Change in a Best Practices Ontology

Daniel E. O'Leary

Marshall School of Business
University of Southern California
Los Angeles, CA 90089-0441
Email: oleary@usc.edu

Abstract

Ontologies can provide a consistent basis for knowledge representation that allows communication between both humans and machines. Best practices ontologies provide a means of facilitating communication about best practices. Further, best practices ontologies can be used to capture and preserve knowledge about best practices to facilitate comparison between different organizational best practices. The focus of this paper is on change of best practice knowledge bases over time. In particular, the focus is on the change of the ontology used to structure the best practice knowledge.

Keywords

Knowledge management, ontologies, best practices

1. INTRODUCTION

Best practices are ways that firms perform particular processes. Increasingly, firms are developing "best practices" knowledge bases as part of their knowledge management systems. Best practices (or leading practices) knowledge bases provide access to knowledge about enterprise processes that appear to define the best ways of doing things. Because of their importance, firms capture those best practices in knowledge bases, indexed using "taxonomies" or "ontologies" (e.g., Gruber 1993). Those best practices are likely to change because business changes and business finds new ways to do business. Since best practices change, it is not unusual to expect that the ontologies used to index those best practices also need to change over time as best practices and business changes. Unfortunately, ontology change has been characterized as one of the primary impediments of the use ontologies (e.g. O'Leary 1997). However, there has been little analysis of the change of knowledge bases, such as these.

1.1 Purpose of this Paper

As a result, the purpose of this paper is two fold. First, the paper initiates an analysis of some of the changes of a best practices ontology to determine some of the ways that an ontology has changed. Second, this paper begins to provide a basis to understand how to analyze change in an ontology. Analyses such as these are critical to understanding how knowledge and knowledge bases change over time, and how to analyze that change.

1.2 Plan of this Paper

This paper proceeds as follows. Section 2 presents the background literature on ontologies, taxonomies, the value chain and best practices knowledge. Section 3 provides a brief review of the "process classification framework," a taxonomy/ontology for enterprise best practices. Section 4 discusses the data used in this paper, including Porter's value chain and different versions of the process classification framework. At the "top level," the best practices ontology analyzed in this paper is based on Porter's value chain. As a result, section 5 investigates change that takes place at the most basic level, going from Porter's version of the value chain to the initial version of the top level of the process classification framework. Section 6 analyzes the change in the process classification framework associated with change from its initial version to a more recent version. Section 7 investigates some extensions, while section 8 summaries the paper.

2. BACKGROUND LITERATURE

The purpose of this section is to briefly summarize background information about ontologies and taxonomies, Porter's value chain, and to review best practices knowledge, each of which is used or discussed later in the paper.

2.1 Ontologies and Taxonomies

Taxonomies have a long history in knowledge management and artificial intelligence (e.g., Chandrasekaran 1983). In addition, taxonomies have played an important role in research on best practices, by providing a basis on which to categorize best practices. Researchers such as Malone et al. (1999) developed a taxonomy for best practices in order to demonstrate " … the basic feasibility of organizing large numbers of activities in a unified specialization hierarchy."

An *ontology* is an explicit specification of a conceptualization (Gruber, 1993). It is a knowledge-based specification that typically describes a taxonomy of tasks that define the knowledge. Ontologies are specifications of discourse between multiple agents in the form of a shared vocabulary. Within artificial intelligence, ontologies are necessary for multiple independent computing agents to communicate without ambiguity, and are the center of much research on reusability of knowledge bases. Ontologies can include taxonomies depending on the particular "specification." In addition to structure, such as that provided in taxonomies, ontologies can include detailed statements of vocabularies.

2.2 Porter's Value Chain

The value chain has been a well-accepted economic concept. As noted by Porter (1985)

> *Every firm is a collection of activities that are performed to design, produce, market, deliver and support its product. All of these activities can be represented using a value chain. .... A firm's value chain and the way it performs those activities are a reflection of its history, its strategy, its approach to implementing its strategy and the underlying economics of the activities themselves.*

As part of the value chain concept Porter (1985) provided some generic descriptions of the wide range of processes that make up the value chain. As we will see below, the value chain provides a model for indexing processes, ultimately becoming part of an ontology for best practices. An illustration of Porter's value chain is given in figure 1.

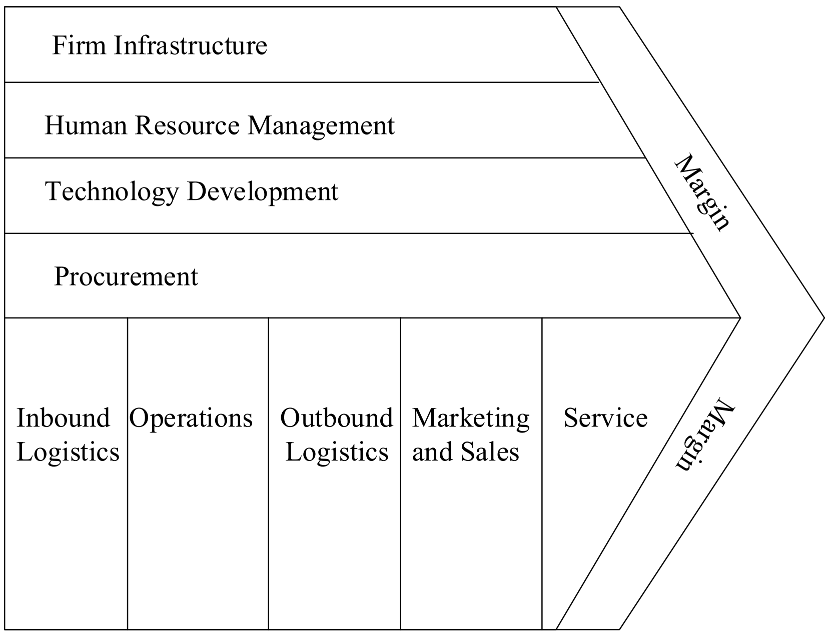

Figure 1: Porter's Value Chain

2.3 What is in Best Practice Knowledge Bases?

Best practice knowledge bases include a range of materials. Typically they include text and or graphic representation of best practice processes. Best practice processes may be generic or designed for specific industries. There may be reference to articles or other descriptions of the processes. Process measurements are also summarized providing a basis for benchmarking. Some best practices knowledge bases include reference materials such as so-called war stories, case studies, and technology enabler information. Finally, the knowledge base may have reference to particular experts on the processes.

3. PROCESS CLASSIFICATION FRAMEWORK

One of the best known ontologies for organizing best practices is the American Productivity and Quality Center's (APQC) Process Classification Framework (APQC 1996).

Arthur Andersen and the American Productivity and Product Center (APQC) developed the "Process Classification Framework" (PCF) starting in 1992, culminating with a published version in 1996. A top-level representation of the framework is seen in figure 2. The PCF was used as the basis to organize the best practices knowledge.

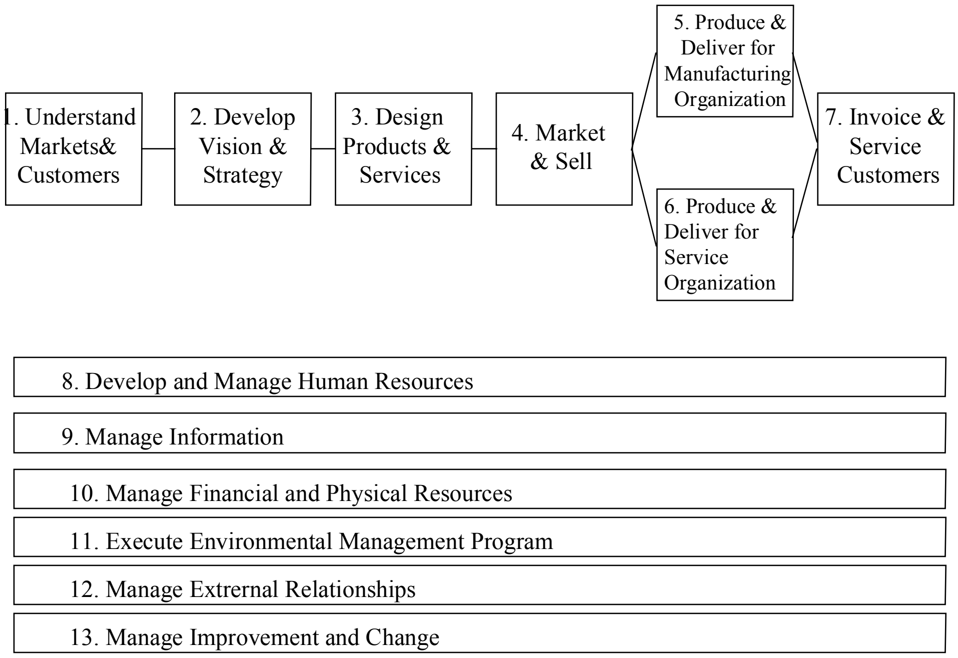

Figure 2: APQC/Arthur Andersen/PWC Best Practices Framework

> *"The process classification framework is the primary framework that Global Best Practices uses to organize best practice tools and information -- it is, in essence, the table of contents for the knowledge base."*
>
> *(http://www.globalbestpractices.com/SiteDocs/default.asp?navid=13)*

The ontology has multiple levels of related concepts. In most of the ontology there are three levels, starting with broad concepts and drilling down to greater detail, e.g.,

3. Design Products and Services

    3.1 Develop new product/service concept and plans

        3.1.1 Translate customer wants and needs into product and/or service requirements.

Best practices and metrics can be stored at each level of the framework, with the framework providing the indexing necessary to categorize the best practices. In addition, summaries of the best practices and case studies can also be indexed the same way.

After its development, Arthur Andersen used the framework as a basis for providing an Internet-based knowledge base of best practices. PriceWaterhouseCoopers (PWC) took over the Internet presence (http://www.globalbestpractices.com/Best_practices/) after the demise of Arthur Andersen. However, that framework is different than the original framework developed by APQC and Arthur Andersen, as seen in figure 3 (PWC 2003). The differences between the Internet version and the original version provide some of the data on which this paper is based.

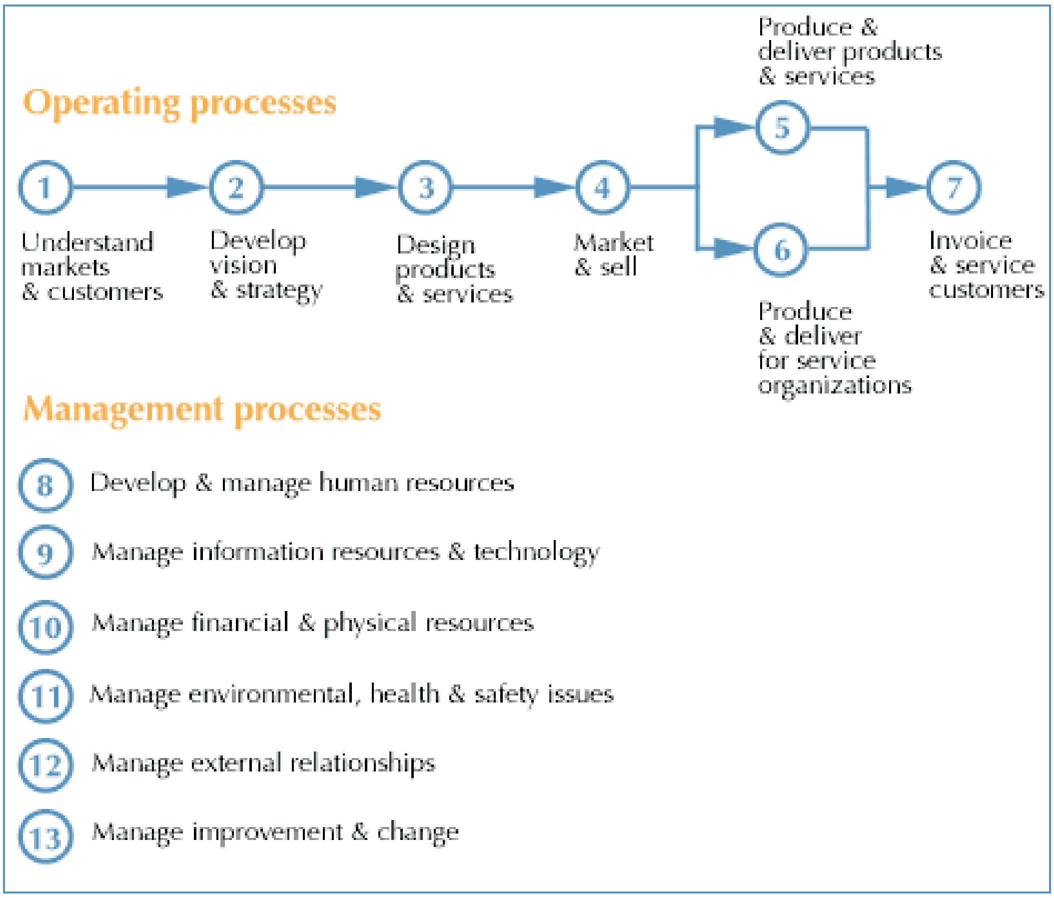

Figure 3: PWC Version

4. DATA

The data for this analysis comes from three sources. First, the APQC study structure is based on Porter's (1985) value chain. Second, APQC developed a revised version of the value chain that they used as the basis for development of the process classification framework. Third, the process classification framework was initially presented in 1996. Fourth, the process classification framework has been updated over time, and the current version of that framework is used as the basis of comparative analysis.

Thus, there are two different sets of comparisons that can be made. We can compare the Porter framework to the initial process classification framework. We also can compare the first process classification framework to the more recent process classification framework.

As a result, the data is qualitative. However, we can still do qualitative analysis of that qualitative data. Because of space limitations, this paper concentrates on counting and comparing different characteristics. However, this data can be analyzed using additional statistical and other approaches.

5. ANALYSIS OF TOP LEVEL DIAGRAMS: PORTER TO INITIAL APQC

In order to understand the initial change associated with the development of the APQC framework, the change from Porter to the APQC version will be analyzed first. As can be seen in an analysis of figures 1 and 2, APQC's structure is different than Porter's value chain. When we compare APQC's best practice knowledge base structures and Porter's value chain, although there are many similarities to each other there also are differences.

Porter's value chain (figure 1) partitions an organization into nine processes. However, the process classification framework (figure 2) partitions an organization into thirteen processes. Porter's value chain differentiates between **primary** activities (e.g, operations, service and marketing and sales) and **support** activities (e.g., human resources). Similarly, the process classification framework differentiates between **operating** processes (e.g., market and sell), and **management and support** processes (e.g., develop and manage human resources).

5.1 Primary Activities

The basic gathering of activities is similar in both. However, Porter puts "primary activities" (Inbound logistics, etc.) on the bottom, and "support activities" on the top, while APQC puts them on the bottom.

Porter uses "Inbound Logistics" and "Outbound Logistics," while APQC uses "Produce and Deliver." APQC does not treat distribution separately.

Porter uses "Operations," while APQC uses "Design Products and Services" and "Produce & Deliver." APQC also expands the notion of operations to include design.

Porter uses "Marketing and Sales," while APQC uses "Perform Marketing and Sales."

Porter uses "Service," while APQC uses "Invoice and Service Customers." APQC explicitly includes an artifact (invoices) in the category definition.

APQC includes "Develop Vision and Strategy" as a primary function, but that function is not represented in the Porter.

5.2 Support Activities.

Porter puts the support activities at the top, while APQC put them on the bottom of their representations.

Porter lists "Firm Infrastructure" as a support function. APQC includes a number of such general infrastructure categories, including, "Manage Information," "Manage Financial and Physical Resources," "Execute Environmental Management Program," "Manage External Relationships," and "Manage Improvement and Change."

Porter uses "Human Resource Management" as a support function, while APQC lists "Develop and Manage Human Resources."

Porter includes "Technology Development." It is not clear where APQC would include this category.

Porter lists "Procurement" as a category, while APQC does not include Procurement as a category, and it is not clear to the author where this function would fall in its current configuration.

5.3 Summary

There are a number of differences between Porter's original value chain and the resulting APQC top level ontology for business best practices. In some cases those differences are just minor differences, but in other cases there are entirely new categories of knowledge.

6. ANALYSIS: CHANGE FROM ORIGINAL APQC VERSION TO CURRENT INTERNET VERSION

The process classification frameworks from the original (APQC 1996) and revised (PWC 2003) version to provide the data used in this section.

In 1996, there were 275 categories and subcategories. There were 13 top-level categories, 71 second level concepts, 181 third level concepts and two fourth level concepts.

In 2003, there were 217 categories and subcategories. There were 13 top-level categories, 60 second-level subcategories and 144 third-level categories in 2003.

In general, there was a substantial reduction in the number of categories, with 58 fewer categories overall. There were 11 fewer second level concepts, 37 third level concepts and no fourth level concepts in the revised version.

6.1 Approach

For each category and subcategory of the revised 2003 version, the corresponding category from 1996, was matched. In a few cases this meant mixing second level in one version and third level with another version. The similarity of the categories was categorized according to one of the following:

Identical: Wording used was identical

Very similar: Wording is virtually the identical, but there are additional descriptors, e.g., "Assess technology innovations" vs. "Assess *new* technology innovations"

Similar: Basic concept is roughly the same, e.g., "Monitor competitive offerings," versus, "Determine customer reactions to competitive offerings."

No matching concept from old to new: No concept in the old framework that matches the new framework.

No matching concept from new to old: No concept in the new framework that matches the old framework.

What was the nature of the matches that were not identical? The "Very Similar" matches would include or exclude a word that would not change the meaning. For example, the original category "2.1.4 Assess new technology innovations," morphed into "2.1.4 Assess technology innovations." The very similar matches also included changes where one word is replaced by another very similar word. For example, the original category "1.3.1 Determine the weaknesses of product/service offerings," morphed into "1.3.1 Determine deficiency of products and services." Further, the "Very Similar" category also included aggregations or dis-aggregations. For example, the original "10.5.1 Ensure Tax Compliance," was broken down into "10.5.1 Manage National Tax Compliance," "10.5.2 Manage Local Tax Compliance," and "10.5.3 Manage International Tax Compliance." For the 15 changes rated "very similar," 4 included a word change, 9 included additional or fewer words and 2 were aggregations or dis-aggregations.

Two categories of non-matches were generated. Those concepts that were only in the old framework were captured as "old only." Those concepts that were only in the new taxonomy were captured as "new only." Redundancies (aggregations and dis- aggregations) from old to new and new to old were also captured "More than one New," and "More than one Old."

The resulting data was analyzed and summarized in a number of different ways (for example by category, changes and matches).

6.2 Concept Level

The results were analyzed by the three major levels of concepts in table 1. There was little change at the top level (Level 1) concepts, with 100% matches and 77% were identical. However, as the concept level decreased from top level to second level to the third level, the percent of identical matches went from 77% to 67% to 50%. In addition, as the concept level went from top to second to third level, the percent of matches went from 100% to 83% to 68%. Accordingly, the results suggest that the more detailed level of the concept, the fewer the percentage matches and the greater likelihood of a change, as compared to higher-level concepts.

	Top Concept	2nd-Level Concept	3rd-Level Concept
Identical	10	40	71
Very similar	2	3	9
Similar	1	7	16
No Match (New Only)	0	10	46
No Match (Old Only)	0	21	87
More than one New	0	0	3
More than one Old	0	0	27
Total *	13	60	142
Percent Identical	0.77	0.67	0.50
Percent Match	1.00	0.83	0.68

* Totals do not match exactly because of one aggregation

Table 1: Quantitative Analysis by Level

6.3 Category Matches

The matches by category are summarized in table 2. Three of the categories had 100% matches, and another 2 had greater than 90% matches. Only three categories had less than 50% matches. Roughly 80% of the "operating processes" categories (1-7) matched, but only 65% of the "management & support processes" categories (8-13) matched. Percent matches are not homogeneous across category.

Category	Identical	Very Similar	Similar	Total Matches	Total	Percent
1	7	3	2	12	12	100.00%
2	10	3	0	13	19	68.42%
3	15	1	1	17	22	77.27%
4	4	0	3	7	15	46.67%
5	13	2	5	20	22	90.91%
6	11	0	1	12	13	92.31%
7	9	0	1	10	10	100.00%
8	10	1	5	16	21	76.19%
9	0	2	1	3	12	25.00%
10	21	1	2	24	31	77.42%
11	0	0	1	1	15	6.67%
12	7	1	0	8	9	88.89%
13	14	0	2	16	16	100.00%
	121	14	24	159	217	

Table 2: Matches by Category

6.4 Statistical Analysis of Proportions of Matches

The data in table 2 can be analyzed using a test of difference in proportions (e.g., Dixon and Massey 1969). The results are summarized in table 3.

Specifically, some of the proportions are statistically significantly different than other proportions, further substantiating the notion of greater change in some portions of the knowledge than in others. For example, the z statistic resulting between the proportion of the first category that stayed the same (100%) and the proportion of the eleventh category that stayed the same (6.67%) is 4.82. The z statistic refers to the cumulative normal distribution.

	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	2.16											
3	1.78	0.63										
4	3.01	1.28	1.91									
5	1.07	1.81	1.23	2.98								
6	0.98	1.61	1.14	2.57	0.14							
7	N/A	1.99	1.64	2.80	0.98	0.89						
8	1.84	0.54	0.08	1.81	1.31	1.19	1.68					
9	3.79	2.36	2.96	1.16	3.93	3.43	3.56	2.86				
10	1.79	0.71	0.01	2.08	1.29	1.17	1.65	0.11	3.19			
11	4.82	3.63	4.22	2.48	5.08	4.53	4.61	4.12	1.33	4.51		
12	1.18	1.17	0.74	2.07	0.17	0.27	1.08	0.79	2.91	0.75	4.02	
13	N/A	2.47	2.04	3.39	1.24	1.13	N/A	2.09	4.21	2.06	5.21	1.36

Table 3: z Statistics for Test of Proportions of Matches for Table 2

6.5 Category Changes

The category changes are presented in tables 4 - 6. There were roughly twice as many old categories without matches (114) as new categories without matches (56), as the number of categories shrank over time. Old categories were eliminated, aggregated and disaggregated.

Category	New Only L-2	New Only L-3	Total New
1	0	0	0
2	0	6	6
3	1	4	5
4	0	8	8
5	0	2	2
6	0	1	1
7	0	0	0
8	1	4	5
9	2	7	9
10	3	2	5
11	3	11	14
12	0	1	1
13	0	0	0
	10	46	56

Table 3: New Categories

As seen in tables 3 and 4, the amount of change at lower levels exceeds the amount of change at higher levels. For example, the greatest number of changes occurred at level 3 while the fewest number of changes occurred at level 1.

Category	Old Only L-2	Old Only L-3	Old Multiples	Old Only L-2+3	All Old
1	0	0	4	0	4
2	0	3	0	3	3
3	0	3	4	3	7
4	0	4	2	4	6
5	0	3	3	3	6
6	0	2	0	2	2
7	0	1	0	1	1
8	3	14	12	17	29
9	7	23	2	30	32
10	0	6	0	6	6
11	10	0	0	10	10
12	0	0	0	0	0
13	1	7	0	8	8
	21	66	27	87	114

Table 4: Old Categories Not Used

Further, as seen in table 5, I found that there were few changes in some parts of the taxonomy, and huge changes in other parts of the taxonomy. The largest number of changes occurred in categories 8, 9 and 11, with 20%, 24.12% and 14.12% of the number of changes occurring in those areas. On the other hand, categories 1, 5, 6, 7, 12 and 13 had less than 5% changes. The non-matches are not homogeneously spread across the 13 categories. Instead, change in knowledge management categories occurs unevenly, with change centered in a few chunks of knowledge.

Category	Total New	Total Old	Total Changes	Percentage of Change
1	0	4	4	2.35%
2	6	3	9	5.29%
3	5	7	12	7.06%
4	8	6	14	8.24%
5	2	6	8	4.71%
6	1	2	3	1.76%
7	0	1	1	0.59%
8	5	29	34	20.00%
9	9	32	41	24.12%
10	5	6	11	6.47%
11	14	10	24	14.12%
12	1	0	1	0.59%
13	0	8	8	4.71%

Table 5: Percentage of Total Change by Category

6.6 Operating Process vs. Management and Support Processes

A list of the total changes is summarized in table 5, by categories. An analysis of those changes indicates that, the changes vary by operating processes versus management processes. In particular, 30% of the change occurred in the seven operating processes (Categories 1-7), whereas 70% of the change occurred in the six management processes (Categories 8-13).

As a result, we might argue that a firm's "operating processes" have been more stable than "management and support processes." For example, changes have had a major impact on information technology over the period 1996 to 2003. Technology, including the Internet, e-business and e-commerce have become routine and integrated into many leading businesses. As a result, the change of the business process in those areas has been substantial.

7. EXTENSIONS

The results of this paper can be extended in many directions. First, although this paper has initiated some statistical analysis, various other statistical tools, such as correlation analysis, regression and entropy analysis could be used to analyze the data in greater detail in order to better understand the underlying changes that are taking place in the data. Second, different characteristics of the knowledge base (e.g., number of metrics in the original version), and outside information, such as information from the World Wide Web, could be brought in to do further analysis. Third, other characteristics of the process classification framework, such as the population of the categories with materials, could be further analyzed.

8. SUMMARY

Best practices ontologies provide a basis to organize information and knowledge about best practices in order to facilitate communication and benchmarking of business processes. APQC, in conjunction with Arthur

Andersen was one of the first to develop a detailed framework for best practices. There has been substantial change over time in the APQC framework. This paper provides a first step at trying to understand how to analyse that change. In addition, it provides insights into the nature of change to that framework.

When APQC developed the initial version of the process classification framework, they changed the top level from Porter's value chain to their own version. In so doing, categories were added, and dropped, while other categories were renamed.

Similarly, as the APQC framework changed over time, there were other changes in the framework. There were qualitative changes to the framework that took multiple basic forms: old categories were dropped; old categories were aggregated; old categories were disaggregated; old categories were rewritten, and new categories were added.

As part of the analysis, I found that change to knowledge is uneven. Statistically, there was greater change in some categories and less change in other categories. In addition, I found that there was greater change in the broad grouping "management and support processes," as compared to the other broad grouping "operating processes."

This paper is a first step in the analysis of knowledge, such as the process classification framework. Additional analysis could be done to the data generated at this step to provide deeper insight into the data. For example, additional statistical analysis of this data could be done to provide additional insights.

REFERENCES

APQC - American Productivity & Quality Center, (1997), Arthur Andersen.

Chandrasekaran, B. (1983). Towards a taxonomy of problem solving types. AI Magazine 4 (1) 9–17.

Dixon, W. and Massey, F., (1969) Introduction to Statistical Analysis, McGraw – Hill, New York.

Gruber, T., (1993) A Translational Approach to Portable Ontologies, Knowledge Acquisition, *5*, 2, pp. 199-220.

Malone, T., Crowston, K., Lee, J., Pentland, B., Dellarocas, C., Wyner, G., Quimby, J., Osborn, C., Bernstein, A., Herman, G., Klein, M., O'Donnell, E., (1999). "Tools for Inventing Organizations: Toward a Handbook of Organizational Processes," Management Science, Vol. 45, No. 3, March, pp. 425-443.

O'Leary, D.E.,(1997). "Impediments in the use of explicit ontologies for KBS development**,"** International Journal of Human Computer Studies, volume 46, pp. 327-337.

Porter, M., (1985) Competitive Advantage, Free Press, New York.

PWC (2003), PriceWaterhouseCoopers , Global Best Practices, http://www.globalbestpractices.com/Best_practices/default.asp?N=S,273,13.02,4751,#13.02

COPYRIGHT