# A DOMAIN-SPECIFIC LANGUAGE APPROACH TO SUPPORT MONITORING AND SURVEILLANCE WITHIN WHOLESALE ELECTRICITY MARKETS

A Dissertation Proposal

by

ERIC CHARLES GRASBY

Submitted to the Graduate School
University of Arkansas at Little Rock
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Computer and Information Sciences (CIS)

Information Quality Track

Chair of Committee, Dr. Daniel Berleant, Professor of Information Science

Committee Members, Dr. Maria Gingras, Lead Market Monitor, Southwest Power Pool

Dr. Elizabeth Pierce, Department Chair, Department of Information Science

Dr. Ningning Wu, Professor of Information Science

September 2025

Department of Information Science

M.S., Information Quality, UA Little Rock, 2020.

G.C., Information Quality, UA Little Rock, 2020.

B.S., Information Science, UA Little Rock, 2018.

This dissertation **proposal**, "A Domain-Specific Language Approach to Support Monitoring and Surveillance Within Wholesale Electricity Markets", by Eric Charles Grasby, is approved by:

Dissertation Advisor:

Dr. Daniel Berleant
Professor of Information Science

Co-Advisor:

Dr. Maria Gingras
Lead Market Monitor, SPP

Dissertation Committee:

Dr. Elizabeth Pierce
Professor of Information Science

Dr. Ningning Wu
Professor of Information Science

Program Coordinator:

Dr. John Talburt
Acxiom Chair of Information Quality

Graduate Dean:

Dr. Brian Berry
Professor of Chemistry

*This page is intentionally retained for inclusion in the final dissertation*

ABSTRACT

This dissertation investigates an intersection between multiple domains of expertise related to electric markets and electric market monitoring. These domains are: 1) data and information quality, 2) language engineering, and 3) knowledge management. Electricity markets present unique opportunities for deriving new value from data and information assets, due to the enormous volume of data generated, each day, by public utilities, transmission and market operators, and regulators. These data assets are critically necessary to both ensuring reliable operation of the North American Bulk Electric System (BES), as well as operating efficient markets for selling energy on a wholesale scale. The outcomes of these markets have tangible, long-term impacts on system reliability, prices for residential and commercial energy consumers, and electricity infrastructure development. Due to the complex nature of energy marketplaces, there are abundant opportunities for market participants to manipulate these markets. Market monitors are the professionals who work to ensure that the markets utilizing the BES remain fair, efficient, and open-access in the face of such manipulative actions. They perform routine surveillance and forensic analysis to identify instances of fraud and market manipulation and refer such cases to regulatory authorities.

This research investigates the efficacy of designing and implementing a layer of abstraction between market monitors and the data that they use to perform their forensic analysis. Using domain-specific programming language engineering and information quality principles, the intent is to develop a system to ease data manipulation for market monitors.

## INDUSTRY ACRONYMS

| | |
|---|---|
| APAC | Asian-Pacific Region |
| AS | Ancillary Services |
| BES | Bulk Electric System |
| BSS | Bilateral Settlement Schedule |
| CRM | Capacity Reserve Margin |
| DA | Day-Ahead Market |
| DSL | Domain-Specific Language |
| DSRM | Design Science Research Methodology |
| DQ | Data Quality |
| EMEA | Europe, the Middle-East, and Africa region |
| ETL | Extract-Transform-Load |
| FERC | Federal Energy Regulatory Commission |
| GPL | General-Purpose Language |
| HHI | Herfindahl-Hirschman Index |
| IQ | Information Quality |
| ISO | Independent System Operator |
| LMP | Locational Marginal Price |
| MMU | Market Monitoring Unit |
| MP | Market Participant |
| MW | Megawatt |
| NERC | North-American Energy Reliability Corporation |
| RTO | Regional Transmission Organization |
| RT | Real-Time Market |
| SME | Subject Matter Expert |
| SOX | Sarbanes-Oxley Act of 2002 |
| TCR / FTR | Transmission Congestion / Financial Transmission Rights |
| TO / TOP | Transmission Owner / Transmission Operator |

TABLE OF CONTENTS

## LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

# 1. INTRODUCTION

This dissertation introduces information quality research into the electricity market monitoring industry, a regulatory component of organized electric markets in the United States [2]. Organized electric markets are a critical component of the North American Bulk Electric System (BES) and its function to reliably and economically deliver power to consumers. Often touted as the "world's largest machine" [4] by electrical industry staff, the BES has become important to the lives of virtually every resident of the United States, Canada, and a portion of Mexico.

Given the high volume and value of transactions across multiple energy products, there are abundant incentives for market participants to manipulate these markets. Market monitors are the group of professionals that work to ensure that the electricity markets utilizing the BES remain fair, efficient, and open-access. They ensure this by identifying and referring behavior that violates established marketplace rules to the Federal Energy Regulatory Commission (FERC). Wholesale markets clear billions of dollars (USD) in revenue each year. Figure 1.1 displays overall energy consumption and cost for one of these markets, Southwest Power Pool's (SPP) marketplace, from 2020 to 2023.[1]

To illustrate the opportunity that superimposing information quality research onto the market monitoring domain presents, this chapter shall introduce the concepts of the BES, electric markets, and market monitoring, and provide detail as to how applying information quality concepts to market monitoring will help address a previously unsolved problem within the industry.

[1] Across the time frame shown in Figure 1.1, the observed market generated an average revenue of over $9.5 billion, serving an average energy load of over 261 GWh. This figure was generated using data combined from multiple public reports [5], [6], [7], [8].

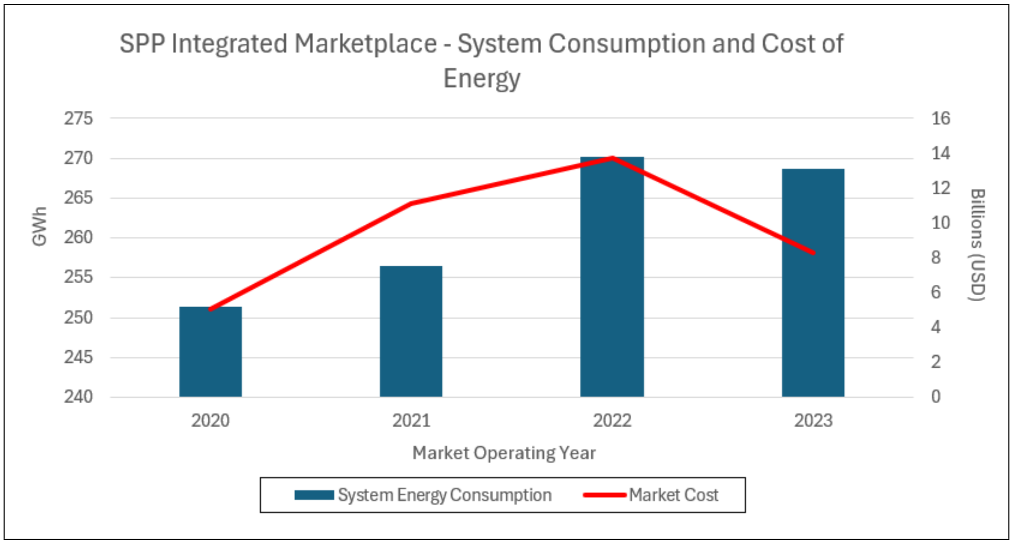


Figure 1.1: Four Year Overall Cost of Energy, SPP Integrated Marketplace

## 1.1 The North American Bulk Electric System

The BES[2] is a highly integrated system that facilitates the transmission of electric power (generalized as "energy") from generation to energy consumption (generalized as a "load") in a real-time and balanced[3] manner. Built with more than 500,000 miles [9] of physical transmission cabling and over 70,000 electrical substations [10], the BES constantly serves energy to North American consumers and industries alike. With few (relatively recent) exceptions[4] , energy is not stored in the system; it is generated (in response to a demand forecast) and then immediately transmitted and consumed by various loads. Figure 1.2 shows a conceptual configuration of some of the components[5] involved in the operation of the BES. [1]

[2]The North American Electric Reliability Corporation (NERC) serves as the Electric Reliability Operator (ERO) for the BES. In their Glossary of Terms (see References section), the BES refers to all high-voltage transmission lines and substations, generation resources, and all other equipment used in both inter/intra-state transmission of electric capacity.

[3]The term "balancing" refers to the need to only generate power as it is needed for consumption, subject to system capacity and economic constraints.

[4]Storage resources are a developing technology that can store energy for deployment during reliability emergencies or intervals of energy scarcity.

[5]An important distinction to note is between the *utility* and *retail* sectors shown in Figure 1.2. While referred to by several names (i.e. wholesale), utility generation and transmission refers to bulk energy (on the scale of MW) that is transmitted on the BES. The retail sector involves distribution of energy to residences and low consumption (on the scale of kW) commercial enterprises. This research focuses on wholesale energy markets, and thus is concerned with the bulk power generation and transmission domains.

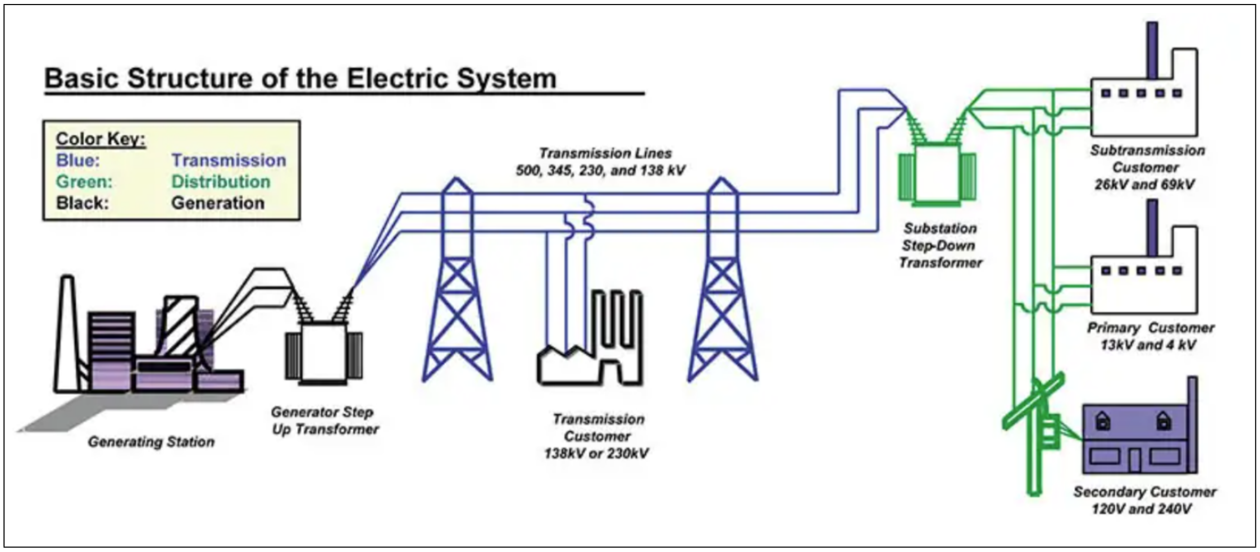


Figure 1.2: A Conceptual Illustration of the Bulk Electric System (source: [1])

The BES must be precisely managed, at all times, to balance power across the system and to exactly serve the load. Deviations can create instances of voltage and frequency instability that have the potential to significantly damage BES infrastructure. Oftentimes, these deviations are caused by larger environmental events, such as periods of extreme hot or cold weather. Demand for electricity increases during these events[6], as do generator outages and fuel delivery problems [12]. No one entity manages the entire BES. Instead, this job is shared across different types of regional entities covering different geographic locales. This dissertation focuses, specifically, on organized markets, which pool resources managed by a Regional Transmission Organization (RTO) or an Independent System Operator (ISO). These entities serve multiple roles; their primary role is to function as grid operators.

These organizations frequently act as Reliability Coordinators (RC) and Balancing Authorities (BA), monitoring power flows, balancing the grid, and acting to restore system reliability during and after emergencies. They also operate energy markets. Using complex optimization, the RTO/ISO dispatches individual resources in economic order[7] while respecting systems con-

[6]Notable weather events that have materially impacted BES reliability include the 2020 Californian heat wave, as well as severe winter weather storms in 2021 and 2022 [11].

[7]*Economic order* refers to the stacking of generator energy offers, in order, from least expensive to most expensive.

straints, a concept termed security-constrained economic dispatch (SCED)[8]. It also serves as a revenue-neutral clearing house for the bulk of energy and transmission-related transactions (as well as various financial products).

Serving these various functions to provide electricity generates enormous amounts of data on an hourly basis. This data is widely diverse in terms of sourcing, formatting, usage, sensitivity, and volume. It is the data assets generated by the operation of the BES that serve as the focus of this dissertation, specifically the data generated by *electricity markets*.

## 1.2 An Introduction to Electricity Markets

As stated above, this research focuses on organized markets. A large portion of the BES includes centrally organized electricity markets. Electricity markets are of great importance to all stakeholders involved in the BES; they allow energy to be economically generated and purchased at the wholesale level.

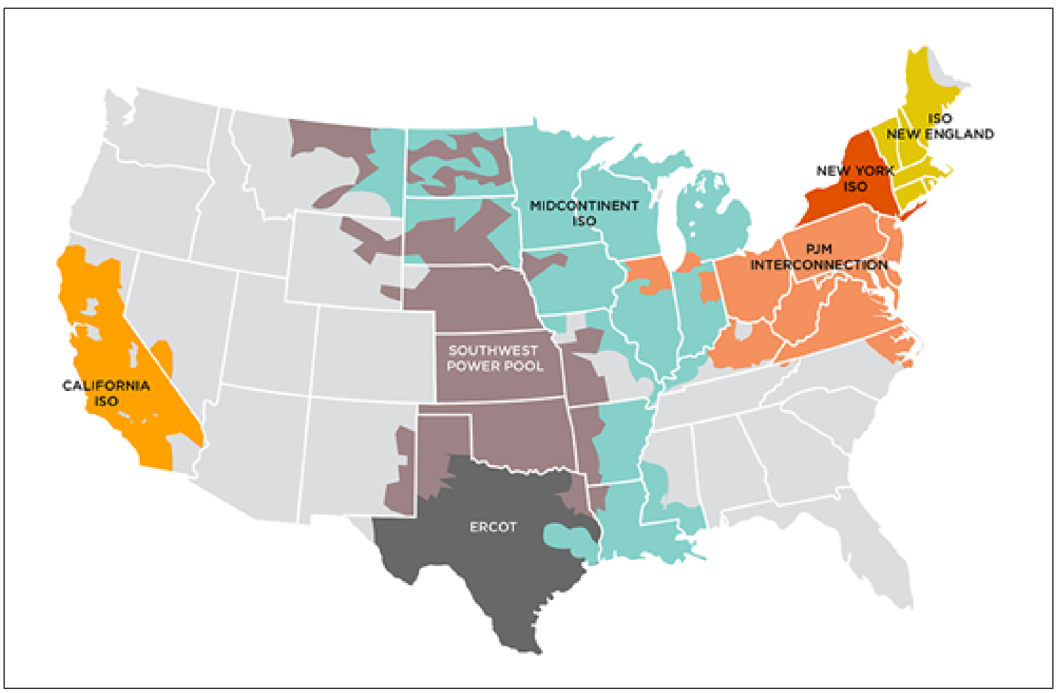


Figure 1.3: Prevailing Energy Market Territories (source: [2])

[8]Economic dispatch ("ED") is a concept that blends economics and the physical reality of the electric grid.

Figure 1.3 displays, at a high level, how the US energy grid is divided into different territories and notes the name of the entity that runs the organized market. These territories are typically governed by FERC, with the exception of Texas[9].

### 1.2.1 Entities

Electricity markets are operated by a small number of operational entities[10] within North America. These entities mostly fall under the jurisdiction of FERC (see Figure 1.3). The prominent entities that are impacted by market monitoring include Independent System Operators (ISO) and Regional Transmission Organizations (RTO). These entities are filed as non-profit businesses and typically perform similar functions. FERC Orders 888 and 889 initiated the formalized incorporation of an ISO, while FERC Order 2000 set forth the criteria that differentiate an ISO from an RTO.

#### *1.2.1.1 Independent System Operator (ISO)*

ISOs are organizations that manage power generation and transmission across the bulk electric system. They operate in an advisory capacity for generation owners (public and investor owned utilities, as well as independent power producers) and transmission owners and operators to effectively generate and trade power in an energy market. A common, helpful analogy that describes the function of an ISO is an air-traffic controller.

Air-traffic controllers merely monitor and route airplane traffic across the airspace over the United States. They own neither physical airplanes nor airlines, and function only as advisors to ensure safe passage of both private and commercial airplanes. ISOs work in a similar fashion. The ISO over the California energy market territory (CAISO) owns none of the transmission infrastructure across which electric power is shipped. However, CAISO manages the flow of energy

[9]ERCOT, the electric system operator for Texas, is not within the jurisdiction of FERC. The Public Utility Commission of Texas, instead, has oversight of the ERCOT market territory. Alaska and Hawaii, additionally, maintain isolated electricity infrastructures.

[10]The entities referred to in this chapter are not limited to the 2 categories listed. They also include Transmission Owners and Operators (TOs / TOPs), Regional Entities (REs), Balancing Authorities (BAs), Reliability Coordinators (RCs), and Federal Power Agencies. Oftentimes, an RTO or ISO will also support one or more of these functions (SPP, for example, hosts operational desks for performing reliability coordination).

across this system, and sends regular dispatch instructions to the generators that inject power into the BES.

#### *1.2.1.2 Regional Transmission Organization (RTO)*

RTOs are another entity that manage organized markets. They are categorized by a group of four core characteristics (and an additional seven broad operational functions) that separate them from independent system operators. RTOs and ISOs share many similarities. Their operations are regulated by the same government entities, namely FERC and NERC, and they both have the same, conceptual goal: to reliably generate and deliver energy at competitive rates.

Both entities influence and deploy energy products and maintain both the configuration of their operating territory as well as the flow of energy across the BES. They also both support development and maintenance of open access transmission tariffs (OATT) that allow for unprejudicial access to energy under US federal law. And, most relevant to this research, they are both required to employ a market monitoring group to screen their organized markets for cases of market manipulation.

Their differences primarily lie in how their corporate structures are filed. Both types of entities must be registered non-profit organizations. However, ISOs have more flexibility where impartiality is concerned. RTOs work to maintain independence from their market participants and strictly focus on system reliability and economic efficiency. While ISOs are also concerned with these goals, they have the ability to lobby government on behalf of their members.

### 1.2.2 Types of Markets

Electricity markets are designed to serve one of several purposes. ISOs typically deploy and operate these different market designs, in concert, as a *marketplace* [13]. Southwest Power Pool's (SPP) Integrated Marketplace is such an example of this. To understand the scope of market monitoring activities, it is important to understand the main energy market designs that exist today. Arguably, there are five main types of electric energy market designs; each design may be deployed with possible variations on its theme (a marketplace that does not operate the function of a capacity market is one such example). Reference Table 1.1 for information on prevailing design types.

| Market Type | Purpose |
|---|---|
| Capacity | Allows a load to purchase an agreement of firm capacity for service in the future, usually on the scale of months to years in the future. |
| Day-Ahead (DA) | Allows buyers and sellers to secure positions a day before the market sells energy to be consumed (operating day). This aids in realistic price formation. |
| Real-Time (RT) | Allows for adjustments between the DA market (what was expected on the day prior) and the system conditions that happen on the operating day. |
| Ancillary Services (AS) | Allows other providers to sell products that support grid reliability. For example, the ability of a generator to quickly "ramp up" to a certain production requirement is such a product sold in an AS market. |
| Transmission Congestion Rights (TCR) | Allows participants to hedge against congestion on the transmission system (adequate power can be produced in an area, but there may be difficulties in transmitting it to where it is consumed). |

Table 1.1: Prevailing Electricity Market Designs

### 1.2.3 BES Complexity and Energy Markets

The BES is a physical system upon which we superimpose a conceptual, economic market. This dichotomy is part of what makes the energy market so complex to both understand and govern with the data that it generates. Physical machines require continuous diagnostics to ensure that they are operating 1) effectively and 2) within adequate maintenance tolerances. The BES is no different.

Thermal generators that produce electricity are monitored for metrics including: wattage output, fuel consumption, and greenhouse gas emissions. High-voltage transmission lines are monitored for temperature (heat is generated via electricity loss across conductors) and short-circuit events. Substations are monitored for transformer faults, tap configurations, and other mechanical system measurements. Each of these values is typically transmitted in a time-series format, resulting in constant streams of monitoring data. These time series data streams are typically parsed into *market intervals* that constitute "units of work" for market activity. Much of the analysis performed on wholesale electricity markets utilizes data secured from market intervals. Depending on the type of market that is under observation, an interval may range from a span of several seconds (dispatch instructions) to "between five and fifteen minutes" [2] (for calculation of real-time market prices).

These diagnostic data are only a fraction of the data that are used in the operation of the BES, and much of it feeds into market systems[11] to calculate a realistic price of power at different locations throughout a market territory. Once the market solves, and locational prices are calculated, the market then must be *settled*[12]. Market settlements are another area of market operations that generates complex and voluminous data. In SPP's Integrated Marketplace, the Marketplace Protocols set forth an extensive settlement plan that determines how charges and credits are calculated and assigned to market participants. These settlement instructions also cover events where the market must be "repriced", or re-settled, due to external factors that occur during the operation of a market, such as equipment outages, data communication errors[13], or extreme weather events.

### 1.3 What is Market Monitoring?

Market monitoring is a discipline that sits at the intersection of two major knowledge domains: *forensics* and *economics*. As such, it takes a scientific approach to testing and analyzing transactions that occur within a marketplace. According to a report published by the United States Energy Association (USEA), Market Monitoring is a necessary component of markets "to ensure that market(s) participants cannot exercise market power, collude or engage in any other behavior that could give them a market share, or higher profits" [3]. This need places the authority to

[11] While market systems are composed of many different products, they are generally referred to as a *market clearing engine*.

[12] Settlements are also facilitated by shadow accounting, a process of maintaining a separate system (in addition to a production accounting system) to track transactions and ensure that accounts are appropriately balanced.

[13] Data elements used to monitor and dispatch generation in the BES are collected and shared in the form of *SCADA* (Supervisory Control and Data Acquisition) [14].

request and analyze records generated by the operation of electricity markets into the hands of market monitoring units (MMUs).

Market monitors will generally screen for MP behavior that matches the following criteria:

- Exercised (or attempted to exercise) positions of market power
- Market gaming through oversights in current market protocols or policy (usually outlined in a governing document, known as a tariff
- Committed cases of market manipulation through acts of collusion, cross-product manipulation, malevolent influence on price formation, or untoward manipulation schemes (economic withholding, uneconomic production, wash trades, "pump and dump" schemes, etc.)

### 1.3.1 The Fall of Enron Corporation

The Enron Corporation, a former American-based commodities and energy broker, is a well-known case study in American corporate governance (or, rather, lack thereof). The fall of Enron (first signaled in late 2000) uncovered deep, systemic problems of fraud and deceptive business practices that directly influenced both the market monitoring and information quality (IQ) disciplines. Outwardly, Enron appeared to be a stable and innovative company. At the turn of the century, Enron's books reflected an ownership of $60 billion in assets. This included an internet-based energy trading desk ("Enron Online") that cleared $2.5 billion in *daily* energy transactions [15]. In reality, Enron management operated under a culture of willful non-compliance, including (but not limited to):

- Energy withholding practices that sometimes caused enormous electricity prices in the California wholesale energy market
- Artificial energy shortages causing regular blackouts[14] within the California service territory

[14] A blackout (an event where power cannot be served) is a breakdown of a portion of grid infrastructure, caused by "an imbalance between power generation and power consumption" [16].

- Scheduling energy transactions to purely cause congestion on California's electric transmission system
- Accounting fraud (enabled by a creative use of "mark-to-market" based accounting) which allowed Enron to report expected revenues before they were realized

These manipulative schemes had widespread impacts, including: 1) the bankruptcy of one of the largest utility companies in California, Pacific Gas and Electric (PG&E), 2) rolling blackouts, 3) increased transmission system congestion, and 4) higher than normal electricity bills for California ratepayers.

Enron's activities constituted both fraud and a behavior known as *market manipulation*; a form of conduct that attempts to willfully circumvent established market rules in pursuit of (usually) large profits. Their usage of market manipulation, especially in California, was a signal to federal regulators that, at the time, there was not enough oversight in the operation of electricity markets. As a result, the United States Congress passed the Energy Policy Act of 2005 to introduce more stringent regulation over the energy industry.

Enron's downfall resulted in a near-complete devaluation of their stock[15], the dismantling of Arthur Andersen (the auditor assigned to ensure that Enron was operating within the confines of financial reporting law), and a slew of legislation[16] serving as the impetus for energy market monitoring in the United States. Additionally, Enron's scandal was a significant contributor to the passing of the Sarbanes-Oxley (SOX) Act of 2002, imposing strict and detailed requirements on data that is used as input for financial reporting. As a result, SOX became another driving force in the implementation of information quality as both an academic discipline and as a business function.

In the years following the disbanding of Enron Corporation, FERC has pursued many more instances of deceptive behavior in organized energy markets. Through the employ of market mon-

[15] As Enron entered into legal proceedings, their stock fell to under $1 per share (from a peak as high as $90 per share in mid-2000) [17].

[16] 18 CFR § 1c.2 (Prohibition of Electric Energy Market Manipulation) is an example of legislation that places monitoring authority in the hands of MMUs [18].

itors and FERC enforcement staff, manipulation cases can be pursued and penalized before they become even more critical threats. Table 1.2 lists several high profile cases of market manipulation that are distinct from the Enron case [19] [20] [21].

| **MP** | **Damages** | **Incident** |
|---|---|---|
| JP Morgan (2013) | $410 mm | Purchased and operated antiquated generation plants and offered them in the market as competitive generators |
| GreenHat Energy, LLC (2021) | >$229 mm | Participated in insider trading to purchase enormous positions in PJM's congestion market |
| Vitol Group (2024) | $2.3mm | Sold physical power at a loss in the DA market to favor their transmission congestion positions |

Table 1.2: Prominent Cases of Market Manipulation

### 1.3.2 Skills of a Market Monitor

Market monitors are one of several groups that function as "unsung heroes" in electricity markets (and within the larger context of the North American bulk electric system). The work that they perform on a routine basis helps to ensure a reliable and economically sound marketplace for buying and selling energy and energy-adjacent ancillary services. Market monitoring units are typically staffed with analysts from a wide variety of both technical and non-technical backgrounds. These include: data analysts, statisticians, economists, accountants, technical writers, engineers, and other corporate staff dedicated to investigating events that occur in their jurisdictional markets.

Figure 1.4 illustrates a cross-section of several of the skill sets that market monitoring analysts might possess, based on the researcher's experience of working as a market monitor within Southwest Power Pool's (SPP) MMU. As one may imagine, it can be difficult to find an individual contributor that possesses an all-encompassing understanding of each of these knowledge areas.

Market monitors have a wide breadth of responsibilities that must be covered by their individual skill sets. Routine market surveillance is a continuous process that must evolve with the

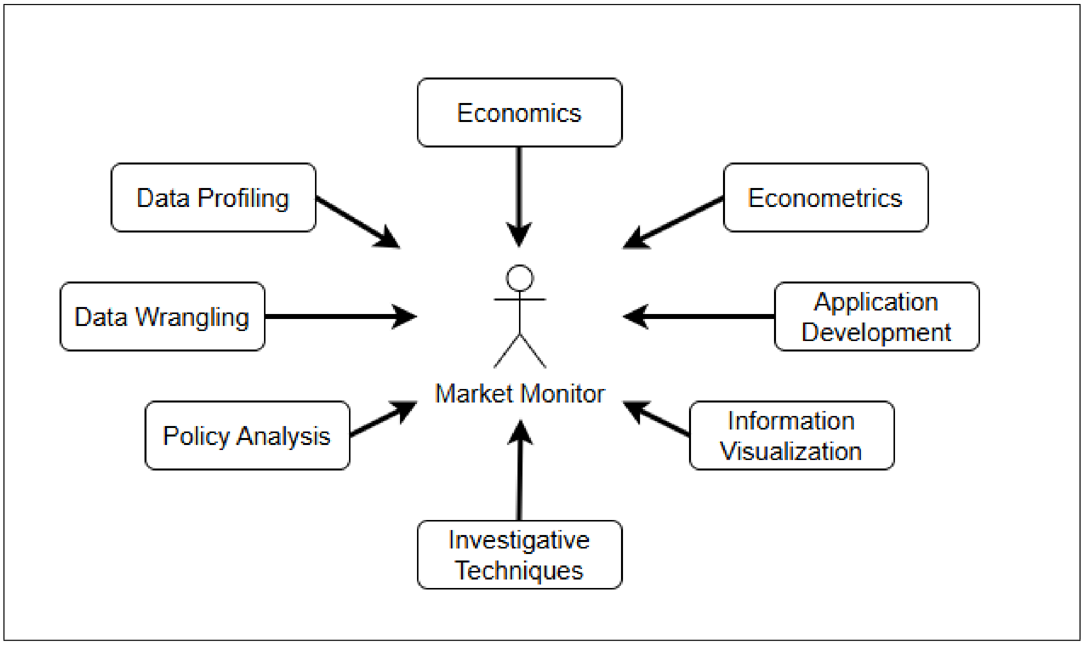


Figure 1.4: Diverse Skill Sets of a Market Monitor

behaviors and trading patterns of market participants. As such, market monitors observe and investigate market bids and offers made each operating day, analyze generator behavior to identify erratic response to dispatch commands, ensure that participants are not covertly colluding or engaging in price fixing[17], review out-of-market transactions that take place through bilateral settlement schedules (BSS)[18] , and ensure that RTOs and ISOs are acting within the rules defined by their FERC-approved tariff[19] documents (excluding ERCOT, which does not fall under FERC jurisdiction).

#### *1.3.2.1 Market Monitoring Firms*

The current groups involved with electricity market monitoring (both in North America and internationally[20]) include the organizations compiled in Appendix E. This list is not intended to be

[17]Energy market prices are typically set on a marginal, spatial basis (representing the cost of injecting another MW of energy, at a location, in a market territory) [22].

[18]Bilateral settlement schedules are agreements between market participants that are used to trade energy outside of normal market operations [23], which help manage credit exposure in the marketplace.

[19]A tariff is a governing document that outlines the rules for operating (and participating in) an energy market.

[20]The US Energy Association (USEA) published a report outlining international, independent market monitoring groups. Many of the international groups on this list come from that report (see Appendix E).

exhaustive, as other nations in the Asian-Pacific (APAC), European, and Middle-Eastern (EMEA) regions of the world may not be reflected here. This roster of organizations with market monitoring authority illustrates the number of individuals that are involved with market analysis on an international scale.

### 1.4 The Case for Information Quality Research in Market Monitoring

As stated above, market monitoring encompasses a wide range of activities and demands a mix of skill sets. These skill sets must be applied with knowledge learned, over time, from others in the industry. Such activities include gathering transactional market data for screening, calculating metrics for market power[21] and concentration analysis (for example, the classic Herfindahl-Hirschman Index [HHI] or the Residual Supply Index [RSI]), mitigating abuses of local market power by market participants, and performing policy analysis to determine efficient market design [25] [26]. It also requires calculating known system metrics, Capacity Reserve Margin (CRM) for example[22], and developing new metrics to measure and track market performance. Data collection, processing, and analysis enables each of these activities, but knowledge on how to make decisions on these data is typically learned through on-the-job training.

In *Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality*, one of the antecedents that can directly affect data quality is "information overload" [28]. This can be especially true in market monitoring, where users must derive a decision by "fusing" data from many different sources. In instances where users are not familiar with the sources, this has the potential to impact both accuracy and trust in the final product.

As an example in an adjacent industry (utilities), Oyoo's work in data validation for power and utilities organizations showed that information quality can contribute to the development and maintenance of Extract-Transform-Load (ETL) processes [29]. In this example, electric utilities may prioritize data templating and formatting over accuracy and trust. Market monitoring cannot afford

[21] Market power is often defined as "the ability of a seller to profitably maintain prices above competitive levels for a significant period of time" [24].

[22] CRM is a metric (computed by taking the difference between total available capacity and peak system demand) that helps identify "how reliable and prepared the [BES] system has been", in normal operations [27].

to make this trade-off. Market monitors work in complex and high-volume data environments. Systems deployed in these environments must be accessible and user-friendly to help market monitors handle the inherent data and information complexity.

### 1.5 Conducting Information Quality Research in Market Monitoring

The case study of Enron Corporation sets both a unique and compelling stage for the role that information quality and information science-based research can play in improving the business impact of an independent market monitor. Included amidst Enron Corporation's numerous failures of governance are a lack of access and an understanding of the complex data generated by the energy industry. The operation of such a large system and marketplace can lead to many siloed data stores, undocumented schemas, and a slew of other information quality-based problems that can also be seen in market monitoring.

Furthermore, the cross-section of skills that market monitors must possess (to effectively perform their job functions) also underscores the importance that data literacy and data quality improvement hold in the market monitoring domain. Both data and system architecture plans, proposed as part of the deliverables of this dissertation, are paramount to designing a system to aid market monitoring staff in their work.

#### 1.5.1 Subject Matter Expert (SME) Discussions

The potential use of this dissertation is driven, in part, by the opinions of subject matter experts in the field of electric energy. In support of this investigation, the researcher has consulted multiple members (in different sectors) of the industry, including two (2) professionals in enterprise asset management, and a separate discussion with a staff member in energy regulation. One such staff member, Kennedy Oyoo, performed information systems research in a different sector of the energy grid, "power and utilities", focusing on the distribution of electricity to end users.

Energy utilities struggle with a similar issue to those faced by market monitors— siloed data exists in different systems, and can exist with vastly different data quality scores (in terms of completeness, accuracy, etc.) Oyoo's research found that a key step, data validation, was missing from the ETL processes of an energy utility. Part of the research completed in his project resulted in an information template-based prototype system that cleansed and validated utility data according to automated assertion rules [29]. Such a methodology has potential impacts for the market monitoring discipline, and may have an influence on the resulting system architecture of this dissertation.

### 1.6 Design Science Methodology

The roadmap of this project is based on the Design Science Research Methodology (DSRM), a research process for developing new product artifacts. The researcher chose DSRM for this project because the main research objective is to design and develop a domain-specific language for market monitors to use in their analysis.

Data and information quality problems exist today in the market monitoring body of knowledge that a DSL-based system can help address:

1. Analysts in different teams that arrive at different values when trying to quantify metrics (market impact, for example)
2. Lack of robust metadata assets such as data dictionaries and well-defined data stewardship roles
3. Lack of a mature change management program for software (which, in turn, can impact report reproducibility)
4. A lack of full automation and meta-automation, causing analysts to spend time performing tedious ETL and archival tasks

Domain-specific languages are powerful models for allowing users to systematically solve problems within a specific domain. As is discussed in the literature review (see Chapter 2), DSLs have only had limited exposure in electricity markets (and even less exposure in market monitoring). This discovered gap in existing research, complemented by the DQ and IQ issues discussed in the review, serves as the inspiration for this project. Figure 1.5 provides more detail on the planned use of design science within this project[23].

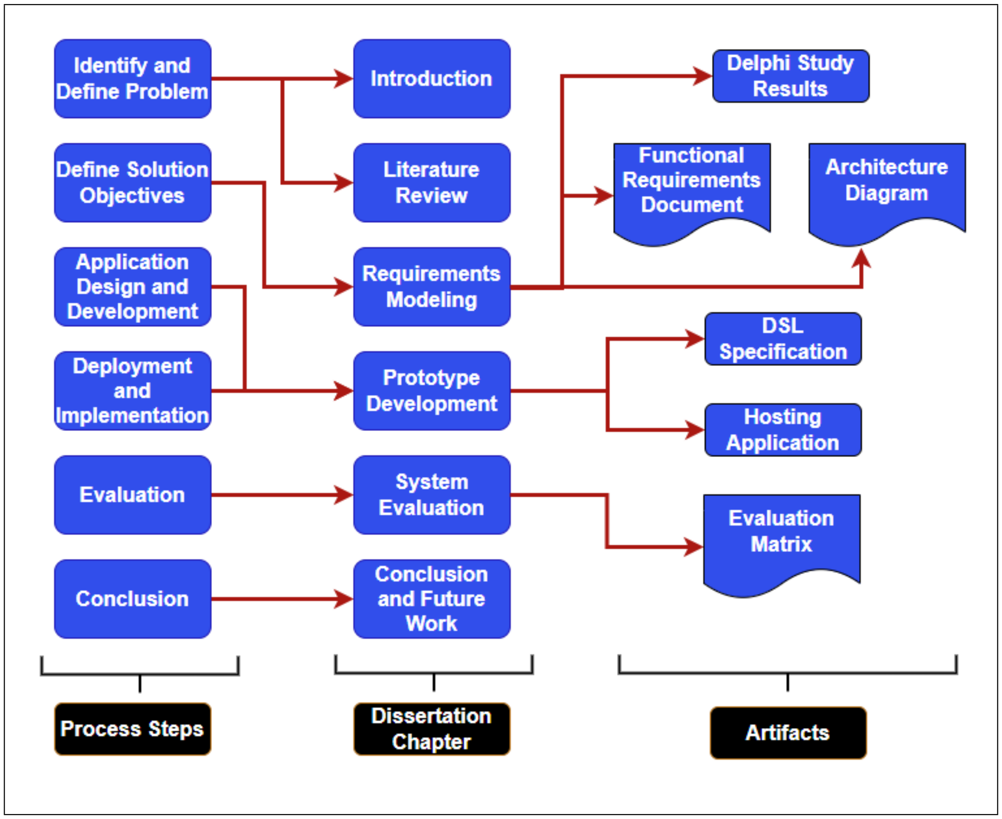


Figure 1.5: Design Science Research Methodology for this Dissertation

### 1.6.1 Proposed Investigation

In *What Design Science Is Not (2008)*, Baskerville states that one view of design science is to be *generative*. That is, design science leads to the generation of artifacts that then influence the formulation of a theory [30]. This prospect of generativity seems to indicate that a design can be updated based upon empirical evidence. Due to this idea, the researcher chose to employ a Delphi study (as the methodology for gathering expert sentiment) to influence the design of the DSL prototype discussed in this dissertation. The requirements modeling chapter shall discuss the setup, conduct, and results of a Delphi study employed in market monitoring. It then explains how

[23] See Appendix C for a schedule of the milestones and expected completion timeline for this project.

the results of this study influenced the design and requirements document for the prototype system.

Overall, the investigation methodology for this dissertation is composed of five (5) main components. These components include 1) a literature review, 2) a Delphi study to gather opinion and synthesize system requirements, 3) a system architecture design, 4) development of the prototype system (and associated tooling and documentation), and 5) an evaluation discussion to determine to what degree the resultant software satisfies the system requirements.

### 1.6.2 Limitations of this Research

The scope of work proposed in this dissertation is limited to the objectives outlined in the Proposed Investigation section. Specifically, it shall ultimately evaluate how well the developed prototype functions as a *minimal viable product* (MVP) for energy market analysis.

The proof-of-concept prototype discussed in this dissertation **is not intended to be a production-ready system**. It is also not expected to be used in a live scenario for market monitors to immediately put into service in support of decision making. The data ingested into the prototype will be based on realistic data scraped from public data systems shown in Appendix D. The evaluation of this system will be based on a matrix (mapping back to the functional requirements document) that "scores" how well the functionality of the prototype supports the intended requirements.

Further system optimization, bug fixes, scalability concerns, and use in live market monitoring scenarios are all currently considered "out of scope" for this dissertation and are reserved for future work. The author, however, may pursue future enhancements to this system as separate research endeavors.

# 2. LITERATURE REVIEW

A review of the literature surrounding information quality research within market monitoring units signals that there is room for improvement in how value is generated from information assets. This gap is highlighted by several major issues (introduced in Chapter 1 and discussed in this literature review), including:

1. Development of technology surrounding domain-specific language engineering

2. An aging workforce in the electric utilities industry that is reaching the tipping point for replacement

3. A lack of existing research on data and information quality within the market monitoring sphere

4. A disruption of traditional energy market design due to advances in technology and the energy regulatory landscape

These issues represent a unique opportunity to obtain important value from the data assets both generated and used by professionals in the market monitoring sphere. The researcher will discuss each of these issues in the context of relevant literature[1], and then introduce possible sources of data that enable the development of a domain-specific language for market monitors.

## 2.1 An Aging Workforce

Many technical industries are experiencing the problem of an aging workforce; trained professionals are reaching retirement age at a rate faster than that of replacement. A 2006 study showed that the largest staff segment, of a surveyed group of energy professionals, is at (or beyond) retirement age. This same study cited "retirements, restructuring, and technology changes" as being the

[1] Appendix B contains a listing of different search phrases used for locating relevant research discussed in this chapter.

primary drivers of staff leaving the energy profession [31]. Given the age of this study, many industries are currently going through a period of training new career professionals. Such a problem impacts market monitoring as well.

MMUs (as previously discussed) were cemented into operation by the US federal government, not long before that same 2006 study. An energy research lab at UC Berkeley shows that staffing across market monitoring groups is relatively low compared to other organizations in the industry. According to Goldman, Lesieutre, and Bartholomew, all operational market monitoring groups in the US had fewer than 20 FTEs (each) in employ by 2004 [25].

While these numbers appear low in comparison to other organizations with energy industry staff[2], this study declares that budgets and headcount (at that time) increased significantly for market monitoring work. As such, this shows a trend of growing demand for market monitoring staff in a workforce that, holistically, employs many individuals on (or beyond) the horizon of retirement. One claim to helping combat this "brain drain" problem is a technology framework that "facilitates knowledge retention" [31], including systems (supported by an IT organization) that capture necessary process knowledge.

### 2.1.1 The Importance of Institutional Knowledge

Given that the energy industry faces career professionals heading into retirement, this research intends to use a domain-specific language-based system to help preserve the institutional knowledge that is lost when market monitors retire (or even simply change positions). As employees make career transitions, the institutional knowledge that they built often leaves with them. In market monitoring, this problem is amplified due to two salient facts: 1) the number of individuals that work in market monitoring is small compared to the rest of the energy and utilities industries, and 2) market monitors often specialize in regards to analysis or policy work. Individuals who primarily work in market surveillance may never directly work in policy development (or vice-versa).

In a personal anecdote, this "brain drain" problem has been seen within the SPP MMU. Several long-term employees have exited the MMU in pursuit of other roles, only to leave undocumented

[2] As of Q1 2025, Southwest Power Pool employs over 800 FTEs (including their internal market monitoring staff).

code behind them that others must support. This code exists in the form of SQL queries and/or SAS scripts that contain embedded business rules (as they were understood by these employees). Deciphering and supporting this code becomes a burden for others that inherit the work of these departed employees.

A 2016 publication *Knowledge Management in Esoteric Management* summarizes this issue well: "...knowledge sharing should be encouraged by placing employees working together closely to create pooled expertise..." [32]. When employees share knowledge, they can keep important information from being destroyed when key individuals leave a department. One of the goals for the system that this dissertation designs is to foster knowledge sharing amongst employees.

### 2.1.2 Application to Market Monitoring Research

Domain-specific programming languages (DSLs) are powerful interfaces that enable subject-matter experts (SME), in any particular domain of expertise, to perform tasks similar to a proper software developer that builds applications in general-purpose languages (GPL). A well-defined DSL can save many hours of development time for specific groups that have IT oriented needs. DSLs also have the ability to lower the "barriers to entry" for SMEs that wish to write applications to handle work within their purview. With support from a DSL, a software engineer does not need to be deeply involved in business rules to understand how to develop an application (a common issue in the energy industry). Likewise, SMEs do not need to heavily codify their business rules into software requirements for an application to be developed when utilizing a domain-specific language.

Following the theme of how DSLs are able to abstract away complexity of big data processing tasks, this research aims to bring such a tool set to bear on tasks performed in energy market monitoring. Commands embedded into a domain-specific language have the potential to speed up many repetitive tasks in the market monitoring domain, especially regarding the generation of report and stakeholder presentations.

## 2.2 An Analysis of Domain-Specific Languages

Consider programming languages on a spectrum from general purpose languages (GPLs) at one end to domain-specific languages (DSLs) on the opposite end [33].[3] DSLs are languages that are limited in purpose, scope, and grammar. Rather than supporting a programming language that can be utilized for constructing almost anything (as GPLs do), DSLs are created to solve issues within a single problem domain.

| **Language** | **Supporting Enterprise** | **Problem Domain** | **Paradigm** |
|---|---|---|---|
| Structured Query Language (SQL) | Various | Data Retrieval | Declarative |
| Terraform / HCL | Hashicorp | Cloud Infrastructure Deployment | Declarative |
| Apache Pig ("Pig Latin") | Yahoo / Apache Software Foundation | Data Retrieval & Processing | Declarative |
| awk | Bell Labs (originally) | Text Processing | Procedural |

Table 2.1: Selection of Well-Known Domain-Specific Languages

*Structured Query Language (SQL)* [34] is one such example of a DSL. SQL queries are designed to retrieve data from databases for extraction, reference, and analysis. SQL dialects have been designed to interface across a broad range of database management systems (for example: PL/SQL for Oracle or T-SQL for Microsoft SQL Server databases).

*Terraform (HCL)* [35], a product of Hashicorp Corporation, is yet another example of a domain-specific language that is prevalent among cloud-based companies. Terraform is known as an "infrastructure-as-code" language, used to interact with numerous cloud platform providers (e.g. Amazon Web Services) to quickly deploy, scale, and spin-down cloud resources. Instead of requiring an engineer to use a console environment to configure cloud resources, that same engineer can write a set of scripts that will precisely configure and deploy all of their desired infrastructure at

[3] Kosar, Tomaz, et al. indicate the "schism" between the two languages by discussing that "GPLs tend to be general, resulting in poor support for domain-specific notation [33]."

once.

*Apache Pig* [36] (and its associated programming language *"Pig Latin"*) is also a DSL. Like SQL, it is a declarative language used for data retrieval and data processing. Developers at Yahoo created Pig as a way to abstract some of the complexity of SQL away from analysts who were working with large datasets. With fewer lines of code, an analyst could use Pig to parse and load a dataset into memory, perform processing tasks, and format and report the results. This was specifically useful to analysts processing large datasets because Pig was designed to take advantage of the high-performance computing facilities of Hadoop.

Yet another powerful example of DSLs can be seen in the ubiquitous Unix program: *awk* [37]. Originally created by Unix developers Aho, Weinberger and Kerninghan[4], awk is a scripting language designed to process textual data— one line at a time. The power of this systems allows individuals to write expressions that are simple to understand, efficient, and portable across different Linux implementations. Countless developers (including the author of this dissertation) have made use of awk's utilities to process and format data in a headless environment.

Each of the languages represented in this review are, on their own, infeasible for writing general purpose applications. SQL, for example, should not be used to write a desktop program. They were, instead, each created to solve specific issues within their respective problem domains. Likewise, Terraform cannot be used to create web applications or develop billing software. One of the "upsides" of these languages are that they can quickly handle tasks within a specific problem domain.

### 2.2.1 DSL Variations

Each of the language examples provided here share a common characteristic in that they are external DSLs. They are interpreted and executed by a runtime program. This trait of DSLs separates them from another subset of domain-specific languages: internal DSLs. These are embedded langauges (sometimes referred to as fluent interfaces) that rely on a (usually) general-purpose language to parse their individual syntactic structures and execute logic. These DSL features are

[4]The name "awk" is an acronym formed from the first letters of the developer's last names.

sometimes referred to as *syntactic sugar* [38], implying that they make the experience of developing solutions in them "sweeter" than without them.

Another common characteristic of DSLs lies in the fact that they are often text-based programming environments. As discussed later in this literature review, graphical domain-specific languages also exist. These are usually seen in the form of either enterprise extract-transform-load (ETL) applications or language workbenches. Jetbrains MPS is a prominent example of the power of graphical DSL representations. Business users are usually positioned to take advantage of these product offerings because they pose abstract interfaces (environments that allow users to graphically model *what they want* instead of having to explicitly write source code in a general purpose language). With basic instruction, users can string together components that construct repeatable workflows without requiring them to possess deep programming knowledge.

## 2.3 Extant Domain-Specific Language Use Cases in the Energy Sector

The literature uncovered in this review supports the claim that DSLs can achieve high efficacy in financial contexts. Krasts, Kleins, and Teilans (2012) illustrates how domain-specific languages have been used in modeling the complex dynamics of financial settlements systems. Financial settlements often encounter system error cases during the settlement of a market. Using a DSL for modeling tasks, in these scenarios, appears to help developers and business users clearly visualize complex systems logic [39].

The literature, even more interestingly, shows that DSLs have been deployed in other areas of the energy sector, as identified by Sobernig, Strembeck, and Beck [40] in *Developing a Domain Specific Language for Scheduling in the European Energy Sector*. While this paper sets precedent for the use of DSLs in the energy industry, it presents an even larger gap that this dissertation addresses. Beck's study explores employing a DSL to assist in tagging and scheduling energy interchange (which is an operational concern that only encompasses a single component of market monitoring). Additionally, and far more importantly, the implementation of their DSL was constrained to a specific market interconnection in Europe, which is significantly less complex than the transmission and congestion[5] issues that are seen throughout the North American Bulk Electric System. Market monitoring, while interested in aspects of transmission scheduling, encompasses a much broader scope of analytics in the energy sector.

## 2.4 The Data Domain in Market Monitoring

The data domain for market monitoring entities is a large ecosystem. It is common for monitoring analysts to synthesize information both from internal and external data sources. This type of ecosystem can quickly become a *data swamp*[6] without careful attention to business rules and the possession of institutional knowledge for how processes are executed on market information.

FERC Order 760, a key piece of legislation that directly relates to this dissertation, is one such example of how market data is sourced and utilized by analysts. It dictates several key categories of data that RTO/ISO organizations must regularly provide to the Commission. Such data includes [42]:

1. Supply offers and demand bids for energy and ancillary services
2. Virtual[7] offers and bids
3. Energy/ancillary service awards
4. Capacity market offers, designations, and prices
5. Resource (electricity generation) output
6. Marginal cost estimates

[5] Current issues in North America show large pockets of stranded load, which have significant economic (and, by extension, reliability) implications. Congestion on the BES constitutes a market (Financial Transmission Rights - FTR) that settles for billions of dollars. This is all due to the fact that monies have to be paid to rights holders for transmitting energy across a congested transmission infrastructure.

[6] Atlan, a leading data governance company, classifies a data swamp as a repository that possesses: poor quality, undocumented code and data schemas, and leads to situations where users must "...spend considerable time on tasks such as data discovery, cleaning, integration, and management" [41].

[7] A "virtual" transaction (as defined by Southwest Power Pool's glossary of terms) is a transaction for purchasing (bid) or selling (offer) energy "at a specified price, settlement location, and period of time ... that is not associated with a physical load" [43].

7. Day-ahead market (DA) shift factors[8]

8. Financial Transmission Right (FTR) data

9. Bilateral Contracts

10. Interchange transactions (otherwise known as "scheduling and tagging")

Understanding how to source and use this information impacts the accuracy of market monitoring analysis and enables FERC to accurately assess and penalize instances of market manipulation. Defined by the industry terms "ex-ante" and "ex-post", assessments of market performance, market power, and market manipulation can be carried out prior to an event, or "after the fact"- when a market event has already occurred [45].

### 2.4.1 Public Data Systems for Wholesale Energy Markets

In a similar manner to how market operators provide regular data to FERC, RTOs and ISOs also regularly publish data that allows the general public[9] a view into the operation and settlement of these markets. Table 2.2 lists the existing systems that this research can utilize for test data. Much of the data contained in these systems match the data requirements for FERC Order 760 (though some identifiers may be omitted or obfuscated to preserve confidentiality).

The researcher intends to source data assets that model market conditions (especially for price formation) from these systems, and use such data for demonstrating the abilities of the prototype system, including: 1) visualization generation, 2) report creation, and 3) shadow calculations. The data assets that may be necessary to produce the demonstration deliverables include public records such as generator names and parameters, generation outages, transmission line data, and component data used to calculate marginal prices (energy, loss, and congestion factors).

[8]A *shift factor* is a percentage based value that determines how much energy a transmission line will receive when a generator injects energy into the transmission grid [44].

[9]See Appendix D for a list of hyperlinks to view more information on these public sources of energy market data.

| **Publishing Entity** | **Data Repository Name** |
|---|---|
| PJM Corporation | Data Miner 2 |
| Monitoring Analytics | Member Information Reporting Application (MIRA) |
| Electric Reliability Council of Texas (ERCOT) | Market Data Transparency Service |
| California Independent System Operator (CAISO) | OASIS & Data Library |
| Midcontinent Independent System Operator (MISO) | MISO Data Exchange |
| Southwest Power Pool (SPP) | Public Data Portal |
| New York Independent System Operator (NYISO) | Operational Data Portal |
| Independent System Operator of New England (ISO-NE) | ISO Express |
| US Energy Information Administration (EIA) | Wholesale Energy Market Reports |

Table 2.2: Public Energy Market Data Systems

# 3. PROPOSAL METHODOLOGY PLAN

Design science shall be used as the primary method of investigation in this dissertation due to its ability to derive insights from the development of new artifacts. However, design science itself is only a process that guides the investigation. To effectuate this process, the researcher has developed the following plan to build and evaluate a software system (powered by a domain-specific programming language) for potential use in the professional domain of energy market monitoring.

This plan includes the following details (illustrated by Figure 3.1):

- A literature review to both introduce the research area and to establish relevance and novelty
- A Delphi-based study using a panel of experts (within market monitoring) to identify the current needs of these experts[1], including:
  - The current landscape of a market monitor's workflow
  - Perspectives from the entire market monitoring community, including market monitors that work outside of the United States
- A comprehensive data and system architecture plan for setting up and operating a DSL-enabled data analysis system
- Development of a prototype system to serve as a proof-of-concept for using a DSL in the context of market monitoring analysis
- Develop a *traceability matrix* [46] that can serve as a method of evaluation, for this system, to ensure that the prototype serves the intended functional requirements
- A discussion of the performance of the system (as this also impacts the system's usability).

[1]The results of this panel shall serve as the basis for constructing a comprehensive software requirements document used during the development of this system.

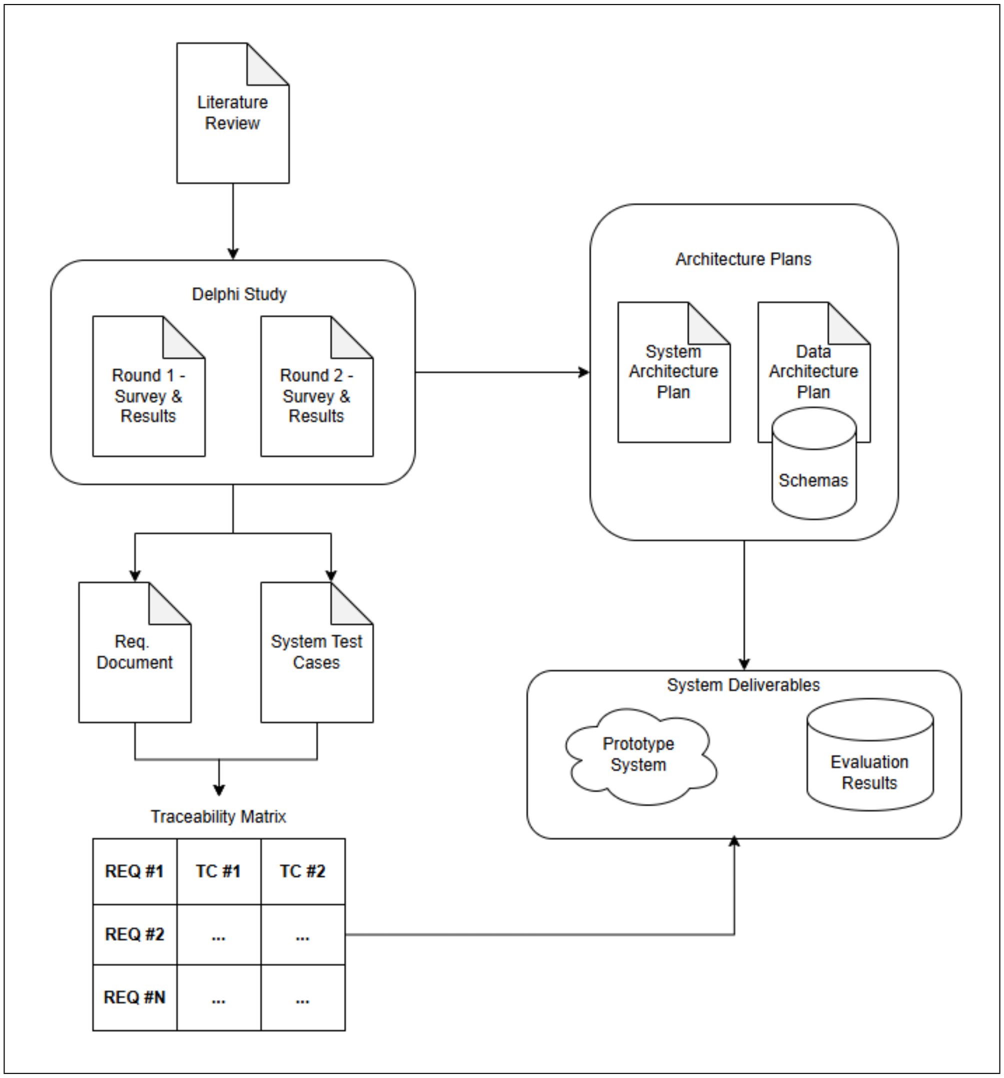


Figure 3.1: Flowchart of Project Deliverables

These deliverables shall be joined together to form a software system (for which a domain-specific language is the primary user interface). To aid in drawing quantitative conclusions from this system, the researcher plans to develop a *traceability matrix* from the requirements contained in the functional requirements document (FRD). This matrix will be joined to the test cases (TC) that are crafted directly from the requirements in the FRD.

There is precedent for the use of a traceability matrix when evaluating prototypical software. Yanbin Ye, a previous student of the UA Little Rock Department of Information Science, utilized this approach in *A Positive Data Control System for the Automation of Data Governance Functions*.

Ye's research identified a table of ten (10) system conditions that should be considered in the development of a data control system [47]. Ye then evaluated this prototype to determine how closely the artifact conformed to those requirements.

Such a scheme for system evaluation is advantageous because it helps to calculate overall system scores— which are useful when determining if such a software application meets the needs of the intended user base. These scores are often used in information quality settings where quality is gauged on a percentage basis.

As an example, a metric that may be computed during the prototype evaluation is a *conformance score* of the number of test cases that pass for a given requirement (e.g. requirement ID *R1.0.0*). This example is illustrated by the following equation:

$$\text{Conformance Score}_{\text{R1.0.0}} = \frac{\text{\# Test Cases Passed}}{\text{\# Total Test Cases}} \times 100\%$$

Another possible evaluation score is the number of test cycles completed to ensure that all test cases pass for a specific requirement. This evaluation scheme is also helpful because the designer can justify why certain requirements may have been changed (or removed) during the design and development phases of this project.

This traceability matrix is intended to be completed in concert with the requirements document, which is ultimately dependent on the completion of the Delphi study. The requirements modeling and traceability matrix are planned to be completed in the Fall 2025 semester.

## REFERENCES


[1] Safe and Sound Security, "Everything You Need to Know about NERC CIP Compliance." https://getsafeandsound.com/blog/what-is-nerc-cip-certification/, 2023. Accessed: 2024-11-01.

[2] Federal Energy Regulatory Commission (FERC), "An Introductory Guide to Electricity Markets regulated by the Federal Energy Regulatory Commission." https://www.ferc.gov/introductory-guide-electricity-markets-regulated-federal-energy-regulatory-commission. Accessed: 2024-11-18.

[3] J. Koolemans-Beynen, "The Role of the Independent Market Monitor around the World." https://www.superservicios.gov.co/sites/default/files/inline-files/johanna_presentation_independent_market_monitor_presentations2.pdf. Accessed: 2024-11-01.

[4] Oak Ridge National Laboratory, "Nils Stenvig: Exploring the world's largest machine | ORNL." https://www.ornl.gov/blog/nils-stenvig-exploring-worlds-largest-machine. Accessed: 2024-11-01.

[5] "2023 annual state of the market report." https://www.spp.org/documents/71645/2023 Accessed: 2025-04-09.

[6] "2022 annual state of the market report." https://www.spp.org/documents/69330/2022 Accessed: 2025-04-09.

[7] "2021 annual state of the market report." https://www.spp.org/documents/67104/2021 Accessed: 2025-04-09.

[8] "2020 annual state of the market report." https://www.spp.org/documents/65161/2020 Accessed: 2025-04-09.

[9] "National transmission analysis maps next chapter of us grid evolution." https://www.nrel.gov/news/features/2024/national-transmission-planning-study.html. Accessed: 2025-04-09.

[10] "Sector Spotlight: Electricity Substation Physical Security." https://www.cisa.gov/sites/default/files/2023-02/Sector

[11] "Explained: Causes of three recent major blackouts and what is being done in response." https://www.nrel.gov/docs/fy24osti/87308.pdf. Accessed: 2025-04-01.

[12] "How it Works: The Role of a Balancing Authority." https://www.energy.gov/sites/default/files/2023-08/Balancing Accessed 2025-03-19.

[13] FERC, "Electric Power Markets | Federal Energy Regulatory Commission." https://www.ferc.gov/electric-power-markets. Accessed: 2024-11-01.

[14] "What is SCADA, Who Uses it and How SCADA Has Evolved." https://inductiveautomation.com/resources/article/what-is-scada, 2018. Accessed: 2025-03-28.

[15] P. Bondarenko, "Enron Scandal." https://www.britannica.com/event/Enron-scandal/Downfall-and-bankruptcy, 2025. Accessed: 2025-03-30.

[16] NERC, "Glossary of Terms Used in NERC Reliability Standards." https://www.nerc.com/pa/Stand/Glossary2018. Accessed: 2024-10-24.

[17] "Enron historical stock price." https://github.com/echarlesgrasby/dissertation/tree/master/ext/. Accessed: 2025-03-30 (archived from https://famous-trials.com).

[18] "Prohibition of Energy Market Manipulation." https://www.ecfr.gov/current/title-18/chapter-I/subchapter-A/part-1c, 2006.

[19] Reuters, “JPMorgan to pay $410 million to settle power market case.” https://www.reuters.com/article/business/jpmorgan-to-pay-410-million-to-settle-power-market-case-idUSBRE96T0NA/. Accessed: 2025-04-10.

[20] Federal Energy Regulatory Commission, “Ferc orders greenhat to respond to market manipulation allegations, proposes penalties.” https://www.ferc.gov/news-events/news/ferc-orders-greenhat-respond-market-manipulation-allegations-proposes-penalties. Accessed: 2025-04-10.

[21] Reuters, “Vitol to pay $2.3 mln to settle California power market manipulation charges.” https://www.reuters.com/business/energy/vitol-pay-23-mln-settle-california-power-market-manipulation-charges-2024-01-05/. Accessed: 2025-04-10.

[22] www.iso-ne.com, “FAQs: Locational Marginal Pricing.” https://www.iso-ne.com/participate/support/faq/lmp. Accessed: 2024-11-07.

[23] S. P. Pool, “Bilateral settlement schedules overview.” https://www.spp.org/documents/28727/bss Accessed: 2025-04-19.

[24] A. F. Rahimi and A. Y. Sheffrin, “Effective market monitoring in deregulated electricity markets,” in *IEEE Transactions on Power Systems*, vol. 18, p. 486–493, 2003. https://doi.org/10.1109/tpwrs.2003.810680.

[25] C. Goldman, B. Lesieutre, and E. Bartholomew, “A Review of Market Monitoring Activities at U.S. Indepdendent System Operators.” https://www.osti.gov/servlets/purl/821332, 2024. Accessed: 2024-12-02.

[26] C. Gao, E. Bompard, R. Napoli, and J. Zhou, “Design of the electricity market monitoring system,” in *2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008. https://doi.org/10.1109/drpt.2008.4523386.

[27] J. Yang, “A Market Monitoring System for the Open Electricity Markets,” in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, IEEE, 2001. https://doi.org/10.1109/pess.2001.970019.

[28] J. Wang, Y. Liu, P. Li, Z. Lin, S. Sindakis, and S. Aggarwal, "Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality," *Journal of the Knowledge Economy*, vol. 74, February 2023. https://link.springer.com/article/10.1007/s13132-022-01096-6.

[29] K. Oyoo and D. Berleant, "An Automated Data Validation Approach to Enterprise Asset Management for Power and Utilities Organizations," in *2021 IEEE Electrical Power and Energy Conference (EPEC)*, 2021. https://doi.org/10.1109/epec52095.2021.9621703.

[30] Richard Baskerville, "What design science is not," *European Journal of Information Systems*, vol. 17, 2008. https://doi.org/10.1057/ejis.2008.45.

[31] D. Ray and B. Snyder, "Strategies to Address the Problem of Exiting Expertise in the Electric Power Industry," in *Annual Hawaii International Conference on System Sciences (HICSS)*, 2006. https://ieeexplore.ieee.org/document/1579804.

[32] Şafak Gündüz, "Knowledge Management in Esoteric Management," in *Proceedings of ADVED 2016 2nd International Conference on Advances in Education and Social Sciences 10-12 October 2016- Istanbul, Turkey*, pp. 273–277, 2016. ISBN: 978-605-64453-8-5.

[33] Kosar, Tomaz, et al., "Comparing General-Purpose and Domain-Specific Languages: An Empirical Study," *Computer Science and Information Systems*, vol. 7, 2010. http://dx.doi.org/10.2298/CSIS1002247K.

[34] "What is sql (structured query language)?." https://aws.amazon.com/what-is/sql/. Accessed: 2025-07-27.

[35] "What is terraform?." https://developer.hashicorp.com/terraform/intro. Accessed: 2025-07-27.

[36] "What is apache pig?." https://pig.apache.org/about.htmlstatus. Accessed: 2025-07-27.

[37] "Awk: A beginners guide for humans." https://dev.to/jpmcb/awk-a-beginners-guide-for-humans-3l25. Accessed: 2025-07-27.

[38] "Domain-specific languages guide." https://martinfowler.com/dsl.html. Published: August 28, 2019.

[39] Ojars Krasts, Arnis Kleins, and Artis Teilans, "Domain specific language for securities settlement systems," in *2012 Second International Conference on Digital Information Processing and Communications (ICDIPC)*, 2012. https://doi.org/10.1109/icdipc.2012.6257291.

[40] S. Sobernig, M. Strembeck, and A. Beck, "Developing a Domain Specific Language for Scheduling in the European Energy Sector," in *6th International Conference, Software Language Engineering 2013, Indianapolis, IN, USA*, pp. 19–35, Springer International Publishing, 2013. https://doi.org/10.1007/978-3-319-02654-$1_2$.

[41] Atlan Pte. Ltd., "Data Swamp: Is It Sinking You In?." https://atlan.com/data-swamp-explained/. Last Updated: 2023-11-21.

[42] W. Sauer and C. Daignault, "Enhancement of Electricity Market Surveillance and Analysis through Ongoing Electronic Delivery of Data from Regional Transmission Organizations and Independent System Operators." https://www.ferc.gov/sites/default/files/2020-06/OrderNo.760.pdf, 2012. Accessed: 2024-12-02.

[43] Southwest Power Pool, Inc., "Integrated marketplace (imp) glossary." https://www.spp.org/glossary/?term=bid. Accessed: 2025-09-10.

[44] Electric Power Engineers, LLC, "Shift Factor Analysis for Unlocking Grid Capacity in Renewable Energy Projects." https://epeconsulting.com/epe-intelligence/publications/grid-capacity-shift-factor-analysis-renewable-projects. Accessed: 2024-06-24.

[45] P. Twomey, R. Green, K. Neuhoff, and D. Newberry (Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology), "A Review of the Monitoring of Market Power: The Possible Roles of TSOs in Monitoring for Market Power Issues in Congested Transmission Systems."

https://ceepr.mit.edu/wp-content/uploads/2023/02/2005-002.pdf, 2005. Accessed: Nov. 02, 2024.

[46] J. Solas, T. Greenberg, “What is a Requirements Traceability Matrix (RTM)?.” https://project-management.com/requirements-traceability-matrix-rtm/. Last Updated: 2025-08-20.

[47] Yanbin Ye, “A positive data control system for the automation of data governance functions,” *University of Arkansas at Little Rock ProQuest Dissertations Theses*, 2021. Retrieved from https://go.openathens.net/redirector/ualr.edu?url=https://www.proquest.com/dissertations-theses/positive-data-control-system-automation/docview/2566109721/se-2.

APPENDIX A

INSTITUTIONAL REVIEW BOARD CERTIFICATES

UA LITTLE ROCK

**INSTITUTIONAL REVIEW BOARD**
UNIVERSITY OF ARKANSAS AT LITTLE ROCK
2801 S. University Ave., Little Rock, AR 72204-1099 | (O) 501.916-6209 | irb@ualr.edu

**TO:** Eric Charles Grasby / Dr. Daniel Berleant, Advisor

**FROM:** UA Little Rock Institutional Review Board

**CC:** Office of Research Compliance

**DATE:** October 4, 2024

**PROTOCOL:** #24-062 “Building an Information Quality-backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market”

---

Thank you for your recent submission to the University of Arkansas at Little Rock Institutional Review Board (IRB). The IRB (FWA #00002205), has reviewed **Protocol # 24-062 “Building an Information Quality-backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market”.** The review was limited to activities described in the protocol narrative. On the basis of the review of this protocol, it has been determined that this research **does not meet the criteria of *human participant research*** based on 45 CFR 46.102.

You are reminded that:

1. The investigator assumes responsibility for notifying the University of Arkansas at Little Rock Institutional Review Board of any changes in the activities described in the protocol. Re-review might be required. Changes include the future use of the data even if that future use was mentioned in the protocol.
2. If absolutely no changes occur to the protocol, then the data does not need to be retained for the requisite three years.

Best of luck with your study.

Figure A.1: Interview Pilot Study

**INSTITUTIONAL REVIEW BOARD (IRB)**
UNIVERSITY OF ARKANSAS AT LITTLE ROCK
2801 S. University Ave., Little Rock, AR 72204-1099 | (O) 501.916-6207 | irb@ualr.edu

**TO:** Eric Charles Grasby / Dr. Daniel Berleant, Advisor

**FROM:** UA Little Rock Institutional Review Board

**CC:** Office of Research Compliance

**DATE:** March 25, 2025

**PROTOCOL:** #25-020-R1 "Building an Information Quality-Backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market (Delphi Study Component)"

---

Thank you for your recent submission to the University of Arkansas at Little Rock Institutional Review Board (IRB). The IRB (FWA #00002205) has reviewed **Protocol #25-020-R1 "Building an Information Quality-Backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market (Delphi Study Component)".** The review was limited to activities described in the protocol narrative. On the basis of the review of this protocol**,** we find that it meets the IRB's criteria for protection of human participants. Your protocol has been **approved** as Human Participant Research. You may now begin to collect data.

You are reminded that:

1. **Your IRB approval end date is March 24, 2026.** If you wish to continue to collect data past the IRB approval end date, you will need to submit a *Request for Continuing Review*. This request must be submitted and approved prior to the end date to allow for a seamless extension. If approval is not received by that date, all research involving human participants connected with this protocol must cease until such time as approval is received.
2. This permission to collect data is restricted to UA Little Rock Principal Investigators (PIs).
3. This approval is limited to **only** the activities described in the approved protocol narrative. Any changes to **Protocol #25-020-R1** (examples include, but are not limited to, a change in PIs, research staff, research design, recruitment, consent forms, or the data that are collected) require a *Request for Protocol Modification*. All requests must be approved by the IRB prior to the initiation of a change.
4. PIs and all research staff are required to follow all FERPA policies. For questions about FERPA policies, contact Malissa Mathis (mktrantham@ualr.edu).
5. All submissions, including modifications and extensions, are reviewed to ensure that the protocol meets current requirements for protection of human participants.
6. Unanticipated problems or new information about risk must be reported within 2 business days to the IRB. The Adverse Event reporting form is located at https://ualr.edu/irb/home/irb-forms/

*Any future research using this data beyond what is described in the protocol requires IRB approval.*

Best of luck with your study.

Figure A.2: Delphi Study

**INSTITUTIONAL REVIEW BOARD (IRB)**
University of Arkansas at Little Rock

2801 S. University Ave., Little Rock, AR 72204-1099 | (O) 501.916-6207 | irb@ualr.edu

**TO:** Eric Charles Grasby / Dr. Daniel Berleant, Advisor
**FROM:** UA Little Rock Institutional Review Board
**CC:** Office of Research Compliance

**DATE:** August 15, 2025

**PROTOCOL:** 25-020-M1-R1 "Building an Information Quality-Backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market (Delphi Study Component)"

---

Thank you for your recent submission to the University of Arkansas at Little Rock Institutional Review Board (IRB). The IRB (FWA #00002205) has reviewed **Protocol 25-020-M1-R1 "Building an Information Quality-Backed Analysis System to Support Monitoring and Surveillance Operations Within a Wholesale Electricity Market (Delphi Study Component)".** The review was limited to activities described in the protocol narrative. On the basis of the review of this protocol, we find that it meets the IRB's criteria for protection of human participants. Your protocol has been **approved** as Human Participant Research. You may now begin to collect data.

You are reminded that:

1. **Your IRB approval end date is March 24, 2026.** If you wish to continue to collect data past the IRB approval end date, you will need to submit a *Request for Continuing Review*. This request must be submitted and approved prior to the end date to allow for a seamless extension. If approval is not received by that date, all research involving human participants connected with this protocol must cease until such time as approval is received.
2. This permission to collect data is restricted to UA Little Rock Principal Investigators (PIs).
3. This approval is limited to **only** the activities described in the approved protocol narrative. Any changes to **Protocol 25-020-M1-R1** (examples include, but are not limited to, a change in PIs, research staff, research design, recruitment, consent forms, or the data that are collected) require a *Request for Protocol Modification*. All requests must be approved by the IRB prior to the initiation of a change.
4. PIs and all research staff are required to follow all FERPA policies. For questions about FERPA policies, contact Malissa Mathis (mktrantham@ualr.edu).
5. All submissions, including modifications and extensions, are reviewed to ensure that the protocol meets current requirements for protection of human participants.
6. Unanticipated problems or new information about risk must be reported within 2 business days to the IRB. The Adverse Event reporting form is located at https://ualr.edu/irb/home/irb-forms/

*Any future research using this data beyond what is described in the protocol requires IRB approval.*

Best of luck with your study.

Figure A.3: Delphi Study - Round 2 Protocol Revision

APPENDIX B

SEARCH QUERIES USED IN LITERATURE REVIEW

*Google Scholar and IEEE Xplore were the primary academic databases that were queried for this dissertation.*

High-level searches used to find seminal articles in the areas of energy markets, market monitoring, and domain-specific languages:

- MARKET MONITORING
- MARKET MONITORING AND DOMAIN SPECIFIC LANGUAGE
- DOMAIN-SPECIFIC LANGUAGE AND BULK ELECTRIC SYSTEM
- DOMAIN SPECIFIC LANGUAGE AND FORENSICS
- DOMAIN-SPECIFIC LANGUAGE

Boolean and grouping operators were also employed to narrow the scope of the relevant articles returned for this literature review. Some of the queries included:

- "MARKET MONITORING" AND ("DOMAIN SPECIFIC LANGUAGE" OR "DOMAIN-SPECIFIC LANGUAGE")
- "MARKET MONITORING" AND ("DOMAIN SPECIFIC LANGUAGE" OR "DOMAIN-SPECIFIC LANGUAGE") AND ("ELECTRIC" OR "ELECTRICITY") -STOCK
- ("INFORMATION QUALITY" OR "DATA QUALITY") ("ELECTRICITY MARKET" OR "POWER MARKET" OR "ELECTRIC MARKET")
- ("INFORMATION QUALITY" OR "DATA QUALITY") (ELECTRICITY OR POWER OR ELECTRIC OR ENERGY) MARKET

- ("INFORMATION QUALITY" OR "DATA QUALITY") (ELECTRICITY OR POWER OR ELECTRIC OR ENERGY) "MARKET MONITORING"
- ("INFORMATION QUALITY" OR "DATA QUALITY") (ELECTRICITY OR POWER OR ELECTRIC OR ENERGY) WHOLESALE "MARKET MONITORING" -LABOR -LABOUR -"FINANCIAL MARKET"
- ALLINTITLE: WHOLESALE "INFORMATION QUALITY" OR "DATA QUALITY" OR ELECTRICITY OR POWER OR ELECTRIC OR ENERGY "MARKET MONITORING" -LABOR -LABOUR -"FINANCIAL MARKET"
- ALLINTITLE: ELECTRICITY "INFORMATION QUALITY" OR "DATA QUALITY" OR "MARKET MONITORING" -LABOR -LABOUR ("DOMAIN-SPECIFIC LANGUAGE" OR "DOMAIN SPECIFIC LANGUAGE") (ELECTRICITY OR ELECTRIC OR POWER OR ENERGY)

# APPENDIX C

## SCHEDULE OF RESEARCH MILESTONES

Note that these milestones are subject to change by the discretion of the principal investigator and research advisor. The timeline may change as milestones are completed.

| Semester | Milestone Description | Associated Deliverables |
|---|---|---|
| FALL 2024 | Article Gathering for Literature Review | *Retrieved Existing Research* |
| | Literature Review Interviews | *Interview Results* |
| | Preparation for Qualifying Exam | *Slide Deck and Summary* |
| | Complete Qualifying Exam | — |
| | Complete Required CITI Modules | *CITI Completion Certificate(s)* |
| SPRING 2025 | Complete Writing Literature Review | *Dissertation Chapter (2)* |
| | Complete Writing Proposal | *Dissertation Proposal* |
| | Schedule Proposal Defense | — |
| | Complete Proposal Defense | — |
| | Develop Initial Set of Delphi Study Questions | *Delphi Study Questions* |
| | Seek IRB Approval | *IRB Approval Letter* |
| | Begin Seating Delphi Panel | *Coordination with Experts* |
| SUMMER 2025 | Complete Dephi Study | *Paper Publication* |
| | Create Requirements Document | *FRD, BPMN Flows, etc.* |
| FALL 2025 | Finish Requirements Chapter | *Dissertation Chapter (3)* |
| | Begin Prototype Development | *Prototype Artifacts* |
| | Begin DSL specification | *DSL Spec* |
| | Begin Implementation Dissertation Chapter | *Dissertation Chapter (4)* |
| SPRING 2026 | Iterative Development | *Prototype Artifacts (cont)* |
| | Track evaluation metrics for prototype | *Evaluation results* |
| SUMMER 2026 | Prototype Setup/Deployment | *PoC environment* |
| | Iterative Development & Bugfixes | *Bug-fixes / version enhancements* |
| | Implementation and Evaluation Chapters | *Dissertation Chapters (4, 5)* |
| FALL 2026 | Finish Conclusion Chapter | *Dissertation Chapter (6)* |
| | Complete Dissertation Formatting | *Dissertation Ready for Defense* |
| | Dissertation Defense | — |

Table C.1: Schedule of Milestones for this Dissertation

# APPENDIX D

## PUBLIC MARKET DATA SOURCES

- Data Miner 2, PJM Corporation

  `https://dataminer2.pjm.com/list`

- Member Information Reporting Application (MIRA), Monitoring Analytics

  `https://www.monitoringanalytics.com/tools/tools.shtml`

- Market Data Transparency Service, ERCOT

  `https://www.ercot.com/services/mdt`

- Open Access Same-Time Information System, California ISO

  `http://oasis.caiso.com/mrioasis/logon.do`

- CAISO Data Library, California ISO

  `https://www.caiso.com/library/market-data`

- MISO Data Exchange, MISO

  `https://data-exchange.misoenergy.org/`

- Marketplace Portal, Southwest Power Pool

  `https://portal.spp.org/`

- Wholesale Energy Market Reports, US Energy Information Administration (EIA)

  `https://www.eia.gov/electricity/`

- Operational Data Portal, New York ISO

  `https://www.nyiso.com/energy-market-operational-data`

- ISO Express, ISO New England

  `https://www.iso-ne.com/markets-operations/iso-express`

# APPENDIX E

## MARKET MONITORING FIRMS

| **Name of Firm** | **Market Affiliation** | **Country** |
|---|---|---|
| Federal Energy Regulatory Commission (FERC) | Jurisdictional oversight for all US-based energy markets | USA |
| SPP Market Monitoring Unit | Southwest Power Pool (Internal) | USA |
| Potomac Economics | MISO, ERCOT, NYISO, ISO-NE (External) | USA |
| ISO-NE Internal Market Monitor | ISO-NE (Internal) | USA |
| CAISO Market Monitoring Unit | California ISO (Internal) | USA |
| Monitoring Analytics | Pennsylvania, New Jersey, Maryland Interconnect (PJM) | USA |
| Market Surveillance Administrator | Alberta | Canada |
| Independent Electricity System Operator (IESO) | Ontario | Canada |
| Canada Energy Regulator | Jurisdictional oversight for all Canadian energy markets | Canada |
| Superintendencia de Electricidad y Combustibles | — | Chile |
| Agencia Nacional de Energia Electricia | — | Brazil |
| Ente Regulador de la Electricidad | — | Argentina |
| Superintendencia de Electricidad | — | Bolivia |
| Regulation on Wholesale Energy Market Integrity and Transparency (REMIT) | European Union Agency for the Cooperation of Energy Regulators | Europe |
| AER | Australian Energy Regulator | Australia |
| NZEA | New Zealand Electricity Authority | New Zealand |
| PEMC | Philippine Electricity Market Corporation | Philippines |

Table E.1: A Roster of Groups with Market Monitoring Authority (source: [3])