

Powering up Workflows with Volunteers and Division of Labour

Paul Flemons,
iDigBio Dried Insect Workshop
April 2013

nature culture **discover**



AM Workflow

- Volunteers – onsite and online
 - image capture, data entry, georeferencing, transcription and validation
- Paid staff – image and data quality, data import
- Specialisation
- Division of Labour

Division of Labour

- Each worker or volunteer can focus on performing one or a few tasks very well
 - benefit of this is both increase in productivity and increase in quality
- Volunteers can choose tasks or workers be allocated to them based on their skills, interests and experience
- Enables flexibility in allocation of tasks both geographically and temporally

Digitisation – complex process

Tasks in the process

1. Selection of taxa
2. Preparation for digitising – eg curation
3. Imaging
4. Data entry
5. Transcription and Interpretation of handwriting
6. Georeferencing
7. Validation of data
8. Selection of correct taxon names
9. Import into database

Typical image of specimen and labels



Why use image based digitisation?

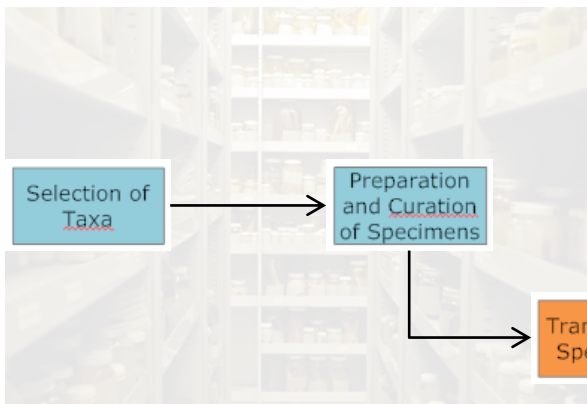
- A readily accessible digital voucher of each specimen and its labels for verification and reference
- Reduced need for specimen handling
- A virtual specimen in the event of collection loss or damage
- Remote access to original label data for review by researchers
- Limited potential for species identification from an image
- Enabling the use of ‘non-experts’ in data entry with the benefit of knowing data quality and enabling dubious data can be checked without having to physically visit a specimen in the collection.
- Decoupling the stages of digitising so that division of labour benefits can be realised.

Workflow

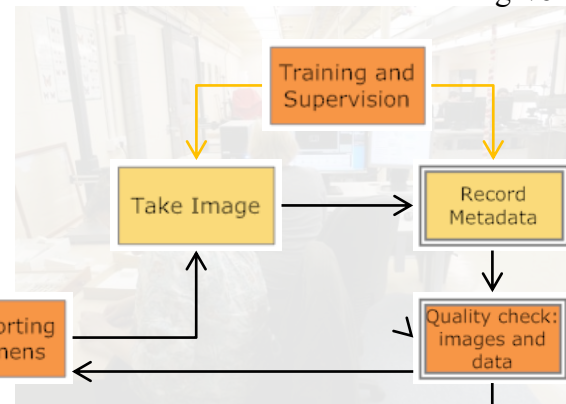


Location

Collection



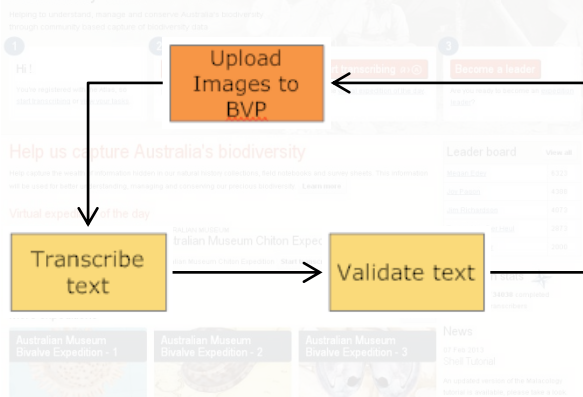
DigiVol Lab



Resource Legend:

- Collection Staff
- Database Specialists
- Volunteer Coordinators
- Volunteers
- MS Access
- KE EMu

Biodiversity Volunteer Portal



Online

Match or create names

Import images and data into EMu

Partial Record

Data cleaning and normalisation

Import into EMu

Full Record

Database

Benefits of this Division of Labour

Benefits

- Standardised consistent process
- Distributed data capture but centralised data import – more control – better data quality
- Use of controlled vocabularies can be used to improve problems associated with misspellings etc
- Speed and accuracy in each step - more records entered and entered more accurately

Issues of using volunteers

Issues

- Volunteers are not free – they require management
- Image capture and metadata entry
- Standardisation of transcription quality
 - Use of picklists
 - Postprocessing using Google Refine
- Recruitment of crowdsourcing volunteers

Links



<http://australianmuseum.net.au/Rapid-Digitisation-Project>

<http://volunteer.ala.org.au/>

<http://www.ala.org.au/>

<https://www.idigbio.org/content/idigbio-public-participation-digitization-biodiversity-specimens-workshop-emu%E2%80%99s-cloud>

<http://www.pensoft.net/journals/zookeys/article/3146/image-based-digitisation-of-entomology-collections-leveraging-volunteers-to-increase-digitization-capacity>

Thank you

www.australianmuseum.net.au

nature culture **discover**

