

Purloined Significance: How Recidivism Algorithms Capture, Transform, and Automate our Intersubjective Unconscious as Data

Macy McDonald
SUNY University at Buffalo

Follow this and additional works at: <https://docs.lib.purdue.edu/clcweb>



Part of the [American Studies Commons](#), [Comparative Literature Commons](#), [Other Arts and Humanities Commons](#), [Other Film and Media Studies Commons](#), [Rhetoric and Composition Commons](#), and the [Social and Behavioral Sciences Commons](#)

Dedicated to the dissemination of scholarly and professional information, [Purdue University Press](#) selects, develops, and distributes quality resources in several key subject areas for which its parent university is famous, including business, technology, health, veterinary medicine, and other selected disciplines in the humanities and sciences.

CLCWeb: Comparative Literature and Culture, the peer-reviewed, full-text, and open-access learned journal in the humanities and social sciences, publishes new scholarship following tenets of the discipline of comparative literature and the field of cultural studies designated as "comparative cultural studies." Publications in the journal are indexed in the Annual Bibliography of English Language and Literature (Chadwyck-Healey), the Arts and Humanities Citation Index (Thomson Reuters ISI), the Humanities Index (Wilson), Humanities International Complete (EBSCO), the International Bibliography of the Modern Language Association of America, and Scopus (Elsevier). The journal is affiliated with the Purdue University Press monograph series of Books in Comparative Cultural Studies. Contact: <clcweb@purdue.edu>

Recommended Citation

McDonald, Macy. "Purloined Significance: How Recidivism Algorithms Capture, Transform, and Automate our Intersubjective Unconscious as Data." *CLCWeb: Comparative Literature and Culture* 24.4 (2022): <<https://doi.org/10.7771/1481-4374.4090>>

This text has been double-blind peer reviewed by 2+1 experts in the field.

The above text, published by Purdue University Press ©Purdue University, has been downloaded 5 times as of 01/30/23.

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](#).

CLCWeb: Comparative Literature and CultureISSN 1481-4374 <<http://docs.lib.purdue.edu/clcweb>>
Purdue University Press ©Purdue University

CLCWeb: Comparative Literature and Culture, the peer-reviewed, full-text, and open-access learned journal in the humanities and social sciences, publishes new scholarship following tenets of the discipline of comparative literature and the field of cultural studies designated as "comparative cultural studies." In addition to the publication of articles, the journal publishes review articles of scholarly books and publishes research material in its *Library Series*. Publications in the journal are indexed in the Annual Bibliography of English Language and Literature (Chadwyck-Healey), the Arts and Humanities Citation Index (Thomson Reuters ISI), the Humanities Index (Wilson), Humanities International Complete (EBSCO), the International Bibliography of the Modern Language Association of America, and Scopus (Elsevier). The journal is affiliated with the Purdue University Press monograph series of Books in Comparative Cultural Studies. Contact: <clcweb@purdue.edu>

Volume 24 Issue 4 (September 2022) Article 10**Macy McDonald,****"Purloined Significance: How Recidivism Algorithms Capture, Transform, and Automate our Intersubjective Unconscious as Data"**<<http://docs.lib.purdue.edu/clcweb/vol24/iss4/10>>

Contents of ***CLCWeb: Comparative Literature and Culture 24.4 (2022)***Special Issue: ***Platform Psychoanalysis***. Edited by **Matthew Flisfeder**<<http://docs.lib.purdue.edu/clcweb/vol24/iss4/>>

Abstract: Ever since ProPublica published their groundbreaking analysis of Northpointe's Correctional Offender Management Profiling for Alternative Sanctions Core Risk and Needs Assessment software (COMPAS) in 2016, this web-based decision support system (DSS) has spawned a wide range of critiques and charges of racial bias. COMPAS provides a full suite of decision support applications to the US prison-industrial complex, including algorithmically derived recidivism predictions that increasingly guide parole decisions. The larger conversation surrounding COMPAS raises the question of how we analyze powerful, and yet opaque, data assemblages. In this article, I model an allegorical analysis of data assemblages. I argue the skills of literary analysis can help us surmount the "trap" of analyzing black-box algorithms. Guided by Poe's "The Purloined Letter," I analyze how COMPAS purloins our mediated ability to give an account of ourselves under the guise of objective, data-driven decision making.

Macy MCDONALD

Purloined Significance: How Recidivism Algorithms Capture, Transform, and Automate our Intersubjective Unconscious as Data

ProPublica's 2016 article "Machine Bias" begins with the tale of two crimes. The first begins as a childish lark. A teenage girl and her friend snag a child's bike and scooter, are spotted by a concerned neighbor, arrested, and charged with burglary and petty theft. The second crime is more habitual. A middle-aged man with a long history of petty crime shoplifts merchandise worth approximately as much the kids' wheels; he's caught and also charged. However, once they are charged, our defendants' fates diverge. The biased machine from the title classifies the teenager as high risk for future crimes whilst it categorizes our career criminal as low risk for reoffending. Time proves the machine false. Two years later, the teen had not been charged with any new crimes while the man was serving an eight-year sentence for subsequent thefts.

"Machine Bias" is not an investigation of those crimes, but rather a public audit of the machine that so drastically misclassified their perpetrators. *ProPublica* investigators, led by Julia Angwin and Jeff Larson, asked why two crimes of similar magnitudes would result in such distinct classifications. For the culprit, all signs pointed to COMPAS, or Northpointe's Correctional Offender Management Profiling for Alternative Sanctions Core Risk and Needs Assessment software. Northpointe Suite Software Systems, operating as Equivant since 2017, markets COMPAS as a web-based decision support system (DSS) that provides a full suite of decision support applications to the US prison-industrial complex. These include algorithmically derived recidivism predictions that increasingly guide parole decisions. In their groundbreaking public audit—later listed as a finalist for the 2017 Pulitzer in Explanatory Reporting—*ProPublica* investigators claimed COMPAS's recidivism results exhibited strong racial bias against Black defendants. In addition to age, gender, and criminal history, our two petty thieves are also racially distinct. The teen is Black, while the man is white. Larson et al. analyzed 10,000 criminal defendants in Broward County, Florida. They compared COMPAS's predicted recidivism rates with the actual rate that occurred over a two-year period and concluded COMPAS's predictions were both inaccurate and racially biased.

The Trap of Black Box Algorithms

Angwin et al. frame their analysis of COMPAS "as part of a larger examination of the powerful, largely hidden effect of algorithms in American life." And it's within that framework that I'd like to begin my own analysis. Since 2016, analyses of COMPAS in particular, and algorithms in general, have proliferated. However, as many of these systems are proprietary and therefore concealed from public analysis, the larger debate has become mired in questions of transparency, accuracy, and bias. And COMPAS, as one of the more widely documented machines, has become a testing ground for algorithmic analyses. Even *ProPublica's* analysis of COMPAS has since received its own share of criticism for assumptions made about the proprietary algorithm, including a recent debate in *Harvard Data Science Review (HDSR)*.

In "The Age of Secrecy and Unfairness in Recidivism Prediction," Rudin et al. partially reconstructed the COMPAS recidivism risk-scoring algorithm. They claim COMPAS does not depend linearly on a defendant's age, which contradicts assumptions about the model made by *ProPublica*, and thus throws their findings of bias under suspicion as well. Further, they claim the larger issue in regards to COMPAS is not fairness but transparency, since more accurate analyses could be made if the proprietary model was made available to researchers. Rudin et al.'s analysis of both *ProPublica's* argument and Northpointe's much critiqued machine received a response from Northpointe representatives titled "Setting the Record Straight: What the COMPAS Core Risk and Need Assessment Is and Is Not." As the title suggests, Northpointe representatives critiqued Rudin et al. for their assumptions about how COMPAS works and concluded with the assurance they're currently pursuing copyrights so that COMPAS's model could be made more transparent. *HDSR* helpfully links both corresponding articles through their web portal, so you can easily see that this conversation sparked at least 14 other responding analyses.

I lay out this condensed version of the continued public debate about COMPAS's predictions, not because I intend to settle the question of bias, but because these ongoing attempts to answer whether or not COMPAS produces racially biased predictions indicate another pressing social question: how do we analyze powerful automated decision systems? And reciprocally, how are these analyses impacting the public? These questions are difficult to answer, especially, as is the case with COMPAS, when they

are "black box" privately owned systems that are not made available for public engagement.¹ As Rudin et al. note in their subsequent response to Northpointe's critique: "It is easy to fall into the trap of explaining a black-box model incorrectly" ("Broader Issues" 13). But what is that trap, exactly? Data scientists like Rudin et al. propose greater transparency as a precursor for increased fairness. However, private companies receive capitalist incentives—namely money and power—by withholding the operating logics of DSS frameworks. Thus, while the evidence these systems negatively impact the public mounts, the specificity of their code remains inscrutable to many positivist modes of critique. So, while we can certainly demand transparency, we also need methods to analyze black-box models like COMPAS that circumvent the "trap" of their opacity.

Algorithms as Allegories

We may never peer into the abyss of COMPAS's specific proprietary operating logics, but we can contextualize COMPAS and other decision supporting machines within the larger operating legacy of sovereign power. In that endeavor, I propose we read algorithms like COMPAS as if they were allegories in order to plot some of the common pitfalls in analyzing proprietary algorithms.² Allegories are stories that say one thing and mean another. Allegories take signs, construct a narrative in which they establish connections between different types of signs, and in doing so transform the symbols they organize through new semantic relationships. In this way, they share a formal resemblance to algorithms. Algorithms, as parts of larger data assemblages, take a (data)set of signs, establish correlations between different classifications of signs, and in doing so transform those signs through new statistical relationships. Given this commonality, the tools of literary scholarship—specifically the skills of close reading, rhetorical analysis, and allegoresis—can help us better understand, engage with, and analyze the proprietary algorithms heralding the reign of data-driven decision making.

My literary frame in this novel approach to analyzing COMPAS is Edgar Allan Poe's "The Purloined Letter." Like COMPAS, the surface tale of Poe's story is the tale of two crimes. The first is filled with political intrigue. Minister D— purloins a letter from the French queen's apartment and proceeds to use the letter, contents never disclosed, to blackmail her. The queen dispatches the Parisian police force to retrieve this mysterious missive, only to fail in their reconnaissance. Discouraged, the Prefect of the police force calls on detective C. Auguste Dupin for advice. Initially, Dupin teases the Prefect for failing to solve such a "simple" problem. Later in the story though, Dupin provides the letter—which he stole from the minister's apartment in near perfect symmetry with the original theft from the queen's chamber—for our bumbling Prefect, in exchange for 50,000 Francs. Of course, like in *ProPublica's* investigation of COMPAS, the two thefts are merely an entry point for a deeper mystery.

As one of Poe's three detective stories, "The Purloined Letter" and its deeper mystery has inspired analysis since its 1841 publication. As Barbara Johnson argues in her exemplary reading of Poe—and Lacan and Derrida's subsequent analyses—, "The Purloined Letter" is an allegory of "the act of analysis of the act of analysis" (213). Though, I would add that "The Purloined Letter" is specifically the act of analysis of the act of analyzing an opponent's future choices.³ In the act of analyzing itself, Poe's story demonstrates that there is no neutral metalanguage with which to perform an analysis and that no analysis can intervene without transforming and repeating elements from the sequence it analyzes (213-14). Poe riddles his tale with mythic references, dubious puns, and satirical rants about empiricism. Further, its rich allegorical structure allows the surface tale of crime and politics to be oddly doubled with a tale of a schoolboy who was able to win most of his peers' marbles by learning to predict his opponents' moves in what should have been a game of chance.

The game of even and odd enters the narrative while Dupin recounts how he recaptured the purloined letter for the captivated Prefect. As Dupin reveals his ploy, he negatively compares the Prefect to a young schoolboy who captured all his peers' marbles through a game of even and odd. The simple game consists of holding marbles in your hands and asking your opponent if you're holding an even or odd number. In the first round, the schoolboy must always guess, but in the second, he assumes that a simpleton will change from even to odd (or *vice versa*) and a foe one degree above a simpleton will stay

¹ This term was coined by Frank Pasquale in *The Black Box Society: The Secret Algorithms That Control Money and Information* (doi:10.4159/harvard.9780674736061) published by Harvard UP in 2015.

² In identifying this formal resonance between algorithms and allegories, I'm preceded by Alexander Galloway's compelling critical trilogy: *Allegories of Control*.

³ A careful reader of psychoanalysis may notice a certain resonance between this statement and Lacan's essay "Logical Time and the Assertion of Anticipated Certainty: A New Sophism." However, as Lacan's prisoners' dilemma immediately dismisses the use of probabilistic reasoning, I have chosen to focus on Lacan's analysis of Poe (in which he analyzes both probabilistic reasoning and traditional logic).

the same. The schoolboy's ability to beat the odds, and his opponents, relied on an assessment of his opponents' intelligence and a prediction of their future choices based on that assessment. The boy assessed his opponents by mimicking the expression on their faces and then examining the "thoughts or sentiments" that arose in relation to his mirrored expression (Poe 258). In his seminar on Poe's story, Lacan also analyzes this game of even and odd. From this game, and Freud's work on the unconscious, Lacan elaborates a theory that the unconscious repeats certain symbolic patterns. In his "Parenthesis of Parentheses," added a decade after he first published the seminar, this leads him to conclude that "it is not unthinkable that a modern calculating machine, by detecting the sentence that, unbeknown to him and in the long term, modulates a subject's choices, could manage to win beyond any usual proportions in the game of even and odd" (45). COMPAS is one such calculating machine, though given the debate surrounding its accuracy, perhaps less prodigious than Poe's schoolboy. Building from Poe and Lacan, we can grasp COMPAS's predictions not as verdicts given in an objective, neutral metalanguage with relative accuracy, but as sentences that repeat and transform the patterns from which they are derived.⁴ And these sentences (to parole or not to parole) are based on an analysis of the current constraints and motivations of potential parolees.

Capturing Significance

COMPAS captures those motivations and constraints through a 137-point questionnaire given to defendants. In their 2020 rebuttal to Rudin et al., Northpointe representatives object to the assumption that COMPAS uses all of their 137-point questionnaire to calculate the risk of recidivism. However, I'm less concerned with how the questions are weighted than with the types of subjective experiences those questions capture. For example, here are three of the questions defendants are asked in the questionnaire, included in documents audited by *ProPublica*:

- 32. If you lived with both parents and they separated, how old were you at the time?
- 38. Was one of your parents (or parental figure who raised you) ever sent to jail or prison?
- 67. In your neighborhood, have some of your friends or family even been crime victims?

Defendants respond using a series of check boxes, and their responses are then quantified, weighted, and serve as input for the software. From there they are algorithmically analyzed and reproduced as bar charts. The stories of how your mother left when you were seven, how your father was in and out of jail, maybe even the death of a friend due to a crime in your neighborhood are captured and mutated into a score that ultimately weighs on your future. These personal narratives become parameters that determine whether or not you'll be released on bail or parole, what treatments would be appropriate for you, and/or how you'll be classified in prison. Like Minister D—, Equivant has captured personal significance and purloined defendants' political choices by transforming that personal significance into a cipher for future behavior.

Of course, COMPAS's process for capturing personal significance is both state sanctioned and more complex than blackmail. Northpointe's "Practitioner's Guide to COMPAS Core" describes their process of identifying defendants' constraints as risk/needs assessments. They claim that statistically based risk/needs assessments are not only accepted methods but are "superior to human judgement" when it comes to "managing offenders." (1-2). They further claim the objective of COMPAS's scales is to provide a "holistic view of the person to address supervision and treatment needs for rehabilitation" (2). They later state people are "complex creatures" and provide brief outlines of six different criminological theories that frame the "salient life events and influences" they assert must be considered to capture a holistic view (5). From these six theories, they have distilled two basic scales: risks and needs. Needs, they emphasize, are simple descriptive constructs such as financial problems meant to describe the "offender." Risks, on the other hand are tied to models of behavior and "discriminate between offenders who will and will not recidivate" (7). Claims of holistic capture aside, the rationales for COMPAS scales are largely presented in simple binaries: risk/needs, will/not recidivate, non/violent recidivism. They, of course, provide additional information about their norm groups and internal assessments for accuracy. Though, as this is not a statistical analysis, I will not assess those claims here. Rather, what interests

⁴ For more on how Poe's short story influenced Lacan's thoughts on cybernetics, I recommend Lydia Liu's unparalleled article, "The Cybernetic Unconscious: Rethinking Lacan, Poe, and French Theory."

me is the framing of "offenders" as both complex and reducible to predictable binaries, and the assertion that this deduction is super human.⁵

The claim of super human judgement comes from the automation of these risk and needs scales via an algorithm. Since the algorithm(s) COMPAS uses are proprietary,⁶ we can't know exactly how they automate the analysis of the data they capture with their questionnaire. However, we can deepen our understanding of the terms "data" and "algorithms" and the power structures these terms alternatively conceal and reproduce. Beyond the opacity of a single proprietary DSS, there are latent snares embedded in our popular understanding of these terms. In particular, this idea that data contain an objective truth beyond the reach of the subjective human is built on a pervasive misconception of what data are. While there's no one source for such a widespread misunderstanding, it is clearly captured by Kenneth Cukier and Viktor Mayer-Schönberger in their popular work *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Cukier and Mayer-Schönberger note that "'data' means 'given' in Latin, in the sense of a 'fact,'" though they emphasize that "[t]oday data refers to a description of something that allows it to be recorded, analyzed, and reorganized" (78). My contention with this definition begins with the idea of "givenness." Data are not given but recorded. And the ways in which data are recorded complicate the idea that "given" data add up to facts or truth. Recording, after all, encodes both human perspectives and motivations in the process of capture. Rob Kitchin addresses the popular misconception of data as "given," in *The Data Revolution*. He writes:

[I]n general use, data refer to those elements that are taken; extracted through observations, computations, experiments, and record keeping. Technically, then, what we understand as data are actually *capta* (derived from the Latin *capere*, meaning 'to take'); those units of data that have been selected and harvested from the sum of all potential data (2).

With the understanding of data as *captured*, we can change the conversation around opaque data assemblages from one concerned with truth of those systems, their accuracy and transparency, to one questioning what has been captured, why, and what frameworks inform the gleaning.⁷

We typically indicate those organizing frameworks of data assemblages with the word "algorithm." Technically, an algorithm is a series of logically ordered steps that shape a body or bodies of measurable input data to produce desired output. While captured data serves as the observational input in a data assemblage, algorithms create statistical relationships between the input data points or sets. Practitioners also refer to algorithms as models or frames. These models, when applied to data sets, establish correlations between different variables. Literary scholars might liken an algorithm to a poetic form like a sonnet or a haiku.⁸ However, in the colloquial usage, "the algorithm" often refers not to a specific model, but to a larger data assemblage. In what follows, I'll adopt Tarleton Gillespie's reframing of "algorithms" as synecdoches. Synecdoches, as Angus Fletcher argues, are an element of allegory (84). As a literary trope, synecdoches allow a part to symbolize a whole, and thus distill a plural concept into a singular object or image. They further imply, as symbols of the whole, that there is a coherent structural whole that can be distilled in a single representative. Algorithms have become a nonvisual symbol for larger data assemblages in which they operate. For example, if an app or website behaves oddly or even discriminatorily, people will say "the algorithm did it." "The algorithm," in that popular connotation, has become a cultural shorthand for the complex web of contingent relationships that make modern information communication technologies possible. As Gillespie notes, this symbolization of data assemblages may be unavoidable in our work, but cannot remain unexamined, lest the synecdoche "algorithm" reify and obscure the processes of which algorithms are only a part (23).

And though I realize it's an odd leap, this is why we can fruitfully critique algorithms through some of the same methods literary scholars employ when analyzing allegories. Allegories have a long tradition of subversively concealing meaning, often to avoid censorship and/or mask a political critique.

⁵ For more on how personal information is abstracted as data for binary categorization, see John Cheney-Lippold's *We are Data: Algorithms and the Making of our Digital Selves*.

⁶ Several scholars have surmised COMPAS primarily uses a linear-regression algorithm, though if that's the case, it is unclear why they would need to seek copyright as linear regressions are common structures that are easily duplicated and have a wide array of applications. Copyrighting a linear regression algorithm would make about as much sense as copyrighting the structure of a sonnet. True, you can claim an individual sonnet, but the form is beyond individual ownership.

⁷ Johanna Drucker also argues for understanding data as *capta* as early as 2011. See works cited.

⁸ For more on the relationship between software and poetic forms (specifically metaphors), see Matthew Flisfeder's *Algorithmic Desire: Toward a New Structuralist Theory of Social Media*.

Etymologically, "allegory" brings together *allos* and *agorei*, which when combined mean to speak publicly otherwise. So, by analyzing algorithms as if they were allegories, we can examine the public claims made by the owners and operators of data assemblages in tension with how they might otherwise function. This emphasis on the tension between claim and process allows us to focus on the relationships these frameworks posit and how those relationships impact the people captured in these systems.

COMPAS as an Allegory

We can easily access COMPAS's publicly stated surface tale through their website. Equivant advertises COMPAS as an "integrated web-based assessment and case management system for criminal justice practitioners who must make decisions regarding community placement, supervision, treatment and case-management of offenders." Equivant claims the software is used across the country, and over the past decade they've expanded their suite of products to include: case management systems for courts, attorneys, and supervision; pretrial risk (the risk someone will fail to appear for their court date), custody management, classification management, problem solving court management, and an Integrated Justice Information Sharing Broker. So, while predicting recidivism is one of the primary applications, the larger software set is a full suite of decision support system meant to manage cases and automate decisions (Equivant). In addition to what the software does technically, how Equivant markets COMPAS is also relevant.

There is a key takeaway from Equivant's branding we should keep in mind before we delve into the deeper mystery beyond the surface tale: their promise of "defensible decisions." Equivant's website boldly advertises its products as "Software for Justice." Thus, justice practitioners—which they use as a shorthand for judges, public attorneys, court clerks, and law enforcement officials—are their target audience. However, in their marketing to this group, Equivant never pitches the software as being more just or fair. Despite their slogan of "Software for Justice," their main appeal to justice practitioners is *defensible* decisions. They imply that Equivant's software can help make juridical decisions that are not only "evidenced-based" and "objective," but also more efficient. Equivant's appeal implies that justice practitioners not only need to defend their decisions, but that data can help them do so. The idea of justice for defendants is notably absent from their marketing; instead, they offer defense for justice practitioners. Keeping an allegorical approach in mind, we can think of Equivant's branding as COMPAS's surface tale.

ProPublica fact checked that publicly stated story in their analysis of alleged racial bias. *ProPublica* found that the software predicted recidivism correctly 61% of the time and correctly predicted violent re-offending only 20% of the time. Beyond the overall accuracy rate, they also found that COMPAS had a strong racial bias, particularly against Black defendants. Observing the two-year period after being granted parole, they found that Black defendants were nearly twice as likely to be misclassified as higher risk compared to their white counterparts (45 percent vs. 23 percent). Using the same two-year window, white defendants were misclassified as low risk almost twice as often as Black re-offenders (48 percent vs. 28 percent). Overall, they found that Black defendants were 45 percent more likely to be classified as higher risk than white defendants, even when controlling for age, gender, future recidivism, and prior crimes (Larson et al.). *ProPublica's* analysis set off a chain of subsequent critiques I referenced in the introduction and will not reiterate here. What we can glean from the larger conversation of analyses of COMPAS is that COMPAS's predictions are themselves an analysis of human behaviors that present as objective and defensible (data-driven) but are based on capta ensnared in subjective (and partially opaque) frameworks. This is why I propose that rather than fall into the trap of empirically analyzing a black-box model like COMPAS, we instead approach it as an allegory and ask why these decisions need to be defended.

As Johana Drucker argues, statisticians and other specialists are exceedingly aware of the interpretive frameworks that inform and shape their analyses. However, these frameworks are often rendered illegible when they are presented in a graphic (as COMPAS's results are) or when used by non-specialists who take numbers at face value. Building from Drucker, I would add that companies like Equivant often further render these interpretive frameworks illegible out of a desire to persuade critics and customers alike of their legitimacy. Data in the popular connotation of "given" lends a guise of objective truth to information. To say a decision was "data-driven" is also to say it was made based on a set of norms that have been rendered rhetorically difficult to dispute. In many ways, claiming a decision was "data-driven" is an attempt to surmount criticism. So, when COMPAS promises justice practitioners "defensible" decisions, they really mean that they have translated the norms of juridical decision-making into a framework that is resistant to critique. This "defends" justice practitioner's decisions and shifts political critique from the US justice system to an automated, opaque proxy.

Of course, judges and hearing examiners have decided defendants' futures through this same exercise of power long before its augmentation by COMPAS. Both pre-trial rulings and parole hearings typically consist of a presentation of past actions, an assumption of future behavior based on that information, and a decision that purloins political choices. This is a function of state power that COMPAS seeks to defend. COMPAS has not changed that these powerful decisions occur, but it has augmented *how* those decisions are made. Equivant claims this augmentation is objective, but we can delve further into this term. "Objective" connotatively implies a fair or unbiased choice. This is the definition of objective *ProPublica* and subsequent analyses have critiqued Equivant for failing to meet. However, "objective" also denotes a realm of sensible experience with a measurable relationship between objects (*OED Online*). This denotation forms the basis of scientific objectivity implied in the connotative use of "objective" as synonymous for fair. Obviously, COMPAS's predictive functions could never meet this sensory definition of objectivity, as they deal with future behaviors that do not yet exist and therefore cannot be sensed. Further, a cursory reading of Equivant's questionnaire reveals that COMPAS measures not the future, but rather defendants' pre-existing constraints and motivations for future actions.

On the surface, COMPAS's process resembles the two games from Poe's short story: the game of even and odd and the battle of wits between Dupin and his dastardly double Minister D—. Both the young school boy and Dupin play the odds at outwitting their opponents by assessing their constraints and aims. However, there is a deeper resonance between COMPAS and "The Purloined Letter" than the improbable games of their surface tales. On a deeper level, both COMPAS and "The Purloined Letter" raise questions about how ownership changes the significance of information. In his seminar on Poe's story, Lacan argues the letter Minister D— captures holds an entirely different significance for the Queen than the significance observed by the Minister (18). From this point, Lacan raises a question of ownership. The meaning of a letter is already intersubjective as it is sent from one person to another, but the Minister's capture of that letter adds another transference and embeds new meaning through recognition. He further transforms that meaning to exercise power. Similarly, COMPAS's questionnaire captures intersubjective experiences and through their analysis transforms those experiences to exercise power. The experience of having a parent arrested is already an intersubjective experience, and there is a question of who that experience belongs to. It is both the parent's and the child's, but also the arresting officer's and anyone connected to the incident and the jail or prison in which the parent was detained. COMPAS observes that experience through the questionnaire and assigns it a different meaning.

As I argued previously, "The Purloined Letter" is the act of analysis of the act of analyzing an opponent's future choices. COMPAS similarly is an automated analysis produced for parole boards as they analyze defendant's future choices. However, contrary to Equivant's claims, there is no neutral metalanguage with which to perform an analysis. And by capturing intersubjective information in their questionnaire, COMPAS transforms the significance of that information and repeats past biases embedded in the justice system. Once combined and put in relation to all the other captured data points, these observations of personal experiences are applied to a defendants' futures with powerful results.

The Purloined Word

While Minister D— purloined a letter, COMPAS purloins a word. *Parole*, as borrowed from the earlier French usage, literally means "word" as in word of honor. To be granted release from prison on parole etymologically implies being released on your own testament or promise. COMPAS, through its questionnaire and subsequent algorithmic analysis of personal data thus purloins our already mediated ability to account for ourselves. As Judith Butler notes in *Giving an Account of Oneself*, ethical and moral inquiry have long assumed an "I" that can act and therefore be held accountable for those actions. When we are brought before a court of law, or even a parole board hearing, we are called upon to account for not only the legality of our actions, but also our circumstances and motivations. As Butler argues, the "I" that acts emerges from an unconscious relationship to linguistic and technological norms, and crucially to other people (8). We are inextricably relational. However, the impact of these relationships on ones' actions are generally opaque to the self. COMPAS attempts to render the opacity of those unconscious relationships that author our actions legible to a juridical assemblage. In asking whether a parental figure was ever sent to jail, COMPAS (in the sense of synecdoche), really asks a defendant to record an unconscious and emerging relationship between past experience and future action without completing the necessary step of performing that allocutionary deed of accounting for ourselves in our own words. As Butler argues, this speech act is essential to the process of self-reflection and transformation (129-30). Parole hearings, and juridical hearings more generally, are moments when we

are called upon to account for ourselves; either to explain our actions or narrativize how we have changed during our incarceration.

Purloining this allocutionary moment is an especially poignant loss because of the stasis it enacts. The patterns that COMPAS automates come from prior arrest and recidivism data. Per the Northpoint Practitioner's Guide, a major function of COMPAS is sorting defendants according to types derived from previous data. Typifying people based on collective abstractions is essentially just stereotyping. When we use systems like COMPAS to type categorize people based on previous patterns, we concretize those patterns and deny people the opportunity to participate in shaping future patterns of behavior. COMPAS arrests an interdependent relationship and detains our ability to mold social patterns. As Antoinette Rouvroy argues, this irrational rationalization attempts to render the world "predictable but insignificant" (222). Though, I must stress this is an *attempt* to render an interdependent unconscious legible, not a success. Just as subjects cannot fully account for their opacity, COMPAS and its practitioners cannot account for the linguistic, technological, relational nature of the *capta* they analyze.

So, recalling COMPAS's branding, what COMPAS promises to defend is judging a subject's word as false. We can never really know if we can take people at their word. Rather than attempt to foster trust or perhaps remove constraints (like systemic poverty) that contribute to recidivism, systems like COMPAS deepen the oppositional lines between our juridical assemblage and those that have "offended" our laws. COMPAS is a method—not unlike Poe's school boy's—of assessing an opponent and trying to guess their next move. Performing our accountability is potentially a moment of self-transformation in which we reflexively recognize and challenge our relationship to a set of norms that Butler, building from Foucault, calls a regime of truth (22). As they both note, these regimes provide a framework for the emergence of self and for moments of self-recognition by framing recognizable forms of being. The legal system in any country is one of these regimes of truth, not because they hold some *a priori* claim to truth, but because they provide the terms of recognizable being or personhood. So, when output from COMPAS enters a parole hearing, it augments a performance and transformation of self in the face of a legal framework by claiming its analysis of our data is truer than our word.

While it's tempting to permit data assemblages a guise of superhuman objectivity, as I've discussed previously, data is nothing more than captured observations rendered quantitatively for analysis of reproducible patterns. As Lacan notes, a calculating machine that can predict unconscious repetition relies on nothing more than the reasoning of "ordinary man" (45). This reading of COMPAS in particular was tested by Julia Dressel and Hany Farid in their article "The Accuracy, Fairness, and Limits of Predicting Recidivism." Dressel and Farid tested COMPAS's accuracy, not by emulating the black box model, but instead by taking the assumed 137 data points available and hiring gig workers via Amazon's Mechanical Turk program to assess whether defendants would reoffend in a two-year period. They found that the ordinary "wisdom of the crowd" outperformed COMPAS in predicting recidivism. COMPAS's analysis of risk and needs is thus neither superior to human judgement nor objective. Though rebranded as a risk/needs assessment, this model of reasoning that predicts actions based on constraints and motivations "lies at the bottom of all the spurious profundity which has been attributed to Rochefoucault, to La Bruyère, to Machiavelli, and to Campanella" (Poe 258).

Systems like COMPAS are often critiqued for their machine bias (as by *ProPublica*) or vaunted for their superiority to human judgement (as in the Practitioner's Guide). However, Dressel and Farid's pithy analysis questions the false binary inherent in distinguishing COMPAS's predictions from human judgement. We should not be too hasty in either vilifying or idolizing systems like COMPAS. Like Dupin, we should recognize the preeminent game at play. Machine bias is a fac-simile of human judgement; they mirror one another as thoroughly as Dupin and Minister D—. And while it's politically necessary to track machine biases, we might treat them more as evidence of our own asymmetries of power rather than technical failings. COMPAS is a product of human judgement, the newest offspring in the line of bureaucratic social control techniques. And beyond the question of its accuracy lie the deeper issues of social reproduction and the legitimacy of power. Here, a return to Poe's tale might yet again prove a fruitful point of comparison.

Illegitimate Reproduction

Poe's tale is especially pertinent for framing an analysis of a recidivism algorithm due to its abiding interest in the repetition and reproduction of crimes as they relate to political legitimacy. Dupin purloins the Queen's letter from the Minister's apartment in a simile of the original crime. And, as Samuel Kimball argues, Dupin's ruse to recapture the letter metonymically indicates the crime concealed in the letter itself (22). Dupin hires a man to fire a musket in the street, startling a crowd of women and children. Though the musket is without balls, the "false report" distracts the Minister long enough for Dupin to

substitute his facsimile. As Kimball notes, Dupin is willing to sacrifice public peace to restore the Queen's peace. Dupin flouts the general law and order in service of an authority of law, the Queen. However, the letter itself, though contents technically unknown, is evidence of transgression on the part of that authority. And, given the Queen's desperation to conceal the contents from a higher authority, the King, we must assume her transgression threatens his own political legitimacy. Reading the clues provided by Poe, we may assume the Queen is having an affair. Such an affair could through the King's ability to reproduce his own power in a viable heir, and thus the legitimacy of legacy, into question.

Poe indicates this deeper breach of trust most significantly in the final line of the story. In revealing how he recaptured the purloined letter, Dupin also reveals his motivations. Rather than leave his facsimile blank, Dupin has copied the following broken line from Crebillon's version of *Atreus and Thyestes*: "*—Un dessein si funeste, S'il n'est digne d'Atree, est digne de Thyeste*" (265). The French roughly translates to "a scheme so baneful, if it is not worthy of Atreus, is worthy of Thyestes." Atreus and Thyestes were twins and grandsons of Tantalus. Though the details shift in the various retellings, as they always do with myth, the basic plot involves Thyestes seducing Atreus's wife, a violation of brotherly trust that sets of a decades long blood feud that ends with Atreus slaughtering Thyestes child/ren and serving them to him at a banquet meant to commemorate the end of their feud just before being deposed as a tyrant. In Crebillon's version, Atreus raises Thyestes' illegitimate son as his own in hopes he will someday murder his biological father. When this plot fails, Atreus murders the son, fills a sacred chalice with his blood, and toasts Thyestes with the blood of the illegitimate offspring in a perverse reconciliation of their blood feud (Sinnet).

As Johnson notes, Dupin rectifies the initial purloining of the letter by repeating it. And the facsimile he leaves in its place references yet another instance of "rectification-as-repetition-of-the-crime (219). Atreus, in serving Thyestes the "fruit" of his illicit union, rectifies Thyestes's disruption of Atreus's lineage by repeating the crime. Of course, on a more literal level, Thyestes is guilty of adultery while Atreus has murdered a boy he raised as his own son, but symbolically the crimes are the same. Each brother has disrupted the other's ability to reproduce their own power. Johnson notes the irony of Atreus forcing Thyestes to literally eat the fruit of his illicit union but leaves a second irony on the table. Atreus and Thyestes can trace their line back to Tantalus; a man who served his own children to the gods to test their omniscience and as punishment will never enjoy the fruits of his crimes.

The repeatedly rectified crime that Poe indicates with his final line is line theft. Dupin tracks a pattern of disrupting the reproduction of lineages of power all the way from Tantalus to the Queen. They have each committed line theft, a crime Dupin repeats with a broken line. The same crime Minister D— conceals in his bid to blackmail the Queen. A crime repeatedly punned upon in the redaction of his name. Minister D—'s name is redacted with a dash, the very same symbol used to break the line Dupin copies from Atreus and Thyestes. Poe literally conceals the name of a man concealing the breakage of a patrilineal line with a line (dash) that breaks a line (of prose). The Queen, as many have surmised, is having an affair. But Poe, in his narrative dexterity, has thoroughly distracted us from the more eminent crime by playing out Dupin and Minister D—'s game.

With this final untranslated line, we are thrown into an intertextual cycle of crimes upon crimes with no clear origin, which is to say, we are in the realm of allegory. As we connect the symbolic thread from Tantalus to Atreus and Thyestes, to "The Purloined Letter," and finally to COMPAS we are left with the overarching question of the legitimacy of power. Tantalus challenges the god's omniscience by sacrificing his own children (the symbol of his ability to reproduce power as a ruler), Atreus and Thyestes repeat their grandfather's crime against their own generation and end up cursing their entire lineage, Dupin and Minister D— both recognize a transgression in the reproduction of power and both profit from it and treat it as a turn in their own vengeful game of powerful oscillations. Finally, COMPAS, though marketed for its objectivity, automates the same symbolic logic that would hold Atreus and Thyestes responsible for the crime of Tantalus.

Thus far, an allegorical reading has revealed that in contrast to its publicly asserted claims, COMPAS's predictions are neither objective nor superior to human judgement. However, this leaves the question of COMPAS's actual objective in purloining defendants' words. COMPAS's branding promises to defend justice practitioners as they judge parolees false, but it is more obscure about what Equivant gets in return for this defense. For this question, we might turn to the motivations of Poe's agonists; namely money and power. When the Prefect initially comes seeking Dupin's advice he begins with a rather vague description of a purloined document. When pressed, he admits that purloining this document "gives the holder of the document an ascendancy over the illustrious personage whose honor and peace was so jeopardized" (251). Purloining the Queen's letter grants ascendancy over her actions to Minister D—. So, the Minister's objective was political power. When Dupin recaptures the altered letter, he does

not rush to return it to the Queen nor notify the Prefect of his success. Instead, he waits for the Prefect to visit again in desperation and only after the Prefect reveals the reward for the letter has doubled and writes Dupin a check for fifty thousand francs does he hand over the letter (255). Dupin claims his objective was not merely financial, but also revenge for an "evil turn" Minister D— once served Dupin. Though, as he delights in the "*facilis descensus Averno*," it becomes clear this revenge is its own form of ascendancy (265). Thus, Dupin's objectives were both money and power. In Poe's allegory, neither of his two oppositional analyzers are properly protagonist or antagonist. They share essentially the same motivations, it's only that one was victor. Further, other than the anticipated fall of the Minister which we cannot be sure will come to pass, none of the characters are significantly moved or transformed in this story. It concludes with them in essentially the same positions as before, only that Dupin has captured some of the Queen's wealth.

Guided by Poe's allegory, we might also ask what COMPAS captures by purloining defendants' words. In the example of the parole hearing, the process has been modified though the figures remain. Parolees are still subject to the judgements of the hearing examiner who represents the interests of the US prison industrial complex. However, the structural inequalities (race, class, and gender) that inform those judgements have become formalized and obscured as data-driven decisions. Defendants' allocutionary roles have also been undermined. The figural positions are basically the same, but their relationships have been augmented by additional oversight. In their "Practitioner's Guide," Equivant notes that justice practitioners may reasonably be expected to deviate from COMPAS's predictions 10% of the time (31). This implies a level of surveillance in how the software is used. Though who exactly monitors the results remains unstated, given the proprietary nature of the system, representatives at Equivant must have at least a mediating role in that surveillance. Surveillance is, of course, a form of power. It grants the overseer ascendancy over the overseen. Further, as this is a proprietary system, that oversight (and all other services of the DSS) comes at a price. Thus, Equivant secures money and power.

Conclusion

Finally, we arrive back at the "trap" of analyzing a black-box algorithm. In Poe's allegory of an act of analysis of an act of analysis, the Prefect begin his search for the purloined letter by using every empirical means at his disposal. Poe includes details of their microscopic investigations of floorboards, how they've measured the parameters of both the rooms and the furniture to find secret drawers, and their examination of the moss between the bricks of the surrounding buildings. When the stymied Prefect first visits Dupin seeking advice for how to find the letter, Dupin first jokes that perhaps the problem is too "simple" for the Prefect and then suggests he "make a thorough re-search of the premises" (255). The Prefect assumes Dupin means the police should re-search the Minister's building to see if they simply missed the letter. However, as is characteristic of Poe's wordplay, Dupin could have also meant the Prefect needed to re-search his ideological premises. Likewise, while critics can devise increasingly complex ways to reconstruct proprietary systems, and perhaps even find errors and the reinforcement of crucial structural inequalities, we should not dismiss the simple premises that frame and motivate these systems.

Rather than cast our analyses on the Procrustean bed of accuracy, we might instead follow Johnson's lead in insisting every analysis repeats and transforms elements of its object. In my own allegorical analysis, I have compared COMPAS to Poe's "The Purloined Letter" to highlight how COMPAS claims to rectify a social issue—that of the possibility of false words given by parolees—by repeating the problem and issuing its own false statements. However, these two transgressions are merely the opening for a deeper mystery: how are algorithms transforming the legitimation of power? As synecdoches for fabled objective decisions, systems like COMPAS attempt to circumvent individual articulation in favor of identifying interpersonal patterns. COMPAS augments our judgment of each others' words by attempting to capture a shared unconscious logic. Of course, assessments of COMPAS's in/accuracy can reveal the enduring legacy of racism and the criminalization of underprivileged groups. This is crucial work. We should also be attentive to the context and lineages of these systems. When we make claims of machine bias or inhuman judgement, we risk mistaking the machine for the social system. This is exasperated by our tendency to mythologize our own judgement as super human simply because it has been automated and rendered numerically. Reading algorithms like COMPAS allegorically will not expose the black-box of their operations, but it can expose the social problems they obscure. In the case of COMPAS, an allegorical reading reveals a deep distrust of taking others at their word and a defensive desire to surpass, or at least conceal, our own inaccurate judgements.

Works Cited

- Angwin, Julia et al. "Machine Bias." *ProPublica*, 23 May 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed 3 Mar. 2019.
- Butler, Judith. *Giving an Account of Oneself*. ProQuest Ebook Central, New York. Fordham UP, 2005.
- "COMPAS Risk Assessment Sample" *ProPublica*, 23 May 2016, <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE>, Accessed 2 Mar. 2019
- Dressel, Julia and Hany Farid. "The Accuracy, Fairness, and Limits of Predicting Recidivism." *Science Advances* 4, no. 1 (2018): DOI: 10.1126/sciadv.aao5580.
- Drucker, Johanna. "Humanities Approaches to Graphical Display," *Digital Humanities Quarterly*, 5:1, 2011, <http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>.
- "Equivalent Home Page," *Equivant*, <https://www.equivant.com/>. Accessed 24 Mar. 2021.
- Fletcher, Angus. *Allegory: the theory of a symbolic mode*. Stanford, CA, Princeton UP, 2012.
- Flisfeder, Matthew. *Algorithmic Desire Toward a New Structuralist Theory of Social Media*. Northwestern UP, 2021.
- Freud, Sigmund. *Beyond the Pleasure Principle*. Translated and edited by James Strachey; Introduction by Gregory Zilboorg; with a Biographical Introduction by Peter Gay. Standard ed, New York, Norton, 1989.
- Gillespie, Tarleton. "Algorithm." *Digital Keywords: A Vocabulary of Information Society and Culture*, edited by Benjamin Peters, Princeton UP, 2016, pp. 18–30, <https://doi.org/10.2307/j.ctvct0023.6>.
- Jackson, Eugenie and Christina Mendoza. "Setting the Record Straight: What the COMPAS Core Risk and Need Assessment Is and Is Not." *Harvard Data Science Review*, 2(1). <https://doi.org/10.1162/99608f92.1b3dadaa>, 2020.
- Johnson, Barbara. "The Frame of Reference: Poe, Lacan, Derrida," *The Purloined Poe: Lacan, Derrida & Psychoanalytic Reading*, edited by John P. Muller & William J. Richardson. Johns Hopkins UP, 1988, pp. 213-251.
- Kimball, A. Samuel. "D-Ciphering Dupin's Fac-Simile Signature: The Infanticidal Implications of a 'Dessein Si Funeste.'" *The Edgar Allan Poe Review*, vol. 6, no. 1, 2005, pp. 20–36.
- Kitchin, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, Los Angeles, SAGE Publications, 2014.
- Lacan, Jacques. *Écrits (Writings). Nouvelle édition augmentée d'une préface*. Paris, Éditions du Seuil, 1970.
- Larson, Jeff et al. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, 23 May 2016, <https://www.ProPublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Accessed 3 Mar. 2019.
- Cheney-Lippold, John. *We Are Data: Algorithms and the Making of Our Digital Selves*. New York UP, 2017, <https://doi.org/10.18574/9781479888702>.
- Liu, Lydia H. "The Cybernetic Unconscious: Rethinking Lacan, Poe, and French Theory." *Critical Inquiry*, vol. 36, no. 2, 2010, pp. 288–320.
- Mayer-Schönberger, Viktor and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, ProQuest Ebook Central, Houghton Mifflin Harcourt Publishing Company, 2013., <https://ebookcentral-proquest-com.gate.lib.buffalo.edu/lib/buffalo/detail.action?docID=3304666>.
- "Objective, adj. and n." *OED Online*, Oxford UP, September 2021, www.oed.com/view/Entry/129634. Accessed 21 October 2021.
- Pasquale, Frank. *The Black Box Society: the Secret Algorithms That Control Money and Information*. Harvard UP, 2015.
- Poe, Edgar Allan. "The Purloined Letter." *Selected Tales*, , Oxford UP, 2008, pp. 249-265.
- "Practitioner's Guide to COMPAS Core." *Northpointe Inc.* 19 Mar. 2015, <http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-031915.pdf>. Accessed 6 Oct. 2021.
- Rouvroy, Antoinette. "Epilogue." *Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology*, edited by Mireille Hildebrandt, and Antoinette Rouvroy, Taylor & Francis Group, 2011. ProQuest Ebook Central, <https://ebookcentral-proquest-com.gate.lib.buffalo.edu/lib/buffalo/detail.action?docID=684096>.
- Rudin, Cynthia, et al., "Broader Issues Surrounding Model Transparency in Criminal Justice Risk Scoring." *Harvard Data Science Review*, 2(1). <https://doi.org/10.1162/99608f92.038c43fe>, 2020.
- . "The Age of Secrecy and Unfairness in Recidivism Prediction." *Harvard Data Science Review*, 2(1). <https://doi.org/10.1162/99608f92.6ed64b30>, 2020.
- Sinnett, Edward. *Atrous and Thyestes: a Tragedy in Five Acts/ altered from the French of Crébillon*. London, C. and H. Baldwin, 1822.

Author's profile: Macy McDonald is the 2022-23 Fisher Center Pre-Doctoral Fellow in Gender and Justice at Hobart and William Smith Colleges and a PhD candidate in English with a specialization in algorithm studies at The University at Buffalo, SUNY. Their research focuses on the increasing power of predictive decision support tools. Their dissertation, "Fictive Analytics," proposes qualitative methods for evaluating and transforming quantitative decision support tools. **Email:** <macymcdo@buffalo.edu>