# Open Data, Open Humans

#### Madeleine Price Ball

PGP Harvard, Director of Research PersonalGenomes.org Bits ↔ Biology, May 1 2014



### Human Biology → Bits → ?

LAL TGGAT CAGAG CTAA CACCT6 .10010001011110 /0111100101011110 10111110101110101 01011011100010101 011111111010101000 16100100000101( )01011100111011 J01011110100001 **11100010111000**1 AGCTGCACGTGG CTGGC CCATC/ **\GTGC IGGCT GCTG ICCA** GGAT \TTA CAT1 iC10 )101( )1000 01111 1110 1111 1011 0111 0101 1111 1111 (111 111101

011

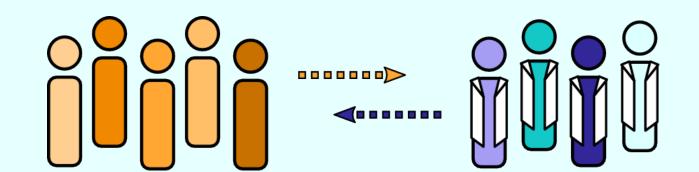
- - 4

ACCTT **`0010100** 00101 J10101111010. 000101010010001 1011010100011101( )01010111000001010 **J01011111110010111** 011101000010101001 0010001111011100 **JAGAACTTCAGGCTC** AAAGAATTCACCCC/ CTAATGCCCTGGCCC AGGTT TTTGT TGGA TGAGC 10001 01101 11101 (1110 J1000 **0001** /10000 01011 01010 10110 J1101 11101 0011<sup>1</sup> . 10000 20010 0001

**CAG** TAGTG AGCT00 .00110 J111010001110 J011110011110101 0111111110001011 0111001001011110 10011101001000010 10100101111010011<sup>+</sup> 11111111001010001( 100111100GCACTGT CAACGTGCTGGTCTG **FGCAGGCTGCCTATC** *<b>TCACTAAGCTCG* TCCAACTACT/ ΑΑΤΑΑΑΑΑΑ 11010 )1010( J1111/ 0010 00101 1011 111101 1101<sup>4</sup> 00011<sup>4</sup> 0010L 010101 111100 111000 )11110 1000111 0011(

#### Human subjects research

#### A very brief review of data sharing.



#### NATIONAL INSTITUTE OF GENEL HUMAN GENETIC INFORMED CO

(Submitters associated with an organization MUST use an Informed Consent Form approv all of the elen

Overview	of the	FILE
----------	--------	------

	SCIENCE
- OF GE	NERAL MEDICAL SCIENCE TIC CELL REPOSITORY CONSENT FORM
INSTITUTE OF GENE	TIC CELL REPOSIT
INFORMEL	CONSE
anizati	ion overseen has an I

sociated with an e	] Mai
	te che
	ו Ind
Diagnosis:	can
Diagnos tof symptoms:	His
Age at onset of symptoms:	'tal
Primary Care Physics testing)	ic
Age at other Geneticist  Primary Care Physician Geneticist  Primary Care Physician Geneticist  Primary Care Physician (prenatal testing) (prenatal testing) (pre	η
ay, benetice animal	
selchair braces orthotics hearing aid service animal	
- braces [] orthotics	
elchair L bias	_
] Other:(able):	
] Other:	-

/	TO BE FILLED IN BY CORIELL	) THE NIGM
	Repository Number:	its, contact > 8:00 AM
National Institute of General Medical Sciences	Date Received:	nigms@r
(NIGMS) Human Genetic Cell Repository Subm		7
Please check or complete all applicable items. Please also attach any records, diagnostic test/laboratory reports or clinic summaries that are	aliniant descriptions, case histories medical	ser
enected rarmy members (if available).		
Submitting Investigator:		
Diagnosis:		
Disease Status: Affected Unaffected Carrier At F	Risk	n.
Sample ID #:	/	
Date of Sample Collection: / /	/	
Age at Sample Collection (if fetal tissue, gestational age in v	NATE NATE	
dividual still living?  Yes No Don't know	NATIONAL INSTIT HUN (Submin I	Time
ed, age at death:	HUN	UTE OF GENERAL IAN GENETIC CEI NFORMED CONT
Male Female Ambiguous	MUST	NFORMED CON
- ie check all that apply):	use an Informed Conse	n organizatio
1 Indian/Alaskan Native 🗆 Asian 🗆 Native F	(Submitters associated with an MUST use an Informed Consen <u>Overview of the NIGMS Human</u> Institutes of Health, collects, store members of data from the Stores	all of approved bus
can American 🗌 White/Caucasian 🔲 Unkr	The NIGHT	on the elements in th
Hispanic Non-Hispanic Unknown	de-identifier Health countrie Cell	Senetic Cell Repositor
'talian, Nigerian, Mexican, German-Japanese, (	members of these individues, stores,	and diation a research
ical Information: (please attach copie:	The NIGMS Human Genetic Cell J Institutes of Health, collects, stores, members of data from inviduals, and make specimens availed inviduals, and clinical genetics iabiable for and in valuable genetics iabiable for antorials processes, and to develop new diano diseases.	with many kinds cell cu
n  Medical Records  Genetics C	processes, and to develop a that can be	research and the peop
Physician report Autopsy Reco		e used by samples to the
nimal Iformation:		
story of this condition?  Yes	I hereby consent	
	D Part -	of.
se describe the relationship c also briefly describe the phe	I hereby consent to the collection Peripheral whole blood Skin bioopy (up to ½ inc of drawing blood are medical rick skin bioops	theck all that app
	of drawing that the	(1-2 tubes: about e
_	skin biopsy are mine	
	I understand that the medical risks of of drawing blood are middle of a start of the skin biopsy are middle ocal pain, som possibility of infection.	t pain and slich spec

Note

astigator.

Dediatrician Deneticist Derin Phenotype (please describe dysmorphology, behaviors and other clin L MEDICAL SCIENCES (NIGMS) ENT FORM braces [] orthoti een by an Institutional Review Board (IRB) their own IRB that includes, at a minimum, : if available): th resource supported by the National i cultures, DNA samples and associated people. The purpose of this collection is to hing and as reference material for use in to the NIGMS Repository may result in results if available): e NIGMS Repository may result in better understand normal and disease ents, and new ways to prevent Sequencing): available) ply): pathology, etc) (please 24 milliliters, or 1-3 teaspoons) and slight possibility h, some bleeding, the e specimens are mini

smole C

of dr. skin b

Possib

Querries of the NIGNS Human Generic Cell Repositor

General Clinical Data Elements Form

[]Specialist:

The NIGALS Human Genetic Call of deside

Sample ID#:

Age at Diagnosis;

Diagnosed by:

e an Informed Consent Form annound overseen by an Institute all of the elements in this form, its

Diagnosis:

Age at onset of

## **Cellulose-based technology**

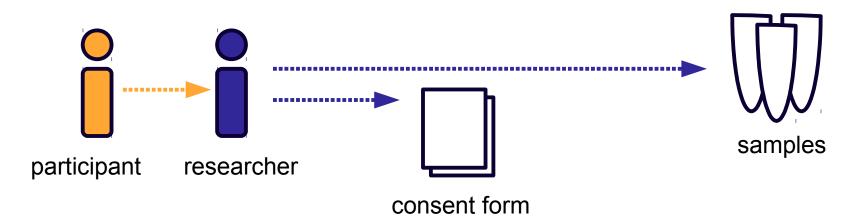
EKG, EEG, DIOPO,			
		eupational therapy	
nt (check all that apply):	physical therapy O	Ecope	
int (check all that apply) speech language therap	py		

that would identify vou, except as explained below. mation that would identify you, except as explained below. The Certificate cannot be used to resist a demand for information from personnel of the United States Government that is used for auditing or evaluation of Federally funded projects or for The Certificate cannot be used to resist a demand for information from personnel of the Unite States Government that is used for auditing or evaluation of Federally funded projects or for Form 1401-63 Rev B 071510 NIGMS Informed Consent Page 1 of 3

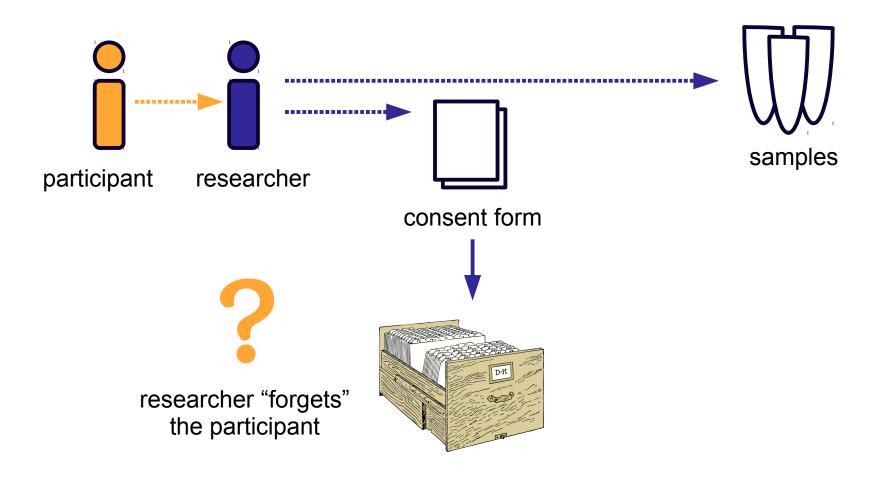
he phenotype.

0110.

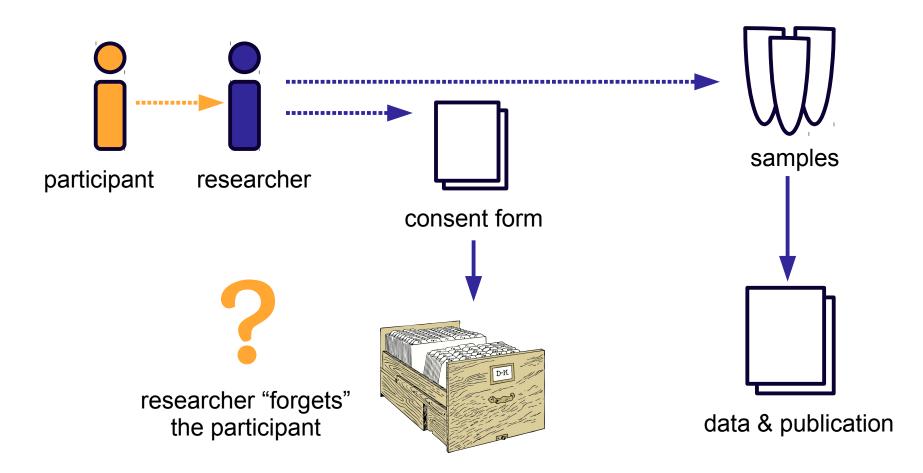








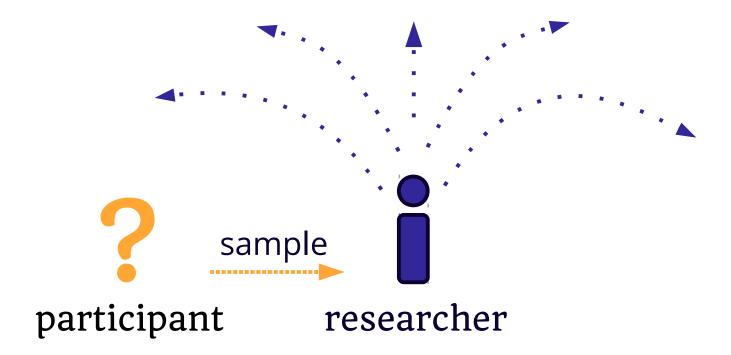




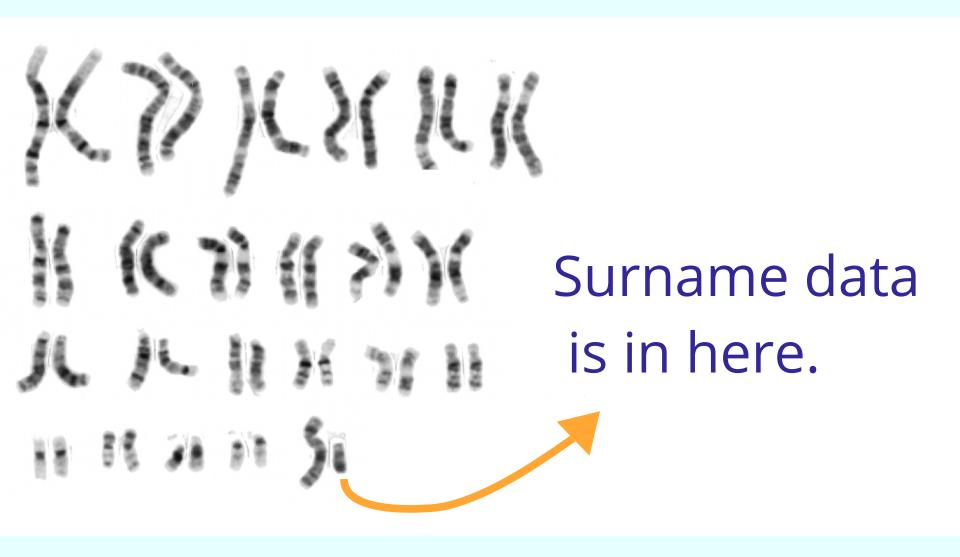
# Human data sharing?



#### public data sharing

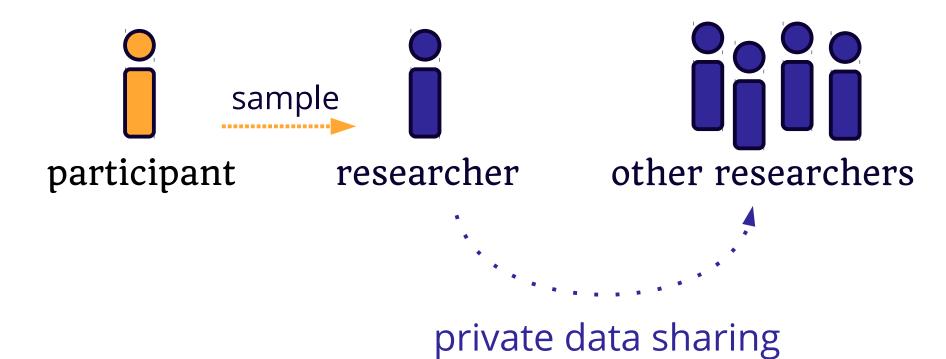


# Problem: The data is identifiable.





#### "Controlled access" sharing



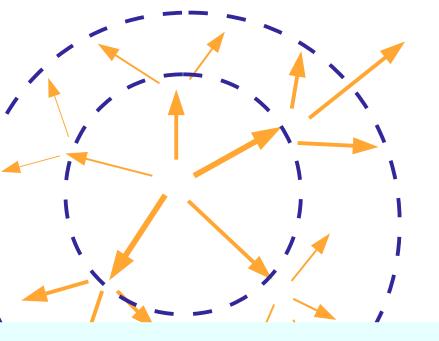
#### But sharing is limited

#### ...and data can leak.

SHINE CE

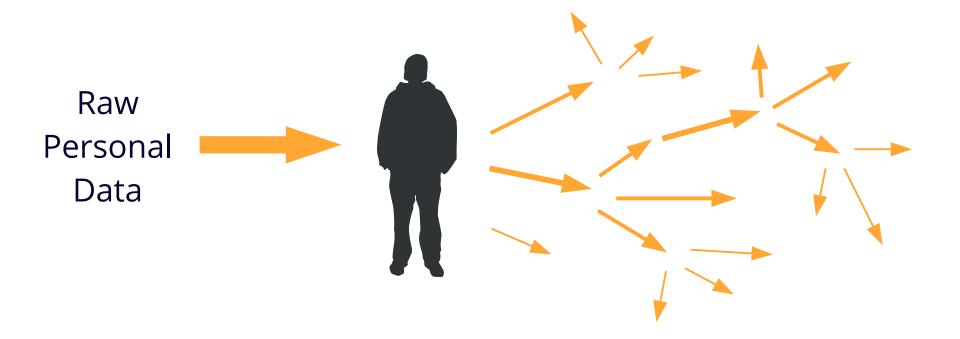
## Change the focus!

Focus less on creating
barriers to sharing...
(which may leak anyway)

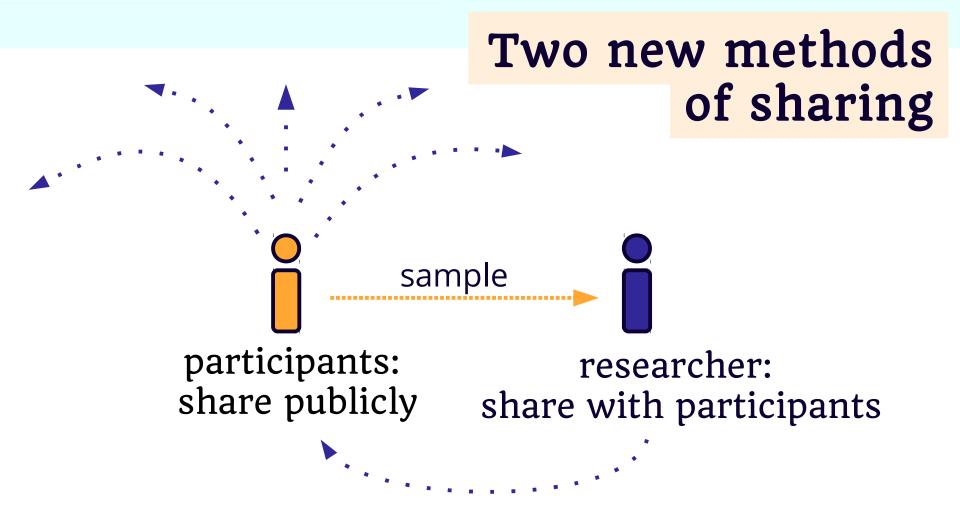


... instead, create new & better types of sharing.

#### Participant-mediated data sharing

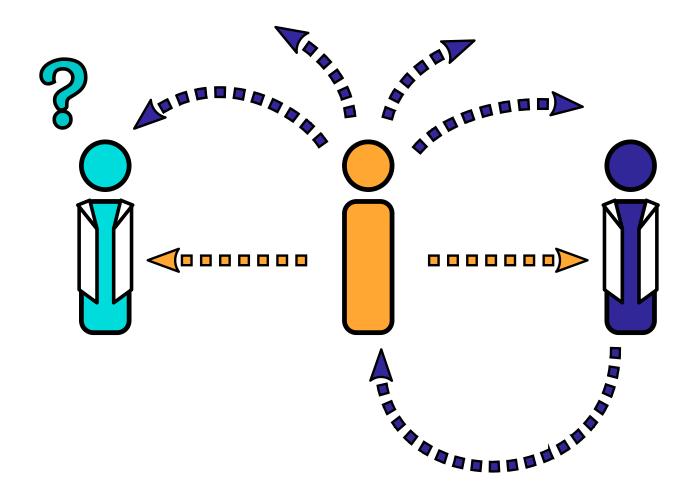


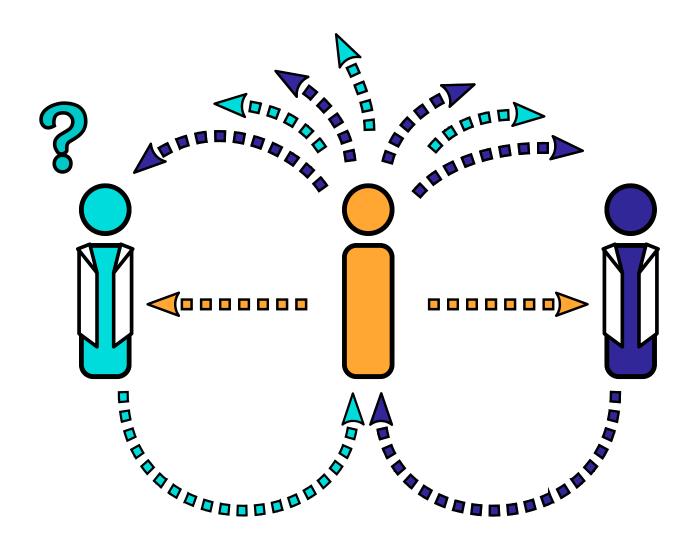
#### **Open Humans**



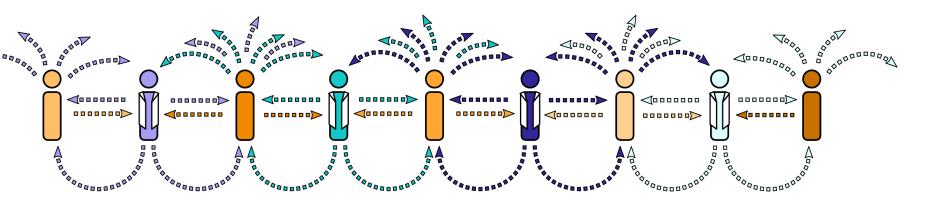
#### Participants can accumulate data







#### Propagate new sharing practices





#### a transformation

#### in research



# www.openhumans.org { mailing list news & information hiring



PersonalGenomes.org









#### **PGP Harvard Participants**

George Church Tom Clegg Jason Bobe Pete Estep Alex Hoekstra

Jeantine Lunshof Ward Vandewege Alexander Wait Zaranek





Robert Wood Johnson Foundation