

Behind the SEI process-maturity assessment

Since its introduction in 1987, the Software Engineering Institute's model for assessing organizations' software processes — the level of management and development maturity — has become widely discussed, since it is the first popularly accepted metric by which to evaluate an organization's capabilities.

The SEI's initial assessments of organizations revealed sobering findings: Most scored at the bottom rung of the five-level maturity framework. From bottom to top, the levels are initial, repeatable, defined, managed, and optimizing. Although the SEI has found several projects at level 3, no company surpassed the second level. At the second level, organizations achieve a stable process with a repeatable level of statistical control within projects, but not across the organization as a whole. (The framework is described in "Characterizing the Software Process: A Maturity Framework," *IEEE Software*, March 1988, pp. 73-79.)

Because the SEI is sponsored by the Defense Dept., the initial reaction from some defense contractors and industry-association members was fear that the process model would make them lose contracts or be forced to make expensive changes in their development processes to keep getting projects. At SEI review meetings, some industry representatives expressed their concerns about the model's potential misuse as a grading system for software contractors.

But a year later, the model seemed to be viewed more positively, as evidenced by frequent praise for it at the 1989 International Conference on Software Engineering, held May 15-19 in Pittsburgh. Fears of its use as a punitive measure seemed to have given way to hopes that the model would help companies identify the steps needed to mature their processes, resulting in better products (and perhaps more contracts for individual companies). More important, it would help the software-engineering community target basic issues in the discipline's ongoing evolution.

How does the SEI conduct the assessments of organizations' process maturity? At the International Software Conference on Software Engineering, Soft News Editor Galen Gruman asked Watts Humphrey, the framework's author and director of the SEI's Software Process Program, to describe how the framework originated and what occurs during an assessment.

Q: *How did you develop this framework?*

A: The basis comes from W. Edwards Deming and J.M. Juran's concepts of statistical repeatability. According to Deming, to improve a process, you first need to stabilize it. Then you need to define it so you can make sensible measurements and informed improvements. Without measurements, it is hard to establish an orderly basis for action. As an IBM Burlington plant manager once said, "You can't manage quality from the end of the line." Without data, that's what we're doing in software.

Q: *What led you to this assessment approach?*

A: Before coming to the SEI, I worked at IBM, where we did technical assessments of both semiconductor and software work. Phil Crosby, author of *Quality is Free* [McGraw-Hill, 1979] and a [former] member of the ITT quality group, had a five-level maturity framework, which I learned while at a course he taught.

At the time, I managed the IBM 370 software-quality work; we did assessments regularly and found that the five-level Crosby model worked but had some limitations since it did not specifically relate to software issues. We could not correlate the assessment results with what people had done to improve their pro-

cesses. It didn't point to successes or establish priorities.

Q: *How did this experience become the SEI framework?*

A: The Air Force asked the SEI for help on how to evaluate software contractor organizations — it wanted a way to know who was the best and most capable one to contract with.

The SEI decided to focus on the process. We started a survey in a joint project with Mitre Corp. and developed some 150 questions trying to figure out a company's software capability.

In one of my many trips, I spent six hours in the Atlanta airport. I had just visited a software group and was flying to Boston to meet with Mitre on the questionnaire. It just clicked. I realized how to apply Deming's concepts of maturity and couple them with Crosby's five levels, so I began sorting the 150 questions accordingly.

I then went back over the new sort and was uncomfortable because the questions were unequal: For example, could an organization be at level 2 without fulfilling some particular question? I marked those elements that really had to be there for each level.

With Mitre, we put together a draft questionnaire and then interviewed about 10 companies, including TRW, Hughes, IBM, and GTE. We got a lot of help from their top people. All were personal contacts and strongly supported the idea.

We then had discussions with people from the Electronics Industry Association and the National Security Industry Association, who were more concerned about what the results would mean to them. Their reaction sensitized us that their concerns are critically important. We don't want the process-assessment model to become bureaucratized, where [the Defense Dept. or any other government agency] blindly uses a checklist to evaluate someone's software efforts.

The SEI then put out a first version of the process-assessment questionnaire.

Q: *Has the government used it properly?*

A: Generally, the [Defense Dept.] people have found that it helps them pinpoint key acquisition risks. Industry has also been positive, except for two cases we have heard of where untrained teams didn't do it properly. When we hear of such cases, we like to straighten them out.

Q: *Besides Defense Dept. efforts, the SEI also evaluates companies' efforts. How do you assess companies?*

A: In the typical assessment, we interview 40 to 60 people. The process starts with a meeting of senior executives — presidents and vice presidents or laboratory directors — where we emphasize that the assessment is confidential and is done to assist their implementation efforts.

As Deming knew, [most] people are doing the best they can within the confines of their process. Since only management can change the process, you have to start with management. I found that you don't have to explain this process to senior management. By and large, these are very sharp people, but they don't know what works [in software development], so they don't know what to do next.

[But] software people often don't like talking to senior management. They need a way to communicate with management other than with just schedules and budgets. A maturity framework gives management a list of things to do outside budgets

We don't want the process-assessment model to become bureaucratized, where it is used as a checklist to evaluate someone's software efforts.

and schedules. [With this list, management] can help the software group address their key issues.

After the senior managers buy off on the assessment proposal, the SEI sets up the schedule. Within 10 days, we get an SEI person to go out to coordinate. Within one month, we have five or six SEI people and two to four people from the company come for a two-day training session at the SEI. [The SEI people train with company people to foster cooperation and help them understand the company's perspective.] We require management to sign an agreement with the SEI committing senior-management involvement and action plans when [the assessment] is through. This underscores the seriousness of the effort.

This training program also builds a team and helps the SEI learn the company's organization and the candidate projects. In these two days, the team defines specific responsibilities and schedules. We then do a team-building exercise to help the team members recognize that they each have information that others need. You often don't know you have information that others need.

The on-site assessment takes four days.

On day 1 of the assessment [typically a Tuesday], everyone involved is present at an opening meeting. The SEI gives a presentation on what the process is and what is to be done (the troops have usually been told beforehand).

Each project leader has filled out a survey a day or two earlier. One SEI person gets there a day early [Monday] to review them. After the [opening] meeting, the assessment team goes over the questionnaires, looking for interesting areas like differences between projects or between levels. Frequently, we see major differences. Managers have a more optimistic view through level 2 than the troops. The troops generally grade themselves a level lower. Rarely — at level 3 — they're much closer, which supports level 3's consistency attribute. Usually, it's like a team where some people are playing football, some basketball, and some hockey!

We interview the project leaders on key areas. We ask for them to come back Thursday [day 3] with documentation for their claims. They often don't have it to the degree that they thought [they did].

We end the day with the assessment team meeting to go over preliminary conclusions. These usually end at 7 p.m. or 8 p.m. We call these "half days" — that's half the 24 hours in a day, right?

On day 2, we conduct four or five sets of interviews each with five to seven people who are experts in particular topics from the company — but not necessarily from the same projects. These meetings generally last one and a half hours. These are very open discussions — we listen to them. In the last half hour, we end with one question: "If you could fix one thing, what would you do?" We get some very powerful ideas. These are the opinion leaders. We get a perspective that is often quite different than from the managers. Surprisingly, though, we rarely find contradictions: We are looking at the same problems from different directions.

That evening, we go through a more intensive effort where we arrive at preliminary findings: key issues and problems. It's a fairly late night: 9 p.m. to 11 p.m.

On day 3, we do not yet have a prepared pitch on the issues, but we know what the issues are. We get the project leaders back one by one for feedback. These meetings last about 45 minutes. We walk through the preliminary findings to see what they

think. We get practically no disagreement on it, although we do get lots of comments and context. We have a lot of data, but no final pitch. We're discussing the bricks to see if we've got the right material to make a complete story.

The project leaders feel they are partners [in the assessment]. They are not defensive, usually. People really want to improve things. They know it [the findings] but they haven't seen it all in one place before.

At 2 p.m., the assessment team starts developing its final findings; it's typically done by 10 p.m. or 11 p.m. Sometimes, the team works all together, and sometimes it forms subteams and regroupings. It depends on the issues and what works for the teams.

On day 4, we take the pitch and walk through it ourselves to clear it up. By 9:30 a.m. or 10 a.m., we get all the project leaders together and go through the whole pitch. There is practically no debate. There is some discussion, which is very positive.

In the early afternoon, all the people are invited — the troops, senior managers of the location, and project leaders — to a meeting. The senior managers often invite corporate executives. After the presentation, the assessment team puts together their recommendations. Meanwhile, I (or whoever leads the assessment) will meet with the senior managers to give them a head's-up view of what might be in store. The typical reaction is "I want to get started right off."

We put together a draft report, which we subdivide into parts due the following week, and we have a review a few weeks later. (This gives us time to go over it.) After about six weeks, we do a formal presentation and provide a written report. The company reviews its action plan with us two or three months later.

Q: *What's the next step?*

A: Now, we're putting together a case-study framework. We're going to go out

to many of the organizations assessed to see what's been done, [to see the] before-and-after results, to get some data on what effect this [assessment process] has.

Several organizations have committed to go to level 4. And they're likely to bring it off. In a couple years, we expect to see measurable effects of this assessment work.

Q: *What does it take to move up in the framework?*

A: At level 1 [in the framework], it's often very individual, so someone in the organization has solved every problem we find, but no one else knows about it. We look for existing solutions to the problems.

At level 2, there's no organizational ethic, just a project ethic. But when a project is disbanded, the project's experience is often lost.

Software is a team activity — and teams don't happen, they're built. What we're finding is that crisis management inhibits teams. We have to get them to act like coaches. The key is management, not managers — *everybody* is involved in the management of the process — the technical people [too].

Q: *This sounds like a simple idea, so why are there so many software-development problems?*

A: We don't have a model for large-scale intellectual activity. Large-scale human behavior is trusting, cooperative behavior. That's not something that's easy — it's simple, [but] it's not easy. It's like baseball: It's easy to understand but it takes a lot of coaching, support, and practice to make champions.

Large-scale human behavior is like baseball: It's easy to understand but it takes a lot of coaching, support, and practice to make champions.
