

# Datathons: An Experience Report of Big Data Hackathons

Craig Anslow, John Brosz, Frank Maurer  
Department of Computer Science  
University of Calgary  
Calgary, Alberta, Canada  
{canslow,jdlbrosz,fmaurer}@ucalgary.ca

Mike Boyes  
Department of Psychology  
University of Calgary  
Calgary, Alberta, Canada  
boyes@ucalgary.ca

## ABSTRACT

Hosting successful hackathons is difficult. With the accessibility of open data there is now an opportunity to host hackathons focused on big data, however these are new and not clearly defined about how to do this process. In this paper we present our experience at hosting three big data hackathons called *datathons* that involved students and members from industry coming together to solve challenging problems with publicly open data and data from not for profits. The resources developed for our datathons about our experience will help inform others who also wish to host big data hackathons.

## Categories and Subject Descriptors

K.3.2 [Computer and Information Science Education]:  
Computer science education

## General Terms

Design

## Keywords

Big Data, Hackathon, Open Data

## 1. INTRODUCTION

## 2. RELATED WORK

Games and hackathons  
Stitch fest paper  
explain what big data is  
explain we are exploring big data hackathons

## 3. DATATHONS

3. Datathons - describe what a datathon is 3.1 Benefits  
3.2 DFG Case Study 3.3 CODE Case Study 4. Discussion  
(Lessons Learned) - community engagement - participants:

recruitment, group dynamics - data: structuring and cleaning, open data, storing - tools: analysis, visualization - logistics: room setup, break out rooms, food - things to avoid  
5. Conclusions

? Intro ? Related Work ? Datathons in general ? benefits  
? Two hackathon case studies ? Code ? Recruitment -  
population of attendees ? Preparing data ? Setup ? Group  
dynamics ? DfG ? Lessons Learned ? How to do, tips ?  
Things to avoid

## 4. PREPARATION AND PLANNING

### 4.1 Recruitment

### 4.2 Venue

### 4.3 Data Preparation

### 4.4 Logistics

## 5. HOSTING THE EVENT

## 6. DISCUSSION

**Participants.** recruitment, group dynamics, community engagement

**Data.** structuring and cleaning, open data, storing

**Tools.** analysis, visualization

**Logistics.** room setup, break out rooms, food

**Things to avoid.**

## 7. CONCLUSIONS

## Acknowledgments

Thanks to the Government of Canada, City of Calgary for providing open data sets. Thanks to the Data For Good Calgary Meetup group for assisting with the datathon event.