

MANAN AHMED ASIF¹

TECHNOLOGIES OF POWER – FROM AREA STUDIES TO DATA SCIENCES

INTRODUCTION

This essay is an attempt to bring together two seemingly divergent trends in the American university of the recent past: first is the disciplinary presence of ‘area studies’ in the US academy since 1958, and the second is the rise of ‘data science institutes’ on US campuses since 2008. The first is responsible for the training of vast numbers of US citizens in languages and cultures (sometimes labeled ‘civilizations’), most frequently, of the places of the world which are of geo-strategic concern to the United States. This training has resulted in the concomitant production of academic scholars of “Near East, East Asia, Middle East, Southeast and South Asia” over the decades with hundreds, perhaps thousands, tenured professorships, monographs etc. The second is the result of a strategic shift of funding away from area studies in 2008 and towards automation, algorithmic capacities, and data analysis which created new offices, new buildings, new faculty positions in data sciences on American campuses. Where the Department of State was the federal funding agency for the first, the *Defense Advanced Research Project Agency* (DARPA) is often the federal funding agency for the second. What combines the two, this essay will argue, is the presence of philology and the primacy of the military concerns of the state – they are both technologies of power, which ought to be collectively studied. In linking ‘area studies’ to ‘data sciences’, I am arguing not for a simple rhetorical framing but to see

¹ I want to thank Tim H. Horning (University Archives) and Sohoni Pushkar (South Asian Studies Bibliographer) at the University of Pennsylvania, Lucas Buresch and Patricia Rosenfield at the Rockefeller Archive Center, and Jennifer S. Comins, Archivist (Carnegie Collections) at the Rare Book & Manuscript Library at Columbia University. Additionally, my gratitude to Durba Mitra, Dennis Tenen, Shahid Amin, David Szanton, and David Lelyveld for their thoughts and questions which helped shape my concerns here.

how the critical philological method was to the accumulation of data about the colonized body, in continental North America and later in the global south. I offer two interventions: first, a re-definition of ‘data’ in order to fold in the history of philology, and a recognition of the grammar and phrase books and the dictionaries not only as those critical tools of colonization but as well data for it. Second, I argue that we need to build upon Bernard S. Cohn’s work on colonial knowledge production, and envisage an ‘algorithmic modality’ within which both the history of philological sciences and data sciences co-exist for the American imperial past and present.

THE STUDY OF AREAS

The philological enterprise lies at the heart of what we understand as “humanistic inquiry”.² It is indeed the case that philology was the critical method employed as a tool to create ‘new’ knowledge of both ‘ancient’ and the ‘new’ worlds – that is from the quixotic search for a biblical, universal language to vocabulary lists of places, flora and fauna in indigenous languages. Yet, we can push that argument further, and say that philology was also the critical method deployed for military gain. In the American context, one can consider the vocabulary lists of native languages acquired by Thomas Jefferson or Alexander von Humboldt in late 18th and early 19th century, which were augmented by grammars, dictionaries and memoirs collected by J.W. Powell and the ethnology scholars.³

The history of area studies in the United States bears out this observation. There is by now a conventional history of area studies. It begins with the coming together of the philanthropic industry, the universities and the American state machinery in the wake of the launch of Sputnik on October 4, 1957 – a date that the northwestern historian L.S. Stavrianos (1913-2004) called the dawn of “the universal age” such that “the entire globe lay open and available”.⁴ It is thought that this new universal age created both an anxiety and a need for American knowledge of the world. The anxiety of knowledge was coupled, and this is understood broadly, by the question of domination – again as Stavrianos asked: who would be the “dominant peoples” that will “comprehend the meaning the universal age, and the organization and

² Edward Said, “The Return to Philology”, in *Humanism and Democratic Criticism*, New York, Columbia University Press, 2004, pp. 57–84.

³ Cp. William C. Sturtevant, “History of Research on the Native Languages of the Southeast”, in Heather K. Hardy and Janine Scancarelli (eds.), *Native Languages of the Southeastern United States*, Lincoln, University of Nebraska Press, 2005, pp. 8–69.

⁴ L.S. Stavrianos, “Is America Ready for the New Age?”, *Northwestern University Tri-Quarterly*, 2, 1959, pp. 15–20.

techniques necessary to exploit the opportunities it offers?” The “who” would have to be the Americans, and the organization and techniques would need to be developed through the “area expert” in the academy.

There are three interlocked structures of support in this conventional history. The first is the legislative – the *National Defense Education Act* (NDEA) which was passed in 1958, as a response to Sputnik, and which created an area studies program focused on East Asia, Near East, South Asia, Middle East and Africa. These programs – first at Cornell, Harvard, Berkeley, University of Pennsylvania, and then at Chicago, Michigan, Columbia, etc. – were funded via the NDEA, the Fulbright-Hays Act, and the Congressional Public Law 480. The second structure are the foundations: such as Ford Foundation, Carnegie Corporation, Rockefeller Foundation. The faculty hired in these area studies programs were further supported by grants from *American Council of Learned Societies* (ACLS), *Social Science Research Council* (SSRC), the *American Institute for India Studies* (AIIS) and others. Finally, the third structure are the university endowments, state legislatures, private endowments which managed students (both graduate and undergraduate), professoriate (language teachers, literary scholars, area specialists) and infrastructure (curriculum). This conventional history recognizes the seeds of area studies’ existence in the state and military apparatus (Sputnik was a *military* crisis), but largely explains it away by stressing their resistance, their independence or their commitment to grander questions of humanistic inquiry.⁵ But this conventional history has a false start.

It is not Sputnik but the long-recognized ‘westward expansion’ that cemented the relationship between colonial knowledge, educational institutes and the technology of power and domination. The origins of area studies should properly be understood in the study of Native American languages and native ‘areas’ of western territories of the United States. The Bureau of Ethnology was established by a Congressional Act, known as the Organic Act of March 3, 1879. The Act established the United States Geological Survey to map, survey and explore the mineral resources and then granted to the Bureau of Ethnology the rights to do systematic ethnographic and philological

⁵ After the critique launched by Said. Cp. Edward Said, *Orientalism*, London, Routledge and Kegan Paul, 1978. See further explication of the military-educational complex in Bruce Cumings, “Boundary Displacement: The State, the Foundations, and Area Studies during and after the Cold War”, in Masao Miyoshi and H. D. Harootunian (eds.), *Learning Places: The Afterlife of Area Studies*, Durham, N.C. and London, Duke University Press, 2002, pp. 261–302; Nicholas B. Dirks, “South Asian Studies: Futures Past” in David L. Szanton (ed.), *The Politics of Knowledge: Area Studies and the Disciplines*, Berkeley, University of California Press, 2004, pp. 341–386. For the humanistic side, cp. Sheldon Pollock, “Philology in Three Dimensions”, *postmedieval: A Journal of Medieval Cultural Studies*, 5 (4), 2014, pp. 398–413.

surveys of the ‘areas’ held by native nations. It stipulated that “the best interests of the public domain require, for the purpose of intelligent administration, a thorough knowledge of its geologic structure, natural resources, and products. The domain embraces a vast mineral wealth in its soils, metals, salines, stones, clays, etc.”⁶

A key role played in this data collection was by the Smithsonian Institute – the first national collecting effort launched by an Act of Congress in 1843.⁷ In 1880, geologist J.W. Powell of the Smithsonian Institute, led a set of academics – linguists and geographers – for the geological survey to document “the lower states of culture exhibited by the tribes of men [...]. Customs, laws, governments, institutions, mythologies, religions, and even arts cannot be properly understood without a fundamental knowledge of the languages”⁸. Powell worked within and outside the nascent academy under the umbrella of the Smithsonian Institute to build ties between the philologists of Hebrew, Assyrian, Latin, Arabic and Sanskrit – foremost among them was Whitney. As the leading expert of Sanskrit in America, Whitney had extensive ties to the Smithsonian Institute. For Powell’s first publication, he designed a phonetic alphabet for the representation of Native American sounds.⁹ These scholars set the foundational relationship between philological inquiry, territorial control and the collecting of national data.

The creation of geological surveys, maps, and access points was embedded with the creation of vocabularies, grammars and language archives of the Native Americans. This was the first gathering of data undertaken by both the governmental and academic enterprises. The

⁶ United States House of Representatives, *The Miscellaneous Documents of the State of the United States for the Third Session of the Forty-Fifth Congress 1878-79*, Washington, Government Printing Office, 1879, p. 51. See also J. W. Powell, *First Annual Report of the Bureau of Ethnology*, Washington, Government Printing Office, 1881, p. xv. For a romanticized treatment of Powell, cp. Wallace Stegner, *Beyond the Hundredth Meridian*, New York, Houghton and Mifflin, 1954.

⁷ Bequeathed by James Smithson in 1829, the law was passed in 1835 under Andrew Jackson. Cp. Pamela M. Henson, “‘Objects of Curious Research’: The History of Science and Technology at the Smithsonian”, *Isis*, 90, 1999, pp. S249–S269. Also see George Brown Goode (ed.), *The Smithsonian Institution 1846-1896: The History of its First Half Century*, Washington, De Vinne Press, 1897.

⁸ Powell, “First Annual Report”, p. xv.

⁹ Whitney expanded the alphabet published in 1861 by the Smithsonian Institute by George Gibbs in “Instructions for research relative to the Ethnology and Philology of America”. Cp. J. W. Powell, *Introduction to the Study of Indian Languages with Words, Phrases and Sentences to be Collected*, Washington, Government Printing Office, 1877, p. 1. Powell also wrote an essay on philology calling it the “science of activities”. Cp. J. W. Powell, “Philology, or the Science of Activities Designed for Expression”, *American Anthropologist*, 2, 1900, pp. 603–637. Whitney also worked with Henry Louis Morgan on building comparative vocabulary lists for Native American and Indic terms of kinship. Cp. Thomas R. Trautmann, *Louis Henry Morgan and the Invention of Kinship*, Berkeley, University of California Press, 1987, pp. 226–229.

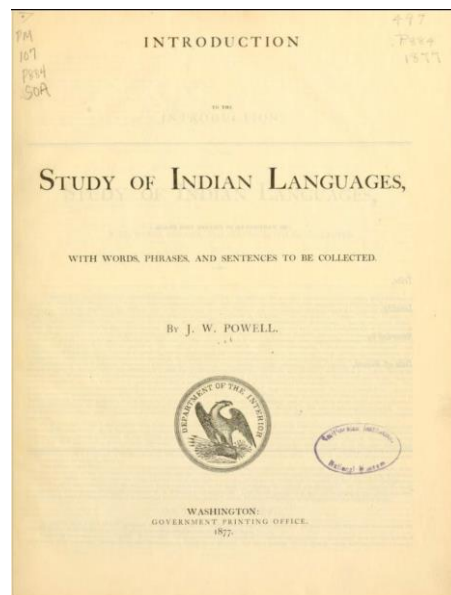
very first handbook for the collection of the study of Indian languages produced by J.W. Powell was already a database. Each handbook began with metadata recording – ‘Tribe’, ‘Locality’, ‘Recorded by’, ‘Date of Record’ – followed by an empty three columned page, under headings like ‘Persons’, ‘Parts of Body’, ‘Relationships and Kinship Structures’, ‘Social Organization’, and so on and ending with ‘Geographic Terms’. The first column had an English term (‘Husband’s father’s elder brother’s daughter’s husband, said by female cousin-in law’) with an empty column for recording the native term and the last column was ‘Remarks’. The handbook would then be filled by the geographers in the ‘field’ with data, to be converted into dictionaries. The production of dictionary (of ‘Oriental’ or ‘Semitic’) languages continued apace as ‘critical’ modern languages came to dominate the funding structures of the university.

What do we learn from this altered history of the origins of area studies? First, that insofar as the ‘area studies’ were a response to political stimuli, it was the colonization, mapping, surveying, and settling of United States which gave rise to them. The various geography departments in the United States instituted ‘area studies’ from the 1880s onwards, focusing not only on continental United States; and on the broader areas such as the Arab world, Philippines, India and China since the early 1900. Second, that surveillance and military structure of US governmental or philanthropic enterprise informed *all* forms of knowledge produced under the ‘area studies’ paradigm. Looking deeper into the history of the philanthropic support not only illuminates the very close working of the US state with these NGOs but the stark revelation that the very foundations of area studies were intellectually shaped by them.

THE PRODUCTION OF (BIG) DATA

If philology was the method linking the study of native peoples to the study of village life in India, then so was the production of data sets across geographies. For a little more than a hundred years, American institutions of higher learning have produced collections of words, their meanings, and commentaries of them as a demonstrable product of investment into the university. At the same campus were scholars engaged in archaeological digs excavating remains of great civilizations of the East. At the same campus were scholars discovering lost manuscripts to create critical editions of histories, philosophies or myths. All of this activity formed a critical part of the greater mission of the university – to train American students for lives and careers deserving of their fullest potentials. What was produced or recovered

was ‘data’, namely manuscripts, texts, artifacts, stories, sounds – all raw materials for understanding the civilizational questions. The investments of the Smithsonian Institute to the British Museum in archiving manuscripts, building collections, creating metadata re-constituted a new cartography of knowledge. The manuscript and museum collections at Yale, Harvard, Chicago, Princeton, Columbia, Johns Hopkins, Berkeley, Michigan are repositories of data collected across the span of the 20th century. These data sets were available to the inhabitants of area studies only as participants in Fulbright or other area studies programs. The collection of these data sets thus sequestered knowledge at Harvard, Yale or Columbia’s libraries until the paradigmatic shift of digital technologies emerged in the late 1990s and early 2000s. These libraries became the source code for the next technology of power: mass digitization.



First Database.

Source: John Wesley Powell. *Introduction to the study of Indian languages: with words, phrases and sentences to be collected*, Washington: Government Printing Office, 1877.

The digitization projects initiated by Google – first as technologies for search – launched in 2004 and soon became the face of the “million books”.¹⁰ The massive scanning projects, initiated at libraries at Oxford, Michigan, Stanford or Harvard are now matters of fact for scholars across all disciplines. These projects, though backward-facing, have had to contend with the ‘born digital’ nature of contemporary world as well – the Library of Congress’ American Memory Project and the Archive

¹⁰ *Google Books History* notes that the founders, Larry Page and Sergey Brin, wrote their first algorithms for Stanford’s Digital Library Technologies Project in 1996. Cp. “Google Books History”, *Google Books*. Available at: <https://www.google.com/googlebooks/about/history.html> [accessed October 17, 2019].

Project are two leading examples of how world as data has entered the collectivization process.

It is beyond my purview here to give a history of the creation of ‘Big Data’ but it is important to note that the data-information-knowledge nexus underwent an ontological shift in the last decade of the 20th century. The availability of massive amounts of personal data – self-created – as well as data on a person shifted us from the statistical ledger to the algorithmic gaze. Google’s algorithms, developed to scan book catalogues and websites, were the first commercially successful articulation of how massive data archives could be queried, displayed and archived. As personal computers and smartphones containing imaging and recording technologies entered the global mass markets, they became instruments of producing self-data. To manage all of this data, either at the personal, scholarly or governmental scale, new algorithms took shape as technologies of power – this time in computer science departments. The history of devices – which record sound, text, video – is also the history of algorithms that index, search, analyze and display.

The ‘database’ which houses Big Data as the repository of structured knowledge shares much of philology’s methodological architecture. Databases at their simplest contain data (fact), relationship between the data, and information about that data (structure). The philologist will easily speak to the database administrator about the necessity of clean, structured data and how the information retrieval can best be optimized and how the commentary on data can be housed. These methods of governing databases evolved from the historical ‘ledger’ but just as importantly from the philological tradition. That methodological connection between data, metadata, retrieval and display creates another link to the future of the area studies program – the data science programs.

The launch of Sputnik provided billions of dollars for the universities and dozens of area studies programs flourished. It also funded the agency which developed and continues to develop new technologies of power for the United States. On February 7, 1958, a Department of Defense Directive created the Advanced Research Project Agency, later known as *Defense Advanced Research Projects Agency* (DARPA).¹¹ DARPA gave birth to the computer science department, as well as to the internet protocol (as ARPANET) which gave us the

¹¹ Cp. DARPA, “Where the Future Becomes Now”, *History and Timeline*. Available at: <http://www.darpa.mil/about-us/darpa-history-and-timeline> [accessed October 17, 2019].

internet.¹² In the structure, functioning and framing of concerns, there was broad consensus and collaboration between the scientists, academics and the military “as the contracting agency” with “an underlying appreciation of security needs”.¹³

DARPA thus entered the university at exactly the same time as the area studies program but little or no attention has been paid to it by scholars studying area studies. It is an un-assailable fact that DARPA’s research collaborations with the university shaped US power in ways far more significant than the products of area studies. However, it would be unwise to de-couple DARPA from area studies. They constitute an organic whole – with funding streams criss-crossing private and governmental sources into departments.

After September 11, 2001, the question of language competency was raised again by the congress and by university administrators. However, this time it took a very different shape than the Army Specialized Training Program during World War II. DARPA led the way with machine learning technologies deployed first in Iraq in 2006. DARPA’s *Spoken Language Communication and Translation System for Tactical Use* (TRANSTAC) and the *Global Autonomous Language Exploitation* (GALE) were the first to use algorithms for detection and elimination of adversaries.¹⁴ The GALE program focused on formally structured data (newspapers, news broadcasts) from Arabic into English; TRANSTAC on un-structured data (encounter speech). In addition, DARPA developed *Autonomous Real-time Ground Ubiquitous Surveillance – Imaging System* (ARGUS-IS) program which present continuous video surveillance from “a revolutionary high-altitude, long-endurance, unmanned helicopter”¹⁵. That same year, in September 2008, DARPA launched the *Multilingual Automatic Document Classification Analysis and Translation Program* (MADCAT) which focused on “processing Arabic hand-writing” from texts into data with collaborations with University of Maryland, LeHigh, Penn State, and Cambridge Computer Science departments.¹⁶ The 1,6 million grant for

¹² Cp. Stephen J. Lukasik, “Why the Arpanet was Built”, *IEEE Annals of the History of Computing*, 33 (3), 2011, pp. 4–21.

¹³ Sandra Braman, “The Framing Years: Policy Fundamentals in the Internet Design Process, 1969–1979”, *The Information Society*, 27, 2011, pp. 295–310, here: p. 302.

¹⁴ Cp. Statement of Director Tony Tether, Subcommittee on Terrorism, Unconventional Threats and Capabilities, House Armed Services Committee, U.S. House of Representatives, March, 2008.

¹⁵ Ibid., p. 8. Also see “A Conversation with Jordan Cohen: Speaking out about speech technology”, *ACM Queue*, 4 (6), 2006. Available at: <http://queue.acm.org/detail.cfm?id=1147528> [accessed October 17, 2019].

¹⁶ *Congressional Record*, 154 (16), p. 21926. Also see Stephanie M. Strassel, “Linguistic Resources for Arabic Handwriting Recognition”, *Proceedings of the Second International Conference on Arabic Language Resources and Tools*, 2009, pp. 37–41.

“Document Analysis and Exploitation” fostered a range of activity – under the guidance of the Linguistic Data Consortium, a number of Arabic “Treebanks” were created, like the *Columbia Arabic Treebank* (CATiB).¹⁷

2008 is a pivotal year for this history of the present, and the year that fractured older models of learning in the university to create new configurations. On October 1, 2008, came the passage of the Emergency Economic Stabilization Bill which gave \$700 billion to purchase troubled assets of Lehman, AIG, Merrill, Goldman-Sachs. The global financial crisis dealt a severe hand to philanthropic organizations and to university endowments. We are all familiar now with the lack of tenured lines, the precarity of adjunct labour and the lack of resources across our disciplines. For area studies, the global financial crisis hit particularly hard – because it also re-aligned the spending priorities of the US government. DARPA funded the *Next Generation Social Science* (NGSS) to determine fundamental measures and causal mechanisms that explain and predict the emergence of “collective identities.”¹⁸ The grants have no upper funding limit. Additionally, each year since 2008, DARPA funds the “Young Faculty Awards” which provide \$1 million over two years to outstanding faculty researching everything from AI to Human-Robot Interactions. A similar program is run by the Office of Naval Research Science and Technology as the “Young Investigator Award”. These awards are supported by all universities which circulate the deadlines and celebrate the recipients.

Yet, since 2008, a new ‘area studies’ has emerged, alongside ‘new’ methods. At Columbia, Rochester, Berkeley, University of Virginia, Cornell, Carnegie-Mellon and many other institutions of higher learning, data science institutes, centres and programs have been launched. Funded by private endowments (often Google, Uber, Tesla, etc.) these new area studies programs work in close synchronization with existing disciplinary programs such as electrical engineering or computer science. The faculty and students in these programs work on critical features like Natural Language Processing, Artificial Intelligence and robotics. The technologies created – such as the “remote viewing” via “unmanned aerial vehicles (UAVs)” developed at Carnegie Mellon’s Robotics Institute are chiefly deployed for surveillance and killing of

¹⁷ Nizar Habash and Ryan M. Roth, “CATiB: The Columbia Arabic Treebank”, *Proceedings of the ACL-IJCNLP*, 2009, pp. 221–224.

¹⁸ DARPA, “Broad Agency Announcement Next Generation Social Science (NGS2) Defense Sciences Office”, *DARPA-BAA-16-32*, March 18, 2016 (Due date is May 18).

terrorists in Pakistan, Yemen and Somalia.¹⁹

The knowledge problems here are uniformly related to the creation of vast amounts of data and their analysis via machines and algorithms.²⁰ The phenomenological space previously occupied by the anthropologist, philologist and the area-specialist is now occupied by data and computer scientists. That which was formally tagged as “humanistic inquiry” is now undertaken by digital humanities where disciplines like history, English, comparative literature, sociology are rushing to become digitally savvy and operational alongside their older siblings economics and political science.

The structures underlying these new area studies are surprisingly familiar – the private foundations, the US military and legislative regimes, and the pressures on the university. The bravura nonchalance of the digital humanities to confront its material past carries the same air of insouciance that the young ethnographer setting out from Chicago to Maharashtra, studying marriage customs, had in the pre-2008 era. The digital humanists are just as convinced that their struggle is for knowledge alone – and it is only for them to articulate an ethics of being in the world.

If the languages and cultures of world’s areas were the focus of inquiry for area studies, the Big Data centres focus on the knowable subject: the consumer, the urban city dweller, the political agent, the business agent, the activist, the criminal and most fundamentally the terrorist. Atomistic in its application, the data science algorithms strive to predict activity. For this, it employs much of the same hermeneutics as area studies did: deploying a mathematical science in order to predict data, oscillating between an idealized social science or civilizational theoretical models and user-generated data.

I will focus on just one example – the drone assassination program. In 2009, the United States Airforce, laid out an ambitious fifty year plan, predicated on the availability of a UAV based camera that can survey enemy territory for a vast period of time, until it needed to act to eliminate the target – they called it “Gorgon Stare”.²¹ This “Find, Fix,

¹⁹ Julie Albright, “Flight of the UAV”, *Carnegie Mellon Today*, January 11, 2016. Available at: http://cmtoday.cmu.edu/robotics_innovation/automated-aerial-vehicle-technology/ [accessed October 17, 2019].

²⁰ A very illuminating video explicating the relationship between DARPA, Data, machine learning and scholarship (“machine will read literature”) is on DARPA’s YouTube Channel. Cp. Paul R. Cohen, “The Big Mechanism Program and the Future of Scholarship”, *DARPAtv*, June 30, 2015. Available at: https://www.youtube.com/watch?v=GCaXSyY_C9Y [accessed October 17, 2019].

²¹ United States Air Force, “United States Air Force Unmanned Aircraft Systems Flight Plan 2009-2047”, *Federation of American Scientists*, Washington, 2009. Available at https://fas.org/irp/program/collect/uas_2009.pdf [accessed October 17, 2019].

Finish” plan rested on keeping afloat a drone (UAV) for long enough to force the “enemy target” to reveal itself. 2008 was a critical turning point for the usage of drones for the assassination program. Under George W. Bush, only 51 known drone strikes came between 2004–2009. Over 500 strikes have happened in Pakistan, Yemen, Somalia since then.²² The logic of remote killing of suspected terrorists via drones operates on a specific tally of metrics – related to locations, to conversations, and to acts – to declare “all military-age males” within a specific pre-determined “ungoverned” space as combatants.²³

In 2013, then President Obama provided the official rationale for the administration’s policy behind drone assassinations, the war against global terrorism, the ways to combat it, and the drone’s rendering of space. In his rationale, Obama accounted for a terrain that was unknowable without technological surveillance:

“But despite our strong preference for the detention and prosecution of terrorists, sometimes this approach is foreclosed. Al Qaeda and its affiliates try to gain foothold in some of the most distant and unforgiving places on Earth. They take refuge in remote tribal regions. They hide in caves and walled compounds. They train in empty deserts and rugged mountains. In some of these places – such as parts of Somalia and Yemen – the state only has the most tenuous reach into the territory. In other cases, the state lacks the capacity or will to take action. And it’s also not possible for America to simply deploy a team of Special Forces to

²² This data is provisional and I am citing the Bureau of Investigative Journalism here. Cp. Jack Serle and Jessica Purkiss, “Drone Wars: The Full Data”, *The Bureau of Investigative Journalism*, 2017. Available at: <https://www.thebureauinvestigates.com/category/projects/drones/drones-graphs/> [accessed October 17, 2019].

²³ For location, cp. Jake Laperruque, “NSA Tracking Location on a Massive Scale”, *Center for Democracy and Technology*, December 16, 2013. Available at: <https://www.cdt.org/blogs/jake-laperruque/1612nsa-tracking-location-massive-scale> [accessed October 17, 2019]; Barton Gellman and Ashkan Soltani, “NSA tracking cellphone locations worldwide, Snowden documents show”, *The Washington Post*, December 4, 2013. Available at: http://www.washingtonpost.com/world/national-security/nsa-tracking-cellphone-locations-worldwide-snowden-documents-show/2013/12/04/5492873a-5cf2-11e3-bc56-c6ca94801fac_story.html [accessed October 17, 2019]. For voice surveillance, cp. Paul Lewis, “Snowden Documents Show NSA Gathering 5bn Cell Phone Records Daily”, *The Guardian*, December 5, 2013. Available at: <http://www.theguardian.com/world/2013/dec/04/nsa-storing-cell-phone-records-daily-snowden> [accessed October 17, 2019]. For targeting, cp. David S. Cloud, “CIA Drones Have Broader List of Targets”, *The Los Angeles Times*, May 5, 2010. Available at: <https://www.latimes.com/world/la-xpm-2010-may-05-la-fg-drone-targets-20100506-story.html> [accessed October 17, 2019]; Jeremy Scahill and Glenn Greenwald, “The NSA’s Secret Role in the US Assassination Program”, *The Intercept*, February 10, 2014. Available at: <https://firstlook.org/theintercept/article/2014/02/10/the-nas-secret-role/> [accessed October 17, 2019]. For un-governed space, cp. Ian Cobain, “Obama’s Secret Kill List – The Disposition Matrix”, *The Guardian*, July 14, 2013. Available at: <http://www.theguardian.com/world/2013/jul/14/obama-secret-kill-list-disposition-matrix> [accessed October 17, 2019].

capture every terrorist. Even when such an approach may be possible, there are places where it would pose profound risks to our troops and local civilians – where a terrorist compound cannot be breached without triggering a firefight with surrounding tribal communities, for example, that pose no threat to us; times when putting U.S. boots on the ground may trigger a major international crisis. [...] So it is in this context that the United States has taken lethal, targeted action against al Qaeda and its associated forces, including with remotely piloted aircraft commonly referred to as drones.”²⁴

Obama’s words echo a material history of violence directed against “unforgiving places” in the actions of Andrew Jackson against native peoples in New Orleans and in Florida. In the defense of Jackson’s brutality by then President James Monroe, in 1818, Jackson’s violence was permissible because “the territory belonged, in a certain sense at least, to the savage enemy who inhabited it.”²⁵ The White House administration used this particular history from 1818 to provide legal sanction for their actions in 2013.²⁶

This linking of legal histories is symbolic of the epistemic linkages that I have argued here between the philological enterprise of early 20th century and the early 21st. The technologies of power that determined the creation of dictionaries and ethnographies of Native Americans relied on the expertise of indologists and the participation of academic institutions and philanthropic organizations. The creation of area studies expanded that nascent empire of knowledge to create a global imperium wherein production of data flourished alongside archivization efforts. Though the mechanics of funding shifted in 2008, the epistemic regimes have a new facade and a stronger base in the American universities – as Artificial Intelligence, Unmanned Aerial Vehicles, and Neural Learning emerged as the academy’s contribution to the war effort in 2008-9, the new technologies of power.

An argument for a conjunction of American university, think-tanks, cultural institutions as sites of knowledge-and-war making prefaces the manifest desire to conquer brown and black bodies. Where the standard histories of “Area Studies” makes oblique its funding structures, the

²⁴ The White House, Office of the Press Secretary, “Remarks by the President at the National Defense University”, *Speeches & Remarks*, March 23, 2013. Available at: <https://www.whitehouse.gov/the-press-office/2013/05/23/remarks-president-national-defense-university> [accessed October 17, 2019].

²⁵ James Monroe, “Second Annual Message”, November 16, 1818.

²⁶ Cp. Brian C. Baldrate, “The Supreme Court’s Role in Defining the Jurisdiction of Military Tribunals: A Study, Critique, & Proposal for Hamdan V. Rumsfeld”, *Military Law Review*, 186, 2005, pp. 1–115

standard histories of Big Data, AI, drones, or even digital humanities, makes oblique the relationship between the wars in Iraq, Afghanistan and the development of key techs as weapons. This essay has argued for a phenomenological reading of the history connecting knowledge practices and imperial practices in the United States. The structures of funding sketched here are emblematic of the deep linkages between knowledge production and the state.

However does this argument fundamentally alter the way in which we can think about the history of mediation between the academy and the military? I offer, in conclusion, one provocation for a methodological framework. It would be an addendum to the intervention of Bernard S. Cohn in his seminal essay collection “Colonialism and Its Forms of Knowledge” where he defined several modalities which governed the making of knowledge for British colonial power – the historiographic, the survey, the enumerative, the travel, the museological, and the surveillance.²⁷

We need a new modality, that I would term, the “algorithmic modality” – based on the foundation of philology – which would assemble the ways in which data has been conceived of with the American imperial paradigm from the 18th century to the present; which would trace how this data was organized into ‘usable’ knowledge through journals, societies, institutions and departments; and how this modality organizes the routes of power through state and imperial functionaries – in this case, academicians and researchers at universities. The algorithmic modality would disallow any triumphal narrative about the data sciences and would make clear that the historians and thinkers of coloniality deserve to be seated at the very tables where automation of our present is being considered. We face a great danger from the automation of the present to the automation of the writing of our past itself – one in which the confluences between power and state would not even merit a footnote.

²⁷ Bernard S. Cohn, *Colonialism and Its Forms of Knowledge*, Princeton, NJ, Princeton University Press, 1996, pp. 1–15.