



Conference paper

## Communal Data Workflow in TaiRON (Taiwan Roadkill Observation Network).

Tyng-Ruey Chuang <sup>1</sup>, Te-En Lin <sup>2</sup>, Yi-Hong Chang <sup>3</sup>, Chih-Yun Chen <sup>2</sup>, Yu-Kai Chen <sup>2</sup>,  
Ping-Keng Hsieh <sup>1,2</sup>, Guan-Shuo Mai <sup>4</sup>

- 1) Institute of Information Science, Academia Sinica, Taiwan.
- 2) Taiwan Endemic Species Research Institute, Council of Agriculture, Taiwan.
- 3) Word Gleaner Ltd., Taiwan.
- 4) Biodiversity Research Center, Academia Sinica, Taiwan.

### Summary

TaiRON (Taiwan Roadkill Observation Network) is a collaborative project in collecting records about road kills and other types of animal mortality incidents in Taiwan. It first started in a social media setting – a Facebook group – where observations of road kills (including images, locations, dates, and other data) were submitted by volunteers. Experts identified species in the observation records, and used the collected records for research. As of 2016-07-10, there are 10,674 volunteers in the Facebook group. The data collection process in TaiRON has been communal in nature – without the volunteers, there will be no data. Furthermore, the data so collected shall be useable by the volunteers as well. Before sharing the data, however, there are a number of issues to address: the quality and period of data releases, the handling of sensitive data (the locations of endangered species, for example) and personal privacy, as well as dataset formats and access APIs. The data acquisition and management workflow in TaiRON has undergone many changes since the project started in 2011. Mobile phone apps are now used for reporting, and the observations are first uploaded to the TaiRON website before they are announced to the Facebook group. User feedbacks, such as species identifications, from the Facebook groups are incorporated in order to make high-quality datasets. In this presentation we will summarize the current data workflow in TaiRON, emphasizing its communal characteristics.

### Collaborative Roadkill Observations at TaiRON

TaiRON (Taiwan Roadkill Observation Network) is a collaborative project in collecting observations about road kills and animal mortality incidents in Taiwan. The project started at the Taiwan Endemic Species Research Institute (TESRI) in 2011. It was one of the first to use the social media, in particular the Facebook, for large-scale ecological surveys. Everyone can participate. They just join the a Facebook user group, and post images of roadkill along with dates, locations, and other information. Researchers, and participants too, identify the species in the Facebook posts. Many research projects have also used Facebook to enlist volunteers to identify species (Sidlauskas 2011), or have built systems to crowdsource classification tasks to volunteers (Raddick 2010). These projects often apply to

Title... Communal Data Workflow in TaiRON (Taiwan Roadkill Observation Network).

pre-existing data collections. In TaiRON, the participants themselves both provide the observations and collaborate with researchers in identifying the species.

The observation records are collected and aggregated for research use (Lin 2012; Mai *et. al.* 2012). Since early 2013, researchers at both the Institute of Information Science and the Biodiversity Research Center, Academia Sinica, have worked with researchers at the Taiwan Endemic Species Research Institute in a joint project to develop novel information systems and services to better support and sustain this collaboration with a large number of participating observers. Among the many issues needed to be address is how to lessen the dependence on the social media, and on the development of a process to derive normalized research datasets from raw observation collections (Chuang 2014; Lin *et. al.* 2105).

## **The Data Workflow at TaiRON**

The TaiRON data workflow involves people of many backgrounds and skills, and they participate in different stages. Foremost the roadkill observation project is a collaboration in which everyone can participate. The current data workflow evolves from the practices used in the project over the years. Ecological scientists and computer scientists work together with voluntary participants and IT experts in deriving the current workflow. As of July 2016, observation data in the TaiRON collaboration passes through and is kept at two major information services: the TaiRON website itself, as well as a Facebook user group (TaiRON: Taiwan Roadkill Observation Network 2013—now; Reptile Road Mortality 2011—now). Three additional tools are employed. The first one is an App (for Android-based smartphones) to record and submit observations (TaiRON APP 2014—now ). Use of the App is optional; participants can just post their observations to the Facebook group (this practice is now depreciated) or upload them directly to the TaiRON website without using the App (encouraged). The second is a crawler that collects posts (along with images, time and location data, and other annotations) on the Facebook Group and put them in an aggregated table of all observations. The third tool is called the Little Helper. It is a browser extension (currently only for Chrome) to allow researchers to make corrections or add annotations to individual records in the aggregated table. We outline the current data collection workflow below.

Participants are encouraged to submit observations using the App, or to upload them to the TaiRON website. The App actually uploads observations to the TaiRON website first. The App and the website form guide the observer into filling data fields that together provide a structured record, including the date and location of the observation, and a preliminary identification of the species and the possible cause of death. Once the TaiRON website receives an observation record (either via App or by direct upload) it produces a digest of the observation, and posts the digest to the Facebook Group. Researchers and all participants can comment on the post, say by providing further information about the species, and engage in discussions on Facebook. There are many benefits for the observation to first go through the TaiRON website: The website now keeps the original image of the roadkill, and gate-keeps sensitive information (by not posting observations about endangered species and their locations to the Facebook Group, for example). The observer can also choose to remain anonymous when TaiRON posts the digest (by not associating the post to any Facebook identity; the default is to tag the observer's Facebook ID in the post), or to not post to the Facebook Group at all (researchers and the observer, however, can still access the record at the TaiRON website). Note that for the above process to work, we need to link the

Title... Communal Data Workflow in TaiRON (Taiwan Roadkill Observation Network).

participants' accounts at TaiRON with their Facebook IDs. This linkage can be set up by asking the participants to login into the TaiRON website by the Facebook IDs.

However, some roadkill observations are still directly posted to the Facebook Group (not by App nor via the TaiRON website). These observations are not in the TaiRON website. Although this practice is depreciated, it is still used. The crawler mentioned earlier will collect these (as well as all other posts) into an aggregated table. As these posts are less structured (e.g. using free text to express dates and locations), the crawler may not be able to normalize information in a post into field values required by a record. This is where the Little Helper comes to help. The researchers use the Little Helper (a browser extension) when viewing Facebook Group posts, and use it to check a post against its respective record in the aggregated table. The Little Helper is also used to correct errors and add missing values to observation records in the table. The aggregated table is further organized into ways that are good for retrieval and visualization at the TaiRON website, and are made fitting for all to download and reuse.

Note that in the above workflow, researchers and participants interact on the Facebook Group, but roadkill observation records are first received and kept at the TaiRON website. A data record flows from TaiRON website (via use of the App or not) to the Facebook Group, and after discussion there, additional information flows back to TaiRON website to enrich the very data record. For observations posted directly to the Facebook Group, they are crawled back to TaiRON and normalized into structured records (again, taking into findings from the discussions on Facebook). This data workflow can be replicated for other participatory ecological observation projects. The TaiRON data workflow is communal in nature: Participants contribute their observations, the observations are discussed openly and aggregated into a collection accessible to all on the TaiRON website. In addition, each participant has access to all the observations records (normalized and enriched) she has contributed at the TaiRON website.

## **Dealing with Sensitive Information**

We are concerned about two issues in the handling of sensitive information in TaiRON, or any large-scale collaborative ecological surveys in which everyone can participate. The first is about information about rare species. The second is about the privacy of the participants. If an observation is about an endangered species, sending it to an open forum (e.g. a Facebook Group) for discussion will not be appropriate. We encourage the practice of sending observations to TaiRON website first, and then have the website post digests to the Facebook Group. On the App and the submission page of TaiRON, an observer can indicate whether an roadkill is likely to be a rare species. Such an observation is kept at the website, and will not be published to the Facebook Group without being approved. The approval process can be programmed into the website so that researchers get notices and review them as they arrive. Also, for aggregated datasets published on the TaiRON website, they do not include records about endangered species.

In the current workflow a participant, by using the App or by uploading to the TaiRON website, has the option of sending an anonymous observation digest to the Facebook Group. The observer's identity, of course, is kept at the TaiRON website and known by the researchers, but is not available to others at the Facebook Group. We feel that this provides some protection of the observer's whereabouts information (as will be revealed by the time

Title... Communal Data Workflow in TaiRON (Taiwan Roadkill Observation Network).

and location of the observation record). When the observations are aggregated into datasets, we remove the observers' IDs in the records before the datasets are made available to others. On the other hand, as many participants have opted to release their observation photos under the Creative Commons Licenses, they must be properly attributed when the image files are released. Participants can have photos attributed to their nicknames, or require no attribution at all (by the use of the CC0 Public Domain Dedication).

## TaiRON: Still Evolving

After five years, the TaiRON project is still evolving. In the beginning in 2011, it was an experimentation in the use of social media to help collect information about reptile mortality. Since then, more than 10,000 people has joined the Facebook Group to participate in roadkill observations. A dedicated website had been set up to reach to the general public to help reduce roadkill. The website is now at the same time a collection of educational resources, a place to find and download data for research, and, perhaps most importantly, where all participants can witness the outcome of their collaborative efforts. In addition, each participant can login into her page where all her observation records have been kept in details and enriched with information contributed by researchers and other participants.

We are currently working to further evolve the TaiRON project on several aspects. The first is on better packaging of the observation records using domain vocabularies (such as Darwin Core) and standardized publishing tools and formats (for example, IPT from GBIF). The second is to develop a new Web App — an App that needs only Web browsers in order to run — for uploading observation records to the TaiRON website. We are moving away from the so-called native Apps as they all depend on specific mobile operation systems and are often costly to maintain. Last but not the least, we are planning to work on a generic software package for various kinds of collaborative ecological surveys. There have been requests to use the software we have developed for TaiRON for other purposes. Currently, the TaiRON project uses a customized setup of information systems and services. As the TaiRON project is still evolving, we will continue to look for and adapt good practices used in all citizen science projects.

## References

**Chuang, T-R** 2014 *Arrangements for Data Sharing and Reuse in Citizen Science Projects*. In: International Conference on Data Sharing and Integration for Global Sustainability (SciDataCon). New Delhi, India, November 2-5, 2014.

**Lin, T-E** 2012 *A Citizen Science Project on Roadkill in Taiwan: Facebook Groups as a Corresponding Platform*. In: The 23rd International CODATA Conference. Taipei, Taiwan, October 28-31, 2012.

**Lin, T-E, Chen, C-Y, Chen, Y-K, Chuang, T-R, Deng, D-P, Hsu, C-H, Lin, L C-H, Mai, G-S, Shao, K-T, and Wang, M-H** 2015 *Collaborative Ecological Observation: Issues in Moving from Social Media to Research Data*. In: Citizen Science 2015 Conference. San Jose, CA, USA, February 11-12, 2015.

Title... Communal Data Workflow in TaiRON (Taiwan Roadkill Observation Network).

**Mai, G-S, Hsu, C-H, Lin, D-E, Lin, H-H, Deng, D-P, and Shao, K-T** 2012 *Harvesting crowdsourcing biodiversity data from Facebook group*. In: The 2nd Asia Regional Conference of the Society for Conservation Biology. Bengaluru, India, August 7-10, 2012.

**Raddick, M J, Bracey, G, Gay, P L, Lintott, C J, Murray, P, Schawinski, K, Szalay, A S, and Vandenberg, J** 2010 *Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers*. *Astronomy Education Review*, 9(1): 010103.

**Reptile Road Mortality** (Facebook Group) 2011—now  
<<https://www.facebook.com/groups/roadkilled/>>.

**Sidlauskas, B, Bernard, C, Bloom, D, Bronaugh, W, Clementson, M, and Vari, R P** 2011 *Life in Science: Ichthyologists Hooked on Facebook*. *Science*, 29;332(6029): 537.

**TaiRON: Taiwan Roadkill Observation Network** (Website) 2013—now  
<<https://roadkill.tw/>>.

**TaiRON APP** (App) 2014—now <<https://play.google.com/store/apps/details?id=roadkill.tw>>.