



# NASPI EATT White Paper: Data Mining Techniques and Tools for Synchrophasor Data

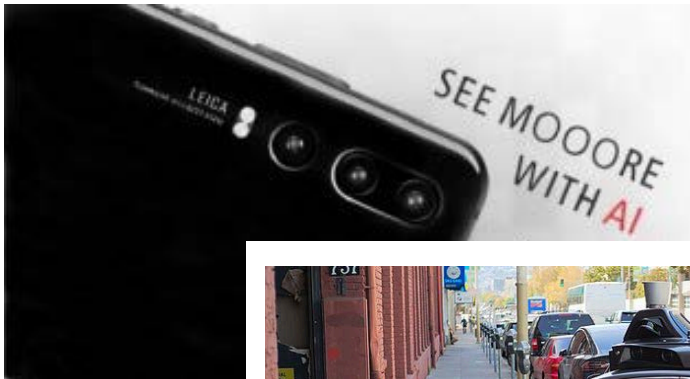
Evangelos Farantatos (EPRI) – EATT Lead

NASPI WG Meeting

April 16 2019

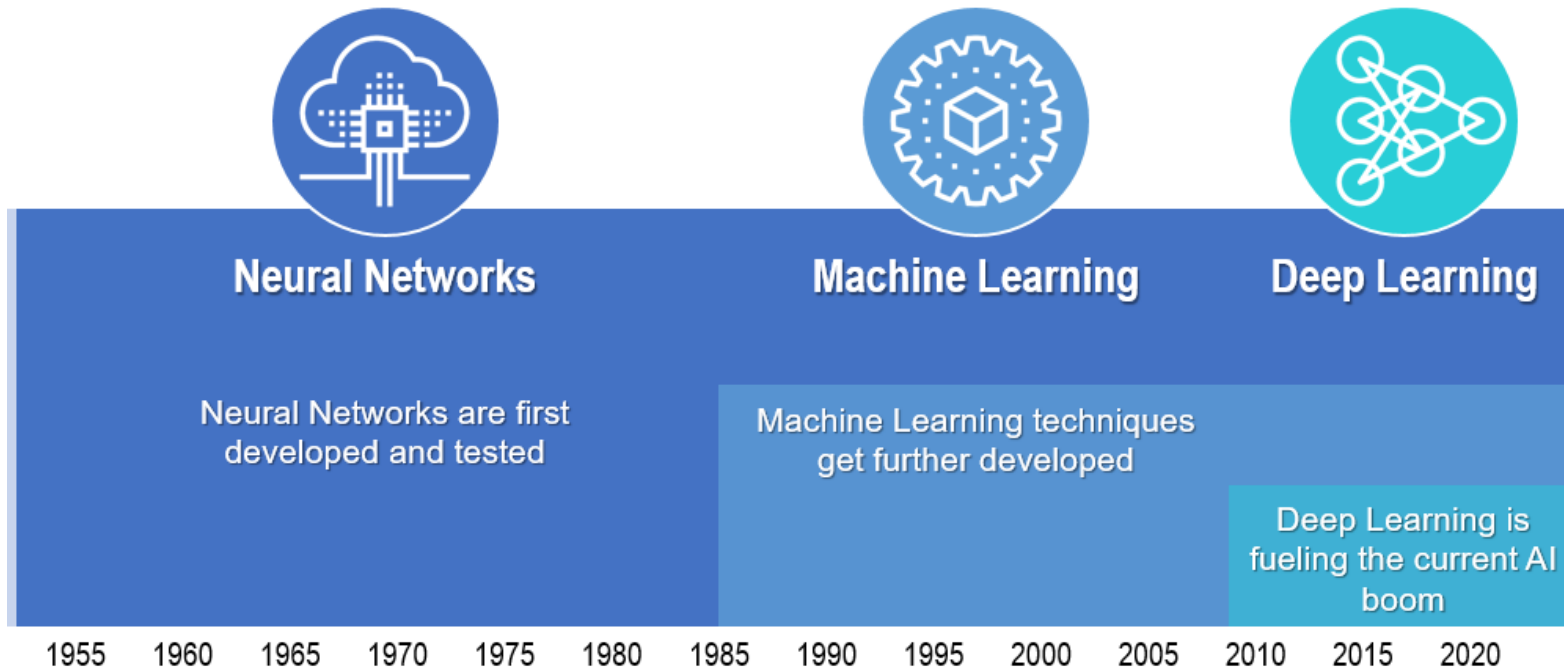
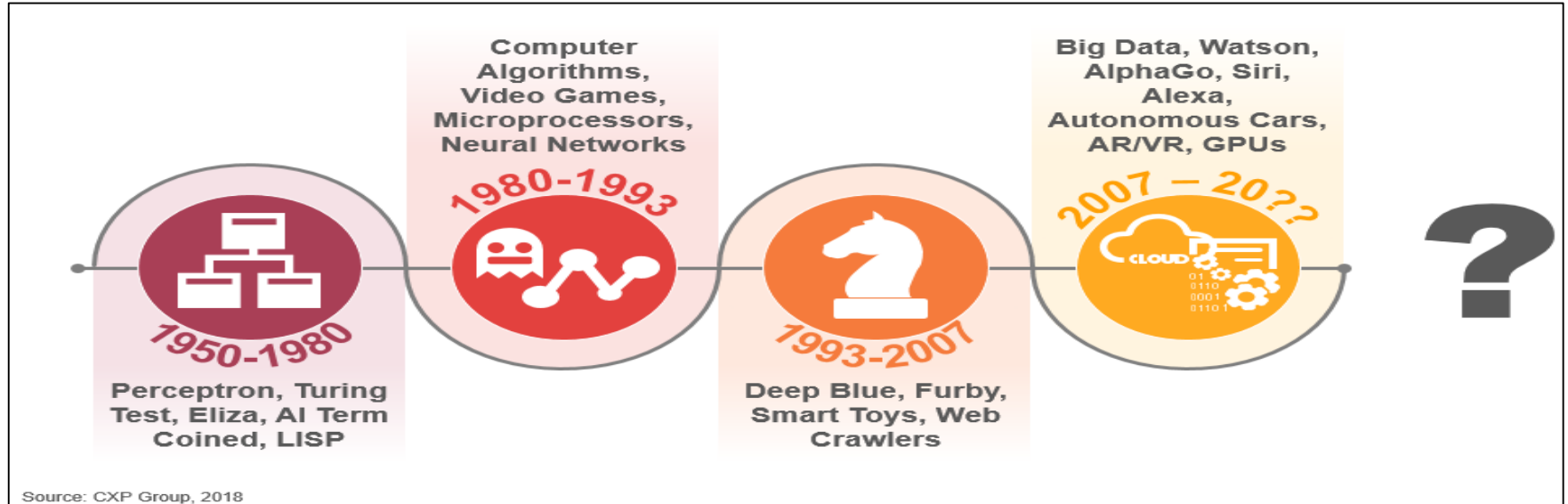
San Diego, CA

# Artificial Intelligence (AI) is Everywhere



**What does AI mean for Power System operations & planning?**

# AI Evolution Timeline



# High Dependence on Data from Many Sources

## System Measurements



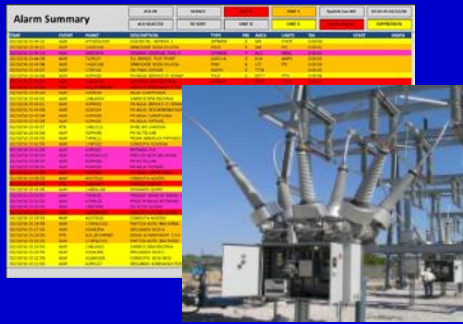
SCADA  
Synchrophasors  
Fault Recorders  
SoE Recorders  
Trend Recorder  
PQ Meters  
Meters & AMI

State Estimation  
Contingency  
Models  
Outages  
Markets  
Simulations  
Operator Action

## Analysis Results



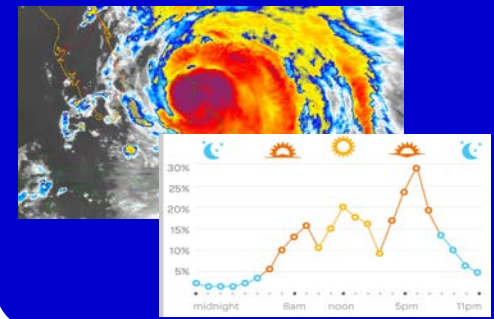
## Equipment & Alarms



Alarms  
Generators  
Transformers  
Breakers  
Relays  
Shunts  
DER, BESS, et. al.

Weather  
GIS  
Geospatial  
GIC  
Satellite Images  
Customer  
Gas, Transport

## Non-Electrical

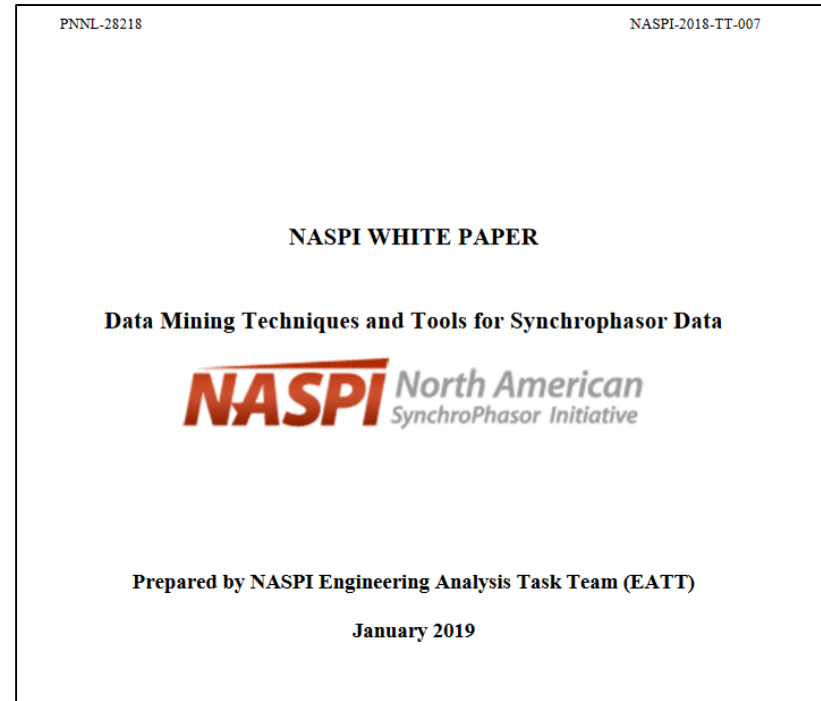


# NASPI EATT White Paper

- **Data Mining Techniques and Tools for Synchrophasor Data**
  - **Lead: Brett Amidan (PNNL)**

## White Paper Focus:

- give a high level overview of data mining
- review how data mining has been used in industry
- present common big data architectures, software languages and tools that facilitate data mining
- provide use cases that show how data mining has been applied in the power systems community
- discuss possible future ways to apply data mining to the power grid and more specifically with synchrophasor data



<https://www.naspi.org/node/743>

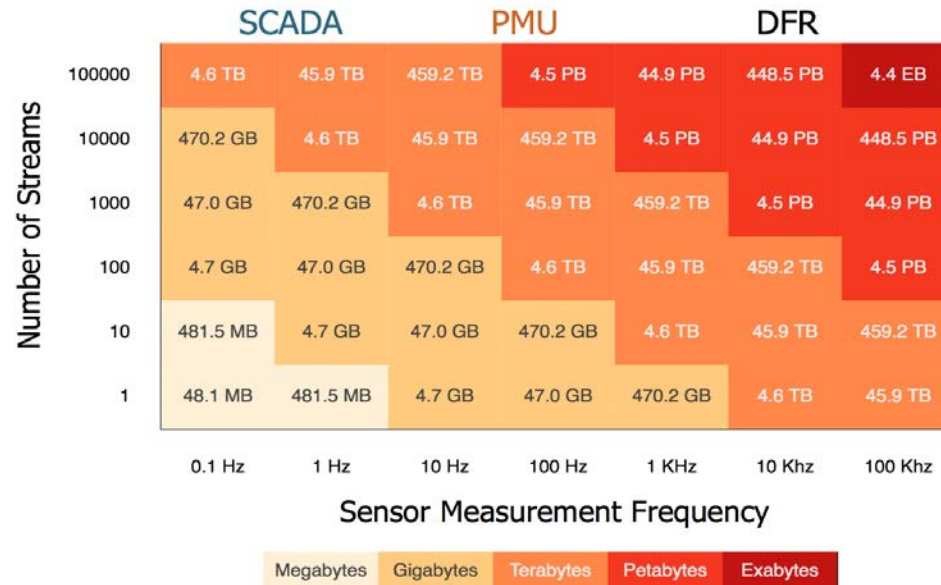
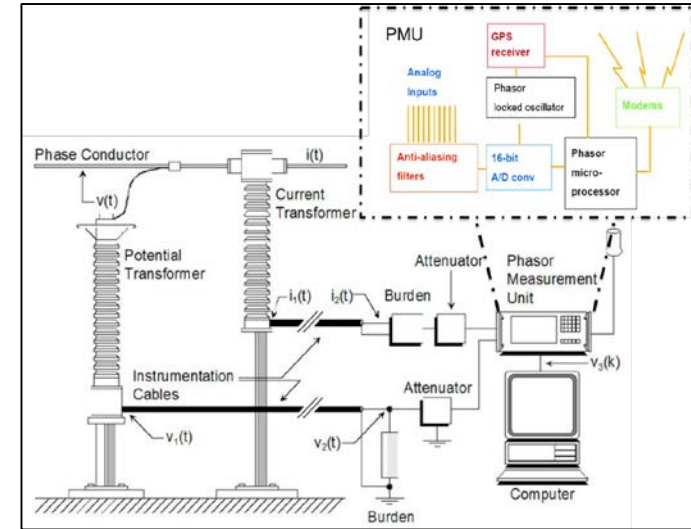
# Data Mining Techniques and Tools for Synchronphasor Data - **Outline**

1. Introduction
2. Data Mining Techniques
3. Software Tools and Big Data Platforms for Data Mining
4. Use Cases
5. Conclusions



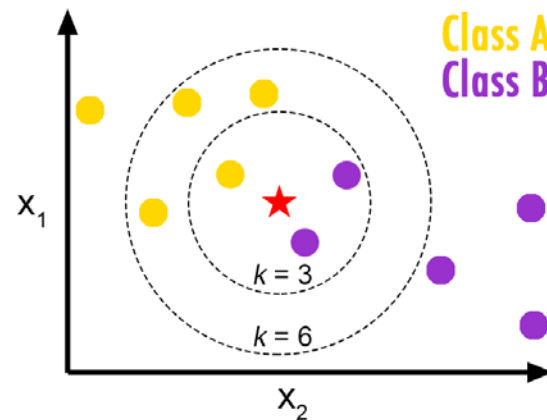
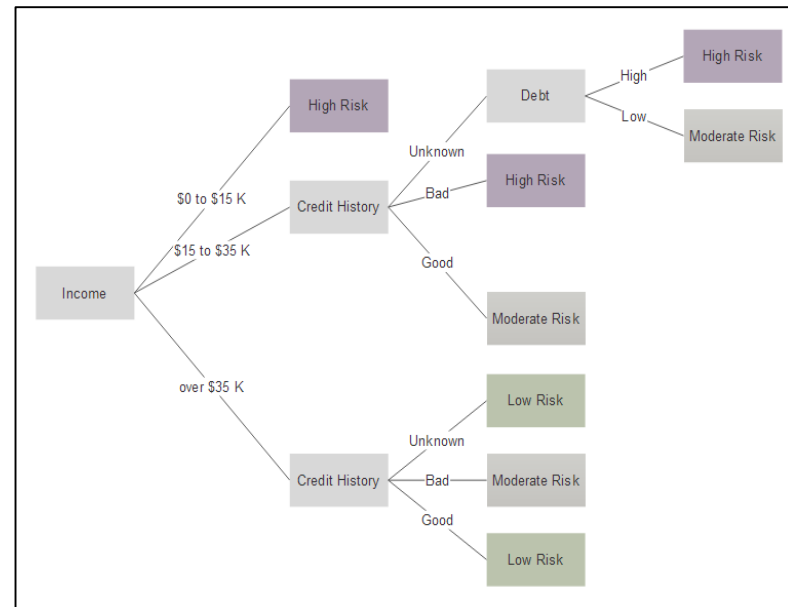
# Data Mining Techniques and Tools for Synchronphasor Data - Introduction

- Synchronphasor Technology Background Information
- Data Mining Background
  - Definition
  - Use of Data Mining in Other Industries
- Big Data Architecture Background
  - Characteristics of Big Data in the Utility Industry
  - How Big Data Architecture Is and Could Be Used in the Power Grid



# Data Mining Techniques and Tools for Synchronphasor Data – Data Mining Techniques

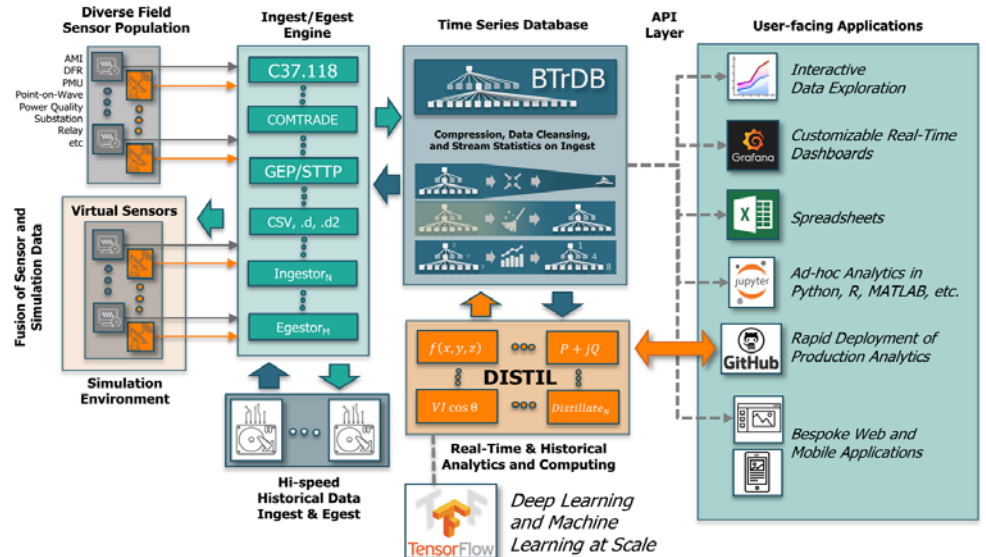
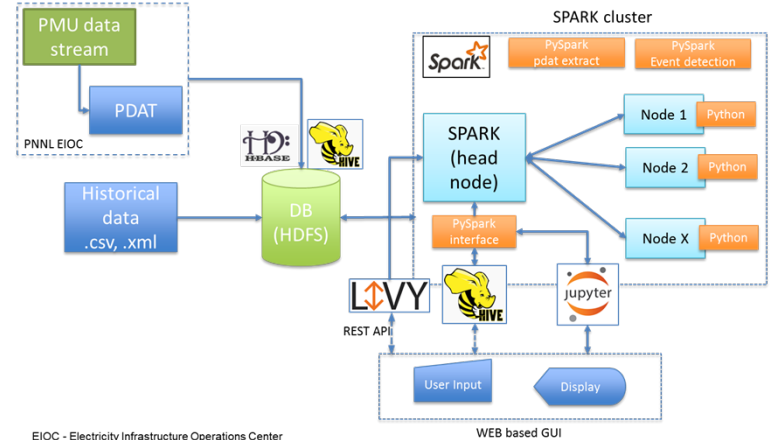
- Feature Extraction
  - Principal Component Analysis
  - Manifold Learning
- Clustering (Unsupervised Learning)
  - K-means
  - Hierarchical Clustering
  - Fuzzy Clustering
  - DBSCAN
- Classification (Supervised Learning)
  - Linear and Quadratic Classifiers
  - Kernel Estimation
  - Decision Trees
  - Support Vector Machines
  - Neural Networks





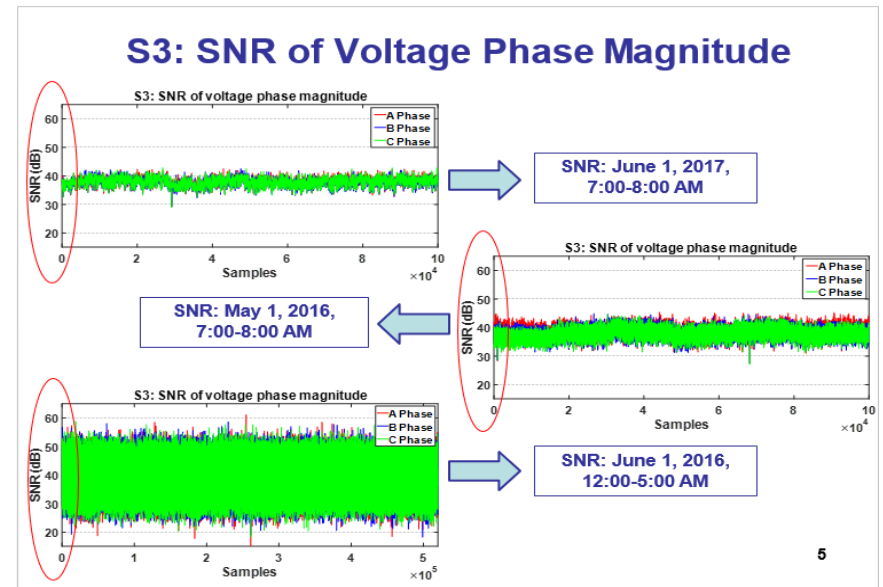
# Data Mining Techniques and Tools for Synchronphasor Data - Software Tools and Big Data Platforms for Data Mining

- Data Mining Tools
- Open Source Languages
- Open Source Data Mining Software
- Commercial Languages
- Commercial Data Mining Software
- Data Stream Processing Software
- Big Data Platforms



# Data Mining Techniques and Tools for Synchrophasor Data – Use Cases

- 11 use cases. Topics cover:
  - Event detection
  - Situational awareness
  - Dynamic security assessment
  - Fault analysis
  - Asset monitoring
  - Load monitoring
- Mostly national labs and academia work



# Acknowledgements

## **Editors:**

Evangelos Farantatos (EPRI) – NASPI EATT Lead

Brett Amidan (Pacific Northwest National Lab) – White Paper Lead

## **Contributing Authors (Alphabetical Order):**

Reza Arghandeh (Florida State University – now in Norway)

Daniel Bienstock (Columbia University)

Pavel Etingov (Pacific Northwest National Lab)

Sean Murphy (PingThings)

Femi Omitaomu (Oak Ridge National Lab)

Matthew Rhodes (Salt River Project)

Tom Rizy (Oak Ridge National Lab)

Anurag K Srivastava (Washington State University)

Kai Sun (University Tennessee Knoxville)

Xiangyang Zhou (American Transmission Company)