

# SUSTAINABLE COASTLINES CHARITABLE TRUST

## LITTER INTELLIGENCE QUALITY ASSURANCE AND QUALITY CONTROLS

### Version 1.0

**Prepared by:** Camden Howitt, Shawn Elise Tierney, Shelley Butt, Ben Knight

**Date:** 15 October 2020

#### **Purpose of document:**

This document provides an overview of the Quality Assurance measures and Quality Controls that have been established to ensure that the Litter Intelligence Citizen Science programme consistently produces high-quality, credible and scientifically rigorous data.

Our **Quality Assurance** measures are proactive, and include the systems and processes we have built into the Litter Intelligence programme, training and technology that aim to prevent and minimise errors, and ensure data quality.

Our **Quality Controls** are reactive and corrective processes that we have put in place to identify and resolve any data entry/user issues or errors, to ensure the data that appears on the Litter Intelligence platform is robust and can be trusted.

These measures are important for providing ongoing confidence in data collected through this programme, and more broadly to instill confidence and trust in Citizen Science data for the widest audience possible including environmental reporting.

The *Litter Intelligence Data Governance Group* will peer review this document and -- pending changes and approval -- this document will be published on the Litter Intelligence website to add to the credibility and transparency of the programme and its data.

#### **Objectives:**

- Maintain a high standard of data rigor by reducing, identifying and resolving any data entry/user issues or errors.
- Ensure a clear and consistent process for all programme administrators / collaborators.
- Provide credibility and transparency for all QA measures and QC processes.

#### **Quality Assurance Measures:**

## 1. Training & Equipment:

- a. All data submitted to the LI platform that is tagged as “Official” has been submitted by Citizen Scientists who have received official training and used standardised equipment.
- b. Training is delivered only by experienced Sustainable Coastlines staff, or by ‘Trainers’ that have received appropriate training in the delivery of this workshop.
- c. Each training workshop follows a standardised format, meaning that each Citizen Scientist receives the same content, while still being flexible enough to cover a range of logistical circumstances.
- d. Each training workshop provides Citizen Scientists with training on how to use a standardised set of litter monitoring equipment. [See equipment list here.](#) Following the workshop, this equipment is provided to Citizen Scientists to use for each survey. Broken or lost equipment is replaced promptly.
- e. Only users who have attended official training workshops are provided with account logins (see QA measures this provides below).

## 2. Technology:

- a. QA measures across the technology:
  - i. Account login functionality provides ‘tagging’ of data as:
    1. **Official:** Submitted by trained Citizen Scientists using standard equipment (logged-in).
    2. **Ad Hoc:** Collected by community members where training & equipment can’t be verified (not logged-in).
  - ii. All critical data fields are required, meaning users cannot move to the next stage of data entry before entering the required information. They also cannot submit data to the database without completing all required fields.
  - iii. The data entry web app user interface is designed for ease of use to avoid errors in data entry. Design elements that provide QA include high contrast colours, large text fields, and responsive design that works on all mobile devices.
  - iv. ‘Tooltips’ provide helpful descriptions and reminders alongside data entry fields to help users enter the right data in the right place.
  - v. The web app functions offline (so long as the web page is loaded on the device prior to leaving wifi or mobile signal), meaning data can be collected and entered in remote areas without mobile reception.
  - vi. The web app caches (temporarily stores) the data on the mobile device, meaning if the user accidentally exits the web page, runs out of battery, turns their phone off, or in some other way is disrupted during the data entry process, the data are saved and can be returned to and submitted later.

- b. QA measures at specific steps in the data entry process:
  - a. Survey Setup (Steps 1 & 2)
    - i. When logged-in to an 'Official' account, some fields are automatically pre-populated, including *Survey Date, Monitoring Group, Citizen Scientist First Name, Last Name, Email and Phone Number*. The *Survey Areas* that can be selected are also limited to those that are assigned to the user. This reduces data entry errors and enhances the user experience.
  - b. Survey Setup (Step 3):
    - i. Users cannot proceed past the selection of an existing Survey Area, nor create a new Survey Area, before accepting the Health & Safety Requirements. This is not recorded in the technology however the user cannot proceed without doing so. This verification action aims to remind users to complete two key H&S tasks: (1) complete a Site Risk Assessment and (2) deliver a Health & Safety Briefing to all participants.
    - ii. Three *Survey Area Photos* must be uploaded by the user, before the user can progress past this page. This presence check ensures that three photos are taken and uploaded by the user. **NOTE:** *While the technology cannot check whether these are the appropriate photos, the photos are checked during the manual verification process via the admin portal.*
    - iii. Values over 10 metres cannot be entered in the *Above High Tide* and *Below High Tide* fields. This range check ensures that users enter only whole numbers 0-10 inclusive, in order to maintain consistency with the methodology.
    - iv. Must identify substrate type to proceed.
    - v. Must enter visual assessment grade to proceed.
  - c. Survey Info:
    - i. End Time must be after Start Time.
  - d. Audit Info:
    - i. Audit Date must be on or after Survey Date.
    - ii. Start Time must be after End Time of Survey Info.
    - iii. End Time must be after Start Time of Audit Info.
  - e. Audit Data:
    - i. Must enter whole numbers for count & weight.
    - ii. Must enter High or Low Confidence.
  - f. Review & Complete
    - i. Must review data before submitting.

## Quality Controls

1. User submits survey, results link automatically emailed to the full Sustainable Coastlines Litter Intelligence (SC LI) operations team.
2. For all submitted surveys: The 'Account Manager' for that Survey Area emails / calls the Lead Citizen Scientist to check the survey results. For 'Ad Hoc' surveys, the designated **Quality Controller** SC team member (person TBD) emails / calls the Lead Citizen Scientist to check results. **NOTE:** *Because the data entry link is available on the Litter Intelligence website, we may not have had any previous contact with those submitting data.* The key points that are checked are:
  - a. Survey Area on map displays in the right place (especially important for the first survey at that area).
  - b. Survey Area length and Survey Width (Above and Below High Tide fields). Survey Length x Survey Width is used to calculate Survey Area, which is used to calculate Litter Density. These are the most common places for data entry error and therefore have a significant influence on our Litter Density figures).
  - c. 'Other' fields under each material category (see point below)
3. If there are any changes, the SC Account Manager records these via the [Survey Verification form here](#). This record allows us to improve our keywords and categories and increase data quality by avoiding items being added into 'Other' categories.
  - a. If any 'Other' field items are identified, the Cit Sci tells us what product/items went into this classification.
  - b. SC Account Managers periodically review these items and use them to make recommended changes to keywords and/or categories, which are reviewed by the Data Governance Group.
  - c. SC Account Managers make any key changes to the data via the administration portal: [admin.litterintelligence.org](http://admin.litterintelligence.org).
4. Once changes are made, SC **Quality Controller** ticks 'Verified' survey.
5. Quality Control process is completed.
6. For 10% of all Survey Areas, an additional Quality Control process is required. This is detailed below.

### **Quality Controls: Process for Re-Surveys and Re-Audits:**

A total of 10% of survey areas sampled should be reassessed, ideally at sites SC staff are present at or when train the trainer can assist. Continue this pilot for 3 months and then review. There are two aspects:

1. Survey Completeness
2. Audit Confirmation

This is not about verifying a single person's quality of work. Instead, it is about understanding the limitations of the methodology as a whole. This work won't do anything to the data - it will provide an error rate. Survey completeness and audit confirmation data will be made available on the LI website.

## 1. Survey Completeness

**Key aim:** Understand and describe the survey error rate.

**Process:** Determine if any litter has been missed. And if so, how many items and much did it weigh. Complete resurveying immediately following the survey at 10% of our Survey Areas in the next 12 months.

1. SC staff to be in attendance initially, with trainers down the line. Introduce that re-survey will be conducted after initial survey - but mention they should survey as normal. And will do a quick re-survey after.
2. Once the original survey has been completed; leave pegs and tape in place, seal/confine the sack of litter and have a quick break.
3. Then as soon as possible, every surveyor is to re-survey the same Survey Area, conducting a minimum of two sweeps and placing found items in a separate sack. Sieve out under 5mm pieces before taking the final count of re-survey items. Any new litter found in re-survey to be counted and weighed (total, not categorised). This total gives an estimate of survey error rate.
4. This count and weight is then recorded in the app (see tech development below).
5. Add re-survey items to the original survey items and audit as per usual process.

**Result/calculation:** The error rate of items is  $\frac{|\text{re-survey value} - \text{original survey value}|}{\text{re-survey value}} \times 100\%$ . For both total item count and total weight.

**Tech consideration:** Add a step into data entry app, at start of the 'Audit Data' section. Allow the user to select 'Re-survey completed' and then add the total item count and weight of items found.

## 2. Audit Confirmation

**Key aim:** Understand and describe the audit error rate (in count, weight).

**Process:** Run "Count Confirmation" on the same 10% of audits at the Survey Areas we are running 'Survey Completeness' studies at.

1. Once original audit item information has been entered into the app, the litter is to stay in the container.

2. Once the data has been entered, under Sustainable Coastlines staff supervision, the litter is to be reaudited using a google sheets form or paper data sheet.
3. The re-audit team will check 3 areas on each sub-category: count & weight, and categorisation at a material class level.
4. The data of the original and re-audit should then be compared in an excel spreadsheet, with formulas designed to give specific errors.

**Result/calculation:** We recommend that categorisation, at a material class level, is assessed separately from count & weight. Review of top 3 contributors to the error rate.

**Tech consideration:** Don't replace original data with the reassessed data. Ultimately in the long run we should move to automate and update error coefficients near real-time via image recognition.

**Considerations:** We suggest that we re-audit all items within the survey immediately.

**General Considerations:**

Choosing sample sites:

- Must be done when SC is at the event, and easier to do ones at Wellington/Auckland
- Will be based on operational, health & safety & logistical considerations.