Conference Abstract

How Artificial Intelligence and Machine Learning can Assist in Collections Curation

Simon Checksfield ‡

‡ CSIRO, National Research Collections Australia (NRCA), Canberra, Australia

Abstract

With increasing pressure on the limited taxonomical expertise in not only Commonwealth Scientific and Industry Research Organisation (CSIRO) but the world, new and innovative ways need to be found to assist in the curation and identification of biological specimens.

CSIRO, through the National Research Collections Australia (NRCA) and Data 61 is hoping to begin a new program of work focused on using Artificial Intelligence (AI) and Machine Learning to build a framework and tools that can help identify a specimen from an image. The framework will include AI models that have been trained by expert taxonomists, thus providing a level of accuracy that has some intrinsic value.

NRCA is also exploring how AI could be linked or cross referenced with another initiative using rapid genetic barcoding to identify all newly collected specimens. Combining genetic and AI determinations will add weight to each, and potentially expose some new AI challenges, such as identifying morphological elements against genomic elements.

Whilst acknowledging challenges still exist regarding standards, acceptance of identification, provenance, accuracy and governance, the NRCA is hoping AI can assist in freeing the time of our researchers and technicians to work on more pressing and complex issues by reducing their time spent on basic identification.

The impact of such a program will also reach into industry and the general public through tools based on the AI models. There is also an opportunity to use this initiative to create global centers of taxonomic expertise, which anyone can use to help identify a specimen.

Keywords

curation, artificial intelligence, identification, provenance, collections, specimen, machine learning, genomics, morphology

Presenting author

Simon Checksfield

Presented at

Biodiversity_Next 2019