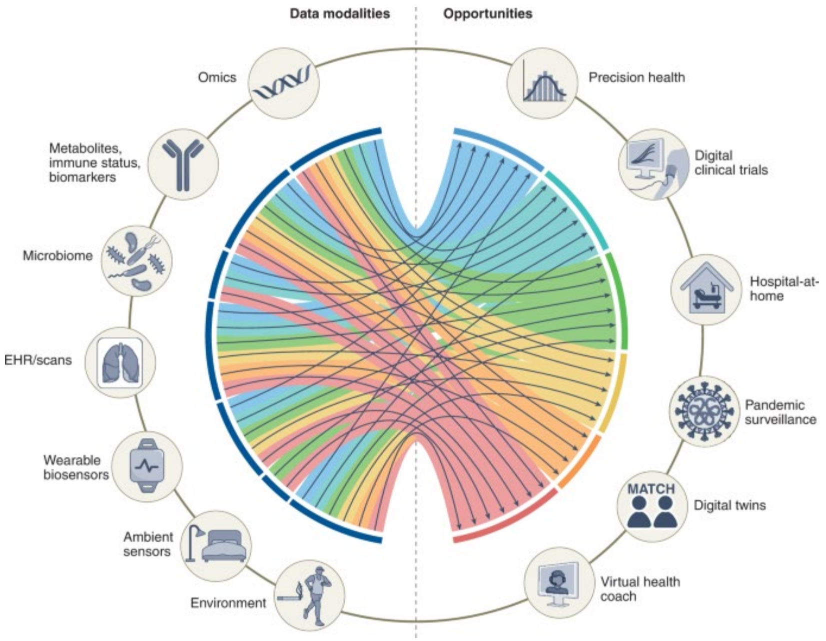
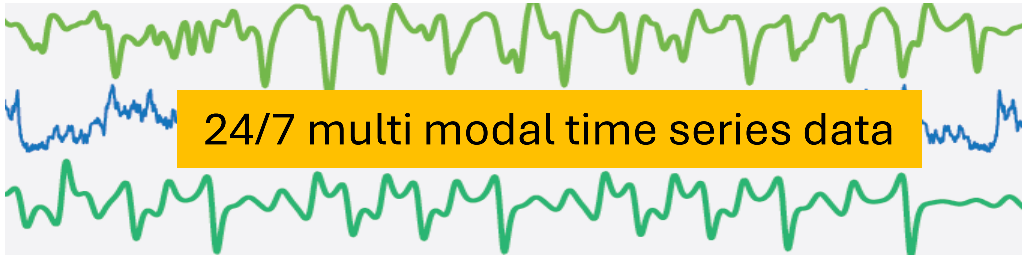
Multimodal sensing and modeling

Akane Sano
Rice University
Department of Electrical and Computer Engineering,
Computer Science, and Bioengineering
Akane.Sano@rice.edu

How can we take advantage of daily life moment-to-moment data + other data?





Personalized feedback system

Sensing

- Physiological
- Behavioral
- Social
- Chemical

- Wearables
- Surveys
- Clinical assessment
- Imaging
- EHR
- Omics

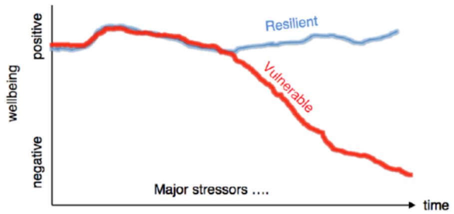
Interpretation Inference

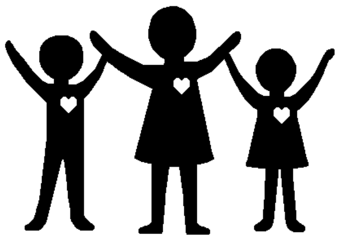
Marker development

Analyses
Classification, regression
Prediction
Causality

Feedback

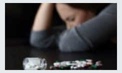
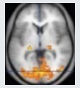
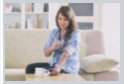
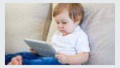
Intervention
Treatment
Decision Support
Stimulus





Patients (e.g., psychiatry, neurology, oncology)
People at higher risk for issues
(e.g., shift workers, caregivers)
Users (e.g.,health professionals)

Remote sensing & intervention

		Target	Wearable	text	imaging	Others	interventions
Patients	Alzheimer's disease patients/caregivers 	Physical and mental health	X 			cytokines 	X
	Substance use disorders 	Affect and craving profiling	X		fMRI 	Genetic markers, gait	
	Cancer patients undergoing chemotherapy 	Symptom detection/ prediction	X			home monitoring	
	Epilepsy 	Seizure prediction	X	X	camera 	Intracranial EEG 	
	Cardiac Disease 	Detection /prediction	X	X	X-ray		
Non-patients	Shift workers 	Sleep	X	X			X
	Office workers 	Productivity	X			work tasks, computer usage	X
	Students 	Mental Health	X			circadian phase	X
	Drivers 	Safety	X		Camera, radar	Vehicle sensors	
	Children 	Screen/TV Use vs weight & development	X		Camera		

Ideal Data?

- Long-term
- Multimodal
- Diversity in demographics and geographic regions
- Critical rare events
- Ideally consistent data formats/streams

will help

- understand individual or subtype differences
- build and test dynamic and personalized machine learning models to capture time-varying users' symptoms/physiology/behavior
  - Classification of human constructs
  - Human construct or risk detection or prediction
  - Intervention/delivery timing/target selection

Challenges

Data Collection

Diversity

User burden

Energy consumption

Storage

Noise

Missing Data

Privacy

Security

User/patient engagement

Modeling

Robustness

Explainablity

Equity

Limited labels/data

Missingness

Individual differences

Multimodal integration

Multiuser data integration

Biobehavioral marker extraction

Feedback loop

What to feedback to users/patients and when/how?

Safety

Trust

Explainability

Actionable feedback

Deployment

Multimodal integration

User Engagement

Adaptability

Scalability

Challenges

Sensing

Energy and memory consumption

Noise/Missingness

Inference

Limited Labels

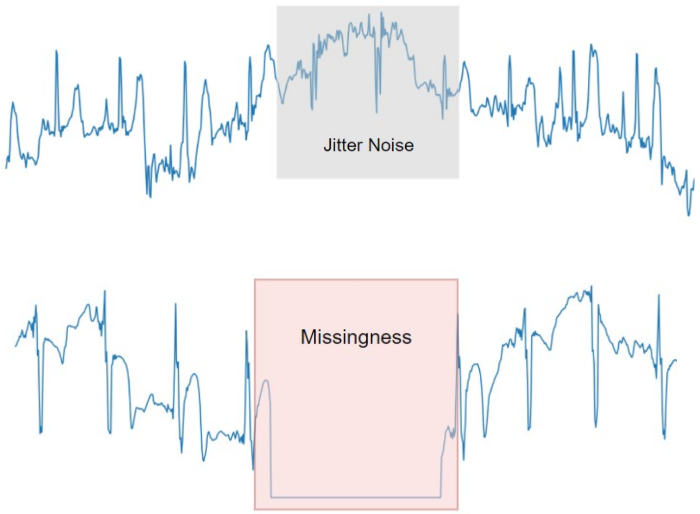
Algorithm Bias

Feedback

Explainability /Actionable Insights

Problem1a: Data Collection is Expensive and Collected Data Can be Noisy

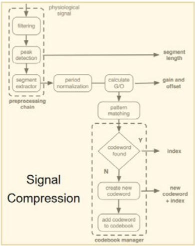
High-frequency signals are informative. However, it requires **higher energy consumption** and **greater hardware demands**.

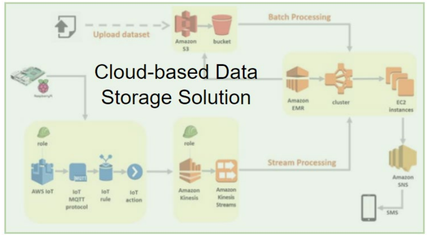


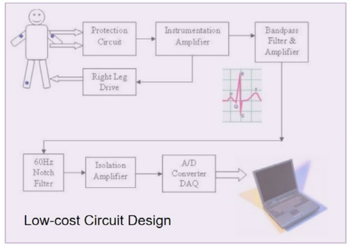
Noise and Missingness

Prior Work:

High frequency data collection

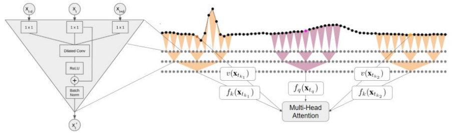


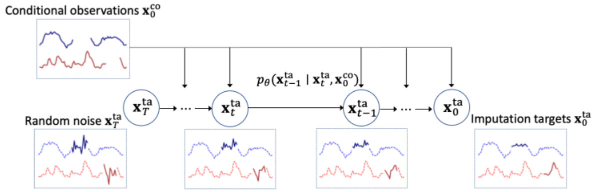




(Hooshmand *et al.*, 2017) (Taher *et al.*, 2019) (Deb *et al.*, 2017)

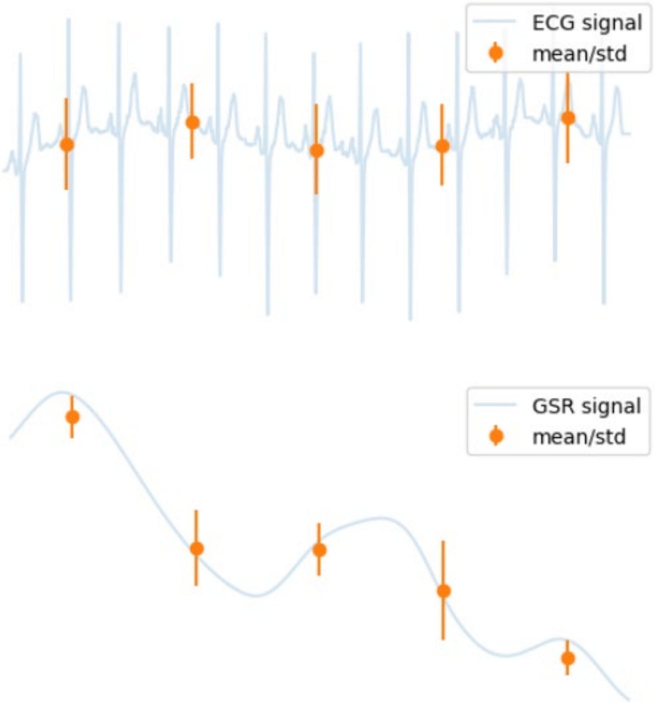
Impute missing/noisy segments





Non-stationarity in Time series data

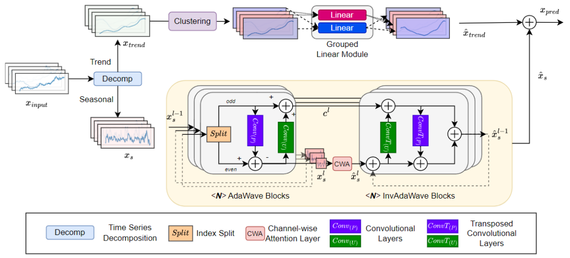
The distributions of time-series data **keep changing over time**



Traditional structures such as convolution network and Transformers **might face challenges in achieving robust generalization**

Research Target:

- **Can we make analytical methods/models to handle multimodal noisy data?**
- **Can we extract information from low sampling rate data?**
- **Can we synthesize realistic multimodal data?**

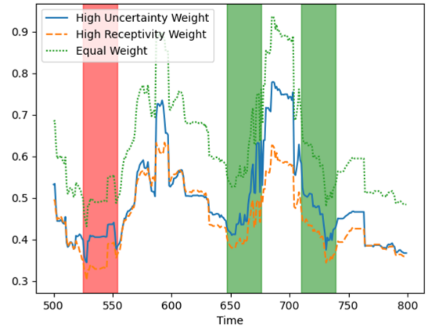


Han Yu, Peikun Guo, Akane Sano, "**AdaWaveNet: Adaptive Wavelet Network for Time Series Analysis**", under review.

Problem 1b: Diverse data collection with higher participant engagement

- When to sample data and labels (e.g. ecological momentary assessments (EMA))?
- When to provide feedback/interventions to users?

- Prior work
  - Reduce questions/frequency to ask
  - Gamification
  - Predicting time points where participants are more likely to respond to ecological momentary assessment (EMA) and Interventions for increasing participant receptivity



- Bias: Participant receptivity is partially influenced by contextual states (i.e., participants who are experiencing stress/negative affect are less likely to respond)

- **Research Target**:

Diversify sampling by taking into account participant context when sending EMAs and Interventions

Zach King, Han Yu, Thomas Vaessen, Iniz Myin-Germeys, Akane Sano, Towards the Understanding of Receptivity and Affect in EMAs using Physiological based Machine Learning Method: Analysis of Receptivity and Affect, 2024 JMIR uhealth and mhealth

Challenges

1. Sensing

Energy and memory consumption

Noise/Missingness

2. Inference

Limited Labels

Algorithm Bias

3. Feedback

Explainability /Actionable Insights

Problem 2a: Limited Labels

- Sensors can collect a large amount of data
- Annotation can be expensive, very labor- and time-consuming

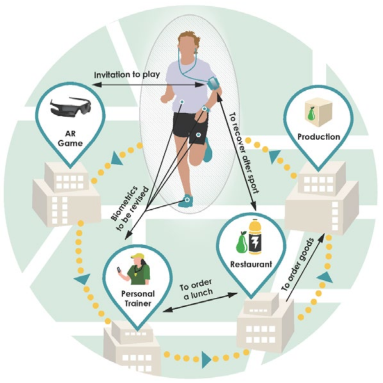


Image credit to Ometov et al. 2017

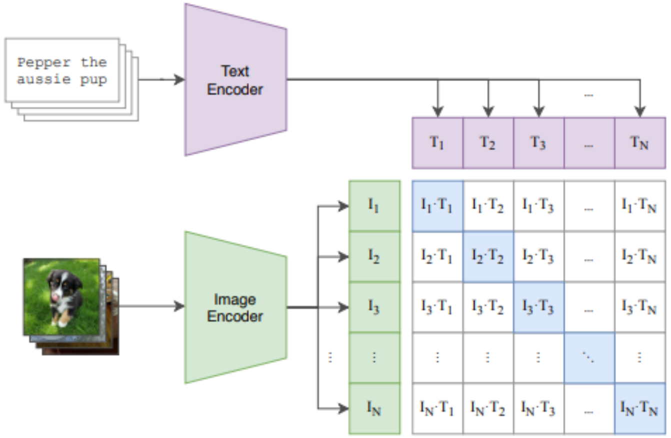
Labeled data Unlabeled data

Research target:

how can we train a robust model with a small amount of labeled data?

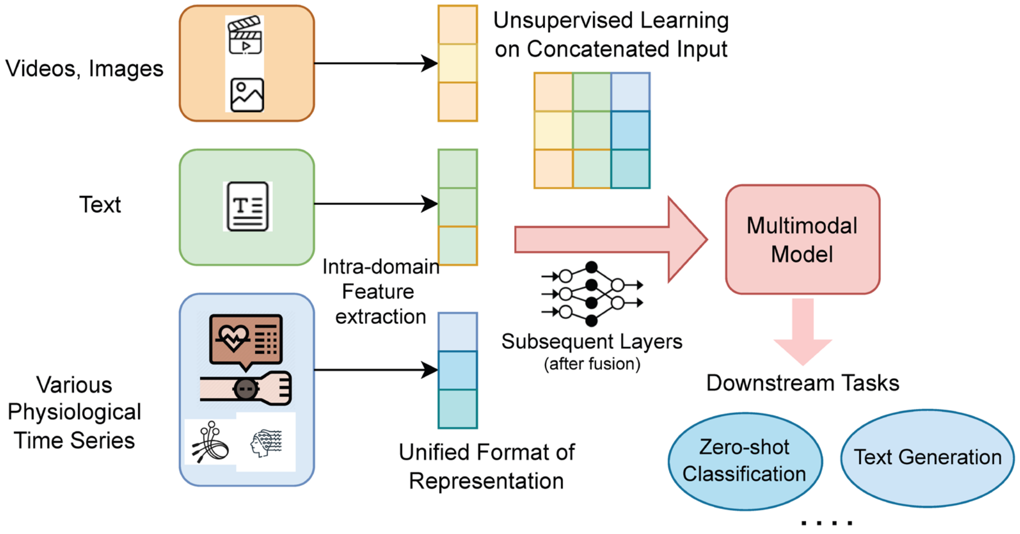
- Data Augmentation (DA)
- Semi-supervised learning + DA
- Self-Supervised Learning (SSL)
- Large Language Models
- Multimodal pre-training

Multimodal Pretraining



(Radford et al., 2021)

?



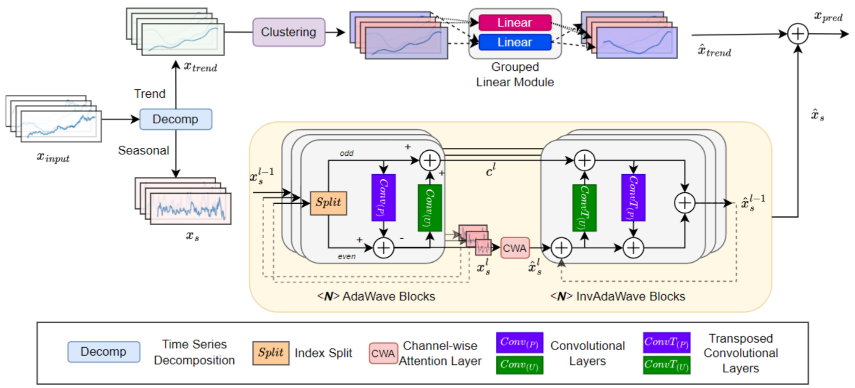
Multimodal pretraining has achieved promising results in field such as image-text pairs.

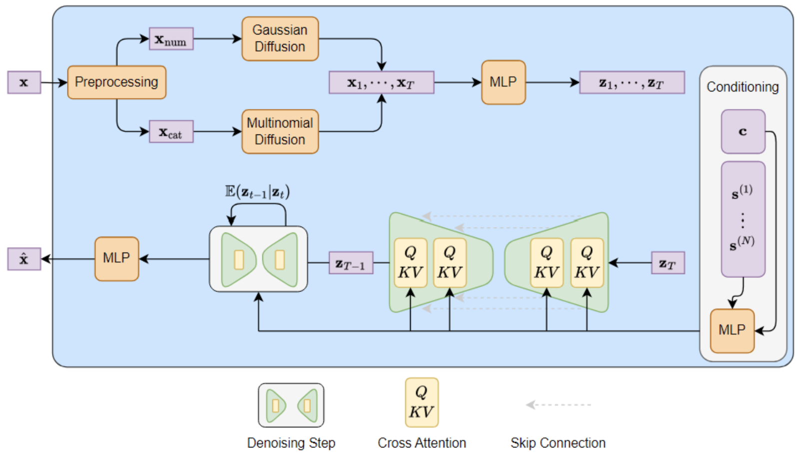
Can we leverage multimodal pretraining for diverse data **to learn representations?**

Requires large sets of data

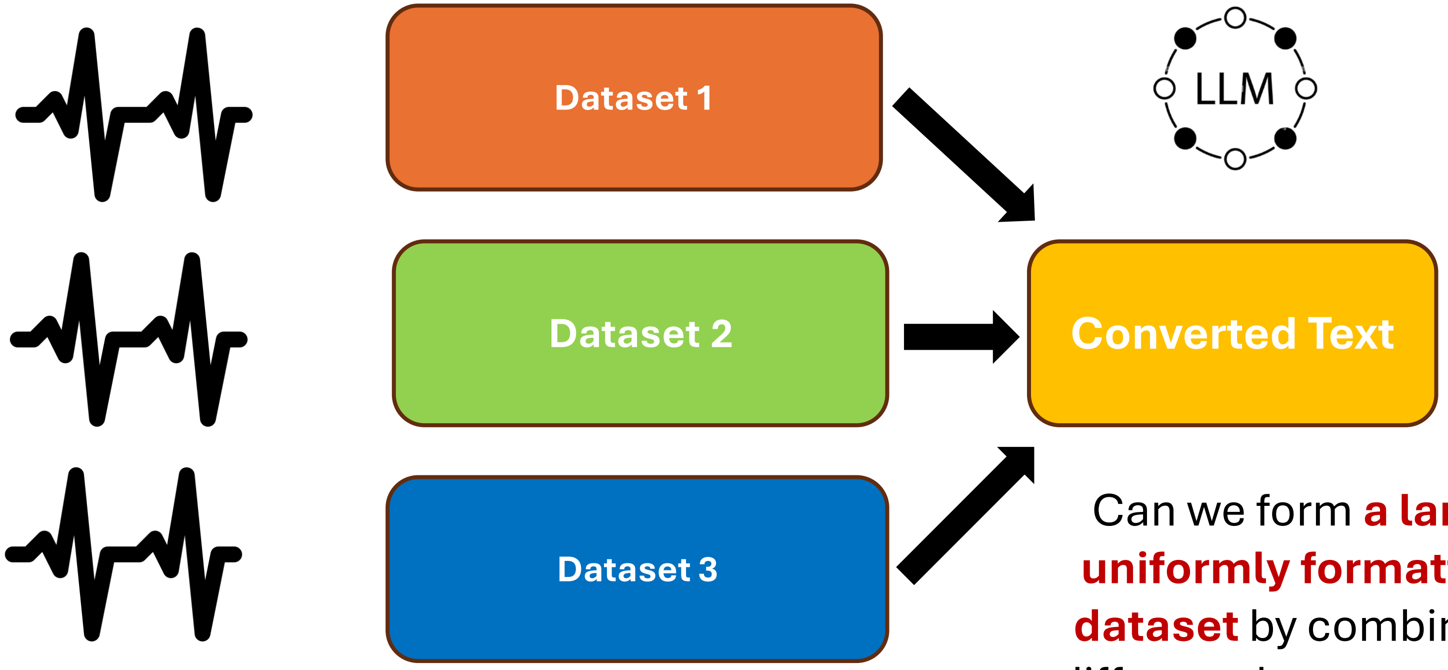
Can we generate synthetic data?

- Generative models
  - Autoencoders
  - Generative adversarial networks (GANs)
  - Transformers
  - Diffusion models
- Advantages
  - Scalability
  - Privacy
- Disadvantages
  - Quality
  - Bias





Can we combine multiple datasets for building foundation models?



Signals in the same format

Various Format of data: difficult for the model to learn a consistent mapping

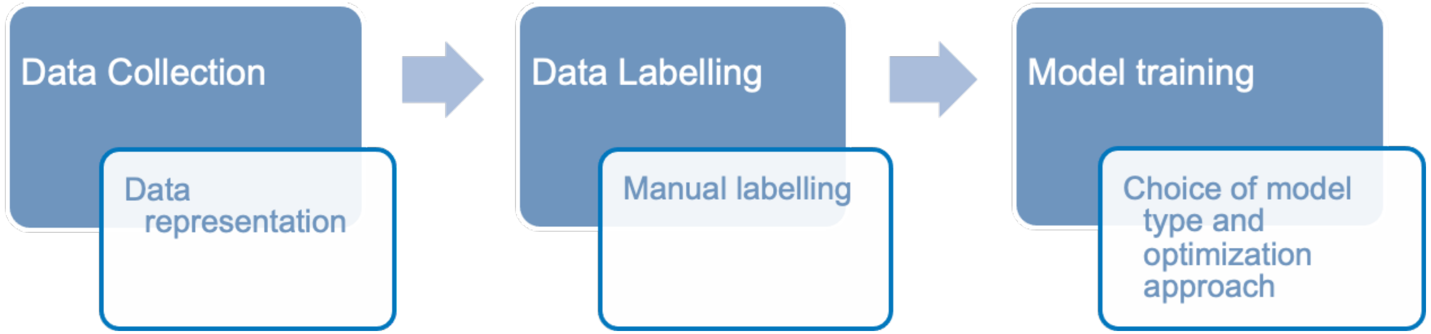
Can we form **a large uniformly formatted dataset** by combining different datasets using LLMs?

H Yu, P Guo, A Sano, ECG Semantic Integrator (ESI): A Foundation ECG Model Pretrained with LLM-Enhanced Cardiological Text, 2024, Transactions on Machine Learning Research

Problem 2b: Bias & Fairness in Machine Learning

An unfair algorithm is one whose decisions are skewed toward a particular group of people.

Bias need to be recognized and treated carefully when models are used in decision making (e.g. social or health systems).

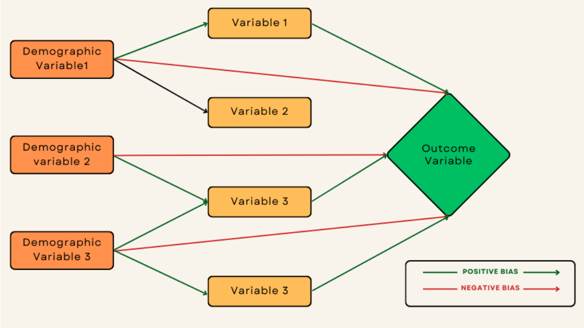


Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, *54*(6), 1-35.

Equitable models for different groups of people

- Bias Mitigation Methods
  - Pre processing (e.g., reweight data for removing imbalances)
  - In processing (e.g., change objective function, add constraint)
  - Post processing (e.g., minimize the ability of adversarial network to predict sensitive attribute)

- Lack of understanding of sources of bias
- Lack of generalizability of bias mitigation methods.



- **Research Targets**
  - **Can we identify/understand where bias come from?**
    - **Causal analyses**
  - **Can we design a generalizable bias mitigation techniques (e.g., design a model or generate data) so that our models work accurately and equally for different group of people?**

Challenges

1. Sensing

Energy and memory consumption

Noise/Missingness

2. Inference

Limited Labels

Algorithm Bias

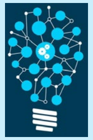
3. Feedback

Explainability /Actionable Insights

Problem 3. Closed-loop feedback



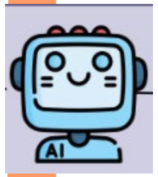

Sensing & Processing /ML Pipeline



Physiology
Behavior
Environment
Profiles
History
User Response
User preference

- Human state models
- Responsiveness models

- Intervention guideline, theories
  (e.g. cognitive behavioral therapy, mindfulness)

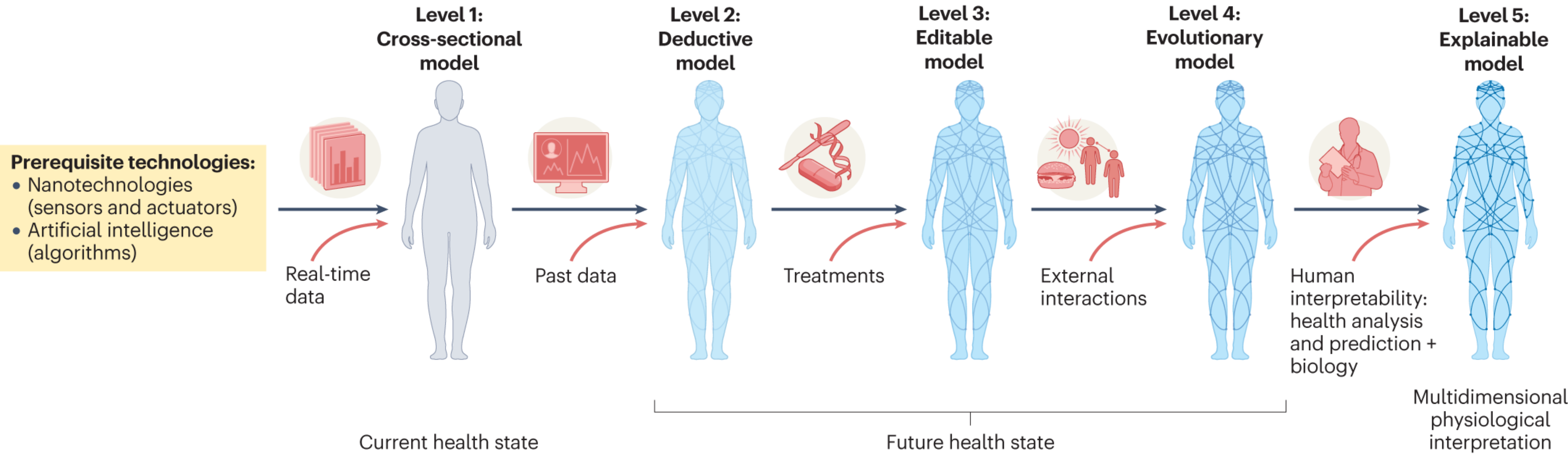


- Interventional messages
- Stimuli
- Treatment plans

Can we design actionable, personalized and adaptive feedback?
When and which intervention strategies to provide?
How do users respond to feedback?
How can we design safe and reliable feedback?
How can we make this feedback loop sustainable?

Digital twins

- How can we model causal relationships between complex input and output?



Challenges and future work in multimodal sensing and modeling approaches

- Missing/noisy data
- Limited data/labels
- Algorithm Bias
- Explainability/Feedback design

- How much/which data do we need to collect or generate?
- How much data/model precision, reliability and interpretability do we need?
- How can we define and test safety?
- Understanding causal relationships
- Require research environments
    - to integrate and compare datasets from multiple studies
    - to accelerate developing and testing the effectiveness and safety of models and findings