
Early version, also known as pre-print

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Citizen Science and Crowd-sourcing
Biological Surveys

Simon Price, Dave Kilbey, Jasper Tredgold, Christopher Bailey,
Mark Van Rossum, Mike Jones and Damian Steer
IT Services R&D / ILRT, University of Bristol
8-10 Berkeley Square, Bristol BS8 1HH, UK
{simon.price,d.kilbey}@bristol.ac.uk

We report on the NatureLocator programme of digital research projects centred around the development and use of mobile applications to collect crowd-sourced data for biological surveys\(^1\). With high-profile coverage by the BBC and other media, the initial NatureLocator project, *Leaf Watch: Conker Tree Science*, used a bespoke smartphone app to gather geo-located observational data from members of the public and created the most comprehensive information on the UK distribution of the Leaf-Mining Moth (Cameraria Ohridella) to-date. Subsequent NatureLocator projects have refined and extended this approach to gather research data and raise awareness of other ecological threats. These include *Plant-Tracker* for recording sightings of invasive plant species, *SealifeTracker* for invasive and climate change indicator marine species, *iRecord Ladybirds* for the UK Ladybird Survey, and *AquaInvaders* for invasive freshwater species. The screenshots below show two screens from iRecord Ladybird, one of the most recently released NatureLocator apps.

Data showing the scale of interest and use for some of the applications produced by NatureLocator projects are given in the table below. We are still gathering data for the most recently launched surveys.

<table>
<thead>
<tr>
<th>NatureLocator Project</th>
<th>App Downloads</th>
<th>Records Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeafWatch</td>
<td>15,000</td>
<td>8,000</td>
</tr>
<tr>
<td>PlantTracker</td>
<td>12,000 (^1)</td>
<td>4,500</td>
</tr>
<tr>
<td>iRecord Ladybirds</td>
<td>6,000 (^1)</td>
<td>2,000</td>
</tr>
</tbody>
</table>

\(^1\) 7,300 iPhone and 4,700 Android.
\(^2\) Since release in May 2013.
By way of comparison with the above figures, before the LeafWatch app was released, the annual Leaf-Mining Moth survey typically gathered around 500 submissions per year. Therefore, in terms of numbers of records submitted, LeafWatch has made a significant improvement to this biological survey, which received around 5,000 submissions in the first year alone (the table shows cumulative figures to-date, which amounts to two years for LeafWatch). There are other advantages to using mobile apps besides increasing the quantity of data. Prior to LeafWatch, submitted records arrived by email, had a multitude of image formats submitted, were poorly geo-located and often required manual follow-up by a researcher to clean the data. The impact of these apps extends beyond the actual survey itself. For instance, following a feature on BBC Countryfile, the Environment Agency has reported that PlantTracker directly contributed to the treatment of many outbreaks of Floating Pennywort nationwide.

All the above projects rely on end-users to manually identify species of interest, using the inbuilt guides, photos and instructions. A recent project, BatMobile, also investigates the feasibility of performing automated species recognition based on audio in a native app. Aimed at the enthusiastic amateur and professional researcher, BatMobile processes audio recorded using an external ultrasound microphone to produce a compressed feature vector that is then used as input to either an on-phone Support Vector Machine classifier or, to a back-end web service Neural Network classifier. Although only a proof-of-concept application, BatMobile has demonstrated that this approach is now viable on a high-end smartphone.

As the NatureLocator programme matures, we reflect on some of the lessons learnt about the practicalities of running such projects. In each project, the development team works with the researchers to build a custom mobile app with a user-centred focus; participants download the app and go out into the field to collect photographic (or audio) geo-located records; and finally, these are verified by nominated experts (or members of the public) using an easy-to-use system, with the aim of producing high-quality validated results. All the NatureLocator projects are intrinsically multidisciplinary and involve collaboration between researchers, information technologists, external agencies and, because these are citizen science projects, amateur biologists, ecologists and conservationists as well as less knowledgeable, but equally enthusiastic, members of the general public. We strongly suspect that, although we have no negative example to support this, a key element in the success of these projects has been the inclusion of a domain expert in each project team.

Another factor has been the continuous effort to generalise and re-use functionality within the software. We rate this factor as more important than the actual implementation technology adopted. Apart from BatMobile, all NatureLocator applications to-date have used the Phonegap open source framework to create cross-platform mobile apps (as opposed to coding native apps for each platform). This approach, although intuitively appealing, has not been without its problems and has resulted in a great deal of effort on occasion to workaround bugs in the framework which we do not have the resources to investigate in the framework source code. Also, the amount of effort required to give a high-quality user experience can be greater when working in a mobile web context, particularly where user expectations of user interface conventions on different platforms may differ. Looking to the future, we would not rule out moving NatureLocator development over to data-driven native apps, although in an environment of diversification of deployment platforms the cost of supporting all platforms may be prohibitive.