“Big Data” In Analysis Of Streaming Physiologic Data: Implications For Health Care

Rishi Kamaleswaran, PhD

Center for Biomedical Informatics,

Division of Critical Care Medicine,

Department of Pediatrics,

The University of Tennessee Health Science Centre, Memphis TN

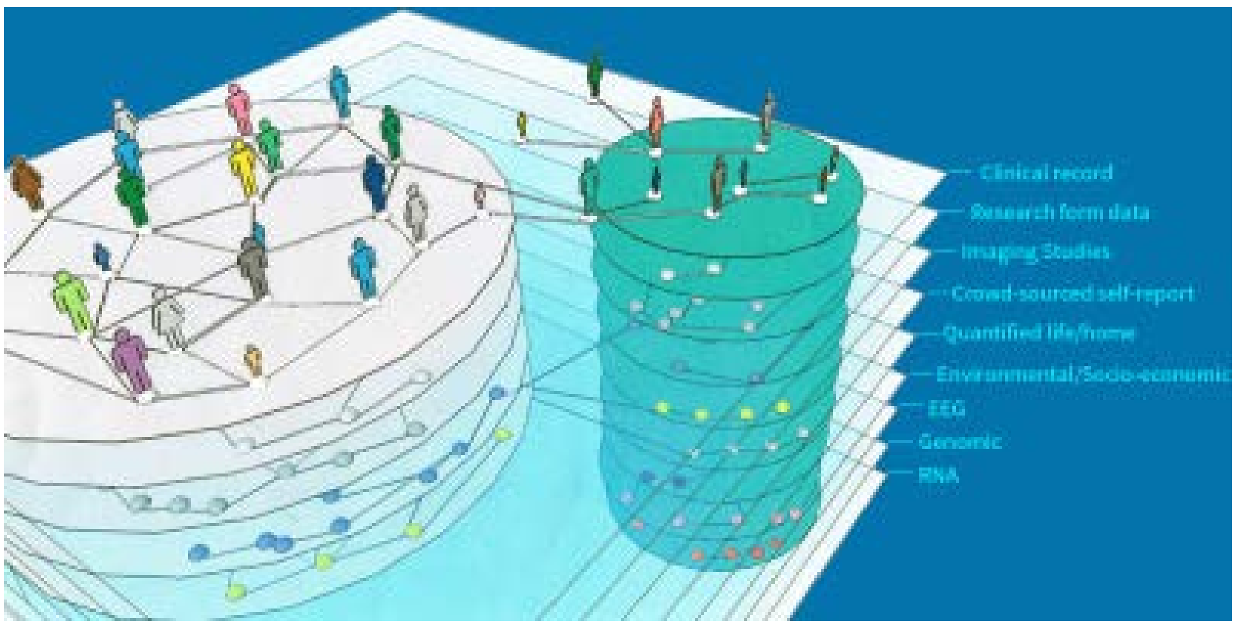
@rkamaleswaran

ADEPT4 Workshop 2017

Big Data to Knowledge Initiative

PIC-SURE Precision Medicine June 2015 Patient Driven meeting (BD2K Center of Excellence, Harvard University)

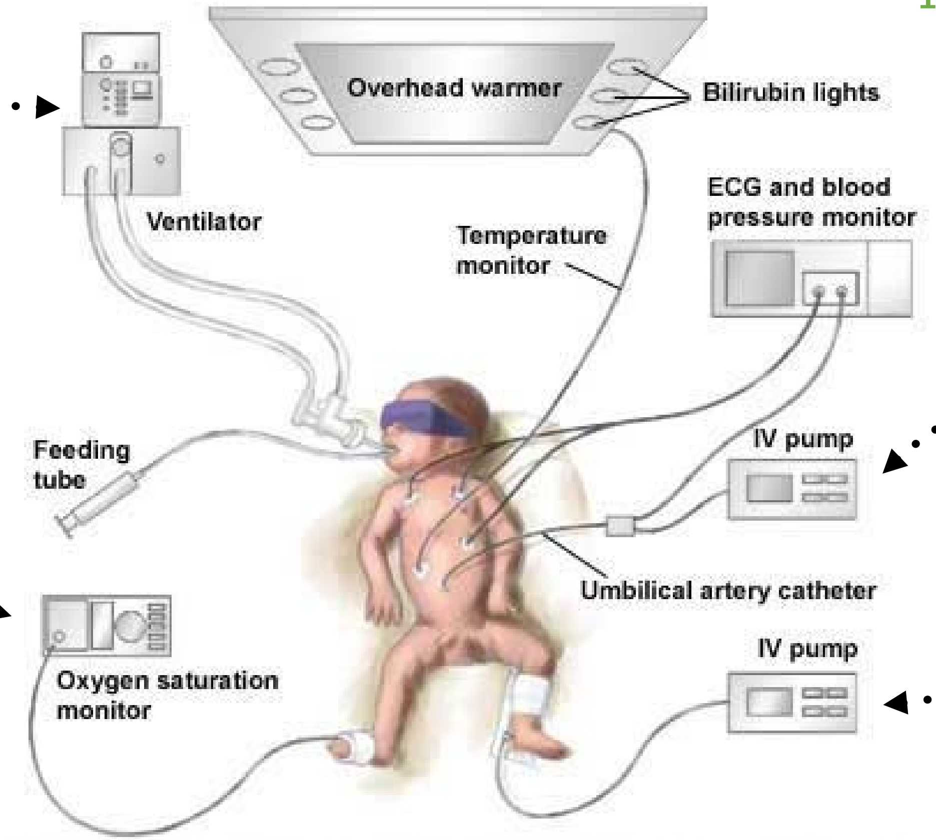
“How do we securely collect and analyze distinct data streams in real-time to guide medical decisions?”



Big Data in the NICU

1000+ Hz

100 - 1000 Hz



1 - 10 Hz

10 - 100 Hz

182 million readings a day per patient.

1 - 5 readings an hour

Characteristics of Physiologic Data

- **Unlike other sensor-based big data sources**
  - Has variances in signal quality, strength and frequently non-stationary
  - Changes over time as the infant grows and matures
  - New normal are established day-by-day
  - Contains numerous artefacts
  - Wasn’t meant to be streamed!

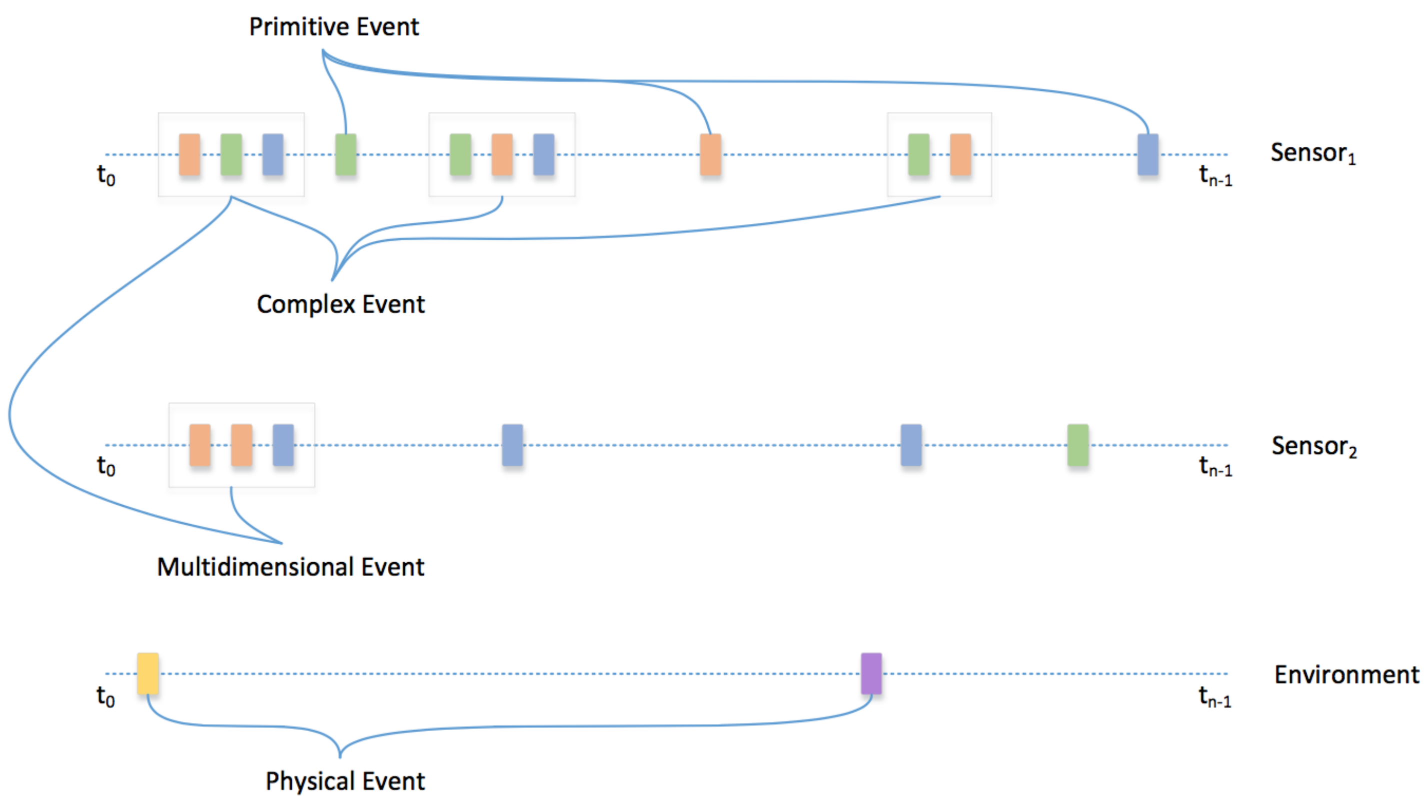
Bringing analytics to the bed-side

- Access the physiologic data to analyze every bit of information that can lead to ***personalized*** and ***earlier interventions***
- The provision of real-time physiologic analysis requires a multidimensional approach:
  - ***Multiple conditions***
  - ***Multiple streams*** of data
  - For which ***multiple behaviours*** can exist
- In addition, integrate of
  - Real-time synchronous medical device data
  - Asynchronous clinical data

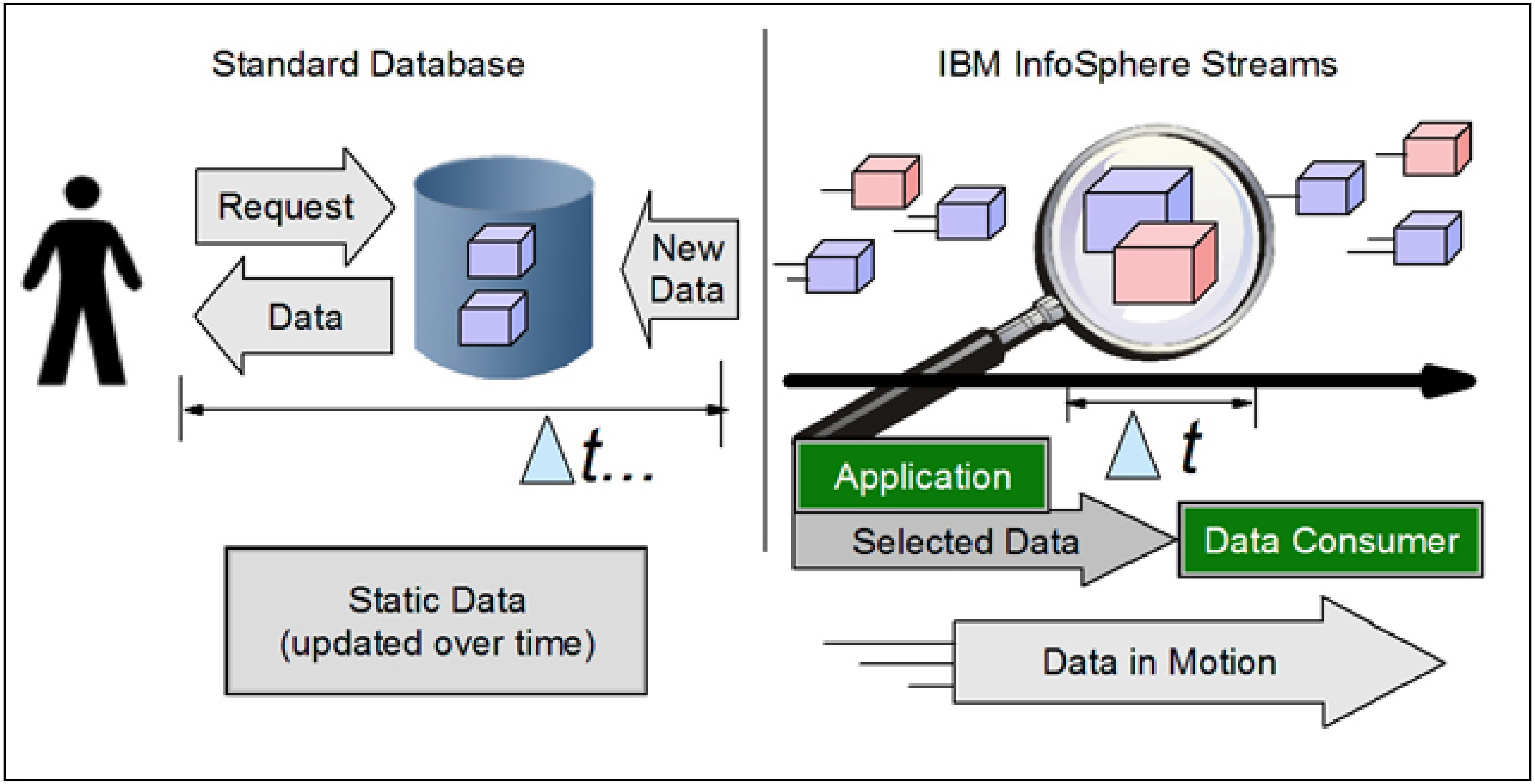
Sensor Analytics- Architectural Requirements and Development

Interactive Exploration of Event Streams

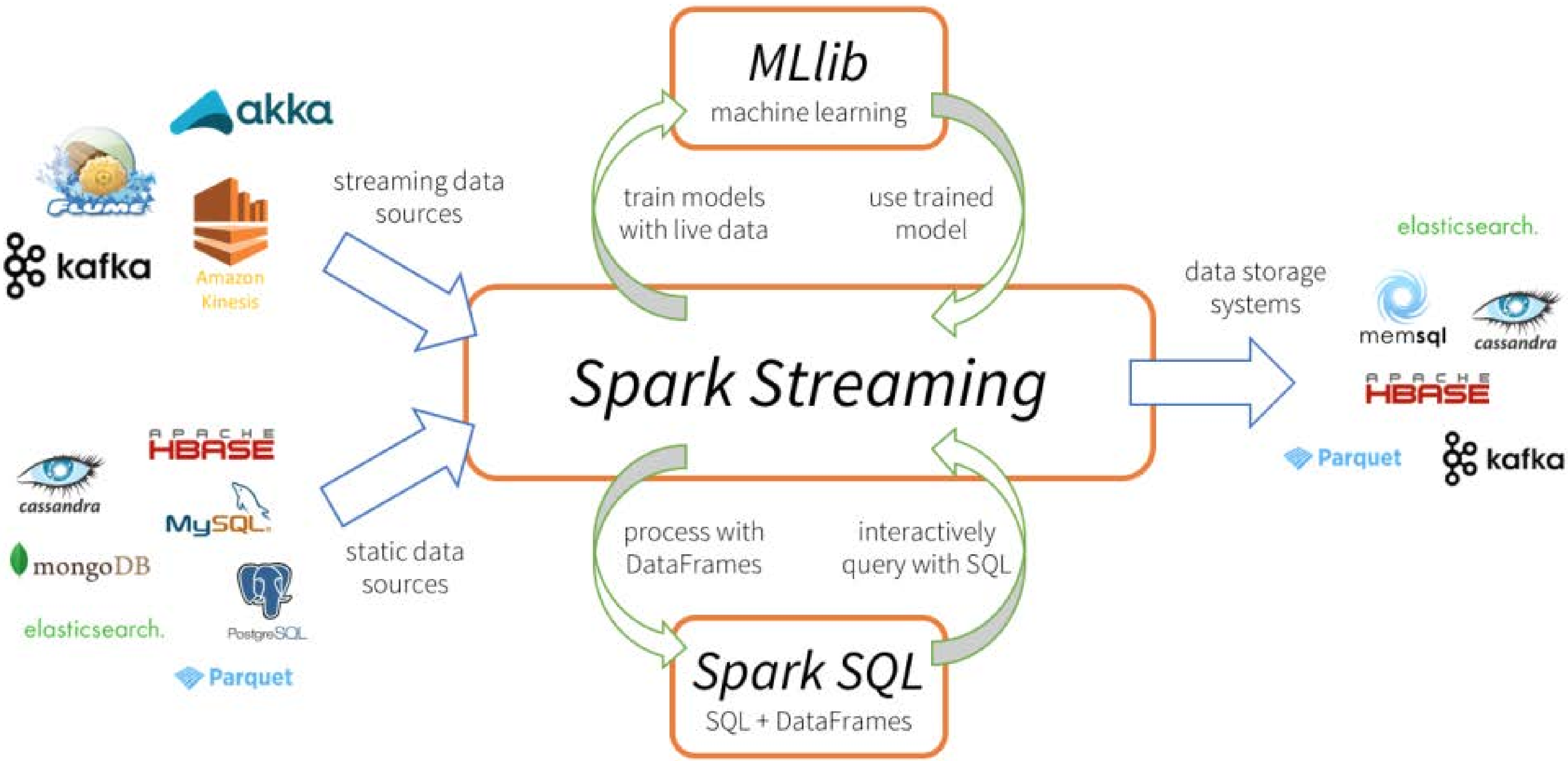
Event Streams– A hierarchy



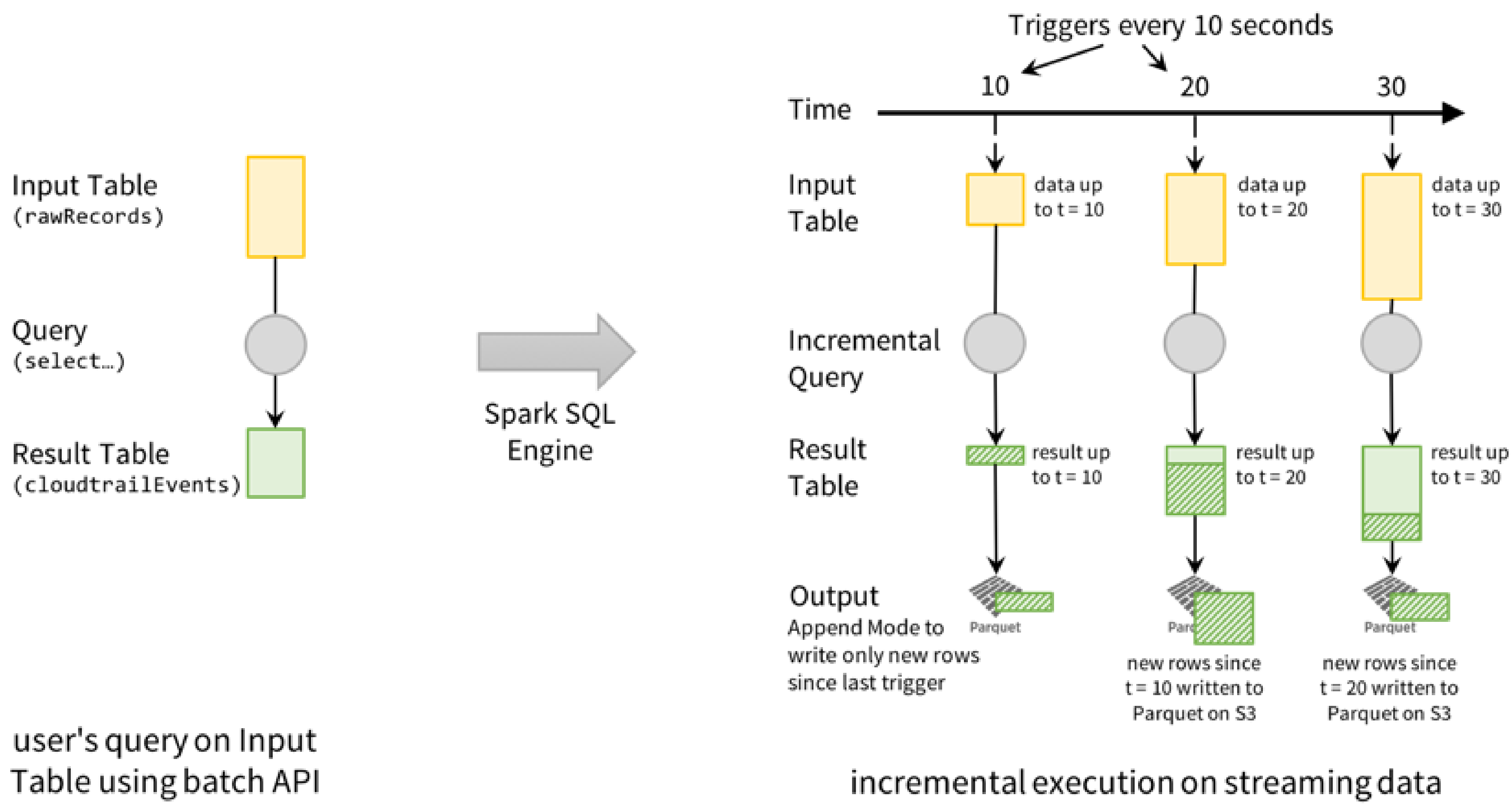
Event Stream Processing (InfoSphere Streams)



Event Stream Processing (Spark Streaming)



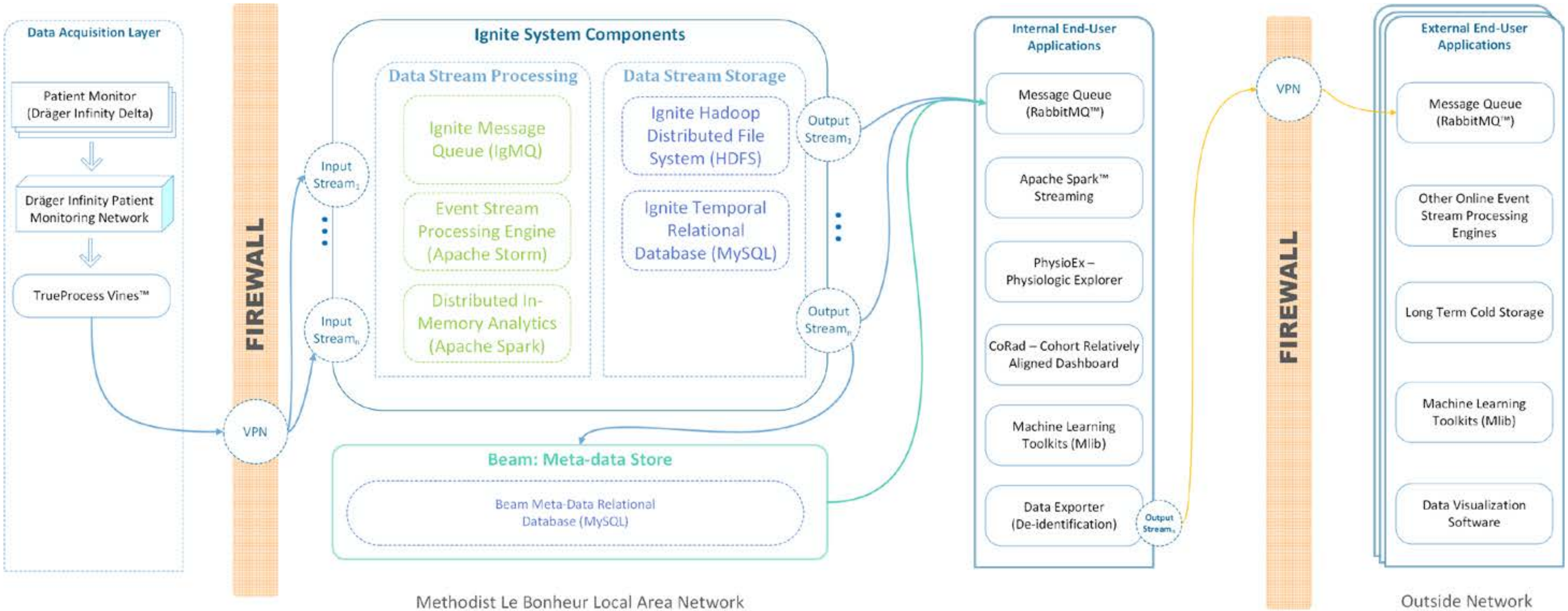
Spark Streaming– Structured Streaming



Structured Streaming Model

users express query on streaming data using a batch API; Spark incrementalizes them to run on streams

System Architecture for Sensor Analytics at Le Bonheur Children's Hospital

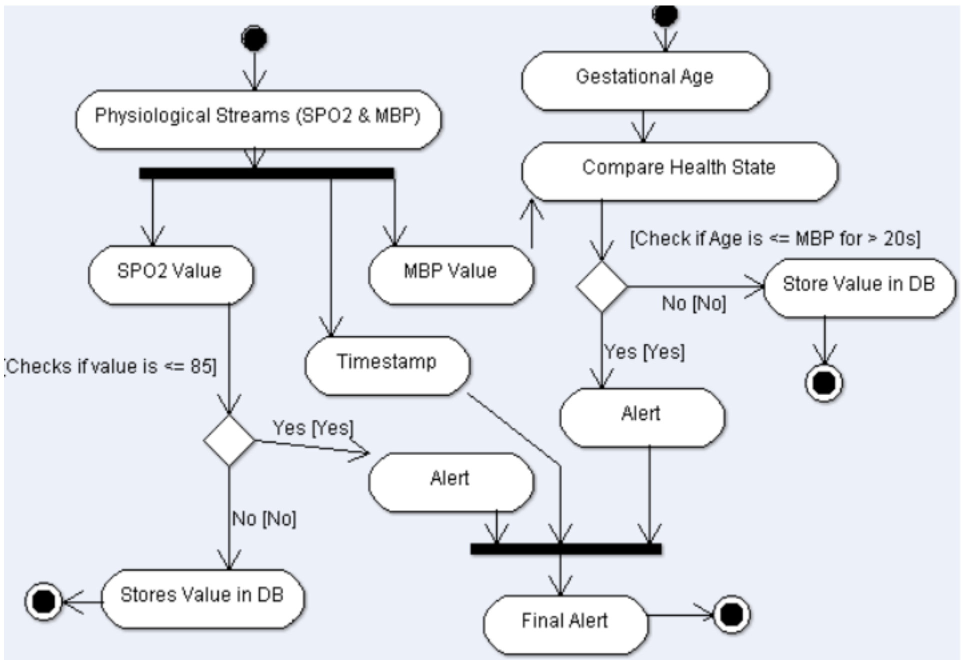


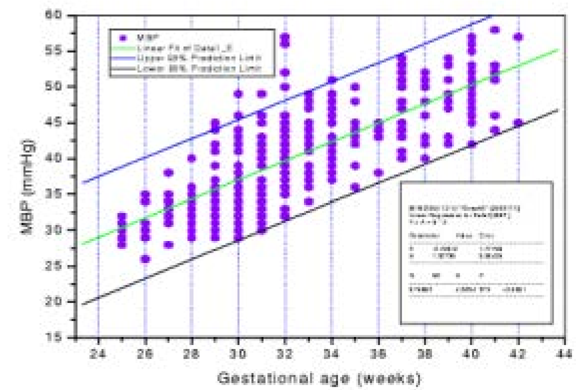
Data Capture | **Signal Processing** | **Data Consumption** | **Secondary Data Consumption**

Analytics at the Point of Care- Past work

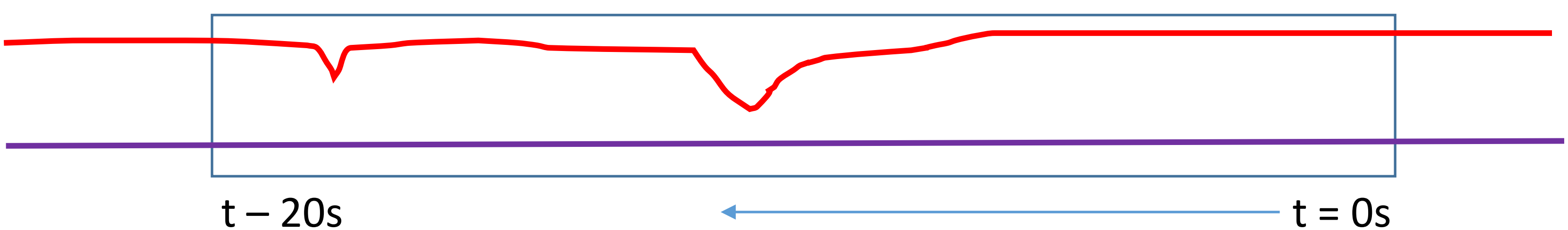
"Right information at the right time"

Kamaleswaran et al., 2010 : A method for clinical and physiological event stream processing



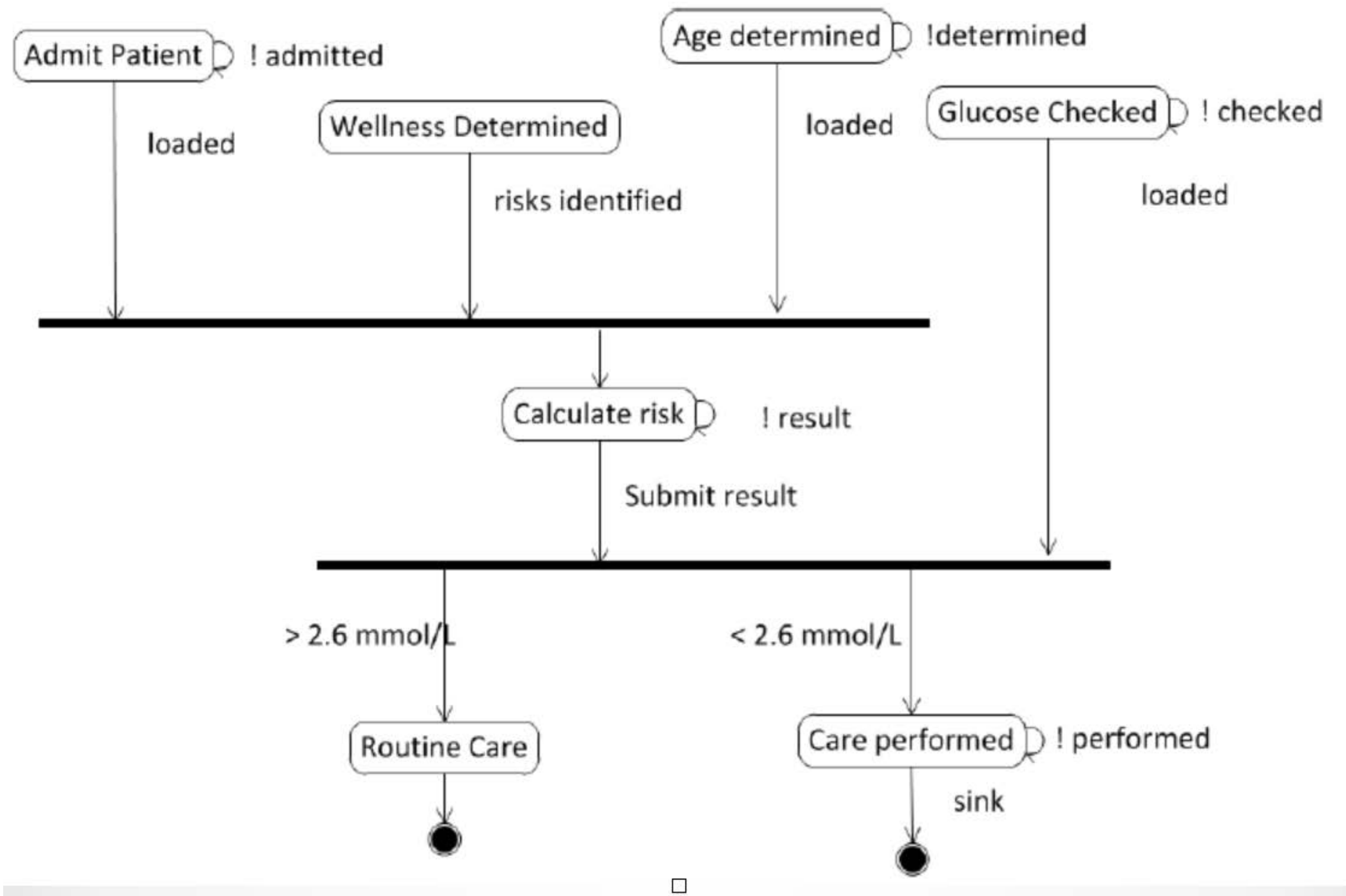


Mean blood pressure
vs Gestational age [1]



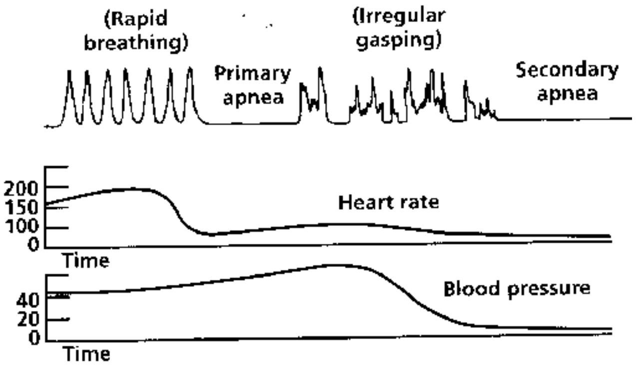
[1] Edward F. Bell, “Blood Pressure in the Newborn”. Iowa Neonatology Handbook, 2012.

Kamaleswaran et al., 2012 : Integrating complex business processes for knowledge-driven clinical decision support systems



Clinical Domain Challenges—Apnea of Prematurity

- Apnoea of prematurity: ***cessation of breathing*** for more than *20 seconds* [2]
- Challenge to ***visually assess*** the etiology and type of apnoea
- Requires extensive monitoring to determine diagnosis
- In lieu of that monitoring, bed-side staff broadly classify any cardiorespiratory event as a "***Neonatal Spell***" [3]

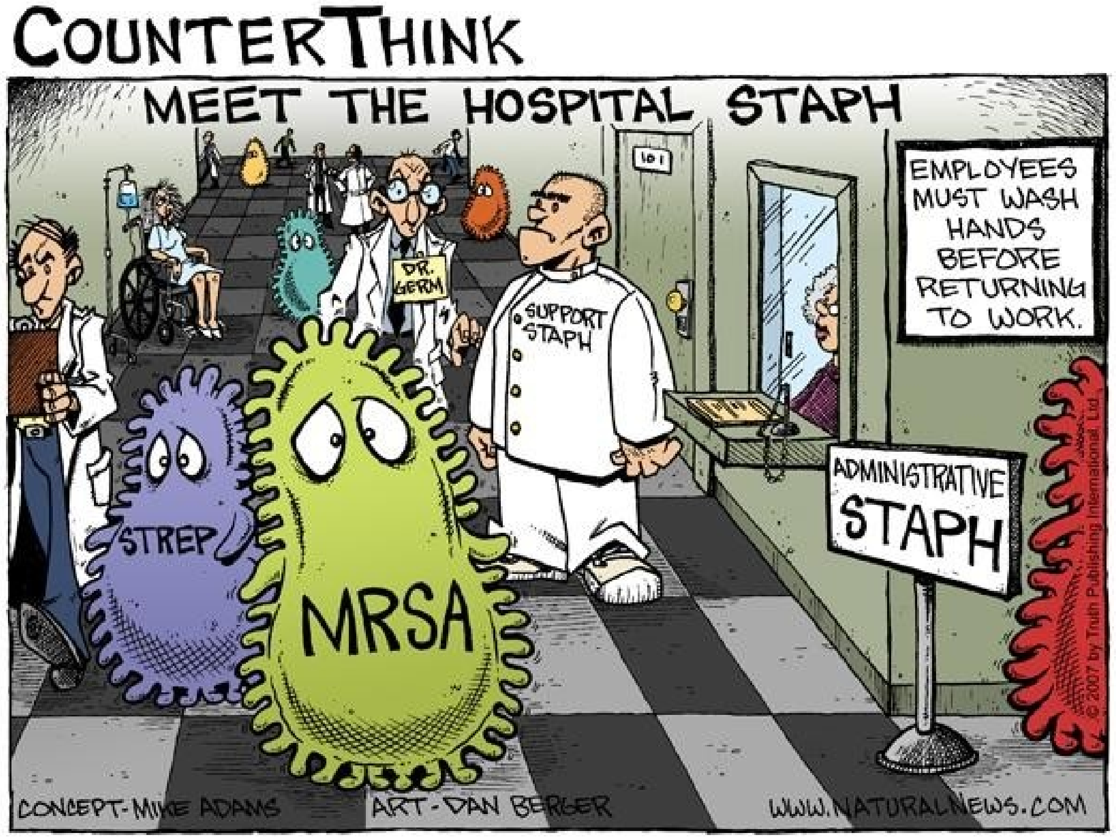


[2] M. J. Miller and R. J. Martin, "Pathophysiology of apnea of prematurity," Fetal and neonatal physiology, vol. 3, pp. 905–917, 1998.

[3] Thommandram A, Pugh JE, Eklund JM, McGregor C, James AG. Classifying neonatal spells using real-time temporal analysis of physiological data streams: Algorithm development. Point-of-Care Healthcare Technologies (PHT), 2013 IEEE. 2013. p. 240–3.

Clinical Domain Challenges—Nosocomial Infection

- ***Nosocomial infection*** (sepsis) is a common hospital-borne infection for babies receiving care in the NICU [4]
- A cocktail of antibiotics are prepared and administered even before diagnosis.
- ***Heart-Rate Variability*** seen as a potential indicator of the onset of neonatal sepsis [5]



[4] Goldmann, Donald A., William A. Durbin, and Jonathan Freeman. "Nosocomial infections in a neonatal intensive care unit." Journal of Infectious Diseases 144.5 (1981): 449-459.
[5] Moorman, J. Randall, et al. "Cardiovascular oscillations at the bedside: early diagnosis of neonatal sepsis using heart rate characteristics monitoring." Physiological measurement 32.11 (2011): 1821.

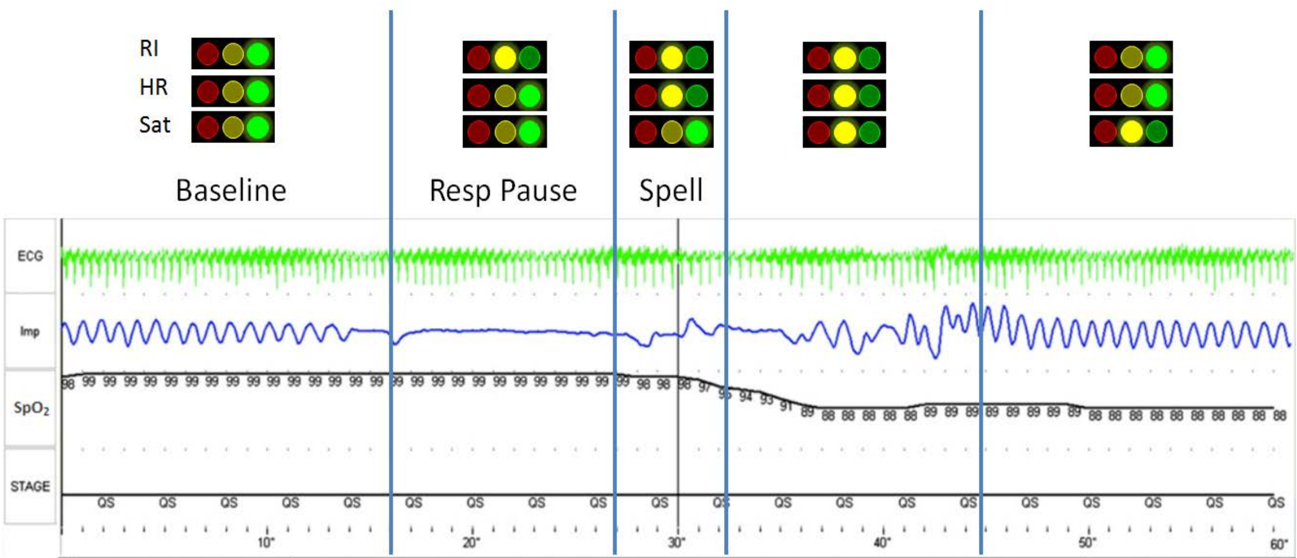
Neonatal Sepsis and Spells

Question: Can we identify neonatal spells prior to onset of nosocomial infection?

- Thommandram et al., built a neonatal ***spells algorithm*** [3] for detecting spells activity in real-time

Classifications:

- central apnoea
- vagal apnoea
- obstructive
- obstructive central
- central obstructive
- isolated blood oxygen desaturation
- isolated bradycardia.

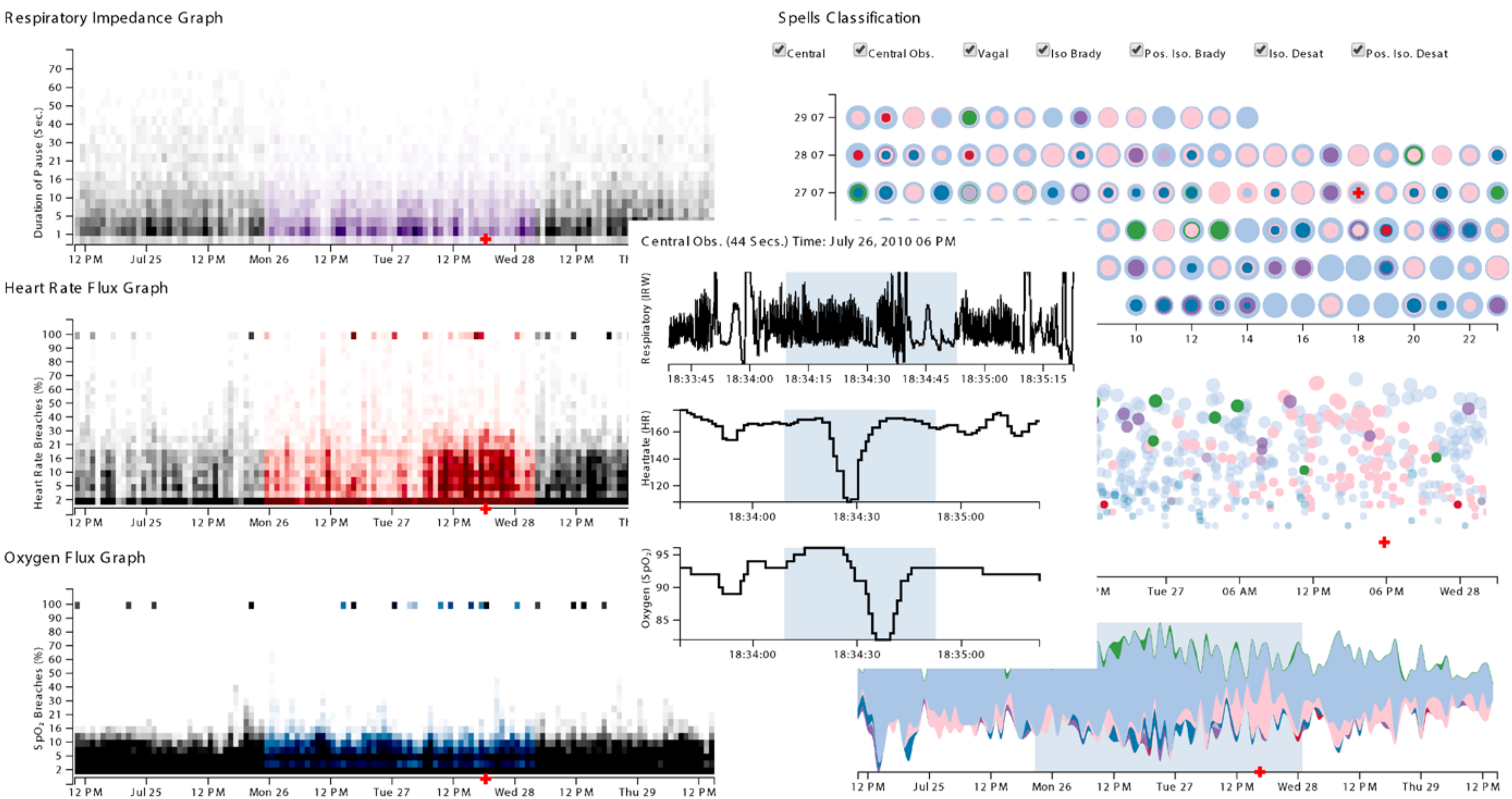


Modified from Sale, 2010

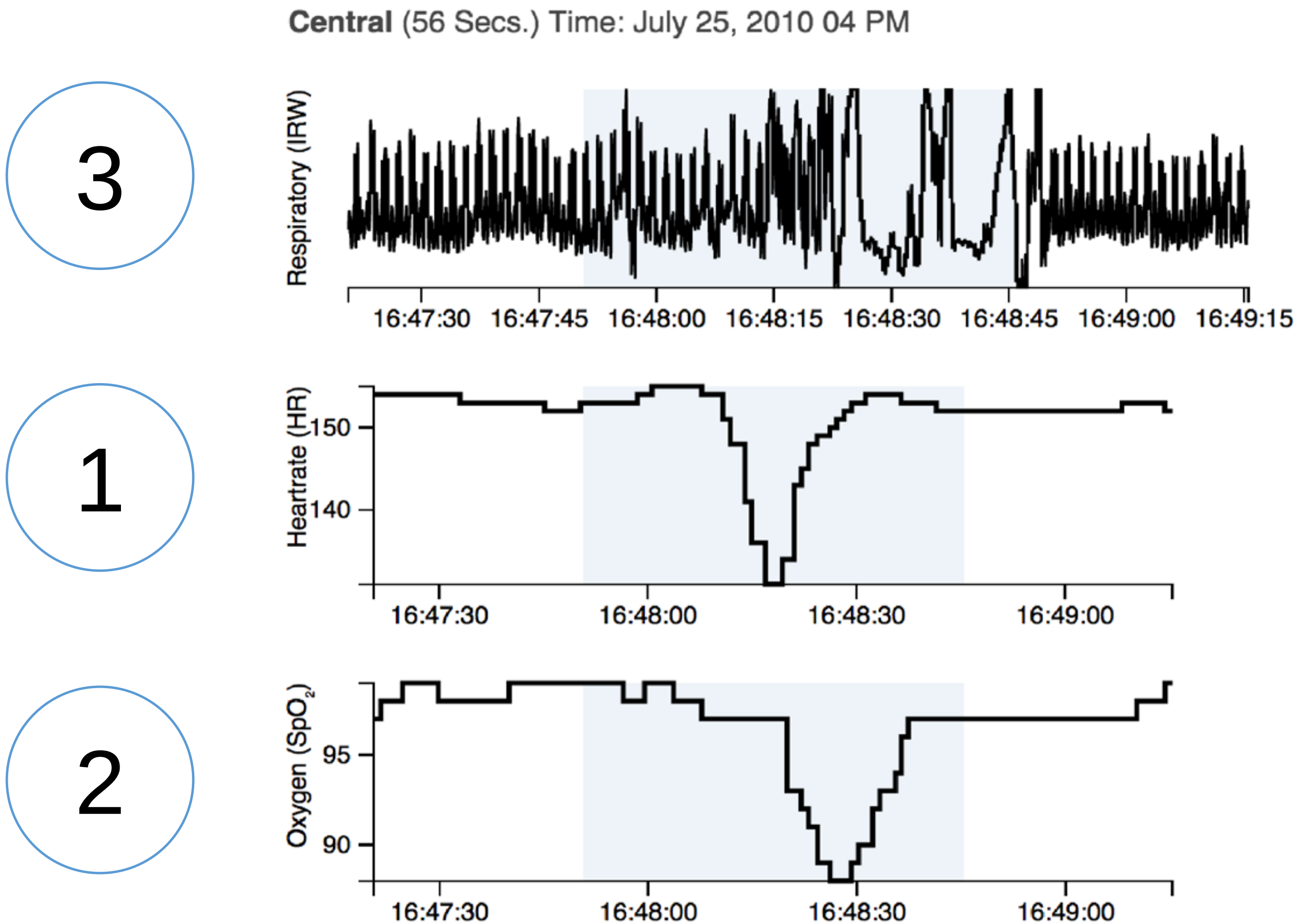
Events as Sequences

	1	2	3	4	5
Central	↓ Resp.	↓ Heart Rate	↓ O_2	↑ Resp.	↑ Heart Rate
Vagal	↓ Resp. ↓ Heart Rate	↓ O_2	↑ Resp. ↑ Heart Rate		
Obstructive	↓ Heart Rate (Incremental)	↓ O_2	↑ Heart Rate		
Obstructive Central	↓ Heart Rate (Incremental)	↓ O_2	↓ Resp.	↑ Resp.	↑ Heart Rate
Central Obstructive	↓ Resp.	↓ Heart Rate	↓ O_2	↑ Resp.	↓ Heart Rate
Desaturation	↓ O_2				
Bradycardia	↓ Heart Rate				

PhysioEx: Visual analysis of physiological event streams [6]



[6] Kamaleswaran, Rishikesan, et al. "PhysioEx: visual analysis of physiological event streams." *Computer Graphics Forum*. Vol. 35. No. 3. 2016.

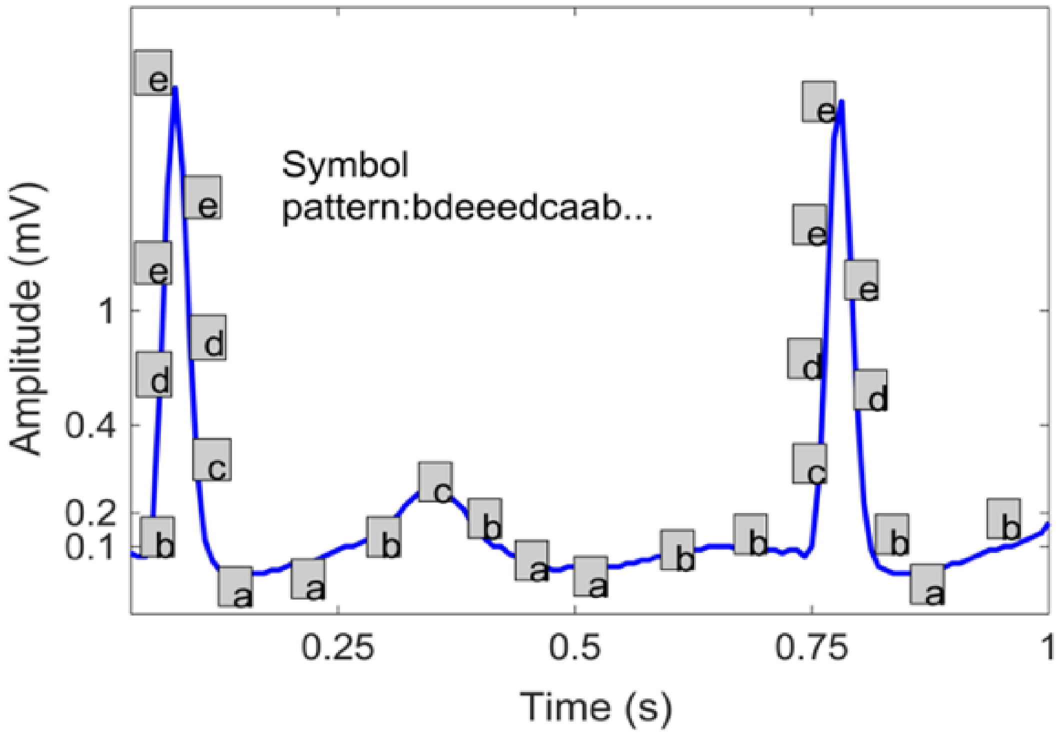

Central (56 Secs.) Time: July 25, 2010 04 PM
3
Respiratory (IRW)
16:47:30 16:47:45 16:48:00 16:48:15 16:48:30 16:48:45 16:49:00 16:49:15
1
Heartrate (HR)
150
140
16:47:30 16:48:00 16:48:30 16:49:00
2
Oxygen (SpO2)
95
90
16:47:30 16:48:00 16:48:30 16:49:00

Data-driven Approaches: Current Work

Knowledge Discovery for Point of Care Applications

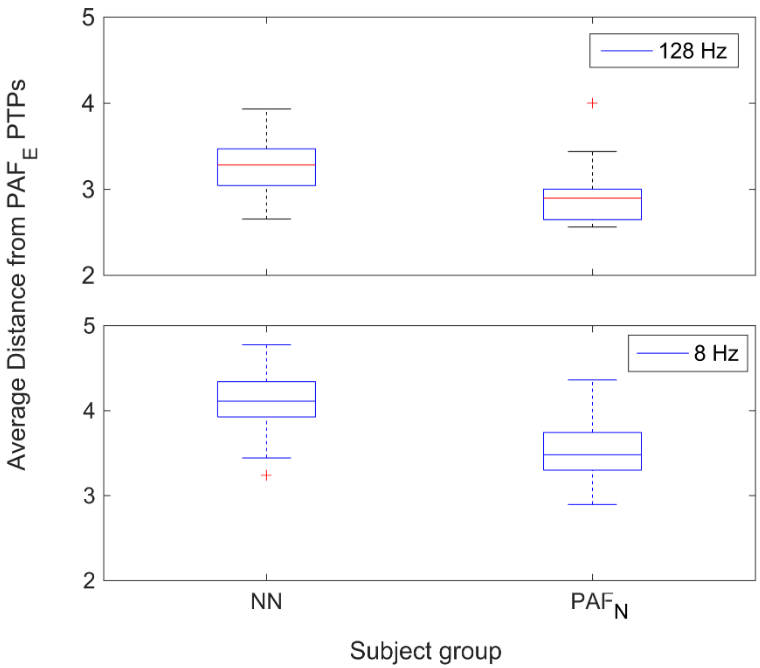
Continuous feature extraction: Probablistic Symbolic Pattern Recognition (PSPR) [7]

Symbol	Range
a	[-∞ 0.10]
b	[0.10 0.20]
c	[0.20 0.40]
d	[0.40 1.0]
e	[1.0 ∞]



[7] Mahajan, R., Kamaleswaran, R., & Akbilgic, O. (2017). Paroxysmal Atrial Fibrillation Screening at Different ECG Sampling Frequencies Using Probabilistic Symbolic Pattern Recognition. In *IEEE International Conference on Biomedical and Health Informatics 2017*. Orlando, FL

Paroxysmal Atrial Fibrillation (PAF) Prediction

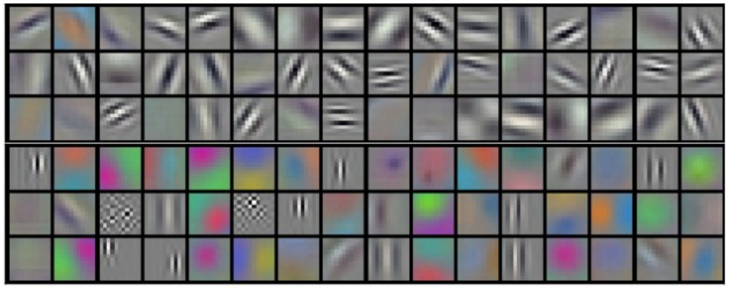


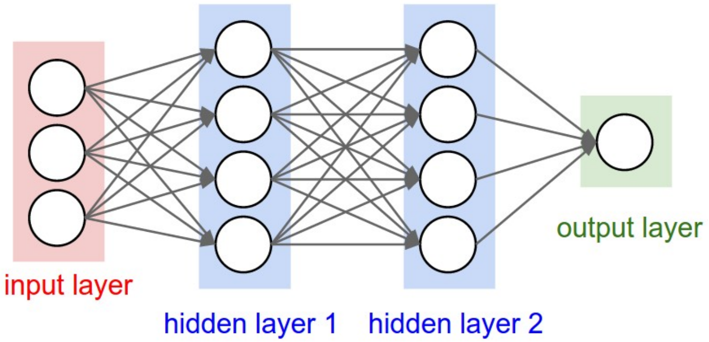
Sampling frequency, f_s (Hz)	Accuracy (%)	
	Training data	*Test data*
128	76.67	73.33
64	81.67	77.33
32	85	74.67
16	85.33	73.33
8	84.33	82.67

Sampling frequency indicates influence on accuracy of predictive model for PAF prediction

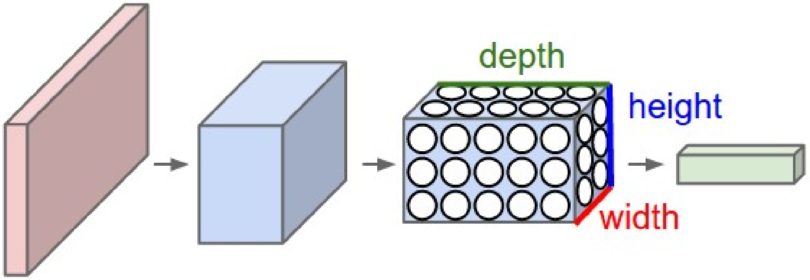
Deep Learning – Convolutional Neural Networks for Sensor Timeseries Analytics

- ConvNets provide a method of automatic feature extraction using a combination of convolutional filters and dimensional reduction
- Accuracies greater than humans have been observed using ConvNets on image data
- Features are extracted using multiple filters applied to the same image





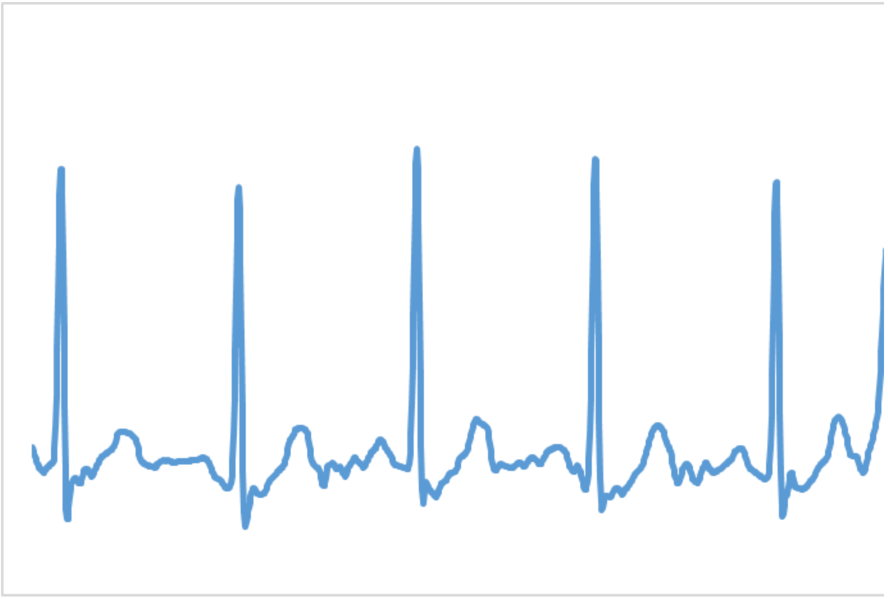
Multilayer perceptron Neural Network



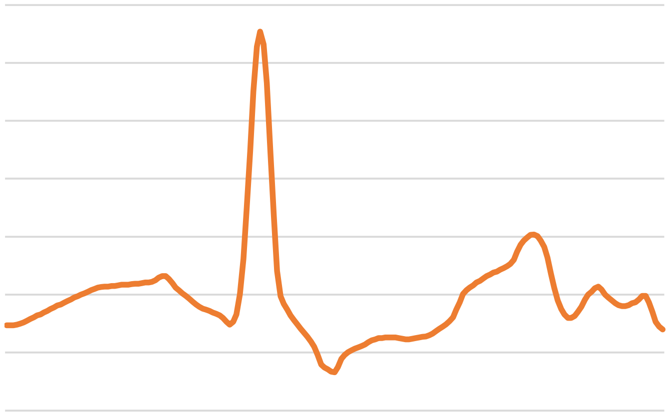
Convolutional Neural Network

Deep Learning – Convolutional Neural Networks

- ECGs are continuously sampled waveform data most remarkably distinguished by the prominent ‘R’ peak.
- An ECG signal can be sampled anywhere from a couple of minutes to hours
- We abstract a continuous waveform and segment it by it’s individual complex



A sample ECG waveform at 300 Hz



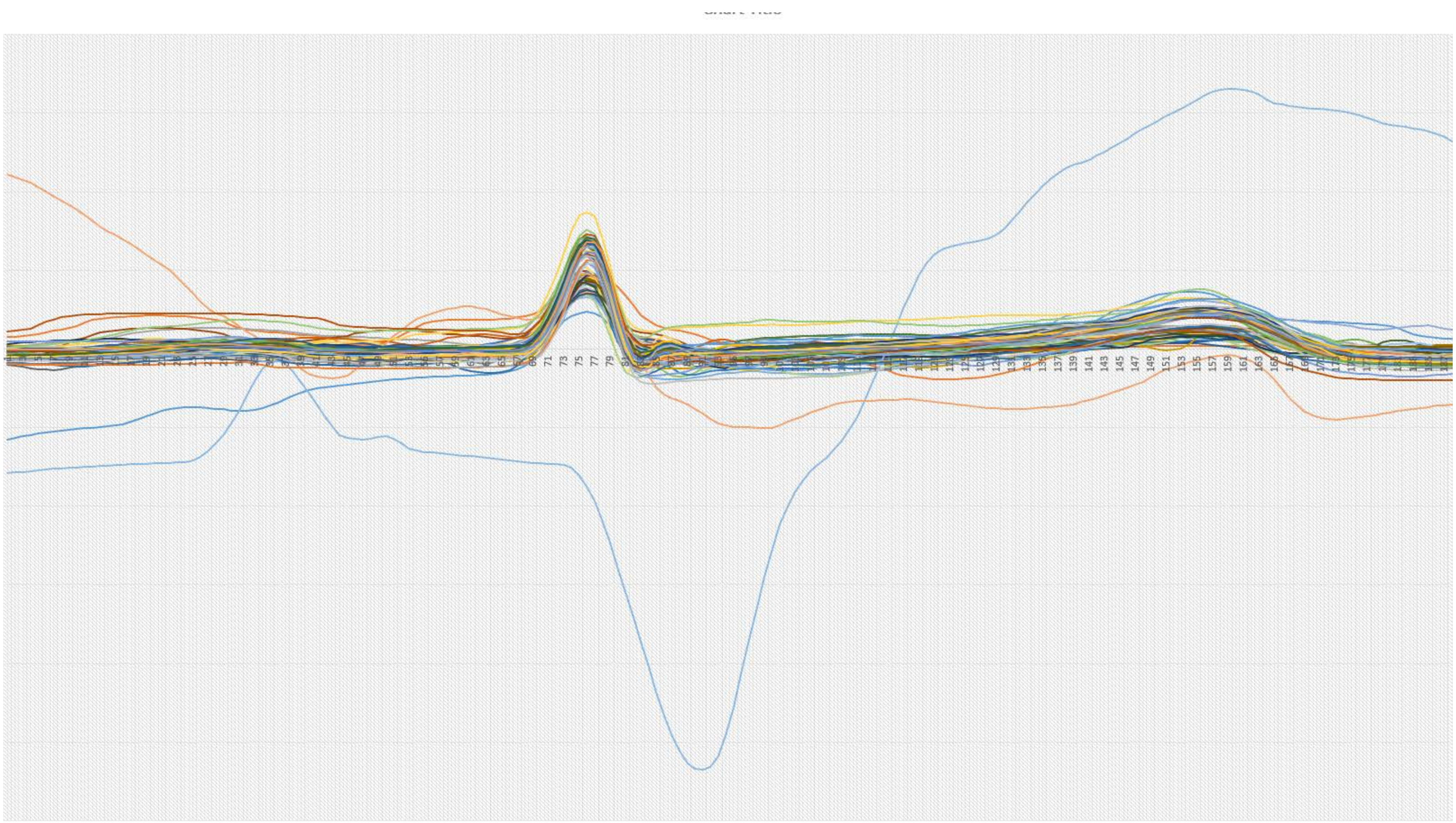
A sample ECG complex at 300 Hz

Deep Learning – Convolutional Neural Networks

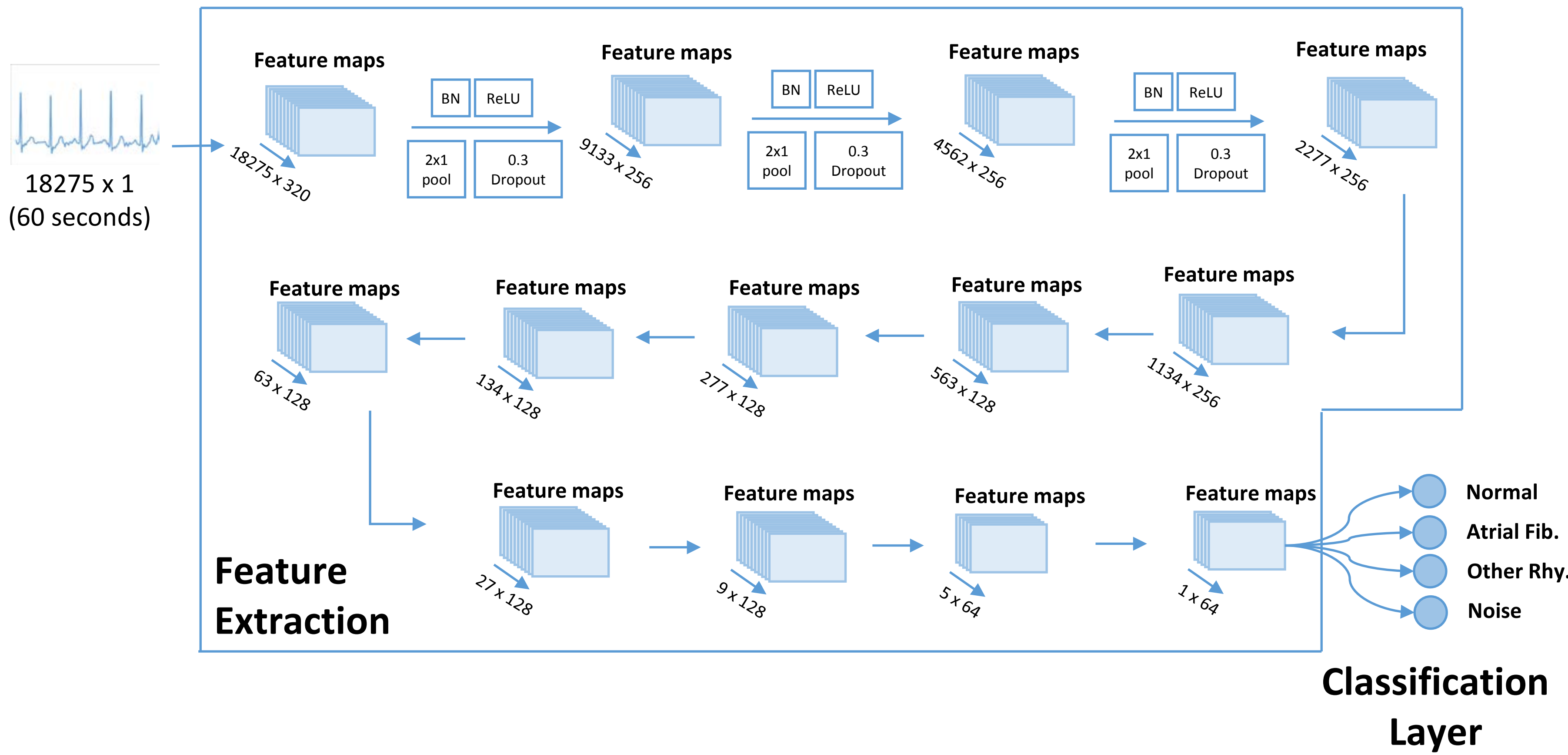
- Each complex is segmented into a square matrix that is then fed into the ConvNet for feature extraction


```
2,"[-185. -193. -200. -207. -212. -214. -217. -219. -222. -224. -226. -228.
 -229. -230. -232. -233. -234. -235. -236. -237. -238. -239. -238. -237.
 -235. -233. -230. -227. -224. -220. -215. -210. -204. -198. -178. -137.
 -101.  -74.  -60.  -49.  -50.  -68.  -92. -127. -160. -184. -199. -213.
 -228. -248. -263. -272. -279. -286. -292. -297. -305. -336. -388. -414.
 -433. -445. -434. -399. -351. -298. -244. -193. -163. -137. -118. -102.
  -90.  -81.  -72.  -62.  -51.  -43.  -35.  -28.  -21.  -15.   -9.   -3.
    3.   10.   16.   23.   30.   36.   40.   44.   47.   50.   53.   57.
   61.   66.   71.   76.   82.   89.   99.  109.  118.  126.  134.  143.
  152.  161.  169.  176.  183.  190.  196.  202.  209.  216.  223.  231.
  241.  254.  266.  278.  286.  292.  298.  302.  304.  304.  301.  296.
  290.  282.  273.  259.  242.  224.  206.  187.  171.  156.  138.  117.
   87.   71.   61.   53.   47.   42.   36.   31.   28.   25.   22.   19.
   14.    8.   -2.  -12.  -21.  -24.  -26.  -28.  -30.  -32.  -34.  -37.
  -38.  -40.  -41.  -42.  -42.  -42.  -41.  -42.  -42.  -43.  -44.  -46.
  -48.  -50.  -53.  -56.  -58.  -61.  -63.  -66.  -68.  -71.  -75.  -81.
  ]"
```

Influences of Noise on Sensor Data

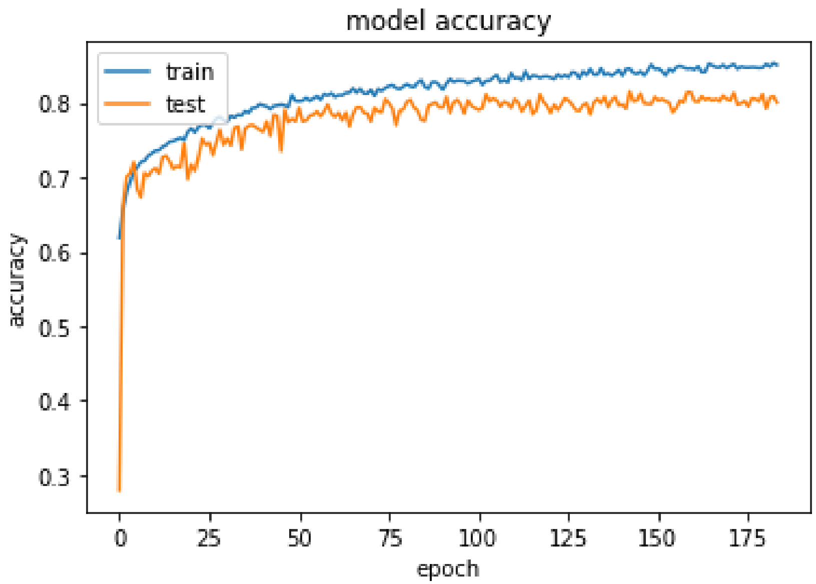


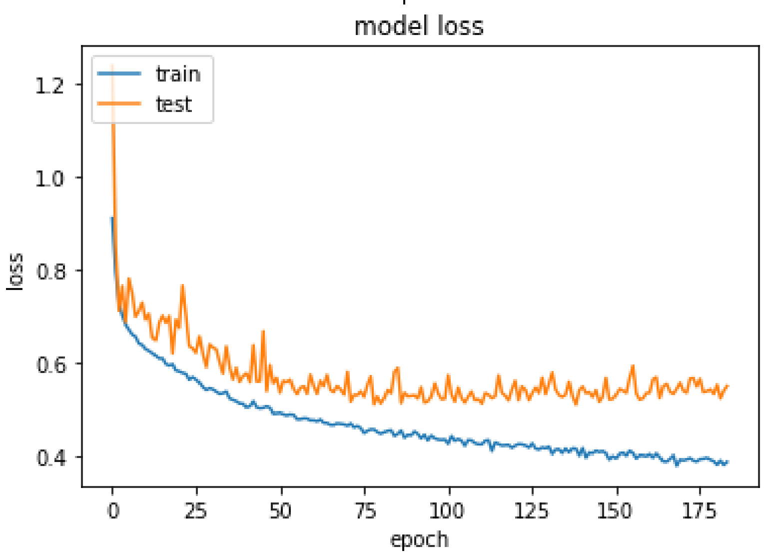
Deep Convolutional Neural Network Topology

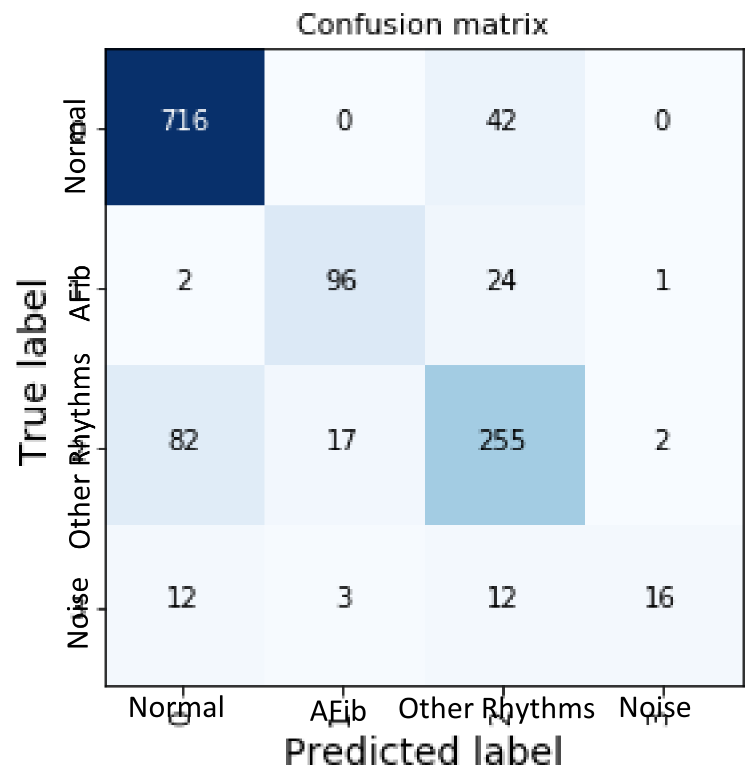


Total params: 2,042,820

Deep Convolutional Neural Network Results







Train Score: acc: 89.14% Test Score: acc: 84.69%

Take home message

- Every breath, every beat –Critical Care
  - Event stream processing a potent tool to uncover new knowledge
  - Collection of high quality data is still a pressing challenge
- Not enough to simply “analyze” the data
  - Knowledge is a holistic experience
  - Feedback loop reinforces algorithm learning
- Deep learning methods offer novel means to classify difficult and very noisy ECG sensor data

Thank You!

Rishi Kamaleswaran

rkamales@uthsc.edu
@rkamaleswaran

Le Bonheur Children's Hospital

The University of Tennessee Health Science Center

Oak Ridge National Laboratory