Role of Data Privacy Towards Safe and Trustworthy Mental Health Assistants

Team 3: Rachit Saini, Ekta Pandey, Surjodeep Sarkar
and Bhavani Shankar Mahamkali

What are Mental Health Apps?

Mental health apps are **mobile apps** designed to **help users** improve their **emotional well-being**, become more mindful, and address common **mental health issues**.

Why Mental Health Apps?

1. Reports released in August 2021 indicated that 1.6 million people in England were on waiting lists to seek professional help with mental health care. Such an overwhelming rise in the number of patients as compared to health practitioners necessitated the use of health apps.

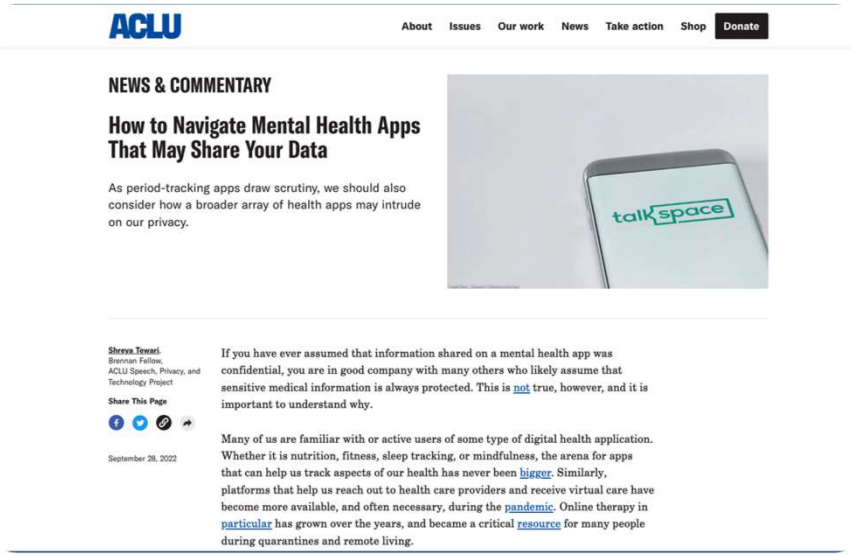
2. The psychological stigma in patients, which even refrained them from seeing a health practitioner.

Source: **Towards Explainable and Safe Conversational Agents for Mental Health: A Survey**
https://arxiv.org/abs/2304.13191

What is the need of the hour?

- The U.S. Department of Health and Human Services (HHS) recently released updated **guidance** on cellphones, health information, and **HIPAA**, confirming that the HIPAA Privacy Rule does **not** apply to most health apps as they are not "**covered entities**" under the law.

- Privacy policies don't always make it clear what kind of data could be shared, and how it could be used.



Reference: https://www.aclu.org/news/privacy-technology/how-to-navigate-mental-health-apps-that-may-share-your-data

https://www.hhs.gov/hipaa/forprofessionals/privacy/index.html

Problems with Health Apps?

The Washington Post
Democracy Dies in Darkness

Help Desk Tech in Your Life | Tech at Work | Your Data and Privacy | Internet Access | What's New | Ethical Issues | Ask a Question

YOUR DATA AND PRIVACY

Health apps share your concerns with advertisers. HIPAA can't stop it.

From 'depression' to 'HIV,' we found popular health apps sharing potential health concerns and user identifiers with dozens of ad companies

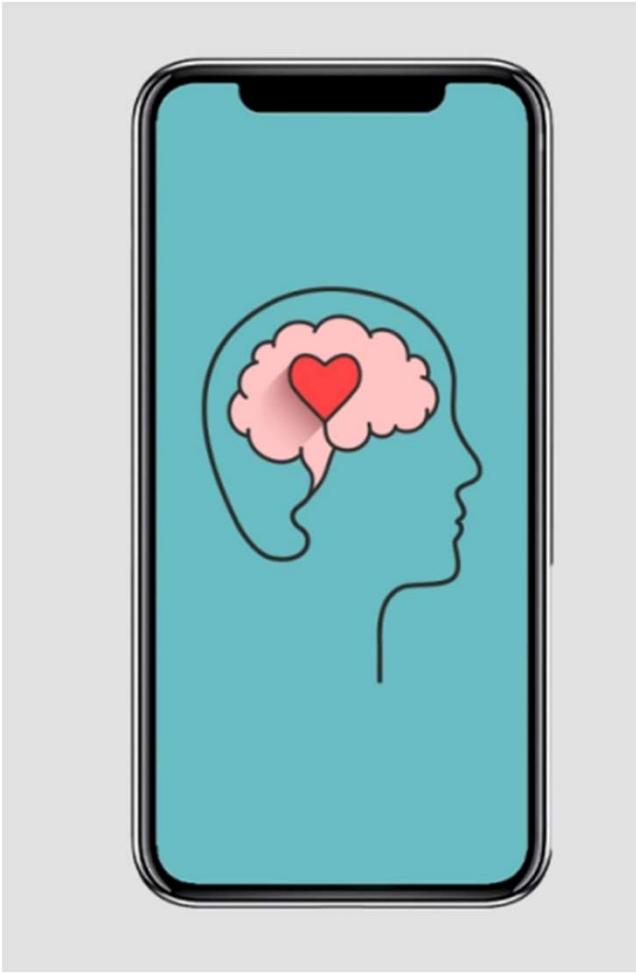
By Tatum Hunter and Jeremy B. Merrill

Updated September 22, 2022 at 10:26 a.m. EDT | Published September 22, 2022 at 7:00 a.m. EDT

Reference: https://www.washingtonpost.com/technology/2022/09/22/health-apps-privacy/

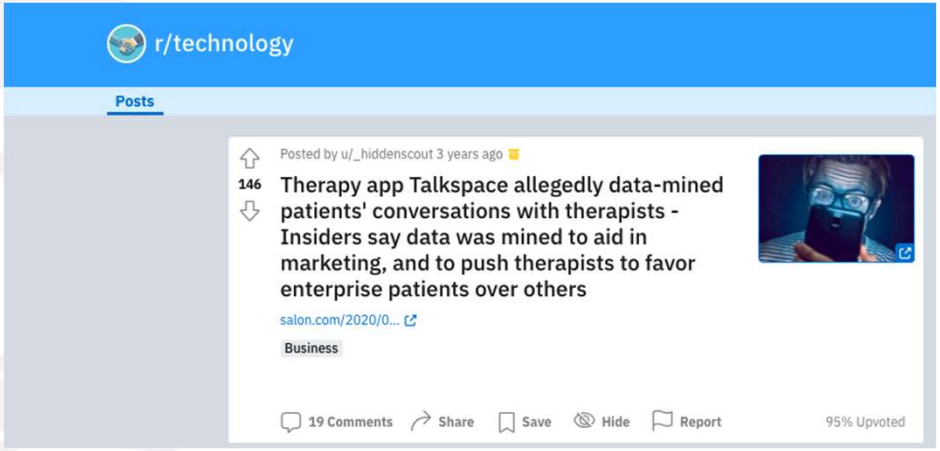
Popular Health Apps

- WoeBot
- **TalkSpace**
- HeadSpace
- BetterHelp
- Calm



Reference: https://formative.jmir.org/2022/6/e36521

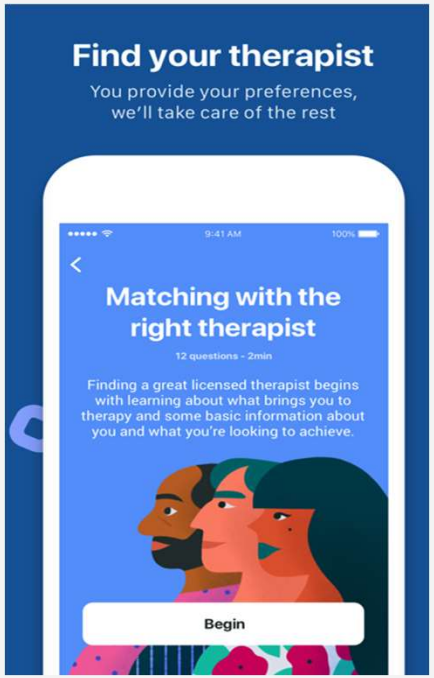
TalkSpace Fails **Privacy Test** and **Data Mines** for Marketing



- https://www.techrepublic.com/article/mozilla-privacy-survey-finds-mental-health-and-prayer-apps-fail-privacy-test-pretty-spectacularly/
- https://www.reddit.com/r/technology/comments/i7h0el/therapy_app_talkspace_allegedly_datamined/

Case Study: TalkSpace

- A teletherapy app that connects users with licensed therapists for online therapy sessions.



Meet our network of licensed providers

Our network of providers cover a range of specialities to meet your specific needs. Get matched today!

- ✓ Depression
- ✓ Relationships
- ✓ Anxiety & Stress
- ✓ Parenting
- ✓ LGBTQIA+
- ✓ Chronic illness
- ✓ Eating disorders
- ✓ Anger management
- ✓ Childhood abuse
- ✓ Mood disorders
- ✓ OCD
- ✓ Trauma & grief
- ✓ Substance abuse
- ✓ Family conflict
- and more...

Source: https://www.talkspace.com/

Features of TalkSpace

- Psychoeducation
- Mindfulness
- Deep Breathing
- Coach/Therapist Connection
- Track Symptoms
- Track Medication
- Goal Setting/Habits

Source: https://mindapps.org/ViewApp

Privacy Policy of TalkSpace

Personal Data Collected	What they do?
Name, Address, Country, DOB, Phone, Gender, Email, Relationship Status, Organization/Employer, Payment Information, Insurance Information, Transaction History, Referral Source	• Provide you with treatment information. • Enrol you in services and administer your account. • Provide announcements, including for marketing purposes. • Permissive reporting of abuse.
• Information you disclose in chat data and your chat sharing preferences (transcripts) • Audio/Video communication • Documents you share with your therapist • Information collected via our symptom tracker • Information collected via chat, telephone, or email support channels	• To provide you with the Services • To conduct clinical and other academic research, internally
Technical information from software or systems hosting the Services, and from the systems, applications and devices that are used to access the Services.	• Provide support to users (therapists and patients) • To develop new products • Monitor performance of our data centers and networks
Data collected via cookies, pixels and other tracking technologies (such as Google Analytics and Google Ads) • Geolocation information • Internet protocol (IP) addresses • Internet service provider (ISP) • Device ID	• To provide you with and to evaluate, improve and develop the Services • Evaluate the success of our marketing campaigns • Marketing, including tailoring advertising

Source: https://www.talkspace.com/public/privacy-policy

Privacy Concern?

- What kind of data are being collected?
- How long the data is being stored?
- Whether the data can be deleted?
- What is the purpose of the data being collected?
- Third Party Access to Data?
- Informed consent?

Major Concerns for Users…

- Once a therapist/client relationship is established, no personally identifiable information is disclosed to third-party service providers about that user, **unless the third party has signed a business associate agreement.**
- They can use personal information for **marketing, tailored advertising**, and research purposes.
- *Note, this mention of potentially using **psychotherapy notes** for **marketing purposes** with permission was removed from the Talkspace's privacy policy with the June 14, 2022, update.*
- Their updated policy now refers to **"chat data"** instead.
- How clear is the authorization process. **A big question mark ?**

Source: https://foundation.mozilla.org/en/privacynotincluded/talkspace/

Retention Policy:

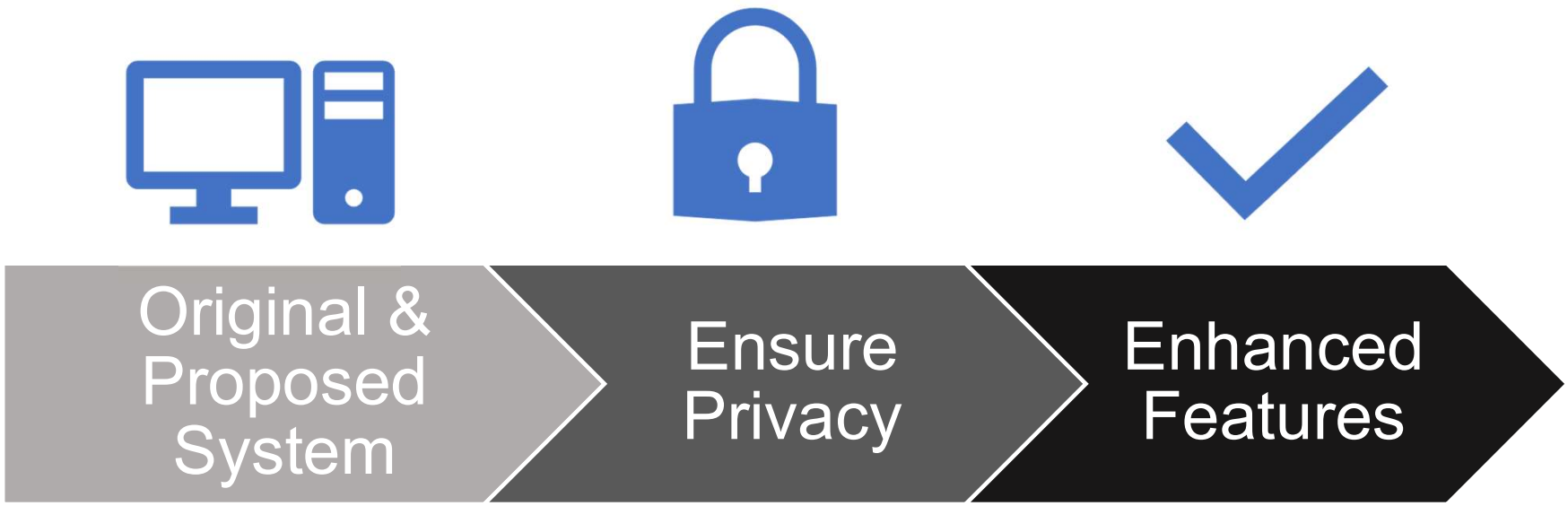
Will retain your information in accordance with the appropriate statutory limitation periods as required by local law, in line with their legitimate business purposes for as long as your account is active or for as long as needed to provide you with the Services, as required in order to comply with Talkspace's legal obligations, a court order or to defend or pursue legal claims, in line with industry codes of practice, to resolve disputes and enforce their agreements.

A very long statement….

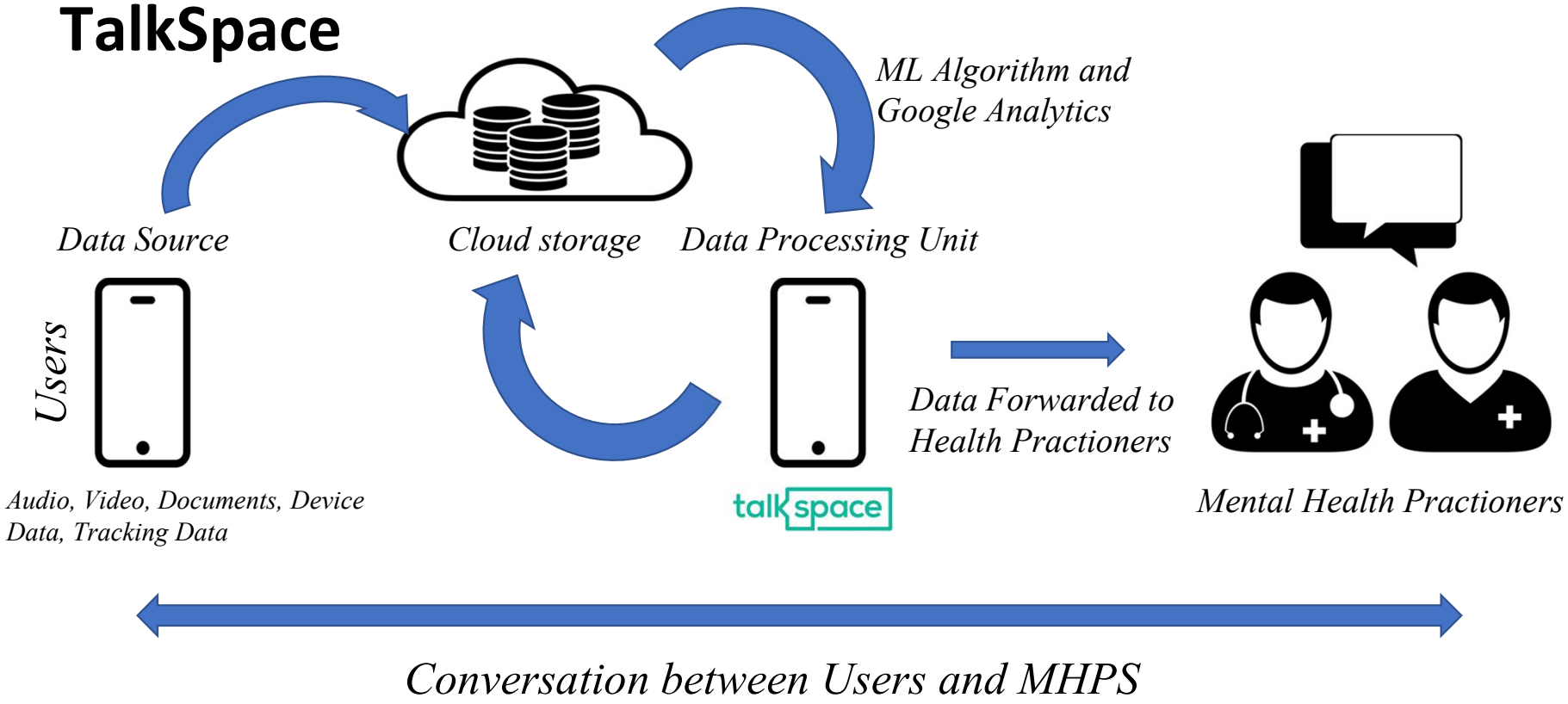
No more retention details are provided in the privacy notice.

Source:https://foundation.mozilla.org/en/privacynotincluded/talkspace/

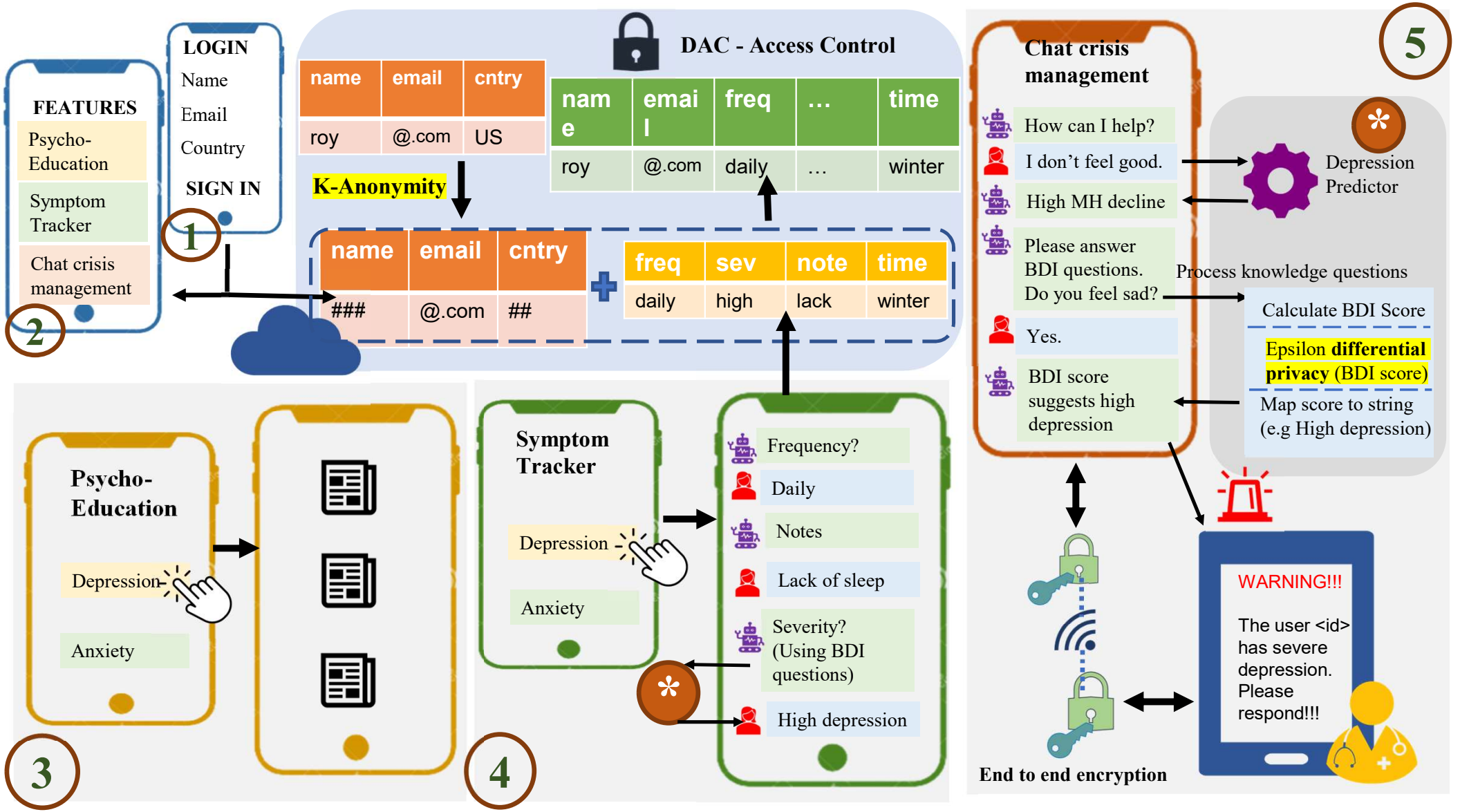
In a Nutshell



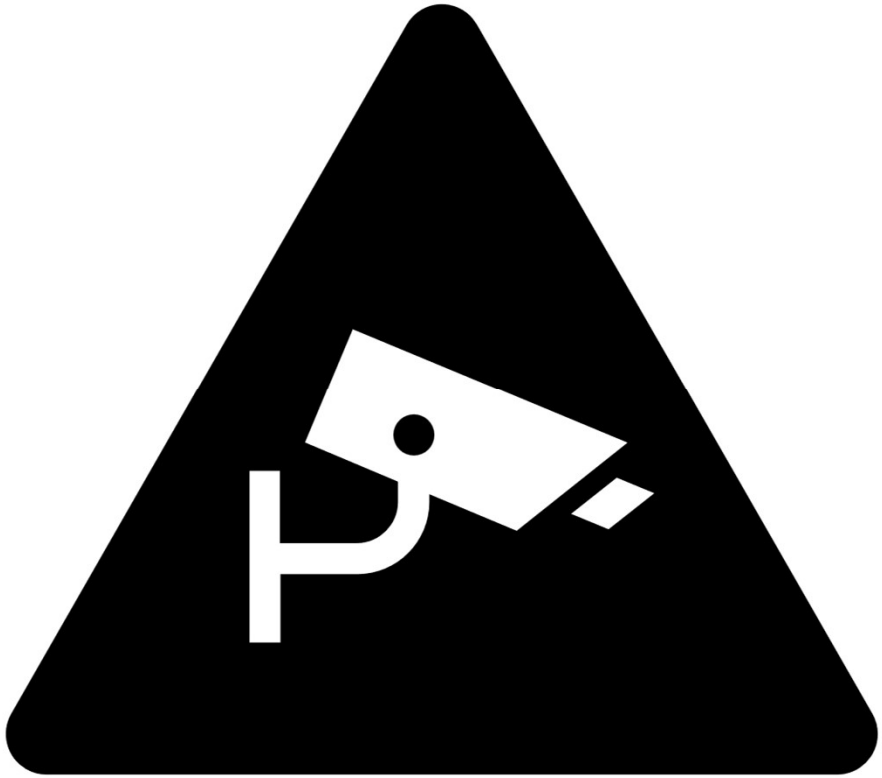
How Does It Work?


TalkSpace
ML Algorithm and
Google Analytics
Data Source
Cloud storage
Data Processing Unit
Users
Data Forwarded to
Health Practioners
Audio, Video, Documents, Device
Data, Tracking Data
talkspace
Mental Health Practioners
Conversation between Users and MHPS

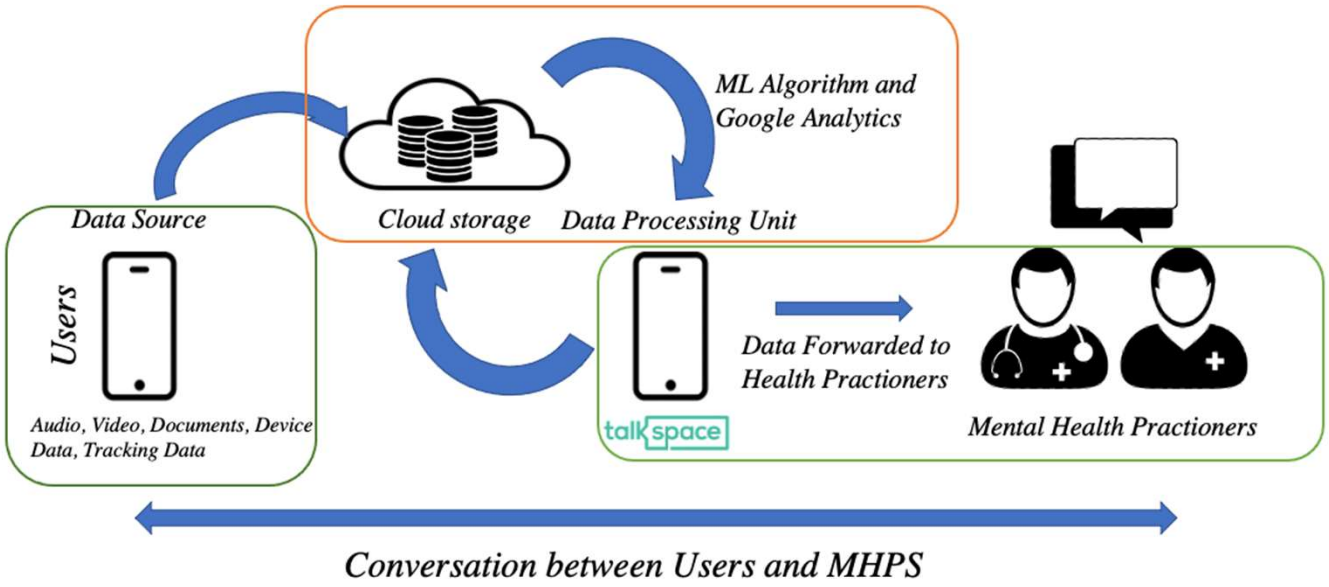
Modified Architecture


LOGIN
Name
Email
Country
SIGN IN
FEATURES
Psycho-Education
Symptom Tracker
Chat crisis management
1
2
DAC - Access Control
name
email
cntry
roy
@.com
US
K-Anonymity
name
email
cntry
###
@.com
##
nam e
emai l
freq
…
time
roy
@.com
daily
…
winter
freq
sev
note
time
daily
high
lack
winter
Chat crisis management
How can I help?
I don't feel good.
High MH decline
Please answer BDI questions. Do you feel sad?
Yes.
BDI score suggests high depression
Depression Predictor
Process knowledge questions
Calculate BDI Score
Epsilon differential privacy (BDI score)
Map score to string (e.g High depression)
5
Psycho-Education
Depression
Anxiety
3
Symptom Tracker
Depression
Anxiety
4
Frequency?
Daily
Notes
Lack of sleep
Severity? (Using BDI questions)
High depression
WARNING!!!
The user <id> has severe depression. Please respond!!!
End to end encryption

How do we minimize Privacy Issues..

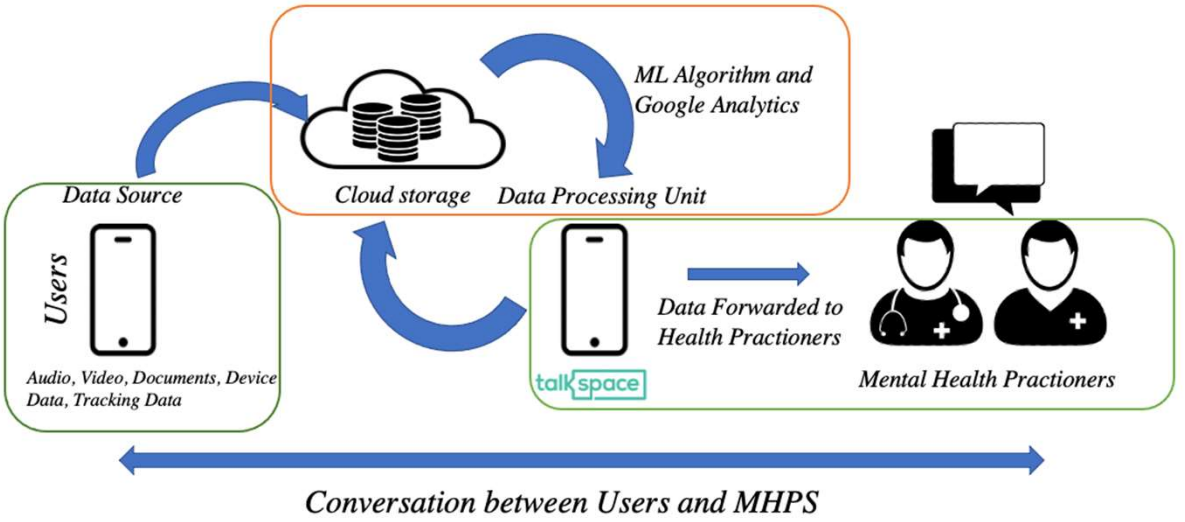


Classify Entities in Domains



- **User Domain:** components under user control (e.g., user devices) [Green Box]
- **Service Domain:** components outside of user control (e.g., backend server at service provider) [Red Box]

Identifying Necessary Data (Data Minimization)



- 1. **Personal Data**(Name, Age, Gender, Country, Email, Relationship Status)
- 2. **Payment Data** -> Email ID
- 3. **Medical Data** -> Symptoms, Severity, Frequency, Notes
- 4. **Device Data** -> Timestamp

Classification of Attributes:

Key Attributes:

- 1. UserID, Name, Email

Quasi-identifiers:

- 1. Gender, Country, Age, Relationship Status

Sensitive Attributes:

- 1. Symptoms, Severity, Frequency, Notes

TalkSpace Features

Psychoeducation

Psychoeducation refers to the process of providing education and information to those seeking mental health services.

Symptom Tracker

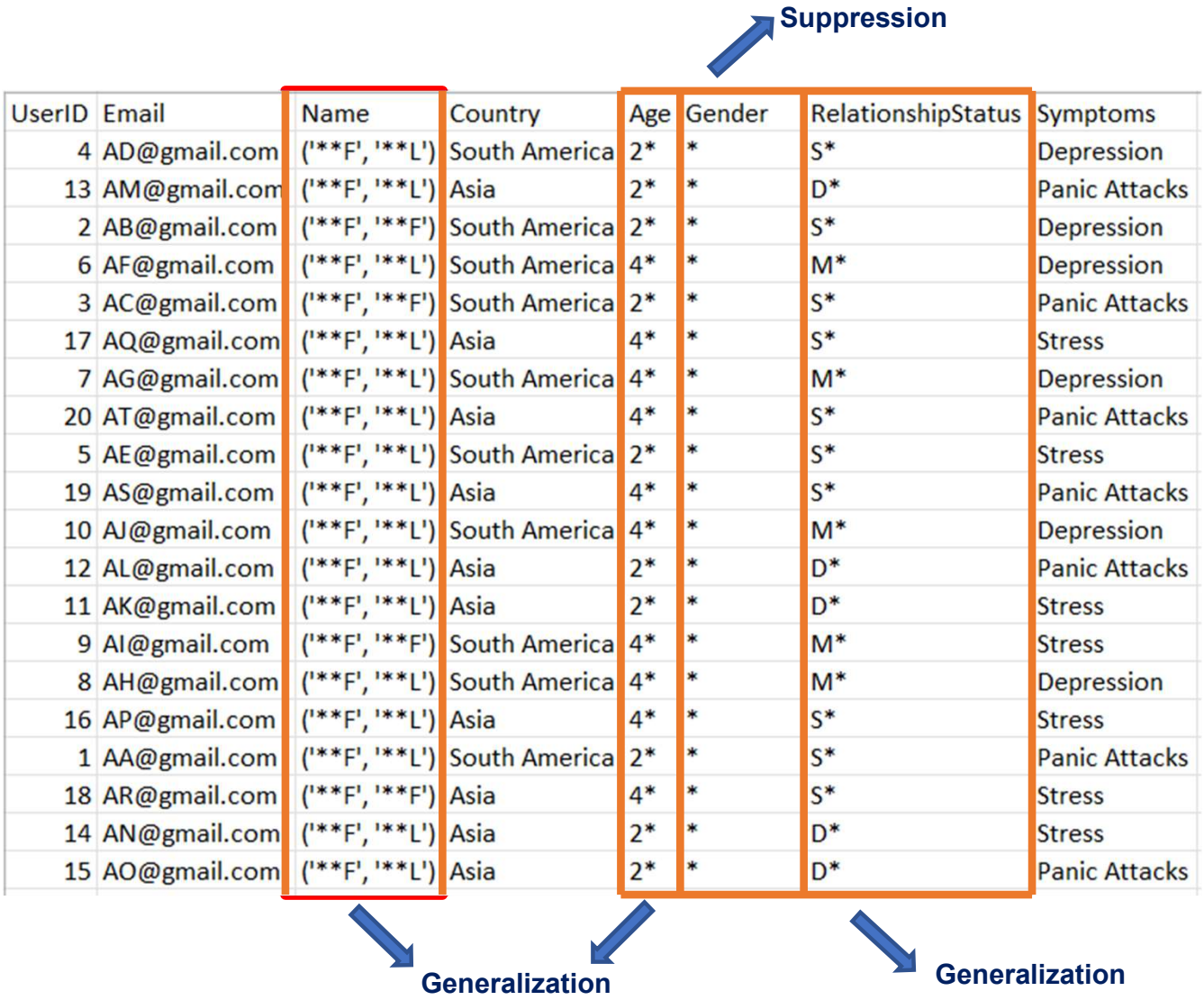
The Symptom Tracker is used to record symptoms of mental health over a period of time..

Original Data (Psychoeducation)

Key Attributes | Quasi-Identifier | Sensitive Attribute

UserID	Email	Name	Country	Age	Gender	RelationshipStatus	Symptoms
4	AD@gmail.com	John Wang	South America	22	Female	Single	Depression
13	AM@gmail.com	Sarah Lee	Asia	28	Male	Divorced	Panic Attacks
2	AB@gmail.com	Emily Lee	South America	29	Female	Single	Depression
6	AF@gmail.com	James Lee	South America	45	Male	Married	Depression
3	AC@gmail.com	John Wong	South America	27	Female	Single	Panic Attacks
17	AQ@gmail.com	Michael Chen	Asia	47	Female	Single	Stress
7	AG@gmail.com	Rachel Brown	South America	48	Male	Married	Depression
20	AT@gmail.com	Robert Lee	Asia	40	Female	Single	Panic Attacks
5	AE@gmail.com	Samantha Brown	South America	26	Male	Single	Stress
19	AS@gmail.com	Robert Davis	Asia	49	Female	Single	Panic Attacks
10	AJ@gmail.com	Emily Johnson	South America	41	Female	Married	Depression
12	AL@gmail.com	Samantha Lee	Asia	27	Male	Divorced	Panic Attacks
11	AK@gmail.com	John Kim	Asia	28	Male	Divorced	Stress
9	AI@gmail.com	James Kim	South America	40	Female	Married	Stress
8	AH@gmail.com	Sarah Brown	South America	43	Other	Married	Depression
16	AP@gmail.com	David Brown	Asia	42	Other	Single	Stress
1	AA@gmail.com	Jessica Kim	South America	20	Male	Single	Panic Attacks
18	AR@gmail.com	Michael Chen	Asia	44	Other	Single	Stress
14	AN@gmail.com	Robert Davis	Asia	29	Male	Divorced	Stress
15	AO@gmail.com	James Brown	Asia	27	Male	Divorced	Panic Attacks

Data Generalization & Suppression



UserID	Email	Name	Country	Age	Gender	RelationshipStatus	Symptoms
4	AD@gmail.com	('**F', '**L')	South America	2*	*	S*	Depression
13	AM@gmail.com	('**F', '**L')	Asia	2*	*	D*	Panic Attacks
2	AB@gmail.com	('**F', '**F')	South America	2*	*	S*	Depression
6	AF@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
3	AC@gmail.com	('**F', '**F')	South America	2*	*	S*	Panic Attacks
17	AQ@gmail.com	('**F', '**L')	Asia	4*	*	S*	Stress
7	AG@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
20	AT@gmail.com	('**F', '**L')	Asia	4*	*	S*	Panic Attacks
5	AE@gmail.com	('**F', '**L')	South America	2*	*	S*	Stress
19	AS@gmail.com	('**F', '**L')	Asia	4*	*	S*	Panic Attacks
10	AJ@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
12	AL@gmail.com	('**F', '**L')	Asia	2*	*	D*	Panic Attacks
11	AK@gmail.com	('**F', '**L')	Asia	2*	*	D*	Stress
9	AI@gmail.com	('**F', '**F')	South America	4*	*	M*	Stress
8	AH@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
16	AP@gmail.com	('**F', '**L')	Asia	4*	*	S*	Stress
1	AA@gmail.com	('**F', '**L')	South America	2*	*	S*	Panic Attacks
18	AR@gmail.com	('**F', '**F')	Asia	4*	*	S*	Stress
14	AN@gmail.com	('**F', '**L')	Asia	2*	*	D*	Stress
15	AO@gmail.com	('**F', '**L')	Asia	2*	*	D*	Panic Attacks

K-Anonymity Applied Table

UserID	Email	Name	Country	Age	Gender	RelationshipStatus	Symptoms
4	AD@gmail.com	('**F', '**L')	South America	2*	*	S*	Depressior
2	AB@gmail.com	('**F', '**F')	South America	2*	*	S*	Depressior
K=5 3	AC@gmail.com	('**F', '**L')	South America	2*	*	S*	Panic Attac
5	AE@gmail.com	('**F', '**F')	South America	2*	*	S*	Stress
1	AA@gmail.com	('**F', '**L')	South America	2*	*	S*	Panic Attac
13	AM@gmail.com	('**F', '**F')	Asia	2*	*	D*	Panic Attac
12	AL@gmail.com	('**F', '**L')	Asia	2*	*	D*	Panic Attac
11	AK@gmail.com	('**F', '**L')	Asia	2*	*	D*	Stress
14	AN@gmail.com	('**F', '**L')	Asia	2*	*	D*	Stress
15	AO@gmail.com	('**F', '**L')	Asia	2*	*	D*	Panic Attac
6	AF@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
7	AG@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
10	AJ@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
9	AI@gmail.com	('**F', '**L')	South America	4*	*	M*	Stress
8	AH@gmail.com	('**F', '**L')	South America	4*	*	M*	Depression
17	AQ@gmail.com	('**F', '**L')	Asia	4*	*	S*	Stress
20	AT@gmail.com	('**F', '**L')	Asia	4*	*	S*	Panic Attac
19	AS@gmail.com	('**F', '**L')	Asia	4*	*	S*	Panic Attac
16	AP@gmail.com	('**F', '**L')	Asia	4*	*	S*	Stress
18	AR@gmail.com	('**F', '**F')	Asia	4*	*	S*	Stress

Original Data (Symptom Tracker)

Key Attribute | Sensitive Attribute | Device Attribute

User ID	Symptom	Severity (BDI score)	Frequency	Notes	Timestamp
4	Depression	57	Daily	Lack of energy	20-02-2021 06:10
13	Panic Attacks	33	Once a week	Difficulty concentrating	03-01-2022 01:53
2	Depression	24	Once a week	Shaking	10-05-2020 02:52
6	Depression	57	Once every two weeks	Feeling overwhelmed	08-01-2020 21:01
3	Panic Attacks	21	Once a week	Feeling overwhelmed	13-07-2019 22:14
17	Stress	17	Once every two weeks	Shaking	05-06-2022 14:29
7	Depression	18	Once a week	Lack of energy	12-01-2022 07:27
20	Panic Attacks	63	Daily	Increased heart rate	30-04-2021 04:03
5	Stress	53	2-3 times per week	Sweating	30-08-2023 17:53
19	Panic Attacks	9	Daily	Difficulty concentrating	18-06-2018 11:51
10	Depression	43	Once a week	Increased heart rate	07-11-2021 13:33
12	Panic Attacks	30	Once every two weeks	Lack of energy	07-01-2023 06:37
11	Stress	16	Monthly	Feeling hopeless	04-02-2023 10:24
9	Stress	22	Daily	Lack of appetite	18-10-2022 17:13
8	Depression	1	Daily	Trouble sleeping	17-08-2023 13:47
16	Stress	41	Monthly	Feeling hopeless	15-01-2021 05:25
1	Panic Attacks	45	Once a week	Lack of energy	26-11-2023 03:37
18	Stress	43	Once every two weeks	Lack of energy	18-12-2023 16:33
14	Stress	26	Daily	Increased heart rate	08-01-2023 20:08
15	Panic Attacks	36	Once a week	Lack of appetite	29-04-2019 02:41

- Data Collected and Severity calculated based on BDI scoring strategy.
- This data is not available to the therapist.

BDI- Beck Depression Inventory

The Beck Depression Inventory (BDI) is a 21-item, self-report rating inventory that measures characteristic attitudes and symptoms of depression.

Source: Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961) An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

BDI-Questionnaire

Question Pattern:

1.

- ❑ 0 I do not feel sad.
- ❑ 1 I feel sad
- ❑ 2 I am sad all the time and I can't snap out of it.
- ❑ 3 I am so sad and unhappy that I can't stand it.

Scoring Pattern:
Total Score _____ Levels of Depression
1-10 _____These ups and downs are considered normal
11-16 _____Mild mood disturbance
17-20 _____Borderline clinical depression
21-30 _____Moderate depression
31-40 _____Severe depression
over 40 _____Extreme depression

Summary:

1. In total we have 21 questions and score ranging from [0-3]/Questions.
2. Total Score is in the range [0-63]

Source: https://www.ismanet.org/doctoryourspirit/pdfs/Beck-Depression-Inventory-BDI.pdf

Why BDI-(Beck Depression Inventory)?

To ensure the severity data captured is safe and trustworthy.

Follows a procedural guideline that is clinically grounded.

Sometimes the patients cannot access their mental state accurately unless they follow some guideline.

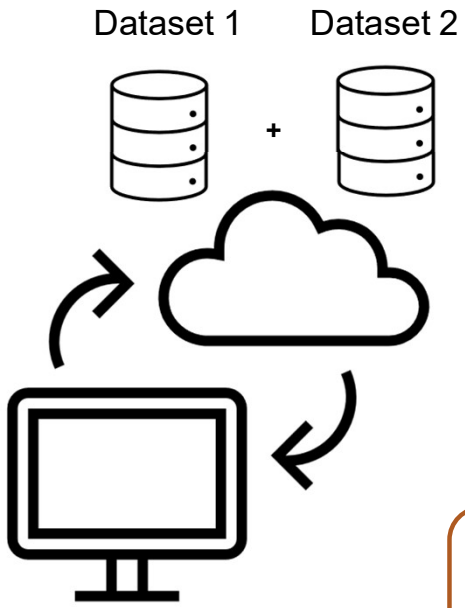
Our First Enhanced Feature

Psychoeducation

+

symptom tracker

How it works?



We will perform a SQL Query to extract data for the final table.

Inner Join

```
SELECT d2.USER_ID, d2.Severity, d2.Frequency, d2.Notes,
d2.Timestamp FROM DATASET1 d1 INNER JOIN DATASET2 d2 ON
d1.USER_ID = d2.USER_ID;
```

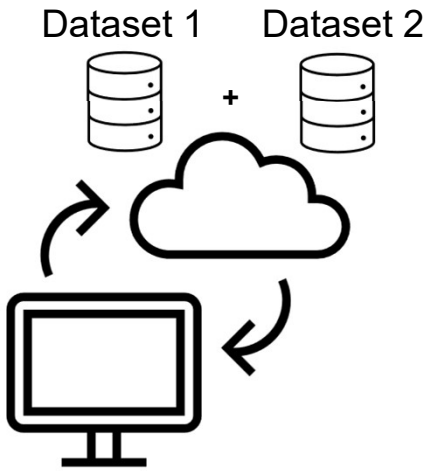
Modified Table

Modified Data

User ID	Symptom	Severity	Frequency	Notes	Timestamp
4	Depression	Extreme depression	Daily	Lack of energy	Winter 2021
2	Depression	Moderate depression	Once a week	Shaking	Spring 2020
6	Depression	Extreme depression	Once every two weeks	Feeling overwhelmed	Winter 2020
7	Depression	Borderline clinical depression	Once a week	Lack of energy	Winter 2022
10	Depression	Extreme depression	Once a week	Increased heart rate	Fall 2021
8	Depression	These ups and downs are considered normal	Daily	Trouble sleeping	Summer 2023
13	Panic Attacks	Severe depression	Once a week	Difficulty concentrating	Winter 2022
3	Panic Attacks	Moderate depression	Once a week	Feeling overwhelmed	Summer 2019
20	Panic Attacks	Extreme depression	Daily	Increased heart rate	Spring 2021
19	Panic Attacks	These ups and downs are considered normal	Daily	Difficulty concentrating	Summer 2018
12	Panic Attacks	Moderate depression	Once every two weeks	Lack of energy	Winter 2023
1	Panic Attacks	Extreme depression	Once a week	Lack of energy	Fall 2023
15	Panic Attacks	Severe depression	Once a week	Lack of appetite	Spring 2019
17	Stress	Borderline clinical depression	Once every two weeks	Shaking	Summer 2022
5	Stress	Extreme depression	2-3 times per week	Sweating	Summer 2023
11	Stress	Mild mood disturbance	Monthly	Feeling hopeless	Winter 2023
9	Stress	Moderate depression	Daily	Lack of appetite	Fall 2022
16	Stress	Extreme depression	Monthly	Feeling hopeless	Winter 2021
18	Stress	Extreme depression	Once every two weeks	Lack of energy	Winter 2023
14	Stress	Moderate depression	Daily	Increased heart rate	Winter 2023

This final data will be available to the therapist.

Access Control (Discretionary Control-DAC)

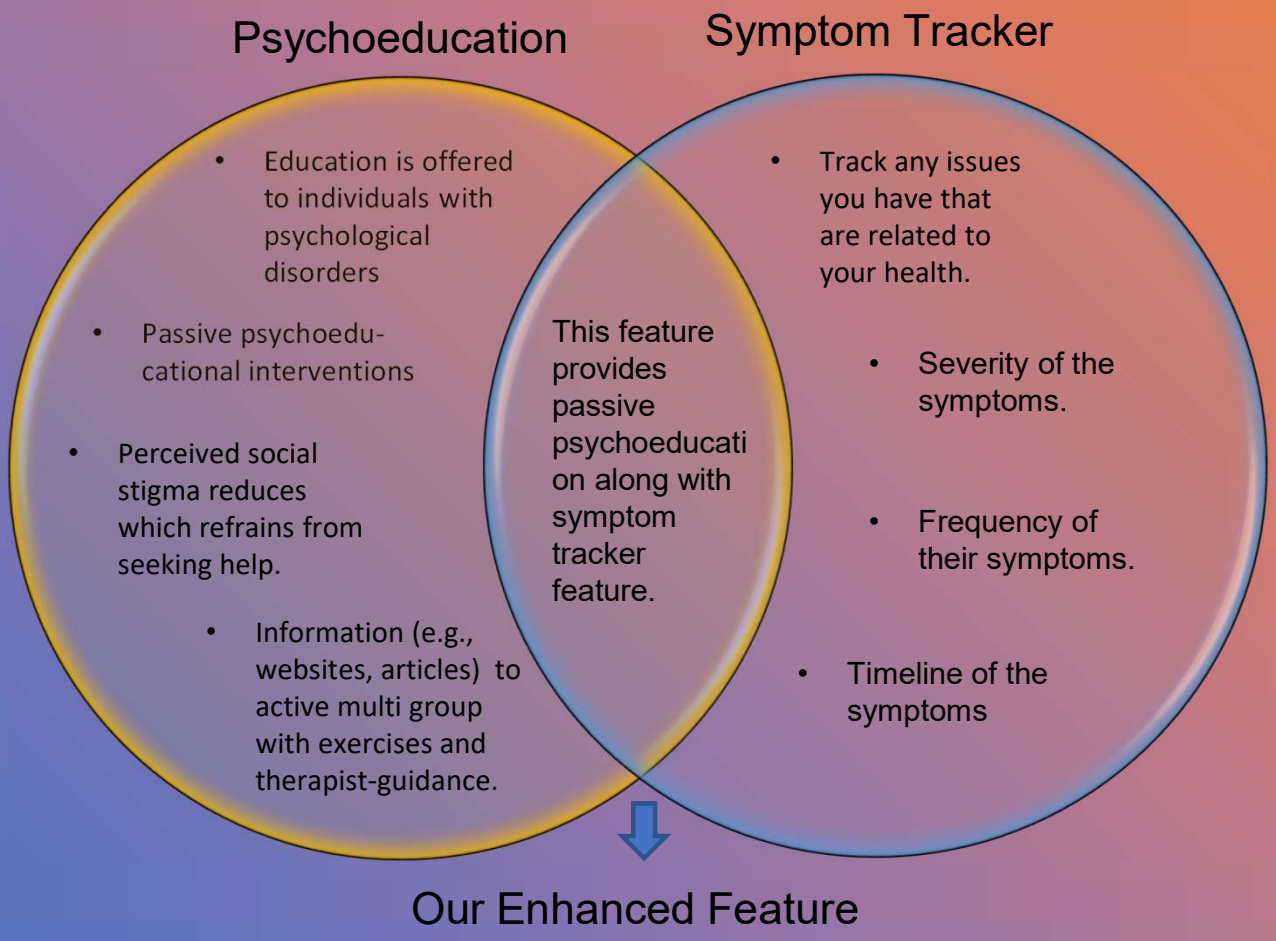


DAC: VIEW

```
CREATE VIEW New_Dataset
AS SELECT UserID, Severity, Symptoms
        FROM Dataset2
        WHERE Symptoms="Stress"
```

Entities			
Subjects	User	Admin/App Developer	Therapist
Objects	All Datasets	Limited view access to dataset 1 (e.g., Email ID for Payment)	View access to dataset 2. (To provide health information)
Access Right	Read, Write, Delete	Read	Read, Search

Overview of Our Enhanced Feature



Our 2nd Enhanced Feature

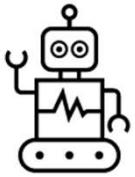
Crisis Management during Chat

Overview

Crisis Management

- Many a times therapists are not available during night time.
- Many patients might require to connect to their therapists.
- Only available choice is AI assistant chatbot in their mental health app.

Process/Medical Knowledge Guidelines



When using AI assistant chatbots; It can record sensitive information provided by the user (e.g., name, relationship status, symptoms).

This process is very unsafe and can lead to privacy violations.

How we resolve?

Solution : Instantiating process knowledge or some kind of clinical guidelines into the system.

In this case we use BDI questioners.

AI Chatbot

How may I help you today?

I am not feeling well today.

AI Chatbot

Please provide a more detailed description of your condition.

Hi doctor, my name is *Ekta*. When I was at my *workplace*, I was *body shamed* by my colleagues *Rohan* and *Ram*. It was so *upsetting* to hear that. I don't feel good at all.

(Contains key attributes and sensitive Information)

Our System

How may I help you today?

I am not feeling well today.

Our System

Please answer the following questions from Beck Depression Inventory:

1. Do You often feel sad or down?
2. Do you feel guilty or worthless?
3. Have you lost interest in doing things?

(Safe questions to ask and gather user's info)

Differential Privacy :

- When the user fills out all the questions, the BDI-score is computed at the backend.
- However, the score still poses a privacy threat about the patient's mental state.

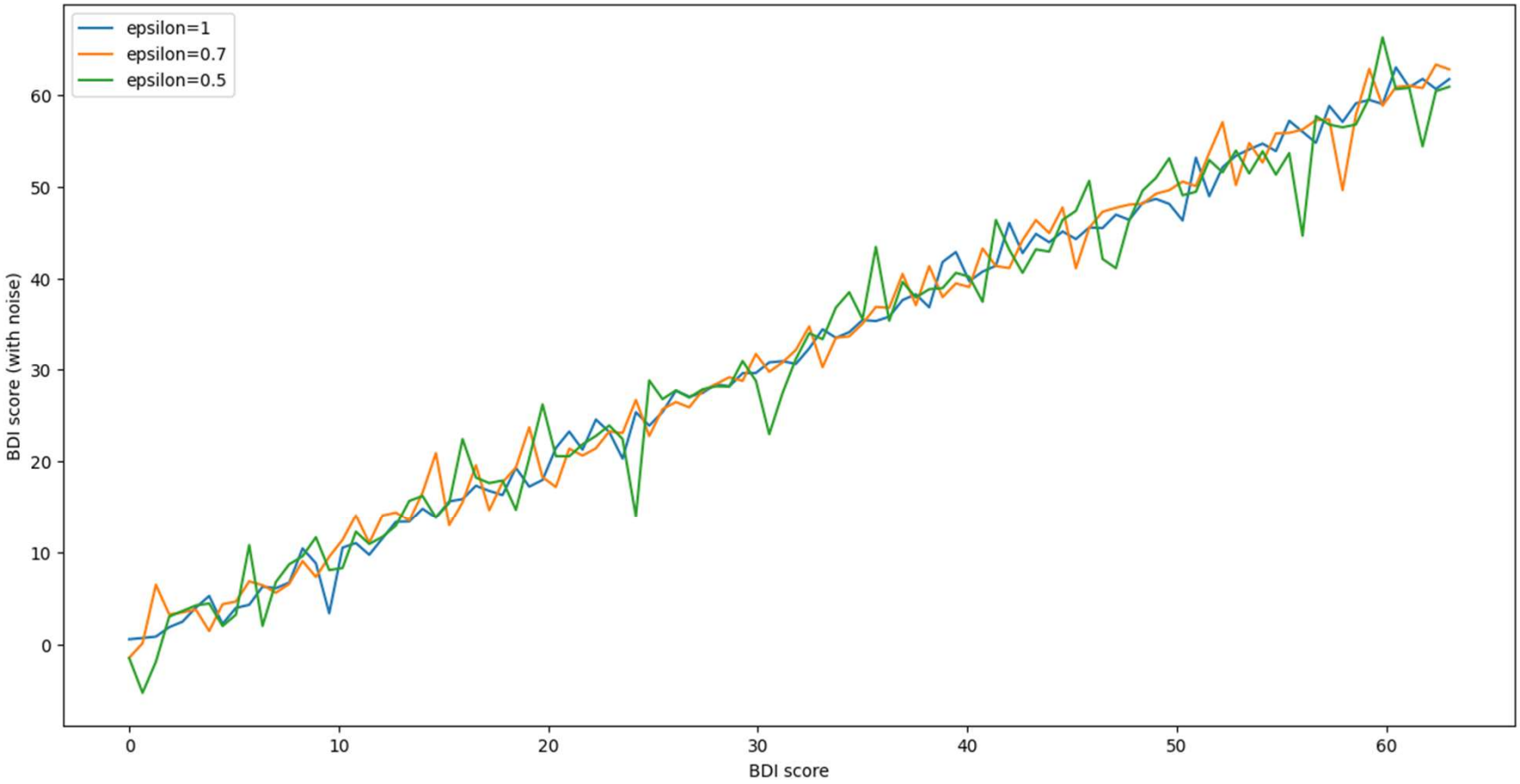
How we resolve?

Solution : Using laplace technique, noise is applied to each score for a specific range of BDI values.

The parameter Epsilon →{0.5,0.7,1}

- As an app developer we also won't have the exact BDI score to make a judgement on patient's confidential mental state.

Result


epsilon=1
epsilon=0.7
epsilon=0.5
BDI score (with noise)
BDI score
0
10
20
30
40
50
60

Unsafe

```
| Welcome to the mental health assessment app.
| How can I help you today?

| I am not feeling well today.

| Based on your input, the probability of a decline in mental health i
s: 77.93%

| Mental health decline probability exceeds 70%!
```

```
| Please provide a more detailed description of your condition.
| Hi doctor, my name is Ekta. When I was at my workplace, I was bodysha
med  by my colleagues Rohan and Ram. It was so upsetting to hear that. I
dont feel good at all. That is why I don't feel good.
```

Safe

```
| Welcome to the mental health assessment app.
| How can I help you today?

| I am not feeling well today.

| Based on your input, the probability of a decline in mental health is: 77.93%

| Mental health decline probability exceeds 70%!
```

```
Please answer the following questions from the Beck Depression Inventory (BDI):
|  Do you often feel sad or down?
| Your response (yes or no): yes
|  Do you feel guilty or worthless?
| Your response (yes or no): yes
| Have you lost interest in things you used to enjoy?
| Your response (yes or no): yes
bdi_score: [3.77929881]
Your BDI score is: [3.4090338]
```

```
| Your BDI score suggests severe depression. Please seek immediate professional help.
| Notification Set
```

Overview of Crisis Management

Based on the computed BDI score an **alarm/notification** will be generated to their therapist for immediate assistance.

DEMO

THREAT MODELING

Source: LINDDUN