

Cracking Open the Black Box of Genetic Ancestry Testing

Joan Donovan, PhD
Data & Society
Research Institute
joan@datasociety.net

Irene Pasquetto
University of California,
Los Angeles
irenepasquetto@ucla.edu

Jennifer Pierre
University of California,
Los Angeles
jp639@g.ucla.edu

Abstract

Stormfront, a well-known online forum for white nationalists, is a place for discussions about race, nation, and biology. We analyzed how members shared and discussed genetic ancestry tests (GATs), which revealed a complicated network of boundary maintenance, identity formation and justification, and biosociality within this online community. Using selection of seventy Stormfront threads discussing GAT results, this study employs primarily digital ethnographic methods to investigate how white nationalists navigate questions of self and community online. Using scientific concepts, genetic data, and multiple databases, white nationalists rely on the ambiguity of genetics and the black boxing of algorithms provided by testing companies to redefine white identity while also remaining committed to biologically-informed conceptions of race. This research raises important questions about the role of scientific data in racial formations.

1. Introduction

In January 2017, Ancestry.com announced that they exceeded three million DNA samples in its AncestryDNA database [1]. In part due to new marketing strategies in 2016, including a Black Friday sale and a series of television commercials, Ancestry.com sold direct to consumer genetic ancestry tests (GATs) in anticipation that consumers would also subscribe to their collection of genealogical materials including family trees, census, and research documents. While some consumers take GATs to learn health information, many also take them in exploration of their genetic pasts or genetic identity. Rabinow [16] explains that genetics, especially advances in personalized genomics, will bring about new social relationships based on shared genetic traits and biological conditions. This “biosociality” has implications for how humans will organize institutions and society as biological understandings of humans surface in all domains of

life including work, the family, education, and our legal, and political systems. Our research attends to these concerns in the online community setting of Stormfront by asking how does one enact a biosociality after genetic ancestry testing?

One way to understand this is to look at how GATs are marketed. A TV commercial titled “Kim” for the AncestryDNA kit, explores how a customer may react to their results. The scene opens with a close-up of Kim and later her test results appear in a pie chart next to her.

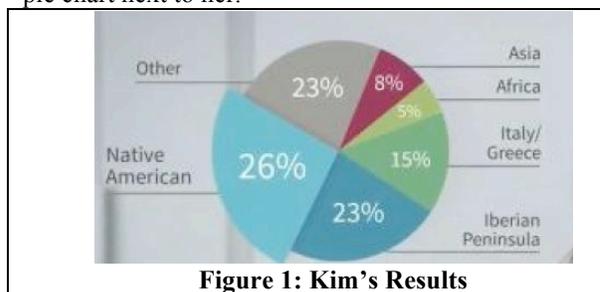


Figure 1: Kim's Results

Kim: “I wanted to know who I am and where I came from. I did my AncestryDNA and couldn't wait to get my pie chart. The most shocking result is that I'm 26% Native America. I had no idea. Just to know this is what I'm made of. This is where my ancestors came from and I absolutely want to know more about my Native American heritage.”

While Kim does not explain why she is “shocked” by some results and not others, the scene cuts to reveal that Kim is inside a Native American museum. The commercial ends with a feminine voiceover: “Discover the story only your DNA can tell. Order your kit at AncestryDNA.com.”

There are many visual metaphors at work in this commercial. The whiteness of the background invokes the sterility of the museum, while also indicates a washing out of Native culture entirely. TallBear's work on DNA and Native Americans illustrates a complex relationship between Natives, colonization, and western medicine [17]. This work describes how American anthropologists and scientists have historically treated Native artifacts and DNA as their own property. In this way

AncestryDNA's commercial is a continuation of the dispossession of Native Americans, both in physical and biological terms, while also provides a biosocial explanation of how to interpret GAT results for yourself. In short, these GAT tests are tools for biosocial colonization.

How does one resist GATs or interpolation into this genetic biosociality? The simple answer is not to take the test. However, this does little to provide social groups with the footing to define their own collective identity. As Reardon and Tallbear show, the many anthropologists and scientists will go on with or without Native consent. As this commercial for AncestryDNA illustrates, so too will the companies who offer direct-to-consumer GATs [17].

Inspired by this research, we shift the lens to describe how a pernicious biosociality emerges from GATs. We ask: in what ways are biological markers used by white nationalists to fashion a collective white identity? Overarchingly, how do these processes particularly unfold in within a digital social media based space? By illustrating how racist white nationalists use the tools and techniques of GATs to define white identity, we hope to illuminate the colonizing logics of GATs as they heighten biological ambiguities, while at the same time reduce culture to percentages.

We begin by illustrating how white nationalists set boundaries around whiteness within their online community. Unexpected assertions arise around the role of phenotypic attributes of whiteness when setting those boundaries. Then, we illustrate how the ambiguity of genetics as a science, the black boxing of commercial databases, and the mutability of genetic data allow white nationalists interpretive flexibility in assessing, accepting, or rejecting their GAT results. In sum, we pose the following overarching question to guide and frame our research: to what end does GAT support white nationalists' beliefs and values about race, nation, and biology? How does this process unfold in a major white nationalist social digital media platform?

2. Literature review

We view Stormfront, our chosen platform for observation, as a digital social media for white nationalists. Social media, though a rather nebulous term, can be most consensually defined through the various key attributes scholars have identified over the years, including an emphasis on relationship building, creation and exchange of user generated content, and the inclusion of profiles in web-based platforms. The forum style presentation of Stormfront

serves as one subcategory of this genre of online communication platform. The emphasis on the use of such platforms for relationship building and friendship maintenance is significant for understanding the role of this particular context in motivating certain boundary setting practices and GAT interpretations [7].

In our analysis of white nationalist use of GATs to enact forms of biosociality, we witnessed a divide in interpretations of results and tactics of identity and community formation. Particularly, the division between reliance on older methods of phenotypical categorization of race clashed with genotypical classifications. We consulted literature in the areas of race and biology and genetics and data to trace the origins of these schools of thought and to frame the biosocial processes we observed.

2.1. Intersecting race and biology

Before GATs, there were multiple other ways for people to define and make sense of race. Today, biological indicators of race are widely adopted in scientific, medical, and forensic research [4]. These indicators developed beliefs in phenotypic traits as ultimate deciding factors of race, and reinforced ideas of phenotypical superiority. For example, in the US race was inscribed in cultural practices through legal standards set by the "one-drop rule" that forbade intermarriages in some states.

Scientific and specifically biological explanations of race became prevalent from the mid 19th century onward. Omi and Winant [12] cite the publication of Darwin's *The Origin of Species* as a key text influencing the development of racial theories based on science and heredity. Gould [6] describes this line of thought as a larger trend of biological determinism. He illustrates that throughout the 20th century racial categories were endemic within biological and bioanthropological work. Anthropology grew in status as a rigorous scientific discipline by specializing in biological inquiry, and in turn perpetuating beliefs in racial hierarchies as rooted in science. Despite UNESCO declaring that race was not biological in 1950, scientists continued to use race as a tool and category, although for different ends.

In response to increased rejection of phenotypical markers as indicators of race and the rise of theories on the social construction of race, sociologists focused on "racial formation theory" which emphasizes how race is real in its consequences [12]. Omi and Winant define race as a concept "which signifies and symbolizes social conflicts and interests by referring to different types

of human bodies” (p.55), and explain that it “has no meaning, but is constructed and transformed sociohistorically through competing political projects through the necessary, yet selective link between the structural and cultural dimensions of race in the U.S. (p.71).”

2.2. GATs, bioregions, and race

Thousands of people purchase GATs per year, and hundreds of companies market these tests around the world. Generally, consumers use the results either to validate genealogical records or to find genetic connections to specific groups, populations or places in Eurasia and Africa [2]. Test takers have a variety of companies to choose from including 23andMe, Family Tree DNA, National Geographic, and Ancestry.com, which provide numerous GATs such as autosomal DNA (your combined family lineages), mtDNA (maternal lineage), and/or Y-DNA tests (paternal lineage, only available to males). While using a quite vast array of different markers and parts of the DNA, all GATs group individuals based on genetic similarities. Indeed, these tests rely on the idea that some individuals have similar mutations on their gene bases that are mapped on to bioregions, locations where these mutations are commonly found. These mutations are referred to as “genetic markers” and can be found in different chromosomes or parts of the DNA. “Genetic markers” can serve as locus of marking differences in racialized terms, but population geneticists have been careful to remain “classificatory ambiguous” when publishing findings, where geneticist shift and blend between race, heredity, region, and migration [13]. Moreover, “dwelling in ambiguity about how to classify and compare populations allows geneticists to do their work without presenting a hard target for criticism” (p.81). Later, we illustrate how white nationalists use classificatory ambiguity to support their interpretation of GATs.

Nelson [10] classified GATs according to the types of “claims” companies deliver to consumers. For Nelson, these technologies fall into three categories: ethnic lineage testing, spatio-temporal testing, and racio-ethnic testing. Nelson describes as *ethnic lineage tests* as the use of both mitochondrial DNA (mtDNA) and Y-chromosome DNA (Y-DNA) to infer ancestral links to contemporary nation-states or cultural groups. In this case, the consumer’s DNA is compared to a reference database of genetic samples. *Spatio-temporal tests* employ consumer’s DNA sample to classify it into a haplogroup according to schemas of ancestral and geographical origins. *Racio-ethnic composite tests* aim to make

claims about one’s ancestry and involves the study of nuclear DNA, which is unique to each person. The necessary assumption behind all GATs is that, because mutations are (partially) genetically inherited, if individuals have specific mutations it could mean they are related to a population or group.

Commercial GATs have been heavily criticized by scientists and social scientists for the low reliability of their scientifically sound results [5, 19]. These critiques mainly focused on showing the uncertainty of the scientific validity of admixture research practices upon which GATs are built [2]. In relation to this set of concerns, researchers highlighted the sampling limitations of GATs. GATs tend to employ samples from limited sub-groups and generalize findings to extended populations (like in the case of West Africans for all African population) [10]. A second limitation is the lack of scientific evidence supporting the biological discreteness of groups of humans. As demonstrated by the HGP project, genetic diversity exists within populations and gene flow occurs between populations [4]. Very few alleles are diagnostic of membership in a specific population. Furthermore, researchers pointed out, racial and ethnic categories are correlated with geography, historical factors, and migrations, and databases of present-day samples that do not properly take these limitations into consideration may therefore provide false leads [9].

A second set of concerns refers to the fact that most of the technologies used in GATs, and, more in general, in admixture research, are black-boxed, i.e. the process by which results are derived is not open to inspection by the public or even by other scientists [14]. While STS scholars widely focused on unveiling the scientific inaccuracy of admixture research and GATs, not much has studied on this black boxing of GATs. In this paper, inspired by work done by critical data studies scholars [3], we aim to bring particular attention to this lack of transparency in relation to the proprietary algorithms and databases used to analyze, visualize, and interpret GATs. Because most algorithms and databases used to interpret GATs results are “proprietary,” which means they are owned by companies, we do not have access to their content, making it impossible to gain a full understanding of how these tests generate results or how these scientific practices get mixed up with racial identities.

As we will show, when white supremacists are not satisfied with their genetic profiles (i.e. they think these are not “white” enough), they tend to move their raw sequence data (provided by the companies) into alternative online databases, such as GEDmatch.com, which provides “DNA and

genealogy tools for comparison and research purposes.” These alternative analyses are based on alternative sample populations and algorithms (called “calculators”). These alternative databases are also black boxed and not available to scrutiny. White nationalists sometimes mistake this lack of transparency as part of the classificatory ambiguity of genetics. While publicly funded research is increasingly made available through open archives and repositories of data, this type of pressure to open databases does not apply to privately owned biomedical data [8].

Scholars showed that while genetic testing is initially perceived by test takers as a possible objective solution to uncover “truer” identities, in practice, once the test takers are exposed to “unwelcomed results” the same individuals employ multiple strategies to repair their newly discovered DNA-based racial identities [9]. Nelson argues [10] that test takers find themselves in a state of “genealogic disorientation,” they feel lost, confused, and numb after they took the test (p.84). As a result, they often start new investigations to re-interpret GATs results by incorporating them into their own collective biographies.

To sum up, many different types of GATs exist, which employ different populations, genetic markers, and calculators to construct racial profiles. GATs are troublesome technologies due (1) to their uncertain scientific validity and classificatory ambiguity; (2) the black boxing of their algorithms, and (3) because they cause genealogic disorientation for the test taker. We show how these conditions grant white nationalists liberty when interpreting genotypical markers of white identity, thus supporting their worldview that races are biologically wedded to nations.

3. Methods

Beginning in 1995, Stormfront is a digital social media for white nationalists. Gaining more notoriety during Trump’s election, Stormfront is a text-based, publicly open site, indexed across many major search engines. It reports over ten million page views per month with over 300,000 registered members and a corpus of 12 million posts. For our study, we selected seventy threads where members posted their GAT results. Then using qualitative content analysis software, we coded the 3,070 posts within these threads for discussions about the meaning of test results, managing reputation after disclosing results, and knowledge about genetic testing companies and their databases. Using posts dating from 2004-2016, we tracked and coded conversations that related to reputation management and open data to understand

how genetic representations of European DNA affect conceptions of white identity. The data presented in this paper are representative of this larger sample.

4. Findings

4.1. Biosociality & boundary work on Stormfront

Establishing, maintaining, and upholding a sense of community is a significant but tenuous process on Stormfront. Practices of community boundary maintenance directly connect to the formation of a Stormfront white nationalist identity, a similarly nebulous concept that requires constant defense, protection, and consensus gathering to sustain. An overarching theme we observed within this space is the divide between strictly phenotypical versus genotypical understandings of race to enforce the boundaries of the community. The debates over which way to best utilize biological concepts and data to reinforce beliefs about race reflect a movement from simplistic phenotypic biological indicators of race to the use of genetic ancestry data as a more abstract connection between ethnicity and race.

In examining the perceived and applied parameters of white nationalist identity on the site, several unexpected patterns emerged around how the community defines itself. These patterns emerge most clearly through the observation of how “trolls,” unwelcome and un-trusted contributors to the site, are identified, reacted to, and eventually exiled. The practices of inclusion and exclusion of trolls on Stormfront reveals the boundaries of the community set by core members, as well as their larger biosocial relation to whiteness. Ultimately, an analysis of the general interrelated process of boundary and identity formation and maintenance in the Stormfront conversations around race and genetics reveals an underlying set of illogics and contradictions that preclude collective community boundary and identity desired. Because phenotypic indicators of race are contested, some white nationalists favor using genetic ancestry data as a stronger body of evidence for reinforcing racial beliefs.

4.1.1. Rejecting “White” Supremacy

Members of Stormfront quickly identify and expel trolls who use stereotypically phenotypic-based extreme racist language on threads, especially in relation to Nazi inspired concepts. Posters who exhibited explicit Nazi-sympathy were met with hostility from other members, who labeled them

trolls for claiming that only Aryans were truly white nationalists. To troll means to derail, insult, or ridicule for amusement [15]. For example, in the “How can I know if I’m racially normal?” thread, the expression of Aryan, German, and Hitler sympathies and a heavy reliance on phenotypic qualities as indicators of race was quickly marked as trolling and met with negative exclusionary responses.

THREAD: How can I know if I'm racially normal?

Original Post: GA

I live in Minnesota, I'd like to think of myself as an Upper Midwesterner from the frozen lakes. But how can I know how well I "mold" into the demographic? I can say with certainty that most of my ancestry is European, probably Anglo Saxon, but I think there is a small part that is from elsewhere, like Asia, I'm not sure how much, I would probably have to go back centuries before I find an ancestor who immigrated to Europe from the Asias. The reason I think that it might be possible I have small percentage of exotic heritage is because I have had 23andme done for my family. When you do 23andme, the results directly from 23andme are bull****, if 23 says you are 100 European that is not enough and you have to look at GEDMATCH for the real answer. (...) Most of my family members had dark hair, which would mean that some of my ancestry had to be from not European people since Anglo Saxons always had brown, red, and blonde hair. Maybe this preceding sentence isn't 100% true, but at school everyone has light hair and darks like me stand out. (...) I also should tell you that many many years I have felt depression as a result of this. I want to be German, and I would rather be German than anything else because Hitler told everyone they are superior to other Europeans.

Response: WBB
Dark hair happen naturally in Europeans.

Response: RHW
You're a troll or an idiot, or both. I was going to try to add more to be sure the mods would let this through, but I think this clearly says exactly what it means to.

Response: HG
Keep in mind that light eyes and hair is a MUTATION, that happened to occur within the European White race. Whites beforehand had dark hair and dark eyes. Just because you have brown eyes and dark hair doesn't mean you aren't White. Also, most southern Europeans are White. I'm a classic example, I am 100% Greek/Italian... It also potentially means that you have a lot of Slavic blood

in you, and Slavs have less Celtic/Germanic features, but are still pure White.

The denouncement of the original poster’s belief in a limited phenotypic definition of white purity is swift, and exposes the belief of a later more fully explained diversity of whiteness as one of the standards of belonging and identity in the white nationalist community. The group in-fighting exposes layered socio-political aims infused in explorations of race and nationhood, where allusions to Hitler’s Germany are considered to be mocking of the real political project of white nationalism. While the goal of the Stormfront white nationalists as a whole is to reinforce ideas around the racial superiority of whiteness, a more detailed hierarchy that imposes ranks within the scientifically determined Caucasoid division elicits anger and rejection of the original poster’s ideas. The responses dismiss the variations in whiteness that the original poster discusses as merely acceptable and normal mutations within the white race, but do not extend this line of logic to claim that the existence of such random mutations resulting in phenotypic variation to undermine the legitimacy of racial superiority. As the authors reveal, the use of scientific concepts to promote racial beliefs is muddled by preexisting socio-political aims, and leaves such arguments unstable. As such, some Stormfront members reject phenotypical markers of whiteness, especially when used as indicators of racial superiority.

4.1.2. Embracing white diversity

In a continued relay of conversation where GermanAmerican1989 responds to the attacks and voices preference for phenotypical purity and German superiority, the community continues to collectively disavow him, confirm their categorization of the poster as a troll, and actively practice boundary setting through continued suggestions that they leave the community. This segment of the thread begins to more specifically expose the tenuous definition of whiteness Stormfronters rely on, as demonstrated by LBG’s post, where he concedes the contradictory nature of his belief in white purity relative to other races versus arguing for the existence of European diversity within “one white race.”

THREAD: How can I know if I'm racially normal?

Response: GA

Then why do so many people have light hair/features and few are dark? I went to Germany this year and could not find people who had dark hair they were all blonde. This must mean that if you have dark hair,

you are getting it from an ancestor who was not German. (...)
Response: LBG No wonder I believe in pan Europeanism. lol Sorry brother, Germans are no more superior than any other whites. We are all the same. :p damn I take that back because that is what anti- whites say. :rofl ...Only one white race so join the party...
Response: DB The OP is a persistent troll who has been posting the same old drivel for months on end. I wish he would find a new hobby like jumping off a cliff without a parachute or something. What a crashing bore he is.
Response: UM Germans hold the wonderful genetic diversity uniquely common to the white race :).
Response: GA Germans are the only race in the world (even among the European races) that have ever contributed anything worthwhile to the world though. And other people are German, so it is the gold standard, if you are below the average for % German heritage you are subhuman.
Response: BC Aww, how cute. Somebody got a hold of mom's computer. Better get to class before your 5th grade teacher calls and says you didn't show up again.
Response: UM It is a definite now that this poster likes playing games with us here. You know the saying folks, don't feed the trolls!
Response: GA Its really what I believe though, and I have a right to share it
Response: DB You're a time waster. F off. :rolleyes:

Here, scientific concepts like genetic diversity are used to refute the original poster's claims about the superiority of white haired, blue eyed, and specifically German people. Similar to the previous example, however, they are applied in a narrow and selective sense in order to make claims of "pan-Europeanism" and the existence of sub-races simultaneously. The understanding of variance in biological phenotypical features is only understood and discussed within the pre-existing framework of racial hierarchy that is rejected by those on Stormfront that seek the unification of whites as a political project. Moreover, the ambiguous classification of whiteness is combined with the tactics of reputation and authority of established posters, such as DB, to label GA as a troll.

The processes shown and analyzed above relays a collective articulation of "white diversity" as a factor of white nationalist identity, demonstrated by the immediate harsh treatment of posters whose views seem to sympathize with a narrower Aryan German-centric definition of white purity. Such a staunch defense of whiteness in a broader sense through a concession of at least some form of diversity in whiteness directly contradicts the collective behavior demonstrated against posters who identify themselves as mixed race. As such, the community and identity boundaries are drawn within those deemed to be fully white nationalities, not just those who appear to look white. This hard line is most clearly defended in the following thread, where an original poster, RO, was criticized for sharing his DNA test results.

THREAD: 61% white as per my DNA test
Response: RO Hello, got my DNA results and I learned today I am 61% European. I am very proud of my white race and my European roots. I know many of you are "whitter" than me, I don't care, our goal is the same. I would like to do anything possible to protect our white race.
Response: FL OP, I've prepared you a drink. It's 61% pure water. The rest is potassium cyanide. I assume you have no objections to drinking it. (You might need to stir it first since anyone can see at a glance that it isn't pure water.) Cyanide isn't water, and YOU are not White.
Response: PWR The accepted background "noise" on DNA tests is 5%, I will give you that. (...) The best thing you can do to help the white race is to never breed with a white person...
Response: TDV Troll or ****tard, it's a mystery.

As demonstrated in the thread above, the "there is only one white race" logic applied in the previous thread no longer applies to RO's mix of European, meaning white, and what is considered distinctly not white, which he will not reveal. Here, the response is exclusion. While GA was perceived as trying to define the white race too narrowly, similar arguments are used the same way against RO who is trying to define the white race too broadly. The white nationalist identity and the tactics of community boundary policing used to confirm it continue here to expose the unequal footing of phenotypical and genotypical expressions of whiteness. The supposed phenotypic diversity in and natural mutation in whites that could support various types of whites does not extend to RO. Because Stormfront is a

message board, biosociality relies on self-disclosure, but much is still up for debate as the social here is mainly political.

4.1.3. The move toward genetic data

The question of white diversity and the discrepancy in interpretation of phenotypic and other forms of biological diversity is also due to a more stable scientific body of evidence provided by GATs, which is presumably able to outweigh reliance on phenotypic descriptions altogether. We argue that a progression towards genetic ancestry data is being used here as a new basis for protecting the white superiority narrative and alleviating the volatile grasp Stormfront members have on the concept of whiteness and their own biosociality. It is a reflection of Omi and Winant's [12] claim that the nature of race and racial projects is ever changing, where the ambiguity genetic data and its multiple meanings are also in flux. This means that GATs do not stand in an objective relation to biology, heredity, and migration, but are the outcomes of racial formations that lend credibility to many different collective identity narratives, be they Black, Native American, or white nationalist and so on.

4.2. Biosociality & identity work on Stormfront

As discussed in the preceding section, for Stormfront forum users racial phenotypic information is a contentious indicator for establishing community boundaries between who can and who cannot be white. This ambiguity justifies a transition toward more "reliable" mechanisms for establishing whiteness and community such as GATs. We were particularly interested in finding out how white nationalists react when GATs return non-white results. We observed how some white nationalists either change their position about the link between biology and race, and consequently to negate that race is simply a matter of genes, or they refute GATs as a reliable method to determine race. However, we identified a group of forum participants who did not engage in either of these strategies. In this section, we show how these test takers rely on the concept of race as genetically determined and how they hold on to it even when their GATs results seem to reveal the opposite.

4.2.1. Moving your data

When faced with non-European test results, this group of test takers looked for alternative

interpretations of GATs by uploading their DNA raw data to third-party databases, which re-calculate racial percentages using black-boxed calculators. Examples of these alternative calculators include deCODEme, GEDmatch, Dodecad K12b, myHeritage, and Dr. McDonald's BGA project. These platforms do not sequence clients' DNA, but they analyze users' genetic profiles by matching them with profiles they collected over time. Unlike proprietary companies, groups of volunteers are behind the organization and maintenance of these databases. Stormfront members consider these calculators "politically independent," and consider them to be trustworthy. In the following examples, posters recommend alternative databases to those who self-disclosed non-European results.

THREAD: 23andme

Response: AC

I suggest looking at Gedmatch also. It lets you upload to their system the raw data file from 23andMe (or some other sites). Once uploaded, you can run your data through a number of calculators, and find DNA matches. (...) I find it is best to use multiple estimations and explore via services like Gedmatch. For me, in the past 500 years, I'm pretty sure I'm 99.7%+ European.

THREAD: Family Tree Myths

Response: ASR

You can try doing a DNA test to see if your ancestor was white or Indian, here is a link to a DNA company <http://refer.dna.ancestry.com/s/2smeh> He should go with a company that support GEDmatch like Ancestry and 23&ME then he can upload Autosomal results with a RAW data file.

The GEDmatch.com site used to be free, but now requires a \$10 dollars monthly-fee. Gedmatch offers a wide range of utilities that supposedly makes it easier for test takers to extract every bit of potentially useful racial information out of their autosomal test results. To use the site, users need to download raw autosomal DNA test results provided by the testing company and then upload them to Gedmatch, by following their directions online. While GEDmatch provides a lot of information about how to upload the data, no information can be found about how these alternative calculations are conducted. GEDmatch does not disclose what kinds of populations are represented in the database or provide information about the algorithms used to link profiles. However, what we do know is that, according to white nationalists, these calculators tend to retrieve results that minimize genetic disorientation and tend to provide "whiter" results than other companies.

4.2.2. Conspicuous consumption of GATs, reference populations, and racial categories

As pointed out earlier in this section, GEDmatch is only one of multiple third-party alternative databases. We tried to understand what drives test-takers to move their data in these alternative databases. We asked: given the vast array of platforms, calculators, and testing companies available, how do white nationalists establish which one to trust? How do they determine accuracy when faced with black boxes?

We found that test takers advocate for different calculators for a variety of reasons. For example, certain platforms are considered better indicators for “racial profiling” based on the *type of test* that they conduct. Some users considered y-DNA a more valid way to define racial categories than autosomal DNA. Significantly, we found arguments in favor and arguments against for all types of tests, y-DNA, mt-DNA, and autosomal DNA. In what follows we provide few examples.

PRO y-DNA, mt-DNA tests
<p>Response: QFT Yes, familytreedna is a good place for DNA testing. They don't use the same technics as ancestrybyDNA. AncestrybyDNA (DNAprint) use autosomal DNA to look for percentages of racial admixture while familytreeDNA only use Y-DNA and mtDNA to trace direct paternal or maternal lines to a haplogroup. Using Y-DNA and mtDNA is very accurate since that DNA never changes from parent to child. From what I read at the FamilytreeDNA form they at one time used autosomal DNA testing provided by ancestrybyDNA, but, then decided to drop the test and only use the yDNA and mtDNA testing.</p>
PRO autosomal DNA test
<p>Response: JC I encourage posters and test-takers here to also do a STR autosomal test along with a SNP autosomal test. with a STR autosomal test you see much more detail, and localized populations who you match. you will also most likely see exactly where that "0.2%" supposed Ashkenazi really come from. It's most likely due to overlap, since the percent is so small, and not from any actual Jewish/Ashkenazi genetic inheritance. It's too small, to be that.</p>

Another recurrent factor taken into consideration by members to evaluate a platform is the *level of diversity of the “reference population”* in the databases. Databases with higher rates of white/Caucasian DNA samples are - not surprisingly

- considered more reliable than databases with more mixed populations.

THREAD: DNA Surprise and Question
<p>Response: GST (...) I notice a lot of white people are showing subsaharan now, that just means because the price went down more blacks are doing the test and it's skewing the results as most black Americans have European ancestry.</p>
<p>Response: SVR The database of people in 23andme is full of Ashkenazi Jewish while deCODEme is more Germanic and central European. They have a great range of reference people from all over the world as well.</p>

White supremacists also want their results as specific as possible. The *level of granularity of the racial categories* retrieved by the platform seems to play an important role as well in the evaluating process. Higher granularity is generally seen as a factor to prefer a calculator to another. Similarly, members have high expectations over technologies the use of GWAS (genome-side association studies) into GATs.

THREAD: DNA testing
<p>Response: OM The company, “Ancestry By DNA,” specializes in finding out what your racial makeup is, using a triangle of 3 races (East Asian, African, European). http://www.ancestrybydna.com. Ancestry By DNA can even find out which European type you are. Through the regular DNA test, you can get another special test, that will analyze your DNA even more, called the EURO-DNA test.</p>
THREAD: 23andme
<p>Response: OT My results: 20.1% Northern European 12.2 % Eastern European 66.1% Nonspecific European :mad: 1.6% Unassigned :mad: I was also disappointed that the UK was pretty much all lumped together. I have a strong Northern Welsh family history, but they don't break it up like that, it's just “UK”. I've been fooling around a bit with gedmatch.com to get more details, but their services are not fully functional at this time. Have you uploaded your file there? I spoke with someone quite a while ago who helped to analyze these results (McDougal?) and he said the "unassigned" portion was probably just garbage DNA.</p>

To sum up, we found that this group of Stormfront users, when faced with unwelcome results about their racial profiles, they either move their data to a different platform and have it recalculated or

they try a different test or testing company. After analyzing over 100 posts on this specific issue, we did not identify a coherent strategy where white nationalists consistently trusted any platform. In sum, all platforms, tools, and calculators had some flaws. As well, not all calculators and tests tell the same racial story, but these calculators and tools often retrieve multiple and contradictory national, ethnic, and racial categorizations.

5. Discussion

As drawn from the discussion of boundary setting and bisociality on Stormfront, the white nationalist identity and the tactics of community boundary policing expose the unequal logical footing of phenotypical and genotypical explanations of race. The contradictions around what or who counts as white and how to define and defend it entreat many questions around whether such consensus actually exists, and further if the community boundaries actually succeed in defining a collective set of standards and expectations of identity. In the expressed opposition to any attempt to too severely limit the definition of pure whiteness and the associated boundaries of appropriate proponents of the white nationalist agenda, an ill-defined threshold between too narrow and too broad on the parameters of white nationalist identity emerges.

It is difficult to pinpoint the exact reasoning for various interrelated views on such boundaries along the spectrum, and to what extent consensus around them exists. However, it is useful to identify the process of crafting and setting parameters and limits to the boundaries of such an unclear concept as white identity, at least as it exists for those who seem white nationalism as a viable political project. By looking at debates about whiteness, the rejection of Nazi-sympathy, and how they expel trolls, we showed that they struggle with particular definitions of white identity and how GATs could potentially fill the void.

Though many questions remain, the observation of patterns around language use and the elicited responses from members of Stormfront are evidence of community and identity boundary formation and more broadly biosociality at play. Additionally, these observations shed light on the consideration of boundary formation and maintenance by Stormfront's members as a significant opportunity for sustained growth and strength of the white nationalist movement more generally. Significantly, these practices of boundary setting are also constantly defended, negotiated, and re-articulated because of the ambiguity at the core of white nationalist identity

and community, and the processes of upholding and maintaining these concepts continue to expose and confirm its instability. As Panofsky and Bliss [13] illustrate, classificatory ambiguity enjoins scientific authority. In our case, it also allows white supremacists the ability to pivot as new racial formations present themselves.

Most importantly, and as the larger argument in our paper reveals, this resistance of more traditional or stereotypical understandings of whiteness and white nationalist identity as most significantly tied to phenotypical elements may be in part due to a new comprehension and assumption of such narratives as too simplistic. This assumption subsequently forges a path for the search for more complex methods and tools for legitimizing beliefs about white identity, for example through scientific methods and tools that can connect racial identity more complexly to genetics.

In the second part of the result section we described how white supremacists reacted when faced with unwelcomed (non-white) GATs results. While we observed some white nationalists react by either modifying their views on genetically-determined whiteness, or by dismissing GATs as an unreliable tool to determine whether genetic material, some of the test takers moved their data in an effort to force different results. By moving their data, they keep their white identity intact and at the same time reaffirm the paradigm of a biologically-determined white race.

By moving their data to an alternative database, white nationalists can obtain a new genetic self, a secondary and alternative racial identity that better fits their expectations and conforms to the biosociality of other white nationalists. In short, our first overarching observation is that different databases, calculators, platforms, and types of tests seem to retrieve different ancestry profiles. Alternative tools enact alternative identities. Social scientists are increasingly invested in showing how population genetics suffers from biases, misinterpretation, and can even be intentionally misused [9]. What is relevant for this paper is that such cloudy research materializes in a variety of forms in different GATs tools, platforms, calculators, and tests. We believe that these two conditions, (1) the uncertainty surrounding admixture research, and (2) the multiplicity of tools, platforms, tests, and calculators available, are what allow white nationalists to repair their non-European GATs results without losing faith in a biologically-determined concept of race or in the validity of GATs. This is how, independently of how we think and may discuss race in scientific setting and

population research, it becomes real in practice, outside the research environment [12].

Because we do not have access to the users' data, or to these alternative calculators, we cannot verify whether different tools retrieve different results. This fact is at the core of the problem that we are discussing in here. GEDmatch.com users' policy recites: "The analysis and comparison results presented on this site are provided "as is" and no representations are made regarding their accuracy or usability. Changes in software and analysis tools may be made from time to time that could change results from those previously provided. The operators of this site are not responsible for the consequences of using the information provided on this site" (<https://www.gedmatch.com/policy.php>). We don't have access to the databases, classificatory systems, and algorithms used to create racial profiles. Black-boxed technologies deprive us from understanding why and whether this is the case. While some Stormfront users reported that GEDmatch, Dodecad K12b, and Dr. McDonald's BGA project retrieved "whiter" results. It is beyond the scope of this paper to assess if this is indeed the case, and, if it is, what are the reasons behind this. However, we can say that we found a clear tendency among Stormfront members to trust and prefer these platforms, tools, and calculators over proprietary companies' tools.

6. Conclusion

In examining how white nationalists use GATs to establish and maintain identity and community, we see a shift from stereotypical phenotypic definitions of whiteness to a greater reliance on complex genetic data to guide biosocial processes. These observations raise significant questions around the role of the scientific community, GAT production companies and marketers, and other stakeholders in maintaining an awareness of how the production and dissemination of GATs can affect pernicious uses of scientific methods and tools.

10. References

- [1] Helft, M. 2017. Ancestry.com DNA Database Tops 3M, Sales Rise To \$850M.
- [2] Bolnick, D.A. et al. 2007. The Science and Business of Genetic Ancestry Testing. *Science*. 318, 5849 (Oct. 2007), 399–400.
- [3] boyd, danah and Crawford, K. 2012. Critical Questions for Big Data. *Information, Communication & Society*. 15, 5 (Jun. 2012), 662–679.
- [4] Epstein, S. 2004. Bodily Differences and Collective Identities. *Body & Society*. 10, 2–3 (Jun. 2004), 183–203.
- [5] Fullwiley, D. 2014. The "contemporary synthesis." *Isis*. 105, 4 (Dec. 2014), 803–814.
- [6] Gould, S.J. 1981. *The Mismeasure of Man*. W. W. Norton & Co.
- [7] Jenkins, H. 2006. *Convergence Culture: Where Old and New Media Collide*. NYU Press.
- [8] Leonelli, S. 2013. Why the Current Insistence on Open Access to Scientific Data? *Bulletin of Science, Technology & Society*. 33, 1–2 (Feb. 2013), 6–11.
- [9] M'Charek, A. 2005. *The Human Genome Diversity Project: An Ethnography of Scientific Practice*. Cambridge University Press.
- [10] Nelson, A. 2008. Bio Science Genetic Genealogy Testing and the Pursuit of African Ancestry. *Social Studies of Science*. 38, 5 (Oct. 2008), 759–783.
- [11] Nordgren, A. and Juengst, E.T. 2009. Can genomics tell me who I am? *New Genetics and Society*. 28, 2 (Jun. 2009), 157–172.
- [12] Omi, M. and Winant, H. 1986. *Racial formation in the United States: from the 1960s to the 1980s*. Routledge & Kegan Paul.
- [13] Panofsky, A. and Bliss, C. 2017. Ambiguity and Scientific Authority: Population Classification in Genomic Science. *American Sociological Review*. 82, 1 (Feb. 2017), 59–87.
- [14] Pasquale, F. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- [15] Phillips, W. 2015. *This Is Why We Can't Have Nice Things*. The MIT Press.
- [16] Rabinow, P. 2005. Artificiality and Enlightenment: From Sociobiology to Biosociality. *Anthropologies of Modernity*. J.X. Inda, ed. Blackwell Publishing Ltd. 179–193.
- [17] Reardon, J. and TallBear, K. 2012. "Your DNA Is Our History" Genomics, Anthropology, and the Construction of Whiteness as Property. *Current Anthropology*. 53, S5 (2012), S233–S245.
- [18] Shriver, M.D. and Kittles, R.A. 2004. Genetic ancestry and the search for personalized genetic histories. *Nature Reviews Genetics*. 5, 8 (Aug. 2004), 611–618.
- [19] TallBear, K. 2014. *The Emergence, Politics, and Marketplace of Native American DNA*. Routledge Handbooks Online.