

A Data Quality Community of Practice

Donnalee Brown, SAS® (Remote Office), Poulsbo, WA

Abstract

Communities of Practice (CoP) are a means by which disparate companies, departments, divisions, and functional areas can share ideas, experiences, and knowledge on a particular area of interest. If this is an intra-company CoP and is corporate-sanctioned, then the area of interest would logically be one that pertains to or would benefit the company goals. However, communities may also expand to include other companies and/or organizations, depending on the focus area. This paper will discuss data quality and Communities of Practice in general but focus primarily on the Data Quality Community of Practice and how having one at your company might provide benefits to your company, your customers, and yourself. Topics covered in this paper are:

- ◆ What is a Community of Practice (CoP)?
- ◆ What is data quality (DQ)?
- ◆ What is the SAS® DQ CoP?
- ◆ How was the SAS DQ CoP formed?
- ◆ What are the steps to forming a DQ CoP at your company?

This paper should be of interest to both technical and business users who are interested in the issues related to data quality within their companies, and in identifying the resources available to address these issues.

Introduction

Communities of Practice are still relatively new to the United States, though they have been in use in Europe since the early 1990s. Major companies such as IBM have conducted studies on the importance of these organizations, and software has been developed specifically dedicated to building “successful online communities.”¹ Identifying (a) why we would want/need a CoP, (b) how to sell the idea for the CoP to upper management, (c) how to set one up, and (d) how to enlist participation was a challenge for the Data Quality CoP. This paper addresses what we did at SAS to form a Data Quality Community of Practice, how we plan on utilizing this to provide benefit for both our company, and ourselves, and how you might create a Community within your own company.

What is a Community of Practice?

Etienne Wenger, born in Switzerland but currently living in California, is a recognized leader in the Communities of Practice field. He and anthropologist Jean Lave coined the term “Community of Practice” in 1991. He states that “Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly.”² He further states that there are three characteristics which are crucial to constitute a community of practice³:

- 1) The domain: a shared interest
- 2) The community: the membership
- 3) The practice: members devoting “time and sustained interaction” to promote their shared interest.

Communities “come in a variety of forms. Some are quite small; some are very large, often with a core group and many peripheral members. Some are local and some cover the globe. Some meet mainly face-to-face, some mostly online. Some are within an organization and some include members from various organizations. Some are formally recognized, often supported with a budget; and some are completely informal and even invisible.”⁴

CoP Roles

“Just as you wouldn’t board a ship that was without a captain and crew, you shouldn’t embark on a journey of supporting communities of practice (CoPs) without a clear sense of who is providing leadership and expertise.”⁵ Understanding the roles in a CoP and having the correct people assigned to those roles is crucial for the successful functioning of a community. Working with several organizations, the Institute for Knowledge Management (IKM) identified 11 formal and informal community roles. The following table lists the roles and their respective functions.⁶

Community Role	Function
1. Subject Matter Experts	Keepers of the community's knowledge domain or practice who serve as centers of specialized tacit knowledge for the community and its members
2. Core Team Members	Looked upon for guidance and leadership before or after a Leader emerges or is selected; guidance includes developing the community's mission and purpose
3. Community Members	Take active ownership in the community by participating in its events and activities and driving the level of commitment and growth of the community
4. Community Leaders	Provide the overall guidance and management needed to build and maintain the community, its relevance and strategic importance in the organization and level of visibility
5. Sponsors	Nurture and provide top-level recognition for the community while ensuring its exposure, support, and strategic importance in the organization
6. Facilitators	Network and connect community members by encouraging participation, facilitating and seeding discussions and keeping events and community activities engaging and vibrant
7. Content Coordinators	Serve as the ultimate source of explicit knowledge by searching, retrieving, transferring and responding to direct requests for the community's knowledge and content
8. Journalists	Responsible for identifying, capturing, and editing relevant knowledge, best practices, new approaches and lessons learned into documents, presentations and reports
9. Mentors	Act as community elders, who take a personal stake in helping new members navigate the community, its norms and policies and their place in the organization

Community Role	Function
10. Admin/Events Coordinators	Coordinate, organize and plan community events or activities
11. Technologists	Oversee and maintain the community's collaborative technology and help members navigate its terrain

Be aware that these roles can, and probably will, be combined such that a single person will serve in multiple roles. The ideal is for the work of the CoP to be distributed reasonably among the community membership. The following table shows the estimated percentage of time devoted to each role.⁷

Role	% Time/Week
Subject matter expert (SME)	21.4%
Core Team	10.0%
Member	6.9%
Leader	20.0%
Sponsor	5.7%
Facilitator	53.4%
Content coordinator	25.0%
Journalist	Unknown
Mentor	5.0%
Admin/events coordinators	26.0%
Technologists	25.0%

Note that some of these percentages are significant and it will be up to the inter-relationship between the company and the Community to ascertain whether this time is justified and/or necessary.

What is Data Quality?

Data quality can be interpreted to mean many things, depending on your area of focus. Searching the web for “data quality” topics will result in everything from manufacturing-specific quality to document standards to Public Law 106-54 of the federal government’s Office of Management and Budget.⁸

For purposes of this paper, I will confine the term “data” to refer to all business data which serves as the basis for a company’s enterprise-wide information reporting, analysis, and decision-making systems. The term “quality”

will refer not only to validity, but also to completeness and relevance. Ted Friedman, principal analyst with Gartner, states “If there’s data but it’s not relevant to the process or project at hand, it’s worthless.”⁹

What are the effects of bad data?

“Companies routinely make decisions based on remarkably inaccurate or incomplete data, a leading cause of the failure of high-profile and high-cost IT projects such as business-intelligence and customer-relationship management deployments... ‘Most enterprises don’t fathom the magnitude of the impact that data quality problems can have,’ said Friedman. According to his research, a quarter of major international companies are working with poor-quality data.”¹⁰

Further, he states: “... [Companies may have collected all the data they need but] it’s not in sync across the enterprise. Synchronization will become an even bigger issue in the near future as they struggle to integrate the enormous amounts of data gathered from RFID.”¹¹

“It’s a problem that can cost companies 10 percent to 20 percent of operating revenue to do what Larry [English, president of Information Impact International] calls ‘information scrap and rework.’ Yet so many companies continue to operate with and share bad data. Not only are hard dollars at stake, but also customer loyalty, reputation, or, in the case of health-care providers, people’s lives. It threatens collaboration, knowledge sharing, productivity, and real-time decision making.”¹²

Why now?

Are we just now beginning to realize we are reporting on bad data? Has no one within your company raised red flags on this issue before? Of course they have. Then why are we now talking about it? It is the “squeaky-wheel” syndrome. In recent years, the corporate focus has been driven by the latest industry buzzwords: Data Warehousing, Customer Relation Management (CRM), and, most recently, Business Intelligence (BI). However, the effect of all this rushing to be at the technological forefront is that the data, which provides the foundation for these systems, becomes more highly visible. This “data visibility” has made glaringly obvious the

quality of that data. Companies now realize that in order to make these wonderful “new” technological tools useful and cost-efficient attention must be focused on something not quite so “jazzy” but much more relevant and critical: the underlying data.

“Some advice from the experts: There has to be accountability for good-quality data at all levels in the organization: from the data-entry clerks all the way to the CEO. Fixing, and eliminating, the problem has to involve a strong methodology and new business processes, not just technology. Line-of-business managers must assume responsibility, not just dump it in the IT department’s lap.”¹³

What is the SAS Data Quality Community of Practice (DQ CoP)?

What is our mission?

We want “to provide a forum for sharing knowledge and tools on topics involving data quality within the SAS environment.”¹⁴

What is our goal?

Our goal is to facilitate knowledge sharing and promote awareness of all topics related to data quality. Community members will:

- 1) Span all functional areas within SAS that have interests or contributions to make in the area of data quality.
- 2) Promote the focus of data quality in both the sales and the consulting functional areas.
- 3) Increase awareness of the importance of data quality and the various SAS solutions and tools available for its implementation.
- 4) Provide a forum for those outside the membership to communicate issues and concerns and to request assistance.
- 5) Develop service standards, along with supporting documents, which will serve as a “template” for assessments and other presales activities related to data quality.¹⁵

What have we accomplished so far?

In the four months we have been operational, we have:

- ◆ Formed a steering committee that sets the direction of the Community and makes proposals to the membership.

- ◆ Created our website and Yellow Pages site.¹⁶
- ◆ Expanded our membership to more than 50 members across the company, including many international members.
- ◆ Conducted our kick-off meeting, which was attended in person by 30+ members and another 5-10 attending by Placeware.
- ◆ Conducted our first CoP-sponsored event, which was a two-day DQ Advanced Topics Workshop.¹⁷
- ◆ Created demos, papers and a reading list, all of which have links from our website.
- ◆ Begun a list of goals which will be further refined and prioritized by the membership

How was the SAS DQ CoP formed?

Every path to a particular goal must be individualized based on the resources available, the complexity of the tasks, the time available to devote to the tasks, and the organizational constraints. In the formation of the SAS Data Quality Community of Practice, we had strong management support and a department (Customer Solutions and Alliances, CSA) already dedicated to these activities. Delfine Welch in this department was extremely helpful in directing the sequence of our tasks. She advised that the first thing we should do was to create our website and Yellow Pages (both using Microsoft® Windows SharePoint Services).¹⁸ Subsequent to this, we took the following steps.¹⁹

- 1) I contacted a core group of people I knew to be interested in data quality and would be active participants. (As self-appointed leader of the CoP, I enlisted the aid of another co-worker, Beth Adams, to be my co-leader.) There were eight of us in all and we would form the core team or steering committee.²⁰
- 2) We made sure to include a team member who would serve as journalist for all our meetings. This is critical not only for the meeting minutes, but also for the preparation of documents that the CoP will produce. For this reason, we asked one of our Documentation Specialists, Susan Springfield, to participate.
- 3) We conducted a face-to-face meeting of a full day devoted to discussing possible goals, additional members, and next steps.²¹
- 4) After this meeting, we issued a formal invitation to the people on our list,

announcing the formation of the DQ CoP, asking for their participation, and informing them how they could join.²²

- 5) At the same time, we began working on our “next steps” list, the first being to organize a kick-off meeting. We had to decide whether this should be a face-to-face, a Placeware session, or some other form of remote meeting. Based on the productivity of our core team meeting, we opted to try for the face-to-face meeting.
- 6) We recognized that, to try and get a lot of people to travel to one spot (our headquarters in Cary, NC, made the most sense, since the majority of people were already on site) and to get management agreement, we should provide some additional incentive. Therefore, we also began planned an Advanced Topics Workshop (ATW), which would be conducted in days adjacent to the kick-off meeting.²³ (Having a technical training member, Mark Craver, on our core team was especially helpful [see reference note 20] in planning the ATW. We needed to know, not only the necessary topics for this, but also the available days and resources for planning this activity.)
- 7) As soon as we had a broad idea of our plans, we sent an email to the full CoP membership (again, using the Yellow Pages as our communication mechanism) that announced the kick-off meeting and the ATW. (We waited until later to work out the details; we wanted to send out the invitation as soon as possible to determine interest and to allow people to begin planning.)
- 8) After the details had been finalized, we sent out dates, agendas, and any other details we felt might be important. In our case, we were also interested in some general information about our membership, so we prepared a survey in which we asked the members to complete.
- 9) The kick-off meeting was held on August 30, 2005, followed by the two-day Advanced Topics Workshop.
- 10) Following these meetings, those of the core team who were able to attend met for a “lessons learned” session in which we summarized the things we think went well and those we feel should be done differently or removed altogether.
- 11) Our next steps will be to summarize the new list of goals that were discussed by the full

membership and to assign priorities to these goals.

- 12) We will begin acting on our high-priority goals while maintaining open lines of communication with our membership. The creation of a newsletter was an obvious next step, since we wanted to present the results of the two meetings to the full membership. This will be an informal email until we can create a more formalized document.

What are the processes involved in forming your own DQ CoP?

Rather than provide a generic list of tasks for forming a CoP (these can be found in the links provided at the end of the paper), I prefer to describe some of the more important lessons we learned in our own CoP formation.

Solicit management support.

First, you must get support from your manager (and ideally, the next level management as well), since a CoP can take considerable time and effort from your workday. I refer back to the role percentage table presented on page 2 of this document. In our case, I was exceptionally fortunate. My direct manager, Steve Becker, suggested the concept of the DQ CoP. I then followed up with my ideas for the CoP formation and its uses to both my direct manager and his manager (Rick Coughenour) and got enthusiastic support from both. Should you encounter resistance or even reluctance from your own management, I would recommend that you determine if, and be prepared to defend, there is a way to demonstrate direct or even indirect benefits to your company.²⁴ Some questions to ask yourself are:

- ◆ Can the CoP help create new sources of revenue?
- ◆ Can it help make employees more productive by sponsoring training or learning materials/tools?
- ◆ Can it help create new marketing tools or collateral that will assist sales?
- ◆ Can it provide information dissemination in an easier fashion than is currently being provided?
- ◆ Can it provide a single point of contact for a specific area of importance? In our case, our data quality experts are scattered throughout various departments, divisions, and locations; the DQ CoP provides a single

place that people can come for general data quality questions and/or assistance.

Get help.

It would be best if you have a department or division within your company whose job it is to help set up communities of practice or other similar organizations. These people will be your best guide for your own next steps. Your next best option would be to find someone who has had experience in creating a community of practice. There are links on the web that you can find with a Google or Yahoo! search. I have also provided a list of links at the end of this section.

Take advantage of the experience of others to avoid repeating mistakes and to leverage their successes. The people I contacted provided me with invaluable assistance in terms of advice, processes, and documentation.

Give care to the formation of the core team.

This is, in my opinion, the single most important item (besides management support): give careful consideration of your core team composition (I assume that, at this point, you have already decided on a specific area of interest for your CoP). As I mentioned earlier, having representatives from your most important functional areas is crucial. These are the people who will provide the guidelines for what will and will not work within your company. They know the most important people to add to your membership list. They will also be the most active in the successful formation and maintenance of your CoP. Choose people appropriate to these tasks. However, do not become too top-heavy; keep this membership number manageable, but have enough of the right kind of people that you will have good exchange of ideas as well as good functional assistance.

It is important to state at this point that leadership is the most important role in the CoP, especially in the start-up phase. Even if you have strong leadership skills among your core team (and I am fortunate to say we do), there needs to be one person to drive the effort so that the core team stays focused. This person may not be the most technical but should have good organizational and management skills. Above all, what this person should do is leverage the

strengths of the other core team members:
delegate, delegate, delegate!

Solicit membership from a wide variety of functional and divisional areas.

In our experience, the more variety you bring to your CoP, the more successful it will be. Cross-functional and cross-divisional membership will provide more diverse:

- ◆ levels of technical experience
- ◆ perspectives
- ◆ field experiences
- ◆ ideas
- ◆ spheres of influence
- ◆ collateral

Decide the communication process that works best for you.

Whether it is through websites, listservs, or some other process, the communication vehicle needs to be put into place early in the process. You will want to disseminate your information and be able to receive feedback as soon as possible and in the most effective way possible.

Maintain a list of action items.

This list should contain everything that needs to be done, from the smallest task to the major ones. In our case, I maintained a task list in an Excel spreadsheet, assigned a responsible person, and designated a due date. I then sorted the list by due date. This was simple to monitor, maintain, and communicate to everyone. As each task was completed, I moved it to a separate worksheet. This is a dynamic document; it changes on a regular basis, sometimes hourly, but is essential for making your CoP run smoothly. It is not only a roadmap, but also a status report; it shows where you are going and where you have been.

Plan your meetings carefully.

There are several considerations for conducting your CoP meetings.

Face-to-face meetings are most helpful for planning sessions where whiteboarding and open discussion are important. However, if this type of meeting venue is desired and travel is involved, try to provide additional incentive that would justify the expense. For example, combine it with training or managerial meetings.

Placeware or some other form of virtual meeting works well if the presentations are not in front of a live audience. As we learned, however, if there is an audience, almost all audience participation will be lost or obscured in this process, making it difficult if not impossible for remote attendees to follow the meeting. Most of our remote attendees dropped off in frustration.

Meeting minutes are important for face-to-face meetings, but are not easily produced in virtual meetings; the speakers cannot be clearly heard nor are they easily identified.

If a virtual meeting is used and the meeting is long, plan the timing of your agenda carefully, many attendees do not want to attend the entire time and will need to schedule based on the agenda topics and times.

Helpful Links

Following are some of the links I have found most relevant to the creation and practice of communities of practice. These and others may be surfaced by using search engines (Google or Yahoo!) and keywords “communities of practice.”

- 1) <http://www.co-i-l.com/coil/knowledge-garden/cop/index.shtml>
- 2) <http://www.tcm.com/trdev/cops.htm>
- 3) <http://www.research.ibm.com/journal/sj/404/less.html>
- 4) http://www.infed.org/biblio/communities_of_practice.htm
- 5) <http://home.att.net/~discon/KM/CoPs.htm>
- 6) <http://www.ewenger.com/>
- 7) <http://www.eccop.com/blogs/public/>
- 8) <http://www.cpsquare.org/>
- 9) <http://www.icohere.com/ccop2005/>
- 10) <http://www.cpsquare.org/edu/foundations/>
- 11) <http://hbswk.hbs.edu/pubitem.jhtml?id=2855&t=knowledge>
- 12) http://www.nelh.nhs.uk/knowledge_management/km2/cop_toolkit.asp

- 13) <http://home.att.net/~discon/KM/CoPReadings.htm>
- 14) <http://www.ala.asn.au/conf/2003/vanasse.pdf>
- 15) http://reframingthefuture.net/Publications_2005/facilitating_workbased_learning_05.pdf
- 16) <http://www.ejkm.com/volume-3/v3i1/v3-i1-art3-bourhis.pdf>

Acknowledgements

SAS is a registered trademark of SAS Institute, Inc. in the USA and other countries. Other brand and product names are registered trademarks or trademarks of their respective companies.

® indicates USA registration..

Contact

Donnalee Brown
18701 10th Ave NE
Poulsbo, WA 98370
206.853.3781
Donnalee.Brown@sas.com

1 <http://www.tomoye.com/>

2 <http://www.ewenger.com/theory/index.htm>

3 Ibid.

4 Ibid.

5 “Keeping Communities of Practice Afloat” by Michael Fontaine, reprinted in Knowledge Management Review, September/October 2001 (originally from the book Creating Value with Knowledge: Insights from the IBM Institute for Business Value, edited by Eric Lesser and Laurence Prusak). [page number?]

6 Ibid.

7 Ibid.

8 “Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies”, Office of Management and Budget, Executive Office of the President. [page number?]

9 <http://www.informationweek.com/story/showArticle.jhtml?articleID=20301106>

10 Ibid.

11 RFID: Radio Frequency IDentification. A wireless data collection technology that uses electronic tags for storing data. See more at http://content.techweb.com/encyclopedia/defineterm.jhtml;jsessionid=2EYYVK2TLQCWQSNDBCCKH0CJUMKJVN?term=rfid&x=16&y=8&_rquestid=115806

12 <http://www.entrepreneurpipeline.com/news/46200037>

13 Ibid.

14 From the SAS Data Quality Community of Practice website.

15 Ibid.

16 There are processes in place at SAS that facilitate the creation of these internet sites. The process for doing this at your company may be more involved.

17 The content/structure of this workshop can be discussed as a separate topic or in individual emails.

18 Yellow Pages, in this instance, refers to the internal SAS Yellow Pages websites through which we communicate information about employees, departments, organizations, etc.

19 It should be emphasized that, aside from the help we received from Delpfine Welsh, we were pretty much following our instincts. There are a number of documents to be found on the internet about the creation of Communities of Practice.

20 It cannot be emphasized enough the importance of the structure of a core team/steering committee. It is vital to have representation from as many departments as possible that will help facilitate the process. For our core team, we enlisted members representing technical training: Mark Craver; DataFlux marketing: Lisa Flint; consulting Donnalee Brown, Beth Adams, Sheryl Lucas; SAS’ internal hardware resources group: Dan Mazza; documentation support: Susan Springfield; and system engineering: Tina Ritterhoff.

21 Membership in a CoP can be opened to the entire company or it can be by invitation only; for various reasons, we chose ours to be the latter, though this may change later.

22 At SAS, this functionality is provided through the Yellow Pages link.

There is a button that can be selected if a person wants to add their name to the membership list.

23 Ibid.

24 There are papers and even formulas dedicated to computing the ROI to a company of a community of practice.