
Designing Online Experiments: Citizen Science Approaches to Research

Andy Pilny

University of Kentucky
andy.pilny@uky.edu

Brian Keegan

Harvard University
bkeegan@gmail.com

Brooke Foucault Wells

Northeastern University
b.welles@neu.edu

Chris Riedl

Northeastern University
c.riedl@neu.edu

David Lazer

Northeastern University
d.lazer@neu.edu

Jason Radford

University of Chicago
jsradford@uchicago.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).

CSCW '16 Companion, February 27 - March 02, 2016, San Francisco, CA, USA
ACM 978-1-4503-3950-6/16/02.

<http://dx.doi.org/10.1145/2818052.2855516>

Katya Ognyanova

Rutgers University
kateto@gmail.com

Leslie DeChurch

Georgia Institute of Technology
leslie.dechurch@psych.gatech.edu

Michael Macy

Cornell University
m.macy@cornell.edu

Noshir Contractor

Northwestern University
nosh@northwestern.edu

Waleed Meleis

Northeastern University
meleis@ece.neu.edu

Abstract

New information technologies allow for new modes of data collection, provide unparalleled computational resources, and facilitate long-distance collaboration. Programs like *Galaxy Zoo*, *Fold.It*, and *ReCAPTCHA* demonstrate the power of recruiting online volunteers as participants and contributors to research. Platforms like *Project Implicit*, *SciStarter*, and *Volunteer Science* confirm that a wide range of social scientific research can be conducted with volunteers in online laboratories.

This workshop will explore the opportunities, challenges, and limitations of online experiments and build a community of scholars performing online experiments. A group of scholars with wide-ranging experience in this area will discuss existing and emerging approaches to online experimentation. Participants will learn about designing and running online studies, and recruiting and managing subjects. Participants will then be given access to the Volunteer Science research platform (volunteerscience.com) where they will work with researchers to develop, test, and deploy their own online studies.

Duration

1 Day

Participants

Number of Participants: 25

Participant Selection

Participants will be encouraged to apply for the workshop in teams of 1 to 5 members. Each team will submit a brief proposal for an online experimental design. Submissions will be evaluated by the organizers based on the nature of the proposed experiment, its innovativeness, relevance, and applicability to the Volunteer Science platform. The proposals would also have to meet the technical and ethical standards specified in advance by the organizers. During the second half of the workshop, each team will work on implementing the proposed experimental project on the platform.

Author Keywords

Online Experimentation, Citizen Science, Social Science, Computer Science

ACM Classification Keywords

Design, Experimentation, Measurement, Performance

Description

New information technologies are changing the way we do science. One promising new development in empirical research methodology is the emergence of online experimentation. Compared to their offline equivalents, online experiments are more customizable, more scalable, and allow for much wider and more diverse participation beyond traditionally WEIRD (Western, Educated, Industrialized, Rich, and Democratic) samples. They allow for designs that minimize experimenter effects, promote replicable methods, and lower the cost of conducting rigorous empirical research at a large scale. Modifying or repurposing existing experimental settings is much easier online, as is the open sharing of technologies, standards, and protocols.

At the same time, online experimentation presents scholars with a host of new challenges. We are still exploring whether being online affects the behavior of individuals in a way that differs from that in an offline setting. Researchers also have to consider problems like self-selection bias, lack of full experimental control, experimental mapping (i.e., extent that manipulations offline translate into virtual worlds), and the increased possibility of participants misunderstanding instructions or engaging in subversive behavior. In addition, online frameworks can present significant technological and ethical challenges, especially with regard to ensuring

individual privacy and data security. Motivating Internet users to participate in a web experiment can also be a resource-intensive and difficult task.

In the first half of our workshop, we will host presentations and discussions on performing online experiments from experts in different social scientific fields. These talks will focus on the process of conducting experiments online, including designing experiments for online environments, the types of experiments that can be conducted online, and how to recruit participants and validate subject behavior. In the second half of our workshop, we will address a series of issues related to the specifics of developing online experiments.

Online experiments require a great deal of technical expertise to create and maintain, significant investments in subject management and recruitment, and present their own limitations as a mode of research. Most researchers ultimately create a single experiment using customized software which integrates only with their specific subject pool. The weakness of this decentralized, ad hoc approach is a continued dependence on narrow subject pools, experiments that remain difficult to replicate or duplicate on other systems, and a continuous re-inventing the wheel as every researcher must solve the same problems like recruiting large numbers of participants.

In response to this, the organizers of the workshop created a platform for developing and deploying online research called Volunteer Science (volunteerscience.com). Volunteer Science provides researchers with an API which enables them to implement common experimental features in an easy-

Participant Recruitment

Calls for participation will be circulated on relevant mailing lists (CSCW, AoIR, SOCNET, CITASA, CRTNET, and others). The workshop will be announced on social media (Twitter, relevant Facebook pages). The organizers will invite social scientists performing experiments to present, participate in a panel discussion, and participate as attendees. The organizers will also mobilize their own social networks to recruit interested participants who may not have created online experiments before.

Equipment

We will request a projector for the discussion part of the workshop. Participants will be encouraged to bring their own laptops.

to-use way. The API includes simplified protocol for randomizing subjects, recording data, and integrating with other platforms like Mechanical Turk and Qualtrics. There are also over two dozen experiment templates researchers can use to learn the API and implement these features without having to create them from scratch. And, the API will soon give researchers the ability to access new types of digital data from volunteers who have consented to donate their information for research. Examples include Facebook data, mobile phone data, and browser data.

In the second half of the workshop, attendees will have the opportunity to work with developers and researchers to create their own experiments using Volunteer Science. Attendees will be given access to the experiment development system, system documentation, experiment templates, and testing platform. The workshop participants will be able to develop and test their own experiments or link survey-based experiments into Volunteer Science and, when possible, pilot their study. After the workshop, interested attendees who have IRB approval and meet basic usability guidelines will be able to post their experiments on VolunteerScience.com and recruit subjects.

Activities and Goals

The workshop will include three interrelated components.

1. Discussion of key challenges in online experiments and citizen science.

The organizers will provide an overview of the cutting edge developments in the field. They will describe the

process of setting up virtual laboratories, recruiting participants, and conducting experimental studies. The organizers will provide useful guidance for future work in that area, describing the obstacles encountered along the way and the solutions that were found. They will discuss reliability and validity in online experimental designs, as well as the extent to which web replications can reproduce results obtained in brick and mortar settings.

2. An overview of Volunteer Science: an open platform for online experiments

After the broad overview of issues related to experimental research on the Internet, workshop attendees will be introduced to one available platform that can support projects of that kind. Volunteer Science provides a scalable, open, and flexible framework for the development and deployment of social experiments. It was designed to facilitate large-scale, multi-person, and synchronous work. The VolunteerScience.com website is built on top of free open source development tools including Django, Bootstrap, HTML5, and JavaScript; enabling it to work on any device with modern web-browsing technologies. The platform uses Amazon Web Services to provide high-performance and scalable on-demand resources.

3. Hackathon

The second half of the day will be dedicated to a hackathon during which attendees will create their own online research. The session will begin with two minute talks from teams in attendance about the studies they intend to develop. Teams will be given a researcher account on Volunteer Science. This will provide them

with access to the experiment development system, API, and testing framework. Developers and researchers will be on hand to answer questions about study design, user experience, the API, or subject recruitment and management. At the end of the session, teams will present on their study. By the end of the workshop, attendees will be aware of the basics issues involved in online social and behavioral research, familiar with the Volunteer Science system, and will have completed a functional prototype of their online study.

Organizers and Bios

Andy Pilny

Andy Pilny is an assistant professor in the Department of Communication at the University of Kentucky. His research is in organizational/small group communication and computational/ network science.

Brian Keegan

Brian Keegan is a research associate at the Harvard Business School, where he works as a data scientist on the HBX online learning platform. His research examines how social media like Twitter and Wikipedia can be used to improve predictive models of electoral success, as well as performing small group experiments using the Volunteer Science platform.

Brooke Foucault Welles

Brooke Foucault Welles is an assistant professor in the Department of Communication Studies at Northeastern University, and a faculty affiliate of the Network Science Institute, NU-Lab for Texts, Maps and Networks, and the program in Women's, Gender, and Sexuality Studies. She studies how social networks,

including online social networks, can be used as a resource to achieve personal, team and social goals.

Chris Riedl

Christoph Riedl is assistant professor for Information Systems at the D'Amore-McKim School of Business at Northeastern University. He also holds an appointment at the College of Computer & Information Science and the Network Science Institute. He is a fellow at the Institute for Quantitative Social Science (IQSS) at Harvard University. He applies and develops novel computational approaches and machine learning to analyze (experimental) field data in order to study the design of contests, crowdsourcing, and collective intelligence mechanisms.

David Lazer

David Lazer is Distinguished Professor of Political Science and Computer and Information Science, Northeastern University, and Co-Director, NULab for Texts, Maps, and Networks. His research focuses on the nexus of network science, computational social science, and collaborative intelligence. He is the founder of the citizen science website Volunteer Science and the political visualization website VisPolitics.

Jason Radford

Jason Radford is a doctoral student in Sociology at the University of Chicago and Manager of the Volunteer Science website. His research focuses on how people create social order in organizations.

Katherine Ognyanova

Katya Ognyanova is an Assistant Professor at the School of Communication and Information, Rutgers University. She does work in the areas of computational

social science and network analysis. Her research interests include social use of technology, mass communication and media effects, political actors and political behavior.

Leslie DeChurch

Leslie DeChurch is an Associate Professor of Industrial and Organizational Psychology at Georgia Institute of Technology, where she is the Director of the DELTA (Developing Effective Leaders, Teams, and Alliances) research group. Her research interests include leadership and teamwork in organizations.

Michael Macy

Michael Macy is the Goldwin Smith Professor of Arts and Sciences, Department of Sociology and Department of Information Science at Cornell University. He is the director of the Social Dynamics Laboratory, which studies the interplay between network topology and dynamics of social interaction using computational models, online data, and experimental designs.

Noshir Contractor

Noshir Contractor is the Jane S. & William J. White Professor of Behavioral Sciences in the McCormick School of Engineering & Applied Science, the School of Communication and the Kellogg School of Management at Northwestern University, USA. He is the Director of the Science of Networks in Communities (SONIC) Research Group at Northwestern University.

Waleed Meleis

Waleed Meleis is an Associate Professor in Electrical and Computer Engineering at Northeastern University. Professor Meleis's research is on developing and

evaluating algorithms and bounds for combinatorially difficult optimization problems.