How to Successfully Implement a Corporate Taxonomy

By Kelly Wibbenmeyer

North central University, United States

Abstract- Purpose: The purpose of this paper is to develop a methodology to follow while implementing a corporate taxonomy.

Design/methodology/approach: The presented results are grounded in both academic literature on taxonomy and qualitative data from two departments within the same organization that implemented separate taxonomy structures.

Findings: The study identifies ten factors to consider when implementing a corporate taxonomy as well as a defined nine step process to implement when implementing a successful corporate taxonomy.

Research limitations/implications: The scope of the literature review and the case study were both limited as finding multiple taxonomy experts in one organization is rare, the account of the research is not considered exhaustive. The paper can assist practitioners in a high level approach to implementing a corporate taxonomy as well as things to invoke to increase the chances of a successful implementation.

Keywords: corporate taxonomy, enterprise content management (ECM), knowledge management, ontology.

GJCST-H Classification: J.4 K.4.2

Strictly as per the compliance and regulations of:
How to Successfully Implement a Corporate Taxonomy

Kelly Wibbenmeyer

Abstract- Purpose: The purpose of this paper is to develop a methodology to follow while implementing a corporate taxonomy.

Design/methodology/approach: The presented results are grounded in both academic literature on taxonomy and qualitative data from two departments within the same organization that implemented separate taxonomy structures.

Findings: The study identifies ten factors to consider when implementing a corporate taxonomy as well as a defined nine step process to implement when implementing a successful corporate taxonomy.

Research limitations/implications: The scope of the literature review and the case study were both limited as finding multiple taxonomy experts in one organization is rare, the account of the research is not considered exhaustive. The paper can assist practitioners in a high level approach to implementing a corporate taxonomy as well as things to invoke to increase the chances of a successful implementation.

Practical implications: Practitioners are provided with an overview of the concepts that are instrumental in achieving a successful corporate taxonomy. The grounded knowledge within the context of this paper is also graphically displayed in a chart that provides detailed information on the importance of all enterprise content management (ECM) constructs which require a taxonomy structure for data retrieval capabilities.

Originality/value: This study is important due to data becoming increasingly important in organizations and a method for extracting as well as finding the right data when it is required is of vital importance within organizations.

Keywords: corporate taxonomy, enterprise content management (ECM), knowledge management, ontology.

I. Introduction

In our fast paced world, data growth is swifter than ever before. Corporations are struggling to keep up with data security while implementing new technologies to stay competitive (Gallagher, 2002). Also, more regulations force organizations to implement data retention practices, which involve more time and resources (Beal & Griffin, 2012). Technology is increasing at a rapid rate which makes it difficult for organizations to retain employees that are not constantly receiving training on new technologies as organizational needs transform as customer demand changes (Moore, 2000). This constant churn in technology is causing employee burnout in IT departments (Moore, 2000). Also, customer demand is changing at a quicker rate, and the expectation for IT modifications is the greatest it has ever been (Moore, 2000). The rationale for the increase of technology innovation is due to the world becoming more technically savvy. IT departments have to find a way to keep up with customer demand while their infrastructure needs, such as updating security patches and ensuring data is available for upper management, increase in demand.

An influx of technology produces an increase of data (JCN Newswire, 2013). Large amounts of data allow organizations to use the information for analysis and analytics that assist in corporate strategy and decision making (JCN Newswire, 2013). An increase in data can also cause issues for organizations (JCN Newswire, 2013). The more data an organization has, the more expensive it is to store and manage the data. Also, data is available in various different formats that it is nearly impossible to place the data in specific classifications for comparative analysis (JCN Newswire, 2013). Data can also be structured (documents, data from databases) or non-structured (website or e-mail), which also adds to the complexity of organizational data ("IDBS transforms ELN," 2015).

Technology innovations and an increase in customer demand for IT services are causing organizations to rethink their past IT strategies. Organizations that have mass amounts of IT customizations throughout the various systems have unintentionally decentralized their data (Gallagher, 2002). Organizations that were known to implement technology customizations in the past are seeking ways to reduce customization and move towards the vendor base strategies to decrease turnaround time for upgrades to meet increasing technology advancements while meeting customer needs (Gallagher, 2002).

II. Enterprise Content Management

Regulations are a primary reason organizations standardize and streamline processes (Beal & Griffin, 2012). The management of data, such as the retention and disposal of data within certain time periods occurs via organizational content management practices (Beal & Griffin, 2012). Content management practices consider all types of media, like audio, visual, and text (Votsch, 2001). Votsch (2001) defined content management as any method for capturing, storing, and retrieving data for usability. The central point of a content management system is the standardization that
Organizations are seeking ways to organize data within enterprise content management (ECM) systems which can handle both structured and unstructured data (Vom Brocke, Simons, & Cleven, 2011). Maican and Lixandroiu (2014) stated that an ECM system comprises the methods to manage and deliver data, both content, and documents, that relate to organizational processes. There are multiple benefits of ECM systems within organizations (Vom Brocke, Simons, & Cleven, 2011). Some of the benefits are the ability to find data quicker and more efficiently as well as being able to manage records management practices in an electronic means, thereby reducing paper processing and storage of hard copy documents (Vom Brocke, Simons, & Cleven, 2011). Additional benefits of ECM systems are improvements in collaboration for both internal and external sources, as well as standardizing work management practices throughout the system (Hullavarad, O'Hare, Russell & Roy, 2015).

ECM systems have multiple features to ensure they work efficiently, such as a corporate taxonomy or content lifecycle aspects (Munkvold, Paivarinta, Hodne, Stangeland, 2006). A corporate taxonomy is a data standard that the organization, as a whole, uses to classify data (Brocke, Simons, Herbst, Derungs, Novotny, 2011). Developing a corporate taxonomy tends to be a large obstacle for organizations as the file systems within various departments are different which causes more data challenges (Brocke et al., 2011). A corporate taxonomy is a vital step in content searchability throughout the organization, which assists in the retrieval of data consistently across the organization (Brocke et al., 2011).

An ECM comprises multiple components. The ECM system manages all of the organizational data. Therefore, multiple systems integrate to present all of the enterprise-wide data. According to Gilbert, Shegda, Chin, Tay, and, Koehler-Kruener, H. (2013), the major aspects of an ECM system are document management, image-processing applications, workflow management, records management, web content management, social content management, and extended components management. All of these applications within systems are imperative in organizational data processing that results in efficient data management.

III. THE IMPORTANCE OF CHANGE MANAGEMENT AND STANDARDIZATION

Per Malek & Yazdanifard (2011) change management is the ability to plan and coordinate organizational modifications to every employee impacted by the change. During change management processing there is a shift from problem identification to a potential future state. An integral aspect of managing the change is to ensure employees are ready to accept the modification by presenting benefits as well as ensuring the employee has an active role in the modification, like being a champion for the prospective change.

Change management is vastly important to the acceptance of new system implementations, especially one that encapsulates the entire organization (Munkvold et al., 2006). One of the major components of ECM implementation is change management (Munkvold et al., 2006). Change management is vital to ensuring the implementation of ECM system and for the ongoing maintenance and support of the ECM system (Munkvold et al., 2006). If the resources are not willing to accept the changes, there is little likelihood that data entry will be standardized and the ECM system will be of no use (Munkvold et al., 2006). Standardization is a key aspect of the implementation of the ECM system (Munkvold et al., 2006). If there is no consistent standard for data, the data will not be reliable for reporting and other needs (Munkvold et al., 2006). Therefore, ensuring a common taxonomy is understood and is implemented throughout the organization is an important aspect of an ECM system.

IV. CORPORATE TAXONOMY

All of the ECM system components are important pieces of the entire corporate taxonomy standard. An enterprise taxonomy standard ensures that no data silos are present (Gilbert et al., 2013). Data management is a complicated process and a workable solution that allows the appropriate users to access the appropriate data at the right time is vital to system viability within the organization (Gilbert et al., 2013). Businesses not only have to worry about how to manage new data but also how to manage legacy data within legacy systems (Gilbert et al., 2013). Determining how to handle legacy system data is an important aspect of building the corporate taxonomy as well. The development of a corporate taxonomy standard allows both new and legacy systems to interact (Gilbert et al., 2013). Data integrations allow the movement of data to interact across both new and legacy systems (Gilbert et al., 2013). Workflows represent the business processes within an organization (Vom Brocke, 2013). Work management processes may need to undergo a redesign to comply with the new corporate taxonomy standards to ensure standardization across the enterprise (Maican & Lixandroiu, 2014).

There are multiple issues when organizations do not implement a standard taxonomy (Munkvold et al., 2006). Some of the issues are data inconsistencies and therefore, reporting and analytics do not present accurate data and therefore, data integration is more difficult as data does not have a consistent naming.
scheme (Munkvold et al., 2006). There are multiple reasons why the creation of data naming standards is beneficial to the organization. Data analytics are more timely and accurate when data is in a federated format and users have a better chance of finding information throughout the organization if the entire enterprise uses the same terminology (Munkvold et al., 2006).

There are multiple references from previously published works stating a consistent taxonomy is the only way to ensure standardization, but the previous studies do not address the method for the taxonomy creation (Barrera, Duran-Limon, Medina-Ramirez, Rodíguez-Rocha, 2012; Munkvold et al., 2006). The primary problem in organizational data standardization is that there is no specific methodology for developing a corporate taxonomy. Some organizations believe that every organization is different and departments within organizations have different needs, therefore it is very challenging, if not impossible to have a corporate taxonomy standard (Eden, 2005; Munkvold et al., 2006). Other articles state that a corporate taxonomy is the best way to manage enterprise data needs (Alexander, 2012; Woods, 2004). Regardless of difficulty, standard corporate taxonomy allows organizations the ability to manage data more efficiently and allows for maximization of information flow due to quick and accurate data availability (Alexander, 2012).

There are multiple things to consider when planning the corporate taxonomy, such as the data the organization uses, compromising strategies between departments on data standards, and ensuring one single unbiased person manages the project to ensure all parties are taken into consideration (Alexander, 2012). Regardless of the methodology, there are steps to take to ensure the various system and user needs are met. The prospective taxonomy implementation plan will not compromise data standardization, but will reduce organizational customization, and increase change management adoption. The primary purpose of the paper is to develop a specific methodology to follow while implementing the corporate taxonomy.

A previous study stated that certain aspects of current work processes will change to accommodate the software package (Votsch, 2001). There are other previous studies that state the taxonomy should be based on national standards to ensure organizational buy-in (Amado-Salvatierra, Hernández, & Hilera, 2012; Hiava, 2014). There is no existing literature regarding a specific process to follow to ensure the taxonomy will fit the needs of the entire organization. The primary goal of this study is to develop a corporate taxonomy implementation plan that any organization can deploy regardless of the software vendor or national standards. Therefore, this article, which is a qualitative grounded theory study addresses the current gap in the existing literature with the following research question:

RQ1: How does the organization ensure the corporate taxonomy will be used by all users of the system?

The research question relates directly to the study, as organizations are unique, and certain questions influence how to shape the organizational data needs such as understanding the current data formats within the organization. Also ensuring the change management and educational aspects of the corporate taxonomy are understood and implemented are important aspects to ensuring the taxonomy adoption occurs throughout the organization. Change management is an important aspect to take into consideration while attempting to adopt a new change throughout the organization.

V. Materials/Methods

There are multiple definitions to comprehend to ensure a total understanding of the important concepts that relate to building a corporate taxonomy. Previous works present different definitions for the major taxonomy components of knowledge management, ECM, ontology, taxonomy, and metadata which adds to the difficulties in comprehension of these terms. Therefore, prior to discussing these concepts any further, the next step is to define these terms and explain how they relate to each other.

Knowledge management is the process of giving the right data to the right people at the right time (Rahman & Somayyeh, 2013). Kotarba (2011) described knowledge management as a system of interconnected processes. The primary processes within knowledge management are resource identification, understanding the data usage within the organization, analysis of organizational needs, and understanding, acquiring, processing, and usage of knowledgeable resources (Kotarba, 2011).

An ECM is a compilation of processes and skills to manage information assets over the entire life cycle (Hullavarad, O’Hare, Russell & Roy, 2015). The primary goal of an ECM system is to streamline tasks by implementing automation that reduces workload, allows for version control, reduces data duplication, and improves search capabilities by presenting one version of the document in one managed location (Hullavarad et al., 2015). ECM systems allow organizations to manage content across the enterprise (Grahlmann, Helms, Hilhorst, Brinkkemper, & van Amerongen, 2012). To comply with regulations, organizations must manage content which in turn fosters a collaborative environment (Grahlmann et al., 2012). When organizational data mapping occurs via an ECM strategy, the organization is more likely to comply with big data standards and also be in compliance with regulations (Hullavarad et al., 2015). The ECM must be complete, generic enough to compare and search, and should always take the future possibilities of the data into consideration (Grahlmann et
Within the Hullavarad et al. (2015) paper, a process to implement an ECM is discussed. The implementation path offered in the Hullavarad et al. (2015) paper is to conduct a strategic roadmap, develop the ECM, deploy the ECM, and implement a support structure to ensure the continual support of the ECM. The high-level process of implementing an ECM is the same fundamental concepts in implementing a knowledge management system within the Kotarba (2011) paper. Therefore, it is vital that the fundamental notions of strategy development, developing the process, deploying the process, governance, and implementing a maintenance plan are vital to implementing both ECM and knowledge management processes.

An ontology uses relationships among attributes and employs rules regarding how the relationships interact (Byrne, 2004). Ontologies are the concepts of how knowledge interacts with a system (Byrne, 2004). The ontology contains the business rules within the organization and is the basis for the taxonomy within the organization (Kotarba, 2011). Ontology practices within organizations provide consistent information regarding roles and duties as well as overall organizational processes (Castillo-Barrera, Duran-Limon, Medina-Ramirez, & Rodriguez-Rocha, 2013). Organizational rules will form the basis for the relationships between various objects within the system as well as constitute the basis for the integrations between systems (Kotarba, 2011). As the ontology undergoes development, consistent data structures, or data class generation occurs, this is known as the taxonomy (Castillo-Barrera et al., 2013). Taxonomy is a standard set of terms that can be hierarchical and represent the organizational content requirements (Byrne, 2004). Metadata or attributes describe the data throughout the lifecycle of the data (Sheriff, Bouchlaghem, El-Hamalawi, Yeomans, 2011).

Document management systems (DMSs) use ontologies and taxonomies to manage structured data within organizations (Castillo-Barrera et al., 2013). DMSs reduce costs as printing and physical file storage are no longer issues as electronic retrieval is available (Castillo-Barrera et al., 2013). Full-text searching and indexing are other features available within a DMS, which reduces time to find documentation (Castillo-Barrera et al., 2013). The taxonomy assists with document retrieval and alleviates parsing through mass quantities of data to find required information. For example, a file management system allows for searching, but the schema for searching retrieves all data with the search term listed, which can take a long time to parse through.

Knowledge management systems influence the financial decisions made within the organization as data extrapolation occurs to make business decisions (Kotarba, 2011). The data that resides in the ECM feeds the knowledge management system to ensure data is available at the appropriate times. The ontology is found within the ECM as it comprises the rules for the data within the ECM. The ontology is the theoretical aspect of the ECM as it represents all of the data models and how they interact (Byrne, 2004). The taxonomy works within the constructs of the ontology and is the system vocabulary of definitions (Byrne, 2004).

Castillo-Barrera et al., (2013) defined an ontology as a method to define terms that represent a particular area of knowledge. The ontology outlines the relationships and theories that describe the organizational data structure (Castillo-Barrera et al., 2013). The knowledge management system takes the information from the ontology and optimizes the data to increase organizational competitiveness (Castillo-Barrera et al., 2013). Therefore, ontologies are foundational to knowledge management systems (Castillo-Barrera et al., 2013).

Knowledge management and ECM coexist in different facets of the organization. Nordhiem and Paivarinta (2006) and Paivarinta and Munkvold (2005) state that ECM is a subcomponent of knowledge management as ECM systems manage data within the knowledge management system. Munkvold et al. (2006) as well as Paivarinta and Munkvold (2005) argue that the fields in an ECM are much broader than what is in the knowledge management systems, such as how scanning occurs within organizations. Other authors state that even though ECM systems support knowledge management functions, both systems are different with some overlapping features (Herschel & Jones, 2005; Kuechler & Vaishnavi, 2006). ECM systems are much broader than knowledge management systems as ECM systems manage both informational and digital information that do not belong to the knowledge management system (Vom Brocke, Simons, & Cleven, 2011). Therefore, the ECM framework and knowledge management functionality represent two different but coinciding systems of thought.

ECM systems also integrate document management, content management (via the web), and record management technologies (Vom Brocke, Simons, & Cleven, 2011). The integrated content concept for an ECM stems from the notion that the management of all organizational data occurs within the ECM (Vom Brocke, Simons, & Cleven, 2011). Besides managing all content within an organization, the ECM must also control versioning of data, searchability of data, and storage of data (Vom Brocke, Simons, & Cleven, 2011). A graphical depiction of the relationship between knowledge management, ECM, ontology, and taxonomy is below in Figure 1.

Understanding the basic concepts of how knowledge management, ECM, ontology, and taxonomy integrate is an important aspect of the research. The
purpose of this article is to propose a specific methodology for composing a corporate taxonomy, but it is vital that the reader understands how all of the concepts relate to each other as that relationship is an important aspect of the creation of the taxonomy proposal.

VI. THE IMPORTANCE OF CORPORATE TAXONOMY

A corporate taxonomy allows data to be searchable (Vom Brocke, Simons, & Cleven, 2011). If the data contains searchability issues, then the system users will have difficulty using the system and user adoption issues will occur (Vom Brocke, Simons, & Cleven, 2011). A corporate taxonomy organizes the data within the system by normalizing data throughout the organization (Vom Brocke, Simons, & Cleven, 2011). Access control of data is of great importance as a poorly designed system can lead to data theft or unintentional data access (Vom Brocke, Simons, & Cleven, 2011). Organizations should understand the access control restrictions and not make the system too restrictive else, it will impede end user usage of the system (Vom Brocke, Simons, & Cleven, 2011). Cybersecurity and big data requirements should also be taken into consideration when dealing with system access and security features (Vom Brocke, Simons, & Cleven, 2011). Access control and other security mandates are important aspects of understanding prior to devising the corporate taxonomy standard for an organization.

Another important concept to understand when creating a corporate taxonomy is the difference between structured versus unstructured data. Structured data is formally defined data usually kept in a database or numerical data (Markham, Kowolenko, & Michaelis, 2015). Structured data uses a classification system via the use of metadata or attributes (Gardner, 2014). Metadata is information that describes the data (Payne, 2013). Some examples of metadata fields are the audience for the data, the language the data is in, and the source of the data. Attributes are specific data fields from a common set of values (Payne, 2013). An example of an attribute field is color, and a set of responses for the attribute would be red, green, blue, and orange. Unstructured data comprises notes, text, and other data that lacks metadata (Gardner, 2014). Structured data uses a standard taxonomy classification system, which value rich metadata and tagging that is inherent in the taxonomy (“Semantic content enrichment,” 2011). There are multiple tools on the market which add metadata tags to add value and structure to unstructured data (“Semantic content enrichment,” 2011). The addition of metadata tags to unstructured data allows for data management within the data analytics tool (“Semantic content enrichment,” 2011). The data analytics tools within organizations provide valuable data to end users and is part of the knowledge management process. Therefore, both structured and unstructured data is of great importance to the implementation of a corporate taxonomy.

Data and workflow management are challenging when attempting to merge systems with structured and unstructured data (Grahlinman et al., 2012). Therefore, interfacing technology is a vital aspect when managing all organizational data (Grahlinman et al., 2012). The ECM system, with the use of the ontology rules and taxonomy, deals with the management of unstructured data (Vom Brocke, Simons, & Cleven, 2011). Multiple other studies state ECM systems combine both structured and unstructured data, which occurs through the integration of applications that contain structured and unstructured data (Chu, Chen, & Chen, 2009; Nordheim & Paivarinta, 2006). Therefore, all data, both structured and unstructured, is centrally located in the ECM system which enables enterprise workflow management to occur.

VII. THEORETICAL PERSPECTIVE

There is one major theory and one concept that relate to the implementation of a corporate taxonomy; Lewin’s change management theory and the theory of Martec’s Law. The goal of Lewin’s change management research was to understand why change occurs, generalize change practices, and improve the planning of change throughout society (Johnson, 2014). Change management is very popular in today’s society due to a rapid pace of technology which promotes constant organizational change (Johnson, 2014). If organizational resources do not embrace change, failure is imminent (Jaffar & Weistroffer, 2012). Developing a corporate taxonomy will require buy-in from all aspects of the organization as well as senior management support to ensure all levels of the organization are implementing a consistent taxonomy across the organization (Jaffar & Weistroffer, 2012). If various departments choose to opt out of the taxonomy, then the data consistency factor is not complete. The rationale for a corporate taxonomy is to streamline structured data for consistency across the organization. Data consistency leads to dependable data, and organizational knowledge becomes more dependable (Munkvold et al., 2006). Therefore, corporate taxonomy is the best way to standardize data across the organization and enhances data analytical output.

Technology is changing at such great rates that organizations will be unable to keep up with the increasing demands (Brinker, 2013). Organizations are reducing complexity to create data standardization and to be able to keep up with customer demand (Wadhwa & Harper, 2014). Therefore, organizations must be
strategic in what organizational changes to implement (Brinker, 2013). Martec’s Law states that organizational change occurs steadily, whereas technology changes occur at an increasingly rapid rate (Brinker, 2013). This concept is another important rationale supporting the creation of a corporate taxonomy. As long as corporate data remains unstructured and has no ontology rules to formalize the data, analytics will not be accurate as data will not have any consistency. A corporate taxonomy adds data consistency to the overall organization and allows for a method for finding and classifying data (Jan, Simons, Herbst, Derungs, & Novotny, 2011).

VIII. STUDY OVERVIEW

The study involves a large U.S. electric utility organization that uses the same electronic document management system in two separate departments that has two separate taxonomy implementations. The qualitative grounded theory design allows the system administrators and end users to present their rationale for the different implementations of two different taxonomy systems that presents the differences and similarities within the taxonomies, and their thoughts on the idea of structuring a corporate taxonomy. Within a grounded theory study, data collection and analysis occurs until a theory emerges (Glaser & Strauss, 1967). Coding of common themes emerge and an extensive literature review occurs to determine if there are similarities in existing data (Glaser & Strauss, 1967). The goal of grounded theory research is to discover basic patterns that evolve into theory generation (Glaser & Strauss, 1967). The theories that evolve from grounded theory research change until all observation is complete (Glaser & Strauss, 1967). Grounded theory studies are useful when trying to develop new theories that are based on existing research (Glaser & Strauss, 1967).

The study involves an organization that has resident taxonomy experts, which deployed two separate taxonomy structures. There are only two departments within the larger organization that currently place their documents into a system that incorporates a taxonomy structure. The rest of the organization is actively looking for ways to structure data to account for the increasing need to provide data analytics and overall enterprise data management. Therefore, a grounded theory approach works well to extrapolate the data from the taxonomy experts to determine the best method for deployment of a corporate taxonomy structure within the organization.

Interviews are the main data collection method. Secondary sources of data were found in documentation and follow up calls to validate the responses. The first organizational business segment implemented their taxonomy in the 1990s, this organizational unit, is classified as department A throughout the rest of the paper. The second organizational segment, which is classified as department B throughout the rest of the paper, reviewed department A’s lessons learned and came up with a preferential method of taxonomy deployment in the late 2000s. A taxonomy specialist was brought in to assist with data collection to enhance the change management principles for department B’s implementation. The organization is a suitable organization to use for the grounded theory study as multiple employees have a thorough understanding of taxonomy benefits and challenges. The selection of study participants was based on users that were well-known taxonomy experts within the organization, end users of the taxonomy system, as well as IT system administrators who manage the data within the system.

The qualitative question is in direct alignment with the primary purpose of the paper, which is to develop a specific method to implement a corporate taxonomy. A total of five people (two from department A and three from department B) were interviewed, with an average length of 60 minutes. The interviews were manually documented during and reviewed after the interview. The interviews focused primarily on the following areas:

1. document management taxonomy current practices and challenges;
2. difficulties implementing taxonomy within the department or organizational segment; and
3. implementing a corporate taxonomy and the perceived challenges and benefits.

IX. STUDY RESULTS

Prior to discussing the results of the study, a general overview of the two separate departments is an important aspect of the study. The departments are vastly different in the methodologies used to implement the taxonomy. After the overview, the discussion continues with the major themes of the study.

Department A, had a very flat taxonomy (over 1,000 classes), due to the limited timeframe to place all of the documentation in the system. Department A decided to migrate the class structure from the mainframe system to the new document management system. The implementation occurred in the early 1990s, and there was no resident taxonomy expert present during the taxonomy implementation. The flat taxonomy made it very difficult to find anything in the system. Department A had approximately 100 data entry clerks who handled data entry in the document management system. Allowing specifically trained groups of users to take responsibility for data entry ensures that the data entry process is consistent, which aids in users searching and finding their documentation. End users were able to find data in the system since the data was consistent, but not without initial challenges.
The data clerks provided assistance to end users who could not find their data, this aided in taxonomy adoption as the experts were on site and easily accessible. After ten years of experts performing data entry, end users were able to quickly add documentation to the system as they understood how to classify the data after ten years of searching within the system.

Department B implemented a high-level class structure, with only 12 classes. The reduction of classes increases the likelihood that end users find their files. Also, finding data was easier and more efficient than ever before. Department B reviewed the lessons learned from department A and spent time interviewing the users of the current document management system of the current issues within the system. There was no existing taxonomy within department B’s document management system and end users were having an extremely difficult time retrieving documents from the system due to the lack of taxonomy. During the implementation of department B’s taxonomy, end users required more efficient and easier access to documents and therefore, end users were more hands on in the implementation. There were controls and workflows put in place to allow end users to create documents, but the documents were not approved until data review occurs with the data taxonomy specialists. This method allows the flexibility to add documents and the controls needed to ensure documents are in the system correctly for later searchability.

The two separate implementations of the taxonomy had some large differences as well as some similarities. Department A, implemented a flat taxonomy due to incorporating the taxonomy structure from legacy mainframe systems whereas department B, implemented a brand new taxonomy from users insights and a migration path to enter legacy data into the new system. Both departments were successful with the taxonomy implementation due to the use of a set of super users who handled data entry and validation.

There were multiple major themes that emerged from the study to ensure a successful taxonomy implementation within an organization. Every study participant discussed two vital aspects to consider while implementing, namely end user concerns and workflow. Therefore, these items will be discussed first. After the end user concerns, workflow, and taxonomy governance discussion, this article changes direction and a discussion of benefits of a taxonomy, issues that occur when implementing the taxonomy, and finally how to guarantee a successful taxonomy implementation is present.

a) End User Concerns

The taxonomy specialist within both departments spoke about the end user needs. End users want to find their data, but do not want to spend the time placing their data in the system accurately to find later. Pincher (2010) states that if organizations want to be successful, all users must understand your content. End users have great difficulty understanding the taxonomy at first. Therefore, the usage of specialist for data entry is a huge plus, if the organization can allocate resources for data validation practices. Pincher (2010) states that content managers and owners are imperative to ensuring content is correct. Content managers approve and edit content and content owners publish content and apply appropriate metadata (Pincher, 2010). Allowing the end users to use the system as a search tool shows the end users how useful the system can be regarding finding their documents quickly. Therefore, when the organization decides to allow end user data entry, the end users will be more cognizant regarding taxonomy to ensure searchability and retrieval ease when finding their documentation.

b) Workflow

Workflow is an important aspect of taxonomy implementation as it determines who is performing what tasks in the organization to ensure data creation and storage is correct. If workflow is not used regularly then it will have a difficult time being accepted by the end users. Minimizing clicks and simplicity is a requirement when dealing with the workflow. Pincher (2010) states that ease of use and user adoption run parallel to each other. Workflow flexibility is a key aspect of workability and user adoption (Pincher, 2010). In department B, the workflow is used one to five times a year and failed because of no consistent usage. The end users did not want to spend the time learning and understanding workflow as they felt it was bothersome. They preferred to work outside the system on the infrequent tasks. In department A, the workflow is in use constantly, and department A has had great success implementing workflow in the organization.

c) Benefits of a Taxonomy

The benefits of implementing a taxonomy were consistent across all interviews. Creating a taxonomy allows for less paper and shipping expenses, as the documents are all in one location, and end users print out their documents. Finding documents is easy and is a huge time saver throughout the organization. Document organization and searchability are two key aspects of any taxonomy (Pincher, 2010). All documentation is in one system, and there are multiple ways to search and find data. Therefore, documents that were once lost can now be found easily. All of the documents are consistent across the organization, therefore if a user changes departments or locations, their rules and standards are the same.

d) Issues with Taxonomy Implementation

There were issues with the taxonomy implementation. Department A implemented a flat
how to successfully implement a corporate taxonomy

the participants spent the majority of the time discussing their current department taxonomy implementation. the taxonomy experts gave their advice regarding things to do to ensure success when implementing a corporate taxonomy. although many topics were present in the research, the items below were consistently present in the interviews with participants.

a) good change management practices

good change management is imperative to taxonomy success. if the organization does not educate and train all members regarding why the taxonomy is important, it will fail. the system will fail if end users do not understand the benefits of the system. therefore, change management is imperative to the implementation of a corporate taxonomy. a good change management practice not only has backing from senior management for the implementation but to ensure the user community is ready to accept the change (decker, durand, mayfield, mccormack, skinner, & perdue, 2012). the implementation should remove as much complexity as possible to ensure a good change management perspective (decker et al., 2012).

b) senior management support

senior management support is crucial to the implementation of a corporate taxonomy. if the senior leaders do not fully support a corporate taxonomy, the implementation will fail. senior management support should drive the effort, ensure appropriate resources are available to support the effort, and ensure other resource requirements are available for input. without senior management support, the taxonomy effort will not be successful as the only way to get all members of the organization consistent focus is via senior management support (janvrin & no, 2012).

c) one person to manage the effort

a specific person should handle the corporate taxonomy effort. having one overall point of contact ensures the data and software silos have one person as a focal point of contact. having one person that is not specifically tied to any one of the department silos also ensures there is no favoritism during the implementation of the taxonomy. this person should have an excellent understanding of taxonomy and the other corporate regulations that must be met after the taxonomy is in place. de koning, de mast, does, vermaat, and simons (2008), state that when implementing any project, one main person should be responsible for the roll-up of the entire plan as this person has an understanding of the total effort and can influence other aspects of the project when necessary. some of the specific regulations or corporate policies that should be considered are data security compliance, data classification standards, and records management practices. the taxonomy must be driven by the tools used within the departments, which means the taxonomy is not driven by software but by organizational need within specific software implementations. the person responsible for the taxonomy effort should also ensure it is understood in every application how to deploy the taxonomy with the application, train users, and have guides and other support documentation to support the effort.

d) limit the taxonomy structure to high-level classes

the biggest reason for taxonomy success within department b was due to limiting the number of classes. if the taxonomy sticks to a high-level class structure, a reduction in the amount of time to structure data in other non-taxonomic systems will occur as it is easier to classify data into groupings of 10 or 20 versus 100. pincher (2010) states to limit the classes to six to twelve high-level classes to ensure success. the taxonomy should also only consist of two or three levels deep to continue the simplistic concept (pincher, 2010). also, training is easier throughout the organization with a reduction in classes. there are fewer disagreements in the data structure and classification when the taxonomy is limited. for example, one of department b’s classes is policy. in another organization, policies were broken down into specific types of policy. instead of adding an attribute stating the policy was a corporate policy versus a department policy, a class was added which led to confusion and disagreement. therefore, implementing a high-level taxonomy and using metadata to add detail...
to the documentation is the best route to ensure corporate taxonomy success.

e) Governance

One of the most important aspects of taxonomy administration is having a team of taxonomy experts decide on taxonomy modifications. Pincher (2010) states that a governance board should define the overall strategy and ensure appropriate content standards are being met. The taxonomy team should also ensure content entry is appropriate as well as developing standards for metadata (Pincher, 2010). The governance team should consist of a minimum of six and a maximum of 12 members (Pincher, 2010). The members in department B state that the number of members on the governance board should be representative of the organizational population, but to ensure there are not too many members else, no decisions will be made, due to lack of agreement. The members should be representative of the organization. Department B had a governance structure in place from the outset of the taxonomy implementation and made few changes to the structure. The taxonomy governance team is very stringent regarding what constitutes a new taxonomy class and what is added as an attribute or metadata. End users are consistently requesting new classes, and the governance team determines if it is a valid request, and if the request is valid, a thorough discussion regarding data integrity ensues. This team over a five year period has only added four new classes, and two of the four classes are system based classes.

f) Work on the Taxonomy First

The taxonomy is the most important aspect of the data classification system and, therefore, should be the primary focus before any data is put into a system. Pincher (2010) states that if corporations start with the taxonomy first, it builds a foundation for organizations to expand their designs. If the organization does not work on the taxonomy first, disorganization occurs and leads to the lack of user adoption issues as well as system confusion (Pincher, 2010). The taxonomy structure should be complete prior to working on any other data aspects of the system, like security, records management, or data classification. The secondary aspects are important and can influence the taxonomy structure, but should not override the overall classification structure. For example, many departments within the organization are working towards records management initiatives and want the taxonomy to follow how the department classifies data. Each department can classify data retention differently and if the organization attempts to create the record management structure and hope that the taxonomy matches will fail greatly. The organizational goal is to have a corporate taxonomy and not a standard for managing records throughout the organization, this is important to remember when working on corporate data initiatives as users tend to be narrow focused when attempting to complete a specific task.

XI. Discussion

In summary, the grounded theory study presented multiple concepts to take into consideration when attempting to establish a corporate taxonomy. The results are summarized in Table 1 below.

Table 1: Things to Consider when Establishing a Corporate Taxonomy

| Keep the taxonomy simple and at a high level |
| — Senior management support is critical |
| Only use workflow if users are going to consistently use it |
| Think about legacy data and clean it up prior to placing into a new system |
| Have a group responsible for data entry (at least at first) |
| Continuously train organization |
| Have great change management practices |
| Have one person responsible for the overall effort especially in large organizations |
| Have a governance board in place to make decisions |
| Work on taxonomy before any other corporate data initiative to reduce rework |

a) Unstructured Data

There were some concepts that were not present as the grounded study was specific to two instances of a document management system and did not involve unstructured data. Much of an organization's data is unstructured data due to the expansion of web pages and media. Participants from department B stated that content that was previously classified in the document management system would be linked to web pages but web pages themselves were not classified. Additionally, study participants noted that e-mail messages could contain important data, and if data was important enough to capture, then it was entered into the document management system manually. Pincher (2010) states that corporations need to determine what data they want to classify and what data does not require classification. Unstructured document management, such as the management of web pages is complicated. Traditional document clustering occurs in a manual form that is not conducive to the rapid rate at which web development occurs (Singh, Hsu & Moon, 2013). New advances in technology offer an on-the-fly assignment of data on web pages, some examples include Clusty (www.clusty.com) or Grokker (www.grokker.com) (Singh, Hsu, & Moon, 2013). The advancements in technology present an important concept regarding data analytics and data storage. All data in the organization is stored somewhere, but not all
data is transformed into data analytics. Organizations should be cognizant that not all data is required for usage.

b) Specific Methodology when Implementing a Corporate Taxonomy

Table 1 discusses the factors to take into consideration when implementing a corporate taxonomy. The development of Table 2 below is based on interview data from respondents in conjunction with the data from Table 1 above. Additional detail for each step is outlined in this section.

Table 2: Steps to Take when Implementing a Corporate Taxonomy

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Obtain senior management support</td>
</tr>
<tr>
<td>2.</td>
<td>Name a responsible person to run the project (this person also is in charge of the governance meetings)</td>
</tr>
<tr>
<td>3.</td>
<td>Obtain contacts from all sub segments of the organization</td>
</tr>
<tr>
<td>4.</td>
<td>Ensure contacts understand and buy into rationale for corporate taxonomy structure (these are the champions for the sub segment of the organization)</td>
</tr>
<tr>
<td>5.</td>
<td>Taxonomy specialist runs through multiple simulations of taxonomy types (see table x for basic dictionary used by department B)</td>
</tr>
<tr>
<td>6.</td>
<td>Simulations occur until there are between six-12 high level classes</td>
</tr>
<tr>
<td>7.</td>
<td>Formulate sub classes if needed ensuring that the levels do not go further than three levels deep</td>
</tr>
<tr>
<td>8.</td>
<td>Test class structure once completed</td>
</tr>
<tr>
<td>9.</td>
<td>Review next steps (security, records management, data classification) to determine the next area of focus</td>
</tr>
</tbody>
</table>

c) Obtain Senior Management Support

As mentioned above, senior management support is crucial to ensuring buy in within the organization. If everyone in the organization is not collaborating on the effort, the taxonomy concept will fail. Per Gunnlaugsdottir (2012) the top three areas that influence a successful taxonomy are user input, training, and senior management support. Communication regarding a corporate taxonomy should also flow from the top management to ensure the organization understands that it is an organizational priority.

d) Name a Responsible Person

The person that is named to run the taxonomy project should have a background in document management, have a clear understanding of organizational standards, and have a background in Information technology (IT). The responsible person should also understand database management that will assist in understanding data structures in the organization. Having a solid background in project management will also assist with the implementation plans and coordination activities. The taxonomy specialist will be running the governance meetings as well as meeting with other organizational contacts that influence the integrations for taxonomy management, such as records management specialists and corporate committees that create standards. The taxonomy specialist handles interoperability that interconnects with end user informational needs. Per Verlag (2011), there are multiple components to ensure the taxonomy is running smoothly across the organization and having someone specifically running the taxonomy project will ensure all organizational units are represented. It is also vital that the responsible person has the authority to make decisions within the organization.

e) Obtain Contacts

The taxonomy contacts should be members of the existing organizations and have background experience with the data within the organization. The contact should be the person able to make decisions in the organization and have great communication skills as this person will handle communication within the subgroup. The contacts should be able to commit themselves to the project and ensure the subgroup has representation at all meetings. A separate change manager should also be in attendance to assist with the success of the project.

f) Ensure Contacts Understand the Process

The contacts are going to be the spokespeople for the process. Therefore, it is vital that they understand the process and have a working vocabulary of taxonomy terms. The simulations should not occur without obtaining all members buy in and support on the process. Having a change manager present will assist with the implementation process as well. Having a workshop to explain the benefits of taxonomy as well as the challenges of implementing a taxonomy is an important aspect of the learning process. This knowledge transfer assists in the understanding of why the taxonomy is important and increases buy-in from the team members. Appropriate training is vital to the success of the taxonomy implementation (Gunnlaugsdottir, 2012).

g) Perform Simulations Until High-Level Structure Emerges

Once all members have a basic knowledge of taxonomy and understand the organizational benefits. Simulations occur when end users present documents in a group setting and everyone classifies the documents. There are multiple ways to perform the simulations. An open forum occurs when all users show and review the documents and judge the documents based on their perceptions. A closed forum occurs when users vote on what they think each document should be. A mix of these procedures can also occur. The taxonomy specialist is in charge of running these
simulations. Pincher (2010) states that obtaining a high-level taxonomy structure is the key to understanding and user adoption of the class structure.

h) Create Sub-Classes

During the simulations, the taxonomy specialist handles running the meetings and continuously voicing the rule of six to 12 top level classes and two to three subclasses. Consensus should dictate the classes. All classes should be generic in nature to fit all aspects of the organization. In an event where participants will not agree, then the taxonomy specialist has the deciding vote. Pincher (2010) encourages organizations to leave the sub-classes at a high level to ensure a high-level structure that is viable within the entire organization.

i) Test Class Structure

Once the class structure is complete, it is important to complete more simulations. Does everyone agree that certain documents fit into certain classes? If not, then it is important that a consensus or understanding is achieved prior to completing the class structure exercises. In this step, it is also important to define terminology for the classes. For example, if one of the high-level classes is a procedure, define procedures. If there are subclasses under the procedure, ensure the high-level class definition makes sense with the lower class structures. Validating the potential class structure is another important way to obtain buy-in from the group (Pincher, 2010).

j) Review Next Steps

To continue the momentum of taxonomy project, it is of great importance to start the project work of determining system alignment. The taxonomy specialist will meet with each of the contacts to determine the systems of impact and how to implement the taxonomy in each system, determine if the system needs to be integrated into another system, or some other method of implementation. Since records management, security, and other mechanisms may be department-centric, these facets can be interwoven into other projects as they emerge. The taxonomy specialist will be a key role in organizational data security measures and information analytics within the organization.

k) Implications

There are multiple aspects of the study to take into consideration when reviewing the best method for implementing a corporate taxonomy. The steps in this paper describe an overall high-level process of implementation. As every organization differs in structure, the method to deploy a corporate taxonomy should fit the specific needs of the enterprise. The grounded theory study is formed from interviews and follow-up conversations with five taxonomy experts within one organization within two different departments. Therefore, the participants were limited to the study. It will be challenging to find multiple taxonomy experts within one organization as it is a unique skillset to encounter within corporations.

There is a need for additional research on the best method to implement a corporate taxonomy to obtain some common ground for practitioners. Understanding how organizations manage unstructured data would also be a benefit to the current foundational literature on the corporate taxonomy subject. Also, organizations that are currently implementing a corporate taxonomy should compare and contrast the method of implementation against the method above to determine if additional insight can be added to the body of research.

XII. Conclusion

In conclusion, the research directly relates to Lewin’s change management theory as the study results show that change management is vital in ensuring organizational implementation success. The planning aspect of Lewin’s change management theory is especially dominant in the grounded theory study above. Multiple participants stated that planning for the implementation and ensuring all parties are a part of the project is the only way to achieve success. Planning is especially important with something as wide-scope as a corporate taxonomy that impacts the entire organization. Martec’s Law is also prominent in the research above as technology is changing at such rapid rates it difficult for organizations to work on foundational data projects while attempting to maintain the current work progress.

The article presented a grounded theory study that reviewed two separate taxonomy structures within one organization based on the timeframe and organizational needs. Multiple similarities and differences between the two department’s taxonomy were present to provide background information. The outcome of the study presented major themes such as end user concerns, workflow management and how to be successful, benefits of taxonomy, issues with taxonomy implementations, and how to ensure a successful corporate taxonomy implementation. In the discussion section, a specific procedure is available which presents an optimal solution to implement a corporate taxonomy. Therefore, the article answers the primary purpose of developing a methodology to follow while implementing the corporate taxonomy in organizations.


