# Observer Report

# ESS 2 Workshop Cairns

# February 2019

Joe El-Hayek

Data Analyst

Plant Health Surveillance

Plant Health Policy Branch | Biosecurity Plant Division

**Summary Table of Useful Reports and Spatial products**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **High Level report Function** | Report Name | Technical Requirement | Users | Measureable benchmark/s | Objective |
| Operational | 1. [Survey Level Data Quality Report](#_Survey_Data_quality) | Brief outline:   * Individual field collector data quality audit reports prior to uploading to the ESS-SIMS database * Opportunity to correct errors, omissions and flag complete cells with no associated values | IPSHP, Animal Health, NBS, Vector | 1. Data Quality 2. Data Completeness 3. Data Accuracy 4. Surveyor level performance 5. Survey level performance 6. Program level performance 7. Department overall performance | Meeting ANAO Recommendation 1:   * 2.23 “…establishing nationally consistent and efficient business processes…” * 2.26 “…implementing surveillance analytics capability…” * Quantitative Performance benchmarking of data quality |
| Operational | 1. [Survey Effort Report](#_Survey_Effort_Report) | Brief outline:   * Internal app or external database calculation of distance over time * Count of specimens over time with future trend predictions in rate of growth | IPSHP, Animal Health, NBS, Vector | 1. Surveyor level efforts 2. Total Survey level efforts 3. Program level efforts 4. Departmental overall effort | Meeting ANAO Recommendation 3:   * 4.16 “…develop a relevant, reliable and complete framework of measure to assess its performance…” * Quantitative Performance benchmarking of survey efforts for resources expended |
| Strategic/Scientific | 1. Pest Free Area supporting maps and table report 2. Pest Status Determination support maps and table report | * Reliability/confidence index against individual records * Refer ISPM 8 for potential reliability evaluation method * Relevant spatial/GIS layers integrated with survey apps and/or SIMS database | Department, Industry, International trading partners | 1. Visual confirmation of pest distribution absence and presence data | * Timeliness in reporting findings and outcomes of survey efforts * Assist in maintaining export market access |

### Background

I had the opportunity to attend in an observer capacity, the ESS 2 Workshop Cairns from the 13th to the 14th February 2019.

In the Stage 4 – Reporting and Analytics discussion I was given the opportunity to provide some commentary and input into potential requirements for report outputs.

The purpose of this document is to clearly outline in detail the concepts that I had envisioned for useful reports that could be extracted from the database system in its current form.

In guiding my thought processes around these reports, I referred to the recommendations made in the ANAO Northern Australia Quarantine Strategy – Follow-on Auditor-General Report No.23 2018-19 specifically:

* Paragraph 2.20 “Data entry activities were not adequately quality assured, and capacity and functionality constraints of the NAQS Database limited the storage and use of survey data and diagnostics results.”
* Recommendation No.3, 4.16 “The department develop a relevant, reliable and complete framework of measures to assess its performance in managing biosecurity risk in northern Australia.”

### Survey Data Quality Report

I have trialled a data entry task with the current EFSA app. The report suggested below should also be possible with the AVOKA app.

* As a Field Collector I would find it useful to be able to generate a data quality report prior to submitting the survey data to the ESS-SIMS database.

This would allow me to undertake a quick overview audit of the data that I have captured.

If for example I had collected 10 records and the structure of the SIMS database required 10 fields to be completed, then there would be a total of 100 cells that potentially had to have data in them (these cells could be mandatory or optional).

As an example of how this could be implemented technically, it would be useful to have a pop-up window that could show me as a percentage, the number of “empty” cells in comparison to the total number of cells.

This would follow with the opportunity to rectify these in field by giving me the opportunity to either:

1. Link back to the empty cell and complete it with the correct value
2. Or simply select all empty cells and apply a “Nil Record” value for text fields or a “0” value for numerical fields

This information could be used in the publication of my final report

* As an Operations or Policy Manager, I may want to aggregate the individual field collector’s survey quality reports into an summary report at the discipline or specific survey level.

This information would assist me in analysing overall data capture performance. If data is being captured inconsistently or inaccurately, it might enable me to identify issues and apply remediation strategies to address them.

Some remediation strategies might include:

* + Individual performance reviews
  + A review of the survey protocols and methods
  + Further user training workshops
* As a Director, it would be useful to further aggregate these reports into an annual summary document with trend analysis looking at the last 5 years of data quality measurements and highlight the hopefully continuing improvement in the efficacy of the individual program’s and department’s data quality improvement initiatives

These benchmarks would be used as evidence based metrics in support of reports on the holistic performance of the ESS system as an integrated, authoritative, point of truth departmental enterprise database.

### Benefits and Outcomes

The reasoning behind this report and some of the functional capabilities outlined earlier, is similar to the approach of placing a “This page is intentionally blank” comment to a document to advise the reader that there is no missing information.

At a minimum this could:

* Improve the overall integrity of the dataset by resolving the issue of ambiguity and uncertainty around dataset completeness
* The benefit would be reduced post data capture cleansing efforts and a reduction in expenditure of departmental resources on these remediation activities

Both components could be readily measured, benchmarked and reported on in meeting any departmental continuous improvement objectives such as ANAO Recommendation 1:

* Paragraph 2.23 “…establishing nationally consistent and efficient business processes…”
* Paragraph 2.26 “…implementing surveillance analytics capability…”
* Quantitative Performance benchmarking of data quality

### Survey Effort Report

Some technical approaches to implementing survey effort metrics to produce a report could be to:

1. Calculate a track file’s distance over time to derive an average km/hr rate.
   1. This can then be compared to an expected performance level set at the policy level, for example in a survey protocol document
   2. I successfully implemented this method to benchmark survey performance of field collectors in the Queensland Department of Agriculture and Fisheries – Biosecurity unit’s Panama Program
   3. This data was used in a multitude of ways, for example:
      1. As evidence to respond to public queries regarding misappropriate use of vehicles when conducting foot based surveys
      2. As evidence that survey protocols developed by the policy team were being followed correctly and data collection objectives were being met
2. Number of target list specimens being collected per survey which could feed into:
   1. Trend line analysis for count of specimens collected annually and over several years historically to predict future trends and feed into decision making processes around future surveillance activity efforts and resourcing requirements
3. Develop survey program metrics to aggregate total distance traversed and possibly hectares covered (derived from a site’s polygon parcel size)
   1. I calculated similar statistics which were subsequently tabled in the Queensland parliament within a Question on Notice to inform ministers of program progress and containment effort successes.
      1. “To date, the Panama TR4 Program has surveyed over 70% of the planted banana industry area in Far North Queensland, walked over 15,600 kilometres conducting surveillance, and collected over 1400 samples for diagnostic testing. The disease is still known only to occur on one property in the Tully Valley.”
   2. <https://www.parliament.qld.gov.au/documents/TableOffice/questionsAnswers/2016/643-2016.pdf>
   3. This might then be linked to costs expended per survey and this might show an efficiency gain as we collect more data, surveyed over larger areas and greater distances for lower input costs.

**Limitations**

For track files:

* This might require some cleansing of the tracks to remove sections driven on road where no surveillance activity took place
* This might be remediated for an automated process by developing protocols and work instructions around collecting track data only when “on site” and not when in transit between sites

For specimen metrics:

* Out of scope for this phase is a desirable link to either BIISA analytical outputs and or directly layering international trade data over the number of specimens collected to provide evidence based analysis in support of correlating future workload trends with existing business resourcing limitations and future resource budgeting

### Benefits and Outcomes

These types of metrics will assist the department in meeting ANAO Recommendation 3:

* Paragraph 4.16 “…develop a relevant, reliable and complete framework of measure to assess its performance…”
* Quantitative Performance benchmarking of survey efforts for resources expended

### Future Opportunities

If at ESS phase 3 of the project, it is decided that either a merger of the existing dual system AVOKA/EFSA apps or an entirely new concept app is proposed, there would be significant benefit to developing a user interface that enforces a chronological data collection method to address some of the current shortfalls of the EFSA app in order to:

* Improve data quality and reliability
* Reduce free text requirements for routine data record entry functions and instead auto-populate the majority of the data fields from a tap of a graphical icon
* Improve in-field data collection efficiency by reducing time to capture a single record, resulting in more scientific records obtained for the same survey effort and expenditure

### Conclusion

Overall the workshop was very productive and the active level of engagement by individuals around the issues of data quality and integrity was a positive development in this process.

With the addition of some useful reporting and performance benchmarking metrics, the ESS and associated apps in their current form can still value-add at an enterprise level to the departments analytical and decision making capabilities.