THE BONE TRADE
STUDYING THE ONLINE TRADE IN HUMAN REMAINS WITH MACHINE LEARNING AND NEURAL NETWORKS

Shawn Graham, Carleton University & **Damien Huffer**, Stockholm University

There is a thriving online trade in anatomical, ethnographic and archaeological human remains that makes ready use of new social media such as Instagram, Facebook, Etsy, and until recently, eBay. The "fetishization" of the 'exotic' dead that underpins this trade by its very nature transforms pieces of the body into material culture: curios, commodities or objets d'art. This practice has deep Colonial-era roots, but today's e-commerce and social media platforms have only expanded collectors' reach and made participation open to anyone with interest and spare finances. The sheer volume of materials being produced, shared, and sold can be overwhelming for a small team to study. The market moves so fast.

Can we teach machines to identify from photographs alone patterns in the 'visual rhetoric' that signal materials for sale? Can 'licit' materials be discerned from 'illicit'? Are there geographical patterns? Can we trace materials back to a source?

We began mapping out the territory of this trade in 2016. We collected thousands of posts and studied the language of the posts - how the collectors and enthusiasts described their engagement with the remains. At that time, we studied only one platform. Our methods were primarily textual.[1]

In this project, we intend to explore the leads suggested in that first study by developing and adapting approaches from machine learning, computer vision, and artificial intelligence (various neural network models) to scale up our ability to study this trade. We are looking at a number of social media platforms and marketplaces.

Building on our previous research, can we marry these insights from machine learning and computer vision, to those generated from text analysis of the posts, and social network analysis of followers? How do particular patterns of display move over the network of participants - are there fads, trends, key players? Finally, what are the ethical, moral, and legal implications of using machine learning in this way?

---

[1] Huffer, D. and Graham, S. 2017 The Insta-Dead: the rhetoric of the human remains trade on Instagram, Internet Archaeology 45. https://doi.org/10.11141/ia.45.5

Our objectives are therefore:

1. To develop and share a trained neural network that can be employed by other researchers interested in this trade in particular;
2. To develop the computational and theoretical tools to allow others to adapt our approach to their own area of interest in humanities' research;
3. To determine the patterns in the visual rhetorics of the trade in human remains online so that this trade can be tracked across social media, monitored, and disrupted;
4. To enable the possibility of sourcing these materials so that they may be repatriated to descendent communities;
5. To build ethical frameworks into our computational approaches;
6. To develop a cohort of highly trained personnel who will take this research forward into other domains.

Schedule

Years 1 and 2:

- Expand and develop a training dataset from the existing corpus of Instagram material, which currently consists of ca. 15 000 photographs, posts, and metadata collected for Huffer and Graham 2017
- Identify salient markers upon which to create a supervised learning model
- Identify appropriate deep-learning unsupervised approaches to create an unsupervised model
- Use Google's Inception v3 trained model to identify clusters based on shared image features and compare these results with our own models
- Evaluate these models against a subset of the original data held in reserve
- Move forward with the appropriate model(s) (objective 1)

Continue to collect posts and metadata over the duration of the project from the targeted social media platforms Instagram and Facebook to keep the research timely (objective 1)

Situate our work in broader context of research into the antiquities trade (objectives 3 & 5)

Years 3 - 5

- Develop and train NN targeting different facets of the material: licit versus illicit, sentiments, provenance, demographics to develop a holistic picture (objectives 1, 2, & 3)
- Develop and release a body of code (under version control) in e.g., Tensorflow and the R statistical computing language, for reproducibility and replicability to other domains of

archaeological or cultural materials. Tensorflow is currently the state-of-the-art in exploring NN; the R language and its associated ecosystem of publication workflows has become a standard for digital humanities work. (objectives 1 & 2)

- Social media platforms evolve quickly. We must continually explore the implications of the changing social media ecosystem, its 'terms of service', and evolving thought on the ethics of such research (objective 5)

Years 4 - 5

- Push the data and the models further: can we identify likely descendent communities o Associated archival and historical research to support this task (objective 4)
- Identify other lines of evidence that support the case for intervention from a legal standpoint (objectives 3, 4, & 5)
- Identify the migration of materials across social media platforms by tracking visually similar images (objectives 3, 4, & 5)
- Develop public-facing tutorials that ethically communicate the results of this research to raise awareness with possible descendent communities, law enforcement, academic and professional audiences, and to advocate for policy changes (objectives 3, 4, & 5)

As students come on board, and new technologies or platforms come online, this schedule may have to be revisited.

Outcomes

We expect to learn

- the strengths and limitations of neural networks as a lens for studying the trade in human remains - a much finer-grained picture of the scope and size of the trade in human remains
- the identification of flows of visual signifiers of what is for sale, where it comes from, and
- where the tastes in this market are set
- the identification of institutions, collections, cultures, or conflicts from which human skeletal remains actively being circulated are sourced
- how to deploy neural networks in the service of humanistic and social science inquiry

We expect to provide

- high quality training to exceptional students who will use this training to launch their careers
- anonymized and ethical exposure of the trade to both the public and to policy makers

As outputs we will have created

- several articles in scholarly and popular journals
- fully cleaned datasets for publication and reuse
- tutorials for use outside our institutions
- seven highly trained personnel
- new course opportunities for undergraduates
- training and research relationships between Carleton University and the University of Stockholm
- a new direction for the study of the trade in illicit and illegal antiquities more generally

At the conclusion of this proposed project, we believe that we will

- be in a position to begin another phase of this research where we will take our results and our method, where digital cultural heritage ethics are foregrounded, to descendent communities at risk, beginning a project to provide training and support in computational methods
- be in a position to influence change at social media companies that currently permit the trade in human remains
- have raised the public consciousness of this trade, including with law-enforcement agencies, policy makers, and the corporations whose platforms currently facilitate the trade

Graduate Training

Graduate students in History/Digital Humanities and History/Data Science will be trained in reproducible computational approaches and open science, as well as the necessary grounding in the antiquities trade, and the human remains subtrade. Digital training will include everything from version control to writing analytical packages, to interactive website design for tutorial-writing and outreach. They will learn state of the art machine learning and neural network approaches with social media data.

Students will be involved in the outset in planning our research design; they will have opportunities to publish as lead authors (especially with regard to mobilizing knowledge through venues such as The Programming Historian and conferences such as DHSITES). MA students will have the opportunity to work in a leadership role with students in Graham's undergraduate digital history research methods courses (HIST3812 and HIST3814); the PhD student will have the opportunity to design and teach undergraduate courses on the intersection of data ethics and

history/archaeology. The graduate students will be encouraged and supported to hold 'unconferences' (see thatcamp.org) on a theme of their choosing; an excellent opportunity exists for international participation and recognition of their work should they choose to hold one (as we will encourage them to do) during the DH2020 Conference in Ottawa.

Their research may include:

- Writing code to generate datasets
- Developing various NN
- Analyzing results
- Ground-truthing training datasets (making sure that training images are properly classified)
- Curating and preparing materials for data publication in appropriate venues
- Research and writing of tutorials
- Research and writing connected with their own research interests as they intersect with this project
- Communicating the results of research with relevant publics at conferences and other venues

Funding