

# Open Data, **Open Humans**

Madeleine Price Ball

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



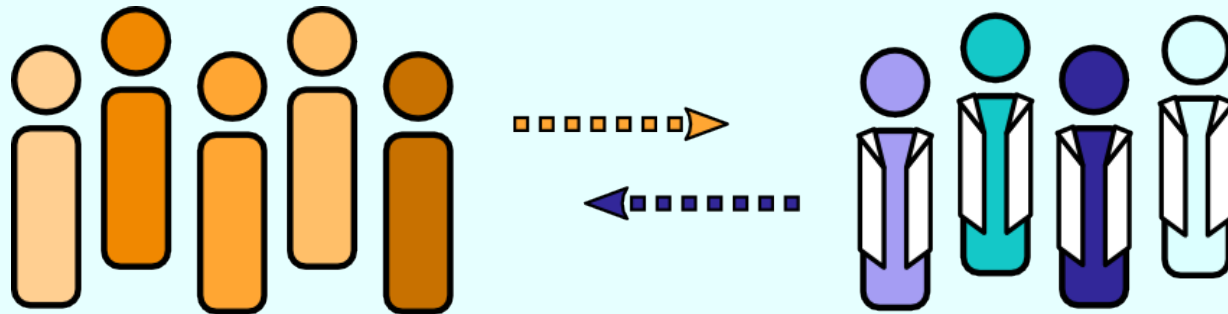
Personal  
Genome  
Project

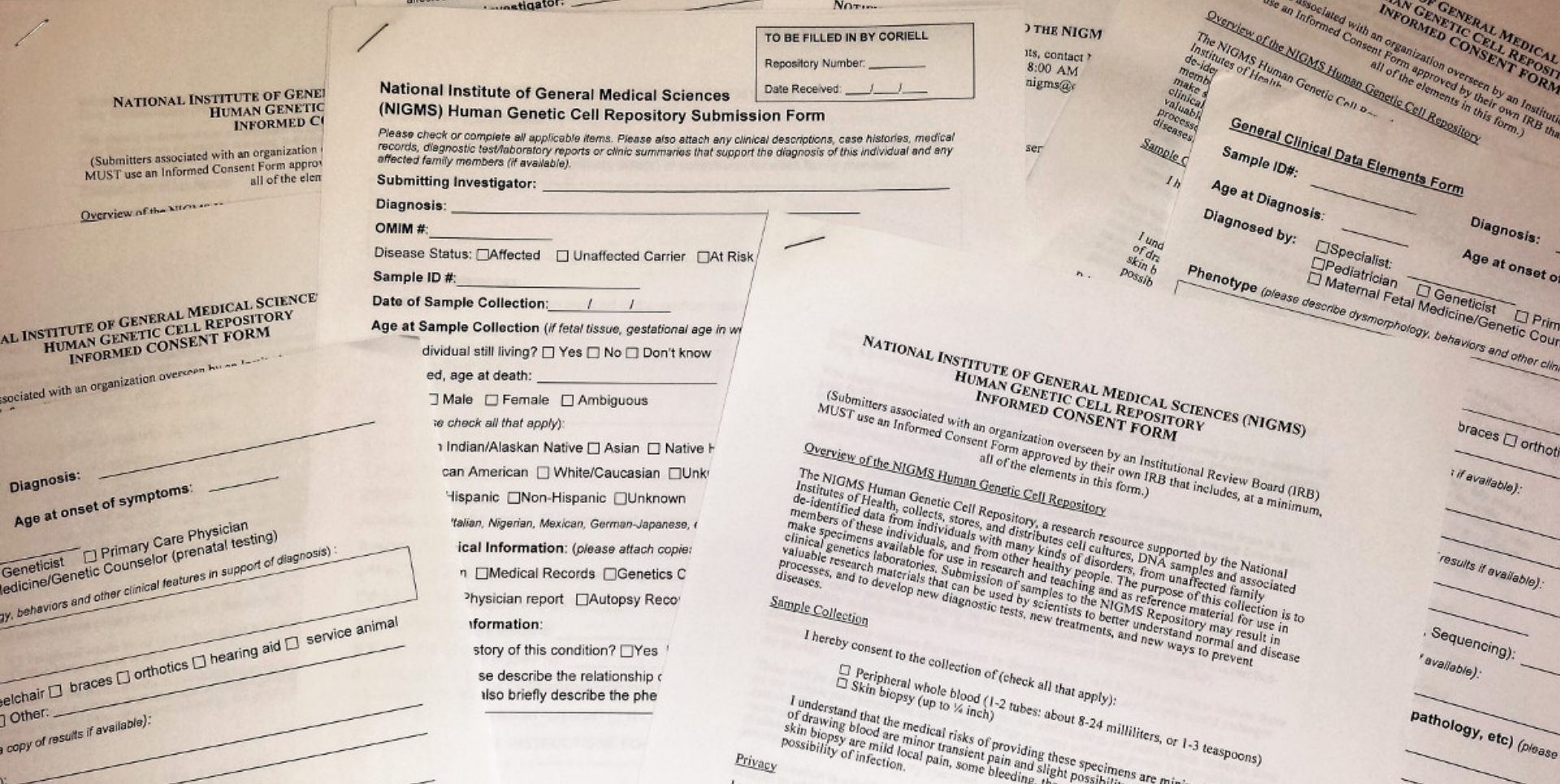
# Human Biology → Bits → ?

ACCTT	ACCTT	ACAG
TGGAT	ACCTT	TAGTG
CAGAG	ACCTT	AGCTC
CTAA	ACCTT	AGCTC
CACCTG	ACCTT	AGCTC
10010001011110	0010100	011010001110
0111100101011110	00101	0111010001110
10111110101110101	010101111010	011110011110101
01011011100010101	000101010010001	011111110001011
01111111101010100	10110101000111010	0111001001011110
1010100100000101	01010111000001010	0011101001000010
01011100111011	01011111110010111	10100101111010011
01011110100001	01110100001010100	11111110010100010
111000101110001	0010001111011100	00111100GCACTGT
AGCTGCACGTGG	0AGAACTTCAGGCTC	CAACGTGCTGGTCTG
CTGGC CATCA	AAAGAATTCACCCC	TGCAGGCTGCCTATCA
GTGG TGGCT	CTAATGCCCTGGCCC	TACTACTAAGCTCG
TGCTG TCCA	TAGGTT TTTGTT	TCCAACTACT/
CGAT ATTA	TGAGC TTGGA	AATAA AAAAC
CATT TC100	0110 10001	11010 1010/
0101 01000	11110 11101	01111/ 0010
01111 1110	01000 00010	00101 1011
1110 1011	10000 01011	11110 1101
01111 0101	01010 10110	00011 00100
111 1111	0110 11101	111100 01010
111 11110	10000 0011	111000 011110
011 100	00100 0001	00110 1000111

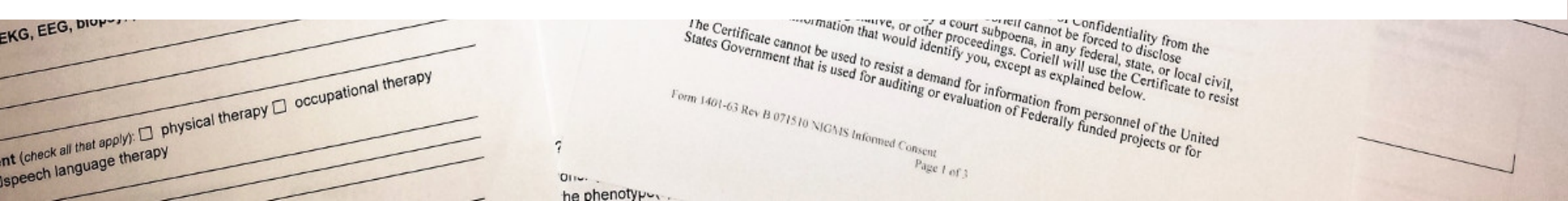
# Human subjects research

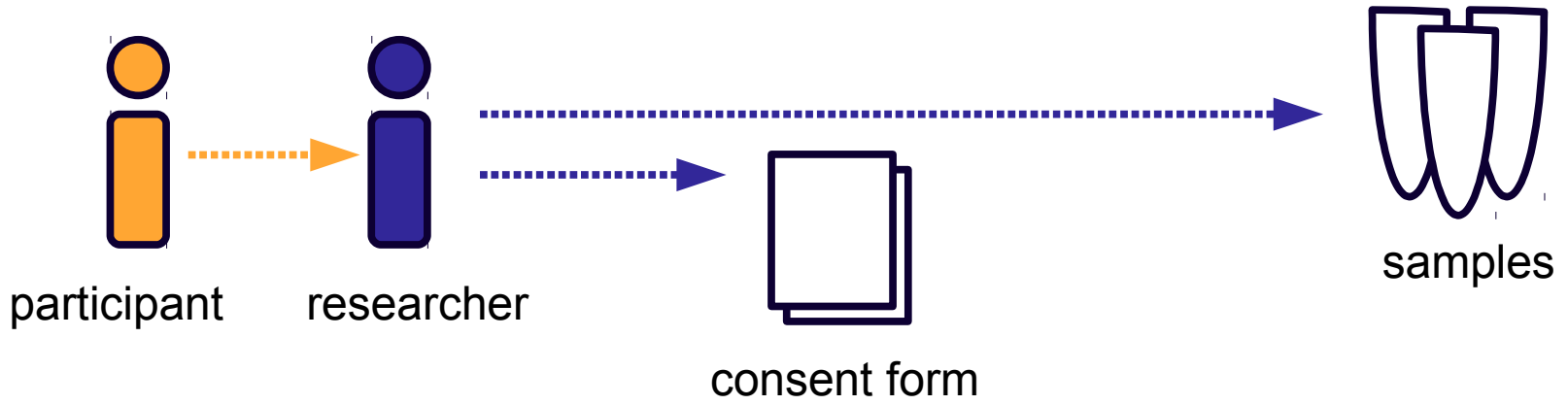
A very brief review of data sharing.

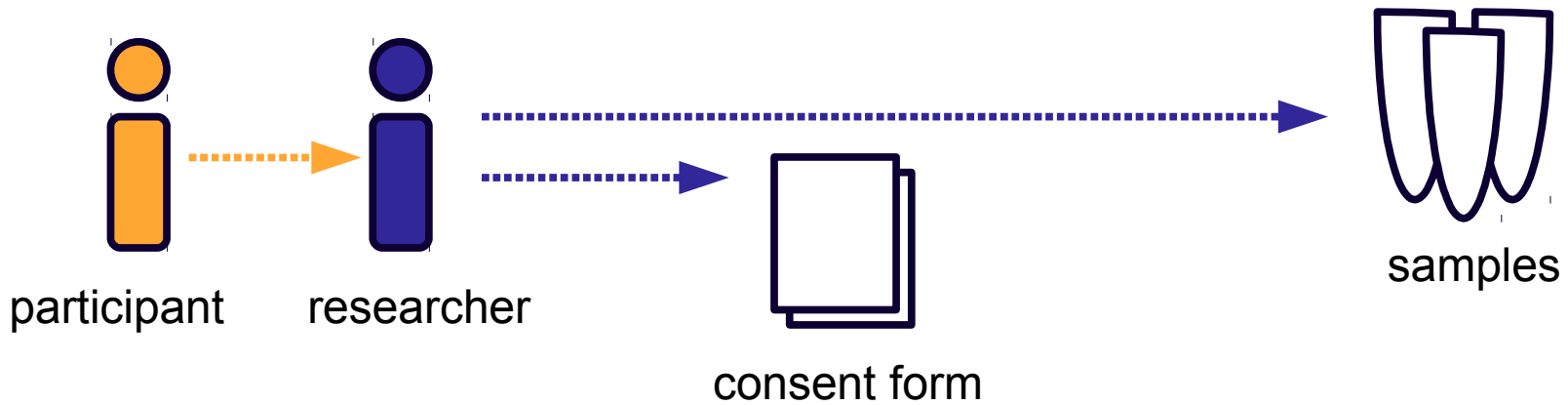




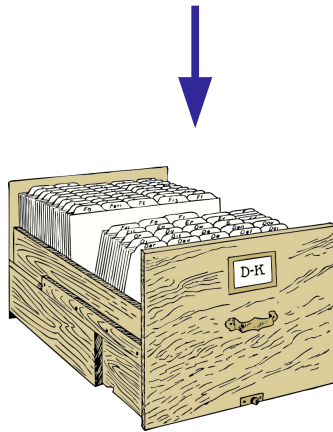
# Cellulose-based technology

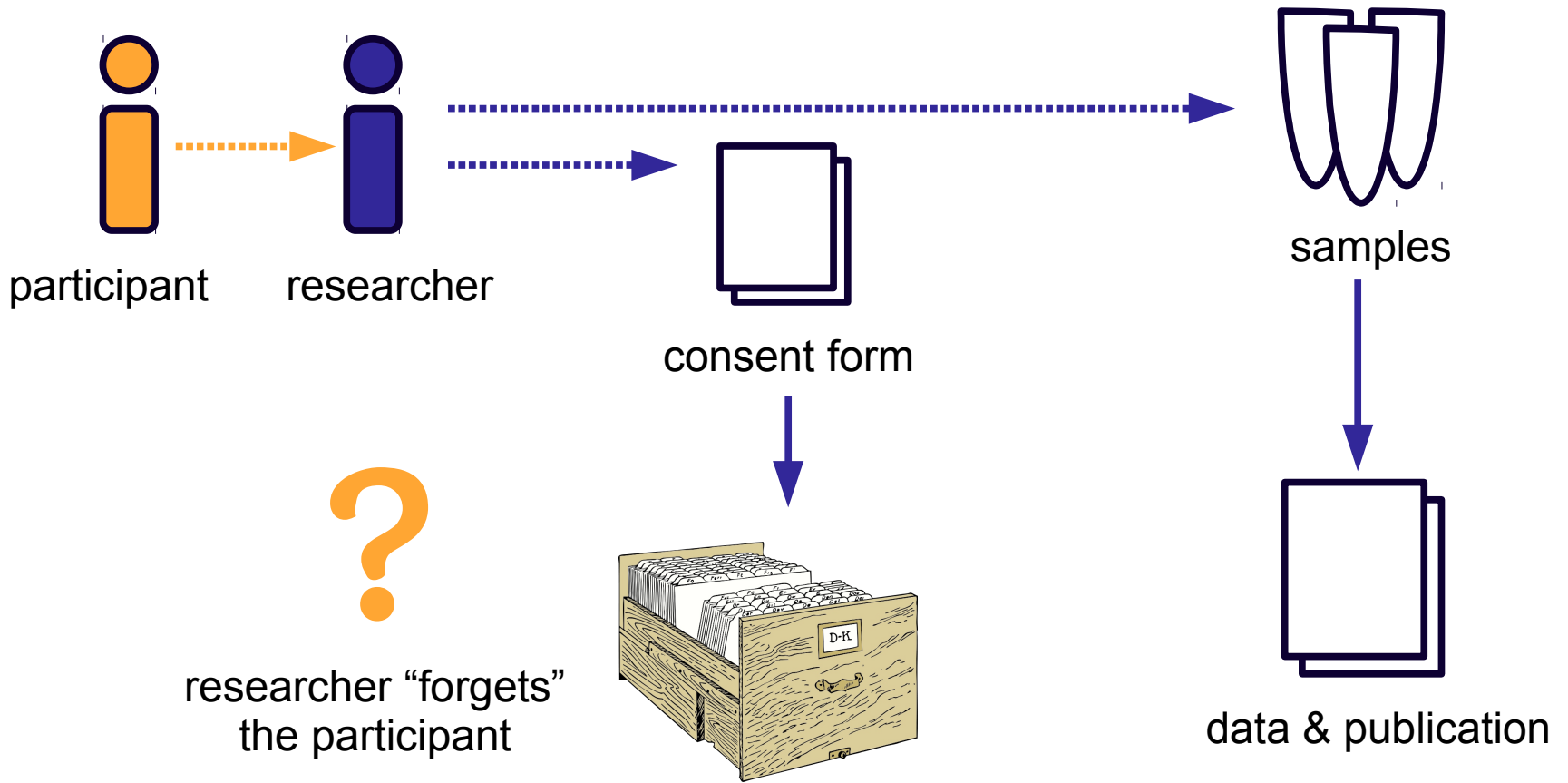






researcher "forgets"  
the participant



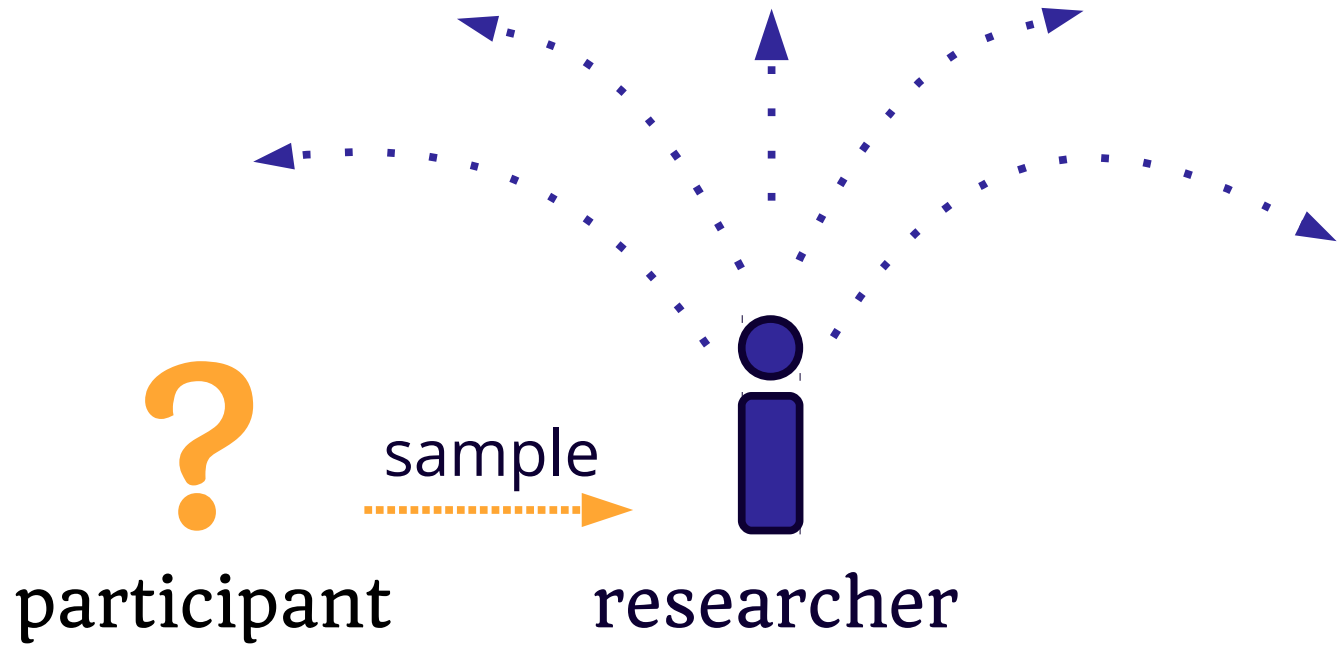


# Human data sharing?





# public data sharing



**WARIN**

**Problem:**

**The data is identifiable.**

**NO**



Surname data  
is in here.

AERATED

# BAND-AID

TRADE MARK

PLASTIC FIRST AID DRESSINGS

*Johnson & Johnson*

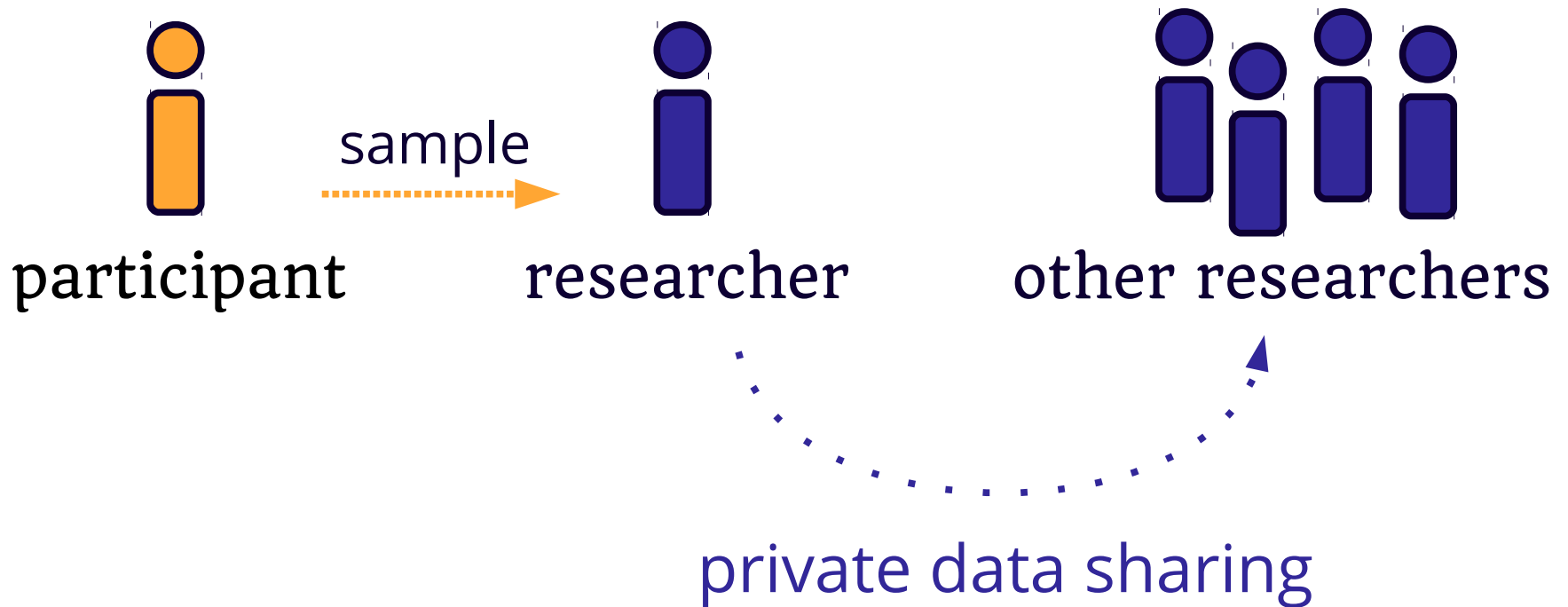
AERATED WATERPROOF



PLASTIC SUPER-STICK



# “Controlled access” sharing





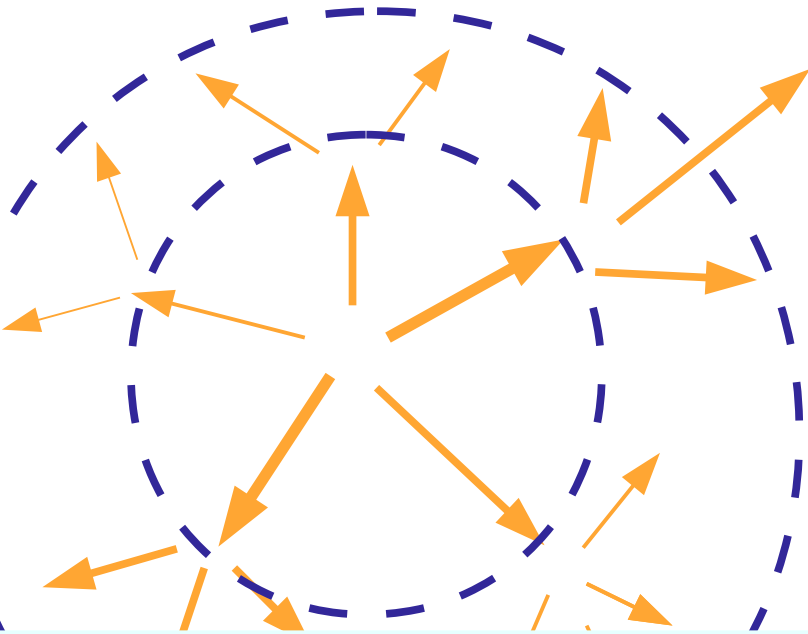
**But sharing is limited**



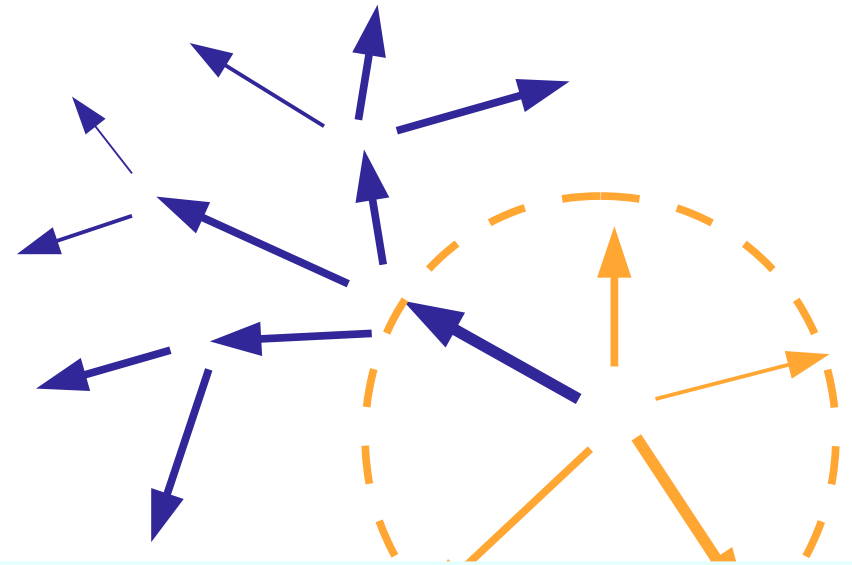
**...and data can leak.**

# Change the focus!

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

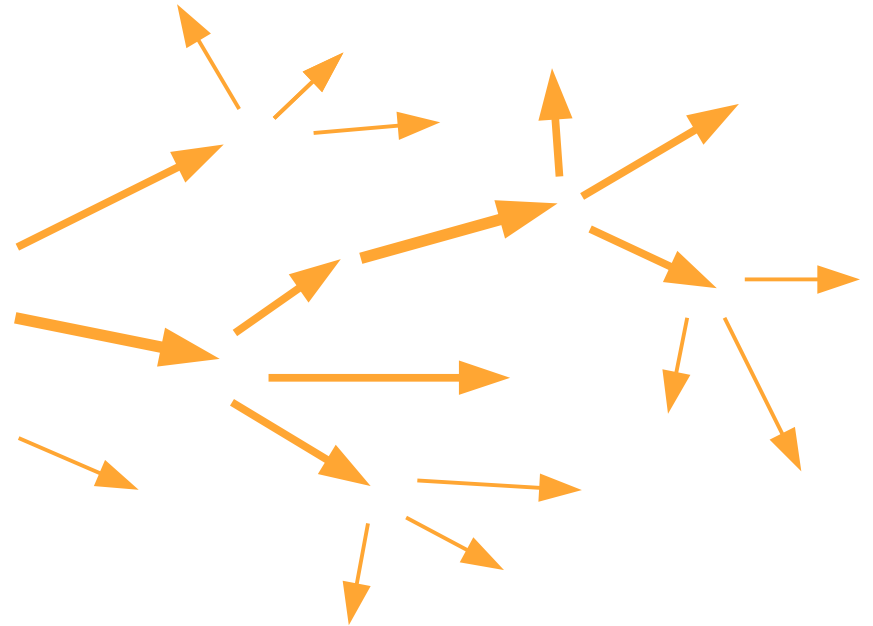


... instead, create **new & better** types of sharing.



# Participant-mediated data sharing

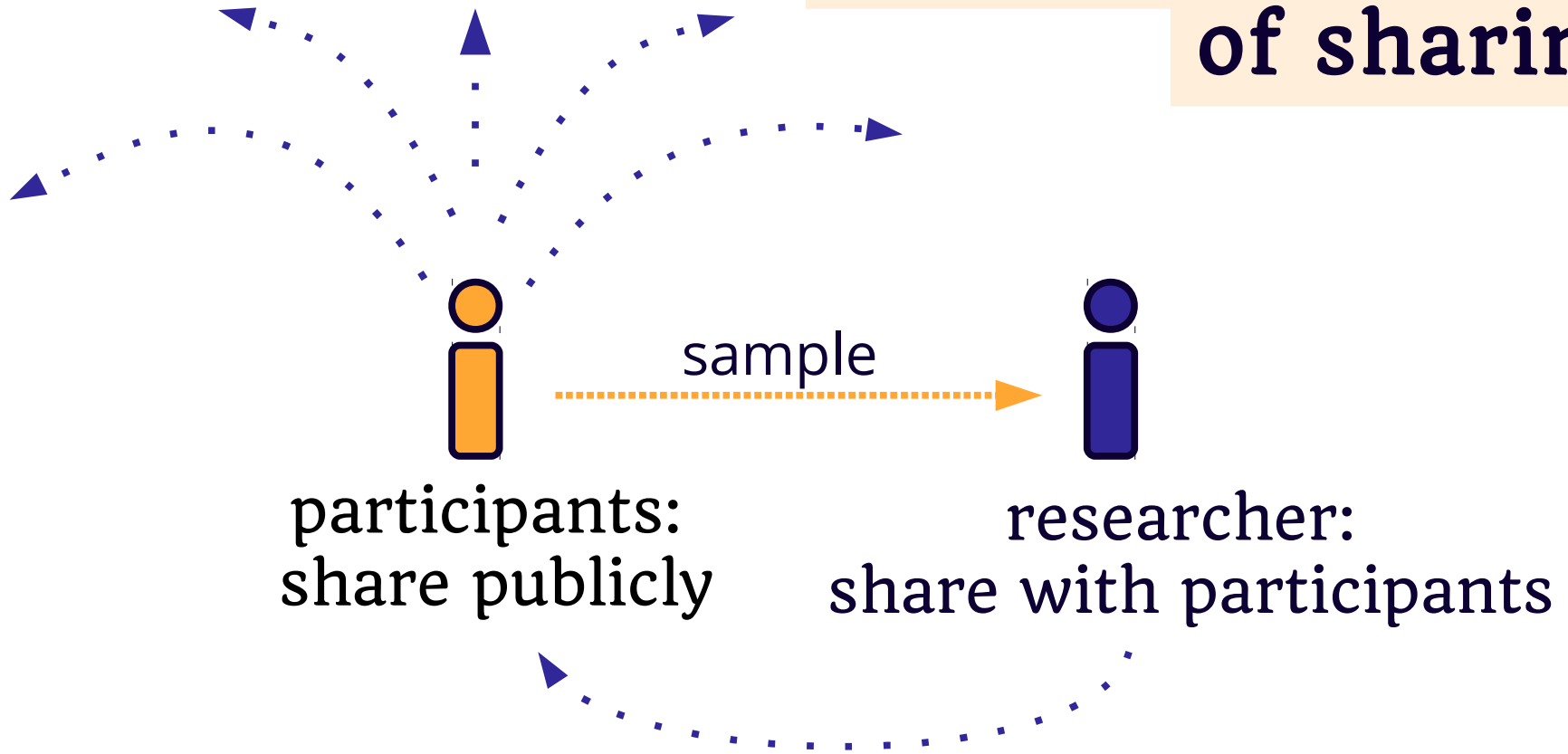
Raw  
Personal  
Data





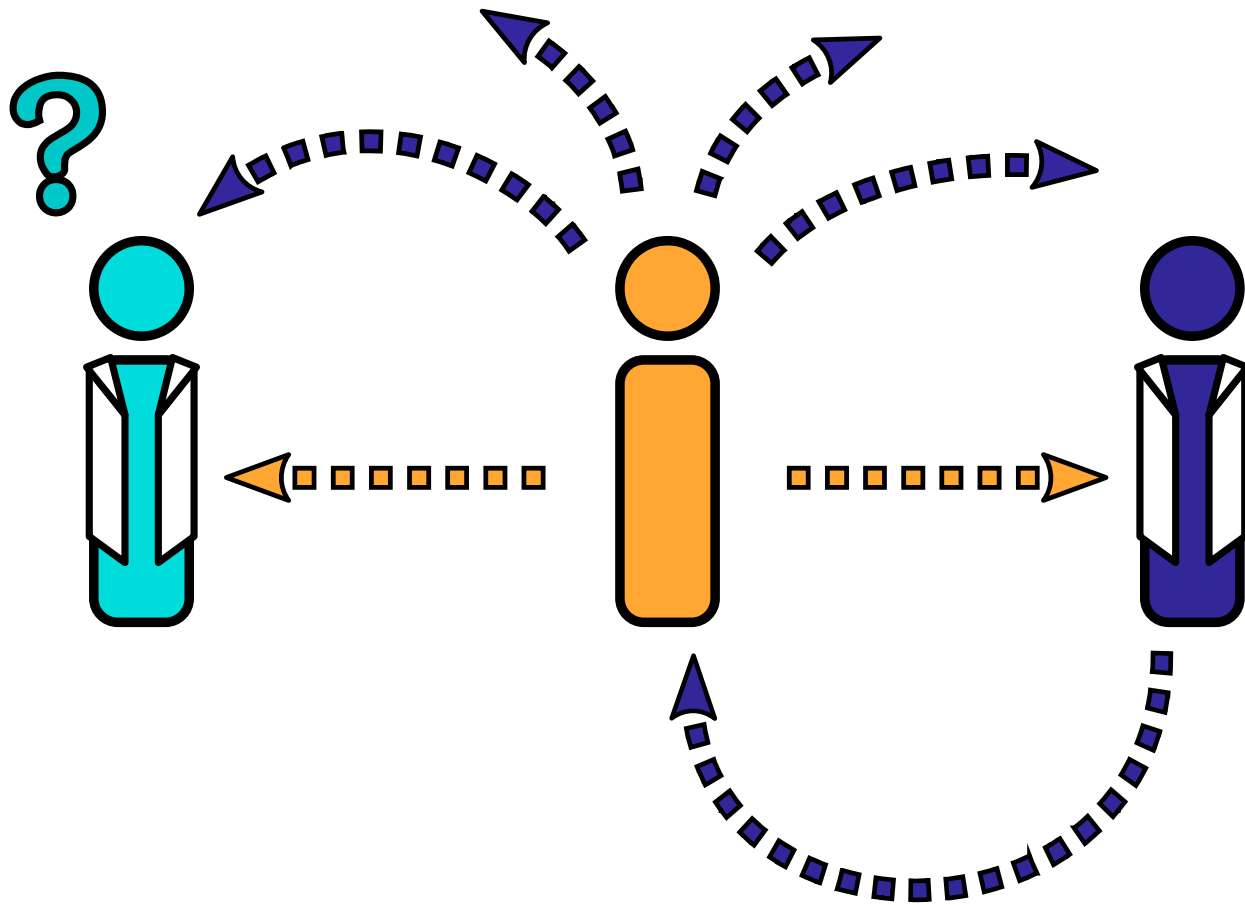
# Open Humans

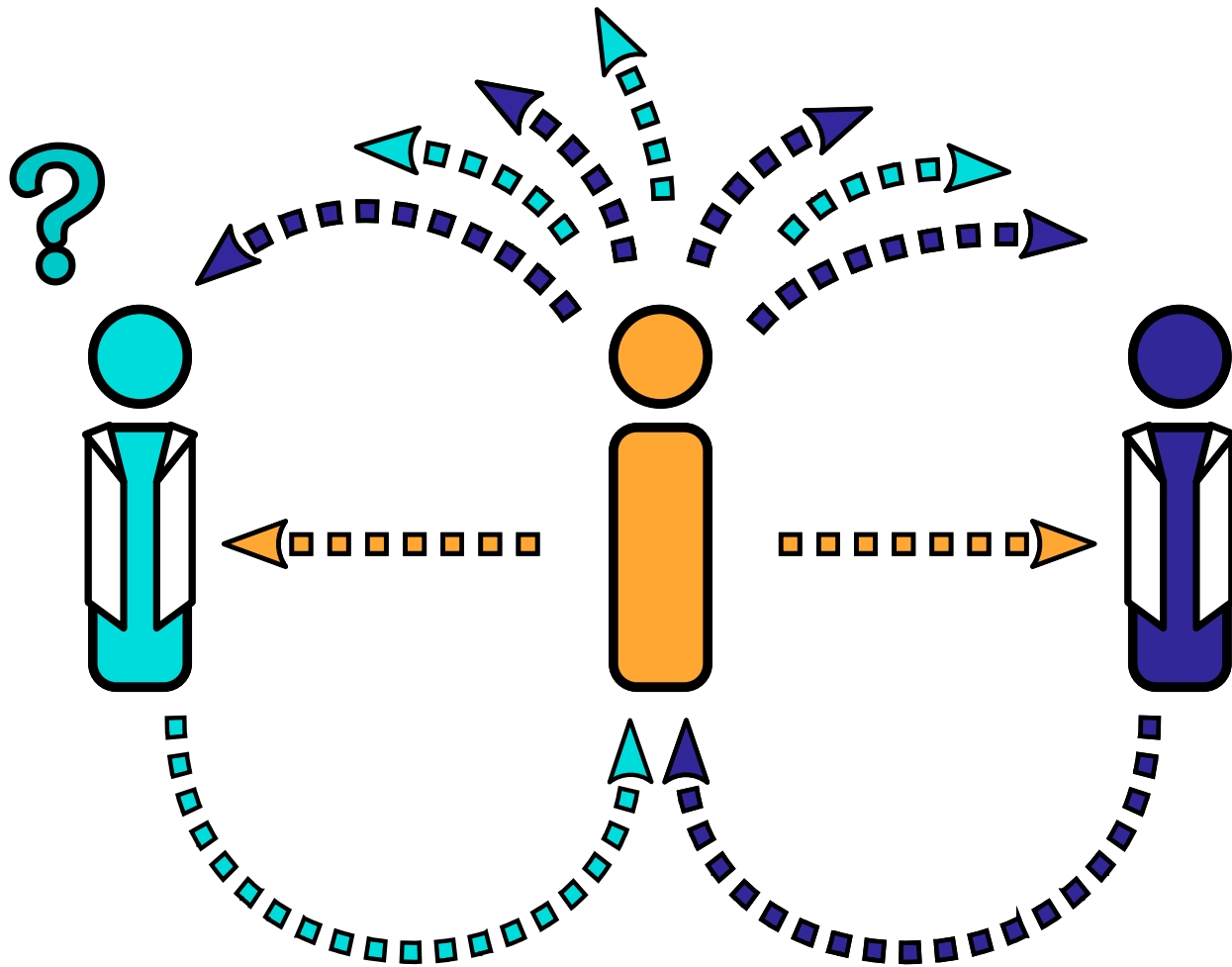
Two new methods  
of sharing



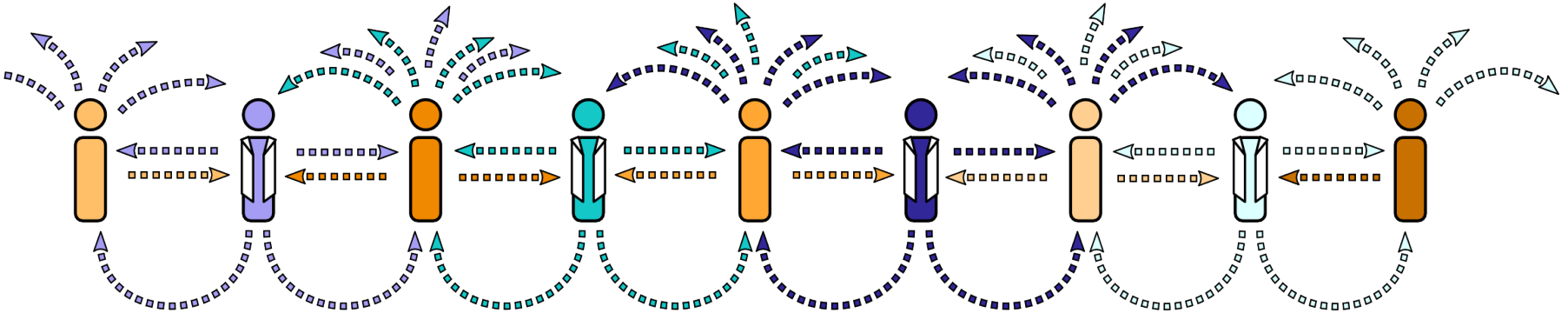
**Participants can accumulate data**







# Propagate new sharing practices



**Spark**

**a transformation**

**in research**





# Open Humans

Researchers and participants, working together to create open data

[www.openhumans.org](http://www.openhumans.org)

{ mailing list  
news & information  
hiring



PersonalGenomes.org



Personal  
Genome  
Project

**GET** Genomes  
Environments  
Traits  
**CONFERENCE**



**flu** near you   
do you have it in you?



## PGP Harvard Participants

George Church  
Jason Bobe  
Tom Clegg  
Pete Estep  
Alex Hoekstra

Jeantine Lunshof  
Ward Vandewege  
Alexander Wait Zaranek

 **Knight Foundation**  
*Informed & Engaged Communities*

  
Robert Wood Johnson Foundation