

## How to Tame Your Data: The Journey from Clinical Telemetry to Research Data

### Abstract:

#### Objective

In this case study we present the hardware, software, and process challenges to develop an efficient method of converting high throughput telemetry data from the patient bedside to a clinical data store.

#### Materials and Methods

The pilot is expected to collect 50TBs of telemetry data over 9 months from 270 beds across two hospitals in an academic medical center. An iterative process was used to optimize the storage and conversion of raw telemetry data into a relational database. The iterations occurred as the result of an investigation of performance limiters. 7-zip, Python, SQL, and BCP were used to compress, extract, transform, and load data into the final SQL database.

#### Results

The time to convert a 125GB benchmark of raw telemetry data reduced from more than 260 hours to 85 minutes.

#### Discussion

Prototyping, network optimization, parallelization, SQL query mapping, and database architecture design were used to optimize this process. Additionally, five heuristics were identified:

1. Start small and scale incrementally.
2. Parallelize everything that should be.
3. Perfect is the enemy of good
4. Design SQL architecture to ensure tables are queryable.
5. Design processes to minimize network reliance.

#### Conclusion

Many of our experiences were unpredictable. Our challenge is one that a growing number of informaticists face as the scale of data grows. Our goal in writing this paper is to provide other informaticists the information we wish we had before starting.

#### About the Speaker:

John Lawrence, MS

BS in Chemical Engineering and Biochemistry Molecular Biology from Rose-Hulman Institute of Technology  
ME in Biomedical Engineering from University of Cincinnati

Epic Systems

The Center for the Advancement of Team Science, Analytics, and Systems Thinking in Health Services and Implementation Science Research (CATALYST) at OSUWMC with Dr Timothy Huerta  
Biomedical Sciences Graduate Program

Active Areas of Research

Infant Mortality in the State of Ohio

Hospital - State Agency Communication and Health Information Exchanges

Telemedicine and eICU availability

FHIR and Healthcare Interfaces

Common Data Models from Primary and Secondary Healthcare Data Sources

Rigor and Reproducibility