

Digital Biomarker Applications in Public Health

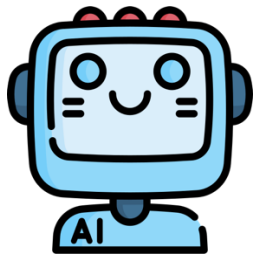
Shravan Aras

The University of Arizona

Center for Biomedical Informatics and Biostatistics

2025-6-12

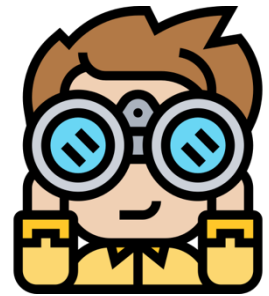
Let's Get Started



AI Assisted



Code Available



Study Data

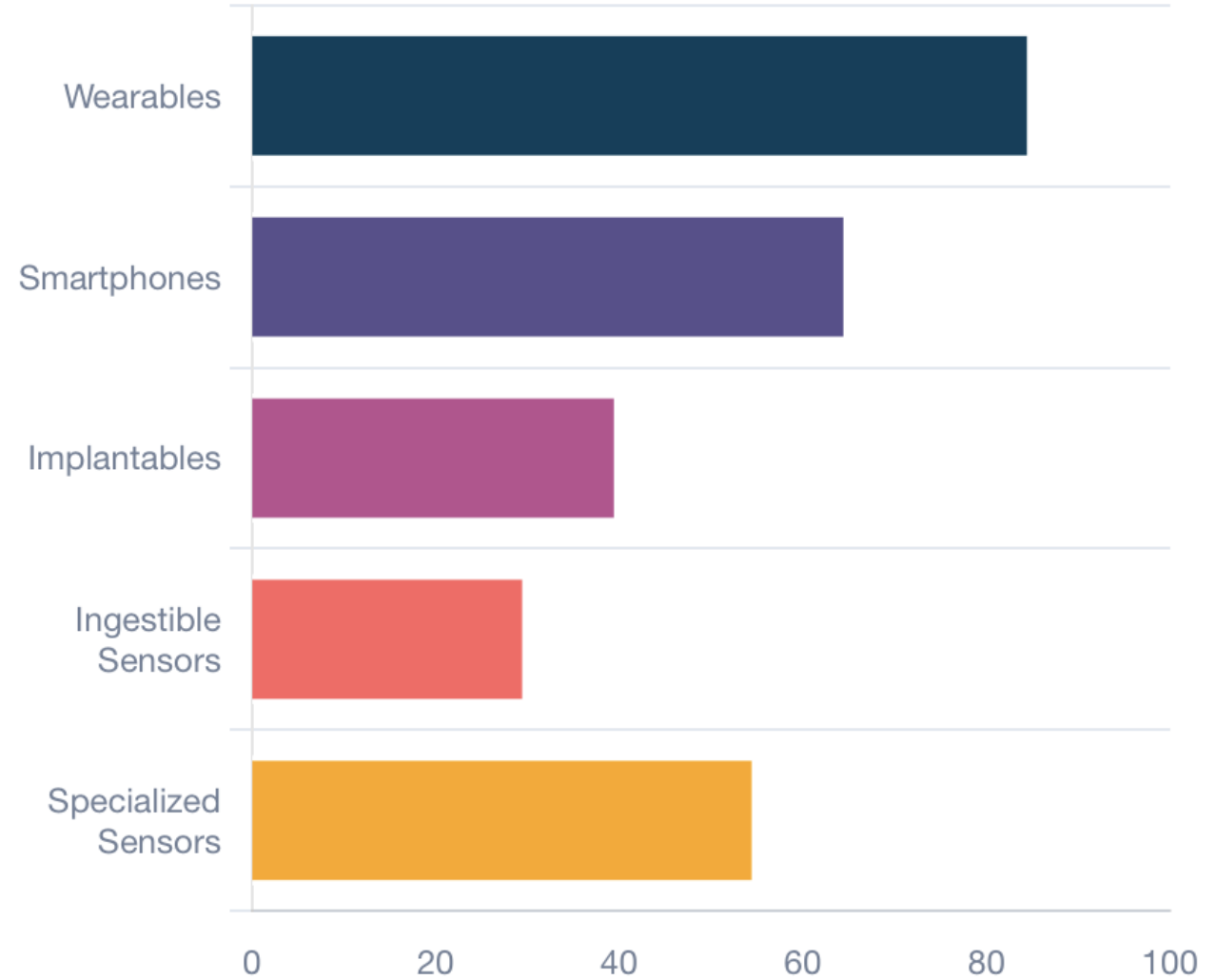
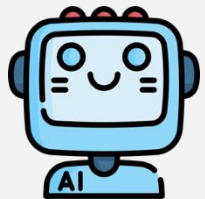


Digital Biomarkers?

One Such Definition

1. Digital biomarkers are quantifiable data points.
2. Collected by digital devices like wearables, smartphones, or implantable sensors.
3. They provide a continuous view of our health.
4. They can be collected remotely.

Different Ways to Collect Them



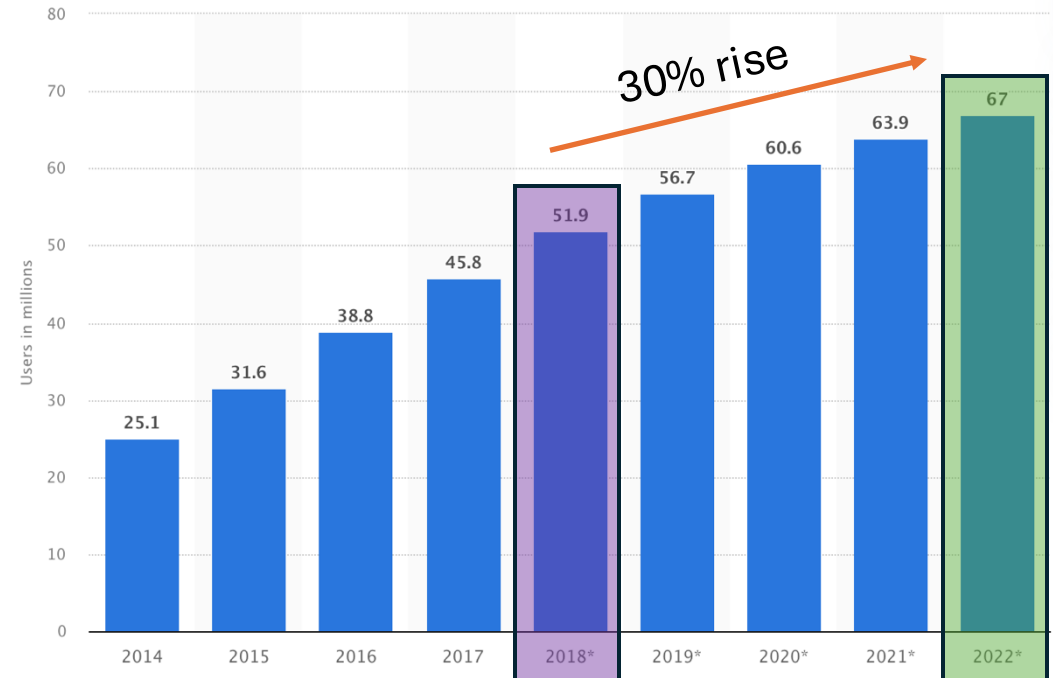


We are going to focus on digital markers coming from wearable devices today!

Why Now?

30% increase in wearable use in last 5 years

Wide range of affordability -
\$100 - \$400



Number of users in millions using wearable fitness devices in U.S.A.

Source : Statista

30% increase in wearable use in last 5 years

Wide range of affordability -
\$100 - \$400

More reliable and
fashionable



Number of users in millions using wearable fitness devices in U.S.A.

Source : Statista

30% increase in wearable use in last 5 years

Wide range of affordability -
\$100 - \$400

More reliable and
fashionable

Increased ease of
operations

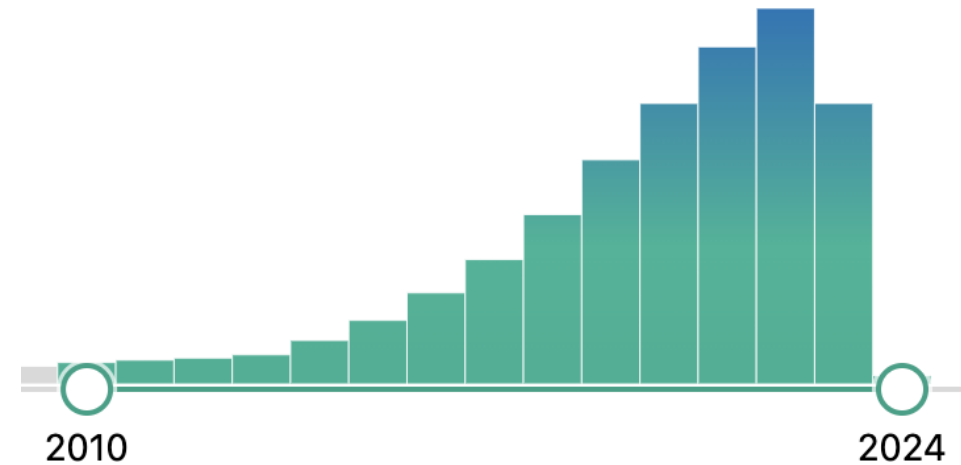


Number of users in millions using wearable fitness devices in U.S.A.

Source : Statista

More than 7,000 publications in last 3 years

34% increase in PubMed publications that involve wearable sensors

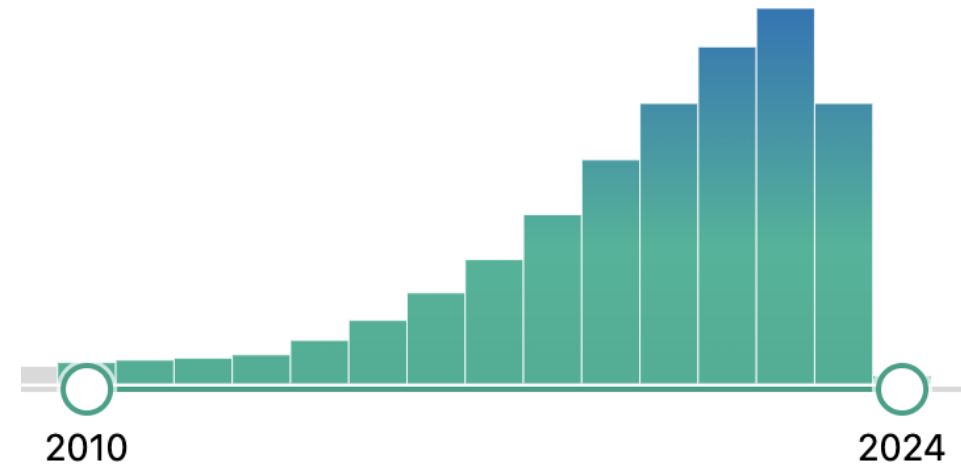


Publications indexed for "wearable sensors" on pubmed

More than 7,000 publications in last 3 years

34% increase in PubMed publications that involve wearable sensors

Over \$58M NIH proposals with wearable sensors funded in last 3 years



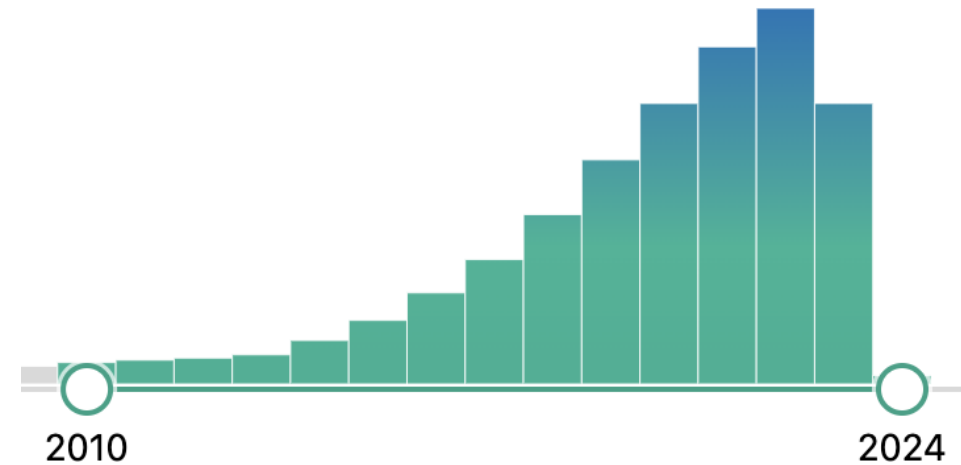
Publications indexed for "wearable sensors" on pubmed

More than 7,000 publications in last 3 years

34% increase in PubMed publications that involve wearable sensors

Over \$58M NIH proposals with wearable sensors funded in last 3 years.

More data points for Idiographic (person specific) and nomothetic (within group)



Publications indexed for "wearable sensors" on pubmed

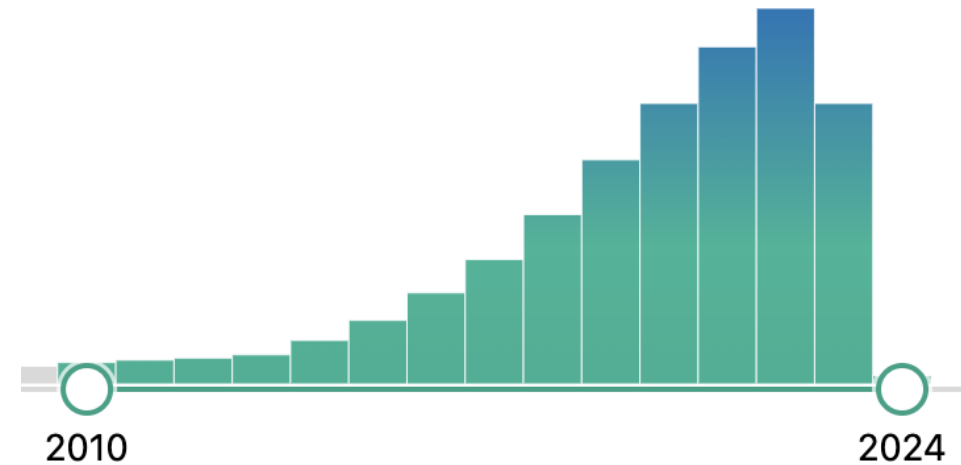
More than 7,000 publications in last 3 years

34% increase in PubMed publications that involve wearable sensors

Over \$58M NIH proposals with wearable sensors funded in last 3 years.

More data points for Idiographic (person specific) and nomothetic (within group)

Inferred vs measured (e.g., physical activity, stress, sleep)



Publications indexed for "wearable sensors" on pubmed

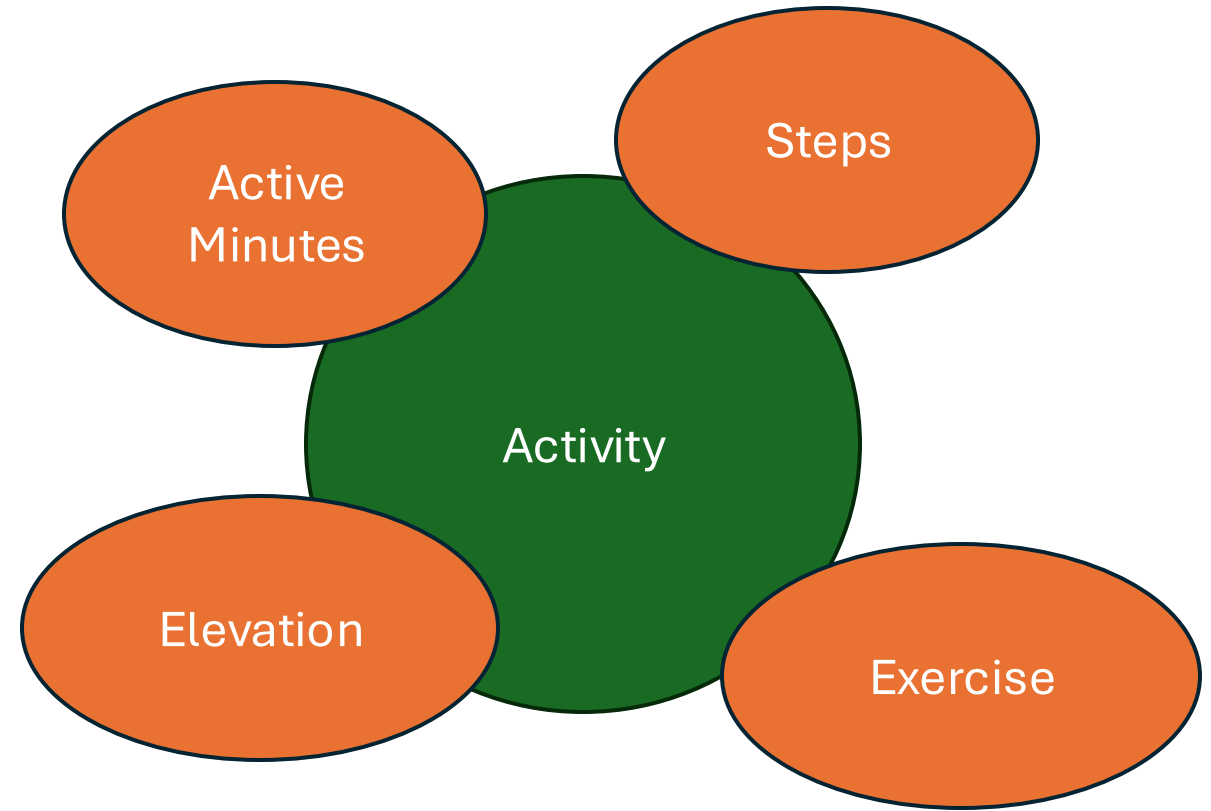
Step 1 – Understanding what to measure



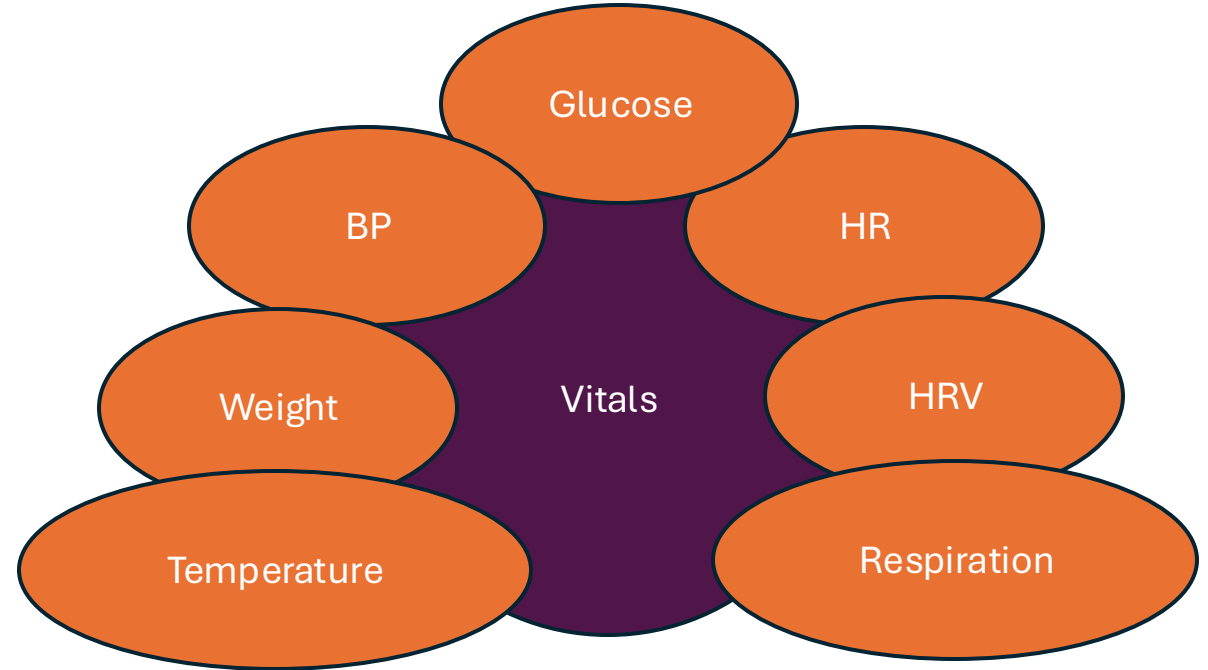
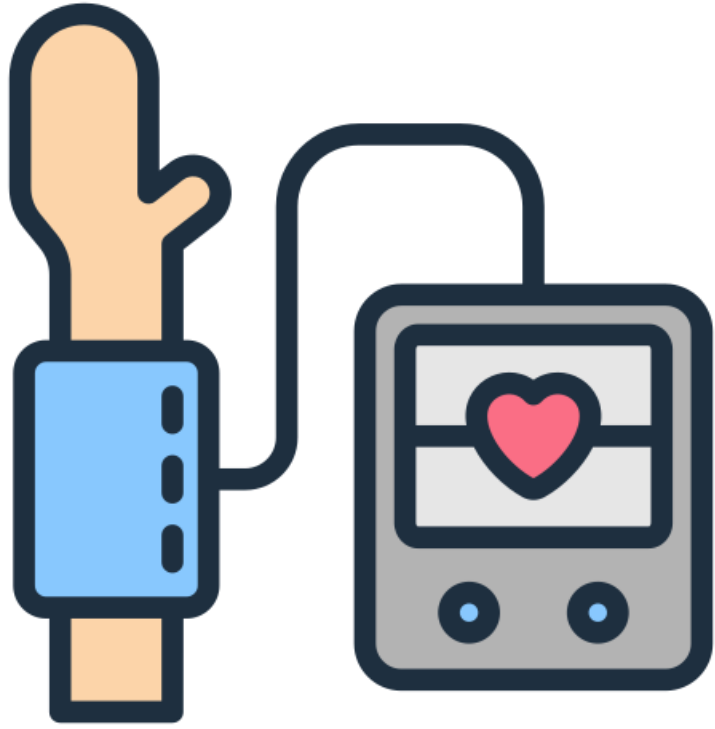
Can feel like a wine menu!

Way too many options out
there

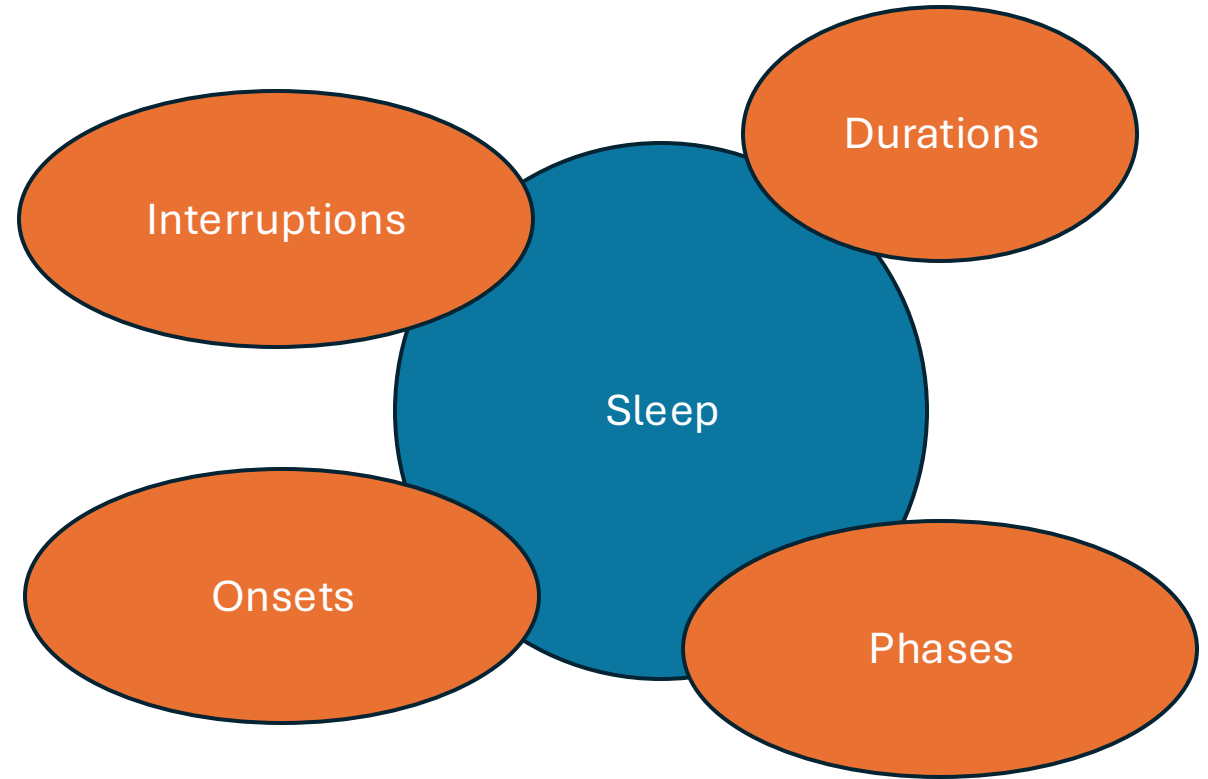
Similar but not the same



What do we want to measure?



What do we want to measure?

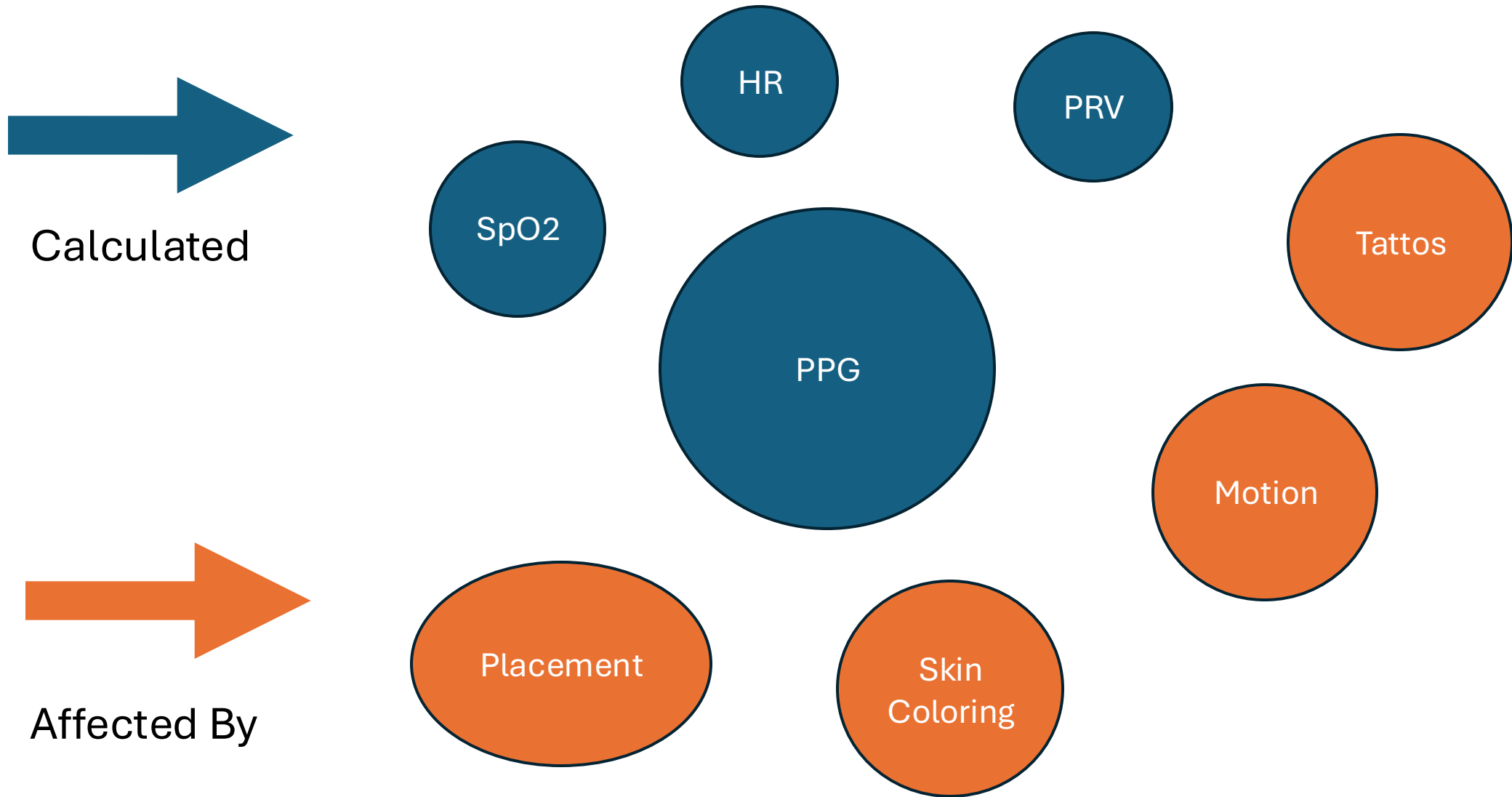


What do we want to measure?

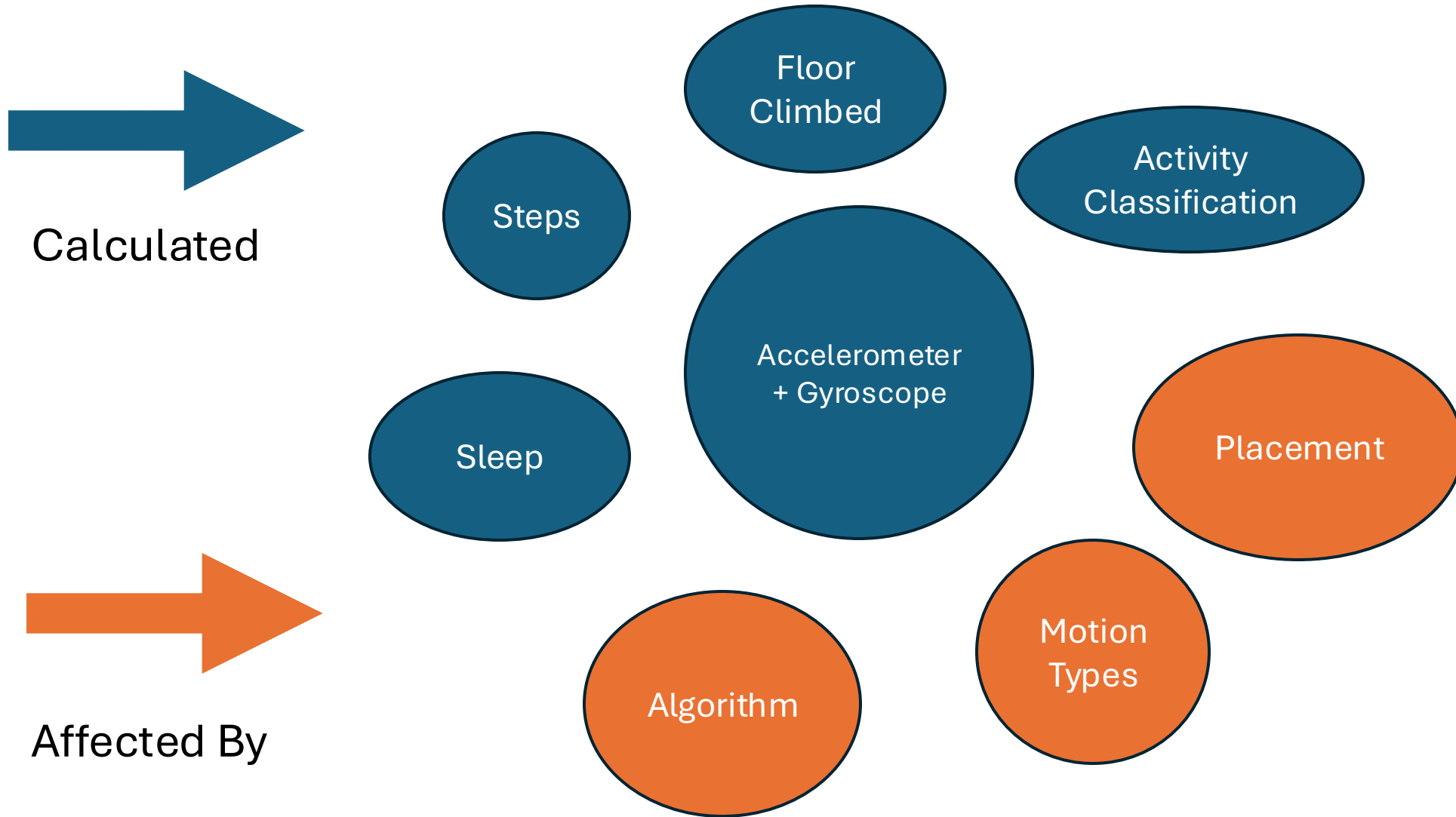
But Wait!!

Not all that you see is calculated, some is estimated!

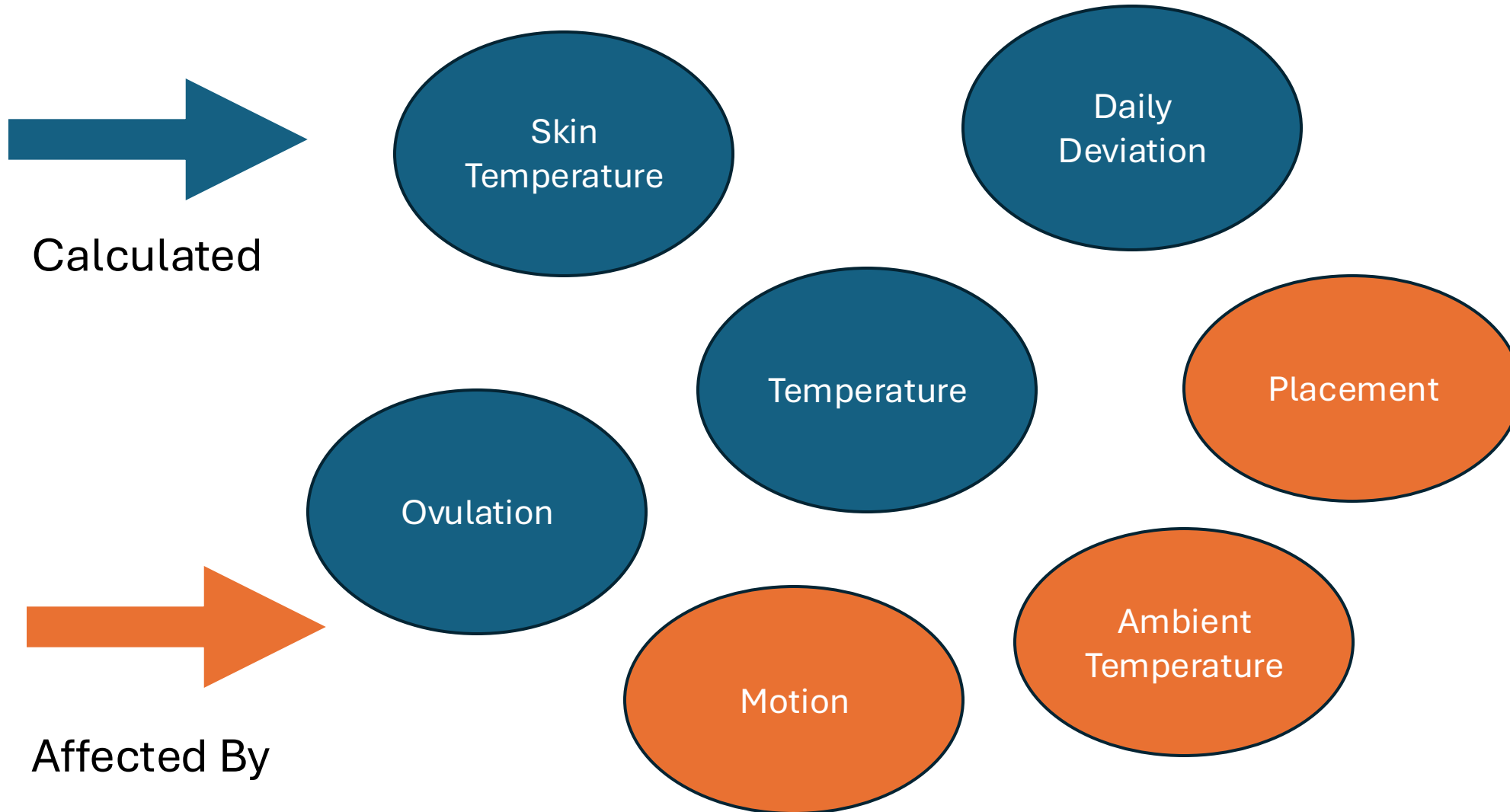
Calculated From PPG Sensor



Calculated From Accelerometer



Calculated From Temperature Sensor



Then what is estimated?

- VO2 Max

Then what is estimated?

- VO2 Max
- Sleep Phases – REM, Deep, etc.

Then what is estimated?

- VO2 Max
- Sleep Phases – REM, Deep, etc.
- Optical Blood Pressure

Then what is estimated?

- VO2 Max
- Sleep Phases – REM, Deep, etc.
- Optical Blood Pressure
- Stress Metric / Body Battery / Recovery Index

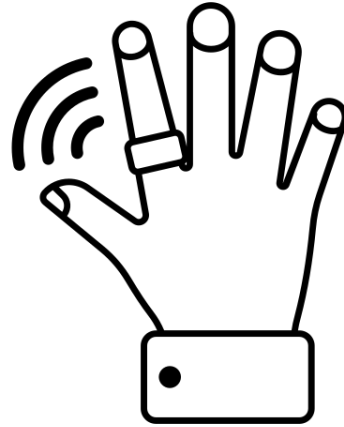
Step 2 – Pick a device

Bring out the roller deck of cards!

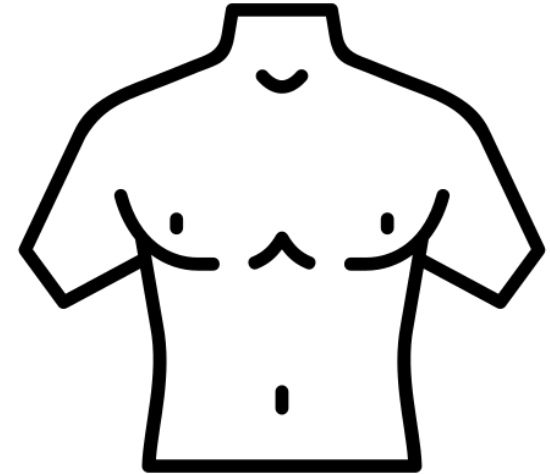
Positioning (Common Ones)



Wrist



Finger



Torso

Choosing a wearable

Measurement	Fitbit (\$)	Apple Watch (\$\$)	Oura (\$\$)	Empatica (\$\$\$)
Daily Activity	Standard	Standard	Standard	Raw Acceleration
Exercise Sessions	Standard	Standard	Standard	Basic
Intraday HR	~ 5 seconds	~ 5 minutes	~ 5 minutes	~ 1minute
Resting HR	Yes	Yes	Yes	Can Calculate
Sleep Duration	Yes	Yes	Yes	Yes
Sleep Phases	Yes	Yes	Yes	No
HRV	Day / Night (summary / intra)	Day / Night (summary / intra)	Night (5 mins)	1 min (RMSSD)
IBI	No	No	Yes	Yes
Raw Acceleration	No	Yes	No	Yes
Temperature	Summary	Summary / ~5s (SensorKit)	Summary / 5m(Research)	1 min
spo2	Yes	Yes	Yes	?

Activity, Exercise, Sleep

Measurement	Fitbit (\$)	Apple Watch (\$\$)	Oura (\$\$)	Empatica (\$\$\$)
Daily Activity	Standard	Standard	Standard	Raw Acceleration
Exercise Sessions	Standard	Standard	Standard	Basic
Intraday HR	~ 5 sec	Everyone supports the basics		~ 1minute
Resting HR	Yes			Can Calculate
Sleep Duration	Yes	Yes	Yes	Yes
Sleep Phases	Yes	Yes	Yes	No
HRV	Day / Night (summary / intra)	Day / Night (summary / intra)	Night (5 mins)	1 min (RMSSD)
IBI	No	No	Yes	Yes
Raw Acceleration	No	Yes	No	Yes
Temperature	Summary	Summary / ~5s (SensorKit)	Summary / 5m(Research)	1 min
spo2	Yes	Yes	Yes	?

Intraday Heart Rate (IHR)

Measurement	Fitbit (\$)	Apple Watch (\$\$)	Oura (\$\$)	Empatica (\$\$\$)
Daily Activity	Standard	Standard	Standard	Raw Acceleration
Exercise Sessions	Standard	Standard	Standard	Basic
Intraday HR	~ 5 seconds	~ 5 minutes	~ 5 minutes	~ 1minute
Resting HR	Yes	Yes	Yes	Can Calculate
Sleep Duration	Yes	Yes	Yes	Yes
Sleep Phases	Yes	Yes	Yes	No
HRV	Day / Night (summary / intra)	Day / Night (summary / intra)	Night (5 mins)	1 min (RMSSD)
IBI	No	No	Yes	Yes
Raw Acceleration	No	Yes	No	Yes
Temperature	Summary	Summary / ~5s (SensorKit)	Summary / 5m(Research)	1 min
spo2	Yes	Yes	Yes	?

Intraday HR support is variable

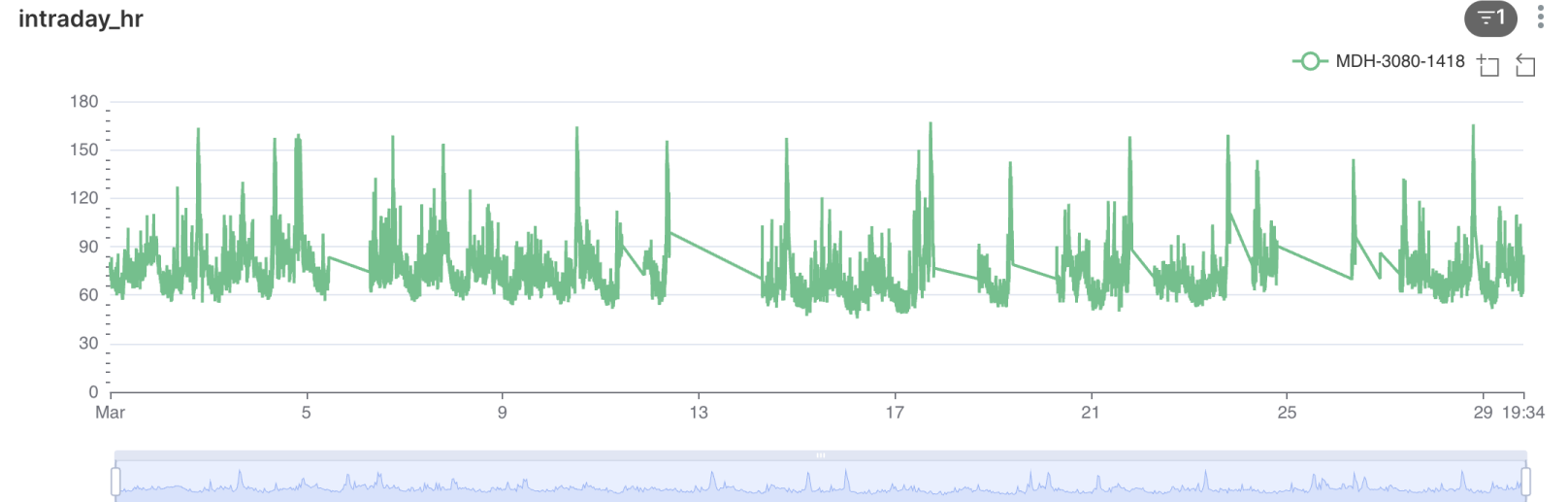
Some devices offer something special

Measurement	Fitbit (\$)	Apple Watch (\$\$)	Oura (\$\$)	Empatica (\$\$\$)
Daily Activity	Standard	Standard	Standard	Raw Acceleration
Exercise Sessions	Standard	Standard	Standard	Basic
Intraday HR	~ 5 seconds	~ 5 minutes	~ 5 minutes	~ 1minute
Resting HR	Yes	Yes	Yes	Can Calculate
Sleep Duration	Yes	Yes	Yes	Yes
Sleep Phases	Yes	Special support		No
HRV	Day / (summary / intra)			1 min (RMSSD)
IBI	No	No	Yes	Yes
Raw Acceleration	No	Yes	No	Yes
Temperature	Summary	Summary / ~5s (SensorKit)	Summary / 5m(Research)	1 min
spo2	Yes	Yes	Yes	?

Fitbit can give more than just summaries

Researchers can request access to Fitbit **intraday data** to unlock a wide range of high frequency metrics

	type
1	activities-heart
2	activities-distance
3	activities-floors
4	activities-steps
5	activities-elevation
6	br
7	activities-calories
8	spo2
9	hrv



Intraday day HR values for a participant

Get in-depth motion data with SensorKit

- **Multiple high frequency metrics**

- Acceleration
- Rotation Rate
- Light Sensor
- Temperature

- **Acceleration & Rotation**

- About 100Hz sampling rate
- Does affect battery – almost half

Does impact battery life

```
{  
  "timestamp": 140941.628781,  
  "acceleration": {  
    "x": -0.30615234375,  
    "y": 0.634033203125,  
    "z": 0.739990234375  
  },  
  "startDate": "2023-08-28T06:43:57-0700",  
  "identifier": 0  
}
```

What about BP, Glucose, Weight, etc.?

- Most BLE / BT smart devices connect Google Fit & Apple Health Kit
- Connecting them can be finicky → right permissions
- Syncing can be a challenge
- Good study material, instructions and monitoring can help!

	type	startdate	date	value	units	sourceidentifier	sourcename	source
1	BloodGlucose	2023-03-12 18:28:25.000	2023-03-12 18:28:25.000	132	mg/dL	com.dexcom.G6	Dexcom G6	17375
2	BloodGlucose	2023-03-12 18:33:25.000	2023-03-12 18:33:25.000	130	mg/dL	com.dexcom.G6	Dexcom G6	17375
3	BloodGlucose	2023-03-12 18:38:25.000	2023-03-12 18:38:25.000	125	mg/dL	com.dexcom.G6	Dexcom G6	17375
4	BloodGlucose	2023-03-12 18:43:25.000	2023-03-12 18:43:25.000	112	mg/dL	com.dexcom.G6	Dexcom G6	17375
5	BloodGlucose	2023-03-12 18:48:25.000	2023-03-12 18:48:25.000	105	mg/dL	com.dexcom.G6	Dexcom G6	17375
6	BloodGlucose	2023-03-12 18:53:25.000	2023-03-12 18:53:25.000	98	mg/dL	com.dexcom.G6	Dexcom G6	17375

Example data from Dexcom G6 CGM from one of our studies.

Gemini Deep Research – Learn about devices

A Universe of Health Insights

Fitbit provides a holistic view of well-being, moving far beyond simple fitness tracking. The API grants access to a rich dataset spanning multiple categories, empowering applications in personal wellness, academic research, and preventative health. This structured access allows for the creation of sophisticated tools that can correlate lifestyle factors with health outcomes.

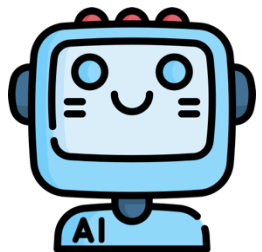
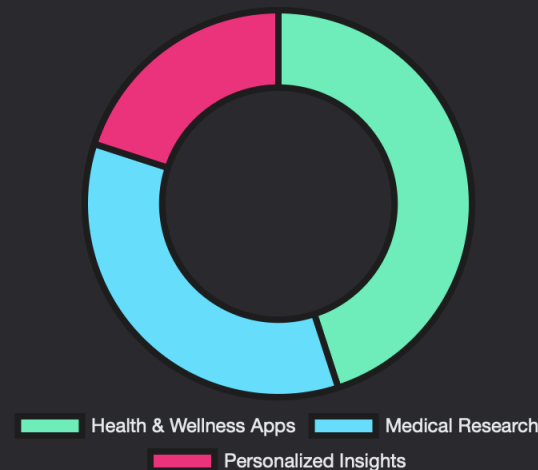
6

Core Data Categories

Activity, Sleep, Heart Rate, Body Composition, Nutrition, and Advanced Vitals.

Primary API Use Cases

The API is a significant platform for a diverse range of applications, with a strong emphasis on research.



Step 3 – Connecting to Devices

Knock Knock, can I get some data from you?

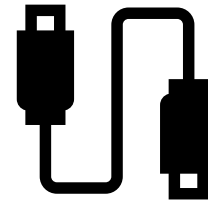
To Code or Not To



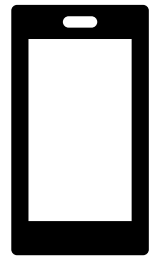
API
(JSON)



Data Dump
(csv, zip, parquet, JSON)



Plug it in

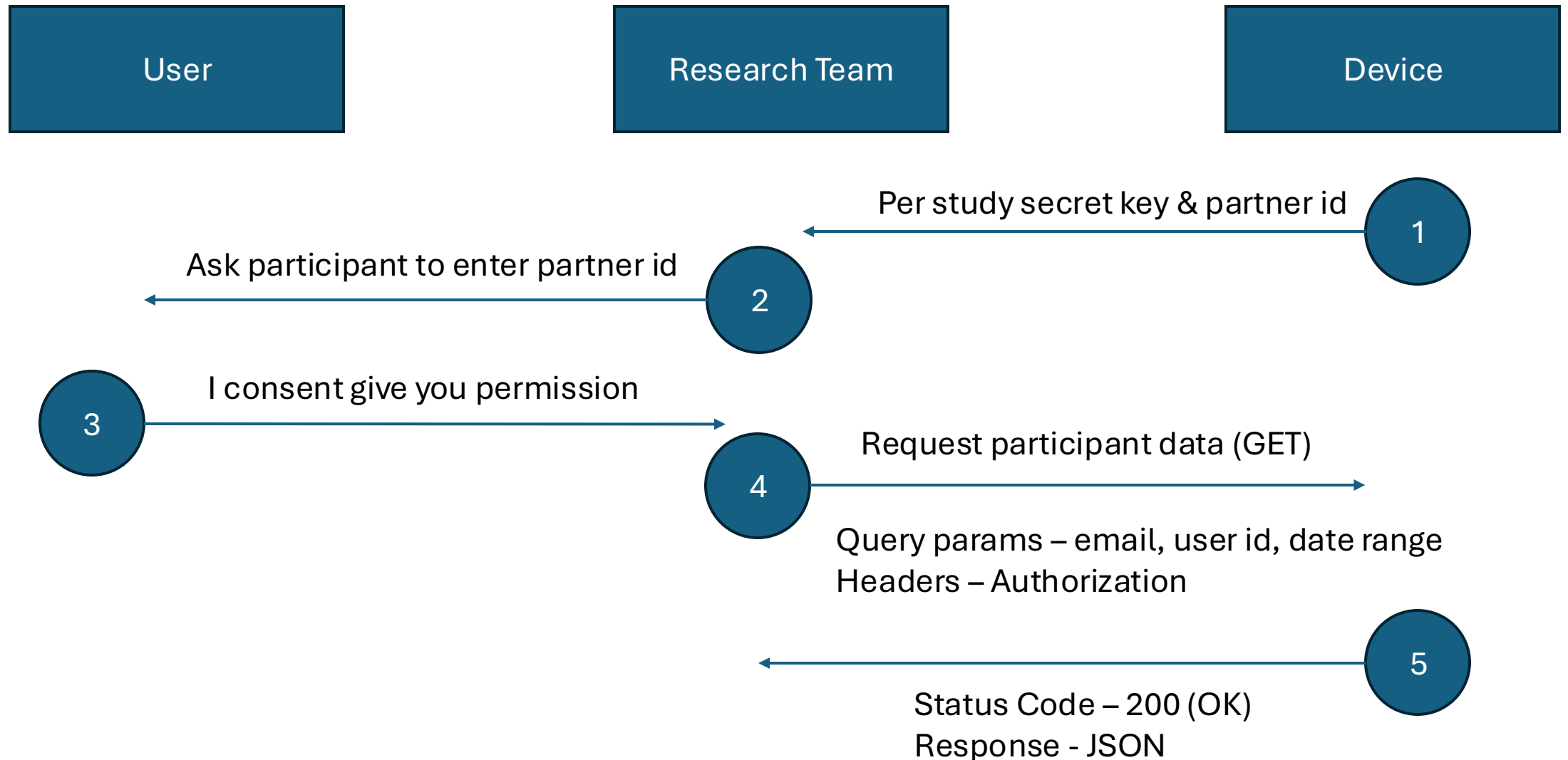


3rd Party
Providers

You can use AI to generate most of the code SO

Let's understand some of the fundamental patterns that you can ask AI to code up

Types of API access – secret key, partner id



Postman

HTTP UH API / getMetrics

GET https://partner.ultrahuman.com/api/v1/metrics?email=shravanaras@arizona.edu&date=2025-06-02 Send

Params Authorization Headers (9) Body Pre-request Script Tests Settings Cookies

Query Params

<input checked="" type="checkbox"/>	Key	Value	Description	...	Bulk Edit
<input checked="" type="checkbox"/>	email	shravanaras@arizona.edu	Email address of the user		
<input checked="" type="checkbox"/>	date	2025-06-02	yyyy-mm-dd		
	Key	Value	Description		

Body Cookies Headers (16) Test Results Status: 200 OK Time: 766 ms Size: 29.93 KB Save as example

Pretty Raw Preview Visualize JSON

```
1 {
2   "data": {
3     "metric_data": [
4       {
5         "type": "hr",
6         "object": {
7           "day_start_timestamp": 1748847600,
8           "title": "Heart Rate",
9           "values": [
10            {
11              "value": 78,
12              "timestamp": 1748876108
13            },
14            {
15              "value": 66,
16              "timestamp": 1748876403
17            },
18            {
19              "value": 81,
20              "timestamp": 1748876688
21            }
22          ]
23        }
24      }
25    ]
26  }
```

Postbot Runner Start Proxy Cookies Vault Trash



Generate Code

UH API / getMetrics

GET https://partner.ultrahuman.com/api/v1/metrics?email=shravanaras@arizona.edu&date=2025-06-02

Send

Params Authorization Headers (9) Body Pre-request Script Tests Settings Cookies

Query Params

<input checked="" type="checkbox"/> Key	Value	Description	...	Bulk Edit
<input checked="" type="checkbox"/> email	shravanaras@arizona.edu	Email address of the user		
<input checked="" type="checkbox"/> date	2025-06-02	yyyy-mm-dd		
Key	Value	Description		

Body Cookies Headers (16) Test Results

Status: 200 OK Time: 766 ms Size: 29.93 KB Save as example

Pretty Raw Preview Visualize JSON

```
1 {
2   "data": {
3     "metric_data": [
4       {
5         "type": "hr",
6         "object": {
7           "day_start_timestamp": 1748847600,
8           "title": "Heart Rate",
9           "values": [
10            {
11              "value": 78,
12              "timestamp": 1748876108
13            },
14            {
15              "value": 66,
16              "timestamp": 1748876403
17            },
18            {
19              "value": 81,
20              "timestamp": 1748876688
21            }
22          ]
6         }
7       }
8     ]
9   }
10 }
```

Code snippet

cURL

- C# - HttpClient
- C# - RestSharp
- ☒ cURL
- Dart - dio
- Dart - http
- Go - Native
- HTTP
- Java - OkHttp
- Java - Unirest
- JavaScript - Fetch
- JavaScript - jQuery
- JavaScript - XHR
- Kotlin - Okhttp
- C - libcurl
- NodeJs - Axios
- NodeJs - Native
- NodeJs - Request
- NodeJs - Unirest
- Objective-C - NSURLSession
- OCaml - Cohttp
- PHP - cURL
- PHP - Guzzle
- PHP - HTTP_Request2
- PHP - pecl_http
- PowerShell - RestMethod
- Python - http.client
- Python - Requests
- R - httr
- R - RCurl

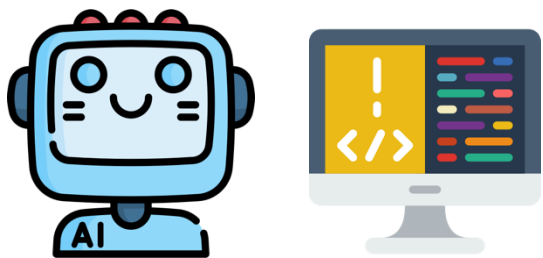
Postbot Runner Start Proxy Cookies Vault



Using your favorite LLM – System Prompt

You are a Python programming assistant. Provide code for user queries, with explanations in the comments. If you are unsure about something, stop and ask the user more follow up questions.

User Prompt



+

Default ▾

API Reponse ▾

Last saved Jun 11 at 5:55 PM

🔗

claude-sonnet-4-20250514

{ }

4

🔗

Examples

🌟 Improve prompt

System Prompt

ⓘ ▾

You are a Python programming assistant. Provide code for user queries, with explanations in the comments. If you are unsure about something, stop and ask the user more follow up questions.

User

📎

You are tasked with writing code to connect to device via REST API. The response will be in JSON.

You will be provided with 4 inputs:

1. `<env>{{ENV API_KEY }}` - The authentication key environment variable name.
2. `<baseurl>{{URL https://partner.ultrahuman. }}` - The endpoint URL.
3. `<endpoint>{{ENDPOINT metrics }}` - The endpoint to information about a specific operation.
4. `<params>{{PARAMS {email:shravanaras@arizon. }}` - JSON object of the paramters.

Now write python code using the Request HTTP library that does the following -

1. Gets the authentication API key from the environment variable `<env>`.
2. Use this as the authentication header in the http request.
3. Combine `<baseurl>` and `<endpoint>` together to form the http request endpoint.
4. Pass the GET parameters from `<params>`

Check to see if we get a http code 200 for success. If not then print out an error message and raise an exception. The response will be returned in JSON. Parse that JSON and convert it into a python object.

🔗 Pre-fill response

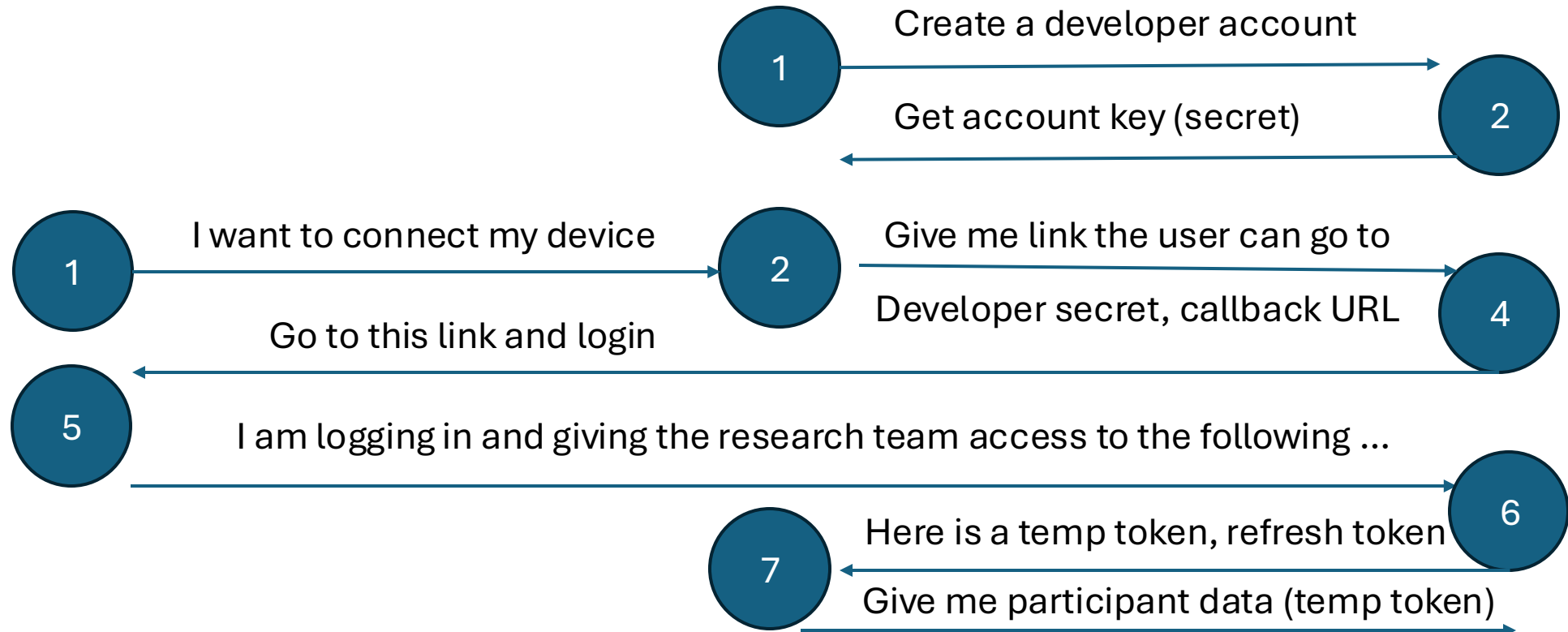
📄 Add message pair

Types of API Access – OAuth 2.0

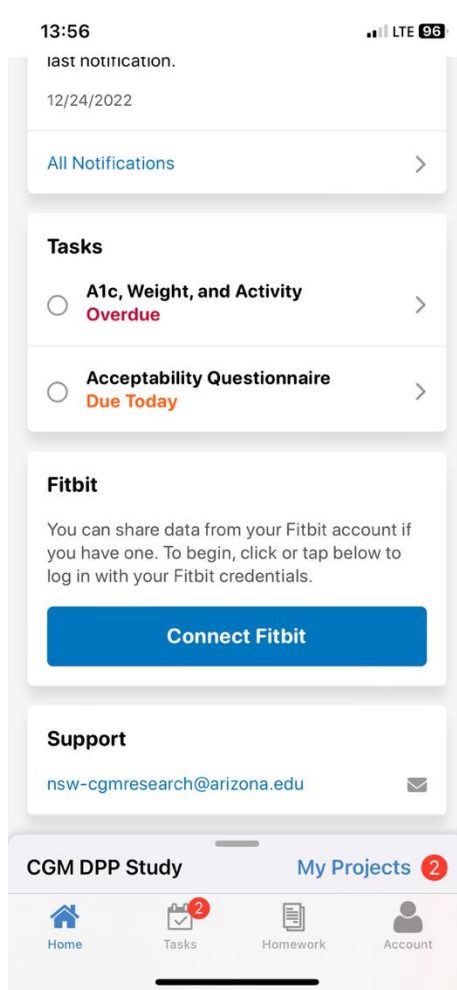
User

Research Team

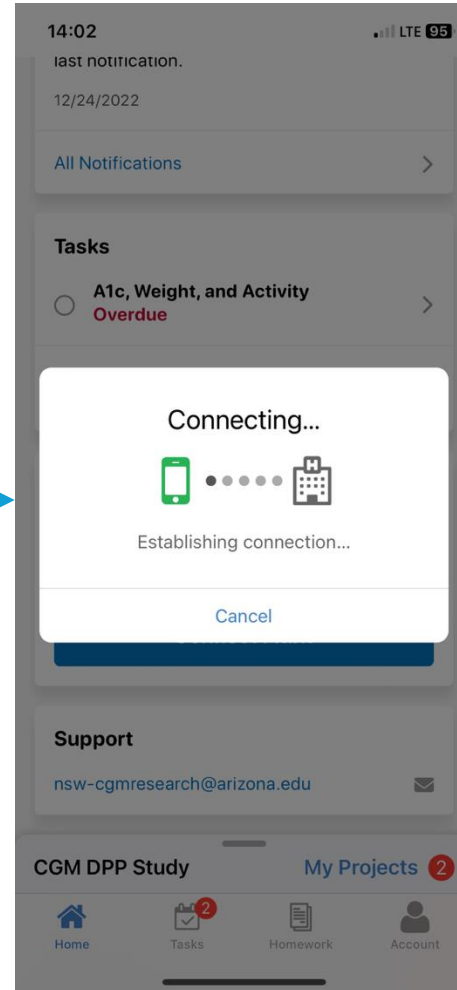
Device



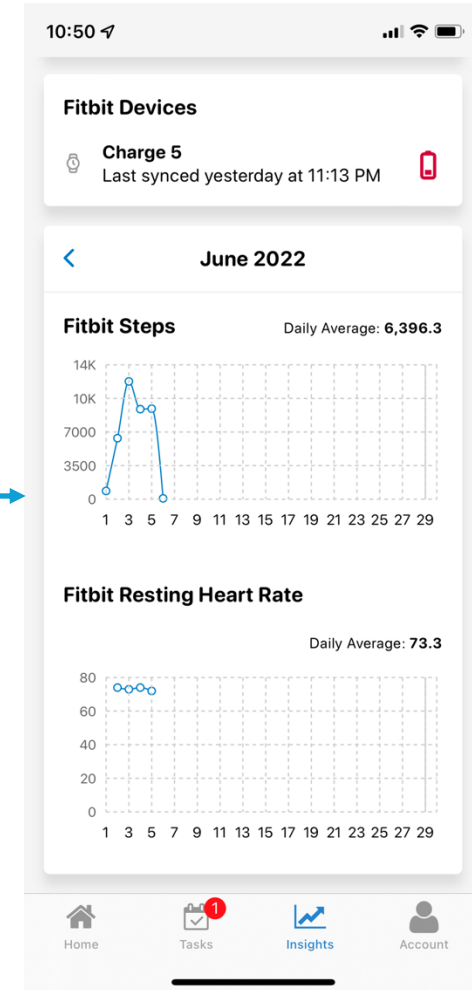
Using 3rd party mHealth Eco-system



User gives
permission



Connected!



Step 4 – Store data and connect

Time to make sure things are easy to query and connect with other rest of the research study instruments

Using a database solution

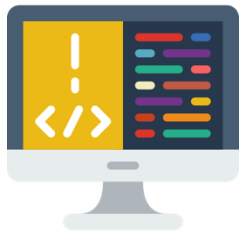
Sensor CSV Exports



Sensor API Connection



Surveys, Interventions, etc



DuckDB / AWS Athena / Postgresql / SQLite ...

SQL

Python
R
PowerBI
Apache Superset

Using AI with SQL Generation

duckdb-nsql

```
ollama run duckdb-nsql
```




↓ 35.4K Downloads ⌚ Updated 1 year ago

7B parameter text-to-SQL model made by MotherDuck and Numbers Station.

7b

Models

[View all →](#)

Name	Size	Context	Input
duckdb-nsql:latest	3.8GB	16K	Text
<u>duckdb-nsql:7b</u> latest 	3.8GB	16K	Text



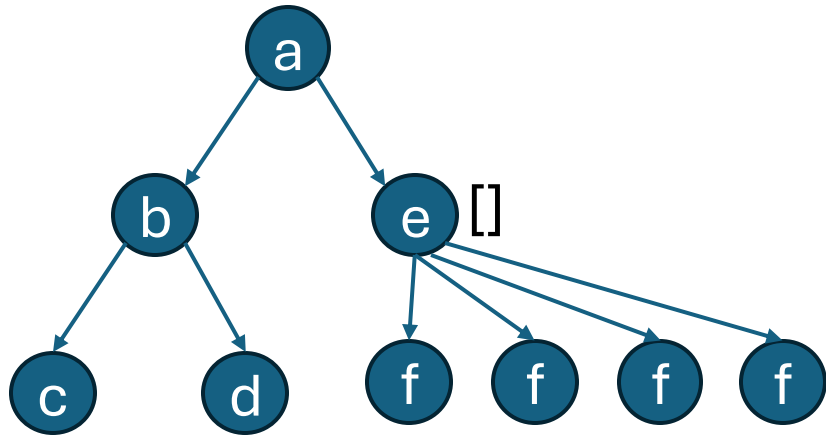
Things to watch out for CSV

- Look at date and time very carefully.
 - Keep things as UNIX timestamp (BIGINT).
 - Convert it all to local UTC ISO timestamp - 2024-01-19T16:00:00-08:00
- Make sure the CSV files have a header.
- Make sure to reduce duplicates in data ingestion.
 - Creating primary keys.
 - Time based logic.

Brings it all together

- Merge data from different endpoints into a single RDS layer
- Use your favorite programming language – R, Python, C++, etc.
- Almost all no code visualization tools support SQ: connectors
- Use LLM models to generate SQL queries

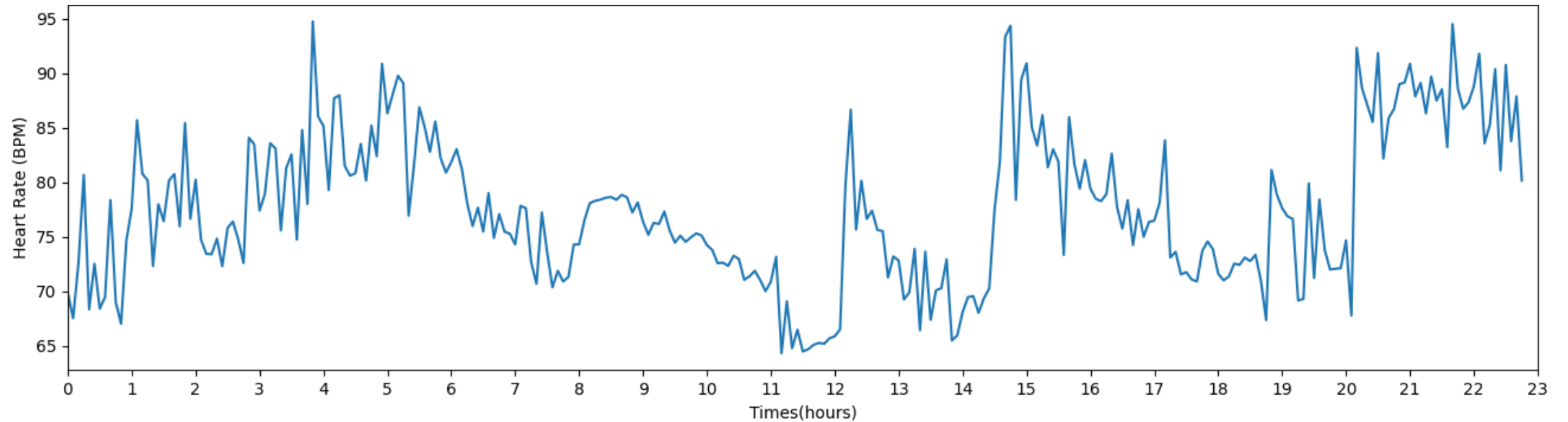
JSON needs a bit more work



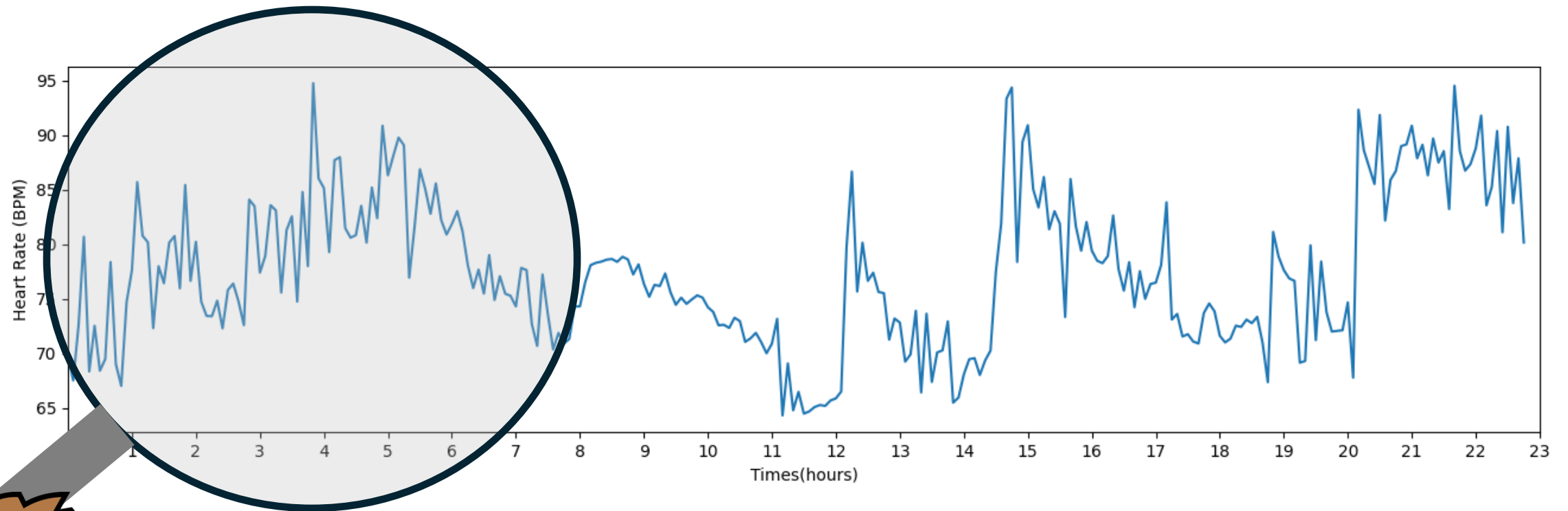
- Columns in tabular structure
 - A
 - A_b_c
 - A_b_d
 - Might need filling
 - A_e_f
 - Each element is a row



Something is Not Right Here!



That HR looks odd for sleeping?

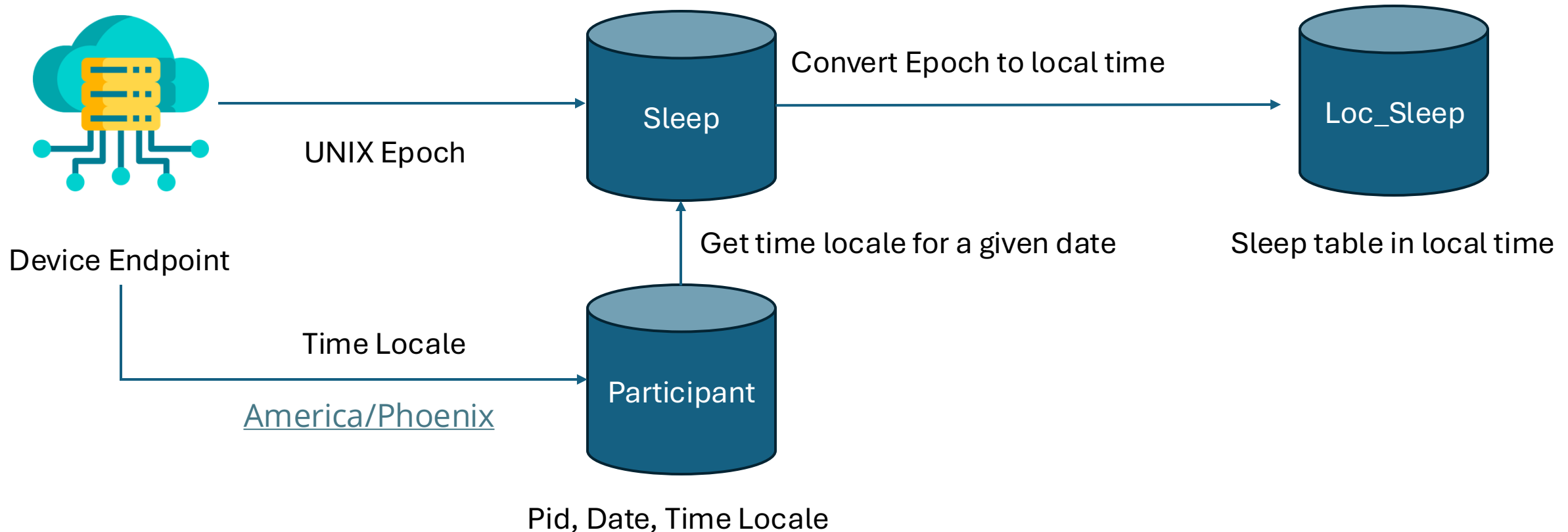


We Forgot Time Zones!

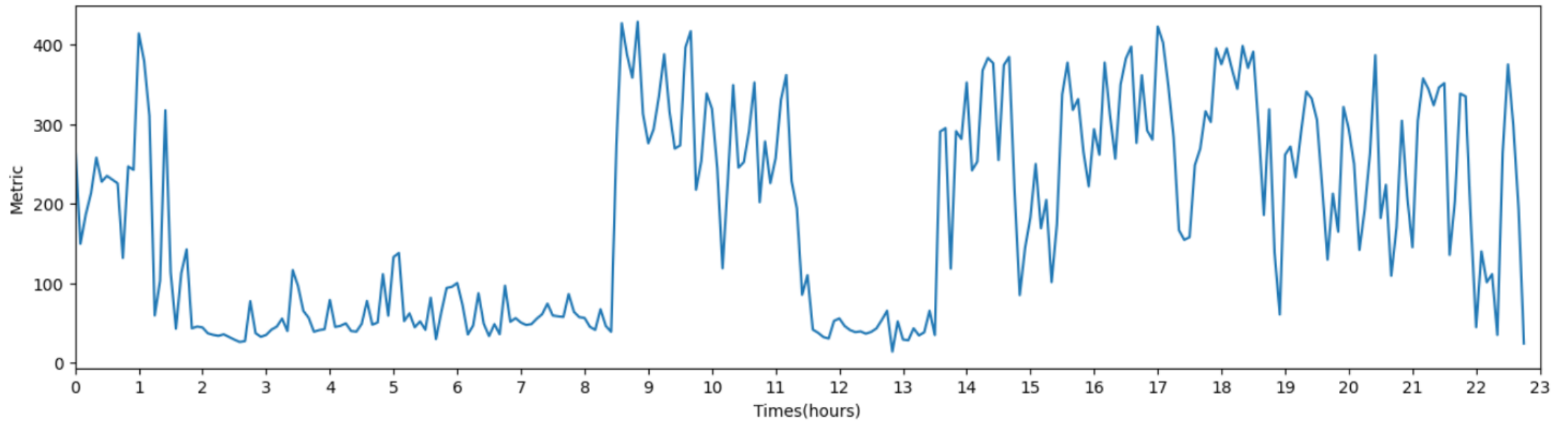
We Forgot Time Zones!

Let's adjust for it now!

Time Zones need a special mention



Much Better!



Example of Our DB

# ▾	timestamp_ns ▾	datetime_utc ▾	datetime ▾	ibi_ms ▾	hr ▾	device_id ▾	timezone ▾	algotversion ▾
1	1706131308920196992	2024-01-24 21:21:48.920	2024-01-24 13:21:48.920	714.853958	83.933227659348		-28800	{major: 8, 'minor': 12, 'patch': 3}
2	1706131309631769194	2024-01-24 21:21:49.631	2024-01-24 13:21:49.631	711.572202	84.32032593650982		-28800	{major: 8, 'minor': 12, 'patch': 3}
3	1706131310294824256	2024-01-24 21:21:50.294	2024-01-24 13:21:50.294	663.055062	90.49022236406665		-28800	{major: 8, 'minor': 12, 'patch': 3}
4	1706131311000879662	2024-01-24 21:21:51.000	2024-01-24 13:21:51.000	706.055406	84.97916663497652		-28800	{major: 8, 'minor': 12, 'patch': 3}

Values returned by Empatica

Per day adjusted local time

Extracted local time offset for reference



Step 4.b – Stay on top of participant compliance

Who is naughty and nice!

Inferring Wear Time

- Looking at HR values
- Checking to see if step counts stay as default
- Looking at resting HR
- BE Creative!

Creating a compliance dashboard

Fitbit Compliance App

Overview		Sleep		Activity		Resting HR		Heart Rate Variability	
Participant	Last Synced	Days Since Sync	Sleep	Activity	Resting HR	HRV			
filter data...									
MDH	2023-07-13	68	Yes	Yes	Yes	Yes			
MDH	2023-05-14	128	Yes	Yes	Yes	Yes			
MDH	2023-03-05	198	Yes	Yes	Yes	Yes			
MDH	2023-03-20	183	Yes	Yes	Yes	Yes			
MDH	2023-06-03	108	Yes	Yes	Yes	Yes			
MDH	2023-05-15	127	Yes	Yes	Yes	Yes			
MDH	2023-06-03	108	Yes	Yes	Yes	Yes			
MDH	2023-07-13	68	Yes	Yes	Yes	Yes			
MDH	2023-07-13	68	Yes	Yes	Yes	Yes			
MDH	2023-06-15	96	Yes	Yes	Yes	Yes			
MDH	2023-07-03	78	Yes	Yes	Yes	Yes			
MDH	2023-08-18	32	Yes	Yes	Yes	Yes			
MDH	2023-06-17	94	Yes	Yes	Yes	Yes			
MDH	2023-07-13	68	Yes	Yes	Yes	Yes			
MDH	2023-07-09	72	Yes	Yes	Yes	Yes			
MDH	2023-05-11	131	Yes	Yes	Yes	Yes			
MDH	2023-02-13	218	No	Yes	Yes	No			
MDH	2023-03-29	174	Yes	Yes	Yes	No			

Number of participants who have -

- Fitbit connected - 18
- Sleep data - 17
- Activity data - 18
- Resting HR data - 18
- HRV data - 16
- Not synced in the last 2 days or more - 18

Legends

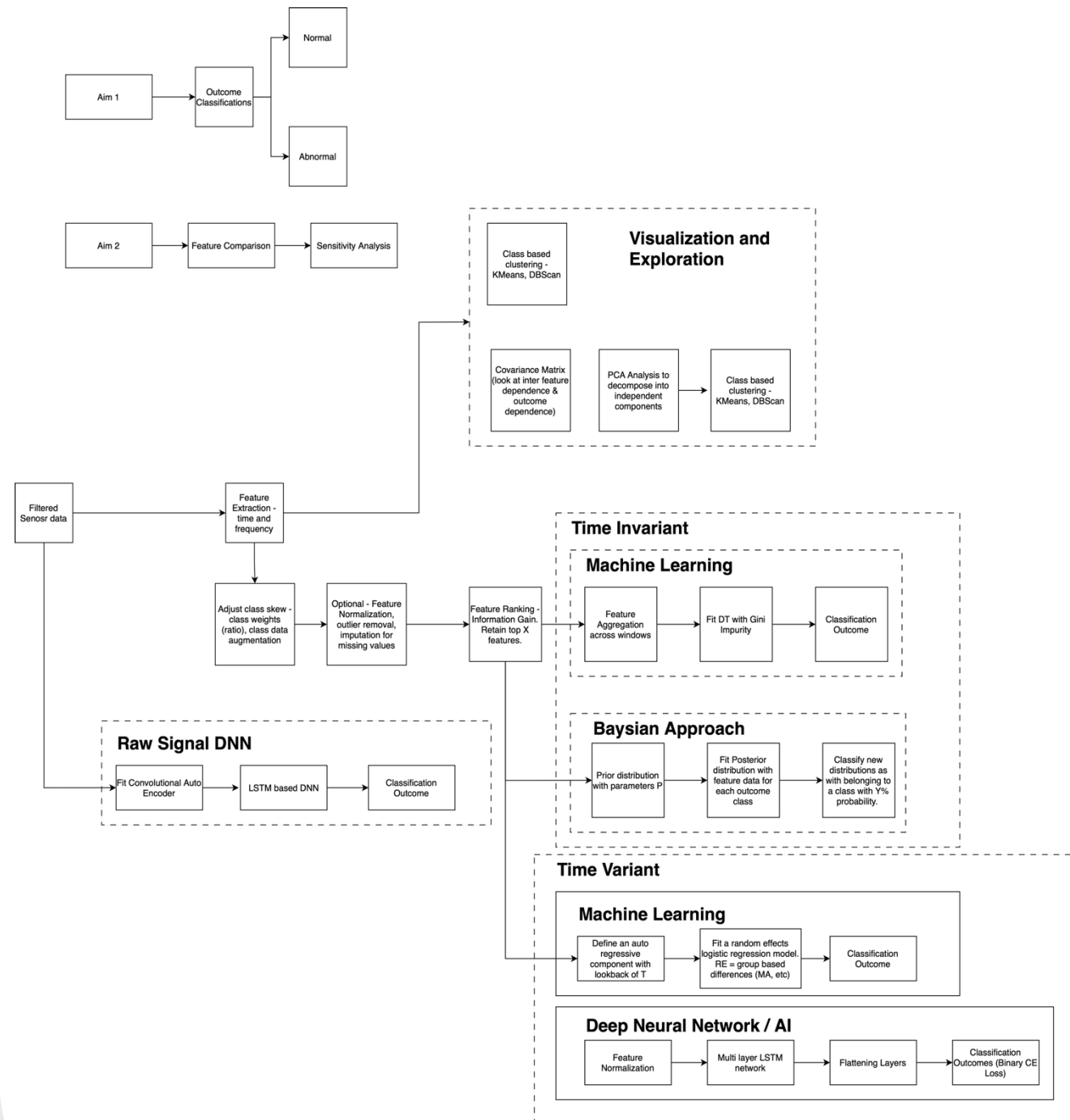
- Green - Data for this metric exists. Refer to the individual tab to learn more about this metric



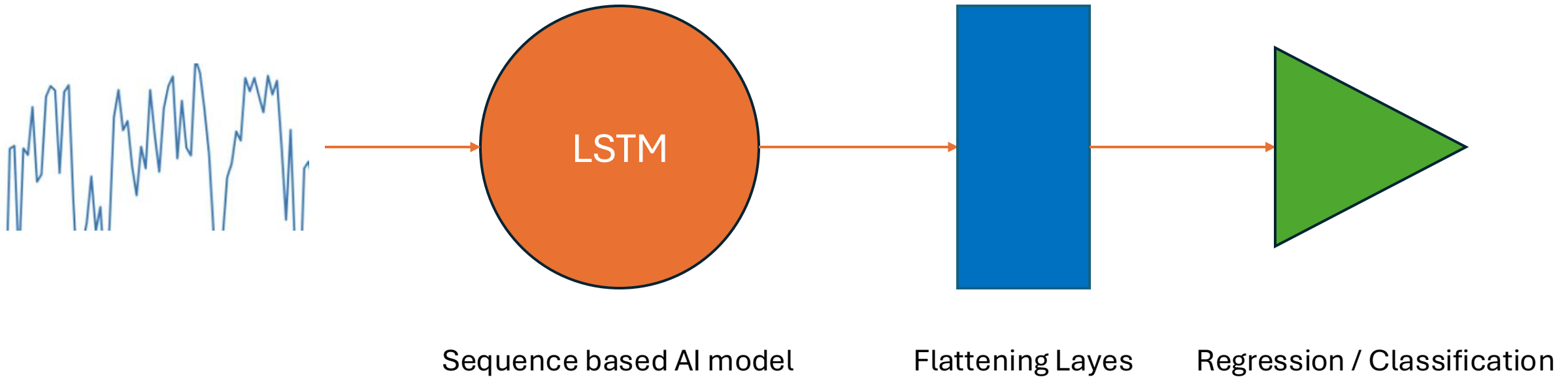
Step 5 – Data analysis pipeline

Alright time to figure out what this data means.

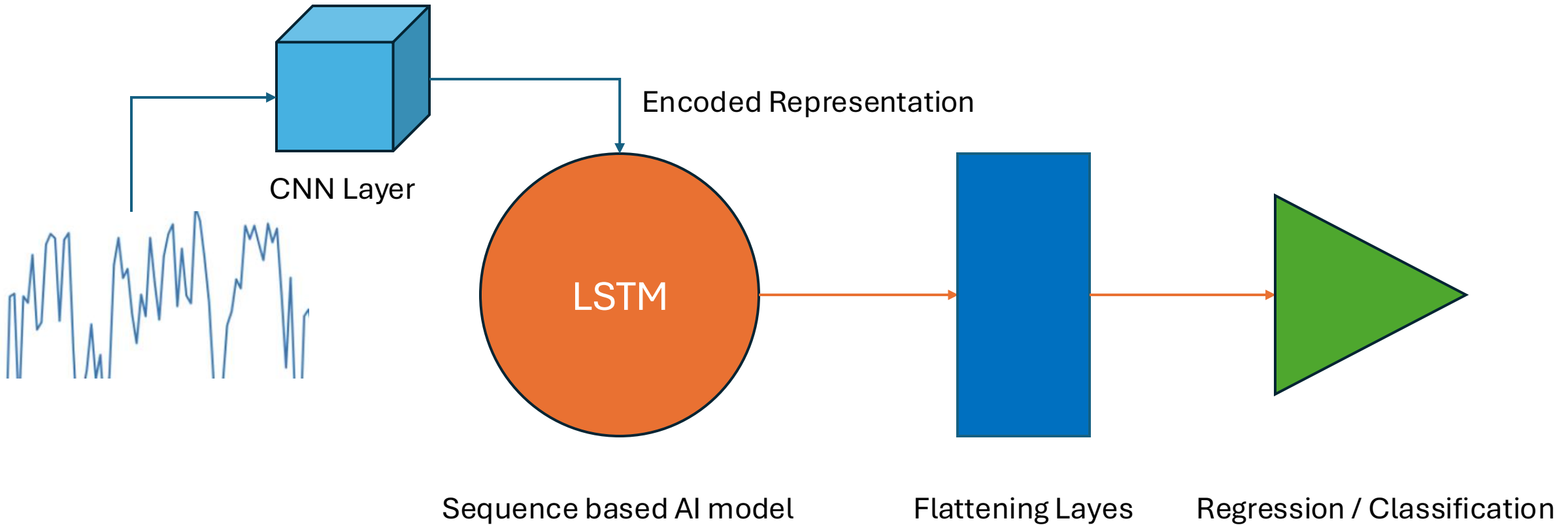
Creating an Analytic Workflow



A Simple Template to Try - 1



A Simple Template to Try - 2



Thank You Let's Stay in Touch!

- Reach out at –
shravanaras@arizona.edu

