Mahalo Button: Building the Network of Gratitude for Sharing the Dataset Usage

*Asanobu Kitamoto[1], Yoko Nakahara[2], Hiroyuki Shimai[2], Toshiyuki Shimizu[3]

1. National Institute of Informatics, 2. Kyoto University, 3. Kyushu University

Open data is a crucial driver for promoting data-driven research and innovation, but experts supporting this trend are often considered underrepresented. To solve this problem, we propose a system called "Mahalo Button" to create the hub of information for the dataset usage with pay-back and pay-forward sharing [1]. 'Mahalo' is a Hawaiian word for 'thank you,' but it has a broader meaning, such as admiration, praise, esteem, regards, and respects. Like data citation, the Mahalo Button also aims to make the contribution of dataset creators more visible but using a different mechanism called the network of gratitude.

The Mahalo Button is a simple button-based system, similar to the Like Button on the Web. Each button has a unique ID, corresponding to the DOI (Digital Object Identifier) or the URL of the dataset landing page. The button accepts a "thank you message" from a user of the dataset with the DOI of research publications, such as papers. As it receives new messages, the button becomes the hub of information that accumulates information about research publications describing the usage of the dataset. The button thereby becomes an inspiring resource for a potential dataset user, in which case studies of the dataset are available.

Here, the "thank you message" corresponds to the pay-back sharing in the sense that it notifies the dataset creator about the actual usage of the dataset with the increased number of messages as a proxy to demonstrate the contribution of the dataset. On the other hand, the button also works as pay-forward sharing in the sense that a message from a dataset user is also helpful for a potential dataset user. If a potential dataset user decides to use the dataset and produce a new research publication to be added to the Mahalo Button, it will also be helpful for future potential users. We aim to create a virtuous cycle of data use and results by using the Mahalo Button as a hub of information to share and visualize the various contributions related to data.

In November 2021, we introduced the Mahalo Button to DIAS (Data Integration and Analysis System) in several steps. First, as a data repository, DIAS collected the dataset usage and registered them to the Mahalo Button. Here the DOI that uniquely identifies the dataset plays a critical role. If dataset users properly cite the DOI of the dataset in their publications, we can collect information about which paper used which dataset using search engines. For example, we already registered 49 papers in the Mahalo Button for the dataset, Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1). On the landing page of the dataset [2], the Mahalo Button shows the number 49, indicating that at least 49 papers used the dataset. You can also browse the list of papers on the "Show Mahalo" page with links to the DOI of papers [3]. Some papers lack proper data citation referring to the DOI of the dataset, but the data curators of the DIAS data repository registered new entries through reading the content and identifying the dataset actually used in the paper.

DIAS already has a system for reporting the usage of the dataset through a Web form, but it is a closed system not designed for sharing information. We believe that the Mahalo Button offers an alternative and highly effective approach for sharing the dataset usage on the data repository. Deployment of the Mahalo

Button to any data repositories is complete with an installation of a few lines of a code snippet. We continue to enhance the functionality of the Mahalo Button to meet the various needs of data repositories.

[1] Asanobu KITAMOTO, "Return on publication (RoP): "DOI return button" for networking data creators and users with pay-back and pay-forward incentives," Abstracts of Japan Geoscience Union (JpGU) Meeting 2018, No. MGI23-P06, 2018.

[2] Hyungjun Kim. (2017). Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set]. Data Integration and Analysis System (DIAS). https://doi.org/10.20783/DIAS.501

[3] Show Mahalo for "Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1)", https://mahalo.ex.nii.ac.jp/button/bdc0f3e5-3747-4fe3-9b8d-2bf5bc1f24aa/mahalo (Accessed February 17, 2022).

Keywords: Mahalo Button, network of gratitude, data citation, dataset usage, data repository, DIAS

