

NBN Conference Workshop Report

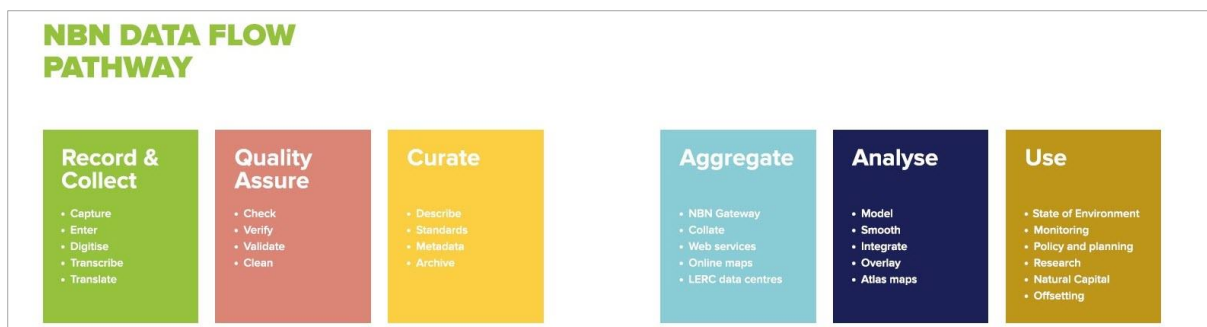
Data Flow (Workshop 3) – Dan Lear & Rachel Stroud

Workshop aim:

To develop an assessment criteria to analyse the current data flows within the Network, to identify where the current blockages are occurring and which stages of the data flow pathway are working efficiently.

This workshop investigated the current data flows and explored the areas that currently work and those that need improving. We went on to discuss what an optimal data flow pathway would look like and how we would measure success.

The groups focussed on species data flows, and produced ‘flow diagrams’ of how they perceive the current flow of data as it moves from generators, through collation and custodianship to publication. We did not consider the ‘Analysis’ and ‘Use’ stages of the current pathway.



One common viewpoint was that new technology is exposing issues and challenges that have always existed but are being magnified. We have made significant progress in streamlining data flow, but further work on standardisation is needed.

The community is still dealing with legacy databases, spreadsheets, formats and processes whilst experiencing a proliferation of new apps, societies, FB groups, etc. This is causing ‘pressure’ on both sides and is coupled with concern over verification ‘black holes’. Applications like iSpot and iRecord make the capture of biodiversity data easier but place increased demands on the verifiers with those records from ‘untrusted’ data providers being of a lower priority and therefore never (or very slowly) verified and released.

A culture of possessiveness still exists, with the (potentially) misguided view that by holding onto data, people are in a stronger position to leverage funding, despite many funders mandating the release of data funded with ‘public’ money.

The groups were asked to consider if we could get a better understanding of data flow by breaking it down by the communities, roles and behaviours at the local, regional and national level, however no group successfully addressed this approach.

What does success look like?

The flow of data should be streamlined *but* tailored to individual users/communities needs and there is a need to remove 'unnecessary' steps and automate where possible. This can be achieved by the adoption/adaption of existing standards for data exchange. These standards would also facilitate the next metric of success; the development of reusable, standards-based tools which can be 'skinned' for different users/communities.

How do we get there?

The common response was that more resources were needed. There is a clear need to 'free' the passionate and expert user community from the necessary but time consuming process of validation and verification. We should also make better use of unverified records with the appropriate flags/caveats to ensure end-users are aware of the provenance of all the data.

It was suggested that we need a step-back. Take the opportunity to review the current landscape and standards and assess their applicability and adopt/adapt where necessary.