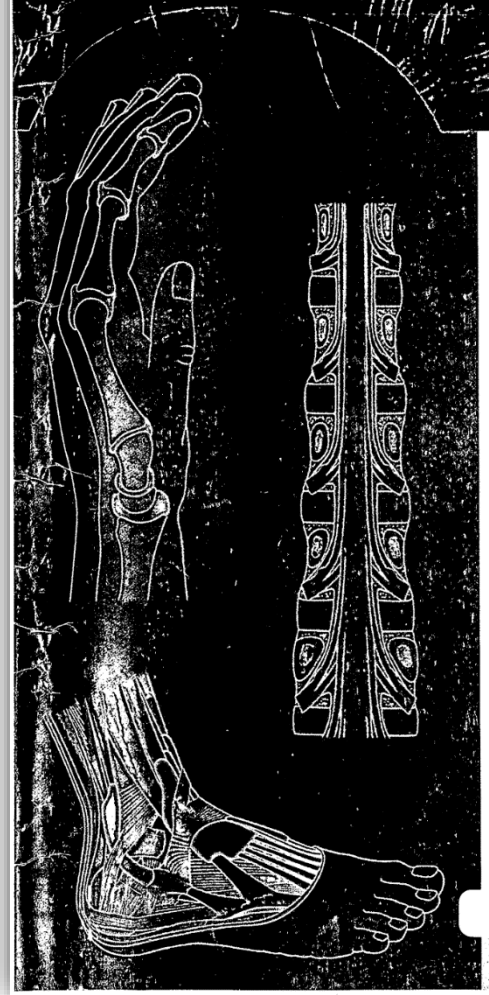


EVALUATING THE LITERATURE & DATA

Volume 2 Issue 1 2003

AUSTRALASIAN JOURNAL of

BONE & JOINT MEDICINE



INTERNATIONAL NEWS 2

REVIEW
Osteoporosis
Meta-analysis of therapies for postmenopausal osteoporosis 4

AUSTRALASIAN RESEARCH 7

LOCAL COMMENTARY
Screening
Dual-energy X-ray absorptiometry versus quantitative ultrasound 8

REVIEW
Screening
US osteoporosis screening guidelines review 9

CASE REPORTS
Iliac bone defects revealing sarcoidosis 16
Systemic lupus erythematosus 18

CONFERENCE HIGHLIGHTS
American Society for Bone and Mineral Research 2002 19

MSD.050.354

MSD.050.354.0001

ELSEVIER

ISSN: 1447-5529

International Journal of Advanced Computer Technology (Online)

http://www.ijact.org

Email:editor@ijact.org,submit_ijact@yahoo.in

==========

TITLE: [redacted]

AUTHOR: David Mazi`eres and Eddie Kohler

Your manuscript has been accepted with minor changes for publication in the *International Journal of Advanced Computer Technology* (IJACT).

SUBMIT THE REVISED VERSION OF YOUR RESEARCH PAPER

to submit@ijact.org,submit_ijact@yahoo.in

by **25/10/14**

In a separate document—the PAPER EVALUATION—you will receive the reviewers' comments and ratings.

IJACT Submission and Publication Process

Please read and follow the instructions below! Failure to complete each step at the appropriate time will likely result in the postponement of the publication of your manuscript, or possibly its rejection altogether. You can find a summary of this step at Paper Template at http://www.ijact.org/papersubmission.htm.

STEP : REFORMAT (two columns) AND RESUBMIT YOUR REVISED MANUSCRIPT based on the reviewers' comments and suggestions.

It is CRITICALLY IMPORTANT that your manuscript conform to the guidelines on the IJACT website noted above. Please review these guidelines and the sample documents before resubmitting.

In the following bulleted items, I have summarized a few additional points that you should check

REVIEW FORM

Paper ID	IJ0350030
Paper Title	[redacted]

NOTE: 1. Excellent 2.Very Good 3. Good 4.Fair 5. Very Poor

1. Appropriateness to publish in IJACT — Option: Excellent
2. Accuracy — Option: Excellent
3. Innovation — Option: Very Good
4. Relevance — Option: Very Good
5. Presentation — Option: Good
6. Quality of writing — Option: Very Good
7. Application of paper as introductory study material for revolutionary new concepts — Option: Very Good
8. Overall Recommendation (Mark the right option):

Strongly Accepted	
Accepted	✓
Accepted with minor changes	
Marginally Accepted	

International Journal of Advanced Computer Technology (Online)

http://www.ijact.org

Email:editor@ijact.org,submit_ijact@yahoo.in

9. Reviewer Comments:

a. Please Use latest references in order to increase your paper quality.
b. Introduction part is precisely explained.
c. Kindly prepare your camera ready version of paper as per IJACT paper format.
d. Use high resolution images in order to increase the appearance of paper.

Get me off Your Fucking Mailing List

David Mazières and Eddie Kohler
New York University
University of California, Los Angeles
http://www.mailavenger.org/

Abstract

Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list.

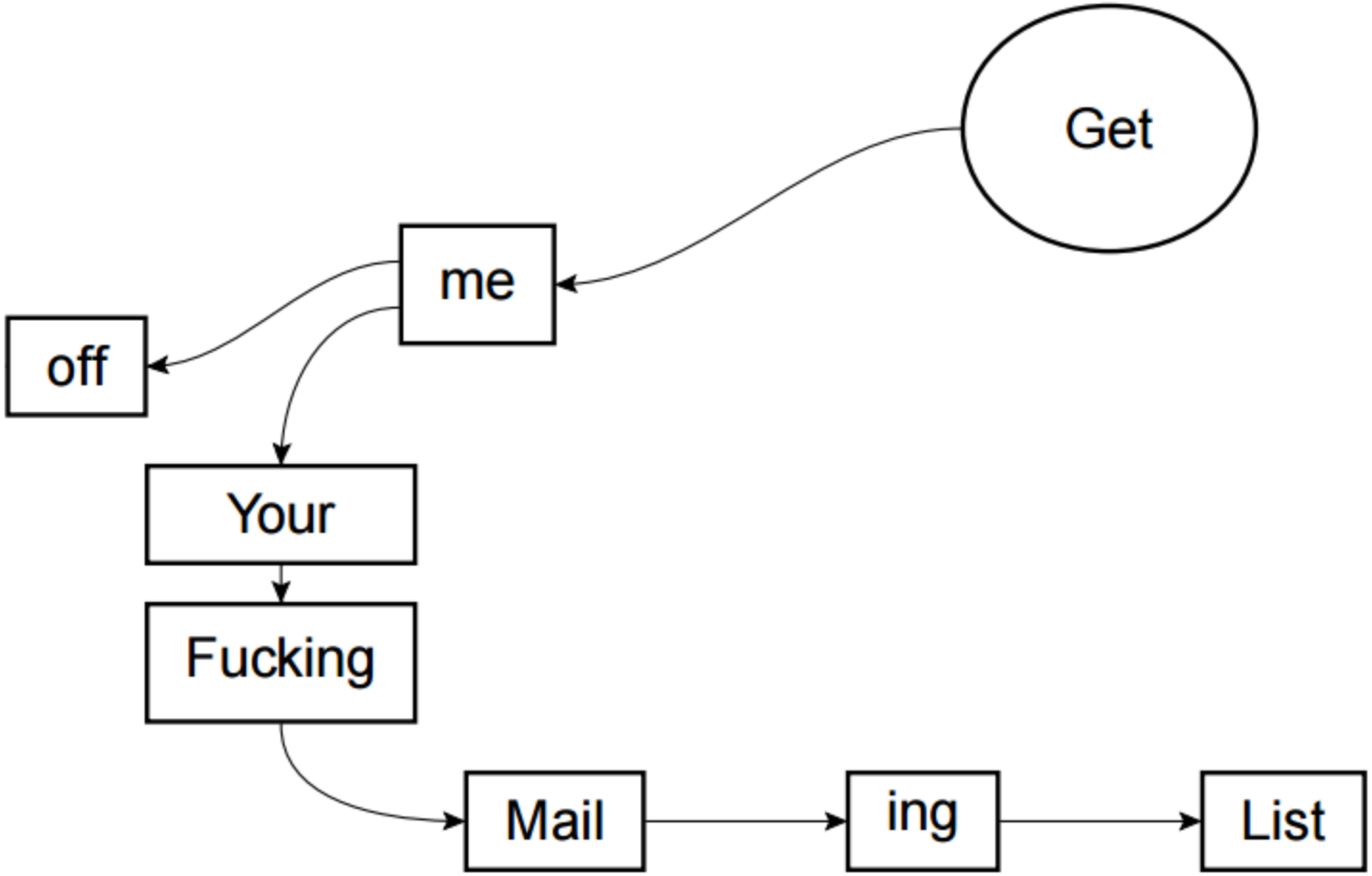


Figure 1: Get me off your fucking mailing list.

The Most Influential Paper Gerard Salton Never Wrote

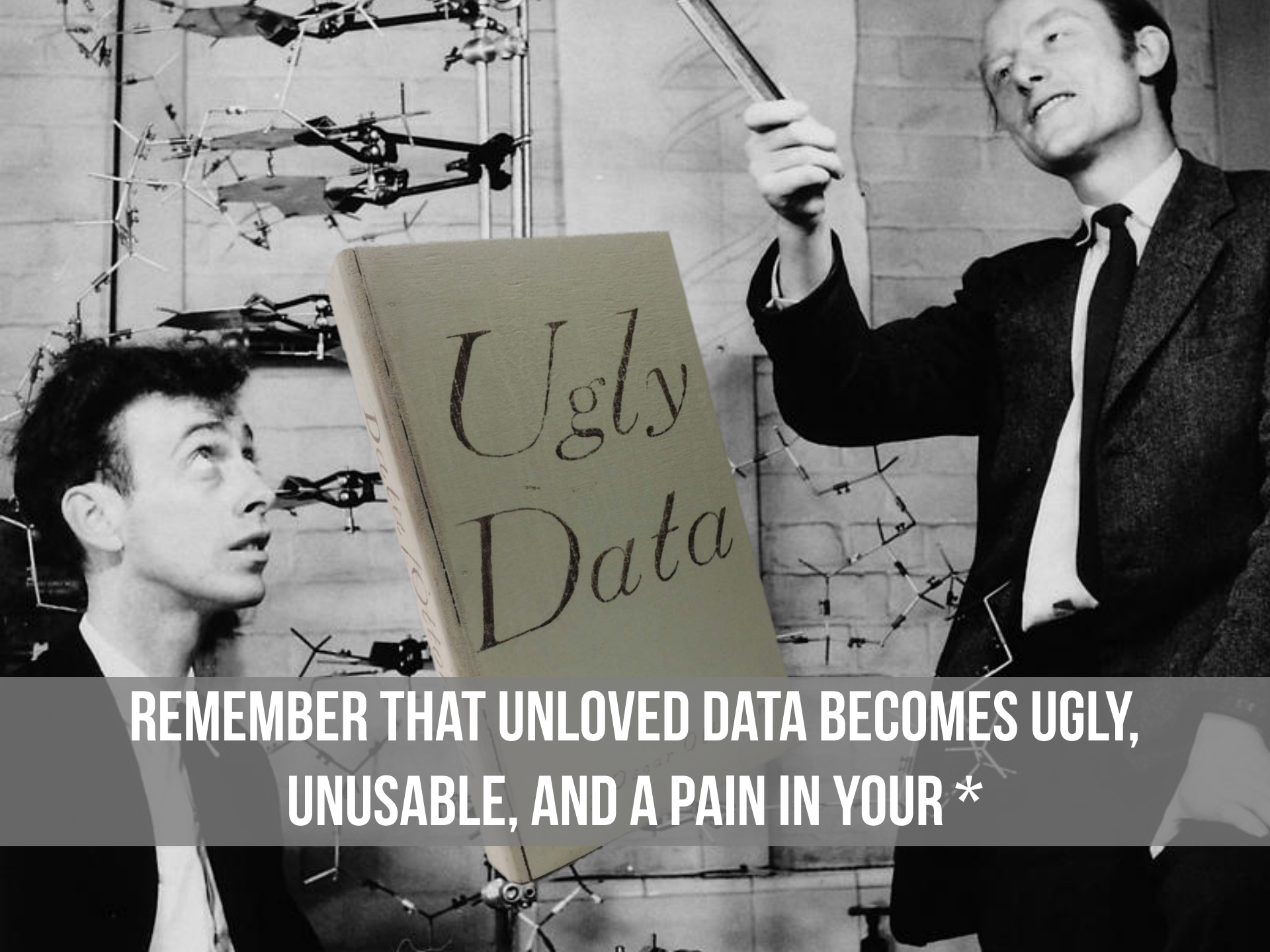
DAVID DUBIN

ABSTRACT
Gerard Salton is often credited with developing the vector space model (VSM) for information retrieval (IR). Citations to Salton give the impression that th
1970 a
longer
model
those a
often c
Retriev
citatio
were o
late 19
but rather as models of specific computations. Citations to the phantom paper reflect an apparently widely held misconception that the operational features and explanatory devices now associated with the VSM must have been introduced at the same time it was first proposed as an IR model.

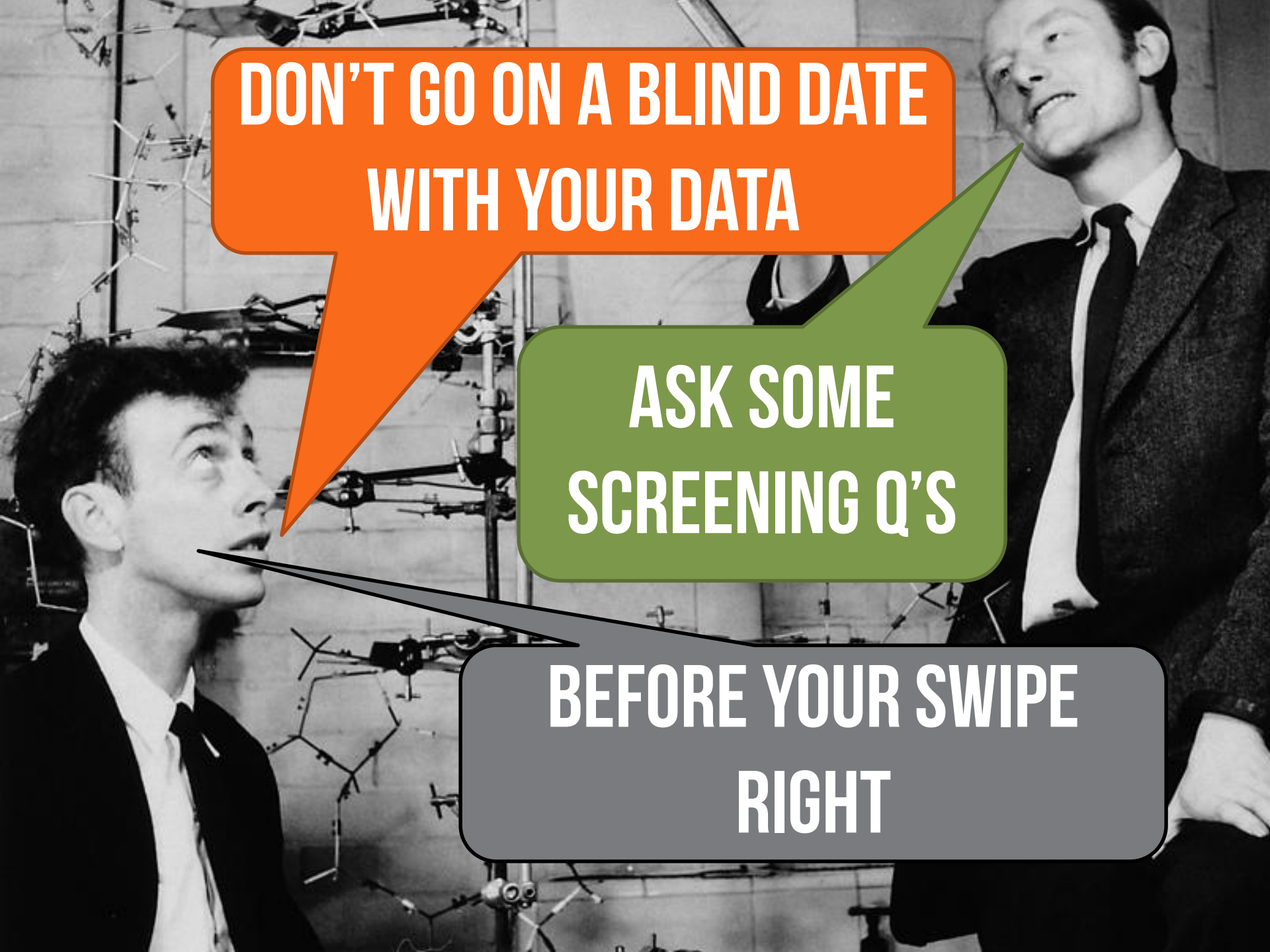
A vector space model for automatic indexing
G Salton, **A** Wong, CS Yang - Communications of the ACM, 1975 - dl.acm.org
... 1. **Vector** representation of document **space**. ... improvements in performance, it would then seem to be sufficient to increase the simi- larity between document **vectors** located in ... of 1/(Fx.F.2) are shown in Table II(b). Here the similarity between documents inside **a** cluster decreases ...
Cited by 5706 Related articles All 39 versions Cite Save More

GERARD SALTON

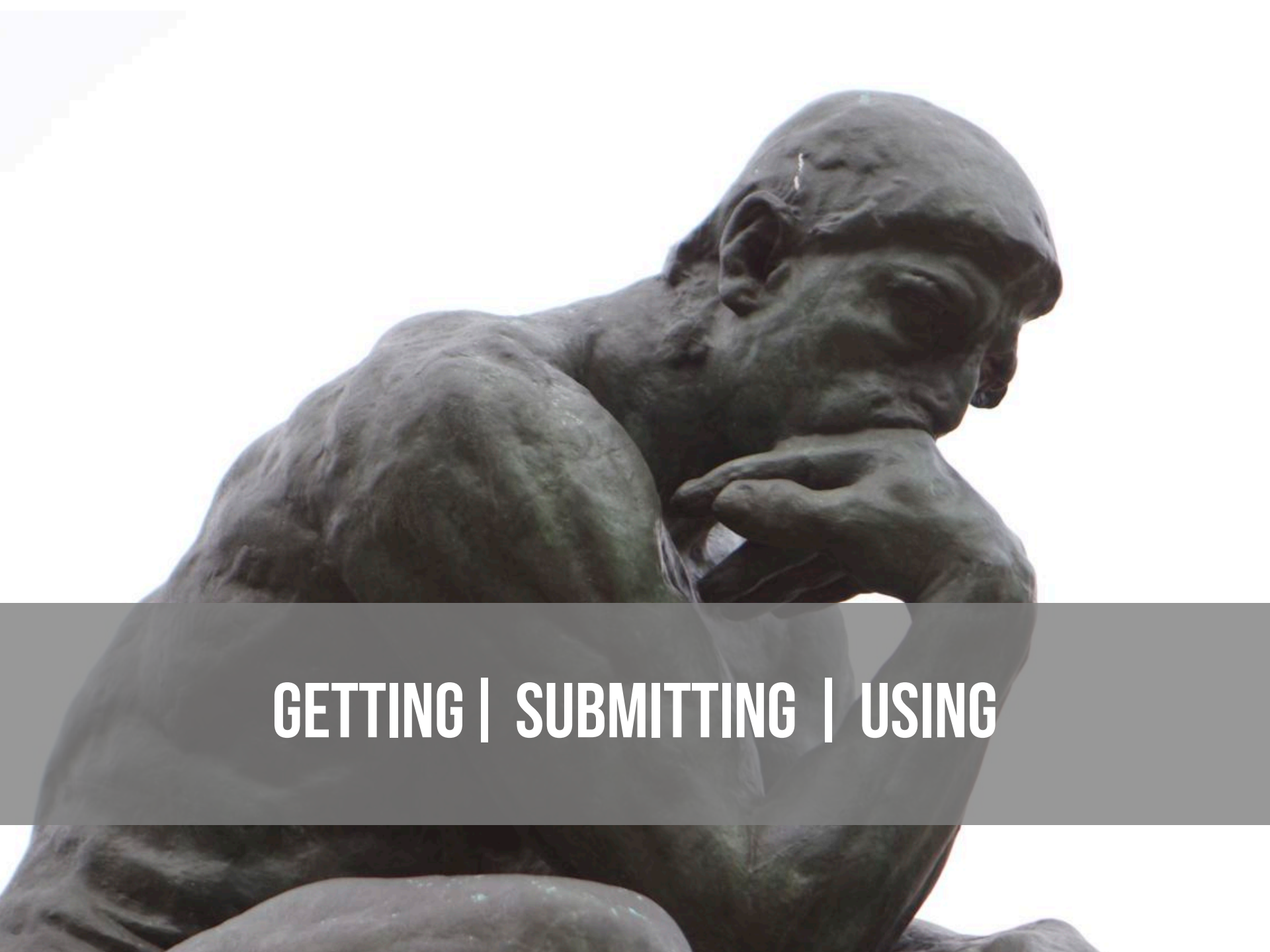
YEAH...
BUT DATA IS JUST... DATA
IT WOULDN'T LIE!!



REMEMBER THAT UNLOVED DATA BECOMES UGLY, UNUSABLE, AND A PAIN IN YOUR *


DON'T GO ON A BLIND DATE WITH YOUR DATA
ASK SOME SCREENING Q'S
BEFORE YOUR SWIPE RIGHT

FINDING DATA EXERCISE


GETTING | SUBMITTING | USING