

Reading Race in Slavic Studies Scholarship through a Digital Lens

Hilah Kohen, Katherine M. H. Reischl, Andrew Janco, Susan Grunewald, and Antonina Puchkovskaia

Introduction

In her widely circulated 2017 article for *The Chronicle of Higher Education*, Sarah Valentine recounts a conversation with her graduate mentor about the isolation and discrimination Black students face while studying in Russia. Valentine was told “that the Slavic field had always been more concerned with the political and cultural dynamics existing between the various Slavic groups...than with ‘outside concerns.’”¹ Along with this narrow definition of our field comes a serious implication: Valentine’s well-being, even as a member of “the Slavic field” itself, was an “outside concern;” Blackness, for *her* field, was and is an “outside concern.”²

Testimonials like Valentine’s indicate that Slavic Studies has a problem with racist exclusion that works by the elision of racial issues.³ Even the 2020 “AATSEEL Statement Concerning Systemic Racism and Police Brutality in the United States,” one of the most productive plans for anti-racist action to emerge from our field’s leadership, opens: “AATSEEL does not generally make statements about public issues unless they directly relate to the Slavic field.”⁴ Framing US-based racism as an exception to this rule is factually incorrect: not only does racism in the US have urgent reverberations in Eurasia, eastern Europe, and Russia, but also nothing relates more directly to “the Slavic field” than the well-being of that field’s own members when their livelihood is hampered by racial discrimination and violence.

These present-day conversations have roots in the history of the field of Slavic Studies and its perceived and real marginalization. Articles traded year after year between luminaries from Roman Jakobson to Ronald Grigor Suny have contained calls for greater attention to Slavic “diversity” as it was threatened by the monolith of a Russo-centric Soviet state and by the crisis of that state’s dispersal.⁵ A redefinition on geographic terms ultimately took hold,

1. Sarah Valentine, “Russian Studies’ Alt-Right Problem,” *The Chronicle of Higher Education*, September 29, 2017 at <https://www.chronicle.com/article/russian-studies-alt-right-problem/> (accessed May 3, 2021).

2. Ibid.

3. See also, Eduardo Bonilla-Silva, *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America* (Lanham, MD, 2017).

4. AATSEEL Executive Council, “AATSEEL Statement Concerning Systemic Racism and Police Brutality in the United States,” June 2020 at https://www.aatseel.org/about/presidents-message/messages/#stop_racism (accessed May 3, 2021).

5. For example, Waclaw Lednicki and Roman Jakobson traded statements in 1954. Waclaw Lednicki, “The State of Slavic Studies in America,” *The American Slavic and East European Review* 13, no. 1 (February 1954); Roman Jakobson, “Comparative Slavic Studies,” *The Review of Politics* 16, no. 1 (January, 1954): 67–90. In Lednicki’s call to the

Slavic Review 80, no. 2 (Summer 2021)

© The Author(s) 2021. Published by Cambridge University Press on behalf of the Association for Slavic, East European, and Eurasian Studies.

doi: 10.1017/slr.2021.78

adding “Eurasia” to the title of our discipline’s organizational body in 2010.⁶ Yet, as Ani Kokobobo recently noted, the majority of PhD-granting institutions in the US have retained the phrase “Slavic Languages and Literatures” in the title of their departments.⁷ We ask: how are these decisions about what the field is reflected in what the field writes?

Anti-racist movements in other predominantly white disciplines point toward a methodology for approaching this question. In recent months, Princeton Classics Professor Dan-El Padilla Peralta has not only highlighted disparities within his field but argued that the very foundation of Classics is racist.⁸ The *New York Times Magazine* framed these arguments as “a crisis of identity” in a field seeking “to shed its self-imposed reputation as an elitist subject overwhelmingly taught and studied by white men.”⁹ As Padilla argues, however, this foundation may be impossible to shed without a radical transformation of the discipline’s foundations. The much newer field of Digital Humanities (DH) has been shaped by internal review that brings methodology and representation to the fore.¹⁰ Tara McPherson has observed that digital media and the Civil Rights era emerged as intertwined responses in a shared Cold War context. This foundational imbrication necessitates the incorporation of “race from the outset...as a ghost in the digital machine.”¹¹

This article offers a first attempt to articulate the “ghosts in our [Slavic] machines” through a new assessment of the shape of our field that stops pretending race is incidental. Current engagement with racism in the United States requires a response that is both immediate and sustained; using what

“state of the field,” Slavic Studies must serve to “explain to the American people the cultural differentiation which exists behind the ‘Iron Curtain’...” Lednicki, 108. Jakobson promoted a definition of “Slavic Studies” on linguistic grounds. Jakobson writes, “Slavic peoples are to be defined basically as a Slavic-speaking peoples.” Questions of ethnic identity are smoothed to encompass a unified classification capable of countering the “subsidiary, marginal” position “Slavs” had in the US scholarly landscape. Jakobson, 67.

6. Dmitry P. Gorenburg and Ronald G. Suny, “Where Are We Going? What is To Be Done?,” *AAASS NewsNet* 46, no.4 (August 2006): 1, 3, available at <https://pitt.app.box.com/s/5jbxq0iql803x6m99g62ki413nvtb5ll> (accessed May 3, 2021). The American Association for the Advancement of Slavic Studies transitioned to become the Association for Slavic, East European, and Eurasian Studies. AATSEEL remains the American Association of Teachers of Slavic and East European Languages.

7. Ani Kokobobo, “What’s in a Name? Are We Slavic, East European, Eurasian, or All of the Above?” *ASEEES NewsNet* (August 2020), 17, available at <https://www.aseees.org/news-events/aseees-blog-feed/what%E2%80%99s-name-are-we-slavic-east-european-urasian-or-all-above> (accessed May 3, 2021).

8. Rachel Poser, “He Wants to Save Classics From Whiteness. Can the Field Survive?,” *New York Times Magazine* (February 20, 2021) at <https://www.nytimes.com/2021/02/02/magazine/classics-greece-rome-whiteness.html> (accessed May 3, 2021).

9. *Ibid.*

10. See Matthew K. Gold, ed., *Debates in the Digital Humanities* (Minneapolis, 2012) at <https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities> (accessed May 3, 2021).

11. Tara McPherson, “Why Are the Digital Humanities So White? or Thinking the Histories of Race and Computation,” in Matthew K. Gold, ed., *Debates in the Digital Humanities* at <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfb1e/section/20df8acd-9ab9-4f35-8a5d-e91aa5f4a0ea#ch09> (accessed May 3, 2021).

Alex Gil and his collaborators call “nimble” digital methodologies or “rapid response research,” we can quickly and broadly assess a field that occupies a unique racial position in the US academy and has found that position to be in need of revision since the field’s very inception.¹² Our approach merges theories of race and DH methods to ask how published articles in “Slavic Studies” do and do not reflect critically on race.¹³ Because the textual field of Slavic Studies has never previously been analyzed as a corpus, the application of these methods is preliminary; however, our results gesture toward concrete circumstances and actionable steps. Namely, in some areas of identity (such as gender), Slavic Studies research has generated conversations that are robust enough to be visible from a digital bird’s-eye view. Although individual articles stretching back two decades demonstrate the importance of race in Russia and Eurasia, the absence of any large, digitally detectable conversations on this topic reinforces Black students’ observations of negligence and ignorance. Meanwhile, works of scholarship that do offer a critical apparatus for thinking about race and Eurasia also demonstrate a great potential for interdisciplinary impact. Such writings can point an anti-racist path forward for the field as a whole.

Methods and Results

Digital Humanities offers the ability to study materials at a scale that would be impossible for any single scholar to grasp. Three specific computational methods allow us to ask research questions of entire fields and disciplines in a sample of over 100,000 scholarly texts: frequency analysis, topic modeling, and perspectival modeling. For predominantly English-language academic sources, we analyzed 41,251 texts, including both articles and books, provided by JSTOR Data for Research within a “Slavic Studies” cluster. For each text, JSTOR provides a list of all the words in the text and their frequency. For scholarship in Russian, we included texts from thirty seven journals. We used each of these samples to ask how our field represents two categories of socially perceived identity: race and gender. Complete information about our methods, corpus, models, and results, as well as additional images are available in this article’s companion GitHub repository.¹⁴

12. “Rapid Response Research (RRR) projects are quickly deployed scholarly interventions in pressing political, social, and cultural crises. Together, teams of researchers, technologists, librarians, faculty, and students can pool their existing skills and knowledge to make swift and thoughtful contributions through digital scholarship in these times of crisis.” *The Nimble Tents Toolkit* at <https://nimbletents.github.io> (accessed May 3, 2021).

13. See Chike Jeffers, “Cultural Constructionism,” in Joshua Glasgow, Sally Haslanger, Chike Jeffers, and Quayshawn Spencer, eds., *What Is Race?: Four Philosophical Views* (New York, 2019).

14. Our project GitHub repository is available at <https://github.com/Russian-NLP/Reading-Racial-Discrimination-in-Slavic-Studies-Scholarship> (accessed May 3, 2021).

Preliminary Approaches

Reading discrimination on a digital scale does not only mean using a computer to scan articles for racist or sexist labels. We initially attempted to screen our samples for outright hate speech using *Hatebase*, a multilingual repository of 3,700 derogatory terms.¹⁵ Term frequency analysis indicated how many times each *Hatebase* term appears in our samples, and in which texts each term appears most often. Pursuing this blunt approach demonstrates that discrimination in Slavic Studies publications is of a different kind. Slurs and derogatory terms do appear in our corpus by the thousands, but they are primarily used in quoted or reported speech with varying degrees of contextualization.

For example, a manual survey of texts that frequently use the word “whore(s)” found that this word (which appears 201 times in our English corpus) often arises in immutable, textually derived phrases (such as “Whore of Babylon”) or in critical analyses of historical archetypes (alongside “virgin” and “mother”). Similarly, the presence of slurs and often-offensive terms about Blackness in our sample indicates that race figures prominently in the sources Slavicists use—perhaps even more prominently than sex work. The word “negro” and its plural appear 732 times in our corpus; the n-word and its plural appear thirty seven times, typically but not always in primary sources (such as the title of a novel by Joseph Conrad).¹⁶ Our Russian corpus yielded comparable ratios, with *shliukha* and *bludnitsa* arising a total of 3,004 times and translations of derogatory terms about Black people appearing 5,597 times.

The English word “whore” is overtly dehumanizing in common parlance; the n-word has been used for centuries to label millions of people as livestock or worse. Any casual repetition of these words should never be publishable. However, frequency analysis is incapable of quantifying how well scholars contextualize these words and address empirical abuses. When discriminatory terms about race and other identity formations appear in our field’s source materials, what matters is whether researchers contextualize broader histories of racism and discrimination.¹⁷ Digital methods can indicate whether such expertise has developed in Slavic Studies scholarship.

15. See *Hatebase* at <https://hatebase.org/> (accessed May 3, 2021). This database contains terms related to nationality, ethnicity, religion, gender, sexual orientation, disability, and class in ninety-five languages. It is based on texts from 175 countries.

16. To filter out the adjective for black in Spanish (“negro”) we identified the language of each text and did not include Spanish-language texts in the count. A regular expression `r“\bnegro\b”` was used to exclude the word Montenegro and other similar terms that might give false positive matches.

17. Frequency analysis can point to articles that use *Hatebase* terms often. This approach is most likely to locate *positive* role models. For example, scholars of southeastern Europe have already demonstrated the feasibility of a mass shift away from the casual repetition of racialized hate speech in their increasing use of the words “Romani” and “Roma” rather than the often-pejorative term “gypsy.” While the terms “gypsy” and “gypsies” were the most common *Hatebase* terms in our corpus, with 1,872 and 2,058 uses respectively (and 4,664 uses for *tsygan[e]* in Russian), we found that their use has decreased noticeably since 2016. Recent pieces that do include these words often use “Romani” as well and contextualize the difference.

When a scholarly field is collectively interested in understanding a topic, it leaves textual traces of that interest beyond individual words. It includes many words about that topic in the same texts, as scholars undertake extended discussions rather than encountering the topic tangentially. A field's areas of interest spawn interconnected citations, as the same names become associated with certain terms. The limited results of our frequency analysis led to a search for such networks of conversation using a different method of text analysis called topic modeling. We applied three rounds of topic modeling, each using a different algorithm, to texts in the JSTOR Slavic Studies sample. Then, we manually examined the subject areas found to be important in this sample, searching for conversations about embodied identity.

Topic modeling is a form of machine learning that searches for statistically significant themes within texts.¹⁸ Each time it runs, the model outputs a set number of “topics,” each of which is actually a long chain of characteristic words. Topic modeling clusters words from its input corpus with no indication from humans as to why the sample is interesting or what the words in it mean.¹⁹ However, a computer can tell that when a Slavic Studies publication contains the word “feminism,” it is also likely related to publications that contain words like “gay,” “girls,” “dowry,” “*zhenotdel*,” and “sex.” This particular set of statistically connected words is sampled from the list of 200 terms that our third round of modeling labeled (arbitrarily) as topic #1.²⁰ A human can look at this “topic” and surmise that *gender relations* are a subject of great interest in Slavic Studies. Humans can also tell how robust this interest is by examining the topic's formative “fingerprint,” that is, the frequently used and highly connected words that models put at the front of each topic. The gender relations “fingerprint” contains not only abstract concepts (like “sexuality,” the 22nd term) and relational terms (“papa,” 4th, and “mama,” 5th) but also historical figures (Maria “*pokrovskaiia*,” 23rd; Sophia “*parnok*,” 29th), and current scholars (Wendy “*rosslyn*,” 25th; Helena “*goscilo*,” 43rd). This network of ideas, histories, and people has played a significant role in allowing conversations about gender to coalesce in our field.

By contrast, on this macroscopic scale, our field writes about racial and ethnic histories without acknowledging the presence of race and ethnicity in those histories. A smaller cluster of scholarship has repeatedly demonstrated the potency of racialization and ethnicity as descriptors for identity in Central

18. In brief, we used the BERT model for TF-IDF. See the GitHub for a full explanation.

19. For an accessible and brief explanation of topic modeling, see Teddy Roland, “Topic Modeling: What Humanists Actually Do With It,” *Digital Humanities at Berkeley*, July 14, 2016 at <https://digitalhumanities.berkeley.edu/blog/16/07/14/topic-modeling-what-humanists-actually-do-it-guest-post-teddy-roland-university> (accessed May 3, 2021).

20. Note that this topic is actually listed second in our results, as there is a topic #0. In topic #1, the first thirty terms in order are as follows: feminist, feminism, trafficking, papa, mama, gay, feminists, girls, lesbian, husband, mothers, sisters, yesterday, dowry, *zhenotdel*, daughter, sex, daughters, married, husbands, divorce, sexuality, *pokrovskaiia*, prostitution, *rosslyn*, motherhood, *khaia*, marriage, *parnok*, heroine. For our full results from this algorithm, see the following file on this article's GitHub repository at https://github.com/Russian-NLP/Reading-Racial-Discrimination-in-Slavic-Studies-Scholarship/blob/main/1_topic_modeling/SlavicStudiesN3_BERTopic.csv (accessed May 3, 2021).

Asia, Russia, and eastern Europe.²¹ Topic modeling makes clear, however, that the broader field has siloed these conversations since they are not statistically detectable in topic modeling alongside gender relations, the Politburo, post-Soviet politics, or individual centuries of Russian literary history. Our algorithm proposed one topic (#11) that used race as a fingerprint term (“racial,” 19th), but the rest of the topic’s fingerprint (including “reich,” 2nd; “himmler,” 3rd; and “volksdeutsche,” 13th) indicated that the term was only significant in relation to scholarship about the Nazi regime.²² Another topic (#79) pointed toward widespread scholarly conversations about nationalities policy without referencing ethnicity.²³ The gender relations topic showed that terms like “girls” and “husband” must be understood through frameworks like “sex” and “feminism”; likewise, it is possible to write about various nations and their histories without a robust understanding of the fundamental role race and ethnicity play in those histories.²⁴

As topic modeling points to gaps in understanding, it also illustrates how fields can understand identity multidimensionally, as a matter of cultural history, political thought, and current scholarship. The intersection between Slavic Studies and gender studies offers one such example. So does the intersection between Slavic Studies and Africana Studies: an algorithm separate from the one described above pinpointed one topic that centered on Black history, including both historical terms and conceptual terms (such as “race” and “racial”) in its fingerprint.²⁵ This was our data’s only reflection of the

21. Recent Anglophone contributions to this line of thought have come from scholars in a wide range of subfields. In Critical Romani Studies, these scholars include Dušan Bjelić, Alaina Lemon, Sunnie Rucker-Chang, Chelsi West Ohuery, Catherine Baker, Alicia Strong, and many more. In Jewish Studies, they include Marina Mogilner, Eugene Avrutin, and Amelia Glaser. In Soviet and socialist history, contributions have emerged from Eric Weitz, Maria Gertrudis van Enckevort, Kate Baldwin, Maxim Matusevich, Joy Gleason Carew, Meredith Roman, Carole Boyce Davies, Steven Lee, Hilary Lynd, Kimberly St. Julian-Varnon, and others. Jennifer Wilson has examined Russian racial ideologies and the Haitian Revolution; Bolaji Balogun and Lenny A. Ureña Valerio have written on race in Poland. For a summative look at some of these developments as well as new directions, see David Rainbow, ed., *Ideologies of Race: Imperial Russia and the Soviet Union in Global Context* (Montreal, 2019).

22. Another topic, #43, included occasional ethnic designations (such as “roma,” 8th) following terms related to World War II (“vojcehovskij,” 2nd; “truman,” 5th; “wartime,” 7th).

23. The fingerprint for Topic #79 included scattered demonyms and place names related to eastern Europe as well as two framing terms: “nationalisation” (3rd) and “postcolonial” (9th). This topic reflects a tendency in our field to use these terms widely without reference to the perceived categories of hereditary physiological or physiognomic difference (namely, ethnicity and race) that often underlay nationalization and colonization, particularly to the east and south of Moscow.

24. For example, topic #0 foregrounded more than a dozen terms historically related to non-Slavic identity, from “estonian” (9th) and “armenia” (15th) to “sakha” (48th), “romanians” (73rd), and “tatars” (77th). However, the topic’s fingerprint as a whole was centered not on ethnicity but on 20th-century literature and policy, with terms like “yeltsin” (8th), “kgb” (10th), “gorky” (11th), “platonov” (16th), and “perestroika” (49th).

25. See Topic #50 in our 2-gram model at https://github.com/Russian-NLP/Reading-Racial-Discrimination-in-Slavic-Studies-Scholarship/blob/main/1_topic_modeling/SlavicStudiesN3_BERTopic.csv (accessed June 11, 2021).

sectors of Slavic Studies that examined what race is doing in racialized histories (beyond Nazism). To examine how studies of gender and Blackness could shape Slavic Studies on a broader scale, it is necessary to move beyond undirected topic modeling and toward a human-supervised method of classifying texts—that is, toward a digital method that looks specifically for developments in Slavic Studies scholarship on race and gender and analyzes those developments over time.

Perspectival Modeling

In his groundbreaking work *Distant Horizons: Digital Evidence and Literary Change*, Ted Underwood offers a novel method called “perspectival modeling” that makes it possible to quantify long-term change in literary genres on a massive scale.²⁶ This method trains classification models for various periods across time, articulating differences among the periodized models in a quantifiable measure of change. In this regard, the term “perspective” is important. Underwood measures, from the perspective of an earlier time, how texts become increasingly dissimilar and unfamiliar while still being clearly identifiable as part of a common genre. We adapted Underwood’s technique in two ways: by using it to understand scholarly fields rather than literary genres and by enabling it to measure granular change over one short period of time within a single model. This new method pinpoints increased linguistic overlap in recent years between Slavic Studies and two fields that focus on identity: Gender Studies and African American Studies.

Our team trained one classification model for each of these fields.²⁷ The model sorting Slavic Studies texts from Gender Studies ones was able to predict the correct discipline in 98% of its samples (40764 correct, 478 incorrect). The African American Studies model was correct 99% of the time (98074 correct, 1070 incorrect). The high accuracy of these models suggests that real and significant differences exist between Slavic Studies and the other two classes. With these models, we ran predictions for every text in our JSTOR Slavic Studies sample. Nearly all the texts were correctly classified as Slavic, but we also recorded a score measuring similarity to other disciplines. This score can be used to analyze individual works, or it can be tracked over time to measure continuous, non-periodized developments in entire fields.

For example, Amanda Bellows’s *American Slavery and Russian Serfdom in the Post-Emancipation Imagination* is a Slavic Studies title that has a very high prediction score for African American Studies (0.999) and a very low score for Slavic (0.0003).²⁸ This text is a comparative study with equal focus on cultural responses to emancipation in Russia and the US. With such a balance, we

26. Ted Underwood, *Distant Horizons: Digital Evidence and Literary Change* (Chicago, 2019), 67. See also: Katherine Bode, “Why You Can’t Model Away Bias” *Modern Language Quarterly* 81, no. 1 (March 2020): 95–124, Matthew Lee Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana, IL, 2013).

27. The classifiers were trained using the spaCy natural language processing library. Full notebooks for this process can be found in the article’s code repository.

28. Amanda Brickell Bellows, *American Slavery and Russian Serfdom in the Post-Emancipation Imagination* (Chapel Hill, NC, 2020).

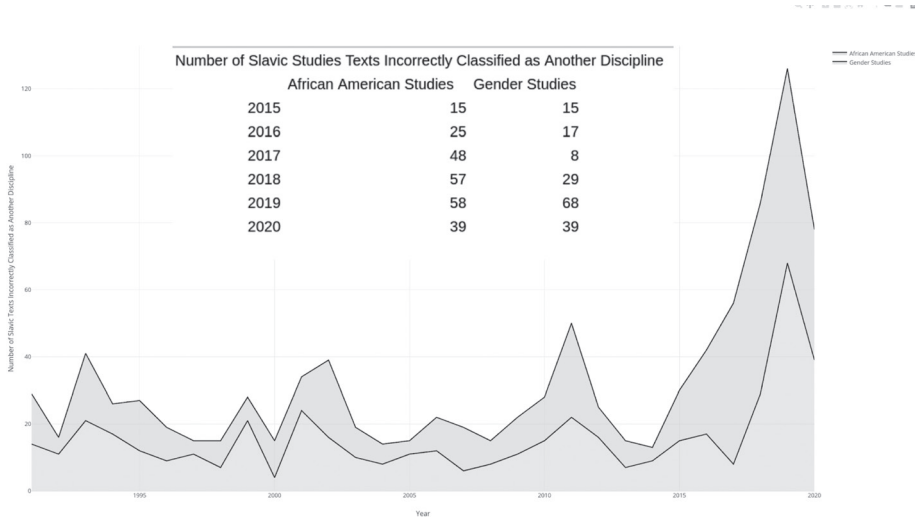


Figure 1: Slavic Studies Texts Incorrectly Classified as Gender Studies or African American Studies

might expect the predictions to be 50% Slavic and 50% African American. However, a prediction score of 0.5 would indicate very high ambiguity. Such a score would show that the text contains nothing very characteristic or distinctive of either discipline. With very few similar works in the corpus, the model has learned that a text with any African American-related subject matter is most likely a work of African American Studies and is highly unlikely to be a work of Slavic Studies.²⁹ (Figure 1).

On the scale of the entire field, however, the ambiguity we might expect from crossover texts does emerge over time. As seen in Figure 2, the model becomes less certain of its classifications for both African American Studies and Slavic Studies between 2015 and 2020.³⁰ (Figure 2). This means overlap with African American Studies has become less of an anomaly in Slavic Studies.

Since 2015, the mean Slavic prediction score for Slavic Studies texts has been declining. This shift occurs at the same time that significantly more texts from Slavic Studies are miscategorized by the models because they bear

29. We can identify terms that most distinguish the African American Studies group from the Slavic Studies group using raw frequency counts or a scaled F-Score. If these terms (e.g., “black,” “tuskegee,” or “liberia”) are included in a work of Slavic Studies, they increase the African American Studies classification score of that text. Similarly, there are terms that are highly distinctive of Slavic Studies. They include any words written in Cyrillic, “dostoevskii,” “aatseel,” and “clitics,” among others. The texts most strongly identified as Slavic (0.999955) are a biography of Dostoevskii, a history of communist Europe, and a history of the Russian steppe.

30. The shift is evident in the African American Studies data as a 3516% increase in the interquartile range (IQR), which is a common measure of statistical spread. In the Gender Studies predictions the IQR changes by 583%. Rather than being clustered near 0 or 1, the predictions between 2015 and 2020 are far more uncertain with more points spread near the middle of the graph (0.5).

Slavic Studies and African American Studies Texts with Classification Score for African American Studies

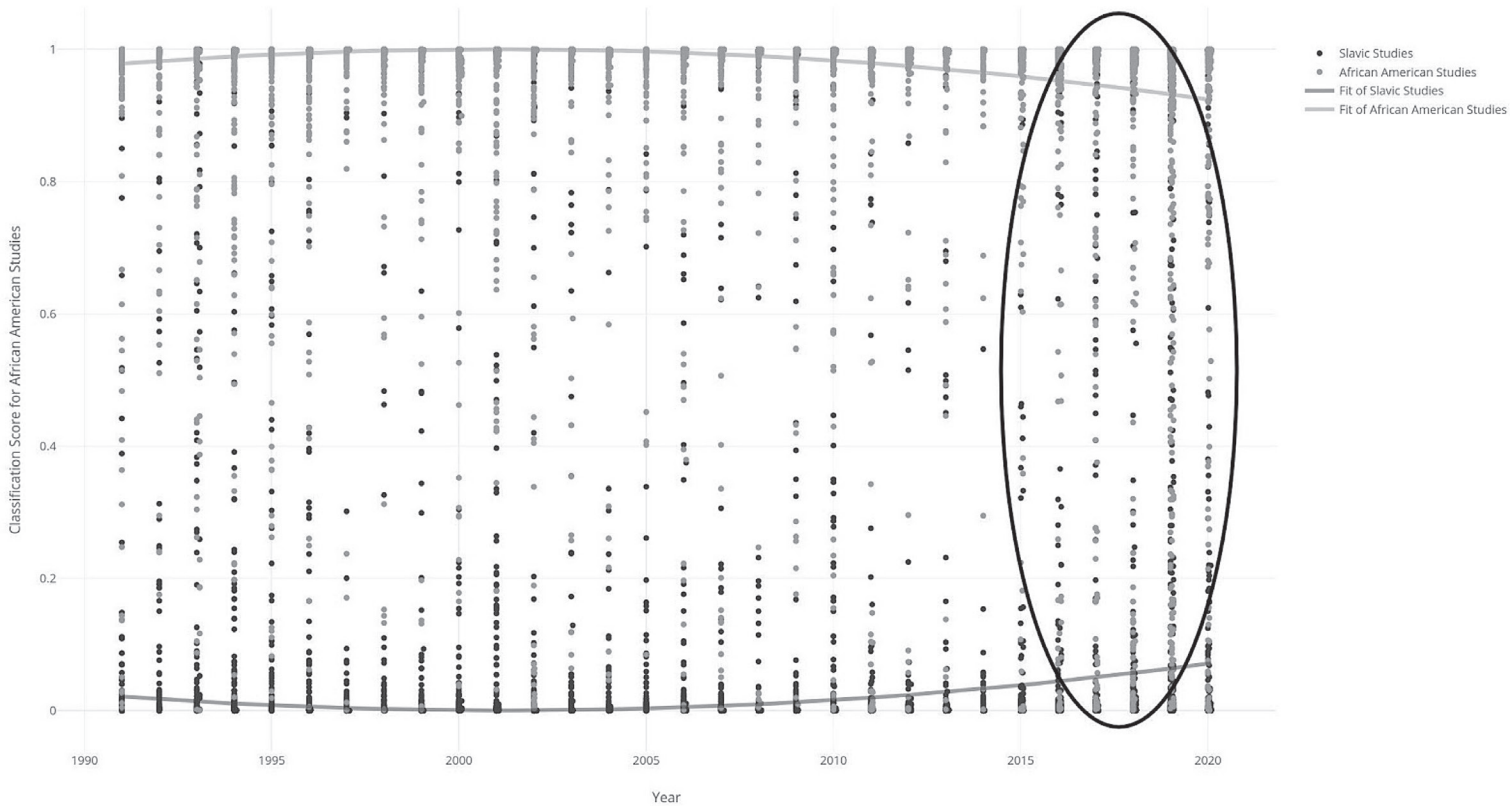


Figure 2: Area of Unusual Ambiguity, 2015–2020

resemblance to texts in other disciplines. This is true for both Gender Studies and African American Studies. Rather than using the model purely as a quantitative measure of change, we found that the models' predictions are most useful as an interpretive tool. If the model has identified a significant change in the field, how might we account for that shift? What has it found?

Looking at the Slavic Studies texts classified most strongly as Gender Studies, there is a clear connection between language and classification. Nearly all the titles are written in Spanish, such as *¿Se puede hablar hoy de populismo en Rusia?* (Can we speak today of populism in Russia?)³¹ Between 1991 and 2020, the Gender Studies sample averages ninety-six titles in Spanish per year. For many years Slavic Studies has had no Spanish titles at all, with an average of 2.8 texts per year.³² In 2001, however, there were thirteen Spanish-language titles; and, in 2018 there were twenty-five Spanish titles in the Slavic sample. Rather than recording "similarity" between fields, the model learned to associate Spanish with Gender Studies to such a degree that any Slavic titles published in Spanish are miscategorized.³³

Perspectival modeling offers a useful method of investigating macroscopic changes over time and to identify significant features of the collection. In this case, a machine learning model has learned, with great accuracy, how to distinguish between works from Slavic Studies and two other disciplines. While the model's internal logic and decision making are not accessible to us, we can nonetheless investigate the model's outputs and interpret its findings. In this case, the model identified a significant change in Slavic Studies scholarship beginning in 2015. The investigation of this shift revealed the significance of language in the model's classifications. Because the machine makes no assumptions about what differentiates scholarly fields, it can arrive at genuine insights that a human researcher might never have noticed.³⁴

Asking where robust discussions about race in the field *do* take place shows our study in yet another light. Students, former students, and non-tenured faculty have written the vast majority of blog posts and articles detailing the structures of racism that constrain who can enter this field and stay in

31. Cf. Vladimir Davydov, *Latinoamérica y Rusia* (Buenos Aires, 2018); Jan Bazant, *Tres prominentes checos: Tomas Masaryk, Eduardo Benes y Alejandro Dubcek: Ensayos biográficos y textos* (Mexico City, 1999); Jean Meyer, "¿Se Puede Hablar Hoy De Populismo En Rusia?" in Guy Hermet, Soledad Loaeza, and Jean-François Prud'homme, eds., *Del populismo de los antiguos al populismo de los modernos* (Mexico City, 2001).

32. The language of each text was predicted using the langdetect Python library at <https://pypi.org/project/langdetect/> (accessed May 3, 2021). Full data can be found in the GitHub repository.

33. There is an equally significant dearth of Slavic languages in the Gender Studies corpus. There is only one Gender Studies article published in Czech, two in Croatian, and one in Romanian. There are no articles in Russian. Given that there is significant Gender Studies scholarship being published in these languages, their absence is likely a sampling problem that reflects the materials available in JSTOR. A keyword search for "Slavic Studies" in Worldcat returns materials in 57 languages. However, 87% of those texts are in English followed by Russian (5%) and German (3%). Fifty of the languages each compose less than 1% of the Worldcat sample. See <https://www.worldcat.org/search?q=kw%3ASlavic+Studies>, accessed on March 17, 2021.

34. Underwood, *Distant Horizons*, 34.

it.³⁵ Moreover, these discussions have taken place almost entirely outside of peer-reviewed scholarly journals, despite our field's affinity for journal-based self-reflection. Perceived bodily difference, race, and Blackness have not been "outside concerns" for *ASEEES NewsNet* or *All the Russias*, either as subject matter for research in Eurasia or as matters that shape Eurasian Studies in the US. These research areas do not show up as prominent subfields in our analysis of journal data, however. Public-facing platforms must be available to scholars in precarious positions; but, when the resulting essays highlight the same patterns of casual discrimination in research advising, study abroad, and other areas for years, it becomes clear that our field cannot keep doing anti-racism and scholarship in two separate forums. The disparity between the robust apparatuses that students and scholars of color have developed for grappling with race through our field's online spaces and the absence of any similarly explicit critical tools in our topic modeling and perspectival modeling results calls for a significant shift in the most established and prestigious institutions of our field. As leading pedagogues and younger researchers pursue anti-racist work, peer-reviewed journals will have the opportunity to welcome and solicit anti-racist research as they have done this year. This shift will not rectify the direct interpersonal aggressions many people of color in the field report, but it will produce a newly proactive and sustained mechanism for grappling with the field's inequities in writing and for recognizing those who do so.

On the path toward that mechanism, we consider our multi-pronged digital approach to be but a first step: a nimble offering in response to our present moment in a time of crisis. Between this article and our accompanying GitHub repository, our hope is that scholars can replicate our dataset and use it for further research. We implore other scholars to build upon our pilot study, analyzing the publishing history of prominent Slavic Studies journals from a range of angles that may help this field reach more ethical practices and norms.

35. See articles by Rachel Stauffer, B. Amarilis Lugo De Fabritz, Amber Casandra Walden, and Kristin Torres in *AATSEEL Newsletter* 58 no. 3 (October 2015) at https://www.aatseel.org/100111/pdf/aatseel_newsletter_october_2015.pdf (accessed May 3, 2021); B. Amarilis Lugo De Fabritz, "Race, Diversity, and Our Students in Russia," *NYU Jordan Center* (blog), August 21, 2013 at <http://jordanrussiacenter.org/news/race-diversity-and-our-students-in-russia/> (accessed May 3, 2021); Aisha Powell, "Black Bread: A Look inside the World of Black Slavic Studies Scholars," *Trumplandia Magazine*, December 8, 2018 at <https://trumplandiamagazine.com/black-bread-a-look-inside-the-world-of-black-slavic-studies-scholars-7579c6b2d5cd> (accessed May 3, 2021); Jennifer Wilson, "Is Slavic ready for Minorities?" *NYU Jordan Center* (blog), July 22, 2014 at <https://jordanrussiacenter.org/news/slavic-studies-rationally-tone-deaf/#.X20jM31Khv1> (accessed May 3, 2021); Kimberly St. Julian-Varnon, "A Voice from the Slavic Studies Edge: On Being a Black Woman in the Field," *ASEEES Newsletter*, September 2020 at <https://www.aseees.org/news-events/aseees-blog-feed/voice-slavic-studies-edge-being-black-woman-field> (accessed May 3, 2021); Sarah Valentine, "The Divine Auditor," *Prairie Schooner* 87, no. 2 (Summer 2013): 91–104; Emily Couch, "Beyond Diversity: Integrating Racial Justice into REECA Studies," *ASEEES NewsNet* 60, no. 5 (October 2020): 11–13 available at <https://www.aseees.org/news-events/aseees-blog-feed/beyond-diversity-integrating-racial-justice-reeca-studies> (accessed May 3, 2021).