Data science shown to expedite return of ancestral Indigenous remains

February 16 2024

Professor Richi Nayak and Dr Md Abul Bashar. Credit: Anthony Weate/Queensland University of Technology

The return of the ancestral human remains of Australian and other Indigenous peoples held in anthropological collections could be sped up using machine-based deep learning according to a new study led by QUT computer scientists.

In [a paper](#) published in *Social Science Computer Review* titled "An Informed Neural Network for Discovering Historical Documentation Assisting the Repatriation of Indigenous Ancestral Human Remains," a research team including Dr. Md Abul Bashar and Professor Richi Nayak, both from QUT's School of Computer Science, outlines how.

"The repatriation of the bodily remains of the ancestors of Australian and other First Nations peoples from mostly Western scientific institutions depends largely on locating information within scientific and other literature published between 1790 and 1970 documenting their theft, donation, sale, or exchange between institutions," Dr. Bashar said.

"Digital library initiatives by national and major research libraries since the mid-1990s have made it possible to systematically search large historical corpora of books, journals, and newspapers. However, finding relevant information within these corpora is difficult."

Supported by the Australian Research Council, Dr. Bashar and Professor Nayak worked with academics from the Australian National University and University of Tasmania. In collaboration with data scientists and social science researchers from the Research, Reconcile, Renew Network (RRR), who have developed and applied text mining techniques, the team have created a machine learning-based solution to automate the process of finding and semantically analyzing relevant texts.

"The RRR researchers face a huge challenge as they work to assist First Nation communities in repatriating their ancestors' remains from Western scientific collections," Professor Nayak said.

"They endeavor to confirm the identity, current location, and other information necessary for a successful reburial. To do this, RRR researchers invest substantial time and intellectual energy in tracing the

movement of the remains after they were stolen through an as yet only partially identified networks of collectors, donors, private sellers, commercial dealers, and scientists associated with museums and other institutions with anthropological collections.

"Current museum catalogues provide some data for analysis but relying purely on the sum of information within them runs the risk of making potentially devastating mistakes."

Dr. Bashar said text mining was a research field that has made significant progress in developing techniques for uncovering and extracting knowledge from large collections of documents.

"Text mining has mostly been applied to pre-existing textual corpora. In the humanities and social sciences, though, there are many pressing research questions related to past human thought and behavior for which relevant collections of historical documents do not exist in machine-readable form," Dr. Bashar said.

"Instead, researchers must use search engines to locate, gather, and analyze documentary content about past people, places, and events. The challenge here is that search engine results can be biased by algorithms and indexing practices and may not accurately reflect all relevant content.

"For RRR researchers, the usefulness of keyword-based search services provided by the creators of corpora really depends on how queries are formed. If the search terms are too specific, few or no relevant documents will be returned. If terms are generic, large numbers of irrelevant documents will be returned."

Professor Nayak said classification models, particularly deep learning-based models, can have low accuracy when trained with small amounts

of labeled (i.e., relevant/non-relevant) documents.

"To improve the accuracy of our detection model, we explored the use of an Informed Neural Network (INN) model that describes documentary content using expert-informed contextual knowledge," Professor Nayak said.

"A set of keywords was provided by RRR researchers, chosen for their highly probability of appearing in documents of interest.

"This provided the basis for a deep learning-based classification model trained with a small portion of labeled documents.

"The results of our search confirm the value of using an INN network model for accurately and quickly identifying relevant documents related to the investigation of the global commercial trade in Indigenous human remains."

It was developed it in 2014 to raise awareness and understanding about repatriation and assist repatriation practitioners and researchers in their efforts to bring Old People home.

Dr. Bashar and Professor Nayak, along with co-authors Professor Cressida Fforde (ANU), Dr. Gareth Knapman (ANU) and Professor Paul Turnbull (University of Tasmania) are all members of projects team with the Research, Reconcile, Renew Network. It was launched in 2014 to raise awareness and understanding about repatriation and assist repatriation practitioners and researchers in their efforts to bring Old People home.

Dr. Knapman is a Research Fellow on the ARC Discovery Project "Profit and Loss, The commercial trade in Indigenous human remains" at the Center for Heritage and Museum Studies at the Australian

National University.

Professor Fforde's primary area of research focus has been the history of the removal and subsequent study of Indigenous human remains by European institutions within the "race" paradigm in the nineteenth century and the campaigns to secure their return.

Professor Turnbull is a historian of anthropology and the biomedical sciences in colonial contexts. He is internationally known for his research on the theft and scientific uses of Australian and other Indigenous peoples' bodily remains.