Medical Data Privacy and Ethics in the Age of Artificial Intelligence

Lecture 13: Data Privacy Protection Models: K-anonymization and others

Zhiyu Wan, PhD (wanzhy@shanghaitech.edu.cn)
Assistant Professor of Biomedical Engineering
ShanghaiTech University

April 18, 2025

Learning Objectives of This Lecture

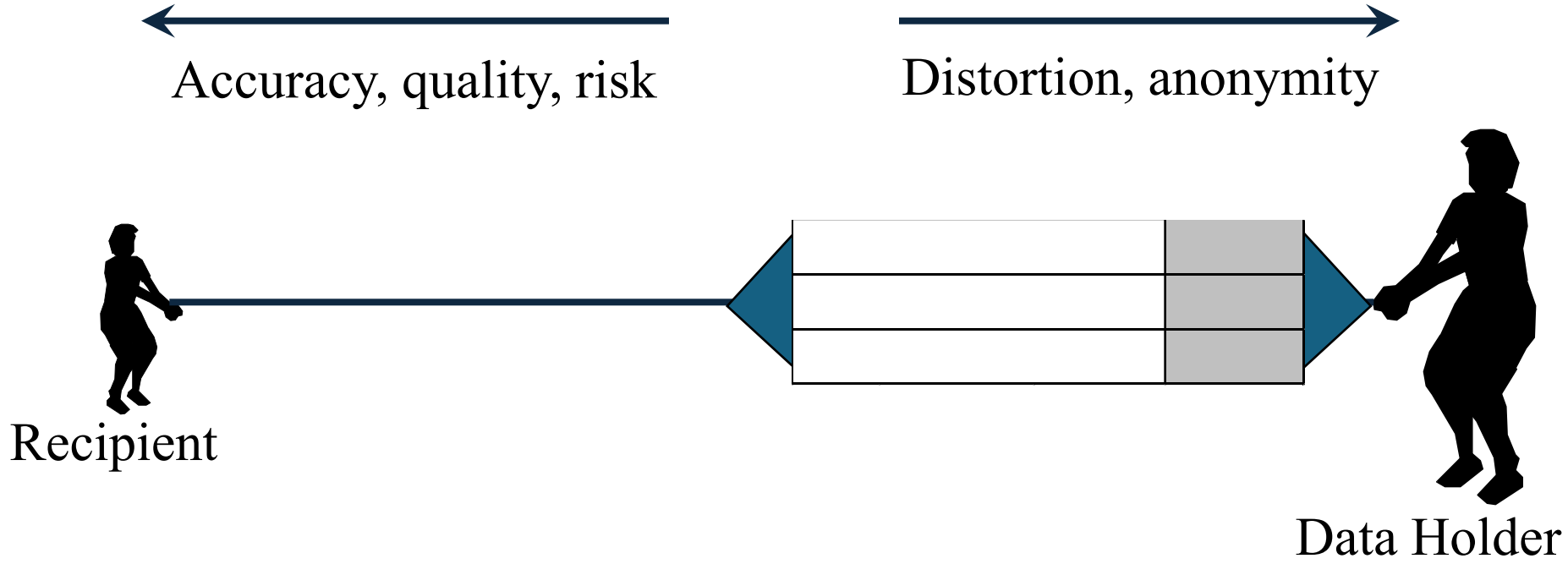
- Understand the concept of K-anonymization
- Know operations to realize K-anonymization
  - Generalization
  - Suppression

Today's Lecture

- Data Sharing Beliefs and Dogmas
- Protection Models
- Models vs. Methods

What Can We Do?

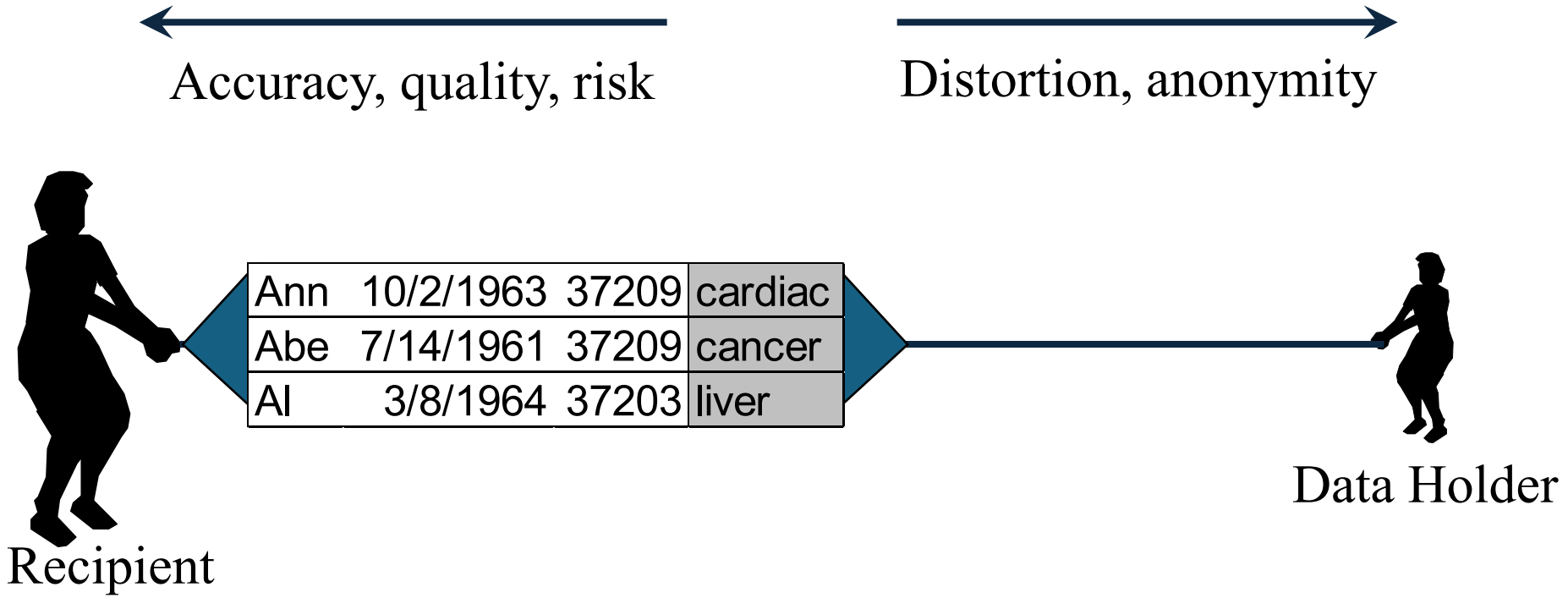
- Some say, “You Can’t Release Any Data”



Adapted from Dr. Malin’s slides

What Can We Do?

- Others say, “Privacy is Dead, Get Over It”



Privacy Is Dead And Most People Really Don't Care

Neil Sahota Contributor
AI
Neil Sahota is a globally sought after speaker and business advisor.

Follow

Listen to article 5 minutes

Are you guarding your data privacy? RAWPIXEL LTD.

Have you read the terms and conditions to use Facebook? Your smart phone? Most people have not, and probably with good reason. They're hundreds, if not, thousands of pages long. In fact, even contract lawyers with thirty years of

a crystal ball.

Categories

- U.S. news
- World news
- Politics
- Business
- Sports
- Entertainment
- Health
- Tech & science
- Space
- Science
- Tech and gadgets
- Games
- Wireless
- Innovation
- Security
- Travel
- Local news
- Weather

Browse

Privacy is dead on Facebook. Get over it.

Cool kids don't care about privacy, claim CEOs. So, neither should you.

By Helen A.S. Popkin
msnbc.com
updated 7:56 a.m. CT, Wed., Jan. 13, 2010

Once upon a time at Facebook, or so the story from an anonymous Facebook employee goes, there was a general password employees could use to access Facebook accounts. For kicks and giggles, some Facebook employees, including the one recently interviewed on the Rumpus Web site, did just that.

Two Facebook employees got fired, says Anonymous Facebook Employee, for manipulating user profile information. Others, such as Anonymous Facebook Employee, just peeked.

Is this story even true? Regarding the veracity of Anonymous Facebook Employee's claims, a company spokesperson stated via e-mail: "This piece contains the kind of inaccuracies and misrepresentations you would expect from something sourced 'anonymously,' and we'll leave it at that."

Most popular

Most viewed | Top rated | Most e

- The 18 best female celebrity tattoos
- UPDATED Jenny Sanford: Husband asked
- affair advice
- UPDATED Obama officials step up pressur
- Toyota
- Michelle Obama: What you see is the rea
- Mom, newborn triplets found dead on fl
- Most viewed on msnbc.com

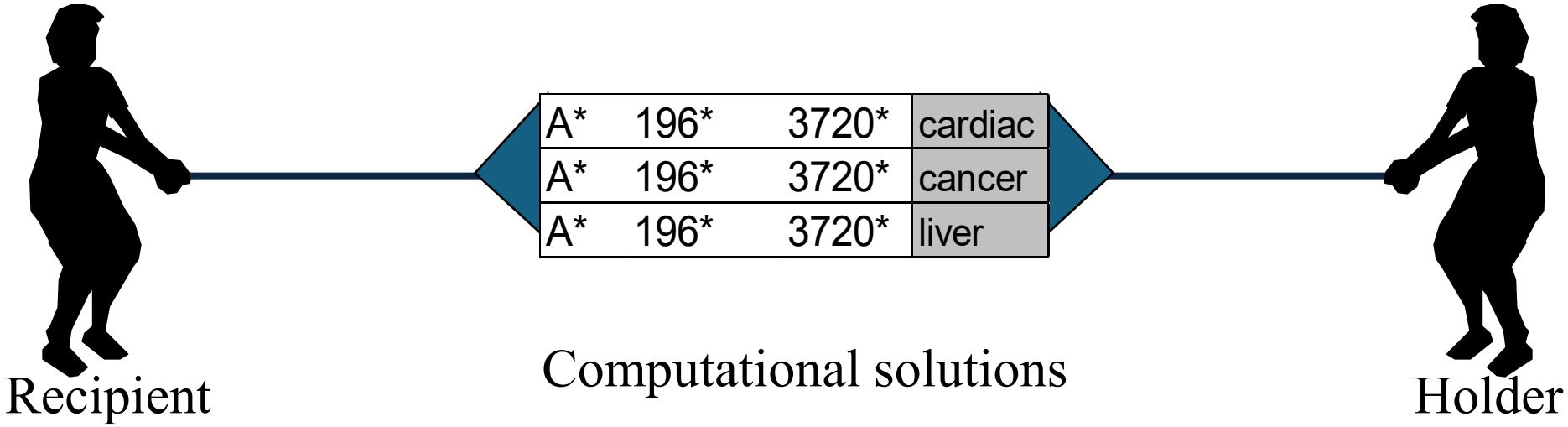
INTERACTIVE

Launch

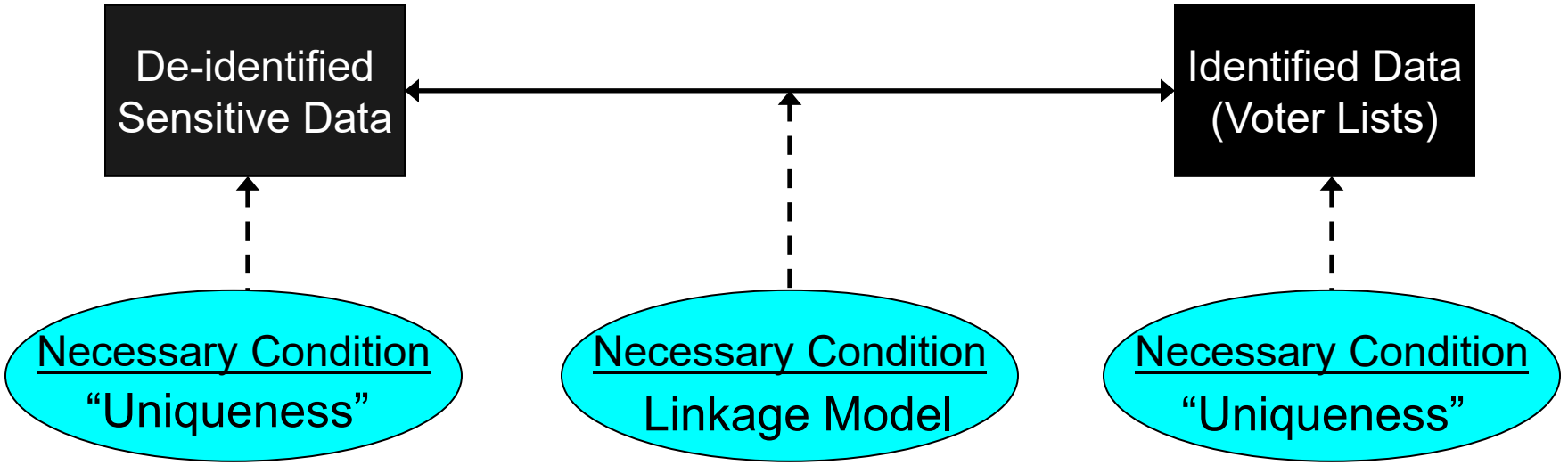
Top 9 Interne
memes of 20
From Crasher S
and Kanye Wes
Balloon Boy, th
the running gag
made the year's
life worth living

What Can We Do?

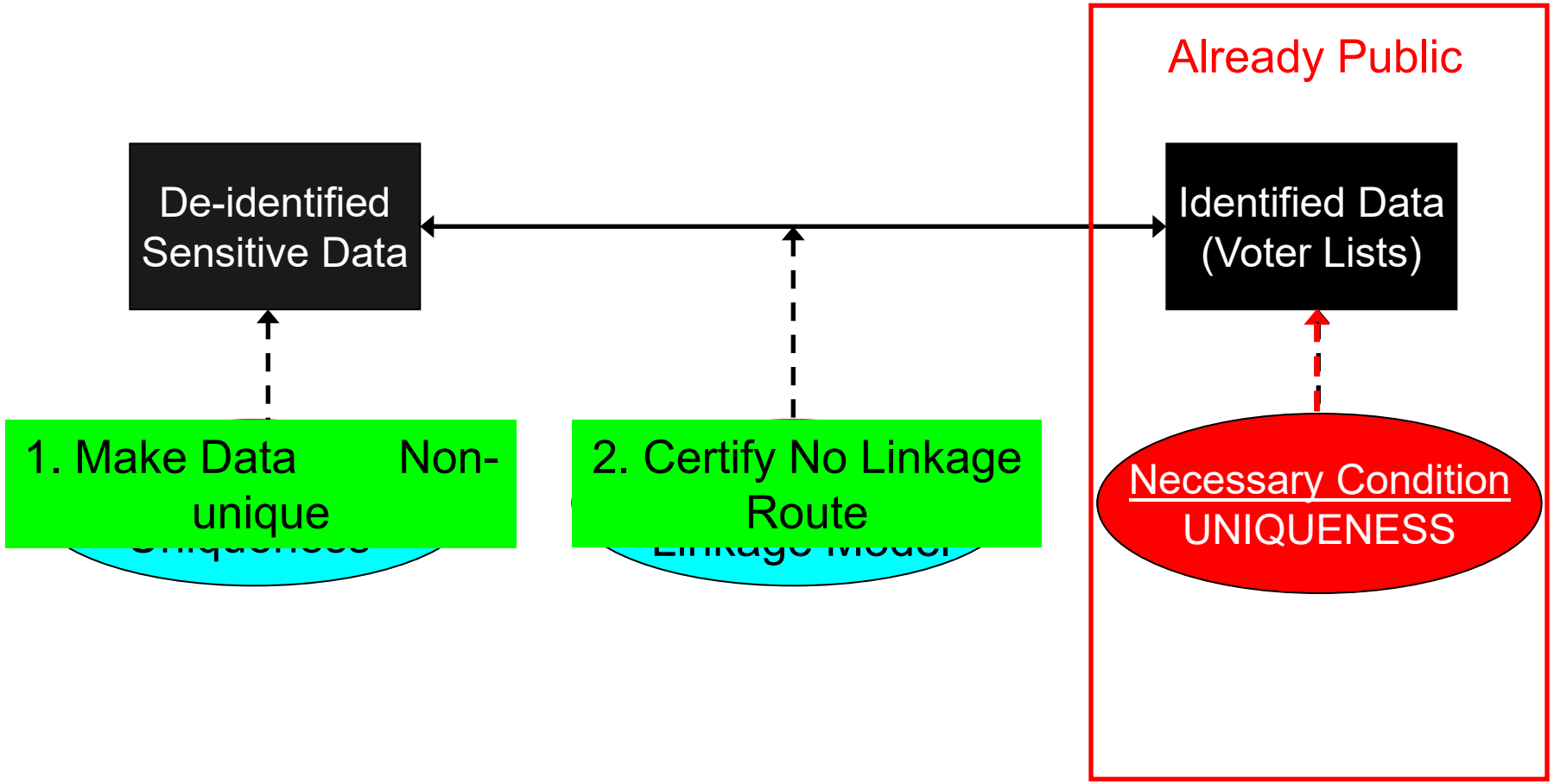
- We say, “Share Data While Providing Guarantees of Anonymity”



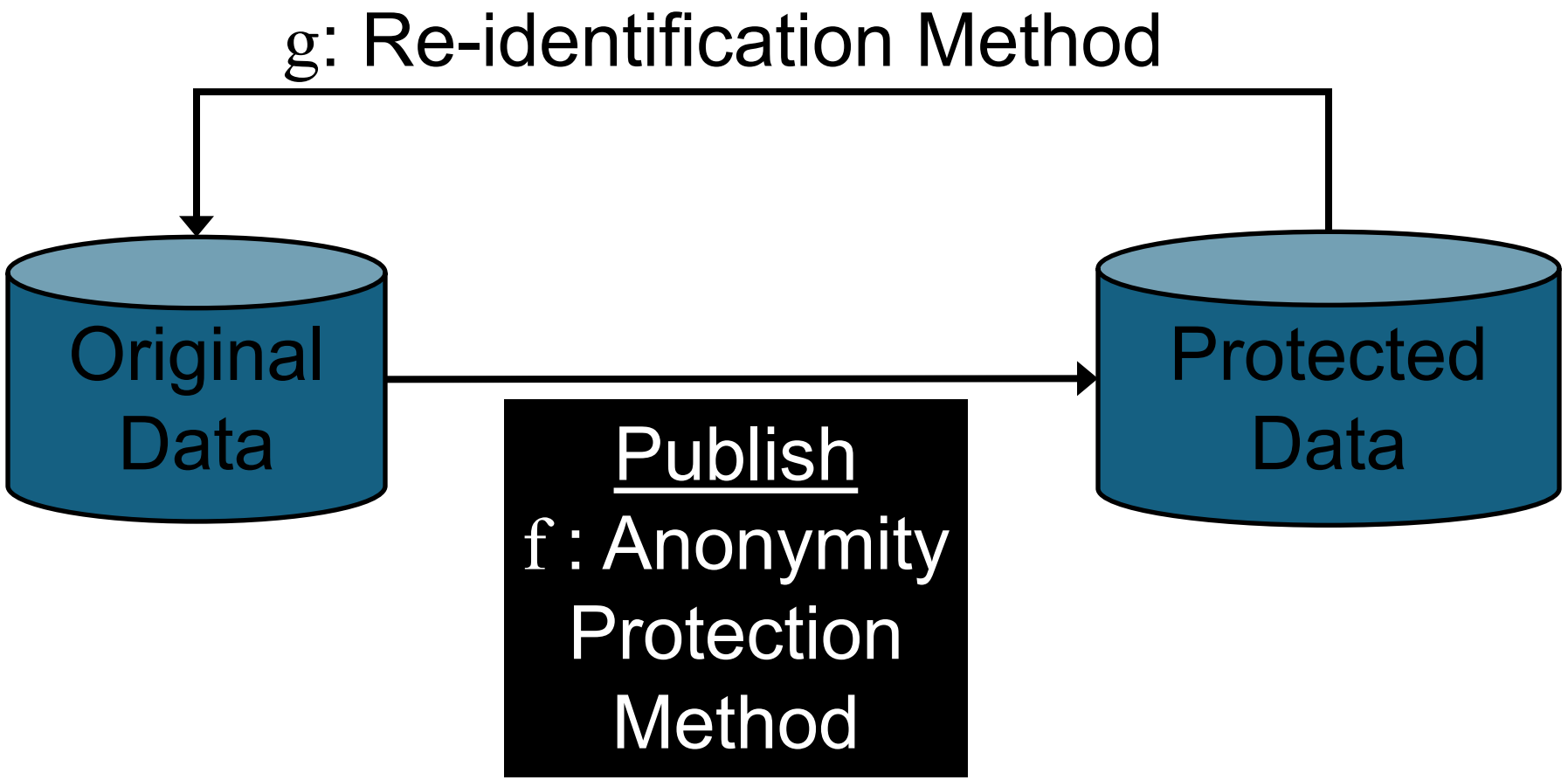
Central Dogma of Re-identification



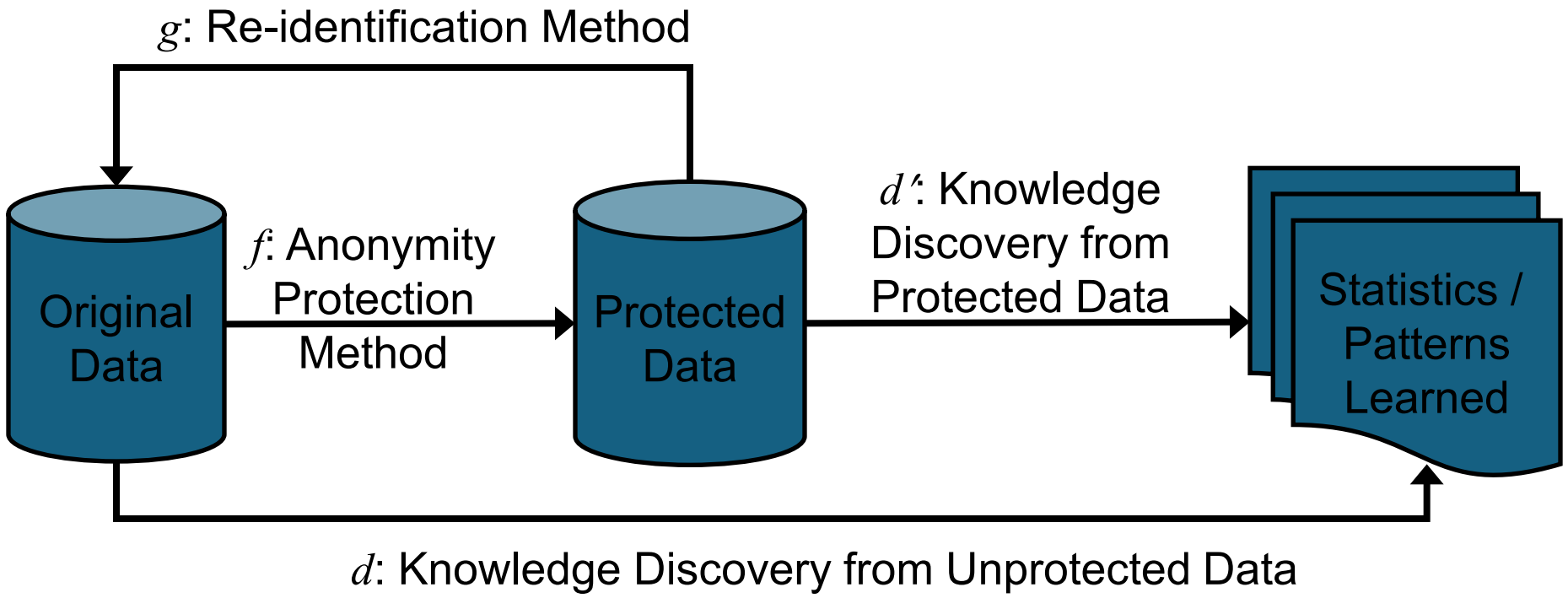
Central Dogma of Anonymity



Formal Protection



Don’t Forget the Bigger Picture



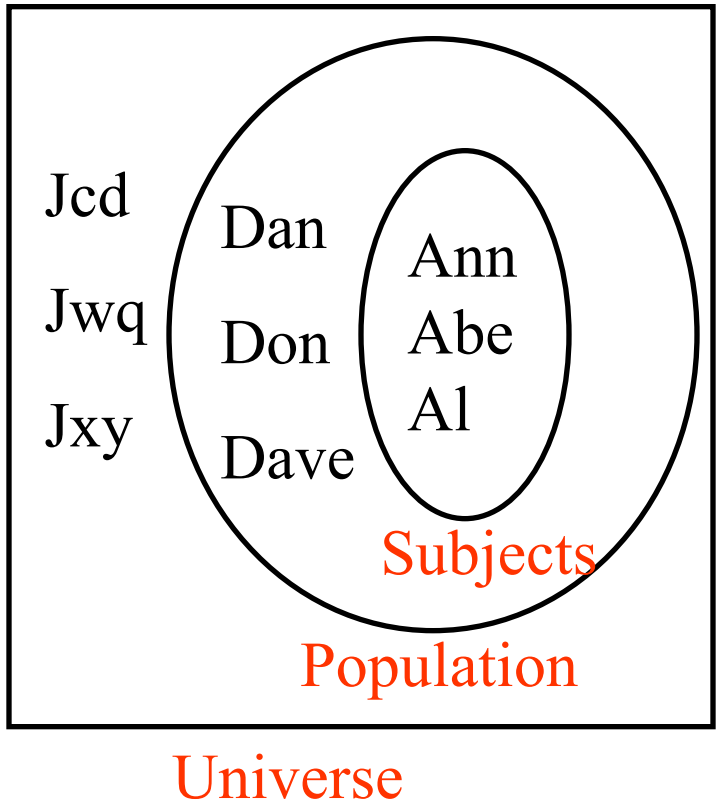
1. Minimize the inversion of f or discovery of g.

2. Minimize difference between d and d'.

Today's Lecture

- Data Sharing Beliefs and Dogmas
- Protection Models
  - Subjects vs. Populations
  - Null-Map
  - Wrong-Map
  - Permutations
  - K-Map
  - K-Anonymity
  - K-Ambiguity
  - K-Unlinkability
- Models vs. Methods

Formal Models of Anonymity

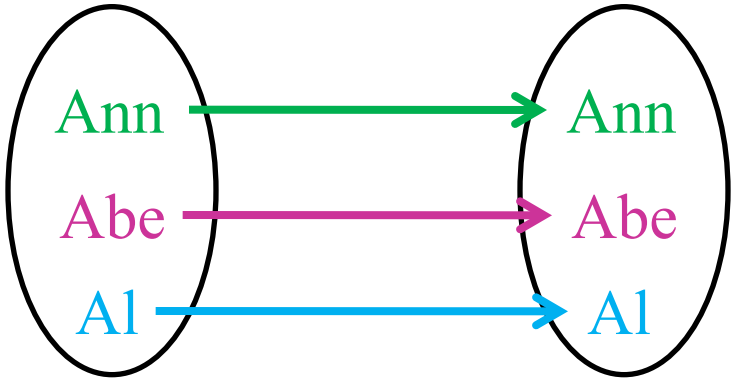


Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

L. Sweeney. Computational Disclosure Control. Ph.D. Thesis Draft, MIT. 2001. http://www-swiss.ai.mit.edu/6095/articles/privacy/sweeney-thesis-draft.pdf

Formal Models of Anonymity



Original Subjects

Original Subjects

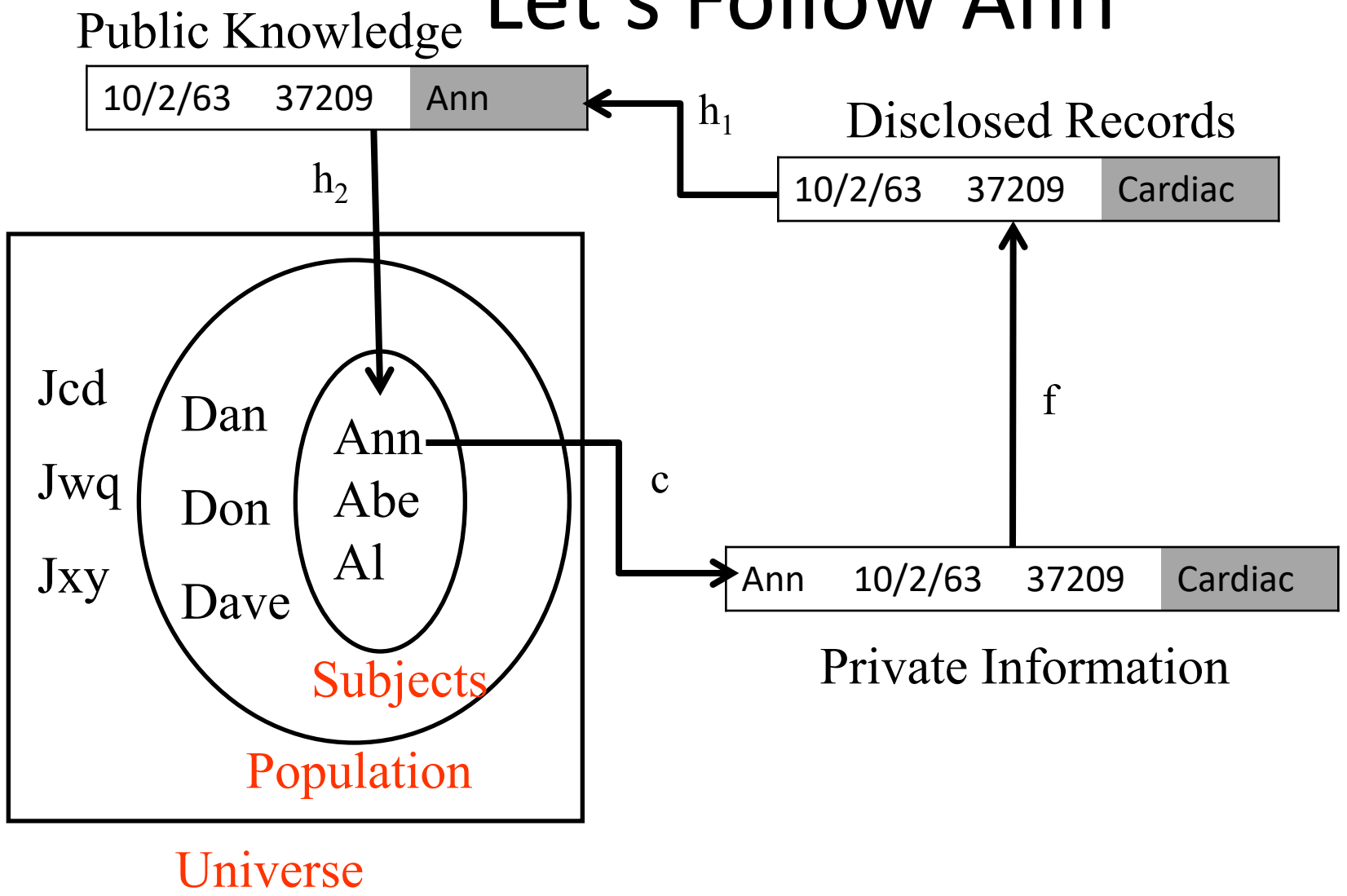
Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

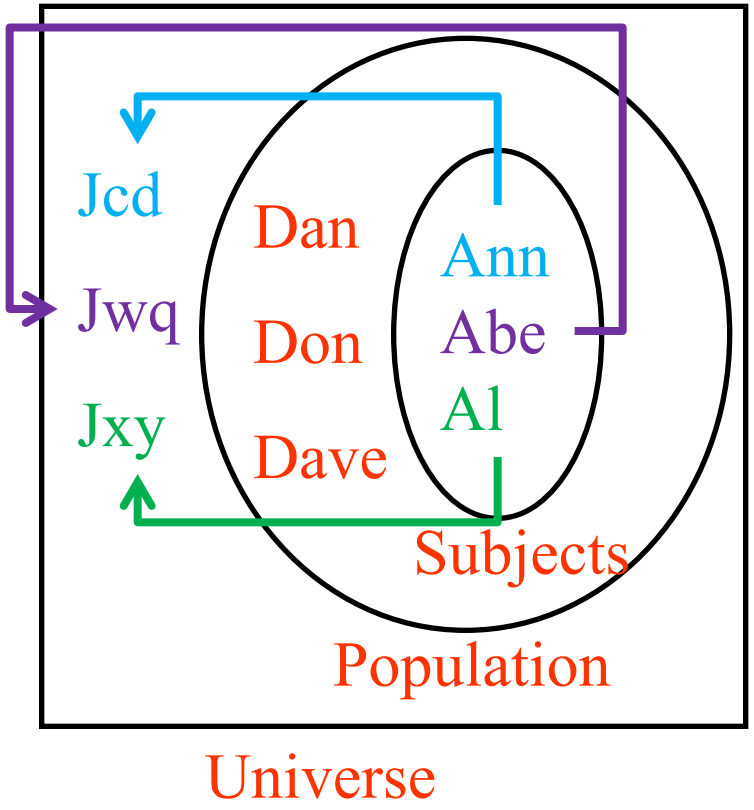
Original mapping.

L. Sweeney. Computational Disclosure Control. Ph.D. Thesis Draft, MIT. 2001. http://www-swiss.ai.mit.edu/6095/articles/privacy/sweeney-thesis-draft.pdf

Let's Follow Ann



Null-Map



Jcd	Cardiac
Jwq	Cancer
Jwy	Liver

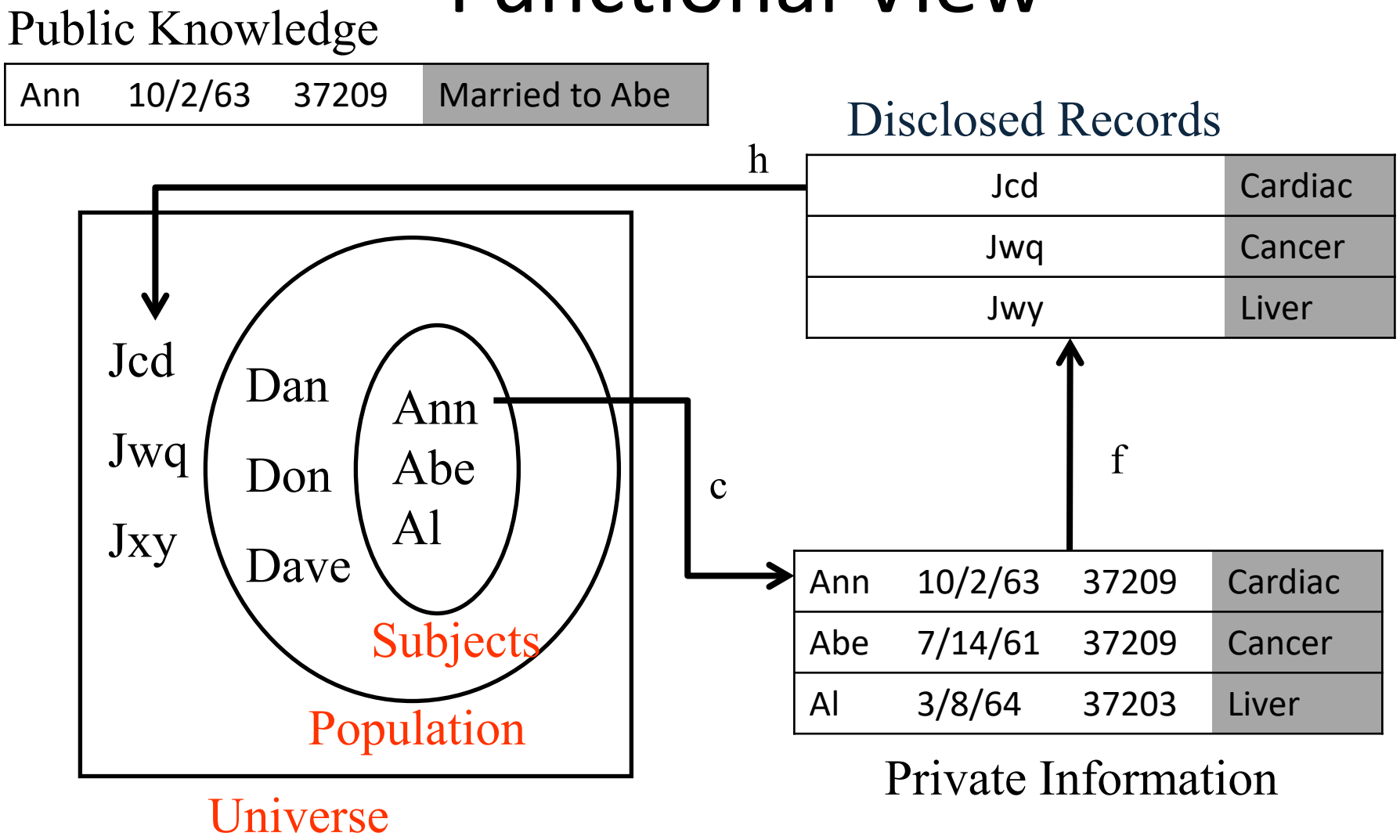
Null-Map

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

Imagine model as a function. NULL(Ann) = Jcd

Functional View

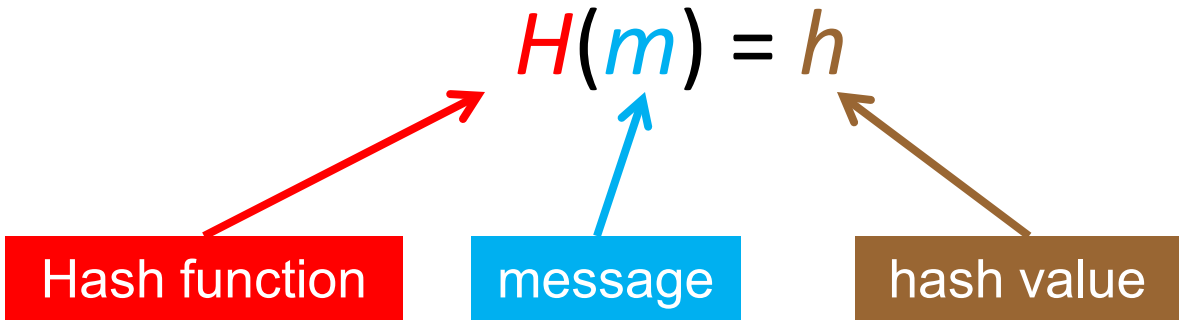


Null?

- Jcd, Jwq, Jxy
- What are these?
- These are not identifiers.
- “ “ “ “ values.
- These are entities, people, things…

Practically Null

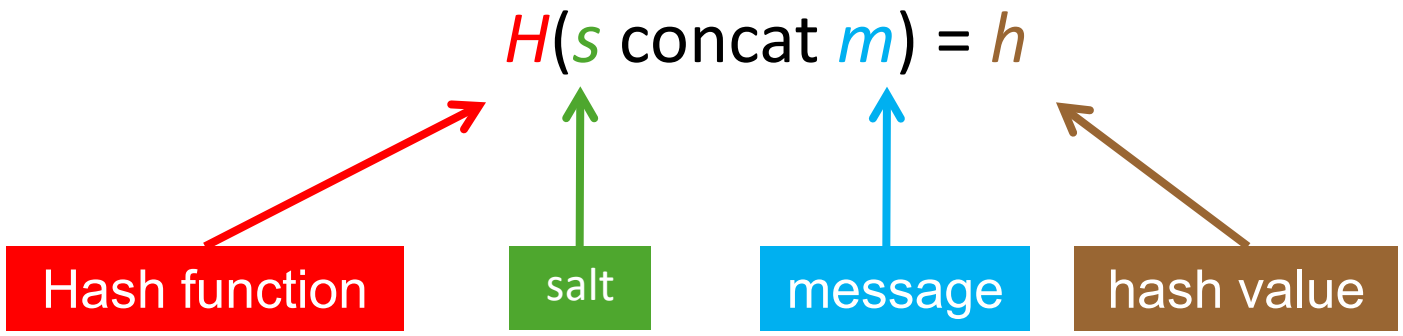
- Hash function: maps “keys” to integers
  - Usually to construct an even distribution on a smaller set of values



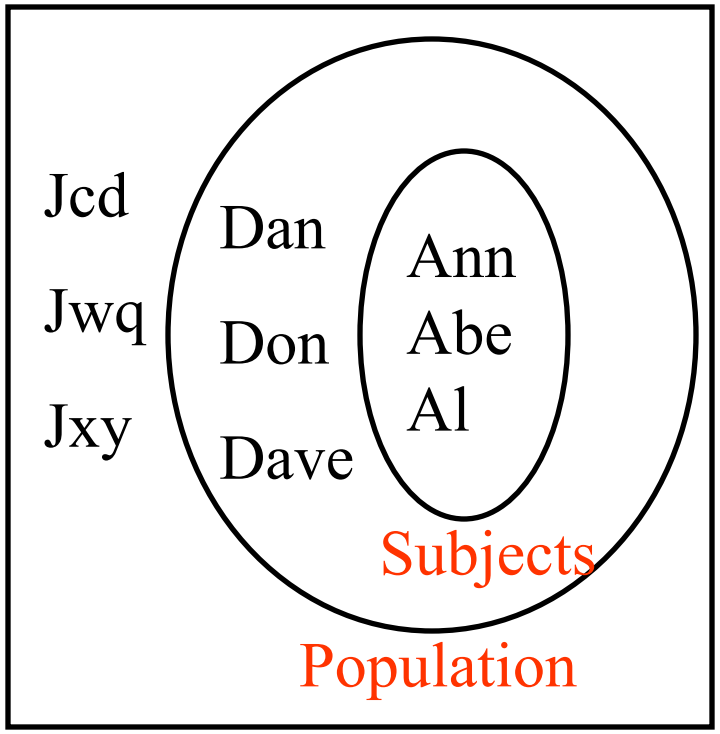
- *HAVAL, MD2, MD4, MD5, PANAMA, RIPEMD, SHA-1, SHA-2, etc.*
- We’ll come back to this if time permits

Salt

- Brute-force possible when attacker has, or can construct, ⟨*message*, *hash*⟩ pairs
- To prevent, or slow down, this attack, we "salt" the hash
- Prior to hashing message, generate a random string and prepend it to your message



Wrong-Map



Al	3/8/64	37203	Cardiac
Ann	10/2/63	37209	Cancer
Abe	7/14/61	37209	Liver

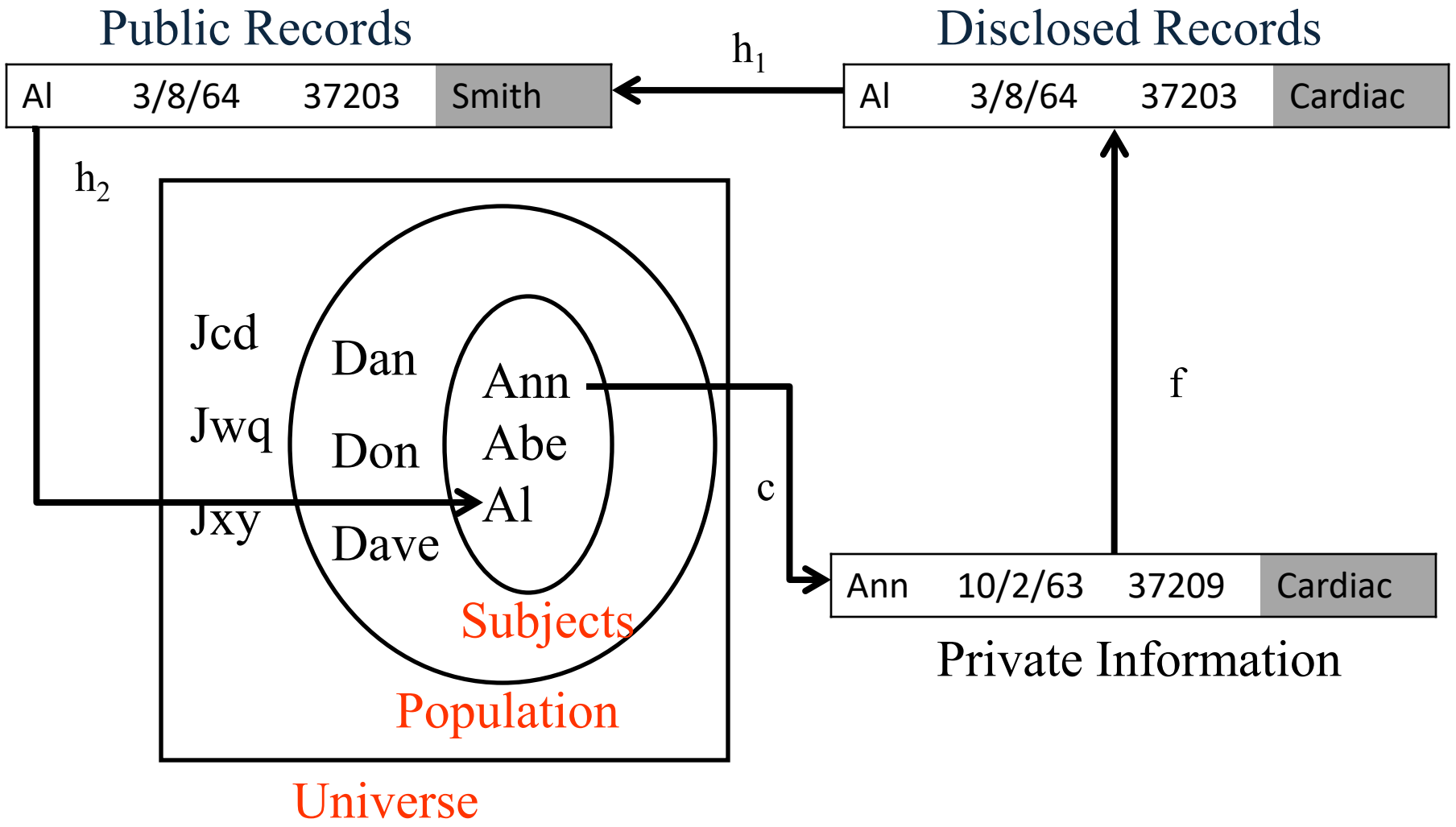
Wrong-Map

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

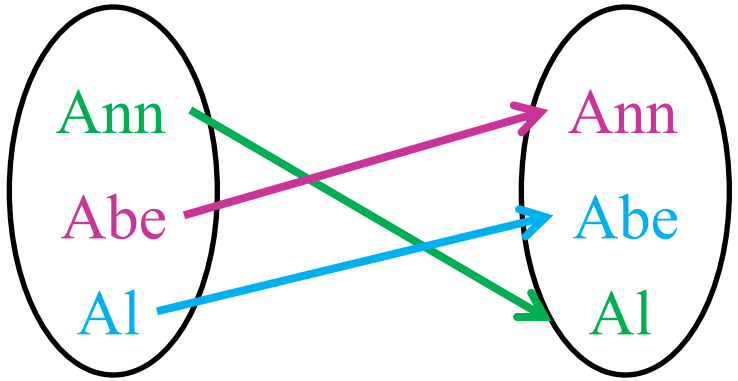
Every record maps to the wrong identity.

Let's Follow Ann!



Wrong-Map

Wrong mapping.



Original Subjects

Disclosed Subjects

Al	3/8/64	37203	Cardiac
Ann	10/2/63	37209	Cancer
Abe	7/14/61	37209	Liver

Wrong-Map

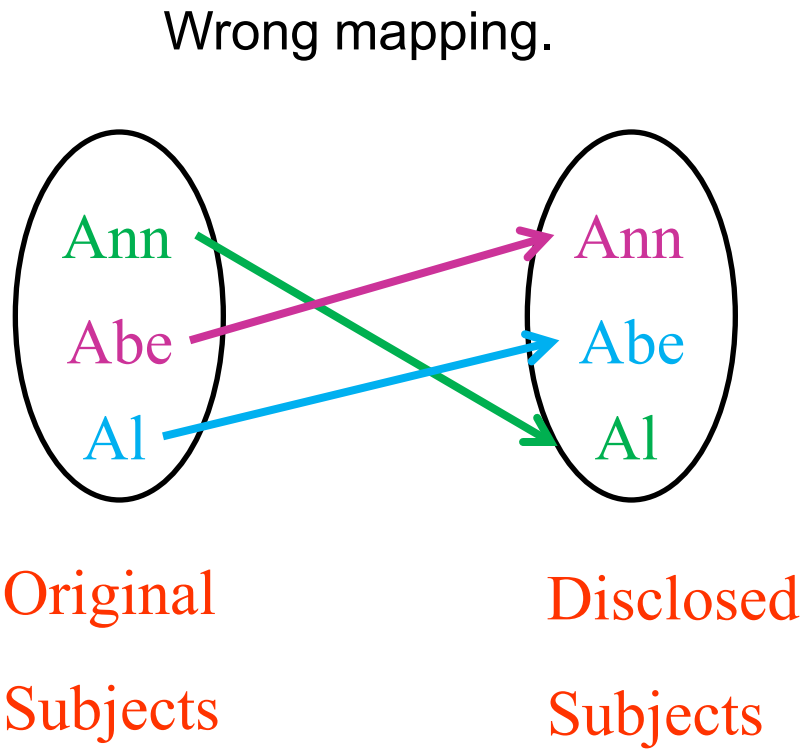
Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

Is this acceptable for privacy? What about utility?

What about a random permutation?

This **is** a Random Permutation



Al	3/8/64	37203	Cardiac
Ann	10/2/63	37209	Cancer
Abe	7/14/61	37209	Liver

Wrong-Map

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

Random Permutations

Original	Al	Ann	Abe
Permutation 1	Al	Ann	Abe
Permutation 2	Al	Abe	Ann
Permutation 3	Ann	Al	Abe
Permutation 4	Ann	Abe	Al
Permutation 5	Abe	Al	Ann
Permutation 6	Abe	Ann	Al

Probability that “Al” really is Al ?

1/3 chance mapped identity to correct record

Wrong-Map

Original	Al	Ann	Abe
Permutation 1	Al	Ann	Abe
Permutation 2	Al	Abe	Ann
Permutation 3	Ann	Al	Abe
Permutation 4	Ann	Abe	Al
Permutation 5	Abe	Al	Ann
Permutation 6	Abe	Ann	Al

Data holder chooses to use permutation 4

Attacker doesn't know if permutation 4 or 5 was applied

An attacker knows that WRONG(Ann) could not be Ann

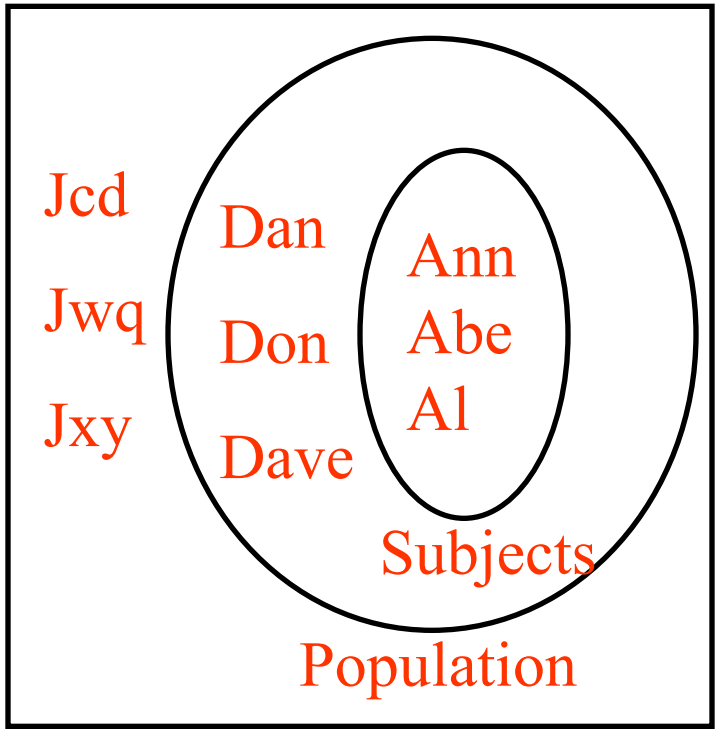
1/2 chance that WRONG(Ann) is really Al

K-Map

- *P*: Population
- *S*: Subjects in Dataset

- K-Map is satisfied, when every record in *S*:
  - “maps” to no less than *k* people in *P*
  - Maps to the correct person from which it was derived

When is this true?



Universe

QUASI-ID

A*	1963	3720*	Cardiac
A*	1961	3720*	Cancer
A*	1964	3720*	Liver

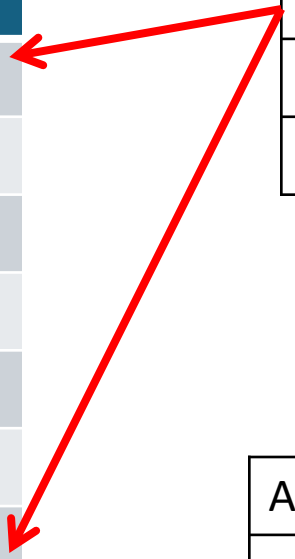
2-Map

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

2-Map

Population		
Ann	10/2/63	37209
Abe	7/14/61	37209
Al	3/8/64	37203
Don	11/2/43	37209
Dan	3/14/87	37209
Dave	1/2/74	37203
Allison	3/4/63	37209
Avery	5/6/61	37209
Arnold	7/8/64	37203



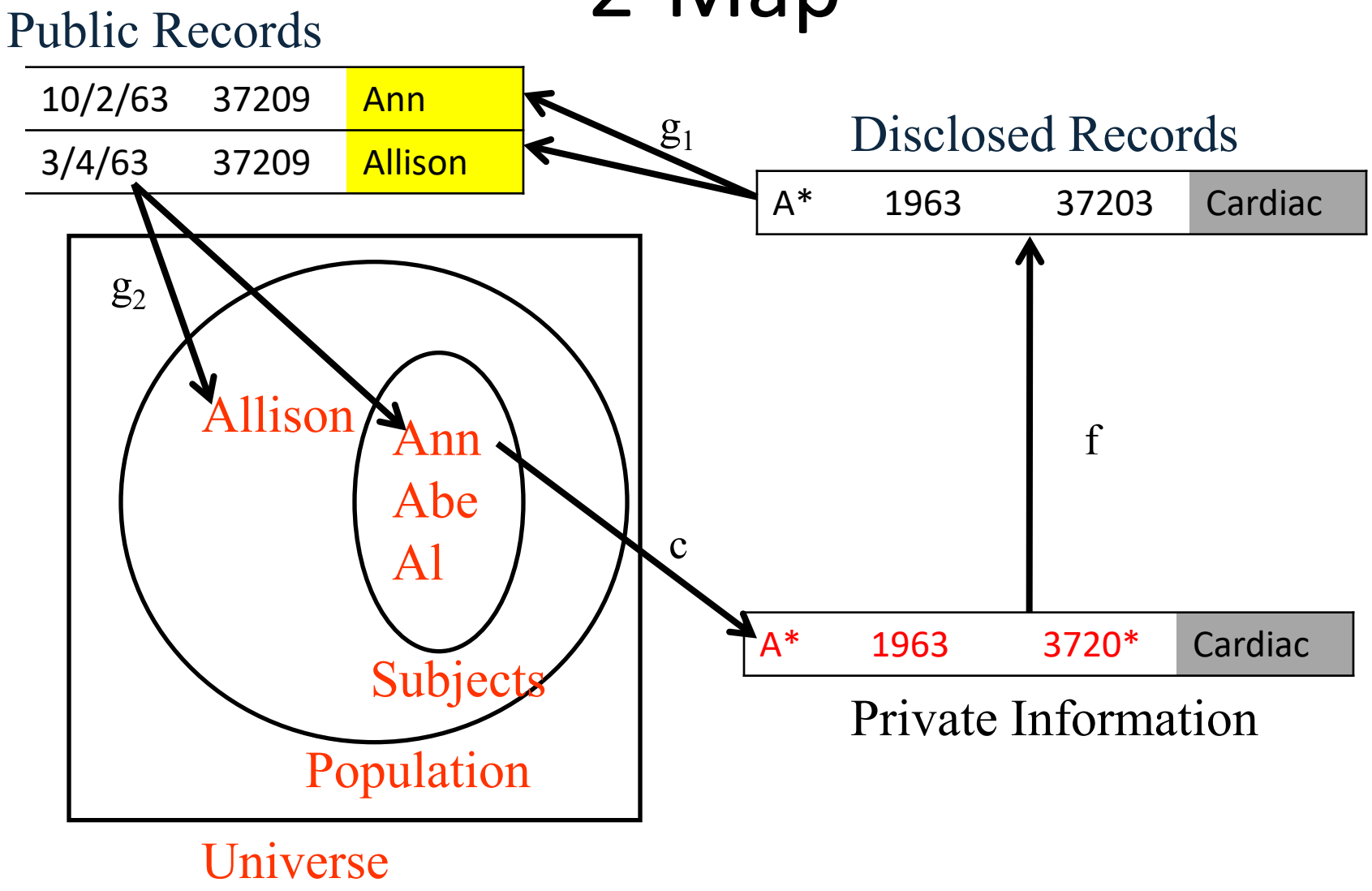
A*	1963	3720*	Cardiac
A*	1961	3720*	Cancer
A*	1964	3720*	Liver

2-Map

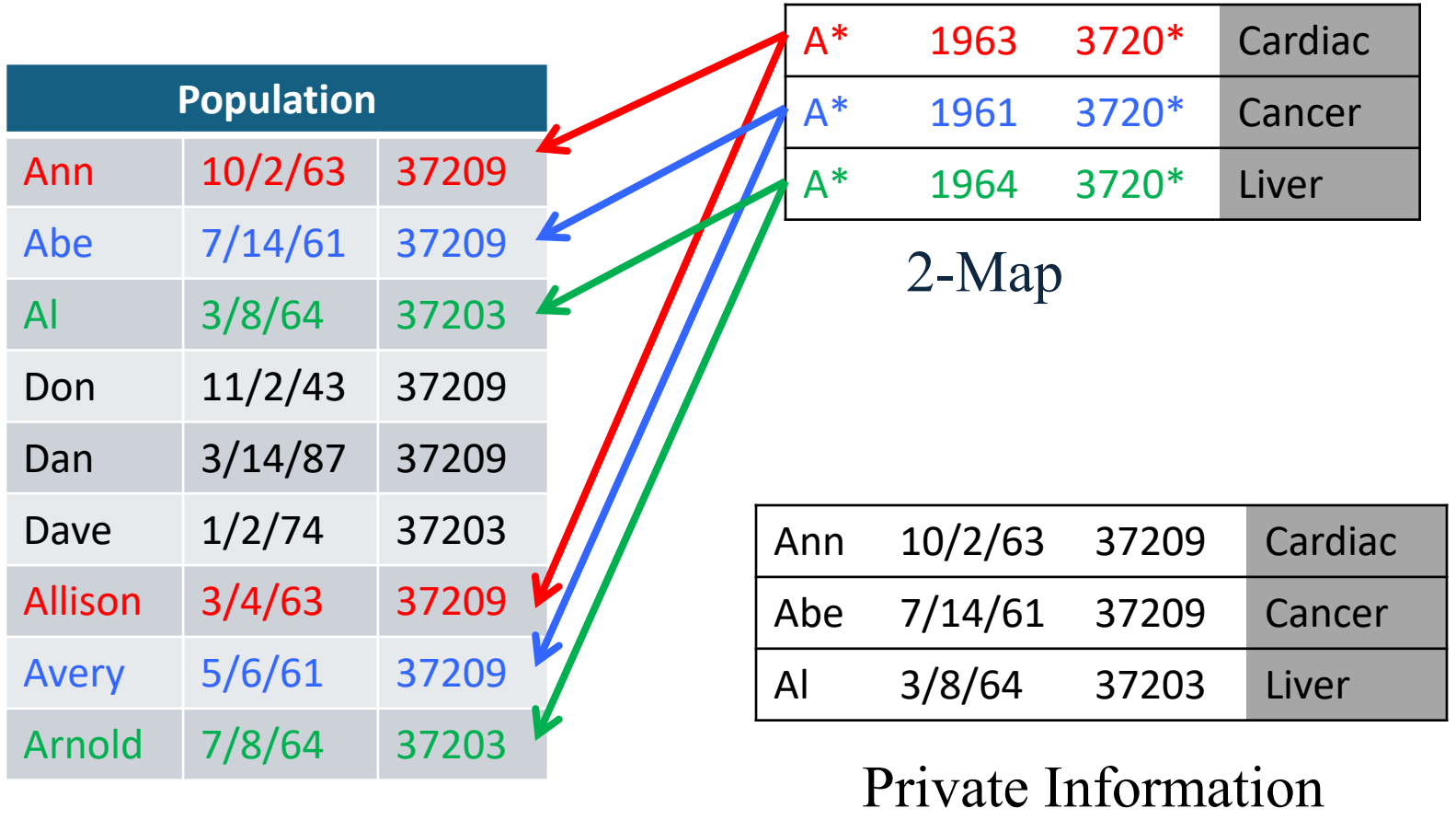
Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

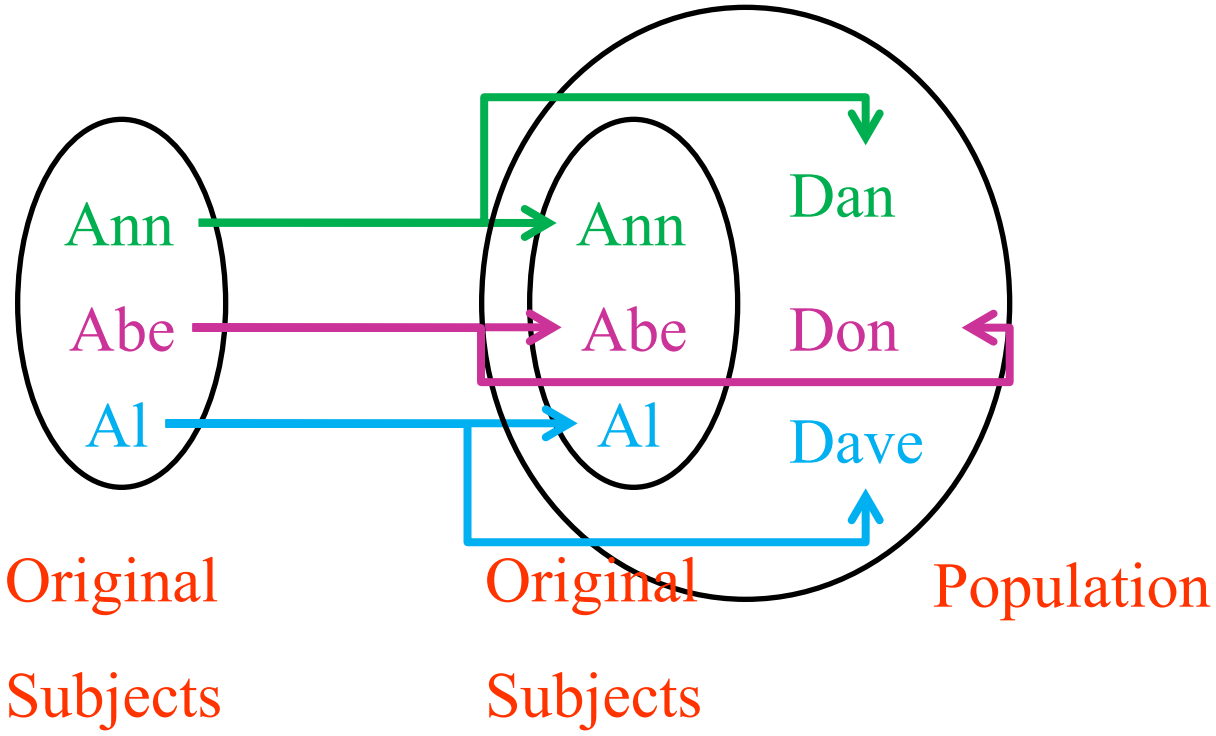
2-Map


Public Records
10/2/63 37209 Ann
3/4/63 37209 Allison
g1
Disclosed Records
A* 1963 37203 Cardiac
g2
Allison
Ann
Abe
Al
Subjects
Population
Universe
c
f
A* 1963 3720* Cardiac
Private Information

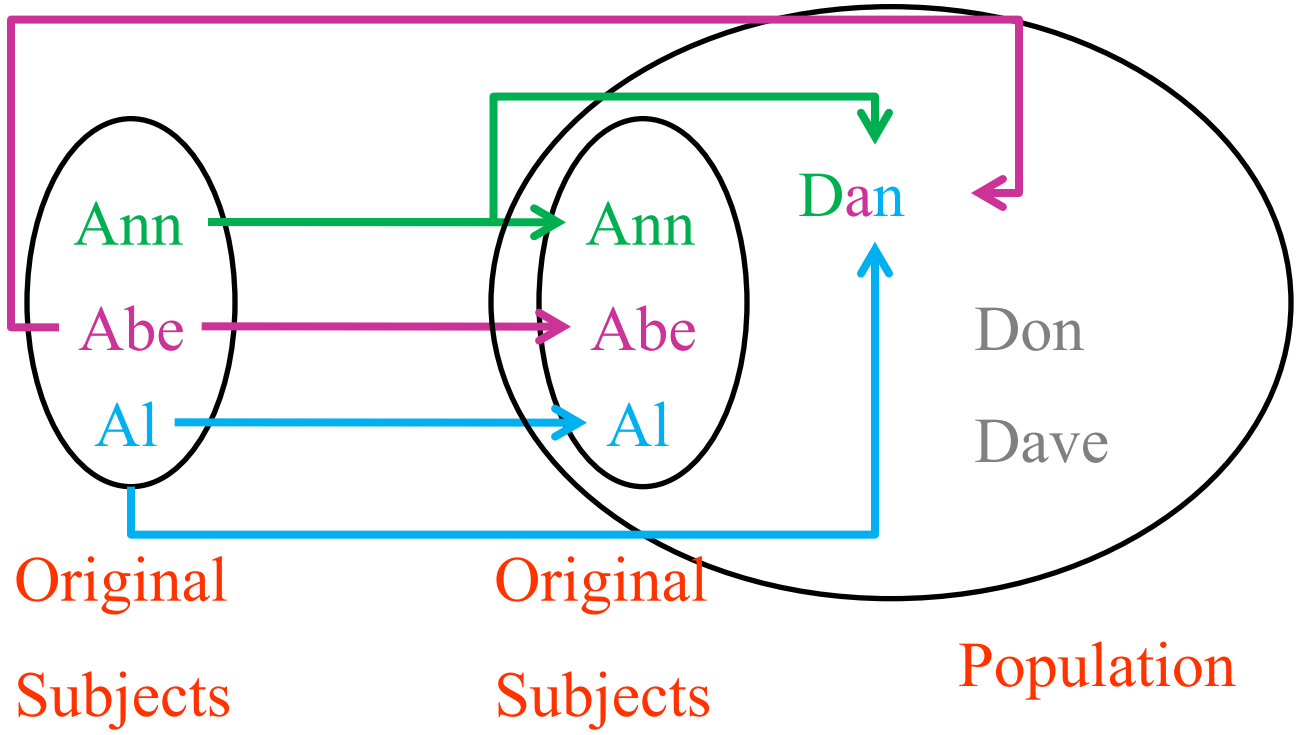
2-Map


Population
Ann 10/2/63 37209
Abe 7/14/61 37209
Al 3/8/64 37203
Don 11/2/43 37209
Dan 3/14/87 37209
Dave 1/2/74 37203
Allison 3/4/63 37209
Avery 5/6/61 37209
Arnold 7/8/64 37203
A* 1963 3720* Cardiac
A* 1961 3720* Cancer
A* 1964 3720* Liver
2-Map
Ann 10/2/63 37209 Cardiac
Abe 7/14/61 37209 Cancer
Al 3/8/64 37203 Liver
Private Information

Another 2-Map Perspective



Now, Does this Satisfy 2-Map?



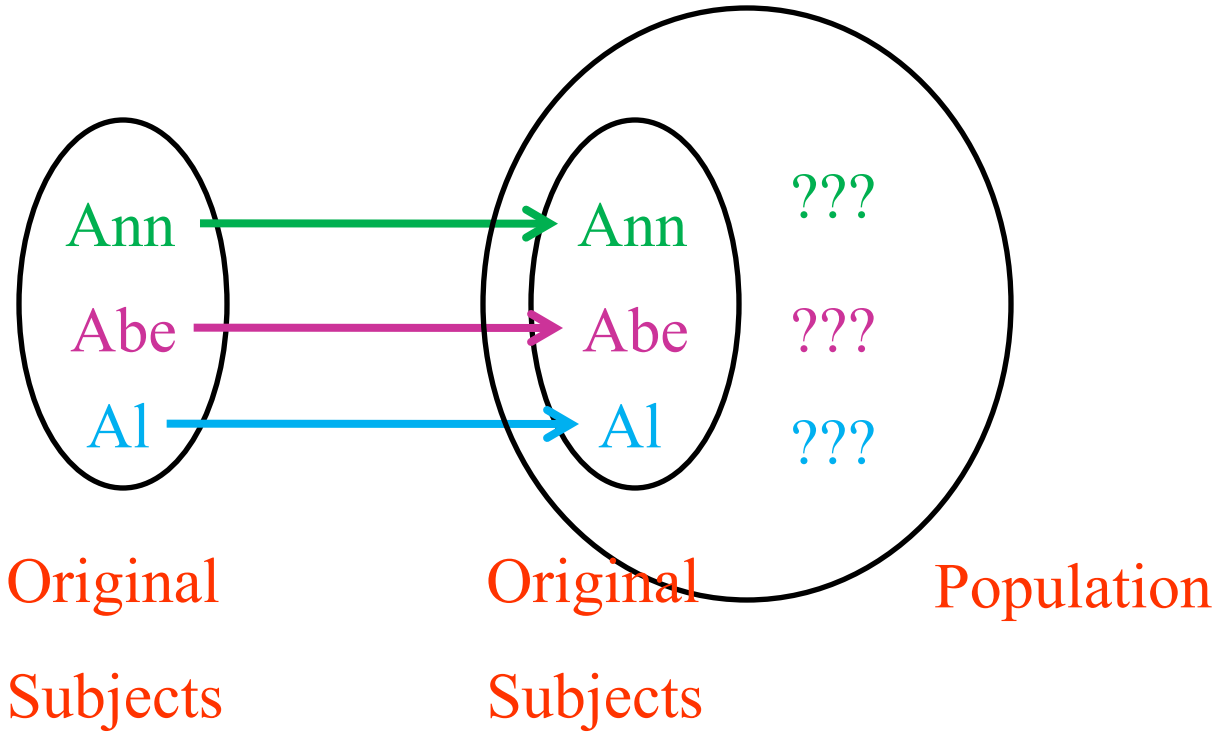
Yes!

Wait a minute

- *K*-Map is a great idea, but what if you don't know who is in, or the features of, the population?
- *K*-Anonymity: Replace population with subjects
- For every record disclosed, there are at least *k*-1 other records that have an equivalent quasi-identifying value

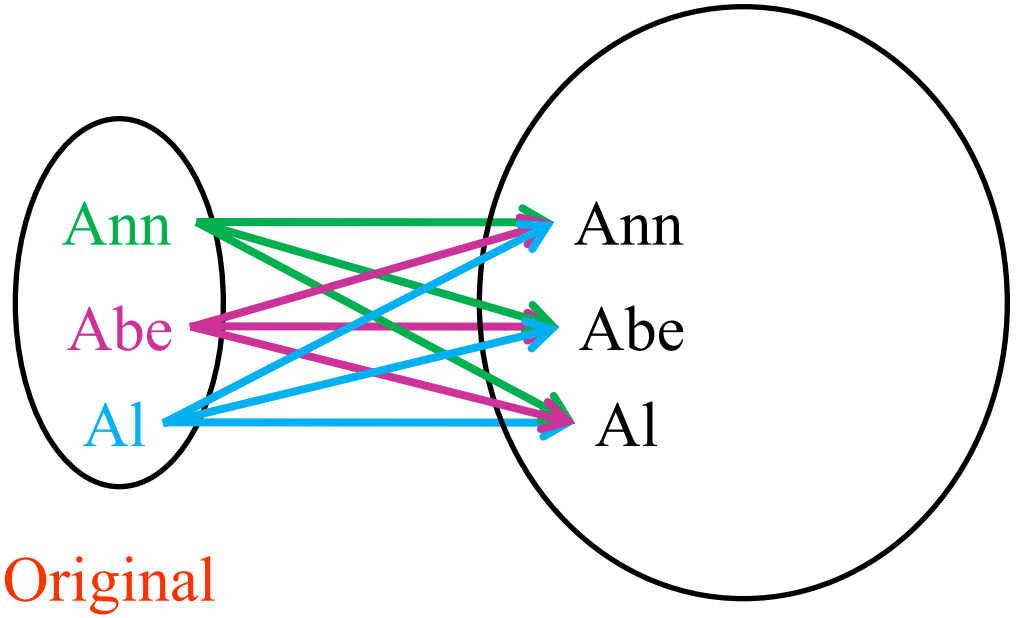
P. Samarati and L. Sweeney. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical Report SRI-CSL-98-04, Computer Science Laboratory, SRI International. 1998.
L. Sweeney. K-anonymity: a model for protecting privacy. International Journal of Uncertainty, Fuzziness, & Knowledge-based systems. 2002; 10(5): 557-570.

2-Anonymity


Ann
Abe
Al
Ann
Abe
Al
???
???
???
Original Subjects
Original Subjects
Population

2-Anonymity

Question: Is this 2-Map?

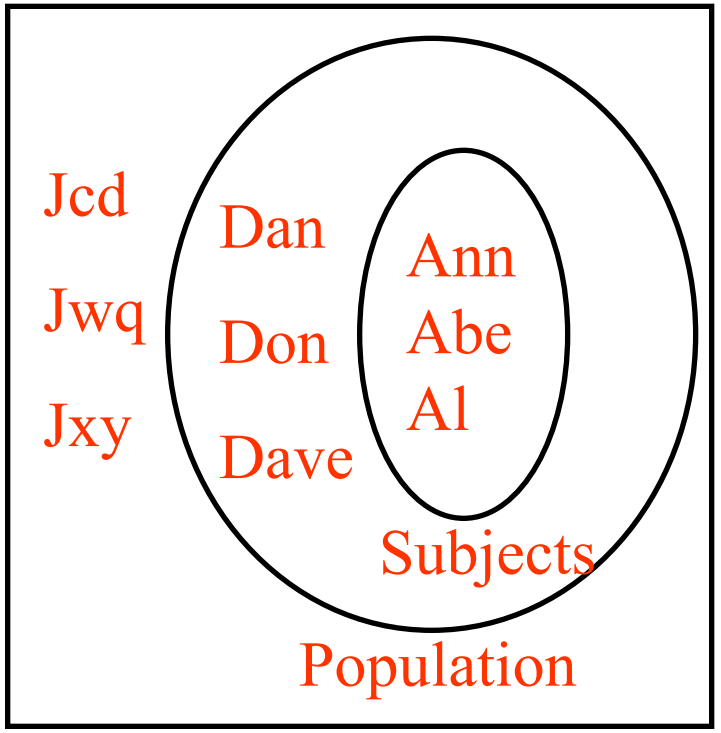


Original Subjects

Original Subjects = Population

Yes!

3-Anonymity



Universe

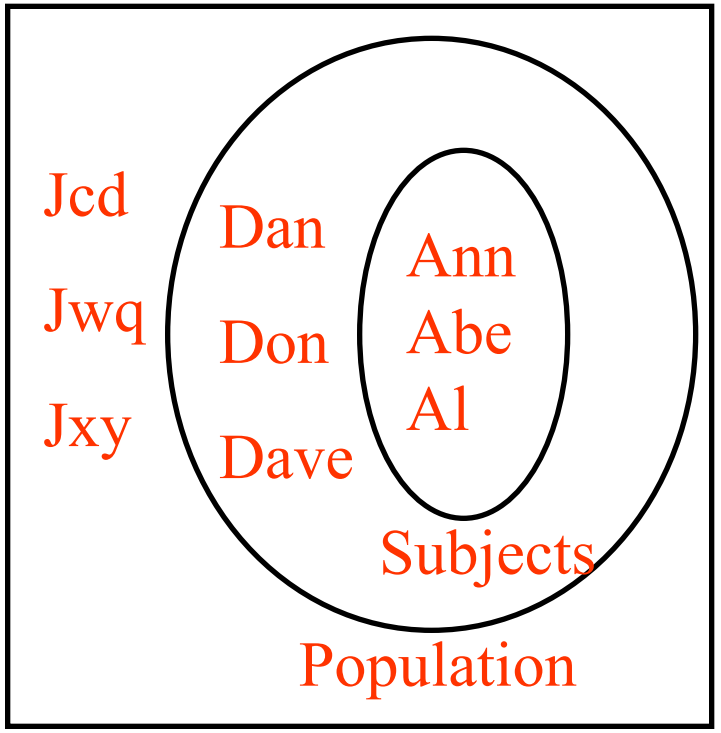
A*	196*	3720*	Cardiac
A*	196*	3720*	Cancer
A*	196*	3720*	Liver

3-Anonymity

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

Question: Is this 2-Anonymity?



A*	196*	3720*	Cardiac
A*	196*	3720*	Cancer
*	*	*	Liver

Ann	10/2/63	37209	Cardiac
Abe	7/14/61	37209	Cancer
Al	3/8/64	37203	Liver

Private Information

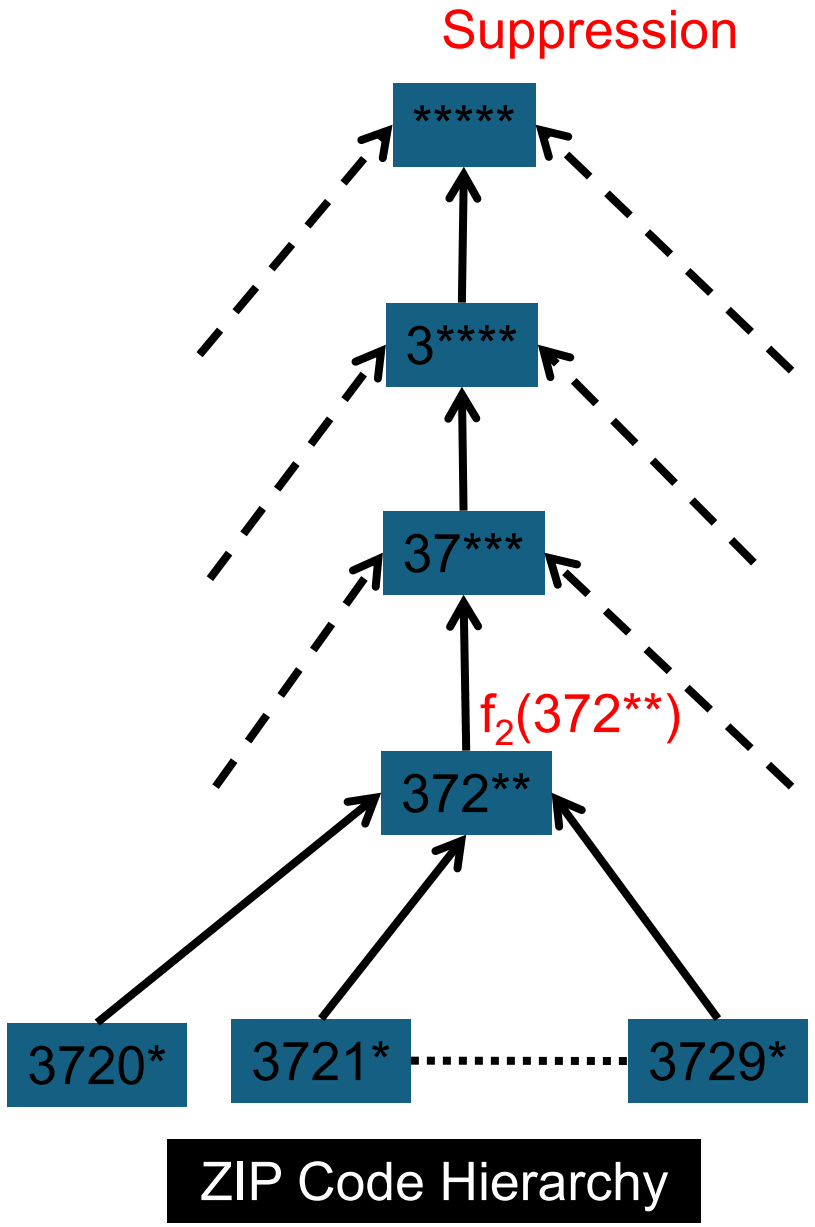
It depends!

K-Anonymity: It's a MODEL

- *K*-Anonymity does not say how the protected data must look
- It only says it must be equivalent
- However, many people choose to follow "generalization" and "suppression"

Generalization

- Substitute less specific values that are semantically **faithful** to the original values
- Tend to talk about "domain generalization hierarchies" (DGH) and generalization functions
- Suppression is a special case – you release no information
  - Actually - saying the value could be anything in attribute's domain



More Rigorous

- *K*-anonymity is based on “indistinguishability” with respect to the shared data T

- Relation ${}_T R_T$
  - Two records $\langle t_i, t_j \rangle$ are **indistinguishable** when every value in the quasi-identifier is equivalent
    - For each attribute in the quasi-ID A_x: $a_{ix} = a_{jx}$

K-Ambiguity

- Based on the notion of **indiscernability***
- Considers the relationship between the values in the disclosed records and the original records
- Protected record $t'_i = [a'_1, \ldots, a'_m]$ and original record $t_i = [a_1, \ldots, a_m]$ are said to be **indiscernable** if
  - For each attribute A_x in the quasi-ID: $a_x \leq a'_x$
- A disclosed table is said to be ***k*-ambiguous**** when each record is indiscernable with *k* records in the original table

*A. Skowron, C. Rauszer. The discernability matrices and functions in information systems. In R. Slowinski, ed. Intelligent Decision Support: Handbook of Applications and Advances of Rough Set Theory, Vol 11, Kluwer, 1992, pp. 331-362.
**S. Vinterbo. On the hardness of the k-ambiguity problem. Technical Report DSGTR-2002-06, Decision Systems Group / Harvard Medical School: Boston, MA. 2002.

Suppression

- Record $t'_i = [a'_1, \ldots, a'_m]$ and

  Record $t_i = [a_1, \ldots, a_m]$ are ***indiscernable*** when

  For all A_x in the quasi-ID: $a'_x \in \{a_x, *\}$

Examples

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213

Private Records

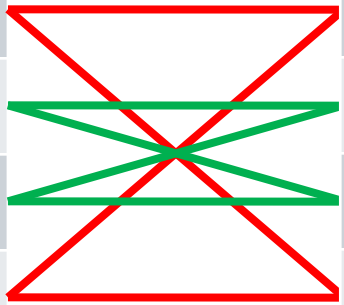
Age	Sex	Zip
30	M	15213
33	*	1521*
33	*	1521*
30	M	15213

2-Anonymous

Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

2-Abiguous

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
33	*	1521*
30	M	15213

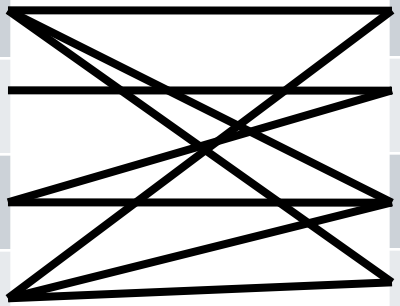
Private Records

2-Anonymous

k-Anonymity forces complete bipartite subgraphs

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

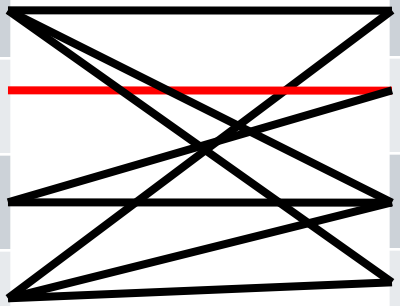
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

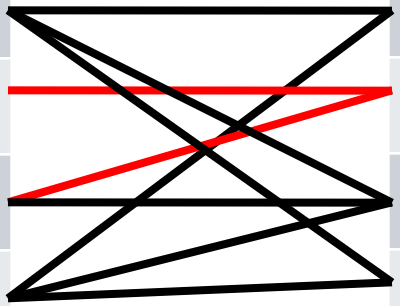
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

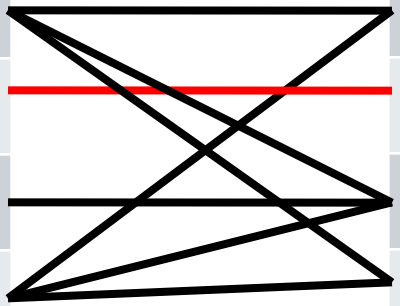
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

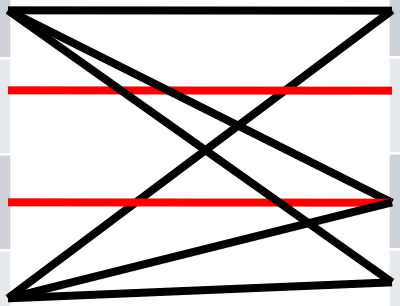
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

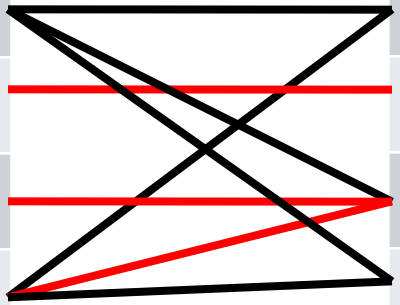
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

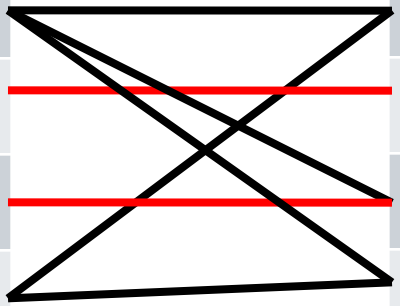
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

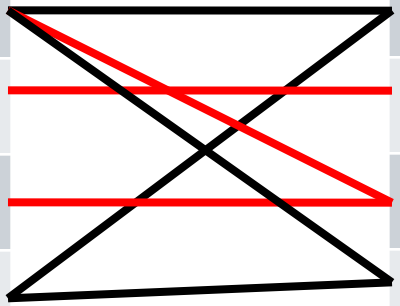
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

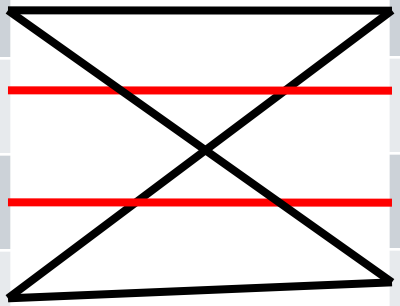
Private Records

2-Abiguous

Notice Anything Strange?

Example

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

Private Records

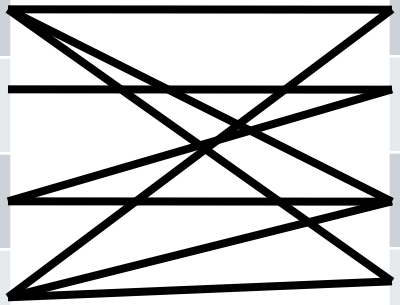
2-Abiguous

Notice Anything Strange?

K-Unlinkability

Age	Sex	Zip
30	M	15213
33	M	15217
33	F	15213
30	M	15213

Population



Age	Sex	Zip
30	M	15213
33	*	1521*
3*	*	15213
30	M	15213

2-Unlinkability

Today's Lecture

- Data Sharing Beliefs and Dogmas
- Protection Models
- Models vs. Methods

True or False

- Models are models
- Models are not methods
- Not all methods are models
- Defining a formal privacy model is “easy”
- Developing a method to transform an existing dataset to satisfy a formal privacy model is “hard”
- What is “hard”?

Final Statement

- k-anonymity is not a method

- k-anonymity is a formal model

https://arx.deidentifier.org/

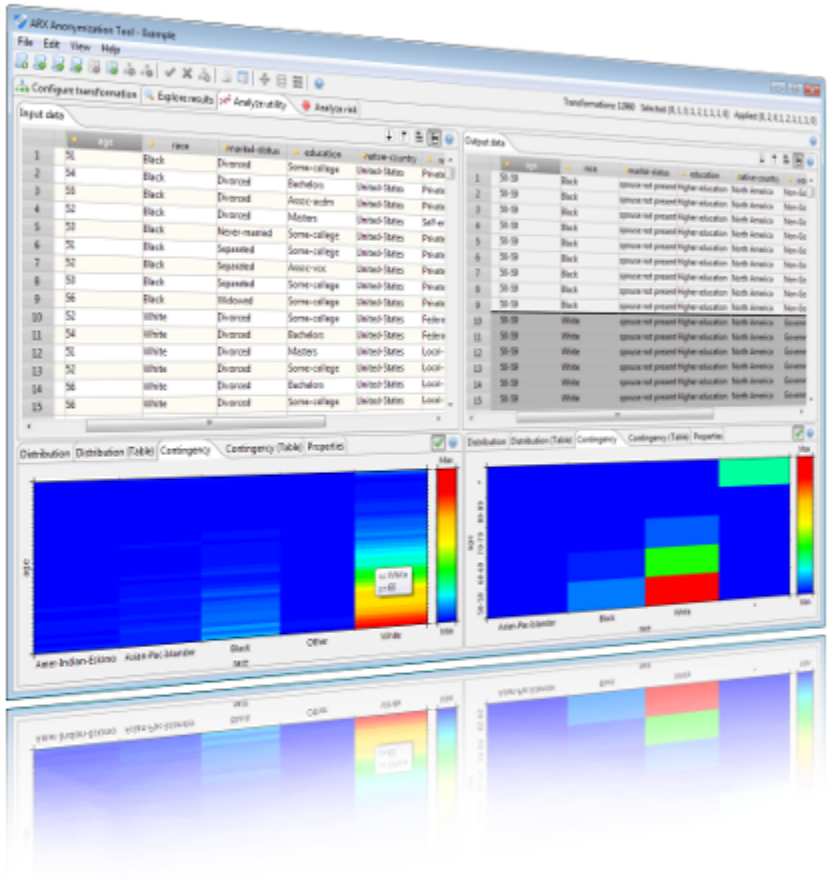
ARX

Data Anonymization Tool

ARX is a comprehensive open source software for anonymizing sensitive personal data. It supports a wide variety of (1) privacy and risk models, (2) methods for transforming data and (3) methods for analyzing the usefulness of output data.

The software has been used in a variety of contexts, including commercial big data analytics platforms, research projects, clinical trial data sharing and for training purposes.

ARX is able to handle large datasets on commodity hardware and it features an intuitive cross-platform graphical user interface. You can find further information here, or directly proceed to our downloads section.



A Note on Hash Functions

- Desirable cryptographic properties
- Length(m) $\in$ [1,∞]
- For any two messages, x and y, Length($H(x)$) = Length($H(y)$)
- $H(m)$ is computationally tractable to compute for any m
- $H(m)$ is one-way
  - In other words, given h, it's difficult to "invert", or find m such that $H(m) = x$
- $H(m)$ is collision-free (easier to say than do)
  - $H(x) = H(y) \rightarrow x = y$ AND $H(x) \neq H(y) \rightarrow x \neq y$

Example: MD5

- Message Digest Algorithm 5 – invented by Ron Rivest in 1991
  - There is an MD2 and MD4 (less secure)
- Converts message into 128-bit hash value
- Step 1: Append Padding Bits
  - Make bit-length is congruent to 448mod(512)
  - Process:
    - Append a “1” bit
    - Append “0” bits until message length is congruent to 448mod(512)

Example: MD5

- Step 2: 64-bit representation of message length is appended
  - If length is < 2^{64}, append only first 64
- Step 3: Initialize a 4 “word” buffer (A, B, C, D), each of which is a 32-bit register:
  - Word A: 01 23 45 67
  - Word B: 89 ab cd ef
  - Word C: fe dc ba 98
  - Word D: 76 54 32 10

Example MD5

- Step 4: Process message in 16-word blocks (512 bits each)
  - Need 4 functions: input 3 32-bit words; output one 32-bit word
    - F(X, Y, Z): $(X \wedge Y) \vee (\neg X \wedge Y)$
    - G(X, Y, Z): $(X \wedge Z) \vee (Y \wedge \neg Z)$
    - H(X, Y, Z): X xor Y xor Z
    - I(X, Y, Z): Y xor $(X \vee \neg Z)$

Example MD5

- For i = 0 to N / 16 – 1 *//Process each 16 word block*
  - For j = 0 to 15 *//Copy block i into X*
    - X[j] = M[i*16 + j]
  - AA = A; BB = B; CC = C; DD = D;
  - 4 Rounds of 16 steps; Each step outputs [abcd k s i] which represents
    a = b + ((a + F(b,c,d) + X[k] + T[i]) <<< s)
    T[i] is the integer part of 4294967296 * abs(sin(i)), i is in radians

Example MD5

- Round 1

[abcd k s i] denotes the operation a = b + ((a + F(b,c,d) + X[k] + T[i]) <<< s).

Remember: F(X, Y, Z): $(X \wedge Y) \vee (\neg X \wedge Y)$

```
[ABCD 0 7 1]    [DABC 1 12 2]    [CDAB 2 17 3]    [BCDA 3 22 4]
[ABCD 4 7 5]    [DABC 5 12 6]    [CDAB 6 17 7]    [BCDA 7 22 8]
[ABCD 8 7 9]    [DABC 9 12 10]   [CDAB 10 17 11]  [BCDA 11 22 12]
[ABCD 12 7 13]  [DABC 13 12 14]  [CDAB 14 17 15]  [BCDA 15 22 16]
```

Example MD5

- Round 2

[abcd k s i] denotes the operation a = b + ((a + G(b,c,d) + X[k] + T[i]) <<< s).

```
[ABCD  1  5 17] [DABC  6  9 18] [CDAB 11 14 19] [BCDA  0 20 20]
[ABCD  5  5 21] [DABC 10  9 22] [CDAB 15 14 23] [BCDA  4 20 24]
[ABCD  9  5 25] [DABC 14  9 26] [CDAB  3 14 27] [BCDA  8 20 28]
[ABCD 13  5 29] [DABC  2  9 30] [CDAB  7 14 31] [BCDA 12 20 32]
```

- Round 3

  Operation a = b + ((a + H(b,c,d) + X[k] + T[i]) <<< s).
- Round 4

  Operation a = b + ((a + I(b,c,d) + X[k] + T[i]) <<< s).

Example MD5

- Finally, increment each of the four registers by the value it had before the block processing began

  A = A + AA; B = B + BB; C = C + CC; D = D + DD

- Output is ABCD

- Implementations
  - http://userpages.umbc.edu/~mabzug1/cs/md5/md5.html
- There are some weaknesses, but we're not going into this today
- Brute force attack is possible, 2^{64} possibilities…

Homework Assignment 1

- 3 Questions (including programming)
- Cover Lectures 9 -12
- Due at 15:00 on April 30, 2025

- Temp. Office Hour:
  - When: 16:40-18:00, April 25, 2025 (Friday)
  - Where: BME 4-228

Readings for the Next Week

- 1. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, Malin B. **Expanding access to large-scale genomic data while promoting privacy: a game theoretic approach**. *The American Journal of Human Genetics*. 2017 Feb 2;100(2):316-22. https://www.cell.com/action/showPdf?pii=S0002-9297%2816%2930526-2
- Optional
  - ❑ 2. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, Ganta R, Heatherly R, Malin BA. **A game theoretic framework for analyzing re-identification risk**. *PloS one*. 2015 Mar 25;10(3):e0120592. https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0120592&type=printable
  - ❑ 3. Wan Z, Vorobeychik Y, Xia W, Liu Y, Wooders M, Guo J, Yin Z, Clayton EW, Kantarcioglu M, Malin BA. **Using game theory to thwart multistage privacy intrusions when sharing data**. *Science Advances*. 2021 Dec 10;7(50):eabe9986. https://www.science.org/doi/pdf/10.1126/sciadv.abe9986

Feedback Survey

- One thing you learned or felt was valuable from today's class & reading

- Muddiest point: what, if anything, feels unclear, confusing or "muddy"

- https://www.wjx.cn/vm/hX0mIro.aspx