

Lessons learned from using data science to empower change agents across data silos

Chia-Kai Liu*
DSP, Inc.
Taipei, Taiwan
ck@dsp.im

T. C. Hsieh
DSP, Inc.
Taipei, Taiwan
johnson@dsp.im

ABSTRACT

To understand or solve complex social problems, it requires close collaboration among multi-sector stakeholders. However, these stakeholders are often divided by their own data silos. The D4SG Initiative (Data for Social Good) adopts a platform strategy to connect and empower change agents across these data silos. After working with government agencies and nonprofit organizations of various levels of data maturity in Taiwan, we have identified multiple success factors and common pitfalls during the process of fostering data-driven social innovation.

Keywords: data thinking; multi-stakeholder partnership; data journey map

I. INTRODUCTION

Public service providers, such as government agencies, nonprofit organizations and social entrepreneurs, usually have the best of intentions to serve society and improve social conditions. But often their solutions fall far short of what they want to accomplish and what is truly needed. Take social work as an example. Social workers are good at case management, but many of them cannot see the forest for the trees and develop the tendency to ignore general patterns or fail to think beyond their lines of responsibility. Consequently, the answers they propose and fund sometimes produce the opposite of what they want over time.

Systems thinking is regarded as a solution to this situation. The systems thinking approach refers to the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them in order to produce desired effects.

When applying systems thinking to tackle the growing complexity of social problems, we often need to connect change agents, the individuals or groups that undertake the task of initiating and managing change in an organization, across organizational silos.

How do we identify the change agents within an organization? How do we identify the change agents located in

various places along the “customer journey” of a social work case or a public policy issue? Once we find these change agents, how do we connect or empower them to make meaningful changes? Furthermore, after a one-time, project-based success, how do we stimulate paradigm shifts and create lasting impact?

The D4SG Initiative (Data for Social Good, D4SG) employs data science methods to enable change agents across data silos. It is our attempt to answer the above questions. After working with dozens of government agencies and nonprofit organizations in Taiwan, we have developed several tools to increase data literacy in organizations that are not tech-savvy, and learned how to use data as a common language for stakeholder engagement, for fostering collaboration across agencies and sectors, and for encouraging development beyond one-time, ad hoc projects.

II. THE D4SG INITIATIVE

A. The Beginning

In 2015, the D4SG Initiative¹ was founded by three entities: DSP, Inc.² (a data consulting company), the Frontier Foundation³ (a nonprofit organization), and ITSA⁴ (an academic research and education program funded by the Ministry of Education).

The original mission of D4SG was simple: to help public service providers use their own data to improve operational efficiency.

After studying the data-driven collaborative consulting methods pioneered by DataKind⁵ and Data Science for Social Good Fellowship at the University of Chicago⁶, we decided to

¹ <http://d4sg.org/>

² <https://dsp.im/>

³ <http://www.frontier.org.tw/>

⁴ <https://www.itsa.org.tw/>

⁵ <http://www.datakind.org/>

⁶ <https://dssg.uchicago.edu/>

take a similar approach: a short-termed, project-based fellowship program.

B. The Pilot Runs

Since it was the first time for this kind of events to take place in Taiwan, many people were not familiar with the concept of using data science for social good. In addition, most nonprofit organizations (NPOs) in Taiwan were not early adopters of information technology or innovative approaches. Therefore, we organized a two-day hackathon as a test drive for NPOs.

Then, in order to understand our limits of execution, we began with two small pilot projects: (1) Understand the nationwide heavy metal contamination in farmland soil⁷ and (2) Work schedule optimization for visually impaired masseurs.⁸

For these pilot projects, we only recruited college and graduate students with background in computer science, statistics and journalism. We also only allocated two months for project execution.

During the process, we followed the basic work principles of data science: problem definition, data inventory, data transformation, data analysis, and post-analysis actions. (Fig. 1)

agriculture and environmental protection agencies announced they would take steps to improve farmland soil monitoring.¹⁰

C. The Working Process

After the pilot runs, we decided to modify our approach, both in the width of offerings and the depth in customer life cycles.

First, we learned that due to the variance of data maturity in each organization, a solution-oriented, fast-paced, time-intensive project may not be the best answer to everyone’s problems. Some organizations can query records from well-organized databases with ease, while others are still using pen-and-paper approach to collect data on the frontline of work. Therefore, instead of offering a one-size-fits-all fellowship program, we provide a wide range of tools and services for each stage of data-driven innovation. (Fig. 2)

Secondly, we discovered that due to the lack of resources or experience, some organizations had difficulties in implementing the results of short-term projects. So, instead of “fire-and-forget”, we created an “after-sales service” to keep in touch with every organization we have worked with.

Following this working process, so far we have completed three tours of fellowship, three hackathons, a data prediction challenge, and seventeen “data mixer” meetups.

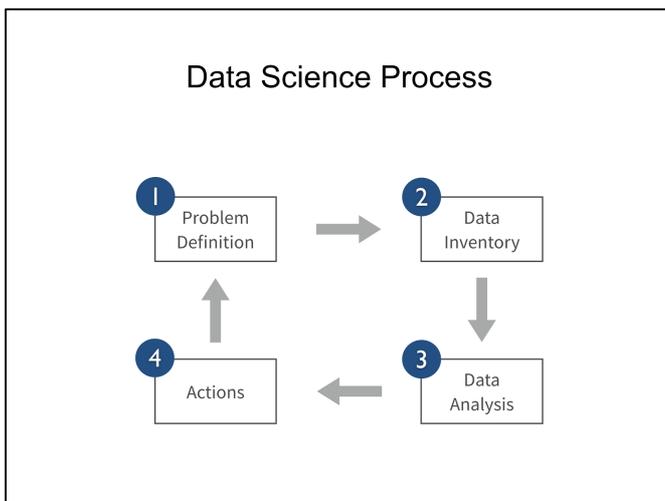


Fig. 1. The Data Science Process.

After project completion, we hosted a public meetup to share the findings of these projects. The audience was consisted of potential data fellows and non-profit organization representatives. The meetup was well-received.

Major news media also reported the findings of the farmland contamination project.⁹ As a consequence, the

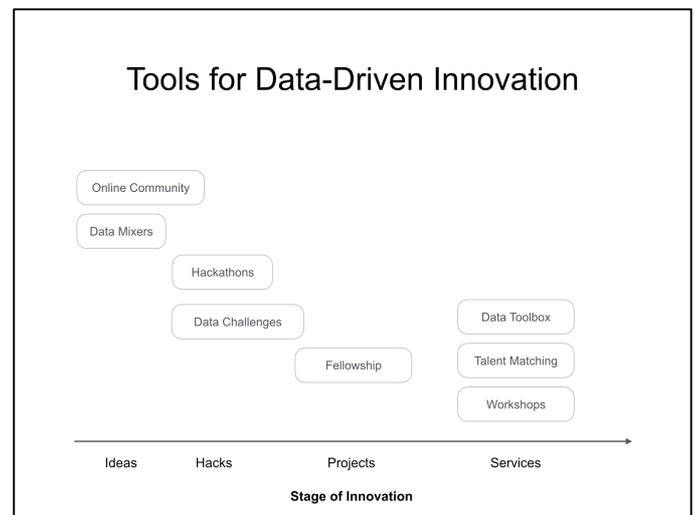


Fig. 2. Tools for Data-Driven Innovation.

III. LESSONS LEARNED

After collaborating with more than two dozens of government agencies and nonprofit organizations, we have discovered quite a few common barriers for starting a data-driven innovation project in public services, including:

- Routine work overload
- Lack of data analysis skills
- Lack of proper data technology

⁷ <http://d4sg.org/farmland-pollution/>

⁸ <http://dspim.github.io/visually-impaired-masseurs/>

⁹ <http://news.pts.org.tw/article/324954>

¹⁰ <https://dsp.im/2016/06/d4sg-data-driven-cooperation/>

