



Who is Measuring What? A Metrics Show & Tell



SNAPSHOT

SPEAKERS

Alex Blanton (Microsoft) & Kelly Schott (TheCR)

COMPETENCIES

Metrics & Measurement

MATURITY PHASES

CMM1, CMM2, CMM3, CMM4



HIGHLIGHTS

1. **Interested in the types of metrics that are most effective?** Alex shared the 4 metrics that he finds most useful.
2. **Looking for formulas?** Alex offered a couple of formulas that help him gauge “net” scores. Learn which 6 net scores he tracks and what they mean, along with what he considers to be a good or bad score.
3. **Ever considered creating a Community Health Index (CHI)?** Alex gave an in-depth overview of the 3 parts of Microsoft’s CHI.

OVERVIEW

- **Background:** Alex has been the Community Manager for Microsoft’s Internal Machine Learning Community for the past 3.5 years.
 - Machine Learning (ML) is software that uses data to make predictions about the future. People are most familiar with it if they use Amazon or Netflix as these companies create recommendations for customers based on their past actions. Other companies use it for price optimization or predictive maintenance (predicting when parts are going to fail in elevators or trains or airplanes). Credit card companies use it to predict fraud and report suspicious or unusual transactions.
 - Alex stressed that he’s not a data scientist or ML engineer. His background is community management, which can be good and bad. It’s good because he doesn’t get hung up on specific debates and he can ask questions without feeling stupid. It’s bad in that he wishes he had more domain knowledge. He feels added domain knowledge would help him to be more effective as a community manager. He turns that to his advantage when recruiting experts to become

CASE STUDY

- **ML Community:** This is a community of practice for people who are interested in learning and scaling-up on ML. It consists of about 6000 members across the world inside of Microsoft (MS) in various roles.
 - The community is part of a broader network called: “*The Worldwide Community Program.*” It’s a set of about 90 technical, domain and role-based communities. The community managers sit within individual product groups or sales groups, etc. They manage the community in addition to their regular job. Alex is an exception. Community management is the main part of his job because the corporate VP strongly believes in the community.
 - The Worldwide Community Program consists of a group of 6 people who manage the program. The Community provides the infrastructure, including metrics.
- **Community goals:** These goals are implicit, but unstated:
 - To provide a venue for people to let their work shine. This is particularly important given the large size of MS. People can sometimes feel buried in their individual teams.
 - To be a testing ground for those who may be able to represent MS publicly.
- **What does MS get from the community?** Alex broke this down into 4 areas:
 - **Workforce enablement:** It is a benefit to MS to help its workforce get better at ML and artificial intelligence (AI). It’s incredibly expensive to buy this kind of talent, so helping people to build it and switch over from other roles is a key reason behind why the community exists.
 - **Improved customer response.** Enabling quick responses from the engineers to the field staff who help customers is vital. Some of the community activity metrics are geared towards driving this goal forward.
 - **Better connections.** MS wants the community to help connect people across the company.
 - **Recruitment and retention.** Currently, there is a competitive talent war happening for candidates with ML and AI skills. MS would like to position itself as a place for these individuals to build their careers.

CASE STUDY, CONT.

- **Community activities.** Providing a diverse set of activities is important to the community to help it feel fresh. Metrics help them decide where to focus their community programming.
 - People can create as many communities as they want at MS. As a result, there are many spin-off communities. For example, communities that use ML to make sporting predictions, ML for security, art with robots and many more.
 - Metrics are not tracked in these spin-off communities as the tools are not available. As a result, the metrics are an incomplete picture of what's happening across the Company in terms of ML. Alex has come to accept that as a fact. That said, he does try to align with them to help them be successful.

- **Types of metrics.** Alex broke down their metrics into 4 areas:
 - **Event related.** This could be a standalone talk or a big conference. They are assessed individually. Assessing the trends helps with community programming.
 - **Attendance.** This refers to online and in person. Online attendance is imperative because of the global nature of the community, as well as the fact that some of the attendance is on-demand. On-demand participants listen to the recording at a time convenient for them.
 - Usually the on demand attendance is 150% or more of the live attendance. Speakers need to be aware of this fact so that they are not disappointed to see low in-person or online attendance.
 - **Trends.** Trends are tracked for topic areas. This is especially useful for Alex who can track when a topic is rising in interest, when it peaks and when it starts to wane. Especially in the ever-changing field of ML, knowing what topics to offer is imperative. For instance, ethics within ML is an up and coming topic.
 - **Repeat attendees.** The Worldwide Community is internal and free. Time is a currency in this instance. Therefore, repeat attendees is a strong metric because people are willing to invest and reinvest their time, meaning it offers them value. This is also a great pool of people to draw from when you need people to take on bigger roles within the community. Alex reaches out to those people who have attended 10 or more sessions a year.
 - **Community activity-related metrics.** This consists of Q&As, conversations and content that is shared. Some of these contribute to a Community Health Index (CHI) assembled by The Worldwide Community Program. A number of these are tracked and reported back to Alex, but they are not counted in the CHI.

CASE STUDY, CONT.

- **Informal.** Alex feels that it's important to pay attention to other activities that can't be reported on formally. If he only paid attention to the activities that he could put a number to, he doesn't believe that he would have a complete picture of what was happening in the community.
 - **Costs.** A necessary evil to track.
- **Why we measure what we measure.** Alex feels that this is a good exercise to go through as it helps when making decisions for the community.
- **Community programming.** This metric helps with understanding the type of content that is attracting people, engaging people and knowing who is good at presenting the material. This helps Alex ensure that he is serving the community well, particularly in the fast-paced ML world.
 - **Engagement.** It's very useful to know what drives engagement and what falls flat.
 - **Identify members to develop.** Some of the activity metrics help Alex to identify members who are active in the community and who would be capable of taking on a bigger role.
 - **Discover new audiences.** Attendance metrics for events and the roles of those members in attendance assists with the discovery of new audiences to the community. Last summer, UX and product designers started entering the community due to their growing interest in ML and AI. Now they have added sessions aimed at designers as a sub-series within the community.
 - **Drive investments.** Metrics help you to make an educated decision on where to invest time, money and resources effectively.
- **Costs.** Alex has a budget that he manages for the conference costs, venue rentals, live streaming/recording, giveaways and food and beverage expenses. Lately, he has had success with giveaways and contests. For one conference, they held a sticker contest. People were asked to create slogans and then stickers were created around the slogans. It was very successful. They had over 100 responses and then asked the community to vote on their 10 favorites. From that, stickers were made out of the top 5.
- At the next conference, they are going to try a t-shirt design contest. The t-shirts will be printed on demand for purchase. It's an experiment, but if it's successful it won't impact their budget and can be used as a contest again.

BEST PRACTICES

- **Consider surveys that provide information on net scores.** Alex uses surveys after each event. They ask a few verbatim questions, but are most interested in the quantitative questions.
 - **Net Satisfaction** = $[VSAT - (DSAT + VDSAT)] + 100$. Satisfied and Neutral are not counted.
 - This is used many ways at MS. It asks the question: “*How satisfied were you with this event?*” There are 5 answers: Very Satisfied (VSAT), Satisfied (SAT), Neutral, Dissatisfied (DSAT) or Very Dissatisfied (VDSAT).
 - Take the percentage of people who answer Very Satisfied and subtract the percentage of people who were Dissatisfied and Very Dissatisfied. Then add 100 so that the number is never negative. Satisfaction and neutral are thrown out from this measurement.
 - As an example, if the survey stated that 60% of people were Very Satisfied, 10% were Satisfied, 10% were Neutral, 10% were Dissatisfied and 10% were Very Dissatisfied the NSAT would be 140. The formula is $60 - 20 + 100 = 140$. Satisfied and neutral are not counted.
 - For an individual speaker event, an NSAT 140 or 150 is Alex’s goal. For a big conference with 100 sessions, multiple tracks and all kinds of different experiences, he aims for an NSAT of 125 because there are so many different things that can effect whether or not people had a really great experience. Therefore, the bar is a little lower.
 - Favorable% = VSAT + SAT
 - Favorable percentage is Very Satisfied plus Satisfied. In the example above, it would be 70%.
 - Alex’s goal for an individual event is closer to 90% and for a big event his goal is 80%.
 - Alex would be disappointed with a favorable percentage of 70%. It’s not a bad score, but it shows that a lot of people were neutral or dissatisfied.
 - Net Effectiveness (“This event was an effective use of my time”)
 - Net Promoter (“I would recommend this community to a colleague”)
 - Net Repeater (“I would attend another ML Community event”)
 - Net Community (“This event made me feel part of the ML Community”)
 - For the last 4 scores, Alex uses the same Net Satisfaction formula, but asks the related question outlined in the brackets above.
 - The Net Repeater score is particularly useful to track. If they held an event that didn’t go well – a bad presenter or the topic didn’t land – satisfaction can be low, but still have a high Net Repeater score. That tells Alex that the community is forgiving of one unsatisfactory session and they would still attend another event.

BEST PRACTICES, CONT.

- If the Net Repeater or the Net Promoter scores start dropping, Alex would take that as a leading indicator that something was wrong in the community because people aren't planning to recommend and they're not planning to attend again.
 - Alex tracks these numbers in a spreadsheet. He recommends paying attention to how many people have replied to a survey. If only 2 people reply and the NSAT is 200, it tells you something, but not much considering it's only 2 people. The same applies to if the score is really low with only a few respondents.
- **Examine trends.** Alex tracks trends for attendance, trends for repeats, trends for how many people have attended X number of the big conferences, etc.
- Alex also pays close attention to attrition because attrition affects attendance trends. The potential repeat audience changes because some of those people have left the company and they can't be repeats. That places context around the number.
- **Consider developing a Community Health Index (CHI).** At MS, the CHI is broken into 3 parts:
- **Membership**
 - Membership growth/decline
 - **Participation**
 - Percentage of members posting
 - Posts per member
 - Posts per community leader
 - **Collaboration**
 - Conversations with +1 reply
 - Median response time (very important when dealing with customers in the field)
 - Posts per day
 - Learning resources published
 - There are other metrics that are tracked, but not counted in the CHI:
 - Roles of the community members, organizations represented, geographic distribution of members, community interconnections (which communities have overlapping members), top threads, top posters and community leadership activity details, i.e. the number of people within the roles of lead, SMEs, champions and time zone lead activities.

BEST PRACTICES, CONT.

- MS tracks the above measurements in Power BI.
 - ❖ The score is measured out of 1000. A score between 400-800 is fine. You cannot score 1000. Some of the metrics are intentional with one another. Anywhere above 400 is good, but you don't want to drop below that score. Seasonality is factored into the score, such as December when many people are on vacation. If the numbers don't come back in January, then you reassess.
 - ❖ There is also a participation element for CHI. Again, the maximum score is 1000 and there is never a perfect score. Obviously, the size of the community will affect the participation scores. When you have 6000 members, unique posts will naturally incur through replies. On the one hand, it's very hard to raise the percentage of members posting because there are a lot of people who are not totally engaged. In the communities that have 200 members, it's going to be harder for them to have a lot of posts per day, but the percentage of members posting might be a lot higher. That's why a score of 1000 is essentially impossible because some of these things would be hard to hit by all the different-sized communities.
- **Think about using TheCR's View of the Community as a communication tool.** This isn't integrated into their reporting system, but Alex really likes TheCR's View of the Community because it helps him to communicate to people who are new to understanding community. The framework breaks the community into the following categories:
 - **Owners:** In The Worldwide Community, it's one person - Alex.
 - **Collaborators:** These are about 150-200 community SMEs, champions, time zone leads, the conference program committee and volunteers, as well as those in the spin-off communities.
 - **Creators:** These are presenters, reviewers and posters who create new IP. About 500 in total.
 - **Contributors:** Contributors are those who add in, but do not participate at a high level. They are repliers, commenters, likers and raters.
 - **Listeners:** These are people on the very outlying spectrum of the community.
 - Alex doesn't track all these categories all the time. The key is to ensure that people are moving inward from listeners all the way to collaborators.
 - There will always be attrition. People will leave the company or change jobs. Volunteers and spin-off communities might dwindle. It's invaluable to understand this pipeline.

LESSONS LEARNED

- Kelly asked Alex to discuss who he reports these metrics to within the organization. Alex does not have to spend a lot of time putting together reports. However, if he had to say, there are 3 people he reports to:
 - **Corporate VP.** This is the executive sponsor of the community. When Alex gives the CVP information, it's often just a 5-minute update. Fortunately, the CVP participates in the community. He will chime in on discussions and help with technical questions. He has his finger on the pulse of what's going on in the community. Therefore, the formal reporting is not that important to him. It allows him to focus on revenue-related reporting.
 - **The director of strategy business management.** This is Alex's direct boss. He reports to him every 6 months. They sit together daily, so again, he is aware of what is going on informally. The bi-annual reporting just has numbers attached to it and is integrated into the employee review process.
 - **The Worldwide Community Program.** This community issues a report on the CHI. It's not so much that Alex reports to them as they report back to him on what has happened in the community. Alex, in turn, brings this information to his boss. There are no lines of authority in this reporting, so it's mostly just information sharing.

- Alex was asked how he decided on which metrics to track and which ones help him to gauge success.
 - For the event-related metrics, the NSAT is the primary metric. That metric came from Alex's learning and development background as they tracked this for the various courses.
 - For the CHI, they looked at what they were trying to accomplish in the community and what could be tracked to help them understand whether they were meeting those goals. For instance, the median response time. If the goal is to connect the field and engineers in a timely manner, a response time of 3 days is an obvious fail. They strive for 90 minutes.
 - Alex's overall advice is to look at what you are trying to accomplish. Then there's an element of feasibility about what you can actually measure. Alex would love to measure time and community – the amount of time people are spending focused on the community in respect to their whole job – but they don't have the right tools at present to measure that effectively.
 - Metrics like the CHI are measured every year and tweaked based on what they see happening in the community relative to their overall goals.

LESSONS LEARNED, CONT.

- Participants discussed how they measure the listening/lurking factor.
 - One participant uses the out-of-the box Jive tool to track what's happening in her community. This participant doesn't necessarily like it or agree with how Jive has it set-up, but it's all they have at the moment. Jive measures all activity on a 30-day cycle. Activity once every 30 days is not what this participant considers an engaged participant. However, they have landed on their participating user rate as their success metric. That looks at the number of users every month and how many participate through commenting or liking.
 - Another member went through an exercise a couple of years ago in which they listed the various behaviors that a member could exhibit and then categorized them. Over the years they've switched. If something didn't seem right, they reviewed where a certain behavior actually belongs. Then they got creative with measuring it. Alex's comments on this resonated with this participant. It's not a clear report that can be pulled that explains everything, but once the buckets are created, this participant can get creative in the pulling of information. Sometimes it's not just using the community platform, but trying to figure out what the listeners are doing. If the listeners are opening weekly email from the marketing service, that's not an inactive person, even though they are a listener. They might not have logged into the community platform, but they've opened an email. That turns them into someone who is reading and paying attention.
 - Another participant liked how Alex uses all the different systems together. That is a much more sensible model long-term vs. the community platform vendors trying to make them push everything into one ecosystem. This participant likes the idea of a bucket approach.
 - Alex does wish that they were on a single platform because this way creates a lot of manual work. When they run conferences, they have to pick rooms for each session. Trying to predict how many people will sign up for each of the sessions is a real challenge. You don't want the session to be overcrowded or too empty. They only just now have enough data to make the right predictions, which has been a painful journey because this is a ML community. Every time a room is overcrowded, they get comments about using ML to predict how many people will show up for each talk. However, they are only just now able to collect enough data to feed into the system to make these kinds of predictions.

LESSONS LEARNED, CONT.

- Alex was asked how he gets people to read the reports. This participant completes reports each month, but never gets feedback. She was looking for advice.
 - Timing is important. Don't try and talk about the report if the executive has something else pressing on his/her mind.
 - Alex believes that there is a great deal of information that comes out of the numbers, but maybe they don't need a formal report. For example, with the attendance numbers. He realized the 150% of the views ultimately come from the recording. That was incredibly worthwhile information to communicate to the speaker. This data was also helpful to share with the speaker in terms of who would be in the audience, i.e. 50% will be developers, 15% data scientists, etc. It helps the speakers to prepare a better presentation because they understand their audience.
 - Hillary shared that they started putting it on people's calendars. Instead of just giving them the report, they took the time to show it to them. She and Kelly worked on this for a long time last year to create a deck. For each slide they made a decision on who would be the best audience and what that audience would potentially need from the Network. Be prepared, however, for feedback. The first time they did this, they heard from people that they just didn't care about the information at all. As hard as that was to hear, it was good to know because now they don't spend time on something that isn't valuable. It was a painful lesson, but at least now the given audience signs off on whether or not the information was helpful.

RESOURCES

- As mentioned in the discussion: <https://network.communityroundtable.com/viewdocument/report-microsoft-worldwide-commun>
- As mentioned in the discussion: <https://network.communityroundtable.com/groups/community-home/digestviewer/viewthread?MessageKey=80254dbb-a9ce-4a76-901c-08239d968945&CommunityKey=7340d4a5-8d4c-4a6e-b1dd-3bfb07f292c3&tab=digestviewer#bm80254dbb-a9ce-4a76-901c-08239d968945>
- As mentioned in the discussion: <https://network.communityroundtable.com/blogs/rachel-happe/2018/01/30/want-insights-the-state-of-community-management-20>