Society Key: Integrating Social Media Data with Governmental Open Data to Encourage Community Wellbeing

Ziad Matni, Rutgers University
Jennifer Sonne, Rutgers University
Dongho Choi, Rutgers University

Abstract
In this paper, we present a service incorporating an electronic online system that integrates data from multiple social media sources and multiple governmental open data sources with the intention of improving awareness of available societal data and to encourage individuals and groups of people to take action towards the wellbeing of their local communities. We call our proposed service/system, "Society Key".

1   Introduction

In this paper, we present a service incorporating an electronic online system that integrates data from multiple social media sources and multiple governmental open data sources with the intention of improving awareness of available societal data and to encourage individuals and groups of people to take action towards the wellbeing of their local communities. We call our proposed service/system, "*Society Key*".

Government open data is a rich and underdeveloped resource that can work to the greater good of its citizenry. One of the difficulties in utilizing government open data at the individual level arises from a lack of awareness as well as an inconvenience for most people to compare information across different datasets. If people were able to visualize and compare their data with government open data, it would increase their awareness of different issues that are important to their daily lives, leading to an increase in their community involvement and conversations. Awareness of these available data sets will help people to interact with them in real time and in the public domain, such as libraries and cafés.

Social media is also a rich information medium and its interfaces with the user are usually designed with social software and user-centric viewpoints and hence are easy to use for query specification and navigation. Social media thus creates low-cost and easy pathways for information seekers to pursue. Social media often relies on user-generated content. We believe that integrating certain user-generated content into government open databases, and vice versa, enhances the relevance and timeliness of the open data and introduces a new value to the usual role of social media. It is our opinion that social media provides the community relationships that government open data needs. In other words, *the marriage of social media with government open data is more valuable than the two entities individually*. In order to better aid this, our system also encourages users to feed data back to the community at large and the open databases as well, thus providing governments with new sources of information about communities and individuals, also contributing to their databases' timeliness and potency.

Finally, it is important to administer this integration of social media and open government databases, not just at the individual, but also at the community level. Good avenues that encourage both awareness and interaction are public spaces and local businesses. We therefore additionally propose to integrate these spaces into our service to allow for further community collaboration.

2   Description

By allowing individuals to craft their own personal experience using government open data, we allow them to become more comfortable with the resources and information available. The familiarity engendered by our proposed service, Society Key, promotes positive community support and wellbeing.

The Society Key system will collect, evaluate, and help the exploration of (i.e. visualize) data from social media in order to compare it against government open data while also enhancing it and helping people to better engage with their communities. It would allow for the forwarding of our collected data,

once we ensure anonymity, to local and national government organizations, as well as non-governmental organizations (NGOs).
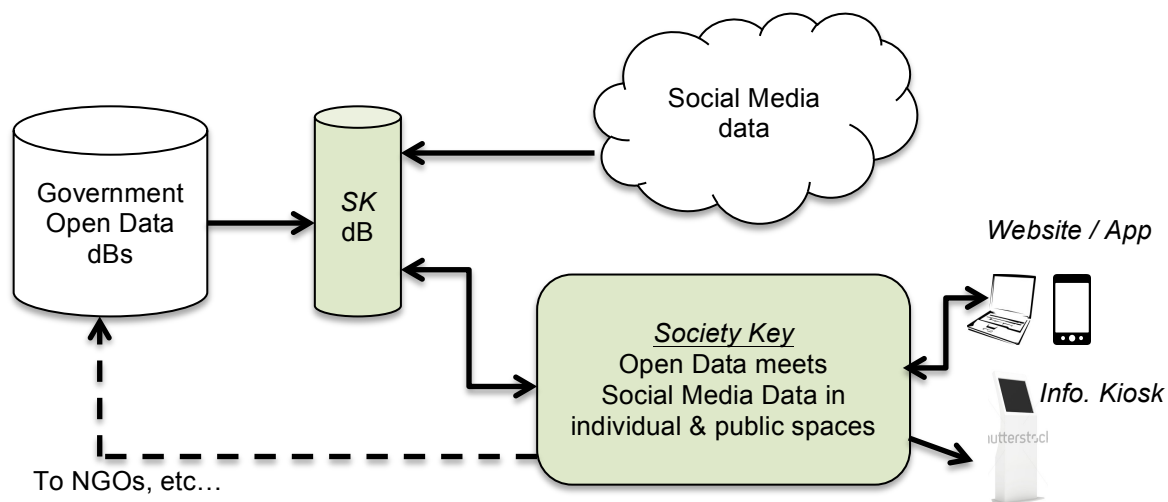

Figure 1. The Society Key System

Society Key can engage users with one another and with local businesses through various mechanisms, such as ones that allow comparing of individuals' health statistics, or ones that help users establish themselves as "top users" of the site, as recognized by "badges" that we issue them when they have attracted enough of their friends to the site, or with "challenges" set forth by businesses that incentivize users to contribute to the site by issuing coupons.

Society Key will be designed as an online Web service for access by individuals and groups. It will thus will have an interactive website optimized for both *large screen* (i.e. for use with regular laptops, desktops, and large tablet computers) and *mobile device* format (i.e. for smart phone and smaller tablet computers). Additionally, we propose to integrate *information kiosks* in and around public spaces as part of our system. This will present people with a physical venue to exchange information with each other (and Society Key) in a public space that encourages them to interact based on common local interests. These public information venues are also an important component that will help people become more aware of these open databases and help them visualize the importance of this information in their lives and their communities.

3   Society Key Users and Their Goals

Our envisioned users are made up of individuals, local organizations and businesses, as well as governmental agencies. All sets of users will enrich the collected Society Key database, contribute to the existing governmental open databases, and foster the well being of their joint communities. Users will interface with our service in differing ways. Individuals will be able to input or view data on various aspects of their personal and their community wellbeing, such as individual health statistics (e.g. BMI, number of minutes of exercise per day), or individual data based on their geographic locales (e.g. their salaries, property taxes, number of elementary schools in their neighborhood/town). Individual users will also have the option to automate some of this data compilation and collection via their mobile devices such as their smart phones.

Local businesses will be able utilize the service by both exploring local data relevant to their commercial interests as well as by interacting with their community members. Interaction with these people can be done both by encouraging them to be aware of their collective wellbeing (for example, by having an information kiosk in the store) and by advertising to them of their own services to motivate them to frequent their business locales (for example, by issuing "community challenges" and coupons). The business locales will thus emphasize a space where community members not only engage in commercial activities, but also have a space for meeting each other and discussing personal or community issues whether on purpose or serendipitously.

As for the users' goals:

3.1   For Individuals:

- Compare and visualize individual's statistics gathered via social networking sites (e.g. tweets with specific hashtags) against averages culled from the government open data (local, state, nationwide, specific demographics, etc.). *Goals: increased awareness and better decision-making*.
- Input individual's data for their own tracking purposes. *Goals: better individual wellbeing*.
- Enter "challenges" and compete against user-base and/or government data averages. *Goals: increased awareness and increased motivation*.

3.2   For Businesses & Non-Governmental Organizations:

- Compare and visualize business' statistics against averages (local, state, nationwide, specific demographics, etc.). *Goals: increased awareness and better decision-making.*
- Fund Challenges. *Goals: better connection to community.*

3.3   For Government Agencies:

- The ability to import data from our database to analyze and potentially integrate into government open datasets. *Goals: increased data and closer connection to citizens.*

4   Suggested System Framework

4.1   Interfacing with Outside Databases

We propose to collect government open data from available sources provided by the United States state and federal governmental bodies, for initial deployment, and having in mind other governments' open data for future use (for example, Canada's available open databases). We would utilize these databases' application programming interfaces (APIs) to collect the data, if technically supported, or gather it using other automatic formats if APIs are not made available (for example, using HTTP commands to get documents and then use automated data mining scripts).

Similarly, we propose using available APIs for social media data collection. Our main website would instruct potential users how they might ensure that their social media data gets integrated into our system. For example, we could suggest a set of unique hashtags to be used on Twitter that make sure that the data gets picked up and used in our system.

4.2   Aggregating Data in our Database

Since there does not currently exist standardized structures for all open data, much of this data on open databases and in social media needs some initialization or scrubbing before it used in our own database. We believe that our system design must include a unified general data structure scheme (for example, using XML or JSON data objects) to ensure an automatic transfer and use of open government data and social media data with our system. This will allow the comparing of heterogeneous data.

4.3   Front-End Software Framework

Our front-end software framework would consist of three components: Check-Ins, Charts & Maps, and Challenges. "Check-Ins" refer to the social networking features for participants. Individual users can register themselves in the system with different levels of access to their social media posts. Then the data the participants generate are automatically collected by our database and organized in general schemes.

"Charts & Maps" refer to the function that compares the collected data from social media against the government open data. Comparison of various perspectives on particular topics can be presented by multiple visualization methods.

"Challenges" are the gamification features of Society Key. A social media platform with the features of gamification would be implemented in order to increase participations, engagements, and activities both of individuals, organizations, and businesses. Businesses and other organizations can issue "challenges" to individuals or groups. If a challenge is accepted and completed, the participating individuals or groups receive discounts or coupons to the business or to the organization. These challenges are meant to promote wellbeing, and can range from health-related ones (e.g. weight-loss challenges) to local urban renewal or enhancement efforts (e.g. submitting the general location of potholes spotted on the streets).
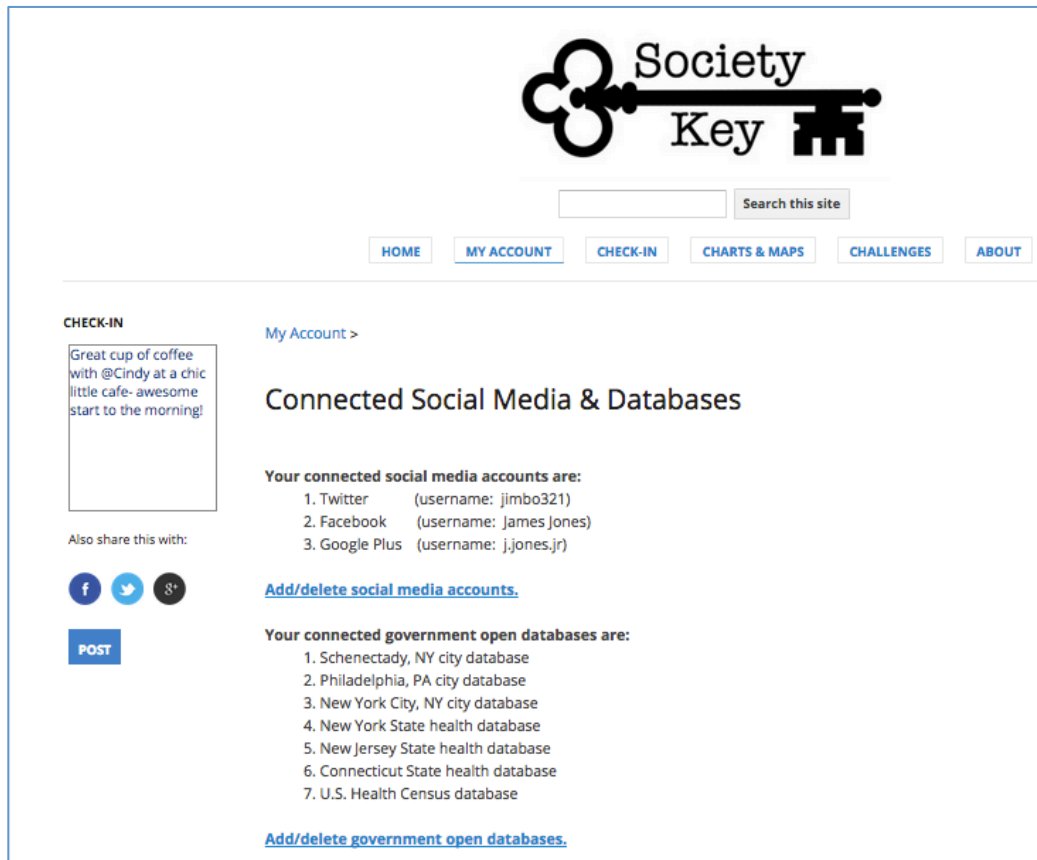
Figure 2: Society Key website snapshots showing examples of the "Check-In" and "View Charts and Maps" features where users can log in, link up with their social media accounts and a variety of government open databases, then create, view, and manage their data visualizations. The prototype website is available at: http://goo.gl/VBTwDm.

5   Future Work

This project is currently a work in progress. We have developed a prototype website at http://goo.gl/VBTwDm and are working on a design for the Society Key database and the information kiosks. We intend to use these exemplars to conduct user studies in order to gauge people's information behavior around tools of varying forms (i.e. laptops, mobile devices, or stationary information kiosks) that incorporate social media features, open databases, and community activity. We foresee that Society Key user studies will likely entail gathering data in both the lab (for example, via the website on laptops or tablets) and in the field (via prototypical information kiosks, for instance). These studies will include monitoring of user behavioral (with the use of eye-trackers and log files in the lab, as examples), user-experience surveys, and semi-structured interviews. Our analyses are thus expected to be both quantitative and qualitative in nature. Additionally, and in every step of our research process, we intend to review and prioritize the best way to protect user privacy.

Building and deploying such a system, at least on a small scale to begin with, would be our next step. Here, the results of our prior research with the initial prototypes should be useful in our re-design efforts of both the websites and the information kiosks. We will then pursue additional user studies, mirroring our prior efforts with the initial prototypes, and new results will be compared against expectations set by the first run of user studies. Of course, we expect that the work to be done in this build/deployment phase will require a greater amount of resources as well as expertise outside of academia. More specifically, we believe we will need input from other experts such as user-experience (UX) designers, programmers (Web services and database coders), and hardware engineers, to name but a few. It is thus incumbent on us, at some point in the future, to plan on how to best acquire and finance these resource requirements.

6   Conclusion

The easy availability of extensive government open data is an underdeveloped resource that can be made use of to better the lives of individuals and communities. This data, among the sea of other publicly available data (like from social media sources) is not always easy to explore or visualize in a meaningful way for most people unless they engage in some kind of data manipulation on a source-by-source basis. In other words, unless one is an information specialist or a computer scientist, making sense of this freely available data is not always practical. We believe that citizens are missing-out on many advantages that knowledge of this data can bring, such as an increased awareness of different important issues that could lead to increased participation in their communities and their general wellbeing.

Our proposed service/system, Society Key, connects people to their community and to their government and promotes positive community support and wellbeing. Society Key allows its users to collect and evaluate data from social media in order to compare against government open data. The service also allows for individuals to directly influence their governments' knowledge of their communities by feeding back data to the government open data sources. Additionally, Society Key allows engagement between individuals and local businesses through several kinds of incentives.