Better ways to manage data

Paul Myrda
Technical Executive

NASPI - Technical Workshop
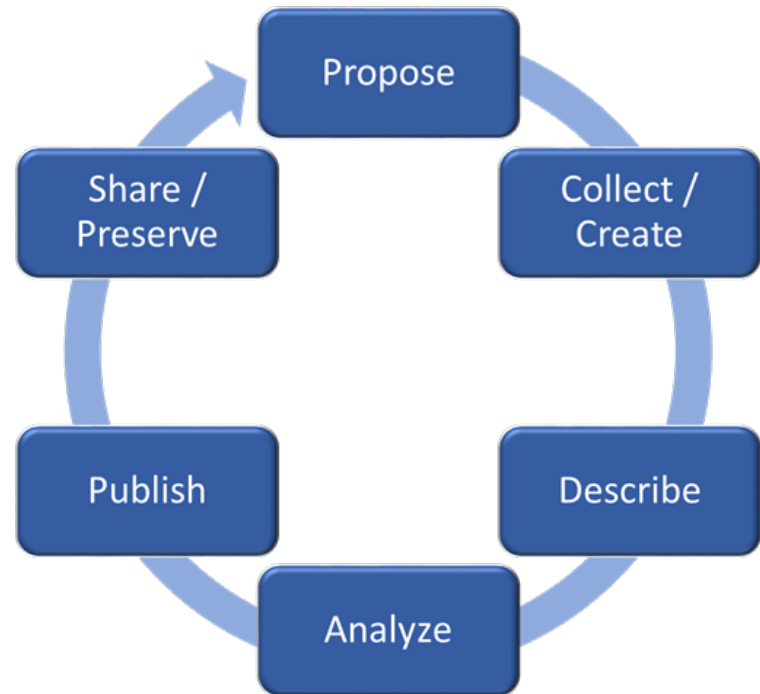**Synchrophasor Data Quality Management**
and Improvement

March 21, 2016

Overview

- Data Life Cycle Phases
  - Propose
  - Collect / Create
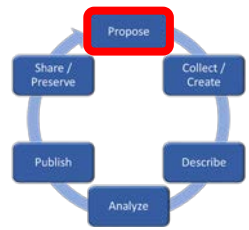 - Describe
  - Analyze
  - Publish
  - Share / Preserve
- 2015 Survey Results
- Business Process Recommendations

Courtesy of UC San Diego Library; http://libraries.ucsd.edu/services/data-curation/data-management/index.html

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Propose Phase

Need to define the question "What are you going to do with the data?" and then establish the data management plan.

- What are the "Use Case(s)" for the synchrophasor data.

- The Use Case(s) establish the basic need for the data.

- Without them utilities are hampered in developing a suitable data management plan.

- Early on data classes were established to provide some guidance.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Collect / Create Phase

Defines how to collect the data and ensure its suitable creation including any data clean-up.

- Installing the PMU at a substation and connecting it to a suitable communication network interface and send the data to a master collection point.

- PMU registry helps capture key information about the PMU and its features.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Describe Phase

It is essential to capture all the background information about a project that seems so obvious at the time but which may very well be forgotten in just a few years.

- Utility design documents that includes information such as any data processing that is performed especially along the data path such as at a phasor data concentrator (PDC).

- Record of PDC configuration settings and also firmware version.

- Use of a data collection template to ensure completeness of data collection and promotes early consideration of data description.

- Purpose as to why the data was chosen for collection. Especially for some of the status information available from the PMU regarding time quality and other such parameters.

- Standards or calibrations used in the installation or subsequent maintenance.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Describe Phase (continued)

When designing the data storage files

- Create a data dictionary that documents the contents of data files.

- Description of the parameters captured including formats, units of measure, and definitions of any coded values such as status codes

- Complete list of field names used and to the extent reasonable use standardized names across the project. The use of industry models such as CIM or IEC 61850 may be of use.

- Description of how the parameter is acquired, what a missing or null values signifies and what data quality parameters mean.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Describe Phase (continued)

Metadata is data about your data.

- Two most useful to synchrophasors are descriptive and technical.

  - **Descriptive metadata** is the most common and familiar kind of metadata that describe an object. Examples include a title, date, and source.

  - **Technical metadata** is information about technical aspects of the data such as format-specific technical characteristics, how an object was created, and storage and location information.

- Creating metadata makes it possible for others to find and use the data properly.

- Without good metadata, a company might not be able to reuse their own data a few years from now! Develop and use a standard approach for metadata.

- A standard, thorough handling and format for metadata enables data exchange with other entities and protects long term data usability

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Analyze & Publish Phases
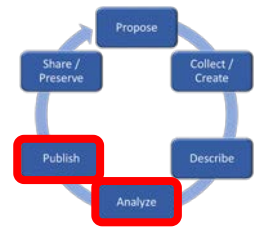
These are not data management functions but rather the role of a variety of synchrophasor applications that are readily available.

Preserve Phase

The preserve process establishes the data archive for subsequent retrieval and use.

- Storing and identifying the data is critical.
  - Storing data for the long-term is easily managed by existing IT processes, however,
  - Identifying all the data needed for long-term is more challenging since a number of different file types that may be stored in different physical places may have to be assembled to effectively recreate a prior study.

- Long term preservation procedures are necessary to maintain the value of post event data.
  - This procedure needs to identify the datasets needed, the location of the data sets and file access permissions.
  - This is best accomplished using a commercial records management tool that can assemble and archive a wide array of file types and create the referential integrity necessary to prevent accidental deletion.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Industry Survey Results on Synchrophasor Data Archiving, Retention, Destruction

This section based on a 2015 survey member companies.

Only about 50% of the participants actually responded to the questions

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Archiving

Data archiving is the process of moving data that is no longer actively used to a separate data storage device for long-term retention. Data archives consist of older data that is still important and necessary for future reference, as well as data that must be retained for regulatory compliance. Data archives are indexed and have search capabilities so that files and parts of files can be easily located and retrieved.

- **Archiving time**
  - Typical data archiving time period is 3 years.
  - Range was daily archiving to not archiving at all.

- **Duration**
  - Shortest duration 30 days.
  - Majority of responses indicated a policy of 3 years
  - Maximum length was 7 years
  - One company indicated that all data was on-line.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Retention

Data Retention is defined as the act of retaining or to keep possession or ownership of individual facts, statistics, or items of information (1)

- Data Retention
  - ~50% indicated that they retain their data for 3 years.
  - ~50% indicated that they retain their data for 30 days.

- Event Data
  - ~50% indicated they retain all events.
  - Many indicated permanent retention of events while others indicated a period of 7 years.
  - One company indicated that all events are retained permanently.

- Lastly one company indicated that all data is retained.

(1) 2013 Merriam-Webster, Incorporated

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Destruction

Data <u>Destruction</u> the act or process of damaging something so badly that it no longer exists or cannot be repaired (1)

- ~50% indicated 3 years
- ~33% indicated after about 30 days.
- One company indicated that no data gets destroyed.

- Various methods of destruction were identified. In some cases the action was for the data to be overwritten.

(1) 2013 Merriam-Webster, Incorporated

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Business Process Recommendations

Synchrophasor Data Performance Metrics

- **Effectively using PMU data for Operations**
  - Need to understand how synchrophasor data can be corrupted or lost between the point of measurement and the final destination,
  - How data issues affect the results of the final application

- **Need to characterize application trustworthiness**
  - application's sensitivity to data errors or loss can allow the creation of confidence intervals
  - Serve as a basis for network upgrades

- **NASPI members have formed the Phasor Applications Requirements Task Force (PARTF) to address this set of challenges.**

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Develop Appropriate PMU Network Architectures

- The impact of PMU Network Architecture on data performance may be significant depending on the applications in use.

- The actual architecture employed and the settings selected for the various assets along the communication channel are critical to successful data collection and may contribute to latency and data drop outs especially in relation to PDCs.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE
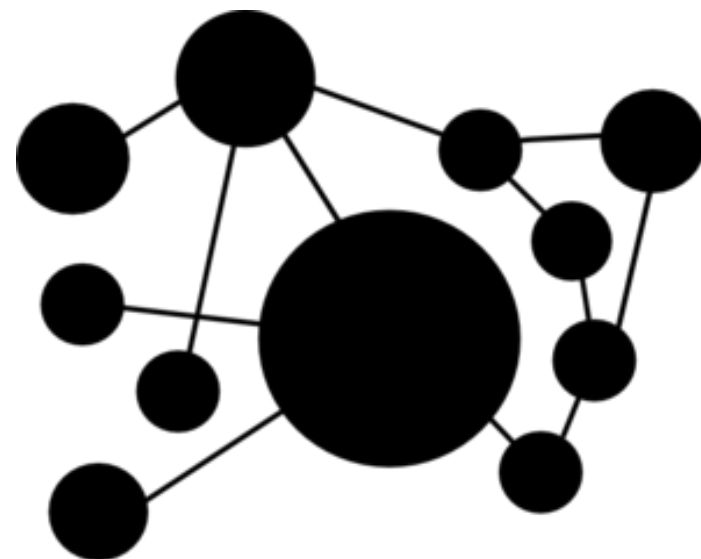
Develop Data Archiving, Retention, and Destruction Policies

- The lack of archiving, retention, and destruction policies is indicative of the immaturity of the synchrophasor business processes.

- Need to instill the rigor of IT business processes on synchrophasor data stores.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Develop Utility Specific Synchrophasor Application Roadmap

- A utility roadmap would help establish the business requirements needed to support the application and also provide guidance to the IT staff on future requirements.

- The roadmap would help in establishing the communication requirements and also the data storage, archiving, retention, and destruction policies.

EPRI | ELECTRIC POWER RESEARCH INSTITUTE
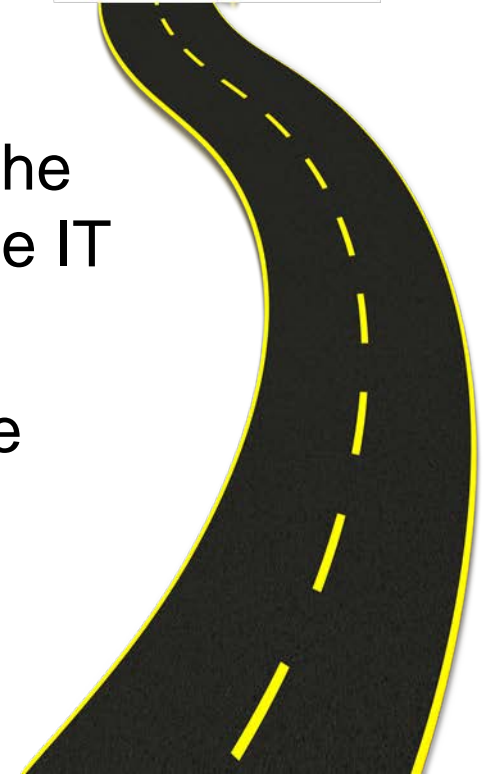
Summary Business Process Recommendations

- Establish synchrophasor data performance metrics

- Develop appropriate PMU network architectures

- Develop data archiving, retention, and destruction policies

- Develop utility specific synchrophasor application roadmap

Together…Shaping the Future of Electricity

EPRI | ELECTRIC POWER RESEARCH INSTITUTE