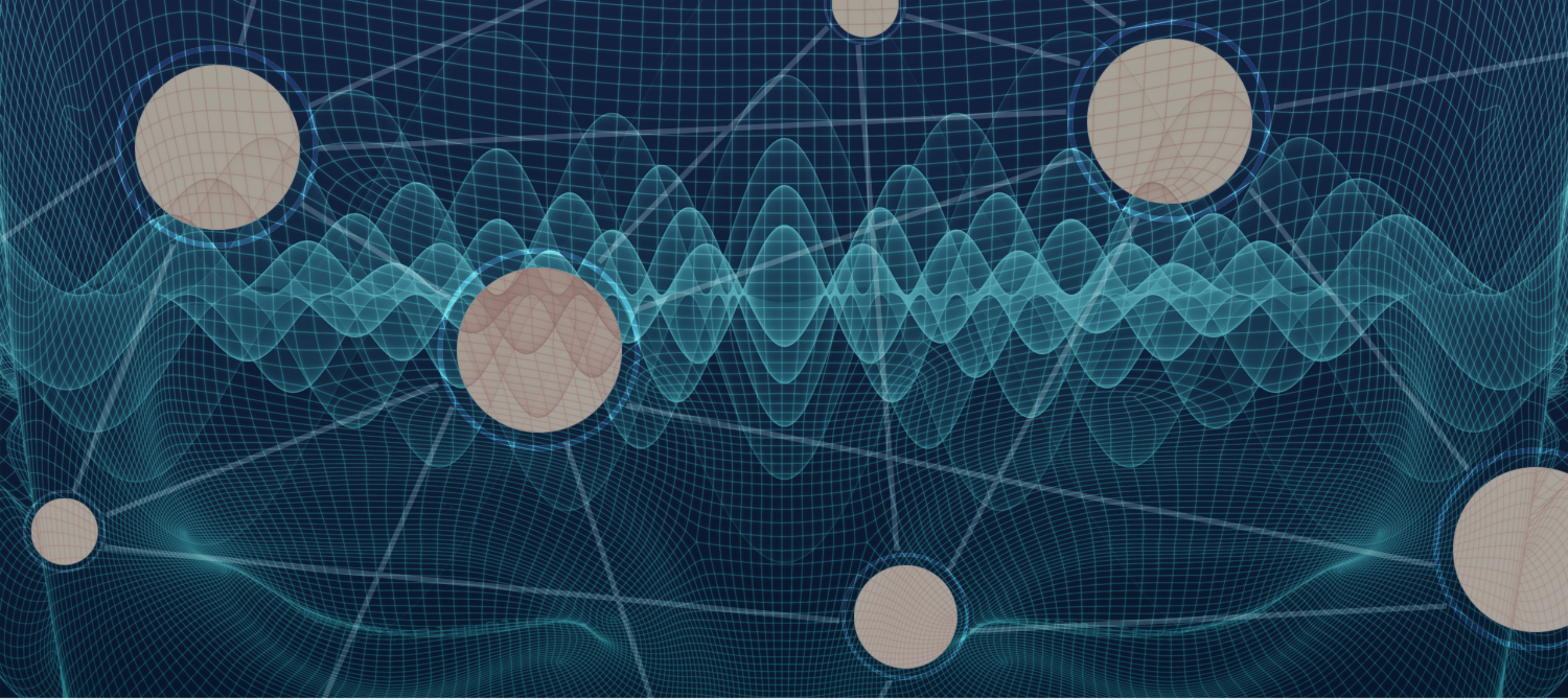


binah.ai
DATA, INSIGHT, DECISIONS

The Missing Link of AI – Signal Processing

OCTOBER 2018

Founders

David Maman

Founder, CEO and CTO

David is a serial entrepreneur: Hexatier (Acquired by Huawei), Precos, Vanadium-soft, GreenCloud, Teridion, Terrasic, Re-Sec and others. Previously a Director in Fortinet's CTO office. 26 year's experience in leadership, AI, security, development and networking. Veteran of an elite IDF unit. David holds a Master's in computer science from the Open University.

Michael Markzon

Co-founder and Chief Scientist

Michael was leading deep learning research at Paypal. Prior to that Michael was Head of R&D at CyActive (acquired by Paypal). More than 15 years experience in developing innovative solutions in machine learning & signal/image processing. Michael holds Master's in applied Mathematics from Bar Ilan University

Konstantin Gedalin

Co-founder and Chief Research Officer

An algorithm phenomenon, he was Chief Scientist of Scifold, developing algorithms for biometrics. Development manager at Silentium, creator of Active Noise Control. 25 year's experience in researching and developing algorithms for wide spectrum of use cases. Konstantin holds a PhD in Biological Physics, Ben-Gurion University and PhD in Applied Mathematics from Tel Aviv University

BINAH's TEAM: 110+ years of development in big data, analytics, signal processing, machine learning, corporate expertise and technology

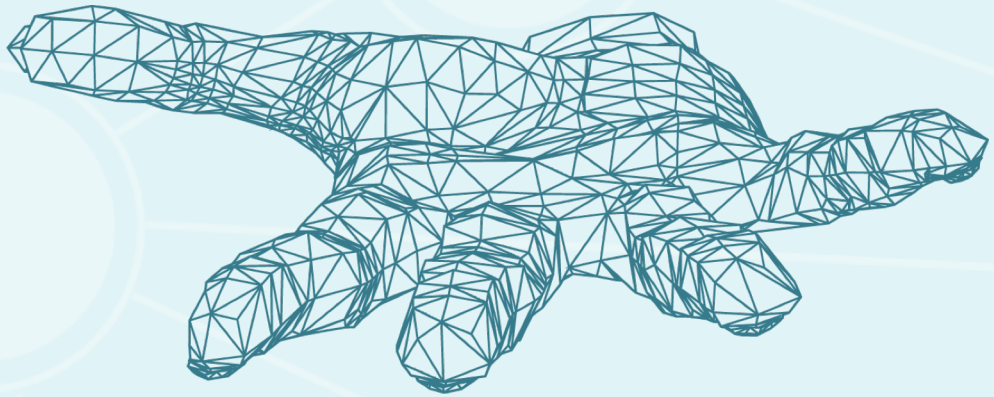
What is **AI?**

Answers to most common data science challenges

Anomalies

Classifications

Predictions



AI is mostly a marketing definition. We don’t create new life forms. When AI will approve or disprove results, we’ll be in AI.

The Problem

#1 Reason AI/Data Science Solutions fail is **lack of Accuracy, Stability and Performance!**

Most common Machine/Deep Learning solutions are based on 20+ years old mathematical libraries which includes many inefficiencies and incorrect implementation.

Researchers and Data scientists have no control or knowledge how the underlying layers operate.

BINAH DECIDED TO CHALLENGE THIS ISSUE!

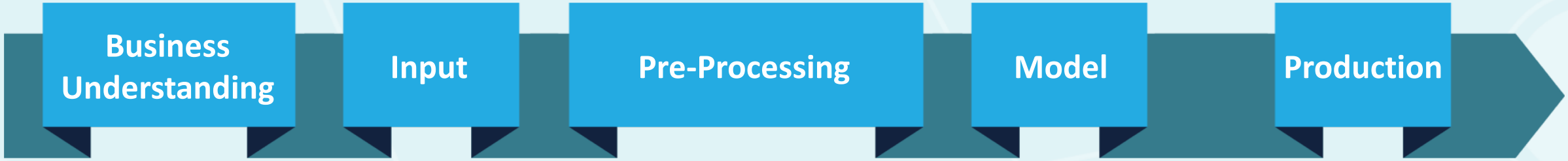
Caffe2

mxnet

TensorFlow

Azure ML

Binah's Data Science Project



- Objectives
- Resources
- Risks
- Data Sources
- Success criteria/KPI

- CSV / Excel
- Images
- Videos
- Live stream
- Structured
- Unstructured
- Etc.

- Data Validation
- Clean-up
- Normalization
- Classification
- Extractions
- Data Selection
- Pre Processing
- Segmentation
- Denoising
- **Signal Processing**
- Feature Extraction

- Model Selection
- Model Tuning
- Model Training
- Model Validation
- Model Testing

- Real time
- Restful API
- Input Interface
- Output Interfaces
- Scalability

The Problem - Data Science Fails! Why?

Gartner: **“84% of Artificial Intelligence projects never gets to production..”**

Business Understanding → **Data** → **Preparation/ Pre-Processing** → **Model** → **Production**

KPMG, McKinsey & Company, Deloitte., BDO, EY Building a better working world, BCG The Boston Consulting Group

ClearStory Data Now You See It, tamr, TRIFACTA, Datameer stay curious, DATAWATCH, pandas, Waterline Data, Paxata

H2O.ai, Keras, alteryx, rapidminer, dataiku, databricks, Open for Innovation KNIME, PYTORCH, Caffe2

Each use case have different challenges

Each Data Science challenge goes through this process which takes months, with no guarantee whether it will deliver the needed results. Multiple stakeholders, throughout the organization.

Lack of strategy, Undefined business goals and inappropriately transforming them to research tasks

Inability to build and apply a uniform data set across organization

Investment in “sexy” algorithms instead of useful ones

Data Scientists are not an engineers

Inappropriate data or infrastructure

Failure to differentiate between academia and the real world

binah.ai

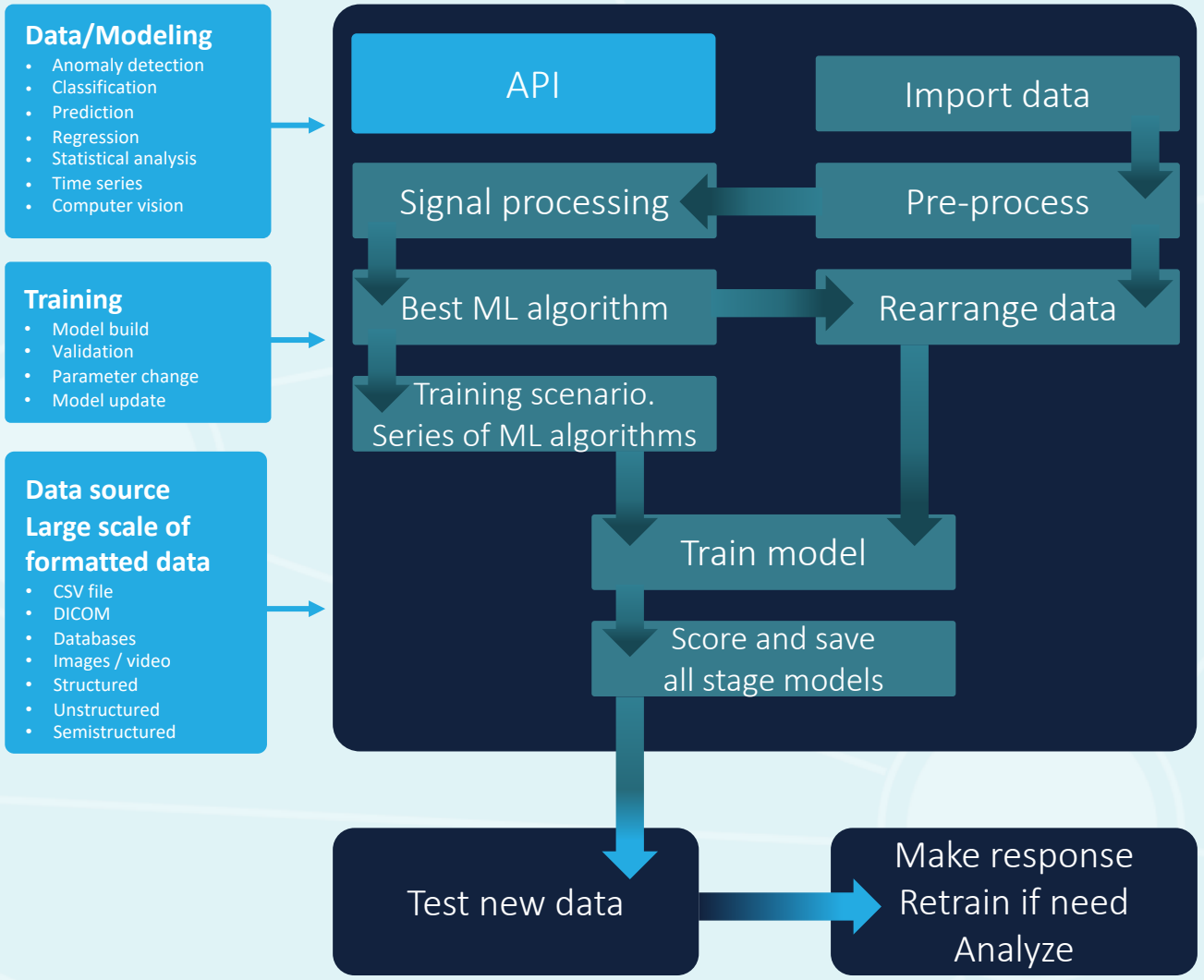
Binah's Custom-built, Multilayered **AI Frameworks**

Linear algebra
- Matrix manipulation
- Tensors
- Decompositions
- Factorization
- Elementary functions

Machine Learning
- k-Means clustering
- Least-angle regression
- Linear regression
- Perceptrons
- PCA
- LDA
- Nearest-neighbor search
- Support vector machines
- Basic neural network definitions
- Back propagation ANN
- Bayesian network

Signal processing
- Filtering
- FFT, Hadamard
- De-noising
- Convolution
- Independent component analysis

Deep Learning
- Simple auto-encoders
- Convolutional neural networks
- Convolutional auto-encoders
- Deep recurrent networks (RNN)
- LSTM and convolutional LSTM
- Boltzmann machines
- Deep Boltzmann machines
- Restricted Boltzmann machines (RBM)
- Deep belief neural networks



Pre-process
Rearrange Data
Signal Processing
- Clean from missing data
- Normalize data
- PCA
- ICA
- Filter data
- Validate data

Best ML- algorithm
Training scenario
- Choose which training algorithm suitable
- Create scenario to create feature extraction
- Create scenario for classifier
- Prepare to parallelize training
- Activate needed data splitting

Signal Processing Strengthens Data Value and Accuracy

Critical to data preparation

Eliminates the outliers

Creates clean data

Reduces the noise

Streamlines data volume

Delivers a better overall data set

USE CASES EXAMPLE

DEAD **RECKONING**

Dead reckoning is a form of prediction that helps determine vehicle position based on various sensor information.

Dead reckoning provides an accurate measure of vehicle's position when GPS is not available over time or its accuracy decreases drifting away from the route.

Sensors mostly being used: position, accelerometer, magnetometer, Linear acceleration, Gyroscope, Light, Atmospheric Pressure and Odometer

STANDARD METHODOLOGY

- State-space model where state defined as position, accelerometer, gyro, Linear acceleration Gyroscope projection to north and east directions
- Measurements of: Light, Atmospheric Pressure, Odometer and magnetometer
- The Standard solution: Extended Kalman Filtering(LQE) or particle filter framework
- More accurate: Recurrent Networks Using same states/measurements definitions

DEAD RECKONING – BINAH'S SOLUTION

Dead Reckoning, why Signal processing

- Different sampling frequencies used for sensors cause "cumulative errors" in position prediction
- All methodologies assume stationary or Gaussian nature (statistically) for noises and high SNR's
- Magnetometer measurements are inaccurate due to magnetic field disturbances
- Longitudinal acceleration errors will yield large position errors

- We've utilized de-noising mechanism based on block matching wavelet and Wiener filtering to remove large amount of noises
- The above resulted in sensor data to be downsampled and additionally filtered to the same sampling rate
- Estimate and remove magnetometer bias

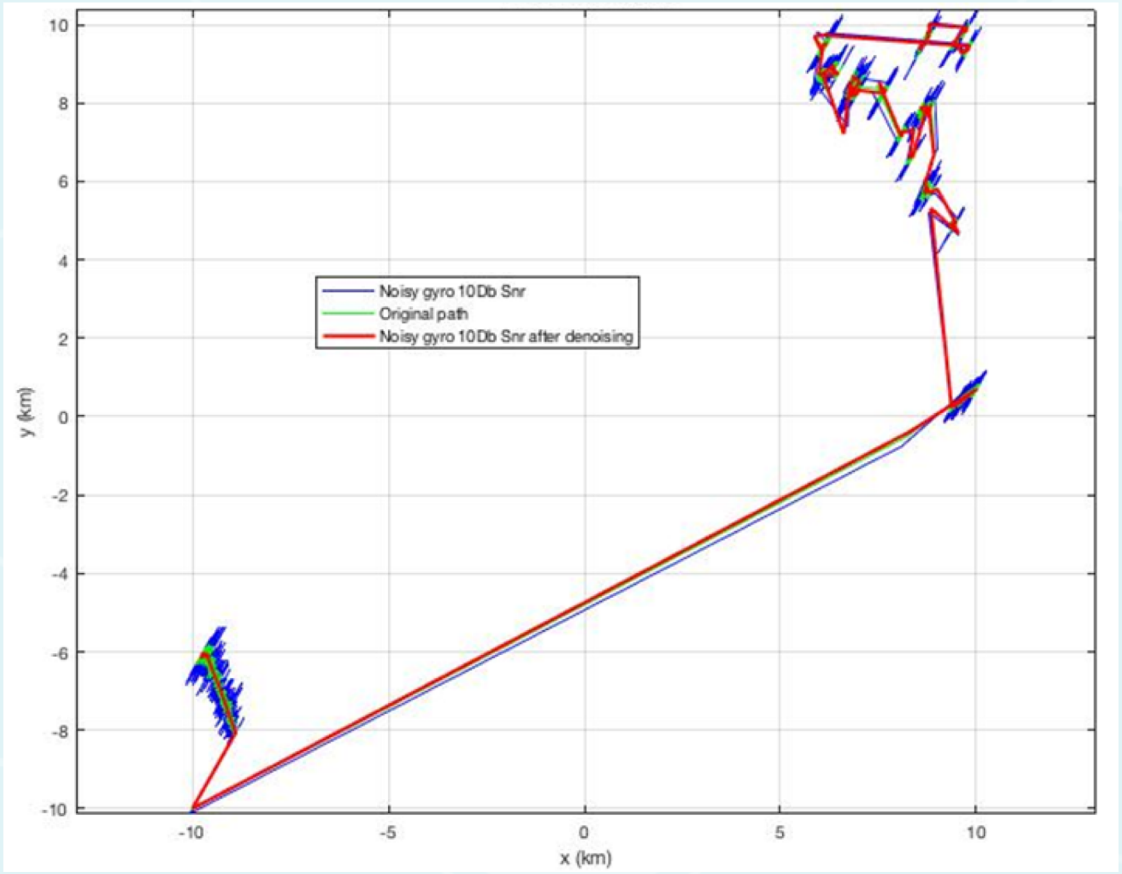
DEAD RECKONING – RESULTS COMPARISON

- Dead reckoning path using gyro, accelerometer and magnetometer for 30 kilometer travel

- **Green**
  The original path

- **Blue**
  10 db gaussian noise added to magnetometer sensor only

- **Red(Binah)**
  clean sensor (without noise) dead reckoning estimation-ideal fit to the real data



Rate Prediction for WM/R- Details

THE CHALLENGE AND GOAL

- 2:30 min prediction at 4 pm for WMR rate for the USD versus G7 currencies that publish at 4:02:30pm
- 10 second intervals prediction in the WMR window

DATA CHALLENGES

- Last-minute data grab creates model challenges (2:30 minutes before 4:02:30 pm close)
- Inherent volatility of the market
- Lack of data(1 prediction point a day as of January 2016)

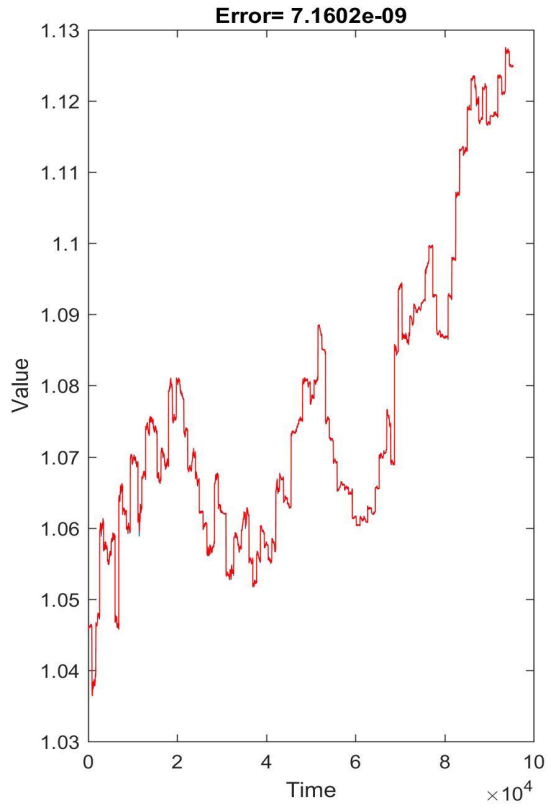
STRATEGY

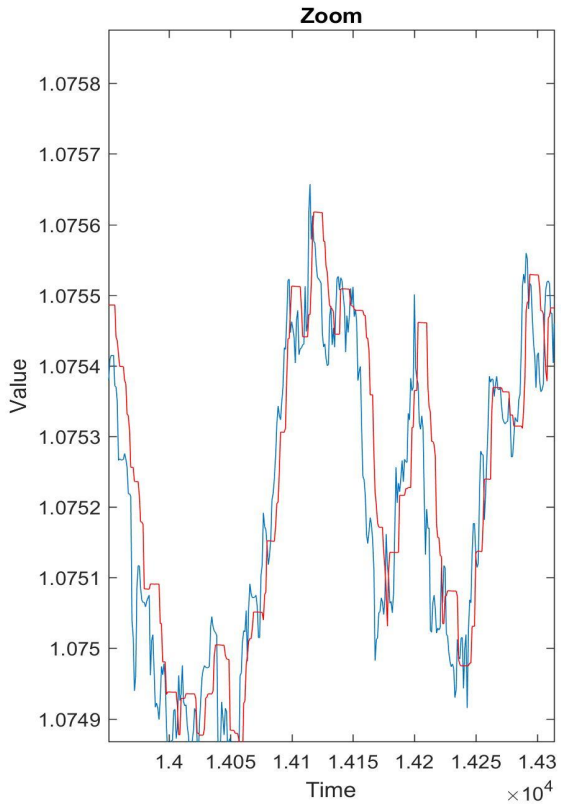
Analyze historical data of specific currency pairs to predict future activities

Rate Prediction for WM/R- Results

RESULTS

- 2.5 minutes forecasting: Mean squared error of 4.9455293983325e-09
- 10 Second windows: Mean squared error of 7.1602e-09





Methodology used to calculate effective Bid/Ask/Spread

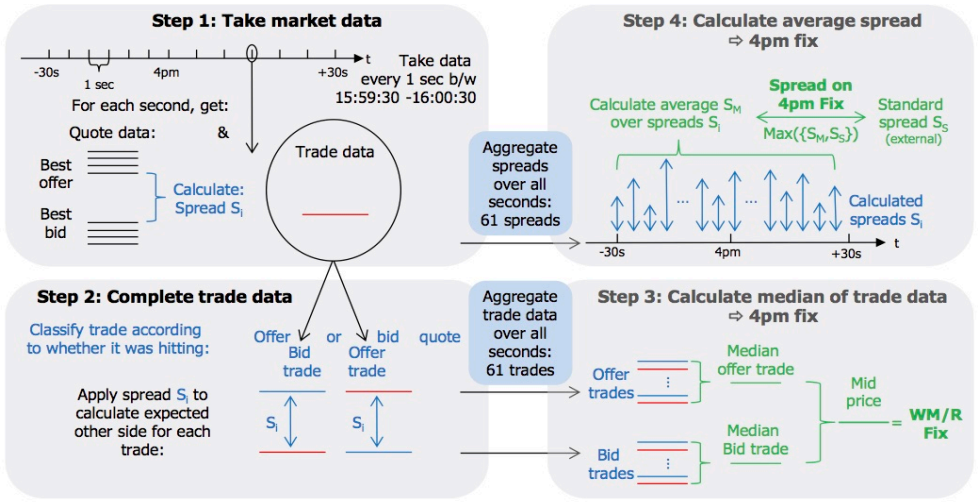


Figure 1: Methodology for calculating the WM/R benchmark rate for trade currency spot rates.

Data validation & preprocessing highlights

- Remove all duplicated and erroneous data
- Transform all ECNs currency pairs simultaneously
- Interpolate data to get 1-millisecond resolution
- Decimation of 1- sec resolution
- Construct Bid/Ask/Spread vectors
- Use AR modeling to construct convenient data

Accuracy is the Name of the Game

Generating an HRV signal requires an extremely accurate heart rate read

Accuracy Goals

- Face detection accuracy goals:
  - Goal: >= 96% Market: <=90%
- Heart rate accuracy goals:
  - Goal: <=2bpm Market: >=3bpm
- Biometric face identification accuracy goals:
  - Goal: >=98% Market: Mostly face recognition

CURRENT SOLUTIONS

- Over 95% of the current solutions in the market are based on OpenCV implementations, although some have evolved from OpenCV
- The current OpenCV implementation of face detection cannot pass the 90% accuracy
- Luminance and stability are key issues with most of the solutions on the market
- Vibration and light influence are critical
- Due to the accuracy levels, HRV cannot be accurately generated

Heart Rate - Remote Photoplethysmography

Heart Rate Estimation is based on correct signal extraction of the skin segments of human faces using video stream.

Estimation of the heart rate in video consists of the following main steps:

- Face detection
- Face tracking
- Motion compensation
- Illumination normalization (removal of external sources)
- Skin region of interest selection

ACCURATE ENOUGH?

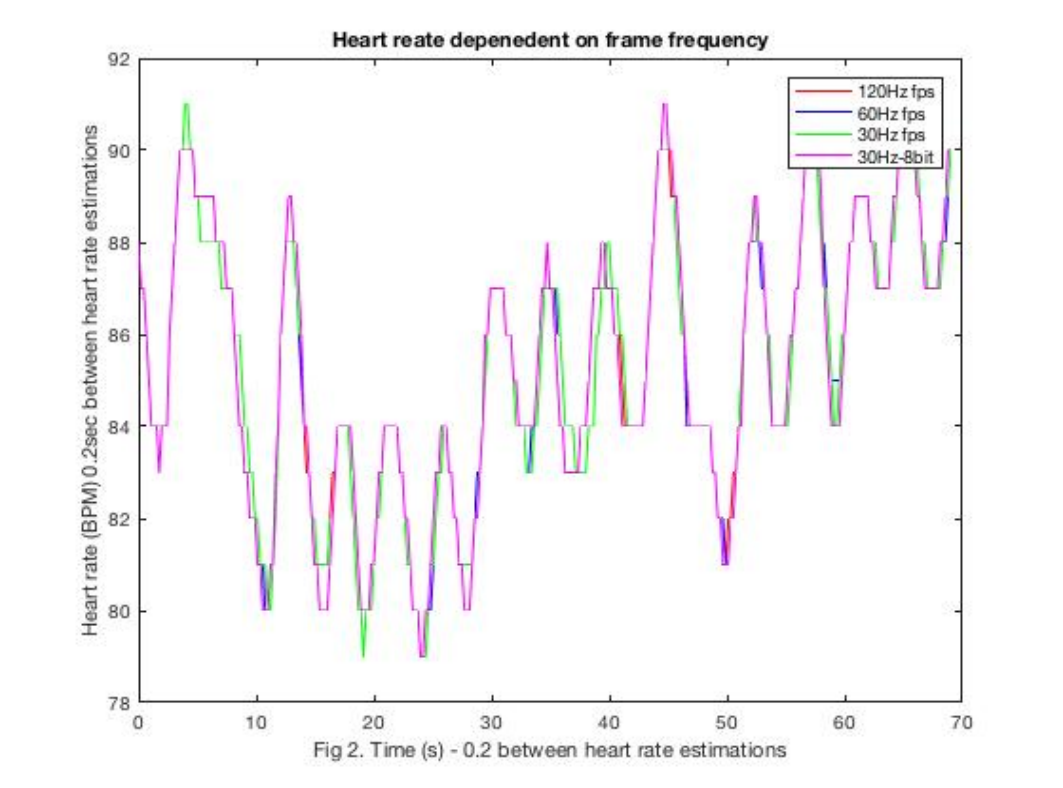
https://www.youtube.com/watch?v=9BSVV2KCHPo

https://www.youtube.com/watch?v=_qLcFxn3PhU

Heart Rate - Remote Photoplethysmography

Heart rate estimation algorithm:

- Creation of photoplethysmography signal (light reflection intensity)
- De-noising signal and removal of artifacts of the rapid intensity changes
- Spectral analysis of the photoplethysmography signal
- Finding best signal approximation - reconstruction of the ECG-like signal
- Signal reconstruction to HRV



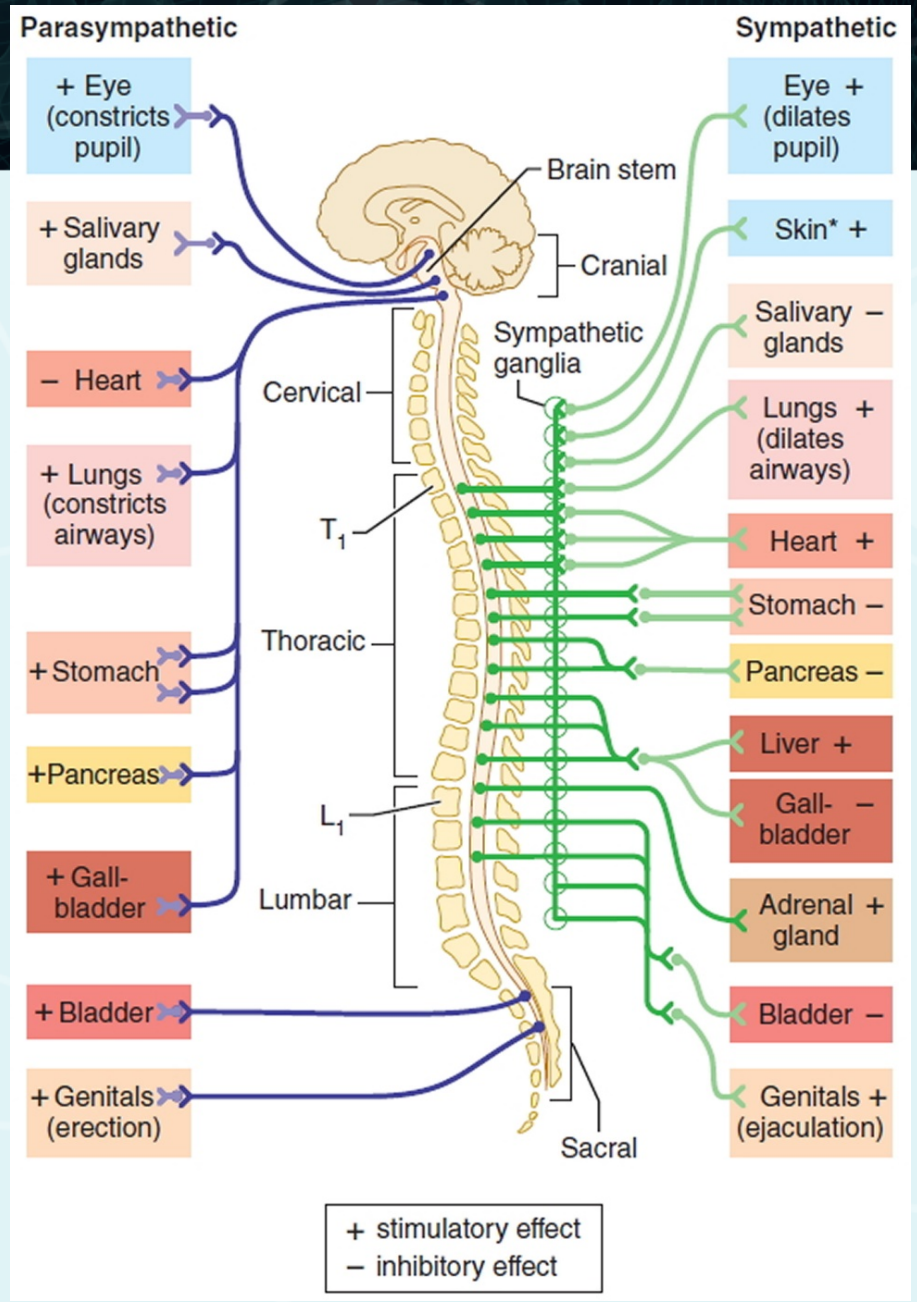
Heart Rate Variability

Once an HRV signal been created, exploring the autonomic nervous system is a matter of time

At a medical service, workplace, interview, family, etc

HRV

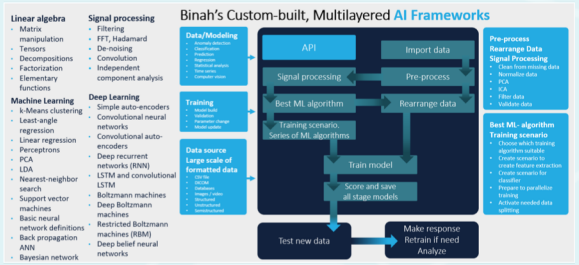
- LIE DETECTION! (approved patent)
- Cortisol levels (Mood swings, muscle issues, etc)
- Heart failure
- High blood pressure
- Fetal distress
- Cognitive and memory performance
- And much more



Binah's **Breakthrough** Solution

Binah took the hard path and implemented it's own mathematical backend!

Binah.ai has built the entire mathematical backend, **without any open source** or third-party development to ensure that binah:



Control the entire underlying layer of each algorithm and model

Provide real-time solutions, everything was implemented in C++

Support the most common operating systems and processing architectures (CPU, GPU, DSP, ARM, FPGA)

Binah's Core IP: **Signal Processing with AI**

The Secret Sauce of AI: Signal Processing!

Delivering Previously Impossible level of Accuracy, Stability and Performance!

By its nature, data isn't clean. Signal processing strengthens data value and accuracy, performance, and the stability of the end results, making it a critical tool in data preparation.

Critical to data preparation

Eliminates the outliers

Creates clean data

Reduces the noise

Streamlines data volume

Delivers a better overall data set

Signal processing helps eliminate the outliers, generate new signals create cleaner data, reduce the noise, streamline data volume, and deliver a better overall data set.

Binah.ai

binah.ai

Founded	**2016**
Location	**Tel Aviv, Israel**
Employees	**24 (13 PhD's)**

Awards

1st Place — FINTECH JUNCTION (2018)

1st Place — TAU INNOVATION CONFERENCE 2018

1st Place — star 一"创业之星"大赛— entrepreneurs

Finalist — BBVA Open Innovation — BBVA Open Talent 2018

Customers	Three out of the world's top 10 banks	Israel's leading financial companies	Largest Automotive Tier-1 supplier
Customer Acquisition	Partners (Deloitte, KPMG, EY)	Ready to use use-cases Including pure SaaS	Added value: **Up to 90% cost and time savings**
Market Size	Total Available Market: Fortune 2000 financial organizations	Serviceable Available Market: North America, and EMEA	Top 20 automotive Vendors

binah.ai

Thank You

info@binah.ai