

Prospectus

Social Networks and Archival Context (SNAC) OpenRefine Extension
(Technical Topic)

Cloud Computing's impact on Game Developers' relationships with users
(STS Topic)

By

Peter Tran

12/10/2019

Technical Project Team Members: Charles Chang, Sandra Gould, Mark Jeong, John Perez,
Victor Shen, Grace Wu, Jessica Xu

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____

Approved: _____ Date _____
Rider Foley, Department of Engineering and Society

Approved: _____ Date _____
Ahmed Ibrahim, Department of Computer Science

Social Networks and Archival Context OpenRefine Plugin

Social Networks and Archival Context (SNAC) is a free, online resource that allows users to discover information about the people and organizations that are documented in primary source documents and the connections between them (Social Networks and Archival Context [SNAC], n.d.). SNAC is used to locate archived collections as well as related resources held around the world. As an international cooperative, SNAC works to “build a corpus of reliable descriptions” of people and artifacts that link to and “provide a contextual understanding” of historical records (SNAC, n.d., para. 1). In order to create these contextual connections, SNAC sources its information from many different libraries and archival institutions. SNAC cooperates with over 4000 institutions to gather and reconcile data (SNAC, n.d.). Each of these institutions has a different structure for storing records. Relationships between different entities, labels for certain types of data, and the hierarchy of the data itself are inconsistent from each outside institution. SNAC needs to reconcile the differences between the outside data and its own data storage structure before importing the data into its database. It is extremely impractical to clean up the data manually or with simple tools (Ham, 2013). The reconciliation of this data is vital to the functionality of an archival organization such as SNAC because it is crucial for efficient and accurate querying (Park, 2008).

The technical project seeks to develop a standalone plugin for Social Networks and Archival Context (SNAC) using OpenRefine. OpenRefine is an open source software that is community-maintained designed specifically for data normalization, transformation, and cleaning (Hill, 2016). It allows users to import and normalize data with a series of pre-existing

default user interfaces after connecting to a target resource. OpenRefine provides a “powerful yet user-friendly interface” for experimenting with and querying data (Hill, 2016, p. 228).

With over 700 edits occurring to its data schema in week, Social Networks and Archival Context (SNAC) is no small data archive (SNAC, n.d.). The current workflow for refining and updating data in SNAC is quite difficult and inaccessible to inexperienced users. It involves users hitting SNAC’s APIs for refining data on their server from the user’s local machine. The technical project aims to greatly simplify this process by creating a streamlined plugin that will have all the functionalities needed to refine and upload data in one location. The logical flow and components needed

for the project are

illustrated in Figure 1.

The plugin will serve

as a connection

between the user’s

local data and SNAC’s

server. It will allow

users to import

external data in the

form of comma-separated

values (CSV) files and make use of APIs provided by SNAC to reconcile and refine that data

with SNAC’s unique JavaScript Object Notation (JSON) data structure. The plugin will have two

main user groups: privileged and unprivileged users. Both types of users will be able to use the

plugin to format any data ported in using SNAC’s organizational schema. Only privileged users

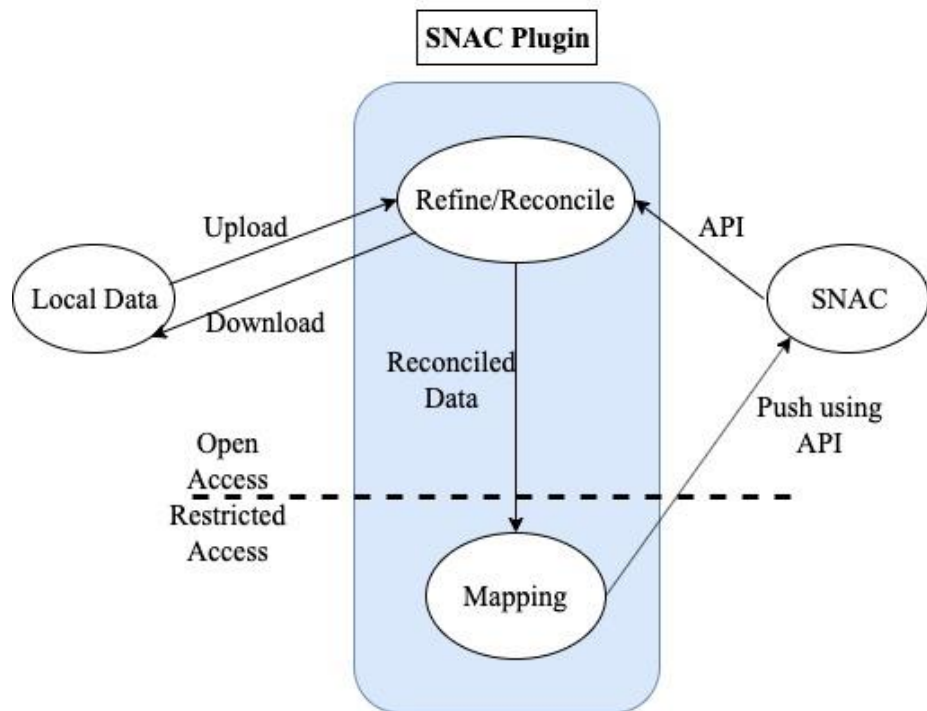


Figure 1: SNAC Plugin Model: An overview of the design of the plugin, depicting the different processes and functions that will be made available by the plugin (Xu, 2019).

will be able to then push the formatted data into SNAC's own database utilizing the APIs provided by SNAC. The technical project will provide an easy way to reconcile outside data with SNAC's existing data in addition with an improved user interfaced for an enhanced user experience.

The development will conduct biweekly customer meetings with the client in order to gather system requirements and get feedback about ongoing work. The minimum requirements for the plugin to be completed by the end of this semester include:

- Allowing users to import CSV data into the plugin
- Connect the data fields with different SNAC IDs
- Search for constellations in SNAC and match them to the imported data
- Allow a human editor to choose from several options to match for when the plugin is unsure
- Reconcile the imported changes based on the connection and matches
- Download the data that is now reconciled with SNAC's structure
- Users with privileges will be able to publish the data to SNAC

Desired requirements include:

- Users will be able to reconcile more complex data items like relationships and geolocations
- Users will be able to edit already existing resources and constellations

So far, no optional requirements have been specified by the client.

The technical project will be developed over the course of the two-semester capstone series led by Professor Ahmed Ibrahim from the Computer Science department, and will result in a technical report. To create this plugin, OpenRefine will be used, as it is a powerful tool for working with disorganized data that can “[transform] it from one format into another; extending

it with web services and external data” (OpenRefine, n.d., para. 1). A similar project exists already for WikiData, but the technical project will create a new implementation specifically for Social Networks and Archival Context (SNAC). The plugin will hopefully provide a faster and more intuitive way for SNAC users to reconcile and update data.

The Future for Cloud Computing and Gaming

While the capstone project is about the SNAC OpenRefine extension, the rest of this prospectus will expand on cloud computing and its impact on game development teams’ relationships with their player-base. The framework of importance here is technological momentum, which basically states that the relationship between technology and society is reciprocal and time-dependent where they both influence each other. The history of evolving systems can be pictured in phases: invention, development, innovation, transfer, growth, consolidation, and competition. However, these phases aren’t simply sequential but overlap and backtrack. An example of backtracking occurs with reverse salients, which are components within a system that have fallen behind or are out of phase with the others (Hughes, 1987). While not necessarily a problem, the constant struggle by video game development companies to have a good relationship with their player-base is a vital component of a successful game. Video games are supported and kept running by both the development team and the player-base in uniquely different roles. Developers in particular rely on users to fund the continued development of their game through microtransactions, initial game purchase, and advertisements. It’s in the best interest of developers to please the players who influence their decisions (Pagano & Brügge, 2013). According to a literature review by Abelein, Sharp, and Paech (2013), there is a 92% positive correlation among 133 studies based on user involvement’s role in influencing the success of a system. More and more tools are coming into fruition which could aid or even

advance the options available to a developer or user to design and develop a game. One of which is cloud computing, which is defined by the National Institute of Standards and Technology to be “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources...” (Mell & Grance, 2011, pg. 2). Reverse salients can occur from outdated computing infrastructures that bottleneck the rest of the system down. However, modern maintenance and technology often prevents this from becoming a huge issue. For the most part, cloud computing is in a cycle between the development, innovation, and growth phase identified in the technological momentum framework. While still being improved and perfected with new ideas and implementations, it is quickly growing and being popularized (Riungu-Kalliosaari, Kasurinen, and Smolander, 2013). It has many capabilities and variations that can be explored in the final paper, but it’s up to development teams to find the ideal fit with their development methodologies and goals to build their products’ momentum. A successful implementation of cloud computing should better the user experience/user satisfaction of a game. Riungu-Kalliosaari, Kasurinen, and Smolander (2013) conducted interviews with several gaming organizations and gathered their views and applications of cloud gaming. The study found five key points: cloud services and cloud gaming are well known in the gaming industry, but are considered too unreliable for main game products; gaming organizations need clear business models and success stories to convince them to adopt cloud gaming; cloud services and cloud gaming business models don’t focus on any platform or organization size; organizations consider cloud services and cloud gaming models to steer products towards services and user groups toward communities (Riungu-Kalliosaari, Kasurinen, and Smolander, 2013). For these reasons, companies are hesitant to fully invest in cloud computing but are aware of the potential and current uses. Despite cloud technology being new, the momentum gained from the back-and-

forth impact between the users and the technology gave way to its current reputation. Current uses are mostly on development side such as code and document hosting rather than game operation. Some mobile games also leave the game processing to the cloud, making the load for a user's mobile console smaller and more manageable (Cai, Leung, and Chen, 2013).

Understanding cloud computing's current uses puts pros and cons into perspective and allows for discussion on whether or not a development team should consider it to engage their players.

Video games, or software in general, act as a medium for players to communicate with other players as well as the development team. There's the developer-user relationship through feedback and developer posts. Then there's the connection between developers and users through the actual game produced. Kaye & Bryce (2012) state that multiplayer games fosters social gaming among players. By playing with other real players, especially people they know, they are experiencing a sense of belonging or companionship with the others in the game. Multiplayer gaming can be highly influential to a player and having this good social gaming experience requires good setup from the development team, potentially through the use of cloud computing. A development team needs to understand how their product can influence their players which in turn affects them because customers pay their bills and provide feedback. By understanding their impact, development teams have more to process and tweak in their design and development processes to benefit both themselves and their customers for their games. Companies need to work towards improving their actual gameplay experience by eliciting user feedback as well as optimizing their game performance wherever possible. Cloud computing can be encouraged by a user-base enjoying the experience it provides for a game. This can incentivize the development teams to invest in a cloud-based future especially if increasing profits are possible. With more investment and reputation, cloud computing can take off and become even more influential than

it already is. It can feature new ways to engage users both inside and outside the game. Enjoyability of a game affects the success of it and how it influences its players. By listening to feedback from users in pre and post development, teams can work towards user satisfaction which nurtures a better relationship between developers and users. It goes beyond a relationship of user and game, but instead is a relationship between users, games, and the developers of said games. Cloud computing can play an immense factor here in how a development team can engage players who determine the fate of how cloud computing is received in the game development world.

Research Question and Methods

The primary research question is: How has the advent of Cloud Computing impacted Game Developers' relationships with users? With the rapid growth of the cloud over the past decade and its current usages, what role could or does cloud play in a market as large as video games? As a result of this potentially adoptable and adaptable technology, how will the relationships developers have formed with users change? How can new cloud supported relationships be formed? Cloud computing is a quickly growing technology and knowing every useful field it could apply to is worth looking into. Data could be composed of but not limited to: company earnings, usage metrics, user feedback, infrastructure costs, case studies, and interviews. Besides providing definitions and analysis on the technology, developers, and players, the research paper will also have data about existing games utilizing cloud computing and how successful they are based on user satisfaction and profits. The same metric can be used for games that are not using cloud computing in order to see how the technology can affect their players and profits. I plan to look at reports, strategic analyses, and case/research studies to gather this information. Interviews are also planned to take place for game development teams

to gather perspective on their views on cloud computing's role in their jobs, how they handle user engagement, and whether or not cloud computing can affect their approach to maintaining their relationships with their users. A combination of thematic and content analysis on the types of cloud computing and which are the most viable to achieve a game developer's goals in the present or future will be made. This is in conjunction with how cloud computing would change the games and the relationship between the developers and players as a result.

Conclusion

My technical topic is the SNAC OpenRefine Extension, which is an extension to Google's OpenRefine data clean up and reconciliation tool that enables the easier upload of data for creating or updating an entity in SNAC's collection of historical entities. The end goal of the project is to have a functional addition to OpenRefine that allows SNAC contributors to have an easier time managing information entry. The result of this will lead to more efficient uploading of information for the public to access. There will also be an improvement to the code base for members of the open source community to learn from. Figure 4 below details the planned work schedule for the team throughout the capstone.

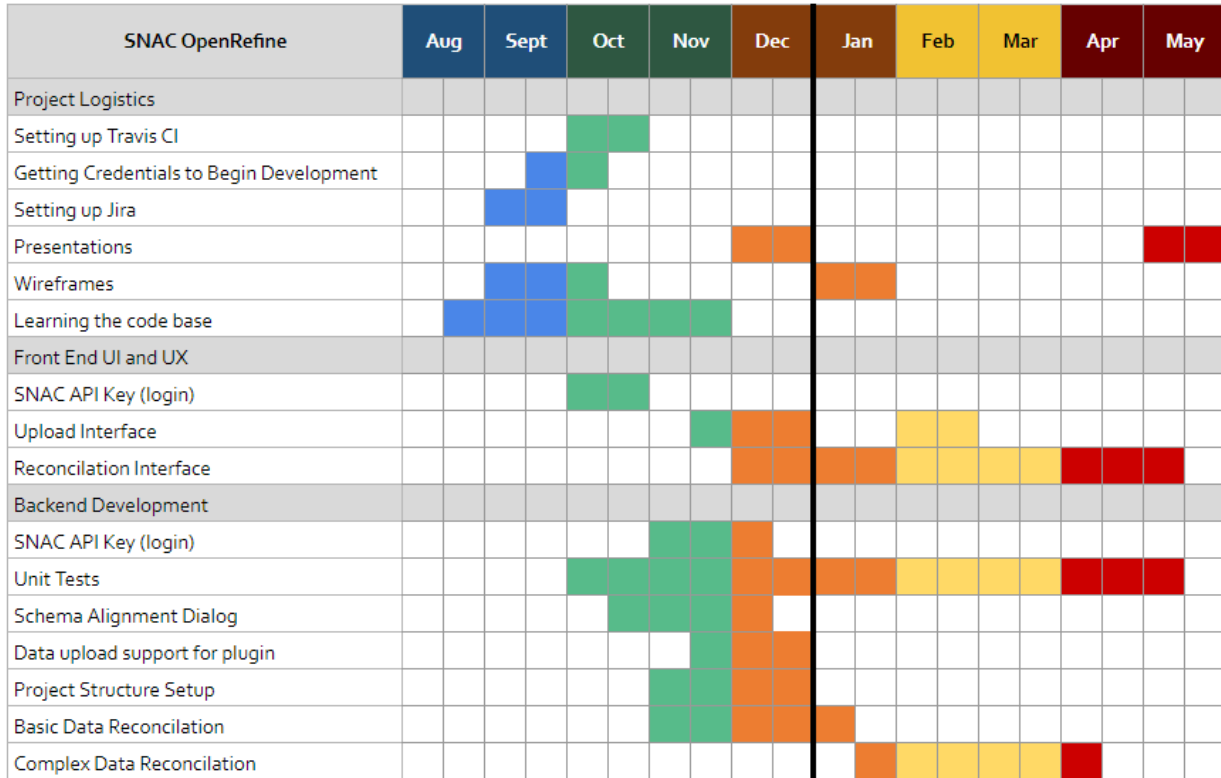


Figure 4. Gantt chart of timeline for technical project. Bolder vertical line indicates the split between semesters.

The hope for the completed STS research paper is that it can provide insight to anyone interested in the research question or anything related to it. At the moment I have no plans on publishing this but the knowledge gained will be valuable in my everyday thinking when I work with technology. As for the STS topic throughout the year, further research and gathering of better sources will occur through the rest of this semester to December. Then I will conduct deeper analysis of said sources in context of the STS framework beginning this semester and heading into most of the second semester roughly around March and April. I'll try to setup interviews with gaming development teams of interest by the end of the start of the second semester. The actual interviewing ideally occurring midway throughout the second semester. Reviewing, cleanup, and refining will occur near the end of the second semester and year.

References List

- Abelein, U., Sharp, H., & Paech, B. (2013). Does involving users in software development really influence system success? *IEEE Software*, 30(6), 17-23.
- Cai, W., Leung, V. C., & Chen, M. (2013, March). Next generation mobile cloud gaming. In *2013 IEEE Seventh International Symposium on Service-Oriented System Engineering* (pp. 551-560). IEEE.
- Ham, K. (2013). Free, Open-source Tool for Cleaning and Transforming Data. *Journal of the Medical Library Association*. Retrieved from <https://www.ncbi.nlm.nih.gov/>
- Hill, K. M. (2016). In Search of Useful Collection Metadata: Using OpenRefine to Create Accurate, Complete, and Clean Title-Level Collection Information. *Serials Review*, 42(3), 222-228. Retrieved from <https://www.sciencedirect.com/journal/serials-review>
- Hughes, T. P. (1987). The evolution of large technological systems. *The social construction of technological systems: New directions in the sociology and history of technology* (pp. 52-80). Cambridge, MA: MIT Press.
- Kaye, L. K., & Bryce, J. (2012). Putting the fun factor into gaming: The influence of social contexts on the experiences of playing videogames. *International Journal of Internet Science*, 7(1), 24-38.
- Mell, P. M., & Grance, T. (2011). The NIST definition of cloud computing. NIST Special Publication 800-145. doi: 10.6028/nist.sp.800-145
- OpenRefine. (n.d.). *Introduction to OpenRefine*. Retrieved from <http://openrefine.org/>
- Pagano, D., & Brügge, B. (2013, May). User involvement in software evolution practice: a case study. In *Proceedings of the 2013 international conference on Software engineering* (pp. 953-962). IEEE Press.

Park, J-R. (2008). Metadata Quality in Repositories: a Survey of the Current State of the Art.
Cataloging & Classification Quarterly, 47(3), 213-228.

doi:10.1080/01639370902737240

Riungu-Kalliosaari, L., Kasurinen, J., & Smolander, K. (2013). Cloud services and cloud gaming in game development. in Proc. the IADIS Game and Entertainment Technologies, 22(24.7), 2013.

Social Networks and Archival Context. (n.d.). *About SNAC*. Retrieved from
<https://portal.snaccooperative.org/about>

Xu, Jessica (2019). Figure 1: SNAC Plugin Model.