



A Practitioner's Perspective

Alec Sharp

Senior Consultant
Clariteq Systems Consulting

asharp@clariteq.com

Some Thoughts on Process Discovery

“Process discovery” means different things to different BPM practitioners. In this Column I’ll illustrate that point, describe the negative consequences of failing at “process discovery” as I define it, and then offer some advice for your next process discovery initiative.

Lesson #1 – Terminology is everything

Anyone who has suffered through one of my educational workshops knows that I emphasize starting *any* project by developing a glossary for the business areas involved. This might be disguised as a data model, but that’s consistent, because the most important function of conceptual data modeling is to agree on terms and definitions, then capture and communicate them.

And of course, like all of you, I practice what I preach. On any project, whether it’s process improvement, requirements definition, application assessment, or whatever, I invest up-front time in clarifying terminology. Most people who have been in this business as long as I have (30+ years) do the same, having been burned when we *didn’t* clarify terminology.

Sometimes, though, outside of a project or classroom environment, I’ll get into some situation or other without stopping to ensure that everyone’s talking the same language. Predictably, misunderstanding, confusion, and general unhappiness ensue. Let’s consider an example from the BPM field.

An illustration

Last year I was contacted by a business development consultant doing market assessment for a tool that would perform automated process discovery. “Automated process discovery” struck me as a dubious idea, but then, dubious ideas are common in “toolspace.” Back in the days of CASE (Computer Aided Software Engineering—“programming without programmers” and all that) we used the phrase “tools for fools.” That still comes to mind when I encounter a team agonizing over the right modeling tool before they know how or what to model, or selecting a <<insert topic here>> repository that is destined to become write-only memory. But I digress. Returning to the market assessment...

As I understood it, the proposed tool would analyze system event logs and not just discover the organization’s business processes, but create process models in some form like BPMN. The expected benefits were faster discovery, standardized results, and elimination of the messy business of getting people together for a discovery workshop. I just couldn’t get my head around the idea that a tool could be turned loose on massive amounts of log file data from all of an organization’s systems, and somehow it would figure out what the business processes were. And what about all the human steps and interactions, or the parts of the process that lived outside the firewall? Besides, process discovery sessions might get messy, but I always found there were major downstream gains from the credibility and buy-in developed during the session—participants invariably learn, first hand, that their processes are larger and more involved than they imagined, and it heads off all those conversations later on that begin “well, those aren’t *really* our processes...”

I gave my feedback to the consultant, but they were persuasive, and sent me some colorful

literature extolling the product's virtues, mostly in 4 point type that I struggled with. "Perhaps a demo would help me see the light?" I suggested, but that was a non-starter because some steps were still done manually using a proprietary methodology. That could have been a red flag, but I was actually happy to hear they had a methodology before starting to code, so I reviewed their literature with some BPM colleagues who were more technology-savvy than me. They felt the same way I did—the technology was unlikely to work other than with carefully selected, "school project" sized datasets. They also agreed that if it's at all possible, there's great value in getting people together for process discovery. One even came up with a study showing that with no process change whatsoever, there were measurable performance improvements after "getting the gang together" to discuss a process. Again, I reported back to the biz dev consultant, and again, it wasn't the news they wanted to hear. I had racked up several pro bono hours, and they weren't getting the kind of assurances that would help them line up capital, so we thanked each other for an interesting exercise and left it at that. Throughout this exercise, I had the nagging feeling that we were somehow talking past each other.

Another, but shorter, illustration

At around the same time, an article made the rounds in the business process community suggesting that during process discovery the focus shouldn't be on the so-called "happy path" but on the problematic cases—the ones that cause grief to the people who work the process. Difficult customers, impossible orders, common failure situations, and so on. "What?!", I thought, "where do these goofy ideas about process discovery come from?" Again, I just couldn't get my head around the idea that anyone who had actually helped an organization discover its business processes could recommend *starting* with all the problem cases. I *get* there, eventually, but have learned the hard way not to *start* there because it leads to confusion, frustration, and a "can't see the forest for the trees" situation.

A revelation!

Fast forward several months, and I'm in Brussels to keynote at Vlerick's 2008 Business Process Innovation Conference. (An excellent conference, by the way.) During a free afternoon I attended a workshop by Prof. Guido Dedene (K. U. Leuven) and Ed Peters (CEO, OpenConnect) on (you saw it coming) "Business Process Discovery, Analytics, and Improvement." The presentation was fascinating, and a revelation with respect to this Column's topic because they had actual examples (what a concept!) of what they meant by "process discovery." I was a bit shocked when I realized what they called "process discovery" was what I called "process mining" or "process analytics." Because I'd been using it for over fifteen years, I was stuck on *my* meaning for "process discovery", which refers to an initial phase in which we determine *what* an organization's processes (or a subset) are, not all the variations on *how* they behave. This isn't my exact wording, but I asked the presenters "Isn't there some necessary "pre-discovery" work to (a) identify the processes that might eventually be studied, and (b) map them so it was known what activities and systems were involved?" The answer, of course, was "of course"—you can't just turn these tools loose on an organization's entire dataset of log files and hope to get something useful within your natural lifetime.

Naturally, this got me thinking about the process discovery described in those earlier illustrations. "Automated process discovery" suddenly was quite reasonable—after all, that's what Dedene and Peters were demonstrating. "Process discovery" by looking at problem cases made a whole lot of sense, and paralleled my usual advice to study specific scenarios. I worried I might have been alone in misinterpreting "process discovery," but was relieved when a few minutes of research uncovered three variations on the use of the phrase, with a fourth showing up just this week in a BPM discussion group. The four variations on process discovery were:

1. As I use it, meaning the initial identification of *what* the processes are, producing a list or simple block diagram;
2. Building a visual representation of process flow (e.g., a swimlane diagram) which is most commonly called "process mapping" or "process modeling;"

3. Tracing or enumerating all the actual execution paths through a process in order to identify the factors that contribute to successful or problematic cases;
4. A more general assessment of what's right or wrong with a process.

The point of these illustrations

All of these are perfectly sensible uses of the term “discovery”—I’m not quibbling that one is more appropriate than others, just that they’re different, and it’s unlikely I’m the only person who has been tripped up by this. The really important point, though, is that the first variation in my list (the one I call process discovery, in which processes are first identified) is seldom even mentioned, much less written about or discussed. I wonder if the thinking is that identifying an organization’s business processes is so straightforward and self-evident that doing it doesn’t even warrant a name—everyone “just knows” what the processes are. My experience is that nothing could be further from the truth, and the number one root cause I encounter when called in to sort out a troubled business process project is poorly defined processes. I’ll get on to that, but first I’d better, as we say in the consulting world, eat my own dog food and go over some of the terminology I’ll be using.

Phases and naming

The default methodology that I use on a business process improvement or redesign initiative includes seven, major activities organized into three phases, each with an overall goal:

I - Frame the process (establish context, scope, goals):

1. Discovery (or identification) of a family of related business processes (a “process area”)
2. Scoping of each process (trigger, results, major cases or variations, etc.)
3. Initial assessment and goal-setting for a selected process

II - Understand the current (as-is) process:

4. Mapping/modeling the as-is process
5. Final assessment of the as-is process (possibly including process mining / analytics)

III - Design the new (to-be) process

6. Characterization of the to-be process
7. Design of the to-be process

Of course, after this are phases for implementation and ongoing management of the process, but this is enough for our purposes.

Figure 1 illustrates the core results of process discovery. The block diagram at the top shows the six business processes that process discovery has identified within the process area “Administer Client Safety Management Programs.” At the bottom of the illustration, we see that five sub-processes comprising the “Grant CSM Program Authorization” program have been identified, which might be done either as part of activity #1 (process discovery) or activity #2 (process scoping.) Effectively, this is a three-level decomposition—a process area into business processes that are decomposed into sub-processes. Even though a decomposition is inherently a top-down depiction of an area, the odds are that the discovery was done in a bottom-up fashion, starting with details that might have been even more granular than the sub-processes. That was the case with this example. We’ll explore the use of bottom-up techniques later in the Column.

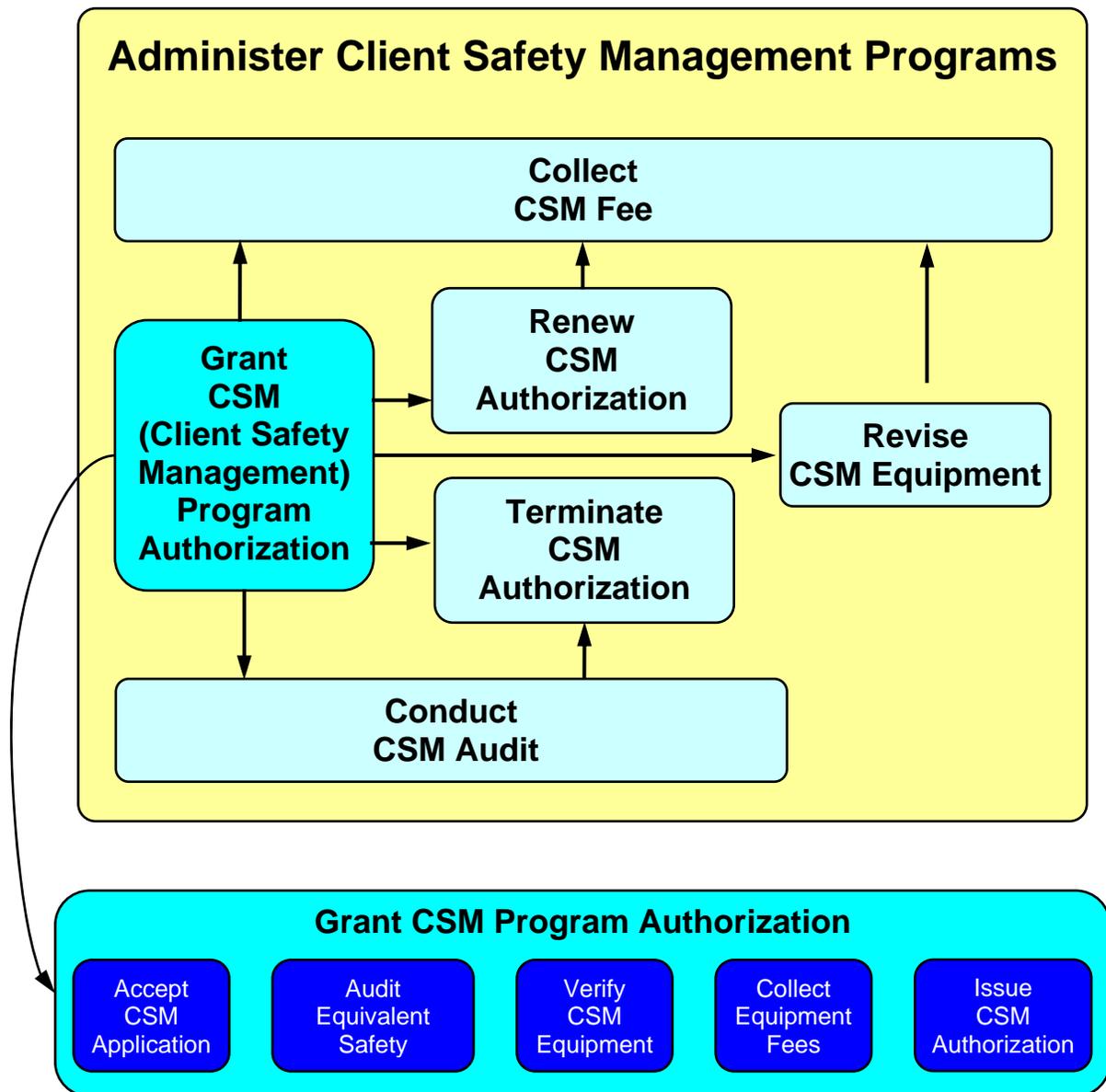


Figure 1: Core results of process discovery

Notes and observations

I wrote an entire book on those seven activities but a few points in particular are relevant to this discussion:

- The result of process discovery is a pure “what, not who or how” identification of business processes. There is no indication of which parts of the organization are involved, how the process is implemented, or whether it is currently done well or poorly; the purpose is purely to establish “what we do” (or should do.)

- “Process discovery,” when it is used to describe the automated enumeration and analysis of actual execution paths, is a technique that fits more or less into activity #5 (final assessment”) in this list. I’ll refer to it as “process mining / analytics” for the rest of this column.
- “Process identification” is another term I use for activity #1 “ but I don’t like it as much because it sounds as if all you have to do is look around and identify or name the processes. The reality is that it’s almost always more difficult than that, so I equate it with a voyage of exploration and discovery. I’ll combine these terms and refer to “process discovery/identification” for the rest of this Column.
- The key point here is that doing a good job of #5 (which might include process mining/analytics) requires that you first complete #1 through #4. What is striking is how little is said about the first couple of activities, especially #1, the supposition seeming to be that “everyone knows what our processes are—let’s get on with fixing them.”

Why process discovery/identification matters

Since the 1990s, a lot of my consulting work has involved troubled projects. (No, not *creating* troubled projects—helping figure out what went wrong and, if possible, getting things back on track.) You don’t have to do this for very long before you notice the same problems showing up again and again, including:

1. Inappropriate or missing sponsorship;
2. Unclear or conflicting business objectives;
3. Poorly defined scope.

When these arise on a business process project, the same *root cause* often underlies them—incorrectly defined processes.

Process identification problems: the usual suspects

Some common variations on incorrect process discovery/identification are:

1. Several “small” processes are identified and addressed independently when they are really each an activity that is part of a single, end-to-end business process.
2. A process is identified that is actually an organizational unit, such as “Sales” or “Logistics,” that participates in multiple business processes.
3. A process is identified that is actually a “process area” or “process domain,” such as “Customer Relationship Management,” that is made up of multiple business processes.

In the first variation, the unfortunate outcome is the “local optimization yielding global suboptimization” phenomenon that Eliyahu Goldratt warned us about. Optimizing each of these “small” processes often proceeds smoothly, because the difficulty of trying to balance different goals and objectives across the organization is avoided. However, when the work of a single “true” business process flows through these separately optimized processes, the activities clash and a business process that behaves *worse* than it did before is the result.

The second and third variations are very similar in outcome because multiple processes are involved, but they haven’t been explicitly identified and remain “hidden” within the larger organization or process area that is the project’s scope. When the team attempts mapping/modeling the current process, what should be the relatively straightforward tracing of a triggering event (starting point) through to its results (end points) becomes an exercise in frustration. Attempts go around and around in circles because there is no start and no end; instead there are multiple flows, all having different timing cycles and frequencies, with everything connected to everything else.

We’ll look at examples shortly, but first, a point should be addressed that may have occurred to you.

The common factor—process vs. organization

Each of the “usual suspects” I just listed exhibits the same factor that is surprisingly common in cases of poor process discovery/identification—business functions (or organizations) are confused with business processes. You’d think that after all these years of hearing about “end-to-end, cross-functional business processes” everyone in a modern enterprise would have internalized the idea. Not so! The majority of staff and management still don’t really “get” what a business process is, and therefore don’t understand what the big deal is or go off in the wrong direction. BPM experts frequently don’t help, because they assume everyone already understands, so they don’t cover the basics. Worse, because they haven’t taken the time to articulate what, precisely, they mean by “business process” they’re unaware that their own definition is so vague as to be of little practical value. The ever-popular “A business process is a linked set of activities that collectively delivers value etc. etc. etc.” applies to everything from a localized task through to an enterprise business area.

As a rough generalization, folks outside of the BPM field think their piece of the pie (“Confirm Prerequisites”) is a complete business process, which is too small, while BPM folks frequently talk about collections of processes (“Curriculum Management,”) which is too large.

At several BPM conferences I’ve presented “Getting Traction for ‘Process’ – What the Experts Forget.” The theme is that when organizations don’t get onboard with business processes, it’s frequently because resident and hired experts have forgotten the fundamentals. The content of the presentation is always changing, but one constant is that the first factor I look at is a lack of clarity on what a “business process” is. (By the way, I’ll be delivering a new version of this presentation at IRMUK’s BPM event in London this September.)

Example 1: Processes—too many, too small

Returning to examples of incorrect process discovery/identification, consider the case of a financial services firm that identified three business processes (among others) as illustrated in Figure 2. They were originally named Prospecting, Solicitation, and Qualification, but following the guideline that you don’t have a process if you don’t have an action verb and a noun, one of my first steps was working with the client to rename them. The fact that all three dealt with the same noun is an important clue to what emerged.



Figure 2: The original “business processes”

The client was trying to improve their business development efforts, including introducing a funnel concept in which each “process” would reduce the number of prospects that the next would have to deal with. The optimization, in isolation, of these three separate processes ended up working against that goal. “Identify Prospect” was expected to come up with as many prospects as possible, and “Solicit Prospect” was expected to make contact with each of these within a set timeframe. Qualify Prospect was expected to sift through any receptive prospects that had made it through solicitation, and then identify those that would most likely become profitable, long-term customers. It was a classic case of the performance targets of individual activities, especially the earlier ones, working against the ultimate goal of the process.

After reviewing some simple guidelines and examples, it was clear to the client that there was only one business process involved, “Acquire Customer.” The original three were some of its sub-processes, as illustrated in Figure 3. (I’ve only shown the relevant parts here.) One guideline: the same “token” (the prospect) flowed through all three, on a 1:1 basis. That is, one identified prospect might become one solicited prospect (1:1) and one solicited prospect might become one qualified prospect (1:1).

In the end, the sequence of the subprocesses was changed, and the performance targets revised to better support the end-to-end process of acquiring a customer. “Identify Prospect” was designed to identify high-quality prospects (not just high numbers of prospects,) and “Qualify Prospects” was designed to identify those that were a good strategic match with the firm. “Solicit Prospect” had a smaller number of better prospects to deal with, and better information, so they were able to spend more time developing their solicitation leading to a higher conversion rate.

It seems so obvious it’s hard to believe this actually happened, but it did, and does, all the time in process improvement efforts. That’s ironic when you consider that the modern era of business processes began when Michael Hammer popularized the idea with stories about “functional silos” and showed us that business processes were cross-functional.

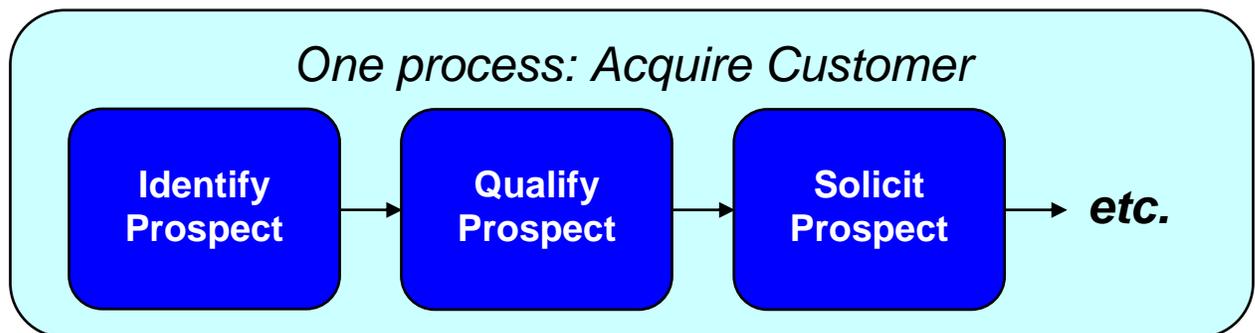


Figure 3: The actual business processes

Example 2: Processes—too few, too big

I used this example in my book, because it’s a classic example of what happens when a team tries to apply process modeling techniques to work that is more than a single process. Like the previous example, it’s heavily edited to fit.

The team for an important, enterprise project brought me in to conduct one of my courses on working with business processes, but using the team’s project as the case study. After going over business process concepts and examples, and introducing some methods, we got to work on identifying the business processes within the team’s scope. It went really well, and I was pleased with the five processes we identified until someone in the class said “But those aren’t our processes.”

I was puzzled, because I thought we had just discovered what their business processes were. Silly me! What emerged was that the team had already worked with a major consulting firm to identify the business processes within their scope, and mapping these processes had been underway for some time. It wasn’t going well at all, which turned out to be why my class was brought in (a little detail I’d like to have known about.)

Somewhat at a loss, I asked what these business processes were, and they listed four for me. From the work we’d just completed, it was instantly obvious that these weren’t processes, they were organizations, but I didn’t want to be quite that blunt. Instead I suggested we work with “their” processes for a while, including cross-referencing each to the four main departments within

their business function. I drew a matrix with each process occupying a row and each column being one of the departments, then asked which departments participated in each process. As Table 1 illustrates, Process 1 took place entirely within Department 1, Process 2 within Department 2, and so on.

	Department 1	Department 2	Department 3	Department 4
Process 1	X			
Process 2		X		
Process 3			X	
Process 3				X

Table 1: "Processes" vs. organization

As I'd hoped, someone in the class pointed out the obvious—they'd been working with *organizations*, not *business processes*—and perhaps this was why they'd had so much trouble. With this as an opening, I felt comfortable pointing out that their C-level executives were expecting the kind of transformational improvements that can only come from cross-functional processes, and theirs were anything but. Sadly, project management decided that it was all just semantics, and directed the team to "stay the course" with the previously identified processes.

The punch line, if you can call it that, came almost exactly a year later. I was working on another assignment at this company when I happened to walk past a conference room that some of the team were working in. "There he is!" I heard someone say, and the next thing I knew I'd been pulled into the room, the door was closed, and they were telling me the whole sad story. Mapping and analyzing their "processes" had become an exercise in frustration for the reasons I explained earlier in this column. After a year of struggle, project management saw the light, and the team was allowed to go back and start over using the real cross-functional business processes we had identified in class as the starting point. The costs of this lost year were tremendous in terms of lowered morale, wasted resources, and lost opportunities.

I think I could go on for a hundred pages with examples of problems caused when business process discovery/identification misses the mark, but this has gone on long enough, so let's wrap up with some points to keep in mind about this critical phase.

Key points for business process discovery/identification

Over time, successful practitioners evolve a set of frameworks and methods that work for them, so I don't expect each of these points to apply to each of you. I do hope you'll think about them, though, because some are just contrary enough to common advice that they might give you an idea for your own practices.

1 – Discover processes in group sessions

One approach to process discovery is to do some research, conduct some fact-finding interviews, then disappear into your office and emerge only after identifying the business processes. You might even be right, but you'd have missed an important opportunity to engage the people who work in the processes and develop their buy-in. When people participate collectively in process discovery/identification, they learn about each other's contributions, how they all affect one another as part of the same "process ecosystem," and take ownership of the identified processes. This is hugely important later on when you get into the difficult aspects of mapping, assessment, and redesign.

2 – Don't expect published reference models to eliminate the effort

It's tempting to skip the messy business of process discovery/identification using the rationale that "others have already done the work, so why reinvent the wheel?" We have seen so many problems from this approach that Roger Burlton and I are talking about writing an article on the subject. In the meantime, here are a few points:

1. Available reference models have a stronger functional orientation than you might expect, and the granularity can get overwhelming.
2. You don't get the same buy-in when you start with an off-the-shelf model.
3. The expected time savings don't materialize, for instance when the effort to adapt a reference model to your industry and organization is greater than the time it would have taken to build one from scratch. (This is akin to taking spending more to modify a purchased application to meet your needs than it would have taken to build your own.)

It might be self-evident, but a reference model is something you refer to, not slavishly adopt. That's how we use them—as a reference to learn about a new area, or as an after-the-fact checklist of activities to see if anything substantial has been missed

3 – Business processes need to be demonstrated and described, not just defined

Because everyone has their own ideas, I have a collection of examples that I use at the beginning of process discovery/identification sessions to illustrate what I mean by a "business process." I've tried to come up with a straightforward definition, but there are inevitably so many ways to interpret it that the only thing that works is to walk through real examples. Like the earlier Example 1, these demonstrations typically have three parts:

1. Show ~5 activities that are reasonable processes. That is, they do real work and deliver a necessary result;
2. Illustrate the problems that follow from optimizing them as separate processes;
3. Show how concrete guidelines reveal that they comprise a single business process. These guidelines include showing how they all deal with the same triggering event, and that the same work item ("token") is moving through the process, being transformed, until the final result is achieved.

Don't use an example from the business area you're working in, or people will want to go with it

4 – A bottom-up method is usually best

Given how much we read about the need for process change to be driven from the top, and the general advisability of top-down methods, recommending a bottom-up approach for process discovery seems odd, but is backed up by the evidence. In a top-down approach, high-level processes are suggested, and if they seem plausible and can be depicted on a one-pager, work proceeds on breaking them down into more granular processes. The problem is that those high-level processes are often business functions, product lines, market segments, or some combination of those, but not a cross-cutting business process. All the problems from confusing process and function that we discussed earlier can then be expected.

In the bottom-up approach I use, a group of business representatives participate in a discovery workshop at which they identify the business activities they participate in or are aware of. This usually starts with some activities that have arisen in pre-workshop interviews. We aim for activities that are *large* enough that completing one is an accomplishment the business might want to count, e.g. "A prospect has been identified" or "A prospect has been qualified." Activities

will typically be *granular* enough that individuals can point to one and say “I do that.” The group then organizes these activities into “chains” or “flows,” and employs guidelines to decide when one process ends and another begins. This is a lot easier when each activity is on a large Post-it.

Not only does this create buy-in, but you also finish with a lot of information about what’s inside of each business process. Once bottom-up discovery has been completed, work proceeds in a top-down fashion, which is a pattern we see repeated in other types of analysis.

Obviously, process discovery/identification is a topic I feel strongly about. I just checked, and two of the longest chapters in my book were chapter 3, “Business Processes: What Are They, Anyway?” and chapter 5, “Discover Business Processes.” This isn’t a theoretical interest—it’s based on many years of seeing what can go wrong in practice, and refining my methods accordingly. I hope this has given you some ideas for your own methods, or perhaps some ammunition to use when you encounter resistance.

Wrapping up

I closed my last Column by noting that “In future Columns, we’ll look at the specifics of modeling business processes at the scope, concept, and specification levels.” That’s still true—process discovery/identification is the essence of scope-level process modeling, and now that the importance has been addressed, we can expect some details and examples.

Enjoy the rest of the summer, or whatever season it is in your part of the world.

From the Trenches
Alec Sharp