CISE: Community Engagement of CEB Cloud Ecosystem in Box

Dr. Benjamin Garlington
Ph.D. Candidate, University of Arkansas at Little Rock, USA

***Abstract*—The explosion of digital and observational data is having a profound effect on the nature of scientific inquiry, requiring new approaches to manipulating and analyzing large and complex data and increasing the need for collaborative solid research teams to address these challenges. These data, along with the availability of computational resources and recent advances in artificial intelligence, machine learning software tools, and methods, can enable unprecedented science and innovation. Unfortunately, these software tools and techniques are not uniformly accessible to all communities, mainly scientists and engineers at Minority Serving Institutions (MSI). Cloud computing resources are natural channels to enhance these institutions' research productivity. However, utilizing cloud computing resources for research effectively requires a significant investment in time and effort, awkward manipulation of data sets, and deployment of cloud-based applications workflows that support analysis and visualization tools.**

Keywords—Collaboration; outreach; engagement; narrowcasting; conceptual; methodological; storage-as-a-service; software-as-a-service; data-as-a-service; infrastructure-as-a-service; platform-as-a-service; cloud-ecosystem; minority serving institutions (MSI)

I. Introduction

The current structure of the STEM education research enterprise highly favors individuals who can spend nearly full time on research and so take an idea from inception to publication at a competitive pace. This model is great for individuals at R1 institutions, while new scholars at smaller or less connected institutions cannot leverage communal connection to information supported by new cyberinfrastructure. The last two decades' research data has become a major defining force for America of its economic, social, and national importance. Over the last two decades, scientists within all sectors of STEM have witnessed an explosive rise of data ecosystems as a research tool, and universities are utilizing these infrastructures and their associated applications as an integral part of scientific discovery. The most commonly use tool is cloud computing environments; these ecosystems have emerged as a central tool in the scientific evolution of U.S. STEM rebirth. Additionally, the U.S. has become aware of the urgent need to educate a workforce in all sectors of STEM to take full advantage of these shared resource to ensure that the U.S. remains competitive in STEM research and scientific innovation [1]-[5]. Complicating this task is the diversity of research data that is being placed in such environments. This data is heterogeneous, and the volumes are unprecedented in scale and complexity. It represents the next frontier for innovation, competition, and productivity in research [19].

Therefore, primary intellectual thrust of our research will come from a vision of providing a holistic pathway integrating multiple stakeholders, data sources, and training models to address both immediate and long-term data needs for researchers at MSIs utilizing cloud environments as their primary medium for research. Our project develops a service-oriented resource that assistances diverse communicate, navigate, connect, and inform the research activities of these communities. As a result, the Cloud Ecosystem in a Box (CEB) hub will have a transformational effect for these institutions to by enhancing and propelling research among minority communities and groups currently underserved.

The focus of our particular aspect of CEB will be around engagement that creates a theoretical framework of a network of practice[8], in which individuals can leverage a knowledge portal or hub to facilitates the diffusion of knowledge among a rather loosely coupled and often disconnected community [9]-[11], in certain disciplines i.e. (1) Cloud computing, (2) Nanoscience, (3) Bioinformatic, and (4) Cybersecurity.

Engaging these institutions with multi-disciplines, puts emphasis on data-enabled areas of research that can be achieved within the hub, which is a pivotal area of research in information science that examines how engagement within the context of educational activities can broaden and enhance community engagement with information sources or data.

II. Problem

Across the US, research has been broad and only been characterized through individual contributions in each field as separate entities with no single solution [1]. Tier1 research institutions and Universities greatly compensated, have the advantage in high caliber research, so primary data and secondary data mitigate and escalate to and from these campuses. Meanwhile low-income Minority Serving Institutions (MSI) that only make up 10% [1] of US institutions do not have the knowledge, capacity nor the funds to deliver top-tier data to research on, nor do they have the resources to properly engage training. Around half of MSI students have no inclination of cloud resources or what cloud resources can provide according to the results of the initial survey. Majority of Minority Serving Institutions (MSI) have no forefront knowledge of future endeavors on transitioning to cloud services provided by CEB.

Therefore, a grant has been funded through the National Science Foundation's (NSF) Directorate for Computer and Information Science and Engineering (CISE) to create an on-ramp to allocate resources for specific community domains. A funded EAGER Cloud-Ecosystem in a Box (CEB)(award:1842679) project serving the principal interface for providers in the CISE research community, the education community, and the public cloud community to contribute support from corporate partners by leveraging the Cloud Ecosystem in a Box (CEB). This grant extends cloud services to complement existing resources centered around completely different computational faucets of research. These recommendations were made at the NSF Cloud for Everyone (NSFC4E) workshop (award:1663794) also recommendations made at Enabling Computer and Information Science and Engineering Research and Education in the Cloud (ECISE-REC) workshop(award:818650).

Researchers all over the United States are over-exerted due to the volume of data being computed and stored over the vast cloud movement of processed data and crowd-sourced learning. This data has increased the number of activities to be hosted in particular with, Minority Serving Institutions (MSI), in which, these facilities has resulted in a great number of domain specific publications. In addition to the volume of this unorganized data, new findings are growing at an exponential rate and it has become more time consuming for researchers to draw conclusions within this area of collection of big data due to the lack of addressing engagement.

A. Dissertation Statement

Minority Serving Institutions (MSI) should have access to services such as building a cloud-based community to enhance research and infrastructure connected together. By using these engagement techniques, we can increase cloud-driven resources by 20%. Measuring analytic collaborations by building state of the art research and educational cloud ecosystems that leverage advance practices and training within a computational sphere.

B. Research Area Problem

How a Hub (CEB) that is an on-ramp onto cloud resources (via AWS, Google, Microsoft, Oracle) could serve Minority Serving Institutions (MSI) community needs and user needs.

With the focus of my research exerting towards Community Engagement by (1)building state of the art research and educational cloud ecosystem (CEB) that leverage advance practices and training within a computational sphere with Cloud providers, and (2)systematically engaging in response time and success of the CEB, plans are put into effect of a more precise nature to establish a mentoring network between Minority Serving Institutions (MSI), through contributing community/user-based needs respectively.

Also maintaining the CEB by being a beacon/liaison for community toolsets to train Minority Serving Institutions (MSI) for cloud-based learning in Cloud *Collaboration* through the life cycles (Phases) of the CEB. This enables superior cloud resources from corporate partners to be at the forefront of separate cloud research tools for a broader research area instead of local centralized research, to accelerate innovation through user-defined computational cloud research while also addressing retention and scalable marketing.

III. Literature Review

This literature review is designed to share similar work corresponding to cloud services provided through an ecosystem like the CEB. It will also compare complexities and pedagogies tied into the various levels of expanding cloud services to areas never incorporated. It will serve as a benchmark for new advancements within cloud services by aiding in filling gaps within the discipline and administering techniques for measuring success through the CEB expounding on the measurement of outreach, engagement, cloud ecosystems**.** I would like to start this literature review with XSEDE (Extreme Science and Engineering Discovery Environment) cloud services, because it is the closest infrastructure related to the CEB that I am most familiar with, also I have extensive expertise within this particular pedological area of cloud resources. I was reluctant enough to be a part of the services provided by XSEDE at the University of Arkansas at Pine Bluff where I practiced as the Campus Champion. My role expanded across a variety of activities such as helping institution researchers, educators and scholars with their computing needs. As the Campus Champion, I was engaged in high-end data intensive research focusing on large scale, high performance computing, distributed high throughput computing environments, visualization and data analysis systems, large-memory systems, data storage, and cloud systems. In addition, I attended/chaperoned the Supercomputing convention.

1) In the paper "XSEDE: Accelerating Scientific Discovery", cloud resources are characterized as a co-existence between multiple cyberinfrastructure communities of scholars, scientists, engineers and researchers to form one ecosystem that addresses the most important and challenging problems [2]. XSEDE's realm of computational technologies and resources divulge in a process that is critical to the success of researchers. These digital services are accessed at a fraction of the cost that they ordinarily would, due to the outsourcing of their complexity to XSEDE. These advanced digital services make up different attributes of technology like supercomputers, storage systems, visualization systems, collections of data, networks and software. The paper emphasizes the motivation to integrate these computational services into a more general form as [2] "judiciously distributed but architecturally and functionally integrated from two arguments": First, the movement of scientific progress within multiple disciplines, which enables a combination of services and resources. Second, that top-tier digital resources are more capable by leveraging multiple smaller institutions, and not being centralized fully at a single institution. XSEDE, as an ecosystem of its on, has an ongoing interface for help-desk support, initiating allocations, and services for sharing other computing digital resources. XSEDE implements tools, policies, and methods that exist with top-tier facilities, campuses, and government institutions as the beginning process for a national e-science infrastructure ecosystem. [3].

2) The paper "open-source cloud management platforms and hypervisor technologies: A Review and Comparison", suggest further implantation of cloud-based solutions and measuring performance in an infrastructural environment [4]. It introduces cloud infrastructure and hypervisor technologies as a means to architect effective cloud solutions, also the need for continues emerging of new tools and technologies used to fit the needs of infrastructure ecosystems. Something very important within cloud infrastructure is what Cloud Framework should be implemented according to needs of the Cloud Service Provider. Software-as-a-Service (SaaS), Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS). It goes into intense discussion about Cloud Management Platforms used to deploy virtual machines and how a major consideration for the success of any cloud platform involves the activity and support of the development community behind the project. Hypervisors are introduced as a special operation system that hosts multiple Virtual Machines on a physical machine that allows simultaneous execution of entire operation system instances at the same time. The paper shows how when you combine several virtual machines onto a single physical server it provides on-demand computing to users in real-time. Discussion took place regarding identifying the best cloud technology through different studies, but in today's infrastructure needs it is enough to sustain several open-source cloud computing platforms that addresses difficulties in cloud environment [4].

3) In the paper "DH Box: A Virtual Computer Lab in the Cloud", it outlines how instruction in Digital Humanities (DH) became more efficient since the creation of the DH Box [3]. Challenges within the classroom before the DH Box consisted hardware and software resources and installment/configure of DH tools running of multiple devices. Run-ins with IT departments played a major part in the attrition of facilitating the classroom and he states that on most campuses, computers were on complete lockdown. The author wanted to develop a portable, reproducible, and approachable way to get people interested in DH, up and running with DH tools with the creation of the DH Box starting out on a Raspberry Pi then graduating to the cloud. From local installs on user hardware to a web application that would be accessible remotely with the infrastructure that consists of a sign-in, front-in website, a back-end that listens for request, takes data and initiates the launch of a script, then the launch script itself, and a simple user menu. The DH Box had cut down on class time devoted to preparing technology, shifting the focus away from maintenance and configuration and toward learning and exploration where students were able to access DH tools from their own devices of their choosing [3].

4) "STEM Outreach: A Literature Review and Definition" is a paper that introduces a pragmatic definition of STEM outreach using a heuristic contextual framework [5]. Without expounding on the ontological results used to complete the definition, this paper concludes the definition as:

"The act of delivering STEM content outside of the traditional student/teacher relationship to STEM stakeholders (students, parents, teachers…) in order to support and increase the understanding, awareness and interest in STEM disciplines."

Outreach delivery methods within this paper include Lecture, Active learning/Inquiry Based learning, Problem Based learning, Workshops and Training Events, and Camps and Events. Examining the College (Undergraduate through Doctoral) target area we see how college students can be both the target audience for STEM outreach programs and also to outreach efforts for younger age groups. Through an extended conference/camp where groups of target criteria can emerge educational support, we see an increasing number of providers and more program support. College students get enhanced networking and more exposure to a particular field of study.

5) In the journal "CARES: Mentoring through University Outreach", characterizes a study which utilizes surveys, campus visits, interviews, and observations for a minority serving middle school [6]. These attributes are at the core of outreach, and within this study proves results to show reluctant change and improvement. The participants were students of the forementioned middle school was paired with a mentor selected and matched by race/ethnicity. The findings in this article articulate three major themes. First, the Cares act increased student's aspirations to attend college by providing access to information that expanded the realms of knowledge and exposed new thoughts and the idea that learning is fun. There are 4 categories that the article demonstrates within the first theme that constitute them for increasing the willingness for higher education: Exposure to new options and ideas, increased sense of purpose, feelings of acceptance and connection with mentors, and recognition of encouragement from mentors. Second, mentors experienced personal growth and benefits from their involvement in the Cares act. This second theme has influenced testimonies from the mentors themselves. All acknowledged the benefits, ranging from rejuvenation and increased motivation to growth in their skills and a sense of emotional satisfaction from seeing their influence on students' futures. Thirdly, the outreach from mentors to the students had decreased the negative socialization that characterized the educational experiences. School policy dictated that education was punitive and boring in which it discouraged them form considering higher education as an option. This outreach shows researchable mitigations in the inequities between the differences of white middleclass students and those of poor African American and Latino students. The policies and procedures within this middle school was evaluated because of the CARES act and staff and faculty came to see how the discipline policy discouraged students.

6) In the journal "Outreach, Engagement, and the Changing Culture of the University" the author narrates that in order to change the culture of an organization there has to be a noticeable crisis [7]. A recognized need for higher education to keep pace of an ever-developing societal change takes a forward standpoint pertaining to the culture of education. Outreach and Engagement are characterized by a reaching out from the university to the people and organizations a university serves, also the exchange between the university and its constituents. The author states that engagement is both outreach and in reach into the university by way of extending scholarly creativity to enhance cultural changes inside and out.

7) In the journal "*Outreach initiatives in academic libraries, 2009-2011*", the purpose of outreach with academic libraries is reprimanded with budget cuts declining, also outreach is still on the rise [8]. This particular study represents comparing outreach initiatives through librarians to a particular project conducted by the author. The approach taken was sending a survey to academic librarians that influenced higher learning with outreach, marketing, and promotion. There were successful outreach initiatives in all categories and only a small response rate reflects the target audience. The findings of the survey conveyed a wide range of initiatives involving outreach that target funding. The authors project exceeded the cost of all other initiatives. Certain limitations identified survey flaws and the implications was just a small reach to the targeted audience. A modified survey to the academic library system involved in outreach and engagement is therefore included for future research. The practical implications hardly coexist with budget restraints, because many higher education programs across the nation have inserted outreach into public services positions. The value this type of outreach brings to the table produces limited research about successful initiatives within the spectrum of the timeline. Outreach from academic librarians will be inspired in the collected project outreach takes to create projects with limited funding.

8) In the paper "Pre-university Outreach: Encouraging Students to Consider Engineering Careers" outreach is discussed by activities being affected in educating students about the challenges and rewards of engineering careers [9],. It discusses the impact of certain programs motivating students to consider engineering educational pathways and career pathways. Originally the program was for female students partaking in a summer day camp. It now includes hands-on activities directly to the classroom presented as a co-ed program. There were multiple workshops ranging multiple years that combined males and females, then ranged with the classifications of each individual. The surveys portray a ripple effect with the percentages of K-12 knowledge awareness of engineering. It represents a convergence of STEM intelligence over the years that outreach becomes more and more successful. Proof of the before and after evaluations of student's knowledge of engineering. The questionnaires and evaluations give direct awareness for essential tools about the engineering profession at the pre-university level. It is still in its early stages to tell the increases in expansion of engineering enrollment, but the numbers don't lie about having a positive impact on women in the field.

Whitney Bouk general manager of enterprise at Box said quote: "Future success for students and educators will be dependent on how well we integrate technology into our modern learning environments, Box helps schools to teach, conduct research and run their operations more creatively and effectively, while also giving students new ways to connect and learn. With our growing network of education partners, we are enabling students and teachers to take ownership of education [10]".

Matthew Self Box VP for Platform Engineering said quote: "We're transitioning from a world where the enterprise, and all of your enterprise content, are within the confines of your network and with applications you control, to this world where you have the cloud, all these different applications and all these different devices, with their content scattered all over the cloud. What Box is introducing is a solution… that will bring all of that content to a single place. By using Box OneCloud, we can leverage all of the applications that businesses want to use to run their business and store the data in a single place [11]".

Liz Herbert, an analyst at Forrester Research said quote: "An interconnected network of add-on applications vendors, consultants and partners centered on an anchor cloud provider — and the benefits of a cloud computing ecosystem are outlined. It's a way for organizations to balance agility and speed with control [12].

IV. Methodology

A. Initial Survey

1) Stage/Plans of adoption for research topics: The cloud computing technology fields provided in this question include Private/Internal Cloud, Storage-as-a-Services Only, Infrastructure-as-a-Services, Platform-as-a-Services, and Software-as-a-Service. The results show a top-down approach from the majority results to less inclined results based on answers given, also the same given the sub results of each category. The percentage equated for each category is calculated as the mean, and the sub results are the actual percentages articulated in the summary.

a) Have no plans to use any of the Cloud Computing Services- 41.33%: This first category ranks top priority among Minority Serving Institutions (MSI) survey results schematic. 41.33% Say they have no plans to use any cloud services. This tells me that research could be non-existent in almost half of the Minority Serving Institutions (MSI) participating in this survey. Having no plans to utilize cloud resources is a sure-fire reason why outreach should be manifested as a key component to the growth of desired infrastructure to move the cloud services via existing cloud resources to new communities.

- 50% Have no plans to use Infrastructure-as-a-Service.
- 43.33% Have no plans to use Private/Internal Cloud.
- 43.33% Have no plans to use Platform-as-a-Service.
- 40 % Have no plans to use Software-as-a-Service.
- 30% Have no plans to use Storage-as-a-Service.

Minority Serving Institutions (MSI) deemed Storage-as-a-Service more of a priority within this spectrum of down-top results. Which means Storage-as-a-Service is a more important field based on the results.

b) Using in Production- 18.66 %: This next category is the next top priority among Minority Serving Institutions (MSI) survey results schematic. 18.66 % Say that these stages/plans of adoption will be used in Production. Within this stage, the design, simulation, production, test, and maintenance of a product provides the whole lifecycle of the concept called cloud manufacturing. It provides a new paradigm developed from existing models under the support of cloud computing consisting of integrated and inter-connected virtualized resources for intelligent and on-demand use of services to provide solutions during the lifecycle of manufacturing [13].

- 33.33% Storage-as-a-Service will be used in Production.
- 23.33% Private/Internal Cloud ranks will be used in Production.
- 13.33% Infrastructure-as-a-Service will be used in Production.
- 13.33% Platform-as-a-Service will be used in Production.
- 10% Software-as-a-Service will be used in Production.

Again, Minority Serving Institutions (MSI) deemed Storage-as-a-Service more of a priority within this spectrum of results in Production.

c) Using for Development and Testing only- 8.67%: The third top category 8.67 % Say that these stages/plans of adoption are used for Development and Testing Only. It can be very costly and time-consuming to build the infrastructure for testing in a local environment. Scaling up resources when needed without investing in the infrastructure is what is utilized with cloud-based software development [14].

- 13.33% use Storage-as-a-Service.
- 10.00% use Private/Internal Cloud.
- 6.67 % use Infrastructure-as-a-Service.
- 6.67 % use Platform-as-a-Service.
- 6.67 % use Software-as-a-Service.

Once again Minority Serving Institutions (MSI) deemed Storage-as-a-Service more of a priority within this spectrum of results for The Development and Testing only Category.

d) Plan to use but not sure when- 18%: 18% Say that these stages/plans of adoption are Planned to be utilized but not sure when.

- 30% Plan to use Software-as-a-Service but not sure when.
- 20 % Plan to use Platform-as-a-Service but not sure when.
- 16.67% Plan to use Infrastructure-as-a-Service but not sure when.
- 13.33% Plan to use Storage-as-a-Service but not sure when.
- 10% Plan to use Private/Internal Cloud but not sure when.

Minority Serving Institutions (MSI) ability to looked forward and distinguished to use Software-as-a-Service as a priority. Maybe this survey is pushing researchers, students, and the public to recognize where resources are needed, and Software-as-a-Service holds grounds for future endeavors.

e) Plan to use within: 3-6 months- 6.67%: 6.67 % Say that these stages/plans of adoption are being Planned to be used within 3-6 months

- 6.67% Private/Internal Cloud.
- 6.67% Storage-as-a-Service.
- 6.67% Infrastructure-as-a-Service.
- 6.67% Platform-as-a-Service.
- 6.67% Software-as-a-Service.

f) Plan to use within: 3 months- 3.33%: 3.33% Say that these stages/plans of adoption are being Planned to be used within 3 months

- 3.33% Private/Internal Cloud.
- 3.33% Storage-as-a-Service.
- 3.33% Infrastructure-as-a-Service.
- 3.33% Platform-as-a-Service.
- 3.33% Software-as-a-Service.

g) Plan to use within: 6-18 months- 3.33%: 3.33% Say that these stages/plans of adoption are being Planned to be used within 6-18 months

- 3.33% Private/Internal Cloud.
- 0% Storage-as-a-Service.
- 3.33% Infrastructure-as-a-Service.
- 6.67% Platform-as-a-Service.
- 3.33% Software-as-a-Service.

2) Applications planned to deploy over a 12-month period: This question makes a relevant claim for CEB Cloud Services. Getting to know what Minority Serving Institutions

(MSI) plan to do as a whole can help mitigate activities involved in computing by transferring from a local machine location to cloud services online. We need to know what Minority Serving Institutions (MSI) have already done, that way we have cloud resources available on-top of already existing local resources.

- 29.87 % Say that Research Activities would be planned to be deployed over a 12-month period.
- 19.48% Say that Collaboration would be planned to be deployed over a 12-month period.
- 16.88% Say that Office productivity (e.g., Google Docks, Zoho) would be planned to be deployed over a 12-month period.
- 11.69% Say that Testing/QA/Staging would be planned to be deployed over a 12-month period.
- 7.79% Say that Web 2.0/Social Networking would be planned to be deployed over a 12-month period.
- 7.79% Say that Other(unspecified) would be planned to be deployed over a 12-month period.
- 3.90% Say that Monitoring/Application Performance Management would be planned to be deployed over a 12-month period.
- 2.60% Say that Security/Compliance would be planned to be deployed over a 12-month period.

3) The Initial adoption of cloud computing within MSI was done by: It is important for students or researchers to know the initial adoption of cloud computing in their institution. This is a question to see how literate the research community is with these services among their institutions. We find out how to distinguish as a whole among Minority Serving Institutions (MSI) where the discrepancies are, when it comes to cloud computing Literacy and direction.

- 29.17% Are not sure how the initial adoption of cloud computing was done on their campus.
- 25.00% Say Faculty for file sharing (i.e., Dropbox, OneDrive, Box.com).
- 16.67% Say A CIO/Senior IT Exec decision.
- 16.67% Say A departmental project.
- 12.50% Say A development team.
- 0% Say Cloud Infrastructure is not used at my institution.
- 0% Say other.

4) Current State of Adoption (12-month plans): This is a follow up question from the previous survey question The Initial adoption of cloud computing within Minority Serving Institutions (MSI). By understanding the initial adoption of cloud computing on your campus, you can give an interpretation of the current state of adoption or a 12-month plan. Here we are interested in a 12 month plan due to the existence of cloud computing resources available via contingent on the results of this survey.

- 52.94% Says the adoption of cloud research remains the same.
- 29.41% Says the adoption of cloud research not sure.
- 11.76% Says the adoption of cloud research includes additional applications within the same department.
- 5.88% Says the adoption of cloud research includes additional applications across departments.
- 0% Says the adoption of cloud research Organizational-wide mandate to adopt cloud computing.
- 0% Says the adoption of cloud research other.

5) How many employees work at each institution?: This question is important because it gives focus on underrepresented Minority Serving Institutions (MSI). The idea of minority-serving institutions Minority Serving Institutions (MSI) having adequate high caliber research opportunities, or state-of-the-art facilities to collaborate research have been researched for years but minimally implemented. The caliber of research for Minority Serving Institutions (MSI) should be fundamentally congruent with resources and funds equal to or of greater depth with tier 1 institutions. The higher the funds and resources allocated; the higher caliber of research accomplished.

- 41.18% Of Minority Serving Institutions (MSI) employee 500-1000.
- 35.29% Of Minority Serving Institutions (MSI) employee 100-500.
- 11.76% Of Minority Serving Institutions (MSI) employee 5001-10000.
- 11.76% Of Minority Serving Institutions (MSI) employee more than 10000.
- 0% Of Minority Serving Institutions (MSI) employee 1001-5000.

6) Mission critical applications for research: This question gives momentum to cloud computing and is the catalyst that drives cloud resources. There are projects that require sophisticated resources for big data. With this type of adaptation researchers can use this process to integrate research-oriented cloud ecosystems that operationalize funded research projects concepts within Minority Serving Institutions (MSI).

- 47.37% Say they do not consider cloud computing mission critical for their research.
- 28.95% Say they do consider cloud computing mission critical for their research.

- 23.68% Say they are unsure to consider cloud computing mission critical for their research.

B. Defining User Communities and User Desires

1) Define User Communities

a) Cloud Computing: Cloud Computing itself encompasses services rendered through the CEB to all the communities as services over a secured network purely online. On-demand services provided are HPC, storage, networking, analytics, databases, software, and AI delivered for faster innovation of particular research to achieve coherence and economies of scale [13]. Thus, *cloud Computing* drastically enhances the economic scale of research, and resources become flexible because the delivery mechanism encapsulates the services provided by overlapping platforms. The services provided are Infrastructure-As-a-Service (IaaS), Platform-As-a-Service (PaaS), Software-As-a-Service (SaaS), and Storage-As-a-Service (STaaS). These services overlap each other for full transparency of chosen delivery methods based on user domains.

b) Nanoscience: The study of extremely small things, the manipulation of materials on an atomic or molecular level to build microscopic devices. The Nanoscience community recognizes support of simulation programs in the cloud as to mitigate away from locally running software. The research tools and methods needed are aimed at fabrication methods and measurement tools along ways of measurement in regard to the characterization, design and synthesis within the ecosystems in place with goals responsible for commercialization of nanotechnology [15].

c) Bioinformatics: Bioinformatics is an interdisciplinary field entailed of Computer Science, Biology, Information Engineering, and Mathematical Science/Statistics that analyzes the interpretations of various data types. The role of *Bioinformatics* in cloud research extends to computation counterparts of analyzing and interpreting data. Large amounts of CPU time for tasks carried out in decomposed parallel development require users to work intermittently and facilitate continuous integration. This platform requires analysis to come to the data and to which the datasets expand in size.

d) Cybersecurity: Cybersecurity research is studying the defense of cloud systems from online attackers for the preservation of integrity, confidentiality, and availability of information in the cloud. These attacks take many forms i.e., Malware, Virus, Trojans, Spyware, Ransomware, Adware, Botnets, etc. Through *Cybersecurity* roles played in the cloud, there lies architecture built on top of layers spread across computing paradigms that hold different defenses on separate platforms. The technology used within this research entails of tools for protection from three main entity endpoint devices like: local/cloud/hybrid computers, routers, and smart devices. Also, technology used include, DNS filtering, antivirus software, malware protection, next-generation firewalls, and email [16]. As increased digitalization of human activities increase to improve lives by transferring data to the cloud, this increases vulnerability to being attached within the cloud.

2) Define User Desires

a) Storage-As-a-Service (STaaS): Reported in the survey results MSI students who had somewhat of a degree of cloud computing or knowledge base on their institution says *Storage-as-a-Service* is what is needed at the current state. This is a top priority on the existence of the CEB resources being allocated through the NSF grant to get the CEB Hub off the ground. Research suggest that storage is a top priority in the cloud for small and medium businesses, universities, and organizations outside of the CEB so there are resources already available to help transition from local computing to cloud computing. They are extremely scalable pay as you go system and easy to manage and maintain.

b) Software-As-a-Service (SaaS): This is the second service that serves the needs of Minority Serving Institutions (MSI) within a 12-month level plan to transition to cloud resources. A way to license hosted software for an extended period of time without purchasing or require a pay-as-you-go per user subscription of actual usage. *Software-As-a-Service* is an infrastructure that is most popular/common service over the internet. It correlates everything covertly that *(Paas)*, (*Iaas)*, and *(STaas)* is made up of infrastructurally which makes it more mature and to some extent more standardized or tailored software providers [17].

c) Platform-As-a-Service (PaaS): Survey conveys Minority Serving Institutions (MSI) plan to use *Platform-As-a-Service* within 6 to 18 months as predominant adoption plan for cloud research. It allows complementary software resources as a development and deployment environment to capitalize on this infrastructure as a set of tools to build on-top of CEB resources. Within the time frame *Software-As-a-Service* will pick up off the ground, and top-tier research can be authenticated with training and collaboration.

C. Understanding Firebase as a Backend Service

1) Firebase Authentication Services: Firebase backend services connected to the CEB authenticate users using username and password, phone numbers and federated identity, linking users accounts to social media providers into a single account on Firebase. Identification processes take an exhausting amount of time with traditional backend development, and Firebase takes that process out so you can focus on your users and not the sign-in infrastructure to support them. By knowing the identity of the user and creating an Authorization state we can customized an experience and keep their data secure, and we give them a choice of presenting login to the user. Either we can build our own interface or take advantage of Google opensource user interface (UI).

a) Firebase Authentication SDK: The identity of our users is crucial for the data to be securely saved in the cloud for a personalized experience across all devices. In order to accomplish this, Firebase Authentication tightly augments with other Firebase services leveraging standards of years of trial and error in the making of google services so it can be easily concatenated with a custom backend.

- Email and password-based Authentication – Create and Manage users through email and password sign-in. Password reset emails.
- Federated identity provider integration - (Google/Facebook/Twitter/GitHub/Apple/Microsoft/Yahoo).
- Phone number Authentication – SMS messages to phones.
- Custom auth system integration – Existing sign-in system to gain access to Firebase Realtime Database/Services.
- Anonymous Authentication - Create temporary anonymous account. If the user later chooses to sign up, you can upgrade to a regular account so the user can continue where they left off.

2) Firebase database system: Firebase's Cloud-Based Database query-based retrieve, sort, and filter data system supports real-time data syncing with no servers to deploy or maintain, it has real-time updates, and pay-as-you-go infrastructure. This is a NoSQL scalable infrastructure to store and sync data for server/client development. Firebase database systems carry mobile-first, real-time SDKs (software development kits) also in support of on-site storage for programs offline. Firebase offers two Database systems: (1) *Realtime Database* and (1) *Cloud Firestore*.

a) Realtime Database: Data is stored as JSON and managed as real-time instances that automatically receive updates within rich applications allowing secure access to the database. It provides a flexible language called the Firebase Realtime Database Security Rules which is an expression-based language used to define how your data access I/O, also defines how it can be structured. When paired with Firebase Authentication you can set up who has direct access to the communities and what services are provided through each community. This simple database system is set up for quick execution unlike an original legacy relational database in that it is set up for the whole user experience serving users of communities building real-time experiences without compromising on responsiveness.

b) Cloud Firestore: Like Firebases Realtime Database which allow for low-latency solutions for user devices requiring synced states across multiple mobile app platforms in real-time. Cloud Firestore builds on top Realtime Database software to create new, and more scaled data model offering rich, faster, and scalable queries that are further implemented based on the data. Following NoSQL pattern of database design, you store the data that support many different types, from simple strings and numbers to complex objects, in fields that are then stored in collections called containers that you can organize the build queries from. Querying in *Cloud Firestore* is expressive, flexible, and efficient. It creates shallow queries from documents without retrieving the entire collection, or subcollections. You are able to sort, limit, and filter into your queries to paginate your analysis in real time, also you are able to protect access to sufficient credentials through Firebase Authentication.

3) Firebase storage: Storage, sync, upload and backup time capsule, and key configurations for Macs, Windows, etc. and different mobile devices, which makes them immediately available, so you can have easier ways to collaborate. Firebase cloud services allow to share important files without the hassle of attachments, but instead gives a link option to download in order to preview, you can set passwords and expiration dates for added security. Having access to be able to preview files before downloading is a perk with the iCloud system for the exception, that, in order to edit them, you will have to download the attachments then upload back to Firebase. Also, the cloud services have the ability to save the latest version of any paper, so you can have the latest document available every time while the older versions are also saved for reference. It gives flexibility to share folders and files with colleagues or the ability to set permissions for security. Basically, having private and public sharing at the same time. As with other familiar services, Firebase can create new documents, but has the ability to assign tasks which stops the need for data task management teams.

4) Firebase machine learning: Machine Learning is a small section of AI which incoherently focuses on training algorithms for data analysis that automates analytical model building. A process where tasks can be created without human interference specifying every step of an algorithm. This is where computers learn on their own from represented data where there are no fully satisfactory algorithms to accomplish a task. Firebase recognizes the need for cloud-based processing, the real-time capabilities of standalone devices, or the flexibility of ML resources geared toward cloud-based or mobile-optimized devices. Cloud-based devices include services like text recognition, landmark recognition, and image labeling application programming interfaces, perform inference in the cloud so they require more computational power for HPC, also require more memory so they can perform with greater accuracy and precision. Mobile devices include services for custom model deployment by uploading to the firebase server creating on-demand capabilities, and AutoML Vision Edge that is an app to help create custom image classification models with user-friendly web interfaces. These services are suitable for real-time and low-latency applications provided by optimized to run on mobile devices that don't require a network connection and runs very quickly. Video processing is best suited for processing frames of video in real-time. Firebase makes it simple with a service called the mobile SDK package to provide insights about the data provided.

a) Firebase Predictions: Firebase Predictions leverage ML analytics data, creating behavior influenced clusters of users so you can make research decisions based on predicted behavior rather than historic behavior. Firebase Predictions is consolidated with other ML resources mentioned above for cloud-based or mobile-device customized user experiences so you can adjust the frequency of ads. You can increase retention by re-engaging a user who has predicted to opt-out of Firebase by sending a one-time campaign or automate sending messages for certain groups predicted to not return. Firebase also gives you the option to create your own predictions by creating a cluster based on prediction of any event defined in the analytics data in the near future. All this data is easily transferred to Googles *BigQuery* for further analysis for pushed into third party services.

b) Natural Language Processing (NLP) / Natural Language Understanding (NLU): Natural Language Processing reluctantly is concerned with finding the interactions between human language computers by analyzing large amounts of language data and processing the contents for the contextual nuances then extracting information and insights from the data then organizing the documents by categorizing them under the pretenses of linguistics. Text recognition applies to automating data entry like credit cards, business cards, and receipts. Also extracting text from pictures for translating documents and increasing accessibility. It is optimized for both photos and documents, recognizes text in hundreds of different languages. Firebase Language Identification is used to determine the language of user provided text, which naturally does not come with any uniform language information of a string of text. The capabilities of text translation are limited for casual and simple translations using dynamic storage for downloading and on-device requirements by managing language packets. Firebases Smart Reply API automatically replies to messages by generating multiple reply suggestions based on the context of strings of sentences through a conversation. This API is intended for just normal conversations. It does this by comparing the messages against a list of sensitive topics, then the model provides up to three suggested responses dependent on a sufficient level of confidence.

V. Preliminary Results

We started out addressing engagement of the CEB by (1) Defining User Communities, and (2) Define user desires based on survey result outcomes. The communities we are building are created from pain points within Minority Serving Institutions (MIS). We will continue address engagement of MSI's by measuring analytics of the CEB through response time through the retention/life cycle.

The results of the survey question Stage/Plans of adoption conclusively show that the overwhelming majority of Minority Serving Institutions (MSI) participating in this initial survey rule that cloud resources are novel. Meaning there is an unfamiliarity of resources beyond each institutions' limits on education. Based on this initial survey results, Storage-as-a-Service is a top priority when utilizing these cloud computing resources for the majority of Minority Serving Institutions (MSI) independently for the immediate moment. Next, Software-as-a-Service and Platform-as-a-Service is considered a priority for the future growth of the majority of Minority Serving Institutions (MSI) independently.

The adoption of infrastructural cloud services Cloud Ecosystem in a Box (CEB) across multiple campuses is integrated through a wide network that not only recognizes the need of increased viable research activities, but dedicated research in the classroom as well. Research Activities is ranked the most important at 29.87% to be utilized within the 12-month time frame. Collaboration, which according to the Survey Outcomes, holds 19.48% will also be an immediate undertaking within Minority Serving Institutions (MSI).

For the initial adoption of cloud computing on an MSI campus, here we intercept the knowledge gap and proceed to fill in the necessary timestamp that accrues when Cloud Computing education is not congruent with Minority Serving Institutions (MSI). There is a 29.17% majority of students that are not sure how the initial adoption of cloud computing was done on their campus. Next, we see that 25% of Minority Serving Institutions (MSI) say the initial adoption of cloud computing was done on their campus by Faculty for file sharing, 16.67 % say a CIO/Senior IT Executive decision, 16.67% also say a department project, and 12.50% say a Development Team. Here we see a portion of Minority Serving Institutions (MSI) collaboration on each individual campus. With the numbers being so low in comparison, we see the discrepancy with not only a shortcoming in cloud resources but an existence of ignorance (lacking knowledge or awareness) on each individual campus as well.

The results are conclusive with The Initial adoption of cloud computing within Minority Serving Institutions (MSI) when we see there is a knowledge gap of more than a 4^{th} of students that are not sure how the initial adoption of cloud computing was done on their campus. As a result, we see 52.94% of Minority Serving Institutions (MSI) say the adoption of cloud research remains the same, and 29.41% says they are not sure. As stated previously in The Initial adoption of cloud computing within Minority Serving Institutions (MSI) conclusions, outreach should be manifested as a key component to the growth of desired infrastructure to move the cloud services via existing cloud resources to new communities.

The majority of our small MSI' 41.18% make up 500-1000 employees. The other majority of 35.29% make up 100-500 employees. Due to this underrepresented institution research sector which consist of mainly teaching institutions, both small and large [1], outreach should be attainable to sustain and utilize these resources by establishing a mentoring network to embody a relationship between mentor and student bodies with the boundaries set within each sector.

47.37% Say they do not consider cloud computing mission critical for their research. In response to nearly half of Minority Serving Institutions (MSI) response that they do not consider cloud computing mission critical for their research, again there is a trend and a consistency within these survey results that points to an education relapse of cloud computing. Previous research regarding this particular domain not only fails to address the logistics of how to properly incorporate an ecosystem, but to share, analyze, and draw insights from mined data.

1) Storage-as-a-Service: Transferring data from a local machine to the cloud can be reluctantly smooth, given the proper guidance and planning. Training as a top priority for outreach will be mandated to take steps to reduce any risks involved. The first thing that needs to happen before any research group starts to transfer, is to have a viable back up stored on a local machine. Reason is you don't know of any (1) security risks, (2) network connection interruptions, or (3) improper file organization. During transferring of data your files are pretty much in limbo or left in the hacker's world. Interruptions related to the network that has the network connection or the strength of the network can play a part in the loss of data. After the successful transfer of data from local to the cloud, you may find your data all in disarray. There are ways to find and organize your data, but keep in mind that keeping a backup will enable you to locate the right ones immediately. Next, research needs to be taken place on who will be a suitable cloud service provider.

CEB Cloud storage allows for offline availability and synchronizing/sharing of files on all API's providing universal access to any and all types of data stored within the CEB Hub. It provides a service layer on top of the storage services already allocated by storing all the data and metadata from the users and provides access from different faucets (i.e., FUSE access, WebDAV, XROOTD, Synchronization client, Web access, Mobile devices). How this is done is by an internal database that keeps track of the migration status for all the users of the system through home directory services. The goal is to consolidate storage solutions (i.e., Batch System, Windows Terminal Servers, Personal computers, Small/Private Experiments) into one service that can be accessed from any API, "incognito" the CEB Hub [18].

2) Software-As-a-Service: How Minority Serving Institutions (MSI) migrate Software-As-a-Service within this 12-month period depends on research capitalization. Collaboration supports the CEB design to support many different users safely and securely with unlimited scalability to support changing research. Coordinated efforts within communities makes the foundation become the anchor for a rich partner ecosystem.

What CEB offers with *Software-As-a-Service* within this 12-month period are resources on many levels of platforms to designated API's. The CEB resources for the communities maintain software that is delivered remotely by providers. The CEB offers ecosystem integrators that use their reputation to convince critical as-a-Service Minority Serving Institutions (MSI) solutions especially in response to security concerns.

3) Platform-As-a-Service: The move to Platform-As-a-Service among Minority Serving Institutions (MSI) will increasingly evolve Big Data correspondence with scalable software delivered in the cloud with pay as you go investment. PaaS is a cloud-based environment for coding and deploying applications that make coding for developers and IT professionals much more efficient. The type of Platform-As-a-Service the CEB offers is called an open-cloud PaaS where it's not tied to a product or operating environment. It allows Minority Serving Institutions (MSI) to use a completely different platform therefore providing flexibility [19].

A. Cloud Ecosystem Design Infrastructure

1) Ecosystem life cycle: The best way to tackle our approach to gather access to cloud services is from a system engineering prospective, less from a developer perspective because the CEB Hub or onramp being built is a system with multiple interacting components performing functions that cannot be achieved by one component alone [20]. Therefore, we move forward with steps toward understanding and improving this multi-component faucet by designing and managing this complex system over its life cycles [21].

2) Bird's eye view: This lifecycle map is a user experience birds-eye view of an ecosystem approach to services rendered through the CEB. This diagram is taken from a source that expounds upon user experience/needs and world-class architecture in system design. This is the first step we must take to fully recognize what users of the CEB need and how to fully engage with them, also recognize future viability of emerging answers to questions and understanding of a cloud Ecosystem. The red circles represent the typical stages a user moves through when engaging with resources tailored to them specifically. The first five stages (Aware, Interest, Research, Consider, Purchase) represent from "prospect" to "user", then continuing until the user becomes aware of new services, capabilities, upgrades, and communities of the system their using [21].

"By looking at clusters, we can begin to see which stages are more important than others on a personal basis. We can look at the time required to move between stages as another element of the user experience, and again we may discover different users (Returning/Expert user or New-User) require different amount of time to transition from one stage to another [21]." Here we discover how the user interacts with specific elements of transition to uncover opportunities such as journey mapping to improving user experience leading to loyalty and returning research.

B. Pedagogy

1) What is behind the CEB: The pedagogy behind the CEB is a cloud service enabled by a complex system of unrelated independent components that work together to form a Symbian relationship of hardware and software integrated with consultants, cloud engineers, cloud customers and partners. In the center of this cloud service is a public provider (*Firebase*) that creates easy access for business or higher education to collaborate and share information with colleagues and clients. It offers storage, sync, upload and backup time capsule, and key configurations for Macs, Windows, etc. and different mobile devices, which makes them immediately available, so you can have easier ways to collaborate. Firebase cloud services allow to share important files without the hassle of attachments, but instead gives a link option to download in order to preview, you can set passwords and expiration dates for added security. Having access to be able to preview files before downloading is a perk with the Firebase Cloud system for the exception, that, in order to edit them, you will have to download the attachments then upload back to *Firebase*. Also, the cloud services have the ability to save the latest version of any paper, so you can have the latest document available every time while the older versions are also saved for reference. It gives flexibility to share folders and files with colleagues or the ability to set permissions for security; basically, having private and public sharing at the same time. As with other familiar services, *Firebase* can create new documents, but has the ability to assign tasks which stops the need for data task management teams. The CEB will provide researchers with tools like RStudio, Jupyter notebooks, Linux OS, Windows, and other applications that enable Research Infrastructure within the CEB hub.

2) Research infrastructure: The service model platform used in this particular CEB cloud framework is utilizing the Infrastructure-as-a-Service (IaaS) model. [4] In this type of architecture, the cloud provider (Firebase) manages IT infrastructures such as storage, high performance computing power, data warehousing, security, server and networking resources, database access, big data computing environments, monitoring/manageability, and more by way of our providers (via AWS, Google, Microsoft, Oracle). There is a way to address this through the way of a hypervisor (CEB), which is a (user friendly) virtual machine monitor that host multiple VM's on a physical machine. This type of monitor allows simultaneous incorporation of different types of operation systems at the same time and provide instances of the diverse machines to users on demand [4]. With the Hub (CEB) portal activated, we see the CEB as an Infrastructure Provision cloud model. This means that the Minority Serving Institutions (MSI) utilize the cloud using the technology as a provisioning tool to supply on-demand virtualized resources. [4] The cloud model used in this project examines different types of characteristics customized for the various needs of our Minority Serving Institutions (MSI).

3) Data mining: Data mining refers to predictive analytics, a field of study recognized in Machine Learning that focuses on exploring data through unsupervised learning to discover previously unknown properties (Entity Resolution) in the data. When dealing with high volume big data projects, the successful implementation of the CEB life cycle become more difficult when dealing with issues related to *Data Governance*. [22]That is why resolving real issues related to the cloud services offered through CEB should identify the highest impact failures that can occur, then using work processes, risk management, and optimization methods to such services. [23].

a) Primary/Secondary data: Using the CEB, resources will be available for designing and populating secondary data resources that researchers can use to integrate already existing data with primary data. With this type of adaptation researchers can use this process to integrate research-oriented cloud ecosystems that operationalize funded research projects concepts within Minority Serving Institutions (MSI). Using the CEB will solve many challenges, mainly the challenge of meta-synthesis where scholars understand how different forms of evidence, variables, and analysis fit together to formulate a plan so the researchers can identify a specific research question, search for, collect, and combine evidence so the research question can be enhanced and to identify which variables from separate data sets may be made comparable to one another.

b) Descriptive metadata: Our descriptive metadata will support numerous STEM researchers' efforts to identify and obtain data for comparative analysis across states/countries fit for compiling details on appropriate data sources, institutional locations and contact information for acquiring this data. The web based descriptive metadata will have the advantage of advanced cloud hypervisor capabilities with the CEB central hub (BOX) allowing researchers to upload further details based on their own use of data resources.

C. Research Focus

1) Collaboration and outreach: Cloud resources enable Collaboration at its highest level, and the CEB will address the cons of cloud computing through Collaboration, by expanding on how access to real-time engagement tools offer strategies of deployment. The benefits of Cloud Collaboration exist within each community and are exercised within each group working on a particular research project also within the maintenance and management of the CEB. As you scale back from a bird's-eye view to see the logistics and the institutional boundaries, questions arise on how Outreach is attainable also how retention is manifested. Most existing national and international data are not designed to focus on the relationships between social groups relevant to our work, but through the CEB's community/user-based discussions will rethink the use of these measures. Therefore, Outreach should be manifested as a key component to the growth of desired infrastructure to move the cloud services via existing cloud resources to new communities.

2) Measuring analytics: Firebase Cloud Messaging offers different ways and ranges of messaging options and capabilities allowing you to send notification messages that are displayed to the user or data messages to determine what happens in your application code used by the client application. Notification messaging is used for user re-engagement and retention. Firebase Cloud Messaging targeting is very versatile, you can distribute messages in three separate ways: to single devices, groups of devices, to user devices subscribed to their different communities by either Message delivery reports, or Notification funnel analysis. You can export your analysis to Googles BigQuery Database infrastructure for further analysis. Firebase Cloud Messaging also makes messaging available through client applications by sending back to your server over its batter-efficient channel set up from settings for relevant engaging messages, targeted messages by community, and flexible custom alert. This helps you engage your applications active users by sending them targeted messages encouraging them to user more of the application features for exploration and discovery.

Mailchimp is a cloud-based marketing automation platform designed for email campaign for a targeted audience. It offers Automated email creation and sending, it is easy to use, it has built in statistics and has a style guide and knowledge base. It is a tool that manages your mailing list for you, and also allows you to create your own custom email templates.

Social Media will be used at a level of networking and outreach to target our particular MSI groups through awareness.

The Backend Functionality configures a serverless framework in response to triggers or HTTPS requests that automatically run in the background. After you write and execute the function it immediately begins to manage as it is triggered.

D. Front-End Web Development

1) Becoming a member: In order for a user to become a member into the CEB portal, you have to get Authentication credentials from the backend set within the Authentication methods noted in the backend text. The credentials are then passed to the Firebase Authentication services, once verified there is a callback response to the client. Once Authenticated by the backend services of Firebase, you can manipulate access to storage, computing power, processing, etc. by modifying the rules set within Realtime Database Services and Storage Security Rules within the Firebase infrastructure.

2) Log In: Users login through the CEB Log in portal either by username and password or by a federated identity, once users are Authenticated through Firebase, the users are placed in their respective communities through the Oauth state, then the community helps them move their research into cloud ecosystem. Firebase backend services connected to the CEB authenticate users using username and password or federated identity, linking users accounts to social media providers into a single account on Firebase.

3) OAUTH: In order for firebase to know which community the logged in user has chosen, there is a script located within the Oauth state that pulls from the database. The Oauth function is also in charge of maintaining state for each user. It fires back at the backend services of firebase so that the user maintains state.

4) Database: At this point the Realtime Database is being used due to the small amount of prototyped user engagement, also its low-latency solution for synced states. This type of database is persisted locally, and able to allocate real-time events offline giving the end user a responsive experience. When online services are regained and the connection is secure, the changes made will merge any conflicts automatically that occurred while the client was offline.

VI. Scope and Timeline (12-Months)

Deliverable 1: Feb 2022

T1: Design methodology to deliver identifying quantitative and qualitative data-driven research and resources to the CEB community.

T2: Measuring data mining using message streams elements, the best tools to engage with them are channels tailored for each group.

o Sentiment Analysis

Deliverable 2: April 2022

T1: Identify content management strategies that best suits this unique community based on user feedback from each group.

T2: Local CEB members and lead correspondence officers will take the lead to develop policy procedures to work with campuses on sustaining CEB's computational research activities and resources for the long term.

o Science Gateways

Deliverable 3: June 2022

T1: Developing a funding strategy for new research hosted in the cloud ecosystem is a top priority for the longevity of the program.

T2: Design branding strategy, user guidelines, and content policies for the ecosystem.

o Code of Conduct

Deliverable 4: Aug 2022

T1: Finalize the CEB Community Groups based on user feedback from survey. Then tailor the users back into their communities.

T2: Test Beta version resources and functionalities

Deliverable 5: Oct 2022

T1: Expanding the first wave of users in the MSI community

T2: Identify outreach opportunities to educate

Deliverable 6: Dec 2022

T1: Promote CEB to the second wave of users in the MSI community.

T2: Consolidate a mentoring network between institutions, corporate partners, and certification organizations to enhance student-mentor relationships throughout the CEB ecosystem

o Certifications from Providers

VII. Implications

A. Where are we now

1) Conceptual: The conceptual contribution conveys focus towards ideas and communication about CEB. Participants from different disciplines and backgrounds within the sphere of different communities: (1) Cloud computing, (2) Nanoscience, (3) Bioinformatic, and (4) Cybersecurity, focused on Minority Serving Institutions (MSI) discussing proposing these research communities to build new research and educational practices to advance leveraging cloud provider ecosystems. These discussions emphasized computational collaboration concepts and introduced new research techniques to using these cloud provider ecosystems. Also, these discussions served to identify critical points of commonality or identify entity resolution within the data in already existing primary data and archived these details in the descriptive metadata to create a sustainable commodity for researchers all around the world, in particular resource challenged institutions Minority Serving Institutions (MSI).

a) Existing Initiative (Existing): Existing Initiative involved public talks, blogs on current topics and publishing in magazine outlets. Meetings convened inviting researchers, policy practitioners and non-profit members, cloud providers. The meetings focused on adapting and refining core ideas within CEB. Participants from different disciplines, and institutions will identify relevant research themes within projects with emphasis on articulating meta-level concepts that are robust and meaningful across national contexts. The first part of this project for year 1 we relied on participants from different disciplines to focus on establishing core ideas within CEB through a virtual meeting broadcasted in real-time via EVO. This meeting emphasized computational concepts on a national level by adapting and refining core ideas within CEB. The main goal of the first part was to showcase the work done and to introduce the community and new research techniques using cloud ecosystems. Within this portion of the project, we developed the conceptual framework for research migration into cloud ecosystems, to include a searchable web-based metadata for quantitative and qualitative data for participating institutions, including primary data from ongoing research specifically designed for researchers to operationalize data-oriented research concepts at these institutions.

b) Education (Current year 2020-2021): Education included classroom innovations, service-learning pedagogies, and graduate webinars to bring together researchers between annual meetings. We invited researchers and postdoctoral students to share their experiences and present individual research activities on both public/private cloud environments. The goal was to both showcase the work done, but also to introduce the community to each other and new research techniques using cloud ecosystems. Provided self-contained learning modules for national certifications, nanodegree programs around cloud computing/cybersecurity resources such as, Cloud Platform-as-a-Service, Cloud Security and Data Protection, and the practical tools used to store, and process access these systems.

2) Methodological: The *methodological* contribution conveys focus on survey result conclusions for user-based needs pertaining to mined research tools that will be mapped as a research Data-as-a-Service (DaaS) Data that already exists with emphasis on data-enabled areas of research by reaching out to corporate partners. This was the pivot point to which educational activities are broadened and enlightened through Data-as-a-Service (DaaS) rather primary or secondary, that have a case-based level on a national scale.

a) Networking (Current year 2020-2021): Networking Video archives of talks, working papers, call for participation in conferences, links to social media, newsletter updates, the descriptive metadata, elaboration of our in-common research design and shared educational resources. All presentation materials are broadcasted in real-time via EVO. Participants are tasked with identifying quantitative and qualitative data-driven research that exist for participating campuses to be leveraged in the new ecosystems. The PIs in collaboration with the core committee members designed and deployed a Hubzero portal (CEB) on Open Educational Resources Commons (OER) to highlight innovative intellectual, methodological elements associated with project activities. This Hubzero portal (CEB) serves as enduring library of unique content, while facilitating collaboration for the MSI audience around deploying research on the cloud. Also, establishing a mentoring network between institutions, corporate partners and certification bodies to foster student-mentor relationships within and across institutional boundaries.

b) Publishing activities (Current year 2020-2021): The second part of this project for year 1 constituted identifying quantitative and qualitative data that exist for participating campuses discussed in part one. The key innovation here is methodological in nature for the shear qualitative case-based data sets and national-level secondary data sets relevant to the CEB. The combined conceptual and methodological work makes the logic keen to realizing the intellectual merit that goes into the CEB, not to mention the intellectual innovations. During this second part we developed an extensible research design for robust mixed-methods research associated with leveraging these ecosystems for (1) bioinformatics, (2) cloud computing, (3) nanoscience, and (4) cybersecurity. Developing the Hubzero portal (CEB) served as a resource for experiential-based learning in cloud ecosystems around cyber infrastructure.

3) Bring it Home: Using the CEB, these Minority Serving Institutions (MIS) will begin to engage computationally and collaboratively within the cloud community it serves. With addressing the methods within this study and an expected scope of future goals, Visualization and Data Analysis System will increasing become more aware within these institutions and faculty will begin to gain value and be better prepared using CEB cloud resources. High Performance and Distributed high Throughput Computing will no longer be accepted into research only for R1 institutions. The CEB will restructure the way research is conducted and will encompass the minority research community, NSF funded ecosystems, and industry partners creating a triple convergence.

Through marketing the CEB, this is the first step to being proactive in outreach for the continuation of services rendered through the CEB. The first plan of action that was already preceded is to target Minority Serving Institutions (MSI) by developing a list of users based on NSF awards on their respective research areas of communities. Then the idea was to narrow this list by geographical location then send out a survey to measure the interest and needs of each community based on gender and demographics. Results from the survey helped us understand the dynamic within each dialect of data-driven research within each respective community of Minority Serving Institutions (MSI). Continuation of marketing throughout the life cycles of the CEB merits different styles of advertising and finding ways to target an indecisive audience. [24].

The CEB will provide at a national level the infrastructure for training program platforms by reinforcing the conceptual framework for research migration into the cloud ecosystems. It will be a direct centralized portal to provide users with: (1) applications to enable them to operate at real-time intervals within parallel computing power, (2) educational material such as links, mailing list, repositories, and YouTube channels pertaining to each Minority Serving Institutions (MSI) research to extend MSI networking through social media engagement, (3) Tools that enable data-driven research: programming interfaces, operating systems over hypervisor VM's, and a liaison/beacon for training.

Cloud services within each of these communities provides stable software platforms Infrastructure-As-a-Service (IaaS), Platform-As-a-Service (PaaS), Software-As-a-Service (SaaS), and Storage-As-a-Service (STaaS) with an associated larger community of providers ((via... AWS, Oracle Academy, Microsoft Azure, Google)) who can provide support and solutions specific to a researchers' domain, using these research specific public cloud providers that make up their own ecosystems [25].

B. *Next Steps*

Building the Communities will be incorporated in the CEB website as tools connected to the cloud providers provided cloud messaging, message bird, trigger emails functions, chat, storage, and machine learning. Upon success of the creation of a conceptual framework to incorporate searchable web-based metadata for computational research, and the developing of a Hubzero portal CEB that will be the center of ongoing research in learning cloud systems communities: (1) Cloud computing, (2) Nanoscience, (3) Bioinformatic, and (4) Cybersecurity. Broadening the scope and intellectual impact of the CEB will be established to build other branches of research. This research includes primary data from current research participating institutions, providing self-contained learning modules like Cloud Platform-as-a-Service, Cloud Security and Data Protection. These include public initiative, education, publishing activities and networking.

Measuring and understanding engagement on any platform is a complex undertaking that requires understanding multiple interconnecting ecosystems. To ensure the project was manageable in the timeframe of our funding, we initially created a survey instrument to select researchers at MSI, that are interested in migrating their research into cloud environments. This will allow us to design discipline level guides to facilitate this movement. In particular to measure engagement, we will capture the following elements, (1) user driven streams from each group i.e. nanoscience, cybersecurity, cloud computing and bioinformatics. (2) The project will be able to store publicly broadcasted messages, which are directed from users, particularly group followers; (3) The project will be able to determine the geographical location of the messages recorded; (4) The project will be able to retrieve user histories to aid in the fabrication of conversation instances; (5) The project will be able to extract all pertinent information from a tweet to aid computational processes; (6) The project will be able 'created_at' stamp which is a more mathematically friendly timestamp of the format MM/DD/YYYY ##:##:##, this will provide context when the message was created; (7) The project will be able to determine the participation of each user involved in a given conversation with a user group; (8) The project will be able to directly query a NOSQL database to keep all data structured and organized by individual user and research groups.

References

[1] Ramirez, G. Fox, A. Kuslikis, K. Barnes, D. Baxter and J. Foertsch, "Engaging Minority-Serving Institutions (MSIs) in Cyberinfrastructure (CI) through CI Days," p. 9.

[2] M. Dahan, J. Towns, T. Cockerill and N. Wilkins, "RG," Sept 2014. [Online]. Available: https://www.researchgate.net/publication/267867365_XSEDE_Accelerating_scientific_discovery/stats.

[3] S. Zweibel, "DH Box: A Virtual Computer Lab in the Cloud," CUNY Academic Works, New York, 2016.

[4] J. Walker, D. Freet, R. Agrawal and Y. Badr, "Open Source Cloud Management Platforms and Hypervisor Technologies: A review and Comparison," IEEE, Norfork, 2016.

[5] R. Tillinghast, D. Appel, C. Winsor and M. Mansouri, "STEM Outreach: A Literature Review and Definition," in 2020 IEEE Integrated STEM Education Conference (ISEC), New Jersey, 2020.

[6] K. K. P. Amy Aldous Bergerson, "CARES: Mentoring through University outreach," Journal of Higher Educaiton Outreach and Engagement, vol. 13, no. Number 1, p. 45, 2009.

[7] J. v. Byrne, "Outreach, Engagement, and the Changing Culture of the University," Journal of Public Service and Outreach, vol. 3, no. Number 2, pp. 3-8, 1998.

[8] M. Dennis, "Outreach initiatives in academic libraries, 2009 - 2011," Emerald, vol. 40, no. 3, pp. 363-383, 2010.

[9] K. A. G. Lisa S. Anderson, "Pre-university Outreach: Encouraging Students to Consider Engineering Careers," Global J. of Engng. Educaiton, Australia, 2003.

[10] D. Knighton, "Box Builds Ecosystem to Modernize Collaboration in Education," Canvas, LOS ALTOS, California, 2013.

[11] S. Fulton, "readwrite," 28 March 2012. [Online]. Available: https://readwrite.com/2012/03/28/box-launches-its-own-enterpris/.

[12] J. Sparapani, "TechTarget," 7 Feb 2018. [Online]. Available: https://searchcio.techtarget.com/blog/TotalCIO/Cloud-computing-ecosystem-Greater-than-the-sum-of-its-parts.

[13] Wikipedia, "Cloud Manufactoring," [Online]. Available: https://en.wikipedia.org/wiki/Cloud_manufacturing.

[14] P. Thomas, "A Sandbox in the Clouds: Software Testing and Development in Cloud Storage," 14 January 2020. [Online]. Available: https://en.wikipedia.org/wiki/Cloud_manufacturing. [Accessed 2 April 2021].

[15] P. University, "Pioneering naotechnology cloud--nanoHUB looks to future," Pioneering naotechnology cloud.

[16] Cisco, "What Is Cybersecurity?," Cisco, [Online]. Available: https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html#~related-topics. [Accessed 22 04 2021].

[17] A. Valkonen, "Cloud Computing Ecosystem: Insights from an Exploratory Study in SaaS and Paas Value Networks," Aalto University, 2013.

[18] H. G. Labrador, G. Alexandropoulous, E. Bocchi, D. Castro, B. Chan, C. Contescu and M. Lamanna, "CERNBox: the CERN cloud storage hub," EDP Sciences, Meyrin, 2018.

[19] A. Cooke, "The 3 Main Types of PaaS," TrustRadius, [Online]. Available: https://www.trustradius.com/buyer-blog/the-3-main-types-of-paas. [Accessed 20 4 2021].

[20] C. o. Engineering, "College of Engineering," [Online]. Available: http://www.bu.edu/eng/departments/se/prospective-students/what-is-systems-engineering/.

[21] Leofrish, "Architecture and User Experience, Part 7: An Ecosystem Approach," Phase 2, 03 02 2017. [Online]. Available: https://www.phaseiidesign.com/posts/architecture-user-experience-part-7-ecosystem-approach/. [Accessed 03 05 2021].

[22] M. Rouse, "TechTarget," [Online]. Available: https://searchitchannel.techtarget.com/definition/cloud-ecosystem.

[23] "Wikipedia," 20 Feb 2019. [Online]. Available: https://en.wikipedia.org/wiki/Systems_engineering.

[24] MailChimp, "MailChimp," [Online]. Available: https://mailchimp.com/.

[25] A. M. O. A. P. H. Hugh P. Shanahan, "Bioinformatics on the Cloud Computing Platform Azure," Plos One, 22 7 2014. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0102642. [Accessed 20 4 2021].