Driving Sustainability of a Gateway with a Customer Relationship Management (CRM) System

Presented at Gateways 2022, San Diego, USA, October 18–20, 2022. https://zenodo.org/communities/gateways2022/

Gerhard Klimeck, Lynn Zentner, Alejandro Strachan
Network for Computational Nanotechnology (NCN)
Purdue University
West Lafayette, USA
gekco@purdue.edu

Michael Zentner
San Diego Supercomputing Center (SDSC)
University of California, San Diego (UCSD)
San Diego, USA
mzentner@ucsd.edu

***Abstract*—The science gateway nanoHUB has operated and supported a growing user base for over 20 years. To move toward future sustainability, nanoHUB must retain existing users, cultivate new users, and foster community involvement. In order to support retention and growth of the community, the nanoHUB team has leveraged their existing user information and analytics along with the commercial customer relationship management (CRM) tool Salesforce to better understand and communicate with current, former, and potential users. This paper details our efforts in that direction.**

Keywords—sustainability, community, growth, CRM, gateway

I. Introduction

As one of the world's leading science gateways, nanoHUB accelerates the NSF vision and strategic goals to 1) empower STEM talent to fully participate in science and engineering, 2) create new knowledge, 3) benefit society by translating knowledge into solutions, and 4) excel at operations and management at both speed and scale. Over 165,000 nanoHUB users have performed millions of simulations, democratizing access in unprecedented ways. Our over two million annual visitors—including users from 140 minority serving institutions—rely on nanoHUB for access to critical research-based resources, such as 700 community-authored and refined simulation tools and apps, 170 courses, and over 6,000 tutorials, lectures, and homework [1,2].

Despite impressive user numbers, we have not yet reached a stage where nanoHUB is sustainable. We have developed 3 specific models for sustainability based on 1) individual subscription, 2) institutional subscription, and 3) industrial sponsorship. The baseline for sustainability within each of these models is simulation user growth for the nanoHUB core service of "Apps in Education" (see Fig. 1). We understand that we need to reduce user churn and re-engage users, specifically faculty members that are teaching with nanoHUB. We also need to recruit new faculty who, through their use in classrooms, will bring new students into nanoHUB. We want to engage faculty in such a way that they are motivated to form a community and contribute content such as homework or project assignments. This in turn increases the nanoHUB content and value. We also need to reach out to adjacent fields of science and engineering and attract new tool and content contributions.

In support of our business models, we have laid out several strategies to drive our sustainability (Fig. 1). The key element is to understand our users and their needs better, as we facilitate community formation and growth. This must be done by constant relationship building in user groups and discussions that are segmented by their needs and requirements. Messaging to different user groups needs to happen via various channels. In any such engagement campaigns we need to be able to measure success in terms of users we retained and new users gained.

nanoHUB knows a lot about its users [1,2]. However, prior to 2020 we had no well-integrated way to look at a user or an institution or into market penetration in a simple way. Data had been processed into somewhat decoupled silos which are to some degree accessible to data scientists, but not to an outreach and marketing team. We realized that we needed to embrace the idea of a Customer Relationship Management (CRM) system to provide a 360-view of nanoHUB users and use such a tool to help drive our sustainability efforts.

Customer Relationship Management (CRM) software and applications have been in use by organizations for decades. CRMs address many areas of interest, including group and targeted communication, marketing, sales, customer success, social media engagement, and gaining a better understanding of online customer behavior. Free CRMs have limited capability while paid versions offer a fully functional suite of options. The Salesforce CRM receives a top rating from Gartner in its CRM category and nanoHUB has an opportunity to use the service in conjunction with other organizations at Purdue.

II. Linking nanoHUB and Salesforce CRM

In Salesforce we have the following major databases that are cross-linked to various degrees as explained below. Each is described in further detail, along with sample report snapshots from our Salesforce implementation.

1. Organizations - linking to users, leads, tools, clusters, and citations
2. Tools (simulation tools, which are our principal value-add service) - linking to nanoHUB users, clusters, and organizations

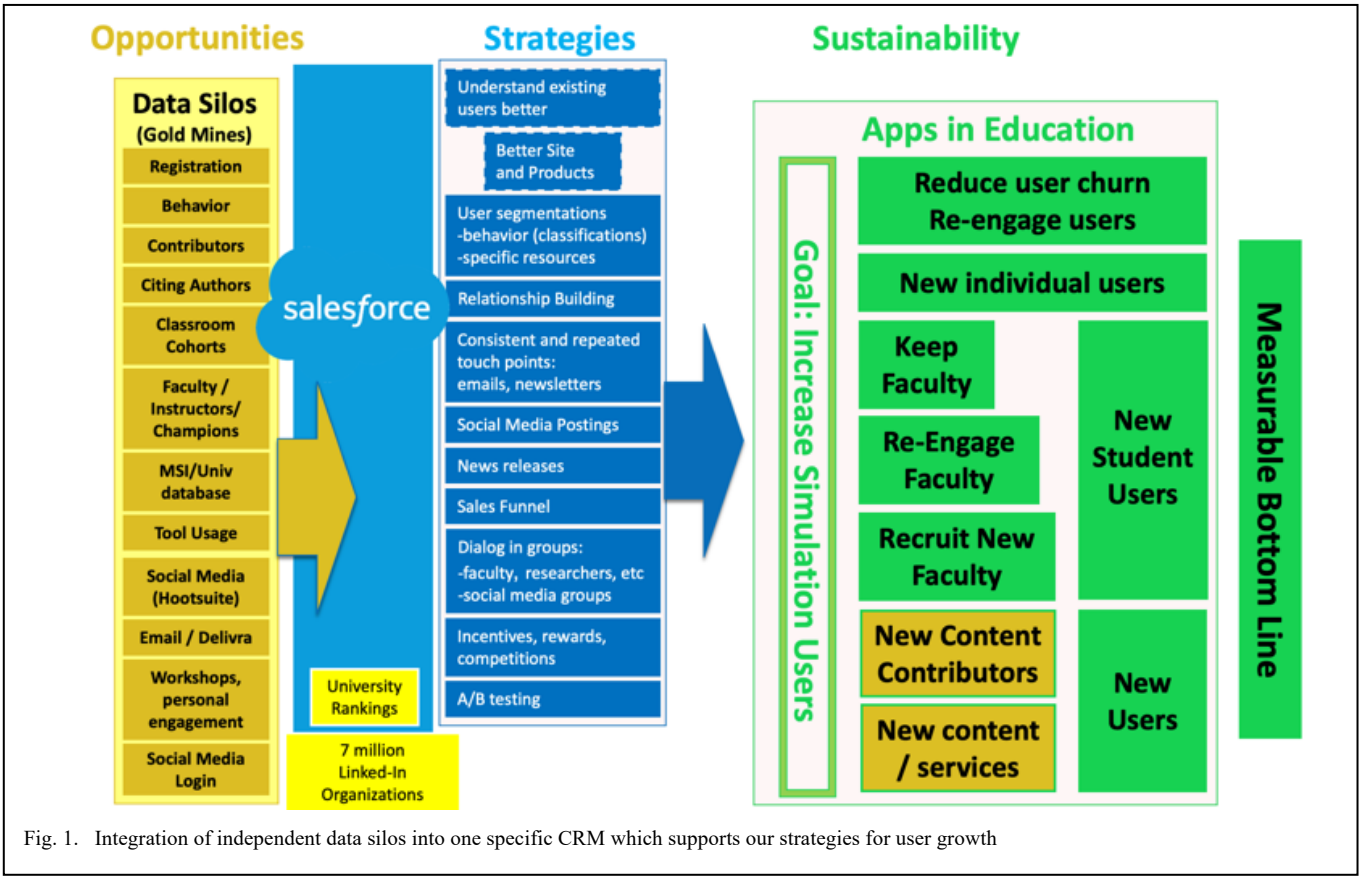

Fig. 1. Integration of independent data silos into one specific CRM which supports our strategies for user growth

3. Clusters (classroom and structured use of nanoHUB by organized cohorts) - linking to tools, users, and organizations
4. Citations (scientific publications that cite nanoHUB) - linking to users, leads, tools, and organizations
5. nanoHUB users - with self-identified information and derived information. We can now derive user information from email addresses, peers in clusters, and strong validation of organizations in citations. Data for each user in Salesforce also includes usage information, cluster participation, and nanoHUB citations.
6. Leads - individuals included in marketing efforts who are not yet registered nanoHUB users

Organizations -

Our business models focus on individual subscriptions, institutional subscriptions, and corporate sponsors. As such we need a solid base of uniquely identified organizations. We imported data and categorized these organizations as described below.

U.S. universities. The National Center for Education Statistics (NCES) [3] gathers a significant amount of data about U.S. education institutions. We downloaded the complete set of 2018 Institutional Characteristics Survey Data. This data set consists of 6857 academic institutions and describes them in regard to a variety of characteristics. All of these institutions are now imported into the nanoHUB Salesforce instance.

1. Basic institution info: Unique numerical identifier, unique name, location, address, state, unique web address
2. Carnegie Classification [4]: The basic classification is an update of the traditional classification framework developed by the Carnegie Commission on Higher Education in 1970 to support its research program. See Fig. 2 for an example of this information within our Salesforce instance.
   - R1: Doctoral Univ. – Very high research activity
   - R2: Doctoral Universities – High research activity
   - D/PU: Doctoral/Professional Universities
   - Master's Colleges and Universities
   - Baccalaureate Colleges
   - Baccalaureate/Associate's Colleges
   - Associate's Colleges
   - Special Focus Institutions- Institutions where a high concentration of degrees is in a single field or set of related fields. Excludes Tribal Colleges.
   - Tribal Colleges - Colleges and univ. that are members of the American Indian Higher Education Consortium, as identified in IPEDS Institutional Characteristics
   - HBCU and MSI classification
3. LinkedIn public data set of over 7 million corporations. We did not include this huge data set in our Salesforce instance, but use this dataset in our pipeline to match nanoHUB user accounts to specific company names or email

addresses. U.S. academic institution names are homogenized with the NCES data set.

4. We imported several of the 2020 U.S. News and World Report university rankings into our organization data set.
5. Not yet included - information on international universities and rankings

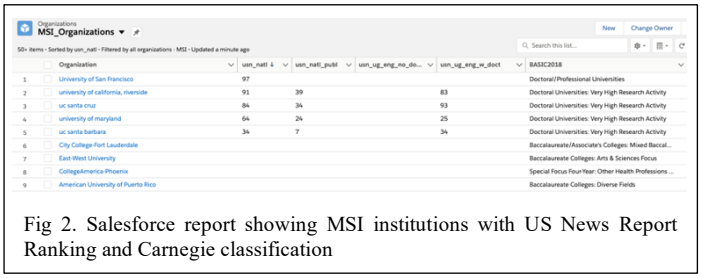

Fig 2. Salesforce report showing MSI institutions with US News Report Ranking and Carnegie classification

Clusters-

We had previously developed a methodology to identify coordinated use by multiple users and vetted such cluster patterns as structured educational use of nanoHUB [1]. We subsequently streamlined our cluster identification and now import this cluster information on a daily basis into Salesforce. These clusters are linked to the organization as well as individual users and examples of cluster data are shown in Fig. 3.

Citations-

We have accumulated a database of over 2,600 papers that cite nanoHUB in the scientific literature. Each of these citations has authors, organizations, and a reference to specific nanoHUB resources (typically tools). nanoHUB users on the author lists are explicitly linked. Not all authors are nanoHUB registered users and these authors are entered into Salesforce as "Leads."

nanoHUB Users (Contacts)-

We have imported all nanoHUB registered users (over 216,000) into Salesforce. This includes their simulation activities, registration date, and profile information. We now associate users with participation in educational clusters, authors of publications, and users of specific tools. Each user's information includes the number of tools used and simulations run.

In the past, many users did not enter information such as their organization/institution into their profile when registering, so it is now a required field. However, users often fail to update that information when it changes. To supplement missing or incorrect information we derive user organizational/institutional affiliations from email addresses, peers in clusters, and citations, thus building comprehensive user profiles in Salesforce, which allow us to focus and customize marketing engagements.

Leads-

Individuals who do not yet have a nanoHUB account, leads, are identified from co-author lists as well as web page scraping.

III. Engagement Campaign Example

Just under 50% of our simulation users are classified as classroom users through our clustering algorithms. We know

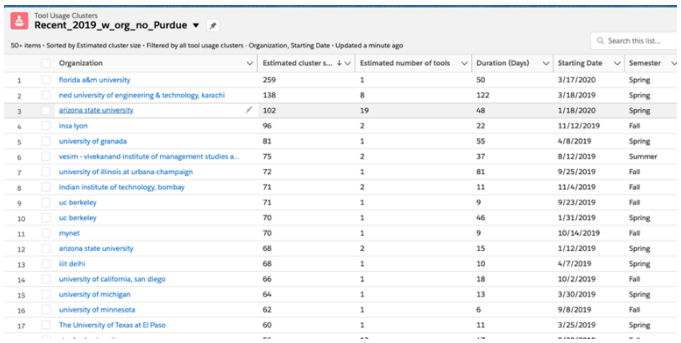

Fig. 3. Salesforce report showing clusters indicative of structured education using nanoHUB tools

that faculty instruct these students with specific exercises. Therefore, each faculty member acts as a user multiplier, and we need to focus on engaging more faculty members into nanoHUB. We want to keep our current faculty members, re-attract faculty that have not been back lately, and market to new, faculty not familiar with nanoHUB.

Inspired by the success of the nanoHUB machine learning (ML) webinars, the team brainstormed on how this approach could be extended towards our most popular toolset, ABACUS. ABACUS is part of our tool-powered curriculum and consists of "all" the simulations needed for teaching a standard Electrical Engineering semiconductor device fundamentals course. Rather than holding a webinar based on a PowerPoint presentation or a lengthy scientific tutorial, we implemented a format consisting of a brief tool demonstration and then provided time for the target audience of faculty to ask questions about the tools and to engage in a discussion. We ultimately homed-in on the term "recitation series," where we assume existing fundamental understanding of the topics, and can focus on the audience's questions.

We executed a marketing campaign for a recitation session series consisting of 7 online meetings where Dr. Gerhard Klimeck reviewed the nanoHUB site, introduced the Semiconductor Device Fundamentals toolkit ABACUS, and then provided in-depth demos of 8 different tools inside ABACUS. These sessions align with the flow of a typical semiconductor device class that includes crystals, electronic band structure, quantum mechanics, carrier distributions, Fermi functions, drift diffusion, PN junctions, bipolar junction transistors, MOS capacitors, and MOSFETs. The 7 sessions were held in early evening US Eastern Time between December 2021 and February 2022.

We created new animations, highlighting one or two specific tools, that were included in each marketing email. We marketed the recitation sessions to 13,324 nanoHUB users who have self-identified as university faculty or staff and to 2,334 leads that we have identified in Electrical and Computer Engineering (ECE) departments. As discussed above, we define a lead as an individual who has not yet created a nanoHUB account and is thus not yet identified as a nanoHUB user. This set of leads were identified through web scraping by two consecutive administrative assistants. The first focused on ECE faculty at MSIs, while the second focused on ECE faculty at R1

institutions. In these campaigns, there were 852 individuals that we identified as being at MSIs (consisting of both existing user contacts and new leads). These marketing emails were sent out originally with a lead time of 3 weeks. Once the recitation series began, we sent marketing emails on the morning following each recitation session, with reminders sent during the following week. The marketing email contained registration links to the WebEx platform in which we held the recitations. Individuals who had registered to attend received a reminder email on the morning of the scheduled recitation. On average, 0.84% of the persons emailed registered for the marketed recitation session.

Fig. 4 summarizes the results of our campaigns, including the number of users reached and the attendance in the recitations. The histogram in Fig. 5 indicates the distribution of repeat attendance by unique individuals. Across the 7 sessions we hosted, we had 316 total attendees, consisting of 192 unique individuals. 137 individuals attended exactly one session and 55 individuals (28%) attended more than one session.

Session	Session Date	Total email marketing	Total Registrants	Total Attendee	New Attendee	Repeat Attendee	Cumulative Unique	Registrants / Marketing (%)	Attendees / Registrants (%)	US - based	US-based (%)
1	12/1/2021	15,658	234	81	81	0	81	1.49%	34.6%	45	55.6%
2	12/8/2021	14,898	99	42	19	23	100	0.66%	42.4%	29	69.0%
3	12/15/2021	14,684	87	31	18	13	118	0.59%	35.6%	16	51.6%
4	1/12/2022	14,567	124	44	23	21	141	0.85%	35.5%	18	40.9%
5	1/19/2022	14,247	99	38	19	19	160	0.69%	38.4%	22	57.9%
6	1/26/2022	14,092	105	39	15	24	175	0.75%	37.1%	21	53.8%
7	2/2/2022	14,032	118	41	17	24	**192**	0.84%	34.7%	23	56.1%
	Average:	14,597	**Total:**	**316**			Average	0.84%	36.92%		55.0%

Fig. 4. Semiconductor device recitation session marketing, registration, and attendance summary as well as engagement percentages for registrants over marketing population and attendees over registrant population.

IV. Summative Engagements in 2021

The efforts discussed above were just one representation of our Salesforce-based engagement. In sum for 2021, we used Salesforce for:

- 35 individual campaigns ranging from individual newsletters to multi-stage invitations to recitations and online workshops
- 1,179,799 total emails sent; 1,112,605 total emails delivered; 239,003 unique email opens; and 22,388 unique clicks in the emails.
- 2,376 total participants in online workshops, recitations and other events hosted by nanoHUB

V. Conclusions

The level of participation by faculty members in the recitation sessions was gratifying and validates this approach as one that resonates with and has value for the community. Given this success and the data gathered about attendees and registrants, we plan to devise additional sessions and follow-up campaigns. For example, in the longer tail of the histogram in Fig. 5 we have 21 (=1+6+5+9) participants that attended at least 4 sessions. These individuals show a very high degree of interest and should certainly be moved into our engagement funnel for direct contacts and follow-up, to help them move nanoHUB into their classrooms. The other individuals who participated should be contacted as well, but perhaps with other messaging.

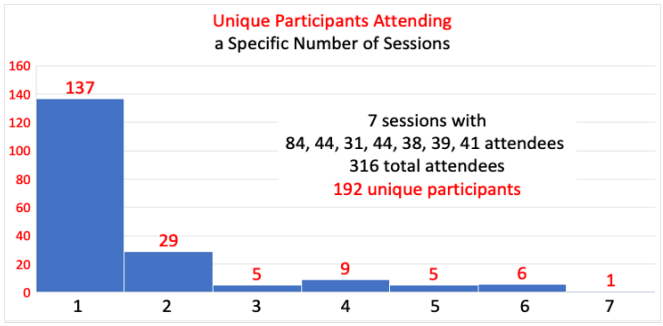

Fig. 5. Histogram of unique participants attending 1 or more recitation sessions. Most participants (137) attended exactly one session. 29 participants attended exactly 2 sessions. A longer tail of repeat attendees is

Another group to contact are the registrants who did not attend. We need to devise a communication strategy with them to learn why they did not attend after registering, and possibly bring them to nanoHUB.

Our registration data show that on average 55% of the attendees were US-based. Going forward, we should consider different time zones in the US, and likely market to Asia separately.

The percentage of registrants compared to the number of people we marketed to seems low at an average of 0.84%. We also need to analyze further if we have cast the net too wide and understand whether there are some specific characteristics in the registrants that we can "see" in our analytics. This might inform future campaigns.

Finally, we have to look at the registration data and attendance data in conjunction with new tool users as well as usage of the recorded recitations that are hosted on YouTube to fully assess the impact of this effort.

Acknowledgment

We would like to thank Travis Merrick on the NCN team for his work in setting up our Salesforce instance and his operational expertise in running campaigns and developing various reports from the Salesforce data. We also thank Praveen Saxena of Appicular for his work on the pipeline connecting nanoHUB data to our Salesforce instance. This work was supported in part by the NSF-funded Network for Computational Nanotechnology Cyber Platform award, EEC-1227110.

References

[1] K Madhavan, M Zentner, G Klimeck, "Learning and research in the cloud", Nature Nanotechnology 8, 786–789 (2013)

[2] K Madhavan, L Zentner, V Farnsworth, S Shivarajapura, M Zentner, N Denny, G Klimeck, "nanoHUB.org: cloud-based services for nanoscale modeling, simulation, and education" Nanotechnology Reviews. Vol. 2, Issue 1, pp 2191-9097

[3] https://nces.ed.gov/ipeds/use-the-data

[4] http://carnegieclassifications.iu.edu/definitions.php