Signal Intelligence: Human-AI Teaming for Time Series Analytics

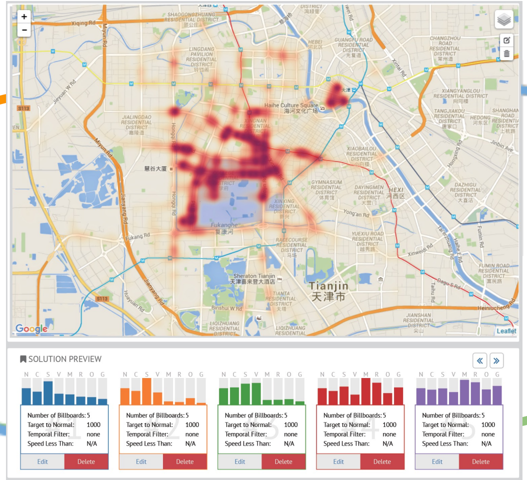
Dongyu Liu

Assistant Professor
Department of Computer Science
University of California, Davis
dyuliu@ucdavis.edu

UCDAVIS
COMPUTER SCIENCE

My Research

AI System
(Human-Centered)



Visual
Interface

Human

AI (ML)

The term is coined by Prof. John McCarthy in 1950s

”The science and engineering of making intelligent machines”

https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

Intelligence might be defined as the ability to learn and perform suitable techniques to solve problems and achieve goals, appropriate to the context in an uncertain, ever-varying world. A fully pre-programmed factory robot is flexible, accurate, and consistent but not intelligent.

Artificial Intelligence (AI), a term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as “the science and engineering of making intelligent machines”. Much research has humans program machines to behave in a clever way, like playing chess, but, today, we emphasize machines that can learn, at least somewhat like human beings do.

Autonomous systems can independently plan and decide sequences of steps to achieve a specified goal without micro-management. A hospital delivery robot must autonomously navigate busy corridors to succeed in its task. In AI, autonomy doesn’t have the sense of being self-governing common in politics or biology.

Machine Learning (ML) is the part of AI studying how computer agents can improve their perception, knowledge, thinking, or actions based on experience or data. For this, ML draws from computer science, statistics, psychology, neuroscience, economics and control theory.

In **supervised learning**, a computer learns to predict human-given labels, such as dog breed based on labeled dog pictures; **unsupervised learning** does not require labels, sometimes making its own prediction tasks such as trying to predict each successive word in a sentence; **reinforcement learning** lets an agent learn action sequences that optimize its total rewards, such as winning games, without explicit examples of good techniques, enabling autonomy.

Deep Learning is the use of large multi-layer **(artificial) neural networks** that compute with continuous (real number) representations, a little like the hierarchically organized neurons in human brains. It is currently the most successful ML approach, usable for all types of ML, with better generalization from small data and better scaling to big data and compute budgets.

An **algorithm** lists the precise steps to take, such as a person writes in a computer program. AI systems contain algorithms, but often just for a few parts like a learning or reward calculation method. Much of their behavior emerges via learning from data or experience, a sea change in system design that Stanford alumnus Andrej Karpathy dubbed **Software 2.0**.

Narrow AI is intelligent systems for one particular thing, e.g., **speech** or **facial recognition**. **Human-level AI**, or **Artificial General Intelligence (AGI)**, seeks broadly intelligent, context-aware machines. It is needed for effective **social chatbots** or **human-robot interaction**.

Human-Centered Artificial Intelligence is AI that seeks to augment the abilities of, address the societal needs of, and draw inspiration from human beings. It researches and builds effective partners and tools for people, such as a robot helper and companion for the elderly.

Text by Professor Christopher Manning, September 2020

AI (ML)

The term is coined by Prof. John McCarthy in 1950s

”The science and engineering of making intelligent machines”

https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

Intelligence might be defined as the ability to learn and perform suitable techniques to solve problems and achieve goals, appropriate to the context in an uncertain, ever-varying world. A fully pre-programmed factory robot is flexible, accurate, and consistent but not intelligent.

Artificial Intelligence (AI), a term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as “the science and engineering of making intelligent machines”. Much research has humans program machines to behave in a clever way, like playing chess, but, today, we emphasize machines that can learn, at least somewhat like human beings do.

Autonomous systems can independently plan and decide sequences of steps to achieve a specified goal without micro-management. A hospital delivery robot must autonomously navigate busy corridors to succeed in its task. In AI, autonomy doesn’t have the sense of being self-governing common in politics or biology.

Machine Learning (ML) is the part of AI studying how computer agents can improve their perception, knowledge, thinking, or actions based on experience or data. For this, ML draws from computer science, statistics, psychology, neuroscience, economics and control theory.

In **supervised learning**, a computer learns to predict human-given labels, such as dog breed based on labeled dog pictures; **unsupervised learning** does not require labels, sometimes making its own prediction tasks such as trying to predict each successive word in a sentence; **reinforcement learning** lets an agent learn action sequences that optimize its total rewards, such as winning games, without explicit examples of good techniques, enabling autonomy.

Deep Learning is the use of large multi-layer **(artificial) neural networks** that compute with continuous (real number) representations, a little like the hierarchically organized neurons in human brains. It is currently the most successful ML approach, usable for all types of ML, with better generalization from small data and better scaling to big data and compute budgets.

An **algorithm** lists the precise steps to take, such as a person writes in a computer program. AI systems contain algorithms, but often just for a few parts like a learning or reward calculation method. Much of their behavior emerges via learning from data or experience, a sea change in system design that Stanford alumnus Andrej Karpathy dubbed **Software 2.0**.

Narrow AI is intelligent systems for one particular thing, e.g., **speech** or **facial recognition**. **Human-level AI**, or **Artificial General Intelligence (AGI)**, seeks broadly intelligent, context-aware machines. It is needed for effective **social chatbots** or **human-robot interaction**.

Human-Centered Artificial Intelligence is AI that seeks to augment the abilities of, address the societal needs of, and draw inspiration from human beings. It researches and builds effective partners and tools for people, such as a robot helper and companion for the elderly.

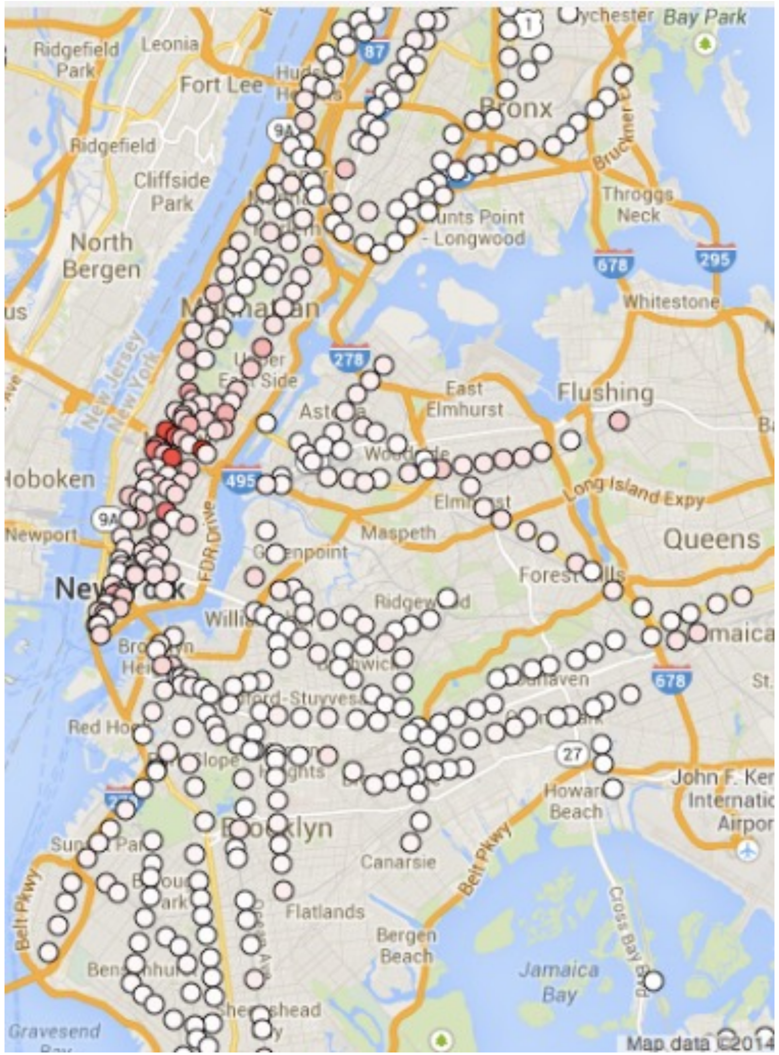
Text by Professor Christopher Manning, September 2020

Data visualization

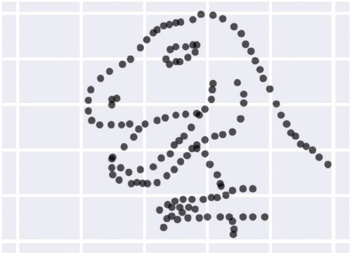
- "Data Visualization is the creation and study of the visual representation of data" - wiki

Input	data
Output	visual form
Goal	insight

MTA FARE @ NY City



Same stats, different graphs



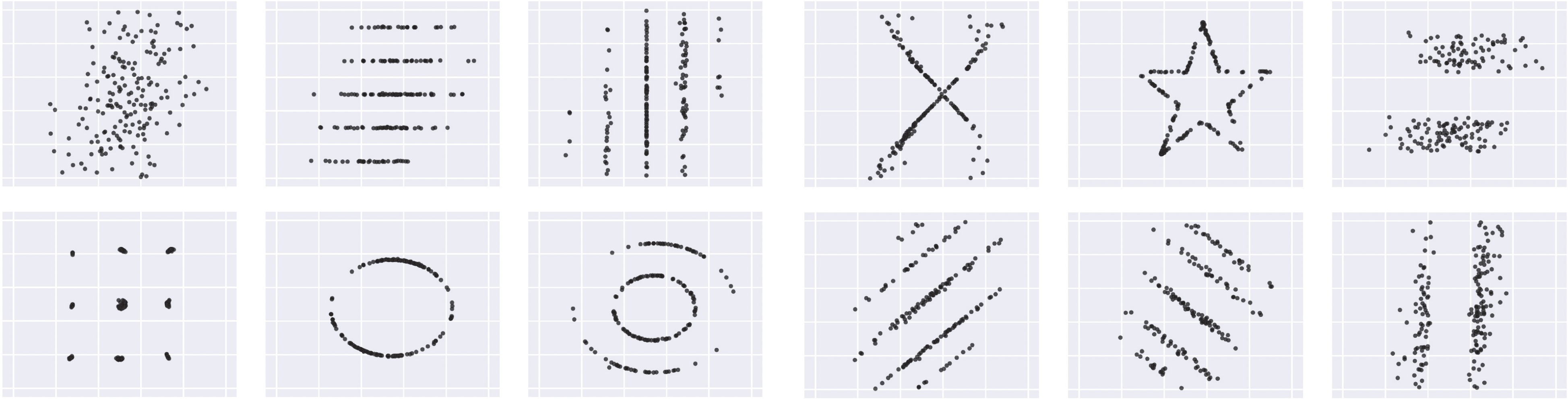
X Mean: 54.26

Y Mean: 47.83

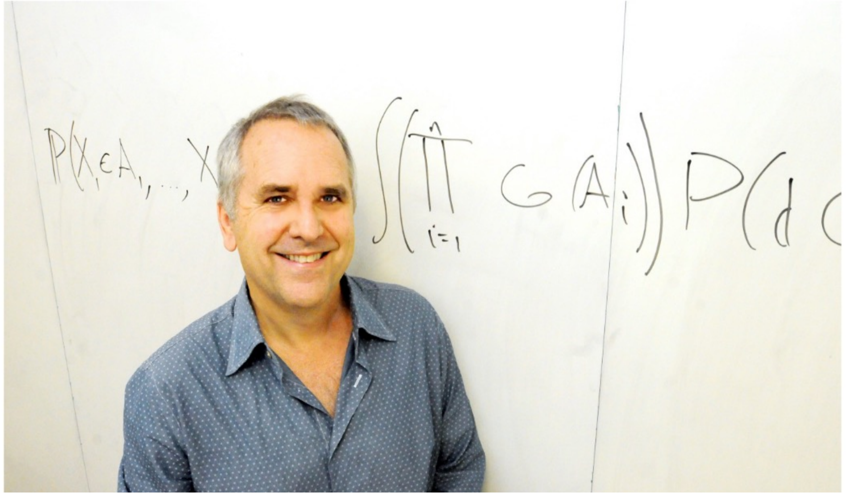
X SD: 16.76

Y SD: 26.93

Corr. : -0.06



Matejka, and Fitzmaurice. Same stats, different graphs: generating datasets with varied appearance and identical statistics through simulated annealing. CHI 2017.



Harvard Data Science Review • Issue 1.1, Summer 2019

Artificial Intelligence—The Revolution Hasn't Happened Yet

Michael I. Jordan[1,2,3]

[1]Berkeley Artificial Intelligence Research Lab, Department of Electrical Engineering and Computer Sciences, University of California Berkeley, Berkeley, California, United States of America,

[2]Department of Electrical Engineering and Computer Sciences, University of California Berkeley, Berkeley, California, United States of America,

[3]Department of Statistics, University of California Berkeley, Berkeley, California, United States of America

Published on: Jul 01, 2019

DOI: https://doi.org/10.1162/99608f92.f06c6e61

Whether or not we come to understand 'intelligence' any time soon, we do have a major **challenge on our hands in bringing together computers and humans in ways that enhance human life.** While some view this challenge as subservient to the creation of artificial intelligence, another more prosaic, but no less reverent, viewpoint is that it is the creation of **a new branch of engineering.** Much like civil engineering and chemical engineering in decades past, this new discipline aims to corral the power of a few key ideas, bringing new resources and capabilities to people, and to do so safely. Whereas civil engineering and chemical engineering built upon physics and chemistry, this new engineering discipline will build on ideas that the preceding century gave substance to, such as information, algorithm, data, uncertainty, computing, inference, and optimization. Moreover, since **much of the focus of the new discipline will be on data from and about humans, its development will require perspectives from the social sciences and humanities.**

AFOSR-3223

Summary Report

AUGMENTING HUMAN INTELLECT: A CONCEPTUAL FRAMEWORK

Prepared for:

DIRECTOR OF INFORMATION SCIENCES
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
WASHINGTON 25, D.C.

CONTRACT AF 49(638)-1024

By: D. C. Engelbart

STANFORD RESEARCH INSTITUTE

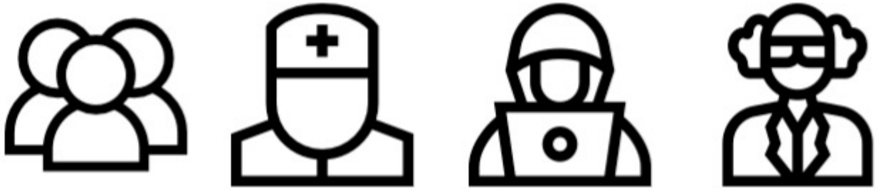
MENLO PARK, CALIFORNIA

SRI

J.C.R. Licklider
March 1960

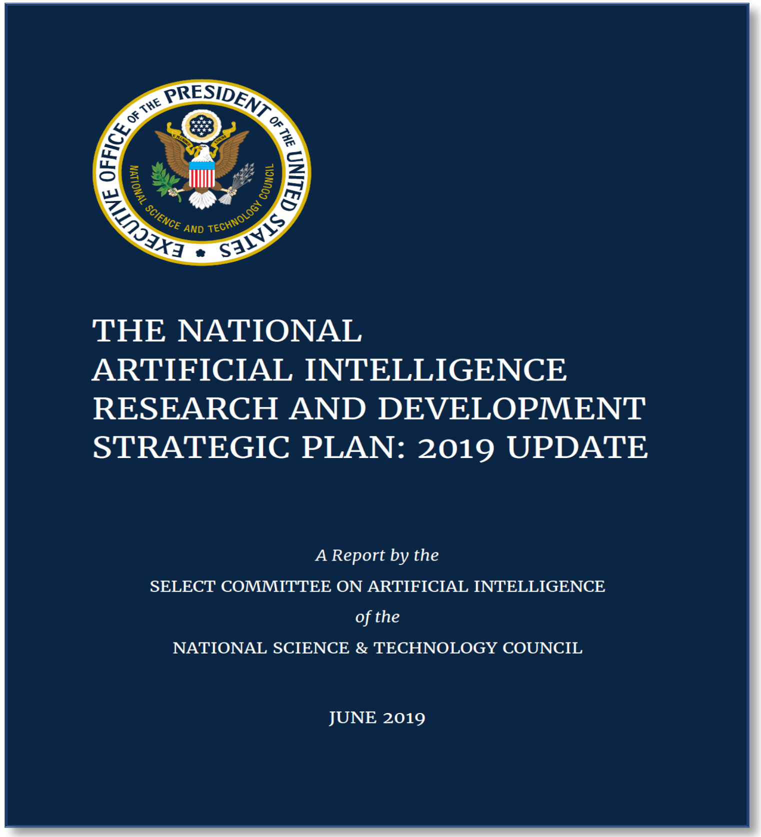
Douglas C. Engelbart
October 1962

Human Users



Three **functional roles** for AI systems:

- AI performs functions alongside the human
- AI performs functions when the human encounters high cognitive overload
- AI performs functions in lieu of a human



Augmenting human intelligence with AI

- Perception
- Attention
- Memory
- Language
- Reasoning
- Problem-solving
- Decision-making
- Creativity

Fine-grained image recognition

http://www.weixiushen.com/project/Awesome_FGIA/Awesome_FGIA.html

Augmenting human intelligence with AI

- Perception
- Attention
- Memory
- Language
- Reasoning
- Problem-solving
- Decision-making
- Creativity

Language Translation

https://www.michigandaily.com/statement/google-translate-and-end-language/

Augmenting human intelligence with AI

- Perception
- Attention
- Memory
- Language
- Reasoning
- Problem-solving
- Decision-making
- Creativity →

Art Design

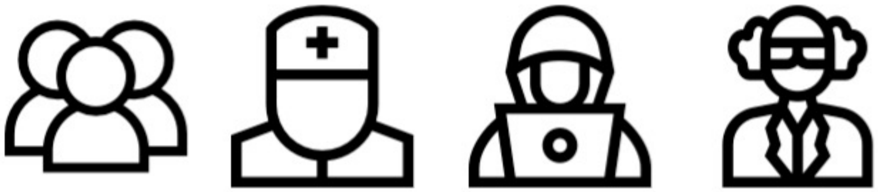
https://www.animaapp.com/blog/design/ai-generated-art-for-product-designers/

Augmenting human intelligence with AI

- Perception
- Attention
- Memory
- Language
- Reasoning
- Problem-solving
- **Decision-making**
- Creativity

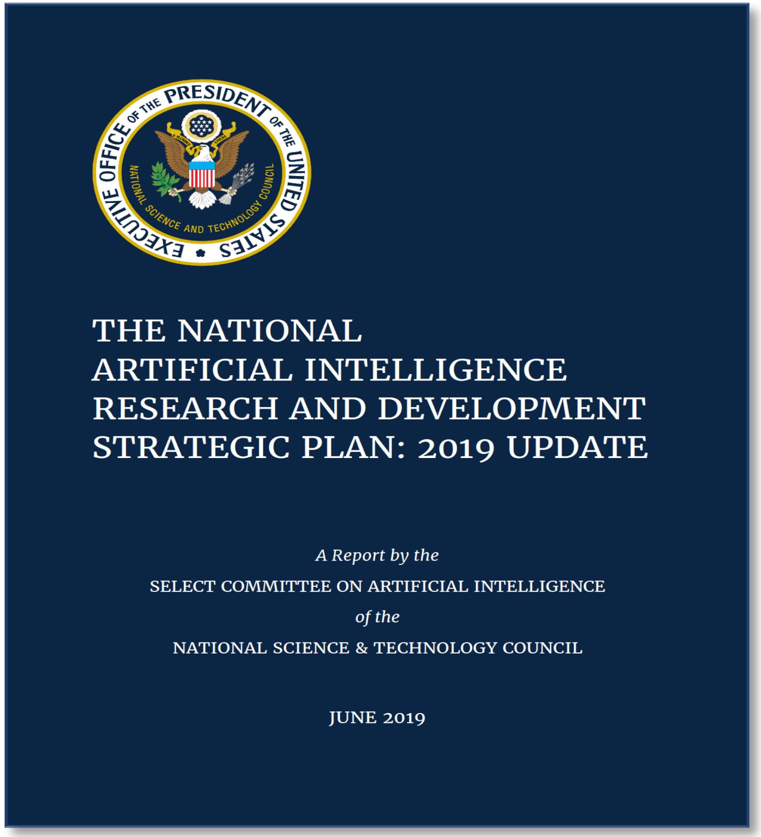
Clinical decision support

Human Users



Three **functional roles** for AI systems:

- AI performs functions alongside the human
- AI performs functions when the human encounters high cognitive overload
- AI performs functions in lieu of a human



OFFICE OF THE PRESIDENT OF THE UNITED STATES · EXECUTIVE · NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN: 2019 UPDATE

A Report by the

SELECT COMMITTEE ON ARTIFICIAL INTELLIGENCE

of the

NATIONAL SCIENCE & TECHNOLOGY COUNCIL

JUNE 2019

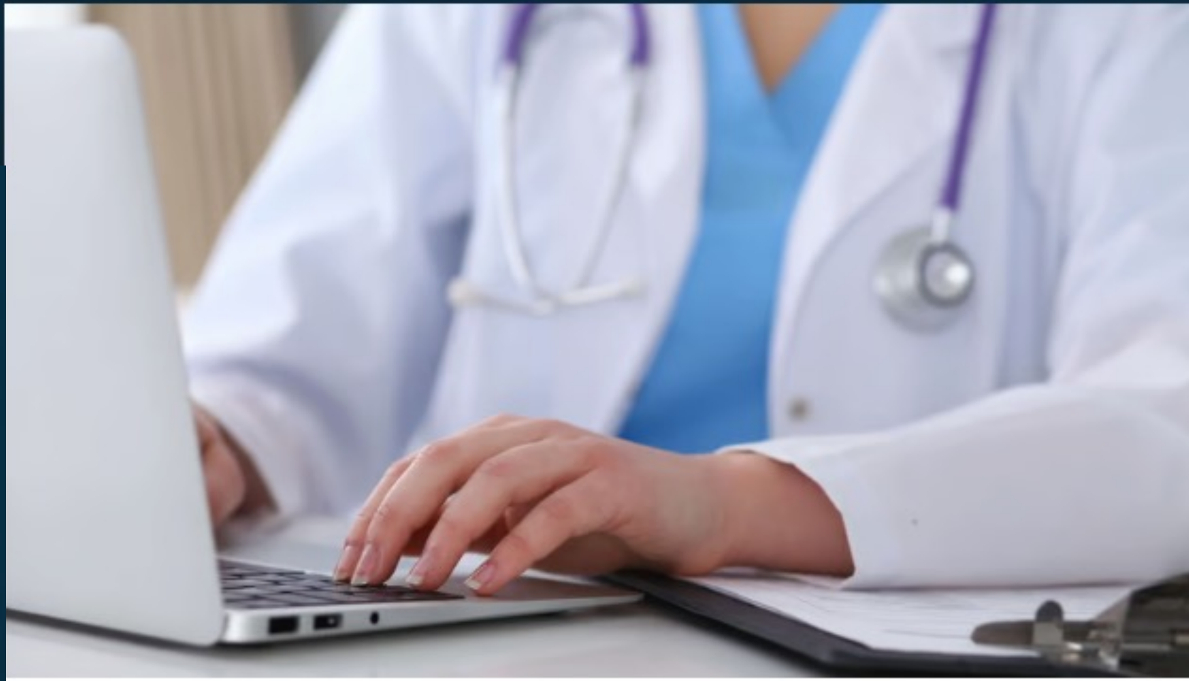
AI is promising, but …

https://www.fiercehealthcare.com/practices/nearly-half-u-s-doctors-say-they-are-anxious-about-using-ai-powered-software-survey

Nearly half of U.S. doctors say they are anxious about using AI-powered software: survey

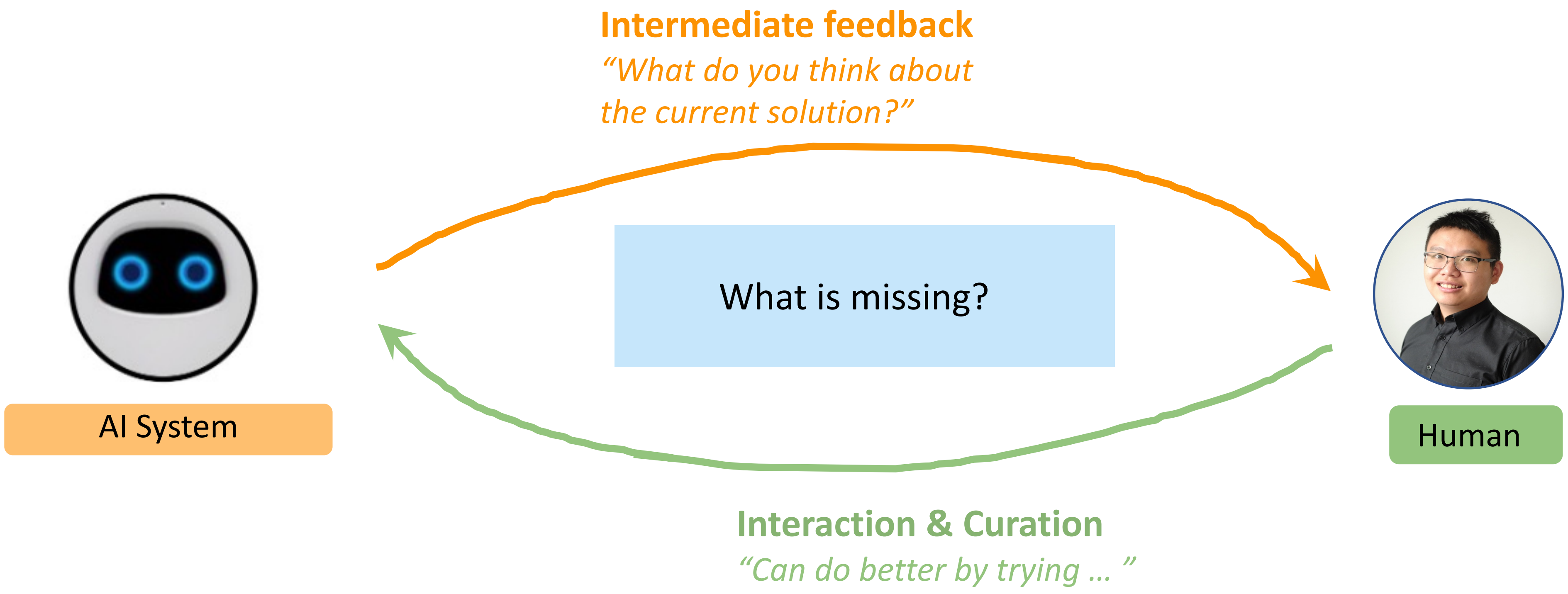
By **Heather Landi** · Apr 25, 2019 10:55am

Artificial Intelligence | Clinical Decision Support | Digital health | machine learning



A new physician survey indicates artificial intelligence applications are still in their infancy and have not affected mainstream physician practice at scale. (Getty/andrei_r)

A general blueprint of Human-AI teaming



General blueprint for a human-in-the-loop interactive AI system. Image modified from:
https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems

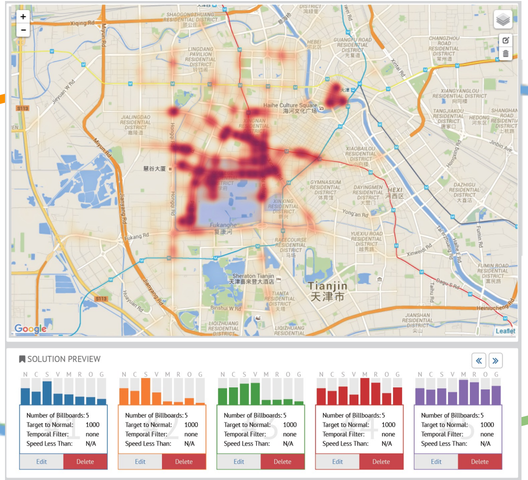
Visualization-powered teaming workflow

Intermediate feedback
"What do you think about the current solution"

AI System (Human-Centered)

"If we want it to play a positive role in tomorrow's world, it must be guided by human concerns"

Feifei Li (Stanford's Human-Centered AI Institute)



Visual Interface

Human

Interaction & Curation
"Can do better by trying … "

Human-AI teaming is essential, when

AI requires significant human knowledge to enhance its performance -> **Ability to learn**

Large Devices
Health Monitoring

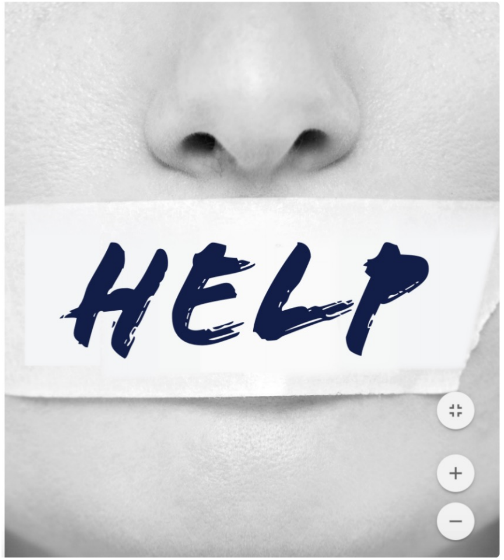
Liu, et al., MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series, CSCW 2022.

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Human-AI teaming is essential, when

AI requires significant human knowledge to enhance its performance

Decisions being made are high-stakes -> **Transparency**



Child Abuse
Hotline Screening

Illness Diagnosis
& Treatment

Zytek, **Liu**, et al., Sibyl: Understanding and Addressing the Usability Challenges of Machine Learning In High-Stakes Decision Making, TVCG (VIS'21).

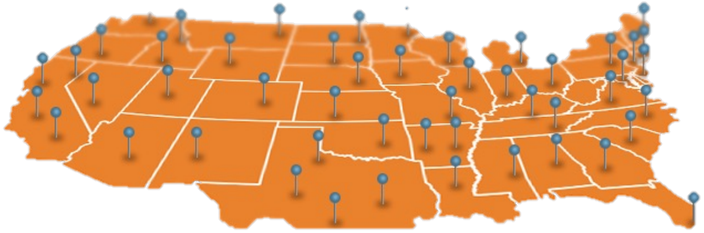
Cheng, **Liu**, et al., VBridge: Connecting the Dots Between Features and Data to Explain Healthcare Models, TVCG (VIS'21).
Best Paper Honorable Mention.

Human-AI teaming is essential, when

AI requires significant human knowledge to enhance its performance

Decisions being made are high-stakes

Decision-making involves multiple criteria and is heavily influenced by the context -> **Steerability**



Advertising Campaign Planning

Liu, et al, SmartAdP: Visual Analytics of Large-scale Taxi Trajectories for Selecting Billboard Locations, TVCG (VAST'16).

Store Operation Optimizing

Liu, et al., TPFlow: Progressive Partition and Multidimensional Pattern Extraction for Large-scale Spatio-temporal Data Analysis, TVCG (VAST'18), Best Paper Award.

Sintel

Signal Intelligence

(ability to learn)

Human-AI teaming for time series data analytics

Sintel https://sintel.dev/

Featurization

Featurize time series with domain knowledge encoded for machine learning uses.

Forecasting

Predict future values by analyzing past trends.

Anomaly Detection

Identify anomalous time series segments.

Classification

Classify time series segments into particular categories.

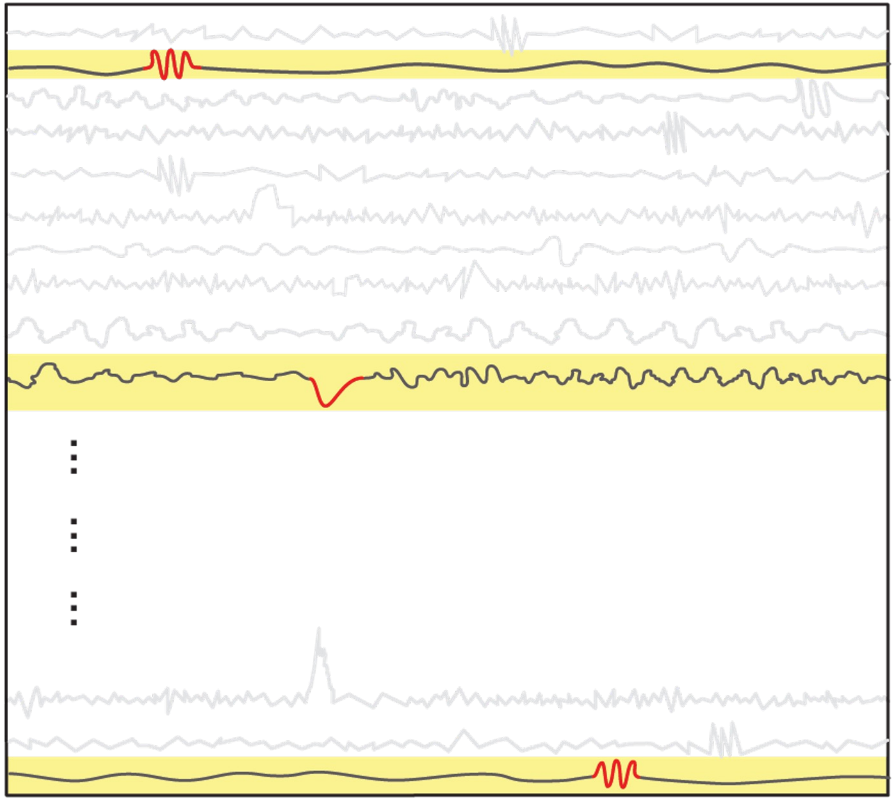
Human-in-the-loop workflow to transfer insights into actions in minutes

Motivation

Wind turbines

Satellites

Air quality monitors



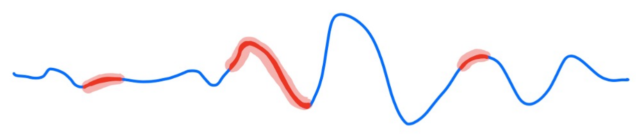
> 30k signals

How can we effectively monitor and analyze **anomalies** facing such massive amount of data?

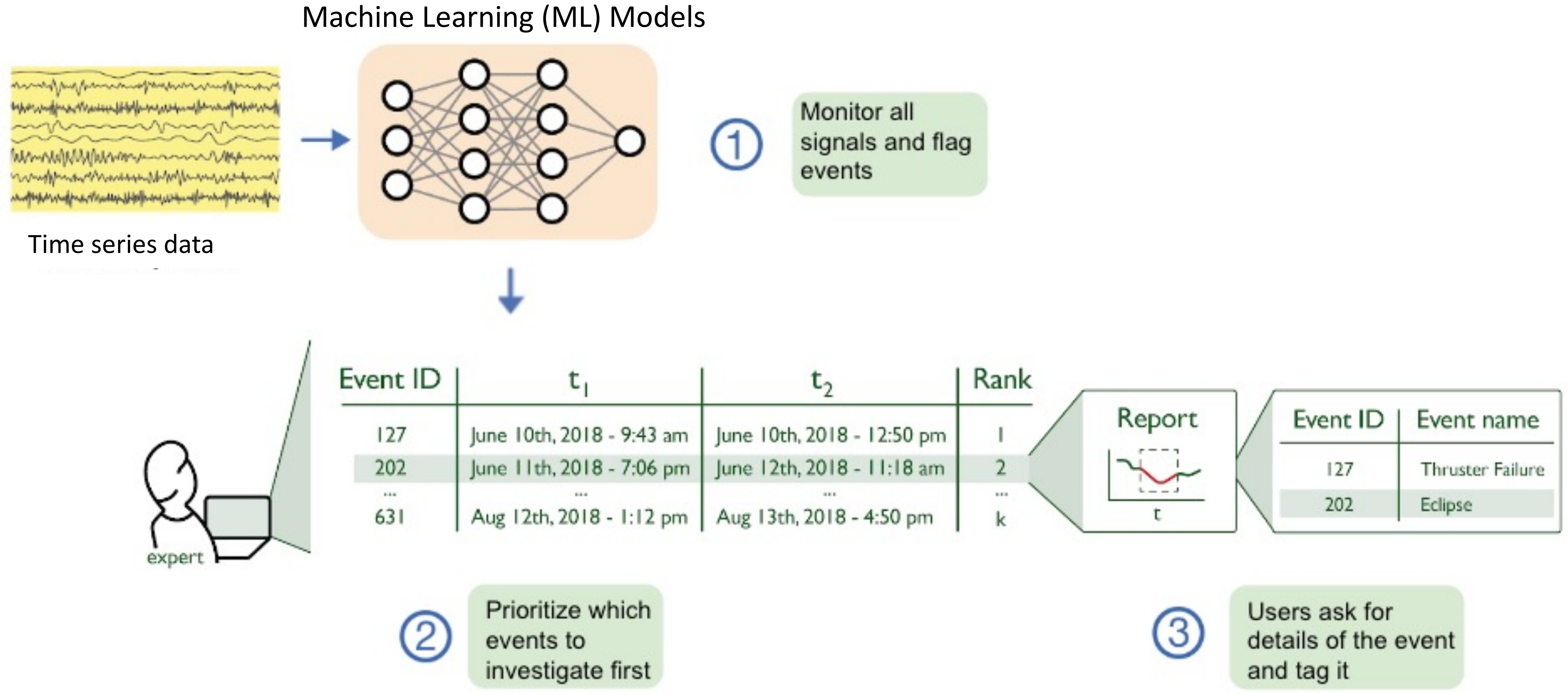
What is time series anomaly detection?

- Given a time series $X = (x^1, x^2, \ldots, x^T)$

- Find $A_{seq} = \{\mathbf{a}_{seq}^1, \mathbf{a}_{seq}^2, \ldots, \mathbf{a}_{seq}^k\}$, where $\mathbf{a}_{seq}^i$ is a **continuous sequence of data points** over time that show anomalous or unusual behavior.



The problem we want to solve

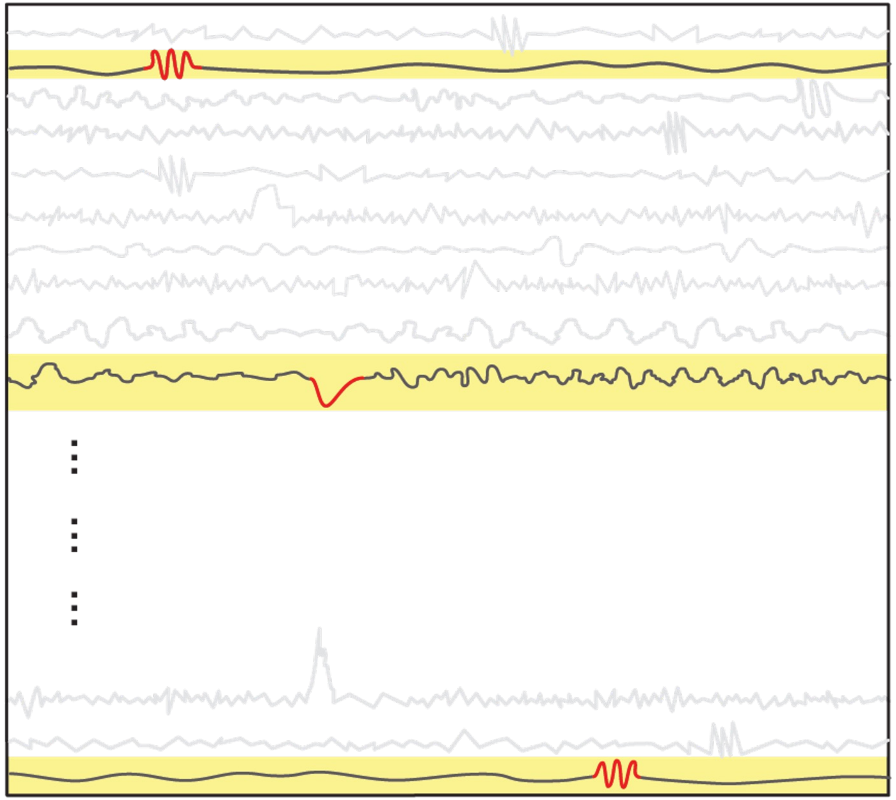


The problem we want to solve

Wind turbines

Satellites

Air quality monitors



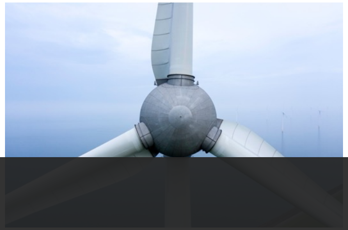
AI System

Machine (AI) challenges:

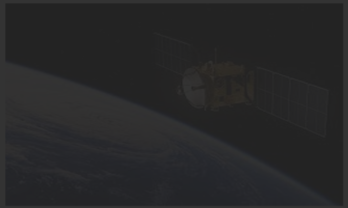
- No labeled data
- No normal baselines

The problem we want to solve

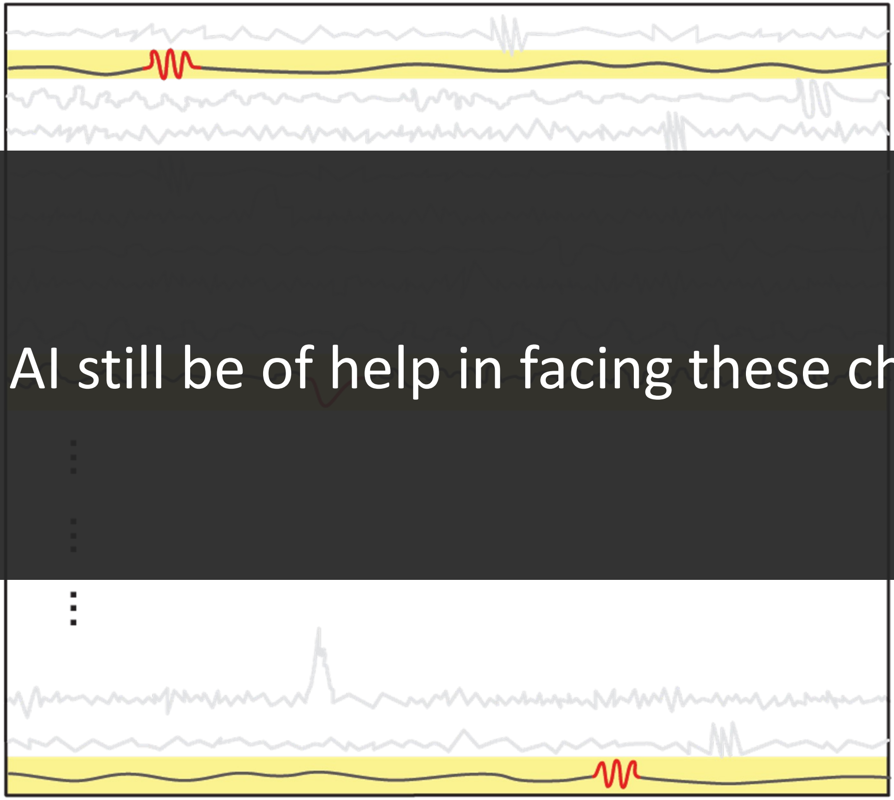
Wind turbines



Satellites



Air quality monitors



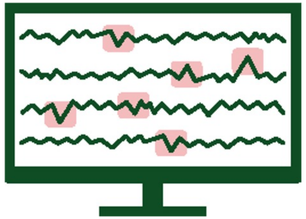
Human

Human factors:

- No ML expertise
- Need to integrate domain knowledge
- No easy way to document findings

Can AI still be of help in facing these challenges?

Human-AI teaming workflow



1 Extract anomalies from massive time series with ML

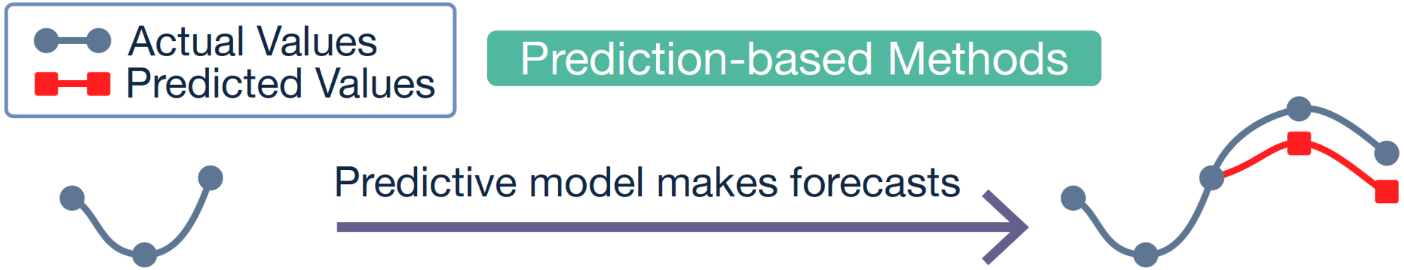
Sintel (*SIGMOD'22*)

6 Enhance ML with human annotations

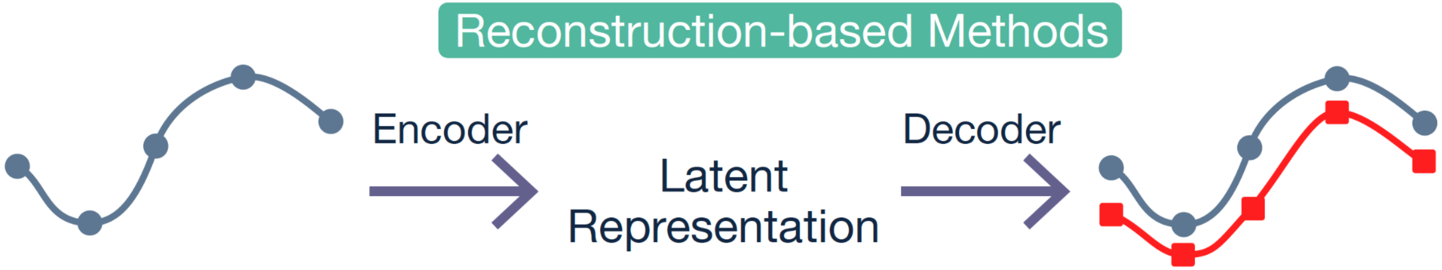
Visualization system keeps collecting user feedback

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Unsupervised anomaly detection



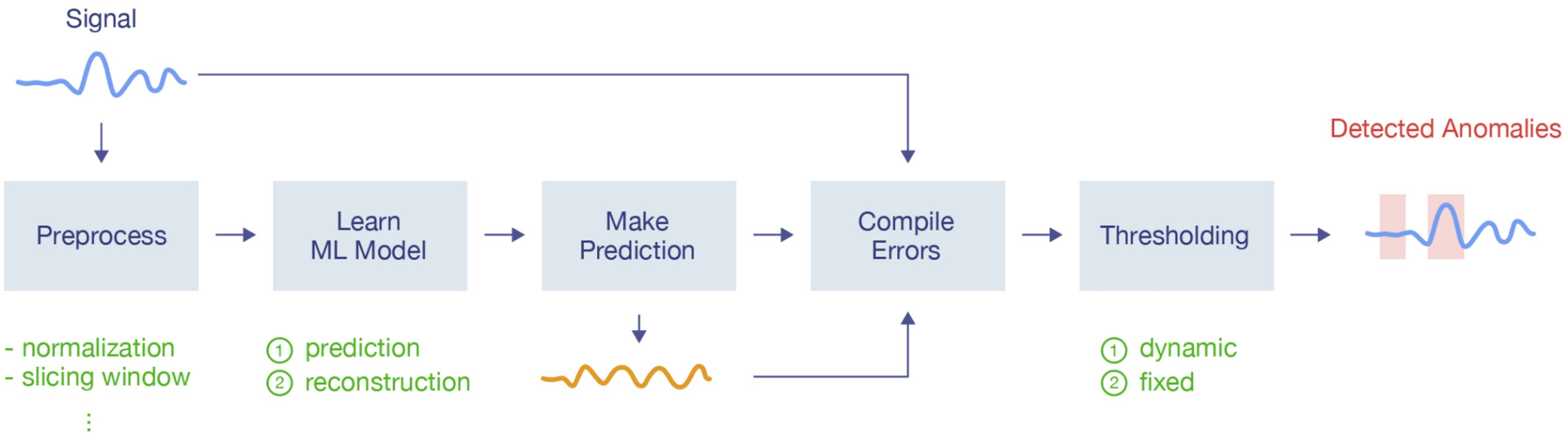
Assumption: Anomalous values cannot be predicted as well as the normal ones.



Assumption: Anomalies cannot be effectively reconstructed since information is lost in the mapping to the latent dimensions.

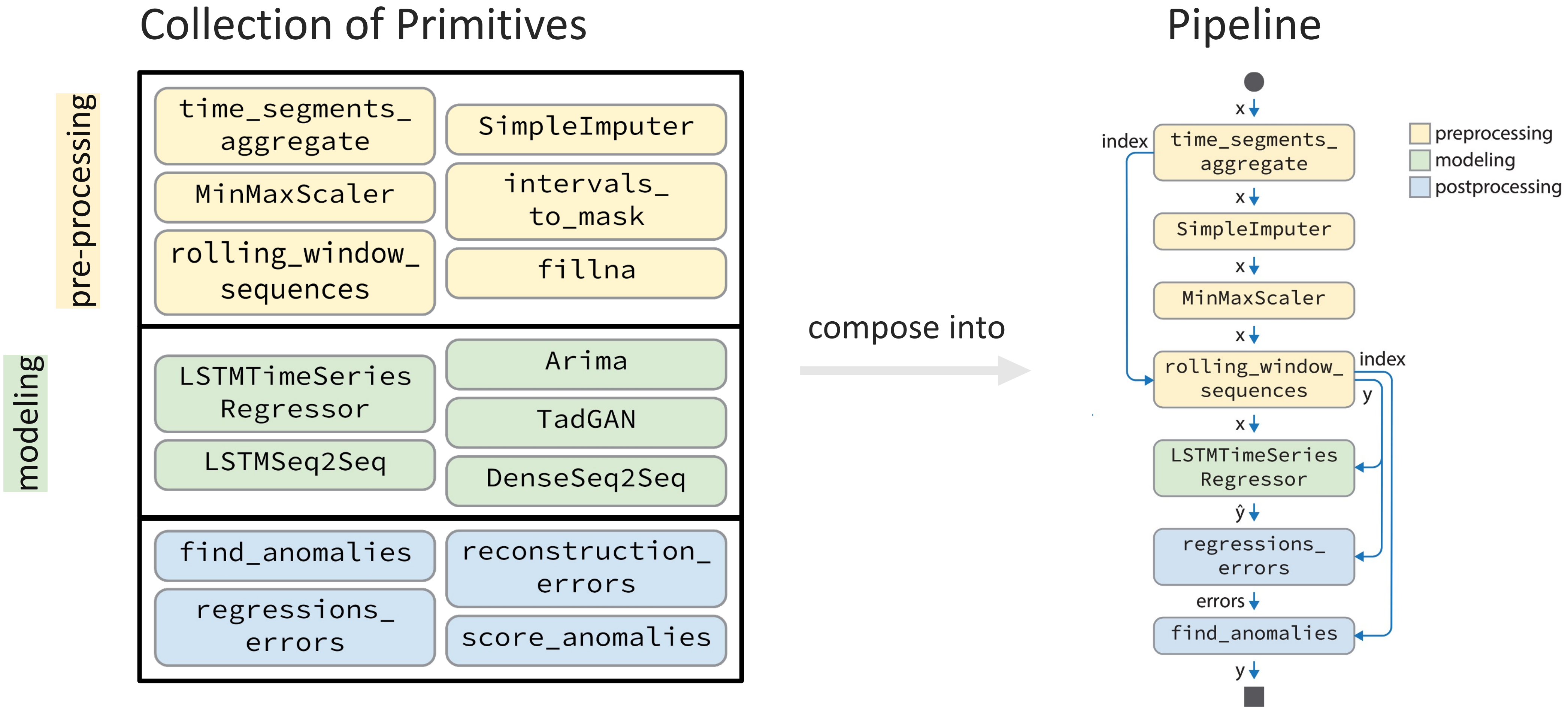
Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Use Sintel for anomaly detection



Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Primitives and Pipelines



Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

What does Sintel achieve?

Integrate domain expertise

pre-processing

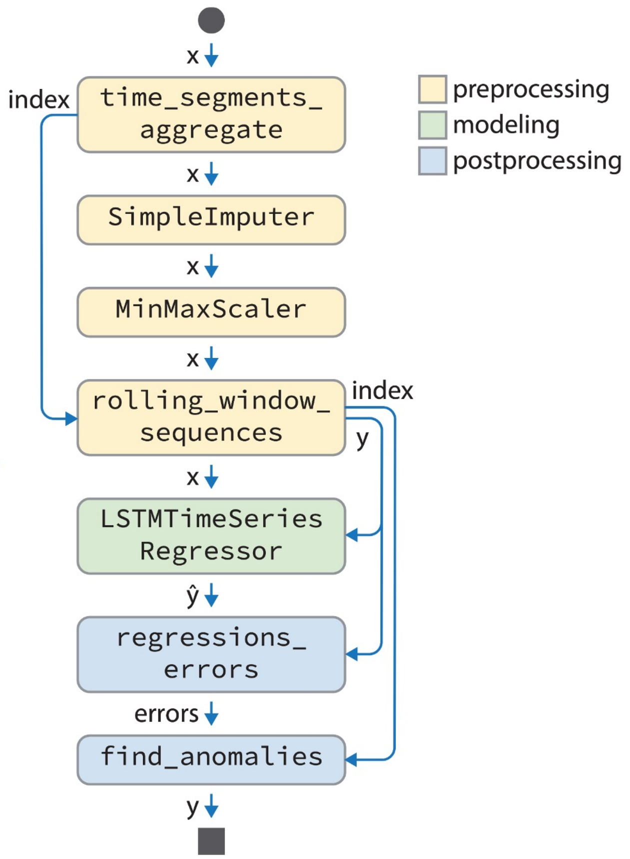
Satellite experts:

- Use *zero-order hold* to impute missing values instead of *mean*

Wind turbine experts:

- Need domain specific aggregation and transformation methods (e.g., *fft*)

Pipeline

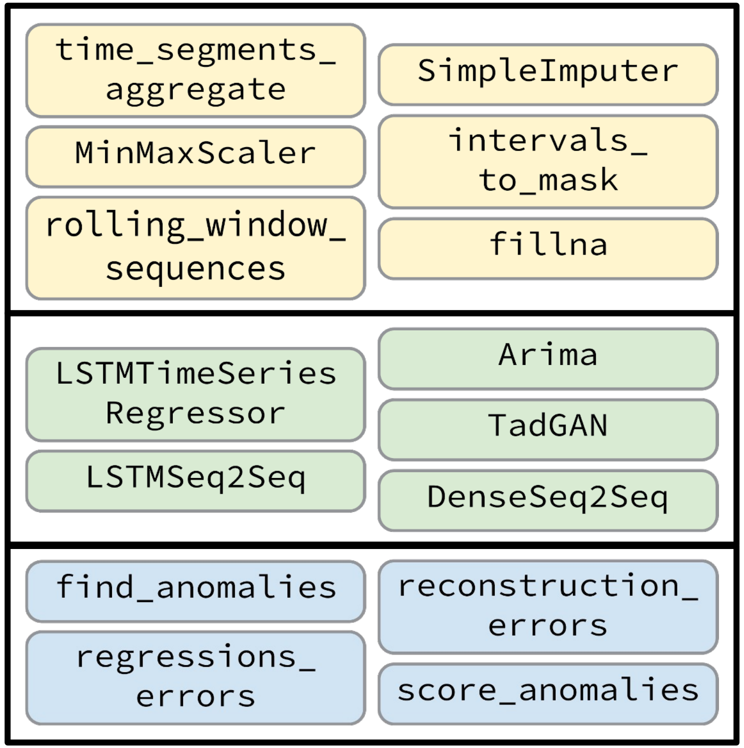


Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

What does Sintel achieve?

Develop better models

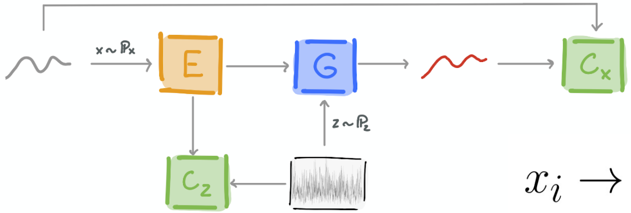
modeling



We now have in total 9 different models integrated:

TadGAN

Liu*, Geiger*, et al., TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks, IEEE BigData 2020



$$x_i \to \mathcal{E}(x_i) \to \mathcal{G}(\mathcal{E}(x_i)) \approx \hat{x}_i$$

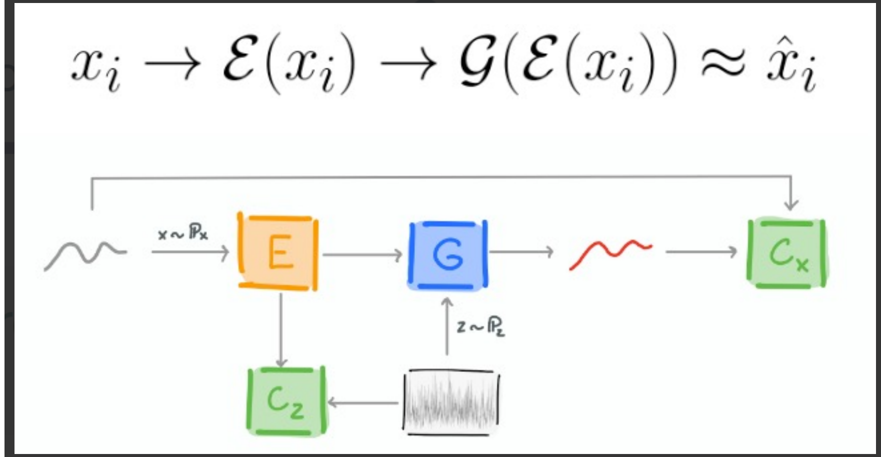
AER

Wong, **Liu**, et al., AER: Auto-Encoder with Regression for Time Series Anomaly Detection, IEEE BigData 2022

$$Loss = \frac{\gamma}{2} V_{pred}(t_{i-1}, r_{i-1}) + \frac{\gamma}{2} V_{pred}(t_{i+n}, f_{i+n}) + (1-\gamma) V_{rec}(t_{i:i+n-1}, y_{i:i+n-1})$$

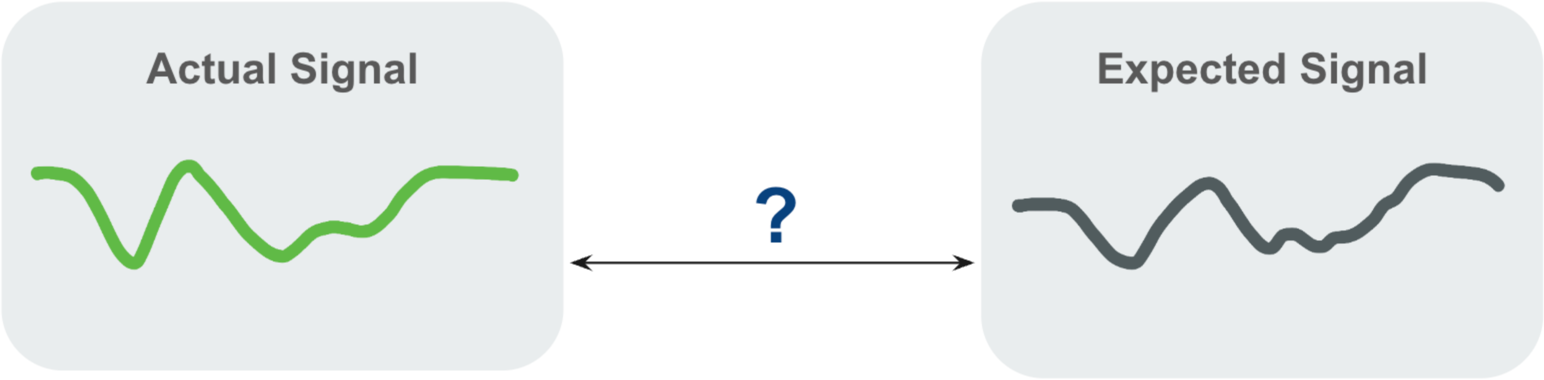
Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

TadGAN



Liu*, Geiger*, et al., TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks, IEEE BigData 2020

Measure the discrepancies



1. **Reconstruction Error**: finding how much deviation there is between the real and the reconstructed signal.
2. **Critic score**: leveraging the trained critic to distinguish between real and reconstructed samples.

Measure the discrepancies

1. **Convex combination.**

$$\mathbf{a}(x) = \alpha Z_{RE}(x) + (1-\alpha) Z_{\mathcal{C}_x}(x)$$

where α controls the relative importance of the two terms (by default alpha = 0.5).

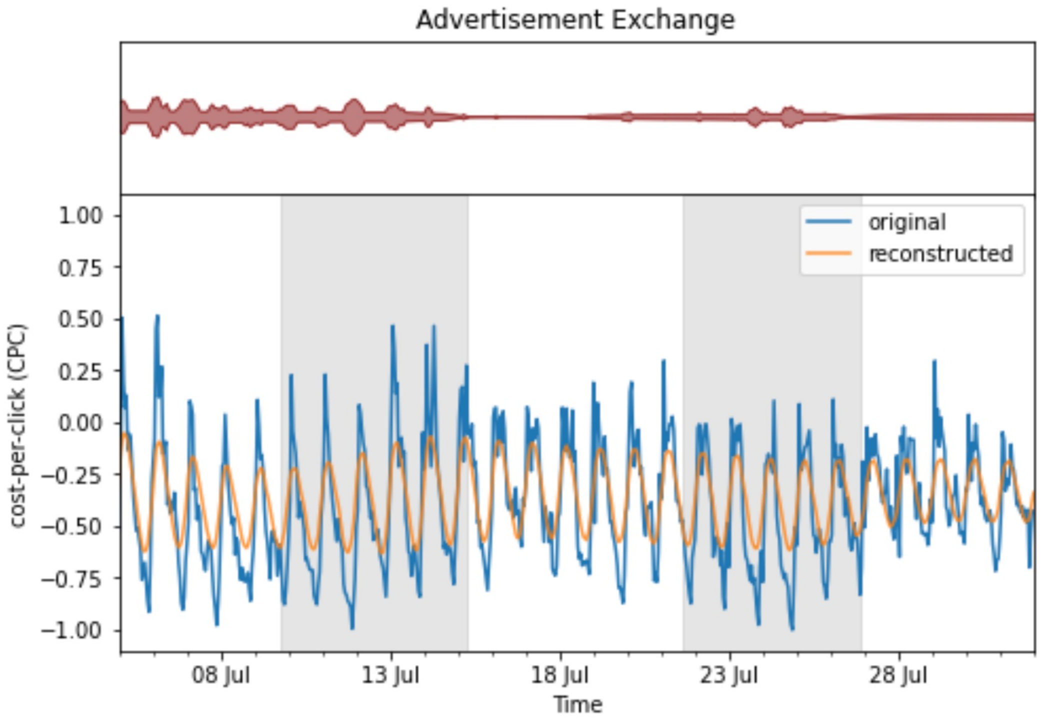
2. **Multiplication.**

$$\mathbf{a}(x) = \alpha Z_{RE}(x) \odot Z_{\mathcal{C}_x}(x)$$

where α defaults to 1.

Measure the discrepancies

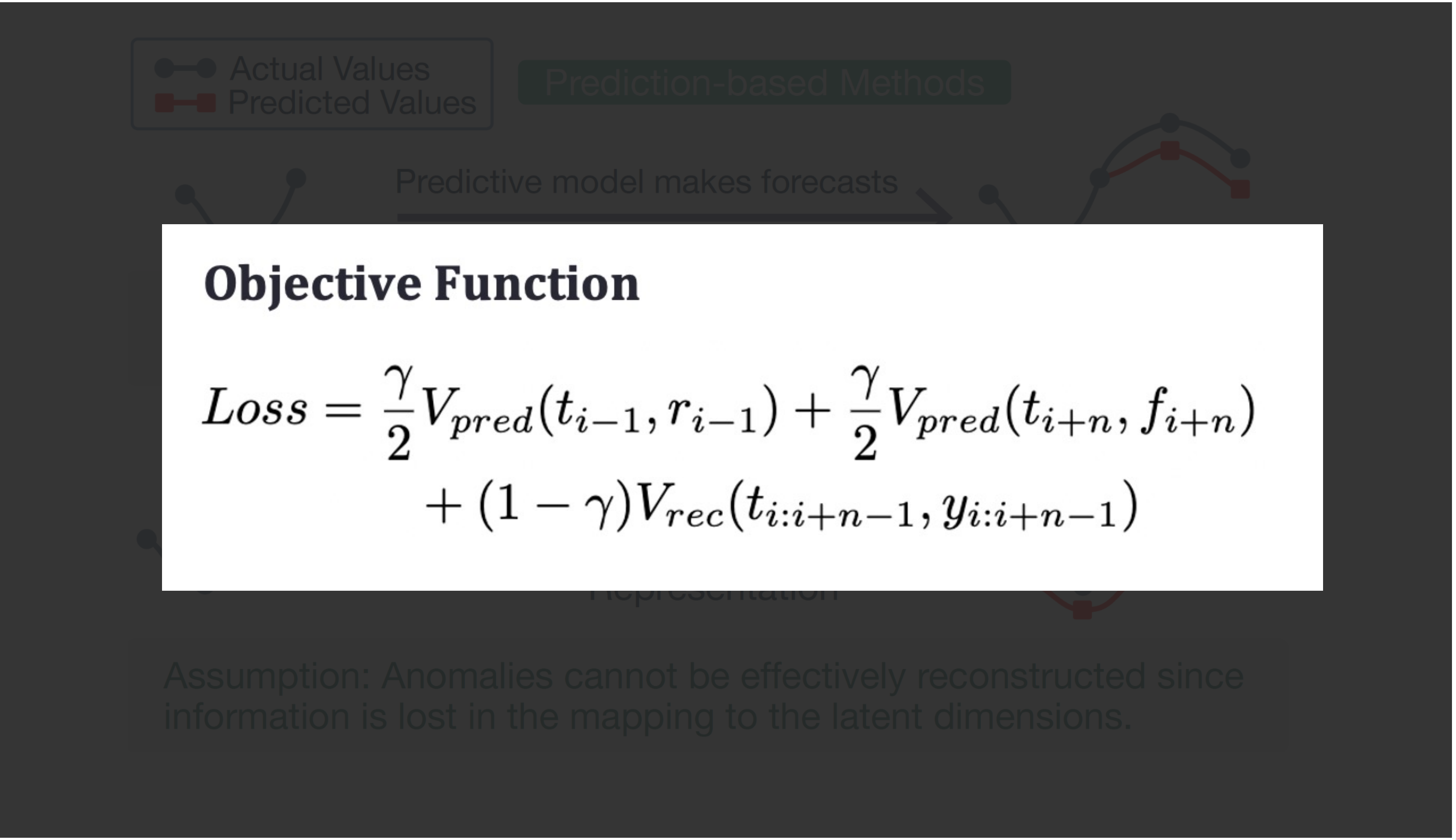
Identify anomalous intervals with **locally adaptive thresholding**



https://arxiv.org/pdf/1802.04431.pdf

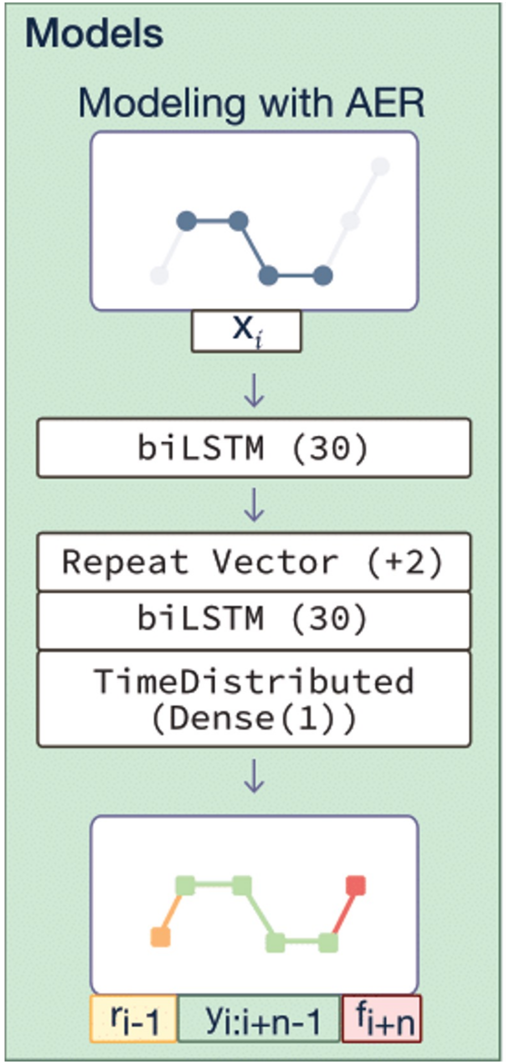
Hundman, K., Constantinou, V., Laporte, C., Colwell, I. and Soderstrom, T., 2018, July. **Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding**. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 387-395).

AER



Objective Function

$$Loss = \frac{\gamma}{2} V_{pred}(t_{i-1}, r_{i-1}) + \frac{\gamma}{2} V_{pred}(t_{i+n}, f_{i+n}) + (1-\gamma) V_{rec}(t_{i:i+n-1}, y_{i:i+n-1})$$



Wong, **Liu**, et al., AER: Auto-Encoder with Regression for Time Series Anomaly Detection, IEEE BigData 2022

What does Sintel achieve?

Develop better models

- One single call to benchmark algorithms and know which one is the best

```
from sintel import benchmark

pipelines = ['arima', 'lstm_dynamic_threshold', '...']
datasets = ['NAB', 'NASA', '...']
metrics = ['f1', 'accuracy', '...']

benchmark(pipelines=pipelines,
          datasets=datasets,
          metrics=metrics,
          rank='f1')

# >>>
#                  pipeline  rank  accuracy  elapsed     f1        precision  recall
# 0  lstm_dynamic_threshold     1  0.986993  915.958132  0.846868  0.879518   0.816555
# 1                   arima     2  0.962160  319.968949  0.382637  0.680000   0.266219
```

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

What does Sintel achieve?

Develop better models

Comparing against ARIMA whose roots are in: Box, George; Jenkins, Gwilym (1970). *Time Series Analysis: Forecasting and Control*

TadGAN achieves the best overall improvement with an over 15% improvement.

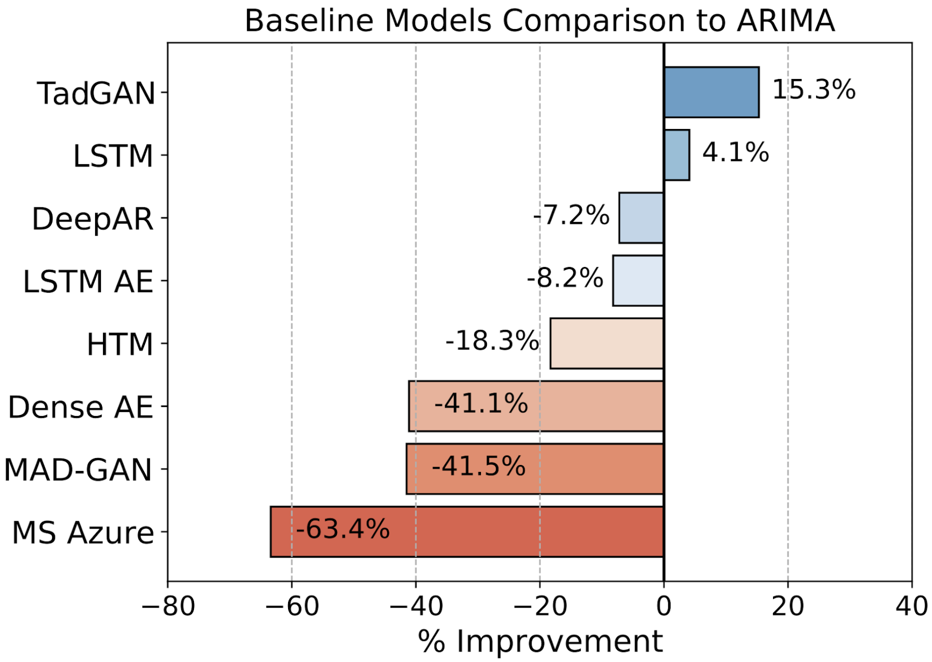


Fig. 3. Comparing average F1-Scores of baseline models across all datasets to ARIMA. The x-axis represents the percentage of improvement over the ARIMA score by each one of the baseline models.

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

What does Sintel achieve?

Develop better models

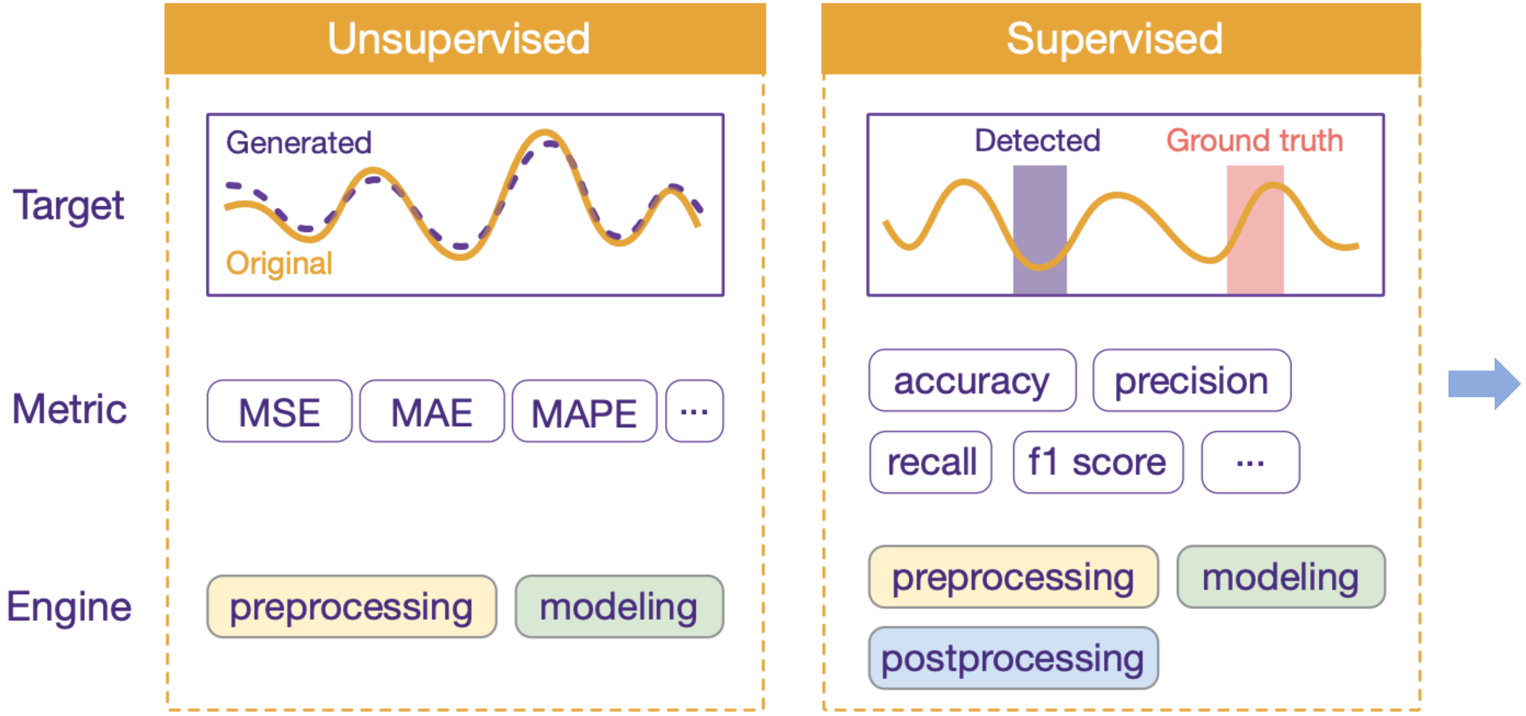
Models	NASA		YAHOO				NAB					UCR	Avg. F1 ($\mu \pm \sigma$)
	MSL	SMAP	A1	A2	A3	A4	Art	AdEx	AWS	Traffic	Tweets	UCR	
ARIMA	0.442	0.333	0.733	0.807	0.818	0.700	0.353	0.518	**0.741**	0.500	0.567	0.124	0.553 ± 0.21
LSTM-DT	0.515	0.707	0.721	**0.980**	0.744	0.638	0.400	0.513	**0.741**	**0.667**	0.580	0.391	0.633 ± 0.16
LSTM-AE	0.500	0.705	0.610	0.866	0.420	0.253	0.545	**0.750**	0.692	0.457	0.483	0.314	0.550 ± 0.17
LSTM-VAE	0.526	0.653	0.575	0.823	0.432	0.240	**0.667**	0.700	0.643	0.483	**0.590**	0.317	0.554 ± 0.16
TadGAN	**0.584**	0.617	0.533	0.842	0.391	0.297	0.571	0.677	0.720	0.581	0.588	0.162	0.547 ± 0.18
AER*	0.541	**0.772**	**0.772**	0.959	**0.896**	**0.722**	0.615	0.635	0.621	0.606	0.585	**0.470**	**0.683 ± 0.14**

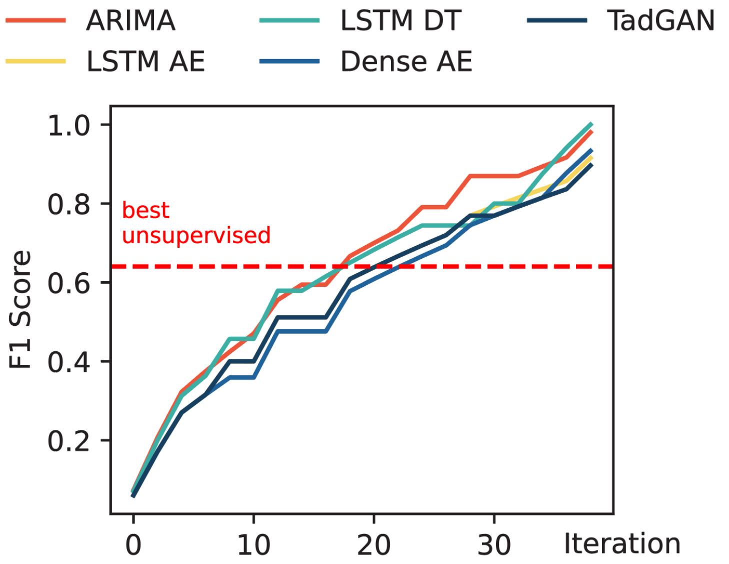
The latest and full results can be found here: https://bit.ly/orion-benchmark

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

What does Sintel achieve?

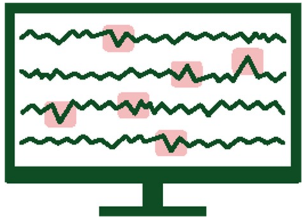
Improve over time





Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Human-AI teaming workflow



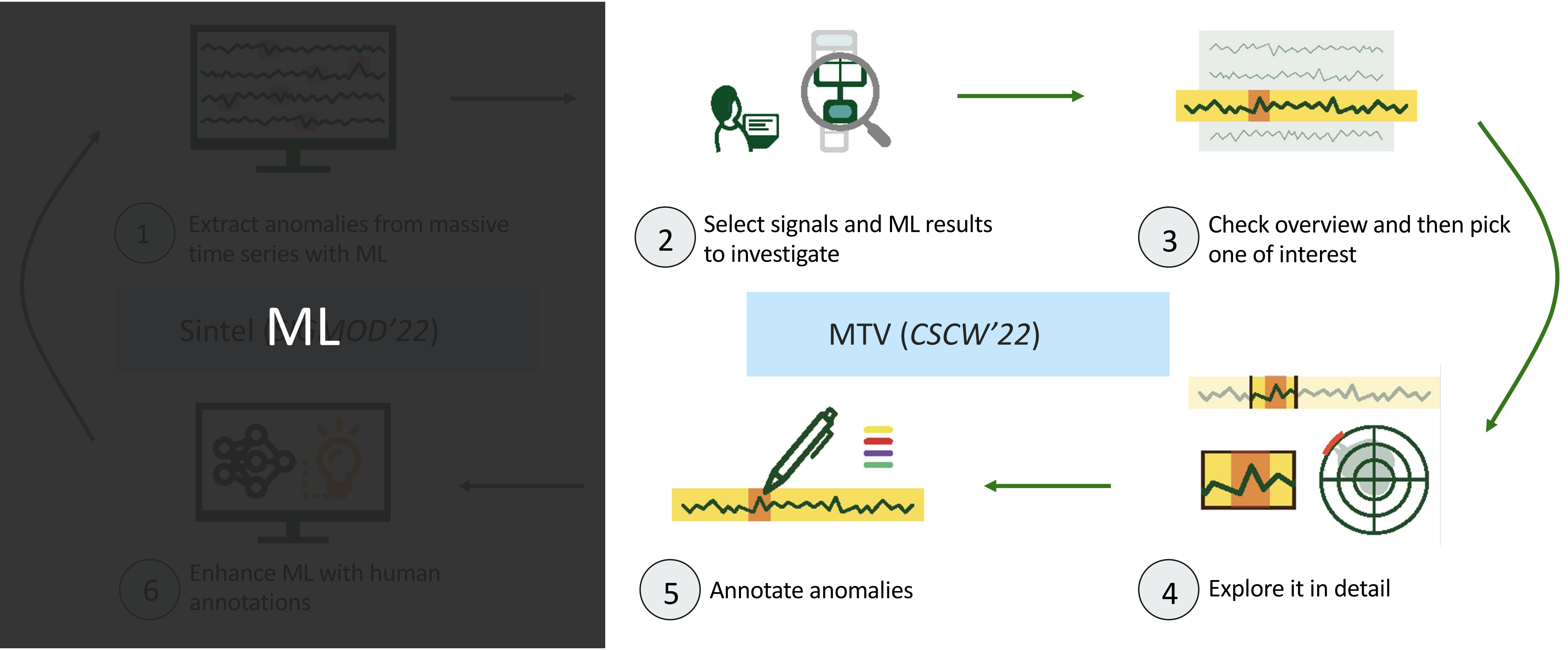
1 Extract anomalies from massive time series with ML

Sintel (*SIGMOD'22*)

6 Enhance ML with human annotations

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Human-AI teaming workflow

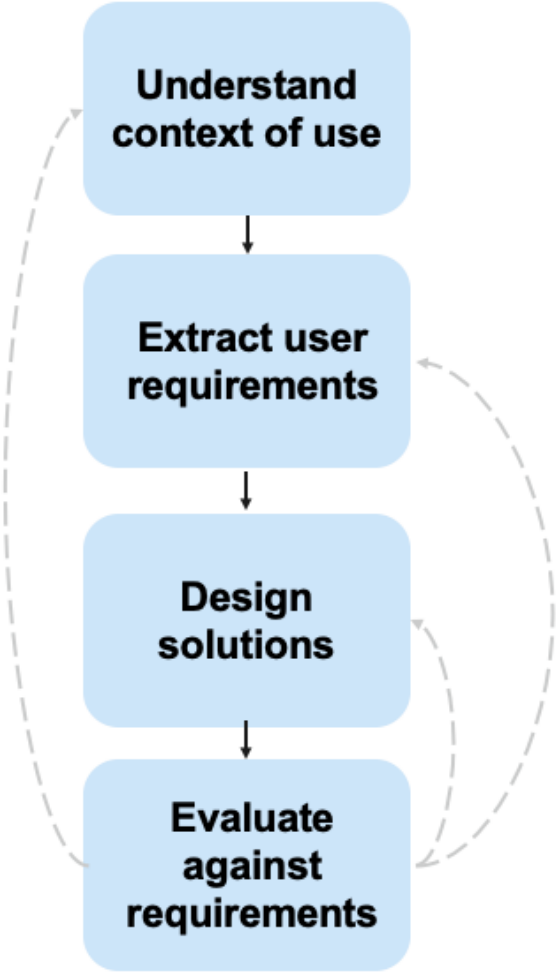


Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Working with domain experts

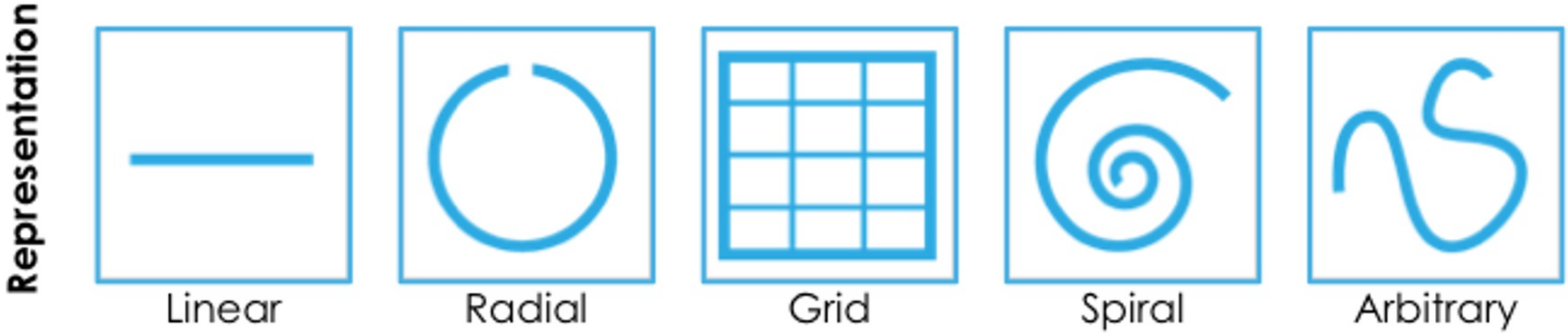
- Collaborated with 9 experts
  - 6 from a satellite operations company
  - 3 from a renewable energy company

- Followed an iterative user-centered design process
  - 6 design requirements



Liu, et al., MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series, CSCW 2022.

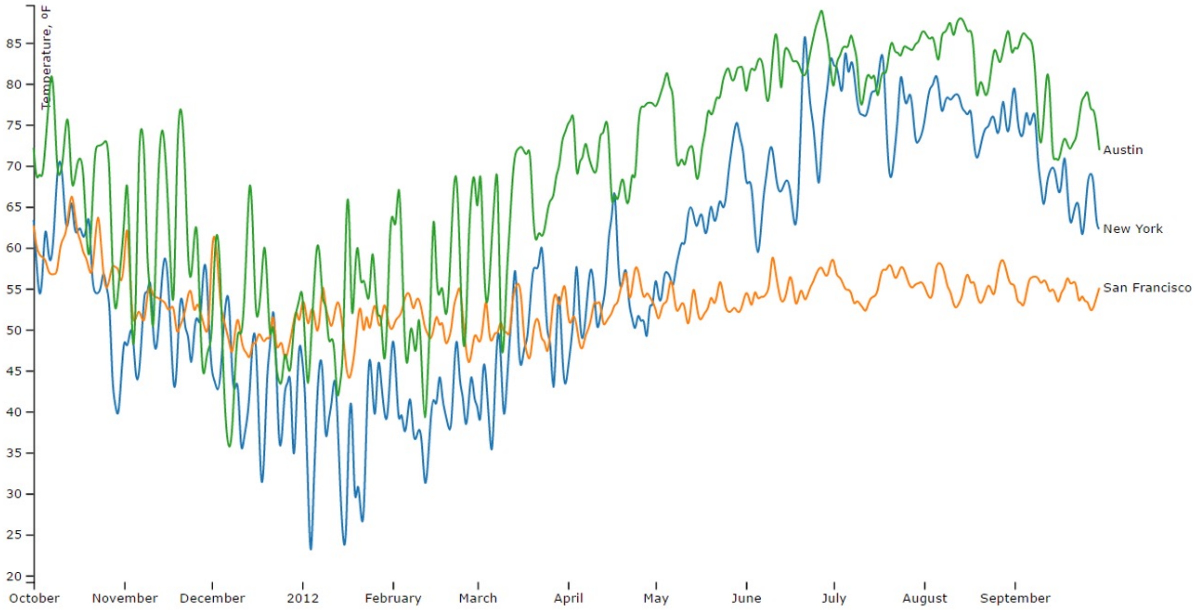
Time Representation



Brehmer, Matthew, et al. "Timelines revisited: A design space and considerations for expressive storytelling." TVCG 23.9 (2016): 2151-2164.

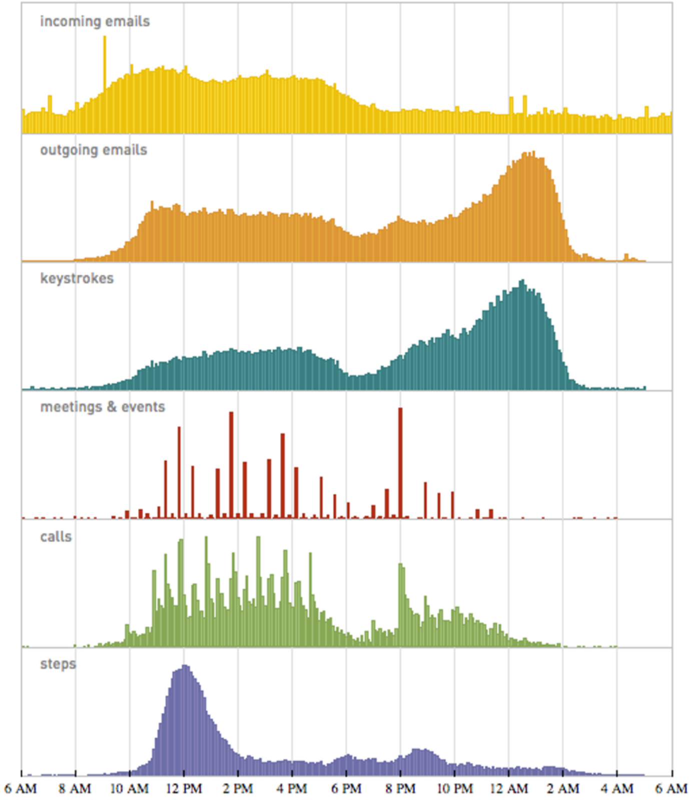
Linear Time

- Present time data as a 2D line graph and
  - Time on x-axis
  - The other variable on y-axis



Linear Time

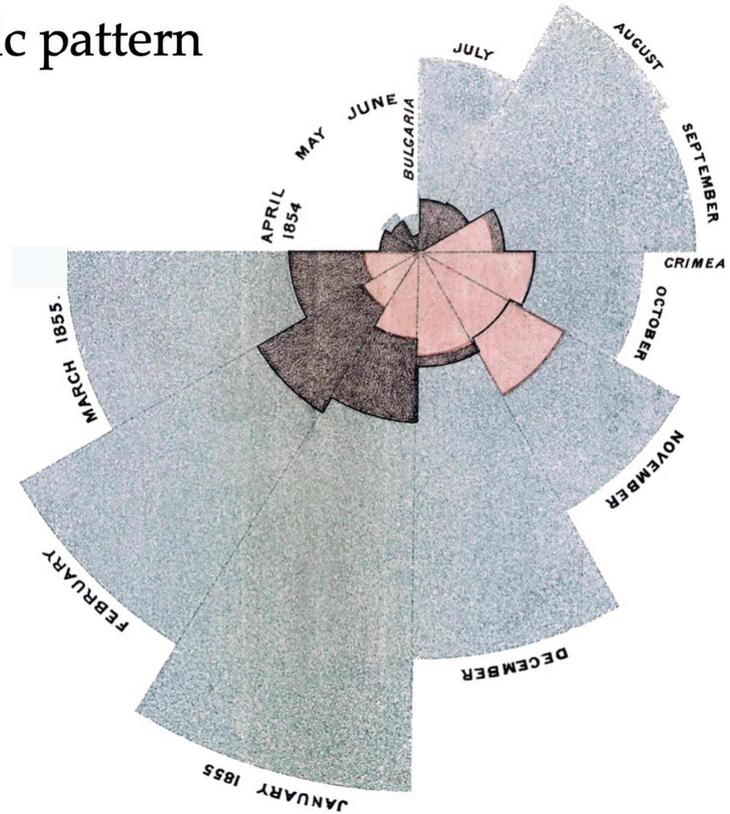
- Stephen Wolfram's *Personal Data Visualization Report*



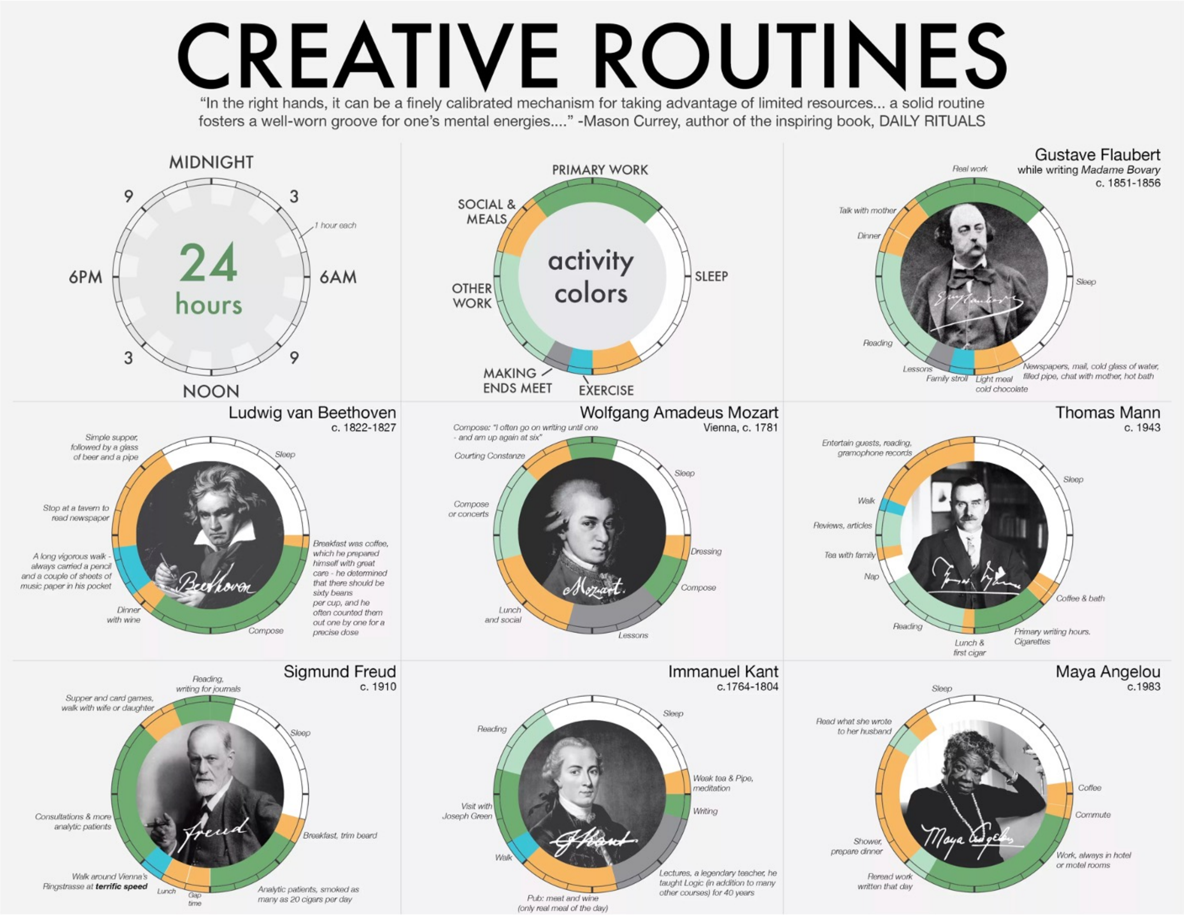
http://blog.stephenwolfram.com/2012/03/the-personal-analytics-of-my-life/

Radial Time

- Data distributed along the spiral
- To reveal cyclic pattern



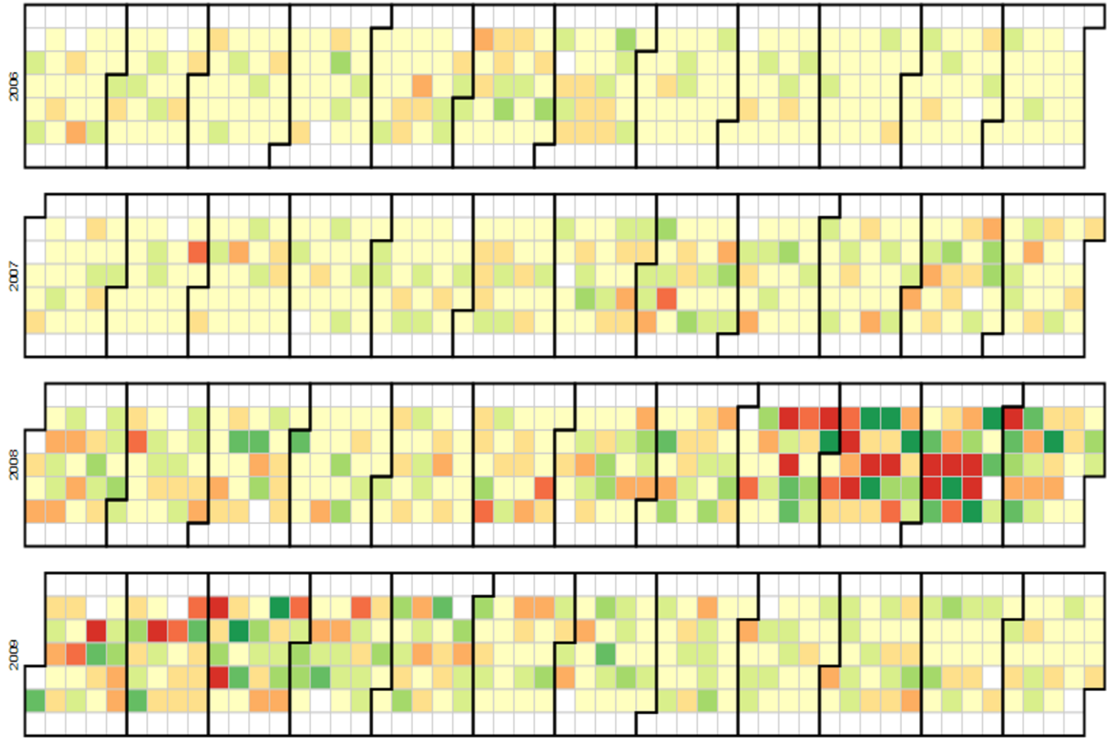
Radial Time



https://infowetrust.com/project/routines

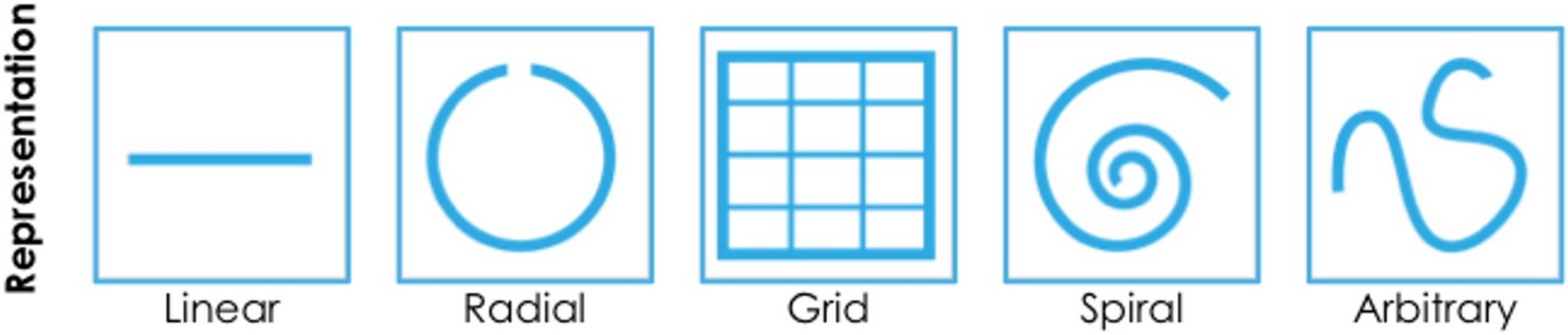
Grid Time

- Dow Jones stock price from 2006 to 2009



http://mbostock.github.io/d3/talk/20111018/calendar.html

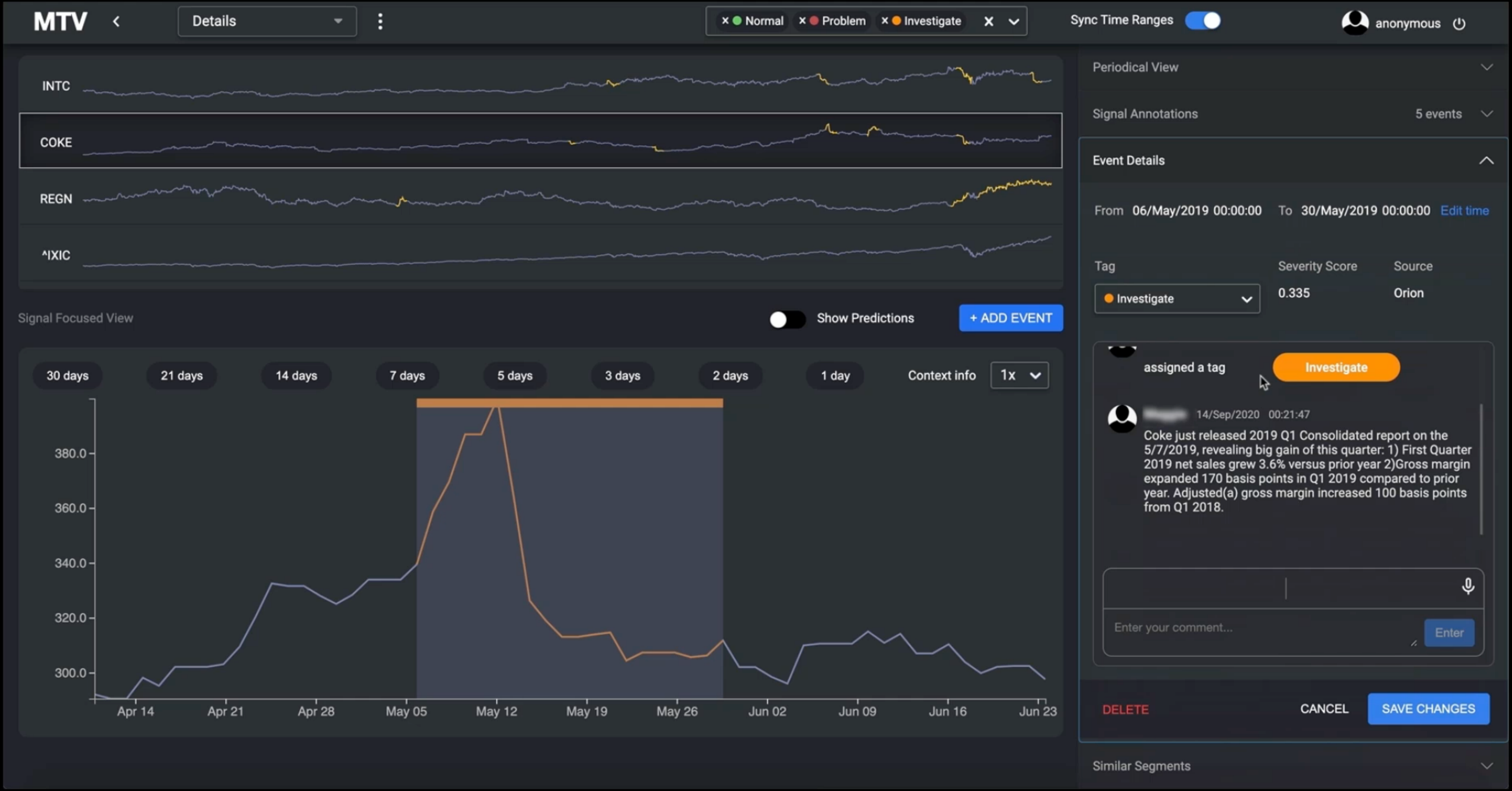
Time Representation



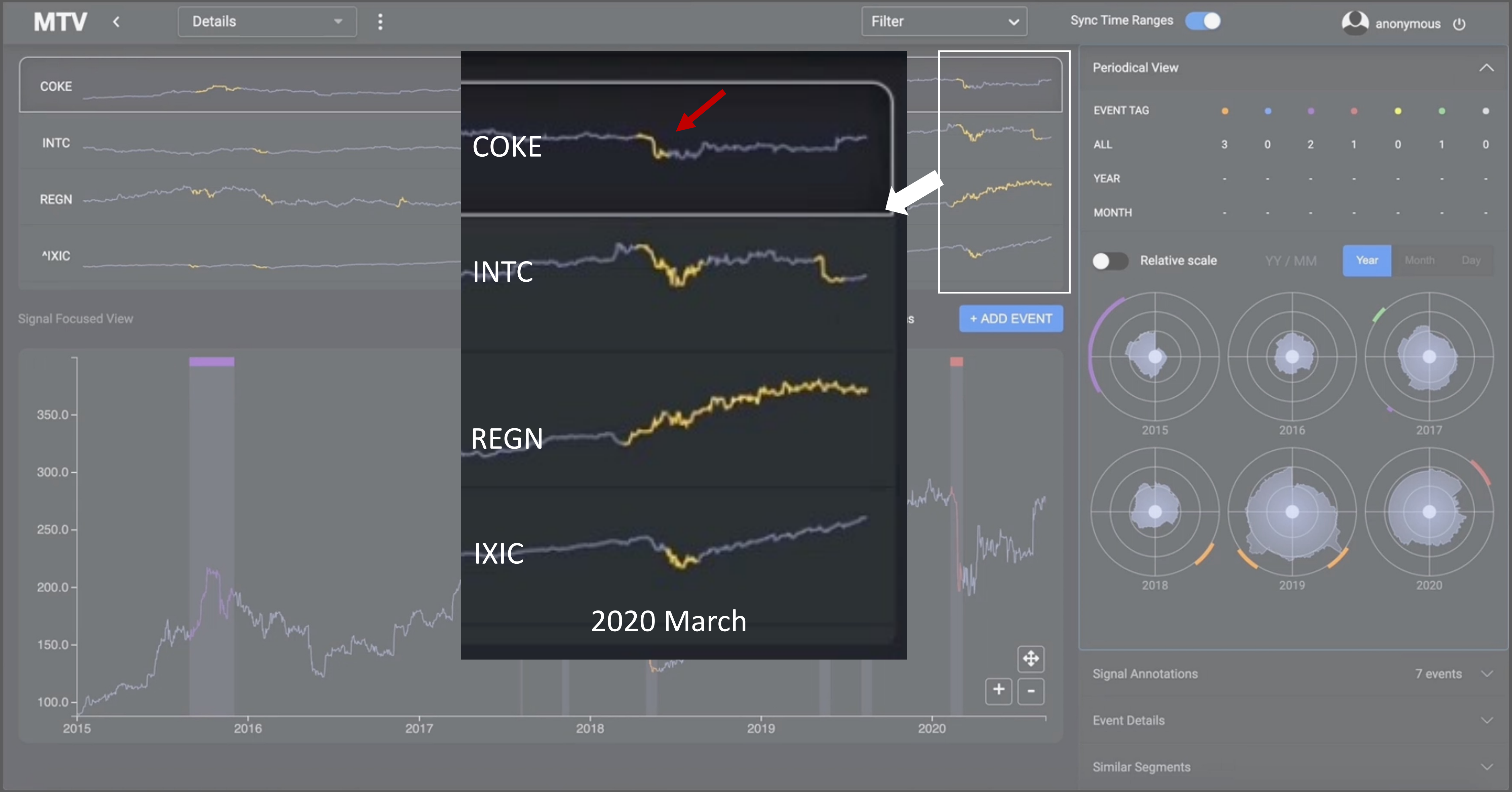
- Among the above four types, linear time representation is more popular


MTV
Details
Filter
Sync Time Ranges
anonymous
COKE
INTC
REGN
^IXIC
Periodical View
EVENT TAG
ALL 3 0 2 1 0 1 0
YEAR
MONTH
Relative scale
YY / MM
Year
Month
Day
2015
2016
2017
2018
2019
2020
Signal Focused View
Show Predictions
+ ADD EVENT
350.0
300.0
250.0
200.0
150.0
100.0
2015
2016
2017
2018
2019
2020
Signal Annotations
7 events
Event Details
Similar Segments


MTV
Details
Filter
Sync Time Ranges
anonymous
COKE
INTC
REGN
^IXIC
Signal Focused View
Show Pred
350.0
300.0
250.0
200.0
150.0
100.0
2015
2016
2017
2018
2019
2020
Periodical View
Day
2015
2016
2017
2018
2019
2020
Signal Annotations
7 events
Event Details
Similar Segments
Advertisement Exchange
original
reconstructed
cost-per-click (CPC)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
08 Jul
13 Jul
18 Jul
23 Jul
28 Jul
Time

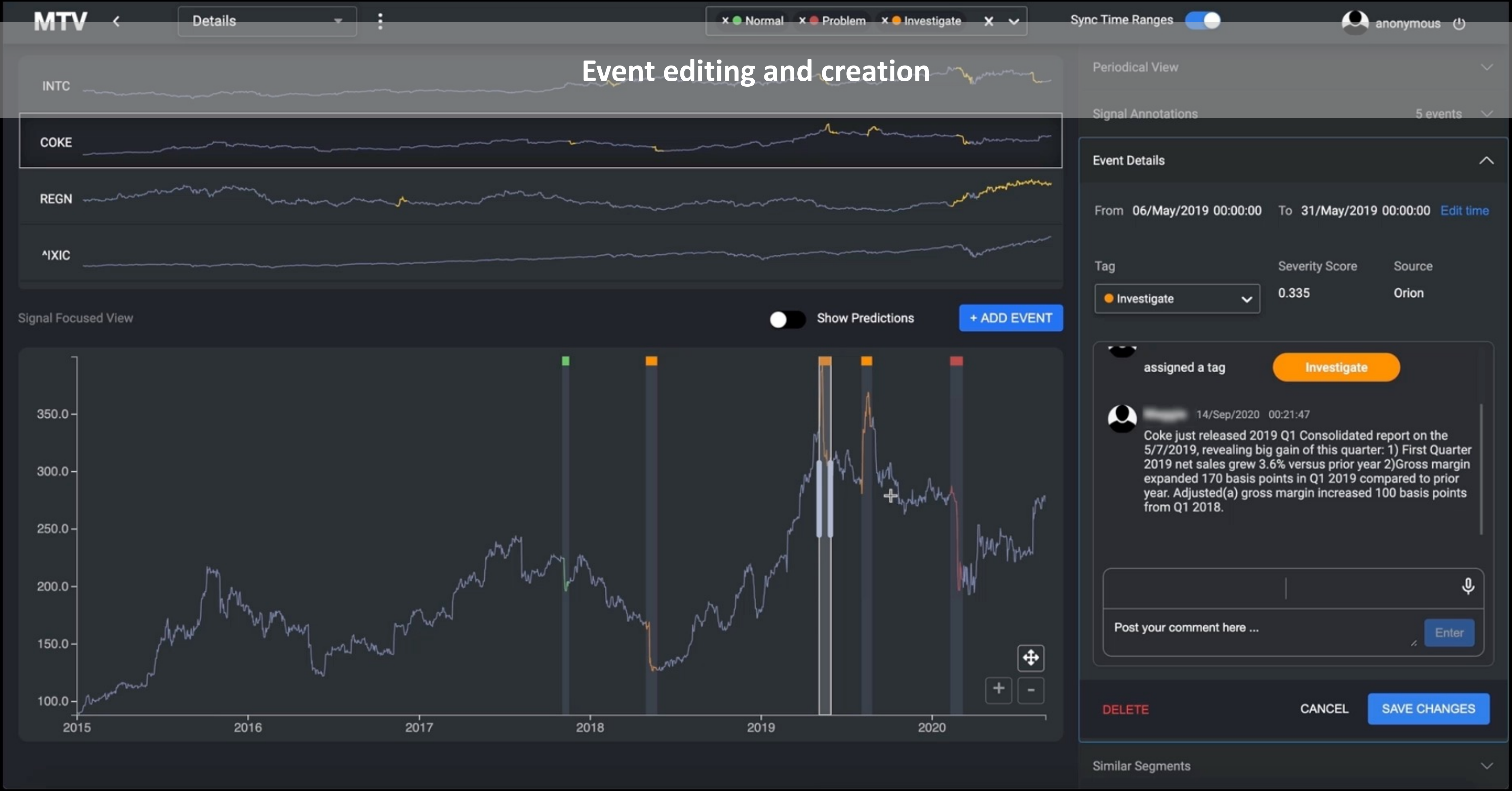

MTV
Details
Normal
Problem
Investigate
Sync Time Ranges
anonymous
INTC
COKE
REGN
^IXIC
Periodical View
Signal Annotations
5 events
Event Details
From 06/May/2019 00:00:00 To 30/May/2019 00:00:00 Edit time
Tag
Severity Score
Source
Investigate
0.335
Orion
Signal Focused View
Show Predictions
+ ADD EVENT
30 days
21 days
14 days
7 days
5 days
3 days
2 days
1 day
Context info
1x
assigned a tag
Investigate
14/Sep/2020 00:21:47
Coke just released 2019 Q1 Consolidated report on the 5/7/2019, revealing big gain of this quarter: 1) First Quarter 2019 net sales grew 3.6% versus prior year 2)Gross margin expanded 170 basis points in Q1 2019 compared to prior year. Adjusted(a) gross margin increased 100 basis points from Q1 2018.
380.0
360.0
340.0
320.0
300.0
Apr 14
Apr 21
Apr 28
May 05
May 12
May 19
May 26
Jun 02
Jun 09
Jun 16
Jun 23
Enter your comment...
Enter
DELETE
CANCEL
SAVE CHANGES
Similar Segments


MTV
Details
Filter
Sync Time Ranges
anonymous
COKE
INTC
REGN
^IXIC
Signal Focused View
Show Predictions
+ ADD EVENT
350.0
300.0
250.0
200.0
150.0
100.0
2015
2016
2017
2018
2019
2020
Periodical View
EVENT TAG
ALL 3 0 2 1 0 1 0
YEAR
MONTH
Relative scale
YY / MM
Year
Month
Day
2015
2016
2017
2018
2019
2020
Signal Annotations
7 events
Event Details
Similar Segments

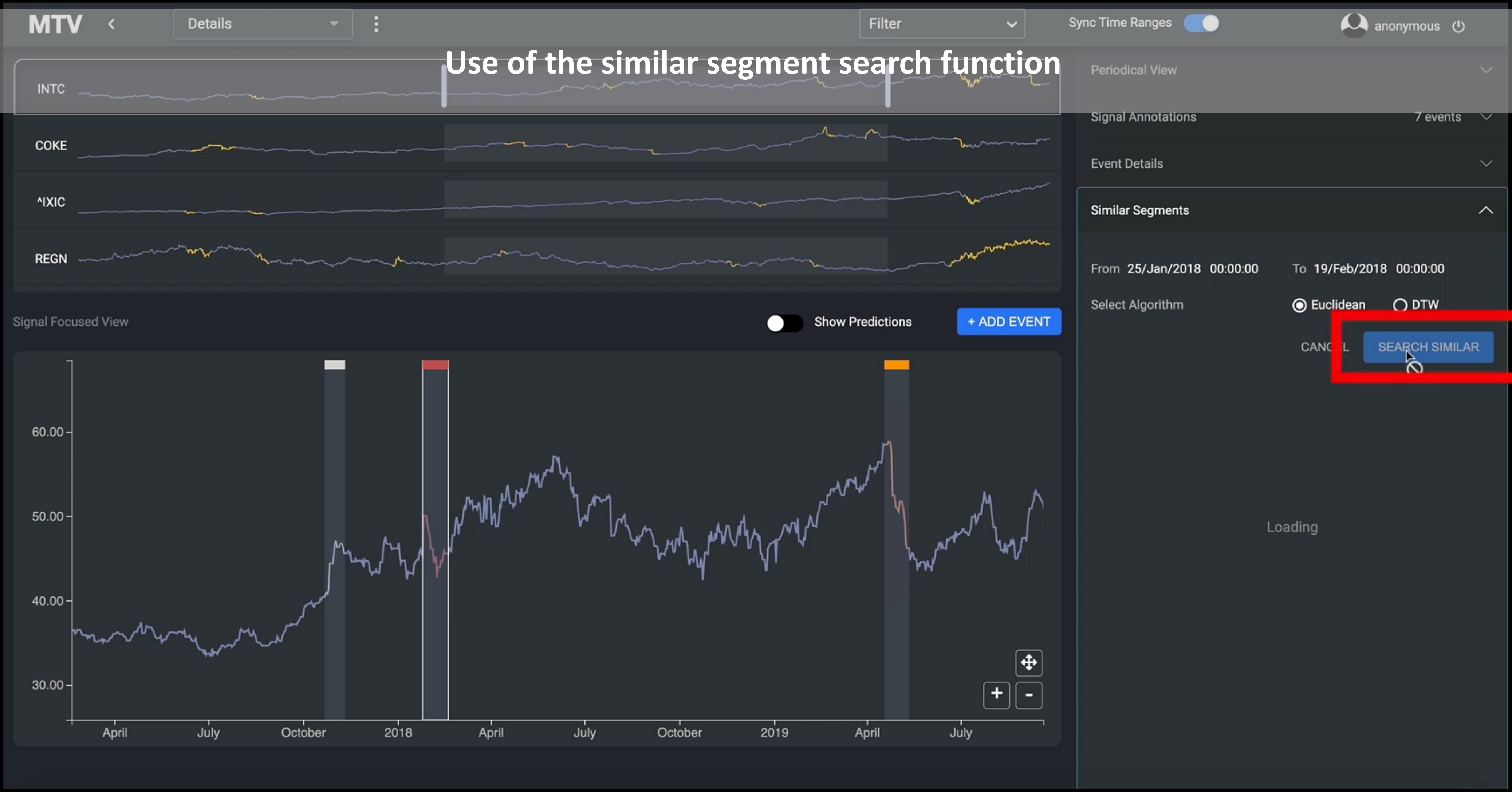


MTV
Details
Filter
Sync Time Ranges
anonymous
COKE
INTC
REGN
^IXIC
COKE
INTC
REGN
IXIC
2020 March
Periodical View
EVENT TAG
ALL 3 0 2 1 0 1 0
YEAR
MONTH
Relative scale
YY / MM
Year
Month
Day
2015
2016
2017
2018
2019
2020
Signal Focused View
+ ADD EVENT
350.0
300.0
250.0
200.0
150.0
100.0
2015
2016
2017
2018
2019
2020
Signal Annotations
7 events
Event Details
Similar Segments

Event editing and creation

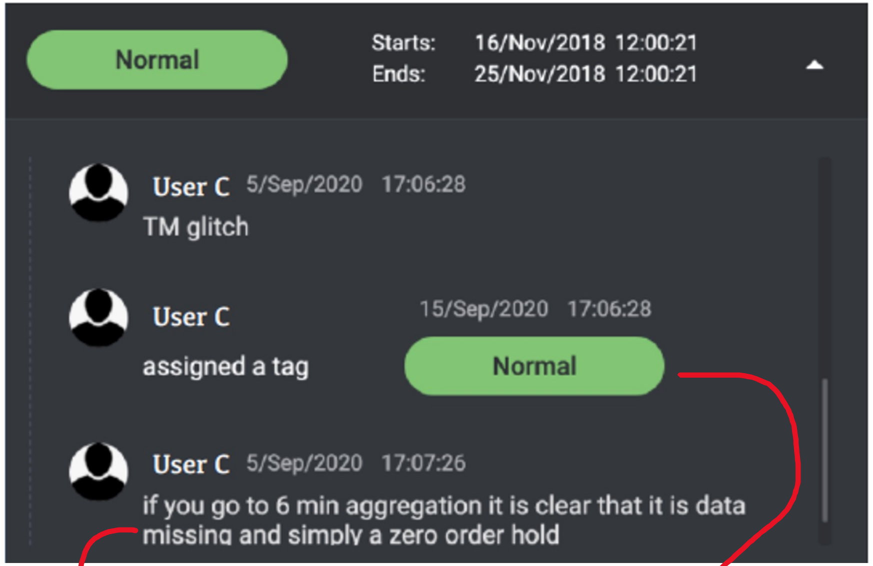


Use of the similar segment search function



System evaluation

- 6 experts from a satellite operations company

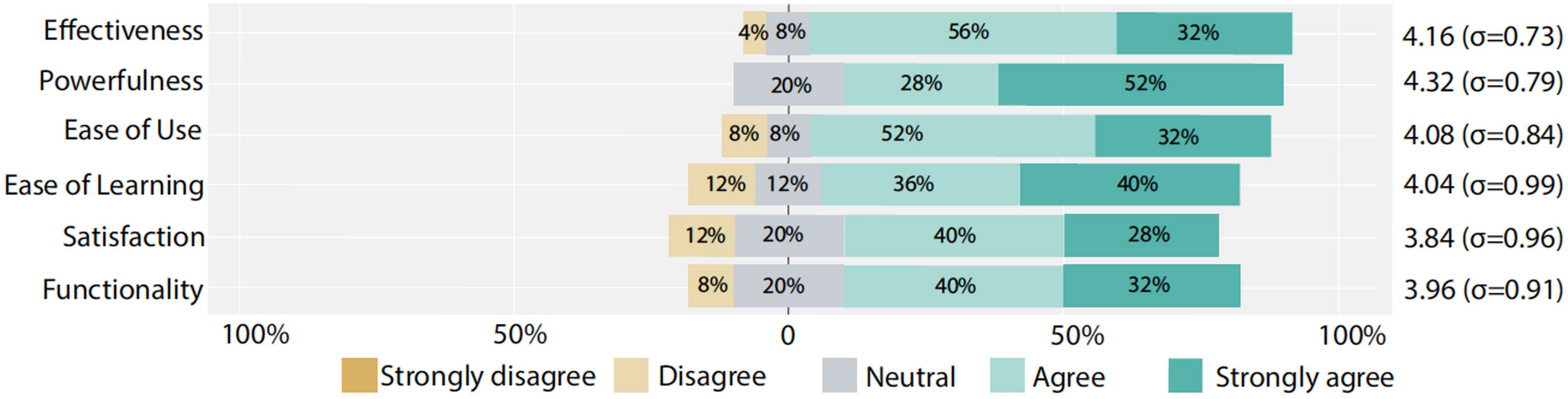


Case	ML Event	User-created Event	Comment (avg.)	Tag (avg.)
1	38	15	162 (3.1)	58 (1.1)
2	45	12	87 (1.5)	60 (1.1)
3	40	8	96 (2.0)	48 (1.0)
4	23	10	66 (2.0)	40 (1.2)

Liu, et al., MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series, CSCW 2022.

System evaluation

- 6 experts from a satellite operations company
- 25 general users using stock data



Liu, et al., MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series, CSCW 2022.

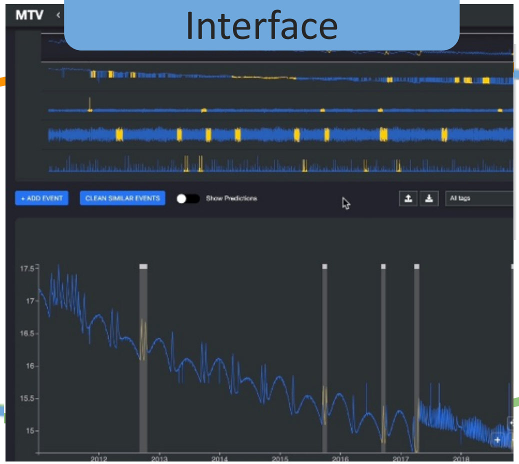
Human-AI teaming workflow

Able to learn

AI System
(Human-Centered)

Need to integrate domain
knowledge to improve performance

Visual
Interface



Human

Cannot monitor massive signals
without AI to pinpoint suspicious locations

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Liu, et al., MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series, CSCW 2022.

Sintel

		MS Azure [30]	ADTK [2]	Luminaire [6]	TODS [21]	Telemanom [17]	NAB [1]	EGADS [22]	Stumpy [24]	GluonTS [2]	*Sintel*
Users	End User	✓	✓	✓	✗	✗	✗	✗	✓	✗	✓
	System Builder	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓
	ML Researcher	✗	✗	✗	✓	✓	✓	✓	✗	✓	✓
Engine	Preprocessing	✗	✓	✓	✓	✗	✗	✗	✓	✓	✓
	Modeling	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓
	Postprocessing	✗	✓	✓	✓	✗	✗	✗	✓	✗	✓
Modular		✗	✓	✓	✓	✗	✗	✗	✓	✓	✓
Comp.	Evaluation	✗	✓	✗	✗	✓	✗	✗	✗	✗	✓
	Benchmark	✗	✗	✗	✓	✗	✓	✗	✗	✓	✓
	Database	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓
API	`lang. specific`	✓	✓	✓	✓	✗	✓	✗	✓	✓	✓
	`RESTful`	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓
HIL		✗	✗	✗	✗	✗	✗	✗	✗	✗	✓

Alnegheimish, **Liu**, et al., Sintel: A Machine Learning Framework to Extract Insights from Signals, SIGMOD 2022.

Anomaly Detection

Orion repository metrics (as of 7/20/23)

https://github.com/sintel-dev/Orion

815
github stars

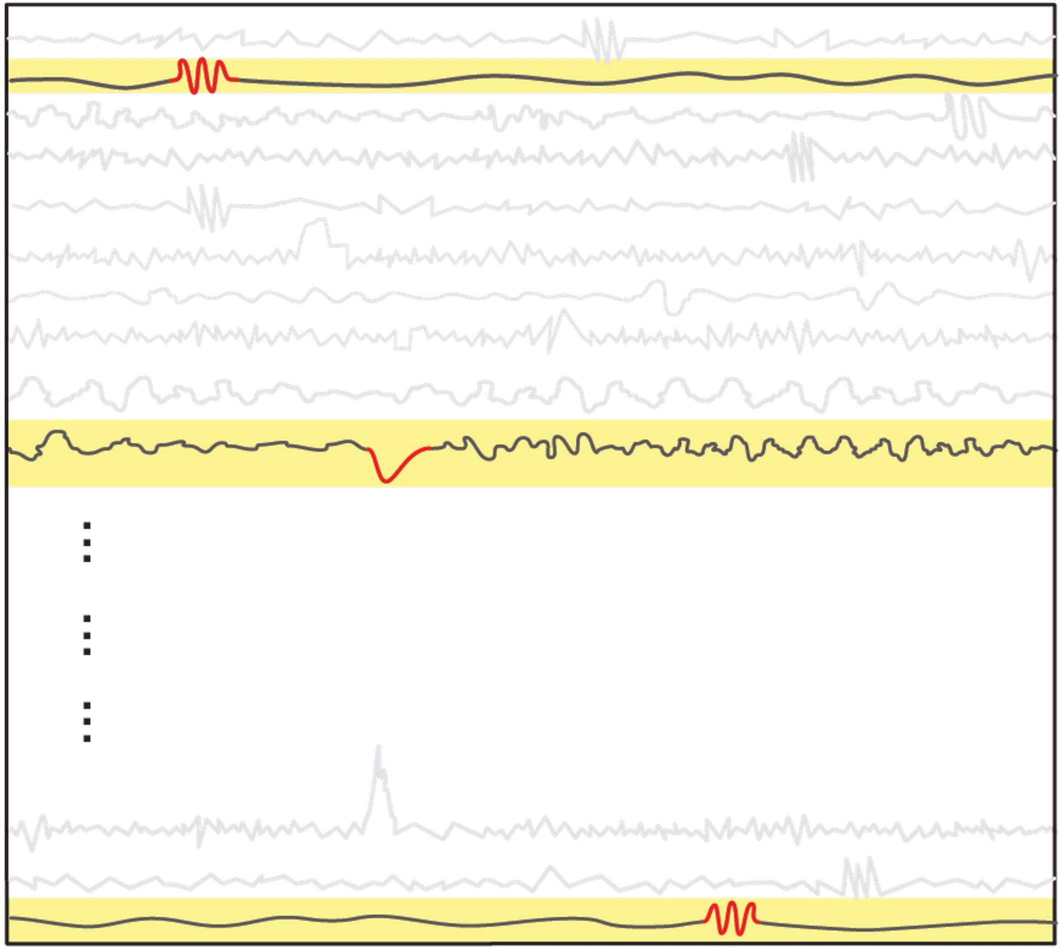
68K
pip downloads

140
forks

358
unique visitors
in 2 weeks

Some open questions

- Priority
- Group analysis

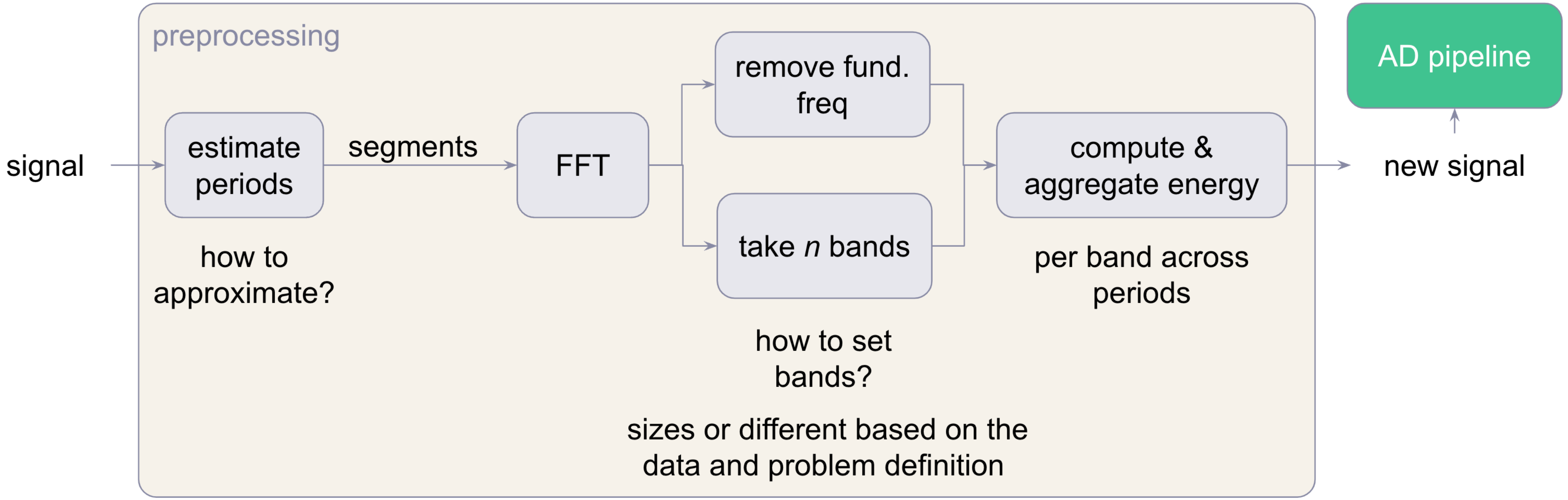


Some open questions

- Propagate / Predict / Suggest Annotations
- Few shot learning

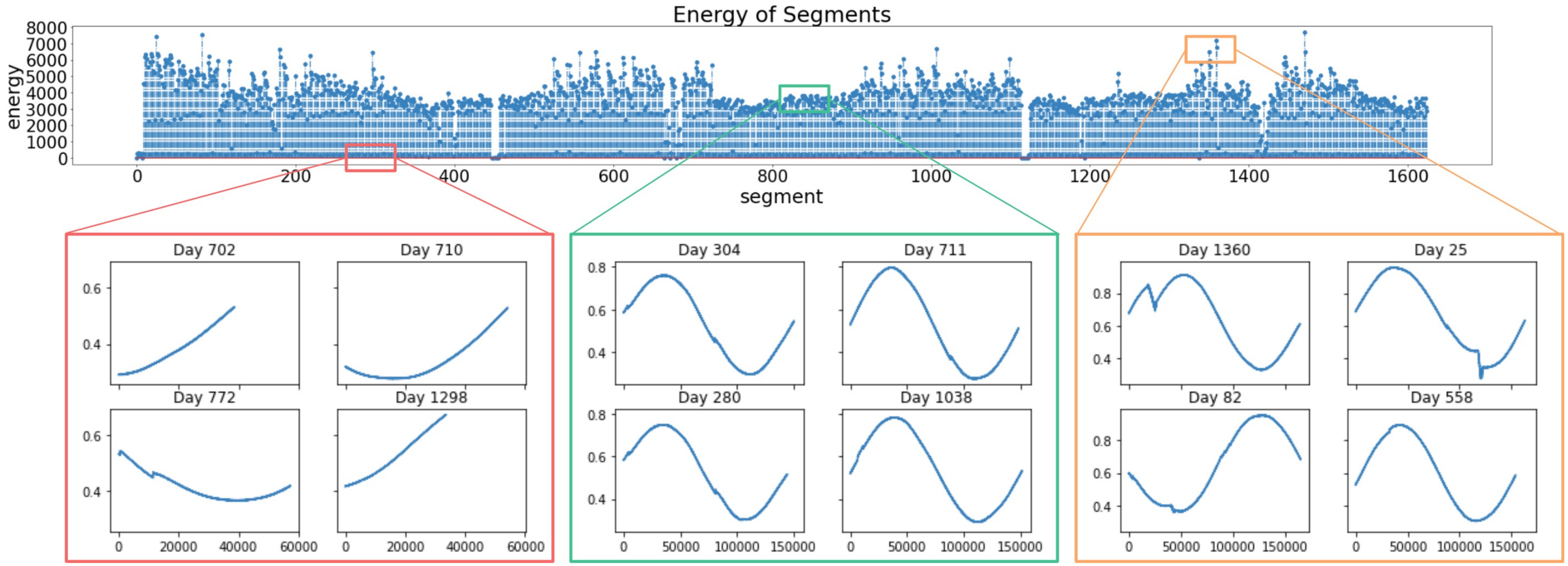
Some open questions

- Is ML necessary?



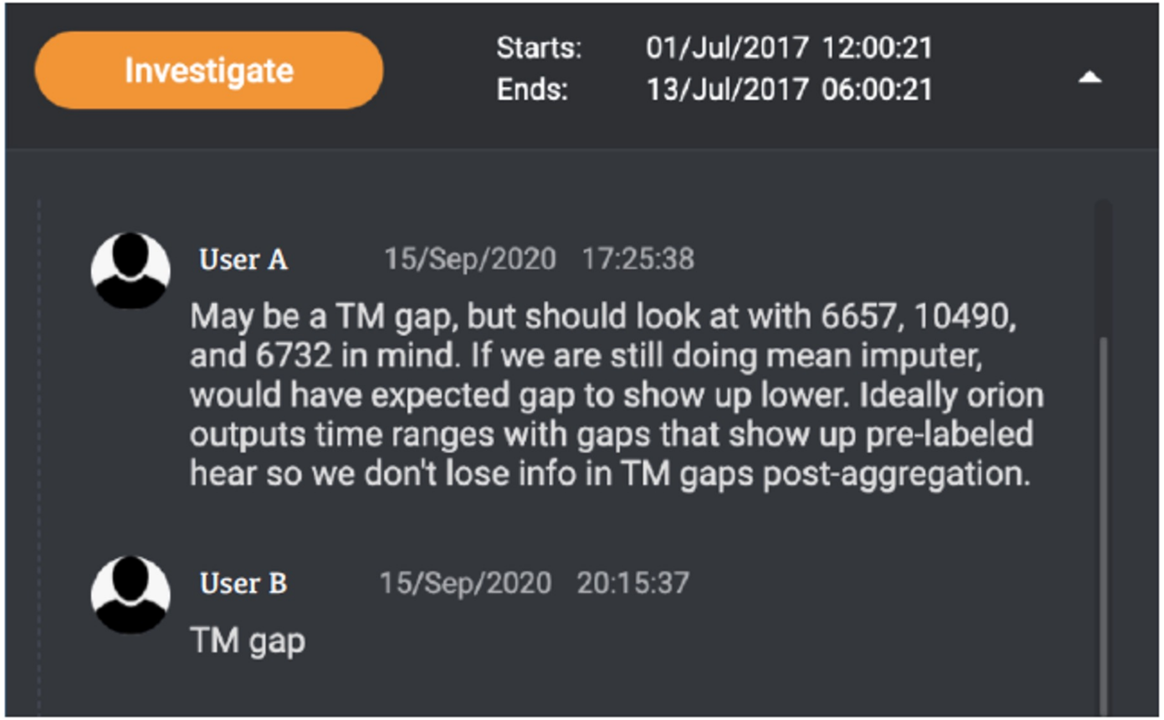
Some open questions

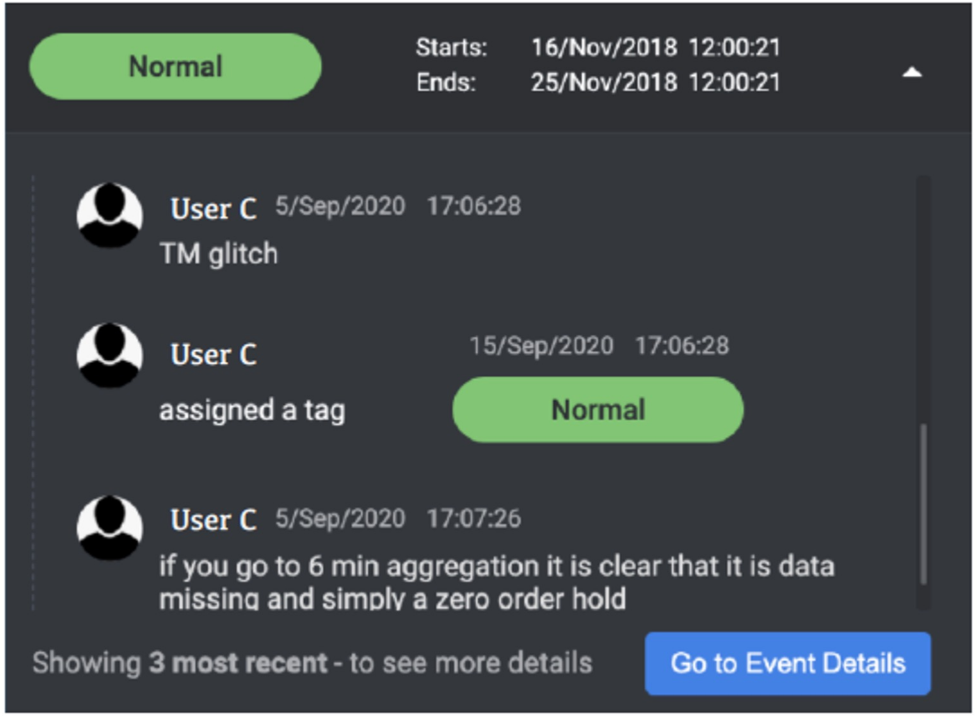
- Is ML necessary?



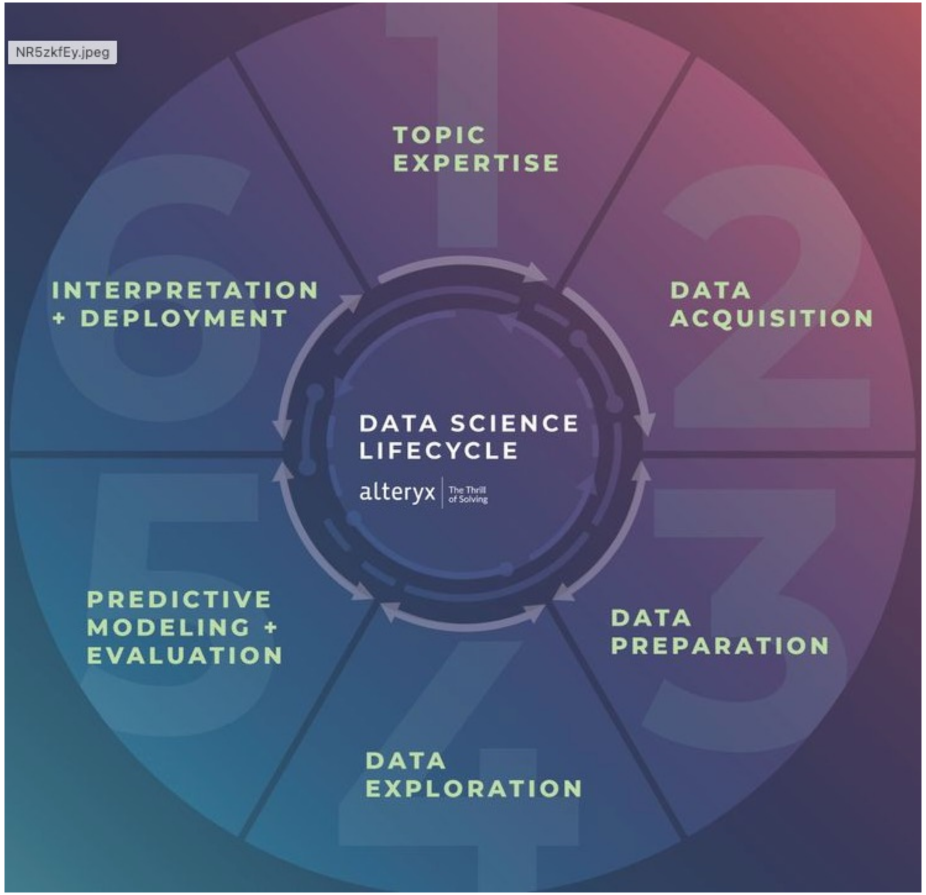
Some open questions

- Actionable decisions





Data science life cycle



Iterate:

- After deployment, new questions, challenges, or insights can emerge, leading to refinements or entirely new cycles of analysis.