11-16-2018

Thinking Computationally about Forensics: Anthropological Perspectives on Advancements in Technologies, Data, and Algorithms

Bridget F.B. Algee-Hewitt
Stanford University, bridgeta@stanford.edu

Jieun Kim
Florida State University, eun3022@gmail.com

Cris E. Hughes
University of Illinois at Urbana-Champaign, hughesc@uiuc.edu

Introduction: Thinking Computationally about Forensics: Anthropological Perspectives on Advancements in Technologies, Data and Algorithms

Bridget F. B. Algee-Hewitt,[1] Jieun Kim,[2]* and Cris E. Hughes[3]

[1]Center for Comparative Studies in Race & Ethnicity, Stanford University, Stanford, California, USA.

[2]Department of Scientific Computing, Florida State University, Tallahassee, Florida, USA.

[3]Department of Anthropology, University of Illinois at Urbana–Champaign, Champaign, Illinois, USA.

*Correspondence to: Jieun Kim, Department of Scientific Computing, Florida State University, 400 Dirac Science Library, Tallahassee, FL 32306-4120 USA. Email: jkim17@fsu.edu.

Short Title: Thinking Computationally about Forensics

KEY WORDS: FORENSIC ANTHROPOLOGY, QUANTITATIVE METHODS, MEDICO-LEGAL CASEWORK, HUMAN IDENTIFICATION, SKELETAL ANALYSIS, GENETIC ANALYSIS, REMAINS RECOVERY, DIVERSE POPULATIONS.

While forensic anthropology is often characterized as an applied science, it is deeply rooted in the larger discipline of biological anthropology. As forensic practitioners, we work to extend the theory of, and methods for, the study of human variation to the medico-legal context. We continue, therefore, to address the fundamental questions of topics critical to biological anthropology, such as the degree and distribution of skeletal and genetic diversity and the effects of environment and life history on morphological expression, just as we seek to infer the demographic parameters of sex, age, and ancestry that allow us to broadly characterize modern peoples. In speaking for the single person, however, forensic anthropologists are uniquely challenged with the issue of scale. We must distill down our approaches (or methods) of biological anthropology for the detection and documentation of populational trends to the level of the individual forensic case as we reconstruct the biological profile and address the personal identity concerns that dominate the forensic anthropological analysis of unknown human remains. In concert with this change in scope, forensic anthropologists must also contend with a refocusing of perspective towards the investigative and judicial system. At once, we are expected to respond to the dynamic needs of individual identification, in the service of human rights, social justice, and the medico-legal community, the increasing demands for scientific rigor in case analysis and reporting, and the changing expectations for admissible evidence and expert testimony in the courtroom (Christensen and Crowder 2009; Grivas and Komar 2008; Lesciotto 2015; Steadman 2009; Steadman et al. 2006; Wiersema et al. 2009).

Not surprisingly, the field of forensic anthropology has evolved considerably over the course of the last half-century. The present state of the discipline is very different from when it was first admitted into the American Academy of Forensic Sciences as the ‘physical anthropology’ section in 1972, yet there are interesting and important parallel moments in its

developmental trajectory. A decade later, Snow (1982) wrote about the new expansion of *physical anthropology into the area of forensics* "at a time when many physical anthropologists are deeply concerned with the need to expand the scope of our field beyond its traditional boundaries […]. (p. 97)." Similarly, we are writing here about how *forensic anthropology is presently undergoing its own transformation*, as it expands its reach to adopt methods that would have been once outside the limits of the forensic anthropologist's traditional expertise. As a field, it stands, therefore, at the crossroads between diverse areas of study beyond skeletal variation– from computational biology, genetics, isotopes, and demography, to cultural anthropology (Algee-Hewitt et al. 2018; Hughes et al. 2017). This point is most obviously supported by the amendment of its AAFS section title to the more inclusive title of 'Anthropology' in 2013. This change was equally born out by the trends observed in publications specific to forensic anthropology. From a Web of Science analysis of the publications of the *Journal of Forensic Science*, when filtered for Anthropology as research area and constrained to the roughly, five-year period of 2013 to 2018, we find that the top ten cited papers (Christensen et al. 2014; Crowder et al. 2013; Edgar 2013; Hackman and Black 2013; Hauther et al. 2015; Hefner and Ousley 2014; Kim et al. 2013; Stephan 2014; Stephan et al. 2013; Tise et al. 2013) have titles that explicitly address the computational topics of advanced statistics, classification algorithms, and imaging methods and speak to the bio-social issues of racial and ancestral identity for peoples of multiple origins with complex population histories. Moreover, in these studies, we see more multicomponent frameworks that embrace the technological advances in complementary, computationally driven, fields, and a recognition of the potential value of computational approaches to the forensic anthropologist's applied work, whether in aid of law enforcement or

in large scale responses to humanitarian crises, human right violations or mass disasters. This movement in forensic anthropology represents swift and major progress forward.

This shift in our research priorities is especially important as it comes at the same time as the wider forensic community is undergoing a period of self-reflection and redefinition of expectations. With this has come many critiques of the rigor of the *science* of forensics, which have acted in several subfields, like fingerprint, bloodstain and bitemark analysis, as a motivating factor for change not only in protocols but also in the perception of the value of certain techniques for casework and the scientific basis for their justification as trial evidence (Cooper 2016; Expert Working Group on Human Factors in Latent Print Analysis 2012; Page et al. 2011a; Page et al. 2011b; Saks et al. 2016; Taylor et al. 2016). We recall how in 2009 the National Research Council (NRC) of the United States (Committee on Identifying the Needs of the Forensic Sciences Community and National Research Council 2009) published a report that admonished the state of the field, writing that "'[t]he bottom line is simple: [i]n a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions…."(p. 53) and how "[m]uch forensic evidence […] is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing…"(p. 107-108). This report also charged the forensic science community to make the fundamental changes necessary to drive its practitioners towards a true transformation in thinking, asking forensic scientists to make the improvements, both "systemic and scientific," that are essential to "…ensure the reliability of the disciplines, establish enforceable standards, and promote best practices and their consistent application" (p. xix). Its coincident release at the peak year for the number of exonerations by DNA in the United States (CNN 2013) has helped us to recognize the miscarriages of justice that

can result from the uncritical admission of unvalidated science, expert testimony or evidence (Garrett 2012). With progressive interest over time, likely owing to the widespread visibility brought to these issues by, for example, the Innocence Project, this NRC report has garnered more attention in very recent years: doubling its readership since 2014, accumulating 21,787 downloads as of April 2018, and generating an Altmetric score of 612, which serves as a measure of the amount of attention the report has received from both social media and the news (The National Academies Press 2018).

We can also look to the 2016 announcement by the Department of Justice (DOJ) on scientific certainty. Building on the momentum of the NRC report and on the recommendations of the National Commission on Forensic Science (NCFS), it formally rejects the phrase "to a reasonable degree of scientific certainty," which has long served in expert testimony as a signifier of factuality (National Commission on Forensic Science (NCFS) 2016). This mandate acknowledges what others in the legal system as well as in forensic anthropology have already noted (Kaye 2010; Steadman et al. 2006): it is a linguistic trope plagued by subjectivity, that may be erroneously equated with certainty at the level of beyond a "reasonable" doubt and that introduces confusion in the presence of actual probabilistic evidence. At the most foundational level, therefore, all forensic sciences need the transparent quantification of uncertainty through error testing and validation studies as well as approaches that reduce error and permit the probabilistic statements on the weight of evidence – these are expectations that can only be met by computational approaches.

Thinking computationally about forensics can significantly strengthen our forensic anthropological practice, yielding more reliable and accurate findings. Computational methods, in particular, offer several advantages to the study of forensic anthropological data. Through the

analysis of large quantities of informational data, they allow us, as researchers, to perform more comprehensive or deeper investigations, effectively overcoming the limitations of cognitive ability and building stronger scientific foundations for our applied casework techniques. For example, recent advancement in various innovative data capturing technologies enable us to collect and examine more nuanced and detailed information from forensic samples and scenes (Claes et al. 2014; Park et al. 2017; Perlin et al. 2015; Stoyanova et al. 2017; van Oorschot et al. 2010; Walsh et al. 2014), and the availability of free-access software and computational algorithms, through open source coding initiatives, allows for widespread access to more complex probabilistic methods (Boldsen et al. 2002; Kim et al. 2013; Konigsberg 2015; Konigsberg et al. 2016), while significantly reducing processing time for big – high dimensional, large sample – data. Most importantly, the growing number of inter-disciplinary collaborations and the high competition in the technology market have lowered the burden of accessibility to new technologies, both in terms of expense as well as practice, making approachable those new methods that were once off limits because of their sizeable learning-curve and need for specialized knowledge or degree of expertise. The ubiquity of technology that is equally relevant to science and society, therefore, pushes us forward towards embracing the new technological developments and adopting the new methodological approaches.

By probing anthropological case data in previously unavailable ways, computational tools also give us a means to reveal latent data trends, identify and explore novel questions, and establish inferential procedures that deliver more satisfying results. Thereby, we improve the quality of our aid to the medico-legal community. Finally, when computational systems are used to represent expert knowledge, they allow us to better capture, distill and interpret complex data (Algee-Hewitt 2016; Konigsberg et al. 2016), while also improving precision and accuracy,

reducing subjectivity, and facilitating the automation of traditional procedures – many of which have been shown to suffer from human, whether, intra- or inter-observer error (Kimmerle et al. 2008; Stoyanova et al. 2017). They permit, therefore, the analysis of data in a standardized way, that promotes objectivity and reproducibility, method- and/or self-assessment and the reporting of rates of error. Incorporating computational methods into the forensic anthropological toolkit helps us to overcome the common criticism that forensic evidence and courtroom testimony, in areas exclusive of, the gold standard, DNA analysis, lack a strong scientific basis (Steadman et al. 2006) To most effectively and responsibly profit from these computational resources, we as both researchers and practitioners must engage in focused discussions of their advantages and limitations. The "hybrid-intelligence of human and machines" (Franke and Srihari 2007) requires the development of best practices for their use, as we contend with the evolving issues of software compatibility and our data management, bioethical concerns over the new kinds of information that we now have or make accessible, and the question of the dissemination of results among our peers, to our students and future colleagues in the classroom, for the medico-legal community, and to the public who we seek to serve and educate.

Now is an exciting and critical time for the future of forensic anthropology. It is ripe for the field and the practitioner to embrace the challenges that increased rigor in case evaluation techniques and heightened expectations for evidence and testimony bring. As current educators, through our dissemination of research, mentorship in the lab and teaching in the classroom, it is also time for us to redefine our educational expectations that reflect the necessary integration of field and new method and, so, meet the needs of the next generation of practitioners. We are charged, therefore, to achieve a synthesis of anthropological and computational thinking. Such interdisciplinary educational opportunities will greatly benefit the field, just as interdisciplinary

research collaborations are already advancing our ability to adopt external approaches and integrate cutting-edge technologies in order to revisit prior assumptions and shed fresh light on forensic anthropological questions (Adserias-Garriga et al. 2017; Algee-Hewitt et al. 2016; Hughes et al. 2017; Villa et al. 2016).

By turning to what Steadman (2018) has called "computational anthropology," we can benefit at present from the many advantages that marrying technology with statistically driven approaches to forensic case analysis offers while also working to propel the theory, methods and practice of forensic anthropology forward. The papers included in the two special issues on "Thinking Computationally about Forensics" that we introduce here are the products of a symposium organized by Algee-Hewitt and Kim for the 87th annual meeting of the American Association of Physical Anthropologists in Austin, TX, in 2018. Its goal was to motivate discussions between those research-focused biological anthropologists, who engage in computational anthropology and those practitioners of forensic anthropology, whose applied work serves the medico-legal community. This symposium operated as a nexus for discussions about the interactions between theoretical contributions to forensic analysis and the practical value of these advancements in the actual casework setting. These special issues embody the spirit of this symposium, as we reflect on the ways in which advanced quantitative approaches to research in biological anthropology can be extended to questions of relevance to forensic anthropology. The contributors to these issues dive deeply into a variety of subjects, from the construction of the biological profile to investigative techniques for human remains recovery, yet they all exemplify the kind of novel work which stands at the point of intersection between theory, method and practice. These articles show the promise of improved methods for ancestry estimation from skeletal metrics and genetics alike, the future of individual identifiability from

linked genetic data, the bioethical concerns over privacy and profiling as technological advancements in predictive modeling move faster than regulatory bodies can sustain their oversight, the application of new imaging technologies for enhanced investigation, whether in the context of the facial reconstruction of human remains from CT scans or the detection of clandestine burials using unnamed aerial vehicles as remote sensing platforms for data acquisition, or the novel fully-computational shape-based estimation of age at death for understudied populations.

The commentaries that accompany each issue address the places we have been, where we stand now and where we must go in service of the field. They provide us with special insight into how computational approaches to forensics are viewed from the perspective of the expert in human identification practice and in scientific computing. Although their point of reference differs, the overarching message is remarkably similar. It is not enough for us to execute computational research in an intellectual vacuum, promising improvement to current methods in forensic casework. We must also deliver on that promise by producing tools – whether as new frameworks of analysis, standards and protocols, executable code, or software – that are conceptually and physically accessible to the wider community of forensic anthropologists. In doing so, we should facilitate, with greater ease, the adoption of the methods of "computational anthropology" by the average practitioner as well their extension by the advanced user, permit the straightforward evaluation of their statistical foundations and rates of error, and encourage their acceptance by the medico-legal system.

The purpose of these special issues is to provide a forum for us to foreground our interests in advancing computational research that has implications for the forensic anthropological sciences, and to enrich current procedures with the potential to change the course

of future human identification practice. Here, we bring together a mix of participants, who engage wide-ranging skeletal, genomic, phenotypic and meta-data analyses, to 1) introduce new algorithmic advances and methodological improvement, 2) present work on the application of computational techniques to understudied populations, novel datasets or new information types, and 3) speak to the challenges that the revolution in data technologies may pose for future scientific investigation as well as the broader social effects on issues of policy, privacy and lay interpretation. We believe that the work of the issues' contributing authors exemplifies successful attempts at making advancements in all steps of the forensic casework analysis process: from locating clandestine graves, to the molecular, chemical, and osteological analysis of the retrieved remains in the forensic laboratory, to the re-fleshing of forensic case to reveal the unique features of the person, and finally to the presentation of our evidence and the communication of our findings beyond the laboratory and courtroom into the wider discourse, here in Human Biology.

We hope that this journal's present commitment to forensic research represents the first of many, with future issues that will showcase the continued development of exciting new techniques, especially for unexplored data types and underserved populations, and support the dissemination of promising research in forensic anthropology to a larger audience.

Received 2 May 2018; accepted for publication 16 May 2018.

Literature Cited

Adserias-Garriga, J., N. M. Quijada, M. Hernandez et al. 2017. Dynamics of the oral microbiota as a tool to estimate time since death. *Mol. Oral Microbiol.* 32:511–516.

Algee-Hewitt, B. F. B. 2016. Population inference from contemporary American craniometrics. *Am. J. Phys. Anthropol.* 160:604–624.

Algee-Hewitt, B. F. B., M. D. Edge, J. Kim et al. 2016. Individual identifiability predicts population identifiability in forensic microsatellite markers. *Curr. Biol.* 26:935–942.

Algee-Hewitt, B. F. B., C. E. Hughes, and B. E. Anderson. 2018. Temporal, geographic and identification trends in craniometric estimates of ancestry for persons of Latin American origin. *Forensic Anthropology* 1:4–17.

Boldsen, J. L., G. R. Milner, L. W. Konigsberg et al. 2002. Transition analysis: A new method for estimating age from skeletons. In *Paleodemography: Age Distributions from Skeletal Samples*, R. D. Hoppa and J. W. Vaupel, eds. Cambridge: Cambridge University Press, 73–106.

Christensen, A. M, and C. M. Crowder. 2009. Evidentiary standards for forensic anthropology. *J. Forensic Sci.* 54:1,211–1,216.

Christensen, A. M., C. M. Crowder, S. D. Ousley et al. 2014. Error and its meaning in forensic science. *J. Forensic Sci.* 59:123–126.

Claes, P., D. K. Liberton, K. Daniels et al. 2014. Modeling 3D facial shape from DNA. *PLoS Genet.* 10:e1004224.

CNN. 2013. Exonerated: Cases by the numbers. Retrieved from: https://wwwcnncom/2013/12/04/justice/prisoner-exonerations-facts-innocence-project/indexhtml.

Committee on Identifying the Needs of the Forensic Sciences Community, and National Research Council. 2009. Strengthening Forensic Science in the United States: A Path Forward. Washington, D.C.

Cooper, S. L. 2016. Challenges to fingerprint identification evidence: Why the courts need a new approach to finality. *Mitchell Hamline Law Review* 42:756–790.

Crowder, C., C. W. Rainwater, and J. S. Fridie. 2013. Microscopic analysis of sharp force trauma in bone and cartilage: A validation study. *J. Forensic Sci.* 58:1,119–1,126.

Edgar, H. J. H. 2013. Estimation of ancestry using dental morphological characteristics. *J. Forensic Sci.* 58:S3–S8.

Expert Working Group on Human Factors in Latent Print Analysis. 2012. Latent print examination and human factors: Improving the practice through a systems approach.

Franke, K., and S. N. Srihari. 2007. Computational forensics: Towards hybrid-intelligent crime investigation. Third International Symposium on Information Assurance and Security, 383–386.

Garrett, B. L. 2012. *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*. Cambridge, MA: Harvard University Press.

Grivas, C. R., and D. A. Komar. 2008. Kumho, Daubert, and the nature of scientific inquiry: Implications for forensic anthropology. *J. Forensic Sci.* 53:771–776.

Hackman, L., and S. Black. 2013. The reliability of the Greulich and Pyle atlas when applied to a modern Scottish population. *J. Forensic Sci.* 58:114–119.

Hauther, K. A., K. L. Cobaugh, L. M. Jantz et al. 2015. Estimating time since death from postmortem human gut microbial communities. *J. Forensic Sci.* 60:1,234–1,240.

Hefner, J. T., and S. D. Ousley. 2014. Statistical classification methods for estimating ancestry using morphoscopic traits. *J. Forensic Sci.* 59:883–890.

Hughes, C. E., B. F. B. Algee-Hewitt, R. Reineke et al. 2017. Temporal patterns of Mexican migrant genetic ancestry: Implications for identification. *Am. Anthropol.* 119:193–208.

Kaye, D. H. 2010. *The Double Helix and the Law of Evidence.* Cambridge, MA: Harvard University Press.

Kim, D. I., U. Y. Lee, S. O. Park et al. 2013. Identification using frontal sinus by three-dimensional reconstruction from computed tomography. *J. Forensic Sci.* 58:5–12.

Kimmerle, E. H., D. A. Prince, and G. E. Berg. 2008. Inter-observer variation in methodologies involving the pubic symphysis, sternal ribs, and teeth. *J. Forensic Sci.* 53:594–600.

Konigsberg, L. W. 2015. Multivariate cumulative probit for age estimation using ordinal categorical data. *Ann. Hum. Biol.* 42:368–378.

Konigsberg, L. W., S. R. Frankenberg, and H. M. Liversidge. 2016. Optimal trait scoring for age estimation. *Am. J. Phys. Anthropol.* 159:557–576.

Lesciotto, K. M. 2015. The impact of Daubert on the admissibility of forensic anthropology expert testimony. *J. Forensic Sci.* 60:549–555.

The National Academies Press. 2018. Stats on Strengthening Forensic Science in the United States.

National Commission on Forensic Science (NCFS). 2016. Testimony using the term "reasonable scientific certainty."

Page, M., J. Taylor, and M. Blenkin. 2011a. Forensic identification science evidence since Daubert: Part I--A quantitative analysis of the exclusion of forensic identification science evidence. *J. Forensic Sci.* 56:1,180–1,184.

Page, M., J. Taylor, and M. Blenkin. 2011b. Forensic identification science evidence since Daubert: Part II--Judicial reasoning in decisions to exclude forensic identification evidence on grounds of reliability. *J. Forensic Sci.* 56:913–917.

Park, C. S., H. P. Jeon, K. S. Choi et al. 2017. Application of 3D laser scanner to forensic engineering. *J. Forensic Sci.* 63:930–934.

Perlin, M. W., J. M. Hornyak, G. Sugimoto et al. 2015. TrueAllele® genotype identification on DNA mixtures containing up to five unknown contributors. *J. Forensic Sci.* 60:857–868.

Saks, M. J., T. Albright, T. L. Bohan et al. 2016. Forensic bitemark identification: Weak foundations, exaggerated claims. *J. Law Biosci*. 3:538–575.

Snow, C. C. 1982. Forensic anthropology. *Annu. Rev. Anthropol.* 11:97–131.

Steadman, D. W. 2009. Personal identification: Theory and applications. The case study approach. In *Hard Evidence: Case Studies in Forensic Anthropology*, 2nd ed. D. W. Steadman, ed. Upper Saddle River, NJ: Prentice Hall, 1–7.

Steadman, D. 2018. Who needs data? I've got experience! Paper presented at the 87th Annual Meeting of American Association of Physical Anthropologists. 2018 April. Austin, TX.

Steadman, D. W., B. J. Adams, and L. W. Konigsberg. 2006. Statistical basis for positive identification in forensic anthropology. *Am. J. Phys. Anthropol.* 131:15–26.

Stephan, C. N. 2014. The application of the central limit theorem and the law of large numbers to facial soft tissue depths: T-Table robustness and trends since 2008. *J. Forensic Sci.* 59:454–462.

Stephan, C. N., E. K. Simpson, and J. E. Byrd. 2013. Facial soft tissue depth statistics and enhanced point estimators for craniofacial identification: The debut of the shorth and the 75-shormax. *J. Forensic Sci.* 58:1,439–1,457.

Stoyanova, D. K., B. F. B. Algee-Hewitt, J. Kim et al. 2017. A computational framework for age-at-death estimation from the skeleton: Surface and outline analysis of 3D laser scans of the adult pubic symphysis. *J. Forensic Sci.* 62:1,434–1,444.

Taylor, M. C., T. L. Laber, P. E. Kish et al. 2016. The reliability of pattern classification in bloodstain pattern analysis, part 1: Bloodstain patterns on rigid non-absorbent surfaces. *J. Forensic Sci.* 61:922–927.

Tise, M. L., M. K. Spradley, and B. E. Anderson. 2013. Postcranial sex estimation of individuals considered Hispanic. *J. Forensic Sci.* 58:S9–S14.

van Oorschot, R. A. H., K. N. Ballantyne, and R. J. Mitchell. 2010. Forensic trace DNA: A review. *Investig. Genet.* 1:14.

Villa, C., J. Buckberry, and N. Lynnerup. 2016. Evaluating osteological ageing from digital data. *J. Anat.*

Walsh, S., L. Chaitanya, L. Clarisse et al. 2014. Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci. Int. Genet*. 9:150–161.

Wiersema, J., J. C. Love, and L. G. Naul. 2009. The influence of the Daubert guidelines on anthropological methods of scientific identification in the medical examiner setting. In *Hard Evidence: Case Studies in Forensic Anthropology*, 2nd ed. D. W. Steadman, ed. Upper Saddle River, NJ: Prentice Hall, 80–90.